AD_____

Award Number: W81XWH-05-1-0391

TITLE: Maintaining Genome Stability: The Role of Helicases and Deaminases

PRINCIPAL INVESTIGATOR: XiaoJiang Chen

CONTRACTING ORGANIZATION: University of Southern California
Los Angeles, CA 90089-1147

REPORT DATE: July 2008

TYPE OF REPORT: Final

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

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-07-2008	Final	1 Jul 2005 – 30 Jun 2008

4. TITLE AND SUBTITLE

Maintaining Genome Stability: The Role of Helicases and Deaminases

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0391

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

XiaoJiang Chen

E-Mail: Xiaojiang.Chen@usc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

University of Southern California
Los Angeles, CA 90089-1147

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 Both helicases and deaminases are enzymes that play an important role in maintaining genomic stability and immune responses. Errors in duplicating DNA can result in genomic instability, leading to various human diseases, such as cancer, immune system disorder, muscle dystrophy, and neurodegenerations. Thus, maintaining genomic integrity is vital to the normal growth of cells and to human health. Maintenance of genome integrity during S phase depends in part on the activity and regulation of the replication enzymes, including replicative helicase called the MCM complex to preserve the genetic information. In contrast, deliberate errors are introduced during maturation of the immune system by modifying enzymes called deaminases. The cells response to different forms of damage is fundamental to its ability to repair itself when challenged by environmental or chemical insults. The goal of this proposal is to investigate mechanisms of genome integrity in cells using a combined genetic, biochemical, and structural approaches. These studies are highly relevant to understanding the development of cancer, including leukemias and breast cancer.

15. SUBJECT TERMS
No subject terms provided

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	90	USAMRMC

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

INTRODUCTION

Genomic DNA stores all the information for living organisms. The faithful duplication the maintanance of DNA are therefore critical for any organisms. Errors in replicating chromosomal DNA can result in genomic instability, leading to various human diseases, such as cancer, immune system disorder, muscle dystrophy, and neurodegenerations. Thus, maintaining genomic integrity is vital to the normal growth of cells and to human health. Maintenance of genome integrity during S phase depends in part on the activity and regulation of the replication enzymes, including replicative helicase called the MCM complex to preserve the genetic information. In contrast, deliberate errors are introduced during maturation of the immune system by modifying enzymes called deaminases. The cells response to different forms of damage is fundamental to its ability to repair itself when challenged by environmental or chemical insults. The major scientific goal of this proposal is to understand mechanisms of maintaning genome integrity in cells by applying a combined genetic, biochemical, and structural methodologies. These studies will contribute to the understanding of the development of cancer, including leukemias and breast cancer.

BODY

Final accomplishments of the approved aims for grant period 07/01/2005-06/30/2008.

Aim 1: *in vitro* function of MCM helicase:

MCM regulator structure/function: While archaeal MCM is functional as a hexamer and double hexamer, the eukaryotic MCM is inactive as a hexemeric complex. We have investigated the in vitro function of an archael MCM function through structure guided mutagenesis and biochemical study, which reveals that the central channel plays a very important role in binding DNA for helicase activity. Our previous crystal structure of *Methanothermobacter thermautotrophicus* MCM (mtMCM) reveals that the dHex has a central channel of 118 Å in length from end to end. The channel has a diameter ranging from 47 Å at the widest point to 23 Å at the narrowest point, as measured from side chain to side chain. There are many positively charged residues located on the inner surface of the channel, generating a strongly positive channel; in contrast, the outside surface is mostly negative. The charge distribution immediately suggests a DNA binding position. For this proposed work, we have mutated six charge residues on the central channel of mtMCM (R116, K117, R122, K178, and R223), all located on the inner central channel surface of the mtMCM. We showed here that mutation of any of these five residues to alanine resulted in lower helicase activity compared with the wt protein. Although all mutants showed ATPase activity comparable to that of wt, all five mutants had consistently reduced ssDNA binding and four out of the five mutants showed reduced dsDNA binding, an observation consistent with the role of these residues in interaction with DNA. This work has been published in ***Biochemistry***, 2008, Aug 29. Epub ahead of print. (see Appendix 1 for details).

With MCM regulation for helicase activity and DNA replication, recent finding reveals that if a complex of GINS (tetrameric complex) binds to MCM in complex with CDC45, the MCM becomes active in unwinding DNA. Thus GINS complex is the "missing" link that regulates and activates MCM helicase activity, and thus DNA replication activity. The GINS complex, which contains the four subunits, Sld5, Psf1, Psf2 and Psf3, is essential for both the initiation and progression of DNA replication in eukaryotes. GINS associates with the MCM2-7 complex and Cdc45 to activate the eukaryotic MCM helicase. It also appears to interact with and stimulate the polymerase activities of DNA polymerase ε and the DNA polymerase α-primase complex. To further understand the functional role of GINS, in collaboration with Dr. Jerard Hurwitz's lab, we determined the crystal structure of the full-length human GINS hetero-tetramer. The four subunits each have a major domain composed of an α-helical bundle-like structure. With the exception of Psf1, other subunits have a small domain containing a three-stranded β-sheet core. Each full-length protein in the crystal has unstructured regions that are all located on the surface of GINS which are probably involved in its interaction with other replication factors. The four subunits contact each other mainly through α-helices to form a ring-like tetramer with a central pore. This pore is partially plugged by a 16-residue peptide from the Psf3 N-terminus which is unique to some eukaryotic Psf3 proteins and is not required for tetramer formation. Removal of this N-terminal 16 residues of Psf3 from the GINS tetramer increased the opening of the pore by 80%, suggesting a mechanism by which accessibility to the pore may be regulated. The structural data presented here indicate that the GINS tetramer is a highly stable complex with multiple flexible surface regions. This work has been published in ***PNAS***, 2007, 104:12685-90. (see Appendix 2).

A model of how the human GINS tetrameric complex interacts and coordinates the activities of its binding partners is proposed in Fig. 6. In this model, we suggest a direct contact between Psf1 and Pol ε because an interaction between *S. cerevisiae* Psf1 and Dpb2 (the second largest subunit of Pol ε) was detected in a yeast two hybrid screen. Because archaeal Gins23 binds to the N terminus of MCM and an interaction of Psf3 with MCM has been detected in the yeast two hybrid system (S. Azuma and H. Masukata, unpublished data, cited in), our model depicts Psf3 in contact with MCM. We assume that

Psf2 contacts the Pol α primase complex based on the report that archaeal Gins23 interacts with primase. The model also allows MCM, Cdc45, and GINS to contact each other based on the isolation of this complex from Drosophila eggs. Finally, we also suggest that GINS must interact with Pol ε and the Pol α-primase complex to coordinate leading and lagging strand synthesis, respectively.

Aim 2: *in vivo* effects of helicase function

1). To examine *in vivo* effects of helicase function and in particular MCM roles in maintaining genome integrity in response to damage.

We performed a screen for *mcm* mutants that are functional for DNA replication, but defective in response to damage. This screen was unexpectedly difficult to implement and generated no new strains. To examine *in vivo* effects of helicase function and in particular MCM roles in maintaining genome integrity in response to damage. We used existing and newly generated mutants in MCM proteins to investigate how MCMs contribution to genome stability during chemical damage. We have initiated a screen for *mcm* mutants that are functional for DNA replication, but defective in response to damage. The screen design is as follows. We mutagenized *in vitro* plasmids containing wild type *mcm2*$^+$, *mcm4*$^+$, and *mcm7*$^+$. These were transformed into strains containing temperature sensitive mutations in *mcm2, mcm4* and *mcm7*, and selected for viability at 36°C, at which temperature the chromosomal MCM is inactivated. Cell viability depends upon the plasmid-borne MCM. This ensures that the plasmid copy is functional for replication. Then, we screened the surviving transformants for sensitivity to hydroxyurea.

Our first screen was conducted under conditions that are lethal for checkpoint mutants such as Δ*rad3*. We did not observe any hydroxyurea-sensitive MCM clones. However, following the old genetics adage "absence of evidence is not evidence of absence", we are re-calibrating our screening conditions to use a concentration of HU where we can see a phenotype in *Δcds1* checkpoint mutants. This strain is more resistant to HU, and our other experiments suggest that the Cds1 kinase may be involved directly in MCM function (see below), which makes it particularly relevant.

Meanwhile, we performed a candidate approach. Recently several new MCM mutations have been identified in the literature, particularly one associated with adenocarcinoma. We constructed this allele into Mcm4 in fission yeast to investigate its effects in our system. We determined that this allele is functional for DNA replication, but causes sensitivity to DNA damage (that is, it fulfills our requirements). We also determined that one of our temperature sensitive alleles, the *mcm4ts-dg* allele, is sensitive to HU and MMS.

Our preliminary studies suggest that the sensitivity for DNA damage may reflect the inability of the checkpoint kinase Cds1 (Chk2) to phosphorylate mutant Mcm4ts-dg protein. This analysis was published in ***Mol. Cell. Biol.*** 2008, 28:1724-38) (See Appendix 3).

2). To determine domains of MCMs responsible for interacting with other proteins.

We identified several molecular associations between MCMs and other factors important for genome stability. We identified an association between Mcm2 and the conserved histone acetyltransferase Mst1 (hTip60)(Gómez et al, 2008 ***Genetics*** in press) (Appendix 4). We also determined that Mcm4 associates the conserved homologous recombination protein Rhp51 (hRad51) and with the checkpoint kinase Cds1 (hChk2) (Bailis et al, 2008 ***Mol. Cell. Biol.*** 28:1724-38) (see Appendix 3). The mutants identified in Fig 1 suggest that at least part of this association is mediated by the Mcm4 amino terminus.

Because the *mst1* mutant causes HU sensitivity, we examined this phenotype more closely and determined that Mst1 interacts genetically with many homologous recombination proteins. Using two-hybrid and co-immunoprecipitation, our data suggest that this may be mediated by interactions with Rad22. We are investigating whether the activity of the Mst1 histone acetyltransferase is required for homologous recombination (Gómez et al, 2008 ***Genetics*** in press). We conclude that Mst1 histone acetyltransferase is required for genome stability during HU arrest, and we are examining how this may reflect MCM association.

Aim 3: Structure-function study of deaminase enzymes

Apolioprotein B (APOB) mRNA-editing enzyme catalytic polypeptide (APOBEC) is a family of cytidine deaminases found exclusively in verterbrates. APOBEC nucleic acid deaminases modify genes by deaminating cytosines in mRNA coding sequences and in ssDNA. Additionally, these enzymes can inhibit the replication of retroviruses, such as the human immunodeficiency virus (HIV) and hepatitis B virus (HBV), and retrotransposons. The APOBEC family is composed of APOBEC-1 (Apo1), Apo2, AID, and APOBEC-3 (3A, 3B, 3C, 3F, 3G, and3H). Apo1, the first member to be characterized, deaminates C^{6666} U in the APOB mRNA thereby creating a premature stop codon, which results in a truncated protein with a different function. Of the APOBEC3 subgroup, APOBEC-3B (A3B), APOBEC-3F (A3F) and APOBEC-3G (A3G) have two cytidine deaminase (CDA) domains and inhibit HIV-1 replication in the absence of the HIV viral infectivity factor (Vif). APOBEC3 proteins also shield the human genome from the deleterious action of endogenous retrotransposons: A3A, A3B, A3C and A3F inhibit LINE 1 and Alu retrotransposition.

Some APOBEC has two homologous cytidine deaminase domains, such as Apo3G, Apo3F, and Apo3B. However, others only have one cytidine deaminase domain, such as Apo2 and AID that are phylogenetically the most ancient members of APOBEC family. AID induces somatic hypermutation (SHM) and class switch recombination (CSR) in activated germinal center B cells. Point mutations in AID are responsible for an immunodeficiency disease, Hyper-IgM-2 (HIGM-2) syndrome, which is characterized by a deficiency in isotype-switching and high affinity antibody formation. Apo2, also known as ARCD-1, is ubiquitously expressed at low levels in both human and mouse and highly expressed in cardiac and skeletal muscle. Pro-inflammatory cytokines, TNF-α and IL-1β, induce Apo2 expression in hepatocytes via the NF-kB response elements identified in the Apo2 promoter region. Apo2 and A3A are the only two APOBEC enzymes that have not yet been reported to deaminate cytosine in ssDNA or RNA *in vitro*. Apo2 is encapsulated into HIV-1 virions when co-expressed with Δvif HIV-1 DNA in 293T cells. However, studies fail to show that Apo2 and nearly all of the single domain APOBEC enzymes inhibit HIV-1 viral replication, except for A3C, which exhibits weak HIV-1 antiviral activity.

A comprehension of the molecular mechanisms of the APOBEC enzymes has been limited by the lack of 3-dimensional structures of APOBEC proteins that use large nucleic acid substrates. We have achieved a major breakthrough through a collaboration of the Chen and Goodman laboratories that resulted in the first high-resolution crystal structure (2.5 Å) for an APOBEC protein, APOBEC-2, reported in Nature, Ref. (Prochnow et al., ***Nature***, 2007, 445, 447-451) (See Appendix 5). The structure is made up of an unusual rod-shaped tetramer that differs significantly from the square tetramer of the free nucleotide C deaminase. Although an enzymatic activity for APOBEC-2 has yet to be identified, its close homology with AID enabled us to use APOBEC-2 as a surrogate to predict how mutations in analogous regions might alter AID activity. This strategy proved to be remarkably effective, where we found that mutations that were predicted to influence dimer and tetramer interface interactions caused significant reductions in AID activity. In an analysis of "APOBEC2 structure – AID function by proxy", we examined 5 HIGM-2 human disease-associated AID mutants that fail to make high-affinity

antibodies. We found that one is mutated at the tetramerization interface, 4 are mutated on an exposed monomer surface and 1 mutation is located proximal to the catalytic active center. These data illustrate the importance of the multimeric interactions in the biochemical function of AID, and, more importantly, in its biological function.

Aim 4: Initiation of Human Hypermutation by the APOBEC Family of Nucleic Acid Deaminases

Almost all mutations are either inconsequential or deleterious. And yet mutations play a central role in evolution, adaptation and fitness from microorganisms to humans. The "trick" is to select and then amplify cells carrying the very few beneficial mutations. Non-patogenic microorganisms, *Escherichia coli* and *Salmonella typhimurium* use error-prone DNA polymerase to enhance cell fitness in highly competitive environments, e.g., limited food resources, whereas their closely related pathogens shut off mismatch repair pathways resulting in "hypermutability". In humans, somatic hympermutation (SHM) of the variable region of immunoglobulin (Ig) genes is responsible for generating high affinity antibodies, and class-switch recombination (CSR), occurring concomitantly with SHM, determines where in the human lymphatic system the high affinity antibodies act to neutralize invading antigens. SHM and CSR occur solely in activated B cells. A different type of hypermutation process, occurring in T cells, is responsible for helping to inactivate HIV-1 AIDS.

We have made substantial progress in dissecting the biochemical basis for the initiation of hypermutation by two members of the APOBEC family of nucleic acid cytidine deaminases, AID (activation-induced cytidine deaminase), and APOBEC3G (A3G) (2). AID is synthesized only in B cells, where tightly regulated synthesis occurs for a period of several hours. AID is required for SHM and CSR of human Ig genes, acting during transcription of V-genes on the non-transcribed strand. A3G is encapsulated along with the genomic RNA in the HIV-1 virion. A3G plays an important role in blocking infection by HIV-1 in non-permissive cells. Both enzymes deaminate cytidine to form uracil (C → U) on ssDNA substrates. Strict regulation of deamination is essential, because unregulated deamination, caused by expression and subsequent action of deaminases in the wrong place at the wrong time, are likely to lead to serious disease, e.g., AID-initiated B-cell lymphomas, by causing excessive levels of mutation (typically a million-fold above normal mutation frequencies).

Our primary goal is to study the biochemical and physical biochemical mechanisms for the APOBEC family of "mutators". The close collaboration of Xiaojiang Chen's laboratory with our laboratory, fostered by the Army Grant support, has enabled us to marry biochemical, physical biochemical analysis with structural analysis (published in ***Nature***, 2007, see Appendix 5). During the grant period two years ago, we obtained the first high-resolution crystal structure for an APOBEC enzyme, APOBEC-2 (Apo2), which is present in heart and skeletal tissue. Based on recent data obtained, where we've successfully cleaved a GST-tag from A3G, we are in an excellent position to obtain crystals of full-length A3G. This result has the potential to initiate breakthrough toward determining structure-function relationships for A3G-catalyzed deamination of HIV-1 and for A3G interaction with the viral infectivity factor (vif), which is responsible the destruction of A3G in T cells. If we can model the A3G-vif interaction *in vitro*, then that would offer clear strategies to protect A3G against apoptotic destruction, which could lead to effective treatment against infection by the AIDS virus.

Biochemical behavior of AID

We have shown previously that AID binds to randomly to ssDNA, irrespective of base composition, and catalyzes deamination of C favoring WRC (W = A/T, R = purine) hot spot motifs and disfavoring SYC (S = G/C, Y = pyrimidine) cold spot motifs. The enzyme scans processively along "naked" ssDNA and tracks processively on the non-transcribed strand of a moving transcription bubble. The deaminations occur in random clusters of between 4 and 10, usually containing at least one WRC triplet, interspersed

with sparsely deaminated regions, i.e., regions with WRC motifs that are "ignored". The question is: what are the biochemical mechanisms? AID is believed to act as a dimer *in vivo* and its found to be phosphorylated at Ser38 in B cells.

Impact of phosphorylation null-mutants on the deamination specificity of AID in relation to human hyper-IgM-2 disease

AID, purified from mouse B-cells has been shown to be phosphorylated on Ser38, Tyr184, and possibly Thr27. Although phosphorylation does not appear to be B-cell specific, AID phosphorylation at Ser38 was suggested to be the modification targeting AID to transcribed dsDNA through interactions with RPA. Since the phosphorylation mutant S38A has differing levels of CSR deficiency in *ex vivo* complementation experiments, we have carried out a comparative biochemical analysis using wild type (WT) AID, mutant S38A and additional phosphorylation defective mutants, to examine the roles of phosphorylation that could account for a loss of AID CSR activity *in vivo*. Also, the mutation of Ser43 to Pro (S43P) was made to characterize the phosphorylation-null HIGM-2 mutation found in a patient.

A key new discovery emerged from the analysis with the phosphorylation-null mutants with respect to human hyper-IgM-2 syndrome, which is characterized by an accumulation of IgM antibody forms owing to an inability to perform CSR. An examination of phosphorylation-null mutant deamination specificities showed the mutants to be much more active in two triplet deamination motifs, a hot spot AGC motif and an intermediate GGC motif, which is published by Pham. et al., ***J. Biol. Chem.*** 283, 17428-17439 (2008) (see Appendix 6). The spectral differences in GGC and AGC for the HIGM-2 AID mutant S43P have the following potentially important human biological consequence. The GGC and AGC motifs occur with exceptionally high frequency in human switch regions. There are 196 GGC motifs in the S-region of IgG4 and 126 in the S-region of IgE, which are by far the most prevalent triplet motifs in these two switch regions, comprising 15% in IgG and 19% in IgE. High levels of C→U deaminations targeted to GGC and AGC motifs in the S-region could result in excessive double-stranded DNA breaks thereby causing aberrant recombination resulting in the elimination of CSR in HIGM-2 patients carrying an S43P mutant AID allele. Here is a "prime" example of how subtle biochemical observations emanating from a simple *in vitro* system can shed important new light on a complex biological disease.

Biochemical behavior of APOBEC3G

Like AID, APOBEC3G (A3G) scans ssDNA processively. We showed, however, that unlike AID, deamination can occur directionally with a strong 3'→5' preference. The observation of directional deamination was entirely unexpected because there is no obvious energy source, such as ATP hydrolysis. During the previous grant period we made significant progress in understanding the biochemical and structural basis for A3G scanning. We showed that the deamination polarity appears to depend on the molecular composition of A3G, which is published by Chelico et al., ***J. Biol. Chem.*** 283, 13780-13791 (see Appendix 7). We found that A3G exists in oligomeric states whose composition is dictated primarily by interactions with DNA, with salt playing an important, yet secondary role. Directional deaminations correlate with presence of dimers, tetramers and larger oligomers of A3G observed by atomic force microscopy, and random deaminations appear to correlate mainly with monomers, which are favored in the absence of salt. Presteady state single-turnover reaction rate measurements revealed a salt-dependent inhibition of C deamination toward the 3'-ssDNA region, offering a molecular basis underlying A3G deamination polarity. Therefore, the data suggest a biochemical basis to account for the observation in human T cells in which HIV-1 infection is restricted by A3G-catalyzed processive C→U deaminations occurring on single-stranded DNA with marked 3'→5' deamination polarity. We suggest that diverse A3G oligomerization occurs *in vivo* and may be responsible for causing HIV-1 proviral DNA mutational bias.

Structure-function analysis of APOBEC enzymes

An initial structural breakthrough was achieved through a collaboration of the Chen and Goodman laboratories that resulted in the first high-resolution crystal structure (2.5 Å) for an APOBEC protein, APOBEC-2, reported in ***Nature***, 2007 (see Appendix 5). We have recently the structural work of APOBEC3-G and structural modeling of APOBEC3. Based on our structural model, we have introduced mutations around the predicted ssDNA substrate groove, and identified residues involved in substrate specificity, ssDNA binding, and deaminase activity. Our results provide an important basis for understanding the underlying mechanisms of substrate specificity for the APOBEC family.

Significance

The humoral response is able to recognize and attack an "unlimited" number of possible infectious agents by radically diversifying the inherited immunoglobulin (Ig) gene. First, the Ig gene is constructed from a variety of inherited gene segments by a Rag-mediated VDJ recombination event which joins together a variable (V), diversity (D) and joining (J) segments to make a functional Ig gene capable of expressing a low affinity antibody. In order to completely clear an infection, B cells undergo affinity maturation to produce antibodies with stronger affinity for antigen and with appropriate effector function. Affinity maturation involves two diversification events, somatic hypermutation and class switch recombination (SHM and CSR). Our principal long-term objective is to focus on SHM and to reconstitute the hypermutational process in an *in vitro* model system. Our initial objective is to explore the action of activation-induced cytidine deaminase (AID), an enzyme required for both SHM and CSR. We have shown that AID-catalyzed deamination of cytosine residues on ssDNA simulates three hallmark properties of SHM – mutational hot and cold spot sequences, broad clonal heterogeneity and transcriptional dependence. Active transcription is also a requirement for SHM, and our current data suggest that AID is able to target C → T mutations preferentially on the non-transcribed strand of a moving transcription bubble. An immediate priority of ours is to develop an *in vitro* human transcription system to investigate how AID is targeted to actively transcribed V- genes but not C genes. A biochemical understanding of AID, whose expression is confined to B cells, and the other APOBEC homologs, whose expression is diverse, should provide important new information on the potential risk factors, which we have referred to as "at risk nucleic acid sequences", which when acted upon inappropriately by APOBEC nucleic acid C deaminases can have serious deleterious consequences resulting from APOBEC-induced genomic instability, leading to cancer and perhaps even to neurodegenerative disease. Progress in understanding the molecular interactions governing SHM, which along with CSR is prerequisite to generating high affinity antibodies, would benefit substantially by the availability of an *in vitro* model system to study AID's role. We intend to provide this invaluable biochemical component.

Future Plans

Although we have progressed towards understanding the biochemical mechanisms responsible for initiating mutations by AID during SHM and by APOBEC3G during HIV-1 inactivation, an in-depth analysis will clearly require high-resolution structural analysis of wild type and mutant APOBEC family enzymes alone and as cocrystals with their nucleic acid substrates. The Army grant has proven invaluable in support of our efforts to meld biochemical and structural analyses, specifically to study how mutations can provide fitness both to microorganisms and higher organisms, and more generally to gain a deeper understanding of the molecular basis of human disease.

KEY RESEARCH ACCOMPLISHMENTS:

- Cloned, expressed, and purified a key MCM regulater, GINS tetramer complex.

- Crystallized and determined the crystal structure of GINS complex and gained a functional insights in knowing how GINS interact with MCM.
- Characterized the functional activity of MCM by mutagenesis based on the structural prediction.
- Demonstrated that MCMs interact with the Cds1 checkpoint kinase (human Chk2, budding yeast Rad53), in which MCM may be the target for the kinase.
- Studied the biochemical function of a deaminase, APOBEC3G.
- Crystallized and determined the atomic structure of a deaminase, APOBEC2.
- Characterized the Structure-function relationship of AID and APOBEC3G

REPORTABLE OUTCOMES (in the form of peer reviewed publications)**:**

1.**Publications:**

Fletcher RJ, Shen J, Holden LG, Chen XS. Identification of Amino Acids Important for the Biochemical Activity of Methanothermobacter thermautotrophicus MCM. ***Biochemistry***. 2008 Aug 29. [Epub ahead of print]. [Appendix 1]

Chang P. Wang, G Bermudez, V. Hurwitz. J, and Chen. X, Structure of GINS complex and the functional insights into DNA replication. PNAS, 104:12685-90 (2007) [Appendix 2]

Bailis JM, Luche DD, Hunter T, Forsburg SL. Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote s-phase genome stability. Mol Cell Biol. 2008 Mar;28(5):1724-38. [Appendix 3]

Go´mez, EB. Nugent, RL. Laria S. and Forsburg SL Schizosaccharomyces pombe Histone Acetyltransferase Mst1 (KAT5) Is an Essential Protein Required for Damage Response and Chromosome Segregation. ***Genetics*** (2008) in press. [Appendix 4]

Prochnaw, C., Bransteitter, R., Klein, M. G., Goodman, M. F., and Chen, X. S. APOBEC2 crystal structure and functional implications for AID. Nature 445, 447-451 (2007) [Appendix 5]

Prochnaw, C., Bransteitter, R., Klein, M. G., Goodman, M. F., and Chen, X. S. APOBEC2 crystal structure Pham, P., Smolka, M. B., Calabrese, P., Landolph, A., Zhang, K., Zhou, H., and Goodman, M. F. Impact of Phosphorylation and Phosphorylation-null Mutants on the Activity and Deamination Specificity of Activation-induced Cytidine Deaminase. J. Biol. Chem. 283, 17428-17439 (2008). [Appendix 6]

Chelico, L., Sacho, E. J., Erie, D. A., and Goodman, M. F. A Model for Oligomeric Regulation of APOBEC3G Cytosine Deaminase-dependent Restriction of HIV. J. Biol. Chem. 283, 13780-13791 (2008). [Appendix 7]

2. Presentation:

- Meeting presentation in Trieste, Italy in 2007, Molecular Biology of Tumor Viruses. Title: Humoral response to viral infection: the structural/functional insights into initial antibody response in CR2 receptor and antibody maturation in APOBEC. By Xiaojiang Chen.
- Meeting presentation in a Keystone meeting 2006, Title: APOBEC enzymes and error prone polymerases, by Myron Goodman

CONCLUSION:

Helicases and their regulators are at the core controlling the faithful DNA replication, and deaminases are important in mutating the particular part of genome for generating high affinity antibodies by radically diversifying the inherited immunoglobulin (Ig) gene. Progress in understanding the molecular interactions governing the assembly, activation/inhibition of helicases and deaminases from are valuable for the understanding of the fundamental basis of human diseases such as leukemia and other types of cancer. Toward this goal, we have made considerable progress in the last two years, as manifested by the high quality publications in this supported direction. In the remaining years, we will based on the solid foundation set in the previous years to further the structures/functional studies of helicases and deaminases *in vitro* system and *in vivo* using yeast genetics.

APPENDICES: (**7 Reprints)**

Appendix 1. Fletcher RJ, Shen J, Holden LG, Chen XS. Identification of Amino Acids Important for the Biochemical Activity of Methanothermobacter thermautotrophicus MCM. ***Biochemistry***. 2008 Aug 29. [Epub ahead of print].

Appendix 2. Chang P. Wang, G Bermudez, V. Hurwitz. J, and Chen. X, Structure of GINS complex and the functional insights into DNA replication. PNAS, 104:12685-90 (2007) [Appendix 2]

Appendix 3. Bailis JM, Luche DD, Hunter T, Forsburg SL. Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote s-phase genome stability. Mol Cell Biol. 2008 Mar;28(5):1724-38. [Appendix 3]

Appendix 4. Go´mez, EB. Nugent, RL. Laria S. and Forsburg SL Schizosaccharomyces pombe Histone Acetyltransferase Mst1 (KAT5) Is an Essential Protein Required for Damage Response and Chromosome Segregation. ***Genetics*** (2008) in press. [Appendix 4]

Appendix 5. Prochnaw, C., Bransteitter, R., Klein, M. G., Goodman, M. F., and Chen, X. S. APOBEC2 crystal structure and functional implications for AID. Nature 445, 447-451 (2007) [Appendix 5]

Appendix 6. Pham, P., Smolka, M. B., Calabrese, P., Landolph, A., Zhang, K., Zhou, H., and Goodman, M. F. Impact of Phosphorylation and Phosphorylation-null Mutants on the Activity and Deamination Specificity of Activation-induced Cytidine Deaminase. J. Biol. Chem. 283, 17428-17439 (2008). [Appendix 6]

Appendix 7. Chelico, L., Sacho, E. J., Erie, D. A., and Goodman, M. F. A Model for Oligomeric Regulation of APOBEC3G Cytosine Deaminase-dependent Restriction of HIV. J. Biol. Chem. 283, 13780-13791 (2008). [Appendix 7]

Biochemistry 2008 Aug 29. [Epub ahead of print]

Identification of Amino Acids Important for the Biochemical Activity of *Methanothermobacter thermautotrophicus* MCM

Ryan J. Fletcher,‡,§ Jingping Shen,‡,| Lauren Holden,§ and Xiaojiang S. Chen*,§

Molecular and Computational Biology, University of Southern California, Los Angeles, California 90089-2910, and
Biochemistry and Molecular Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262

Received January 7, 2008; Revised Manuscript Received May 16, 2008

ABSTRACT *Methanothermobacter thermautotrophicus* minichromosomal maintenance protein (mtMCM) is a 75 kd protein that self-assembles into a double hexamer structure. The double hexamer formed by the N-terminal region of mtMCM has a highly charged (overwhelmingly net positive) inner channel. Here we investigate the effects of point mutations of some of these charged residues on the biological activities of mtMCM. Although all of the mutants were similar to the wild type in protein folding and complex assembly, we found that mutations impaired helicase activity. The study of the DNA binding and ATPase activities of these mutants revealed that the impairment of the helicase activity was highly correlated with a decrease in DNA binding, providing evidence consistent with the role of these charged residues of the inner channel in interactions with DNA.

The progression from G_1 to S phase during the cell cycle is highly regulated. A group of proteins essential for the successful transition and the completion of S phase are the minichromosomal maintenance proteins (MCM). The MCM proteins in eukaryotes are a subgroup of six proteins (MCM2-7) within the much larger AAA+ ATPase family (*1*). MCM2-7 share a highly similar stretch of amino acids referred to as the MCM box (*2-4*) located in the C-terminal half of the protein. The MCM box contains the Walker A and Walker B motifs responsible for the binding and hydrolysis of ATP.

MCM proteins play an important role in the initiation of replication, as these proteins serve as replication helicases for fork unwinding. During G_1 phase, origins with the potential to serve as sites for the initiation of DNA replication are marked by an origin recognition complex (ORC). Subsequent binding by Cdc6 and Cdt1 promotes the recruitment of MCM proteins to form the pre-replication complex (pre-RC). Failure to recruit MCM proteins results in loss of origin firing and G1-phase arrest. Ensuing activity by both Dbf4 dependent kinase (DDK) and cyclin dependent kinase (CDK) promotes the binding of Cdc45 and recruitment of a number of replication proteins to the pre-RC, ultimately leading to the initiation of replication. The presence of all six MCM proteins is necessary for entry and completion of S-phase (reviewed in (*5-7*).

Methanobacterium thermoautotrophicum ΔH has a single MCM (mtMCM) protein that has the ability to self assemble into a double hexamer structure (*8-13*). Purified mtMCM protein contains ATPase, DNA binding, and helicase activities *in vitro*. The crystal structure of the N-terminal portion of double-hexameric mtMCM (N-mtMCM) revealed an overwhelmingly net positively charged inner channel surface (*10*). Here we investigated the functional role of five residues located on the inner channel surface of the double hexameric N-mtMCM structure. Mutational and biochemical studies revealed that

these residues play important roles in helicase function.

EXPERIMENTAL PROCEDURES

Cloning and mutagenesis. The gene for *Methanobacterium thermoautotrophicum ΔH* MCM was cloned into the pGEX-2T expression vector using the restriction enzymes *Xho I* and *Xba I* for expression as a GST-MCM fusion protein. Purified wild-type mtMCM protein was always severely degraded, resulting in a heterogeneous population. We identified the two major cleavage sites for this degradation by N-terminal sequencing and mass spectrometry of the degraded protein fragments (data not shown). These two cleavage sites occurred prior to R275 and R338. Both of these residues are located in the linker region between the N-terminal domains and C-terminal helicase domain. To decrease the degradation and provide a more homologous protein product, we generated a mutant with the two arginines mutated to alanines. The degradation of the mutant was greatly reduced compared with the unmodified wild type protein. At the same time, this mutant protein retained helicase, ATPase, and DNA binding activities at levels similar to those of to the wt protein (data not shown). All the channel surface mutants reported in this study were constructed to also contain these two mutations to eliminate differences in degradation between different constructs. In this study, the construct containing the two arginine to alanine mutations is referred to as wild type (wt). Five charged residues on the channel surface, R116, K117, R122, K178, R223 (Figure 1A, 1B), were all mutated to alanine through standard site-directed mutagenesis methods. The sequences of all clones were verified by sequencing the DNA region encoding the entire MCM protein.

Protein purification. All proteins were grown in *E. coli* cells at 37˚C until OD_{600}=0.2 and then the temperature was reduced to 24˚C. Isopropyl-β-D-thiogalactoside (IPTG) was added to 0.2 mM when OD_{600}= 0.5. The cells were shaken at 24˚C overnight. Cell pellets were re-suspended in purification buffer (50 mM Tris, pH 8.0, 250 mM NaCl, 10 mM DTT) and sonicated on ice with protease inhibitors pepstatin, leupeptin, and phenylmethylsulfonyl fluoride (PMSF). The GST-MCM fusion protein in the supernatant was separated from other proteins using glutathione resin (Amersham) at 4˚C. After thrombin cleavage of the GST-fusion protein, the free MCM protein was further purified using Superose-6 gel filtration column chromatography (16/60 cm). Peak fractions were collected and purity was analyzed by SDS-PAGE. Protein was concentrated using Amicon Ultra (Millipore) centrifuge devices and quantified using the Bio-Rad protein assay and Coomassie blue staining of an SDS-PAGE gel.

Heat stability of mutants. A protein samples (10 μg) were heated at 55˚C in 15 μL of helicase buffer for 30 minutes. The heat-treated protein was analyzed by Superose-6 column chromatography and on a 10% SDS-PAGE gel.

Helicase assay. A 60 nucleotide primer $^{5'}$TTT TTT TTT TTT TTT TTT TTT TTT CGC GCG GGG AGA GGC GGT TTG CGT ATT GGG CGC C$^{3'}$ (Operon) was purified and radio-labeled as previously described for DNA binding assays. The labeled primer was mixed in a 2:1 ratio with M13mp18 circular single-stranded DNA in 300 mM NaCl and 20 mM Tris (pH 8.0), boiled in 900 μL water for 3 minutes and allowed to cool to room temperature over 5 hours. This helicase substrate containing 34 complementary bases and a 26 dT non-complimentary overhang was purified from excess primer using 700 μL Sephacryl 300HR resin in a 2-mL glass column (Bio-Rad). Fractions containing helicase substrate were pooled and stored at 4˚C. Reactions of 20 μL containing 20 mM Tris (pH 7.8), 10 mM $MgCl_2$, 1 mM DTT, 5 mM ATP, 0.1 mg/mL BSA, helicase substrate, and MCM protein were assembled on ice. The reactions were then incubated at 50˚C for 30 minutes, 5x stop solution (10 mM EDTA, 0.5% SDS, 0.1% xylene cyanol, 0.1% bromophenol blue, and 50% glycerol) was added and the sample was placed on ice. Reactions were analyzed on a 12% polyacrylamide gel in 1X TBE run on ice for 50 min at 150 volts. The gel was dried and

then exposed to film and a phosphorimaging plate for quantification.

DNA binding assay. All primers (Operon) were resuspended in 10 mM NaOH and purified by FPLC using a Mono-Q 10/100 column (Amersham). Purified DNA was concentrated and quantified using OD_{260}. For the double-stranded DNA (dsDNA) binding assay, complementary primers were mixed, boiled for 10 min in 900 μL of water, and allowed to cool to room temperature overnight. The annealed dsDNA was then purified by FPLC over a Superdex-75 16/60 (Amersham) column in 50 mM NaCl. The DNA used for the single-strand binding was $^{5'}$AAA GCG CTG ACC TAT CGC GTA TAG CTC GAG GA$^{3'}$ and the complementary oligonucleotides used for the double strand binding assay were $^{5'}$GCG CTG ACC TAT CGA CCT ATA CGG TTA GCC$^{3'}$ and $^{5'}$GGC TAA CCG TAT AGG TCG ATA GGT CAG CGC$^{3'}$. The primers were labeled using T4 polynucleotide kinase (NEB) and γ-^{32}P-ATP (Amersham, 3000 Ci/mmol). The labeled primer was purified using Microspin G-25 columns (Amersham) and stored at 4˚C. Reactions of 20 μL were assembled at room temperature with 50 mM NaCl, 5 mM $MgCl_2$, 50 mM Tris (pH 7.8), 1 mM DTT, 0.1 mg/mL BSA, mtMCM protein (between 0.5-1.0 μM), and 15 nM labeled DNA. Reactions were incubated at 25˚C for 30 min and then 5 μL loading buffer (50% glycerol, 0.1% xylene cyanol, and 0.1% bromophenol blue) was added. Reactions were analyzed by gel electrophoresis and by filter binding assay. For gel analysis, the reactions were loaded onto a 2% agarose gel and electrophoresed for 1 hour at 80 volts in 1X TBE buffer. Gels were dried and exposed to film and phosphorimaging plates for quantification. For filter binding assay, 10 ul reaction mixtures were applied to nitrocellulose filter (Millipore, HA 0.45 mm). The filter was washed with 15 ml reaction buffer. The radioactivity retained on the filter was quantified by liquid scintillation counting.

ATPase assay. Reactions (10 μL) containing 50 mM NaCl, 5 mM $MgCl_2$, 50 mM Tris (pH 7.8), 1 mM DTT, 0.1 mg/mL BSA, γ-^{32}P-ATP (Amersham 3000 Ci/mmol), 15 μM cold ATP, and various amounts of MCM proteins were assembled on ice. Reactions were incubated at 50˚C for 30 minutes, plunged into ice water and 10 mM EDTA was added. Aliquots (5 μL) from each reaction were placed onto a pre-washed PEI cellulose TLC plate (EMD Chemicals INC) and run for two hours in 2 M acetic acid and 0.5 M LiCl. Plates were dried and exposed to film and phosphorimaging plates for quantification.

RESULTS

Location of mutated residues on the mtMCM double-hexamer structure. The crystal structure of the N-terminal region of mtMCM (N-mtMCM) shows that the five residues (R116, K117, R122, K178, R223) studied here are located close to the central channel surface (Fig. 1A, 1B). K117, R122, and K178 are highly conserved among eukaryotic MCMs (Figure 1C). R223 is not strictly conserved, but there appear to be few positively charged residues in the vicinity of R223 among the aligned MCM sequences and R116 appears to be unique to mtMCM (Figure 1C). R116, K117, and R122 are located on the same β-strand and K178 is on the adjacent anti-parallel β-strand (Figure 1A, 1B). R223 is located on the previously characterized DNA binding β-hairpin finger (*10*) and is close to the cluster of the other mutated amino acids.

Mutant protein assembly and stability. We first performed experiments to test the structural integrity of the mutant mtMCM proteins in order to rule out the possibility of disruption of protein folding or assembly by the mutations. All mutant proteins and the wild-type (wt) protein were analyzed by FPLC using an analytical Superose-6 gel filtration column before and after the proteins were heat-treated at 55 ˚C. With or without heat treatment, all mutants had similar chromatographic profiles as the wt, characterized by a peak (P2) with an apparent molecular weight of a double hexamer of mtMCM (Figure 2A). SDS-PAGE of fractions collected across this peak revealed a

predominant band consistent with the size of the full-length protein (Figure 2B). N-terminal sequencing showed that other minor bands were degradation products of the full-length protein that could not be completely eliminated even after extensive purification. As all mutants and wt showed similar extents of degradation, the degradation should not interfere with the comparison and interpretation of biochemical results. The chromatography analysis indicated that none of the mutations affected mtMCMs folding, oligomerization state, or heat stability in solution.

Helicase assay. A helicase assay showed that the mutants had various degrees of reduced helicase activity when compared with the wt protein (Figure 3A, 3B). Over the tested protein concentration range, helicase activity of R223A was 2-4 fold lower than the activity of the wt protein and the helicase activities of R116A, K117A, R122A, and K178A were 3-7 fold lower that that of the wt (Figure 3A, 3B). This suggests that these charged residues are important for the helicase activity of mtMCM.

DNA binding. The ability of wt and mutants to bind single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) was determined by native gel shift using a radio-labeled oligonucleotide and four different concentrations for each protein. All mutants showed reduced ssDNA binding compared to wt (Figure 4A). At the protein concentration of 1 μM, the ssDNA binding activity for K117A and K178A was approximately 4 fold lower than that for the wt and for the other three mutants (R116A, R122A, R223A), it was roughly 2 fold lower than wt. dsDNA binding assay results were quite similar to those of the ssDNA binding analysis (Figure 4B). Four out of the five mutants (K117A, R122A, K178A, R223A) showed lower dsDNA binding (approximately 2-4 fold) than the wt (Figure 4B). However, R116A bound to dsDNA at a level similar to the wt protein.

ATPase activity. The effect of each mutation on the ability of mtMCM to hydrolyze ATP was measured using standard ATPase assays. The mutations (R116A, K117A, R122A, K178A, R223A) had little effect on the level of ATP hydrolysis when compared to wt (Figure 5A, 5B), suggesting that none of the residues are directly involved in ATP binding or hydrolysis.

DISCUSSION

The charged residues under investigation (R116, K117, R122, K178, and R223) are all located on the inner central channel surface of the N-mtMCM, a location potentially important for interacting with DNA, which binds within the central channel, and for helicase function. We showed here that mutation of any of these five residues to alanine resulted in lower helicase activity compared with the wt protein. Although all mutants showed ATPase activity comparable to that of wt, all five mutants had consistently reduced ssDNA binding and four out of the five mutants showed reduced dsDNA binding, an observation consistent with the role of these residues in interaction with DNA.

Even though the positively charged residues mutated in this study are all located on the inner channel of the double hexamer, the five residues have different degrees of exposure to the channel surface. For example, R122, R116, and K117 are located on the side of the monomer that faces the next monomer within the ring (Figure 1A, 1B), forming the edge around the wide-open side of the channel. These three residues are not as exposed to the channel surface as R122 and K178. In order to directly contact DNA, a modest conformational change would be required to expose these residues more to the inner channel surface. Alternatively, DNA may interact with these residues by binding or passing the side channel, as suggested in the looping model of SV40 LTag (*14-16*).

Furthermore, the location of the residues near the interface between two adjacent subunits (R122, R116, K117) may influence the interactions between two neighboring monomers during conformational changes necessary for helicase function. Of the residues evaluated, R122 has been shown to have direct contact with

the next neighbor through the β-hairpin (Figure 1B). Indeed, of the mutants analyzed, R122A, R116A, and K117A showed the most significant reduction in ssDNA and dsDNA binding (Figure 4) as well as in helicase function (Figure 3). R116, K117, and R122 are conserved among most *S. cerevisiae* and human MCMs (Figure 1C), suggesting the importance of these residues for MCM function.

Based on the N-mtMCM structure, R223 sits on the previously characterized DNA binding β-hairpin structure that has been shown to be important for helicase function (*10, 17*). Previously, double or triple mutations of three other positively charged residues on this β-hairpin, R226, R229, and K231, were shown to completely disrupt the DNA binding and helicase function. Intriguingly, point mutation of R223 to alanine only partially disrupted MCM DNA binding and helicase activity (Figure 3, Figure 4), suggesting that the multiple positively charged residues on the β-hairpin may be partially complementary to each other.

In summary, this mutational study clearly demonstrated that the positively charged residues on the inner central channel, R116, K117, R122, K178 and R223, are critical for the functional activities of mtMCM. Mutation of any of these residues resulted in decreased helicase activity, which was strongly associated with reduced binding to both ssDNA and dsDNA. Even though these residues are located outside of the highly conserved MCM box (helicase domain), the critical residues for classifying MCM proteins, four of the five residues are highly conserved in eukaryotes, underscoring the functional significance of these charged residues of the hexameric channel surface of the N-terminal region of MCM.

Figure 1: Location of the five positively charged residues on the inner channel surface of N-mtMCM. A: N-mtMCM hexamer structure showing the five mutated residues, R116 (yellow), K117 (pink), K178 (green), R122 (blue), and R223 (cyan). **B:** A close-up of the N-mtMCM view shown in panel A. R223 is on one side of the β-hairpin, R122 is adjacent to the β-hairpin of another monomer. K117 is also on one side of a monomer, around the side channel formed between two neighboring subunits. **C:** Sequence alignments of mtMCM (methanoMCM) and MCM2-7 from *S. cerevisiae* (scMCM) and humans (humanMCM). Amino acids mutated in this study are marked by asterisks. Red boxes highlight aligned residues that are mutated in this study.

Figure 2: Gel filtration analysis of wt and mutant mtMCM proteins. A: FPLC overlay of OD_{280} curves from wt (blue) and the five mutant proteins run on a preparative grade Superose 6 column. The profiles of all proteins overlapped. Peak 1 (P1) is in the void volume, P2 contains protein that has an apparent molecular weight (Mwt) of approximately 960 kD, peak 3 (P3) contains degraded protein that cannot oligomerize. The peak P2 has an apparent molecular weight consistent with that of a double hexamer. B: SDS-PAGE analysis of the peak fraction of P2 for the wt and each mutant protein. Lanes 1-6 are wt, R116A, K117A, K178A, R122A, and R223A, respectively. An asterisk marks the position of the full-length protein. The minor bands below the full-length protein band were the result from the small amount of degradation of the full-length protein.

Figure 3: Helicase activity of wt and mutant proteins. A: Gel analysis of helicase assay. Increasing concentrations of wt (lane 3-6), R223A (lane 7-10), R116A (lanes 11-14), K117A (lanes 15-18), R122A (lanes 19-22), and K178A (lanes 23-26) were incubated with radio-labeled helicase substrate at 50˚C for 30 minutes and samples were then analyzed by 12% polyacrylamide gel electrophoresis. Boiled and no protein controls are in lanes 1 and 2, respectively. **B:** Quantitation of unwinding in panel A, expressed as percentage of ssDNA released from the dsDNA substrate. Protein concentrations were calculated based on double hexamer molecular weights. Error bars (s.e.m.) are from a minimum of three independent experiments.

Figure 4: DNA binding of wt and mutants. A: Quantitation of ssDNA binding of wt and mutants at the indicated protein concentrations. **B:** Quantitation of dsDNA binding of wt and mutants at the indicated protein concentrations. Protein concentrations were calculated based on double hexamer molecular weights. Error bars (s.e.m.) are from a minimum of three independent experiments.

Figure 5: ATPase activity of the wt and mutants. A: TLC analysis of ATPase assay. Four protein concentrations, 16, 32, 64, and 128 nM, were used for wt (lanes 2-5 and 19-22), R116A (lanes 6-9), K117A (lanes 10-13), R122A (lane 14-17), K178A (lanes 23-26), and R223A (lanes 27-28). No protein control is in lanes 1 and 18. **B:** TLC analysis of ATPase assay of the mutants in the presence of dsDNA (10 ng/ul 1.0 kb PCR DNA product), showing that all mutants responded to dsDNA stimulation of ATPase activity similarly to the wt. **C:** Quantitation of results from panel A**.** Error bars (s.e.m.) are from a minimum of three independent experiments. Protein concentrations were calculated based on double hexamer molecular weights.

References:

1. Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) *AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes*, *Genome Res 9*, 27-43.
2. Tye, B. K. (1999) *Minichromosome maintenance as a genetic assay for defects in DNA replication*, *Methods 18*, 329-34.
3. Lei, M., and Tye, B. K. (2001) *Initiating DNA synthesis: from recruiting to activating the MCM complex*, *J Cell Sci 114*, 1447-54.
4. Kelman, Z., and White, M. F. (2005) *Archaeal DNA replication and repair*, *Curr Opin Microbiol. 8*, 669-76. Epub 2005 Oct 20.
5. Tye, B. K. (1999) *MCM proteins in DNA replication*, *Annu Rev Biochem 68*, 649-86.
6. Bell, S. P., and Dutta, A. (2002) *DNA replication in eukaryotic cells*, *Annu Rev Biochem 71*, 333-74. Epub 2001 Nov 9.
7. Forsburg, S. L. (2004) *Eukaryotic MCM proteins: beyond replication initiation*, *Microbiol Mol Biol Rev 68*, 109-31, table of contents.
8. Shechter, D., Ying, C., and Gautier, J. (2000) *The intrinsic DNA helicase activity of M. thermoautotrophicum delta H minichromosome maintenance protein*, *J Biol Chem 275*, 15049-59.
9. Chong, J., Hayashi, M., Simon, M., Xu, R., and Stillman, B. (2000) *A double-hexamer archaeal mini-chromosome maintenance protein is an ATP-dependent DNA helicase*, *Proc Natl Acad Sci U S A 97*, 1530-5.
10. Fletcher, R. J., Bishop, B. E., Leon, R. P., Sclafani, R. A., Ogata, C. M., and Chen, X. S. (2003) *The structure and function of MCM from archaeal M. Thermoautotrophicum*, *Nat Struct Biol. 10*, 160-7.
11. Kelman, Z., Lee, J., and Hurwitz, J. (1999) *The single minichromosome maintenance protein of M. thermoautotrophicum DeltaH contains DNA helicase activity*, *Proc Natl Acad Sci U S A 96*, 14783-8.
12. Fletcher, R. J., Shen, J., Gomez-Llorente, Y., Martin, C. S., Carazo, J. M., and Chen, X. S. (2005) *Double hexamer disruption and biochemical activities of Methanobacterium thermoautotrophicum MCM*, *J Biol Chem. 280*, 42405-10. Epub 2005 Oct 11.
13. Kasiviswanathan, R., Shin, J. H., Melamud, E., and Kelman, Z. (2004) *Biochemical characterization of the Methanothermobacter thermautotrophicus minichromosome maintenance (MCM) helicase N-terminal domains*, *J Biol Chem. 279*, 28358-66. Epub 2004 Apr 20.

14. Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J. A., Fanning, E., Jochimiak, A., Szakonyi, G., and Chen, X. S. (2003) *Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen*, *Nature. 423*, 512-8.
15. Gai, D., Zhao, R., Li, D., Finkielstein, C., and Chen, X. (2004) *Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen*, *Cell 119*, 47-60.
16. Sclafani, R., Fletcher, R., and Chen, X. (2004) *Two heads are better than one: regulation of DNA replication by hexameric helicases*, *Genes Dev 18*, 2039-45.
17. McGeoch, A. T., Trakselis, M. A., Laskey, R. A., and Bell, S. D. (2005) *Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism*, *Nat Struct Mol Biol 14*, 14.

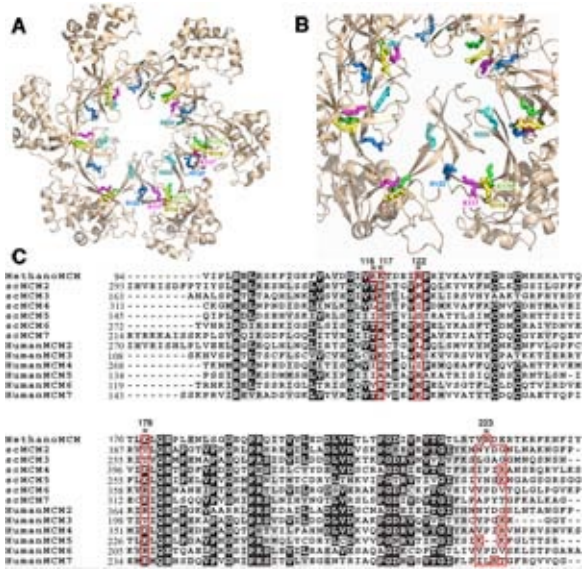
A
B
C

Fig.1

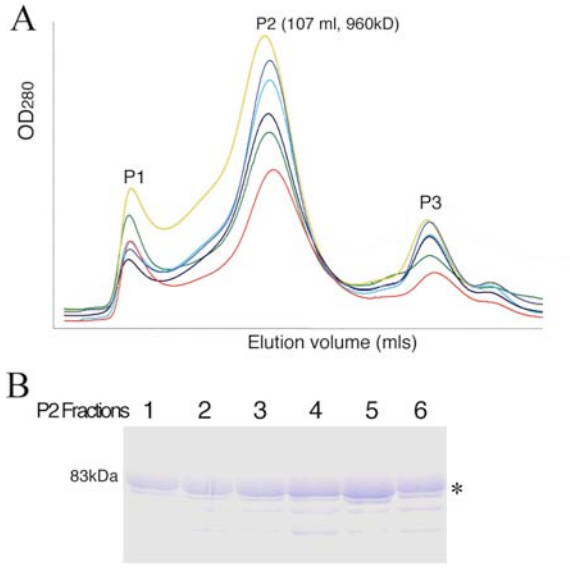
A
P2 (107 ml, 960kD)
OD280
P1
P3
Elution volume (mls)
B
P2 Fractions 1 2 3 4 5 6
83kDa
*

Fig.2

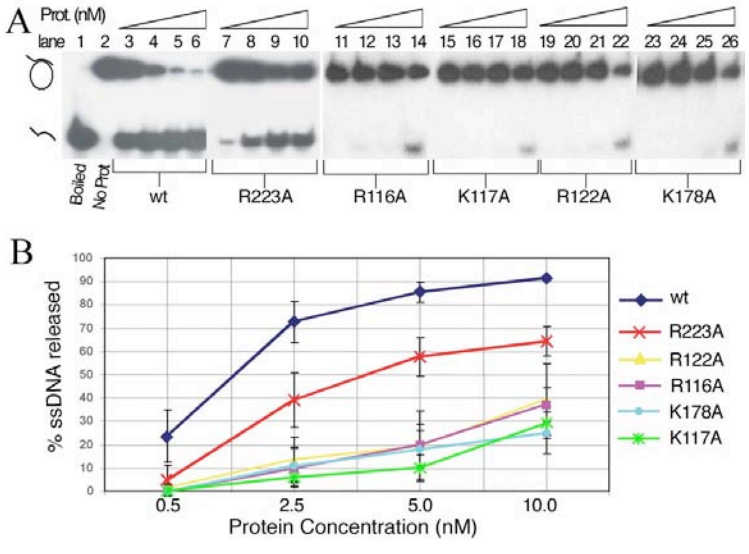
A
Prot (nM)
lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Boiled
No Prot
wt
R223A
R116A
K117A
R122A
K178A
B
% ssDNA released
Protein Concentration (nM)
0.5
2.5
5.0
10.0
wt
R223A
R122A
R116A
K178A
K117A

Fig.3

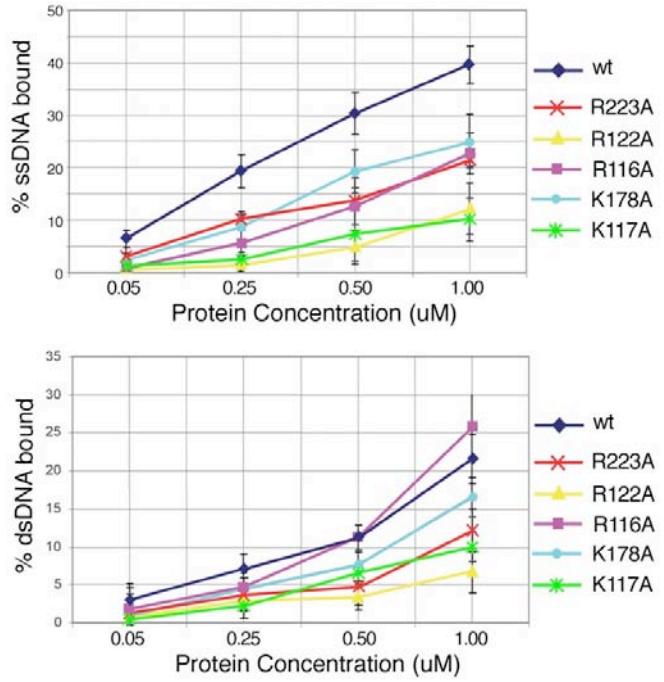
% ssDNA bound
Protein Concentration (uM)
0.05
0.25
0.50
1.00
wt
R223A
R122A
R116A
K178A
K117A
% dsDNA bound
Protein Concentration (uM)
0.05
0.25
0.50
1.00
wt
R223A
R122A
R116A
K178A
K117A

Fig.4

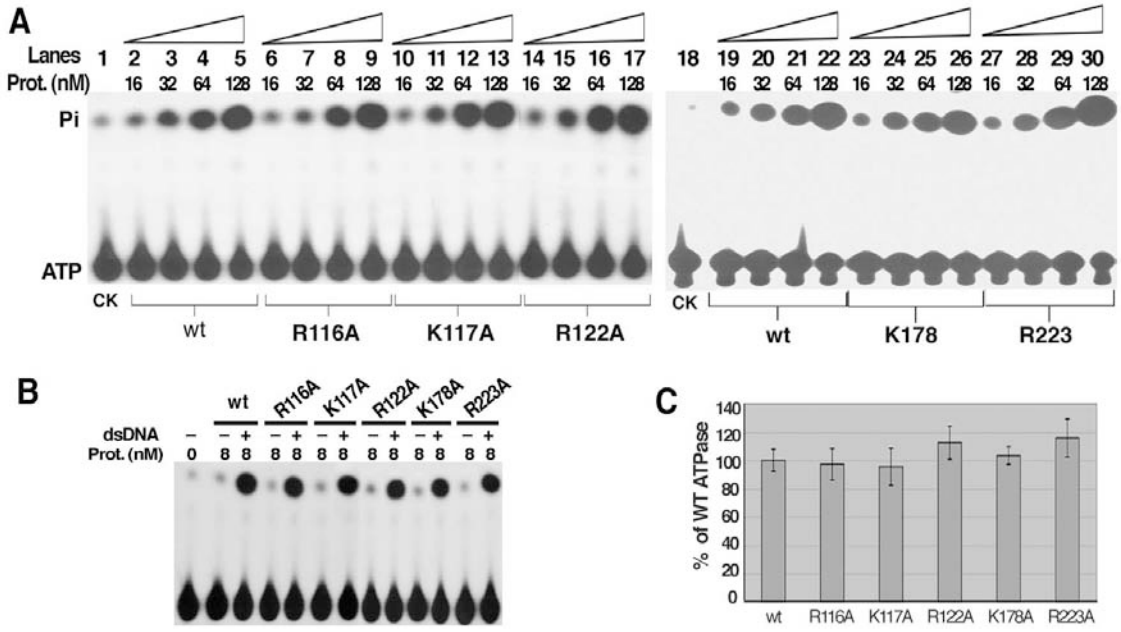
A
Lanes 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Prot. (nM) 16 32 64 128 16 32 64 128 16 32 64 128 16 32 64 128 16 32 64 128 16 32 64 128 16 32 64 128
Pi
ATP
CK
wt
R116A
K117A
R122A
CK
wt
K178
R223
B
wt
R116A
K117A
R122A
K178A
R223A
dsDNA – – + – + – + – + – + – +
Prot. (nM) 0 8 8 8 8 8 8 8 8 8 8 8 8
C
% of WT ATPase
140
120
100
80
60
40
20
0
wt
R116A
K117A
R122A
K178A
R223A

Fig5

Crystal structure of the GINS complex and functional insights into its role in DNA replication

Y. Paul Chang*†, Ganggang Wang†, Vladimir Bermudez‡, Jerard Hurwitz‡§, and Xiaojiang S. Chen*†§

*Graduate Program in Genetic, Molecular, and Cell Biology, and †Section of Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089; and ‡Molecular Biology Program, Memorial Sloan–Kettering Cancer Center, 1275 York Avenue, Box 97, New York, NY 10021

Contributed by Jerard Hurwitz, June 13, 2007 (sent for review May 21, 2007)

The GINS complex, which contains the four subunits Sld5, Psf1, Psf2, and Psf3, is essential for both the initiation and progression of DNA replication in eukaryotes. GINS associates with the MCM2-7 complex and Cdc45 to activate the eukaryotic minichromosome maintenance helicase. It also appears to interact with and stimulate the polymerase activities of DNA polymerase ε and the DNA polymerase α-primase complex. To further understand the functional role of GINS, we determined the crystal structure of the full-length human GINS heterotetramer. Each of the four subunits has a major domain composed of an α-helical bundle-like structure. With the exception of Psf1, each of the other subunits has a small domain containing a three-stranded β-sheet core. Each full-length protein in the crystal has unstructured regions that are all located on the surface of GINS and are probably involved in its interaction with other replication factors. The four subunits contact each other mainly through α-helices to form a ring-like tetramer with a central pore. This pore is partially plugged by a 16-residue peptide from the Psf3 N terminus, which is unique to some eukaryotic Psf3 proteins and is not required for tetramer formation. Removal of these N-terminal 16 residues of Psf3 from the GINS tetramer increases the opening of the pore by 80%, suggesting a mechanism by which accessibility to the pore may be regulated. The structural data presented here indicate that the GINS tetramer is a highly stable complex with multiple flexible surface regions.

Cdc45 | DNA helicase | minichromosome maintenance complex | DNA polymerase

Eukaryotic DNA replication is controlled by a series of ordered and regulated steps (1–3) that commence with the binding of the six-subunit origin recognition complex (ORC) to replication origins. During the G1 phase of the cell cycle, Cdc6, and Cdt1 are recruited to the origin, and together with ORC, support the loading of the heterohexameric MCM2-7 complex (minichromosome maintenance, MCM) to form the prereplication complex (pre-RC). Although a substantial amount of data suggest that MCM acts as the replicative helicase, MCM present in the pre-RC (as well as isolated MCM) is devoid of helicase activity (summarized in ref. 4). At the G_1/S transition of the cell cycle, it appears that the MCM helicase activity is activated by a complex and an as yet poorly understood series of modifications that require the action of two protein kinases, DDK (Cdc7-Dbf4) and CDK (cyclin-dependent), as well as the participation of at least eight additional factors, including Mcm10, Cdc45, Dpb11, GINS, synthetic lethal with *dpb11* mutant-2 (Sld2), and Sld3 (4). Two of these components, Cdc45 and GINS, appear critical for helicase activation because DNA unwinding is observed (3, 5), concomitant with their loading at origins. In accord with these findings, a complex containing near-stoichiometric levels of MCM, Cdc45, and GINS was isolated from *Drosophila* and shown to possess DNA helicase activity (6). Studies with *Xenopus* extracts revealed that a complex that included MCM, Cdc45, and GINS was found at sites at which replication forks were halted artificially by a streptavidin–biotin complex (7). In *Saccharomyces cerevisiae*, GINS was shown to play a critical role in supporting interactions between MCM and Cdc45, as well as a number of key regulatory proteins. They together formed a large replisome progression complex that migrated with the replication fork. Upon selective degradation of the Psf2 subunit of GINS, replication was halted and Cdc45 was no longer associated with MCM. These findings suggest that interactions between MCM and other key replication factors might be mediated by GINS. Collectively, they indicate that GINS is an essential component of the replicative machinery that moves with the replication fork.

GINS is a heterotetrameric complex consisting of Sld5, Psf1 (partner of Sld5-1), Psf2, and Psf3 and was first discovered by using a variety of genetic screens in *S. cerevisiae* (8). The four GINS subunits are paralogs, among which the specific subunit pairs Psf1–Sld5 and Psf2–Psf3 are more closely related (9). Each of the subunits is relatively small ($\approx$200 aa) and highly conserved in all eukaryotes. In archaea, only two homologues, Gins15 and Gins23, have been identified that appear to interact and form a dimer of the heterodimer, suggesting that, like its eukaryotic counterpart, it is a tetramer (10). Direct interactions between the archaeal GINS complex and the archaeal MCM, as well as primase, have been reported (10). Other reports have recently appeared suggesting that GINS may serve as an accessory factor for eukaryotic DNA polymerases, including DNA polymerase (Pol) ε (11) and the DNA Pol α-primase complex (12).

Despite the essential role of GINS in DNA replication, how GINS interacts with MCM, Cdc45, and other protein factors at the replication fork remains unclear. To understand the structural/functional roles of GINS in replication, we crystallized the human GINS complex and determined its crystal structure. This complex included the full-length proteins of each of the four subunits. During the preparation of this work, the structure of the human GINS complex containing a truncated Psf1 subunit appeared (13). The tetramer structure that we have obtained is basically the same as that reported by Kamada *et al.* (13); the folds of each subunit and the interactions between the four subunits are essentially the same with slight variations found for certain loops and β-strands. However, our crystal structure revealed certain features not reported for the structure of the truncated complex reported by Kamada *et al.* (13) that may have important functional implications. The structural and muta-

Author contributions: Y.P.C., J.H., and X.S.C. designed research; Y.P.C., G.W., and V.B. performed research; Y.P.C., G.W., and X.S.C. analyzed data; and Y.P.C., G.W., V.B., J.H., and X.S.C. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: MCM, minichromosome maintenance; ts, temperature-sensitive; Pol, polymerase.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2Q9Q).

§To whom correspondence may be addressed. E-mail: hurwitz@ski.mskcc.org or xiaojiang.chen@usc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0705558104/DC1.

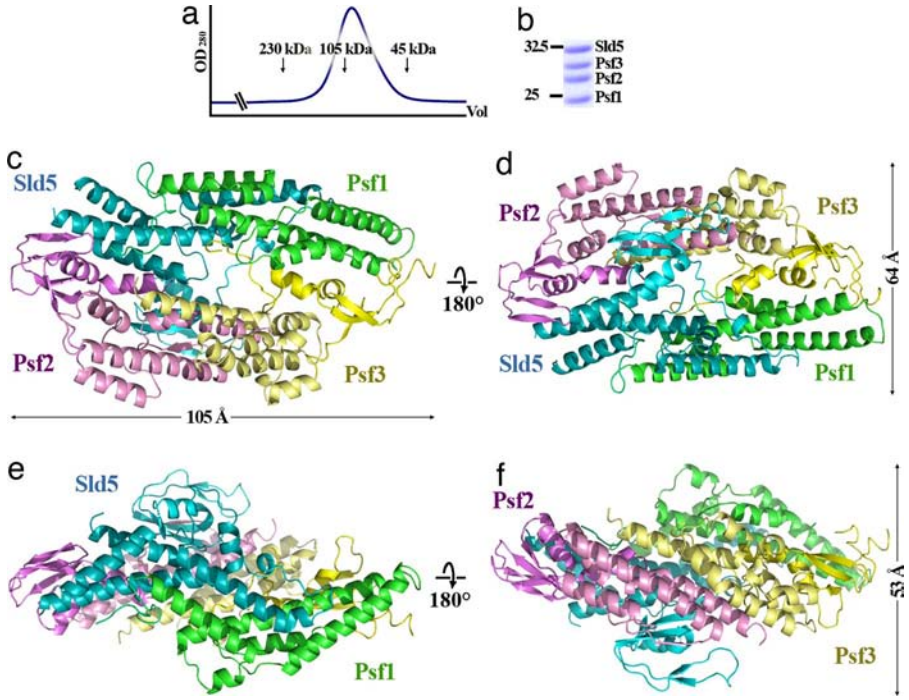

Fig. 1. The overall structural features of human GINS complex. (*a*) Superdex-200 gel-filtration profile of the human GINS complex crystallized. (*b*) The SDS/PAGE analysis of the peak fraction from *a*, showing that all four subunits are present as full-length proteins. (*c* and *d*) Two views of the wider faces of the GINS complex structure, showing the spindle-shaped structure and part of the central pore opening. The tetramer is composed of predominantly α-helices that are arranged in α-helical bundles. (*e* and *f*) Two views of the narrower sides of the GINS complex. The four subunits are shown in different colors as indicated.

tional data we obtained suggest that the dimension of a central pore in GINS appears to be regulated by a short N-terminal peptide of Psf3. The positions of disordered regions in our structure, including the C-terminal 51 residues of Psf1, colocalize on the surface of the GINS complex as patches and likely serve as interaction sites for the binding of GINS to its replication protein partners.

Results

Overall Structural Features of the GINS Complex. The four full-length subunits of GINS were coexpressed in *Escherichia coli* and the complex purified to homogeneity. The isolated complex had an apparent molecular mass of ≈90 kDa as estimated from gel filtration chromatography (Fig. 1*a*) and glycerol gradient sedimentation (data not presented) consistent with a 1:1:1:1 molar ratio of the four different proteins in the tetrameric complex (Fig. 1*b*). We crystallized the GINS complex as described in *Materials and Methods*, and SDS/PAGE and mass spectrometric analyses confirmed that all four proteins present in the crystals were full-length. We determined the x-ray structure of the GINS complex to 2.36 Å resolution [statistics presented in supporting information (SI) Table 1]. Each asymmetric unit in a crystal cell contained two GINS heterotetramers with identical conformations. The overall morphology of the GINS tetramer complex resembles a slightly elongated spindle (Fig. 1 *c–f*) with a visible central hole (Fig. 1 *c* and *d*). The body of the tetramer is composed of α-helices with few peripheral short β-strands. The gross structural features are essentially the same as those recently reported for the structure containing a truncated Psf1 subunit (13); in the reported structure, 47 residues of the C-terminal region of Psf1 had to be deleted for crystallization. Surprisingly, even though we crystallized GINS with all full-length proteins, only the first 145 residues of Psf1 were ordered. These residues were present in the structure containing truncated Psf1 reported by Kamada *et al.* (13). The C-terminal 51 residues of Psf1 are not visible in our structure, indicating that this region is intrinsically disordered.

Structures of Individual GINS Subunits. Each of the four subunits contains a major domain composed of α-helices (α-domains, Fig. 2 *a–d*). The folds of the α-domains of all four subunits are similar; each contains four to five helices arranged more or less in a parallel fashion to form a partial three-helix bundle structure. In three of the four subunits (Sld5, Psf2, and Psf3), there is a small β-sheet composed of three antiparallel β-strands (β1, β2, and β3) near one end of the α-domain. Around the β-sheet are two helices and a β-hairpin or loop, forming a small but definable β-domain in these three subunits. The monomeric structures overlap well with the reported GINS structure (13), especially within the α-domains. However, conformational differences are noted in the β-domains of each subunit (Fig. 2 *e–g*). The differences in the β-domains and in a few loops are all located on the surface of the GINS complex, suggesting a certain degree of plasticity in the surface structures.

Psf1 has only an α-domain (residues 1–145), whereas all of its C-terminal 51 residues are disordered. There is substantial space to accommodate the C-terminal 51 residues of Psf1 in at least one of the two GINS tetrameric complexes in the asymmetric unit. Nonetheless, no definable electron density can be seen for the C-terminal 51 residues of Psf1 (residues 146–196), despite the fact that sequence alignment suggests a fold similar to the β-domain that is present in the other three subunits (SI Fig. 7*a*). Kamada *et al.* (13) reported that their GINS complex crystallized only when a Psf1 mutant lacking the C-terminal 47 residues (14) was used, suggesting that the presence of this β-domain inhibited crystal packing of the GINS complex. Kamada *et al.* (13) proposed that the deleted region of Psf1 folds into a β-domain structure and that the correct positioning of this domain on the surface of the GINS complex is critical for function. However, we crystallized the GINS complex with all full-length proteins under

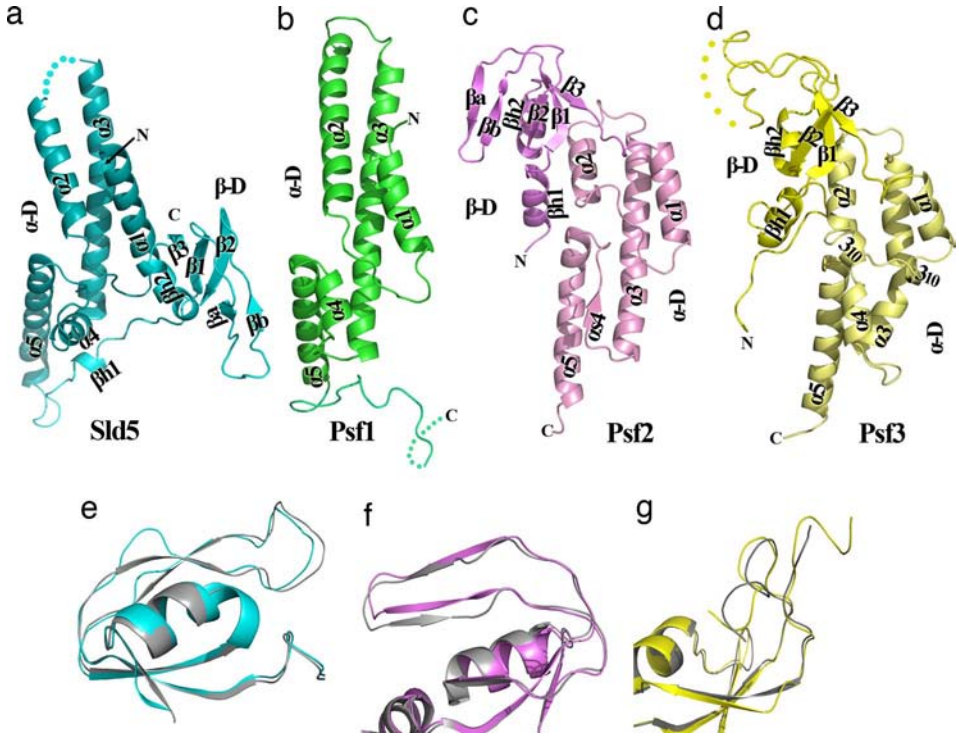

Fig. 2. The structural folds of individual subunits. (*a–d*) The detailed structures of each of the four subunits. N and C indicate the position of the N and C termini in each subunit, whereas dashed lines indicate the location of disordered fragments. The α-domains (α-D) and β-domains (β-D) for each subunit are indicated. Labels for the secondary structures in the α-domains start with α and labels for the β-domains start with β. (e–*g*) A comparison of the β-domains of our structure (in color) and that reported for the structure determined by Kamada *et al.* (13), showing conformational differences in Sld5 (e), Psf2 (*f*), and Psf3 (*g*).

different crystallization conditions but in the same space group with similar unit cell dimensions. Our structure revealed that the C-terminal 51 residues of Psf1 are not folded in the heterotetrameric GINS complex and, thus, are not anchored to the GINS surface. These findings suggest that the Psf1 residues proposed to anchor the β-domain on the GINS surface by Kamada *et al.* (13) could, instead, play a role in binding other protein partners.

Even though the GINS subunits have similar α- and β-domains, the relative arrangements of the two domains differ among the subunits, as predicted in ref. 9. In Sld5 and Psf1, the larger N-terminal portion forms the α-domain, and the smaller C-terminal fragment corresponds to the β-domain (SI Fig. 7*a*). In contrast, the order of α- and β-domains of Psf2 and Psf3 is reversed (i.e., the β-domain is at the N terminus and the α-domain is at the C terminus; SI Fig. 7*b*). In addition, the space between the α- and β-domains for Psf2/Psf3 is only 6 residues, but is 21 residues in Sld5 and possibly about the same length in Psf1, based on sequence alignment (SI Fig. 7 *a* and *b*). Despite the differences in spacer length, the β-domains present in Sld5, Psf2, and Psf3 appear anchored to their respective α-domains through direct contacts (Fig. 2 *a*, *c*, and *d*).

Tetramer Formation. The structure of the GINS tetramer was reported to be formed by an arbitrarily assigned "vertical" interface formed through the α-domains and a "horizontal" interface mediated through the β-domains (13). Despite the similarities of our GINS tetramer structure to that reported, we interpret the interactions that support tetramerization in our crystal structure somewhat differently, as described below. In our structure, each of the four subunits interacts with two other molecules to form a ring-like structure with a central hole (Fig. 3*a*). The four subunits are arranged around the ring in the order Sld5, Psf1, Psf3, and Psf2, such that Psf2 contacts Sld5 to complete the circle. Within the pore created by this ring, there are no direct bonding contacts between Sld5 and Psf3 or between Psf1 and Psf2. The intersubunit interactions are mainly through the sides of the α-helical bundles to generate extensive contacts between subunits with buried interface areas between two neighboring subunits ranging from 2,900 to 4,000 Å^2. These large interface areas presumably provide strong bonding forces at the subunit interfaces, which explains why the tetrameric complex is stable in solution. Major bonding forces are provided through helix–helix interactions between adjacent α-domains of different subunits (Sld5–Psf1, Psf1–Psf3, Psf3–Psf2, and Psf2–Sld5), mediated mostly by hydrophobic residues. However, the β-domains at the interface between Psf2–Sld5 also play a role in stabilizing the tetramer. The β2 of the Sld5 β-domain interacts with a small β-strand (αs4) from the Psf2 α-domain, expanding the Sld5 β-domain to a four-stranded β-sheet (Fig. 3*b*). Additionally, a β-hairpin from the Psf2 β-domain contacts an α-helix of the Sld5 α-domain (Fig. 3*c*).

Possible Roles of the Unstructured Regions of Sld5, Psf1, and Psf3. Although our crystal structure of the GINS complex included full-length proteins of all four subunits, portions of Sld5, Psf1, and Psf3 have no visible electron density (represented by dashed lines in Fig. 2 *a–d* and SI Fig. 7 *a* and *b*), likely due to the flexibility of these regions. Comparisons between our structure and that reported by Kamada *et al.* (13) reveal that the disordered regions in Sld5 and Psf3 are basically the same in both structures. Because our GINS constructs and crystallization conditions differed from those of Kamada *et al.* (13), the fact that similar regions are missing from the electron density maps of both structures suggests that these regions are highly flexible and unstructured in the GINS complex. The last visible C-terminal residue of Psf1 (S145) before the disordered C-terminal domain is adjacent to the disordered fragment of Sld5 (residue 65–71) and to the disordered C-terminal residues of Psf3 (residues 194–216); these disordered regions are shown as spheres in Fig.

BIOCHEMISTRY

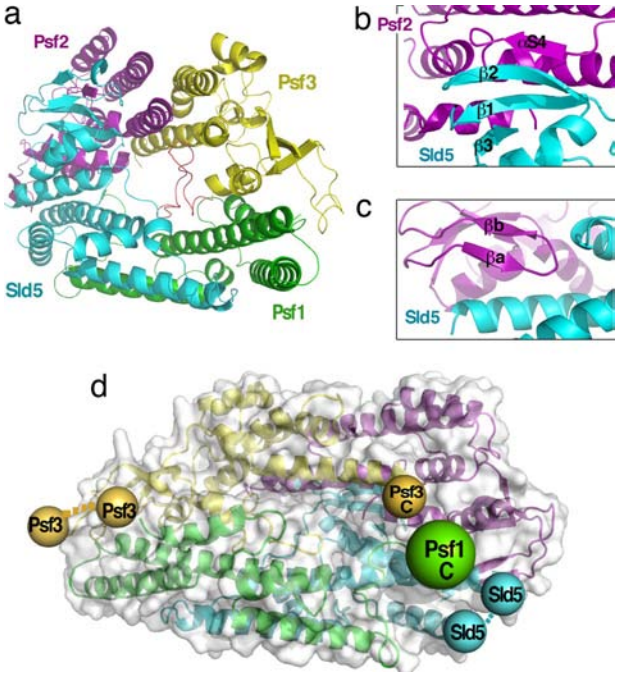

Fig. 3. The ring structure of the GINS complex and colocalization of disordered regions on the surface. (*a*) The ring structure of the GINS complex can be visualized along the long axis (the axis of the α-helical bundles) of the tetramer. From this view, it is clear that most of the intersubunit interactions are mediated through α-helices. The central opening is also apparent. (*b* and *c*) These views show interactions between Sld5 and Psf2, which involve β-strands at two locations on intersubunit interfaces. (*d*) Colocalization of the disordered regions on the tetramer surface. Each sphere indicates the location of the residue immediately adjacent to the disordered fragment. The subunit (and sphere) color scheme is the same as that used in *a*. The disordered fragments from three subunits colocalize on one end of the tetramer. As shown, the disordered region of Psf3 is on the same face but at the opposite end. These disordered regions may become structured when bound to interacting partners.

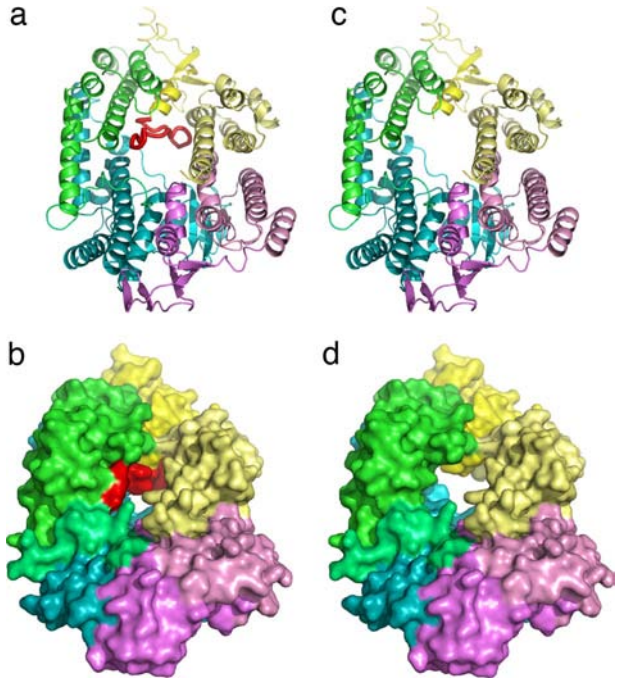

Fig. 4. The structure of the central pore of GINS and its accessibility. (*a*) A view of the central pore of the GINS tetramer. The red loop inside the pore is formed by the N-terminal 16 residues of Psf3. The loop is not tightly bonded to the pore surface. (*b*) The surface representation of the view in *a*. (*c*) The same view as in *a* but without the N-terminal residues of Psf3. (*d*) The surface representation of the view shown in *c*.

3d. The colocalization of these disordered parts of three different subunits on the GINS surface suggests that this site may bind partner proteins in the replication complex. On the same side, but located at the other end of the tetramer, is the disordered region within the β-domain of Psf3 (Fig. 3*d*), which may also serve as a protein-binding site. These disordered regions located on the surface of GINS may become structured upon binding to proteins known to interact with GINS, such as MCM, Cdc45, and DNA polymerases. Most of these are large proteins or complexes and the location of disordered sites on opposite sides of the tetramer may allow binding of more than one of these factors at the same time. In keeping with this idea, previous studies on *Xenopus* reported that the chromatin loading of Cdc45 and GINS were mutually dependent (15). Furthermore, Kamada *et al.* (13) observed that the human GINS complex containing the Psf1 subunit with the flexible C-terminal region truncated (labeled Psf1 C in Fig. 3*d*) failed to bind to the *Xenopus* pre-RC and failed to load Cdc45. These findings support our notion that the flexible regions present on the GINS surface are important for its binding to replicative proteins.

Accessibility of the Central Pore. When visualized by negatively stained EM, the recombinant *Xenopus* GINS complex has a ring-like structure with a central hole (15). Kamada *et al.* (13) suggested that the ring-like shape observed by EM was an artifact resulting from low-density distribution of electrons at the center of the complex. Examination of our GINS structure, however, revealed that it does have ring-like structure when viewed along the central pore, as shown in Fig. 4*a*. Based on EM measurements, the ring diameter was reported to be ≈9.5 nm (95 Å), with the central pore ≈4 nm (40 Å) in diameter (15); in contrast, the diameter of the ring in our crystal structure is, at maximum, ≈78 Å (measured from edge to edge in the large dimension) and the central pore is 10 Å in diameter (Fig. 4 *a* and *b*). These quantitative differences in dimensions could be due to flattening (deformation) and dehydration effects intrinsic to negative staining methods used to prepare EM samples.

Detailed examination of the central pore present in our GINS crystal structure revealed that a 16-residue peptide loop from the N terminus of Psf3 appeared to fit loosely into the pore, effectively restricting the opening (Fig. 4 *a* and *b*, red). Interactions of this peptide loop with the surface of the pore are limited, suggesting that the peptide may enter and leave the pore without a significant energy barrier. Interestingly, sequence alignment of this 16-aa sequence at the N terminus of Psf3 revealed that it is present in human and some higher eukaryotes but not in many other organisms (SI Fig. 8*a*). This sequence data suggest that the first 16 N-terminal residues of Psf3 may not be needed for the structural integrity of the GINS tetramer. A speculative role of this 16-residue peptide may be to regulate the dimensions of the central pore, and thus its accessibility, by plugging and unplugging this cavity.

To test the notion that this N-terminal 16-residue peptide is not required for tetramer formation and stability of the GINS complex, we generated Psf3 constructs lacking either 10 or 18 residues of the N terminus and examined whether such constructs supported tetramer formation. Both Psf3 constructs formed stable tetramers under high and low salt conditions and behaved similarly to the GINS complex containing full-length Psf3 in gel filtration analyses (SI Fig. 8*b*). This finding indicates

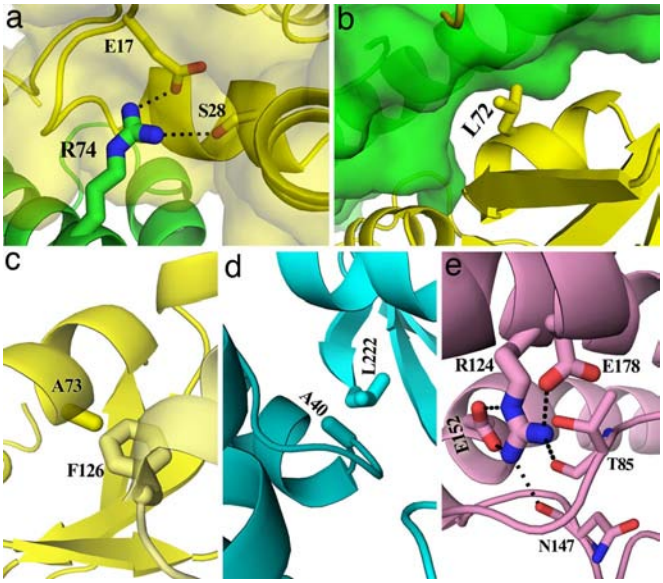

Fig. 5. The mapping of the yeast ts mutants on the equivalent sites present in human GINS structure. (*a* and *b*) The class I ts mutant residues involved in the intersubunit interactions between Psf1 R74 and Psf3 (*a*) and between Psf3 L72 and Psf1 (*b*). (*c*–*e*) The class II ts mutant residues Psf3 A73, Sld5 L222, and Psf2 R124 are involved in intrasubunit interactions. Mutation of these residues, as indicated, would be expected to affect the folding and stability of individual subunits within the GINS complex.

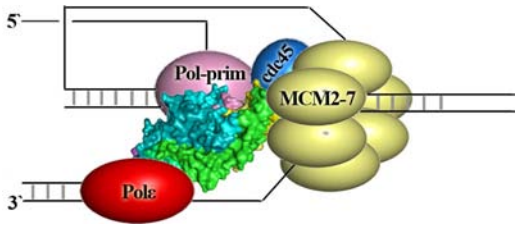

Fig. 6. A model for the GINS complex coordinating MCM, Cdc45, Pol ε, and Pol α-primase complex at the replication fork. The GINS complex is shown in surface representation. All other components (drawn as cartoons) are shown as bound directly to the GINS complex.

that the N-terminal 16-residue loop of Psf3 is not needed for folding and stability of the GINS complex.

The diameter of the central pore is ≈10 Å at its narrowest point, but the opening is increased to 18 Å upon removal of the N-terminal 16-residue loop of Psf3 (Fig. 4 *c* and *d*). This finding may also partially explain the larger pore size observed in EM (15) because the acidic negative staining solution used for EM may dislodge the N-terminal peptide of Psf3 from the pore. Thus, the 16-residue peptide at the N terminus of human Psf3 might regulate pore opening and closing. Although a functional role for the central pore is as yet unknown, we speculate that one possible role is to bind and hold a fragment from MCM, Cdc45, or DNA polymerases during DNA synthesis. It is also possible that the pore could bind ssDNA because its dimension is sufficiently large to accommodate ssDNA even with the N-terminal peptide of Psf3 situated within the pore. In support of this possibility, there are 10 Arg and Lys residues, as well as several Asn and Gln amino acids distributed on the surface of the pore (SI Fig. 9), despite the overall negatively charged outer surface of the tetramer. This charge distribution is similar to that of the proliferating cell nuclear antigen (PCNA) and T7 gp4 helicase, both of which have positively charged and polar residues around the central openings even though they have an overall negatively charged outer surface (14, 16). The GINS complex might require an additional factor(s) to bind ssDNA because DNA-binding activity has not as yet been detected for the GINS complex isolated from human (data not shown) or other organisms (6, 11).

Understanding Temperature-Sensitive (ts) Mutants of GINS. A total of nine different ts mutations have been identified in the four yeast GINS proteins (8, 17, 18). These mutated residues are highlighted in the sequence alignments shown in SI Fig. 7 *a* and *b*, eight of which are highly conserved across eukaryotes. Our GINS structure provides a rationale for all of these ts mutations. The nine ts mutations can be divided into three classes according to their possible defect(s) at nonpermissive temperatures. The mutants in class I involve residues important for intersubunit interactions. Psf1-R74G (Fig. 5*a*) and Psf3-L72P (Fig. 5*b*) belong to this class. Psf1 R74 (yeast Psf1–1 R84G) forms hydrogen bonds with Psf3 E17 and S28 (Fig. 5*a*); a Gly substitution at this position should abolish the two hydrogen bonds and significantly weaken interactions between Psf1 and Psf3, which may lead to instability of the complex at nonpermissive temperatures. Similarly, the Psf3 L72P mutation (yeast Psf3-21 L46P) should reduce the hydrophobic packing with Psf1 F64 (Fig. 5*b*). The ts mutants in class II involve residues important for intrasubunit interactions. Psf3 A73 (yeast Psf3-21 A47P) and Sld5 L222 (yeast Sld5-13 L293P) are both involved in β-domain packing, as well as α- and β-domain interactions (Fig. 5 *c* and *d*). Psf2 R124 (yeast Psf2-209 R133K) forms five hydrogen bonds within three helices and two loops in its α-domain (Fig. 5*e*). Class III ts mutants involve residues buried in the hydrophobic core structures of Sld5 and Psf3. Mutations of these buried residues would be expected to reduce the thermodynamic stability of the structure, providing a plausible molecular explanation for the phenotypes of these mutants. Surprisingly, ts mutations have not as yet been identified on the surface of the GINS. Because GINS interacts directly with the MCM complex, Cdc45, and other replicative protein in different species (4), it is likely that the binding sites for these proteins are conserved. We speculate that mutations on the surface of GINS that result in temperature sensitivity may help identify their binding sites.

Structural Conservation of GINS in Archaea. The archaeon *Sulfolobus solfataricus* has two GINS homologs, Gins15 and Gins23 (10). Sequence analysis indicates that eukaryotic Psf1/Sld5 and Psf2/Psf3 are close paralogs of Gins15 and Gins23, respectively (9), suggesting that these two GINS proteins can form a tetrameric complex through the dimerization of the heterodimer. The two independently derived crystal structures of human GINS reveal that the residues involved in the interfaces between subunits are conserved among archaeal and eukaryotic GINS proteins (data not shown), providing further support for a structural and functional conservation between archaeal and eukaryotic GINS complexes. In keeping with these considerations, it was shown that the GINS complex formed in the archaeon is a tetramer formed through the dimerization of the Gins15–Gins23 heterodimer (10).

A Model for GINS Role at the Replication Fork. A speculative model of how the human GINS tetrameric complex interacts and coordinates the activities of its binding partners is proposed in Fig. 6. In this model, we suggest a direct contact between Psf1 and Pol ε because an interaction between *S. cerevisiae* Psf1 and Dpb2 (the second largest subunit of Pol ε) was detected in a yeast two-hybrid screen (8). Because archaeal Gins23 binds to the N terminus of MCM (10) and an interaction of Psf3 with MCM has been detected in the yeast two-hybrid system (S. Azuma and H. Masukata, unpublished data in ref. 18), our model depicts Psf3

in contact with MCM. We assume that Psf2 contacts the Pol α-primase complex based on the report that archaeal Gins23 interacts with primase (10). The model also allows MCM, Cdc45, and GINS to contact each other based on the isolation of this complex from *Drosophila* eggs (6). Finally, we also suggest that GINS must interact with Pol ε and the Pol α-primase complex to coordinate leading- and lagging-strand synthesis, respectively (11, 12).

In summary, the structural and biochemical data presented both here and by Kamada *et al.* (13) suggest that GINS functions as a tight heterotetrameric complex. The molecular interactions between the subunits of GINS are mediated mostly through helix–helix interactions that amplify the helix-bundle-like structure of each individual subunit. In agreement with an EM study of the *Xenopus* GINS complex (15), we find an open pore along the long axis of the ring-like tetramer structure. The pore size appears to be influenced by the position of the N-terminal 16–20 aa residues of Psf3, providing a possible mechanism for regulating the accessibility of the central opening for its binding to a peptide from an interacting factor or to ssDNA. The conservation of the GINS sequence among archaea and eukaryotes suggests a fundamental functional role for this complex in DNA replication; it is likely that GINS serves as a scaffold for the assembly and maintenance of an active helicase and replication complex at the fork. This structure should provide a framework for future studies directed at how the GINS complex interacts with other replication proteins that jointly support both the initiation of DNA synthesis as well as fork progression.

Materials and Methods

Cloning and Protein Expression, and Purification GINS. hSld5 (GenBank accession no. NM_032336), hPsf2 (GenBank accession no. BC062444), and hPsf3 (GenBank accession no. BC005879) cDNAs were amplified from a HeLa cDNA library. hPsf1 (GenBank accession no. BC012542) cDNA was derived from Integrated Molecular Analysis of Genomes and their Expression (IMAGE) clone 4333095. All of these cDNAs were subcloned into pGEX6P-1 (GE Healthcare, Chalfont St. Giles, U.K.) in the order GST-Psf2, Sld5, Psf1, and His-8–Psf3 to produce a polycistronic mRNA. The N-terminal deletion mutants of Psf3 (N-terminal 10 and 18 residues deleted, respectively) were constructed as a C-terminal His-8 fusion (Psf3–His-8). The four subunits of human GINS were coexpressed in *E. coli* cells by isopropyl β-D-thiogalactoside (Sigma–Aldrich, St. Louis, MO) induction at 18°C. After cells were lysed by passage through a French Press, the GINS complex was purified by nickel-affinity chromatography (Qiagen, Hilden, Germany) and a glutathione resin affinity column (GE Healthcare) in a buffer containing 50 mM Tris·HCl (pH 8.0), 250 mM NaCl (buffer A), and 5 mM 2-mercaptoethanol (Sigma–Aldrich). The GST and His-8 tags were subsequently removed by PreScission protease treatment in buffer A containing 1 mM DTT (Sigma–Aldrich). The GINS complex with 1:1:1:1 molar stoichiometry was obtained by using Resource Q ion-exchange chromatography with a 50–500 mM NaCl gradient elution and gel-filtration chromatography through a Superdex-200 column (GE Healthcare) in buffer containing 50 mM Tris·HCl (pH 8.0) and 50 mM NaCl. The typical yield from 24-liter culture was ≈30 mg.

Crystallization and Structure Determination. Crystals were grown by the hanging drop vapor diffusion method with 20 mg/ml GINS complex against a solution containing 60 mM Mes (pH 5.5), 2% (vol/vol) isopropanol, and 34 mM calcium chloride. Multiple anomalous diffraction data from Se-Met crystals were collected using the synchrotron at Argonne National Laboratory (Argonne, IL) from plate crystals ≈200 × 100 × 20 μm in size (SI Table 1). Data were processed with HKL2000. A total of 52 selenium sites were located by the SHELXD program by using multiple anomalous diffraction data between 30 and 3.5 Å resolution range. The SHARP program was used to calculate the experimental phases by using the multiple anomalous diffraction data in the resolution range of 50–2.5 Å. RESOLVE was used for density modification, resulting in a high-quality electron-density map for model building with O refined with CNS to 2.36 Å resolution. The final refinement statistics and geometry as defined by Procheck were in good agreement and are summarized in SI Table 1.

Note. As this work was being reviewed, another study reporting the GINS complex structure with truncated Psf1 by Choi *et al.* was published (19).

We thank Dr. R. Zhang at APS 19id in Argonne National Laboratory for assistance with data collection and M. Klein and other X.S.C. laboratory members for their help and discussion. This work was supported in part by a Cellular, Biochemical, and Molecular Sciences training grant (to Y.P.C.) and U.S. Army Medical Research and Materiel Command and National Institutes of Health Grants R01AI055926 (to X.S.C.) and GM 34559 (to J.H.).

1. Forsburg SL (2004) *Microbiol Mol Biol Rev* 68:109–131.
2. Blow JJ, Dutta A (2005) *Nat Rev Mol Cell Biol* 6:476–486.
3. Takahashi TS, Wigley DB, Walter JC (2005) *Trends Biochem Sci* 30:437–444.
4. Labib K, Gambus A (2007) *Trends Cell Biol* 17:271–278.
5. Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, Labib K (2006) *Nat Cell Biol* 8:358–366.
6. Moyer SE, Lewis PW, Botchan MR (2006) *Proc Natl Acad Sci USA* 103:10236–10241.
7. Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC (2006) *Mol Cell* 21:581–587.
8. Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H (2003) *Genes Dev* 17:1153–1165.
9. Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV (2005) *Nucleic Acids Res* 33:4626–4638.
10. Marinsek N, Barry ER, Makarova KS, Dionne I, Koonin EV, Bell SD (2006) *EMBO Rep* 7:539–545.
11. Seki T, Akita M, Kamimura Y, Muramatsu S, Araki H, Sugino A (2006) *J Biol Chem* 281:21422–21432.
12. De Falco M, Ferrari E, De Felice M, Rossi M, Hubscher U, Pisani FM (2007) *EMBO Rep* 8:99–103.
13. Kamada K, Kubota Y, Arata T, Shindo Y, Hanaoka F (2007) *Nat Struct Mol Biol* 14:388–396.
14. Gulbis JM, Kelman Z, Hurwitz J, O'Donnell M, Kuriyan J (1996) *Cell* 87:297–306.
15. Kubota Y, Takase Y, Komori Y, Hashimoto Y, Arata T, Kamimura Y, Araki H, Takisawa H (2003) *Genes Dev* 17:1141–1152.
16. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB (2000) *Cell* 101:589–600.
17. Gomez EB, Angeles VT, Forsburg SL (2005) *Genetics* 169:77–89.
18. Yabuuchi H, Yamada Y, Uchida T, Sunathvanichkul T, Nakagawa T, Masukata H (2006) *EMBO J* 25:4663–4674.
19. Choi JM, Lim HS, Kim JJ, Song OK, Cho Y (2007) *Genes Dev* 21:1316–1321.

MOLECULAR AND CELLULAR BIOLOGY, Mar. 2008, p. 1724–1738 Vol. 28, No. 5
0270-7306/08/$08.00+0 doi:10.1128/MCB.01717-07

Minichromosome Maintenance Proteins Interact with Checkpoint and Recombination Proteins To Promote S-Phase Genome Stability▽†

Julie M. Bailis,[1]‡ Douglas D. Luche,[2] Tony Hunter,[1] and Susan L. Forsburg[1,2]*

Molecular and Cell Biology Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037,[1] and Molecular and Computational Biology Section, University of Southern California, 1050 Childs Way RRI 201B, Los Angeles, California 90089-2910[2]

Received 18 September 2007/Returned for modification 19 October 2007/Accepted 19 December 2007

The minichromosome maintenance (MCM) complex plays essential, conserved roles throughout DNA synthesis: first, as a component of the prereplication complex at origins and, then, as a helicase associated with replication forks. Here we use fission yeast (*Schizosaccharomyces pombe*) as a model to demonstrate a role for the MCM complex in protecting replication fork structure and promoting recovery from replication arrest. Loss of MCM function generates lethal double-strand breaks at sites of DNA synthesis during replication elongation, suggesting replication fork collapse. MCM function also maintains the stability of forks stalled by hydroxyurea that activate the replication checkpoint. In cells where the checkpoint is activated, Mcm4 binds the Cds1 kinase and undergoes Cds1-dependent phosphorylation. MCM proteins also interact with proteins involved in homologous recombination, which promotes recovery from arrest by ensuring normal mitosis. We suggest that the MCM complex links replication fork stabilization with checkpoint arrest and recovery through direct interactions with checkpoint and recombination proteins and that this role in S-phase genome stability is conserved from yeast to human cells.

Replicative stress is a major source of mutations that contribute to chromosome instability and the initiation of cancer (5, 26). A primary defense against genome instability is S-phase checkpoints, which recognize DNA damage, arrest the cell cycle, and preserve essential replication structures (10, 38, 62). In response to S-phase agents such as the ribonucleotide reductase inhibitor hydroxyurea (HU), replication forks stall and activate checkpoint kinases (ATR-Chk1 in mammals and Rad3-Cds1 and Mec1-Rad53 in the yeasts *Saccharomyces pombe* and *Saccharomyces cerevisiae*, respectively) (10, 38, 62). The primary function of this replication checkpoint is to retain the assembled replisome at stalled forks during arrest: checkpoint defects lead to catastrophic replication fork collapse and cell lethality. However, recent studies have shown that proteins directly involved in DNA replication as well as recombination are required for cells to repair DNA damage, restart replication, and recover from arrest and must interact with the checkpoint to facilitate these functions (10, 38, 62).

The minichromosome maintenance (MCM) complex is an essential replicative helicase that consists of six related subunits (MCM2 to MCM7) and is required for the initiation and elongation phases of DNA synthesis (reviewed in reference 22). The MCM helicase is therefore centrally positioned to monitor and maintain genome stability at the replication fork. In budding yeast, the use of degron alleles of MCM subunits, which degrade the protein upon a shift to high temperature, has demonstrated that MCM function is required for fork progression throughout S phase as well as for synthesis to resume after forks have stalled in HU (36). In cells lacking the Rad53 protein kinase or its S-phase activator Mrc1, treatment with HU results in excessive DNA unwinding, catastrophic replication fork collapse, and an inability to restart synthesis (12, 15, 21, 55, 73). Under these conditions, MCM proteins remain chromatin bound, but the leading and lagging strands of the replication fork become uncoupled (12, 34, 55). In metazoans, MCM subunits are phosphorylated by the ATM/ATR kinases (17, 33, 67, 82), suggesting that the MCM complex may be a target or an effector of the replication checkpoint.

However, MCM proteins may also promote S-phase genome stability through checkpoint-independent roles (reviewed in references 3 and 22). Although cellular levels of MCM proteins are estimated as 10- to 40-fold excess over the number of replication origins, yeast temperature-sensitive *mcm* mutants (*mcm-ts*) that partially reduce MCM protein levels exhibit increased recombination, chromosome loss, and checkpoint sensitivity (32, 40, 41). In addition, increased expression or amplification of MCM genes is associated with many types of human cancers that are characterized by genomic instability (reviewed in reference 39) and the mutation or the dysregulation of MCM subunits can induce skin carcinoma and mammary carcinoma in mouse models (30, 68). These results suggest that the MCM complex plays a central role in protecting genome stability during S phase.

The stabilization of replisome structure and its recovery from replication arrest also depend upon proteins involved in homologous recombination (HR). In budding yeast, the Rad51 strand exchange enzyme and the Sgs1 recombination helicase are required to maintain DNA polymerase ε at stalled replication forks (16). Sgs1 is additionally involved in checkpoint

* Corresponding author. Mailing address: Molecular and Computational Biology Section, University of Southern California, 1050 Childs Way RRI 201B, Los Angeles, CA 90089-2910. Phone: (213) 740-7342. Fax: (213) 740-8631. E-mail: forsburg@usc.edu.

† Supplemental material for this article may be found at http://mcb.asm.org/.

‡ Present address: Amgen, 1201 Amgen Court West, Seattle, WA 98119.

▽ Published ahead of print on 7 January 2008.

activation (6). In fission yeast, the SpRad22 (ScRad52) recombination protein is recruited to nuclear foci immediately as cells are released from HU arrest (48). Although fission yeast *rad22* and *rhp51* (*rad51*) mutants are proficient for activation of the replication checkpoint, they still show sensitivity to HU (48, 84). Together, these results suggest that components of the replisome interact with recombination proteins to maintain and restart replication forks.

Previously, we and others have shown that most fission yeast *mcm-ts* mutants accumulate approximately 2C DNA content, but undergo a lethal, checkpoint-dependent arrest in late S phase or G_2 (18, 23, 40, 41). In contrast, a degron allele of *mcm4ts* blocks cells at or near replication initiation (42). Here, we use these different *mcm* alleles to examine the consequences to cells of inactivating MCM function during S phase in the absence or the presence of forks stalled by HU. We demonstrate that the loss of MCM function generates DNA breaks, cell cycle arrest, and a loss of viability similar to that observed in mutants that undergo replication fork collapse. Consistent with these results, we find that Mcm4 interacts with the checkpoint protein kinase Cds1 and undergoes Cds1-dependent phosphorylation in cells treated with HU. This result suggests that MCM proteins act to maintain replication fork structure both during normal S phase and during S-phase arrest.

Our data additionally suggest that MCM function is required for proper recovery from replication arrest induced by HU. We observed an interaction between the MCM complex and the HR protein Rhp51 (Rad51). Although Rhp51 and other HR proteins are not required to activate the replication checkpoint, to maintain fork structure in HU, or to restart DNA synthesis, we find that the loss of HR function results in chromosome missegregation following release from HU-induced replication arrest (this work; 48). We suggest that MCM proteins modulate replication fork progression, arrest, and restart and that they couple these functions with the repair of DNA damage to protect S-phase genome stability. Our analysis in fission yeast is complemented by studies of mammalian cells, suggesting that the role of the MCM complex in S-phase genome stability is conserved.

MATERIALS AND METHODS

Yeast strains and media. *S. pombe* strains (Table 1) were constructed and maintained according to standard procedures (53). Cells were treated with 15 or 20 mM HU (Sigma) for the indicated times or with 0.0025% methyl methanesulfate (Sigma) or 5 μU/ml bleomycin (Sigma) for 1 h or were irradiated with 100 Gy using a ^{60}cobalt source. Asynchronous cultures or cultures synchronized by nitrogen starvation (53) were shifted to the restrictive temperature (36°C) for 4 h; cells pulse-labeled with bromodeoxyuridine (BrdU) were shifted to the restrictive temperature for 3.5 h and then incubated with 200 μg/ml BrdU (Sigma) for an additional 30 min at 36°C.

Cell culture and short interfering RNA (siRNA). HeLa cells were grown in Dulbecco's modified Eagle's medium containing 10% calf serum (Invitrogen), with appropriate antibiotics. HeLa cells were synchronized in G_1/S with 2 mM thymidine (Sigma) for 18 h and then released into S phase for 3 to 4 h in the absence of added DNA damage. Cells were blocked in early S phase with 2.5 mM HU for 18 h.

For siRNA experiments, synthetic RNA duplexes were transfected into HeLa cells to a final concentration of 20 nM using Oligofectamine (Invitrogen) according to the protocol provided by Dharmacon. siRNA duplexes used were CGACAGCTAGAGTCATTAA for MCM4 and ATCGGATTGTGAAGA TGAA for MCM7, synthesized by Integrated DNA Technologies, Inc. Pools of four additional siRNA duplexes against MCM4 or MCM7 were obtained from Dharmacon. The siRNA negative control duplex was from Invitrogen. Relative

TABLE 1. Yeast strains used in this study

Strain	Genotype	Source
FY7	*h⁻ 972*	Our stock
FY72	*h⁻ ade6-M210 ura4-D18 leu1-32*	Our stock
FY243	*h⁻ cdc19-P1 ade6-M210 ura4-D18 leu1-32*	Our stock
FY261	*h⁺ ura4-D18 leu1-32 ade6-M216 can1-1*	Our stock
FY364	*h⁻ cdc21-M68 leu1-32 ura4-D18 ade6-M210 can1-1*	Our stock
FY522	*h⁻ cdc22-M45 ade6-M210 leu1-32 ura4-D18*	Our stock
FY641	*h⁻ cdc6-23 ura4-D18 leu1-32 ade6-M210*	Our stock
FY865	*h⁻ Δcds1::ura4⁺ ura4-D18 leu1-32*	Our stock
FY866	*h⁺ Δcds1::ura4-D18 leu1-32*	Our stock
FY961	*h⁺ mis5-268 ura4-D18 leu1-32 ade6-M210 can1-1*	Our stock
FY969	*h⁻ cdc27-K3 ura4-D18 leu1 ade6-M210*	Our stock
FY986	*h⁺ hsk1-1312 ura4-D18 leu1-32 ade6-M216*	Our stock
FY999	*h⁺ hsk1-1312 Δcds1::ura4⁺ ura4-D18 leu1-32*	Our stock
FY1106	*h⁺ Δrad3::ura4⁺ ura4-D18 leu1-32 ade6-M210*	Our stock
FY1167	*h⁻ mcm4-HA::leu1⁺ ura4-D18 leu1-32 ade6-M210*	Our stock
FY1201	*h⁻ mcm7-98 ura4-D18 leu1-32 ade6-M210*	Our stock
FY1209	*h⁺ Δrad22::[hisG ura4⁺ hisG] ade6-M210 ura4-D18 leu1-32*	A. Pastink
FY1313	*h⁻ cdc17-K42 leu1-32 ade6-M210 can1-1*	Our stock
FY1374	*h⁻ cdc19-P1 Δrad22::[hisG ura4⁺ hisG] ade6-M210 leu1-32 ura4-D18*	This study
FY1375	*h⁺ cdc21-M68 Δrad22::[hisG ura4⁺ hisG] ade6-M210 leu1-32 ura4-D18*	This study
FY1390	*h⁻ smt-0 Δrhp51::ura4 ura4-D18*	Our stock
FY1413	*h⁻ cdc21-M68*	Our stock
FY1507	*h⁻ smt-0 Δrhp51::ura4⁺ cdc21-M68 ade6-M210 ura4-D18*	This study
FY1617	*h⁻ cdc21-M68 Δcds1::ura4⁺ ura4 leu1 ade6-M210*	Our stock
FY1884	*h⁻ smt-0 Δrhp51::ura4⁺ ade6-M210 leu1-32 ura4-D18*	This study
FY1963	*h⁻ cdc21-M68 chk1::ura4⁺ ura4-D18*	Our stock
FY2234	*h⁺ Δtel1::ura4⁺ ade6-210 ura4-D18 leu1-32*	Our stock
FY2317	*h⁺ leu1-32::hENT1-leu1⁺*(pJAH29) *his7-366::hsv-tk-his7⁺*(pJAH31) *ura4-D18 ade6-M210*	31
FY2356	*h⁻ smt-0 Δrhp51::ura4⁺ cdc21-HA::leu1⁺ ade6-M210 leu1-32 ura4-D18*	This study
FY2423	*h⁺ cdc21-HA::leu1⁺ cds1::ura4⁺ ade6-M210 leu1-32 ura4-D18*	Our stock
FY2514	*h⁺ leu1-32::hENT1-leu1⁺*(pJAH29) *his7-366::hsv-tk-his7⁺*(pJAH31) *cdc21-M68 leu1-32 ura4-D18 ade6-M210*	This study
FY2728	*h⁺ Δrad3::ura4⁺ Δtel1::ura4⁺ ade6-M210 leu1-32 ura4-D18*	This study
FY2878	*h⁻ rad22:YFP:kanMX4 ade6-M210 ura4-D18 leu1-32*	G. Freyer
FY2887	*h⁻ mcm4 (cdc21-M68)-ts-td::ura4⁺ ade6-M210 leu1-32 ura4-D18*	This study
FY2906	*h⁺ rad22-YFP:kanMX4 cdc21-M68 leu1-32 ura4-D18*	This study
FY3201	*h⁻ cdc19-P1 Δrad32::kanmx ade6-M210 leu1-32 ura4-D18*	This study
FY3203	*h⁻ cdc19-P1 Δrhp51::ura4⁺ ade6-M210 leu1-32 ura4-D18*	This study
FY3204	*h⁻ cdc21-M68 Δrad32::kanmx ade6-M210 leu1-32 ura4-D18*	This study
FY3396	*h⁺ cdc21-M68-ts-dg::ura4⁺ ura4-D18*	This study
FY3650	*h⁺ Δcds1::ura4 Δrhp51::ura4 ura4-D18 leu1-32*	This study
FY3651	*h⁻ smt-0 Δcds1::ura4 Δrhp51::ura4 ura4-D18*	This study
JBY235	*h⁻ mcm4-HA::leu1⁺ cds1-13myc::kanmx leu1-32 ura4-D18*	This study
JBY267	*h⁻ cds1-13myc::kanMX ade6-M210 leu1-32 ura4-D18*	This study

viability was determined as the percentage of viable cells at 48 h compared to the percentage of viable cells at the time of transfection (0 h).

Immunofluorescence. Spread nuclei were prepared from fission yeast cells according to previous procedures (29). Briefly, cells were spheroplasted in phosphate-buffered saline (PBS) with 0.5 mg/ml zymolyase 20T (Seikagaku) and 0.5 mg/ml lysing enzymes (Sigma). Cells were then washed in MES [(2-*N*-morpholino)ethane sulfonic acid] buffer (0.1 M MES, 1 M sorbitol, 1 mM EDTA, 0.5 mM $MgCl_2$, pH 6.4), resuspended in 100 μl MES plus 360 μl 4% paraformaldehyde, and poured onto Colorfrost microscope slides (Thermo Fisher Scientific). Slides were rinsed with 0.4% Photo-Flo (Kodak) and fixed with low heat.

S. pombe chromatin fibers were prepared by a modification of a protocol used to visualize yeast artificial chromosome chromatin fibers. The protocol generates fibers of approximately 1.8 kb/μm (65). Cells were spheroplasted in zymolyase mix (1 M sorbitol, 60 mM EDTA, 100 mM sodium citrate, pH 7, 0.5 mg/ml zymolyase 20T, 100 mM β-mercaptoethanol), and then spheroplasts were air dried onto Colorfrost slides. Fibers were stretched by pipetting 50 μl lysing solution (50 mM Tris-Cl, pH 7.4, 25 mM EDTA, 500 mM NaCl, 0.1% [wt/vol] Nonidet P-40, 1% [wt/vol] sodium dodecyl sulfate [SDS], 3 mM β-mercaptoethanol) onto the spheroplasts, waiting 1 to 3 min, and then holding the slides vertically to drain the liquid. Chromatin fibers were fixed in 4% formaldehyde in PBS for 20 min, rinsed in PBS, and then heat fixed before being immunostained with the indicated antibodies.

Slides were blocked in PBS containing 5% calf serum and 0.05% Tween 20 for 1 h at room temperature. Primary antibody staining was carried out overnight at 4°C in a humid chamber. Antibodies against phosphorylated histone H2AX (phospho-H2AX; Millipore) and BrdU (Becton Dickinson) were used at 1:50. Immunostaining for BrdU was performed as described previously (29), with sequential immunostaining with phospho-H2AX antibodies. Anti-Rhp51 antibodies (14) and antihemagglutinin (anti-HA) antibodies (16B12; Covance) were used at 1:100. Primary antibodies were detected by using Alexa Fluor 488 or Alexa Fluor 594 secondary antibodies (Invitrogen) at 1:500 in PBS with 5% calf serum. DNA was counterstained with 4′,6-diamidino-2-phenylindole (DAPI).

For staining by terminal deoxynucleotidyltransferase-mediated dUTP-biotin nick end labeling, spread *S. pombe* chromosomes were labeled with dUTP-digoxigenin as described previously (4) and then immunostained with an anti-digoxigenin antibody (Invitrogen) and detected by incubation with an anti-digoxigenin secondary antibody coupled to rhodamine (Invitrogen). Slides were immunostained sequentially with antibodies to detect terminal deoxynucleotidyltransferase-mediated dUTP-biotin nick end labeling and phospho-H2AX.

HeLa cells used for indirect immunofluorescence experiments were grown on glass coverslips. Cells on coverslips were permeabilized with PBS that contained 0.1% Triton X-100 for 7 min and then fixed with 3.7% formaldehyde in PBS for 15 min. Cells were then incubated with blocking buffer (PBS with 5% bovine serum and 0.5% Triton X-100) for 30 min. Stretched chromatin fibers were prepared from HeLa cells as described previously (7); cells collected by trypsinization were resuspended in 75 mM KCl and pipetted onto Colorfrost slides. Slides were incubated in lysis buffer (25 mM Tris-Cl, pH 7.5, 0.5 M NaCl, 1% Triton X-100, 0.5 M urea) for 10 min and then lifted slowly out of the buffer and fixed with 3.7% formaldehyde in PBS for 15 min. Chromatin fibers generated by this method are approximately 2 kb/μm (49). HeLa cells and chromatin fibers were immunostained overnight at 4°C in a humid chamber with antibodies against phospho-H2AX (Millipore), MCM7 (Santa Cruz), and Rad51 (Santa Cruz), each used at 1:100 dilution; BrdU (Becton Dickinson) used at 1:50 dilution; or cleaved caspase 3 (Cell Signaling) used at 1:500 dilution. Primary antibodies were detected with Alexa Fluor 594 or Alexa Fluor 488 secondary antibodies used at 1:500. DNA was counterstained with DAPI.

Images were visualized and acquired using a Leica DMR microscope and Improvision software or with a Leica 6000 microscope and Leica FW4000 software.

Flow cytometry. *S. pombe* cells were fixed in 70% ethanol and then treated with RNase in 50 mM sodium citrate. Cells were resuspended in Sytox green dye (Invitrogen) at 1:2,500 in 50 mM sodium citrate as described previously (2) for both flow cytometry and examination of nuclear morphology.

HeLa cells were collected by trypsinization and then fixed with 70% ethanol for 30 min. Cells were then washed in PBS, treated with RNase in PBS, and then resuspended in PBS with 4 μg/ml propidium iodide.

Cells were analyzed on a FACScan cytometer (Becton Dickinson) by using CellQuest software.

Viability assays. *S. pombe* cells were grown asynchronously at the permissive temperature for 16 h, and then HU was added for 4 h. For each condition, 5,000 cells were counted and plated for absolute viability. The remaining cells were then washed and transferred to the same original volume of medium, either with or without HU, at the permissive or restrictive temperature. For each condition, 5,000 cells were plated to assess relative viability. Colonies were counted after 4 days of growth on yeast extract plus supplements (YES) solid medium at the permissive temperature. Cellular DNA content was analyzed by flow cytometry as described above.

Immunoprecipitation and immunoblotting. We prepared soluble protein lysates from fission yeast cells by vortexing cells with glass beads in B88 lysis buffer (20 mM HEPES pH 7.0, 50 mM potassium acetate, 5 mM magnesium acetate, 100 mM sorbitol, 0.1% Triton X-100) to which 2 mM dithiothreitol, protease inhibitor cocktail (Sigma), 50 mM sodium orthovanadate, and 50 mM β-glycerophosphate were added (2). A total of 0.5 to 1.0 mg protein was used per immunoprecipitation. Proteins were analyzed by 10% SDS-polyacrylamide gel electrophoresis (PAGE), followed by immunoblotting with anti-HA (16B12; Covance), anti-myc (9e10), antitubulin (4A1) (31), and anti-Rhp51 antibodies (14), each at a 1:1,000 dilution. Antibodies were covalently coupled to protein A beads as described previously (2).

To assay the depletion of MCM protein and the coprecipitation of MCM proteins with Rad51, HeLa cells were lysed in modified CSK buffer [10 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid), pH 6.8, 100 mM NaCl, 300 mM sucrose, 1 mM $MgCl_2$, 1 mM EGTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 10 U/ml aprotinin, 50 μM NaF, 50 μM β-glycerophosphate, 0.5% Triton X-100] for 20 min at 4°C. Lysates were centrifuged at a low speed for 5 min and then analyzed by 10% SDS-PAGE and immunoblotting with anti-MCM4 (Becton Dickinson), anti-MCM7 (Santa Cruz), anti-PCNA (Santa Cruz), or anti-Rad51 (Santa Cruz) antibodies (all used at 1:1,000).

Mcm4 mobility shift. Whole-cell protein extract was prepared with modifications from Foiani et al. (21a). Cells were disrupted with glass beads by vortexing in 10% trichloroacetic acid, washed with acetone, and resuspended in Laemmli buffer. A total of 150 μg of protein was analyzed by 8% SDS-PAGE, followed by immunoblotting with anti-MCM4 antibodies (66). For phosphatase treatment, soluble protein lysates (described above) were incubated for 20 min at 30°C with lambda protein phosphatase (New England Biolabs). Reaction products were analyzed by 8% SDS-PAGE, followed by immunoblotting with anti-MCM4 (66) and anti-HA antibodies (16B12; Covance).

RESULTS

***mcm-ts* mutants cause irreversible DNA damage.** Although MCM activity is crucial throughout S phase, most *mcm-ts* mutants appear to complete bulk DNA replication (18, 23, 40). However, these *mcm-ts* mutant cells activate a Chk1-dependent checkpoint and arrest irreversibly in late S phase or G_2 (41, 47), raising the possibility that the cells accumulate DNA damage during replication termination or in specific chromosomal regions. To test this idea, we prepared spread nuclei from wild-type and *mcm-ts* mutant cells grown at permissive and restrictive temperatures and immunostained for phosphorylated histone H2A (phospho-H2A), a marker for DNA damage. In mammalian cells, the histone H2A variant H2AX is rapidly and specifically phosphorylated in response to DNA double-stranded breaks (DSBs) and this signal spreads to regions of flanking chromatin megabases away from the break sites (64). A phosphorylation-specific H2AX antibody (phospho-H2AX) recognizes a phospho-SQE amino acid sequence in the carboxyl terminus of the H2AX protein that is highly conserved in other species (64, 69), including the *S. pombe* histone H2A homologues Hta1 and Hta2 (54). Phospho-H2A foci in yeast form in response to chemical agents that cause DNA DSBs, localize to DNA break sites, and are dependent on checkpoint protein kinases, indicating that they also mark DSBs (54, 83) (see Fig. S1A, S1B, and S1C in the supplemental material).

Wild-type cells and *mcm4ts* cells grown asynchronously at the permissive temperature contain very little DSB damage: we detected phospho-H2A foci in only a small fraction of spread nuclei, and there were <5 foci per nucleus (Fig. 1A and data not shown). In contrast, we observed extensive DNA damage in elongated, arrested *mcm4ts* cells that had been held at the

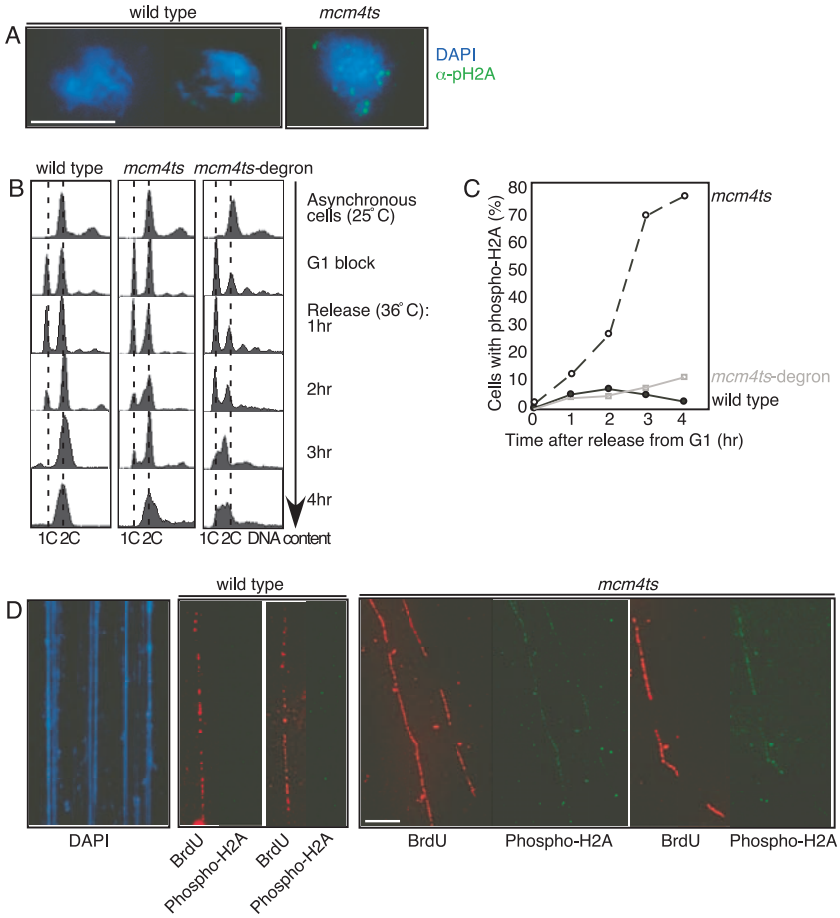

FIG. 1. *mcm-ts* mutants generate DNA damage during replication. (A) *mcm4ts* cells generate DNA damage. Spread nuclei prepared from wild-type (FY72) and *mcm4ts* (FY364) cells shifted to the restrictive temperature (36°C) for 4 h were immunostained for phospho-H2A (green) to mark DNA breaks and counterstained with DAPI (blue) to detect the DNA. α-pH2A, anti-phospho-H2A antibody. Scale bar, 10 mM. Two examples of wild-type nuclei are shown; while most nuclei did not contain phospho-H2A foci, a small number of S-phase nuclei displayed a few phospho-H2A foci. (B and C) DNA damage in *mcm4ts* cells requires S-phase progression. Wild-type (FY72), *mcm4ts* (FY364), and *mcm4ts-degron* (FY2887) cells grown at 25°C were blocked in G_1 by nitrogen starvation and then released into S phase at 36°C. At 1-h intervals, (B) DNA content was measured by flow cytometry, and (C) DNA damage was examined by immunostaining spread nuclei for phospho-H2A. (D) DNA damage in *mcm4ts* cells corresponds to sites of new synthesis. Wild-type (FY2317) and *mcm4ts* (FY2514) cells were shifted to the restrictive temperature for 3.5 h and then pulse-labeled with BrdU for 30 min. Stretched chromatin fibers were then prepared from these cells and visualized with DAPI (left). Chromatin fibers were immunostained with antibodies to BrdU and phospho-H2A to detect sites of replicating DNA and DNA breaks, respectively (right). Scale bar, 10 μm (approximately 18 kb).

restrictive temperature for 4 h, with more than 80% of the spread nuclei containing >20 phospho-H2A foci per nucleus (Fig. 1A). In addition, the *mcm4ts* mutant cells also displayed nuclear foci of the DNA repair proteins Rhp51 and Rad22 (the fission yeast homologues of Rad51 and Rad52) (see Fig. S1D and S1E in the supplemental material; data not shown), which have previously been used as molecular markers for DNA damage in fission yeast (20, 48), supporting the conclusion that DNA breaks occur in these cells.

We also examined other mutants affecting the MCM complex, and additional replication mutants, for the formation of DNA damage at the restrictive temperature. All mutants analyzed contained Rhp51 foci, which mark single-stranded breaks and DSBs (13, 60, 76). Mutations affecting subunits of DNA polymerase δ (*cdc6* and *cdc27*) and the *cdc22* ribonucleotide reductase mutant, all of which retain viability upon a shift to high temperature, displayed few phospho-H2A foci. In contrast, *mcm-ts* mutants and the DNA ligase *cdc17* mutant all contained extensive phospho-H2A foci, indicating DNA DSBs, and suffered severe losses of viability. Thus, phospho-H2A foci correlate with inviability (see Fig. S1D in the supplemental material; Table 2). Together, these data suggest that stability of the MCMs in the replisome is important for cell viability following replication arrest.

DNA damage in *mcm-ts* cells requires replication progression. While *mcm-ts* alleles complete bulk DNA synthesis (18, 23, 40, 41), a degron allele of *mcm4ts* blocks initiation, resulting in limited or no DNA synthesis (42). To determine whether the lethal DNA damage in *mcm-ts* cells occurred as a result of defects in initiation or replication elongation, we compared phospho-H2A focus formation in spread nuclei from wild-type, *mcm4ts*, and *mcm4ts-degron* cells following release from G_1/S arrest into S phase at the restrictive temperature. We observed a dramatic accumulation of phospho-H2A foci in *mcm4ts* cells

TABLE 2. DNA damage in fission yeast replication mutants[a]

Characteristic or mutation	Strain	Wild-type gene function	Arrest point	Relative viability (%)	% Spread nuclei with foci at 36°C		
					Phospho-H2A	Rhp51	Phospho-H2A and Rhp51
Wild type	FY72				4	4	4
mcm2 (*cdc19*)	FY243	MCM complex: replicative helicase	Late S	16.3	84	84	84
mcm4 (*cdc21*)	FY364		Late S	29.5	94	94	94
mcm6 (*mis5*)	FY961		Late S	18.8	75	75	75
mcm7	FY1201		Late S	15.2	85	85	85
cdc6	FY641	DNA polymerase	Late S	111.1	17	38	17
cdc17	FY1313	DNA ligase	Late S	29.9	90	90	90
cdc22	FY522	Ribonucleotide reductase	Early S	102.3	17	55	17
cdc27	FY969	DNA polymerase	Late S	131.3	10	46	10

[a] Yeast replication mutants that arrest irreversibly at the restrictive temperature generate phospho-H2A foci. Cultures were grown to early log phase at 25°C and then shifted to 36°C for 4 h. Arrest points were determined by flow cytometry and analysis of cell morphology (data not shown). At 0 and 4 h after shift to 36°C, aliquots of cells were plated and incubated at the permissive temperature. Relative viability was determined as the number of viable colonies from the 4-h time point compared to the number at the 0-h time point. Spread nuclei prepared from arrested cells were immunostained for phospho-H2A and Rhp51. At least 300 spread nuclei per strain were examined.

as they progressed through S phase, and these persisted in cells that arrested (Fig. 1B and C). In contrast, the *mcm4ts-degron* cells displayed few phospho-H2A foci, and the inefficient DNA synthesis coupled with cell elongation resulted in apparent delayed and incomplete S-phase progression by flow cytometry. This observation suggests that the DNA damage in *mcm4ts* cells occurs during replication elongation, which would be consistent with replication fork collapse. In agreement with this suggestion, the phospho-H2A foci observed in *mcm4ts* cells were suppressed by the addition of HU (<5% of cells with foci [data not shown]), which delays cells in early S phase with replication initiated from only a subset of origins and active forks stabilized by the checkpoint. In comparison, a low percentage of wild-type cells exhibited a few phospho-H2A foci during S phase, but these disappeared as the cells completed S phase (Fig. 1C), suggesting that low levels of spontaneous DNA damage or fork collapse may occur during replication but that such damage is normally recognized and repaired.

Genome-wide replication fork breakage can be monitored using pulsed-field gel electrophoresis gels that detect fragmented DNA (for an example, see reference 51). However, we found that the majority of chromosomes prepared from *mcm-ts* cells at the restrictive temperature did not enter pulsed-field gel electrophoresis gels (41; data not shown), suggesting that they contain topologically complex chromosomal structures, such as replication bubbles or other S-phase intermediates that impair electrophoretic mobility. The same result has also been observed for other replication mutants that arrest S-phase progression (for an example, see reference 81). Although replication fork breakage at specific sites can be monitored by two-dimensional gel electrophoresis, the DNA breaks in *mcm-ts* cells occur late in S phase and are unlikely to be associated with replication origins or other defined sequences. Therefore, as an alternative method to examine DNA breaks associated with replicating DNA in *mcm4ts* mutants, we pulse-labeled cells with BrdU and then prepared stretched chromatin fibers which we costained for BrdU, phospho-H2A, and DAPI (Fig. 1D). The BrdU incorporation pattern on fibers from both wild-type and *mcm4ts* cells was discontinuous and punctate, suggesting that DNA synthesis had initiated from multiple, discrete origins (Fig. 1D). For chromatin fibers prepared from wild-type cells pulse-labeled with BrdU (Fig. 1D), just 9% of BrdU-labeled fibers displayed phospho-H2A staining, indicating that the chromatin-stretching procedure itself and the short incubation with BrdU did not cause DNA DSBs. In contrast, 63% of BrdU-stained chromatin fibers prepared from *mcm4ts* cells also exhibited extensive phospho-H2A staining and these damage foci overlapped with regions of BrdU incorporation (Fig. 1D). These results suggest that *mcm-ts*-induced breaks are associated with regions of newly synthesized DNA, consistent with the model that they result from replication fork collapse. Importantly, these fibers are chromatinized and fully proteinated; the DNA break sites may be held together by protein-protein interactions, such as those suggested for the recombination complex Mre11-Rad50-Nbs1 (45), or may represent DSBs in the process of repair.

MCM proteins are required for replication restart after arrest. These observations suggest that one function of the MCM complex is to protect replication fork structure during normal S phase. Because MCM subunits can be phosphorylated by checkpoint kinases (17, 33, 67, 82), we considered the possibility that the MCM complex also contributes to fork stability during replication arrest. To test this idea, we treated wild-type, *mcm4ts*, or *mcm4ts-degron* mutant cells with HU, which stalls replication forks and activates the replication checkpoint. We next inactivated the Mcm4 subunit by shifting to the restrictive temperature while maintaining HU arrest. We then asked whether this treatment generated DNA DSBs and whether the cells could recover from the arrest (Fig. 2A). Wild-type cells with normal Mcm4 function arrested in HU with a primarily 1C DNA content, and the cells completed S phase and resumed cycling upon release from HU (Fig. 2B and D). Few phospho-H2A foci were observed at either the permissive or the restrictive temperature, consistent with our previous observations that suggest limited, spontaneous DNA damage can occur as a normal part of S-phase progression (Fig. 2C; see Fig. S1 in the supplemental material). In a manner similar to that of wild-type cells, *mcm4ts-degron* cells also arrested in HU, with few phospho-H2A foci (Fig. 2B and C). However, upon a shift to the restrictive temperature, we observed that *mcm4ts-degron* cells generated numerous phospho-H2A foci, even though the cells were maintained in HU and

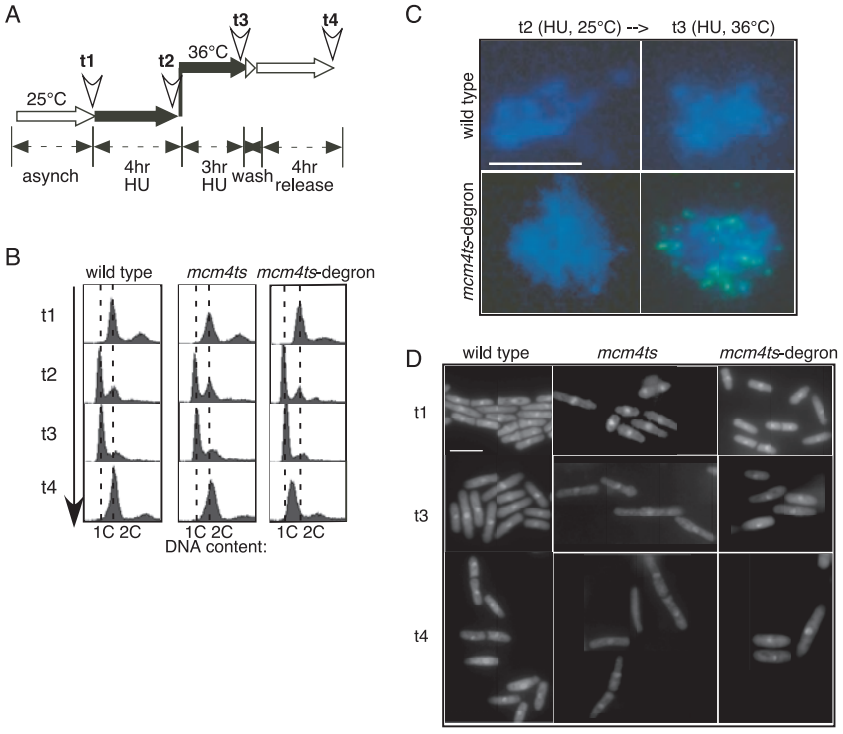

FIG. 2. MCM proteins are required for recovery from HU. (A) Schematic of the experimental protocol. Wild-type (FY72), *mcm4ts* (FY364), and *mcm4ts-degron* (FY2887) cells were grown asynchronously (white arrows) and then arrested in HU for 4 h (black arrows). Cells were shifted to the restrictive temperature of 36°C and then released from HU. Samples were taken at time points t1, t2, t3, and t4. (B) Cellular DNA content was analyzed by flow cytometry. All strains contained approximately 1C DNA content in HU. The wild-type and *mcm4ts* cells, but not the *mcm4ts-degron* cells, completed bulk synthesis upon release from HU. (C) Inactivation of Mcm4 resulted in DNA breaks in HU. Wild-type and *mcm4ts-degron* mutants were arrested in HU at 25°C (t2) and shifted to 36°C while maintained in HU (t3). DNA DSBs were detected by phospho-H2A immunostaining as described in the legend to Fig. 1. (D) Cellular morphology was examined by staining with Sytox green to detect the DNA. Scale bars, 10 μm.

remained arrested (Fig. 2B and C). This result suggests that the loss of Mcm4 function after replication forks have formed and stalled causes extensive DNA damage. Upon release from HU at the restrictive temperature, the *mcm4ts-degron* mutants showed little additional DNA accumulation and no sign of division, consistent with the essential role for the MCM complex throughout DNA replication (36).

Unexpectedly, we found that *mcm4ts* cells blocked in HU and then shifted to a high temperature could recover from arrest, enter mitosis, and remain viable, with an increase in cell number similar to that of the wild type (178.7% ± 11.3% [mean ± standard deviation] relative viability for *mcm4ts* compared to 178.0% ± 6.4% for the wild type) (Fig. 2A). This result was surprising because *mcm4ts* mutants shifted asynchronously to the restrictive temperature, or released from nitrogen starvation to the restrictive temperature, accumulate phospho-H2A foci and lose viability (Fig. 1). The *mcm4ts* mutant contains a single amino acid substitution (Leu238 to Pro), a relatively modest lesion that may cause loss of function at a high temperature through a conformational change in the protein (42). We hypothesized that activation of the replication checkpoint during HU arrest might rescue the defect associated with *mcm4ts* by stabilizing the Mcm4ts mutant protein.

Mcm4 is a target of the replication checkpoint. We therefore tested whether Mcm4 interacts with the Cds1 checkpoint kinase and whether Mcm4 is a target of Cds1-dependent phosphorylation. We found that Mcm4 and Cds1 coimmunoprecipitate in the presence of HU (Fig. 3A). We also observed a mobility shift of the Mcm4 protein in extracts from HU-treated cells (Fig. 3B to E). This shifted band could be collapsed by phosphatase treatment, suggesting that it represents the phosphorylation of Mcm4 (see Fig. S2A in the supplemental material). Identical results were observed using a Δ*cds1* strain with untagged Mcm4 (FY865) (data not shown). Although we observed the Mcm4 mobility shift in both wild-type and *mcm4ts* strains, it did not occur in *mcm4ts-degron* strains, even under permissive conditions. Interestingly, the *mcm4ts-degron* mutant lost viability in HU, prior to any temperature shift, suggesting that it has a defect in the replication checkpoint response even at the permissive temperature (see Fig. S2B, S2C, and S2D in the supplemental material).

Consistent with the possibility that Mcm4 is a target of the replication checkpoint, we found that the Mcm4 mobility shift depends upon Cds1 (Fig. 3E) as well as on the Hsk1 (ScCdc7) kinase (see Fig. S2B, S2C, and S2D in the supplemental material). Hsk1 is required for full Cds1 activation (75) and is itself a potential Cds1 target (72). Simple overexpression of Cds1 was unable to rescue *mcm4ts* inviability (data not shown), indicating that activation of the checkpoint is required. These data suggest that Mcm4 undergoes Cds1-dependent phosphorylation that may contribute to replication checkpoint function

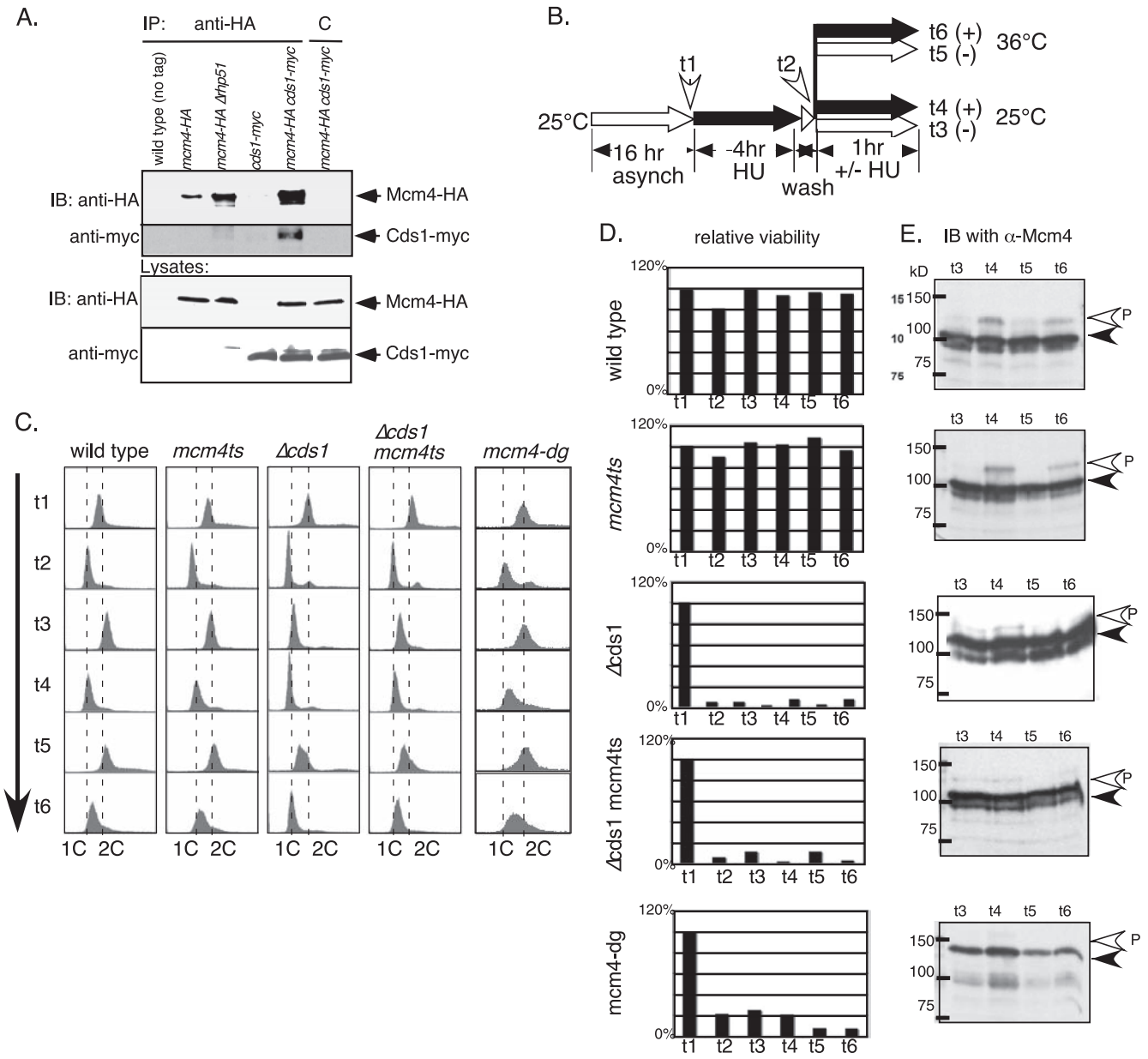

FIG. 3. Cds1-dependent phosphorylation of Mcm4. (A) Mcm4 coimmunoprecipitates Cds1. Protein lysates prepared from wild-type (FY72), *mcm4-HA* (FY1167), *mcm4-HA* Δ*rhp51* (FY2356), *cds1-myc* (JBY267), and *mcm4-HA cds1-myc* (JBY235) strains incubated with 15 mM HU for 4 h were immunoprecipitated (IP) with anti-HA antibodies or control (C) antibody (antitubulin); immunoprecipitates were then immunoblotted (IB) to detect Cds1-myc and Mcm4-HA. (B). Schematic of the experimental protocol. Cells were grown asynchronously (white arrow) and arrested in HU for 4 h (black arrow). Cells were then released into medium with or without HU, either at 25°C or 36°C. Samples were taken at time points t1, t2, t3, t4, t5, and t6. (C) Cellular DNA content was analyzed in wild-type (FY7), *mcm4ts* (FY1413), Δ*cds1* (FY2423), Δ*cds1 mcm4ts* (FY1617), and *mcm4ts-degron* (FY3396) cells by flow cytometry. (D) Relative viability of the strains in panel C was determined as described in Materials and Methods. (E) Mcm4 gel mobility shift in HU. Trichloroacetic acid-precipitated lysates from the strains indicated in panel C were immunoblotted (IB) for Mcm4 protein with anti-Mcm4 antibody (α-Mcm4). The shifted (phosphorylated [P]) form of Mcm4 is indicated by the white arrowhead. The Mcm4-degron protein exhibits reduced mobility relative to wild-type Mcm4 because of the large N-terminal degron tag. Strain FY2423 contains Mcm4HA, but identical results were observed using FY866, which has untagged Mcm4 (data not shown).

and further imply that the Mcm4ts protein is stabilized through Cds1-dependent modification in HU.

MCM proteins link checkpoint activation to recovery. In prokaryotes, it is well established that replication forks that arrest or break can restart and resume synthesis through recombination (reviewed in references 27 and 50). However, evidence for a similar pathway in eukaryotic cells remains controversial. We observed that *mcm-ts* Δ*rhp51* and *mcm-ts* Δ*rad22* double mutants were unable to grow at intermediate temperatures that were permissive for either single mutant (Table 3). This observation suggests that HR is required to repair the fork-dependent DNA breaks in *mcm-ts* cells or that

TABLE 3. Genetic interactions between mutants of the MCM complex and DNA repair mutants[a]

Strain	Genotype	Phenotype
FY1507	*mcm4ts* Δ*rhp51*	Slow/reduced growth at 29°C, 32°C, and 34°C
FY1375	*mcm4ts* Δ*rad22*	Slow/reduced growth at 29°C, 32°C, and 34°C
FY3203	*mcm2ts* Δ*rhp51*	Slow/reduced growth at 29°C, 32°C, and 34°C
FY1374	*mcm2ts* Δ*rad22*	Slow/reduced growth at 29°C, 32°C, and 34°C
FY3204	*mcm4ts* Δ*rad32*	No genetic interaction
FY3201	*mcm2ts* Δ*rad32*	No genetic interaction

[a]HR mutants decrease the ability of *mcm4ts* to grow at semipermissive temperatures. Double mutant strains were compared against the single mutants for growth on solid rich medium (YES) incubated at 25°C, 29°C, 32°C, and 34°C for 4 days.

HR proteins cooperate with the replication machinery to protect replication fork structure. Given that both MCM proteins and HR proteins are predicted to act at or near replication forks, we tested for physical association of these proteins. Intriguingly, we found that Rhp51 (ScRad51) coimmunoprecipitated with Mcm4 from cells treated with HU when both proteins were expressed at endogenous levels (Fig. 4A; see Fig. S3 in the supplemental material). The interaction between Rhp51 and Mcm4 was still observed when lysates were pretreated with DNase (data not shown), indicating that it was not an indirect effect of the association of both proteins with the same DNA fragments. In contrast, little or no Rhp51 bound to Mcm4 in asynchronous cultures (Fig. 4A) or in cells released synchronously from a G_1 block into S phase in the absence of HU (see Fig. S3 in the supplemental material). It remains a possibility that Rhp51 and Mcm4 do associate during normal S phase in the subset of cells that generate Rhp51 or phospho-H2A foci, which may reflect sites of spontaneous DNA damage. We also observed that Mcm4 coimmunoprecipitates overexpressed Rad22-yellow fluorescent protein (YFP) from cells treated with HU (data not shown), suggesting that the MCM complex may interact with repair complexes. These interactions could link fork protection and replication restart with recovery from arrest. Therefore, we examined the dynamics of recovery from replication arrest in recombination mutants in more detail.

When cells are treated with HU, replication forks stall, leading to Cds1 activation and replication arrest. In the Δ*cds1* mutant, this checkpoint cannot be activated to stabilize fork structure and the forks collapse, leading to rapid loss of viability. This is also the case in Δ*rad3* mutants, because the Rad3 kinase is required for Cds1 phosphorylation and activation. In contrast, while the Δ*rhp51* and Δ*rad22* recombination mutants are sensitive to growth on solid medium that contains HU, they display only a modest loss of viability during continuous HU exposure (48, 84) (Fig. 4B and data not shown). The viability loss in Δ*rhp51* and Δ*rad22* cells occurs after 3 to 4 h in HU, which corresponds to the time point when wild-type cells begin to adapt to the drug and resume cell cycle progression (Fig. 4B). Because the loss of viability in HR mutants is modest compared to that in checkpoint mutants and occurs at a later time point, we infer that HR proteins have a distinct, nonessential role in recovery from HU. We next assessed phospho-H2A focus formation in checkpoint and recombination mutant cells that were blocked in HU and then released. Nearly all of the Δ*cds1* and Δ*rad3* checkpoint mutant cells exposed to HU contained substantial numbers of phospho-H2A foci (Fig. 4C), consistent with previous evidence that replication forks collapse in the absence of replication checkpoint activation (48, 56, 73). In the recombination mutant cells, however, a much lower percentage of spread nuclei displayed phospho-H2A foci and there were fewer foci per nucleus (<5 foci/nucleus) than in the checkpoint mutant cells (>20 foci/nucleus) (Fig. 4C). These results suggest that replication forks do not collapse in HR mutants, leading us to conclude that HR proteins are not required for replication fork protection during HU-induced arrest.

We also examined recruitment of the HR proteins themselves by using a Rad22-YFP fluorescent fusion protein described previously (20, 48) as well as by using indirect immunofluorescence to detect Rad22 and Rhp51 (Fig. 4D and E). In asynchronous cultures, a low frequency of wild-type cells displayed nuclear foci of Rhp51 and Rad22 (Fig. 4A and data not shown), a result similar to what we observed for phospho-H2A; these foci may represent sites of spontaneous DNA damage or stalled replication forks in S-phase cells (37). However, the percentage of cells with Rad22 or Rhp51 foci was increased in cells arrested in HU and released. In asynchronous cultures, 8.3% ± 0.57% of cells displayed one Rad22-YFP focus, while fewer than 1% of cells exhibited more than one Rad22-YFP focus (Fig. 4D). After 2 h in HU, 24% ± 2.0% of cells displayed at least one Rad22-YFP focus. By 4 h in HU, when nearly all of the cells were elongated and wild-type cells were starting to adapt and reenter the cell cycle, 82.0% ± 3.2% of cells exhibited at least one Rad22-YFP focus (Fig. 4D). Nuclear foci of Rhp51 were detected in cells treated with HU with timing similar to that for Rad22-YFP (Fig. 4E). While just 6.0% ± 2.8% of spread nuclei prepared from cells from asynchronous cultures displayed Rhp51 foci, 18.5% ± 2.1% of spread nuclei prepared from cells incubated with HU for 2 h and 82.0% ± 2.8% of spread nuclei prepared from cells incubated with HU for 4 h contained multiple Rhp51 foci (Fig. 4C). Consistent with results of previous reports (44, 48), we found that the localization of recombination proteins to nuclear foci did not occur immediately upon treatment with HU but occurred during the recovery stage, consistent with the timing of viability loss in HR mutants (Fig. 4B).

To investigate how recombination contributes to survival of replication arrest, we blocked cells in HU for 4 h, released the cells into fresh medium without HU, and assayed DNA content, cell morphology, viability, and DNA damage during arrest and recovery. Wild-type cells elongated and arrested with a ~1C DNA content on treatment with HU and completed bulk DNA replication and retained normal cell morphology and viability when released from HU (Fig. 5A). Similar results were observed for *mcm4ts* cells (Fig. 3C and data not shown) as described above. In contrast, Δ*rad3* checkpoint mutant cells were unable to restrain cell division in the presence of HU, resulting in immediate chromosome missegregation and cell lethality; Δ*cds1* mutant cells initially arrested due to activation of the Chk1-dependent checkpoint, but then they also divided in the presence of DNA damage (8, 48) (Fig. 5). Significantly, we observed that 78% and 77% of the missegregation events in

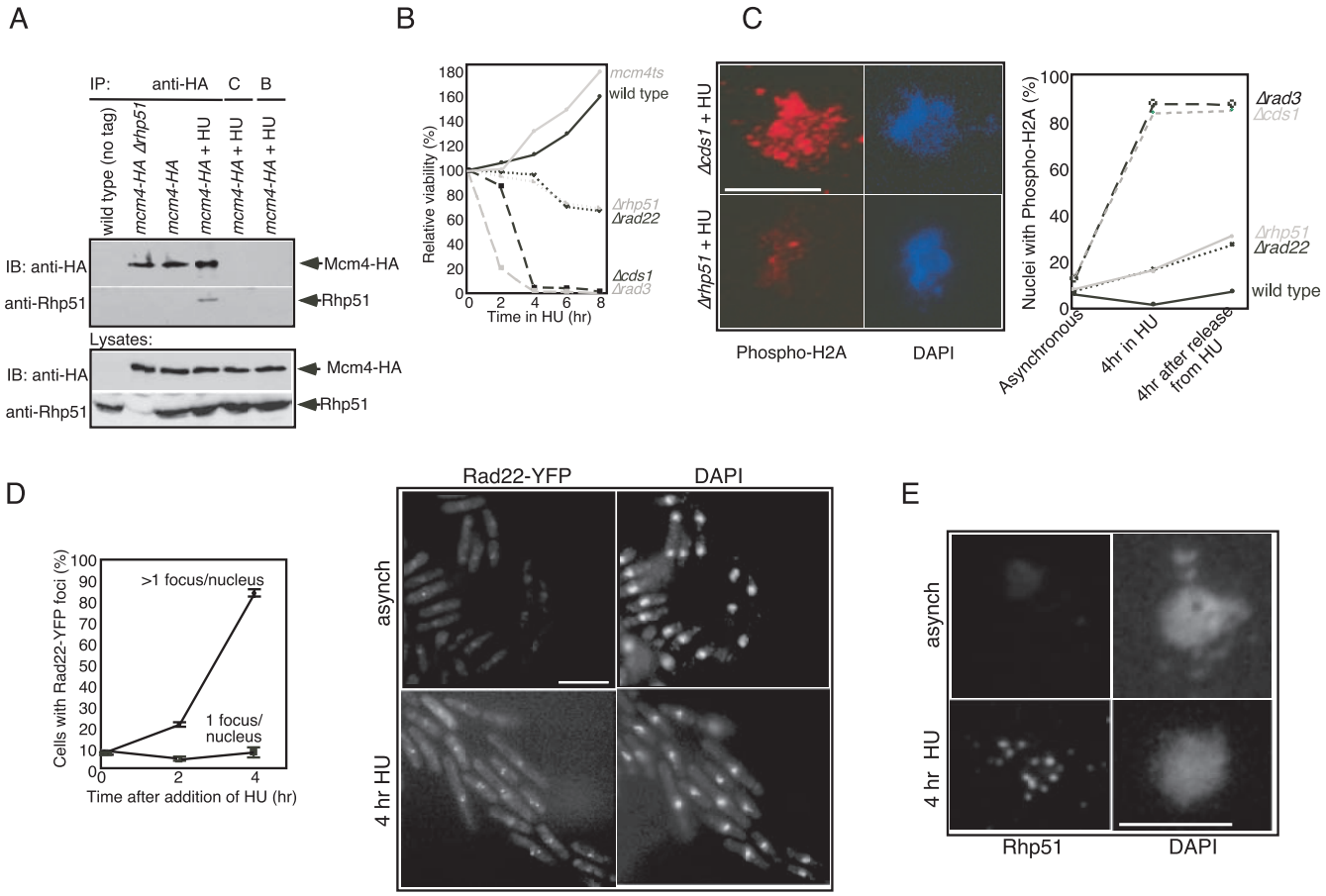

FIG. 4. Recombination proteins in the HU response (A) Mcm4 coprecipitates Rhp51 in HU. Protein lysates prepared from wild-type (FY72), *mcm4-HA* Δ*rhp51* (FY2356), and *mcm4-HA* (FY1167) cells untreated or treated with 15 mM HU for 4 h (+HU) were immunoprecipitated (IP) with anti-HA antibodies, an irrelevant control antibody (lane C), or protein A beads only (lane B). Immunoprecipitates and whole-cell lysates were analyzed by 10% SDS-PAGE, followed by immunoblotting (IB) with antibodies to detect Mcm4-HA (anti-HA) and Rhp51 (anti-Rhp51). (B) HR mutants display a modest loss of viability in HU compared to checkpoint mutants. Wild-type (FY72), *mcm4ts* (FY364), Δ*cds1* (FY865), Δ*rad3* (FY1106), Δ*rad22* (FY1209), and Δ*rhp51* (FY1884) strains were grown in liquid cultures at the permissive temperature (25°C) in the presence of 15 mM HU. Aliquots of cells were plated to YES medium at the indicated time points and incubated at 25°C for 4 days; relative viability was calculated as the number of colonies at each time point relative to the number at the zero time point. (C) HR mutants generate fewer phospho-H2A damage foci than checkpoint mutants, suggesting that most replication forks are stabilized. Asynchronous cultures of wild-type (FY72), Δ*cds1* (FY865), Δ*rad3* (FY1106), Δ*rad22* (FY1209), and Δ*rhp51* (FY1884) cells grown at 25°C were treated with 15 mM HU for 4 h and then released into fresh medium without HU. At the indicated time points, spread nuclei were prepared and immunostained for phospho-H2A. Phospho-H2A focus formation in the Δ*rad3* mutant is due to activity of the Tel1 protein kinase, which acts redundantly with Rad3 (54, 83) (see Fig. S1 in the supplemental material). (D) Rad22-YFP foci form in cells recovering from HU-induced arrest. Asynchronous (asynch) cultures of wild-type (FY72) and Rad22-YFP (FY2878) cells were incubated with 20 mM HU for 4 h at 25°C. At 0, 2, and 4 h after the addition of HU, Rad22-YFP localization was examined and quantitated in live cells (FY2878). (E) Rhp51 localization was assayed by anti-Rhp51 immunostaining of spread nuclei prepared from wild-type cells (strain FY72) treated with HU for 0, 2, or 4 h. The percentage spread nuclei that contained Rhp51 nuclear foci is indicated in the text. DNA was counterstained with DAPI. Scale bar, 10 μM.

Δ*rad3* and Δ*cds1* mutant cells, respectively, displayed the "cell untimely torn" (cut) phenotype, where the cells divided before replication was completed. Cell division was confirmed by immunostaining for tubulin to detect the mitotic spindles (data not shown).

In contrast, the Δ*rhp51* and Δ*rad22* cells elongated and arrested in HU and completed bulk DNA synthesis upon release from HU (Fig. 5B), consistent with the findings of a recent report (48). These results indicate that the recombination mutant cells were competent to activate the checkpoint as well as to resume replication during recovery. However, we found that approximately 25% of HR mutant cells missegregated chromosomes in the subsequent mitosis (Fig. 5B). In contrast to the checkpoint mutants, the main type of chromosome missegregation in HR mutants was unequal division of chromosomes or chromatids (82% of Δ*rhp51* cells and 79% of Δ*rad22* cells), suggesting a defect in spindle attachment, sister chromatid cohesion, or segregation. Interestingly, Meister et al. (48) reported that the deletion of *rhp51*$^+$ allowed some Δ*cds1* cells to complete bulk DNA replication, suggesting that one function of HR proteins during replication arrest and recovery is the resolution of nonproductive DNA structures that form as a consequence of replication fork stalling and collapse. We observed that Δ*cds1* Δ*rhp51* double mutants were heterogeneous in colony size. During acute HU exposure and release, the double mutant was as inviable as the Δ*cds1* single mutant

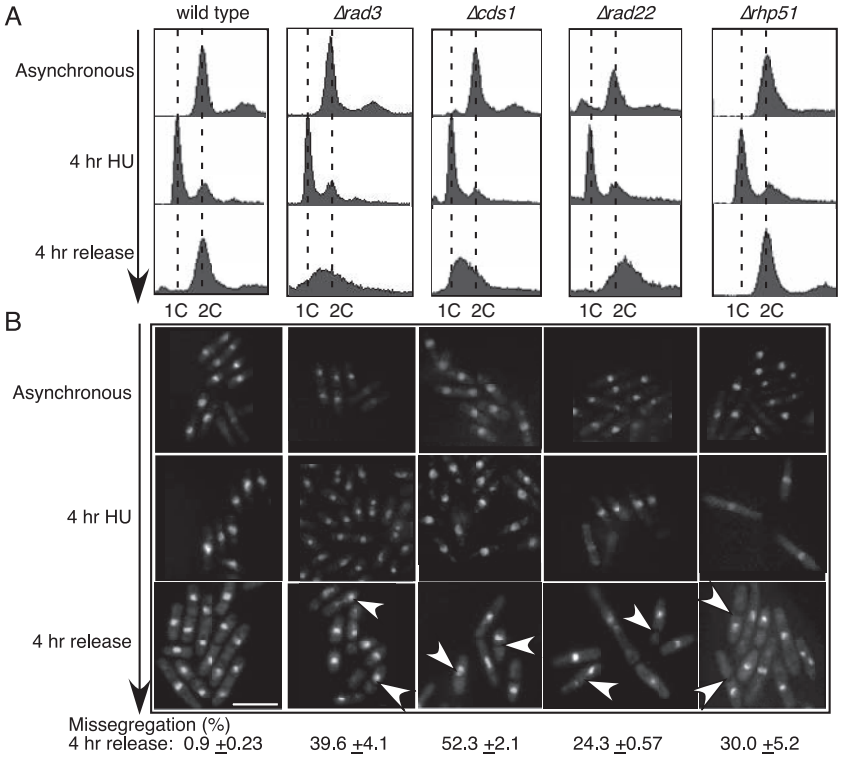

FIG. 5. HR proteins have a role in recovery from HU distinct from that in the checkpoint. Asynchronous cultures of wild-type (FY72), Δ*cds1* (FY865), Δ*rad3* (FY1106), Δ*rad22* (FY1209), and Δ*rhp51* (FY1884) cells grown at 25°C were treated with 15 mM HU for 4 h, then released into fresh medium without HU, and grown for another 4 h at 25°C. Cells were analyzed for (A) DNA content by flow cytometry and (B) nuclear morphology by staining with Sytox green. Arrowheads indicate examples of chromosome missegregation events (detailed in the text); averages and standard deviations were calculated from three independent experiments. Scale bar, 10 μM.

alone, but during chronic exposure, the double mutant was more sensitive to HU, indicating that continued lack of *rhp51* reduces the fitness of the Δ*cds1* mutant, consistent with independent effects on recovery (see Fig. S4 in the supplemental material).

Down-regulation of human MCM proteins by siRNA results in DNA damage and cell death. MCM proteins are highly overexpressed in many human cancers and have also been implicated more directly in cancer etiology (63, 77). We therefore tested whether MCM proteins promote S-phase genome stability in human cells by using siRNA to reduce the levels of MCM4 and MCM7 proteins in HeLa cells (Fig. 6A). Cells that were mock transfected continued to grow and divide, with a 265.0% ± 60.2% increase in cell viability by 48 h. In contrast, cells transfected with MCM siRNA exhibited a modest loss of viability, with just 41.6% ± 19.3% (MCM4) or 53.2% ± 13.1% (MCM7) of the cells remaining viable after 48 h. In experiments where MCM protein levels were not as strongly reduced, cell growth was still inhibited relative to that of cells treated with either mock siRNA or nontargeting control siRNA, although the overall viability of the population of cells did not decrease (data not shown). This result is consistent with observations in yeast that cells with a modest reduction in MCM protein levels still undergo bulk DNA replication, whereas complete depletion of MCM proteins is needed to prevent DNA synthesis (24, 41). In our experiments, MCM4 or MCM7 siRNA did not substantially alter the distribution of G_1, S, and G_2/M cells as analyzed by flow cytometry (Fig. 6B), although a higher proportion of cells displayed abnormal nuclear morphology, such as multinucleate cells or micronuclei (about 30% of cells treated with MCM4 siRNA or MCM7 siRNA compared to about 10% of mock-treated cells) (Fig. 6C). We also found that the siRNA oligonucleotides used to reduce MCM7 levels also partially decreased levels of MCM4, suggesting that loss of function of one of the MCM subunits may affect the stability of other components of the MCM complex. We conclude that the loss of MCM function in HeLa cells, as in yeast, leads to changes in chromosome structure or cell cycle progression that result in cell inviability.

To determine whether the loss of viability resulted from DNA damage, we assessed the focus formation of phosphorylated histone H2AX (phospho-H2AX), the mammalian counterpart to yeast phospho-H2A (64). While just 3.0% ± 1.0% of mock-treated cells contained phospho-H2AX foci after 48 h (Fig. 6D), 78.0% ± 2.8% of cells treated with MCM4 siRNA and 75.0% ± 3.0% of cells treated with MCM7 siRNA contained phospho-H2AX foci (Fig. 6D). To investigate whether this DNA damage localized to replication sites, we pulse-labeled siRNA-treated cells with BrdU, prepared stretched chromatin fibers, and immunostained for Rad51 (the Rhp51 homologue) and BrdU (Fig. 6E). We found that 67% of BrdU-stained chromatin fibers from cells treated with MCM siRNA also displayed Rad51 foci and that the Rad51 foci overlapped with BrdU foci. In contrast, just 8% of BrdU-stained chroma-

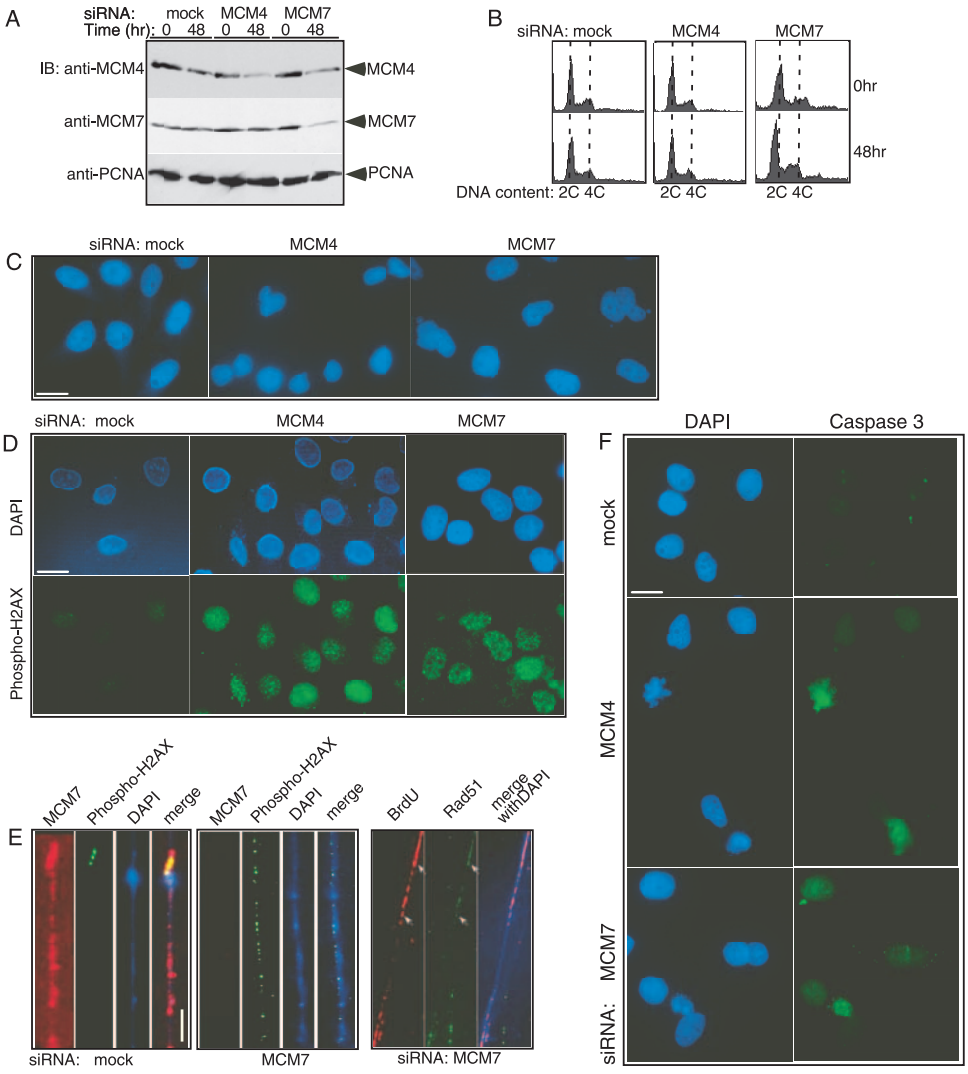

FIG. 6. siRNA against MCM4 or MCM7 in HeLa cells generates DNA damage and cell inviability. HeLa cells were transfected with buffer only (mock) or with siRNA against MCM4 or MCM7, and the cells were analyzed at 0 and 48 h after transfection. (A) Proteins were analyzed by 10% SDS-PAGE, followed by immunoblotting (IB) with antibodies to detect MCM4 or MCM7, and subsequently immunoblotted with anti-PCNA antibody as a loading control. (B) DNA content was assessed by flow cytometry. HeLa cells mock treated or treated with MCM4 siRNA or MCM7 siRNA for 48 h were stained with (C) DAPI to visualize the DNA, and (D) phospho-H2AX and DAPI to detect DNA damage. (E) The DNA damage correlates with loss of MCM localization and corresponds to sites of new DNA synthesis. Stretched chromatin fibers prepared from HeLa cells 48 h after siRNA treatment were immunostained for MCM7 and phospho-H2AX and counterstained with DAPI (left); chromatin fibers prepared from cells pulse-labeled with BrdU for 1 h were immunostained for BrdU and Rad51 and counterstained with DAPI. Scale bar, 10 μm (approximately 20 kb). (F) MCM siRNA results in cell death by apoptosis. Cells mock treated or treated with siRNA against MCM4 or MCM7 were immunostained for cleaved caspase 3 and counterstained with DAPI. Scale bars, 10 μm.

tin fibers from mock-treated cells also contained Rad51 foci. We suggest that as in yeast, the loss of MCM function in HeLa cells is associated with DNA breaks at sites of DNA synthesis. Moreover, the DNA damage resulted in cell death by apoptosis, as cells treated with siRNA against MCM4 or MCM7, but not mock-treated cells, stained for cleaved caspase 3, a marker for activation of the apoptotic pathway (Fig. 6F). We observed similar phenotypes of DNA damage and cell death by using a pool of four additional siRNA duplexes directed against either MCM4 or MCM7, but not with a nontargeting siRNA duplex (data not shown), indicating that the phenotypes we observed are due to specific depletion of these proteins. These findings support a role for the MCM proteins in S-phase genome stability that is conserved from yeast to human cells.

Interactions between MCM proteins and Rad51 in human cells. In fission yeast cells treated with HU, MCM proteins interact with the Rad51 homologue Rhp51 (Fig. 5) and this association may be important for cell survival of replication arrest. To test whether this interaction is conserved in human cells, we examined whether endogenous MCM proteins can coimmunoprecipitate with Rad51 in HeLa cells treated with HU. We found that the interaction between MCM proteins and Rad51 in HeLa cells was not specific to cells treated with HU; it also occurred in asynchronous cultures, as reported recently (70), as well as in cultures of cells released synchro-

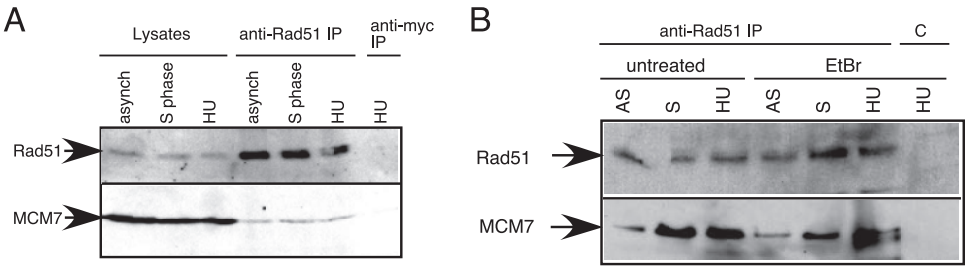

FIG. 7. MCM proteins interact with Rad51 in HeLa cells. (A) Whole-cell lysates prepared from asynchronous cultures of HeLa cells (asynch), cells blocked in 2 mM thymidine for 18 h and then released into S phase for 3 h (S phase), and cells blocked in early S phase with 2.5 mM HU for 18 h (HU) were immunoprecipitated with anti-Rad51 antibodies or with an irrelevant, control antibody (anti-myc). The anti-myc immunoprecipitation (IP) was performed on cells blocked in HU. Rad51 and associated proteins were analyzed by 10% SDS-PAGE, followed by immunoblotting to detect Rad51 and Mcm7. (B) Coimmunoprecipitation of MCM7 with Rad51 from HeLa cells is not dependent on DNA. Anti-Rad51 antibodies immunoprecipitate Rad51 and MCM7 from both untreated HeLa cell lysates and lysates that were treated with ethidium bromide (EtBr). AS, asynchronous cells; S, cells released from thymidine block into S phase; HU, cells arrested in hydroxyurea. C, control IP (anti-myc).

nously into S phase from a thymidine block (Fig. 7A). Furthermore, the association between Rad51 and Mcm7 was still observed when immunoprecipitations were carried out in the presence of ethidium bromide (Fig. 7B) or DNase (data not shown), indicating that the interaction was due to protein-to-protein associations and not mediated by DNA. However, while we observed nuclear MCM7 localization in S-phase cells both in the absence and the presence of HU, Rad51 nuclear foci were found primarily in S-phase cells that had been treated with HU (data not shown). Thus, while MCM proteins interact with Rad51/Rhp51 in both human and yeast cells, the interaction in human cells does not appear to be driven by activation of the replication checkpoint. However, it is possible that the interaction between MCM proteins and Rad51 in untreated cells might reflect a role in response to spontaneous S-phase DNA damage or stalled replication forks that may occur more frequently in the human genome.

DISCUSSION

MCM proteins promote replication fork stabilization and recovery. In response to DNA damage or blocks to DNA replication, checkpoint activation promotes replication fork arrest and also stabilizes fork structure so that the leading and lagging DNA strands and associated proteins remain coupled (12, 15, 34, 46). In this study, we analyze the role of the MCM complex in maintaining genome stability both during normal S phase and in response to replication fork stalling and checkpoint activation caused by treatment with HU. Using fission yeast as a model system, we provide evidence for distinct steps in the recovery process. Once the replication checkpoint is activated, MCM proteins are required to maintain the integrity of arrested replication forks. MCM proteins are also needed for cells to resume synthesis following S-phase arrest. We and others have identified a distinct role for recombination proteins in the recovery process (this work; 1, 48). Here we show that recombination is important for accurate chromosome segregation in the subsequent mitosis, and we suggest that recombination repairs DNA breaks that form spontaneously during replication fork arrest and restart. Thus, the checkpoint, MCM activity, and recombination define separable events that contribute to S-phase genome stability and the survival of replication arrest.

Fission yeast *mcm-ts* mutants undergo lethal Chk1-dependent arrest even though cells appear to complete bulk DNA synthesis (23, 40, 41, 47). Previously, we showed that wild-type MCMs are delocalized from the chromatin and lost from the nucleus when *mcm-ts* mutants are shifted to the restrictive temperature (61). Here, we demonstrate that the inactivation of *mcm-ts* also generates extensive DNA breaks during S phase identified by extensive phospho-H2A staining, and we observed similar results for human cells where MCM function is inhibited by siRNA. That the breaks are linked to active DNA synthesis and replication forks is suggested because the breaks are not observed in a *mcm4ts-degron* mutant that blocks replication elongation and because the DNA damage foci are found in regions of newly synthesized DNA. Thus, we propose that the loss of MCM function causes replication fork breakdown that generates DNA DSBs. This damage may be limited to certain regions of the genome or to regions of replication termination, and genome-wide analysis will be required to distinguish where it occurs.

Interestingly, while the DNA damage in *mcm-ts* mutants activates a DNA damage checkpoint and results in the recruitment of recombination proteins, such as Rhp51 and Rad22, to nuclear foci, the DNA breaks apparently remain unrepaired. In contrast, in prokaryotes, broken replication forks are a substrate for repair by recombination, which reassembles the replisome at the fork and promotes restart of synthesis (reviewed in references 27 and 50). One possibility is that the amount of DNA damage generated in *mcm-ts* mutants is extensive and may exceed the repair capacity of the cell. In support of this model, we observed that the overproduction of Rad22 in *mcm-ts* cells improved cell viability; however, the overproduction of Rad22 also caused a delay in cell cycle progression, suggesting that fewer DNA breaks were formed at any given time (J. M. Bailis and S. L. Forsburg, unpublished data).

MCM proteins have checkpoint-dependent and -independent roles in genome stability. We also find that MCM proteins contribute to fork stability and repair during replication arrest induced by HU. While cells blocked in HU display little

or no DNA DSBs, as assessed by phospho-H2A focus formation, extensive DNA breakage results when MCM function is subsequently lost from stalled forks using the *mcm4ts-degron* allele. This result indicates that the MCM complex is required to maintain fork stability in HU arrest, even after the replication checkpoint kinase Cds1 has been activated. Unexpectedly, we found that the Mcm4ts (nondegron) protein retained its function following HU arrest, correlating with Cds1-dependent phosphorylation. In contrast, the *mcm4ts-degron* strain is HU sensitive even at the permissive temperature and shows reduced modification of the Mcm4ts-degron protein. Together, these data raise the possibility that Mcm4 is a Cds1 substrate. Two motifs for Cds1 family kinases have been identified, RXX-S/T (58) and S/T-F (71), and both motifs exist in the N terminus of SpMcm4. In vertebrates, MCM subunits are substrates of the checkpoint protein kinases ATR and ATM (17, 33, 67, 82). Consistent with a role for checkpoint kinases in regulating MCM proteins, in *S. cerevisiae rad53* or *mrc1* mutants, MCM proteins are delocalized from the replication fork in HU (16). Taken together, these data suggest the MCM complex is a proximal effector of the Cds1 checkpoint, is required for replication fork stability, and may be a direct Cds1 substrate.

An alternative possibility is that *mcm4ts* mutants recover from HU because all replication origins have already fired. In this case, Mcm4 might normally regulate the timing of replication origin firing and no additional origins would need to initiate after the cells are released from HU. However, preliminary data suggest that some origins that normally fire early do not fire in *mcm4ts* (D. D. Luche and S. L. Forsburg, unpublished data). Moreover, when the number of active origins is reduced in budding yeast, such as by mutation of the Cdc7 kinase or MCM subunits, cells can still complete S phase, although more slowly (9, 19, 28). In addition, our comparison of *mcm4ts* mutants and the *mcm4-degron* mutant shifted to the restrictive temperature suggests that the damage does not occur during replication initiation (Fig. 1).

In many respects, *mcm-ts* mutants resemble Δ*cds1* mutants treated with HU: both result in the formation of phospho-H2A foci and the loss of viability, suggesting that both cause replication fork collapse (Fig. 1 and 4) (11, 48, 57). However, there are important differences. Although Δ*cds1* mutants treated with HU activate Chk1 and the DNA damage checkpoint, cell cycle arrest is transient (11) and the cells undergo a lethal mitosis without completing S phase (11, 48, 57) (Fig. 4). Moreover, Δ*cds1* cells in HU retain the capacity to resume synthesis, because at least some cells can complete bulk DNA replication if *rhp51*$^+$ is deleted (48). In contrast, *mcm-ts* mutants undergo Chk1-dependent arrest that is irreversible (23, 41, 47) and the arrest is not suppressed by the deletion of *rhp51*$^+$ or *rad22*$^+$ (J. M. Bailis and S. L. Forsburg, unpublished data). These data suggest that while Cds1 activity is essential to protect the replication fork, possibly in part by stabilizing the MCM proteins, MCM proteins may play additional downstream roles in recovery and reentry into the cell cycle that are Cds1 independent.

Interactions between MCM proteins and recombination during recovery. We found that fission yeast MCM proteins interact with Rhp51 specifically during HU arrest, while human MCM proteins associate with Rad51 both in the presence and the absence of HU (70; this work), suggesting that one such additional role in recovery may be the recruitment of recombination factors to forks that break during recovery. An analysis of Δ*rhp51* or Δ*rad22* mutants following release from HU suggests that cells are proficient to resume DNA synthesis, but defective in chromosome segregation. While the precise defect has not been determined, this result is reminiscent of those of recent studies of budding yeast that suggest recombination proteins, such as the MRN complex and Rad52, are required to tether broken ends together to facilitate proper segregation in mitosis (35, 45) and suggests that the role of HR proteins in HU recovery may be a structural function linked to replication fork restart that is required for proper chromosome segregation. In support of this idea, *S. cerevisiae* Rad51 is needed to maintain the association of DNA polymerase ε with the replisome during arrest (6).

However, recombination could also promote recovery by resolving DNA intermediates. While Sp Δ*rhp51* mutants complete bulk DNA replication after release from HU, as assessed by flow cytometry (48; this work), the chromosomes fail to enter pulsed-field gels (52), suggesting that they contain nonlinear DNA intermediates. While at least some of these intermediate structures are associated with DSBs (this work), they do not activate a G_2/M damage checkpoint. Although a similar phenotype was reported for other recombination mutants, such as *rad60* (52) and *smc6* (1), those mutants may not have activated the checkpoint because the cells arrested prior to mitosis. Thus, while the DNA damage in HR mutants does not activate a checkpoint leading to G_2/M-phase arrest, the breaks may cause uneven separation and segregation of the DNA during mitosis. It remains to be determined whether the role of HR proteins in recovery represents a requirement for active recombination or a structural role in maintaining chromosome integrity as observed in DSB repair mutants in budding yeast (35, 45). Alternatively, DNA breaks in Δ*rhp51* cells may not be in a chromatin context that can be recognized by the checkpoint. This model is supported by evidence that some Δ*rhp51* Δ*cds1* cells can restart replication and suggests that recombination proteins may function both to repair damaged forks and to resolve DNA intermediates formed when forks stall (reviewed in references 10 and 38).

The number of Rhp51 and Rad22 foci that we observe in cells treated with HU is far fewer than the predicted number of active forks after early firing of origins in HU (21), consistent with the possibility that recombination proteins only associate with or accumulate at a subset of replication forks or that damaged forks are assembled into repair centers as suggested for multiple DSBs in *S. cerevisiae* (43). We suggest that the interaction of Rhp51 with MCM proteins during replication fork arrest and recovery may be important to maintain fork stability during replication restart. However, in *mcm-ts* cells, the recruitment of Rhp51 to DNA breaks may instead reflect an MCM-independent attempt to repair forks that have already broken but cannot be processed appropriately. Thus, we speculate there are two forms of HR protein recruitment during S phase: fork recovery, via MCM proteins, and DSB repair, which is MCM independent.

MCM proteins may also contribute to an analogous function as Swi1 and Swi3, which have been proposed to form a fork protection complex (57). We observed that fission yeast *mcm2ts* is synthetically lethal with mutants of *swi1* (W. P. Dolan and S. L. Forsburg, unpublished data). In *S. cerevisiae*,

MCM proteins physically associate with the Swi1 homologue, Tof1 (55), and with Mrc1, a Cds1-adaptor protein which localizes to replication forks and also functions as a replication processivity factor (55, 59, 74, 79). In HU-treated Δ*mrc1* or Δ*tof1* cells, replication forks do not collapse, but the MCM helicase becomes separated from the polymerase on the chromatin as the leading and lagging DNA strands become uncoupled (16, 34, 78). Although Mrc1 contributes to checkpoint activation, the replisome decoupling in *mrc1* mutants is distinct from the irreparable fork collapse of *rad53* (or Sp *cds1*) checkpoint mutants (16, 79). Because MCM proteins interact with both Mrc1 and Cds1/Rad53, we suggest that the MCM complex contributes to S-phase genome stability at multiple levels.

The dysregulation of checkpoint, replication, and repair pathways is strongly linked with the capacity for proliferation and development of cancer (reviewed in references 25 and 80). The MCM complex has recently been used as a molecular marker for malignant cells and has also been proposed as a target for cancer therapy (reviewed in reference 25). The central role of the MCMs is further supported by the recent identification of an allele of Mcm4 linked to mouse mammary adenocarcinoma (68). Additionally, overexpression of the Mcm7 subunit alone can accelerate oncogene-induced tumorigenesis (30). It will be important to determine the mechanisms by which interactions between the MCM, checkpoint, and recombination proteins modulate fork arrest and restart with repair and to regulate whether S-phase arrest is survivable or lethal to cells.

ACKNOWLEDGMENTS

We thank Greg Freyer, Wolf Heyer, Stephen Kearsey, Albert Pastink, Paul Russell, Jill Meisenhelder, Andrew Dillin, and Beth Sullivan for reagents, protocols, and use of equipment. Elyse Rodgers-Vieira, Ji-Ping Yuan, and Suzanne Simon provided technical assistance and support. We thank past and present members of the Forsburg and Hunter laboratories, Walter Eckhart, and Matthew Weitzman for discussion and critical reading of the manuscript.

J.M.B. was supported by the Damon Runyon Cancer Research Foundation (fellowship DRG-1634) and the National Institutes of Health (training grant CA009370). T.H. is a Frank and Else Schilling American Cancer Society Research Professor. S.L.F. was a Stohlman Scholar of the Leukemia and Lymphoma Society. This work was supported by Public Health Service grants CA14195 and CA80100 from the National Cancer Institute (T.H.), DOD W81XWH-05-1-0391 (X. Chen and S.L.F., co-principal investigators), and GM059321 from the National Institutes of General Medicine (S.L.F.).

REFERENCES

1. **Ampatzidou, E., A. Irmisch, M. J. O'Connell, and J. M. Murray.** 2006. Smc5/6 is required for repair at collapsed replication forks. Mol. Cell. Biol. **26:**9387–9401.
2. **Bailis, J. M., P. Bernard, R. Antonelli, R. Allshire, and S. L. Forsburg.** 2003. Hsk1/Dfp1 is required for heterochromatin-mediated cohesion at centromeres. Nat. Cell Biol. 1111–1116.
3. **Bailis, J. M., and S. L. Forsburg.** 2004. MCM proteins: DNA damage, mutagenesis, repair. Curr. Opin. Gen. Dev. **14:**17–21.
4. **Bailis, J. M., A. V. Smith, and G. S. Roeder.** 2000. Bypass of a meiotic checkpoint by overproduction of meiotic chromosomal proteins. Mol. Cell. Biol. **13:**4838–4848.
5. **Bartkova, J., Z. Horejsi, K. Koed, A. Kramer, F. Tort, K. Zieger, P. Guldberg, M. Sehested, J. M. Nesland, C. Lukas, T. Orntoft, J. Lukas, and J. Bartek.** 2005. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature **434:**864–870.
6. **Bjergbaek, L., J. A. Cobb, M. Tsai-Pflugfelder, and S. M. Gasser.** 2005. Mechanistically distinct roles for Sgs1p in checkpoint activation and replication fork maintenance. EMBO J. **24:**405–417.
7. **Blower, M. D., B. A. Sullivan, and G. H. Karpen.** 2002. Conserved organization of centromeric chromatin in flies and humans. Dev. Cell **2:**319–330.
8. **Boddy, M. N., B. Furnari, O. Mondesert, and P. Russell.** 1998. Replication checkpoint enforced by kinases cds1 and chk1. Science **280:**909–912.
9. **Bousset, K., and J. F. X. Diffley.** 1998. The Cdc7 protein kinase is required for origin firing during S phase. Genes Dev. **12:**480–490.
10. **Branzei, D., and M. Foiani.** 2007. Interplay of replication checkpoints and repair proteins at stalled replication forks. DNA Repair **6:**994–1003.
11. **Brondello, J. M., M. N. Boddy, B. Furnari, and P. Russell.** 1999. Basis for the checkpoint signal specificity that regulates Chk1 and Cds1 protein kinases. Mol. Cell. Biol. **19:**4262–4269.
12. **Byun, T. S., M. Pacek, M. C. Yee, J. C. Walter, and K. A. Cimprich.** 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes Dev. **19:**1040–1052.
13. **Caspari, T., J. M. Murray, and A. M. Carr.** 2002. Cdc2-cyclin B kinase activity links Crb2 amd Rqh1-topoisomerase III. Genes Dev. **16:**1195–1208.
14. **Catlett, M. G., and S. L. Forsburg.** 2003. Schizosaccharomyces pombe Rdh54 (TID1) acts with Rhp54 (RAD54) to repair meiotic double-strand breaks. Mol. Biol. Cell **14:**4707–4720.
15. **Cobb, J. A., L. Bjergbaek, K. Shimada, C. Frei, and S. M. Gasser.** 2003. DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1. EMBO J. **22:**4325–4336.
16. **Cobb, J. A., T. Schleker, V. Rojas, L. Bjergbaek, J. A. Tercero, and S. M. Gasser.** 2005. Replisome instability, fork collapse, and gross chromosomal rearrangements arise synergistically from Mec1 kinase and RecQ helicase mutations. Genes Dev. **19:**3055–3069.
17. **Cortez, D., G. Glick, and S. J. Elledge.** 2004. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. Proc. Natl. Acad. Sci. USA **101:**10078–10083.
18. **Coxon, A., K. Maundrell, and S. E. Kearsey.** 1992. Fission yeast cdc21$^+$ belongs to a family of proteins involved in an early step of chromosome replication. Nucleic Acids Res. **20:**5571–5577.
19. **Donaldson, A. D., W. L. Fangman, and B. J. Brewer.** 1998. Cdc7 is required throughout the yeast S phase to activate replication origins. Genes Dev. **12:**491–501.
20. **Du, L. L., T. M. Nakamura, B. A. Moser, and P. Russell.** 2003. Retention but not recruitment of Crb2 at double-strand breaks requires Rad1 and Rad3 complexes. Mol. Cell. Biol. **23:**6150–6158.
21. **Feng, W., D. Collingwood, M. E. Boeck, L. A. Fox, G. M. Alvino, W. L. Fangman, M. K. Raghuraman, and B. J. Brewer.** 2006. Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. Nat. Cell Biol. **8:**148–155.

21a. **Foiani, M., F. Marini, D. Gamba, G. Lucchini, and P. Plevani.** 1994. The B subunit of the DNA polymerase alpha-primase complex in *Saccharomyces cerevisiae* executes an essential function at the initial stage of DNA replication. Mol. Cell. Biol. **14:**923–933.

22. **Forsburg, S. L.** 2004. Eukaryotic MCM proteins: beyond replication initiation. Microbiol. Mol. Biol. Rev. **68:**109–131.
23. **Forsburg, S. L., and P. Nurse.** 1994. The fission yeast *cdc19*$^+$ gene encodes a member of the MCM family of replication proteins. J. Cell Sci. **107:**2779–2788.
24. **Forsburg, S. L., D. A. Sherman, S. Ottilie, J. R. Yasuda, and J. A. Hodson.** 1997. Mutational analysis of Cdc19p, a *Schizosaccharomyces pombe* MCM protein. Genetics **147:**1025–1041.
25. **Gonzalez, M. A., K. E. Tachibana, R. A. Laskey, and N. Coleman.** 2005. Innovation: control of DNA replication and its potential clinical exploitation. Nat. Rev. Cancer **5:**135–141.
26. **Gorgoulis, V. G., L. V. Vassiliou, P. Karakaidos, P. Zacharatos, A. Kotsinas, T. Liloglou, M. Venere, R. A. J. Ditullio, N. G. Kastrinakis, B. Levy, D. Kletsas, A. Yoneta, M. Herlyn, C. Kittas, and T. D. Halazonetis.** 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. Nature **434:**907–913.
27. **Heller, R. C., and K. J. Marians.** 2006. Replisome assembly and the direct restart of stalled replication forks. Nat. Rev. Mol. Cell Biol. **7:**932–943.
28. **Hoang, M. L., R. P. Leon, L. Pessoa-Brandao, S. Hunt, M. K. Raghuraman, W. L. Fangman, B. J. Brewer, and R. A. Sclafani.** 27 August 2007. Structural changes in Mcm5 protein bypass Cdc7-Dbf4 function and reduce replication origin efficiency in *Saccharomyces cerevisiae*. Mol. Cell. Biol. doi:10.1128/MCB.00997-07.
29. **Hodson, J. A., J. M. Bailis, and S. L. Forsburg.** 2003. Efficient labeling of fission yeast Schizosaccharomyces pombe with thymidine and BUdR. Nucleic Acids Res. **31:**e134.
30. **Honeycutt, K. A., Z. Chen, M. I. Koster, M. Miers, J. Nuchtern, J. Hicks, D. R. Roop, and J. M. Shohet.** 2006. Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. Oncogene **25:**4027–4032.
31. **Huang, H. K., J. M. Bailis, J. D. Leverson, E. B. Gomez, S. L. Forsburg, and T. Hunter.** 2005. Suppressors of Bir1p (Survivin) identify roles for the chromosomal passenger protein Pic1p (INCENP) and the replication initiation factor Psf2p in chromosome segregation. Mol. Cell. Biol. **25:**9000–9015.
32. **Hyrien, O., K. Marheineke, and A. Goldar.** 2003. Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. Bioessays **25:**116–125.
33. **Ishimi, Y., Y. Komamura-Kohno, H. J. Kwon, K. Yamada, and M. Nakanishi.**

2003. Identification of MCM4 as a target of the DNA replication block checkpoint system. J. Biol. Chem. **278:**24644–24650.
34. **Katou, Y., Y. Kanoh, M. Bando, H. Noguchi, H. Tanaka, T. Ashikari, K. Sugimoto, and K. Shirahige.** 2003. S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. Nature **424:**1078–1083.
35. **Kaye, J. A., J. A. Melo, S. K. Cheung, M. B. Vaze, J. E. Haber, and D. P. Toczyski.** 2004. DNA breaks promote genomic instability by impeding proper chromosome segregation. Curr. Biol. **14:**2096–2106.
36. **Labib, K., J. A. Tercero, and J. F. X. Diffley.** 2000. Uninterrupted MCM2-7 function required for DNA replication fork progression. Science **288:**1643–1647.
37. **Lambert, S., and A. M. Carr.** 2005. Checkpoint responses to replication fork barriers. Biochimie **87:**591–602.
38. **Lambert, S., B. Froget, and A. M. Carr.** 2007. Arrested replication fork processing: interplay between checkpoints and recombination. DNA Repair **6:**1042–1061.
39. **Lau, E., T. Tsuji, L. Guo, S. H. Lu, and W. Jiang.** 2007. The role of pre-replicative complex (pre-RC) components in oncogenesis. FASEB J. **21:**3786–3794.
40. **Liang, D. T., and S. L. Forsburg.** 2001. Characterization of *S. pombe mcm7*$^+$ and *cdc23*$^+$ (*MCM10*) and interactions with replication checkpoints. Genetics **159:**471–486.
41. **Liang, D. T., J. A. Hodson, and S. L. Forsburg.** 1999. Reduced dosage of a single fission yeast MCM protein causes genetic instability and S phase delay. J. Cell Sci. **112:**559–567.
42. **Lindner, K., J. Gregan, S. Montgomery, and S. Kearsey.** 2002. Essential role of MCM proteins in pre-meiotic DNA replication. Mol. Biol. Cell **13:**435–444.
43. **Lisby, M., U. H. Mortensen, and R. Rothstein.** 2003. Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat. Cell Biol. **5:**572–577.
44. **Lisby, M., R. Rothstein, and U. H. Mortensen.** 2001. Rad52 forms DNA repair and recombination centers during S phase. Proc. Natl. Acad. Sci. USA **98:**8276–8282.
45. **Lobachev, K., E. Vitriol, J. Stemple, M. A. Resnick, and K. Bloom.** 2004. Chromosome fragmentation after induction of a double-strand break is an active process prevented by the RMX repair complex. Curr. Biol. **14:**2107–2112.
46. **Lucca, C., F. Vanoli, C. Cotta-Ramusino, A. Pellicioli, G. Liberi, J. Haber, and M. Foiani.** 2004. Checkpoint-mediated control of replisome-fork association and signalling in response to replication pausing. Oncogene **23:**1206–1213.
47. **Maiorano, D., G. Blom van Assendelft, and S. E. Kearsey.** 1996. Fission yeast cdc21, a member of the MCM protein family, is required for onset of S phase and located in the nucleus throughout the cell cycle. EMBO J. **15:**861–872.
48. **Meister, P., A. Taddei, L. Vernis, M. Poidevin, S. M. Gasser, and G. Baldacci.** 2005. Temporal separation of replication and recombination requires the intra-S checkpoint. J. Cell Biol. **168:**537–544.
49. **Michalet, X., R. Ekong, F. Fougerousse, S. Rousseaux, C. Schurra, N. Hornigold, M. van Slegtenhorst, J. Wolfe, S. Povey, J. S. Beckmann, and A. Bensimon.** 1997. Dynamic molecular combing: stretching the whole human genome for high-resolution studies. Science **277:**1518–1523.
50. **Michel, B., H. Boubakri, Z. Baharoglu, M. LeMasson, and R. Lestini.** 2007. Recombination proteins and rescue of arrested replication forks. DNA Repair **6:**967–980.
51. **Michel, B., S. D. Ehrlich, and M. Uzest.** 1997. DNA double strand breaks caused by replication arrest. EMBO J. **16:**430–438.
52. **Miyabe, I., T. Morishita, T. Hishida, S. Yonei, and H. Shinagawa.** 2006. Rhp51-dependent recombination intermediates that do not generate checkpoint signal are accumulated in *Schizosaccharomyces pombe rad60* and *smc5/6* mutants after release from replication arrest. Mol. Cell. Biol. **26:**343–353.
53. **Moreno, S., A. Klar, and P. Nurse.** 1991. Molecular genetic analysis of the fission yeast *Schizosaccharomyces pombe*. Methods Enzymol. **194:**795–823.
54. **Nakamura, T. M., L. L. Du, C. Redon, and P. Russell.** 2004. Histone H2A phosphorylation controls Crb2 recruitment at DNA breaks, maintains checkpoint arrest, and influences DNA repair in fission yeast. Mol. Cell. Biol. **24:**6215–6230.
55. **Nedelcheva, M. N., A. Roguev, L. B. Dolapchiev, A. Shevchenko, H. B. Taskov, A. F. Stewart, and S. S. Stoynov.** 2005. Uncoupling of unwinding from DNA synthesis implies regulation of MCM helicase by Tof1/Mrc1/Csm3 checkpoint complex. J. Mol. Biol. **347:**509–521.
56. **Noguchi, E., C. Noguchi, L. L. Du, and P. Russell.** 2003. Swi1 prevents replication fork collapse and controls checkpoint kinase Cds1. Mol. Cell. Biol. **23:**7861–7874.
57. **Noguchi, E., C. Noguchi, W. H. McDonald, J. R. Yates III, and P. Russell.** 2004. Swi1 and Swi3 are components of a replication fork protection complex in fission yeast. Mol. Cell. Biol. **24:**8342–8355.
58. **O'Neill, T., L. Giarratani, P. Chen, L. Iyer, C. H. Lee, M. Bobiak, F. Kanai, B. B. Zhou, J. H. Chung, and G. A. Rathbun.** 2002. Determination of substrate motifs for human Chk1 and hCds1/Chk2 by the oriented peptide library approach. J. Biol. Chem. **277:**16102–16115.
59. **Osborn, A. J., and S. J. Elledge.** 2003. Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53. Genes Dev. **17:**1755–1767.
60. **Osman, F., M. Adriance, and S. McCready.** 2000. The genetic control of spontaneous and UV-induced mitotic intrachromosomal recombination in the fission yeast *Schizosaccharomyces pombe*. Curr. Genet. **38:**113–125.
61. **Pasion, S. G., and S. L. Forsburg.** 1999. Nuclear localization of Schizosaccharomyces pombe Mcm2/Cdc19p requires MCM complex assembly. Mol. Biol. Cell **10:**4043–4057.
62. **Paulsen, R. D., and K. A. Cimprich.** 2007. The ATR pathway: fine-tuning the fork. DNA Repair **6:**953–966.
63. **Ren, B., G. Yu, G. Tseng, K. Cieply, T. Gavel, J. Nelson, G. Michalopoulos, Y. P. Yu, and J.-H. Luo.** 2005. MCM7 amplification and overexpression are associated with prostate cancer progression. Oncogene **9:**1090–1098.
64. **Rogakou, E. P., C. Boon, C. Redon, and W. M. Bonner.** 1999. Megabase chromatin domains involved in DNA double-strand breaks in vivo. J. Cell Biol. **146:**905–916.
65. **Rosenberg, C., R. J. Florijn, F. M. Van de Rijke, L. A. Blonden, T. K. Raap, G. J. Van Ommen, and J. T. Den Dunnen.** 1995. High resolution DNA fiber-fish on yeast artificial chromosomes: direct visualization of DNA replication. Nat. Genet. **10:**477–479.
66. **Sherman, D. A., S. G. Pasion, and S. L. Forsburg.** 1998. Multiple domains of fission yeast Cdc19p (MCM2) are required for its association with the core MCM complex. Mol. Biol. Cell **9:**1833–1845.
67. **Shi, Y., G. E. Dodson, P. S. Mukhopadhyay, N. P. Shanware, A. T. Trinh, and R. S. Tibbetts.** 2007. Identification of carboxyl-terminal MCM3 phosphorylation sites using polyreactive phosphospecific antibodies. J. Biol. Chem. **282:**9236–9243.
68. **Shima, N., A. Alcaraz, I. Liachko, T. R. Buske, C. A. Andrews, R. J. Munroe, S. A. Hartford, B. K. Tye, and J. C. Schimenti.** 2007. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. Nat. Genet. **39:**93–98.
69. **Shroff, R., A. Arbel-Eden, D. Pilch, G. Ira, W. M. Bonner, J. H. Petrini, J. E. Haber, and M. Lichten.** 2004. Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break. Curr. Biol. **14:**1703–1711.
70. **Shukla, A., V. M. Navadgi, K. Mallikarjuna, and B. J. Rao.** 2005. Interaction of hRad51 and hRad52 with MCM complex: a cross-talk between recombination and replication proteins. Biochem. Biophys. Res. Commun. **329:**1240–1245.
71. **Smolka, M. B., C. P. Albuquerque, S. H. Chen, and H. Zhou.** 2007. Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases. Proc. Natl. Acad. Sci. USA **104:**10364–10369.
72. **Snaith, H. A., G. Brown, and S. L. Forsburg.** 2000. *Schizosaccharomyces pombe* Hsk1p is a potential Cds1p target required for genome integrity. Mol. Cell. Biol. **20:**7922–7932.
73. **Sogo, J. M., M. Lopes, and M. Foiani.** 2002. Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. Science **297:**599–602.
74. **Szyjka, S. J., C. J. Viggiani, and O. M. Aparicio.** 2005. Mrc1 is required for normal progression of replication forks throughout chromatin in *S. cerevisiae*. Mol. Cell **19:**691–697.
75. **Takeda, T., K. Ogino, K. Tatebayashi, H. Ikeda, K. Arai, and H. Masai.** 2001. Regulation of initiation of S phase, replication checkpoint signaling, and maintenance of mitotic chromosome structures during S phase by Hsk1 kinase in the fission yeast. Mol. Biol. Cell **12:**1257–1274.
76. **Tashiro, S., J. Walter, A. Shinohara, N. Kamada, and T. Cremer.** 2000. Rad51 accumulation at sites of DNA damage and in postreplicative chromatin. J. Cell Biol. **150:**283–291.
77. **Todorov, I. T., B. A. Werness, H. Q. Wang, L. N. Buddharaju, P. D. Todorova, H. K. Slocum, J. S. Brooks, and J. A. Huberman.** 1998. HsMCM2/BM28: a novel proliferation marker for human tumors and normal tissues. Lab. Investig. **78:**73–78.
78. **Tourriere, H., and P. Pasero.** 2007. Maintenance of fork integrity at damaged DNA and natural pause sites. DNA Repair **6:**900–913.
79. **Tourriere, H., G. Versini, V. Cordon-Preciado, C. Alabert, and P. Pasero.** 2005. Mrc1 and Tof1 promote replication fork progression and recovery independently of Rad53. Mol. Cell **19:**699–706.
80. **van Gent, D. C., J. H. Hoeijmakers, and R. Kanaar.** 2001. Chromosomal stability and the DNA double-stranded break connection. Nat. Rev. Genet. **2:**196–206.
81. **Waseem, N. H., K. Labib, P. Nurse, and D. P. Lane.** 1992. Isolation and analysis of the fission yeast gene encoding polymerase δ accessory protein PCNA. EMBO J. **11:**5111–5120.
82. **Yoo, H. Y., A. Shevchenko, A. Shevchenko, and W. G. Dunphy.** 2004. Mcm2 is a direct substrate of ATM and ATR during DNA damage and DNA replication checkpoint responses. J. Biol. Chem. **279:**53353–53364.
83. **You, Z., C. Chahwan, J. Bailis, T. Hunter, and P. Russell.** 2005. ATM activation and its recruitment to damaged DNA require binding to the C terminus of Nbs1. Mol. Cell. Biol. **25:**5363–5379.
84. **Zolezzi, F., J. Fuss, S. Uzawa, and S. Linn.** 2002. Characterization of a Schizosaccharomyces pombe strain deleted for a sequence homologue of the human damaged DNA binding 1 (DDB1) gene. J. Biol. Chem. **277:**41183–41191.

REPRINT ORDER FORM — GENETICS
(ALL CURRENCY IN U.S. DOLLARS)

REFERENCE: **GENETICS** MONTH: _____ Article #: _____

ALWAYS USE OUR ORDER FORM to list your requirements and specifications. Remember to supply a photocopy to your purchasing department, and keep a photocopy for your files.

DO NOT DELAY ORDERING YOUR REPRINTS: Return this form even if no reprints are desired.

DELIVERY AND INVOICES: Invoices are mailed at the time of shipment of reprints, normally within 4 weeks after the issue mailing.

Reprints for delivery to U.S. destinations are routinely shipped by Parcel Post, and the shipping cost is included in the reprint price.

Reprints for delivery to foreign destinations are shipped by surface mail (10–12 weeks delivery time), or you may choose expedited air shipment. In both cases, the cost of shipping to foreign destinations is charged to the author and will be billed on the reprint invoice.

REFERENCE INFORMATION MUST APPEAR ON ALL CHECKS AND DRAFTS.

AUTHOR(S): _____

TITLE OF ARTICLE: _____

Please return this form even if no reprints are ordered.

Order form should be returned with author-corrected proofs.

Split shipments will be charged an additional shipping and handling fee of $29.80 per additional address.

Non-members will be charged $75/100 reprints for color, while members will be charged $52/100.

No purchase order is needed with the order form.

BASIC REPRINT PRICES

No. of Reprints

PAGES	25	50	100	200	300	400	500	ADDL 100
1–4	$128	$137	$155	$192	$232	$266	$303	$30
5–8	247	256	282	334	390	445	498	54
9–12	355	373	408	479	547	619	693	78
13–16	468	490	535	625	711	795	884	102
17–20	572	599	654	764	869	973	1081	125
21–24	694	724	785	906	1029	1150	1277	150
25–28	844	876	937	1060	1209	1369	1501	178
29–32	922	963	1045	1209	1369	1535	1694	199

Canadian authors must add 7% GST to total basic reprints charges before publication charges are added. Orders are accepted in multiples of 100 copies only.

SHIPPING INSTRUCTIONS

NOTE: Shipping instructions must be completed on this form. Type or print all information below:

Ship: _____ copies

To: _____

_____

_____

_____

Phone or FAX: _____

Ship: _____ copies

To: _____

_____

_____

_____

Phone or FAX: _____

Ground Freight within the continental United States is included in "Basic Reprint Prices," except for orders in excess of 1000 copies.

PAGE CHARGES ESTIMATE

(you will be invoiced for exact amount)

Page charges will be billed at $65 per page for members, $80 per page for non-members.

Members of the Genetics Society of America will be charged $100 for each each color figure in an article, while non-members will be charged $400 for the first color figure in an article and $250 for each subsequent color figure.

REPRINT CHARGES

_____ Copies $ _____

Color: add $52/100 reprints (members) $ _____

$75/100 reprints (non-members) $ _____

Split Shipment Fee: $29.80 per additional address $ _____

SUBTOTAL (before GST) $ _____

Canada GST (7% for Canadian destinations only) $ _____

A. TOTAL ORDER REPRINTS $ _____

PUBLICATION CHARGES

Page Charges (if article consists of an odd number of pages, round up to the next even number)

$ _____ per page × _____ pages $ _____

Color Charges

Members $100 per color figure

× _____ figures $ _____

Non-members $400 for first color figure

plus _____ figures at $250 each $ _____

B. TOTAL PUBLICATION CHARGES $ _____

TOTAL CHARGES (A + B) $ _____

INSTRUCTIONS FOR REPRINT INVOICING

(Author must complete)

Send invoice to:

Name of Person _____

Institution _____

Department _____

Street _____

City _____ State _____ Zip _____

Country _____ Phone no. _____

SHIPPING TO FOREIGN DESTINATIONS

Authors pay for reprint shipments to foreign destinations. The cost of shipping will be included on the offprint invoice. Check one of the following:

☐ Send via surface mail (allow 10–12 weeks for delivery).

☐ Send via expedited air service.

Proofreading instructions

1. Mark all corrections in the margins of the master proof, clearly indicating where the insertion or deletion should be made.
2. Carefully check the reproduction, orientation, and sizing of figures; confirm consistency between figures and the respective legends. Check the location of all figures and tables. If a figure requires alteration, you must provide a quality hard copy and an electronic copy (TIF or EPS, 1000 dpi for line art and 500 for color or grayscale) of the revised figure with your corrected proof.
3. Pay particular attention to accuracy of references.
4. Answer all queries on the query page.
5. Corrections should be kept to a minimum. Authors are charged for any changes (other than corrections of printer or editor errors), including answers to queries regarding information that is incomplete or incorrect in the original manuscript. At this stage of publication, even the smallest change to the master proof will incur a significant fee. These alteration charges cannot be waived.
6. Corrected proofs must be faxed to (802) 244-1462 *within 48 hours or publication may be delayed.* Please return every page of the proof and query page. Printing of late changes cannot be guaranteed; we can accept only *one* round of corrections. Please include your article number in *all* correspondence and on the top of proof page 1.

Reprint instructions

- Always use the Reprint Order Form to list your request and specifications; supply a copy to your institution's purchasing department. Please complete both the Shipping Instructions and Instructions for Reprint Invoicing, even if the information is the same. Also, please complete the Reprint Charges section, but leave the Publication Charges section blank.
- Do not delay ordering your reprints! Write "No reprints ordered" on the Reprint Order Form if you do not wish to receive any reprints.
- Fax the Reprint Order Form with your master proof to Dartmouth Journal Services (see fax number above) even if no reprints are requested. Please do not send payment with the Reprint Order Form; you will be billed after publication.
- Reprints are shipped approximately 3 weeks after publication of the journal. With questions concerning reprint orders only, please call (802) 244-1457, ext. 349.

Signature Required

Signature _____

Name (print) _____

By signing this form the author agrees to accept responsibility for the payment of the reprints and/or all charges described and tabulated on the Reprint Order Form. If such charges are to be billed to an institution, the author assumes the responsibility for making the necessary arrangements for the issuance of a formal institutional purchase order. Otherwise, it is understood that the author will bear the cost of this reprint order.

DOI: 10.1534/genetics.107.085779

Schizosaccharomyces pombe Histone Acetyltransferase Mst1 (KAT5) Is an Essential Protein Required for Damage Response and Chromosome Segregation

Eliana B. Gómez,*,† Rebecca L. Nugent,* Sebastián Laria† and Susan L. Forsburg*,†,[1]

*Molecular and Computational Biology Section, University of Southern California, Los Angeles, California 90089-2910 and †Molecular and Cell Biology Laboratory, Salk Institute for Biological Studies, La Jolla, California 90089-2910

Manuscript received December 12, 2007
Accepted for publication April 3, 2008

ABSTRACT

Schizosaccharomyces pombe Mst1 is a member of the MYST family of histone acetyltransferases and is the likely ortholog of *Saccharomyces cerevisiae* Esa1 and human Tip60 (KAT5). We have isolated a temperature-sensitive allele of this essential gene. *mst1* cells show a pleiotropic phenotype at the restrictive temperature. They are sensitive to a variety of DNA-damaging agents and to the spindle poison thiabendazole. *mst1* has an increased frequency of Rad22 repair foci, suggesting endogenous damage. Two-hybrid results show that Mst1 interacts with a number of proteins involved in chromosome integrity and centromere function, including the methyltransferase Skb1, the recombination mediator Rad22 (Sc Rad52), the chromatin assembly factor Hip1 (Sc Hir1), and the Msc1 protein related to a family of histone demethylases. *mst1* mutant sensitivity to hydroxyurea suggests a defect in recovery following HU arrest. We conclude that Mst1 plays essential roles in maintenance of genome stability and recovery from DNA damage.

MANY events in DNA metabolism, including replication and repair, are modulated by the pattern of different histone modifications in the chromatin, leading to the speculation that a "histone code" provides epigenetic information that facilitates numerous chromatin functions (reviewed in Jenuwein and Allis 2001; Kouzarides 2007). Histone acetylation is associated with diverse chromatin functions (reviewed in Kouzarides 2007; Lee and Workman 2007). Histone acetylation changes association of DNA with the underlying nucleosomes (Shogren-Knaak *et al.* 2006) and also creates specific binding sites for proteins involved in a variety of DNA transactions (reviewed in Kouzarides 2007; Lee and Workman 2007). Histone acetyltransferase (HAT) enzymes can be separated into several sequence-specific subfamilies that differ in their substrates and in the composition of their associated complexes. The MYST family of HATs (reviewed by Utley and Cote 2003) has been extensively characterized in budding yeast, which has three members (reviewed in Lafon *et al.* 2007). We have reported two MYST family proteins in fission yeast (Gómez *et al.* 2005): Mst2, the likely ortholog of *Saccharomyces cerevisiae* Sas2, is a nonessential gene required for normal telomere silencing (Gómez *et al.* 2005). Mst1, the likely ortholog of *S. cerevisiae* Esa1, is an essential gene and is the subject of this report.

[1]*Corresponding author:* Molecular and Computational Biology Section, 1050 Childs Way, RRI 201, University of Southern California, Los Angeles, CA 90089-2910. E-mail: forsburg@usc.edu

Work from several experimental systems shows that ScEsa1 histone acetyltransferase is a catalytic subunit of a large protein complex called NuA4 and a smaller subcomplex, Piccolo, that is part of NuA4 (Allard *et al.* 1999; Boudreault *et al.* 2003; Eberharter *et al.* 2005; Selleck *et al.* 2005). The NuA4 complex acetylates histone H4 and H2A (Smith *et al.* 1998; Allard *et al.* 1999; Galarneau *et al.* 2000). It may also be associated with acetylation of the histone variant H2AZ (Keogh *et al.* 2006; Millar *et al.* 2006). Esa1 family HATs contain a chromodomain, a motif associated with binding to methylated histones (reviewed in de la Cruz *et al.* 2005). While this suggests that some aspect of Esa1 function is mediated by binding methylated histone(s), the target for such binding remains elusive. The mammalian homolog is Tip60 (reviewed in Squatrito *et al.* 2006). Under a new initiative to regularize nomenclature in the chromatin field, Mst1, Esa1, and Tip60 are defined as the KAT5 family of histone acetyltransferases (Allis *et al.* 2007).

Sc*ESA1* is an essential gene, and loss of function results in cell-cycle-specific arrest after replication but before mitosis (Smith *et al.* 1998; Clarke *et al.* 1999), although a recent report suggests that it may also be required for cell cycle entry (Early *et al.* 2004). The roles of *Sc*Esa1 are diverse, and recent studies have also linked it to the spindle assembly checkpoint (Le Masson *et al.* 2003) and the morphogenesis checkpoint (Le Masson *et al.* 2003; Ruault and Pillus 2006). Genome-level analyses in budding yeast indicate that *Sc*Esa1 activity is associated with active gene expression

(Galarneau *et al.* 2000; Vogelauer *et al.* 2002). Paradoxically, however, H4 acetylation is also associated with decreased gene expression (Deckert and Struhl 2001). Recent studies suggest that *Sc*Esa1 may also repress activity of specific gene regions, particularly the heterochromatic domains of the telomeres and the rDNA (Clarke *et al.* 2006).

Histone acetylation by the NuA4 complex has a role in DNA double-strand break (DSB) repair (reviewed in van Attikum and Gasser 2005b; Squatrito *et al.* 2006). Increased histone acetylation near the damage site occurs (Yu *et al.* 2005) and acetylation of H4 is linked to activation of homologous recombination and nonhomologous end-joining (Bird *et al.* 2002; Tamburini and Tyler 2005). *Sc*Esa1 is recruited to DNA breaks (Bird *et al.* 2002; Tamburini and Tyler 2005). One of the major responses to DNA double-strand breaks is phosphorylation of the histone variant H2AX by the damage checkpoint kinases ATM and ATR (reviewed in Foster and Downs 2005; Morrison and Shen 2005). H2AX phosphorylation and recruitment of NuA4 occur with similar timing (Downs *et al.* 2004). Interestingly, two chromatin-remodeling complexes, INO80 and SWR1, are also recruited to H2AX, at least in part through the common Arp4 subunit that they share with NuA4 (Downs *et al.* 2004; Robert *et al.* 2006). NuA4, INO80, and SWR1 are proposed to facilitate exchange of phospho H2AX, ultimately downregulating the damage signal (Morrison *et al.* 2004; Tsukuda *et al.* 2005; van Attikum and Gasser 2005b; Papamichos-Chronakis *et al.* 2006). Recent studies have implicated the NuA4 complex in regulation and modification of an additional histone H2A variant called H2AZ (Babiarz *et al.* 2006; Keogh *et al.* 2006; Millar and Grunstein 2006). It has been suggested that the downregulation of the H2AX phosphorylation signal reflects exchange of phospho-H2AX with H2AZ (reviewed in van Attikum and Gasser 2005a).

Tip60, the human ortholog of Mst1 and *Sc*Esa1, acetylates nonhistone proteins independently of the NuA4 complex with substrates including Notch (Kim *et al.* 2007), ATM (Sun *et al.* 2005; Jiang *et al.* 2006), p53 (Sykes *et al.* 2006; Tang *et al.* 2006), and DNA-PKcs (Jiang *et al.* 2006). Tip60-deficient cells exposed to ionizing radiation have defects in chromosome repair and reduced apoptosis (Ikura *et al.* 2000). This leads to the important conclusion that HATs are not limited to histones as substrates. These data suggest that KAT5 family HATs contribute to DNA damage response in two ways: first, by affecting DNA repair and recovery (*e.g.*, through histone modifications that recruit specific DNA damage response proteins) and, second, by affecting the checkpoint response (*e.g.*, through direct modification of p53 or ATM).

To gain further insight into the diverse function of the KAT5 protein family in DNA metabolism, we investigated the role of *Schizosaccharomyces pombe* Mst1 in the tractable fission yeast system. In this report, we present initial genetic characterization of this gene. Mst1 is essential for viability, and mutants are sensitive to a variety of DNA-damaging agents, including hydroxyurea (HU), ultraviolet (UV) irradiation, methyl methanesulfonate (MMS), and bleomycin. Mutants are also sensitive to the spindle poison thiabendazole. Using a two-hybrid screen, we identify a number of potential interaction proteins involved in DNA dynamics, including chromatin proteins, transcription factors, and the recombination protein Rad22 (ScRad52). Analysis of the HU response suggests that Mst1 is required for normal recruitment of Rad22 in HU. Thus, Mst1 has multiple roles in maintenance of genome integrity in *S. pombe*.

MATERIALS AND METHODS

Strains, media, and manipulations: Strains used in this study are listed in Table 1. Strains were grown and maintained on yeast extract plus supplements (YES) or Edinburgh minimal media with appropriate supplements, using standard techniques (Forsburg and Rhind 2006). Matings were performed on synthetic sporulation agar plates for 2–3 days at 25°. Temperature-sensitive (ts) cells were grown at 25° and non-ts cells at 32°. Transformations were carried out by electroporation. Double mutants were constructed by standard tetrad analysis or random spore analysis. G418 plates were YES supplemented with 100 μg/ml G418 (Sigma, St. Louis).

Plasmids and constructions: Disruption and tagging of *mst1+* is described in Gómez *et al.* (2005). Site-directed mutants were constructed using the Quick-Change mutagenesis kit (Stratagene, La Jolla, CA).

Primers used were L-S/F (CATACAATTTTCTTATGAATC GACCAAGCGTGAGCACAAAC) and L-S/R GTTTGTGCTCA CGCTTGGTCGATTCATAAGAAAATTGTATG and L-P/F GG TGTAGGAATATTTGTCTTCCGTCCAACCTTTTTCTGGATC AC and L-P/R GTGATCCAGAAAAAGTTTGGACGGAAGA CAAATATTCCTACACC. Complementation and temperature sensitivity were assessed by transforming the expression plasmid into the heterozygous *mst1*$^+$/Δ*mst::ura4+* diploid FY1754 and by isolating haploids that contain both Δ*mst1::ura4+* and the plasmid markers. An integrant was constructed by moving the *Pnmt-mst1-L344S* cassette from plasmid pEBG99 into plasmid pJK148 to create plasmid pEBG98 and integrating this construct at the *leu1-32* locus in the diploid. Haploid temperature-sensitive segregants containing both Δ*mst1::ura4+* and *leu1*$^+$*::nmt-mst1-L344S* were isolated by random spore analysis.

Silencing was assessed by examining growth in the absence of uracil and low adenine in a strain with *ura4*$^+$ integrated into the centromere and *ade6*$^+$ integrated into the telomere of a minichromosome.

FACS: Samples were collected after 5 and 10 hr, ethanol fixed, and analyzed by flow cytometry as described previously (Gómez and Forsburg 2003).

Immunoprecipitation: Samples were prepared as described in Gómez *et al.* (2002). For co-immunoprecipitation with MCM proteins, we used anti-HA antibody 12CA5 (Abcam) and antiV5 (Invitrogen, San Diego). For histone acetylation assays, the dried beads were treated with 20 μl of a histone acetylation mix containing 5 μM C14-Acetyl CoA (ICN), 0.7 μg of core histones, 50 mM HEPES (pH 7.9), 10% glycerol, and 1 mM DTT. The reaction was incubated 1 hr 30 min at 30°. Beads

TABLE 1

Strains used in this study

Strain	Genotype	Source
11	*h⁻ ade6-M210*	Our stock
261	*h⁻ ura4-D18 leu1-32 ade6-M210 can1-1*	Our stock
865	*h⁻ Δcds1::ura4 ura4-D18 leu1-32*	T. Wang
1104	*h⁻ Δrad3::ura4⁺ ura4 leu1-32 ade6-M210 (?)*	Our stock
1107	*h⁻ Δrad3::ura4⁺ ura4-D18 leu1-32 ade6-M216*	Our stock
1256	*h⁻ Δmad2::ura4⁺ ade6M210 leu1-32 ura4D18*	S. Sazer
1552	*h⁻ Δbub1::ura4⁺ ade6M210 leu1-32 ura4D18*	K. Hardwick
1754	*h⁺/h⁻ Δmst1::ura4⁺/⁺ ura4-D18/ura4-D18 leu1-32/leu1-32 ade6-M216/ade6-M210*	Gómez *et al.* (2005)
2180	*h⁺ cbh1::GFP-pBSIIKS-ura4⁺-chh1C leu1-32 ura4-D18*	Baum and Clarke (2000)
2192	*h⁻ Δchk1::ura4⁺ ura4-D18 ade6-M216*	Our stock
2217	*h⁻ Δcbh1::LEU2 ura4-D18 ade6-M216 leu1-32*	Y. Murakami
2339	*h⁺ Δmst1::ura4⁺ leu1⁺::mst1-L344S ura4-D18 ade6-M216 (#7)*	This study
2396	*h⁻ Δmst1::ura4⁺ leu1⁺::mst1-L344S ura4-D18 ade6-M210*	This study
2449	*h⁺ Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 ade6-M210*	This study
2450	*h⁻ Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 ade6-M210*	This study
2495	*h⁺ Δtaz1::ura4⁺ Δmst1::kanMX6 leu1⁺::mst1-L344Sura4-D18 leu1-32 ade6-M210*	This study
2498	*h⁻ Δrhp51::ura4⁺ Δmst1::kanMX6 leu1⁺::mst1-L344Sura4-D18 leu1-32 ade6-M210*	This study
2499	*h⁺ Δrad50::kanMX6 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ura4-D18 leu1-32 ade6-M21?*	This study
2501	*h⁺ Δrad32::ura4⁺ mst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 leu1-32 ade6-M21?*	This study
2505	*h⁺ mst2::ura4⁺ mst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 leu1-32 ade6-M210*	This study
2506	*h⁻ mst2::ura4⁺ mst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 leu1-32 ade6-M210*	This study
2508	*h⁺ swi6::ura4⁺ mst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 leu1-32 ade6-M210*	This study
2525	*h⁻ hsk1-1312 mst1::kanMX6 leu1⁺::mst1-L344S⁺ leu1-32 ade6-M21?*	This study
2535	*h⁺ Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18*	
2562	*h⁺ Δrad12::ura4⁺ Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18*	This study
2566	*h⁻ Δclr4::ura4⁺ Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18 (otrR(SphI)::ade6⁺) ?*	This study
2569	*h⁺ orp1-4 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18 (can1-1)?*	This study
2570	*h⁺ Δclr3::kanMX6 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18 (can1-1)?*	This study
2572	*h⁺ mcm7-98 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ ade6-M216 ura4-D18 (can1-1)?*	This study
2574	*h⁺ cdc19-P1 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18*	This study
2576	*h⁺ pol1-1 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18*	This study
2578	*h⁺ cdc21-M68 Δmst1::ura4⁺ leu1⁺::mst1-L344S⁺ ade6-M216 ura4-D18 (can1-1)?*	This study
2601	*h⁺ Δskb1::ura4⁺ mst1::kanMX6 leu1⁺::mst1-L344S⁺ ura4-D18 ade6-M210*	This study
2535	*h⁺ Δmst1::kanMX6 leu1⁺::mst1-L344S ade6-M210 ura4-D18*	This study
3135	*h⁻ ade6-M210 leu1-32 ura4-D1 Δhip::ura4⁺*	S. Whitehall
3286	*h⁻ Δcds1::ura4⁺ rad22YFP::KanMX6 ura4-D18 leu1-32 ade6?*	This study
3369	*h⁺ Δmst1::ura4⁺ leu1⁺::mst1-L344S ura4-D18 ade6-M21? rad22-YFP:kanMX4*	This study
3436	*h⁻ rad22-YFP:kanMX4 ade6-M210 ura4-D18 leu1-32*	This study
3479	*h⁺ Δmsc1::kanMX6 leu1-32*	N. Walworth
3480	*h⁺ msc1-HA-kanMX6 leu1-32*	N. Walworth
3553	*h? Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ otr1L(dh/HindIII)::ura4⁺ leu1-32 ura4ΔS/E ade6ΔN/N [ch16 M23::LEU2⁺ tel::ade6⁺]*	This study
3563	*h? Δcbh1::LEU2 Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ade6-M21? ura4-D18*	This study
3562	*h? Δmst1::kanMX6 leu1⁺::mst1-L344S⁺ ade6-M210 ura4-D18 Δhip::ura4⁺*	This study

were spun down, and supernatant was recovered and boiled in SDS–Laemmli loading buffer before electrophoresis in 15% SDS–polyacrylamide gel electrophoresis (PAGE). The gel was fixed and treated with fluorography-enhancing solution (NEN Life Science) before drying and autoradiography.

For co-immunoprecipitation with Rad22-YFP or Cbh1-GFP (strains 3436, 3480), we prepared soluble protein lysates from fission yeast cells in B88 buffer as in Moreno *et al.* (1991) A total of 1.0 mg protein was used per immunoprecipitation. Lysates were precleared with sepharose A beads (Repligen) for 1 hr. One microliter of anti-GFP (Abcam 290) was added and left to rotate overnight at 4°. Proteins were analyzed by 8% SDS–PAGE, followed by immunoblotting with anti-V5 (Invitrogen) at 1:2500 dilution and with anti-GFP (living colors, Clontech) at a 1:1000 dilution. Crosslinking immunoprecipitations were done as previously described (Auger *et al.* 2008). Briefly, cells were crosslinked with 1% formaldehyde (Sigma) for 50 min on ice. Harvested cells were vortexed with glass beads in IP buffer (50 mm HEPES–KOH, pH 1.6, 500 mm NaCL, 1 mm EDTA, 1% Triton X-100, 0.1% sodium deoxycholate). Extracts were treated with DNase I to reverse crosslink. A total of 1.0 mg protein was used per immunoprecipitation. Whole-cell extracts and immunoprecipitates were incubated with SDS sample buffer for 1 hr at 90° before Western blot analysis.

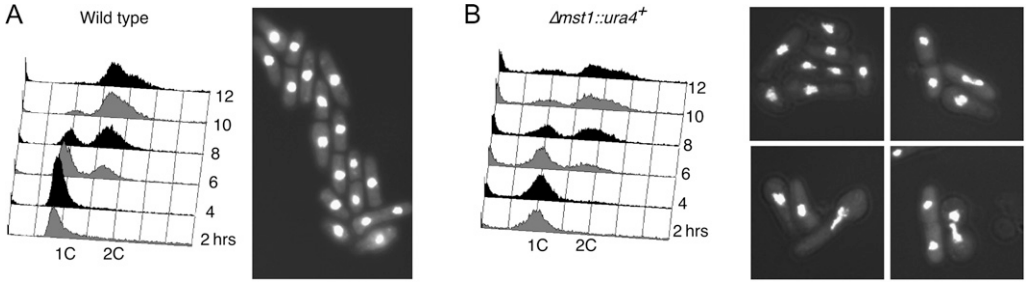

FIGURE 1.—Spore germination. (A) Wild-type diploid strain (*ura4*$^+$/*ura4-D18*) and (B) heterozygous disruption mutant (*mst1*$^+$/Δ*mst1::ura4*$^+$) were sporulated and processed for random spore analysis. Spores were inoculated into media lacking uracil and sampled for flow cytometry to monitor entry into S phase and the process of DNA replication. Cell morphology was visualized by staining for DNA with DAPI.

Microscopy: Whole fixed cells were stained as described for DAPI and Calofluor (GÓMEZ and FORSBURG 2003). Cells were visualized with a Leica DMR microscope. Objectives used were Leica 100X/1.30, PLFL, NA = 1.30, and Leica 63X/1.32, PLApo, NA = 1.32. Images were captured with a Hamamatsu digital camera and Improvision Openlab software (Improvision, Lexington, MA). Images were assembled using Canvas (ACD/Deneba) and adjusted for contrast. Nuclear spreads were performed as in PANKRATZ and FORSBURG (2005) with the following modifications. Rad22yfp was detected using Abcam 290-50 rabbit anti-GFP, 1:2000 in 5% BSA (Sigma), and secondary donkey anti-rabbit cy3 1:500 DNA was stained with 1× Hoechst (Invitrogen). The objective used was Leica 63X/1.32, PLApo, NA = 1.32. Images were captured and assembled as above.

Damage assays: Cells were grown to $A_{595} = 1$, serially diluted fivefold, and spotted onto YES plates or YES plates containing 10 μg/ml thiabendazole (TBZ), 5 mM HU, or 0.005, 0.007, or 0.01% MMS, and incubated for 3–5 days at 25°, 29°, or 32°, as indicated. For UV irradiation assay, 1000 cells were plated in duplicate onto YES and immediately treated with the indicated doses of UV irradiation using a Stratalinker (Stratagene). Control plates were left untreated. Plates were incubated at 32° for 2–3 days. Colonies were counted and the percentage survival was obtained relative to the number of colonies grown on the untreated control plates.

Two-hybrid strains: The *mst1*+ open reading frame was isolated by PCR from plasmid pEBG72 with primers HAT-3 forward (AGATCTCGCCAAAAACTTGCTCAATCTTCTTCC) and HAT-3 reverse (GGTACCCGACATAGACTGGCACATTA CATCC) and cloned into the two-hybrid vector pDBLeu cleaved with *Nco*I and *Not*I to create plasmid pEBG91. Budding-yeast two-hybrid strain AH109 (Clontech) with *HIS3*, *ADE2*, and *lacZ* reporters downstream of heterologous *GAL4*-responsive promoter elements was transformed with an *S. pombe* Matchmaker cDNA library (Clontech) and screened for interactions as described in LEVERSON *et al.* (2002). Plasmids from positive clones were isolated and verified by retransformation, and the identity of the interacting open reading frame was determined by sequence.

RESULTS

Δ*mst1* cells undergo disordered mitosis: Previously, we identified two members of the MYST family of histone acetyltransferases in fission yeast. Our analysis indicated that Mst1 is the likely ortholog of *S. cerevisiae* Esa1, and like *Sc*Esa1, is a chromatin-bound protein that is essential for viability (GÓMEZ *et al.* 2005). Data from metazoans suggest that the Hbo1 histone acetyltransferase, which is also a MYST family protein, may function in DNA replication (IIZUKA and STILLMAN 1999; BURKE *et al.* 2001; IIZUKA *et al.* 2006). To investigate whether *S. pombe* Mst1 contributes to DNA replication, we examined the phenotype of cells containing the Δ*mst1* allele. Mutant diploids heterozygous for Δ*mst1::ura4*$^+$ and wild-type diploids heterozygous for *ura4*$^+$ were induced to sporulate. We inoculated purified spores from both diploids into medium lacking uracil to carry out bulk spore germination. Only spores containing the *ura4*$^+$ gene are able to germinate in these conditions, and their DNA content was determined by flow cytometry. Spores containing Δ*mst1::ura4*$^+$ synthesized DNA to an approximately 2C DNA content, with timing similar to that of wild type (Figure 1). Half of each population contains nongerminating *ura4-D18* spores, which persist as a 1C population. Upon microscopic examination, we observed that many of the Δ*mst1::ura4*+ cells showed highly disordered mitosis, including a high fraction of *cut* phenotypes. This suggests defects in chromosome segregation, which might result from mitotic defects or from replication defects and checkpoint disruption.

Isolation and characterization of a temperature-sensitive allele: To better characterize Mst1, we constructed a temperature-sensitive allele based on temperature-sensitive alleles in *S. cerevisiae ESA1* (CLARKE *et al.* 1999). We tested two mutations: L271P (corresponding to L254P in *Sc*Esa1) and L344S (corresponding to L267S in *Sc*Esa1). Constructs containing these were placed in a *LEU2* plasmid under control of the *nmt1* promoter and transformed into the Δ*mst1::ura4*$^+$ diploid. Following sporulation, we assessed recovery of spores at 25° containing both the *ura4*$^+$ marker (marking Δ*mst1*) and the *LEU2* marker (marking the plasmid). Both mutants were able to complement the disruption at 25°. However, upon shift to 36°, only the *mst1-L344S* strain was temperature sensitive.

We attempted to replace the endogenous *mst1*$^+$ allele with the *L344S* allele, without success. We reasoned that, when integrated under the native promoter, a single

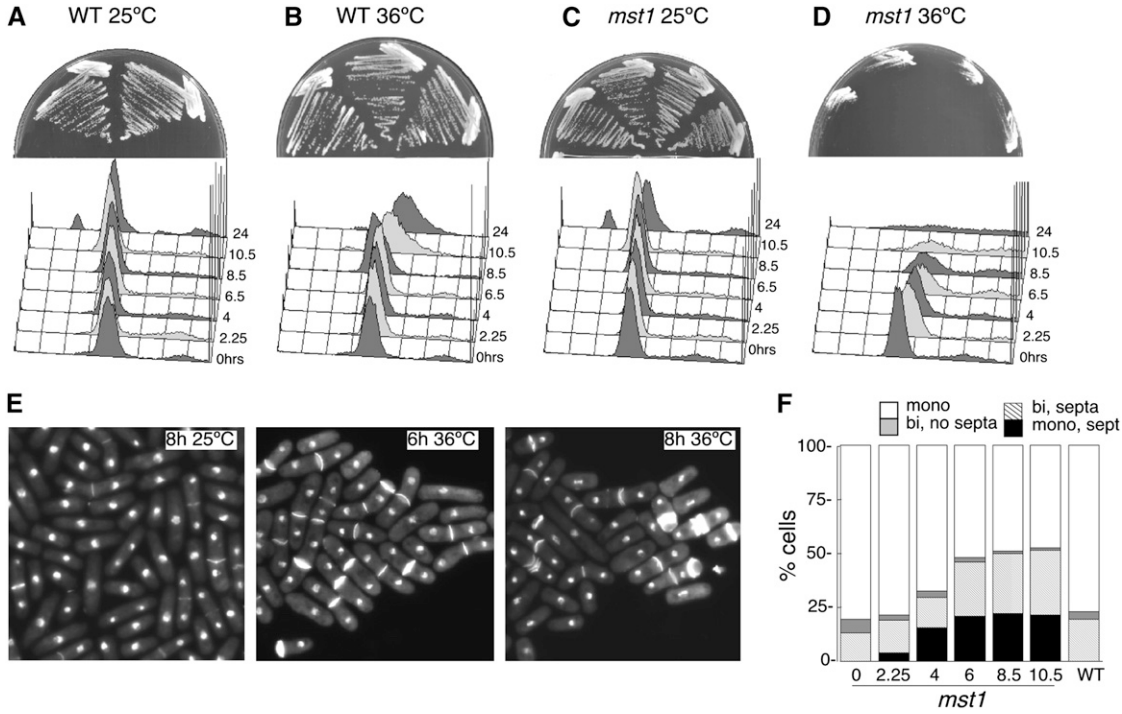

FIGURE 2.—Characterization of temperature-sensitive allele. (A–D) Comparison of growth rate and FACS for wild-type (FY11) and *mst1-L344S* cells (FY 2396) grown in the presence of thiamine at 25° or 36°. (Top) Colony formation on plates, where colony size is proportional to growth rate. (Bottom) FACS analysis following shift to 36°. (E) *mst1-L344S* cells at 25° (left) or 36° at 6 (middle) and 8 (right) hr after shift. Cells were stained with DAPI and Calcofluor. (F) Fraction of cells with indicated phenotypes during a temperature-shift experiment.

copy of *mst1-L344S* might be insufficient for viability. Therefore, we integrated *mst1-L344S* under control of the *nmt1*$^+$ promoter into the *leu1-32* locus of the heterozygous diploid and sporulated to create the haploid strain Δ*mst1::ura4*$^+$ *leu1*$^+$*::P*$_{nmt1}$*-mst1-L344S*. Upon shift to 36° and addition of thiamine to shut off the *nmt1* promoter, this strain was unable to grow. However, cells were able to grow in the presence of thiamine at 25° or 32°, indicating that the protein is functional at lower temperatures. There is sufficient background expression of the *nmt1*$^+$ promoter even the presence of thiamine for stable proteins to be maintained (Figure 2, A–D). An equivalent strain with a *kanMX* marker at Δ*mst1* behaved similarly (data not shown).

We examined the phenotype of cells containing *mst1-L344S* by shifting asynchronous cells to the restrictive temperature. The increase of cell number was blocked within 2 hr. However, cell mass measured by A_{595} continued to increase for ~8 hr. DNA content, measured by flow cytometry, showed no change from the 2C peak characteristic of exponentially growing wild-type fission yeast, although later time points showed a flattening of the profile and an increase of cells with higher DNA content. This is common in heterogeneous, elongated cells and is typical of most *cdc* mutants.

We examined cells for nuclear and septum morphology by staining with DAPI and Calcofluor, respectively, and observing under fluorescence (Figure 2E, quantified in Figure 2F). Strikingly, the number of cells staining with Calcofluor significantly increased in the *mst1-L344S* strain at restrictive temperature, with nearly 50% of cells arrested with septa by 6 hr. Some of these were mononucleate, showing asymmetric nuclear segregation. Additionally, there was an increase in the overall number of binucleate cells relative to wild type. *S. pombe* has an unusual cell cycle in which cytokinesis is completed during the subsequent nuclear cell cycle. Thus, cells with a 2C DNA content by FACS may be mononucleate cells in G_2 or M phase or binucleated cells in which the nuclei are in G_1 or S phase and cytokinesis is not complete. Therefore, arrest as a binucleate septated cell may represent a defect entering DNA replication in the ensuing cell cycle. Alternatively, it could represent a defect in cytokinesis that delays entry into the next cell cycle.

This phenotype is different from the disordered mitosis seen in the Δ*mst1* mutant spores. Because germinating spores arise from stationary or G_1 cells, while exponentially growing *S. pombe* are predominantly G_2 with a 2C DNA content, we reasoned that the different phenotypes could represent the effect of inactivating Mst1 at different cell cycle stages. Alternatively, they could suggest that the *mst1-L344S* allele creates a hypomorph rather than a complete inactivation of the gene. We attempted to resolve whether the mutation affects S phase by synchronizing the *mst1-L344S* cells in G_1,

TABLE 2

Genetic interactions between *mst1-L344S* and other mutants

Mutation	Protein affected	Double-mutant phenotype
	DNA replication	
mcm2ts	MCM helicase	Reduced growth temperature
mcm4ts	MCM helicase	Reduced growth temperature
mcm7ts	MCM helicase	Reduced growth temperature
orp1ts	ORC subunit	Reduced growth temperature
pol1ts	DNA polymerase α	Reduced growth temperature
hsk1ts	Cdc7 kinase	Reduced growth temperature
Δ*rad2*	FEN endonuclease	No effect
	DSB repair	
Δ*rad32*	MRE11	No effect
Δ*rad50*	RAD50	Reduced growth temperature
Δ*rhp51*	RAD51	Reduced growth temperature
Δ*rqh1*	SGS1/Blm helicase	No effect
Δ*rti1*	Rad52 homolog	No effect
	Centromere/heterochromatin	
Δ*cbh1*	CENP-B homolog	No effect
Δ*clr3*	Histone deacetylase	Reduced growth temperature
Δ*clr4*	Histone methyltransferase	Reduced growth temperature
Δ*hip1*	HIRA homolog	No effect
Δ*msc1*	Putative demethylase	No effect
Δ*mst2*	MYST acetyltransferase	No effect
Δ*swi6*	HP1 homolog	Reduced growth temperature
	Other	
Δ*skb1*	Arginine methyltransferase	No effect
Δ*taz1*	Telomere-binding protein	No effect

Double mutants were constructed and growth rate relative to the single mutants was assessed by streaking at different temperatures.

using nitrogen starvation (supplemental Figure 1). When released from nitrogen starvation into complete medium, *mst-L344S* cells completed replication with similar timing to wild type, also suggesting that inactivation of *mst1* by the *mst1-L344S* allele or by the Δ*mst1* allele is not sufficient for cells to block DNA replication in the first cell cycle.

***mst1-L344S* genetically interacts with replication, recombination, and heterochromatin mutants:** We examined the phenotypes of double mutants, which contained *mst1-L344S* as well as mutations in genes required for DNA replication, double-strand break repair, and heterochromatin assembly (Table 2, Figure 3). We tested the double-mutant strains for their temperature sensitivity. We defined a genetic interaction as occurring when a double-mutant strain was unable to grow at temperatures where both parents were growth proficient ("reduced growth temperature"). There were modest genetic interactions with many replication initiation mutants. Additionally, we observed that the maximum growth temperature was reduced when *mst1* was combined with mutations Δ*rhp51* (ScRad51), Δ*rad50*, and Δ*rad22*, which affect homologous DNA recombination. The Δ*rad22 mst1-L344S* strain grew extremely slowly even at 25°; however, because of the high frequency of spontaneous suppressors in Δ*rad22* backgrounds (Osman *et al.* 2005), we cannot be confident that this strain is suppressor free.

Because *mst1-L344S* cells show defects in chromosome segregation and are sensitive to thiabendazole (see Figure 5), we examined interactions with genes required for centromere heterochromatin assembly (Table 2, chromatin). We observed that maximum-growth temperature was strikingly lower than that of the parents when *mst1-L344S* was combined with Δ*swi6*, Δ*clr3*, and Δ*clr4*, all of which define genes required to establish and maintain centromeric heterochromatin (reviewed in Ekwall 2007). However, we observed no defects in centromere or telomere silencing in *mst1-L344S* cells (data not shown). This suggests that the heterochromatin itself is intact. We examined additional chromatin-related proteins. No effect is observed when *mst1-L344S* is combined with mutation of histone acetyltransferase Δ*mst2* or the telomere-binding protein Δ*taz1*.

Identification of potential Mst1 interaction partners by two-hybrid screening: SpMst1 interacts with the actin-related protein SpAlp5/ScArp4, a known subunit of the NuA4 complex (Minoda *et al.* 2005). Because evidence suggests that Tip60 may function independently of the NuA4 complex (see Introduction), we sought to identify additional proteins that interact with Mst1, using a two-hybrid screen. We isolated a number of interacting gene products, summarized in Table 3.

We isolated Rad22, the *S. pombe* ortholog of homologous DNA recombination protein Rad52. Rad22 promotes the strand-invasion step of homologous recombination. Fission yeast has an additional Rad22 family member, called Rti1. We constructed plasmids to test whether Mst1 interacted in the two-hybrid system with Rti1 or with Rhp51. Although Rhp51 and Rad22 themselves interact in two-hybrid screens (*e.g.*, Catlett and Forsburg 2003), we observed no evidence for interaction between Mst1 and Rhp51 or between Mst1 and Rti1, suggesting that this interaction is specific (Figure 4A). Immunoprecipitation of Rad22-YFP using a cross-reacting anti-GFP antibody precipitates Mst1-V5 (Figure 4B); however, immunoprecipitation of Mst1-V5 does not precipitate Rad22-YFP (data not shown).

We isolated three chromatin-related proteins as potential Mst1 partners. The first, Hip1, is a putative histone chaperone most closely related to ScHir1 and human

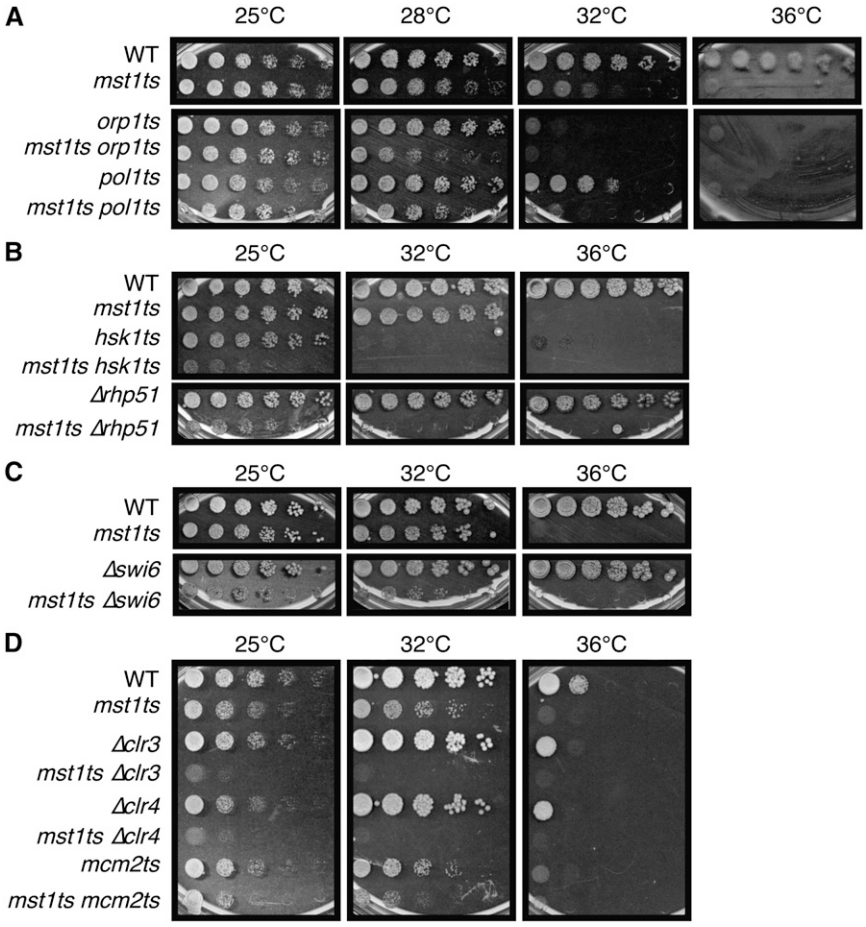

FIGURE 3.—Genetic interactions. Double mutants were constructed and serial dilutions were performed at different temperatures to characterize any synthetic interactions. Equivalent numbers of exponentially growing cells were diluted fivefold and incubated at the indicated temperatures.

HIRA (BLACKWELL *et al.* 2004). The second, Cbh1, is one of the three CenpB homologs in fission yeast associated with centromere assembly (BAUM and CLARKE 2000; NAKAGAWA *et al.* 2002). We tested all three CenpB homologs for interaction with Mst1. Only Cbh1 showed a specific interaction (Figure 4). We observed no co-immunoprecipitation between Cbh1 and Mst1 from soluble or crosslinked lysates (data not shown).The third gene that we isolated was Msc1. Msc1 was identified as a high-copy suppressor of the checkpoint mutation Δ*chk1*; it has E3 ligase activity and homology to a family of demethylases, and mutation disrupts histone acetylation (AHMED *et al.* 2004; DUL and WALWORTH 2007). We constructed double mutants containing *mst1-L344S* with Δ*hip1*, Δ*cbh1*, or Δ*msc1* and tested growth at different temperatures. There was no difference in the growth characteristics of the double mutants compared to *mst1-L344S* cells alone. We were unable to detect Msc1 in soluble extracts and our crosslinking protocol was unsuccessful on this protein (data not shown).

The most frequently isolated two-hybrid clone was *skb1*$^+$, which encodes an arginine methyltransferase, an ortholog of *S. cerevisiae* Hsl7 (FUJITA *et al.* 1999; SHULEWITZ *et al.* 1999; LEE *et al.* 2000). We observed no obvious genetic interactions between Δ*skb1* and *mst1-L344S*, in contrast to budding yeast where *hsl7* and *esa1* mutants have been shown to have synthetic interactions (RUAULT and PILLUS 2006). Two putative transcription factors were also isolated. The *res2/pct1*$^+$ gene has been studied extensively and it is required for G_1/S-specific transcription (TANAKA *et al.* 1992; ZHU *et al.* 1994). *taf111*$^+$ (also known as *taf1*$^+$) is a putative TFIID subunit required for induction of genes during nitrogen starvation and sexual development (UENO *et al.* 2001). The remaining two-hybrid isolates included *bud6*$^+$, which is also associated with actin function (GLYNN *et al.* 2001), and a *SEC18* homolog likely involved in vesicular transport (SPAC1834.11c). The only previously described fission yeast Mst1 interactor, Alp5 (MINODA *et al.* 2005), was not isolated in our screen, nor were any other subunits of the NuA4 complex.

We observed very little Mst1 (or Mst2) protein in the soluble fraction (supplemental Figure 2a). This may explain why our attempts to confirm the two-hybrid interactions using co-immunoprecipitation in fission yeast were for the most part unsuccessful. Upon overproduction of Mst1, it is possible to observe some interactions with other proteins; for example, we found that Mcm2 can interact with Mst1 or Mst2 but only when both partners are overproduced (supplemental Figure 2b). However, we observed no interactions with endogenous levels of MCM proteins (data not shown), even

TABLE 3

Genes isolated by two-hybrid screening

β-Galactosidase	Gene	Function
+	*bud6/fat1*$^+$	Actin-interacting protein (ScBUD6/AIP3 homolog) (Glynn *et al.* 2001)
+	*cbh1*$^+$	One of three fission yeast CENP-B proteins. Required for heterochromatin function. No *S. cerevisiae* homolog (Baum and Clarke 2000; Nakagawa *et al.* 2002).
+	*hip1*$^+$	HIRA histone chaperone (ScHIR1 homolog) (Blackwell *et al.* 2004)
++	*msc1*$^+$	High-copy suppressor of Δ*chk;* E3 ubiquitin ligase and putative demethylase (Ahmed *et al.* 2004; Dul and Walworth 2007). No *S. cerevisiae* homolog.
++	*rad22*$^+$	Homologous recombination protein (ScRAD52 homolog) (Ostermann *et al.* 1993)
+	*res2/pct1*$^+$	Transcription factor; forms a complex with Cdc10 (ScMBP1 homolog) (Miyamoto *et al.* 1994).
+	*sec18*$^+$	Putative vesicle membrane fusion protein; uncharacterized *Sc*Sec18 homolog. ORF: SPAC1834.11c.
++	*skb1*$^+$	Arginine methyltransferase (ScHSL7 homolog) (Gilbreth *et al.* 1996)
++	*taf111*$^+$	Transcription factor; putative TFIID subunit required for sexual development SPAC2G11.14 (Ueno *et al.* 2001).

β-Galactosidase is a qualitative measure of the degree of activation in the two-hybrid system.

though MCMs themselves can be coprecipitated under those conditions (*e.g.*, Sherman *et al.* 1998; Gómez *et al.* 2002). We completed our initial characterization of Mst1 protein by showing that acetylation of the four core histones can be observed in an IP–acetylation assay (supplemental Figure 3).

***mst1-L344S* cells are damage sensitive:** Data suggest that the NuA4 complex in other species is recruited to double-strand breaks through association with phosphorylated histone H2AX (Downs *et al.* 2004; Robert *et al.* 2006). Our data showing an interaction both genetically and physically with Rad22 suggested that Mst1 may contribute to repair of DSBs by influencing homologous recombination. Rad22 has also been implicated in recovery from HU (Meister *et al.* 2005; Bailis *et al.* 2008). Therefore, we examined *mst1-L344S* mutants for evidence of a role in the damage response.

We first examined growth of *mst1-L344S* in response to chronic exposure to low levels of DNA-damaging agents (Figure 5A). We examined MMS (alkylating agent delays replication), HU (inhibits ribonucleotide reductase and causes fork stalling), and camptothecin (CPT; blocks topoisomerase and causes DSBs). *mst1-L344S* cells were sensitive to HU and MMS, but not to CPT at 29°. *mst1-L344S* was also sensitive to UV irradiation (200 J/m^2), but not to γ-irradiation (400 Gy) at 25° (not shown). We observed that sensitivity to these agents increased with temperature (data not shown). The *mst1-L344S* mutant was extremely sensitive to bleomycin at 32° (radio-mimetic; Figure 5B). These results are consistent with known roles for Mst1 family acetyltransferases in some forms of DSB repair. We also observed substantial sensitivity to thiabendazole, a microtubule-destabilizing drug (Figure 5C). Mutations affecting centromere or kinetochore function are often TBZ sensitive; this may relate to the synthetic interactions that we observed with the heterochromatin mutants (Table 2) and to the physical interactions with Cbh1 and Msc1, which affect centromere function (Table 3).

Mst1 is required for recovery from HU arrest: HU causes nucleotide depletion, which results in replication fork stalling early in S phase. Defects in HU response may reflect a failure to arrest the cell cycle during the acute treatment or an ability to restore replication and recover after HU is removed (reviewed in Branzei and Foiani 2007). As shown in Figure 6, A, B, and D, *mst1-L344S* mutants arrest with a 1C DNA content and modestly elongated morphology, as do wild-type and Δ*cds1* cells, indicating that checkpoint-mediated cell cycle arrest is intact. As seen in the time course in Figure 6, A and B, wild-type cells complete S phase within 30 min of removing HU, while Δ*cds1* mutants are severely delayed, consistent with their recovery defect. Wild-type cells also proceed to division and enter the next cell cycle, as seen in the increase of 4C population (reflecting binucleate cells that complete S phase before completing the previous cytokinesis). This population is not seen in the Δ*cds1* mutants, which are defective for recovery and show a broad, heterogeneous peak of DNA content at the later time points. The *mst1* cells resemble wild type, suggesting that the mutants maintain and restart the replication fork normally. This suggests that the sensitivity that we observed in *mst1* cells in HU reflects defects in a later stage of recovery. Similar results have been seen for Δ*rad22* and Δ*rhp51* (Meister *et al.* 2005; Bailis *et al.* 2008). Interestingly, although *mst1-L344S* cells were sensitive to HU during chronic treatment (Figure 5),

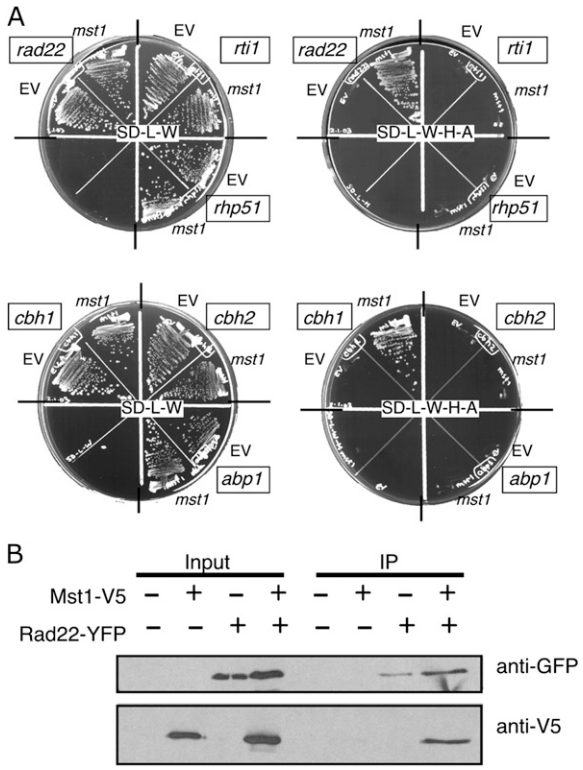

FIGURE 4.—Two-hybrid interactions. (A) A prey construct containing Mst1 was tested for interaction with full-length Rad22, Rti1, Rhp51, Cbh1, Cbh2, and Abp1 (see MATERIALS AND METHODS). *S. cerevisiae* cells containing the indicated plasmids were tested for growth on SD media lacking histidine and adenine (right), compared to control media containing both supplements. Growth on the selective media indicates a successful two-hybrid interaction. (B) Physical interaction between Mst1-V5 and Rad22-GFP. One milligram of protein was immunoprecipitated using 1 μl of anti-GFP antibody. Input lanes: 50 μg of total protein extract. GFP was detected by ECL+ incubation for 30 min. After 30 min, V5 blots were exposed to film after ECL+ (Pierce) incubation for 30 sec.

viability was relatively unaffected by an acute treatment in the HU block-and-release protocol (Figure 6C). This suggests that prolonged exposure to HU leads to an accumulation of growth defects, whereas transient exposure for a single generation causes relatively little problem.

One measure of DNA damage is provided by recruitment of recombination proteins to repair centers (CASPARI *et al.* 2002; DU *et al.* 2003; MEISTER *et al.* 2005). During HU treatment, Rad22 focus formation can distinguish two functions. In wild-type cells, Rad22 is recruited in a burst during the recovery phase after cells are released from HU. In contrast, in checkpoint mutants that fail to respond properly to HU, collapsed replication forks cause DNA breaks and Rad22 is recruited during HU treatment (CASPARI *et al.* 2002; DU *et al.* 2003; MEISTER *et al.* 2005; BAILIS *et al.* 2008). To determine whether reduced Mst1 function causes DNA damage, we examined Rad22YFP focus formation in wild-type and *mst1-L344S* cells at 32° (Figure 6, E and F). Due to background fluorescence observed in *mst1-L344S* mutant cells, we had to detect Rad22-YFP using indirect immunofluorescence on nuclear spreads. We observed a modest increase in Rad22 focus formation in *mst1-L344S* cells without HU treatment, relative to wild-type cells (compare 0-hr time point of *mst1* and wild type in Figure 6F). This suggests that the cells suffer some amount of endogenous damage. During acute HU treatment, Δ*cds1* cells show a dramatic increase in Rad22 YFP focus formation, consistent with replication fork collapse and double-strand breaks (MEISTER *et al.* 2005; BAILIS *et al.* 2008). There is no change in the Rad22YFP foci in *mst1-L344S*, again consistent with the model that replication forks are intact.

DISCUSSION

The MYST family of histone acetyltransferases is associated with diverse functions (reviewed in UTLEY and COTE 2003). Previously, we showed that *S. pombe* *mst1*$^+$ is an essential gene that encodes a protein constitutively bound to the chromatin (GÓMEZ *et al.* 2005). In this report, we have characterized the phenotypes associated with deletion of *mst1*$^+$ and a temperature-sensitive allele.

S. pombe Mst1 is the ortholog of *S. cerevisiae* Esa1 and human Tip60, the catalytic subunit of the NuA4 family of histone acetyltransferases, which is now referred to as

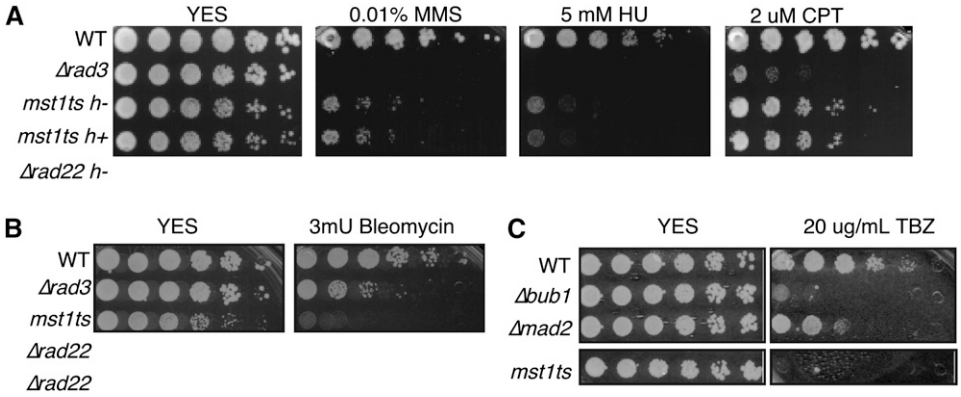

FIGURE 5.—Damage sensitivity of *mst1-L344S*. (A) *mst1* sensitivity relative to DNA-damaging agents. Exponentially growing cells were diluted fivefold on YES plates with the indicated drugs and grown 4 days at 29°. (B) *mst1* is sensitive to bleomycin. Exponentially growing cells were diluted fivefold on YES plates and grown 3 days at 32°. (C) *mst1* is sensitive to the spindle poison TBZ. Exponentially growing cells were diluted fivefold on YES plates and grown 3 days at 32°.

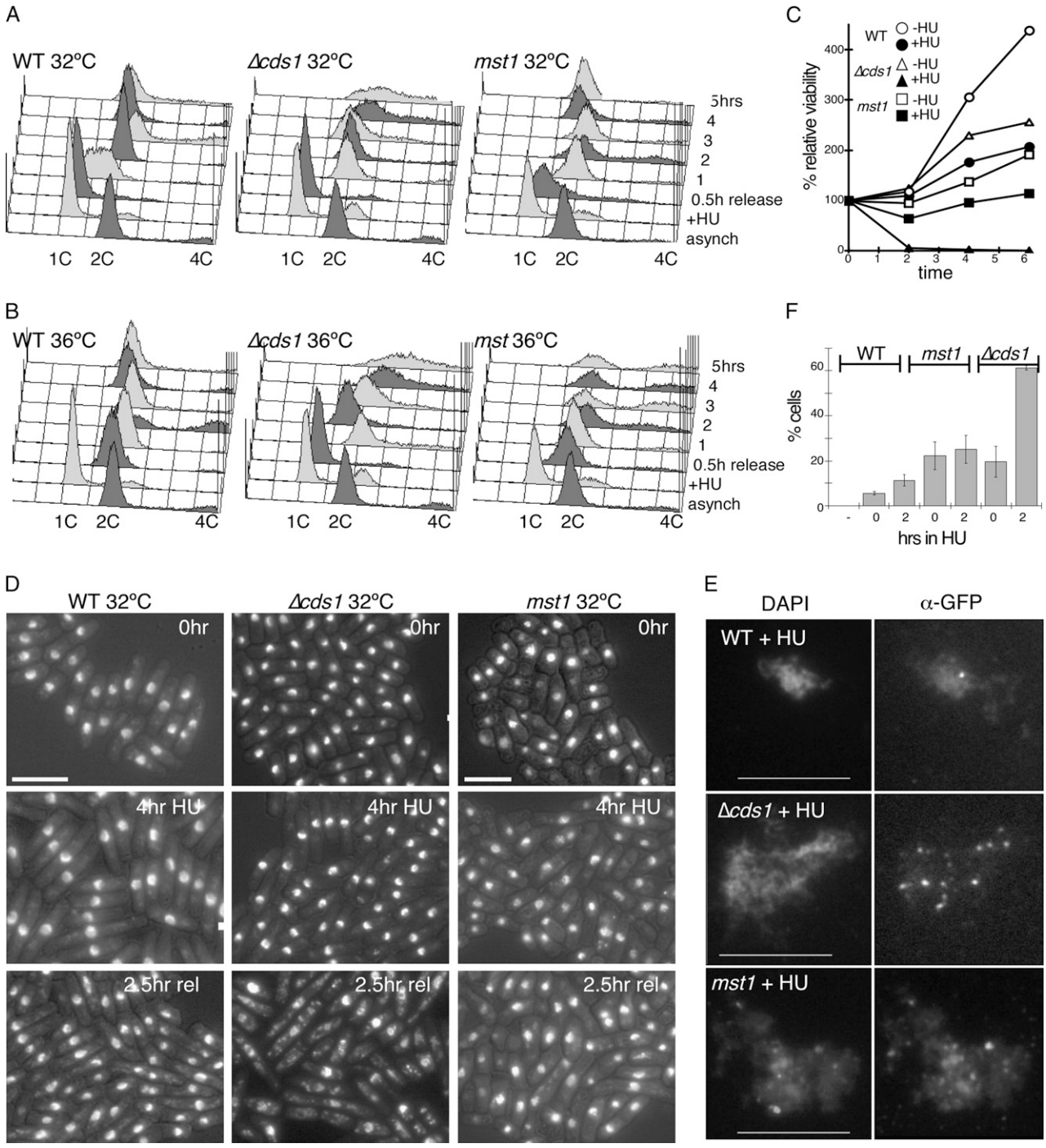

FIGURE 6.—*mst1-L344S* has HU recovery defect. (A and B) Time course of DNA content following release from HU at 32° (A) and 36° (B) for cells in YES. Strains: wild type (FY11), *mst1-L344S* (FY2396), Δ*cds1* (FY865). (C) Viability of cells following treatment with 12 mM HU at 32°. Circles, wild type; triangles, Δ*cds1*; squares, *mst1*. Open symbols, no HU; solid symbols, 12 mM HU. (D) Morphology of cells in C released before (−HU), during (4 hr HU), and after (2.5 hr rel). Cells were fixed in ethanol and stained with DAPI. Bar, 10 μm. (E) Rad22YFP foci in wild-type, Δ*cds1*, or *mst1* cells after 2 hr in HU. Rad22YFP was detected by immunofluorescence with anti-GFP antibody, which recognizes the YFP variant. (F) Fraction of cells with Rad22YFP foci at 0 or 2 hr of incubation in HU. -, negative control with no Rad22YFP. Strains: wild type (FY261), wild type with Rad22YFP (FY3436), *mst1* Rad11YFP (FY3369), Δ*cds1* Rad22YFP (FY3286).

the KAT5 family (reviewed in SQUATRITO *et al.* 2006; ALLIS *et al.* 2007). In addition to its presumed effects in transcriptional regulation (SMITH *et al.* 1998; ALLARD *et al.* 1999; CLARKE *et al.* 1999, 2006; DURANT and PUGH 2006), this family is implicated in checkpoint control by direct acetylation of ATM and p53 (SUN *et al.* 2005; JIANG *et al.* 2006; SYKES *et al.* 2006; TANG *et al.* 2006), as well as in repair by direct recruitment to phosphorylated histone H2AX at double-strand breaks (BIRD *et al.* 2002; DOWNS *et al.* 2004; TAMBURINI and TYLER 2005;

TABLE 4

Summary of processes associated with *mst1*+ function

Associated genes	Function	*mst1ts* genetic interactions	Mst1 physical interactions
	Process: DNA double-strand break repair		
	Phenotypes: *mst1ts* mutant sensitivity to DNA-damaging agents MMS, bleomycin, and CPT		
rhp51+	Homologous recombination	Yes	None (Y2H)
rad22+	Homologous recombination	Yes	Yes (Y2H, IP)
	Process: replication fork stability/recovery		
	Evidence: *mst1ts* mutant is sensitive to chronic HU exposure. Increased Rad22 foci even under permissive conditions		
mcm2+	MCM helicase, required for replication fork assembly and elongation, and recovery from HU	Yes (all)	Mcm2 (coIP)
mcm4+			
mcm7+			
hsk1+	MCM kinase, required for replication initiation and HU recovery	Yes	ND
pol1+	DNA polymerase α	Yes	ND
	Process: centromere function		
	Evidence: Δ*mst1, mst1ts* mutants have chromosome segregation defects. *mst1ts* mutant is TBZ sensitive. No silencing defect suggests that the heterochromatin is intact in *mst1ts*.		
clr3+	Pericentromere heterochromatin assembly, silencing, centromere cohesion, assembly of Cnp1 histone in central core	Yes (all)	ND
clr4+			
swi6+			
cbh1+	Pericentromere heterochromatin assembly, assembly of Cnp1 histone in central core	None	Yes (Y2H)
msc1+	Assembly of Cnp1 histone in central core, kinetochore assembly	None	Yes (Y2H)
hir1+	Pericentromere heterochromatin silencing, kinetochore assembly	None	Yes (Y2H)

18 IP, co-immunoprecipitation. Y2H, yeast two-hybrid interaction. ND, not done.

Robert *et al.* 2006). Data also suggest that this family of HAT proteins acetylates histone H2AZ (Babiarz *et al.* 2006; Keogh *et al.* 2006; Millar *et al.* 2006); this histone variant is paradoxically associated with both regions of active gene transcription, as well as with centromeres and silent heterochromatin (reviewed in Guillemette and Gaudreau 2006).

We characterized two mutations affecting *mst1*$^+$ function. As we reported previously, the Δ*mst1* mutation causes lethality (Gómez *et al.* 2005). In this report, we show that spores bearing the disruption are able to germinate and complete DNA replication. However, they show a highly disordered mitosis. We also constructed the temperature-sensitive mutant *mst1-L344S*, which contains the same lesion as a temperature-sensitive budding-yeast mutant. At the restrictive temperature, this strain is also proficient for replication, whether it is shifted asynchronously (Figure 2), released from nitrogen starvation in G_1 (supplemental Figure 1), or released from hydroxyurea in S phase (Figure 6). We observed pleiotropic phenotypes, including disordered mitosis, irregular segregation, and a high fraction of cell cycle arrest with septa. Cells arrest within one to two cell cycles with this mixed phenotype.

Our genetic analysis and subsequent two-hybrid screen suggest that Mst1 impacts three pathways (Table 4): damage response, which is consistent with the physiological analysis of the *mst1-L344S* mutant and data from other systems; centromere function; and control of mitotic entry.

Even at permissive temperatures, *mst1-L344S* cells are sensitive to a variety of agents that disrupt the cell cycle, including bleomycin and MMS, which cause DNA damage and double-strand breaks; hydroxyurea, which causes replication fork arrest; and thiabendazole, which causes spindle depolarization. The damage sensitivity was expected because data from numerous sources link members of this family to the double-strand break response. The recruitment of KAT5 orthologs to double-strand breaks is thought to modulate the break response and perhaps downregulate the signal associated with H2AX phosphorylation by recruiting the remodelers INO80 and SWR1 and the alternative histone H2AZ (Morrison *et al.* 2004; Tsukuda *et al.* 2005; van Attikum and Gasser 2005b; Papamichos-Chronakis *et al.* 2006).

When *mst1-L344S* is combined with the mutations Δ*rhp51* (Rad51), Δ*rad50*, or Δ*rad22*, cells have a severe synthetic growth phenotype in which they grow more poorly than either parent. This is particularly apparent in Δ*rad22 mst1-L344S*; however, because Δ*rad22* alone is prone to accumulate suppressors in the *fbh1*$^+$ gene (Osman *et al.* 2005), we cannot be sure that our double mutant is not also carrying a suppressor. It is possible that Δ*rad22 mst1-L344S* will be synthetically lethal but we have not been able to test this convincingly. These

genetic data suggest that *mst1-L344S* requires a functional recombination pathway for viability. To assess whether *mst1-L344S* causes damage itself, we examined focus formation of Rad22YFP in *mst1-L344S* cells. There is an increase in foci in *mst1* compared to wild type at 32°, even in the absence of any damaging treatment (Figure 6F). This suggests that attenuation of Mst1 function results in some form of damage that requires Rad22 for repair, but also suggests that Rad22-YFP can be recruited to sites of damage in the absence of Mst1.

We observed a physical association between Mst1 and Rad22 in two-hybrid analysis and in co-immunoprecipitation. This interaction is specific to Rad22, as neither the Rad22 homolog Rti1 nor the Rad51 homolog Rhp51 associate with Mst1 in two-hybrid tests (Figure 4). This could reflect a role in recruitment. Another possibility is that Mst1 acetylates Rad22 itself, as Tip60 does to ATM or p53, to modulate its behavior, and we will investigate this in the future.

Unexpectedly, *mst1-L344S* cells are also sensitive to hydroxyurea, which causes replication fork arrest. The mutant cells arrest properly with a 1C content without Rad22-YFP foci and appear to complete S phase following release, indicating that the checkpoint functions and the replication fork are intact; instead, they have modestly reduced viability reminiscent of Δ*rhp51* or Δ*rad22* mutants (Meister *et al.* 2005; Bailis *et al.* 2008). We observe that HU sensitivity of *mst1-L344S* to HU differs if cells are exposed to low levels over a prolonged time, which causes cell death (chronic exposure, Figure 5), or whether they are exposed to higher levels over a single generation (acute exposure, Figure 6). Similar results have been observed for the fork protection mutant Δ*swi1* (ScTof1), which survives acute exposure but dies during chronic exposure (Sommariva *et al.* 2005). It is also possible that this HU sensitivity is related to the interaction between *mst1* and mutations in recombination genes; however, Rad22-YFP recruitment appears largely normal in *mst1* cells treated with HU.

A second broad functional area defined by our two-hybrid and genetic interactions is centromere assembly. Fission yeast centromeres are large and complex elements, with a central core surrounded by flanking repeats that form peri-centromeric heterochromatin (reviewed in Ekwall 2007; Grewal and Jia 2007). Mutations disrupting the heterochromatin, the central core, or the kinetochore typically show sensitivity to spindle poisons such as TBZ, and a high fraction of chromosome loss and disordered mitoses, as we observed (Figures 1, 2, and 5). Our genetic analysis suggests that mutations that disrupt normal pericentromeric heterochromatin function such as Δ*clr3* (histone deacetylase), Δ*clr4* (methyltransferase), or Δ*swi6* (heterochromatin protein 1) reduce the growth temperature of *mst1-L344S* quite substantially. The simplest interpretation of this epistasis experiment is that $mst1^{+}$ affects a separate pathway from Clr4 and Swi6 in centromere function, leading to a synthetic effect in the double mutant. We see no loss of silencing at centromeres or telomeres in a *mst1* mutant, which suggests that Swi6-formed heterochromatin is largely intact, and some other aspect of centromere assembly and function is disrupted. A good candidate is the central core. Three of the proteins that we identified in the two-hybrid tests are linked to central core and kinetochore function:

1. We isolated Msc1, which was first identified as a high-copy suppressor of a strain lacking the checkpoint kinase Chk1 (Ahmed *et al.* 2004). A recent study shows that overexpression of Msc1 rescues mutations of *cnp1* (the specialized H3-Cen/CenpA of the centromere central core) in addition to Δ*chk1* and that this suppression requires the H2AZ histone variant encoded by *pht1*+ (Ahmed *et al.* 2007). H2AZ is linked to heterochromatin assembly and centromere function (reviewed in Guillemette and Gaudreau 2006) and H2AZ is a known substrate of KAT5 family HATs (Babiarz *et al.* 2006; Keogh *et al.* 2006; Millar *et al.* 2006; Auger *et al.* 2008).
2. We isolated Cbh1, which is one of the three CenpB homologs in fission yeast (Baum and Clarke 2000; Nakagawa *et al.* 2002). The CenpB homologs are proposed to affect centromere structure (Nakagawa *et al.* 2002) and, in metazoans, CenpB is involved in recruitment of H3-Cen/CenpA (reviewed in Mellone and Allshire 2003).
3. We isolated Hip1, the HIRA/Hir1 homolog. HIR proteins are histone chaperones required for assembly of chromatin outside of S phase (reviewed in Gunjan *et al.* 2005). In fission yeast, deletion of Hip1 causes chromosome loss, TBZ sensitivity, and disruption of centromere silencing (Blackwell *et al.* 2004).

HIR proteins in *S. cerevisiae* are implicated in normal kinetochore function (Sharp *et al.* 2002). In contrast to the heterochromatin mutants, we see no synthetic growth phenotypes when *mst1* is combined with mutants Δ*cbh1*, Δ*msc1*, or Δ*hip1*. This would be consistent with $mst1^{+}$ acting in a pathway with these genes. Together with the segregation defects of *mst1-L344S*, these two hybrid interactions suggest that Mst1 may affect the centromere central core. We are examining mitotic phenotypes of *mst1* in more detail to determine how this may occur and the role of H2AZ in this process.

There is indirect evidence that Mst1 may affect mitotic entry. We isolated the Skb1 methyltransferase in our two-hybrid screen. Budding yeast Hsl7, the ortholog of Skb1, was identified because mutations of *hsl7* are synthetically lethal with histone tail mutants (Ma *et al.* 1996); genetic interactions between ScΔ*hsl7* and Sc*Esa1* suggest that histone tail modifications are important for normal mitosis when Swe1 kinase is activated (Ruault and Pillus 2006). While we observed no genetic interaction between SpΔ*skb1* and our

temperature-sensitive allele of *mst1*, the physical interaction that we observed, together with the genetic data from budding yeast, suggests that there is a role for histone tail modification in the proper regulation of mitotic entry in both species. Another link to CDK activation is provided by our isolation of Hip1 as a putative interactor. The Δ*hip1* mutation is synthetically lethality with *cdc25-22*, which is a temperature-sensitive mutation in the mitotic inducer Cdc25 that opposes Wee1 (ScSwe1) (Blackwell *et al.* 2004; reviewed in Karlsson-Rosenthal and Millar 2006).

We did not isolate any of the NuA4 complex proteins in our two-hybrid screen, although the previously characterized SpAlp5/ScArp4 is known to interact with Mst1 by co-immunoprecipitation (Minoda *et al.* 2005). This is not surprising as the two-hybrid screen identifies binary associations, and many of the interactions in the large NuA4 complex are likely to be indirect (Auger *et al.* 2008). We speculate that the structure of the two-hybrid bait construct, in which the amino-terminus of Mst1 is fused to the DNA-binding domain of Gal4, may interfere with the assembly of NuA4 subunits with Mst1. Conversely, we did not observe co-immunoprecipitation of epitope-tagged Mst1 with most of the two-hybrid clones. This could reflect chromatin association of the proteins involved, temporally or spatially limiting association, or could indicate that the C-terminal epitope tag on Mst1 is not available when the protein is in a complex with other factors. A more detailed structure–function analysis and localization studies are in progress to resolve these observations.

The interactions that we observed, in addition to the phenotypes associated with the *mst1* mutants, suggest that Mst1 affects a variety of cell functions. Importantly, the recent discovery that Tip60 can acetylate nonhistone proteins in the absence of the NuA4 complex (*e.g.*, Sun *et al.* 2005; Sykes *et al.* 2006; Tang *et al.* 2006; Kim *et al.* 2007) indicates that the identification of Mst1 as a histone acetyltransferase may be too limiting a definition. We speculate that the pleiotropic effects of *mst1* mutants indicate a variety of functions that may be directly modulated by Mst1 catalytic activity, not necessarily linked to histone acetylation. Our study suggests that Mst1 may have functions in replication fork stability and in centromere assembly, and experiments are underway to determine whether these effects are mediated by histone modifications or by other aspects of Mst1 function.

We thank Louise Clarke, Greg Freyer, Stevan Marcus, Paul Russell, Nancy Walworth, and Simon Whitehall for strains. We thank Vanessa Angeles and Paula Tran for assistance with genetics and growth assays; Su Lam for help with the two-hybrid analysis; Marc Green for microscopy assistance; and Jeff Hodson, John Marlett, and Ji-Ping Yuan for lab support. We thank Joaquin Espinosa for assistance with histone acetyltransferase assays. We are grateful to Ruben Petreaca for careful reading of the manuscript and to members of the Forsburg lab for many helpful comments throughout the course of the study. We thank Lorraine Pillus and the Pillus lab at the University of California, San Diego, for their hospitality to E.B.G. during part of this work. R.L.N. was supported by DOD W81XWH-05-1-0391 (X. Chen). This work was supported by National Institutes of Health grant R01 GM059321 to S.L.F.

LITERATURE CITED

Ahmed, S., C. Palermo, S. Wan and N. C. Walworth, 2004 A novel protein with similarities to Rb binding protein 2 compensates for loss of Chk1 function and affects histone modification in fission yeast. Mol. Cell. Biol. **24:** 3660–3669.

Ahmed, S., B. E. Dul, X. Qiu and N. C. Walworth, 2007 Msc1 acts through histone H2A.Z to promote chromosome stability in *Schizosaccharomyces pombe*. Genetics **177:** 1487–1497.

Allard, S., R. T. Utley, J. Savard, A. Clarke, P. Grant *et al.*, 1999 NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. EMBO J. **18:** 5108–S5119.

Allis, C. D., S. L. Berger, J. Cote, S. Dent, T. Jenuwien *et al.*, 2007 New nomenclature for chromatin modifying enzymes. Cell **131:** 633–636.

Auger, A., L. Galarneau, M. Altaf, A. Nourani, Y. Doyon *et al.*, 2008 Eaf1 is the platform for NuA4 molecular assembly that evolutionarily links chromatin acetylation to ATP-dependent exchange of histone H2A variants. Mol. Cell. Biol.

Babiarz, J. E., J. E. Halley and J. Rine, 2006 Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae. Genes Dev. **20:** 700–710.

Bailis, J. M., D. D. Luche, T. Hunter and S. L. Forsburg, 2008 MCM proteins interact with checkpoint and recombination proteins to promote S phase genome stability. Mol. Cell. Biol. (in press).

Baum, M., and L. Clarke, 2000 Fission yeast homologs of human CENP-B have redundant functions affecting cell growth and chromosome segregation. Mol. Cell. Biol. **20:** 2852–2864.

Bird, A. W., D. Y. Yu, M. G. Pray-Grant, Q. Qiu, K. E. Harmon *et al.*, 2002 Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. Nature **419:** 411–415.

Blackwell, C., K. A. Martin, A. Greenall, A. Pidoux, R. C. Allshire *et al.*, 2004 The *Schizosaccharomyces pombe* HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. Mol. Cell. Biol. **24:** 4309–4320.

Boudreault, A. A., D. Cronier, W. Selleck, N. Lacoste, R. T. Utley *et al.*, 2003 Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin. Genes Dev. **17:** 1415–1428.

Branzei, D., and M. Foiani, 2007 Interplay of replication checkpoints and repair proteins at stalled replication forks. DNA Rep. **6:** 994–1003.

Burke, T. W., J. Cooks-Gowen, M. Asano and J. R. Nevins, 2001 Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. J. Biol. Chem. **276:** 15397–15408.

Caspari, T., J. M. Murray and A. M. Carr, 2002 Cdc2-cyclin B kinase activity links Crb2 amd Rqh1-topoisomerase III. Genes Dev. **16:** 1195–1208.

Catlett, M. G., and S. L. Forsburg, 2003 Schizosaccharomyces pombe Rdh54 (TID1) acts with Rhp54 (RAD54) to repair meiotic double-strand breaks. Mol. Biol. Cell **14:** 4707–4720.

Clarke, A. S., J. E. Lowell, S. J. Jacobson and L. Pillus, 1999 Esa1p is an essential histone acetyltransferase required for cell cycle progression. Mol. Cell. Biol. **19:** 2515–2526.

Clarke, A. S., E. Samal and L. Pillus, 2006 Distinct roles for the essential MYST family HAT Esa1p in transcriptional silencing. Mol. Biol. Cell **17:** 1744–1757.

Deckert, J., and K. Struhl, 2001 Histone acetylation at promoters is differentially affected by specific activators and repressors. Mol. Cell. Biol. **21:** 2726–2735.

de la Cruz, X., S. Lois, S. Sanchez-Molina and M. A. Martinez-Balbas, 2005 Do protein motifs read the histone code? BioEssays **27:** 164–175.

Downs, J. A., S. Allard, O. Jobin-Robitaille, A. Javaheri, A. Auger *et al.*, 2004 Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. Mol. Cell **16:** 979–990.

Du, L. L., T. M. Nakamura, B. A. Moser and P. Russell, 2003 Retention but not recruitment of Crb2 at double-strand breaks requires Rad1 and Rad3 complexes. Mol. Cell. Biol. **23:** 6150–6158.

Dul, B. E., and N. C. Walworth, 2007 The plant homeodomain fingers of fission yeast Msc1 exhibit E3 ubiquitin ligase activity. J. Biol. Chem. **282:** 18397–18406.

Durant, M., and B. F. Pugh, 2006 Genome-wide relationships between TAF1 and histone acetyltransferases in Saccharomyces cerevisiae. Mol. Cell. Biol. **26:** 2791–2802.

Early, A., L. S. Drury and J. F. Diffley, 2004 Mechanisms involved in regulating DNA replication origins during the cell cycle and in response to DNA damage. Philos. Trans. R. Soc. Lond. B Biol. Sci. **359:** 31–38.

Eberharter, A., R. Ferreira and P. Becker, 2005 Dynamic chromatin: concerted nucleosome remodelling and acetylation. Biol. Chem. **386:** 745–751.

Ekwall, K., 2007 Epigenetic control of centromere behavior. Annu. Rev. Genet. 12

Forsburg, S. L., and N. Rhind, 2006 Basic methods for fission yeast. Yeast **23:** 173–183.

Foster, E. R., and J. A. Downs, 2005 Histone H2A phosphorylation in DNA double-strand break repair. FEBS J. **272:** 3231–3240.

Fujita, A., A. Tonouchi, T. Hiroko, F. Inose, T. Nagashima *et al.*, 1999 Hsl7p, a negative regulator of Ste20p protein kinase in the Saccharomyces cerevisiae filamentous growth-signaling pathway. Proc. Natl. Acad. Sci. USA **96:** 8522–8527.

Galarneau, L., A. Nourani, A. A. Boudreault, Y. Zhang, L. Heliot *et al.*, 2000 Multiple links between the NuA4 histone acetyltransferase complex and epigenetic control of transcription. Mol. Cell **5:** 927–937.

Gilbreth, M., P. Yang, D. Wang, J. Frost, A. Polverino *et al.*, 1996 The highly conserved skb1 gene encodes a protein that interacts with Shk1, a fission yeast Ste20/PAK homolog. Proc. Natl. Acad. Sci. USA **93:** 13802–13807.

Glynn, J. M., R. J. Lustig, A. Berlin and F. Chang, 2001 Role of bud6p and tea1p in the interaction between actin and microtubules for the establishment of cell polarity in fission yeast. Curr. Biol. **11:** 836–845.

Gómez, E. G., and S. L. Forsburg, 2003 Analysis of *Schizosaccharomyces pombe* cell cycle, pp. 93–111 in *Cell Cycle Checkpoint Protocols*, edited by H. G. Lieberman. Humana Press, Tatawa, NJ.

Gómez, E. G., M. G. Catlett and S. L. Forsburg, 2002 Different phenotypes *in vivo* are associated with ATPase motif mutations in *Schizosaccharomyces pombe* minichromosome maintenance proteins. Genetics **160:** 1305–1318.

Gómez, E. B., J. Espinosa and S. L. Forsburg, 2005 *S. pombe mst2+* encodes a MYST-family histone acetyltransferase required for telomere silencing. Mol. Cell. Biol. **25:** 8887–8903.

Grewal, S. I., and S. Jia, 2007 Heterochromatin revisited. Nat. Rev. Genet. **8:** 35–46.

Guillemette, B., and L. Gaudreau, 2006 Reuniting the contrasting functions of H2A.Z. Biochem. Cell Biol. **84:** 528–535.

Gunjan, A., J. Paik and A. Verreault, 2005 Regulation of histone synthesis and nucleosome assembly. Biochimie **87:** 625–635.

Iizuka, M., and B. Stillman, 1999 Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. J. Biol. Chem. **274:** 23027–23034.

Iizuka, M., T. Matsui, H. Takisawa and M. M. Smith, 2006 Regulation of replication licensing by acetyltransferase Hbo1. Mol. Cell. Biol. **26:** 1098–1108.

Ikura, T., V. V. Ogryzko, M. Grigoriev, R. Groisman, J. Wang *et al.*, 2000 Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. Cell **102:** 463–473.

Jenuwein, T., and C. D. Allis, 2001 Translating the histone code. Science **293:** 1074–1080.

Jiang, X., Y. Sun, S. Chen, K. Roy and B. D. Price, 2006 The FATC domains of PIKK proteins are functionally equivalent and participate in the Tip60-dependent activation of DNA-PKcs and ATM. J. Biol. Chem. **281:** 15741–15746.

Karlsson-Rosenthal, C., and J. B. Millar, 2006 Cdc25: mechanisms of checkpoint inhibition and recovery. Trends Cell Biol. **16:** 285–292.

Keogh, M. C., T. A. Mennella, C. Sawa, S. Berthelet, N. J. Krogan *et al.*, 2006 The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes Dev. **20:** 660–665.

Kim, M. Y., E. J. Ann, J. Y. Kim, J. S. Mo, J. H. Park *et al.*, 2007 Tip60 histone acetyltransferase acts as a negative regulator of Notch1 signaling by means of acetylation. Mol. Cell. Biol. 13

Kouzarides, T., 2007 Chromatin modifications and their function. Cell **128:** 693–705.

Lafon, A., C. S. Chang, E. M. Scott, S. J. Jacobson and L. Pillus, 2007 MYST opportunities for growth control: yeast genes illuminate human cancer gene functions. Oncogene **26:** 5373–5384.

Lee, J. H., J. R. Cook, B. P. Pollack, T. G. Kinzy, D. Norris *et al.*, 2000 Hsl7p, the yeast homologue of human JBP1, is a protein methyltransferase. Biochem. Biophys. Res. Commun. **274:** 105–111.

Lee, K. K., and J. L. Workman, 2007 Histone acetyltransferase complexes: one size doesn't fit all. Nat. Rev. Mol. Cell Biol. **8:** 284–295.

Le Masson, I., D. Y. Yu, K. Jensen, A. Chevalier, R. Courbeyrette *et al.*, 2003 Yaf9, a novel NuA4 histone acetyltransferase subunit, is required for the cellular response to spindle stress in yeast. Mol. Cell. Biol. **23:** 6086–6102.

Leverson, J. D., H. K. Huang, S. L. Forsburg and T. Hunter, 2002 The Schizosaccharomyces pombe aurora-related kinase Ark1 interacts with the inner centromere protein Pic1 and mediates chromosome segregation and cytokinesis. Mol. Biol. Cell **13:** 1132–1143.

Liang, D. T., and S. L. Forsburg, 2001 Characterization of *Schizosaccharomyces pombe mcm7*$^+$ and *cdc23*$^+$ *(MCM10)* and interactions with replication checkpoints. Genetics **159:** 471–486. 14

Ma, X. J., Q. Lu and M. Grunstein, 1996 A search for proteins that interact genetically with histone H3 and H4 amino termini uncovers novel regulators of the Swe1 kinase in *Saccharomyces cerevisiae*. Genes Dev. **10:** 1327–1340.

Meister, P., A. Taddei, L. Vernis, M. Poidevin, S. M. Gasser *et al.*, 2005 Temporal separation of replication and recombination requires the intra-S checkpoint. J. Cell Biol. **168:** 537–544.

Mellone, B. G., and R. C. Allshire, 2003 Stretching it: putting the CEN(P-A) in centromere. Curr. Opin. Genet. Dev. **13:** 191–198.

Millar, C. B., and M. Grunstein, 2006 Genome-wide patterns of histone modifications in yeast. Nat. Rev. Mol. Cell Biol. **7:** 657–666.

Millar, C. B., F. Xu, K. Zhang and M. Grunstein, 2006 Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. Genes Dev. **20:** 711–722.

Minoda, A., S. Saitoh, K. Takahashi and T. Toda, 2005 BAF53/Arp4 homolog Alp5 in fission yeast is required for histone H4 acetylation, kinetochore-spindle attachment, and gene silencing at centromere. Mol. Biol. Cell **16:** 316–327.

Miyamoto, M., K. Tanaka and H. Okayama, 1994 res2+, a new member of the *cdc10*$^+$*/SWI4* family, controls the 'start' of mitotic and meiotic cycles in fission yeast. EMBO J. **13:** 1873–1880.

Moreno, S., A. Klar and P. Nurse, 1991 Molecular genetic analysis of the fission yeast *Schizosaccharomyces pombe*. Methods Enzymol. **194:** 795–823.

Morrison, A. J., and X. Shen, 2005 DNA repair in the context of chromatin. Cell Cycle **4:** 568–571.

Morrison, A. J., J. Highland, N. J. Krogan, A. Arbel-Eden, J. F. Greenblatt *et al.*, 2004 INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. Cell **119:** 767–775.

Nakagawa, H., J. K. Lee, J. Hurwitz, R. C. Allshire, J. Nakayama *et al.*, 2002 Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. Genes Dev. **16:** 1766–1778.

Osman, F., J. Dixon, A. R. Barr and M. C. Whitby, 2005 The F-Box DNA helicase Fbh1 prevents Rhp51-dependent recombination without mediator proteins. Mol. Cell. Biol. **25:** 8084–8096.

Ostermann, K., A. Lorentz and H. Schmidt, 1993 The fission yeast rad22 gene, having a function in mating-type switching and repair of DNA damages, encodes a protein homolog to rad52 of Saccharomyces cerevisiae. Nucleic Acids Res. **21:** 5940–5944.

Pankratz, D. G., and S. L. Forsburg, 2005 Meiotic S-phase damage activates recombination without checkpoint arrest. Mol. Biol. Cell **16:** 1651–1660.

Papamichos-Chronakis, M., J. E. Krebs and C. L. Peterson, 2006 Interplay between Ino80 and Swr1 chromatin remodeling enzymes regulates cell cycle checkpoint adaptation in response to DNA damage. Genes Dev. **20:** 2437–2449.

Robert, F., S. Hardy, Z. Nagy, C. Baldeyron, R. Murr *et al.*, 2006 The transcriptional histone acetyltransferase cofactor TRRAP associates with the MRN repair complex and plays a role in DNA double-strand break repair. Mol. Cell. Biol. **26:** 402–412.

Ruault, M., and L. Pillus, 2006 Chromatin-modifiying enzymes are essential when the *Saccharomyces cerevisiae* morphogenesis checkpoint is constitutively activated. Genetics **174:** 1135–1149.

Selleck, W., I. Fortin, D. Sermwittayawong, J. Cote and S. Tan, 2005 The Saccharomyces cerevisiae Piccolo NuA4 histone acetyltransferase complex requires the Enhancer of Polycomb A domain and chromodomain to acetylate nucleosomes. Mol. Cell. Biol. **25:** 5535–5542.

Sharp, J. A., A. A. Franco, M. A. Osley and P. D. Kaufman, 2002 Chromatin assembly factor I and Hir proteins contribute to building functional kinetochores in *S. cerevisiae*. Genes Dev. **16:** 85–100.

Sherman, D. A., S. G. Pasion and S. L. Forsburg, 1998 Multiple domains of fission yeast Cdc19p (MCM2) are required for its association with the core MCM complex. Mol. Biol. Cell **9:** 1833–1845.

Shogren-Knaak, M., H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie *et al.*, 2006 Histone H4–K16 acetylation controls chromatin structure and protein interactions. Science **311:** 844–847.

Shulewitz, M. J., C. J. Inouye and J. Thorner, 1999 Hsl7 localizes to a septin ring and serves as an adapter in a regulatory pathway that relieves tyrosine phosphorylation of Cdc28 protein kinase in Saccharomyces cerevisiae. Mol. Cell. Biol. **19:** 7123–7137.

Smith, E. R., A. Eisen, W. Gu, M. Sattah, A. Pannuti *et al.*, 1998 ESA1 is a histone acetyltransferase that is essential for growth in yeast. Proc. Natl. Acad. Sci. USA **95:** 3561–3565.

Sommariva, E., T. K. Pellny, N. Karahan, S. Kumar, J. A. Huberman *et al.*, 2005 *Schizosaccharomyces pombe* Swi1, Swi3, and Hsk1 are components of a novel S-phase response pathway to alkylation damage. Mol. Cell. Biol. **25:** 2770–2784.

Squatrito, M., C. Gorrini and B. Amati, 2006 Tip60 in DNA damage response and growth control: many tricks in one HAT. Trends Cell Biol. **16:** 433–442.

Sun, Y., X. Jiang, S. Chen, N. Fernandes and B. D. Price, 2005 A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. Proc. Natl. Acad. Sci. USA **102:** 13182–13187.

Sykes, S. M., H. S. Mellert, M. A. Holbert, K. Li, R. Marmorstein *et al.*, 2006 Acetylation of the p53 DNA-binding domain regulates apoptosis induction. Mol. Cell **24:** 841–851.

Tamburini, B. A., and J. K. Tyler, 2005 Localized histone acetylation and deacetylation triggered by the homologous recombination pathway of double-strand DNA repair. Mol. Cell. Biol. **25:** 4903–4913.

Tanaka, K., K. Okazaki, N. Okazaki, T. Ueda, A. Sugiyama *et al.*, 1992 A new cdc gene required for S phase entry of *Schizosaccharomyces pombe* encodes a protein similar to the *cdc10*$^{+}$ and *SWI4* gene products. EMBO J. **11:** 4923–4932.

Tang, Y., J. Luo, W. Zhang and W. Gu, 2006 Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. Mol. Cell **24:** 827–839.

Tsukuda, T., A. B. Fleming, J. A. Nickoloff and M. A. Osley, 2005 Chromatin remodelling at a DNA double-strand break site in Saccharomyces cerevisiae. Nature **438:** 379–383.

Ueno, M., R. Kurokawa, H. Renauld, K. Watanabe, T. Ushimaru *et al.*, 2001 Schizosaccharomyces pombe taf1+ is required for nitrogen starvation-induced sexual development and for entering the dormant GO state. Curr. Genet. **38:** 307–313.

Utley, R. T., and J. Cote, 2003 The MYST family of histone acetyltransferases. Curr. Top. Microbiol. Immunol. **274:** 203–236.

van Attikum, H., and S. M. Gasser, 2005a ATP-dependent chromatin remodeling and DNA double-strand break repair. Cell Cycle **4:** 1011–1014.

van Attikum, H., and S. M. Gasser, 2005b The histone code at DNA breaks: A guide to repair? Nat. Rev. Mol. Cell Biol. **6:** 757–765.

Vogelauer, M., L. Rubbi, I. Lucas, B. J. Brewer and M. Grunstein, 2002 Histone acetylation regulates the time of replication origin firing. Mol. Cell **10:** 1223–1233.

Yu, Y., Y. Teng, H. Liu, S. H. Reed and R. Waters, 2005 UV irradiation stimulates histone acetylation and chromatin remodeling at a repressed yeast locus. Proc. Natl. Acad. Sci. USA **102:** 8650–8655.

Zhu, Y., T. Takeda, K. Nasmyth and N. Jones, 1994 *pct1*$^{+}$, which encodes a new DNA-binding partner of p85^{cdc10}, is required for meiosis in the fission yeast *Schizosaccharomyces pombe*. Genes Dev. **8:** 885–898.

Communicating editor: E. Jones

Genetics June 2008
Author query sheet Gómez (85779)

1 "reviewed" as meant here and throughout for "rev."?

2 "Two-hybrid strains" as meant in heading?

3 The term "chromatin" does not appear in Table 2. Please revise text if necessary.

4 "interact in two-hybrid screens" as meant?

5 "an arginine methyltransferase, an ortholog of *S. cerevisiae* Hsl7" as meant?

6 The term "chronic exposure" does not appear in Figure 5 and "acute exposure" does not appear in Figure 6. Please revise text or figures if necessary.

7 "the two-hybrid screen" as meant?

8 Please replace DOD with its definition.

9 Please verify the page range in Allard *et al.* 1999.

10 Please update Auger *et al.* 2008 by providing volume no. and page range.

11 Please update Bailis *et al.* 2008 by providing volume no. and page range.

12 Please provide volume no. and page range for Ekwall 2007.

13 Please provide volume no. and page range for Kim *et al.* 2007.

14 Please cite Liang and Forsburg 2001 in the text or delete the reference.

15 Please define "rel" in "2.5 hr rel."

16 "h^-" (with superscript minus sign) as meant throughout Table 1?

17 "two-hybrid screening" (not "2 hybrid") as meant?

18 "yeast two-hybrid interaction" as meant?

LETTERS

The APOBEC-2 crystal structure and functional implications for the deaminase AID

Courtney Prochnow[1]*, Ronda Bransteitter[1]*, Michael G. Klein[1], Myron F. Goodman[1] & Xiaojiang S. Chen[1]

APOBEC-2 (APO2) belongs to the family of apolipoprotein B messenger RNA-editing enzyme catalytic (APOBEC) polypeptides, which deaminates mRNA and single-stranded DNA[1,2]. Different APOBEC members use the same deamination activity to achieve diverse human biological functions. Deamination by an APOBEC protein called activation-induced cytidine deaminase (AID) is critical for generating high-affinity antibodies[3], and deamination by APOBEC-3 proteins can inhibit retrotransposons and the replication of retroviruses such as human immunodeficiency virus and hepatitis B virus[4–7]. Here we report the crystal structure of APO2. APO2 forms a rod-shaped tetramer that differs markedly from the square-shaped tetramer of the free nucleotide cytidine deaminase, with which APOBEC proteins share considerable sequence homology. In APO2, two long α-helices of a monomer structure prevent the formation of a square-shaped tetramer and facilitate formation of the rod-shaped tetramer via head-to-head interactions of two APO2 dimers. Extensive sequence homology among APOBEC family members allows us to test APO2 structure-based predictions using AID. We show that AID deamination activity is impaired by mutations predicted to interfere with oligomerization and substrate access. The structure suggests how mutations in patients with hyper-IgM-2 syndrome inactivate AID, resulting in defective antibody maturation.

We crystallized APO2, which contains amino acid residues 41–224, with four monomers in each asymmetric unit that form a tetramer with an atypical elongated shape (Fig. 1a). This tetramer

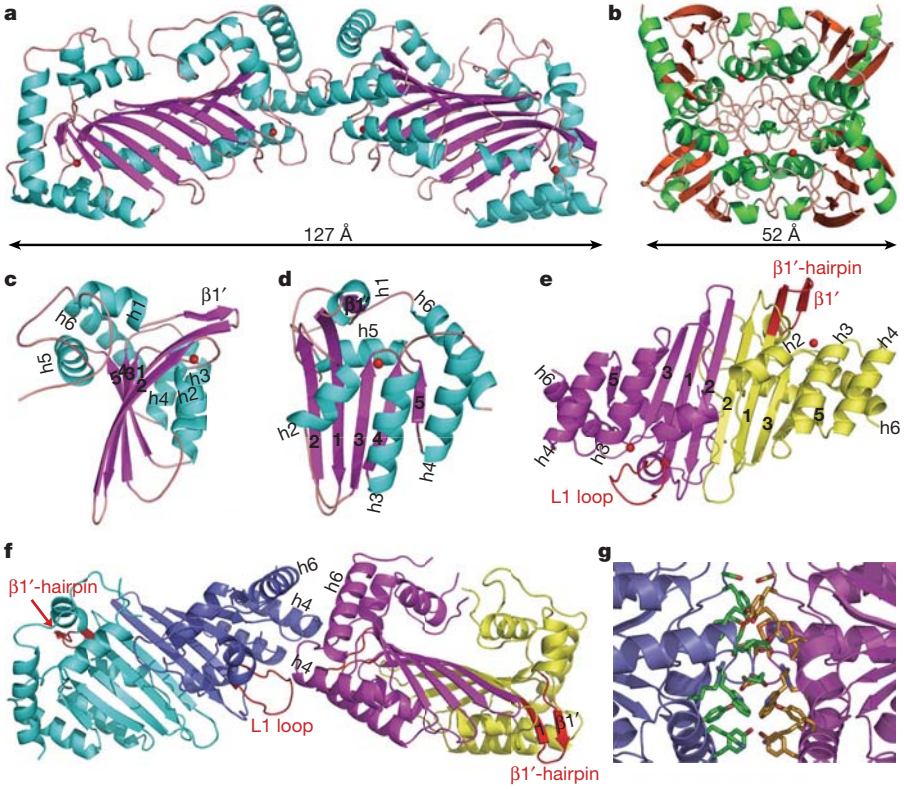

Figure 1 | The structure of APO2. a, The APO2 tetramer structure. It has an end-to-end span of ~126.9 Å. Zn atoms in the active centres are shown as red spheres. **b**, The square-shaped structure of human cytidine deaminase (PDB accession number: 1MQ0), a fntCDA. **c**, **d**, The APO2 monomer structure rotated by 90 degrees, showing the unique features of APO2: the short β1′ strand and helices h4 and h6. h4 and h6 dictate how APO2 oligomerizes. **e**, The APO2 dimer formed by two monomers (in purple and yellow). Each has a different conformation for the h1/β1-turn (in red): a loop (L1) and a hairpin. **f**, The tetrameric interface, showing the extensive interactions mediated through h4, h6 and L1. **g**, The stick model (hydrophobic, polar and charged amino acids in h4, h6 and L1) of the interactions at the tetramer interface.

[1]Molecular and Computational Biology, University of Southern California Los Angeles, California 90089, USA.
*These authors contributed equally to this work.

assembles through two different monomer–monomer interfaces, in contrast to the canonical square shape of the free nucleotide cytidine deaminase (fntCDA) tetramer (Fig. 1b), in which all four monomers interact with each other[8]. The elongated APO2 tetramer has the shape of a butterfly (Fig. 1a) with an end-to-end span of approximately 126.9 Å.

The APO2 monomer appears to adopt the typical core fold of the fntCDAs with a five-stranded β-sheet flanked by helices on both sides (Figs 1c, d). However, one new attribute is the additional α-helices surrounding the core β-sheet (Figs 1c, d); six long helices are present in the APO2 monomer whereas only three or four are observed in the fntCDA monomer (excluding the shorter 3_{10} helices)[8–11]. Helices h3 and h6 make extensive contacts with h4, stabilizing the position of the helices within the monomer subunit. On the basis of the close sequence homology of APO2 with other APOBEC proteins, the long helix (h4) probably serves as a structural signature of this family (Figs 1c, d).

The APO2 dimer is formed by pairing two long β-strands (β2) (Fig. 1e), joining two β-sheets sideways to form one wide β-sheet that resembles a ribcage (Fig. 1a, e). Twelve residues (residues 82–93) on each β2 strand form 12 hydrogen bonds through main-chain atoms, providing the principal bonding force between the two monomers. The dimer interface is reinforced by the side-chain interactions occurring through the loops and helices located on both sides of the β-sheet. Ordered water molecules also help to stabilize this interface.

The dimer is nearly symmetrical (Fig. 1e) with six helices (h2, h3 and h4 of both molecules) located on one side of the augmented β-sheet and four helices (h1 and h5 of both monomers) on the other side. Capped on both edges of the β-sheet are h4 and h6. However, one part of the dimer shows obvious asymmetry at the turn between h1 and strand β1 (h1/β1-turn). This h1/β1-turn (residues 57–68) assumes a hairpin structure (β1′-hairpin) in one monomer, and a loop conformation (L1) in the other monomer (Fig. 1e, f).

The APO2 tetramer is formed by two dimers joining through head-to-head interactions. The two dimers make extensive contacts via the residues from h4 and h6, as well as the loop L1 at the h1/β1-turn (Fig. 1f). Residues Y61, F155, M156, W157, P160, Y214 and Y215 from each side of the interface form extensive hydrophobic packing interactions, and residues R57, S62, S63, R153, E158, E159 and E161 establish salt bridges and hydrogen bonds (Fig. 1g). Some charged residues even use their aliphatic side chains to interact with hydrophobic residues. Thus, hydrophobic, polar and charged amino acid side chains are all involved in the tetramerization interactions. The total buried area is 1,745 $Å^2$ within the tetramer interface, where h4 and h6 play a major role forming the interface (Fig. 1f). h4 and h6 also sterically hinder the formation of the square-shaped fntCDA-type tetramer by occupying the space where another monomer would need to be. Therefore, h4 and h6 appear to determine directly the elongated tetramer formation.

A prominent feature of the APO2 tetramer distinctive from the fntCDA tetramer is that the active sites are accessible for large RNA or

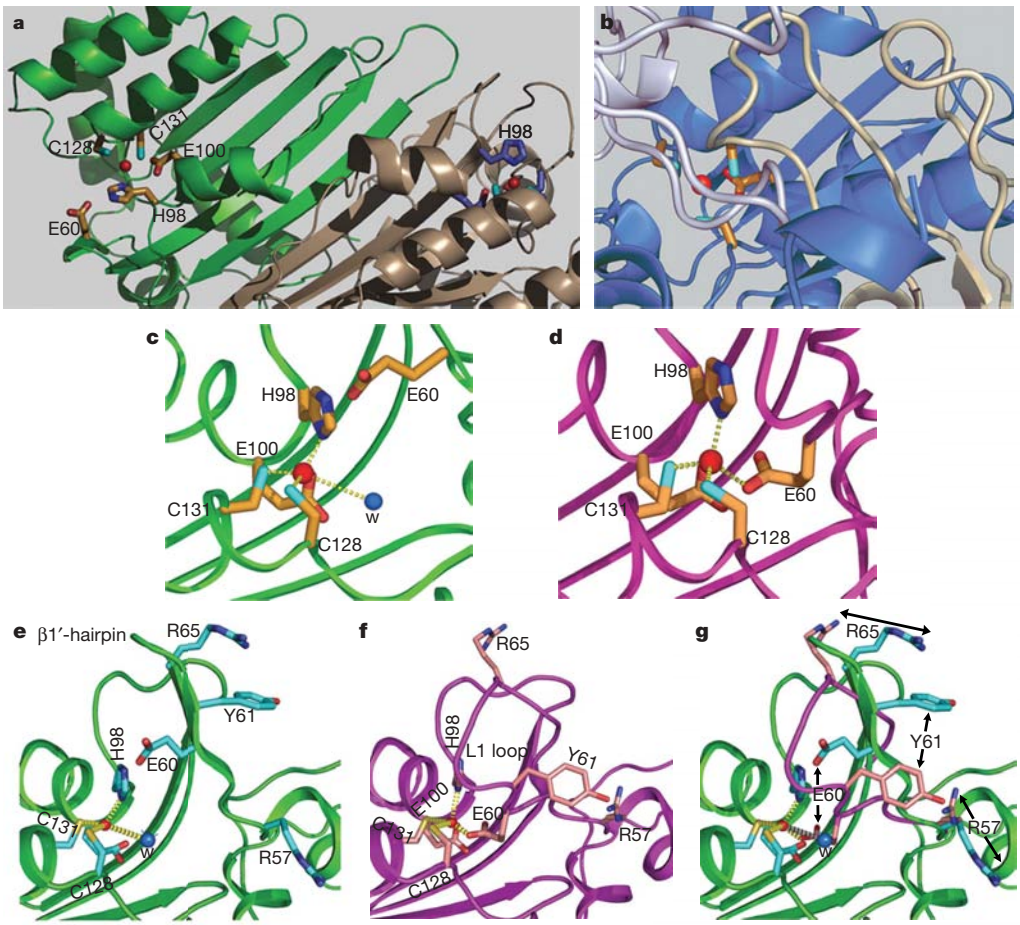

Figure 2 | The APO2 active site. **a**, The APO2 active sites are accessible to DNA/RNA. Red spheres represent Zn. **b**, The fntCDA active site is accessible only to free nucleotides. **c**, The outer APO2 active sites show Zn coordination (yellow dashed lines) by three residues (H98, C128, C131) and a water molecule (blue sphere). **d**, The middle APO2 active centre sites show Zn coordination by a fourth residue, E60. **e**, In the β1′-hairpin structure, the hydrophobic ring of Y61 interacts with the guanidine group of R65, stabilizing the conformation. **f**, In the h1/β1 loop, the E60 coordinates with Zn. Y61 now rotates away from R65 and interacts with R57, facilitating the disruption of the β1′-hairpin and stabilizing the loop conformation. **g**, Superimposed monomers show that the h1/β1 loop (purple) is pulled down ~8.5 Å towards the active site owing to the E60–Zn bond formation.

DNA substrates (Fig. 2a). In the square-shaped fntCDA tetramer, loops from two neighbouring monomers cover the active sites so that only small free nucleotides can bind to the buried sites (Fig. 2b). Although the yeast fntCDA, CDD1, has been reported to deaminate the apolipoprotein B mRNA *in vitro*, its known biological substrate *in vivo* is a free nucleotide and the CDD1 structure is a canonical square-shaped tetramer[9].

In fntCDAs, the active centre Zn atom is coordinated by three residues (either three cysteines, or two cysteines + one histidine) and forms a fourth bond with a water molecule with a bond distance of ~3.0 Å (ref. 10). This type of Zn coordination is also present in APO2 (Fig. 2c), but only in the two outer monomers of the tetramer. Surprisingly, the active sites for the other two monomers in the middle of a tetramer contain an E60 residue, which replaces the water molecule and makes the fourth coordination bond with the Zn (Fig. 2d). This coordination of Zn by four residues is unexpected given that all known fntCDA structures have only three amino acid residues participating in Zn coordination[8–13].

A closer examination of the structure reveals a 'built-in' mechanism for a conformational switch between the two types of Zn coordination. The switch is mediated by sequences contained in the h1/β1-turn, in which E60 is located. The h1/β1-turn can adopt either a hairpin (β1′-hairpin, Fig. 2e) or a loop (L1) conformation (Fig. 2f), which controls whether or not E60 coordinates with Zn. The E60 is located 6 Å from the Zn when the h1/β1-turn is a β1′-hairpin (Fig. 2e). The β1′-hairpin is stabilized by main-chain hydrogen bonds within the β1′-hairpin and reinforced by interactions between Y61 and the guanidine group of R65. In the two middle monomers of an APO2 tetramer, the h1/β1-turn folds into a loop (L1, Fig. 1f). In this conformation, Y61 rotates its side chain to interact with R57 instead of the R65 (Fig. 2f, g). The new pairing of Y61 with R57 destabilizes the β1′-hairpin while stabilizing the loop. In the loop conformation, the E60 is 2.2 Å from the Zn (Fig. 2f).

The hairpin–loop switch may have two important consequences. First, switching to the loop and forming the fourth Zn coordination by E60 prevents coordination by water and subsequent hydroxylation of Zn necessary for deamination (Fig. 2d, f). Second, Zn coordination by E60 pulls the h1/β1-turn approximately 8.5 Å towards the active centre (Fig. 2g), which could restrict substrate access to the active centre. On the other hand, breaking of the fourth coordination of E60 may allow the loop to move away from the active centre to form the β1′-hairpin as observed in the outer monomers. The E60 would no longer prevent the Zn hydroxylation and nucleic acid substrate access to those active sites. Thus, the hairpin–loop switch can be a molecular mechanism for regulating substrate access and enzyme activity mediated through Zn coordination.

The APO2 fragment in the structure shares a 33.3% amino acid identity (44.6% homology) with AID, and the buried residues in APO2 share a 75% identity (96% homology) with AID (Fig. 3a). The highly conserved residues buried inside the structure and those located at the dimer/tetramer interfaces strongly suggest a structural conservation of AID with APO2. Thus, the APO2 crystal structure should provide functional insights for AID and other APOBEC family members, despite of the lack of the known biological activity of APO2. For this reason, we use AID as a surrogate to test how mutations guided by APO2 structure affect AID deamination activity.

We generated glutathione *S*-transferase (GST)–AID mutants with amino acid substitutions located at the tetrameric interface (Fig. 3b), and showed that the mutants either had no detectable or significantly reduced deaminase activity compared to wild-type GST–AID (Fig. 3d–f). Mutants R112C and Y114A/F115A were inactive (Fig. 3e, f), while mutants K16A and C116A had a 3.3-fold reduction in activity (Figs 3e, f).

AID mutations within the predicted dimerization domain, F46A/Y48A (Fig. 3c), displayed a fourfold decrease in deamination activity

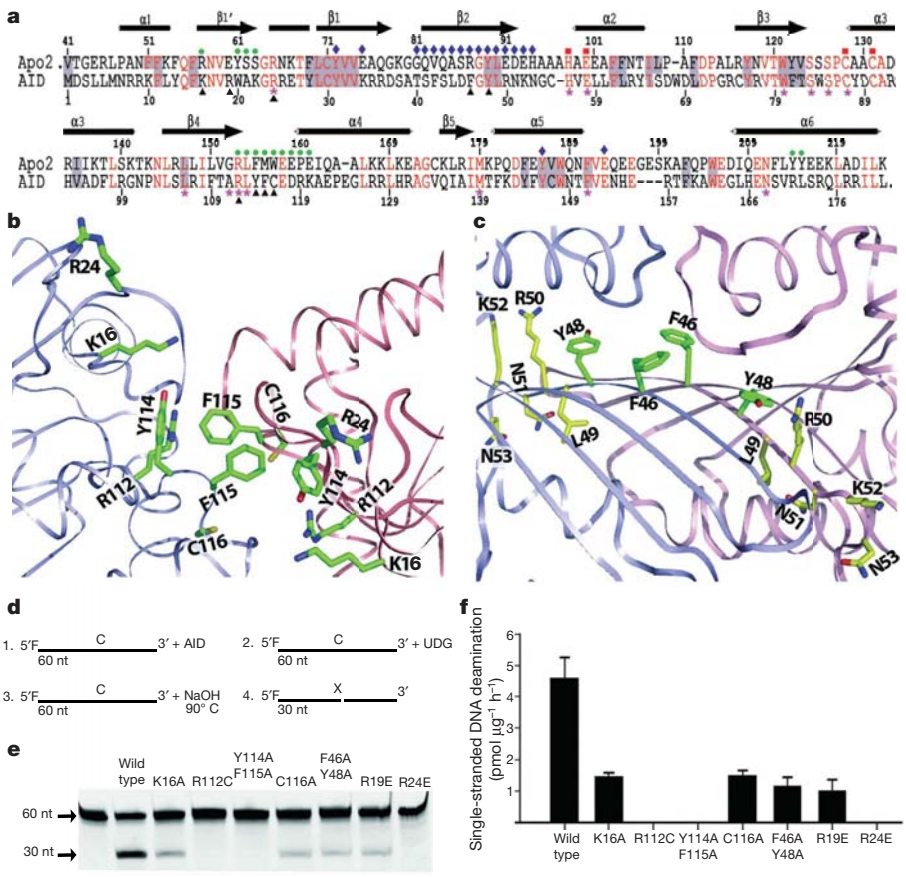

Figure 3 | Structurally guided mutagenesis of AID impairs deamination activity. **a**, A sequence alignment of APO2 and AID, showing significant homology. Red, identical residues; grey shading, buried residues; red squares, active centre residues; green dots, tetrameric interface residues; blue diamonds, dimeric interface residues; purple stars, HIGM mutated residues; and black triangles, mutated AID residues. **b**, Mutated AID residues (in green) at the tetramer interface as modelled based on the APO2 structure. **c**, Mutated AID residues (in green) in the dimer interface as modelled based on the APO2 structure. **d**, A sketch describing the cytidine deamination assay. F, fluorescein; UDG, is uracil DNA glycosylase. **e**, Denaturing PAGE analysis of the deamination activity for wild-type and mutant AID proteins. The 30-nucleotide (nt) band indicates deamination activity. **f**, Bar representation of the specific activities for wild-type and mutant AID proteins.

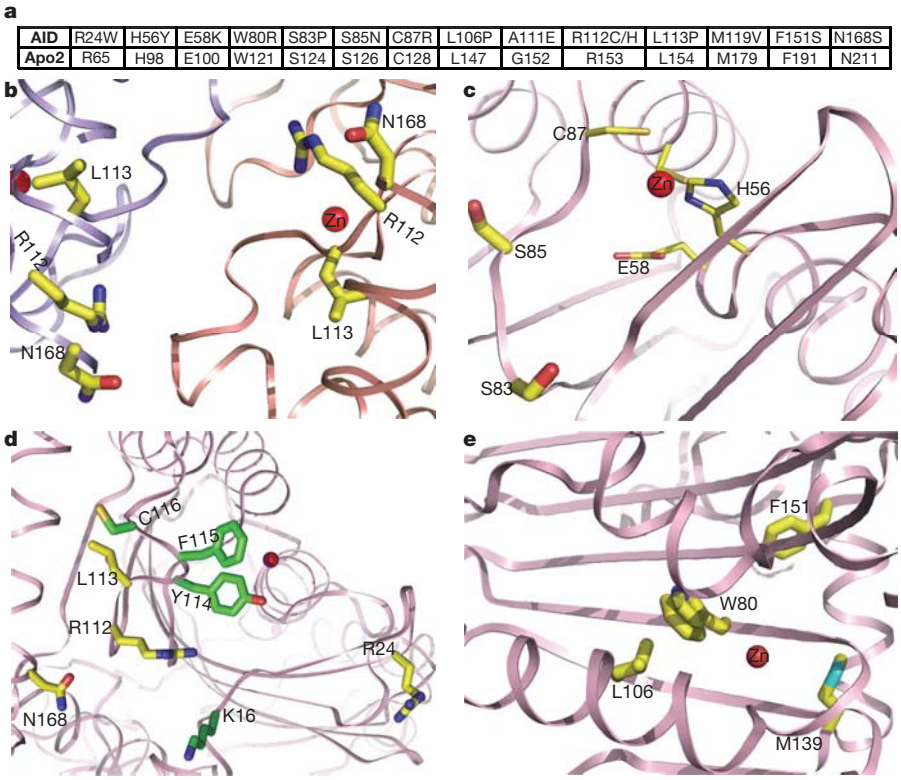

Figure 4 | AID HIGM-2 mutations. a, Alignment of mutated residues of AID from HIGM-2 patients with the corresponding residues in APO2, showing high sequence conservation. **b**, Mapping the residues in AID HIGM-2 mutations (R112, L113, N168) to the tetramer interface as modelled from the APO2 structure. **c**, Mapping the AID HIGM-2 mutations, S83 and S85, near the active site. **d**, Mapping the AID mutations, K16, Y114/F115 and C116 (in green), to the exposed surface of an outer monomer. The HIGM-2 AID residues (R112, L113, N168, in yellow), which are at the tetramer interface (**b**), are also located on this exposed surface. **e**, Mapping of AID HIGM-2 mutations, W80, L106, M139 and F151, to the interior core structure.

(Fig. 3e, f). The dimer interface is extensive, so two mutations should not completely disrupt dimeric AID. This explains why weak deamination activity was observed with this double AID mutant. These mutational results suggest that the residues within the predicted dimeric and tetrameric interfaces are important for deamination activity. One caveat is that the residues on the tetramer interface are also present on the exposed surface of the outer ends of the tetramer and thus could also be involved in an additional role beside tetramerization (Figs 3b, 4d). We noticed that in gel filtration assays the tetramer was a minor species when compared with the dimer, suggesting a stronger dimeric interaction.

AID has an arginine (R19) at the equivalent position of the APO2 E60 residue (Fig 3a) that may have a negative regulatory role for APO2 activity by blocking Zn hydroxylation and substrate access (Fig. 2f). We showed that an AID R19E mutant mimicking APO2 E60 had a significantly decreased deamination activity (about 4.6-fold less than the wild type, Figs 3e, f). Similarly, the AID R24 residue is equivalent to APO2 R65, which interacts with Y61 of APO2 to stabilize the open β1′-hairpin conformation. We predicted that the disruption of the R65–Y61 interaction would collapse the β1′-hairpin into the closed loop conformation to block substrate access and impair deamination activity. Indeed, the AID R24E mutant was completely inactive on single-stranded DNA (Fig. 3e, f).

Mutations in human AID cause hyper-IgM-2 (HIGM-2) syndrome, characterized by an impaired production of high-affinity antibodies[14,15]. The mutated AID residues of HIGM-2 patients are highly conserved in APO2 (Fig. 4a). A plausible explanation for why and how HIGM-2 mutations disrupt AID function is given by the structure of APO2 (Fig. 4b–e). On the basis of the crystal structure, HIGM-2 AID mutations can be divided into four classes. The first mutant class (A111, R112, L113 and N168) occurs at the tetramerization interface (Fig. 4b). The second mutant class includes residues in and near the active centre (Fig. 4c), H56, E58, S83, S85 and C87, which are conserved among all APOBEC enzymes. The AID R24 residue is also mutated in HIGM-2 patients. As previously discussed, R24 may stabilize the β1′-hairpin, which keeps the active site open for DNA/RNA access. The third class consists of residues located on the enzyme surface (Fig. 4d), including those residues located at the tetramer interface (A111, R112, L113, N168; see Fig. 4d). A fourth class of HIGM-2 AID mutations are those with large hydrophobic side chains buried within the core (Fig. 4e), including W80, L106, M139 and F151. Three of these residues are located near the active centre. Mutating these residues should disrupt the folding and stability of AID.

Since many of the APOBEC enzymes are reported to form dimers and multimers, the APO2 structure may shed light on how these enzymes oligomerize[16–23]. The elucidation of the APO2 structure, fortified by the structure-guided predictions for the activity of specific AID mutants, provides a structural basis to pursue further functional studies of APOBEC proteins with an eye towards developing therapeutic strategies to deal with deficiency in deaminating cytidine and to restrict retroviral replication.

METHODS

Crystallography statistics can be found in the Supplementary Information.

Protein purification and crystallization. Human APO2 containing residues 41–224 was cloned and expressed in *Escherichia coli* as a recombinant GST fusion protein. Following GST cleavage by thrombin, further purification of APO2 was achieved using Superdex-75 gel filtration chromatography. Native and selenium-methionine labelled protein was concentrated to 15 mg ml^{-1} in a buffer containing 25 mM Hepes, pH 7.0, 50 mM NaCl and 10 mM dithiothreitol. Crystals were grown at 18 °C by hanging-drop vapour diffusion from a reservoir

solution of 85 mM Na-citrate, pH 5.6, 160 mM $LiSO_4$, 24% (weight/volume) polyethylene glycol monomethyl ether and 15% glycerol.

Structure determination and refinement. Native and selenium-multiwavelength anomalous diffraction (MAD) data were collected at the synchrotron and processed using HKL2000[24] (Supplementary Table 1). An initial solution was obtained at 3.5 Å using the program Solve[25] and located eight selenium atoms using the peak wavelength data set. A search with SHELXD[26] found four additional selenium atoms (totalling twelve) and, subsequently, the program SHARP identified four additional weaker anomalous scattering atoms, which were recognized as Zn atoms. Density modification schemes of solvent flattening and fourfold non-crystallographic symmetry (NCS) averaging were applied using the program RESOLVE[27]. Additionally, phase extension in RESOLVE was performed with the native data set to 2.5 Å resolution using the two-wavelength MAD phases calculated in SHARP. The molecular model was built based on this experimental map using the program "O" and was refined with the Crystallography and NMR System (CNS). A twofold NCS constrain was applied during the initial simulated annealing, but the final refinement was carried out without NCS constrain. The protein geometry is excellent when examined using the program PROCHECK.

Construction of AID mutants. Mutant AID proteins were contructed by site-directed mutagenesis using the pGEX-KG-AID vector as the PCR template and primers specific for the respective mutations (5′-CTG AGG ATC TTC ACC GCG TGC CTC TAC TTC TGT GAG GAC-3′ (R112C), 5′-ATC TTC ACC GCG CGC CTC GCC GCC TGT GAG GAC CGC AAG GCT-3′ (Y114/Y115), 5′-GCG CGC CTC TAC TTC TGT GCG GCC CGC AAG GCT GAG CCC GAG-3′ (E117/E118A), 5′-AAG TTT CTT TAC CAA TTC GCA AAT GTC CGC TGG GCT AAG-3′ (K16A), 5′-ACC GCG CGC CTC TAC TTC GCT GAG GAC CGC AAG GCT GAG-3′ (C116A), 5′-TAC CAA TTC AAA AAT GTC GAG TGG GCT AAG GGT CGG CGT-3′ (R19E), and 5′-ACA TCC TTT TCA CTG GAC GCT GGT GCT CTT CGC AAT AAG AAC GGC-3′ (F46A/Y48A). Mutant constructs were verified by DNA sequencing.

Deamination reactions. Deamination experiments were performed as previously described[28] with the exception that the substrate used was a fluorescein-dT incorporated single-stranded DNA substrate (5′-taa agg fluorescein-dTga aga gag gag aga gaa gta agC tga aga gag aga agg aag aga gtg aag gag-3′). Reaction products were visualized on a BioRad FX scanner.

Received 10 October; accepted 28 November 2006.
Published online 24 December 2006.

1. Pham, P., Bransteitter, R. & Goodman, M. F. Reward versus risk: DNA cytidine deaminases triggering immunity and disease. *Biochemistry* **44,** 2703–2715 (2005).
2. Conticello, S. G., Thomas, C. J. F., Petersen-Mahrt, S. K. & Neuberger, M. S. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol. Biol. Evol.* **22,** 367–377 (2004).
3. Bransteitter, R., Sneeden, J. L., Allen, S., Pham, P. & Goodman, O. M. F. First AID (activation-induced cytidine deaminase) is needed to produce high affinity isotype-switched antibodies. *J. Biol. Chem.* **281,** 16833–16836 (2006).
4. Chiu, Y. L. & Greene, W. C. Multifaceted antiviral actions of APOBEC3 cytidine deaminases. *Trends Immunol.* **27,** 291–297 (2006).
5. Cullen, B. R. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J. Virol.* **80,** 1067–1076 (2006).
6. Franca, R., Spadari, S. & Maga, G. APOBEC deaminases as cellular antiviral factors: a novel natural host defense mechanism. *Med. Sci. Monit.* **12,** RA92–RA98 (2006).
7. Bonvin, M. *et al.* Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. *Hepatology* **43,** 1364–1374 (2006).
8. Johansson, E., Mejlhede, N., Neuhard, J. & Larsen, S. Crystal structure of the tetrameric cytidine deaminase from *Bacillus subtilis* at 2.0 Å resolution. *Biochemistry* **41,** 2563–2570 (2002).
9. Xie, K. *et al.* The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1. *Proc. Natl Acad. Sci. USA* **101,** 8114–8119 (2004).
10. Teh, A. *et al.* The 1.48 Å resolution crystal structure of the homotetrameric cytidine deaminase from mouse. *Biochemistry* **45,** 7825–7833 (2006).
11. Chung, S. J., Fromme, J. C. & Verdine, G. L. Structure of human cytidine deaminase bound to a potent inhibitor. *J. Med. Chem.* **48,** 658–660 (2005).
12. Betts, L., Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W. Cytidine deaminase. The 2.3 Å crystal structure of an enzyme: transition-state analog complex. *Curr. Biol.* **235,** 635–656 (1994).
13. Smith, A. A., Carlow, D. C., Wolfenden, R. & Short, S. A. Mutations affecting transition-state stabilization by residues coordinating zinc at the active site of cytidine deaminase. *Biochemistry* **33,** 6468–6474 (1994).
14. Durandy, A., Peron, S. & Fischer, A. Hyper-IgM syndromes. *Curr. Opin. Rheumatol.* **18,** 369–376 (2006).
15. Minegishi, Y. *et al.* Mutations in activation-induced cytidine deaminase in patients with hyper IgM syndrome. *Clin. Immunol.* **97,** 203–210 (2000).
16. Anant, S. *et al.* ARCD-1, an apobec-1-related cytidine deaminase, exerts a dominant negative effect on C to U RNA editing. *Am. J. Cell Physiol.* **281,** C1904–C1916 (2001).
17. Jarmuz, A. *et al.* An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79,** 285–296 (2002).
18. Shindo, K. *et al.* The enzymatic activity of CEM15/Apobec-3G is essential for the regulation of the infectivity of HIV-1 virion but not a sole determinant of its antiviral activity. *J. Biol. Chem.* **278,** 44412–44416 (2003).
19. Wiegand, H. L., Doehle, B. P., Bogerd, H. P. & Cullen, B. R. A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J.* **23,** 2451–2458 (2004).
20. Opi, S. *et al.* Monomeric APOBEC3G is catalytically active and has antiviral activity. *J. Virol.* **80,** 4673–4682 (2006).
21. Navarro, F. *et al.* Complementary function of the two catalytic domains of APOBEC3G. *Virology* **333,** 374–386 (2005).
22. Wang, J. *et al.* Identification of a specific domain required for dimerization of activation-induced cytidine deaminase. *J. Biol. Chem.* (in the press).
23. Teng, B. *et al.* Mutational analysis of apolipoprotein B mRNA editing enzyme (APOBEC1): structure-function relationships of RNA editing and dimerization. *J. Lipid Res.* **40,** 623–635 (1999).
24. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
25. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr.* **55,** 849–861 (1999).
26. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr.* **58,** 1772–1779 (2002).
27. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr.* **56,** 965–972 (2000).
28. Bransteitter, R., Pham, P., Scharff, M. D. & Goodman, M. F. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl Acad. Sci. USA* **100,** 4102–4107 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Chen for comments on the manuscript. We also thank G. Wang from Chen lab and the staff at ALS LBL8.2.1, BL8.2.2 and APS 19id in the Argonne National Laboratory for assistance in data collection. This work was supported in part by National Institutes of Health grants to M.F.G. and X.S.C. and an NIH-NIA Predoctoral Traineeship to R.B.

Author Information The structure of APO2 has been uploaded to the Protein Data Bank under accession number 2NYT and to the Research Collaboratory for Structural Bioinformatics (RCSB) under accession number RCSB040471. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to X.S.C. (xiaojiang.chen@usc.edu).

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 25, pp. 17428–17439, June 20, 2008

Impact of Phosphorylation and Phosphorylation-null Mutants on the Activity and Deamination Specificity of Activation-induced Cytidine Deaminase*[S]

Received for publication, March 17, 2008 Published, JBC Papers in Press, April 16, 2008, DOI 10.1074/jbc.M802121200

Phuong Pham[‡], Marcus B. Smolka[§], Peter Calabrese[‡], Alice Landolph[‡], Ke Zhang[¶], Huilin Zhou[§], and Myron F. Goodman[‡1]

From the [‡]Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, California 90089-2910, the [§]Ludwig Institute for Cancer Research, Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California 92093, and the [¶]Division of Clinical Immunology/Allergy, Department of Medicine, UCLA School of Medicine, Los Angeles, California 90095

Activation-induced cytidine deaminase (AID) initiates somatic hypermutation and class switch recombination in B cells by deaminating C → U on transcribed DNA. Here we analyze the role of phosphorylation and phosphorylation-null mutants on the biochemical behavior of AID, including enzyme specific activity, processivity, deamination spectra, deamination motif specificity, and transcription-dependent deamination in the presence and absence of RPA. We show that a small fraction of recombinant human AID expressed in *Sf*9 insect cells is phosphorylated at previously identified residues Ser[38] and Thr[27] and also at Ser[41] and Ser[43]. S43P AID has been identified in a patient with hyper-IgM immunodeficiency syndrome. Ser-substituted phosphorylation-null mutants (S38A, S41A, S43A, and S43P) exhibit wild type (WT) activity on single-stranded DNA. Deamination of transcribed double-stranded DNA is similar for WT and mutant AID and occurs with or without RPA. Although WT and AID mutants catalyze processive deamination favoring canonical WRC hot spot motifs (where W represents A/T and R is A/G), their deamination spectra differ significantly. The differences between the WT and AID mutants appear to be caused by the replacement of Ser as opposed to an absence of phosphorylation. The spectral differences reflect a marked change in deamination efficiencies in two motifs, GGC and AGC, which are preferred by mutant AID but disfavored by WT AID. Both motifs occur with exceptionally high frequency in human switch regions, suggesting a possible relationship between AID deamination specificity and a loss of antibody diversification.

AID, a B-cell specific protein, is required for tightly regulated mechanisms of Ig antibody diversification, somatic hypermutation, class switch recombination, and gene conversion. SHM[2] is characterized by an exceptionally high mutation rate $\sim 10^{-3}$ to 10^{-4} per base pair per cell division within the V(D)J rearranged Ig genes (1). CSR is a unique region specific event in which the donor switch region upstream of the IgH C_μ exon (coding for an IgM antibody) recombines with one of the switch regions that is 5′ to each of the downstream C exons, such as C_γ, C_α, or C_ϵ, producing IgG, IgA, or IgE antibody isotypes, respectively (2).

Typically, under tight transcriptional control, AID is induced in activated B cells within germinal centers, causing deamination of cytosine residues in the variable and switch regions (V- and S-regions, respectively) of transcribed Ig loci. Deaminated DNA is subsequently replicated or repaired by different cellular repair mechanisms to give rise to diversified isotype-switched and antigen-specific high affinity antibodies (for reviews, see Refs. 3–5).

Biochemical studies have shown that purified native B-cell AID or recombinant AID protein, expressed in *Sf*9 insect cells or *Escherichia coli*, deaminates C residues on ssDNA but not dsDNA, single-stranded RNA, or DNA/RNA hybrid molecules (6–9). AID acts processively on ssDNA and has a distinctive deamination specificity, favoring C targets in WRC hot spot motifs (where W represents A or T, and R is purine) while avoiding SYC cold spots (where S represents C or G and Y is pyrimidine) (10, 11). These intrinsic biochemical properties of AID are likely to be significant determinants of SHM hallmark features in B-cells, favoring mutations at WRC hot spot motifs and broad clonal heterogeneity of mutations within V-regions (12–14).

Transcription is required for both SHM and CSR to occur (15–17). Since AID does not appear to act on dsDNA, transcription probably plays an essential role in generating transient ssDNA, its principal substrate. The ability of AID to deaminate C residues on dsDNA undergoing active transcription has been demonstrated in cells and cell-free assays. For example, ectopically expressed AID can induce mutations on highly transcribed target gene in fibroblasts (18) and *E. coli* (9,

* This work was supported, in whole or in part, by National Institutes of Health Grants ES013192 and R37GM21422 (to M. F. G.) and Grant GM080469. This work was also supported by the Ludwig Institute for Cancer Research (to H. Z. and M. B. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1 and 2.

[1] To whom correspondence should be addressed: Depts. of Biological Sciences and Chemistry, RRI-201, University of Southern California, Los Angeles, CA 90089-2910. Tel.: 213-740-5190; Fax: 213-821-1138; E-mail: mgoodman@usc.edu.

[2] The abbreviations used are: SHM, somatic hypermutation; AID, activation-induced cytidine deaminase; WT, wild type; CSR, class switch recombination; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; HIGM-2, hyper-IgM syndrome; MI, mutability index; nt, nucleotide(s).

19), and purified AID expressed in various sources has been shown to act on dC targets on linear dsDNA and closed circular dsDNA model substrates when transcribed by a prokaryotic RNA polymerase (7, 10, 11, 20). AID has been shown to interact with human RNA polymerase II (21), although there are currently no biochemical data for AID acting in conjunction with the human transcription machinery to deaminate dsDNA.

Native AID isolated from stimulated primary B-cell nuclei is phosphorylated on Ser^{38} and Tyr^{184} residues (22, 23), where Ser^{38} phosphorylation occurs on about 5–15% of total AID protein (23). Phosphorylation does not appear to be B-cell-specific, since a similar or greater level of Ser^{38}-phosphorylated AID was observed when expressed in 293T and 3T3-NTZ cells (23). Protein kinase A (PKA) was found to interact with AID (22, 24) and appears to be a primary kinase responsible for phosphorylation at Ser^{38} and possibly at Thr^{27} residues *in vivo* and *in vitro* (22, 24).

The potential role of phosphorylation of AID in SHM and CSR was deduced primarily from studies of AID mutations at phosphorylated residues Thr^{27}, Ser^{38}, and Tyr^{184}. A phosphorylation-defective mutant, Y184A, did not affect CSR (22). In contrast, a mutation at S38A reduced CSR dramatically to no more than 20% of the wild type (22, 24) or moderately (35–80%) (23, 25) in *ex vivo* stimulated B-cells. AID S38A caused a significant decrease of SHM in B-cells (23) and SHM and gene conversion in chicken DT40 cells (26). Phosphorylated Ser^{38} was reported to be required for interaction with the 32-kDa subunit of RPA (replication protein A) (27). AID S38A did not interact with RPA; nor did it catalyze deamination of C on dsDNA in a model T7 RNA polymerase transcription system (22, 27).

There are data, however, indicating a lesser role of Ser^{38} phosphorylation in regulating AID activity. Zebrafish AID, which lacks a serine at a position corresponding to human or mouse Ser^{38}, is fully active in supporting SHM and CSR (28, 29) and gene conversion (26). Recombinant AID expressed in *Sf*9 cells and *E. coli* was found to be active on transcribed dsDNA substrates in the absence of RPA (9–11, 30, 31). Thus, although it is clear that a minority of AID molecules are phosphorylated at Ser^{38}, it is considerably less clear how the biochemical behavior of AID is influenced by either the location of phosphorylated residues or whether or not a particular residue is phosphorylated.

The roles of phosphorylation on the properties of AID *in vitro* are amenable to a comparative biochemical analysis using wild type (WT) AID and phosphorylation-defective mutants. Here we have investigated how replacing individual "phosphorylation-active" Ser residues influences AID specific activity, processivity, deamination specificity, transcriptional-dependent deamination, including its interactions with RPA, and mutational spectra. One such mutant, S43P, has been identified in humans diagnosed with hyper-IgM (HIGM-2) syndrome (32), characterized by the absence of CSR. A comparison of deamination motif specificities favored by S43P in relation to WT AID suggests a possible connection of AID deamination specificity with human immunodeficiency disease.

EXPERIMENTAL PROCEDURES

Enzymes and Substrates—Mutant AID proteins (S38A, S38D, S41A, S41D, S43A, and S43P) were constructed by site-directed mutagenesis (QuikChange site-directed mutagenesis kit, Stratagene) using the pAcG2T-AID vector (6, 10) as the template. Recombinant baculoviruses encoding WT and mutant AID were generated according to the recommended protocol (BD Bioscience). WT and mutant GST-AID proteins were expressed and purified as described previously (10, 11), with an additional of Halt phosphatase inhibitor mixture (Pierce) in the lysis buffer. AID proteins were dialyzed in a buffer containing 20 mM Tris-HCl (pH 7.5), 25 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, and 10% glycerol and stored at −80 °C. *E. coli* single-stranded binding protein and recombinant human RPA were overexpressed in *E. coli* and purified according to published protocols (33, 34). T7 RNA polymerase was purchased from Promega, and ultrapure NTP was purchased from Amersham Biosciences. The active, recombinant catalytic subunit of human protein kinase A was purchased from Calbiochem. M13mp2 gapped DNA and M13mp2T7 covalently closed circular dsDNA substrates were prepared as described (10, 11).

Mass Spectrometry Analysis—Approximately 20 μg of purified AID were reduced with 5 mM dithiothreitol in the presence of 1% SDS and subsequently alkylated with 20 mM iodoacetamide. Protein was then precipitated by the addition of 3 volumes of a solution containing 50% acetone, 49.9% ethanol, and 0.1% acetic acid; the sample was kept on ice for 15 min and then centrifuged for 10 min at maximum speed on a bench top centrifuge. The pellet containing precipitated AID protein was resuspended in 40 μl of 8 M urea, 100 mM Tris-HCl, pH 8.0. The suspension was then diluted by adding 160 μl of 150 mM NaCl, 50 mM Tris-HCl, pH 8.0, and the protein was digested with 2 μg of trypsin (Promega) overnight. After digestion, 0.4% trifluoroacetic acid was added, and the peptides were desalted using a 50-mg C18-Sep-Pak cartridge (Waters). Phosphopeptides were purified by IMAC as previously described (35). Purified phosphopeptides were analyzed by microcapillary-liquid chromatography-electrospray ionization-tandem mass spectrometry on a Thermo Finnigan LTQ quadrupole ion trap mass spectrometer, as described (35). For data analysis, the SEQUEST (version 3.4 beta 2) program running on a Sorcerer system (SageN, San Jose, CA) was used for peptide identification. A data base search was performed using a sub-data base consisting of 50 budding yeast proteins in addition to the human AID sequence. The following variable modifications were considered: +80 Da (phosphorylation) for serine, threonine and tyrosine residues; +16 Da (oxidation) for methionine residues. Up to four variable modifications were allowed per peptide, and the peptide mass tolerance used was 3 Da. A semitryptic restriction was applied, and only the top-matched peptides with a probability score above 0.9 were subsequently considered for close inspection. Each MS/MS spectrum that led to a phosphopeptide identification was manually verified to confirm that all significant ions were accounted for and then validated.

Measurements of Deamination-specific Activity on ssDNA—Specific activities of WT AID and phosphorylation mutants were

measured using ^{32}P-labeled 36-nt ssDNA 5′-AGAAA-AGGGGAAAGCAAAGAGGAAAGGTGAGGAGGT-3′.

Reactions were carried out in a buffer containing 50 mM HEPES (pH 7.5), 1 mM dithiothreitol, 10 mM $MgCl_2$ in the presence of 500 fmol of the substrate DNA, 200 fmol of recombinant GST-AID (*Sf*9-expressed) or 6400 fmol of *E. coli* expressed GST-AID, and 20 ng of RNase A. Following incubation at 37 °C for 5 min, the reactions were quenched by a double extraction with phenol/chloroform/isoamyl alcohol (25:24:1), and the deamination product was analyzed as described previously (6). Specific activities were calculated as amount (fmol) of deaminated substrate/min/μg of enzyme.

Deamination Activity of WT and Mutant AID on Transcribed dsDNA—Transcription-dependent AID deamination on the nontranscribed strand of dsDNA was measured using a linear dsDNA, undergoing active transcription by T7 RNA polymerase (10). In a typical reaction (50 μl), GST-AID (7 pmol of *Sf*9-expressed WT or mutant AID or 50 pmol of *E. coli*-expressed AID), RNase A (200 ng), dsDNA substrate (1 pmol), and T7 RNA polymerase (1 μl) in a reaction buffer containing HEPES (50 mM, pH 7.5), dithiothreitol (5 mM), $MgCl_2$ (10 mM), and NTPs (250 μM) were incubated at 37 °C for 30 min. *E. coli* SSB or human RPA (3, 6, or 12 pmol), if present, were added to the reactions as indicated. The reactions were stopped by extracting twice with phenol/chloroform/isoamyl alcohol (25:24:1). The deaminated products were analyzed using a primer elongation-dideoxynucleotide termination assay (10). Thermo Sequenase (U.S. Biochemical Corp.) was used to extend an 18-mer ^{32}P-labeled primer annealed to the target strand in the presence of three dNTPs and either ddATP or ddGTP (80 μM each). The reactions were carried out for seven cycles (95 °C for 30 s, 55 °C for 45 s, 72 °C for 1 min) and terminated by adding an equal volume of stop solution containing 95% formamide and 20 mM EDTA. The reaction products were resolved by 19% polyacrylamide denaturing gel electrophoresis and analyzed by phosphorimaging. Deamination efficiencies were calculated from extension reactions with the ddA mix as a ratio of the band intensity opposite the converted U template site compared with integrated band intensities opposite and beyond the C template site. The efficiencies were also calculated from extension reactions with ddG as a ratio of integrated band intensities beyond the C template site to the integrated band intensities opposite and beyond the C template site.

Mutation Analysis of AID-targeted C Deamination in Vitro—Deamination specificities of WT and mutant AID were measured using the following reaction conditions: 30-μl volume, 50 mM HEPES (pH 7.5), dithiothreitol (1 mM), $MgCl_2$ (10 mM), gapped DNA (500 ng), RNase A (200 ng), and WT or mutant AID (50 to 100 ng of *Sf*9-expressed AID or 1 μg of *E. coli*-expressed AID). Following incubations for 2.5, 5, and 10 min at 37 °C, the reactions were quenched by a double extraction with phenol/chloroform/isoamyl alcohol (25:24:1). Conversions of C → U on the DNA substrate were detected as white or light blue plaques, indicating C → T mutations in a *lacZα* target gene after transfection into uracil glycosylase-deficient (*ung*$^-$) *E. coli*, as previously described (10).

WT and mutant AID deamination spectra were compared using χ^2 tests with the null hypothesis that the distribution of deaminations at the nucleotide positions is the same for two enzymes. We consider a contingency table where the two columns are the WT and mutant AID, the rows are the nucleotide sites, and the numbers in the cells are the deaminations corresponding to that column and row. In order to deal with the problem of small denominators, we only consider those sites for which the sum of the deaminations for the two enzymes is 5 or greater (we have tried other thresholds and reached the same conclusions).

In order to determine whether a motif is deaminated differently by the WT or mutant enzymes, we performed the following test. For each WT or mutant enzyme (S38A, S41A, S43A, or S43P), we calculated the "site mutability index" by dividing the number of deaminations at each site by the total number of deaminations. For the set of sites sharing the motif, we compared the WT index with each of the four mutant AID indexes by the paired Wilcox test. We used all four mutants, since their spectra are similar to each other (but not the WT), and there are a relatively small number of sites for each motif. Since we repeated this test for each of the 16 motifs, the *p* values in Table 2 have been corrected for multiple tests. The reported value is the probability that if 16 independent tests have been performed, then at least one will be as extreme as observed: $c = 1 - (1 - p)^{16}$, where *c* is the correction of *p* from the Wilcox test.

Processivity Deamination Analysis of WT and Mutant AID—An 85-nt ssDNA substrate with two identical hot spot AGC motifs, containing a fluorescein tag located between the two motifs, was used. Deamination reactions (45-μl volume) were carried out in the presence of 15 pmol of ssDNA substrate, 5–15 pmol of WT or mutant GST-AID, and 200 ng of RNase A for 2.5, 5, and 10 min. After treatment with 4 units of uracil DNA glycosylase (New England Biolabs) and alkali, the deamination products were separated on a 16% denaturing polyacrylamide gel, visualized, and quantified using a FX fluorescence scanner (Bio-Rad). Analysis of processive deamination by a single WT and mutant AID enzyme was carried out as described (36, 37). The single deamination rates are approximately equal at the 5′-C and 3′-C target motifs. The correlated double deamination efficiency is calculated as the probability that a single AID molecule deaminates both 5′ and 3′ target motifs on one ssDNA substrate during a single enzyme-DNA encounter.

RESULTS

Human AID Expressed in Sf9 Cells Is Partially Phosphorylated at Ser Residues 38, 41, and 43 and at Thr^{27}—Mass spectral analysis demonstrates that WT recombinant AID expressed in *Sf*9 insect cells is phosphorylated at four predicted protein kinase motifs; two motifs are the previously identified Ser^{38} and Thr^{27} residues (22, 24), and two newly identified residues are at Ser^{41} and Ser^{43} (Fig. 1). Individual peptide fragments of AID showed that just one of the four residues was phosphorylated. Multiple phosphorylations of individual proteins were not detected despite using an isolation technique that strongly favored the detection of multiply phosphorylated peptide fragments (Fig. 1). These three Ser residues 38, 41, and 43 along with Thr^{27} are predicted to be phosphorylation sites for protein kinase PKA or CAMK2, protein kinase Cδ, CK1, and PKA, respectively (Web Scansite 2.0, medium stringency scan (38)).

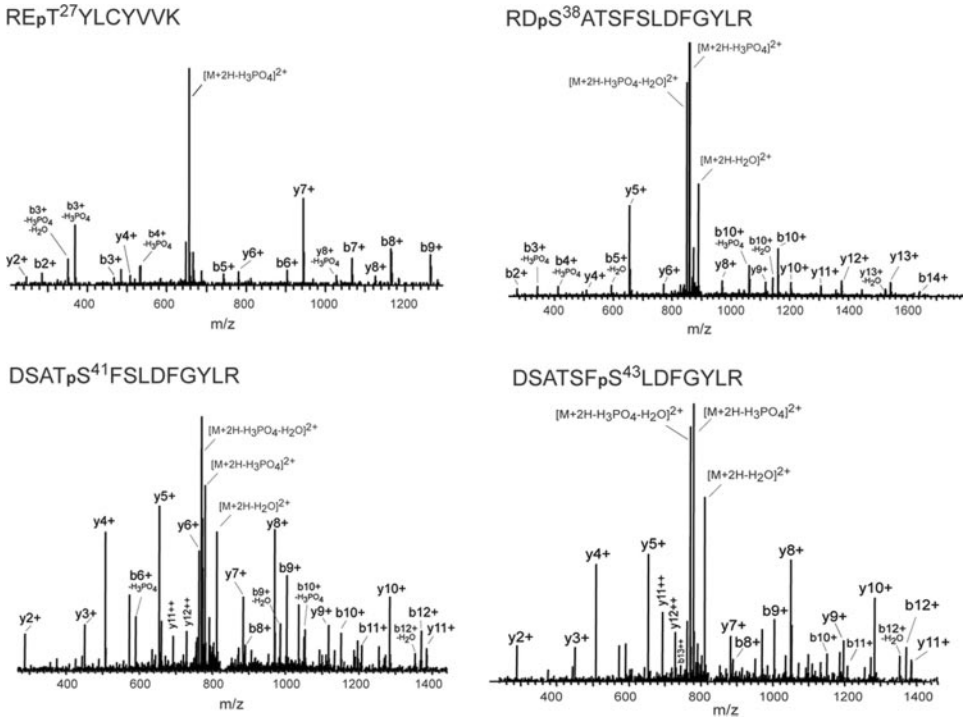

FIGURE 1. **Human AID expressed in *Sf*9 cells is phosphorylated at four sites.** Individual phosphorylation at threonine 27 and serines 38, 41, and 43 was identified by mass spectrometry of tryptic phosphopeptides purified by IMAC. MS/MS spectra of four distinct doubly charged phosphopeptides of AID are shown. In each case, the phosphorylated residue identified is indicated in the sequence (*pS* or *pT*) and is followed by a *number* referring to the amino acid position in the protein. In the spectrum, the peaks corresponding to the detected b-ions (N-terminal fragment) and y-ions (C-terminal fragments) are labeled accordingly. Also, peaks corresponding to the neutral loss of water or phosphate are labeled. Only monophosphorylated peptides were identified.

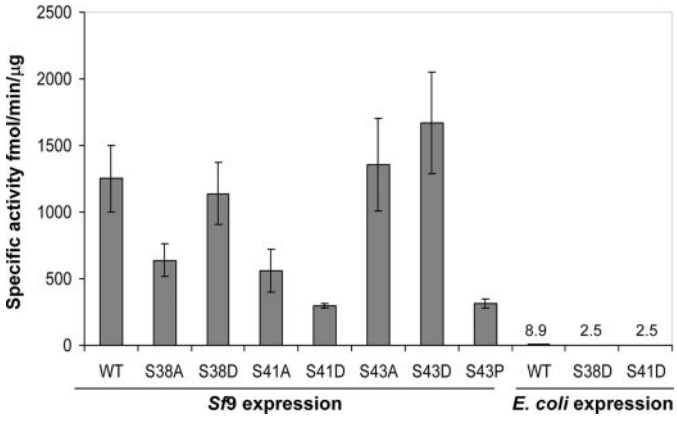

FIGURE 2. **Specific activity of WT and mutant AID proteins expressed in *Sf*9 insect cells or in *E. coli*.** Deamination-specific activities were measured using 36 nt of ^{32}P-labeled ssDNA containing a single C target in an AGC hot spot motif. Reactions were carried out for 5 min at 37 °C in the presence of ssDNA (500 fmol) and recombinant GST-AID (200 fmol) expressed in *Sf*9 cells or with of GST-AID (6400 fmol) expressed in *E. coli*. Due to its extremely low specific activity, the *E. coli*-expressed AID was present at ~32-fold higher concentration than *Sf*9-expressed AID. The *error bars* represent ±1 S.E. from at least three independent measurements.

Based on the amount of protein used in the analysis, we estimate that the phosphorylated species represents less than a few percent of the total AID. The detection of phosphorylated AID required that the enzyme isolation and purification be carried out in the presence of a mixture of phosphatase inhibitors (see "Experimental Procedures"). When phosphatase inhibitors are omitted from the lysis and purification buffers, we were unable to detect any AID phosphorylation by mass spectral analysis (data not shown). Phosphorylation was also not detected for human AID expressed in *E. coli* (data not shown).

Deamination of ssDNA by Phosphorylation-defective AID Mutants—To address whether phosphorylation might regulate AID activity, we constructed single mutants at residues 38, 41, or 43 by replacing Ser with Ala (*i.e.* "phosphorylation null" mutants) or by replacing Ser with Asp, which confers a −1 charge. Deamination measurements were made in the linear range of reaction times and protein concentrations, using the same protein purification protocol (see "Experimental Procedures").

Each AID mutant expressed in insect cells is active on ssDNA, behaving similarly to WT AID within a factor of 3 (Fig. 2). S43A has about the same specific activity as WT AID, whereas S38A and S41A are roughly 2-fold less active. The S38D and S43D mutants are similar to WT AID, whereas S41D and S43P are about 2.5-fold less active. In contrast, human AID expressed in *E. coli* is ~100-fold less active than *Sf*9-expressed AID (Fig. 2). The much lower specific activity of *E. coli*-expressed AID is not likely to be caused by the absence of phosphorylation. Notably, *Sf*9-expressed AID purified in the absence of phosphatase inhibitors showed no detectable amount of phosphorylation but behaved indistinguishably from phosphorylated AID with regard to specific activity and in each biochemical measurement performed below (data not shown). We conclude that the ability of AID to deaminate ssDNA does not depend on phosphorylation; nor does the activity appear to require charged amino acid residues at positions 38, 41, and 43.

Substantial Changes in WT AID Deamination-induced Mutation Spectra Are Caused by Amino Acid Substitutions at Serines 38, 41, and 43—AID was incubated *in vitro* with phage M13mp2, a circular DNA substrate containing a *lacZα* reporter target sequence in a 365-nt single-stranded gapped region (Fig. 3*a*). Deaminations are detected as C → T mutations occurring in DNA isolated from individual phage progeny following transfection of AID-treated DNA into *ung*$^-$ *E. coli* (10). The ratio of mutant phage (clear or light blue plaques) to wild type phage (dark blue plaques) is <5–7%, which ensures that each mutant phage reflects the action of a single AID molecule acting on M13 DNA (Table 1). Conditions that limit M13 DNA targeting to a single AID molecule are satisfied at short (2.5-min) and long (10-min) incubations with AID. We have verified that C → T mutations are absent in DNA isolated from wild type

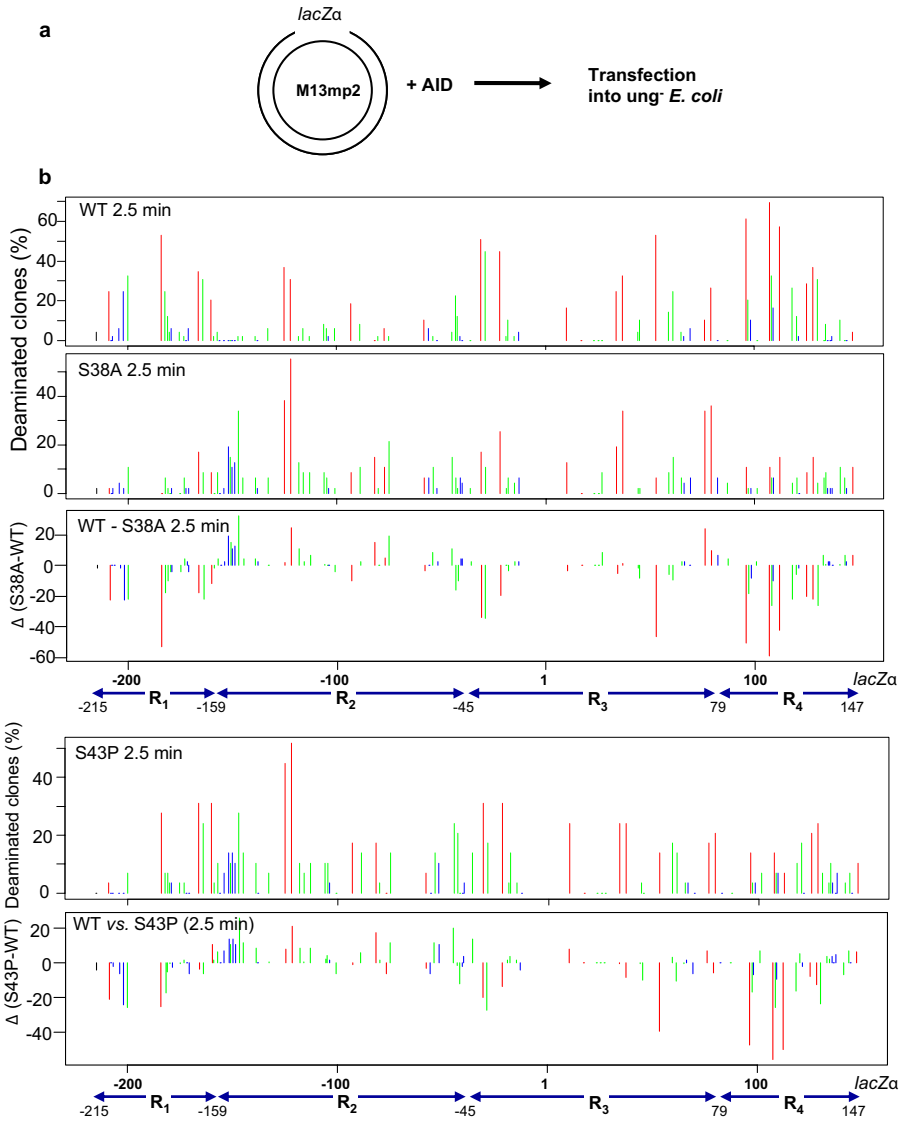

FIGURE 3. **Deamination spectra of WT AID, S38A, and S43P phosphorylation-null mutant AID.** *a*, *sketch* of a protocol to detect C → T mutations resulting from AID-catalyzed C deaminations occurring on individual ssDNA substrate molecules in a gapped M13 DNA construct. The ssDNA gap region is 365 nucleotides long and contains 230 nucleotides of a *lacZα* reporter sequence. Mutations in *lacZα* are identified as *clear* or *light blue plaques*; nonmutated phage appear as *dark blue plaques* (see "Experimental Procedures"). *b*, C → U deamination spectra and the identification of C → T mutations in DNA clones after incubation with WT, S38A, and S43P AID for 2.5 and 10 min were obtained by sequencing about 50 individual mutant clones. Each *colored bar* represents a percentage of mutated phage clones with a C → T mutation at the indicated position on the *lacZα* target sequence (−215 to +149). *Red bars* identify C deaminations occurring in 5′-WRC hot spot motifs, *blue bars* represent 5′-SYC cold spot motifs, and *green bars* represent intermediate motifs (neither WRC nor SYC). WT − S38A and WT − S43P "difference spectra" at 2.5 min and 10 min represent side-by-side comparison of deamination spectra of WT and S38A and WT and S43P, respectively. The *bars* in the difference spectra represent the differences in deamination frequencies at each deaminated site by in the mutant AID spectrum compared with the ones of WT (deamination frequencies in the mutant AID spectrum minus deamination frequencies of WT). *Bars below 0* represent excess deaminations for WT AID, and *bars above 0* denote excess mutant AID deaminations.

phage by direct sequencing of individual DNA clones isolated from dark blue plaques (10).

There are pronounced differences in C → T mutational spectra engendered by differences in the deamination patterns comparing WT AID with S38A and S43P AID mutants, whereas the mutants closely resemble each other (Fig. 3*b*). These differences occur although the WT and mutant enzymes clearly favor deaminations in WRC hot spot motifs and disfavor deaminations in SYC cold spot motifs (Table 2). For both WT and mutant AID, the shorter incubation period (2.5 min) shows a combination of strongly favored deaminations in WRC hot spot motifs (Fig. 3*b* (*red bars*), Table 2, and supplemental Fig. 1) relative to SY*C* cold spot and intermediate deamination motifs (Fig. 3*b*, *blue bars* and *green bars*, respectively). A comparison of WT and mutant AID spatial deamination patterns can be made by dividing the 365-nt ssDNA gapped region into four regions: R_1 (nucleotide positions −215 to −159), R_2 (−158 to −45), R_3 (−44 to +79), and R_4 (+80 to +147). There is a large disparity in the clonal deamination patterns in R_1, R_2, or R_4, comparing mutant S38A and S43P with WT AID (Fig. 3*b*).

To evaluate the spatial differences in deamination quantitatively, we have computed a difference spectrum (Fig. 3*b*, Δ) by subtracting the frequency of deaminated clones in the WT AID spectra from those of S38A and S43P AID at every target C site in the 365-nt gap (Fig. 3*b*). *Lines* extending *below 0* represent excess deaminations for WT AID, and *lines above 0* denote excess S38A and S43P mutant AID deaminations (Fig. 3*b*). The difference spectra reveal a much higher probability of WT AID-catalyzed deaminations, especially in WRC motifs, occurring in R_1, somewhat less so in R_3, and much more so in R_4. In contrast, mutant AID-catalyzed deaminations are clearly favored in R_2 (Fig. 3*b*). A χ^2 test (see "Experimental Procedures") provides strong statistical support that the WT and mutant spectra are significantly different: at 2.5 min, S38A ($p < 2.2 \times 10^{-16}$), S43P ($p < 2.6 \times 10^{-9}$), S41A ($p < 2.2 \times 10^{-16}$), S43A ($p < 2.2 \times 10^{-16}$); at 10 min, S38A ($p < 2.2 \times 10^{-16}$), S43P ($p < 7.4 \times 10^{-4}$), S41A ($p < 4.6 \times 10^{-7}$), S43A ($p < 2.2 \times 10^{-16}$).

For WT AID, there are numerous mutations in WRC motifs (Fig. 3*b*, *red bars*) in the R_1 region, with about 50% of the clones having mutations at C template position −184 and ~18–22% of the clones with mutations at positions −209, −166, and −160 (Fig. 3*b*). Notably, S38A and S43P AID poorly deaminate WRC motifs at sites −209 and −184, which correspond to the two favored hot spot sites in R_1 in the wild type AID spectrum. A second difference is that in the beginning of R_2, a cluster of

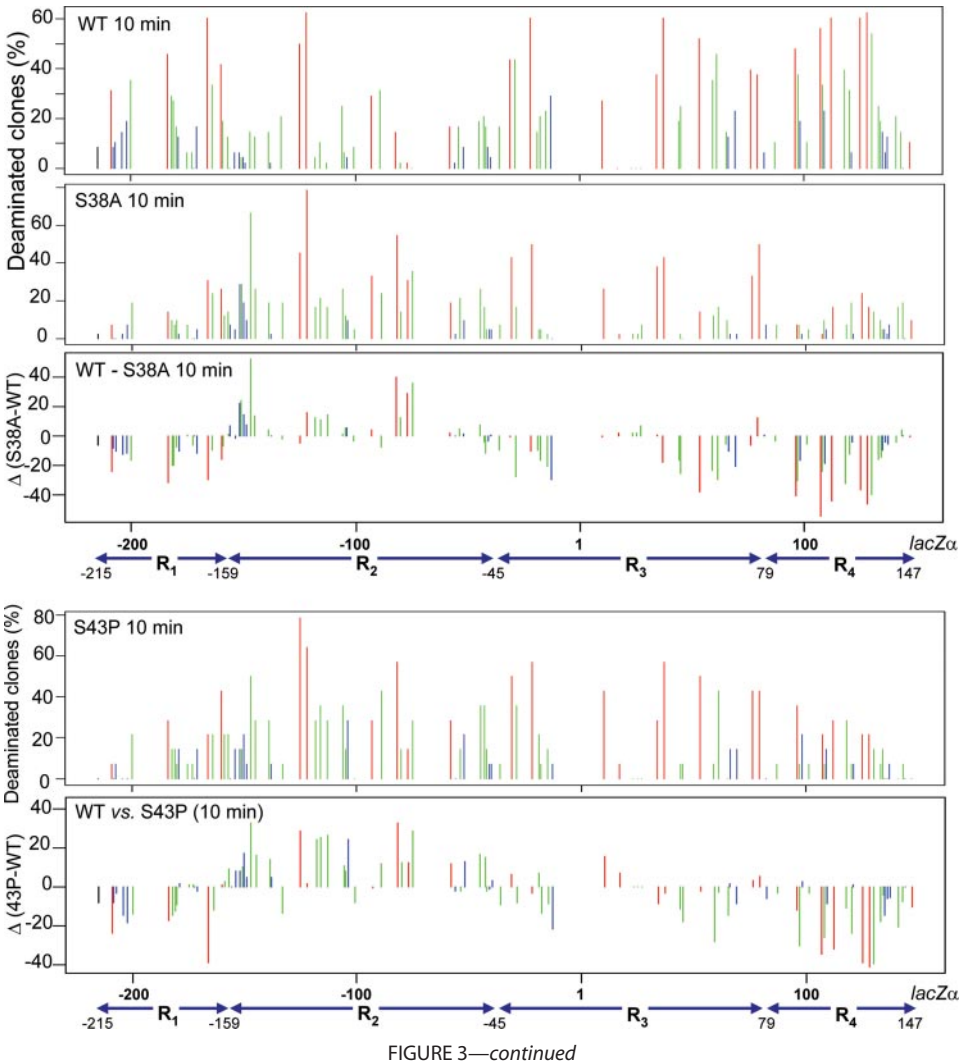

FIGURE 3—*continued*

S38A and S43P (Fig. 3*b*). Differences in WT *versus* S38A and S43P AID spectra also occur in R_3 but are far less pronounced.

A mutability index (MI), which denotes the experimentally measured mutation frequency for a particular sequence motif compared with the expected unbiased mutation frequency (39), has been used as a quantitative measure for AID deamination specificity (10). MI values for the 12 trinucleotide motifs for WT, S38A, and S43P AID have been tabulated (Table 2). On average, WT and mutant AID exhibit the same general deamination preferences, strongly favoring hot spots (WRC) over cold spots (SYC), with intermediate motifs in between (Table 2, MI average).

The large disparity in the WT and mutant AID spectra may be caused by differences in the distribution of individual intermediate triplet motifs in the *lacZα* regions and their mutability by WT and mutant AID. For example, in the R_2 region, which strongly favors S38A and S43P over WT AID deaminations, the intermediate motifs CGC and GGC are each represented five and four times, respectively. These two intermediate motifs have higher mutability indexes for S38A and S43P compared with WT AID by factors of ~2.7 ($p = 0.004$) and 3.4 ($p = 0.0001$), respectively (Table 2). In contrast, ACC, which is the intermediate motif most strongly favored by WT over mutant AID by a factor of 3.7 ($p = 0.02$) (Table 2), is absent from R_2.

TABLE 1
Time course analysis of deamination by wild-type and mutant AID on ssDNA gap substrate

	Average number of mutations per clone		Mutant frequency ($\times 10^{-2}$)[a]		
	2.5 min	10 min	2.5 min	5 min	10 min
***Sf*9-expressed**					
WT AID	13	20	3.8	5.1	5.7
S38A	9	16	3.7	4.7	6.1
S41A	11	15	2.7	3.4	4.3
S43A	16	24	4.6	4.7	5.5
S43P	10	20	1.7	2.2	2.5
E. coli-expressed WT AID	ND[b]	17	0.9	1.4	1.5

[a] The background mutant frequency for gapped substrate in the absence of AID is $\sim 0.07 \times 10^{-2}$. *E. coli*-expressed AID was present at 10–20-fold higher levels (1 μg) compared with *Sf*9-expressed AIDs (50–100 ng).
[b] Not determined.

The spectral differences for WT and mutant AID persist for the longer (10-min) incubation period (Fig. 3*b*). The frequency of deaminated clones at each site increases in an approximately spatially uniform manner in all regions from 2.5 to 10 min for WT and mutant AID. The difference spectra comparing WT with mutant AID for 10-min incubations (Fig. 3*b*) show that WT AID is far more active in R_1, R_3, and R_4, and S38A, S41A, S43A, and S43P AID mutants are more active in R_2, in accord with the 2.5 min incubation data (Fig. 3*b* and supplemental Fig. 1).

A compilation of individual trinucleotide motifs present in Ig switch regions reveals that two of the most highly represented motifs are much more strongly favored for deamination by the phosphorylation-null mutants compared with WT AID. The GGC motif (MI 1.0 S43P, MI 0.23 WT; Table 2) represents 33% of IgA, 19% of IgE, and 15% of IgG4 motifs (Table 3). The AGC motif (MI 2.0 S43P, MI 1.3 WT; Table 2) represents 40% of IgA, 14% of IgE, and 11% of IgG4 (Table 3). Notably, S43P AID is

template Cs with SYC cold spots (Fig. 3*b*, *blue bars*) and intermediate motifs (Fig. 3*b*, *green bars*) are deaminated at significant frequencies (10–30%). Deaminations at these sites are virtually absent in the WT AID spectrum (Fig. 3*b*). A major difference between WT and S38A and S43P AID occurs in R_4, where deaminations in WRC motifs (Fig. 3*b*, *red bars*) are present in a much higher fraction in the WT spectra compared with

TABLE 2

Three-nucleotide motif MI for WT and mutant AID (S38A and S43P) at 2.5 min

The MI is defined as the number of times a given trinucleotide motif within a segment of DNA contains a mutation, divided by the number of times the oligonucleotide would be expected to be mutated for a mechanism with no sequence bias.

Motif	MI			
	WT	S38A	S43P	*p* value[a]
Hot spots				
AAC	2.5	1.5	1.6	0.41
AGC	1.3	2.1	2.0	0.06
TAC	3.4	2.0	2.2	0.22
TGC	2.6	2.2	2.6	0.99
MI average (WRC)	2.5	1.9	2.1	
Cold spots				
CCC	0.28	0.52	0.44	0.65
CTC	0.45	0.76	0.53	0.99
GCC	0.05	0.18	0.05	0.07
GTC	0.32	0.67	0.00	0.99
MI average (SYC)	0.28	0.53	0.26	
Intermediates				
ACC	1.2	0.32	0.35	0.02
ATC	1.0	0.63	0.71	0.97
CAC	1.4	0.91	0.96	0.31
CGC	0.5	1.3	1.4	0.004
GAC	0.24	0.5	0.35	0.79
GGC	0.23	0.78	1.0	0.0001
TCC	0.48	1.1	1.1	0.75
TTC	0.53	0.71	0.71	0.46
MI average	0.70	0.78	0.82	

[a] *p* values were calculated by the paired Wilcox test (see "Experimental Procedures").

TABLE 3

Distribution of trinucleotide motifs in Ig switch regions

The number of trinucleotide motifs was calculated for the nontranscribed strand of IgA (mouse germ line IgA switch region, accession number gi:194415), IgE (human IgE switch region, accession number gi:32983), and IgG (human Ig γ-4 switch region, accession number gi:31510).

	Number of motifs		
	IgA	IgE	IgG
Hot spots			
AAC	6	37	24
AGC	79	90	148
TAC	4	10	7
TGC	7	30	83
Cold spots			
CCC	0	49	122
CTC	1	31	106
GCC	2	54	135
GTC	5	17	44
Intermediates			
ACC	4	42	74
ATC	2	20	21
CAC	2	34	97
CGC	3	11	30
GAC	15	62	79
GGC	65	126	196
TCC	2	30	108
TTC	0	22	57

present in a patient with hyper-IgM syndrome characterized by the absence of CSR (32).

Wild Type and Mutant AID Act Processively on ssDNA—The average numbers of C → T mutations were between 9 and 16 for a 2.5-min incubation of the M13 gapped substrate with either WT AID or each of the four AID mutants and were between 15 and 24 for a 10-min incubation (Table 1). At the 10-min time point, roughly two-thirds of the clones acted on by WT AID had more than 10 mutations, with some clones containing 60–70 mutations (supplemental Fig. 2). A similar fraction of clones containing >10 mutations is observed for S38A AID (supplemental Fig. 2), although compared with WT AID, this mutant produces many more clones, about 50%, with 10–20 mutations and a smaller number of clones having >20 mutations. Since more than 90% of the DNA exposed to AID has no mutations, as determined by sequencing DNA isolated from individual blue plaques (see Ref. 10), both WT and mutant AID act processively.

An ssDNA oligonucleotide, containing a fluorescein tag (*F*) located between the two identical AGC hot spot motifs, has been used to detect multiple deaminations occurring on the *same* ssDNA molecule (36, 37) (*i.e.* a single conversion of C → U in either the 3′-motif or 5′-motif or double C → U conversions occurring in both 3′- and 5′-motifs) (Fig. 4*a*). AID acts bidirectionally, resulting in roughly equal deamination frequencies at the 3′ and 5′ targets on the same DNA (Fig. 4*a*).

The "processivity factor" is calculated as the ratio of the observed fraction of double deaminations occurring at both 5′-C and 3′-C on the same ssDNA substrate to the predicted fraction of independent double deaminations (36). Processivity factors between 3.5 and 9 for WT AID and AID mutants (Fig. 4, *a–c*) confirm that processive double deamination is well in excess of double deamination attributable to random encounters with the same ssDNA substrate by two AID molecules. The reduction in processivity factors with time reflects an increased probability of multiple enzyme-DNA encounters when a larger percentage of substrates are deaminated (Fig. 4, *b* and c, 5- and 10-min incubations). The average correlated double deamination efficiency for WT and mutant AID is ~30% (Fig. 4*d*) (*i.e.* 70% of deaminations catalyzed by a single AID molecule occur just once, at either the 3′ or 5′ target motif, and 30% occur processively at both 3′ and 5′ target motifs). The data demonstrate that the WT and mutant forms of AID are processive to about the same extent. Significantly, the processive action of AID seems unaffected by amino acid substitutions that eliminate the possibility of phosphorylation at sites 38, 41, and 43.

AID-catalyzed Transcription-dependent C Deamination Occurs in the Absence of Phosphorylation and RPA—We have examined how phosphorylation of AID affects its ability to deaminate C during active transcription of dsDNA (Fig. 5). T7 RNA polymerase is used to transcribe linear dsDNA having a single AGC hot spot motif on the nontranscribed strand. The deamination of C to U on the nontranscribed strand is detected by measuring the termination of primer elongation caused by incorporation of ddAMP opposite U and ddGMP opposite C (see "Experimental Procedures"). Two primer elongation reactions are performed simultaneously, where ddATP is substituted for dATP (Fig. 5*a*, *ddA lanes*) and ddGTP is substituted for dGTP (Fig. 5*a*, *ddG lanes*). The conversion of C → U is detected by the appearance of a lower doublet termination band corresponding to the incorporation of ddAMP, just *below* the C template position, in the *ddA lanes* (Fig. 5*a*, *arrow*, *ddA lanes 5*, *7*, *9*, *11*, *13*, *15*) and by the presence of bands migrating past the C template site in the *ddG lanes* (Fig. 5*a*, *bracket* to the *right* of the gels, *ddG lanes 6*, *8*, *10*, *12*, *14*, and *16*). The upper doublet band in the *ddA lanes* occurs by incorporating dGMP opposite residual C moieties not acted on by AID (Fig. 5*a*). AID-catalyzed deamination occurs only when the dsDNA undergoes

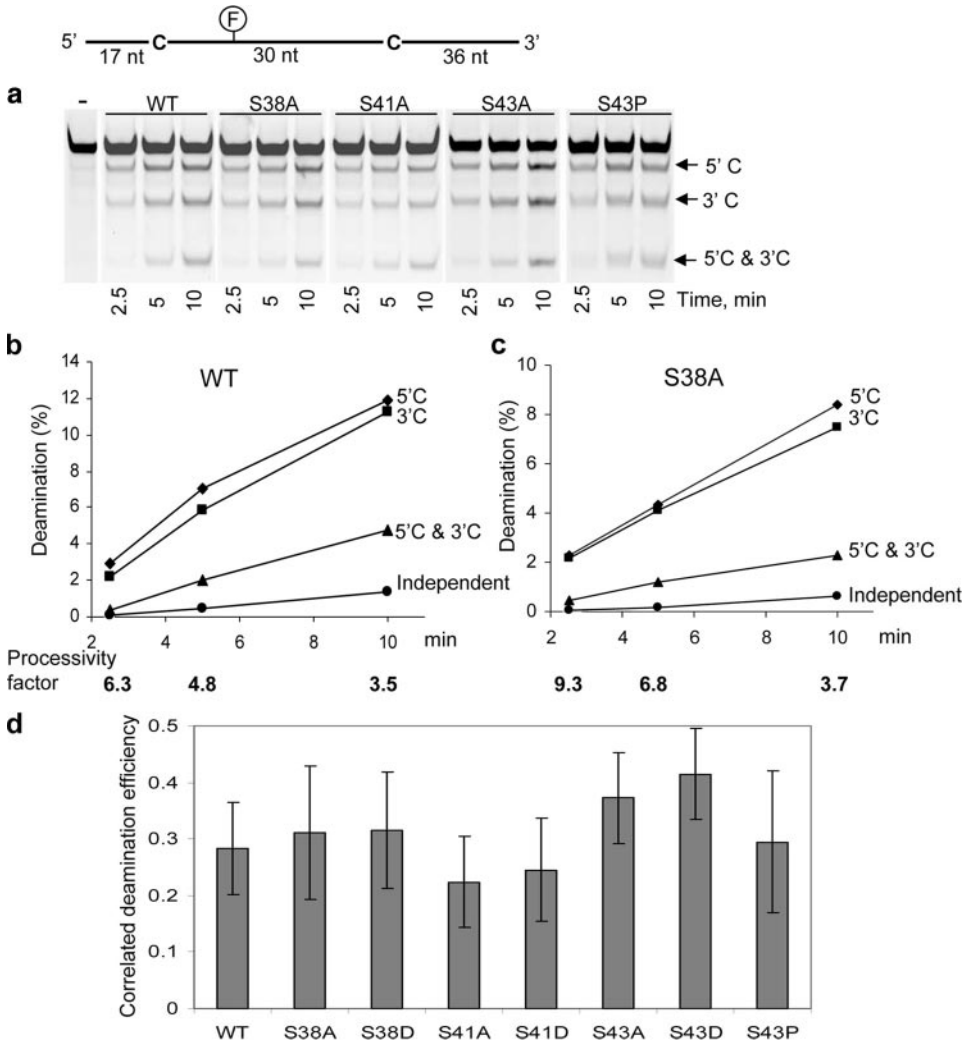

FIGURE 4. **Analysis of processive deamination by WT and mutant AID on the same oligonucleotide ssDNA molecule.** *a*, deamination by WT AID and S38A, S41A, S43A, and S43P phosphorylation-null AID mutants on an ssDNA substrate with two C targets. An ssDNA substrate (85 nt) with two identical C deamination hot spot motifs (AGC), containing a fluorescein (*F*) reporter molecule located between the two motifs (*sketch* at the *top*), is used to measure single and double deaminations occurring on an individual ssDNA substrate. Single deaminations of the 5′-C and 3′-C are detected as fluorescein-labeled 67- and 48-nt fragments, *5′C* and *3′C*, respectively; double deaminations of both Cs on the same molecule are detected as a 30-nt fragment (*5′C & 3′C*). The integrated gel band intensities are used to determine the fraction of correlated double deaminations, generated by the processive action of a single AID molecule (see "Experimental Procedures"). *b* and *c*, deamination (percentage) of substrates containing single deaminations at either 5′-C (♦) or 3′-C (■) and double deaminations at both 5′-C and 3′-C (▲) obtained from the integrated gel band intensities in *a* for WT AID and for S38A mutant AID. ●, the calculated fraction of independent double deaminations (*Independent*), caused by two molecules of AID attacking each site independently. The processivity factor, defined as the ratio of the observed double deaminations occurring at both 5′-C and 3′-C on an individual ssDNA to the calculated fraction of independent double deaminations (36), is shown at the *bottom* of each *graph*. The processivity factor for both WT and mutant AID is significantly greater than 3, indicating that the majority of double deaminations are caused by a single AID molecule deaminating both target C residues. *d*, calculated average correlated deamination efficiency for WT AID and AID mutants. The "correlated deamination" efficiency is the probability for a single enzyme to perform double correlated deaminations at both 3′-C and 5′-C upon a single encounter with an ssDNA substrate. *Error bars*, ±1 S.E.

active transcription in the presence of T7 RNA polymerase (Fig. 5*a*, lanes 5 - 16), whereas no deamination is detected in the absence of either AID or T7 RNA polymerase (Fig. 5*a*, *lanes 1–4*).

The activity of human WT AID expressed in *Sf*9 cells is either similar or slightly less when compared with Ser → Ala phosphorylation-null mutants or with Ser → Asp (−1 negatively charged) mutants (Fig. 5*b*). The activity of the HIGM-2 S43P mutant is indistinguishable from WT AID (Fig. 5*b*). WT AID expressed in *E. coli* requires about a 7-fold higher protein concentration to attain a similar deamination efficiency on transcribed dsDNA as *Sf*9-expressed human AID (Fig. 5, *a* (*lanes 15* and *16*) and *b*). Since the *E. coli*-expressed AID is ~100-fold less active than *Sf*9-expressed AID on ssDNA (Fig. 2), a smaller difference in the relative activities of *Sf*9 and *E. coli* AID on transcribed dsDNA might reflect the possibility that the measured deamination activity in this assay may be dependent on a rate-limiting formation of the transcription bubble needed to provide ssDNA (*i.e.* the nontranscribed strand) as a substrate for AID.

AID-catalyzed transcription-dependent deamination on linear dsDNA occurs in the absence of RPA (Fig. 5). Deamination is inhibited when RPA is present at substoichiometric and higher levels (3–12 pmol) compared with AID (7 pmol) (Fig. 6*a*). We observed ~35–70% inhibition of transcription-dependent deamination, depending on the level of RPA in the reaction, which is shown by a reduction in the intensity of ddA termination bands (Fig. 6*a*, *U arrows*) and by a concomitant reduction in the intensities of bands migrating beyond the target C site in the presence of ddGTP (Fig. 6*a*, *ddG lanes 2*, *4*, *6*, and *8*). RPA inhibits the activities of S38A and S43P AID mutants to about the same extent as WT AID (Fig. 6*b*). A similar inhibitory effect of RPA on WT AID deamination is observed using supercoiled circular DNA, containing a *lacZ* target gene containing numerous hot spot, cold spot, and intermediate target motifs (Table 4).

We investigated the effect of human RPA and *E. coli* SSB on AID activity on a linear DNA oligonucleotide (36-mer, Fig. 6*c*) and gapped (M13 DNA, Fig. 6*d*) substrates. Both single-stranded binding proteins exert a significant inhibitory effect on AID activity. Increasing levels of either RPA or SSB strongly inhibit AID activity on the oligomer substrate (Fig. 6*c*). There is a monotonic reduction in C → T mutant frequencies on the gapped substrate, decreasing to background mutation levels (mutant frequency ~0.07–0.1%) at about 40 pmol of RPA or SSB, which corresponds to a saturating concentration of sin-

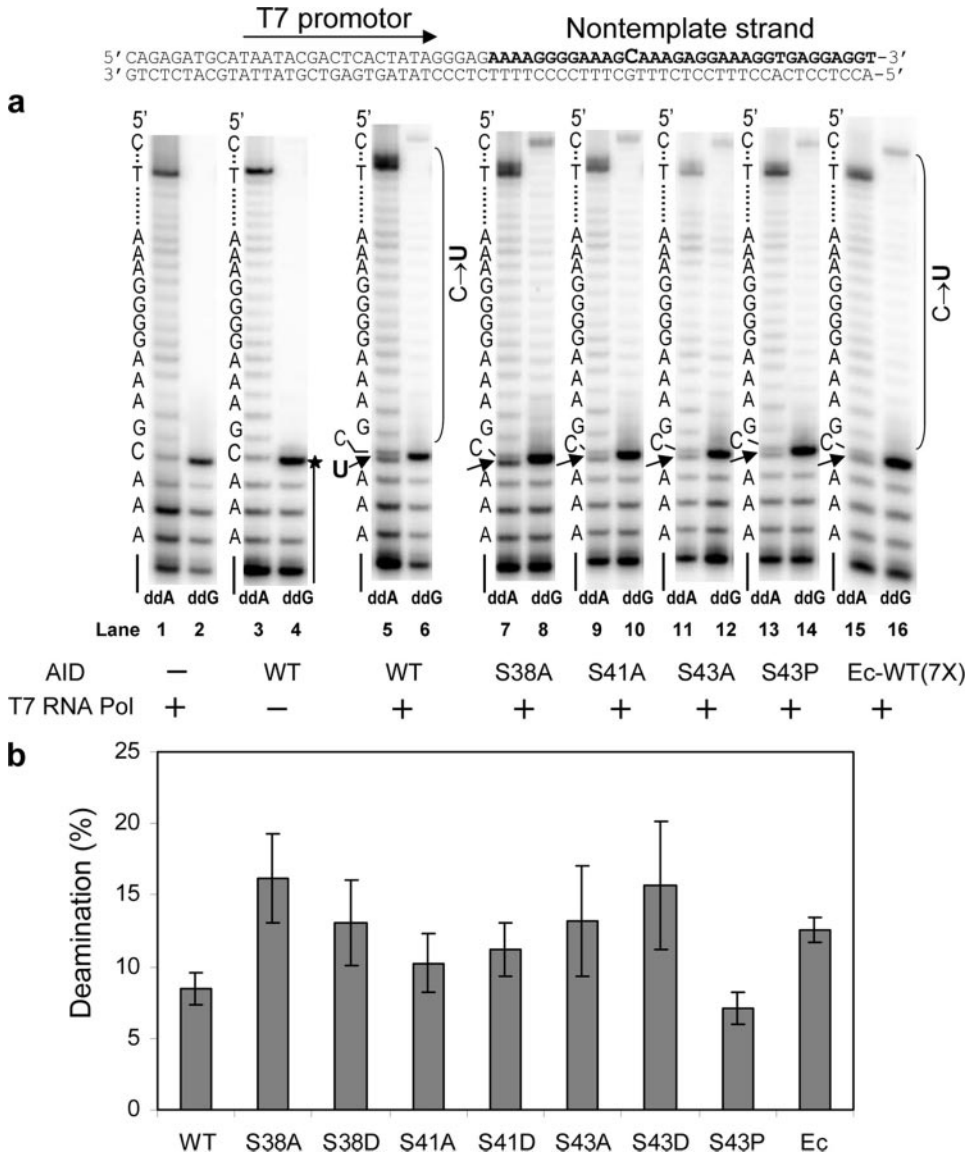

FIGURE 5. **Deamination activity of WT and mutant AID proteins on a transcribed linear dsDNA.** *Top*, dsDNA template with an incorporated T7 promoter for transcription by T7 RNA polymerase. A single target C residue located on the nontranscribed strands is indicated in *boldface type* by a *larger font size*. *a*, polyacrylamide gel electrophoresis data showing C → U deamination on the nontranscribed DNA strand for WT and mutant AID proteins expressed in *Sf*9 cells (*lanes 5–14*) and WT protein, expressed in *E. coli* (*Ec-WT*, *lanes 15* and *16*). A primer elongation dideoxynucleotide termination assay (10) (see "Experimental Procedures") was used to measure the conversion of C → U. Deamination is measured by the appearance of an intense *termination band* just *below* the *template C band* (indicated with an *arrow* on the *left* of each gel) when ddATP is used in place of dATP in the extension reactions (*ddA*, *lanes 5*, *7*, *9*, *11*, *13*, and *15*). Concomitant with the conversion of C to U band in the *ddA lanes* is the presence of bands migrating past the C template site, denoted by the *bracket* to the *right* of the gel, when ddGTP is used instead of dGTP (*ddG lanes 6*, *8*, *10*, *12*, and *14*). Deamination on the dsDNA substrate is not detected in the absence of either AID or T7 RNA polymerase (*lanes 1–4*). The presence of intact C template is observed by the absence of the lower band just below the C template position in the ddA lanes (*lanes 1* and *3*) and by the absence of bands migrating beyond the C template site in the *ddG lanes* (*lanes 2* and *4*). *b*, deamination efficiency of WT and mutant AID on transcribed dsDNA. The deamination reactions were carried out in the presence of 7-fold more (50 pmol) *E. coli*-expressed WT AID (*Ec-WT*) compared with *Sf*9-expressed AID (7 pmol).

gle-stranded binding protein (one RPA trimer or one SSB tetramer to 30-nt ssDNA; see Refs. 40 and 41).

PKA-phosphorylated AID Activity on Single-stranded DNA and on Transcriptionally Activated Double-stranded DNA—Since phosphorylation at Ser[38] has been implicated in efficient CSR *in vivo* (22–24, 27) and in transcription-dependent deamination *in vitro* (22, 27), we have incubated AID with the catalytic subunit of human PKA in the presence of γ-[32]P-labeled dATP, to determine if more extensive phosphorylation at Ser[38], which has a PKA motif, might affect the biochemical properties of AID. We observed a strong phosphorylation signal corresponding to the electrophoretic migration of AID that is dependent on incubation with PKA (Fig. 7*a*, *right-hand gel*), and we estimate that about 50% of the AID is phosphorylated based on the intensities of Coomassie-stained bands corresponding to phosphorylated and nonphosphorylated AID (Fig. 7*a*, *left-hand gel*). Excess phosphorylation has no significant effect on the specific activity of AID acting on ssDNA (Fig. 7*b*). The same conclusion holds for transcription-dependent deamination of dsDNA in the presence or absence of RPA (Fig. 7*c*).

DISCUSSION

AID resides principally in the cytoplasm of activated B cells (42, 43) but is transported into the nucleus, where it catalyzes deamination of C → U on Ig V-gene regions undergoing active transcription. Approximately 10–15% of AID expressed in B cells is reported to undergo phosphorylation at Ser[38] (23). We have found using *Sf*9 insect cells to overexpress human GST-AID that a small fraction of the enzyme is phosphorylated at Ser[38] and Thr[27] and at two residues not reported previously, Ser[41] and Ser[43] (Fig. 1). The data suggest that phosphorylation occurs on just one of these four positions in a single AID molecule. There is considerable interest in understanding the biological roles of phosphorylation of AID, especially with respect to the ability to target transcriptionally active V-regions and switch regions selectively (22, 27). The replacement of Ser[38] with Ala delays the onset of CSR in B-cells (23, 44) and has been reported to abrogate transcription-dependent deamination on linear dsDNA using T7 RNA polymerase *in vitro* (22, 27). The replacement of Ser[43] with Pro engenders hyper-IgM syndrome in humans (32).

In this paper, we aimed to distinguish between how phosphorylation-null mutants *versus* phosphorylation *per se* affects the

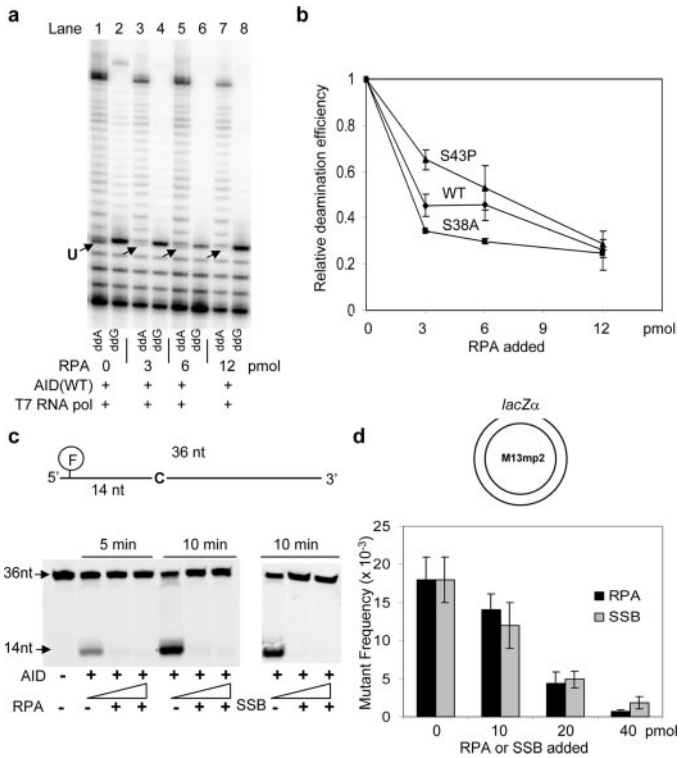

FIGURE 6. **Effect of single-stranded binding proteins (human RPA and *E. coli* SSB) on AID deamination activity.** *a*, deamination of WT AID on the transcribed linear dsDNA in the presence of increasing amounts of RPA. Experiments were carried out as described in the legend to Fig. 5*a*. *b*, relative deamination efficiencies on transcribed dsDNA of WT and mutant S38A and S43P AID in the presence of RPA. Deamination activities in the presence of RPA were normalized to the activity of the enzymes in the absence of RPA. *c*, RPA and SSB completely inhibit activity of WT AID on a fluorescein (*F*)-labeled 36-nt ssDNA substrate with a single hot-spot C (AGC) target. Reactions were carried out in the presence of 7.5 pmol of substrate DNA (255 pmol as nucleotides) and 3 pmol of WT AID for either 5 or 10 min. RPA or SSB, when present, were added at saturating concentration of 10 or 20 pmol. *d*, WT AID-catalyzed deamination on the gapped ssDNA substrate in the presence of RPA and SSB. Deamination activity is monitored by the increase in frequency of mutant (white or light blue color) M13 phage caused by C to U conversion in the *lacZα* target reporter gene. The background mutant frequency for the gapped DNA in the absence of AID is typically 0.7×10^{-3} (10). At 40 pmol, RPA and SSB saturate ssDNA gapped substrates (one RPA trimer or one SSB tetramer to ~30 nt of ssDNA). *Error bars*, S.D. values.

TABLE 4

Effect of RPA on AID activity on transcribed closed circular dsDNA substrate

AID deamination activity was measured using a covalently closed circular dsDNA M13mp2T7 substrate, undergoing active transcription by T7 RNA polymerase (see Ref. 11).

	Mutation frequency ($\times 10^{-3}$)
No AID	0.8 ± 0.3
WT AID	3.7 ± 0.5
WT AID + RPA (60 ng)	4.7 ± 0.9
WT AID + RPA (300 ng)	2.5 ± 0.5
WT AID + RPA (1 μg)	2.2 ± 0.4

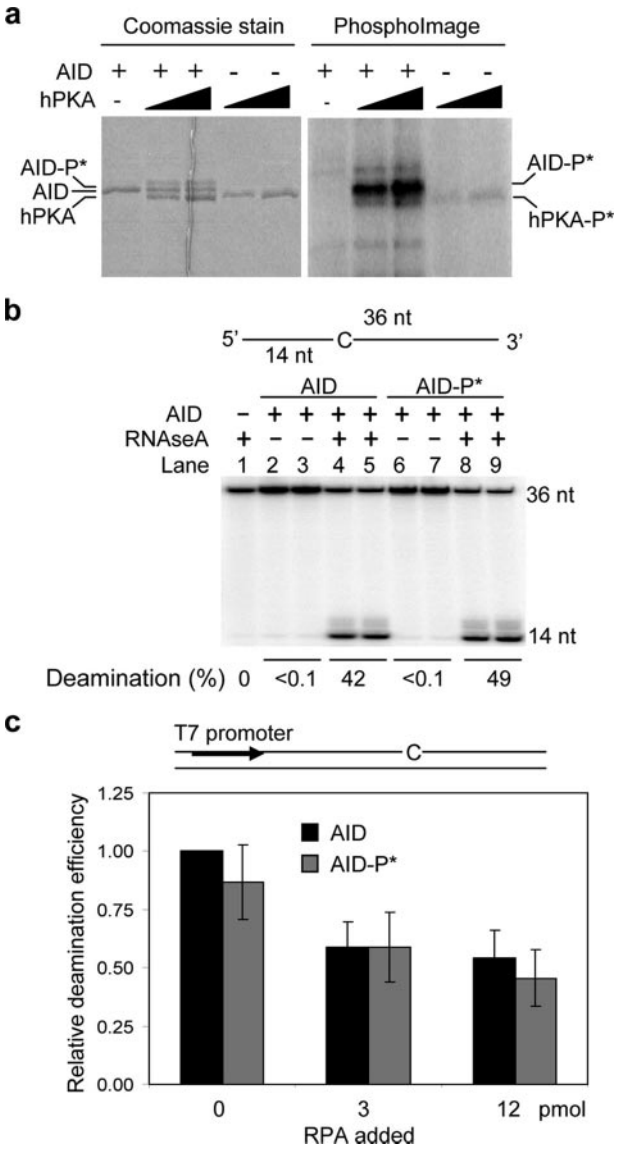

FIGURE 7. **Effect of *in vitro* phosphorylation on AID.** *a*, efficient phosphorylation of human AID by recombinant catalytic subunit of human PKA (*hPKA*). Human PKA and AID were incubated at a 1:2 or 1:1 ratio in the presence of 100 μM ATP and [^{32}P]ATP. About 50% of AID is phosphorylated (*AID-P**) by human PKA (at a 1:1 ratio) as shown by the presence of a slower migrating band on a Coomassie-stained SDS-polyacrylamide gel. The catalytic subunit of hPKA alone is autophosphorylated in the presence of ATP (45). *b*, phosphorylation of AID does not alter the requirement for activation by RNase (*AID-P**; *lanes 6–9*) or change AID activity on ssDNA. *c*, phosphorylation of AID has no stimulating effect on the activity of AID on the transcribed dsDNA even in the presence of RPA. RPA exhibits similar inhibitory effect on both low phosphorylated AID and highly phosphorylated AID (*AID-P**) in a T7-driven transcription deamination assay. The deamination efficiencies were normalized to the activity of low phosphorylated AID in the absence of RPA.

biochemical behavior of AID. The biochemical properties investigated include enzyme specific activity, processivity, deamination spectra, deamination motif specificity, and transcription-dependent deamination, including the influence of RPA. The behavior of insect cell-expressed human AID acting on ssDNA and transcribed dsDNA *in vitro* mimic characteristics of SHM in B-cells, including an approximate simulation of hot spot and cold spot C → T mutational patterns (10). Consequently, a comparative analysis of WT and mutant AID could provide biologically relevant insights into the role of phosphorylation in initiating SHM and CSR.

A comparison of WT AID with phosphorylation-null AID mutants shows major differences in deamination spectra (Fig. 3) and trinucleotide motif specificity (Table 2). However, there appear to be no significant differences in any of the other biochemical properties of AID. The specific activities of WT and mutant *Sf*9-expressed human AID are similar to within a factor of 3 (Fig. 2). Given the inherent variability in enzyme prepara-

tions, these differences are probably not important. Much more important, however, is that phosphorylation of AID at any of the three serine residues is not essential for AID activity. With relevance to human immunodeficiency disease, the S43P mutant associated with HIGM-2 syndrome shows essentially normal activity on ssDNA (Fig. 2) and transcribed dsDNA (Figs. 5 and 6) but generates a significantly different deamination-dependent C → T mutation spectrum compared with WT AID (Fig. 3*b*).

A hallmark property of WT AID is its ability to deaminate ssDNA *in vitro* with high processivity, using a combination of random sliding and jumping modes of action (36, 37). The phosphorylation-null AID mutants are as processive as WT AID; the mutants and WT AID generate about the same average numbers of mutations/clone (Table 1), exhibit similar processivity factors (*e.g.* see Fig. 4*b*, comparing WT AID and S38A mutant AID), and catalyze about the same fraction of correlated double deaminations (Fig. 4*d*).

When T7 RNA polymerase was used to transcribe linear dsDNA (Figs. 5 and 6) or supercoiled circular dsDNA (Table 4), there were no discernible differences in the deamination activity of WT AID compared with phosphorylation-null AID mutants, either in the presence or absence of RPA (Fig. 6). These data are in contrast to a previous report showing that mouse AID obtained from B-cells requires phosphorylation at Ser^{38} and the presence of RPA to deaminate linear dsDNA during transcription by T7 RNA polymerase (22, 27). Although we cannot explain the apparent discrepancy between the two results, our data with phosphorylation-null mutants suggest that neither phosphorylation *per se* nor RPA is needed for transcription-dependent deamination on either linear or circular dsDNA, using GST-tagged human AID expressed either in insect cells (Figs. 5 and 6 and Table 4) or in *E. coli* (Fig. 5) (see also Ref. 9). There is also no apparent increase in transcription-dependent deamination for AID that has been about 50% phosphorylated by the catalytic subunit of PKA, irrespective of whether or not RPA is included in the transcription reaction (Fig. 7).

The presence of a single-stranded binding protein (human RPA or *E. coli* SSB) at a saturating concentration (one RPA trimer or one SSB tetramer per about 30 nt of ssDNA) leads to an almost complete inhibition of AID activity on both short oligonucleotide and gapped ssDNA substrates (Fig. 6, *c* and *d*). We suggest that by binding to ssDNA, RPA and SSB shield templates from deamination by AID. The similar inhibitory effects of RPA and SSB suggest that selective binding of RPA to AID is not occurring, nor is phosphorylation of AID necessary (Fig. 6), at least for the case of human AID expressed in insect cells. The absence of detectable deamination after a relatively lengthy (10 min) incubation time suggests that AID is unable to displace RPA or SSB from ssDNA (Fig. 6*c*). Perhaps RPA might act to protect B cells from a potential "global" mutator activity of AID when DNA becomes transiently single-stranded during normal replication and repair processes.

In contrast to the absence of discernible differences in specific activity, processivity, and transcription-dependent deamination, the replacement of Ser^{38} with Ala or Ser^{43} with Pro has a significant effect on deamination motif specificity and C → T mutation spectra that distinguish the phosphorylation-null mutants from WT AID (Fig. 3*b*). The spectral changes are most pronounced in *lacZα* region R_2 (Fig. 3*b*, Δ difference spectra), where the AID mutants are much more active than WT AID, and in regions R_1 and R_4, where WT AID-catalyzed mutations predominate (Fig. 3*b*, Δ difference spectra). The spectra for all of the phosphorylation-null AID mutants at amino acid residues 38, 41, and 43 are similar to each other (Fig. 3*b* and supplemental Fig. 1) and are substantially distinct from WT AID (*p* values of $\sim 10^{-7}$).

WT AID and phosphorylation-null AID mutants share the same general motif preferences, where WRC hot spot motifs are generally favored over intermediate motifs and SYC cold spot motifs (Table 2). However, the average MI for WRC hot spot motifs is higher for WT AID (average MI ~2.5) (Table 2) compared with S38A and S43P AID mutants (average MI ~2.0) (Table 2), thus providing a possible explanation for excess WT AID deaminations in regions R_1 and R_4. An analysis of individual DNA clones used to compile the mutational spectra shown in Fig. 3*b* reveals that mutations occur most frequently in clusters (11). We have attributed mutational clustering to a processive bidirectional sliding of AID along ssDNA (10, 11). Mutations initiated in hot spot motifs lead to additional nearby mutations, thus amplifying differences between WT and mutant AID spectra in R_1 and R_4 (Fig. 3*b*, Δ difference spectra).

Favored deaminations occurring in R_2 for the phosphorylation-null AID mutants (Fig. 3*b*) appear to have a more subtle origin that may reflect differences in individual intermediate motif deamination efficiencies. The phosphorylation-null mutants are much more apt than WT AID to attack two intermediate motifs, CGC and GGC (Table 2). The possible relevance of this observation is that CGC and GGC appear five and four times, respectively, in R_2, each having a mutability index favoring mutant over WT deaminations by factors of 2.7 ($p = 0.004$) and 3.4 ($p = 0.0001$), respectively, whereas ACC, which is the favored intermediate motif for WT relative to mutant AID (Table 2, MI ratio ~3.6, $p = 0.02$), is absent from R_2. A clustering of mutations proximal to the intermediate motifs could amplify differences that would favor mutant AID-catalyzed deaminations in R_2.

Studies of AID mutations (T27A and S38A) have led to a suggestion that phosphorylation of AID *in vivo* plays an important role in CSR and SHM. But what that role is has yet to be established. Since phosphorylated AID is preferentially enriched in the chromatin fraction (23), perhaps phosphorylation is needed for transporting AID from the cytoplasm into the nucleus, or to target AID to actively transcribed V-genes and switch regions. Whatever the role turns out to be, it is essential to determine how phosphorylation modulates the biochemical properties of AID. Complicating matters, however, is that zebrafish AID, which has no equivalent Ser^{38} residue, is fully competent in both CSR and SHM (26, 28, 29). Therefore, it is imperative to distinguish between effects attributable to the absence of phosphorylation at a specific Ser amino acid residue *versus* replacement of the Ser residue by a different amino acid (*e.g.* Ala or Pro).

Our data show that the biochemical characteristics of AID, including specific activity, processivity, and transcription-de-

pendent deamination, are not affected either by the absence of phosphorylation on serine residues 38, 41, and 43 or by excess phosphorylation. The replacement of each of these residues, however, results in a significant change in mutant *versus* WT AID deamination spectra. The AID mutants show a high level of deamination notably at two non-hot spot motifs, GGC and CGC, where WT AID is relatively inactive (Table 2).

Switch region DNA sequences contain remarkably high frequencies of GGC motifs. There are 196 GGCs in the S-region of IgG4 and 126 in the S-region of IgE, which are by far the most prevalent triplet motifs in these two switch regions, comprising ~15% in IgG and ~19% in IgE (Table 3). Since WT AID deamination spectra measured on ssDNA and on transcribed dsDNA *in vitro* agree well with C → T SHM spectra in humans (10, 11), by analogy it seems possible, perhaps even likely, that S43A and S43P mutant AID might introduce high levels of U at GGC switch region motifs. Thus, if aberrant recombination were to result from excess deamination targeted to noncanonical switch region motifs, that might contribute to the delay in S38A-initiated CSR in cultured B cells (23) and perhaps to the elimination of CSR in HIGM-2 patients having a S43P mutant form of AID (32).

In a general sense, we have observed that the motifs that WT AID favors *in vitro* are strikingly similar to hot spot motifs of SHM *in vivo* (10), although the enzyme must contend with chromatin, transcription, and assuredly with AID-associated proteins in B cells (3–5). The ability to simulate biochemically several important biological functions of AID suggests that many characteristics of SHM and CSR are determined by its intrinsic properties. Our data suggest the prospect that altered deamination specificity caused by amino acid replacements at Ser^{38}, Ser^{41}, or Ser^{43} could perhaps account for significant deficiencies in SHM and CSR in B-cells.

Acknowledgments—We thank Linda Chelico for preparation of the 85-nt substrate for processivity measurements and Matthew Scharff for a critical reading of the manuscript.

REFERENCES

1. Rajewsky, K., Forster, I., and Cumano, A. (1987) *Science* **238,** 1088–1094
2. Stavnezer, J. (2000) *Curr. Top. Microbiol. Immunol.* **245,** 127–168
3. Goodman, M. F., Scharff, M. D., and Romesberg, F. E. (2007) *Adv. Immunol.* **94,** 127–155
4. Di Noia, J. M., and Neuberger, M. S. (2007) *Annu. Rev. Biochem.* **76,** 1–22
5. Muramatsu, M., Nagaoka, H., Shinkura, R., Begum, N. A., and Honjo, T. (2007) *Adv. Immunol.* **94,** 1–36
6. Bransteitter, R., Pham, P., Scharff, M. D., and Goodman, M. F. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4102–4107
7. Chaudhuri, J., Tian, M., Khoung, C., Chua, K., Pinaud, E., and Alt, F. W. (2003) *Nature* **421,** 726–730
8. Dickerson, S. K., Market, E., Besmer, E., and Papavasiliou, F. N. (2003) *J. Exp. Med.* **197,** 1291–1296
9. Sohail, A., Klapacz, J., Samaranayake, M., Ullah, A., and Bhagwhat, A. S. (2003) *Nucleic Acids Res.* **31,** 2990–2994
10. Pham, P., Bransteitter, R., Petruska, J., and Goodman, M. F. (2003) *Nature* **423,** 103–107
11. Bransteitter, R., Pham, P., Calabrese, P., and Goodman, M. F. (2004) *J. Biol. Chem.* **279,** 51612–51621
12. Rogozin, I. B., and Kolchanov, N. A. (1992) *Biochim. Biophys. Acta* **1171,** 11–18
13. Smith, D. S., Creadon, G., Jena, P. K., Portanova, J. P., Kotzin, B. L., and Wysocki, L. J. (1996) *J. Immunol.* **156,** 2642–2652
14. Dorner, T., Foster, S. J., Brezinschek, H.-P., and Lipsky, P. E. (1998) *Immunol. Rev.* **162,** 161–171
15. Peters, A., and Storb, U. (1996) *Immunity* **4,** 57–65
16. Fukita, Y., Jacobs, H., and Rajewsky, K. (1998) *Immunity* **9,** 105–114
17. Maizels, N. (1995) *Cell* **83,** 9–12
18. Yoshikawa, K., Okazaki, I. M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H., and Honjo, T. (2002) *Science* **296,** 2033–2036
19. Ramiro, A. R., Stavropoulos, P., Jankovic, M., and Nussenzweig, M. C. (2003) *Nat. Immunol.* **4,** 452–456
20. Shen, H. M., and Storb, U. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12997–13002
21. Nambu, Y., Sugai, M., Gonda, H., Lee, C., Katakai, T., Agata, Y., Yokota, Y., and Shimizu, A. (2003) *Science* **302,** 2137–2140
22. Basu, U., Chaudhuri, J., Alpert, C., Dutt, S., Ranganath, S., Li, G., Schrum, J. P., Manis, J. P., and Alt, F. W. (2005) *Nature* **438,** 508–511
23. McBride, K. M., Gazumyan, A., Woo, E. M., Barreto, V. M., Robbiani, D. F., Chait, B. T., and Nussenzweig, M. C. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8798–8803
24. Pasqualucci, L., Kitaura, Y., Gu, H., and Dalla-Favera, R. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 395–400
25. Shinkura, R., Okazaki, I. M., Muto, T., Begum, N. A., and Honjo, T. (2007) *Adv. Exp. Med. Biol.* **596,** 71–81
26. Chatterji, M., Unniraman, S., McBride, K. M., and Schatz, D. G. (2007) *J. Immunol.* **179,** 5274–5280
27. Chaudhuri, J., Khuong, C., and Alt, F. W. (2004) *Nature* **430,** 992–998
28. Barreto, V. M., Pan-Hammarstrom, Q., Zhao, Y., Hammarstrom, L., Misulovin, Z., and Nussenzweig, M. C. (2005) *J. Exp. Med.* **202,** 733–738
29. Wakae, K., Magor, B. G., Saunders, H., Nagaoka, H., Kawamura, A., Kinoshita, K., Honjo, T., and Muramatsu, M. (2006) *Int. Immunol.* **18,** 41–47
30. Shen, H. M., Ratnam, S., and Storb, U. (2005) *Mol. Cell Biol.* **25,** 10815–10821
31. Besmer, E., Market, E., and Papavasiliou, F. N. (2006) *Mol. Cell Biol.* **26,** 4378–4385
32. Zhu, Y., Nonoyama, S., Morio, T., Muramatsu, M., Honjo, T., and Mizutani, S. (2003) *J. Med. Dent. Sci.* **50,** 41–46
33. Lohman, T. M., Green, J. M., and Beyer, R. S. (1986) *Biochemistry* **25,** 21–25
34. Henricksen, L. A., Umbricht, C. B., and Wold, M. S. (1994) *J. Biol. Chem.* **269,** 11121–11132
35. Smolka, M. B., Albuquerque, C. P., Chen, S. H., Schmidt, K. H., Wei, X. X., Kolodner, R. D., and Zhou, H. (2005) *Mol. Cell Proteomics* **4,** 1358–1369
36. Chelico, L., Pham, P., Calabrese, P., and Goodman, M. F. (2006) *Nat. Struct. Mol. Biol.* **13,** 392–399
37. Pham, P., Chelico, L., and Goodman, M. F. (2007) *DNA Repair* (*Amst.*) **6,** 689–692
38. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) *Nucleic Acids Res.* **31,** 3635–3641
39. Shapiro, G. S., Aviszus, K., Murphy, J., and Wysocki, L. J. (2002) *J. Immunol.* **168,** 2302–2306
40. Lohman, T. M., and Ferrari, M. E. (1994) *Annu. Rev. Biochem.* **63,** 527–570
41. Wold, M. S. (1997) *Annu. Rev. Biochem.* **66,** 61–92
42. Rada, C., Jarvis, J. M., and Milstein, C. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7003–7008
43. Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., and Honjo, T. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 1975–1980
44. Basu, U., Chaudhuri, J., Phan, R. T., Datta, A., and Alt, F. W. (2007) *Adv. Exp. Med. Biol.* **596,** 129–137
45. Skalhegg, B. S., and Tasken, K. (2000) *Front Biosci.* **5,** 678–693

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 20, pp. 13780–13791, May 16, 2008

A Model for Oligomeric Regulation of APOBEC3G Cytosine Deaminase-dependent Restriction of HIV*[S]

Received for publication, February 7, 2008, and in revised form, March 21, 2008 Published, JBC Papers in Press, March 24, 2008, DOI 10.1074/jbc.M801004200

Linda Chelico[‡], Elizabeth J. Sacho[§], Dorothy A. Erie[§¶], and Myron F. Goodman[‡1]

From the [‡]Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, California 90089-2910, the [§]Department of Chemistry and [¶]Applied and Materials Sciences Curriculum, University of North Carolina, Chapel Hill, North Carolina 27599

APOBEC3G (A3G) restricts HIV-1 infection by catalyzing processive C → U deaminations on single-stranded DNA (ssDNA) with marked 3′ → 5′ deamination polarity. Here we show that A3G exists in oligomeric states whose composition is dictated primarily by interactions with DNA, with salt playing an important, yet secondary, role. Directional deaminations correlate with the presence of dimers, tetramers, and larger oligomers observed by atomic force microscopy, and random deaminations appear to correlate mainly with monomers. The presence of a 30-nt weakly deaminated "dead" zone located at the 3′-ssDNA end implies the presence of a preferred asymmetric direction for A3G catalysis. Single turnover reaction rates reveal a salt-dependent inhibition of C deamination toward the 3′-ssDNA region, offering a molecular basis underlying A3G deamination polarity. Presteady state analysis demonstrates rapid diffusion-limited A3G-ssDNA binding, a slower salt-dependent conformational change, possibly indicative of DNA wrapping, and long (5–15 min) protein-DNA complex lifetimes. We suggest that diverse A3G oligomerization modes contribute to the human immunodeficiency virus, type 1, proviral DNA mutational bias.

In 2002, Sheehy *et al.* (1) determined that APOBEC3G (A3G), originally called CEM15, a proposed cytidine deaminase based on sequence analysis, is the nonpermissive host factor that blocks virion infectivity factor-defective (Δ*vif*) HIV-1 infection of T cells. The experimental determination that CEM15 was a cytidine deaminase belonging to the APOBEC family followed soon after with experiments demonstrating G → A-induced hypermutation of proviral DNA in a Δ*vif* HIV-1 virion (2, 3). A3G has a duplicated deaminase domain structure, but only the C-terminal domain is responsible for the single-stranded DNA (ssDNA)[2] deamination activity (4, 5).

Apart from A3G-catalyzed deamination, A3G may also have the capacity for blocking reverse transcription, (+)-DNA synthesis, and provirus formation either by interacting with RNA or DNA of HIV-1 (6–8) or by possibly interacting with HIV-1 proteins (9, 10). However, these noncatalytic effects on HIV inhibition may be attributable to the overexpression of A3G and may not be occurring during normal infection (11–13). It is important to bear in mind that any actions of A3G on DNA, catalytic and possibly noncatalytic, are balanced *in vivo* against cellular RNA binding, which forms a high molecular mass A3G-RNA complex, which may prevent A3G incorporation into virions (7, 14–16), and against HIV RNA binding, which forms an intravirion A3G complex in which A3G must be activated by HIV RNase H for DNA deamination to ensue (16).

We have shown that A3G-catalyzed deamination occurs processively while exhibiting a 3′ → 5′ polarity favoring deamination toward the 5′-region of ssDNA (15). Directional deamination is an intrinsic property of A3G, occurring in the absence of an obvious source of energy (*e.g.* ATP or GTP) (15) and can in principle contribute to the HIV-1 G → A mutational bias, increasing in a 3′-direction in the RNA genome (17, 18). A3G processivity appears consistent with a three-dimensional facilitated diffusion process involving sliding and jumping motions along the DNA (15), where jumping involves microscopic dissociation and reassociation events occurring with the same DNA molecule, rather than diffusion into the surrounding bulk solution (19–21). Three-dimensional processive movement has been documented for a variety of site-specific DNA-targeting proteins (*e.g.* restriction endonucleases) (21–24), but in contrast to A3G, catalysis along the DNA is random. Random processive deaminations also occur with activation-induced cytidine deaminase (AID), an APOBEC family member with a single deaminase domain (25).

We have previously suggested that A3G deamination is strongly favored toward the 5′-direction, perhaps because of a catalytic orientation specificity and structural asymmetries in the enzyme (15). The presence of multiple catalytic and binding domains in A3G contained in an elongated rod-shaped oligomeric structure found in this deaminase family (26, 27) has the potential to impose catalytic directionality. A3G has been structurally characterized in solution as a tail-to-tail dimer (26), but monomeric forms are also active (28–30). When bound to DNA, A3G may exist in different oligomeric states, as implied by multiple A3G-DNA bands in gel shift analyses (15, 31).

* This work was supported, in whole or in part, by National Institutes of Health Grants ESO13192 and R37GM21422 (to M. F. G.) and Grant GM79480. This work was also supported by American Cancer Society Grant RSG-03-047 (to D. A. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1 and Figs. S1–S6.

[1] To whom correspondence should be addressed: Dept. of Biological Sciences, Molecular and Computational Biology Section, 1050 Childs Way, Los Angeles, CA 90089. Fax: 213-821-1138; E-mail: mgoodman@usc.edu.

[2] The abbreviations used are: ssDNA, single-stranded DNA; AFM, atomic force microscopy; dsDNA, double-stranded DNA; HIV, human immunodeficiency virus; nt, nucleotide(s).

In the present work, using human A3G produced from baculovirus-infected *Sf*9 insect cells, we have applied presteady state kinetic analyses in combination with atomic force microscopy (AFM) to investigate the dynamic and structural determinants of A3G processivity and catalytic polarity. We provide evidence showing that the 3′ → 5′ bias in deamination of C → U appears to be determined by the oligomeric state of the enzyme, with populations of dimers, tetramers, and perhaps larger oligomers favoring directionally biased deamination and most likely monomers catalyzing random deamination. Both the directional and random deamination modes occur with high processivity, readily observable on a time scale of a few seconds.

EXPERIMENTAL PROCEDURES

Substrates—DNA oligonucleotides were synthesized on an Applied Biosystems 3400 DNA/RNA synthesizer and are listed in supplemental Table S1. Fluorescein-dT and pyrrolo-dC were purchased from Glen Research. The partially dsDNA substrates were constructed by annealing with a 1.4-fold molar excess of complementary ssDNA. A circular DNA substrate was constructed by 5′-end phosphorylation of a 118-nt ssDNA substrate with T4 polynucleotide kinase (New England Biolabs) and subsequent ligation using a linker DNA oligonucleotide. The circular product was then purified after resolution on a 10% denaturing polyacrylamide gel.

A3G and AID Expression and Purification—A baculovirus-expressed GST-A3G protein was constructed as described previously (15). A3G recombinant virus was obtained as described previously (32). *Sf*9 cells were infected with recombinant A3G virus at a multiplicity of infection of 1 and harvested after 72 h. Cells were lysed as described previously (15) in the presence of 50 μg/ml RNase A. Cleared lysates were then incubated with glutathione-Sepharose resin (GE Healthcare) and subjected to a series of salt washes (0.25–1 M NaCl) before elution in buffer as previously described (15). A3G was treated with 0.02 units/μl thrombin (GE Healthcare) for 16 h at 21 °C to release the glutathione *S*-transferase tag, and subsequently the sample was diluted to 50 mM NaCl and loaded onto a DEAE FF column (GE Healthcare) equilibrated with 50 mM Tris-HCl, pH 8.9, 50 mM NaCl, 1 mM dithiothreitol, and 10% glycerol. A3G was eluted at ~200 mM NaCl by using a linear gradient from 75 to 1000 mM NaCl. Fractions were collected and stored at −70 °C. A3G is ~95% pure. AID was purified as previously described (32).

Deamination Assays—A3G activity on ssDNA substrates was measured in reactions containing 50 nM ^{32}P-labeled DNA, 5 or 500 nM A3G in buffer (50 mM HEPES, pH 7.3, 1 mM dithiothreitol, and varied amounts of $MgCl_2$ or other salts where indicated) and incubated for 2.5–60 min at 37 °C. ssDNA substrates (0.3–0.5 μM) with an internal fluorescein label were incubated with A3G or AID (20–300 nM), and deaminations were carried out as described above. All reactions with partial dsDNA substrates also contained 30 mM NaCl to stabilize the duplex region. Reactions were terminated by an extraction with phenol/chloroform/isoamyl alcohol (25:24:1), and deaminations were detected by the uracil DNA glycosylase assay, as described previously (15). Circular substrates were linearized before PAGE resolution by annealing to a 30-nt complementary oligonucleotide containing an MseI site and incubating with 5 units of MseI for 60 min at 37 °C. Gel band intensities were visualized by phosphorimaging or with a FX fluorescence scanner (Bio-Rad) and then measured with ImageQuant software (Amersham Biosciences).

Presteady state reactions were mechanically sampled and quenched using a KinTek quench-flow instrument. A solution of DNA and A3G at twice the final intended concentration was rapidly mixed in equal volumes and allowed to react for a given amount of time (0.1–180 s) before quenching with phenol/chloroform/isoamyl alcohol (25:24:1). Samples were then treated as described above.

Analysis of Processive Deamination by a Single A3G or AID—The analysis of processive deamination described previously (15, 25) was used. This entails the quantification of integrated band intensity of fractions of substrates with a single deamination at the 5′ C or 3′ C or a double deamination at both the 5′ C and 3′ C. The predicted fraction of independent double deaminations is then calculated by multiplying the fraction of all deaminations occurring at the 5′ C (sum of the integrated gel band intensities from deaminations occurring at 5′ C and double deaminations (5′ C and 3′ C)) to the fraction of all deaminations occurring at the 3′ C (sum of the integrated gel band intensities from deaminations occurring at 3′ C and double deaminations (5′ C and 3′ C)). The ratio of the observed fraction of double deaminations (occurring at both 5′ C and 3′ C on the same molecule) to the predicted fraction of independent double deaminations is the "processivity factor" and was used to determine processive (*i.e.* correlated double) deamination of two target motifs. A ratio larger than 1 indicates that a majority of double deaminations are caused by the same A3G or AID molecule acting processively on both target motifs.

Presteady State Anisotropy and Fluorescence Assays—Reactions were at 37 °C in buffer containing 50 mM HEPES, pH 7.3, 1 mM dithiothreitol, and 0 or 5 mM $MgCl_2$.

Stopped flow anisotropy measurements were performed using an Applied Photophysics Π*180 spectrofluorometer equipped for anisotropy by exciting with vertically polarized light at 494 nm (2.5-nm slit width) from a xenon lamp and monitoring emissions using a 515-nm cut-off filter. To determine on-rates, 0.8–2 μM A3G was incubated in one syringe, and 200 nM fluorescein-labeled 69-nt ssDNA substrate was incubated in a second syringe, before rapidly mixing 50 μl from each syringe and monitoring the change in anisotropy for 0.2 s. The anisotropy change as a function of time was fit to a double exponential equation to obtain on-rates. To determine off-rates, in one syringe, 100 nM fluorescein-labeled 69-nt ssDNA substrate was preincubated with 400 nM A3G. The second syringe contained 10 or 20 μM unlabeled ssDNA substrate. The reaction was initiated by mixing equal volumes of the syringes (50 μl) and monitoring the change in anisotropy for 500–1000 s. The anisotropy change as a function of time was fit to a double exponential decay to obtain off-rates.

Stopped flow fluorescence measurements were performed using an Applied Photophysics Π*180 spectrofluorometer equipped for fluorescence. The ssDNA containing pyrrolo-C was excited at 335 nm (2-nm slit width) with a xenon-mercury lamp, and emissions were monitored using a 435-nm cut-off filter. To determine fluorescence changes, measured in volts,

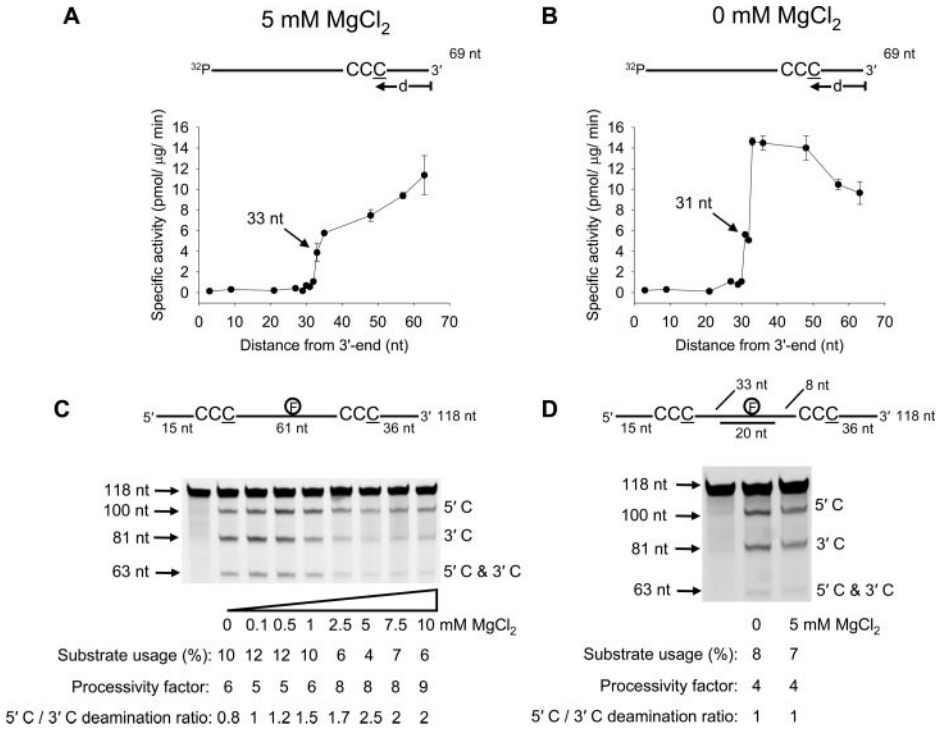

FIGURE 1. **A3G deamination activity, polarity, and processivity on ssDNA is dependent on $MgCl_2$ and the presence of a dsDNA region.** *A* and *B*, specific activity of A3G (5 nM) is plotted as a function of the distance (*d*) of the CCC motif from the 3′-end of a 69-nt ssDNA substrate (50 nM) deaminated in the presence of 5 mM $MgCl_2$ (*A*) or absence of $MgCl_2$ (*B*). *Error bars*, ±1 S.E. from three independent experiments. *C* and *D*, A3G deamination of a 118-nt internally fluorescein-labeled ssDNA substrate at an enzyme/DNA ratio of 1:20 in the absence (*C*) or presence (*D*) of a 20-nt complementary oligonucleotide annealed between the two C targets. Single deaminations of the 5′ C and 3′ C are detected as the appearance of labeled 100- and 81-nt fragments, respectively; double deamination of both Cs on the same molecule results in a 63-nt labeled fragment (5′ C and 3′ C). Substrate usages (%), processivity factors, and ratios of 5′ C/3′ C deaminations are shown *below* the *gels*. For deamination reactions with partially dsDNA, 30 mM NaCl was present to stabilize the duplex region. The reaction on ssDNA shown in *C* is unaffected by the presence of 30 mM NaCl.

2 μM A3G was incubated in one syringe, and 200 nM pyrrolo-C-labeled 69-nt ssDNA substrate was incubated in a second syringe, before rapidly mixing 50 μl from each syringe and monitoring the change in fluorescence emission for 0.6 s. The fluorescence change as a function of time was fit to a double exponential equation to obtain rates.

Atomic Force Microscopy—A3G imaging buffer (50 mM HEPES, pH 7.3) either with or without 5 mM $MgCl_2$ was heated to 65 °C and allowed to cool slowly to dissolve salts. A3G was then diluted to 100 nM in A3G imaging buffer, and 0.1 mM dithiothreitol (final concentration) was added to the sample. Ten microliters of the protein sample were then deposited onto freshly cleaved mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water, excess water was wicked from the surface using filter paper, and the surface was then dried using a stream of nitrogen before being placed in the AFM for imaging. For experiments done in the presence of DNA substrate, 200 nM (final concentration) of ssDNA substrate was added to the A3G sample before depositing as for the protein alone.

All AFM images were captured in air using either a Nanoscope III or IIIa (Digital Instruments, Santa Barbara, CA) microscope in tapping mode. Pointprobe Plus tapping mode silicon probes (Agilent, Tempe, AZ) with resonance frequencies of ~170 kHz were used for imaging. Images were collected at a speed of 2–3 Hz with an image size of 1 × 1 μm at 512 × 512 pixel resolution. Each experiment was repeated at least twice. Volume analysis was done as previously described (33, 34). The program Kaleidagraph (Synergy Software, Reading, PA) was used to generate statistical plots.

RESULTS

Electrostatic Influence on A3G Deamination Polarity and Processivity—We previously observed that A3G deamination on linear ssDNA exhibits a marked deamination bias, where the rate of deaminating the third C in the 5′-CCC target to U is enhanced significantly as the motif is located further away from the ssDNA 3′-end (15). The earlier experiments were performed in the presence of salt ($MgCl_2$) (15). To study A3G deamination bias at higher resolution, with and without salt, we used 11 ssDNA substrates, each with a 5′-A(G/T)A CCC AAA hot spot motif in a different location along a 69-nt ssDNA molecule (Fig. 1, *A* and *B*).

A "dead" zone containing 30 nt was present at the 3′-end of the ssDNA substrate, where the specific activity was <1 pmol/μg/min irrespective of target position or salt concentration (Fig. 1, *A* and *B*). We speculate that the 30-nt dead zone may be the optimal space A3G needs at a 3′-ssDNA end to encounter a CCC motif in an orientation allowing catalysis. There was a monotonic increase in A3G specific activity, in the presence of 5 mM $MgCl_2$, as the target motif was moved out of the dead zone toward the 5′-end, achieving a maximum specific activity of 11 pmol/μg/min (Fig. 1*A*). In the absence of $MgCl_2$ the activity of A3G reached a peak near the center of the ssDNA (~15 pmol/μg/min) and then remained roughly constant before decreasing from about 50 to 69 nt from the 3′-end (0 mM $MgCl_2$) (Fig. 1*B*).

A3G processivity and deamination polarity can be measured simultaneously by placing two CCC motifs on the same ssDNA substrate and having a fluorescein label in between (15, 25) (Fig. 1*C*). The processivity factor (see "Experimental Procedures") measures the fraction of correlated double deaminations occurring at two CCC motifs located on the same ssDNA substrate by a single enzyme molecule compared with uncorrelated double deaminations catalyzed independently by two different A3G molecules (15, 25). The analysis requires that less than ~15% of the substrate be used to maintain single-hit kinetics (35). A3G processivity did not change significantly as a function of $MgCl_2$ concentration, where the processivity factors are in the range of 5–9 (Fig. 1*C*). A measurement of A3G deamination

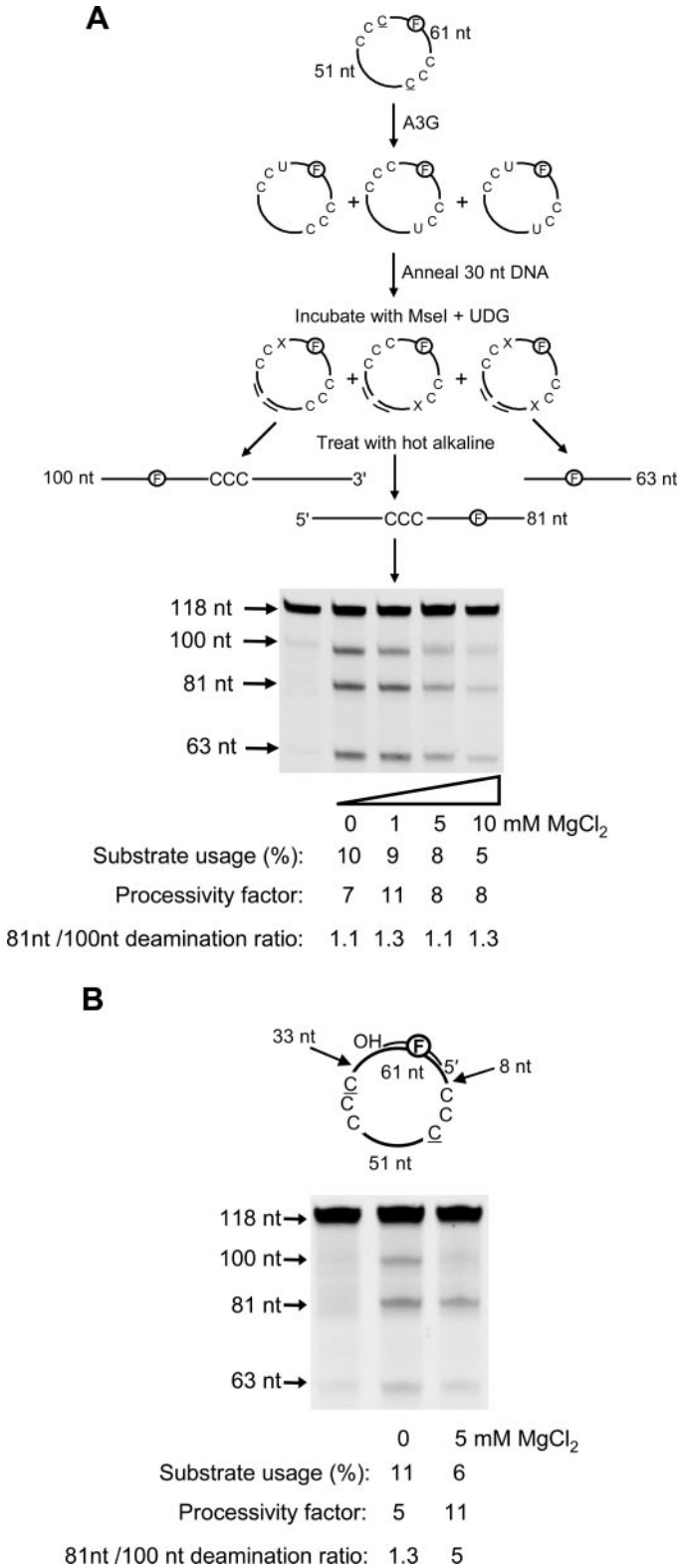

FIGURE 2. **A3G directional bias is absent on circular ssDNA and present on partially double-stranded circular DNA.** A3G deamination bias and processivity were characterized on circular fluorescein-labeled DNA substrates containing two CCC motifs using an enzyme/DNA ratio of 1:20. *A* (*sketch*), deamination events on a 118-nt circular ssDNA are visualized after a 30-nt complementary oligonucleotide is annealed to the circular DNA to enable cutting by MseI (*sketch*). This procedure linearizes the circular DNA so that all possible single and double deamination events can be detected using the uracil DNA glycosylase (UDG) assay (see "Experimental Procedures"). An *X* represents an abasic site. *A*, A3G exhibits processive but not directionally biased deamination on circular DNA as a function of $MgCl_2$. *B*, A3G deamination of circular substrate in the presence of a 20-nt complementary oligonucleotide annealed between the two C targets is processive, and the directional bias is restored with 5 mM $MgCl_2$. Substrate usages (%), processivity factors, and ratios of 81-nt fragment to 100-nt fragment are shown *below* the *gels*. For deamination reactions with partially dsDNA, 30 mM NaCl was present to stabilize the duplex region. The reaction on circular ssDNA shown in *A* is unaffected by the presence of 30 mM NaCl.

polarity on the two C targets, which are out of the deamination dead zone, as a function of salt concentration, shows that deamination at the 5′ C target (100 nt from the 3′-end) is favored over the 3′ C target (36 nt from the 3′-end) by roughly 2-fold at $MgCl_2$ concentrations of >5 mM (5′ C/3′ C deamination ratio; Fig. 1*C*). The deamination polarity decreased with less salt until, in the near absence of $MgCl_2$, A3G deaminated the C targets with no preference (5′ C/3′ C deamination ratio ~1; Fig. 1*C*).

The directional bias of A3G deamination was clearly electrostatic in origin as opposed to a specific interaction with Mg^{2+}, because favored deamination toward the 5′-end also occurred with $CaCl_2$, $MnCl_2$, and NaCl (supplemental Fig. S1). In the case of monovalent NaCl, a 20-fold higher concentration was required to obtain a similar 5′-end deamination bias observed for divalent salts.

We reexamined the polar and processive properties of A3G on the 118-nt ssDNA (Fig. 1*C*) but with a 20-nt oligonucleotide annealed between the deamination targets (Fig. 1*D*). Since A3G bound dsDNA with a low GC content at ~5-fold lower affinity than ssDNA (supplemental Fig. S2) and does not bind dsDNA with a high GC content (15), a loss in the sliding component of the processive motion is expected (15, 31). However, the jumping component of processivity appeared to be retained, since the processivity factor was reduced from 8 to 4 (Fig. 1*D*). The lack of deamination bias in the absence of $MgCl_2$ (5′ C/3′ C deamination ratio is 1; Fig. 1*D*) was expected (see Fig. 1*C*, 0 mM $MgCl_2$). However, the annealed dsDNA region between the two deamination targets also resulted in a loss of deamination bias in the presence of 5 mM $MgCl_2$ (5′ C/3′ C deamination ratio is ~1; Fig. 1, compare *C* and *D*), indicating that the dsDNA between the deamination sites decoupled the two sides of the substrate. The deamination in each single-stranded region depended on the target size for each region, the 36-nt target length from the 3′-end of the 118 nt substrate and the 33-nt target length region between the 5′ C and the dsDNA (Fig. 1*D*). There is, however, an important distinction to be made regarding the occurrence of the dead zone (Fig. 1, *A* and *B*). The ~30-nt dead zone occurs only at the 3′-end of the ssDNA; there is no internal dead zone created by the dsDNA region, but instead deamination is dependent of the nucleotide target size (15).

Deamination Polarity on Circular ssDNA Requires an Intervening dsDNA Region—To examine what effect the ssDNA ends have on A3G deamination bias, ends were eliminated by circularizing the molecule depicted in Fig. 1*C* (Fig. 2*A*, *sketch*). A3G shows essentially no deamination bias either with or without $MgCl_2$ on the circular molecule (81 nt/100 nt deamination ratio ~1.1–1.3; Fig. 2*A*). Although deamination was slightly inhibited in the presence of $MgCl_2$, the apparent binding constants of the circular DNA compared with its linear counterpart

were the same (K_d ~ 160 nM). A3G was processive on the circular substrate with processivity factors similar to those on the linear substrate (compare Fig. 2*A* with Fig. 1*C*).

We examined the effect on deamination of an oligonucleotide annealed to the circular DNA substrate to form a partially dsDNA circle (Fig. 2*B*). To ensure that A3G had enough room to deaminate both C targets equally (see Fig. 1*B*), we placed the dsDNA at least 33 nt from the 3′-side of each CCC motif. A3G was able to act processively on partially dsDNA circles (processivity factor between 5 and 11; Fig. 2*B*). Without $MgCl_2$, A3G deaminated each C target with about equal preference (81 nt/100 nt deamination ratio of 1.3; Fig. 2*B*). However, with 5 mM $MgCl_2$, A3G regained the ability to deaminate with a directional bias (81 nt/100 nt deamination ratio of 5; Fig. 2*B*).

The dsDNA segment, in the presence of $MgCl_2$, may reinstate the deamination bias by limiting the available nucleotide target size. That A3G now favors the C target with the largest target size for binding (*i.e.* that furthest from the 3′-end of the dsDNA segment) (87 nt; Fig. 2*B*) implies that there is an intrinsic 3′ → 5′ directional catalytic constraint for A3G deamination so that deamination occurs predominantly when A3G approaches a C target with a 3′ → 5′ orientation on ssDNA, consistent with our previous data (15). A3G showed no deamination bias on a completely single-stranded circle (Fig. 2*A*), presumably because the enzyme can encounter either of the two C targets in a 3′ → 5′ orientation without a nucleotide target size limitation imposed by linear ssDNA ends. A3G did not use a terminal 3′-OH for directional orientation (supplemental Fig. S3). Deamination assays with a biotin/streptavidin block at the 3′- or 5′-end showed no difference in the deamination activity or directional bias, which suggests that preferential binding at an ssDNA end is probably not occurring (data not shown).

Temporal Analysis of A3G Deamination Polarity and Processivity—To make a finer tuned analysis of the two deamination modes of A3G, directionally biased or random, we used rapid-quench kinetics to investigate deamination polarity and processivity on short time scales (several seconds). We had previously established that during "long" (10-min) incubations, A3G-catalyzed C deaminations occur processively and with 3′ → 5′ polarity in the presence of salt (15) (Fig. 1*C*). Notably, the same degree of processivity and deamination polarity in the presence of salt was observed during "short" incubations (Fig. 3, *A* and *B*). The reactions with 5 mM $MgCl_2$ (Fig. 3*A*) show that at the earliest time that deaminations are detectable (~5 s), deamination of the 5′ C is favored over the 3′ C by a factor of 9. In the absence of $MgCl_2$, A3G deaminated each C target with no preference; the "3′ C" was 36 nt from the 3′-end and was fully accessible for deamination (Fig. 3*B*). The processivity factors decreased substantially, approaching ~1 (no processivity) as substrate usage went from 4 to >50% (Fig. 3, *A* and *B*). The loss of processivity at high substrate usage probably resulted from multiple encounters of A3G with the ssDNA substrates, leading to independent double deaminations (15), whereas, at low substrate usage, each ssDNA substrate is expected to encounter A3G only once, giving rise to processive correlated double deaminations.

To characterize the mechanism of deamination polarity and to obtain A3G single turnover deamination rates at different target motif locations along the ssDNA substrate, we employed three of the 69-nt ssDNA molecules used previously (see Fig. 1, *A* and *B*) with one 5′-A(G/T)A CCC AAA hot spot motif either 57, 36, or 27 nt from the 3′-DNA end in reactions with a large excess of A3G over substrate DNA (A3G/DNA ~ 10:1) (Fig. 3, *C* and *D*). The ssDNA molecules were bound by A3G with the same apparent K_d (100 nM) (supplemental Fig. S4).

The single turnover rate for a reaction containing 5 mM $MgCl_2$ for a C target 57 nt from the 3′-end fits well to a single exponential and yields a deamination rate of 0.25/s (Fig. 3*C*), with a maximum of ~100% substrate deamination within 20 s (Fig. 3*C*, *inset*). In contrast, the C targets 36 or 27 nt from the 3′-end required 1500 and 3600 s to reach 100% substrate deamination, respectively (Fig. 3*C*). These data are not fit well by single exponentials and require double exponential fits with a fast and a slow phase. For ssDNA with a C target near the middle (36 nt from the 3′-end), the rate of the fast phase is 0.12/s but only accounts for 23% of the deaminations taking place (Fig. 3*C*, *inset*), whereas the slower rate, 0.004/s (one deamination every 250 s) accounts for 77% of the deamination (Fig. 3*C*). The C target in the ssDNA 3′-end dead zone, 27 nt from the 3′-end, was deaminated the slowest, with a fast rate of 0.02/s occurring for about 24% of the deaminations (Fig. 3*C*, *inset*) and a slower rate of 0.0007/s for the remainder (Fig. 3*C*). Biphasic deamination rates under single turnover conditions suggest that multiple deamination modes of A3G exist that catalyze deaminations at different rates.

In the absence of $MgCl_2$, A3G showed greater ability to deaminate at the center of the ssDNA (Fig. 3*D*), which is consistent with specific activities determined at various locations along the DNA substrate (Fig. 1*B*). The deamination rates for substrates with C targets 57 or 36 nt from the 3′-end were essentially the same, with deaminations occurring every 14 s (~0.08/s; Fig. 3*D*, *inset*). On the substrate with a C target in the ssDNA 3′-end dead zone, 27 nt from the 3′-end, the initial rate (0.03/s) limited deaminations to ~30 s for each and was far slower than the other C targets (Fig. 3*D*, *inset*). The second reaction rate representing 45% of the amplitude for a C target located 27 nt from the 3′-end was 0.001/s (Fig. 3*D*).

When we probed deaminations occurring at individual C sites within each motif, it was evident that the third C in the 5′-CC<u>C</u> was favored for deamination, the middle C (5′-C<u>C</u>C) could also be deaminated on a longer time scale, and the first C (5′-<u>C</u>CC) was barely deaminated (Fig. 3*E*). The general pattern of deaminations within a motif (Fig. 3*E*) was consistent with A3G 3′ → 5′ directional polarity in which triplet motifs undergo more rapid deamination nearer the 5′-end of the ssDNA substrate (Fig. 3, *A* and *C*).

Rates of A3G Binding and Release of ssDNA are Biphasic—It is essential to view A3G deamination time scales (Fig. 3) within the broader milieu of time scales required for A3G-ssDNA complex formation and dissociation. We used presteady state kinetic analyses of rotational anisotropy and fluorescence emission to measure A3G on-rates and off-rates and to

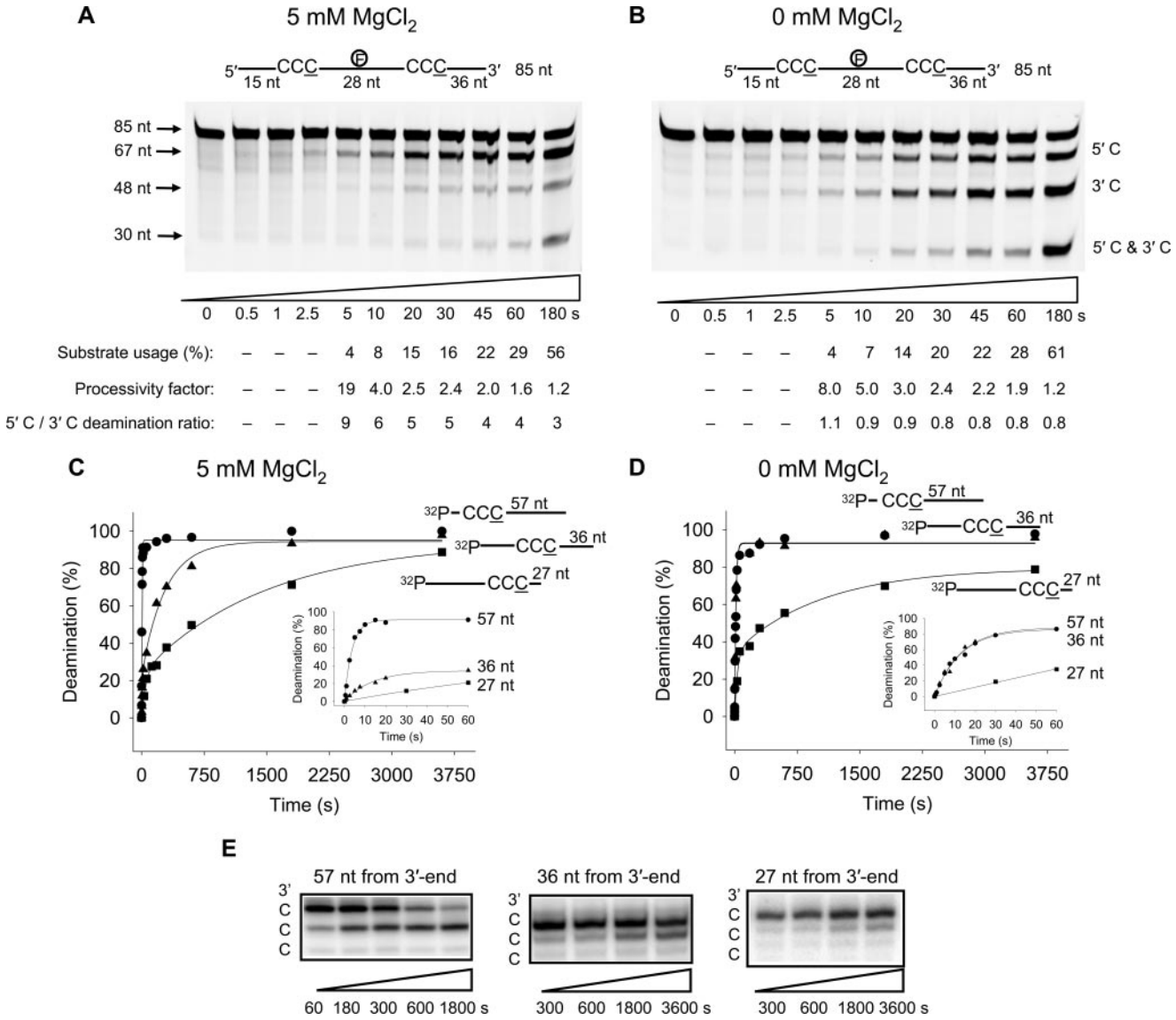

FIGURE 3. **Presteady state analysis of polarity, processivity and deamination rates.** *A* and *B*, processive deamination occurs on a presteady state time scale. A3G was incubated with an 85-nt internally fluorescein-labeled ssDNA substrate with two CCC motifs 28 nt apart and at an enzyme/DNA ratio of 1:1. Single deaminations of the 5′ C and 3′ C are detected as the appearance of labeled 67- and 48-nt fragments, respectively; double deamination of both Cs on the same molecule results in a 30-nt labeled fragment (5′ C and 3′ C). Processive deamination was determined in the presence of 5 mM $MgCl_2$ (*A*) and absence of $MgCl_2$ (*B*). Substrate usages (%), processivity factors, and the ratios of 5′ C/3′ C deaminations are shown *below* the *gels*. *C* and *D*, single turnover reaction rates for A3G were determined on ssDNA substrates with CCC motifs 57 (●), 36 (▲), or 27 nt (■) from the 3′-end of a 69-nt substrate under enzyme saturating conditions (500 nM A3G, 50 nM DNA). *C*, with 5 mM $MgCl_2$, the deamination of a CCC motif 57 nt from the 3′-end, measured as a function of time, was fit to a single exponential equation to determine a single turnover rate of 0.25/s. The deamination of a CCC motif 36 nt from the 3′-end measured as a function of time was fit to a double exponential equation to determine a minor "fast" rate component of 0.12/s (23% amplitude) and a major "slow" rate component of 0.004/s (72% amplitude). The deamination of a CCC motif 27 nt from the 3′-end measured as a function of time was fit to a double exponential equation to determine a minor "fast" rate component of 0.02/s (24% amplitude) and a major "slow" rate component of 0.0007/s (70% amplitude). *Inset graph*, the first 60 s of the reaction. *D*, reaction rates determined as in *C* but in the absence of $MgCl_2$. The data for substrates with CCC motif 57 and 36 nt from the 3′-end were fit to a single exponential equation to obtain rates of 0.07/s and 0.08/s, respectively. The data for substrate with CCC motif 27 nt from the 3′-end were fit to a double exponential equation with an initial "fast" rate of 0.03/s (34% amplitude) and a second "slow" rate of 0.001/s (45% amplitude). *Inset graph*, the first 60 s of the reaction. *E*, examination of how CCC motifs are deaminated. Deamination of the 3′ and middle C residues in the motif 5′-CCC-3′ is influenced by distance from the 3′-end of the ssDNA. Deamination of the 5′ C residue in the motif 5′-CCC-3′ is barely observed. The gel reaction buffers contained 5 mM $MgCl_2$.

investigate possible conformational changes of the ssDNA substrate (Fig. 4).

The on-rate of A3G, determined with a 10-fold excess of A3G to DNA in the presence and absence of $MgCl_2$, was extremely rapid, exhibiting an initial fast phase and slower second phase (Fig. 4, *A* and *B*). About half of the ssDNA was bound within the instrument dead time (1.1 ms). The estimated on-rates under these conditions are 140/s (Fig. 4*A*; 5 mM $MgCl_2$) and 91/s (Fig. 4*B*; without $MgCl_2$). These measurements were also made as a function of A3G concentration (0.4–0.8 μM) to calculate the bimolecular rate constants, which approach the diffusion limit, with or without $MgCl_2$ (5 mM $MgCl_2$, 1×10^8/M/s; 0 mM $MgCl_2$, 9×10^7/M/s). Each on-rate experiment, in the presence or absence of $MgCl_2$,confirmed that the A3G on-rate is biphasic with an initial rapid phase (~60% amplitude) and secondary slow phase (~40% amplitude). It may be that the two distinct

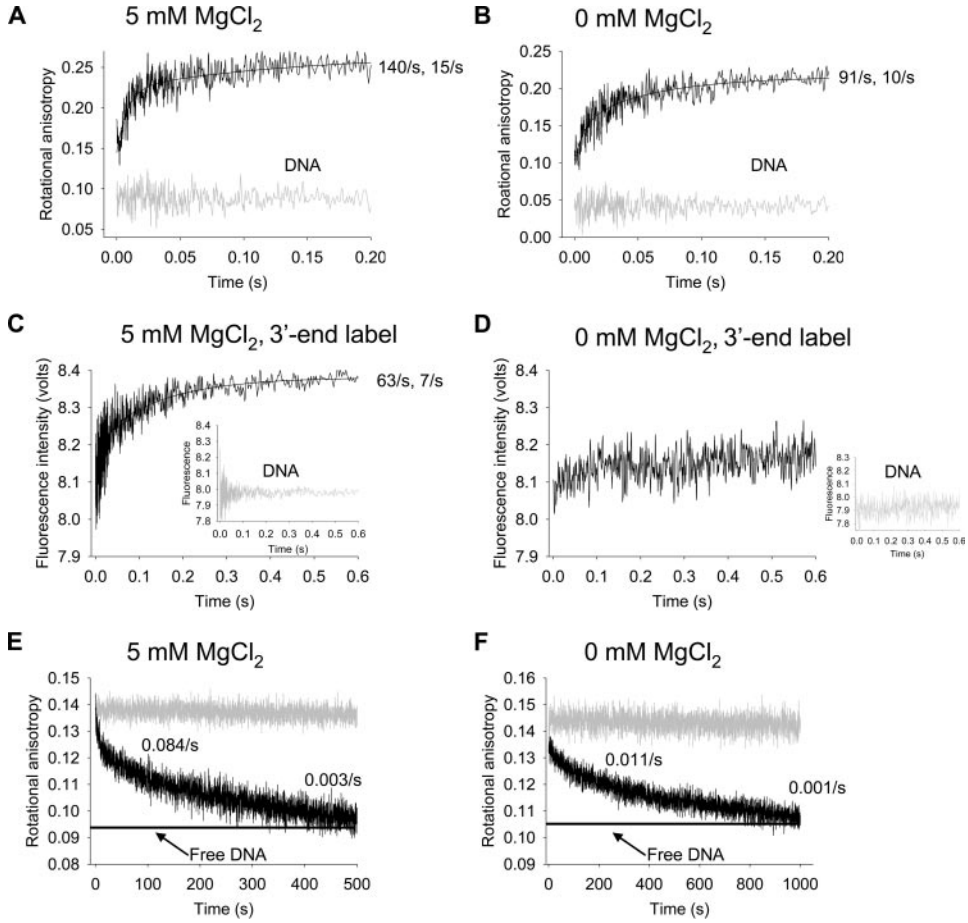

FIGURE 4. **Stopped-flow fluorescence presteady state interactions of A3G with ssDNA.** *A* and *B*, the on-rate of A3G (1 μM) to a 69-nt fluorescein-labeled ssDNA substrate (100 nM) was determined by monitoring changes in rotational anisotropy of free ssDNA (*gray line*) following the addition of A3G (*black line*) as a function of time. The on-rate was fit to a double exponential equation. *A*, for A3G in buffer with 5 mM $MgCl_2$ the rapid rate of 140 ± 14/s (62% amplitude) was followed by a slower rate of 15 ± 2/s (38% amplitude). *B*, for A3G in the absence of $MgCl_2$, the rapid rate of 91 ± 10/s (53% amplitude) was followed by a slower rate of 10 ± 1/s (47% amplitude). *C* and *D*, changes in fluorescence of 100 nM pyrrolo-C-labeled ssDNA (*inset graph*) upon the addition of 1 μM A3G (*black line*) were monitored as a function of time. The pyrrolo-C label was placed 10 nt from the 3′-end of the 69 nt ssDNA. The fluorescence change was fit to a double exponential equation. *C*, under conditions with 5 mM $MgCl_2$ the fluorescence change began with a rapid rate of 63 ± 8/s (40% amplitude) and was followed by a slower rate of 7 ± 0.3/s (60% amplitude). *D*, in the absence of $MgCl_2$, no fluorescence change was detected. *E* and *F*, the off-rate of A3G (0.2 μM) from a 69-nt fluorescein-labeled ssDNA substrate (50 nM) was determined by preincubating A3G with ssDNA (*gray line*) and monitoring changes in rotational anisotropy following the addition of 100–200-fold excess unlabeled 69-nt ssDNA as trap (*black line*). The anisotropy of free fluorescein-labeled ssDNA mixed with unlabeled ssDNA trap is denoted with a *horizontal line*. The off-rate was fit to a double exponential decay. *E*, in the presence of 5 mM $MgCl_2$, the reduction in the rotational anisotropy signal, representing A3G dissociating from fluorescein-labeled DNA and binding the unlabeled trap DNA, had an initial rate of decay fitting to 0.084 ± 0.005/s (20% amplitude) and a second decay rate of 0.00279 ± 0.00006/s (80% amplitude). *F*, in the absence of $MgCl_2$, the reduction in the rotational anisotropy signal had an initial rate of decay fitting to 0.011 ± 0.001/s (25% amplitude) and a second decay rate of 0.0010 ± 0.0002/s (75% amplitude).

on-rates reflect different interactions of a single A3G form with DNA or that multiple oligomeric forms of A3G exist and interact differently with DNA.

It is likely that when an A3G molecule, which has two DNA binding domains per monomer, binds to ssDNA, there will be multiple binding interactions along the substrate, which may cause a conformational change in the structure of the ssDNA. To detect conformational changes in the ssDNA, such as base unstacking, we measured the presteady state change in fluorescence intensity of a 2′-deoxycytidine analogue, pyrrolo-dC, which was located 10 nt from the 3′-end of a 69-nt ssDNA (Fig. 4, *C* and *D*). Notably, biphasic changes in fluorescence were observed with 5 mM $MgCl_2$ (Fig. 4*C*), having a fast rate of 63/s (40% amplitude) and slower rate of 7/s (60% amplitude). The data at 5 mM $MgCl_2$ suggest that there was a conformational change in the ssDNA occurring at a considerably slower rate (63/s; Fig. 4*C*) than A3G-ssDNA binding (140/s; Fig. 4*A*). In contrast, there was almost no detectable fluorescence change in the absence of $MgCl_2$ of a pyrrolo-C residue located 10 nt from the 3′-end (Fig. 4*D*).

An estimate of the spatial properties of the conformational change can be made by measuring the change in steady state fluorescence intensity for pyrrolo-C residues placed at different locations along the ssDNA (supplemental Fig. S5*A*). There was a significant increase in fluorescence at all locations along the DNA in the presence of $MgCl_2$, with maxima occurring at 10 nt from the 5′- and 3′-ends, corresponding to the locations of the fluorophores in the presteady state analysis (Fig. 4*C* and supplemental Fig. S5, *A* and *B*). In the absence of $MgCl_2$, there was essentially no change in fluorescence intensity along the DNA except when the pyrrolo-C was located 10 nt in from the 5′-end (supplemental Fig. S5, *A* and *C*). The presteady state and steady state fluorescence intensity measurements are mutually supportive and suggest that a conformational change in the DNA, perhaps indicative of DNA wrapping, occurs most strongly with 5 mM $MgCl_2$ (supplemental Fig. S5*A*) at ~15 ms after A3G binds to DNA (Fig. 4, *A* and *C*).

A3G must remain bound to the same ssDNA substrate long enough for it to act processively (see Fig. 3, *A* and *B*). To measure the residence time of A3G bound to a single DNA substrate molecule (Fig. 4, *E* and *F*), A3G was prebound to fluorescein-labeled ssDNA, and a large excess (100–200-fold) of the identical unlabeled ssDNA was added as a trap to bind A3G once it dissociated, observed as a decrease in anisotropy from the fully bound starting point. We verified that the off-rates were independent of the trap concentration, suggesting that the unlabeled DNA acts as a passive trap, not as a competitor for A3G binding.

The off-rate of A3G was biphasic with a rapid initial off-rate occurring during the first few seconds (~20% amplitude), fol-

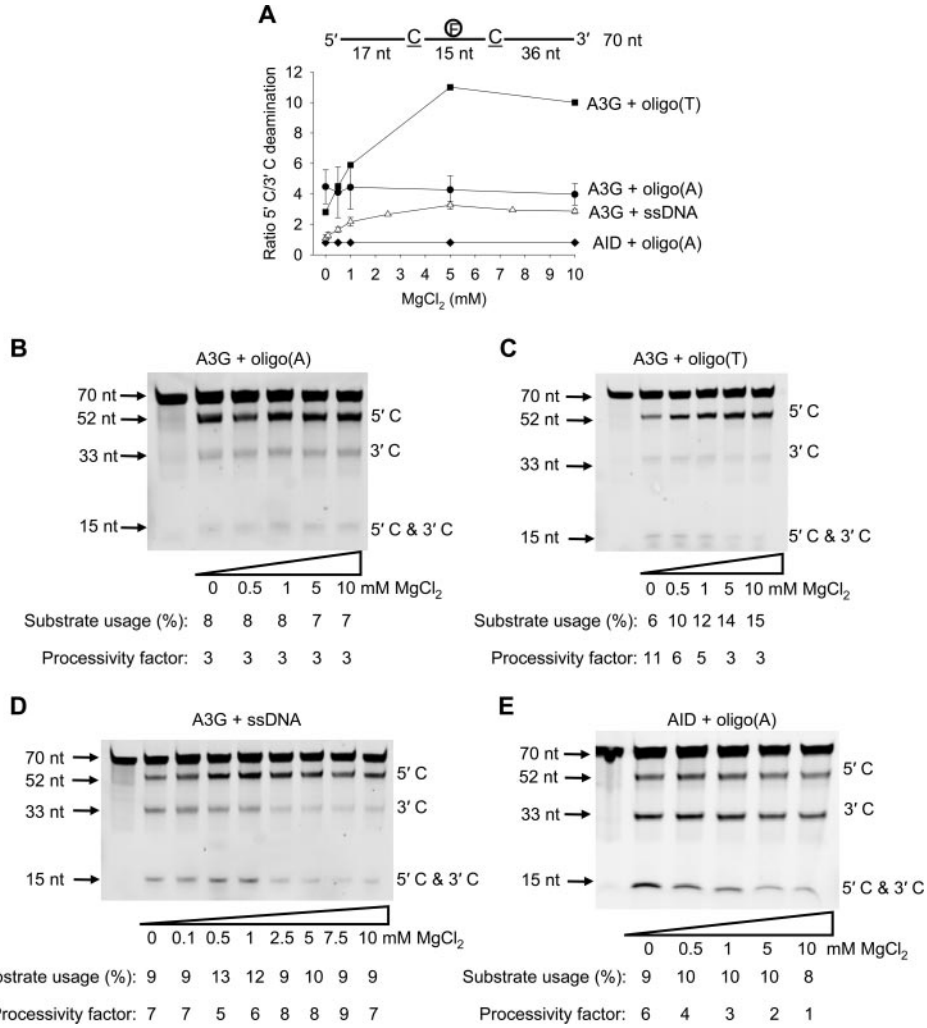

FIGURE 5. **A3G and AID deamination of oligo(A) and oligo(T) substrates containing two target motifs.** A3G or AID was incubated with a 70-nt internally fluorescein-labeled ssDNA substrate with two CCC or AGC motifs 13 nt apart and at an enzyme/DNA ratio of 1:20. Single deaminations of the 5′ C and 3′ C are detected as the appearance of labeled 52- and 33-nt fragments, respectively; double deamination of both Cs on the same molecule results in a 15-nt labeled fragment (5′ C and 3′ C). *A*, A3G deamination bias is affected by the sequence composition. The ratio of 5′ C/3′ C deamination is shown as a function of $MgCl_2$ concentration for each variation of the 70-nt internally fluorescein-labeled ssDNA substrate tested. *Error bars* show ±1 S.E. from three independent experiments. Shown is A3G deamination on an oligo(A) sequence ($A_{15}CCCA_6T_1A_6CCCA_{36}$) (*B*), an oligo(T) sequence ($T_{15}CCCT_{13}CCCT_{36}$) (*C*), and an ssDNA (*D*). *E*, AID deamination on an oligo(A) sequence ($A_{16}GCA_7T_1A_6GCA_{36}$). Substrate usages (%) and processivity factors are shown *below* the *gels*.

lowed by a much slower off-rate occurring on a time scale of minutes (~80% amplitude) (Fig. 4, *E* and *F*). The residence time (reciprocal of the off-rate) for the A3G rapidly dissociating small amplitude phase was ~12 s (0.084/s) with salt (Fig. 4*E*) and about 8-fold longer (~90 s) without salt (0.011/s; Fig. 4*F*). The second much more stable binding mode was characterized by residence times of about 300 s and 1000 s, in the presence and absence of $MgCl_2$, respectively (Fig. 4, *E* and *F*). It is likely that the two distinct off-rates reflect the stability of different A3G-DNA complexes. The key point, however, is that A3G can remain bound to the DNA on the order of minutes, which provides ample time to catalyze the correlated deaminations that take place within 5 s of a single A3G-DNA encounter (Fig. 3, *A* and *B*).

A3G Deamination Polarity Dependence on DNA Topology—Target CCC motifs were embedded in homopolymer A and T regions to investigate the effect of DNA topology on A3G deamination bias (Fig. 5). Since A3G can cause a conformational change in DNA in the presence of $MgCl_2$ (Fig. 4*C*) we expected that A3G would be sensitive to DNA topology. Oligo(A) ssDNA substrates are known to have more rigid conformations than oligo(T) or ssDNA (36). When acting on an oligo(A) substrate, ($A_{15}CCCA_6T_1A_6CCCA_{36}$, where T contains fluorescein label), A3G deaminated the 5′ C with the same 5′ C/3′ C deamination ratio (~4) regardless of $MgCl_2$ concentration (Fig. 5, *A* and *B*). This result is in contrast to ssDNA, where A3G showed no deamination bias in the absence of $MgCl_2$ (Fig. 5, *A* and *D*; see Fig. 1*C*). When acting on the more flexible oligo(T) substrate, ($T_{15}CCCT_{13}CCCT_{36}$), A3G exhibited a deamination bias in the presence or absence of $MgCl_2$, but in contrast to oligo(A) and ssDNA, there was a marked increase in the deamination bias with increasing $MgCl_2$, going from a 5′ C/3′ C deamination ratio of 2.8 (no salt) to 10 (10 mM $MgCl_2$) (Fig. 5, *A* and *C*). In contrast, the APOBEC family member, AID, deaminated ssDNA in an unbiased manner in the presence and absence of $MgCl_2$ on ssDNA (25) or on the oligo(A) substrate ($A_{16}GCA_7T_1A_6GCA_{36}$; Fig. 5, *A* and *E*).

The Oligomerization State of A3G Depends on $MgCl_2$ and DNA—The presteady state kinetic data revealed biphasic interactions of A3G with DNA (Fig. 4) and for some conditions biphasic deamination rates (Fig. 3, *C* and *D*). This could mean that a homogenous population of A3G dimers (26) interacts with the DNA in two modes or that different oligomeric states of A3G are present and exhibit different kinetic rates.

AFM, which yields topographic images of molecules deposited on a surface, can be used effectively for examining the oligomerization states of proteins (33, 34, 37–42). Specifically, it has been demonstrated that the volume of proteins (*V*) measured by AFM exhibits a linear dependence on molecular weight (*MW*) with $V = 1.2 \times MW - 14.7$ (33, 34). Accordingly, we used AFM to characterize the oligomeric state of A3G (45.6 kDa) on ssDNA, oligo(A), and oligo(T) substrates, in the presence and absence of $MgCl_2$ (Fig. 6).

A3G alone (*i.e.* in the absence of ssDNA) in the presence of salt exists as a mix of monomers (~40%) and a considerably

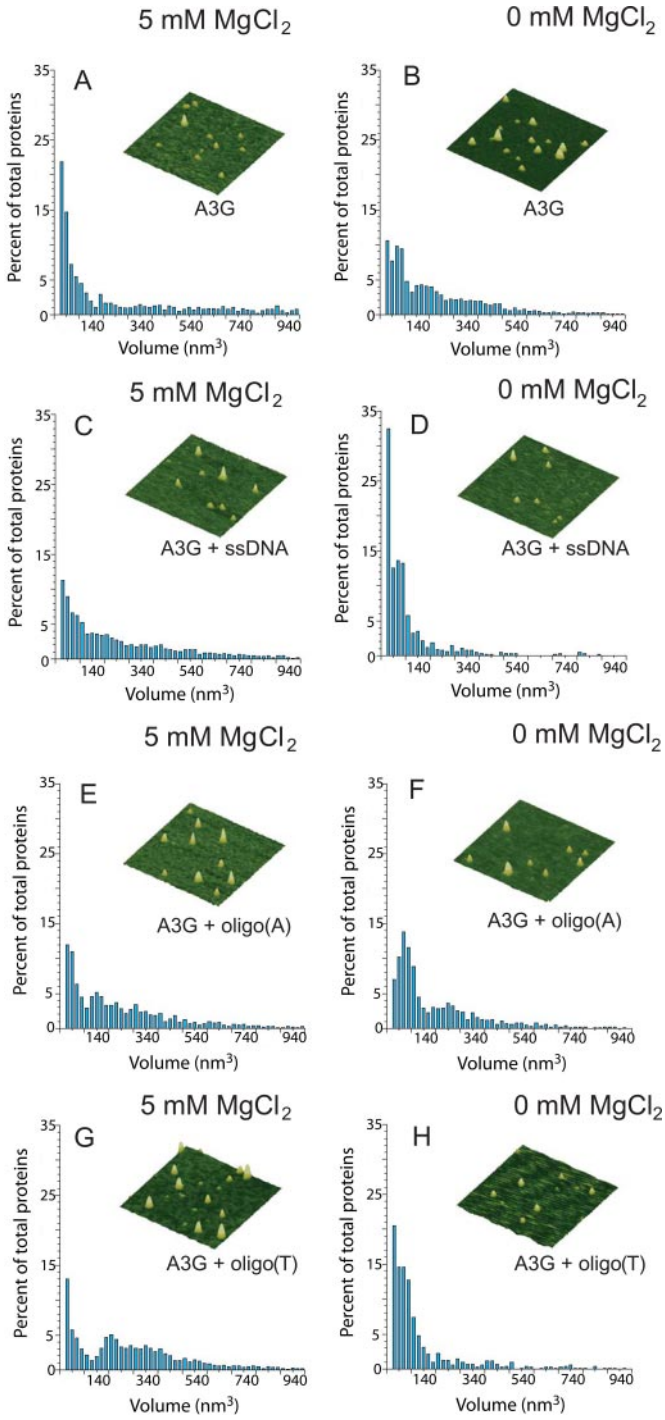

FIGURE 6. **$MgCl_2$ and DNA binding modulate the oligomerization state of A3G.** AFM was used to determine the oligomeric state of A3G alone in the presence of 5 mM $MgCl_2$ (*A*), alone in the absence of $MgCl_2$ (*B*), in the presence of ssDNA and 5 mM $MgCl_2$ (*C*), and in the presence of ssDNA and absence of $MgCl_2$ (*D*). The oligomeric state of A3G was also determined in the presence of oligo(A) and 5 mM $MgCl_2$ (*E*), in the presence of oligo(A) and absence of $MgCl_2$ (*F*), in the presence of oligo(T) and 5 mM $MgCl_2$ (*G*), and in the presence of oligo(T) and absence of $MgCl_2$ (*H*). The 70-nt ssDNA was added to 200 nM final concentration. A3G was added to 100 nM final concentration. Volume distributions of A3G are plotted against the percentage of total proteins. Total proteins counted were as follows: 1354 (*A*), 3168 (*B*), 4231 (*C*), 663 (*D*), 1563 (*E*), 1224 (*F*), 6738 (*G*), and 706 (*H*). Representative AFM images of A3G are shown for each condition. Images are 300 × 300 nm with a height scale of 5 nm.

TABLE 1

Distribution of A3G oligomerization states in response to $MgCl_2$ and DNA

Data are summarized from AFM experiments shown in Fig. 6.

AFM conditions	Distribution of oligomerization states		
	Monomers (0–70 nm^3)	Dimers (70–120 nm^3)	Higher order oligomers (>120 nm^3)
	%	%	%
A3G alone, 5 mM $MgCl_2$	40	14	46
A3G alone, 0 mM $Mg Cl_2$	23	19	58
A3G + ssDNA, 5 mM $Mg Cl_2$	23	15	62
A3G + ssDNA, 0 mM $Mg Cl_2$	51	26	23
A3G + oligo(A), 5 mM $Mg Cl_2$	26	11	63
A3G + oligo(A), 0 mM $Mg Cl_2$	24	27	49
A3G + oligo(T), 5 mM $Mg Cl_2$	21	7	71
A3G + oligo(T), 0 mM $Mg Cl_2$	43	28	29

smaller population of dimers (~14%), along with higher order oligomers (~46%) (Fig. 6*A* and Table 1). A3G alone in the absence of salt behaved differently, showing about a 2-fold reduction in monomers and an increase in the combined fractions of dimers and higher order oligomers (Fig. 6*B* and Table 1). These AFM data, which visualize A3G in the absence of DNA, are consistent with gel filtration data showing that A3G exists in multiple forms (7, 15). A3G dimers were observed with small angle x-ray scattering in the presence of NaCl and $MgCl_2$, and we speculate that the absence of higher order oligomers may reflect differences in buffer conditions (26).

A3G in the presence of salt and ssDNA resulted in about a 2-fold reduction in monomers (~23%) and an increase in higher order oligomers (62%), relative to the absence of DNA (Fig. 6, compare *A* with *C*, and Table 1). These data (Fig. 6*C*) are consistent with equilibrium binding data for A3G association with ssDNA, which show a slight cooperativity in A3G binding to DNA (15) (supplemental Fig. S4). The polydispersed nature of A3G on DNA (Fig. 6*C*) is also in agreement with previous data showing multiple bands in gel shift analyses (15, 31). In contrast, in the absence of $MgCl_2$, the addition of ssDNA caused a shift in the A3G distribution, resulting in 2-fold more monomers (51%) than dimers (26%), whereas in the absence of salt and DNA, the monomer/dimer ratio was ~1 (Fig. 6, *B* and *D*, and Table 1). The 2-fold increase in monomer/dimer ratio was accompanied by a concomitant 2-fold reduction in higher order oligomers. These data indicate that in the presence of salt, binding to ssDNA influences A3G to associate into higher order oligomers, whereas, in the absence of salt, binding of ssDNA causes the A3G to dissociate from higher order oligomers to predominantly monomers and dimers.

It appears that different oligomeric states may be responsible for the biphasic ssDNA binding and deamination kinetics. Taken together with Fig. 5*D*, the AFM data (Fig. 6, *C* and *D*) suggest that the different oligomeric states present when A3G binds ssDNA under different salt conditions influence whether A3G acts in either a directional or random manner, outside the 30-nt dead zone (Fig. 1, *A* and *B*). We confirmed that this correlation is also found when the AFM conditions, A3G/DNA ratio of 1:2, are used for the deamination reaction (supplemental Fig. S6). We further speculate that oligomers promote directional deamination and monomers promote random deamination (compare Fig. 6, *C* and *D*, with Fig. 5*D*; see Table 1).

The observation that the oligo(A) and oligo(T) ssDNA substrates promote the deamination polarity of A3G without the addition of $MgCl_2$ (Fig. 5*A*) suggests that binding to these DNA substrates may influence the oligomeric state of A3G. To test this idea, we investigated the oligomeric states of A3G upon binding to both oligo(A) and oligo(T) ssDNA substrates in the presence and absence of $MgCl_2$. The distributions of A3G with oligo(A) were similar in the presence and absence of $MgCl_2$ (Fig. 6, *E* and *F*, respectively). The monomeric species were similar, as were the sum of dimeric and higher order species (Table 1). The data are consistent with A3G exhibiting comparable deamination polarities on oligo(A) in both the presence and absence of salt (Fig. 5*B* and supplemental Fig. S6). A comparison of the data in the absence of $MgCl_2$ for no DNA (Fig. 6*B*), oligo(A) (Fig. 6*F*), and ssDNA (Fig. 6*D*) reveals that ssDNA, but not oligo(A), promotes the dissociation of A3G to a predominantly monomer population (Table 1) (2-fold fewer monomers with oligo(A) than ssDNA) in accord with the lack of polarity for A3G deamination of ssDNA in the absence of $MgCl_2$ (Fig. 5*D*). In contrast to the oligo(A) data, the distributions of A3G with oligo(T) in the presence and absence of $MgCl_2$ are different (Figs. 6, *G* and *H*, respectively), with a far higher proportion of molecules being present in higher order oligomeric states in the presence of salt (71%) compared with the absence of salt (29%) (Table 1). These relative distributions agree with the large increase in deamination polarity for A3G on oligo(T) with increasing $MgCl_2$ concentration (Fig. 5*A*, supplemental Fig. S6). In the absence of $MgCl_2$, the distribution for A3G binding to oligo(T) DNA is similar to that for A3G binding to ssDNA, except that the monomer peak is reduced for oligo(T) relative to ssDNA (43% *versus* 51%, Table 1; Fig. 6, compare *H* and *D*). This shift to higher order oligomeric states is consistent with A3G showing polarity on oligo(T) in the absence of $MgCl_2$; the ratio of 5′ C/3′ C deamination is ~2.8 (Fig. 5*A*).

DISCUSSION

We have previously shown that A3G binds randomly to ssDNA and is able to deaminate C residues processively with a 3′ → 5′ bias (15). In this paper, we have investigated the mechanisms of processivity and favored deamination in the 5′-ssDNA direction using a combination of kinetics and AFM analysis. Both attributes of A3G, its polarity in the presence of salt and its processivity in the presence and absence of salt, occur at the shortest times (~5 s) for which deamination can be detected (Fig. 3, *A* and *B*).

Aside from the importance of understanding the mechanisms underlying A3G processivity and especially polarity, which occur in the absence of an external energy source such as ATP or GTP, the intrinsic biochemical properties of A3G are important biologically. The time scales of A3G-DNA complexes, which rapidly form and remain for 5–15 min (Fig. 4, *A*, *B*, *E*, and *F*) suggest that *in vivo* A3G has the potential to catalyze sizable numbers of processive deaminations on (−)-cDNA within the limited time that it is single-stranded. Notably, *in vivo* data reveal twin 5′ → 3′ G to A mutational gradients of the HIV-1 genome, which means that deaminations are increasing toward the 5′-end of the (−)-cDNA in two regions marked by reverse transcription primer sites on the (−)-cDNA (17, 18). It has been suggested that the mutational gradients are attributable to the temporal effects of reverse transcription acting in combination with RNase H, so that regions that remain single-stranded over the longest time frame are most susceptible to A3G-catalyzed deaminations (17, 18). A recent analysis of edited HIV-1 genomes, however, suggests that the intrinsic 3′ → 5′ deamination bias of A3G (15) makes a substantial contribution to local regions of HIV-1 mutational polarity *in vivo* (17).

Our AFM results demonstrate that interactions of A3G with ssDNA and salt can determine A3G oligomeric forms. Binding of A3G to ssDNA at high salt promotes the formation of higher order oligomers, whereas binding at low salt favors the presence of monomers (Table 1). Monomeric forms of A3G observed by AFM in the presence of ssDNA and absence of salt (Fig. 6*D* and Table 1) correlate with random deamination (Fig. 5*D*), whereas oligomeric forms that occur in the presence of ssDNA and salt (Fig. 6*C* and Table 1) correlate with a 3′ → 5′ bias (Fig. 5*D*). However, the polydispersed nature of the larger oligomers of A3G, at 5 mM $MgCl_2$ (Fig. 6*C* and Table 1), precludes the identification of a single oligomeric state responsible for the deamination polarity. Nevertheless, the role of dimers, tetramers, and higher oligomeric forms of A3G in governing the deamination bias is strongly supported by deaminations of CCC motifs embedded within oligo(A) or oligo(T) sequences. A3G deaminates with a 3′ → 5′ deamination bias even in the absence of salt on the oligo(A) and oligo(T) (Fig. 5, *B* and *C*) but not on ssDNA substrates (Fig. 5*D*), and the higher oligomeric states of A3G are more stable on oligo(A) and oligo(T) than on the ssDNA in the absence of $MgCl_2$ (Fig. 6, compare *F* and *H* with *D*; see Table 1). It is conceivable that larger oligomeric states form concatamers of the dimeric form (26) and probably involve C-terminal as well as N-terminal interactions, in accord with the APOBEC2 crystal structure (27), but we cannot resolve such oligomeric interactions with AFM data.

The single turnover deamination rates, determined with A3G in a 10-fold molar excess over ssDNA, show that the $MgCl_2$ concentration dictates the intrinsic catalytic rates of C targets at different locations along an ssDNA molecule. Virtually identical intrinsic deamination rates exist in the absence of salt for target motifs located outside of the deamination dead zone at 57 and 36 nt from the 3′-end (Fig. 3*D*). In the presence of salt, however, the single turnover deamination rates differ substantially for target motifs located 57 and 36 nt from the 3′-end (Fig. 3*C*, *inset*). Comparing the deamination rates with the slowest off-rate measured in the presence of salt (0.003/s; Fig. 4*E*), a C target 57 nt from the 3′-end of ssDNA, having a deamination rate of 0.25/s (Fig. 3*C*) can easily be deaminated before A3G falls off the DNA. However, this is not generally true for the C target 36 nt from the 3′-end of an ssDNA, which is characterized by biphasic deamination kinetics (Fig. 3*C*). Roughly 25% of the substrates are deaminated at 0.12/s, whereas 75% are deaminated at 0.004/s (Fig. 3*C*), the latter rate being about the same as the slowest off-rate (Fig. 4*E*). Thus, in the presence of salt, there is a reduction in intrinsic rates of deamination occurring in targets nearer the 3′-end that seems to account for the 3′ → 5′ polarity.

In the presence of salt, the slower catalytic rate for C-targets closer to the 3′-ssDNA end (Fig. 3*C*) may stem, in part, from

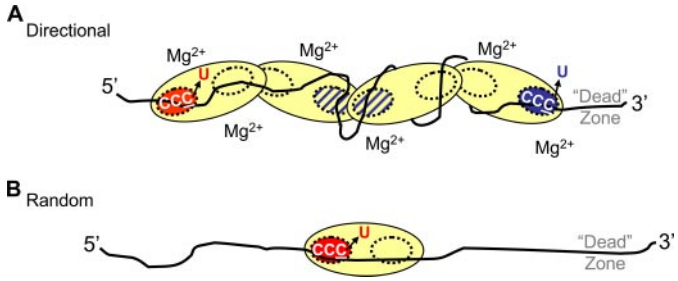

FIGURE 7. **Model for oligomeric regulation of A3G directional and random deamination.** *A*, directional 3′ → 5′ deamination in the presence of "large" A3G oligomers. Larger oligomers, shown here as a tetramer, although dimers may also be present, are stable when binding ssDNA in the presence of Mg^{2+} (Fig. 6*C*). In the presence of Mg^{2+}, there is a reduction in the intrinsic rates of C-deamination toward the 3′-DNA end, which accounts for the 3′ → 5′ polarity (Fig. 3, *C* and *D*). The catalytically inactive N-terminal domains are depicted as *yellow dashed circles*. The hot active site (*red circle*) is aligned in the catalytically favored 3′ → 5′ orientation and closest to the preferential C for deamination in a 5′-CC<u>C</u> target motif, whereas the cold active site (*blue circle*) is aligned in the opposite orientation nearest the barely deaminated C in a 5′-<u>C</u>CC target motif. Active sites located internally (*cross-hatched circles*) cannot effectively access CCC motifs and are essentially noncatalytic. The combination of inaccessible CCC motifs and the cold active site (*blue circle*) located toward the 3′-DNA end results in an inhibition of deamination relative to the easily accessible CCC motifs toward the 5′-DNA end. CCC motifs located within the ~32-nt dead zone at the 3′-DNA end are deaminated with at least a 10-fold lower activity compared with motifs outside the dead zone located near the 5′-DNA end (Fig. 1*A*). We envision that although the A3G oligomer can slide and jump bidirectionally along the ssDNA, multiple A3G-DNA binding interactions may wrap the DNA around A3G (Fig. 4*C*) and restrict free diffusion of the oligomer, thus reducing the possibility for resampling nondeaminated C targets. *B*, random deamination performed by a monomer. Larger oligomers are disrupted upon binding ssDNA in the absence of Mg^{2+} (Fig. 6*D*). The A3G monomers can diffuse freely along the ssDNA and thereby deaminate each CCC motif with roughly similar efficiencies outside the dead zone at the 3′-DNA end. The monomer retains partial catalytic asymmetry by deaminating with at least a 15-fold lower activity on CCC motifs within the ~30-nt dead zone at the 3′-DNA end compared with motifs outside the dead zone (Fig. 1*B*). The dead zone at the 3′-DNA end appears to provide an optimal binding region, allowing A3G to align in a catalytically favored 3′ → 5′ orientation to access the preferred C of a 5′-CC<u>C</u> motif.

reduced access of the CCC motifs to enzyme active sites. The reduced catalytic rate toward the 3′-DNA region occurs in conjunction with the formation of larger oligomers in the presence of salt (Fig. 6*C* and Table 1). We speculate that active sites situated within an oligomer, such as a tetramer, are unable to access and deaminate C targets efficiently. However, we hypothesize that A3G oligomers, including dimers, are directionally biased. If multiple binding interactions were to wrap the DNA around A3G (Fig. 4*C*) and restrict free diffusion of the oligomer, then a second level of bias is added. The intact oligomer would still be able to diffuse bidirectionally, encompassing sliding and jumping motions, but three-dimensional diffusion would become increasingly constricted as the oligomer size increased, limiting sampling of the ssDNA to find all possible C targets. We have incorporated these ideas in a schematic model (Fig. 7*A*). The elongated shape of the A3G protein contained in the sketch is based on small angle x-ray scattering of A3G in aqueous solution (26).

We further suggest that there is a polarity of the C-terminal active sites at each end of the A3G oligomer so that they do not deaminate with equal rates (Fig. 7*A*, *red circle* (high rate) and *blue circle* (low rate)). There are three observations, in addition to previous data (15), that support an asymmetric catalytic orientation of A3G so that active sites facing the 5′-DNA end (Fig. 7*A*, *red circle*) catalyze C → U deaminations much more efficiently than those facing the 3′-DNA end (Fig. 7*A*, *blue circle*). First, there is a 30-nt region that is deaminated with extremely low efficiency (deamination dead zone) located at the 3′-end of linear ssDNA (Fig. 1, *A* and *B*). The 30-nt dead zone may provide an optimal space for A3G to encounter a CCC motif in an orientation allowing catalysis. Second, directionally biased deamination occurs on a partially dsDNA circle in the presence but not absence of salt, whereas directional deamination does not occur on an ssDNA circle either with or without salt (Fig. 2), implying that A3G needs to approach a CCC motif with a specific orientation. Third, there is a biased orientation in each 5′-CCC motif, where A3G prefers to deaminate the 3′-C residue (5′-CC<u>C</u>) much more than the 5′-C residue (5′-<u>C</u>CC) even under single turnover conditions (Fig. 3*E*). Therefore, we have depicted the "hot" active site facing the 5′-DNA end and aligned closest to the favored C for deamination in a 5′-CC<u>C</u> motif (Fig. 7*A*). Conversely, the "cold" active site is facing the 3′-DNA end and aligned closest to the barely deaminated C in a 5′-CCC motif (Fig. 7*A*). Although A3G has not been previously modeled with C-terminal domains external to an oligomer (26), there is evidence that a "monomeric" C-terminal deaminase domain of A3G is catalytically active (29, 30), which suggests that *trans* subunit complementation common to the cytidine deaminase family (43) is not required. We favor this model, because it seems to be the most probable asymmetric arrangement that can account for all aspects encompassing the biochemical characteristics of A3G 3′ → 5′ deamination polarity (Fig. 7*A*).

Random deaminations take place in the absence of salt. Concomitantly, A3G oligomers that form in the absence of salt (Fig. 6*B*) favor the formation of monomers in the presence of ssDNA (Fig. 6*D* and Table 1). A3G is illustrated as a monomer in the absence of salt (Fig. 7*B*), although A3G dimers may also form (Fig. 6*D* and Table 1). It is likely that the A3G monomers are free to diffuse randomly on the ssDNA substrate and deaminate CCC motifs essentially irrespective of target location, since there is not a strong conformational change indicative of DNA wrapping, in the absence of $MgCl_2$ (Fig. 4*D*), which would otherwise limit resampling of undeaminated C targets (Fig. 7*A*).

For our study, the high salt state has provided mechanistic insights into the determinants for 3′ → 5′ deamination polarity. *In vivo*, there are likely to be multiple forms of A3G in higher and lower order oligomeric states. There is evidence that A3G exists in high molecular mass or intravirion A3G complex and low molecular mass states when associated with and when dissociated from RNA, respectively (7, 16, 26). Although A3G is incorporated into virions as an intravirion A3G complex, unable to deaminate (−)-cDNA, RNase H releases the inhibitory HIV RNA during reverse transcription to form the active low molecular mass form, characterized as monomers and/or dimers by gel filtration (16). However, the form that the low molecular mass takes when binding DNA has not been functionally characterized apart from our study, and it is likely that *in vivo*, the dimeric and larger oligomeric states formed with A3G, DNA, and salt (Fig. 6*C*) are relevant.

In vivo, the intrinsic polarity of A3G deamination makes a substantial contribution to local regions of HIV-1 mutational

polarity (17), which suggests that higher order oligomers of A3G are acting on (−)-cDNA. Apart from deamination activity, A3G may also have the capacity for blocking reverse transcription. A recent *in vitro* study has shown that A3G can block reverse transcriptase DNA elongation by competitively binding RNA or ssDNA segments of the HIV-1 genome or (−)-cDNA with a higher affinity than reverse transcriptase (8), a finding supported by A3G inhibition of HIV-1 in resting CD4+ T cells (7). Our data suggest that a potentially key aspect of the mechanism inhibiting reverse transcriptase would involve the formation of higher oligomeric states of A3G. Higher oligomeric forms of A3G can presumably interact with many regions on the DNA, which is likely to inhibit access by other enzymes, including reverse transcriptase. Here the assembly of variably sized oligomeric forms offers two distinct ways that A3G can inhibit viral propagation when binding to (−)-cDNA, a catalytic mode creating biased "hot spots" for deamination (17) to ensure mutational inactivation of HIV and a noncatalytic mode involving higher ordered structures, which can block the action of reverse transcriptase (8).

Acknowledgments—We thank Drs. Phuong Pham, Jeffrey Bertram, and Peter Calabrese for insightful criticisms and discussions concerning this study.

REFERENCES

1. Sheehy, A. M., Gaddis, N. C., Choi, J. D., and Malim, M. H. (2002) *Nature* **418,** 646–650
2. Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C., and Gao, L. (2003) *Nature* **424,** 94–98
3. Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003) *Nature* **424,** 99–103
4. Navarro, F., Bollman, B., Chen, H., Konig, R., Yu, Q., Chiles, K., and Landau, N. R. (2005) *Virology* **333,** 374–386
5. Hache, G., Liddament, M. T., and Harris, R. S. (2005) *J. Biol. Chem.* **280,** 10920–10924
6. Li, X. Y., Guo, F., Zhang, L., Kleiman, L., and Cen, S. (2007) *J. Biol. Chem.* **282,** 32065–32074
7. Chiu, Y. L., Soros, V. B., Kreisberg, J. F., Stopak, K., Yonemoto, W., and Greene, W. C. (2005) *Nature* **435,** 108–114
8. Iwatani, Y., Chan, D. S., Wang, F., Maynard, K. S., Sugiura, W., Gronenborn, A. M., Rouzina, I., Williams, M. C., Musier-Forsyth, K., and Levin, J. G. (2007) *Nucleic Acids Res.* **35,** 7096–7108
9. Luo, K., Wang, T., Liu, B., Tian, C., Xiao, Z., Kappes, J., and Yu, X. F. (2007) *J. Virol.* **81,** 7238–7248
10. Guo, F., Cen, S., Niu, M., Yang, Y., Gorelick, R. J., and Kleiman, L. (2007) *J. Virol.* **81,** 11322–11331
11. Miyagi, E., Opi, S., Takeuchi, H., Khan, M., Goila-Gaur, R., Kao, S., and Strebel, K. (2007) *J. Virol.* **81,** 13346–13353
12. Schumacher, A. J., Hache, G., Macduff, D. A., Brown, W. L., and Harris, R. S. (2008) *J. Virol.* **82,** 2652–2660
13. Mbisa, J. L., Barr, R., Thomas, J. A., Vandegraaff, N., Dorweiler, I. J., Svarovskaia, E. S., Brown, W. L., Mansky, L. M., Gorelick, R. J., Harris, R. S., Engelman, A., and Pathak, V. K. (2007) *J. Virol.* **81,** 7099–7110
14. Chiu, Y. L., Witkowska, H. E., Hall, S. C., Santiago, M., Soros, V. B., Esnault, C., Heidmann, T., and Greene, W. C. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 15588–15593
15. Chelico, L., Pham, P., Calabrese, P., and Goodman, M. F. (2006) *Nat. Struct. Mol. Biol.* **13,** 392–399
16. Soros, V. B., Yonemoto, W., and Greene, W. C. (2007) *PLoS Pathog.* **3,** e15
17. Suspene, R., Rusniok, C., Vartanian, J. P., and Wain-Hobson, S. (2006) *Nucleic Acids Res.* **34,** 4677–4684
18. Yu, Q., Konig, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J. M., and Landau, N. R. (2004) *Nat. Struct. Mol. Biol.* **11,** 435–442
19. Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981) *Biochemistry* **20,** 6929–6948
20. von Hippel, P. H., and Berg, O. G. (1989) *J. Biol. Chem.* **264,** 675–678
21. Halford, S. E., and Marko, J. F. (2004) *Nucleic Acids Res.* **32,** 3040–3052
22. Stanford, N. P., Szczelkun, M. D., Marko, J. F., and Halford, S. E. (2000) *EMBO J.* **19,** 6546–6557
23. Jack, W. E., Terry, B. J., and Modrich, P. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 4010–4014
24. Terry, B. J., Jack, W. E., and Modrich, P. (1985) *J. Biol. Chem.* **260,** 13130–13137
25. Pham, P., Chelico, L., and Goodman, M. F. (2007) *DNA Repair* (*Amst.*) **6,** 689–692
26. Wedekind, J. E., Gillilan, R., Janda, A., Krucinska, J., Salter, J. D., Bennett, R. P., Raina, J., and Smith, H. C. (2006) *J. Biol. Chem.* **281,** 38122–38126
27. Prochnow, C., Bransteitter, R., Klein, M. G., Goodman, M. F., and Chen, X. S. (2007) *Nature* **445,** 447–451
28. Opi, S., Takeuchi, H., Kao, S., Khan, M. A., Miyagi, E., Goila-Gaur, R., Iwatani, Y., Levin, J. G., and Strebel, K. (2006) *J. Virol.* **80,** 4673–4682
29. Chen, K. M., Martemyanova, N., Lu, Y., Shindo, K., Matsuo, H., and Harris, R. S. (2007) *FEBS Lett.* **581,** 4761–4766
30. Chen, K. M., Harjes, E., Gross, P. J., Fahmy, A., Lu, Y., Shindo, K., Harris, R. S., and Matsuo, H. (2008) *Nature* **452,** 116–119
31. Iwatani, Y., Takeuchi, H., Strebel, K., and Levin, J. G. (2006) *J. Virol.* **80,** 5992–6002
32. Bransteitter, R., Pham, P., Calabrese, P., and Goodman, M. F. (2004) *J. Biol. Chem.* **279,** 51612–51621
33. Ratcliff, G. C., and Erie, D. A. (2001) *J. Am. Chem. Soc.* **123,** 5632–5635
34. Yang, Y., Wang, H., and Erie, D. A. (2003) *Methods* **29,** 175–187
35. Creighton, S., Bloom, L. B., and Goodman, M. F. (1995) *Methods Enzymol.* **262,** 232–256
36. Goddard, N. L., Bonnet, G., Krichevsky, O., and Libchaber, A. (2000) *Phys. Rev. Lett.* **85,** 2400–2403
37. Bao, K. K., Wang, H., Miller, J. K., Erie, D. A., Skalka, A. M., and Wong, I. (2003) *J. Biol. Chem.* **278,** 1323–1327
38. Tessmer, I., Moore, T., Lloyd, R. G., Wilson, A., Erie, D. A., Allen, S., and Tendler, S. J. (2005) *J. Mol. Biol.* **350,** 254–262
39. Brar, S. S., Sacho, E. J., Tessmer, I., Croteau, D. L., Erie, D. A., and Diaz, M. (2008) *DNA Repair* (*Amst.*) **7,** 77–87
40. Xue, Y., Ratcliff, G. C., Wang, H., Davis-Searles, P. R., Gray, M. D., Erie, D. A., and Redinbo, M. R. (2002) *Biochemistry* **41,** 2901–2912
41. Schlacher, K., Leslie, K., Wyman, C., Woodgate, R., Cox, M. M., and Goodman, M. F. (2005) *Mol. Cell* **17,** 561–572
42. Verhoeven, E. E., Wyman, C., Moolenaar, G. F., and Goosen, N. (2002) *EMBO J.* **21,** 4196–4205
43. Xie, K., Sowden, M. P., Dance, G. S., Torelli, A. T., Smith, H. C., and Wedekind, J. E. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 8114–8119