Spring 6-5-2018

Completion of DNA Replication in *Escherichia coli*

Brian Michael Wendel
Portland State University

Completion of DNA Replication in *Escherichia coli*

by

Brian Michael Wendel

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Biology

Dissertation Committee:
Justin Courcelle, Chair
Jeffrey D. Singer
Rahul Raghavan
Michael Bartlett
Dirk Iwata-Reuyl

Portland State University
2018

Abstract

To maintain genomic integrity, all cells must accurately duplicate their genetic material in order to provide intact and complete copies to each daughter cell following cell division. Successful inheritance of chromosomal information without changing even a single nucleotide requires accurate and robust DNA replication. This requires that cells tightly control replication initiation from the origin(s), processive elongation of the replisome, and the completion of DNA replication by resolving convergent replication forks ensuring that each sequence is duplicated without alteration. Unlike initiation and elongation, the process by which replication forks converge and are resolved into two discrete, inheritable DNA molecules is not well understood. This process must be remarkably efficient, occurring thousands of times per cell division in human cells, and is likely to be a fundamental step in regulating genome stability in all cells.

In this dissertation I address how DNA replication completes in the model system *Escherichia coli.* To achieve this, I examined candidate mutants for impairments in the completion of DNA replication. By evaluating growth, viability, chromosomal copy number, and plasmid stability I identified a requirement for the proteins RecBCD, ExoI, and SbcCD in the completion reaction. SbcCD and ExoI act before RecBCD in the completion reaction and process the DNA intermediates arising as replication forks converge. These enzymes act in the completion reaction without recombination or RecA, but in the absence of the normal process recombination is required to complete DNA replication via an aberrant pathway that results in genomic instability.

Table of Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

Completion of DNA replication

Fundamental to life is the ability of all cells generate two inheritable copies of their genetic information. The accuracy of the process is essential, ensuring that the complete DNA sequence is inherited equally by each daughter cell following cell division. In humans, impairment in any stage of DNA replication can contribute to a range of varied disease states resulting from the unequal inheritance of genetic information. To ensure accurate DNA replication, cells tightly regulate all stages of replication.

Cells control the replication initiation on their chromosomes by tightly regulating origin recognition, replisome assembly and licensing enzymatically. Once replication has begun, cells dedicate even more proteins to maintain replication, ensuring it remains processive even when it encounters a damaged template. In bacteria, more than 30 proteins have been identified which function to facilitate DNA replication initiation and elongation (reviewed in (1)). In Eukaryotes, the number of protein players in these first stages is even greater (for reviews see (2, 3)). Given this extensive cellular control devoted to the early steps of DNA replication, it seems likely that cells would devote comparable enzymatic control toward ensuring that DNA replication completes accurately and efficiently.

The completion of DNA replication must be occurring at frequencies that would require tight enzymatic regulation to efficiently maintain genomic stability. In bacteria, replication forks initiate from a single bidirectional origin of replication and then progress around the circular chromosome (4, 5). In Eukaryotes, replication forks progress from multiple bidirectional origins along each linear chromosome (6). In both cases, the forks must converge and resolve at the point where all sequence has precisely doubled, without gaining or losing any sequence. This reaction is likely to be highly efficient, as it must happen hundreds of times per cell division in Eukaryotes without changing even a single nucleotide (7).

The completion reaction could be argued to be more critical to faithful genomic duplication than initiation. In Eukaryotes, only half of all replication origins initiate in any given cell cycle, yet in the absence of any particular initiation event, replication from neighboring origins can compensate by simply continuing to replicate through those regions (6, 8, 9). Both prokaryotic and eukaryotic cells tolerate variations in origin number without severe phenotypic consequences (9–12). However, a failure to accurately limit or join even a single event where two replication forks converge would be expected to result in duplications, deletions, rearrangements, or even a loss of viability.

DNA replication is essential to nearly every aspect of biology and developing an understanding of how DNA replication completes will greatly advance our understanding of how genetic information is inherited in normal cells. Perhaps even more significant, understanding the normal process of DNA replication completion could lead to the development of therapeutic treatments for the aberrancies associated with aging or in diseases as diverse as cancer. Thus, how DNA replication is completed is an incredibly

important and fundamental aspect of cellular metabolism that remains to be characterized. In this dissertation, I examine the final steps of DNA replication, identify gene products required for the process to occur, and begin to characterize the mechanism by which they act to complete DNA replication in the model organism, *Escherichia coli*.

Completing DNA replication is distinct from the segregation of chromosomes

Remarkably little is known about how DNA intermediates arising from head-on replication fork convergence are processed and accurately resolved. Many studies have shown that after DNA replication is completed however, several pathways are involved in actively partitioning two newly formed chromosomes (13–16). While important for genomic inheritance, the segregation process is temporally and spatially separate from that of DNA replication. Furthermore, several sequence elements of chromosomes have been proposed to contribute to completion, but upon closer examination appear to be at best performing an inessential supportive role.

After replication is completed, chromosomes can remain linked, or concatenated, and require resolution of the chromosome concatemers before segregation can occur. The concerted action of XerCD can resolve chromosomal concatemers or chromosome dimers by site-specific recombination at the *dif* site located in the terminus region in *E. coli* (17–19). The idea that segregation occurs independently from and after the completion of DNA replication has occurred is supported by observations that the addition of a second origin on the bacterial chromosome does not affect the site or accuracy of segregation at the endogenous *dif* site (10). Similarly, in the archaea *Sulfolobus solfataricus*, which

possesses three discrete origins, segregation occurs at a single identified *dif* site spatially far removed from any of the fork convergence zones and well after replication has completed (20, 21). Finally in Eukaryotes, the majority of replication forks initiate, elongate, and complete replication far from the centromere where the chromosomes are segregated while origins near the centromere have been shown to replicate early (22–25). These observations illustrate the temporal and, in some cases, spatial separation of chromosome concatemer resolution and segregation from the completion reaction.

Ter sequences are part of another mechanism that appears to encourage replication to finish proximal to where the chromosome segregates, but is not directly involved in the completion reaction. These sequences were identified in early studies as sequence-specific sites that caused replication forks to pause (26). In *E. coli* these sites are called termination (*ter*) sequences which encompass a ~400 kilobase region in the terminus of the chromosome. When bound by the protein Tus, these protein-DNA complexes impede the progress of replication forks in an orientation- specific manner (27–30). The global arrangement of these *ter* sequences combined with the orientation-specific character of the Tus/*ter* complex serve to create a "replication fork trap", containing these highly processive replication forks to a defined region in the terminus (*Figure* 1.2 and (31, 32)). Interestingly, cells lacking either *ter* sites or Tus replicate normally, indicating that while this mechanism may be dictating roughly where replisomes converge, it is not acting directly in the molecular pathway of completion (28, 33). Furthermore, termination systems are not found in all bacteria, and no analogous systems are present in Archaea or Eukaryotes. While these structural components of chromosomes could serve to enhance the efficiency of completion, the separation of these

processes from the head-on convergence of replication forks strongly indicates that they are not directly involved in the completion reaction.

Identifying mutants impaired for completing replication

In order to identify gene products involved in the resolution of convergent replication forks during replication completion, one could look to the strategies used to identify proteins involved in the initiation or elongation phases of DNA replication. Since DNA replication is essential for viability, isolating mutants can be challenging. Early attempts to identify genes involved in DNA replication employed screens to isolate conditional mutants that were healthy at low temperatures but displayed impaired growth and viability at high temperatures. Several of these temperature-sensitive mutants were found to impair the cell's ability synthesize DNA (34–36). In some mutants, DNA synthesis was found to arrest rapidly following a shift from the permissive temperature to the restrictive temperature, while other mutants were observed to continue replicating for a time prior to arrest. A rapid arrest after temperature shift was interpreted to suggest that the mutation involved a gene product that played a role in elongation, while a gradual arrest suggested that the mutated gene product was normally involved in initiation and could continue to replicate using already proceeding replication forks but could not initiate new ones (37, 38). Similar to mutants impaired for elongation or initiation, mutants impaired in the ability to complete replication might be expected to exhibit reduced DNA synthesis or a loss of viability.

A second strategy used to screen and isolate mutants in DNA replication was based on their inability to maintain minichromosomes, or plasmids (39–42). Plasmids are selfish genetic elements that replicate separately from the chromosome. They use much of the same cellular machinery to replicate as the chromosome, yet are often maintained at a higher copy number and therefore undergo more replication events per cell. Mutants can be scored for their efficiency at replicating plasmids more easily than the chromosome due to the plasmid being nonessential for viability and the enhanced phenotypes observed in replication-impaired mutants due to increased DNA replication events per cell duplication. Mutants impaired in either the initiation or elongation steps of replication demonstrate difficulty in the maintenance of minichromosomes, or plasmids. For mutants impaired in the ability to complete replication, a similar plasmid instability might be expected.

E. coli has been widely used a model organism. *E. coli* genes involved in DNA replication were identified by evaluating the gene products found to be mutated in the genetic screens described above. Many of these mutants were implicated in DNA synthesis, but others were associated with other aspects of DNA metabolism. Among these genes, several were associated with the repair of double-strand breaks by homologous recombination, and their mutants confer phenotypes that suggest they play a broader role in replication of the chromosome (Table 1.1).

Double-strand break repair proteins as completion candidates

Double-strand break repair in *E. coli* occurs by homologous recombination. This process involves several genes that were identified by the observation that cells deficient in those gene products were unable to generate stable recombinant genomes following conjugation (43–45). These mutants were then subsequently found to be hyper sensitive to various forms of DNA damage (46, 47). Based primarily on biochemical studies of the identified gene products, the prominent models currently in the literature for how these enzymes function *in vivo* was established and is summarized below.

RecA

RecA is essential for all homologous recombination and double-strand break repair in *E. coli* (48–50). Mutants lacking RecA activity exhibit no recombination as monitored by conjugation, transduction, or transformation, and are hyper-sensitive to X-ray and other DNA damaging agents (46, 49, 51–53). *In vitro*, RecA forms long filaments along single-stranded DNA ends, and facilitates ATP-driven strand invasion with homologous double-stranded DNA templates (Figure 1.1 and (54–56)). In current models of double-strand break repair, this RecA-mediated strand invasion is often proposed to be initiated by the RecBCD helicase/nuclease. In these models, RecBCD processes double-strand breaks by degrading the 5'-DNA strand generating 3'-ends then actively loading recA onto the resulting single-stranded tail ((57, 58) and Figure 1.1).

RecBCD

The current models of RecBCD function, originally identified as ExoV, were derived primarily from the interpretations of initial genetic screens for recombination proficiency and the subsequent biochemical activities of the purified enzyme (59–61). In the absence of RecB or RecC, recombination is reduced by two to three orders of magnitude, as measured by conjugation or transduction (48). Additionally, mutants lacking either of these gene products are hyper-sensitive to X-ray and other DNA damage (57, 62, 63). *recB, recC,* and *recD* gene products form a heterotrimeric complex that contains processive dual helicase and nuclease activities capable of degrading up to one kilobase of DNA per second (Figure 1.3a and (64–69)).

Purified RecBCD preferentially binds blunt double-strand ends, using the opposing polarity RecB and RecD helicase motors to translocate and degrade both strands of the DNA duplex (70, 71). Upon encountering a Chi site, 5'-GCTGGTGG-3', the RecD helicase activity is attenuated and the nuclease activity located on the RecB subunit is altered, primarily degrading the 5'-end leaving a recombination-capable 3'-end. The exact nature of this alteration in nuclease activity appears dependent on the Mg^{2+} concentration (Figure 1.3b and (72, 73)). Chi recognition is also thought stimulate recruitment and loading of RecA onto the newly created single-stranded 3'-end (58). Loss of all enzymatic activities results from inactivation of RecB or RecC. Inactivation of RecD results in a lack of nuclease activity and Chi recognition as well as an alteration of helicase activity (74, 75). The loss of nuclease activity and Chi recognition without the RecD gene product, despite the activities originating from the RecB and RecC subunits,

suggests that a conformational shift inactivates RecD and alters the helicase and nuclease activities upon Chi recognition (76, 77). The preference *in vitro* for a blunt double-strand end and hypersensitivity to agents that cause double-strand breaks, led several researchers to propose that RecBCD initiates recombination by processing double-strand ends and recruiting RecA to 3'-single-stranded ends useful in recombination only after degrading through a Chi site (78–80).

However, there are several phenotypes associated with *recB, -C,* and *–D* mutants that are inconsistent with this model and suggest it may have a broader, more fundamental role in DNA replication. *recBC* mutants are less sensitive to DNA damage and retain some capacity for recombination relative to *recA* mutants (81, 82). However, in the absence of DNA damage, *recBC* mutants grow poorly and cultures contain high numbers of inviable cells (44, 83). If the poor viability in *recBC* mutants were due to an inability to repair double-strand breaks, then *recA* mutants should have an even lower viability, yet *recA* mutants grow nearly as well as wild-type cells (44, 83, 84).

Additionally, and in contrast to *recA*, *recBC* and *recD mutants* exhibit plasmid instabilities (40, 85, 86). Plasmids replicating within *recD* mutants in particular have been shown to develop multimeric species (87). This multimerization is likely a cause of plasmid instability in these mutants (88). If these plasmid instabilities were due to a lack of repair of double-strand breaks arising on the plasmids then plasmids should also be more unstable in *recA* mutants, yet plasmids in *recA* mutants are hyper-stable (88, 89).

Furthermore, *in vivo* studies from several research groups done early in the characterization of RecBCD found that RecBCD acts after RecA during the recombination pathway, rather than to before it as has been proposed based on *in vitro*

biochemical activities of the enzyme (49, 90, 91). During conjugation between Hfr and F- cells in *recBC* mutants the recovery of stable recombinants is greatly reduced relative to wild type cells and abolished in cells lacking RecA (49, 82). To examine the role these enzymes played in recombination during conjugation, a study by Birge and Lowe used Hfr and F- strains containing two different mutations in *lacZ* and performed crosses between them (92). Similar to what had been previously reported, recovery of heritable recombinants was greatly reduced in *recB* and *recC* mutants relative to wild type cells. Yet during these crosses transcription and translation of a functional β-Galactosidase was transiently observed and detected at levels comparable to wildtype in those same mutants. By contrast, no transcribable intermediates of protein products were ever observed in *recA* mutants. This demonstrated that the *recA* mutant blocked recombination at a point upstream of RecBCD action, and that *recB* and *C* mutants are capable of initiating recombination and forming a joint transcribable wild-type *lacZ* sequence. The relative lack of stable recombinants from this recombination event despite the presence of intact recombinant transcript suggest that cells devoid of RecBCD are able to accurately initiate but can't complete the resulting recombination events to form discrete, heritable molecules.

RecBCD is a complex enzyme with varied and potent capabilities that may suggest a more sophisticated role than simply degrading ends for recombination. RecBCD has been shown to act on several DNA substrates, including cruciform DNA, gapped DNA, and single-strand DNA. On cruciform DNA, purified RecBCD will proceed along one of the arms ultimately nicking the DNA at the cruciform junction (Figure 1.3c and (93)). Double-strand DNA containing gaps or single-stranded regions

have also been shown to be unwound by RecBCD (94). Together these phenotypes are worth consideration when trying to modify the models for RecBCD to encompass some of the phenotypes described that appear to be inconsistent with a role in initiating double-strand break repair.

ExoI

Other enzymes have also been identified that exhibit properties suggestive of a potential role in processing replication intermediates. Exonuclease I is a powerful exonuclease, the first to be identified in *E. coli* and has since been characterized as playing a role in homologous recombination (95, 96). It rapidly degrades single-stranded DNA from 3' to 5' at a rate of up to ~10 kilobases per minute, yet dissociates upon encountering double-stranded DNA (95, 97, 98). It was independently identified as the product of the gene *sbcB*, named as a suppressor of the *recBC* phenotype (99). *sbcB* was a mutation that spontaneously arose in *recBC* mutants and was able to restore the growth, viability, resistance to DNA damage and recombination proficiency (84, 100–103). It was presumed to suppress these phenotypes in *recBC* mutants through the persistence of recombinogenic 3'-ends in the absence of ExoI. Yet it was later found that mutations in *sbcB* and *sbcC* were both required for suppression (102–104). Based on this striking phenotype, ExoI was characterized as a homologous recombination protein and models were proposed that suggested it activated an alternative, backup recombination pathway for RecBCD to initiate recombinational events (96).

Yet like RecBCD, many phenotypes observed in ExoI mutants were suggestive of a more fundamental role in the cell cycle. ExoI is 3' to 5' single-strand nuclease, an activity that is objectively anti-recombinogenic as it would remove free 3'-ends essential for strand invasion and replication. Its inability to degrade 3'-ends when double-stranded contrasts sharply with the activities of other 3' to 5' exonucleases like ExoII, the exonuclease of Polymerase I. Unlike ExoII however, ExoI is highly active on a 3'-flap displaced from duplex DNA (98). This activity would again suggest a role for ExoI in inhibiting homologous recombination. Furthermore, ExoI has a strong binding affinity for single-strand DNA binding protein (SSB), which serves to recruit ExoI to a potential substrate and stimulates enzymatic activity (105). RecA is also thought to interact with ExoI based on its co-purification over several columns, suggesting a tight *in vivo* association (106, 107). Interestingly, ExoI effectively displaces SSB polymerized on a 3'-strand to degrade it, but not RecA (106). The association of ExoI with RecA *in vitro* poses some interesting questions about its *in vivo* role that suggest it may have a broader function in genomic stability.

SbcCD

SbcCD was identified as a spontaneous mutation arising in *recBC* mutants that when combined with mutations in ExoI restores the growth, viability, resistance to DNA damage and recombination proficiency of *recBC* mutants (84, 101–103). The only major phenotype of *sbcCD* mutants is that palindromic sequences, which are unstable in wild type cells, are stabilized in the absence of SbcCD and can be propagated *in vivo* (108,

109). This is directly related to the structure-specific endonuclease activity of this protein. SbcCD has been shown to target DNA hairpins for cleavage (110, 111). It has been hypothesized that during DNA replication of palindromic sequences the regions become single-stranded and form hairpins that are incised by SbcCD generating the instability (112, 113). Yet its role in suppressing the phenotypes in recBC mutants is largely unknown. Interestingly, based on a genetic assay, SbcCD is targeted for inhibition by the λ bacteriophage protein Gam during λ phage infection, similar to RecBCD (114–116). It was hypothesized that RecBCD was targeted for inactivation to prevent degradation of the phage DNA, but why SbcCD is targeted is not well understood.

SbcC and D are widely conserved across evolutionarily-divergent organisms. The eukaryotic homologs of SbcCD have been identified and characterized as the double-strand break repair proteins Rad50 and Mre11, and mutants of these gene products are associated with genomic instability in cells and are essential for normal development (117, 118). These eukaryotic homologs for SbcCD have been identified based upon similarities in overall structure and biochemical activities, and form complexes with many other gene products in many aspects of DNA metabolism (117, 119–121). This high degree of conservation of SbcCD in a diversity of organisms, growth defects exhibited by human homologs, and the enzymes association with *recBC* mutants all suggest that SbcCD may also be associated with a more fundamental role in the cell cycle (122–125).

Conceptual methodology for evaluating completion in mutant strains

In vitro, convergent replication forks lead to over-replication of the region where forks meet. Following reconstitution of bidirectional replication on plasmid substrates, studies have shown that when replication forks converge *in vitro,* one replisome will displace the other and continue replicating using the newly synthesized daughter strand as a template (126, 127). Using a long 640bp single-stranded template, the Ullman group was able to induce template strand switching by hybridizing the template strand to an oligonucleotide of ~80 base-pairs with 40 base-pairs of homology downstream of the primer binding site. Following replication from this substrate *in vitro,* the authors were able to detect hybrid sequences that were consistent with the replication fork switching templates at the junction of the bound oligonucleotide.

Similarly, convergent replication forks may over-replicate at sites where DNA replication completes. As the proceeding replication forks meet, one replisome would displace the other proceeding with replication on to the newly replicated daughter strand generating an over-replicated intermediate. This would have the inherent advantage of multiple copies of DNA in the terminus region, reducing the risk of losing genetic information in the case of aberrant segregation.

An alternative possibility is that during completion replication forks stop short of reaching their chromosomal doubling point and the remaining gaps could be filled-in and ligated. In either case, mutants deficient in completing DNA replication would exhibit detectable abnormal DNA copy number at the point where replication forks converge. Differentiating between these potential models is one of the primary goals of this work.

Bacteria make an ideal model organism in which to perform this analysis characterizing the completion reaction due to bacterial chromosome structure. *E. coli* contains a circular chromosome with a single bidirectional origin (4, 5). This means that the completion reaction can be localized to a region directly opposite the origin of chromosomal replication. Furthermore, the replication fork trap which is not directly involved in the completion reaction, serves to contain converging replication forks to a specific area, further defining the location of the reaction (128).

Some evidence suggest that over-replication may be occurring at sites where completion occurs. In early studies, "dormant replication origins" *oriX*, *oriK*, and *oriM* were identified in *E. coli* that could be observed under conditions of cellular stress, and mapped to the terminus region of the chromosome (129, 130). Considering that these "dormant origins" were detected based on elevated copy numbers in the region, an alternative interpretation could be that these areas of elevated copy number resulted from replication continuing beyond the doubling point which would produce elevated copies of DNA in this region. While these over-replicated regions in *E. coli* have been well-documented, a similar phenomenon has been observed in Eukaryotes, making it tempting to speculate that a similar process may be occurring (24, 131). These observations taken together with the numerous studies identifying gene products with replication-defect phenotypes in *E. coli* make it an excellent candidate for the study of completion.

In this dissertation, I examine candidate mutants exhibiting phenotypes associated with poor growth and/or plasmid instability for their ability to complete replication on the chromosome using genomic profiling. For this work I employ this strategy to determine if mutants are impaired in their ability to complete replication based upon the observation

of abnormal copy number of sequences in the terminus region, relative to wild type cells. To this end, total genomic DNA from replicating cultures of wild type and mutant strains is purified and analyzed by high-throughput sequencing technology to generate replication profiles. The replication profile of the cultures represents the replication-specific copy number of each sequence around the genome.

In an exponentially growing wild type culture, all cells are replicating. As shown in *figure 1.4a*, in a randomly replicating population, most cells contain at least two copies of the origin (which replicates first) and one copy of the terminus (which replicates last). The sequences between these locations have proportionally fewer copies as the genetic distance from the origin increases. Thus, by randomly sequencing fragments of the chromosome and then plotting their frequency as a histogram, a replication profile similar to that shown in *figure 1.4b* can be obtained.

In cultures of cells with an impaired ability to complete replication, abnormalities in the copy number can be detected in the terminus region. If replication in a mutant fails to reach the terminus or fails to join the DNA ends at the end of the replication cycle, those ends may be missing or subject to degradation. This would result in a replication profile that lacks, or has reduced amount of DNA in the terminus region. Alternatively, if a mutant fails to complete replication at the doubling point and continues beyond where the replication forks meet, a third copy of the sequence would be generated and would be observed as an increased amount of DNA within the terminus region of the chromosome.

In the following chapter, I identify gene products involved in the completion reaction by establishing candidate gene products whose mutants exhibit replication-deficient phenotypes (Table 1.1). I evaluate those candidate mutants by examining their

growth, plasmid stability, and replication-specific DNA copy number at the point where replication forks converge. This work demonstrates that RecBCD is required for the completion reaction in a manner distinct from its role in homologous recombination, as the *E. coli* recombinase RecA is not required for the reaction.

Throughout the third chapter of this dissertation, I investigate the roles that the bacterial orthologs of Mre11-Rad50, SbcCD and ExoI play in the completion reaction. The data presented in this chapter show that SbcCD and ExoI are required for the completion reaction and act before RecBCD in the completion pathway. I further demonstrate that completion can occur by the normal pathway involving RecBCD, SbcCD, and ExoI and that in the absence of these gene products completion becomes dependent on homologous recombination. This aberrant recombinationally-mediated pathway is associated with genomic rearrangements and instabilities.

In the fourth chapter of this dissertation, I use bidirectionally-replicating plasmids, or plasmids containing two replisomes, to better understand the process of completion. By evaluating how two-replisome plasmids are transformed and maintained overtime in mutant strains deficient in completing DNA replication, I further the understanding of how these proteins function during completion. The data presented here show that RecBCD is required to process convergent replication forks and that homologous recombination cannot rescue completion without RecBCD. Furthermore, even in wild type cells the aberrant completion pathway, decreases the efficiency of the completion reaction and without recombination the efficiency of the reaction is enhanced.

The results contained within these chapters define a novel and fundamental process for resolving convergent replication forks during the completion of DNA

replication. This process involves gene products originally characterized as double-strand break repair proteins, demonstrating novel roles for these proteins during completion. These protein players act to complete DNA replication in a manner that does not require recombination and distinct from their roles in double-strand break repair by homologous recombination. Yet in the absence of the normal completion process, an aberrant recombinationally-mediated pathway becomes required to rescue the convergence of replication forks.

Table 1.1 Candidate genes to be examined, functions and known phenotypes

Genes	Known Functions	Growth Abnormalities	Plasmid Instability	DNA Damage Sensitivity
recBC	Forms a heterotrimer with potent helicase/nuclease activity,	yes[2]	yes[4]	yes[3]
recD	nuclease activity altered by Chi-site recognition[1]	yes[2]	yes[4]	no[10]
sbcCD	Structure-specific endonuclease, targets DNA hairpins[5], some exonuclease activity	yes[5]	yes[5]	yes[6]
sbcB (xonA)	3' to 5' single-stranded exonuclease[8]	yes[7]	yes[7]	yes[6,7]
recJ	5' to 3' single-stranded exonuclease[9]	yes[11]	-	yes[6]
recG	Junction-specific DNA helicase, branch migration[13]	no[13]	-	yes[13]
xth	3' to 5' double-stranded exonuclease, Mg^{2+} dependent[8]	-	-	yes[12]
recN	SOS-induced, involved in transformation, involved in recombination in *recBCsbcBC* mutants[14]	no[14]	-	yes[14]
recQ	Helicase that loads onto single- or double-stranded DNA ends with low processivity[15]	-	yes[16]	yes[17]

1-(66, 71, 72)**, 2**-(44)**, 3**-(48)**, 4**-(85, 88, 132)**, 5**-(101, 111, 118)**, 6**-(133)**, 7**-(102, 103, 103)**, 8**-(134)**, 9**-(135)**, 10**-(57, 75)**, 11**-(136)**, 12**-(137)**, 13**-(138, 139)**, 14**-(46, 51, 61, 140) **15**-(141) **16**-(103, 142) **17**-(143)

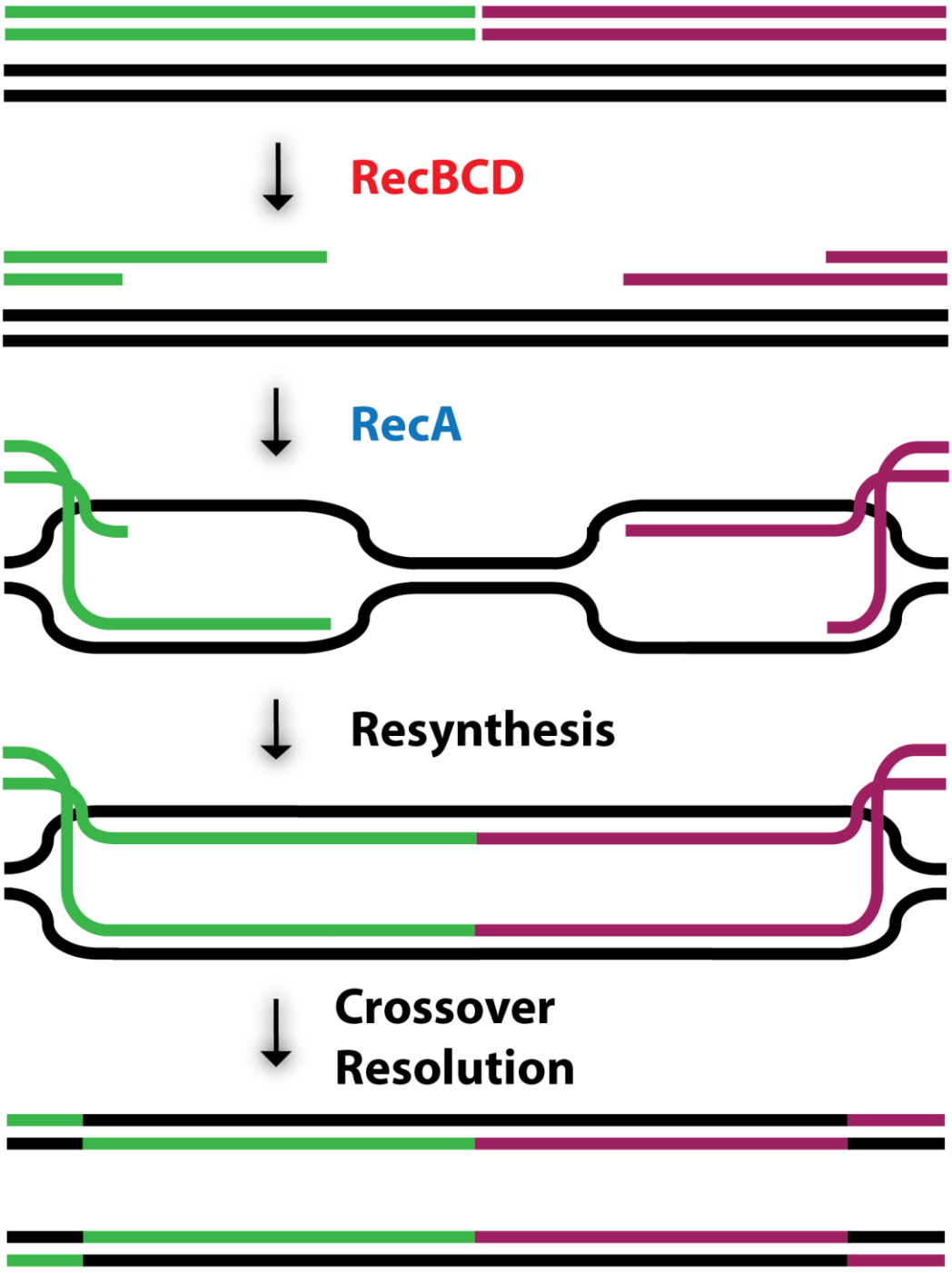
RecBCD
RecA
Resynthesis
Crossover
Resolution

Figure 1.1. Proposed model for double-strand (ds) break repair via homologous recombination in *E. coli*. RecBCD binds a (ds) DNA end and unwinds the duplex while digesting the 5'-strand resulting in a single-stranded (ss) 3'-tail. As RecBCD translocates along the DNA it is simultaneously loading the protein RecA. RecA protects the (ss) DNA end and facilitates strand invasion. Resynthesis of the gaps results in holiday junctions which are subsequently resolved. Adapted from (144).

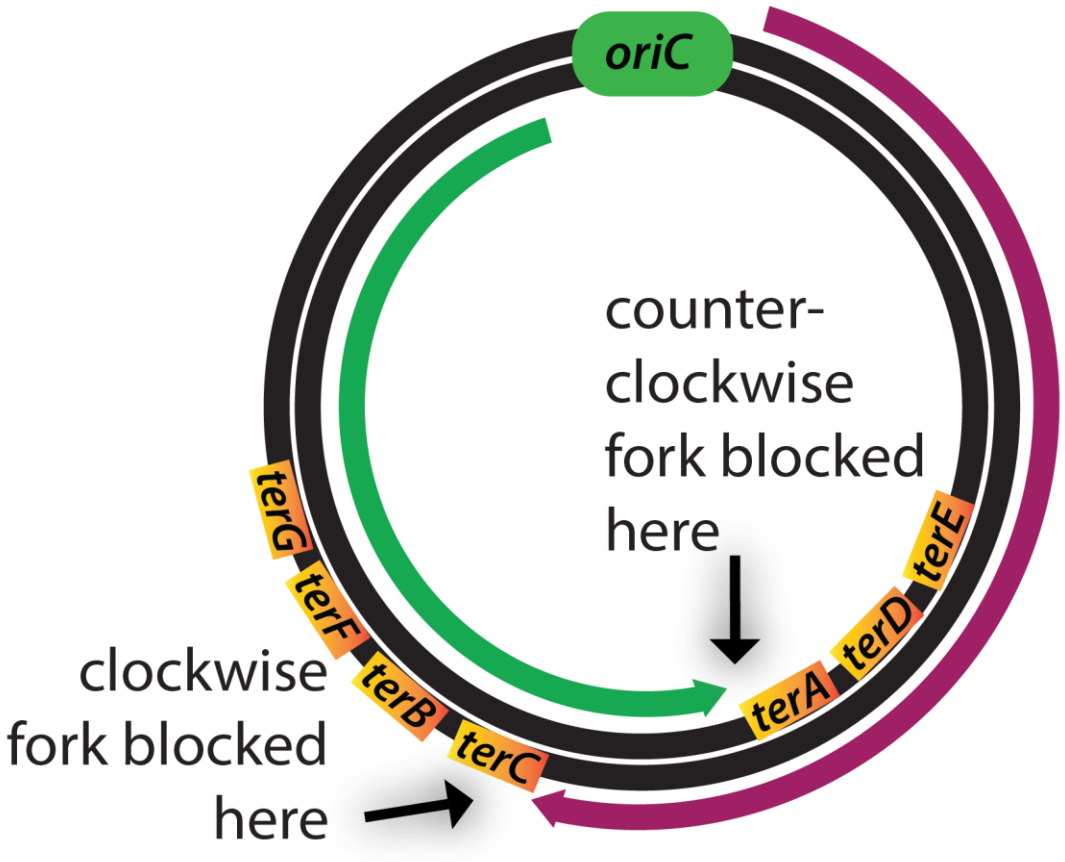

Figure 1.2. *E. coli* is an excellent model system for the study of completion. This *E. coli* chromosomal diagram illustrates the location and orientation of the replication forks emanating bidirectionally from *oriC* as the polar replication pause sites, *ter* sites bound by the protein Tus, serve to contain the convergent replication forks to the terminus region. Adapted from (144).

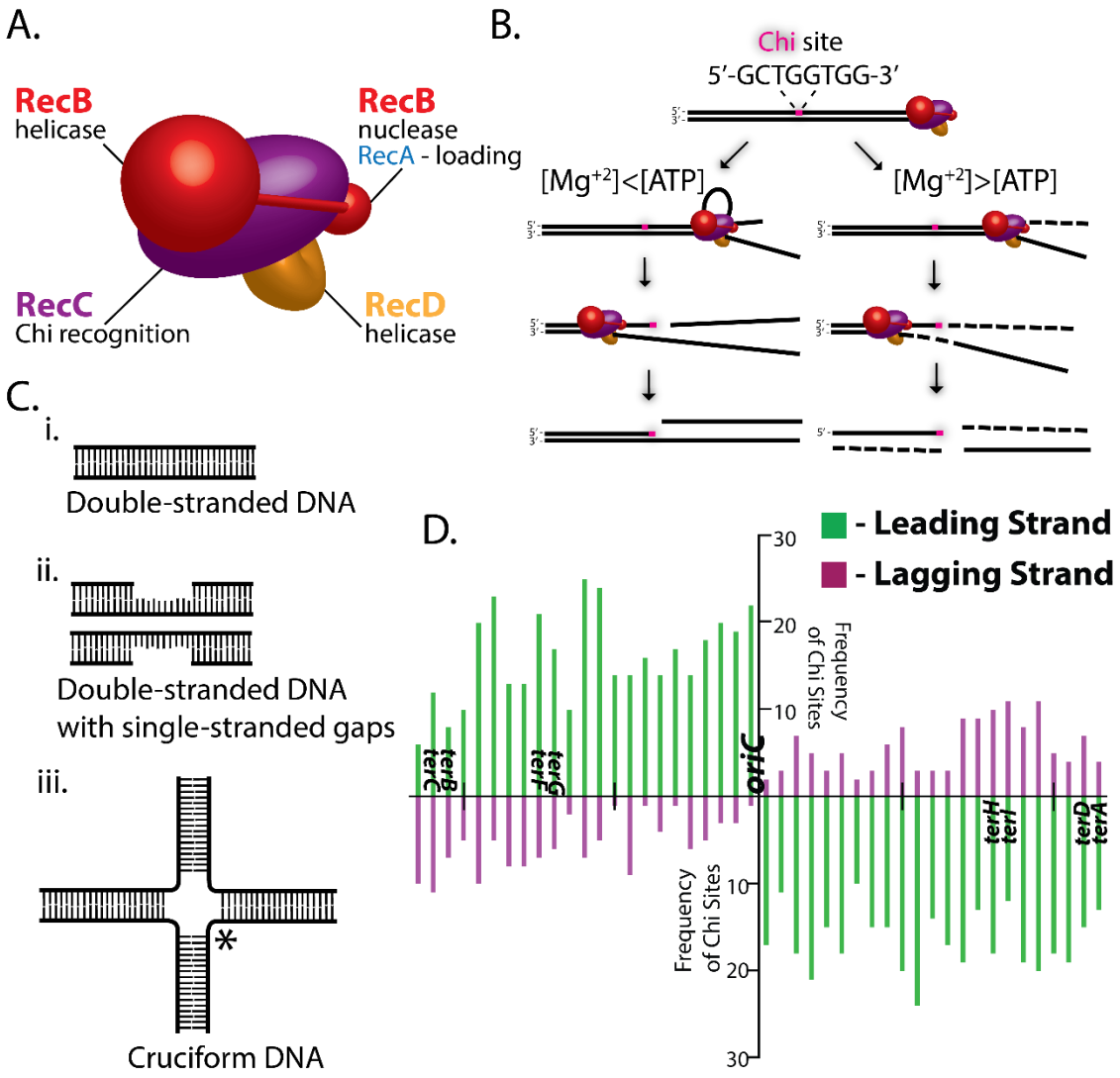

Figure 1.3. RecBCD is complex enzyme A. RecBCD is heterotrimer with potent helicase and nuclease activities as well as an inducible RecA loading activity. Helicase activity is specific to the RecB and RecD subunits, while RecB contains nuclease activity and constitutive RecA loading. **B.** Upon encountering a Chi site, 5'-GCTGGTGG-3', RecBCD attenuates is helicase activity and alters cleavage pattern. This alteration depends on concentrations of Mg^{2+} and ATP *in vitro*, and the *in vivo* concentrations have been difficult to measure. **C.** RecBCD has been shown to act on a variety of DNA structures. It is known to unwind and degrade double-stranded DNA, unwind double-stranded DNA with single-stranded gaps, and to unwind an arm of a cruciform molecule and cleave at the branch point (*). **D.** Chi appears to be oriented with chromosomal replication as the sequences are enriched in the leading strand emanating from the origin and that enrichment flips in the terminus region where replisomes converge. Adapted from (144).

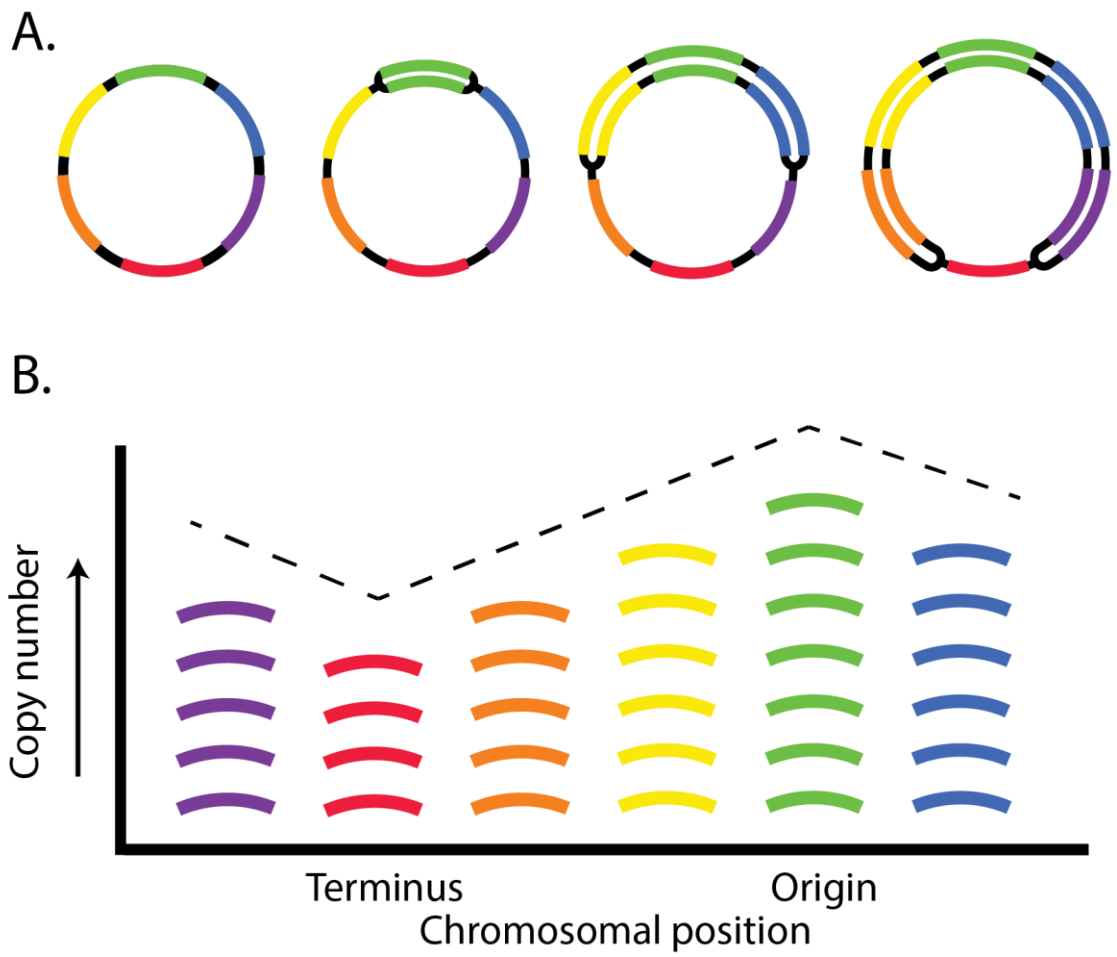

Figure 1.4. Conceptual diagram of a replication profiling. A. A population of replicating cells will have more copies of the origin region ( ) than the terminus region ( ). **B.** By fractionating the genomes, sequencing the population's fragments, mapping where those sequence match the chromosomal sequence, and comparing that to a population in stationary phase, a replication profile can be generated. In mutants impaired in completion, abnormalities would be expected in the terminus region, where replication forks are converging. Adapted from (145).

References

1. Marians KJ (1992) Prokaryotic DNA Replication. *Annual Review of Biochemistry* 61(1):673–715.

2. Coverley D, Laskey RA (1994) Regulation of Eukaryotic DNA Replication. *Annual Review of Biochemistry* 63(1):745–776.

3. Bell SP, Dutta A (2002) Dna Replication in Eukaryotic Cells. *Annual Review of Biochemistry* 71(1):333–374.

4. Masters M, Broda P (1971) Evidence for the Bidirectional Replication of the *Escherichia coli* Chromosome. *Nature* 232(31):137–140.

5. McKenna WG, Masters M (1972) Biochemical Evidence for the Bidirectional Replication of DNA in *Escherichia coli*. *Nature* 240:536–539.

6. Beach D, Piper M, Shall S (1980) Isolation of chromosomal origins of replication in yeast. *Nature* 284(5752):185–187.

7. Heichinger C, Penkett CJ, Bähler J, Nurse P (2006) Genome-wide characterization of fission yeast DNA replication origins. *The EMBO Journal* 25(21):5171–5179.

8. Wu P-YJ, Nurse P (2009) Establishing the Program of Origin Firing during S Phase in Fission Yeast. *Cell* 136(5):852–864.

9. Gros J, Devbhandari S, Remus D (2014) Origin plasticity during budding yeast DNA replication in vitro. *EMBO J* 33(6):621–636.

10. Wang X, Lesterlin C, Reyes-Lamothe R, Ball G, Sherratt DJ (2011) PNAS Plus: Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proceedings of the National Academy of Sciences* 108(26):E243–E250.

11. Hawkins M, Malla S, Blythe MJ, Nieduszynski CA, Allers T (2013) Accelerated growth in the absence of DNA replication origins. *Nature* 503(7477):544–547.

12. Ge XQ, Jackson DA, Blow JJ (2007) Dormant origins licensed by excess Mcm2–7 are required for human cells to survive replicative stress. *Genes Dev* 21(24):3331–3341.

13. Gordon GS, et al. (1997) Chromosome and Low Copy Plasmid Segregation in E. coli: Visual Evidence for Distinct Mechanisms. *Cell* 90(6):1113–1121.

14. Yamaichi Y, Niki H (2000) Active segregation by the Bacillus subtilis partitioning system in *Escherichia coli*. *PNAS* 97(26):14656–14661.

15. Kadoya R, Hassan AKM, Kasahara Y, Ogasawara N, Moriya S (2002) Two separate

DNA sequences within oriC participate in accurate chromosome segregation in Bacillus subtilis. *Molecular Microbiology* 45(1):73–87.

16. Huo Y-J, et al. (2015) MreBCD-associated Cytoskeleton is Required for Proper Segregation of the Chromosomal Terminus during the Division Cycle of *Escherichia Coli*. *Chin Med J* 128(9):1209–1214.

17. Colloms SD, Sykora P, Szatmari G, Sherratt DJ (1990) Recombination at ColE1 cer requires the *Escherichia coli* xerC gene product, a member of the lambda integrase family of site-specific recombinases. *Journal of bacteriology* 172(12):6973–6980.

18. Blakely G, et al. (1993) Two related recombinases are required for site-specific recombination at dif and cer in E. coli K12. *Cell* 75(2):351–361.

19. Ip SCY, Bregu M, Barre F-X, Sherratt DJ (2003) Decatenation of DNA circles by FtsK-dependent Xer site-specific recombination. *EMBO J* 22(23):6399–6407.

20. Duggin IG, Dubarry N, Bell SD (2011) Replication termination and chromosome dimer resolution in the archaeon Sulfolobus solfataricus. *EMBO J* 30(1):145–153.

21. Popławski A, Bernander R (1997) Nucleoid structure and distribution in thermophilic Archaea. *J Bacteriol* 179(24):7625–7630.

22. Knott SRV, et al. (2012) Forkhead Transcription Factors Establish Origin Timing and Long-Range Clustering in S. cerevisiae. *Cell* 148(1–2):99–111.

23. Luo H, Li J, Eshaghi M, Liu J, Karuturi RK (2010) Genome-wide estimation of firing efficiencies of origins of DNA replication from time-course copy number variation data. *BMC bioinformatics* 11(1):247.

24. Xu J, et al. (2012) Genome-wide identification and characterization of replication origins by deep sequencing. *Genome Biology* 13(4):R27.

25. Pohl TJ, Brewer BJ, Raghuraman MK (2012) Functional Centromeres Determine the Activation Time of Pericentric Origins of DNA Replication in Saccharomyces cerevisiae. *PLoS Genet* 8(5):e1002677.

26. LaDuca RJ, Fay PJ, Chuang C, McHenry CS, Bambara RA (1983) Site-specific pausing of deoxyribonucleic acid synthesis catalyzed by four forms of *Escherichia coli* DNA polymerase III. *Biochemistry* 22(22):5177–5188.

27. Lee EH, Kornberg A, Hidaka M, Kobayashi T, Horiuchi T (1989) *Escherichia coli* replication termination protein impedes the action of helicases. *PNAS* 86(23):9104–9108.

28. Hill TM, Tecklenburg ML, Pelletier AJ, Kuempel PL (1989) tus, the trans-acting gene required for termination of DNA replication in *Escherichia coli*, encodes a

DNA-binding protein. *Proceedings of the National Academy of Sciences* 86(5):1593–1597.

29. Neylon C, et al. (2000) Interaction of the *Escherichia coli* Replication Terminator Protein (Tus) with DNA: A Model Derived from DNA-Binding Studies of Mutant Proteins by Surface Plasmon Resonance†. *Biochemistry* 39(39):11989–11999.

30. Mulcair MD, et al. (2006) A Molecular Mousetrap Determines Polarity of Termination of DNA Replication in E. coli. *Cell* 125(7):1309–1319.

31. Hill TM, Henson JM, Kuempel PL (1987) The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc Natl Acad Sci U S A* 84(7):1754–1758.

32. Sharma B, Hill TM (1995) Insertion of inverted Ter sites into the terminus region of the *Escherichia coli* chromosome delays completion of DNA replication and disrupts the cell cycle. *Molecular Microbiology* 18(1):45–61.

33. Iismaa TP, Wake RG (1987) The normal replication terminus of the Bacillus subtilis chromosome, terC, is dispensable for vegetative growth and sporulation. *J Mol Biol* 195(2):299–310.

34. Carl PL (1970) *Escherichia coli* mutants with temperature-sensitive synthesis of DNA. *Molec Gen Genet* 109(2):107–122.

35. Wechsler JA, Gross JD (1971) *Escherichia coli* mutants temperature-sensitive for DNA synthesis. *Molec Gen Genetics* 113(3):273–284.

36. Hirota Y, Ryter A, Jacob F (1968) Thermosensitive Mutants of E. coli Affected in the Processes of DNA Synthesis and Cellular Division. *Cold Spring Harb Symp Quant Biol* 33:677–693.

37. Beyersmann D, Messer W, Schlicht M (1974) Mutants of *Escherichia coli* B/r Defective in Deoxyribonucleic Acid Initiation: dnaI, a New Gene for Replication. *J Bacteriol* 118(3):783–789.

38. Filip CC, Allen JS, Gustafson RA, Allen RG, Walker JR (1974) Bacterial Cell Division Regulation: Characterization of the dnaH Locus of *Escherichia coli*. *J Bacteriol* 119(2):443–449.

39. Kingsbury DT, Helinski DR (1973) Temperature-Sensitive Mutants for the Replication of Plasmids in *Escherichia coli*: Requirement for Deoxyribonucleic Acid Polymerase I in the Replication of the Plasmid ColE1. *J Bacteriol* 114(3):1116–1124.

40. Niki H, et al. (1988) Chromosomal genes essential for stable maintenance of the mini-F plasmid in *Escherichia coli*. *J Bacteriol* 170(11):5272–5278.

41. Bassett CL, Kushner SR (1984) Exonucleases I, III, and V are required for stability of ColE1-related plasmids in *Escherichia coli*. *Journal of bacteriology* 157(2):661–664.

42. Maine GT, Sinha P, Tye B-K (1984) Mutants of S. CEREVISIAE Defective in the Maintenance of Minichromosomes. *Genetics* 106(3):365–385.

43. Willetts NS, Clark AJ, Low B (1969) Genetic Location of Certain Mutations Conferring Recombination Deficiency in *Escherichia coli*. *J Bacteriol* 97(1):244–249.

44. Capaldo-Kimball F, Barbour SD (1971) Involvement of recombination genes in growth and viability of *Escherichia coli* K-12. *Journal of bacteriology* 106(1):204–212.

45. Barbour SD, Clark AJ (1970) Biochemical and genetic studies of recombination proficiency in *Escherichia coli*, I. Enzymatic activity associated with recB+ and recC+ genes. *Proceedings of the National Academy of Sciences* 65(4):955–961.

46. Sargentini NJ, Smith KC (1986) Quantitation of the Involvement of the recA, recB, recC, recF, recJ, recN, lexA, radA, radB, uvrD, and umuC Genes in the Repair of X-Ray-Induced DNA Double-Strand Breaks in *Escherichia coli*. *Radiation Research* 107(1):58–72.

47. Youngs DA, Bernstein IA (1973) Involvement of the recB-recC Nuclease (Exonuclease V) in the Process of X-Ray-Induced Deoxyribonucleic Acid Degradation in Radiosensitive Strains of *Escherichia coli* K-12. *J Bacteriol* 113(2):901–906.

48. Clark AJ, Margulies AD (1965) Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12. *Proceedings of the National Academy of Sciences of the United States of America* 53(2):451.

49. Wilkins BM (1969) Chromosome transfer from F-lac+ strains of *Escherichia coli* K-12 mutant at recA, recB, or recC. *Journal of bacteriology* 98(2):599–604.

50. Willetts NS, Mount DW (1969) Genetic Analysis of Recombination-Deficient Mutants of *Escherichia coli* K-12 Carrying rec Mutations Cotransducible with thyA. *J Bacteriol* 100(2):923–934.

51. Sargentini NJ, Smith KC (1983) Characterization of an *Escherichia coli* Mutant (radB101) Sensitive to γ and uv Radiation, and Methyl Methanesulfonate. *Radiation Research* 93(3):461–478.

52. Witkin EM (1976) Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol Rev* 40(4):869–907.

53. Kuzminov A (1999) Recombinational Repair of DNA Damage in *Escherichia coli* and Bacteriophage λ. *Microbiol Mol Biol Rev* 63(4):751–813.

54. Pugh BF, Cox MM (1988) General mechanism for RecA protein binding to duplex DNA. *Journal of Molecular Biology* 203(2):479–493.

55. Chow SA, Chiu SK, Wong BC (1992) RecA protein-promoted homologous pairing and strand exchange between intact and partially single-stranded duplex DNA. *J Mol Biol* 223(1):79–93.

56. van der Heijden T, et al. (2008) Homologous Recombination in Real Time: DNA Strand Exchange by RecA. *Molecular Cell* 30(4):530–538.

57. Chaudhury AM, Smith GR (1984) A new class of *Escherichia coli recBC* mutants: implications for the role of RecBC enzyme in homologous recombination. *Proc Natl Acad Sci U S A* 81(24):7850–7854.

58. Anderson DG, Kowalczykowski SC (1997) The Translocating RecBCD Enzyme Stimulates Recombination by Directing RecA Protein onto ssDNA in a χ-Regulated Manner. *Cell* 90(1):77–86.

59. Oishi M (1969) An ATP-dependent deoxyribonuclease from *Escherichia coli* with a possible role in genetic recombination. *PNAS* 64(4):1292–1299.

60. Lloyd RG, Thomas A (1984) A molecular model for conjugational recombination in *Escherichia coli* K12. *Mol Gen Genet* 197(2):328–336.

61. Picksley SM, Attfield PV, Lloyd RG (1984) Repair of DNA double-strand breaks in *Escherichia coli* K12 requires a functional recN product. *Mol Gen Genet* 195(1–2):267–274.

62. Clark AJ (1973) Recombination Deficient Mutants of E. Coli and Other Bacteria. *Annual Review of Genetics* 7(1):67–86.

63. Kushner SR (1974) In Vivo Studies of Temperature-Sensitive recB and recC Mutants. *J Bacteriol* 120(3):1213–1218.

64. Amundsen SK, Neiman AM, Thibodeaux SM, Smith GR (1990) Genetic dissection of the biochemical activities of RecBCD enzyme. *Genetics* 126(1):25–40.

65. Taylor AF, Smith GR (1985) Substrate specificity of the DNA unwinding activity of the RecBC enzyme of *Escherichia coli*. *Journal of Molecular Biology* 185(2):431–443.

66. Roman LJ, Kowalczykowski SC (1989) Characterization of the helicase activity of the *Escherichia coli* RecBCD enzyme using a novel helicase assay. *Biochemistry* 28(7):2863–2873.

67. Roman LJ, Eggleston AK, Kowalczykowski SC (1992) Processivity of the DNA helicase activity of *Escherichia coli* recBCD enzyme. *J Biol Chem* 267(6):4207–4214.

68. Bianco PR, Kowalczykowski SC (2000) Translocation step size and mechanism of the RecBC DNA helicase. *Nature* 405(6784):368–372.

69. Bianco PR, et al. (2001) Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* 409(6818):374–378.

70. Taylor AF, Smith GR (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* 423(6942):889–893.

71. Dillingham MS, Spies M, Kowalczykowski SC (2003) RecBCD enzyme is a bipolar DNA helicase. *Nature* 423(6942):893–897.

72. Taylor AF, Smith GR (1992) RecBCD enzyme is altered upon cutting DNA at a chi recombination hotspot. *Proceedings of the National Academy of Sciences* 89(12):5226–5230.

73. Zaman MM, Boles TC (1994) Chi-dependent formation of linear plasmid DNA in exonuclease-deficient recBCD+ strains of *Escherichia coli*. *Journal of bacteriology* 176(16):5093–5100.

74. Howard-Flanders P, Theriot L (1966) Mutants of *Escherichia coli* K-12 defective in DNA repair and in genetic recombination. *Genetics* 53(6):1137–1150.

75. Amundsen SK, Taylor AF, Chaudhury AM, Smith GR (1986) recD: the gene for an essential third subunit of exonuclease V. *Proceedings of the National Academy of Sciences* 83(15):5558–5562.

76. Myers RS, Kuzminov A, Stahl FW (1995) The recombination hot spot chi activates RecBCD recombination by converting *Escherichia coli* to a recD mutant phenocopy. *PNAS* 92(14):6244–6248.

77. Dixon DA, Churchill JJ, Kowalczykowski SC (1994) Reversible inactivation of the *Escherichia coli* RecBCD enzyme by the recombination hotspot chi in vitro: evidence for functional inactivation or loss of the RecD subunit. *PNAS* 91(8):2980–2984.

78. Roman LJ, Dixon DA, Kowalczykowski SC (1991) RecBCD-dependent joint molecule formation promoted by the *Escherichia coli* RecA and SSB proteins. *PNAS* 88(8):3367–3371.

79. Anderson DG, Churchill JJ, Kowalczykowski SC (1999) A Single Mutation, RecBD1080A, Eliminates RecA Protein Loading but Not Chi Recognition by RecBCD Enzyme. *J Biol Chem* 274(38):27139–27144.

80. Churchill JJ, Kowalczykowski SC (2000) Identification of the RecA protein-loading domain of RecBCD enzyme1. *Journal of Molecular Biology* 297(3):537–542.

81. Lloyd RG, Evans NP, Buckman C (1987) Formation of recombinant lacZ + DNA in conjugational crosses with a recB mutant of *Escherichia coli* K12 depends on recF, recJ, and recO. *Mol Gen Genet* 209(1):135–141.

82. Lloyd RG, Buckman C, Benson FE (1987) Genetic Analysis of Conjugational Recombination in *Escherichia coli* K12 Strains Deficient in RecBCD Enzyme. *J Gen Microbiol* 133(9):2531–2538.

83. Capaldo FN, Barbour SD (1975) The Role of the rec Genes in the Viability of *Escherichia coli* K12. *Molecular Mechanisms for Repair of DNA*, Basic Life Sciences., eds Hanawalt PC, Setlow RB (Springer US), pp 405–418.

84. Templin A, Kushner SR, Clark AJ (1972) Genetic analysis of mutations indirectly suppressing *recB* and *recC* mutations. *Genetics* 72(2):205–215.

85. Biek DP, Cohen SN (1986) Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in *Escherichia coli*. *Journal of bacteriology* 167(2):594–603.

86. Seelke R, Kline B, Aleff R, Porter RD, Shields MS (1987) Mutations in the recD gene of *Escherichia coli* that raise the copy number of certain plasmids. *Journal of bacteriology* 169(10):4841–4844.

87. Silberstein Z, Cohen A (1987) Synthesis of linear multimers of OriC and pBR322 derivatives in *Escherichia coli* K-12: role of recombination and replication functions. *Journal of bacteriology* 169(7):3131–3137.

88. Summers DK, Sherratt DJ (1984) Multimerization of high copy number plasmids causes instability: Cole 1 encodes a determinant essential for plasmid monomerization and stability. *Cell* 36(4):1097–1103.

89. Benito A, Vidal M, Villaverde A (1993) Enhanced production of pL-controlled recombinant proteins and plasmid stability in *Escherichia coli* RecA+ strains. *Journal of Biotechnology* 29(3):299–306.

90. Willetts NS, Clark AJ (1969) Characteristics of Some Multiply Recombination-Deficient Strains of *Escherichia coli*. *J Bacteriol* 100(1):231–239.

91. Hall JD, Howard-Flanders P (1972) Recombinant F′ factors from *Escherichia coli* K-12 strains carrying recB or recC. *Journal of bacteriology* 110(2):578–584.

92. Birge EA, Low KB (1974) Detection of transcribable recombination products following conjugation in Rec+, RecB− and RecC− strains of *Escherichia coli* K12. *Journal of Molecular Biology* 83(4):447–457.

93. Taylor AF, Smith GR (1990) Action of RecBCD enzyme on cruciform DNA. *Journal of Molecular Biology* 211(1):117–134.

94. Xie F, Wu CG, Weiland E, Lohman TM (2013) Asymmetric Regulation of Bipolar Single-stranded DNA Translocation by the Two Motors within *Escherichia coli* RecBCD Helicase. *J Biol Chem* 288(2):1055–1064.

95. Lehman IR, Nussbaum AL (1964) The Deoxyribonucleases of *Escherichia coli* V. On the specificity of Exonuclease I (phosphodiesterase). *J Biol Chem* 239(8):2628–2636.

96. Kushner SR, Nagaishi H, Templin A, Clark AJ (1971) Genetic Recombination in *Escherichia coli*: The Role of Exonuclease I. *PNAS* 68(4):824–827.

97. Richardson CC, Lehman IR, Kornberg A (1964) A Deoxyribonucleic Acid Phosphatase-Exonuclease from *Escherichia coli* II. Characterization of the exonuclease activity. *J Biol Chem* 239(1):251–258.

98. Brutlag D, Kornberg A (1972) Enzymatic synthesis of deoxyribonucleic acid. 36. A proofreading function for the 3’ leads to 5’ exonuclease activity in deoxyribonucleic acid polymerases. *J Biol Chem* 247(1):241–248.

99. Wang TC, Smith KC (1985) Mechanism of sbcB-suppression of the recBC-deficiency in postreplication repair in UV-irradiated *Escherichia coli* K-12. *Mol Gen Genet* 201(2):186–191.

100. Kushner SR, Nagaishi H, Clark AJ (1972) Indirect suppression of recB and recC mutations by exonuclease I deficiency. *Proceedings of the National Academy of Sciences* 69(6):1366–1370.

101. Lloyd RG, Buckman C (1985) Identification and genetic analysis of sbcC mutations in commonly used recBC sbcB strains of *Escherichia coli* K-12. *Journal of bacteriology* 164(2):836–844.

102. Phillips GJ, Prasher DC, Kushner SR (1988) Physical and biochemical characterization of cloned sbcB and xonA mutations from *Escherichia coli* K-12. *Journal of bacteriology* 170(5):2089–2094.

103. Luisi-DeLuca C, Lovett ST, Kolodner RD (1989) Genetic and physical analysis of plasmid recombination in recB recC sbcB and recB recC sbcA *Escherichia coli* K-12 mutants. *Genetics* 122(2):269–278.

104. Berger I, Cohen A (1989) Suppression of recA deficiency in plasmid recombination by bacteriophage lambda beta protein in RecBCD- ExoI- *Escherichia coli* cells. *J Bacteriol* 171(6):3523–3529.

105. Lu D, Keck JL (2008) Structural basis of *Escherichia coli* single-stranded DNA-

binding protein stimulation of exonuclease I. *PNAS* 105(27):9169–9174.

106. Bedale WA, Inman RB, Cox MM (1993) A reverse DNA strand exchange mediated by recA protein and exonuclease I. The generation of apparent DNA strand breaks by recA protein is explained. *Journal of Biological Chemistry* 268(20):15004–15016.

107. Butland G, et al. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 433(7025):531–537.

108. Chalker AF, Leach DR, Lloyd RG (1988) *Escherichia coli sbcC* mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71(1):201–205.

109. Bzymek M, Lovett ST (2001) Evidence for two mechanisms of palindrome-stimulated deletion in *Escherichia coli*: single-strand annealing and replication slipped mispairing. *Genetics* 158(2):527–540.

110. Naom IS, Morton SJ, Leach DR, Lloyd RG (1989) Molecular organization of sbcC, a gene that affects genetic recombination and the viability of DNA palindromes in *Escherichia coli* K-12. *Nucleic Acids Res* 17(20):8033–8045.

111. Connelly JC, de Leau ES, Leach DR (1999) DNA cleavage and degradation by the SbcCD protein complex from *Escherichia coli*. *Nucleic acids research* 27(4):1039–1046.

112. Pan X, Leach DRF (2000) The roles of mutS, sbcCD and recA in the propagation of TGG repeats in *Escherichia coli*. *Nucleic Acids Res* 28(16):3178–3184.

113. Darmon E, et al. (2010) E. coli SbcCD and RecA control chromosomal rearrangement induced by an interrupted palindrome. *Mol Cell* 39(1):59–70.

114. Kulkarni SK, Stahl FW (1989) Interaction between the Sbcc Gene of *Escherichia coli* and the Gam Gene of Phage ? *Genetics* 123(2):249–253.

115. Murphy KC (1991) Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of *Escherichia coli* RecBCD enzyme. *Journal of bacteriology* 173(18):5808–5821.

116. Marsić N, Roje S, Stojiljković I, Salaj-Smic E, Trgovcević Z (1993) In vivo studies on the interaction of RecBCD enzyme and lambda Gam protein. *J Bacteriol* 175(15):4738–4743.

117. Sharples GJ, Leach DRF (1995) Structural and functional similarities between the SbcCD proteins of *Escherichia coli* and the RAD50 and MRE11 (RAD32) recombination and repair proteins of yeast. *Molecular Microbiology* 17(6):1215–1217.

118. Connelly JC, Kirkham LA, Leach DR (1998) The SbcCD nuclease of *Escherichia coli* is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proceedings of the National Academy of Sciences* 95(14):7969–7974.

119. Paull TT, Gellert M (1999) Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev* 13(10):1276–1288.

120. Bressan DA, Baxter BK, Petrini JHJ (1999) The Mre11-Rad50-Xrs2 Protein Complex Facilitates Homologous Recombination-Based Double-Strand Break Repair inSaccharomyces cerevisiae. *Mol Cell Biol* 19(11):7681–7687.

121. Ghodke I, Muniyappa K (2013) Processing of DNA double-stranded breaks and intermediates of recombination and repair by Saccharomyces cerevisiae Mre11 and its stimulation by Rad50, Xrs2, and Sae2 proteins. *J Biol Chem* 288(16):11273–11286.

122. Koonin EV, et al. (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biology* 5:R7.

123. Yoshida T, Claverie J-M, Ogata H (2011) Mimivirus reveals Mre11/Rad50 fusion proteins with a sporadic distribution in eukaryotes, bacteria, viruses and plasmids. *Virology Journal* 8:427.

124. Badugu SB, et al. (2015) Identification of Plasmodium falciparum DNA Repair Protein Mre11 with an Evolutionarily Conserved Nuclease Function. *PLOS ONE* 10(5):e0125358.

125. Connelly JC, Leach DRF (2002) Tethering on the brink: the evolutionarily conserved Mre11–Rad50 complex. *Trends in Biochemical Sciences* 27(8):410–418.

126. Hiasa H, Marians KJ (1994) Tus prevents overreplication of oriC plasmid DNA. *Journal of Biological Chemistry* 269(43):26959–26968.

127. Patel R, et al. (1996) Formation of chimeric DNA primer extension products by template switching onto an annealed downstream oligonucleotide. *PNAS* 93(7):2969–2974.

128. Duggin IG, Wake RG, Bell SD, Hill TM (2008) The replication fork trap and termination of chromosome replication. *Molecular Microbiology* 70(6):1323–1333.

129. Demassy B, Patte J, Louarn J, Bouche J (1984) oriX: A new replication origin in E. coli. *Cell* 36(1):221–227.

130. Magee TR, Asai T, Malka D, Kogoma T (1992) DNA damage-inducible origins of DNA replication in *Escherichia coli*. *EMBO J* 11(11):4219–4225.

131. Santocanale C, Sharma K, Diffley JFX (1999) Activation of dormant origins of DNA replication in budding yeast. *Genes Dev* 13(18):2360–2364.

132. Cohen A, Clark AJ (1986) Synthesis of linear plasmid multimers in *Escherichia coli* K-12. *Journal of bacteriology* 167(1):327–335.

133. Thoms B, Borchers I, Wackernagel W (2008) Effects of Single-Strand DNases ExoI, RecJ, ExoVII, and SbcCD on Homologous Recombination of recBCD+ Strains of *Escherichia coli* and Roles of SbcB15 and XonA2 ExoI Mutant Enzymes. *J Bacteriol* 190(1):179–192.

134. Lovett ST (2011) The DNA exonucleases of *Escherichia coli*. *Ecosal Plus* 2011. doi:10.1128/ecosalplus.4.4.7.

135. Lovett ST, Kolodner RD (1989) Identification and purification of a single-stranded-DNA-specific exonuclease encoded by the recJ gene of *Escherichia coli*. *PNAS* 86(8):2627–2631.

136. Lovett ST, Luisi-DeLuca C, Kolodner RD (1988) The genetic dependence of recombination in recD mutants of *Escherichia coli*. *Genetics* 120(1):37–45.

137. Cunningham RP, Saporito SM, Spitzer SG, Weiss B (1986) Endonuclease IV (nfo) mutant of *Escherichia coli*. *J Bacteriol* 168(3):1120–1127.

138. Lloyd RG, Buckman C (1991) Genetic analysis of the recG locus of *Escherichia coli* K-12 and of its role in recombination and DNA repair. *J Bacteriol* 173(3):1004–1011.

139. Whitby MC, Ryder L, Lloyd RG (1993) Reverse branch migration of holliday junctions by RecG protein: A new mechanism for resolution of intermediates in recombination and DNA repair. *Cell* 75(2):341–350.

140. Lloyd RG, Porton MC, Buckman C (1988) Effect of recF, recJ, recN, recO and ruv mutations on ultraviolet survival and genetic recombination in a recD strain of *Escherichia coli* K12. *Mol Gen Genet* 212(2):317–324.

141. Umezu K, Nakayama K, Nakayama H (1990) *Escherichia coli* RecQ protein is a DNA helicase. *Proc Natl Acad Sci USA* 87(14):5363–5367.

142. Kusano K, Nakayama K, Nakayama H (1989) Plasmid-mediated lethality and plasmid multimer formation in an *Escherichia coli* recBC sbcBC mutant. Involvement of RecF recombination pathway genes. *J Mol Biol* 209(4):623–634.

143. Chow K-H, Courcelle J (2007) RecBCD and RecJ/RecQ Initiate DNA Degradation on Distinct Substrates in UV-Irradiated *Escherichia coli*. *Radiation Research* 168(4):499–506.

144. Courcelle J, Wendel BM, Livingstone DD, Courcelle CT (2015) RecBCD is required to complete chromosomal replication: Implications for double-strand break frequencies and repair mechanisms. *DNA Repair* 32:86–95.

145. Wendel BM, Cole JM, Courcelle CT, Courcelle J (2017) SbcC-SbcD and ExoI process convergent forks to complete chromosome replication. *Proceedings of the National Academy of Sciences*:201715960.

Chapter 2

Completion of DNA replication in *Escherichia coli*

*****This chapter has been published

Wendel BM, Courcelle CT, Courcelle J (2014) Completion of DNA replication in *Escherichia coli*. *Proc Natl Acad Sci*:201415025

Abstract

The mechanism by which cells recognize and complete replicated regions at their precise doubling point must be remarkably efficient, occurring thousands of times per cell division along the chromosomes of humans. Yet this process remains poorly understood. Here we show that in *Escherichia coli*, the completion of replication involves an enzymatic system that effectively counts pairs and limits cellular replication to its doubling point by allowing converging replication forks to transiently continue through the doubling point before the excess, over-replicated regions are incised, resected, and joined. Completion requires RecBCD, and involves several proteins associated with repairing double-strand breaks including, ExoI, SbcDC, and RecG. However, unlike double-strand break repair, completion occurs independently of homologous recombination and RecA. In some bacterial viruses, the completion mechanism is specifically targeted for inactivation in order to allow over-replication to occur during lytic replication. The results suggest a primary cause of genomic instabilities in many double-strand break repair mutants arises from an impaired ability to complete replication, independent from DNA damage.

Significance Statement

All phases of DNA replication are tightly regulated to ensure that daughter cells inherit a precise copy of the genomic DNA. While the mechanisms regulating initiation and elongation have been well characterized, the process of how cells recognize replicated regions and complete replication at the precise doubling point remains a fundamental question yet to be addressed. Here we show that the completion of replication involves a transient over-replication of the region where forks converge before the excess regions are incised, resected, and joined. Completion requires several proteins associated with repairing double-strand breaks, but unlike break repair, it occurs independently of homologous recombination and is targeted for inactivation by some bacterial viruses during the transition to lytic replication.

Introduction

During chromosomal replication, cells tightly regulate the processes of initiation, elongation, and completion to ensure that each daughter cell inherits an identical copy of the genetic information. While the mechanisms regulating initiation and elongation have been well characterized (reviewed in (1, 2)), the process of how cells recognize replicated regions and complete replication at the precise doubling point remains a fundamental question yet to be addressed. Whether this event occurs once per generation as in *E. coli* or thousands of times per generation as in human cells, the failure to efficiently carry out this function would be expected to result in a loss of genomic stability. Considering the large number of proteins that cells devote to ensuring the fidelity of replication initiation and elongation, it seems highly probable that the final critical step in this process will be also be tightly regulated and controlled enzymatically.

In some aspects, one could argue that the efficiency of completion is likely to be more critical to the faithful duplication of the genome than that of initiation. When replication origins fail to initiate efficiently, elongation of replication forks from neighboring origins is often able to compensate (3, 4), and both prokaryotic and eukaryotic cells are able to tolerate variations in their origin number without severe phenotypic consequences (5–7). However, a failure to accurately limit or join any event where forks converge would be expected to result in duplications, deletions, rearrangements or a loss of viability depending upon how the DNA ends are resolved at segregation.

A number of studies suggest that an ability to sense when all sequences in the genome have doubled will be critical to genomic replication. *In vitro*, converging

replisomes continue through their meeting point as one replisome displaces the other, resulting in over-replication, or a third copy, of the region where the forks meet (8). Complicating the process of genomic doubling a even further, several studies have suggested that illegitimate initiations of replication frequently occur at single strand nicks, gaps, D-loops, and R-loops throughout the genomes of both prokaryotes and eukaryotes (9–14). Similar to when replication forks continue through a previously replicated template, each of these events would generate a third copy of the chromosomal region where the event occurs. Thus, over-replication may be inherent and promiscuous during the duplication of genomes. If true, then to ensure that each sequence of the genome replicates once, and only once per generation, cells must encode an enzymatic system that is essentially able to count in pairs, and efficiently degrade odd or over-replicated regions until the two nascent end pairs of replication events can be joined.

The model organism *E. coli* is particularly well-suited to dissect how this fundamental process occurs. In *E. coli*, the completion of replication occurs at a defined region on the genome, opposite to the bidirectional origin of replication (15). Most completion events can be further localized to one of six termination (*ter*) sequences within the 400-kb terminus region due to the action of Tus, which binds to *ter* and inhibits replication fork progression in an orientation-dependent manner, in effect stalling the replication fork at this site until the second arrives (16, 17). Although Tus confines converging replication forks to a specific region, it does not appear to be directly involved in the completion reaction since *tus* mutants have no phenotype and complete replication normally (18). Furthermore, plasmids and bacteriophage lacking *ter* sequences are maintained stably (19).

Many mutants impaired for either replication initiation or elongation were initially isolated based on their growth defects or an impaired ability to maintain plasmids (20–22). We reasoned that mutants impaired for the ability to complete replication might be expected to exhibit similar phenotypes and initially focused our attention on the properties of *recBC* and *recD* mutants. RecB-C-D forms a helicase-nuclease complex that is required for homologous repair of double-strand breaks in *E. coli* (23, 24). The enzyme utilizes specific DNA sequences, termed Chi sites, to initiate recombination between pairs of molecules. Loss of RecB or C inactivates the enzyme complex, whereas loss of RecD inactivates the nuclease and Chi recognition, but retains helicase activity (23, 24). Here, we show that inactivation of RecBCD leads to a failure to recognize and join replicating molecules at their doubling point. Although the completion process requires RecBCD, it is distinct from double-strand break repair and does not involve a double strand break intermediate, homologous recombination, or RecA.

Materials and Methods

Bacteria and Plasmids. Strains and plasmids used in this study are presented in Table 1.

Growth rates. Fresh overnight cultures were 10-fold serially diluted in Luria-Bertani medium supplemented with 10 µg/ml thymine (LBthy), and 0.2-ml aliquots then were plated in duplicate into the wells of a sterile 96-well microtiter dish. The microtiter cultures were then agitated at 37°C, and the absorbance at 630nm for each culture was measured over time using a BIO-Whittaker ELx808 plate reader. The number of viable colonies per ml in each overnight culture was determined at the start of every experiment.

Plasmid Stability. Cultures containing the plasmid pBR322 were grown for 30 generations in LBthy medium at 37°C with aeration. Ten-µl aliquots of serial 10-fold dilutions were then spotted on LBthy plates in the presence and absence of 100 µg/ml ampicillin. Viable colonies were counted following overnight incubation at 37°C.

Total Genomic and Plasmid DNA extraction. 200µl of a fresh overnight culture grown in LBthy medium supplemented 100 µg/ml ampicillin was pelleted and used to inoculate 20 ml of LBthy medium. Cultures were grown without ampicillin selection in a shaking incubator at 37°C to an OD600 of 0.5 (~ 5 x 10^8 cells/ml). 0.75 ml samples were then placed into 0.75 ml cold 2X NET (100 mM NaCl, 10 mM Tris, pH 8.0, 10 mM EDTA). Each sample was pelleted, resuspended in 150 µl of 1 mg/ml lysozyme and 0.2 mg/ml RNaseA in TE (10 mM Tris, pH 8.0, 1 mM EDTA), lysed at 37°C for 20 min. At this time, proteinase K (10µl, 10mg/ml) and sarcosyl (10µl, 20%) was added and incubated at 37°C for 1 hr. Samples were then extracted with 4 volumes of phenol/chloroform (1/1), and dialyzed for 1 hour on 47mm Whatman 0.05 µm pore disks (Whatman #VMWP04700) floating on a 250 ml beaker of TE.

Southern Analysis of Plasmid Replication intermediates. Total genomic DNA samples were digested with Sac II (New England Biolabs), which is not found in pBR322, extracted with chloroform, and equal volumes were loaded onto a 1.0 % agarose, 1X TAE at 4V/cm. Alternatively, to resolve plasmid sizes, samples were run in 0.5 % agarose, 1X TAE at 1V/cm. Gels were transferred to Hybond N+ nylon membranes and

probed with pBR322 that had been labeled with ^{32}P by nick translation according to the protocol supplied by Roche using alpha [^{32}P]dCTP (PerkinElmer). Radioactivity was visualized using a Storm 840 and its associated ImageQuant Software (Molecular Dynamics).

Sequencing, assembly and copy-number analysis. Fresh overnight cultures were diluted 1:250 in LBthy media and grown at 37°C with aeration to an OD_{600} of 0.4. Total genomic DNA was extracted as described above. Stationary phase cultures were grown for 36 hours prior to genomic DNA extraction. Library preparation and sequencing of the genomic DNA samples were performed using NexteraXT and Illumina HiSeq2000 (Illumina). The SR108 parent strain sequence was determined using single-end 51 base pair bar-coded reads to assemble contigs with the Velvet 1.2.10 *De Novo* assembler with a k-mer value of 31 a minimum coverage depth >34 (25). Contigs were then aligned to *E. coli* K12 W3110, as a scaffold, using CONTIGuator 2.7.3 (26). The original Illumina sequences reads for all subsequent strains were then aligned and assembled using Bowtie 1.0.0 (27) with the SR108 parent as reference. The aligned reads were then analyzed for nucleotide frequencies at each position and the copy number of sequences per kilobase was determined using Perl scripts. Relative copy number values for each strain were normalized to values obtained for stationary phase cultures to correct for any differences in read depth, and then plotted against their location on the genome (28).

Density labeling and CsCl analysis. Fresh overnight cultures were diluted 1:100 in 20 ml of Davis medium (2.0 g KH_2PO_4, 7.0 g K_2HPO_4, 0.5 g $Na_3C_6H_5O_7$, 0.1 g $MgSO_4$, 1.0

g $(NH_4)_2SO_4$ per liter, pH 7.0) supplemented with 0.4% glucose, 0.2% cas-amino acids, and 10 µg/ml thymine (DGCthy) supplemented with 0.1 µCi/ml of [^{14}C]thymine (53 mCi/mmol) and were grown to an OD_{600} of 0.5 at 37°C with aeration. At this time, half the culture was UV irradiated with 25 J/m^2 and the other half was mock irradiated. Cultures were then filtered onto FisherBrand general filtration 0.45-µm membranes, washed with NET buffer, resuspended in 10 ml DGC medium supplemented with 20 µg/ml 5-bromouracil in place of thymine and 0.5 µCi/ml [^{3}H]thymine (60.5 Ci/ mmol), and allowed to recover for a period of 1 hr at 37°C with aeration. Two volumes of ice-cold NET buffer were added to the 10-ml cultures, and the cells were then pelleted, resuspended in 150 µl TE (10 mm Tris, 1 mm EDTA, pH 8.0), and lysed in 170 µl of 0.5M H_2KPO_4/KOH, pH 12.5, and 1.25% Sarkosyl. Isopycnic alkali CsCl gradients composed of 0.3 g of a DNA lysate solution, 2.23 g CsCl, and 3.31 g of a 0.1 M H_2KPO_4/KOH, pH 12.5, solution (refractive index 1.4055) were centrifuged to equilibrium at 80,000g for 96 hr at 20°C. Gradient fractions were collected onto Whatman no. 17 paper, washed in 5% TCA, and then washed in 95% ethanol. The quantity of ^{3}H and ^{14}C in each fraction was determined by liquid scintillation counting.

Total DNA accumulation. Fresh overnight cultures were diluted 1:100 in 40 ml DGCthy medium supplemented with 0.1 µCi/ml [^{3}H]thymine (60.5 Ci/mmol) and grown to an OD_{600} of 0.4 at 37°C with aeration. At this time, half the culture was UV irradiated with 25 J/m^2 and the other half was mock irradiated. At the indicated times, duplicate 200-µl aliquots were precipitated in 5 ml of 5% trichloroacetic acid (TCA) and filtered onto

Fisherbrand glass fiber filters. The amount of ^{3}H-labeled DNA on each filter was determined by liquid scintillation counting.

Results

Similar to other mutants that are involved in replication initiation or elongation, *recBC* and *recD* mutants each exhibit growth abnormalities and plasmid instabilities. These phenotypes are unique from other recombination mutants, and suggest they have a broader, more fundamental function in replicating cells. Relative to wild type cultures, *recBC* cultures grow poorly and produce large numbers of small, nonviable cells, whereas *recD* cultures grow for a longer time period and reach a higher cell density (Figure 2.1a) (29–31). By comparison, cultures lacking either RecF or RecA, which is essential for all homologous recombination and RecBCD-mediated double-strand break repair, grow comparatively well, arguing that some function of RecBCD is unique from homologous repair and DNA damage.

Mutations inactivating RecBC or RecD also affect the stability of plasmid minichromosomes, a feature that is again distinct from other recombination mutants (Figure 2.1b) (32, 33). Plasmids grown in *recD* mutants continue to replicate past the doubling point, producing large quantities of multimeric circles as well as long linear multimers (Figure 2.1c). The over-replicated products observed in *recD* mutants are distinct in that they contain both odd- and even-numbered multimeric products (Figure 2.1d). By contrast, in other recombination mutants or in wild type cultures, the few multimeric products that are detected occur as paired, or even-numbered multimers. *recBC* mutants are also less able to retain plasmids relative to wild type cultures,

although overall cell viability is similarly reduced (Figure 2.1b). The unstable phenotype in *recBC* is distinct from *recD* mutants, and involves an elevated level of gapped molecules and dimer plasmids, rather than extensive over-replication (Figure 2.1c). We interpret these observations to suggest that during plasmid replication, the RecD and RecBC subunits of the enzyme are required for cells to recognize and resolve those ends at the doubling point, respectively.

If the plasmid instability in *recBC* and *recD* mutants arose from an inability to process double-strand breaks, these mutants would be expected to accumulate broken intermediates. Yet, as *figure 2.1c* and *f* demonstrate, the proportion of broken, linear plasmids is actually lower in *recBC* or *recD* cultures relative to wild type or other recombination mutants. Additionally, double-strand breaks are estimated to arise in vivo at frequencies ranging from 0.01-1 break per 4.5-Mb of replicated genome (34), making it unlikely that these account for the instability of a 4.5-kb plasmid. Finally, plasmids remain stable and replicate normally in *recA* mutants which are defective in all homologous recombination and RecBCD-mediated double-strand break repair (Figure 2.1b and c). Taken together, these observations argue strongly against the idea that the growth and minichromosome abnormalities in *recBC* and *recD* mutants arise from defective processing of double-strand breaks. However, these phenotypes are all consistent with those expected of mutants that have an impaired ability to recognize and complete replication.

Other phenotypes associated with *recBC* mutants also suggest the gene products play a role at the end of the cell cycle. Following UV-irradiation, many hypersensitive recombination mutants, including *recA* and *recF*, cease DNA synthesis immediately after

replication encounters the DNA damage (35, 36). However, *recBC* mutants are unusual in that they initially recover and continue to replicate similar to wild type cells. The replication continues normally for a short period before DNA synthesis ceases at a point when the DNA has approximately doubled ((37) and 2.2), indicating that the defect in *recBC* mutants is distinct from RecA and arises at the final stages of replication. Consistent with this interpretation, *ter* sequences are hot spots for RecBCD-mediated recombination (38, 39), implying that the region where replication completes contains substrates frequently recognized by RecBCD in vivo.

To directly examine whether RecBCD functions in completing replication on the chromosome, we profiled the genomes of replicating wild type and mutant cultures using high-throughput sequencing. In replicating wild type cultures, the copy number of sequences is highest surrounding the bidirectional origin, then gradually decreases until it reaches the terminus where replication completes (Figure 2.3a). In mutants lacking RecBC, there is a marked decrease in the copy number of sequences specifically in the terminus region. The terminus sequences in *recBC* mutants are underrepresented by up to two fold, relative to wild type cultures. Assuming that greater than half of the sequence reads correspond to the parental DNA, one can infer that the majority of cells in the population have difficulty replicating or maintaining sequences in this region. Conversely, an increase in the copy number of sequences within the terminus region is observed in *recD* mutants, which inactivates the exonuclease activity of the enzyme complex (Figure 2.3b). Consistent with the observations on plasmids, the results indicate that the RecBCD complex is required to allow the efficient and accurate completion of replication on the chromosome. The presence of the over-replicated intermediate inside

the boundary of the ter sites in *recD* mutants implies that converging forks transiently pass each other before the nuclease activity of RecBCD resects these over-replicated intermediates back to the doubling point. The lack of sequences at the termination region in *recBC* mutants reveals that the enzyme complex is required to resolve and join the convergent forks at the doubling point. In its absence, the DNA ends of the converging forks remain subject to nucleolytic attack and are degraded.

Importantly, the completion of replication on the chromosome occurs normally in *recF* and *recA* mutants, indicating that the completion reaction catalyzed by RecBCD does not require homologous recombination or involve the repair of double-strand breaks (2.3B). We are aware of no recombination models for repairing collapsed forks that do not involve RecA, nor do any known recombinational processes require RecBC, but not RecA. Thus, the lack of the terminus region DNA in *recBC* mutants is inconsistent with the idea that the intermediates are associated recombination defect or collapsed replication forks occurring in this region. We infer that the impaired ability to complete replication in *recBC* mutants is independent from its role in double-strand break repair, and likely accounts for the poor growth of these cells relative to *recA* or other recombination mutants.

Additional genes associated with double-strand break repair are also involved in completion. SbcDC, a structure-specific helicase-nuclease, and ExoI, a prominent 3'-5' exonuclease, suppress the growth defects of *recBC* mutants when mutated, and lead to plasmid instability similar to *recD* (29, 40, 41). Mutations in human homologs of these proteins are associated with genetic instabilities and impaired double-strand break repair (42). In replicating *sbcDC xonA* mutants, a similar over-replication of the terminus

region is observed (Figure 2.3c), indicating that these genes play a role in processing or resolving the transient over-replicated regions. Over-replication was less pronounced in the single mutant, possibly suggesting either functional redundancy or cooperativity between these gene products.

A recent study has shown that mutants lacking RecG, a translocase important for dissolving mis-primed events after DNA damage, also over-replicates its terminus region (Figure 2.3c and (43)). In both *recD* and *recG* mutants, the DNA ends from unresolved completion events lead to over-replication that can also be observed on plasmids. However, as shown in *figure 2.4*, the over-replication that occurs in these mutants is distinct in several aspects. The aberrant long linear-multimeric intermediates that accumulate in *recD* mutants do not appear in *recG* mutants. In addition, *recD* mutants are unique in that they contain prominent odd numbered circular plasmid multimers, suggesting that RecD contributes to efficient pair recognition prior to resolution. In contrast, the over-replicated species in *recG* mutants predominately consist of even-numbered circular multimers (Figure 2.4), suggesting that these mutants retain the ability to recognize and resolve molecules as pairs. We interpret these results to suggest that although RecG plays a role in preventing illegitimate re-initiations from occurring, it is not directly involved in recognition or joining of the linear DNA ends at the doubling point. Consistent with this interpretation, *recG* mutants grow normally and plasmids are stably maintained (Figure 2.1d and e). *recG* mutants are also constitutively induced for SOS expression (44), which may contribute to the over-replication that occurs on plasmids and the chromosome in these strains (10, 11). Interestingly, the over-replication that occurs in both *recD* and *recG* mutants depends on RecA (Figure 2.4),

demonstrating that recombination can lead to aberrant re-initiation events when the efficiency of the completion reaction is compromised.

Many lytic viruses, including bacteriophage lambda, have two modes of replication, an early phase in which its genome doubles similar to the bacterial chromosome, and a late phase in which the viral genome is amplified before packaging and release from the cell (45). Late phase replication in phage lambda requires expression of the phage Gam protein, which targets and inactivates RecD and SbcDC in the host (46). Similar to the amplification of phage and plasmid DNA (41), we observed that gam expression results in an over-replication of the terminus region (Figure 2.5). Thus, to initiate genomic amplification during lytic infection, the phage targets and inactivates the cellular mechanism that limits replication to the doubling point in order to allow over-replication to occur.

Discussion

Taken together, the plasmid and chromosomal data presented here indicate that RecBCD is directly involved in limiting replication events and resolving them at points where sequences have doubled. This process is distinct from double-strand break repair and occurs efficiently in the absence of RecA or homologous recombination on both plasmids and the chromosome. However, when one considers the mechanism by which double strand breaks are repaired, it becomes clear how these two processes may be related (Figure 2.6). Double-strand break repair in *E. coli* requires both RecA and RecBCD function. RecA is believed to pair the severed strands with intact homologous duplex DNA (23, 24). Once this occurs, the sequences between the opposing strands are replicated and joined using the second molecule as a template. A structurally similar

process must also occur whenever two replication forks converge. However, in the case of completion, the opposing nascent strands have been brought together by replication forks and should be independent of RecA.

During double strand break repair, RecBCD is proposed to process the DNA ends prior to strand invasion. In most models, this processing is restricted to the early stages of the reaction (23, 24). However, *in vivo* experiments have suggested that strand invasion can occur in the absence of RecBCD, but that its function is still required if viable recombinants are to be recovered (47, 48), arguing that RecBCD enzyme function acts late in the recombination process, perhaps by actively resolving the re-replicated regions at the doubling point.

Considering the chromosomal phenotypes of *recBC* and *recD* mutants, it is tempting to speculate that monomeric linear plasmid species, which are diminished or absent in these strains, represent an incised intermediate of over-replicated products (Figure 2.1d). However, the precise substrates RecBCD, or SbcDC, XonA, and RecG act upon in the completion process remains to be determined, as does the presumed role that a polymerase and ligase must play in joining the DNA ends.

Several observations favor a mechanism involving a transient over-replication when forks converge. The location of *ter* sequences on the chromosome are positioned to allow overreplication of the terminus region to occur before replication is blocked by the action of the Tus protein (15). A number of early studies found that under various stress conditions, the copy number of sequences surrounding the *ter* regions increased and speculated that these represented cryptic origins of replication, termed *oriX*, *oriK*, or *oriM* (9, 11, 49). However, it is also reasonable to consider that these new "origins"

actually represent replication continuing through the terminus, since both events would result in elevated copy numbers in this region. Consistent with this, chromosomal over-replication is generally observed to occur under the same conditions as when it is seen on plasmids (Figure 2.1&2.3), arguing against the idea of cryptic origins in the terminus. *In vitro*, converging replisomes continue through their meeting point as one replisome displaces the other, resulting in over-replication of the point where the forks meet (8). Finally, transient over-replication has the intuitive advantage of buffering against any potential loss of genetic information and may prevent the loss of genetic information should cell division occur prematurely.

Mutations in a number of human double strand break repair genes also exhibit growth defects and genetic instabilities in the absence of exogenous DNA damage, suggesting that some of these mutants may have an impaired ability to complete replication.

Data deposition

The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo accession no. SRP047195)

Acknowledgements

This work was supported by NIH/NIEH Grant-R15ES021594.

Table 2.1. *E. coli* strains and plasmids used in this study

Strain	Genotype	Source or Construction
SR108	*λ-, thyA, deo, IN(rrnD-rrnE)*	(5)
HL921	SR108 D(*srlR-recA*)306::*Tn*10	(6)
HL922	SR108 *recB21C22 argA81*::*Tn*10	(6)
HL923	SR108 *recD1011 argA81*::*Tn*10	(6)
HL946	SR108 *recF332*::*Tn*3	(7)
CL008	SR108 *recG258*::*Tn*5	(8)
HL1034	SR108 *xonA*::*Cat300*	(9)
CL826	D(*lacU169*) *nadA*::*Tn*10, *gal490*, Lambda *cI857*, D(*cro-bioA*), *sbcCD*::*Cat*	Recombineering strain DY329 (10) was transformed with primers 5'TCCTGCTGAATAGTTATTTCACTGCAAACGTACTTTCCAGCTTTCGAATTTCTGCCATTC 5'AGGGAACCGTTATGCGCATCCTTCACACCTCAGACTGGCAATGAGACGTTGATCGGCAC to replace codon10 of *sbcD* through codon1040 of *sbcC* with a chloramphenicol resistance cassette.
CL835	SR108 *sbcCD1040*::*cat*	P1 transduction of *sbcCD1040*::*cat* from CL826 parent into SR108 recipient
CL2357	SR108 *xonA*::*Cat300 sbcCD*::*Gm*	P1 transduction of *sbcCD*::Gm from KM135 (11) into recipient HL1034
pBADGam	Arabinose inducible expression plasmid, Ampicillin resistance, ColE1 origin	Phage Lambda gam ORF clones into the multiple cloning site of pBAD/Myc-HisA (Invitrogen)
pBR322	Ampicillin, Tetracycline resistance, ColE1 origin	(12)

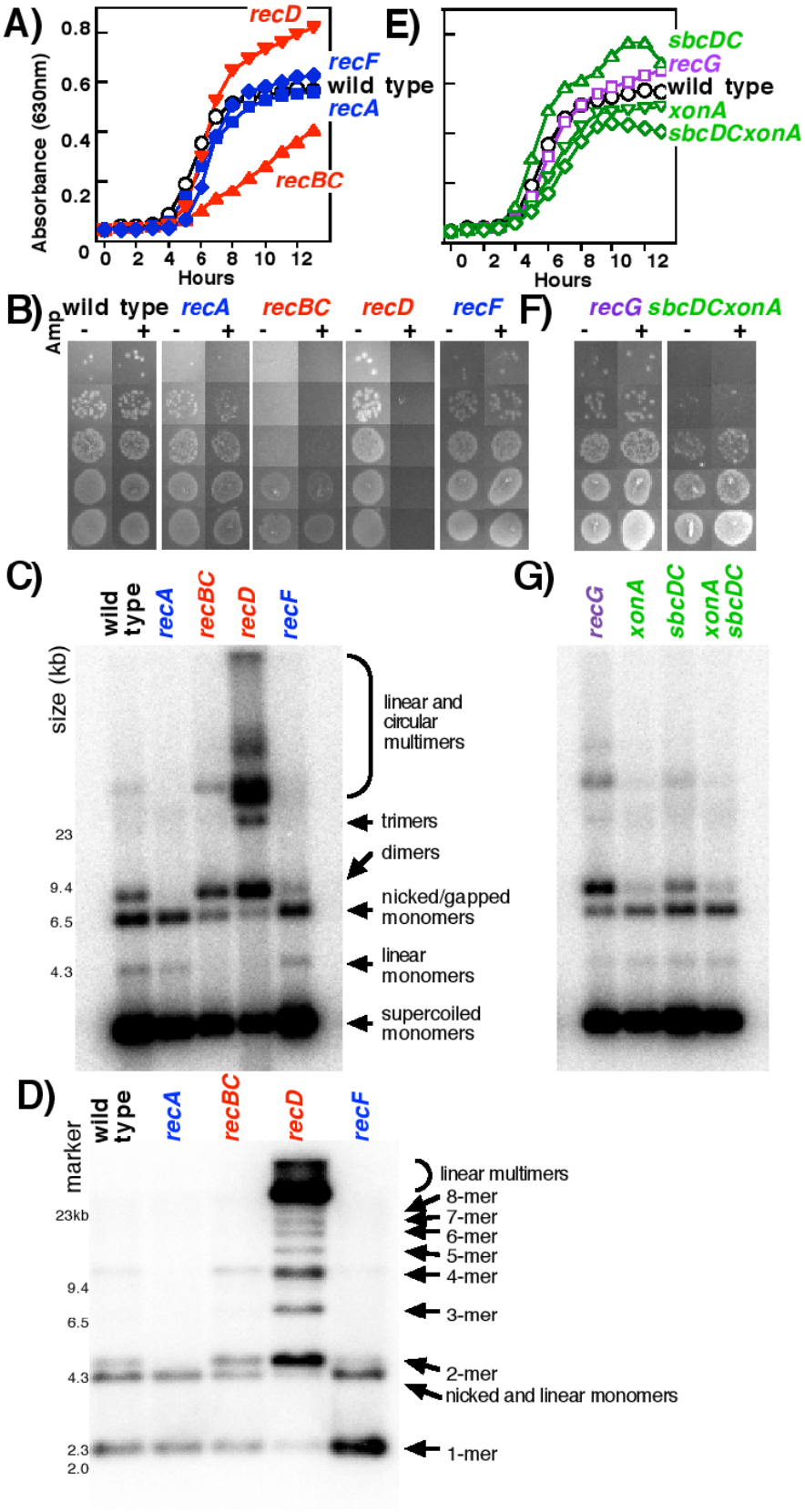
A)
Absorbance (630nm)
Hours
recD
recF
wild type
recA
recBC
E)
sbcDC
recG
wild type
xonA
sbcDCxonA
B)
Amp
wild type
recA
recBC
recD
recF
F)
recG
sbcDCxonA
C)
wild type
recA
recBC
recD
recF
size (kb)
23
9.4
6.5
4.3
linear and circular multimers
trimers
dimers
nicked/gapped monomers
linear monomers
supercoiled monomers
G)
recG
xonA
sbcDC
xonA sbcDC
D)
marker
wild type
recA
recBC
recD
recF
23kb
9.4
6.5
4.3
2.3
2.0
linear multimers
8-mer
7-mer
6-mer
5-mer
4-mer
3-mer
2-mer
nicked and linear monomers
1-mer

Figure 2.1. *recBC* and *recD* mutants exhibit growth abnormalities and an impaired ability to maintain monomeric plasmids. A) The growth of *recBC* mutants is impaired, whereas *recD* mutants grow for a longer period of time and reach a higher density relative to cultures of wild type, *recA* or *recF* mutants. The absorbance at 630nm of cultures grown at 37°C is plotted over time. B) *recBC* mutants and *recD* mutants exhibit plasmid instability. Cultures containing the plasmid pBR322 were grown for 30 generations before 10µl drops of 10fold serial dilutions were plated with and without ampicillin to determine the fraction of cells that retained the plasmid in each strain. C) Plasmids replicating in *recBC* cultures accumulate dimer plasmids, whereas *recD* cultures accumulate circular and linear multimers. Linear monomers, indicative of double-strand breaks, are reduced in both *recBC* and *recD* mutants relative to wild type cultures. Total genomic and plasmid DNA was prepared from replicating cultures containing pBR322 and examined by Southern analysis using 32P-labelled pBR322 as a probe. DNA was electrophoresed through a 1.0% agarose gel in 0.5X TAE at 4 V/cm^2 D) Unlike other mutants, replication of plasmids in *recD* mutants leads to multimeric circles that contain both odd and even numbers of plasmid copies. Samples were analyzed as in (C) except the DNA was electrophoresed through a 0.5% agarose gel in 0.5X TAE at $1V/cm^2$. Resolution under these conditions resolves molecules primarily based on the molecule's size and reduces the impact that shape has on the migration rate of the molecule. E) The growth, F) plasmid stability, and G) plasmid intermediates for *recG*, *xonA*, *sbcDC*, and *xonAsbcDC* mutants were analyzed as in (A) (B) and (C).

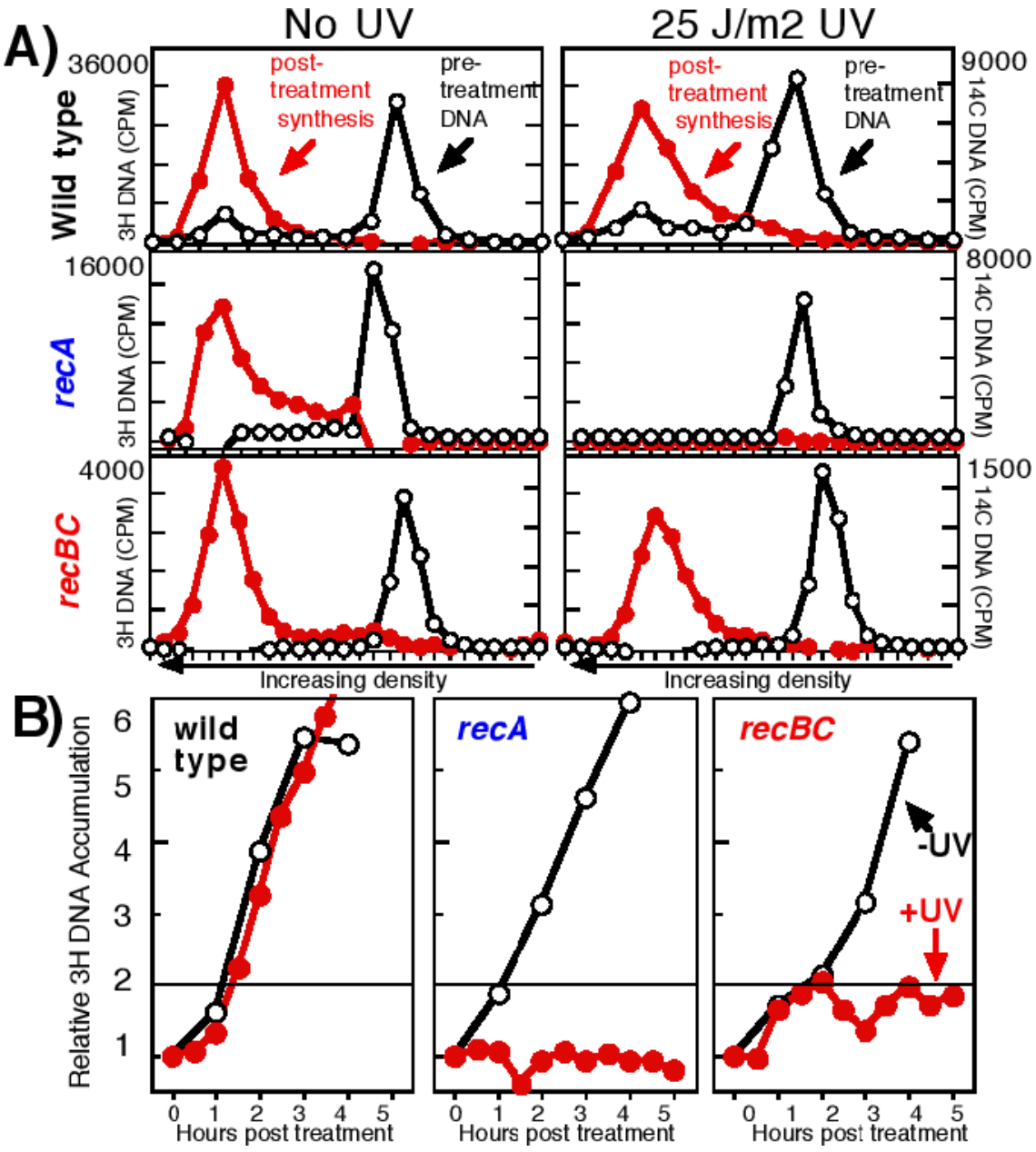
A)
No UV
25 J/m2 UV
Wild type
recA
recBC
3H DNA (CPM)
14C DNA (CPM)
36000
16000
4000
9000
8000
1500
post-
treatment
synthesis
pre-
treatment
DNA
Increasing density
B)
Relative 3H DNA Accumulation
wild
type
recA
recBC
-UV
+UV
Hours post treatment

Figure 2.2. Following UV-irradiation, *recBC* mutants initially recover replication, but then replication arrests after an approximate doubling of their genomic material. A) *recBC* mutants initially recover replication. [^{14}C]thymine prelabeled cultures were UV irradiated or mock treated, and resuspended in media containing [^{3}H]5-bromodeoxyuridine for 1 hour to density label the replication occurring during this period. The denser replicated DNA was then separated in alkaline CsCl density gradients and quantified. Both wild type and *recBC* mutants restore replication equally well during the 1st hour after UV treatment. *recA* mutants do not recover. B) Replication arrests in *recBC* mutants after an approximate doubling of the DNA. Cultures growing in ^{3}H-thymine were UV irradiated or mock treated and sampled at various times to determine the total amount of [^{3}H]DNA accumulated. Wild type cells recover replication and continue to grow following irradiation. *recA* mutants do not recover replication. *recBC* initially recover replication, but then arrest replication once the DNA has approximately doubled. Initial [^{3}H]DNA counts were between 1057 and 2610 cpm for all experiments. Plots represent the average of duplicate samples.

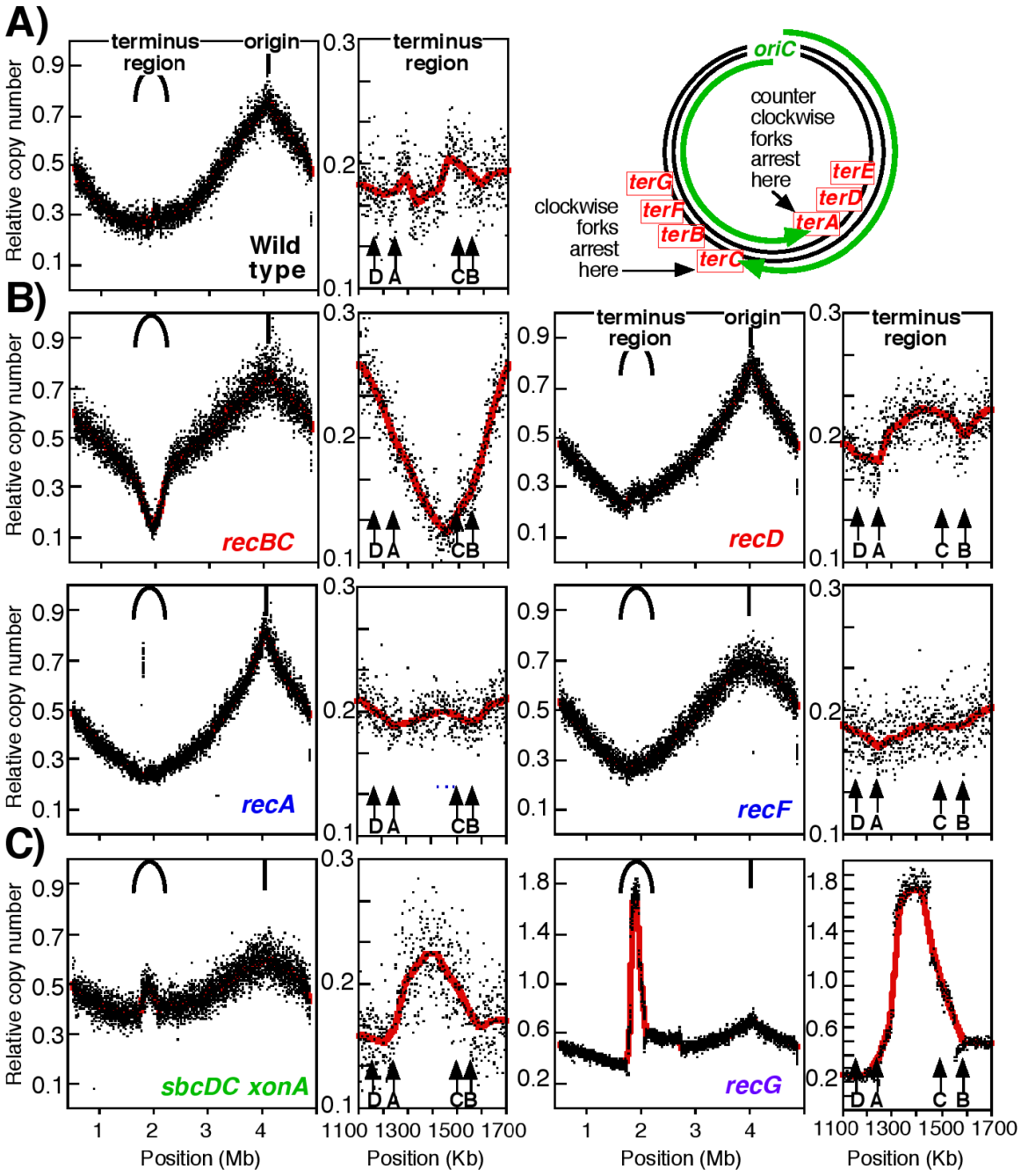
A)
B)
C)
terminus region
origin
Relative copy number
Wild type
recBC
recA
recD
recF
sbcDC xonA
recG
oriC
counter clockwise forks arrest here
clockwise forks arrest here
terG
terF
terB
terC
terE
terD
terA
D A
C B
Position (Mb)
Position (Kb)

Figure 2.3. RecBCD resolves and completes replication at the doubling point on the chromosome, independent of homologous recombination. A) In wild type cultures, replication proceeds bidirectionally from the origin and completes in the terminus region. Genomic DNA from replicating cultures was purified, fragmented, and profiled using high-throughput sequencing. Sequence read frequencies, normalized to stationary phase cells, are plotted relative to their position on the genome. The terminus region of the chromosome, containing *terD*, *A*, *C*, and *B*, is shown next to each plot. An 8kb floating average of the sequence frequency is plotted in red. B) *recBC* mutants fail to complete replication, leading to degradation of the terminus region. *recD* mutants fail to resect and limit replication to the doubling point, leading to over-replicated regions in the terminus. Completion occurs normally in *recF* and *recA* mutants. C) Over-replicated regions persist in *sbcDC xonA* mutants. Illegitimate re-initiations of replication occur in *recG*. Note the different scale for *recG*.

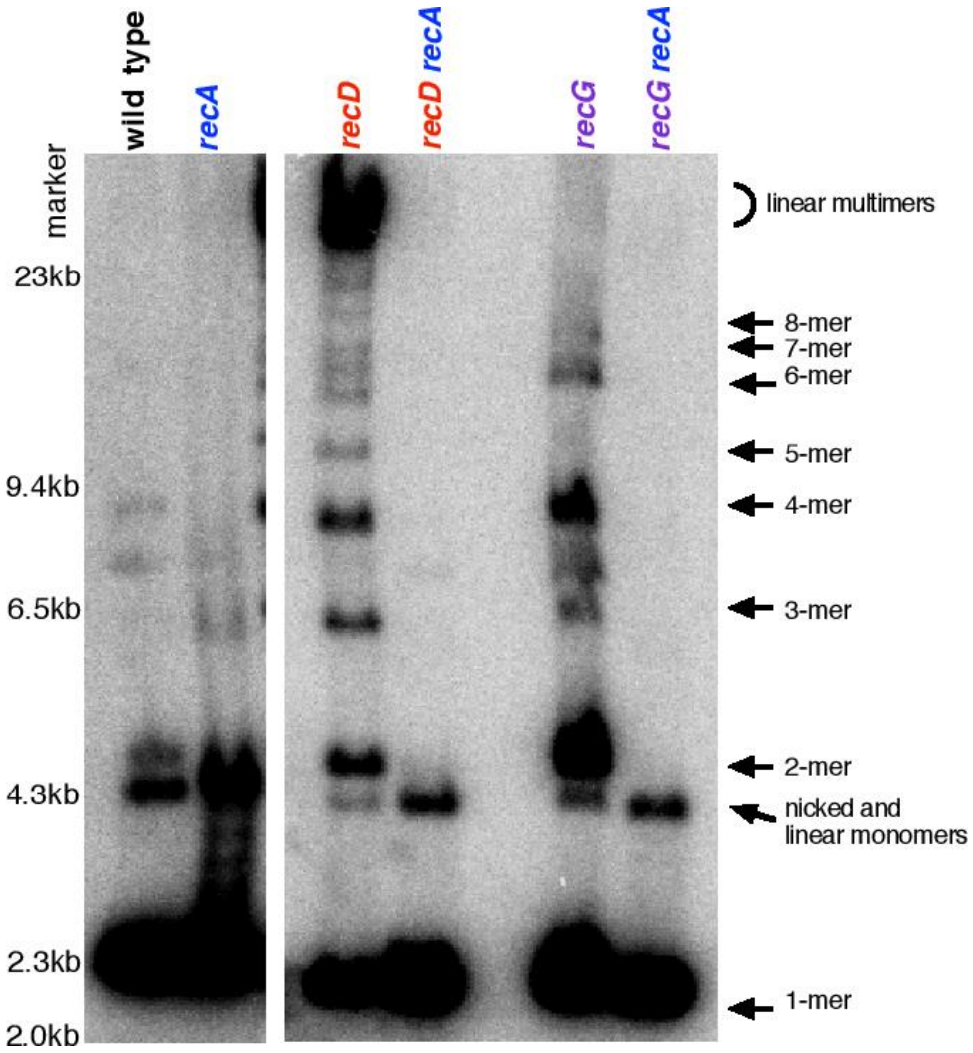

Figure 2.4. DNA ends from unresolved completion events lead to distinct recombination-dependent, over-replication intermediates on plasmids in *recD* and *recG* mutants. *recD* mutants accumulate long linear multimers, as well as both odd- and even-numbered multimeric circles. In *recG* mutants, the over-replicated products consist of predominantly even-numbered, multimeric circles. In both *recD* and *recG* mutants, the illegitimate re-initiations of replication depend on RecA. Cultures and DNA were prepared and analyzed as in *figure 1c*.

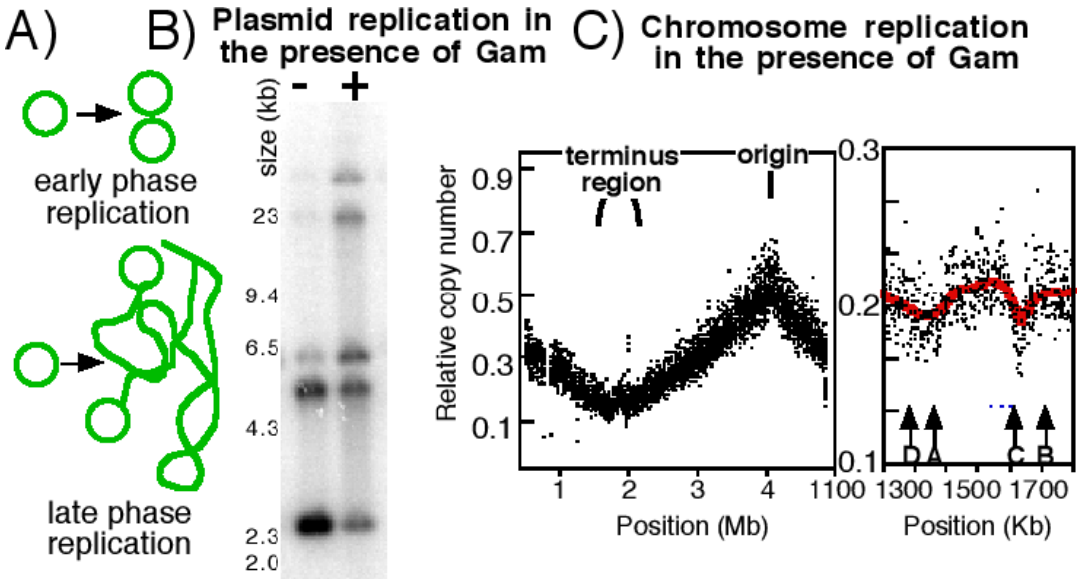

Figure 2.5. Induction of the bacteriophage *gam* gene inactivates the cellular mechanism that limits replication to the doubling point. A) Lambda late phase replication requires *gam* induction. B) *gam* induction leads to over replication on plasmids and C) the chromosome. Cultures containing a plasmid with an arabinose-inducible *gam* gene were grown with 0.4% glucose (-) or 0.4% arabinose (+) and prepared as in *figure 1b* and *c*.

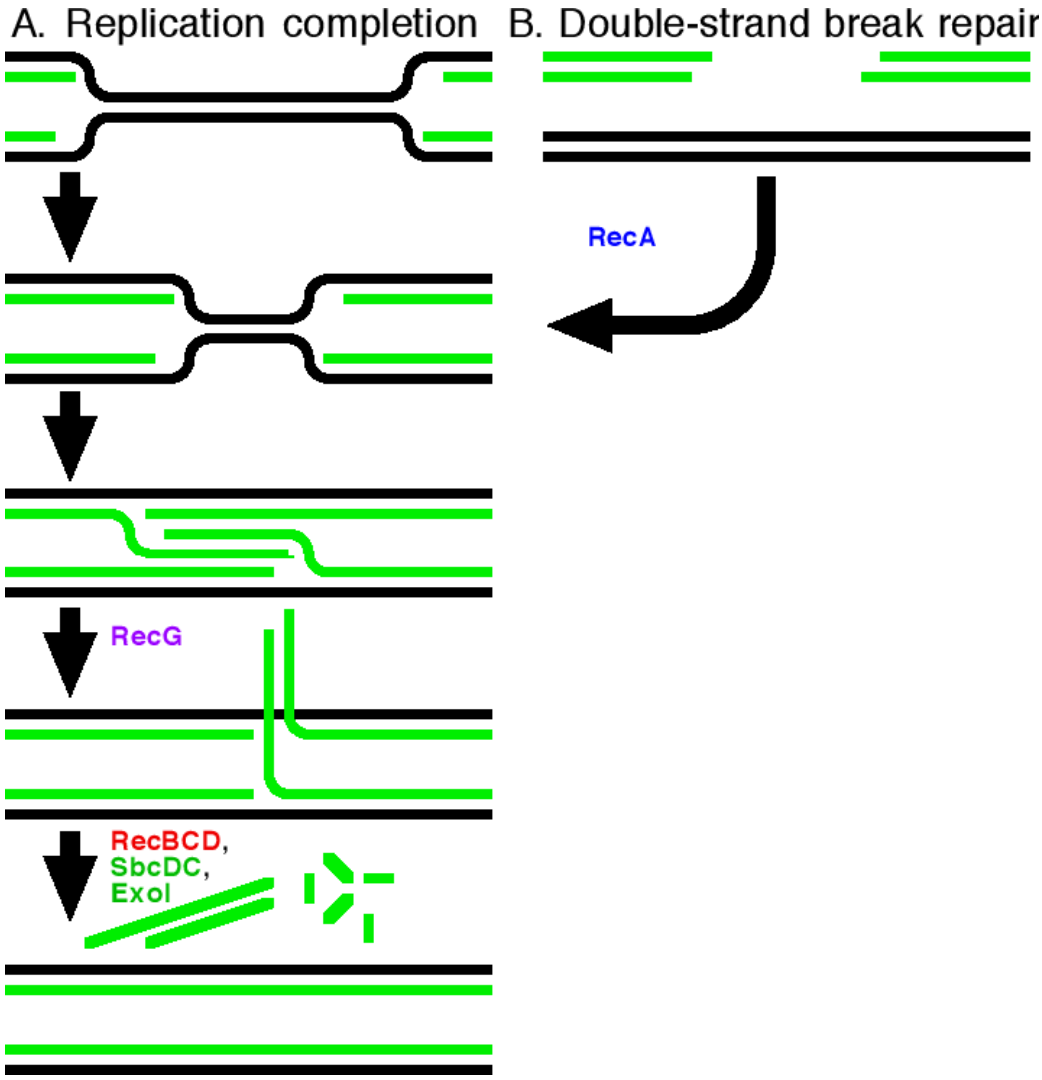

Figure 2.6. Model for the completion of replication and its relationship to homologous double-strand break repair. A) Converging replication forks pass each other, leading to a transient over-replicated intermediate. RecBCD promotes the degradation and resolution of the over-replicated regions at the doubling point. RecG facilitates unwinding of the over-replicated intermediates to reduce re-initiation events and illegitimate replication. SbcDC and ExoI also participate in the degradation of the over-replicated regions to limit replication. B) RecA initiates homologous double-strand break repair by pairing DNA ends with a homologous double-stranded template, generating an intermediate that can be repaired by completing the replication of the intervening sequences.

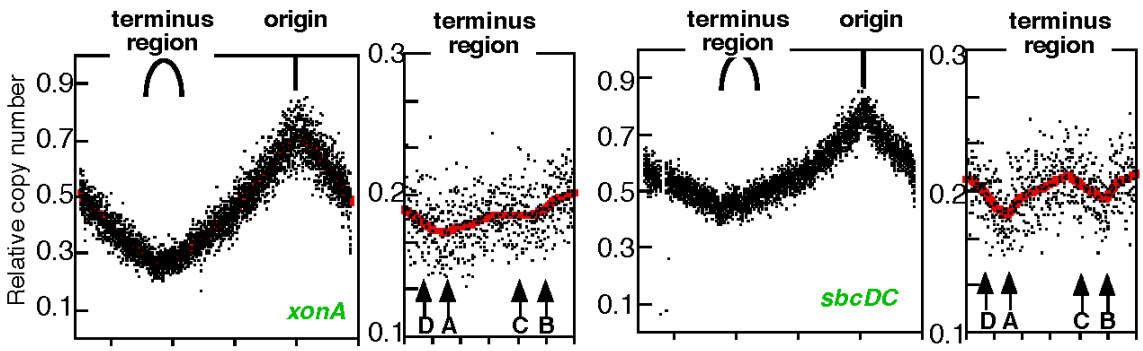

Figure 2.7. Over-replication of the terminus region is less pronounced in *xonA* and *sbcCD* mutants relative to the *xonA sbcDC* double mutant. Genomic DNA from replicating cultures was purified, fragmented, and profiled using high-throughput sequencing. Sequence read frequencies, normalized to stationary phase cells, are plotted relative to their position on the genome. The terminus region of the chromosome, containing *terD*, *A*, *C*, and *B*, is shown next to each plot. An 8kb floating average of the sequence frequency is plotted in red.

References

1. O'Donnell M, Langston L, Stillman B (2013) Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb Perspect Biol* 5
2. Costa A, Hood IV, Berger JM (2013) Mechanisms for initiating cellular DNA replication. *Annu Rev Biochem* 82:25–54.
3. Aladjem MI (2007) Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* 8:588–600.
4. Wu PY, Nurse P (2009) Establishing the program of origin firing during S phase in fission Yeast. *Cell* 136:852–864.
5. Dubey DD, Zhu J, Carlson DL, Sharma K, Huberman JA (1994) Three ARS elements contribute to the ura4 replication origin region in the fission yeast, Schizosaccharomyces pombe. *EMBO J* 13:3638–3647.
6. Wang X, Lesterlin C, Reyes-Lamothe R, Ball G, Sherratt DJ (2011) Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proc Natl Acad Sci U S A* 108:E243–50.
7. Hawkins M, Malla S, Blythe MJ, Nieduszynski CA, Allers T (2013) Accelerated growth in the absence of DNA replication origins. *Nature* 503:544–547.
8. Hiasa H, Marians KJ (1994) Tus prevents overreplication of oriC plasmid DNA. *J Biol Chem* 269:26959–26968.
9. de Massy B, Fayet O, Kogoma T (1984) Multiple origin usage for DNA replication in sdrA(rnh) mutants of *Escherichia coli* K-12. Initiation in the absence of oriC. *J Mol Biol* 178:227–36.

10. Asai T, Kogoma T (1994) D-loops and R-loops: alternative mechanisms for the initiation of chromosome replication in *Escherichia coli*. *J Bacteriol* 176:1807–1812.
11. Magee TR, Asai T, Malka D, Kogoma T (1992) DNA damage-inducible origins of DNA replication in *Escherichia coli*. *Embo J* 11:4219–4225.
12. Bhatia V, Barroso SI, Garcia-Rubio ML, Tumini E, Herrera-Moyano E, Aguilera A (2014) BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* 511:362–365.
13. Hamperl S, Cimprich KA (2014) The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst)* 19:84–94.
14. Donnianni RA, Symington LS (2013) Break-induced replication occurs by conservative DNA synthesis. *Proc Natl Acad Sci U S A* 110:13475–13480.
15. Hill TM (1992) Arrest of bacterial DNA replication. *Annu Rev Microbiol* 46:603–633.
16. Hill TM, Tecklenburg ML, Pelletier AJ, Kuempel PL (1989) tus, the trans-acting gene required for termination of DNA replication in *Escherichia coli*, encodes a DNA-binding protein. *Proc Natl Acad Sci U S A* 86:1593–157.
17. Kobayashi T, Hidaka M, Horiuchi T (1989) Evidence of a *ter* specific binding protein essential for the termination reaction of DNA replication in *Escherichia coli*. *Embo J* 8:2435–241.

18. Roecklein B, Pelletier A, Kuempel P (1991) The tus gene of *Escherichia coli*: autoregulation, analysis of flanking sequences and identification of a complementary system in Salmonella typhimurium. *Res Microbiol* 142:169–75.

19. Duggin IG, Wake RG, Bell SD, Hill TM (2008) The replication fork trap and termination of chromosome replication. *Mol Microbiol* 70:1323–1333.

20. Kohiyama M, Cousin D, Ryter A, Jacob F (1966) [Thermosensitive mutants of *Escherichia coli* K 12. I. Isolation and rapid characterization]. *Ann Inst Pasteur (Paris)* 110:465–486.

21. Wechsler JA, Gross JD (1971) *Escherichia coli* mutants temperature-sensitive for DNA synthesis. *Mol Gen Genet* 113:273–284.

22. Maine GT, Sinha P, Tye BK (1984) Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics* 106:365–385.

23. Smith GR (1988) Homologous recombination in procaryotes [published erratum appears in Microbiol Rev 1988 Jun;52(2):304]. *Microbiol Rev* 52:1–28.

24. Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol Rev* 58:401–465.

25. Templin A, Kushner SR, Clark AJ (1972) Genetic analysis of mutations indirectly supressing recB and recC mutations. *Genetics* 72:205–215.

26. Capaldo FN, Barbour SD (1975) The role of the rec genes in the viability of *Escherichia coli* K12. *Basic Life Sci* 5A:405–418.

27. Capaldo-Kimball F, Barbour SD (1971) Involvement of recombination genes in growth and viability of *Escherichia coli* K-12. *J Bacteriol* 106:204–212.

28. Niki H, Ichinose C, Ogura T, Mori H, Morita M, Hasegawa M, Kusukawa N, Hiraga S (1988) Chromosomal genes essential for stable maintenance of the mini-F plasmid in *Escherichia coli*. *J Bacteriol* 170:5272–5278.

29. Biek DP, Cohen SN (1986) Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in *Escherichia coli*. *J Bacteriol* 167:594–603.

30. Pennington JM, Rosenberg SM (2007) Spontaneous DNA breakage in single living *Escherichia coli* cells. *Nat Genet* 39:797–802.

31. Courcelle J, Carswell-Crumpton C, Hanawalt PC (1997) recF and recR are required for the resumption of replication at DNA replication forks in *Escherichia coli*. *Proc Natl Acad Sci U S A* 94:3714–379.

32. Jones C, Holland IB (1985) Role of the SulB (FtsZ) protein in division inhibition during the SOS response in *Escherichia coli*: FtsZ stabilizes the inhibitor SulA in maxicells. *Proc Natl Acad Sci U S A* 82:6045–609.

33. Horiuchi T, Nishitani H, Kobayashi T (1995) A new type of *E. coli* recombinational hotspot which requires for activity both DNA replication termination events and the Chi sequence. *Adv Biophys* 31:133–147.

34. Horiuchi T, Fujimura Y, Nishitani H, Kobayashi T, Hidaka M (1994) The DNA replication fork blocked at the *Ter* site may be an entrance for the RecBCD enzyme into duplex DNA. *J Bacteriol* 176:4656–4663.

35. Lloyd RG, Buckman C (1985) Identification and genetic analysis of sbcC mutations in commonly used recBC sbcB strains of *Escherichia coli* K-12. *J Bacteriol* 164:836–44.

36. Silberstein Z, Cohen A (1987) Synthesis of linear multimers of OriC and pBR322 derivatives in *Escherichia coli* K-12: role of recombination and replication functions. *J Bacteriol* 169:3131–3137.

37. Sharples GJ, Leach DR (1995) Structural and functional similarities between the SbcCD proteins of *Escherichia coli* and the RAD50 and MRE11 (RAD32) recombination and repair proteins of yeast. *Mol Microbiol* 17:1215–1217.

38. Rudolph CJ, Upton AL, Stockum A, Nieduszynski CA, Lloyd RG (2013) Avoiding chromosome pathology when replication forks collide. *Nature* 500:608–611.

39. Salzman LA, Weissbach A (1967) Formation of intermediates in the replication of phage lambda DNA. *J Mol Biol* 28:53–70.

40. Sakaki Y, Karu AE, Linn S, Echols H (1973) Purification and properties of the gamma-protein specified by bacteriophage lambda: an inhibitor of the host RecBC recombination enzyme. *Proc Natl Acad Sci U S A* 70:2215–2219.

41. Kulkarni SK, Stahl FW (1989) Interaction between the sbcC gene of *Escherichia coli* and the gam gene of phage lambda. *Genetics* 123:249–253.

42. Lloyd RG, Thomas A (1984) A molecular model for conjugational recombination in *Escherichia coli* K12. *Mol Gen Genet* 197:328–336.

43. Lloyd RG, Evans NP, Buckman C (1987) Formation of recombinant lacZ+ DNA in conjugational crosses with a recB mutant of *Escherichia coli* K12 depends on recF, recJ, and recO. *Mol Gen Genet* 209:135–141.

44. de Massy B, Patte J, Louarn JM, Bouche JP (1984) oriX: a new replication origin in *E. coli*. *Cell* 36:221–27.

45. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829.

46. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A (2011) CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* 6:11.

47. Langmead B (2010) Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* Chapter 11:Unit 11.7.

48. Muller CA, Nieduszynski CA (2012) Conservation of replication timing reveals global and local regulation of replication origin activity. *Genome Res* 22:1953–1962.

49. Mellon I, Hanawalt PC (1989) Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 342:95–98.

50. Courcelle J, Crowley DJ, Hanawalt PC (1999) Recovery of DNA replication in UV-irradiated *Escherichia coli* requires both excision repair and recF protein function. *J Bacteriol* 181:916–22.

51. Donaldson JR, Courcelle CT, Courcelle J (2004) RuvAB and RecG Are Not Essential for the Recovery of DNA Synthesis Following UV-Induced DNA Damage in *Escherichia coli*. *Genetics* 166:1631–1640.

52. Courcelle J, Hanawalt PC (1999) RecQ and RecJ process blocked replication forks prior to the resumption of replication in UV-irradiated *Escherichia coli*. *Mol Gen Genet* 262:543–51.

53. Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* 97:5978–5983.

54. Poteete AR, Rosadini C, St Pierre C (2006) Gentamicin and other cassettes for chromosomal gene replacement in *Escherichia coli*. *Biotechniques* 41:261–2, 264.

55. Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW, Crosa JH, Falkow S (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2:95–113.

Chapter 3

SbcC-SbcD and ExoI process convergent forks to complete chromosomal replication

*****This chapter has been published

Wendel BM, Cole JM, Courcelle CT, Courcelle J (2017) SbcC-SbcD and ExoI process convergent forks to complete chromosome replication. *Proceedings of the National Academy of Sciences*:201715960.

Abstract

SbcC-SbcD ExoI are the bacterial homologs of Mre11-Rad50 and CtIP, a nuclease complex essential for genome stability, normal development, and viability in mammals. In vitro, these enzymes function together to degrade long DNA palindromic structures. However, long DNA palindromes are generally not found in bacterial or human genomes, leaving the cellular substrates and function of these enzymes unknown. Here, we show that during the completion of DNA replication, convergent replication forks form palindrome-like intermediates that require nucleolytic processing by SbcC-SbcD and ExoI before chromosome replication can be completed. Inactivation of these nucleases prevents completion from occurring, and under these conditions, cells maintain viability by shunting the reaction through an aberrant recombinational pathway. The results identify replication completion as an event critical to maintain genome integrity and cell viability, demonstrate SbcC-SbcD-ExoI act prior to RecBCD

and is required to initiate the completion reaction, and reveal how defects in completion result in genomic instability.

Significance statement

SbcC-SbcD and ExoI belong to a class of highly conserved nucleases that are critical to genome stability, but whose cellular function remains poorly understood. In humans, the homologs are essential for viability and normal development, and lead to severe developmental abnormalities and cancer predisposition when mutated. Here we show that these enzymes process DNA intermediates at sites where replication forks converge and are required for chromosome replication to complete normally. Cells lacking these gene products are unable to complete replication normally and rely on an aberrant recombinational mechanism to maintain viability that leads to genomic instability and amplifications at these sites similar to that seen in human cancers where these genes have been inactivated.

Introduction

Accurate replication of the genome requires that cells tightly regulate the processes of initiation, elongation, and completion to ensure that each daughter cell inherits an identical copy of the genetic information. While the mechanisms regulating initiation and elongation have been extensively characterized (1, 2), the process of how DNA replication is completed has, until recently, remained unknown. Accurately completing replication necessitates that cells recognize replicated regions and join the strands of converging replication forks at the precise point where all sequences have doubled. This event occurs thousands of times per division along the chromosomes of human cells and therefore must be remarkably efficient. Failure of even a single completion event would be expected to result in duplications, deletions, rearrangements, or cell death depending upon how the convergent strands are processed. Cells devote a large number of enzymes to maintain the fidelity of both replication initiation and elongation, and recent studies have shown that this final step, required to maintain genomic integrity, similarly requires enzymatic control and is tightly regulated (3, 4).

The completion step of DNA replication has been challenging to study in eukaryotic cells, in part because multiple origins are utilized with varying efficiencies and timing, making the regions where forks meet highly variable between cells and cell cycles (5, 6). By comparison, *Escherichia coli* is well-suited to dissect this fundamental aspect of cellular metabolism since the event can be localized to a single ~400 kb region of the chromosome, opposite to its bidirectional origin of replication (7). This region is flanked by *ter* sequences which bind the protein Tus, blocking

replication forks in an orientation specific manner (8). Although *ter* ensures that completion events occur within this region, they do not appear to be directly involved in the reaction, as chromosomes and plasmids lacking *ter* replicate normally and are stably maintained in the cell (9, 10).

In *E. coli*, the completion of DNA replication requires the RecBCD helicase-nuclease complex and occurs through a reaction that does not involve recombination or RecA (3). In the absence of RecBCD, cultures either fail to replicate, or fail to maintain the chromosome region where replication forks converge, leading to large segments of the chromosome remaining unreplicated in this region. The inability of *recBCD* mutants to replicate or maintain this region of the chromosome severely compromises growth and viability in these cultures (3). *In vitro*, RecB and -C interact with RecD to form a dual helicase-nuclease complex that unwinds and degrades double-strand DNA ends (11, 12). The enzymatic complex is capable of rapidly translocating and degrading DNA for tens to hundreds of kilobases at a rate of ~1000 bp/s, a processivity and rate that approaches that of the replisome's ability to synthesize these sequences (13–15). Loss of RecB or -C inactivates the complex, whereas loss of RecD inactivates the nuclease, but retains helicase activity (16–18). How RecBCD promotes the completion of DNA replication, and what intermediate substrates are involved in this reaction, remain to be characterized.

Other enzymes that compromise the ability to complete replication have also been identified and include SbcC-Sbc and ExoI. In the absence of these enzymes, the region where replication forks converge is over-replicated or amplified (3, 19). Curiously, although chromosome replication is clearly abnormal in these mutants, the

over-represented, amplification regions of the chromosome do not appear to compromise growth and viability of cells in culture.

SbcC-SbcD and ExoI belong to a class of nucleases that are critical to genome stability but whose cellular function remains poorly understood. They are the bacterial orthologs of Mre11-Rad50 in humans and yeast (20, 21). They are both structurally and functionally conserved in all three domains of life (22, 23). All SbcC-SbcD orthologs form a heterotetrameric complex with a remarkably unique architecture, containing two long coiled-coil regions that extend up to 60nm within the SbcC (Rad50) dimer where the ATPase and SbcD (Mre11) interaction site reside (22). When bound with the SbcD dimer, these complexes contain a prominent endonuclease activity that incises DNA hairpins, as well as single-strand endonuclease and double-strand 3'-5' exonuclase activities (24). Human CtIP, yeast Sae2, and *E. coli* ExoI appear to be conserved at the functional level. Inactiviation of these nucaleases enhances the palindrome amplifications and genetic instabilities that arise in mutants lacking SbcC-D or Rad50-Mre11 (25–27). In humans, these genes are essential for viability and normal development, whereas hypomorphic mutations lead to severe developmental abnormalities and cancer predisposition (28–31). Mutations render cells in culture hypersensitive to double strand breaks and perturb the normal progression through the cell cycle (29, 30, 32–34). These phenotypes are often attributed to defective repair of double strand breaks, and the enzymes are most commonly proposed to function as initiators of recombinational repair, processing DNA ends before the Rad51 recombinase can load and facilitate recombination (35, 36). However, several enigmatic phenotypes of these enzymes don't fit this model.

Rad51 foci form normally in the absence of Mre11 or Rad50 (37–39), implying that processing by these enzymes is not required to initiate recombinational Rad51 filaments. Curiously, the preferred substrate of the *E. coli*, yeast, and human enzymes is long hairpin or palindromic DNA structures (40–42). Inactivation of Mre11 and CtIP *in vivo* leads to widespread genomic instabilities that include the appearance of palindromic duplications (25, 43). Similarly in *E. coli*, palindromic sequences can only be maintained in mutants lacking SbcCD ExoI (44, 45), indicating that, somehow, palindromic sequences are relevant to its *in vivo* function. Yet, long palindromic sequences are not encoded in eukaryotic or bacterial genomes, making it unclear why palindrome-specific nucleases would be essential in mammalian cells or how palindromic duplications would rapidly appear in their absence. Thus, the substrates and essential function that these enzymes have in the cell has remained an elusive question.

In this study, we show that SbcC-SbcD and ExoI are required to initiate the completion of DNA replication on the chromosome, and act at a step prior to RecBCD in the process. The enzymes process a 'palindrome-like' intermediate that forms at sites where replication forks converge. SbcCD and ExoI processing likely creates a substrate that allows RecBCD to bind and promote joining of convergent fork strands, either directly or indirectly. In the absence of SbcCD and ExoI, these structural intermediates persist and the normal completion of replication is prevented. Furthermore, we show that under these conditions, cells maintain viability by processing these intermediates through an aberrant recombinational mechanism that accounts for the widespread genomic instabilities and amplifications

that are observed in mutants lacking these nucleases.

Materials and methods

Bacteria. All strains used in this work are derived from SR108, a *thyA deoC* derivative of W3110 and listed in table 3.1.

Growth Rates. Equal numbers of viable cells were grown in 0.1-mL cultures in LB medium supplemented with 10µg/mL thymine (LBthy), at 37°C with agitation in a 96-well microtiter dish. Absorbance at 630 nm was measured over time using a BIO-Whittaker ELx808 plate reader **(3)**.

Plasmid Stability. Cultures containing the plasmid pBR322 were grown for 30 generations in LBthy medium at 37°C with aeration. Ten-microliter aliquots of serial 10-fold dilutions were then spotted on LBthy plates in the presence or absence of 100 µg/mL ampicillin. Viable colonies were counted following overnight incubation at 37°C (3).

Total Genomic DNA Extraction. 0.75 mL cultures grown in LBthy medium supplemented with 100 µg/ml ampicillin were taken and placed into 0.75 Ml of cold 2× NET (100 mM NaCl, 10 mM Tris, pH 8.0, 10 mM EDTA). Each sample was pelleted, resuspended in 140 µl of 1 mg/mL lysozyme and 0.2 mg/mL RNaseA in TE (10 mM Tris, pH 8.0, 1 mM EDTA), and lysed at 37 °C for 30 min. At this time, proteinase K

(10 µl, 10 mg/ml) and Sarkosyl [10 µl, 20% (wt/wt)] were added and incubated at 37 °C for 30 min. Samples were then extracted with 4 vol of phenol/chloroform (1/1) and dialyzed for 1 h on 47 mm Whatman 0.05-µm pore disks (Whatman #VMWP04700) floating on a 250-mL beaker of TE(1 mM Tris, pH 8.0, 1 mM EDTA) (3).

Southern Analysis of Plasmid Replication Intermediates. Total genomic DNA samples were digested with Sac II (New England Biolabs), pBR322 lacks Sac II restriction sites, and extracted with chloroform. Equal volumes were loaded onto a 0.5% and 1.0% agarose gel containing 0.5× TBE (220 mM Tris, 180 mM Borate, 5 mM EDTA, pH 8.3) and electrophoresed at 1 and 2.5 V/cm respectively. Gels were transferred to Hybond N+ nylon membranes and probed with pBR322 that had been labeled with 32P by nick translation according to the protocol supplied by Roche using [α-32P]dCTP (PerkinElmer). Radioactivity was visualized using a Storm 840 and its associated ImageQuant Software (Molecular Dynamics) (3).

Sequencing, Assembly, and Replication Profiling. Fresh overnight cultures were diluted 1:250 in LBthy media and grown at 37 °C with aeration to an OD600 of 0.4. Stationary-phase cultures were grown for 36h. Total genomic DNA was prepared from cultures. Sequencing of the genomic DNA samples was performed using NexteraXT and Illumina HiSeq2000 (Illumina) using single-end, 51-bp, bar-coded reads according to the manufacturer's instructions. SR108 parent sequence was determined using *breseq* to identify structural variations between SR108 and its W3110 parent genome, and differences were hand annotated to generate the SR108 reference genome (46).

The original Illumina sequence reads for all subsequent strains were then aligned using Bowtie 1.0.0 (47), using SR108 as reference. Aligned reads were then analyzed for nucleotide frequencies at each position, and the copy number of sequences per kilobase was determined using custom Perl scripts. Copy number values were normalized to those of stationary phase cells to determine replication-specific copy number frequencies and eliminate any sequencing bias. These relative copy number values were then plotted against their location along the genome to generate replication profiles for each strain of interest.

Results

SbcC-SbcD and ExoI are required for replication to complete normally, and act upstream of RecBCD in the reaction

The ability to complete chromosome replication can be observed by profiling the copy number of DNA sequences across the genome of replicating cultures. In this approach, genomic DNA is purified from replicating cultures, fragmented, and sequenced. The replication profile is then determined by counting sequences that align to each segment of the chromosome (Figure 3.1a). In wild type cultures, sequences surrounding the bidirectional origin replicate first, and are observed at higher frequencies relative to chromosome regions that replicate later. Overall, sequence frequencies decrease inversely with their distance from the origin, until reaching the terminus region where replication forks converge and replication completes ((3) and

Figure 3.1B). In *recBCD* mutants, a dramatic loss of sequences is observed in the region where replication forks converge (Fig3.1C). An identical result is observed in *recBC* mutants alone (3). The copy number of sequences in the terminus region is reduced ~two fold relative to wild type cultures. Assuming that more than half of all sequence reads correspond to parental DNA strands, one can infer that nearly all cells in the population have difficulty replicating or maintaining this region of the chromosome. This severely impairs growth and viability in *recBC* cultures (3). The slope of the progressing replication forks away from the origin is similar in both *recBCD* and wild type profiles (-30.0 vs. -28.8 normalized reads/Mb, respectively) before reaching the terminus region, where the *recBCD* slope rapidly deteriorates, relative to wild type cultures (-133.8 vs. -13.9 normalized reads/Mb, respectively). The observation argues against the idea that overall fork progression is impaired or impeded in *recBCD* mutants and indicates that the replication defect localizes specifically to the region where replication forks converge in these mutants. In contrast, completion has previously been shown to occur normally in *recA* mutants (3, 4), demonstrating that the RecBCD-mediated reaction occurs independently of homologous recombination or double strand repair and likely represents an intramolecular reaction.

The inability of *recBCD* mutants to maintain the terminus region could arise either because RecBCD is required for replication forks to reach the terminus, or because RecBCD is somehow required to allow DNA ends from convergent replication forks to be joined. In the latter possibility, persistent DNA ends from convergent replication forks would remain susceptible to nucleolytic attack, leading to

extensive degradation and loss of DNA in the region where forks converge.

To differentiate between these two general mechanisms, we examined the replication profiles of *recBC* mutants that also lacked SbcCD and ExoI nucleases. ExoI contains a processive 3'-5' single-stranded exonuclease (48, 49) while SbcCD is a double-stranded exonuclease that also has endonucleolytic activity at DNA palindromes (24, 44). We reasoned that if RecBCD facilitates replication forks reaching the terminus, then the absence of exonucleases would have little effect on the depleted terminus region. Alternatively, if RecBCD action is required before DNA ends from convergent forks can be joined, then the absence of exonucleases may reduce or prevent the degradation occurring, and partially restore the terminus region. We initially focused upon SbcCD and ExoI as candidates because inactivation of these gene products has been shown to suppress recombination defects in *recBC* mutants during sexual processes, such as conjugation, transformation, or transduction (50, 51). In addition, SbcCD and Exo I appear to participate in the completion reaction, as mutants lacking these gene products exhibit an increased copy number of sequences in the terminus region (Figure 3.1b and (3)).

We found that inactivation of SbcCD and ExoI was sufficient to restore the DNA to the terminus region of *recBC* mutants. Surprisingly, loss of these nucleases resulted in over-replication of the terminus region, similar to that seen in *sbcCD xonA* mutants alone (Figure 3.1e). The observation demonstrates that replication forks reach the terminus region normally in the absence of RecBCD and implies that convergent fork ends remain unjoined and susceptible to degradation in *recBCD* mutants. Notably, over-replication in *sbcCD xonA* is phenotypically dominant to degradation in

recBC mutants, indicating that converging replication forks bypass each other to create an over-replicated intermediate that must be processed by SbcCD-ExoI before RecBCD can facilitate resection and joining during the completion reaction.

The sbcCD and xonA mutations appear additive in their effect on overreplication in the terminus (Figure 3.2). The observation could suggest that processing of the overreplicated region occurs by each of these enzymes, but that a double-stranded end required for RecBCD loading is only generated after both processing events have occurred.

Maintaining the region where replication forks converge correlates with the growth and viability of cells in culture.

The inability of *recBC* mutants to maintain the terminus region of the chromosome correlates with the slow growth and reduced viability of these cultures relative to wild type cells. Curiously, although the completion reaction is clearly abnormal in mutants lacking SbcCD and ExoI, the excess copy number or amplification of the terminus region does not impair the growth or viability of these mutants (Figure 3.1b, c, and d). We noted that inactivation of SbcCD and ExoI, which prevents degradation of the terminus region in *recBC* mutants, also restores the growth and viability of these cultures (Figure 3. 1b, c, and d). The observation further supports the interpretation that the impaired growth and viability of *recBC* mutants is caused by the inability to maintain the terminus region on the chromosome. Importantly however, although SbcCD and ExoI inactivation restores growth and viability in *recBC*

mutants, chromosomal abnormalities and amplifications persist in the region where forks converge, indicating that completion in the absence of these nucleases occurs through an aberrant, alternative pathway.

Abnormalities in completing replication can also be observed on plasmids replicating in these mutants.

In *recBC* mutants, plasmids have a reduced copy number relative to wild type cultures, suggesting that they replicate with a lower frequency of success. In *recD* mutants, which lack nuclease but retain helicase activity, plasmids continue to replicate past the doubling point, producing large quantities of both odd- and even-numbered multimeric circles as well as long linear multimers. These plasmids are not maintained and rapidly lost in culture (3, 52, 53). By contrast, in the absence of RecA, which is required for all recombination and double strand break repair, plasmids remain stable and replicate normally. The observations argue that the plasmid instabilities in *recBC* and *recD* mutants are associated with an impaired ability to complete replication, rather than defects in recombination or double strand break repair. We next examined how SbcCD and ExoI affected plasmid stability in *recBC* and *recD* mutants. We found that similar to what is observed on the chromosome, inactivation of SbcCD and ExoI alleviates the low copy number phenotype in *recBC* mutants and improves plasmid stability in both *recBC* and *recD* mutants (Figure 3.3a). Also similar to what is seen on the chromosome, although stability is restored, large amounts of multimeric amplification products are produced and persist in these populations, indicating that,

under these conditions, the plasmids are maintained through an alternate, aberrant mechanism (Figure 3.3b).

When normal completion is impaired, the ability to maintain cell growth and the region where forks converge becomes dependent on an aberrant recombinational mechanism.

To further investigate how cell viability and the terminus region are maintained under these aberrant conditions, we examined whether this alternative process depended on recombination. The completion of replication on the chromosome normally occurs through a mechanism that is independent of recombination and RecA (3). In mutants lacking RecA, the region where forks converge is maintained and cultures grow at rates comparable to wild type cultures (Figure 3.4). However, we observed that in mutants where the normal mechanism of completion is prevented, the cell's ability to maintain growth and the chromosome region where forks converge becomes dependent on recombination and RecA (Figure 3.4, Figure 3.5). In the absence of RecA, *sbcCD xonA* mutations no longer suppress the defects in *recBC* cultures, as the terminus region is not maintained, and growth is severely compromised. Similarly, in *sbcCD xonA recD* mutants, the ability to maintain the terminus region and grow depends entirely on RecA. Thus, the normal RecBCD-mediated completion reaction in the absence of recombination. However, when the normal completion reaction is prevented, the reaction is shunted through an aberrant recombination-dependent mechanism that prevents loss of this chromosome region but is associated with amplification and

genomic instabilities at these locations. Strains that lose the ability to maintain the region where replication forks converge are no longer able to grow and viability is severely reduced.

Inactivation of SbcCD and ExoI, in an otherwise wild type background, leads to amplification and genomic instability in the terminus region (Figure 3.1). To determine if the absence of these nucleases, by themselves, prevent replication from completing normally, we asked whether maintaining the terminus region and cell growth of *sbcCD xonA* mutants depended on recombination. We found that the absence of SbcCD ExoI, alone, prevented replication from completing normally and that in the absence of recombination, cells were unable to maintain the region where replication forks converge, and growth was severely impaired (Figure 3.6). Thus, in the absence of the nucleases that process the intermediate created by converging replication forks, normal completion of replication is prevented, and the reaction is shunted through an aberrant recombinational process that prevent loss of these chromosome regions but leads to genetic instability and amplifications at these sites.

Discussion

Taken together, the data presented here demonstrate that the completion of DNA replication occurs through a transient over-replication of the sequences where replication forks converge. This over-replication can be observed both in vivo and in vitro (3, 54). SbcCD and ExoI are required to process the intermediate produced at sites where replication forks converge before RecBCD can resect these regions and facilitate joining of convergent DNA ends (Figure 3.7a). In the absence of SbcCD-

ExoI, completion cannot occur and cell growth becomes dependent on an aberrant recombination mechanism to resolve these chromosomes that leads to genomic instability and amplification in the region. We additionally show that replication forks reach the terminus normally in the absence of RecBCD, but that joining of the nascent convergent forks fails to occur, leading to extensive degradation in these regions. SbcC-SbcD, and its human homolog, Mre11-Rad50, specifically incise long DNA palindromes both in vitro and in vivo (24, 25, 44, 55). Yet long palindromic sequences are not found in these genomes, and until now, identifying the cellular substrate or function of these enzymes has remained elusive. When considering the substrate created by the transient over-replication at sites where replication forks converge, it is apparent that intermediate bears a striking 'palindrome-like' structure that is known to be cleaved by SbcC-SbcD and Mre11-Rad50 in vitro (Figure 3.7b). Here, we show that these sequences are substrates for SbcCD and ExoI *in vivo*. Furthermore, SbcCD-ExoI processing of this intermediate is essential for the normal completion of replication to occur, and results in genomic instabilities and amplifications when processing is impaired or prevented.

Several observations suggest the conserved Mre11-Rad50 and CtIP function similarly in humans and other eukaryotes (20, 21). When processing of convergent replication fork intermediates is prevented in *E. coli*, abnormalities are observed on the chromosome. Although only one completion event occurs per cell cycle in *E. coli*, human cells require that this reaction occurs thousands of times along the chromosome each time the cell divides, and may explain why these genes are essential, and lead to severe developmental defects and cancer when their function is impaired (25, 29, 30).

Consistent with this interpretation, lethality in mutants with a compromised Mre11-Rad50-Nbs1 function arises not during but after replication through S-phase, due to a failure to resolve 'DNA bridges' that persist between chromosomes attempting to separate during mitosis (33). Based on the results presented here, we would propose that these bridges represent unresolved completion events.

We show that when processing of the convergent replication forks is prevented, over-replicated regions create homologous substrates that can alternatively be resolved through recombinational mechanisms (Figure 3.8). However, when completion occurs under these conditions, it produces chromosome instabilities and amplifications that are readily apparent in the replication profiles of mutants. In human cells, where completion events occur throughout the genome, this compromised fidelity would be expected to provide cancerous cells with a broad selective mechanism by which these cells may achieve a growth advantage through amplification. Direct and palindromic amplifications are correlated with tumor initiation, metastasis, and the development of drug resistance, sometimes in regions of known oncogenes (56–62). These mutational signatures match those appearing on the *sbcCD xonA* chromosome, and are similar to the widespread palindromic amplifications and gross chromosome rearrangements arising in yeast following inactivation of Mre11 and Sae2 (25, 43).

Finally, the aberrant, recombination-mediated mechanism that operates to maintain the chromosome in *E. coli* when the normal mechanism of completion is prevented has a number of similarities to mechanisms that maintain chromosomes in immortalized eukaryotic cells. Similar to bacteria, telomerase-negative immortalized cells maintain these regions are restored and maintained through a recombinational

process that is required for growth and viability in these populations, but is associated with reduced fidelity, chromosome rearrangements, and amplifications across the genome (25, 33, 43, 63–66). SbcC-SbcD, ExoI and Mre11-Rad50, CtIP are conserved at the structural, biochemical, and phenotypic levels, making it highly likely that the role of these bacterial enzymes in completing replication, reported here, is likely to be similar for the mammalian enzymes and account for their essential role in the cell.

Data deposition

Sequence data reported in this paper have been deposited in the Sequence Read Archive (SRA), www.ncbi.nlm.nih.gov/Traces/sra (accession #s SRP107355 and SRP047195).

Acknowledgments

This work was supported by the National Science Foundation, grant MCB0130486 and National Institute of Environmental Health Sciences grant R15 R15ES025953.

Table 3.1 Strains used in this study

Strain	Relevant Genotype	Source, Reference, and/or construction
SR108	λ - *thyA deo IN(rrnD-rrnE)*	(67)
HL922	SR108 *recB21C22 argA81::*Tn*10*	(68)
HL923	SR108 *recD1011 argA81::*Tn*10*	(68)
CL542	SR108 *recA::cam*	(69)
CL1056	SR108 D(*recC ptr recB recD*)::*cam*	P1 transduction of D(*recC ptr recB recD*)::*cam* from KM78 (gift from KC Murphy)
CL2357	SR108 *xonA::Cat300 sbcCD::Gm*	(3)
CL3539	SR108 *xonA::Cat300 sbcCD::Gm D(srlR-recA)306::Tn10*	P1 transduction of *D(srlR-recA)306::*Tn*10* from HL921 (68) into CL2359
CL851	SR108 *recB21C22 argA81::*Tn*10 recA::cam*	(69)
CL726	SR108 *recD1011 argA81::*Tn*10 recA::cam*	(69)
CL2542	SR108 *xonA::Cat300 sbcCD::Gm recB21C22 argA81::*Tn*10*	P1 transduction of *recB21C22 argA81::*Tn*10* from HL922 (68) into CL2357
CL2539	SR108 *xonA::Cat300 sbcCD::Gm recD1011 argA81::*Tn*10*	P1 transduction of *recD1011 argA81::*Tn*10* from HL923 (68) into CL2357
CL2575	SR108 *xonA::Cat300 sbcCD::Gm recB21C22 argA81 D(srlR-recA)306::*Tn*10*	CL2542 was first cured of Tn*10* (70). This was followed by P1 transduction of *D(srlR-recA)306::*Tn*10* from HL921 (69) into CL2542
CL2576	SR108 *xonA::Cat300 sbcCD::Gm recB21C22*	CL2542 was first cured of Tn*10* (70). This was followed by P1 transduction of *D(srlR-*

	*argA81 D(srlR-recA)306::*Tn*10*	*recA)306::*Tn*10* from HL921 (69) into CL2542 (separate isolate of 2575)
CL2577	SR108 *xonA::Cat300 sbcCD::Gm recD1011 argA81 D(srlR-recA)306::*Tn*10*	CL2539 was first cured of Tn*10* (70). This was followed by P1 transduction of *D(srlR-recA)306::*Tn*10* from HL921 (69) into CL2539.
HL1034	SR108 *xonA::Cat300*	(71)
CL2344	SR108 *sbcCD::Gm*	P1 transduction of *sbcCD::Gm* from KM137 (72) into SR108
CL718	SR108 *xonA::Cat300 D(srlR-recA)306::*Tn*10*	P1 transduction of *D(srlR-recA)306::*Tn*10* from HL921(68) into HL1034
CL3535	SR108 *sbcCD::Gm D(srlR-recA)306::*Tn*10*	P1 transduction of *D(srlR-recA)306::*Tn*10* from HL921(68) into CL2344

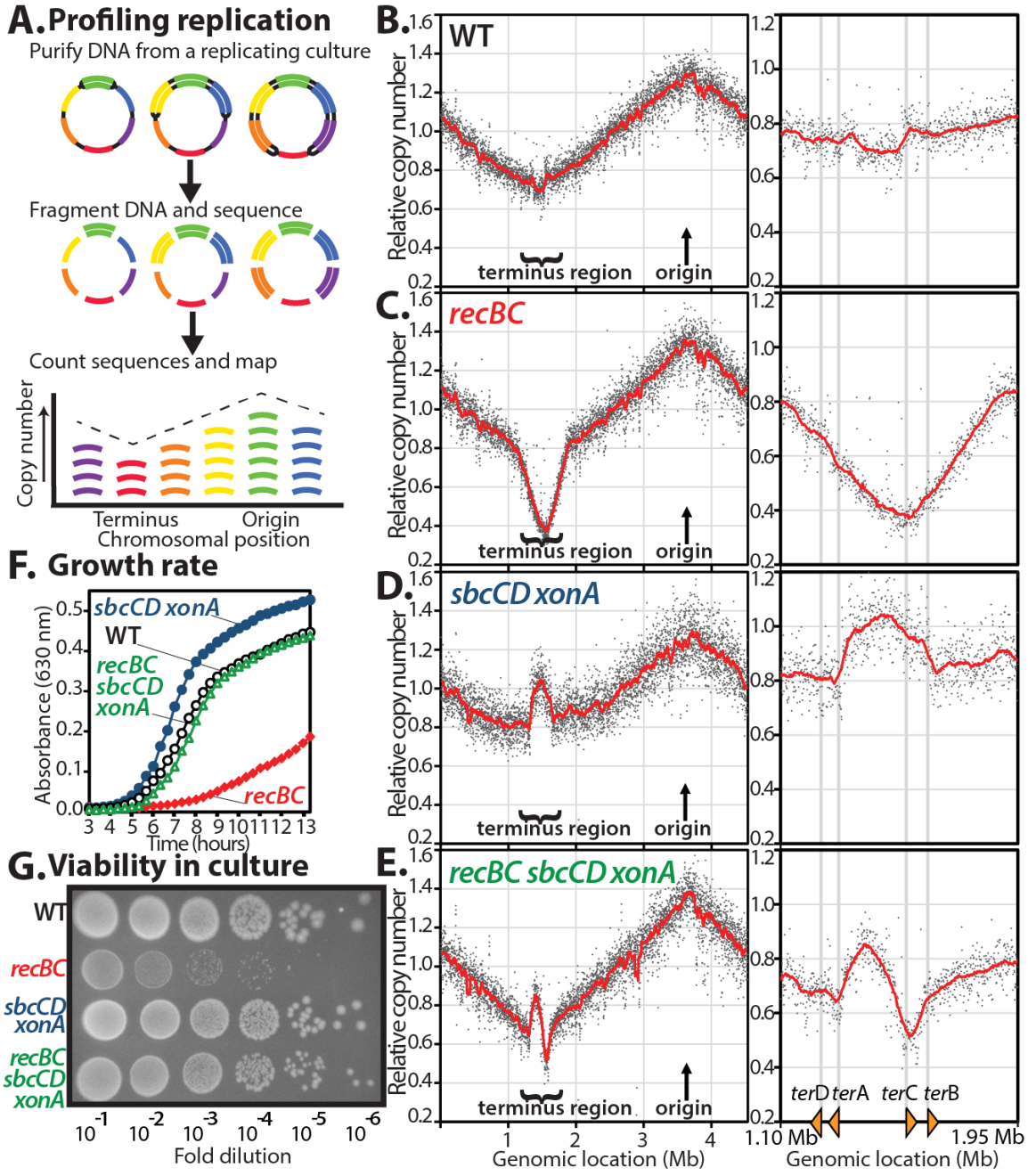
A. Profiling replication
Purify DNA from a replicating culture
Fragment DNA and sequence
Count sequences and map
Copy number
Terminus
Origin
Chromosomal position
B.
WT
C.
recBC
D.
sbcCD xonA
E.
recBC sbcCD xonA
Relative copy number
terminus region
origin
Genomic location (Mb)
terD terA terC terB
1.10 Mb
1.95 Mb
F. Growth rate
Absorbance (630 nm)
sbcCD xonA
WT
recBC sbcCD xonA
recBC
Time (hours)
G. Viability in culture
WT
recBC
sbcCD xonA
recBC sbcCD xonA
Fold dilution

Figure 3.1. Inactivation of SbcCD ExoI nucleases prevents processing of convergent replication forks and restores the terminus region in *recBC* mutants, indicating that SbcCD ExoI is required to process the over-replicated intermediate before RecBCD-mediated completion can occur. A) *Diagram of replication profile methodology*. B) *Inactivation of the SbcCD ExoI nucleases restores the region where replication forks converge in recBC mutants.* Genomic DNA from replicating cultures was purified, fragmented, and profiled using high- throughput sequencing. Normalized sequence read frequencies are plotted relative to genome position along with a 30kb floating average in red. The terminus region of the chromosome, containing *terD*, *-A*, *-C*, and *-B*, is shown next to each plot. *sbcCD xonA* read frequencies were replotted from (3) C) and D) *The ability to maintain the chromosome's terminus region correlates with growth and viability in culture*. Absorbance (630nm) of cultures grown at 37C is plotted over time (C). 10µL drops from 10fold serial dilutions of overnight cultures were plated. Colonies were observed following overnight 37C incubation (D).

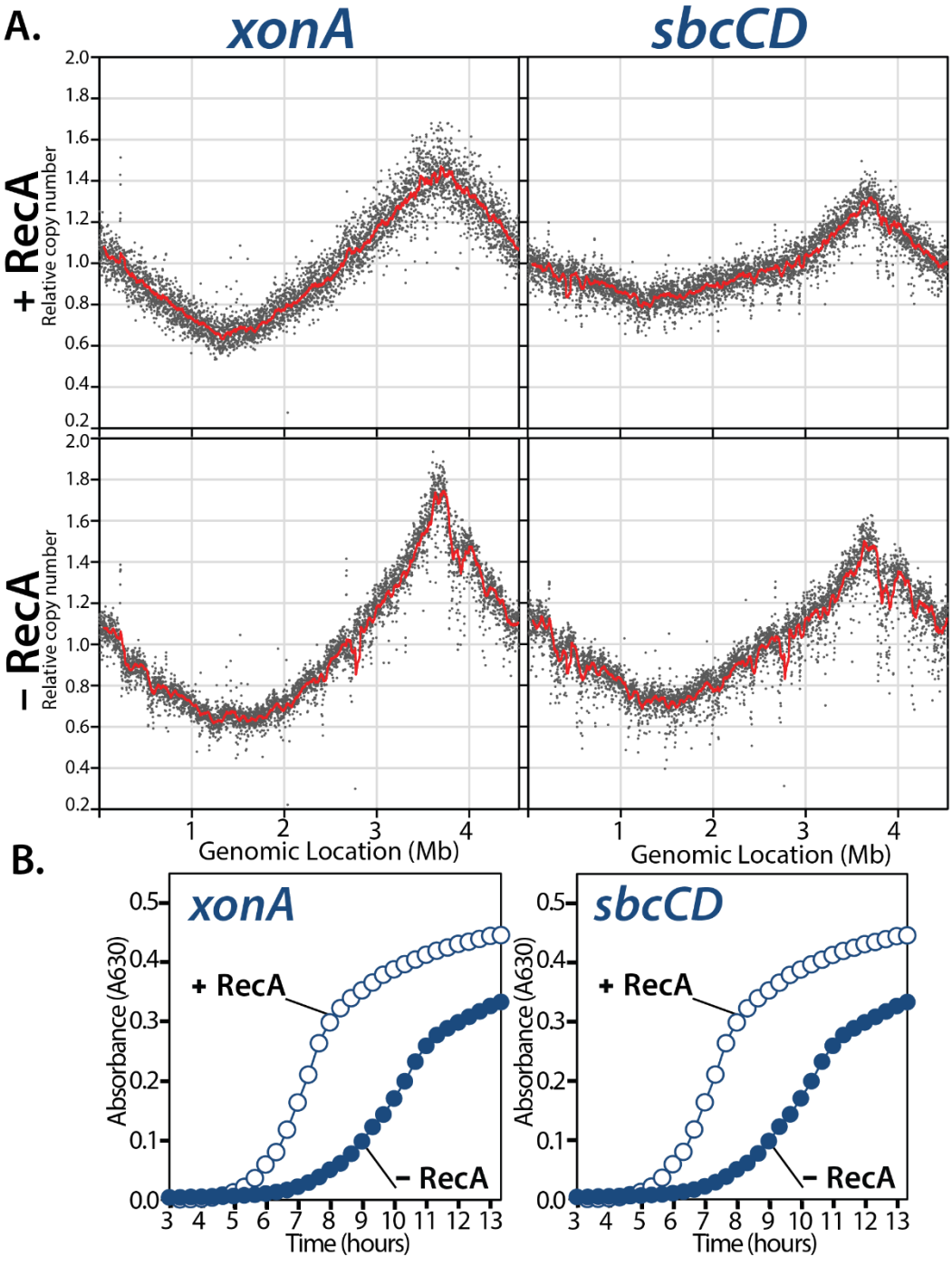
A.
xonA
sbcCD
+ RecA
– RecA
Relative copy number
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
1
2
3
4
Genomic Location (Mb)
B.
xonA
sbcCD
+ RecA
– RecA
Absorbance (A630)
0.5
0.4
0.3
0.2
0.1
0.0
3 4 5 6 7 8 9 10 11 12 13
Time (hours)

Figure 3.2. *xonA* and *sbcCD* mutations are additive in their effect on over-replication in the terminus and RecA-dependent growth. A) Replication profiles for *xonA* and *sbcCD* in the presence of absence of RecA are plotted as in *figure 3.1*. B) Growth rates for *xonA* and *sbcCD* in the presence or absence of RecA are plotted as in *figure 3.1*.

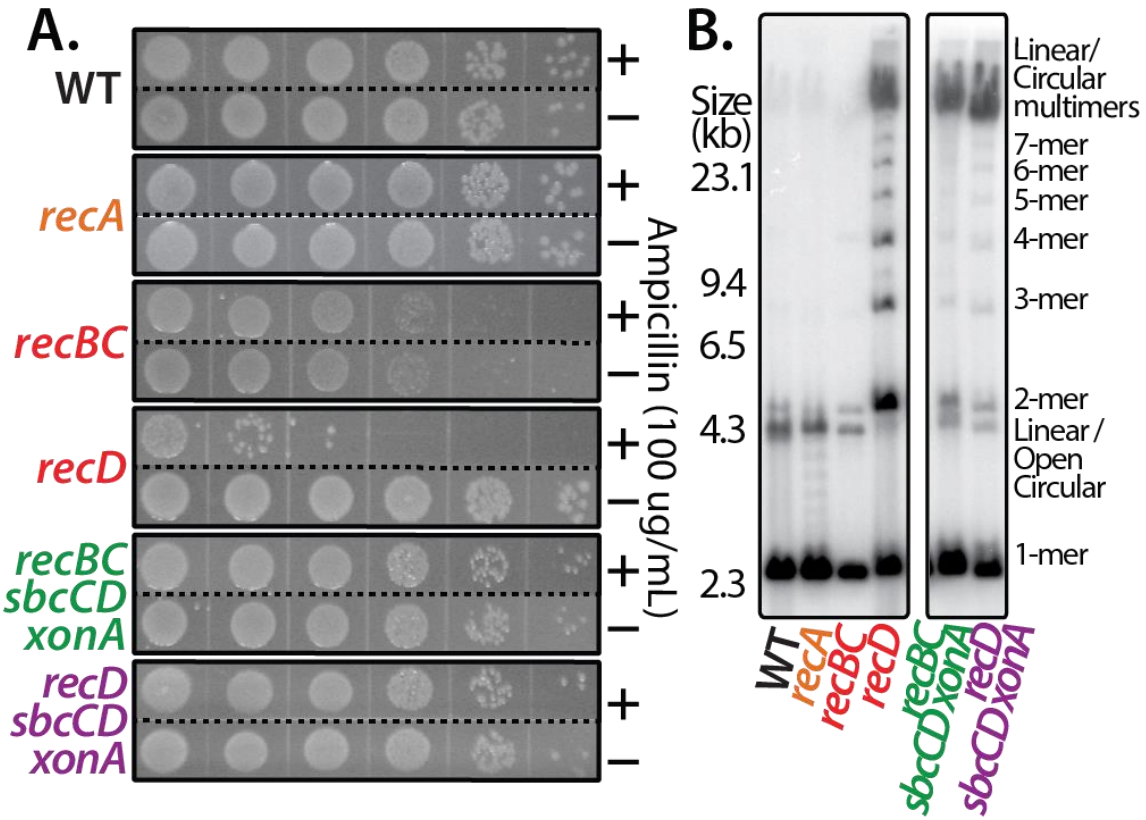

Figure 3.3. Inactivation of SbcCD-ExoI improves plasmid stability in *recBC* and *recD* mutants, although aberrant plasmid species accumulate and persist. A) Inactivation of SbcCD-ExoI improves the stability of plasmids in recBC and recD mutants. Cultures containing plasmid pBR322 were grown for 30 generations without selection before plating 10µL drops of 10X serial dilutions, with and without ampicillin selection, to determine the fraction of cells that maintained the plasmid in each strain. B) Plasmids in *recBC* and *recD* mutants lacking SbcCD-Exo I are maintained through an aberrant alternative pathway. Total DNA was prepared from replicating cultures containing pBR322 and examined by Southern analysis following agarose gel electrophoresis using 32P-labeled pBR322 as a probe.

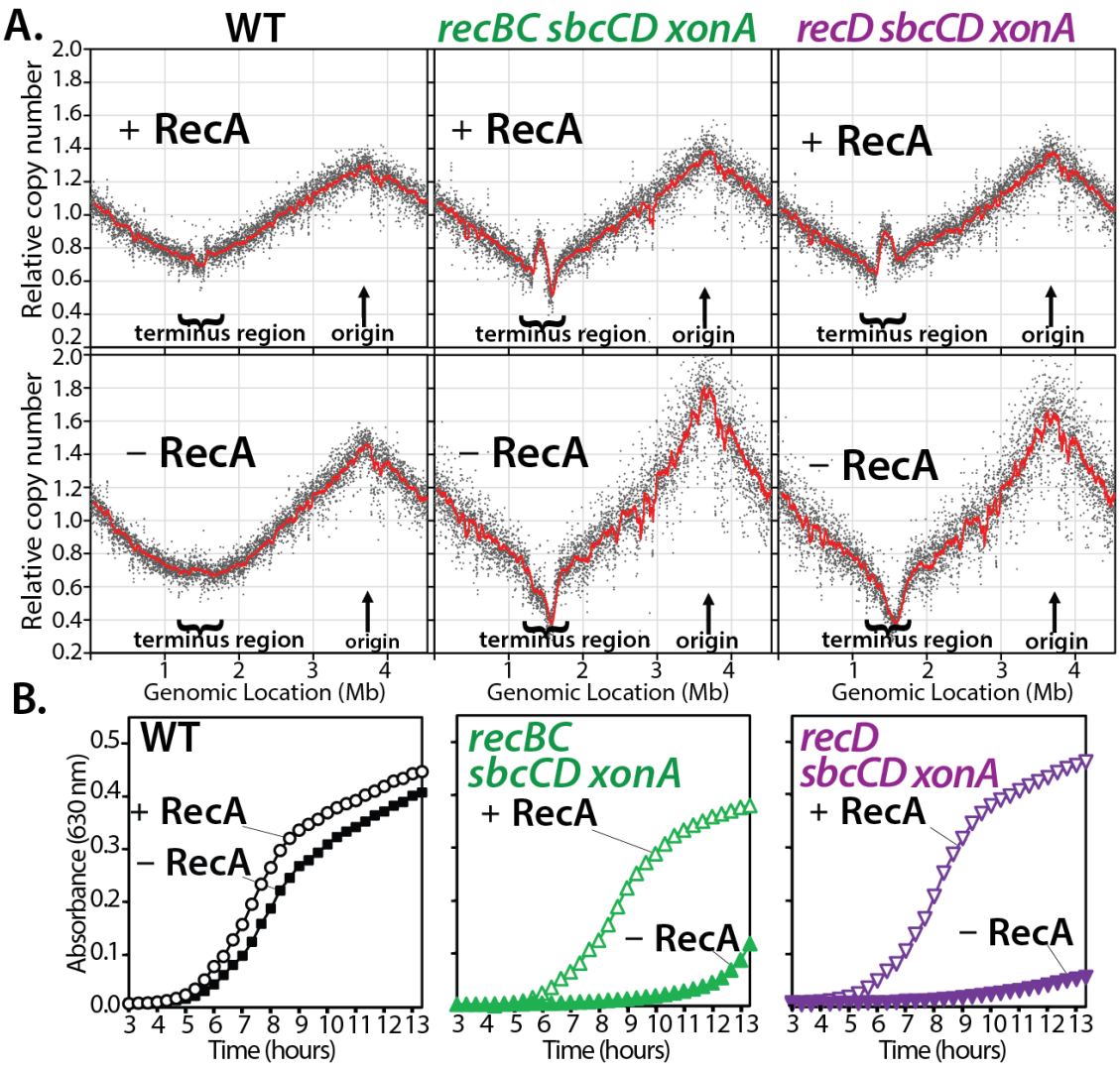

Figure 3.4. In the absence of SbcCD-ExoI processing, normal completion is prevented, and maintaining the region where forks converge becomes dependent on recombination. A) Replication completes normally in the absence of RecA. However, following SbcCD-ExoI inactivation, recBC and recD mutants become dependent on RecA to complete replication. B) Maintaining the chromosome region where forks converge correlates with the ability of cells to grow in culture. Replication profiles and growth were performed as in *figure 3.1*.

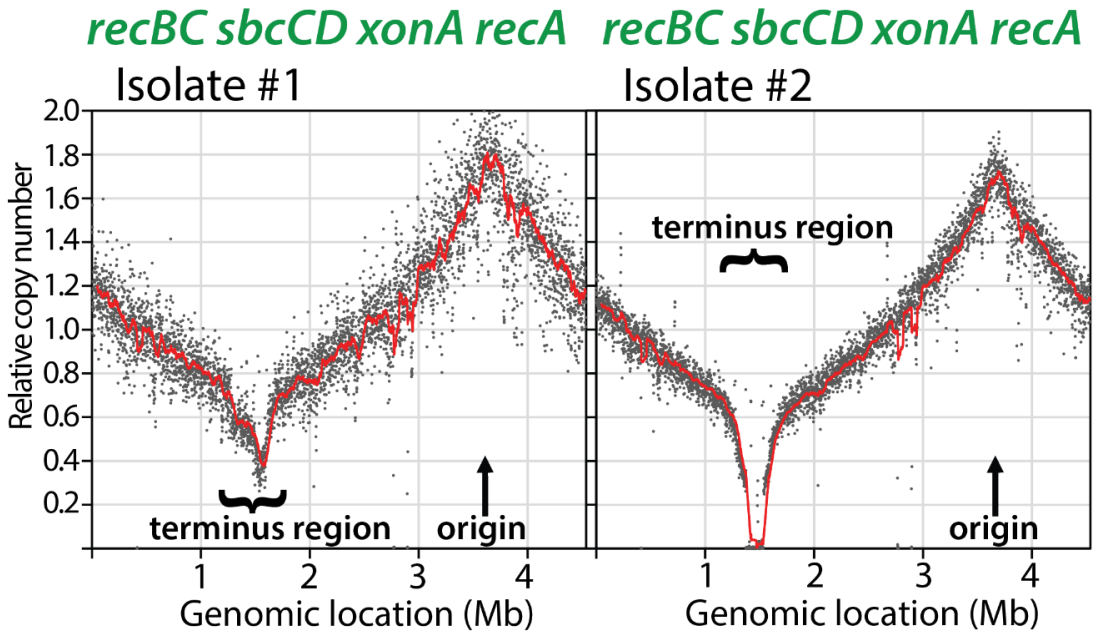

Figure 3.5. When the normal mechanism of completing replication is prevented, maintaining the region where forks converge becomes dependent on RecA and recombination. Replication profiles from two separate isolates of *recBC sbcCD xonA* mutants are shown. Both fail to maintain the region where replication forks converge. In the second isolate, the terminus region was deleted entirely. Replication profiles were performed as in *figure 3.1*. Isolates correspond to strains CL2576 and CL2575, respectively

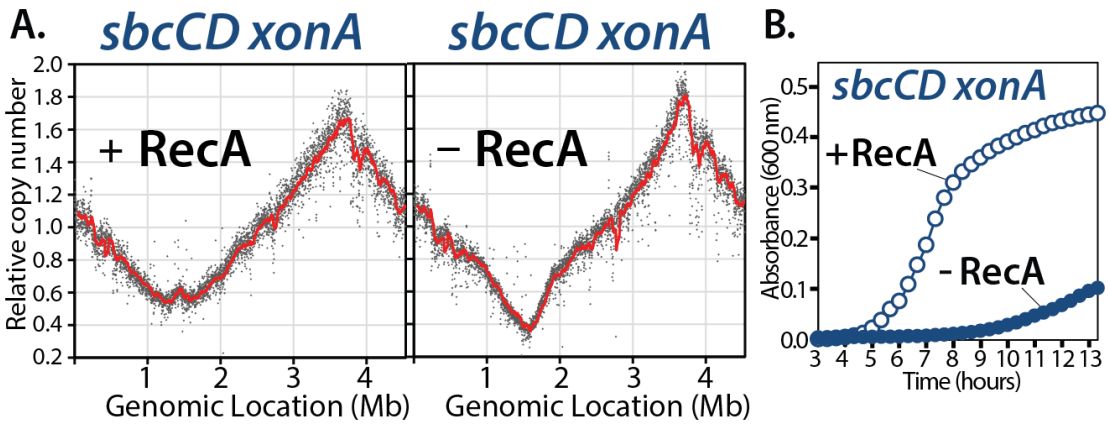

Figure 3.6. In the absence of SbcCD-ExoI processing terminus maintenance, growth, and viability become dependent on recombination. A) In an SbcCD-ExoI deficient strain, the aberrant recombinational pathway facilitates completion B) Growth and viability of the *sbcCD-xonA* mutant correlate with terminus maintenance. Replication profiles and growth were performed as in *Figure 3.1*.

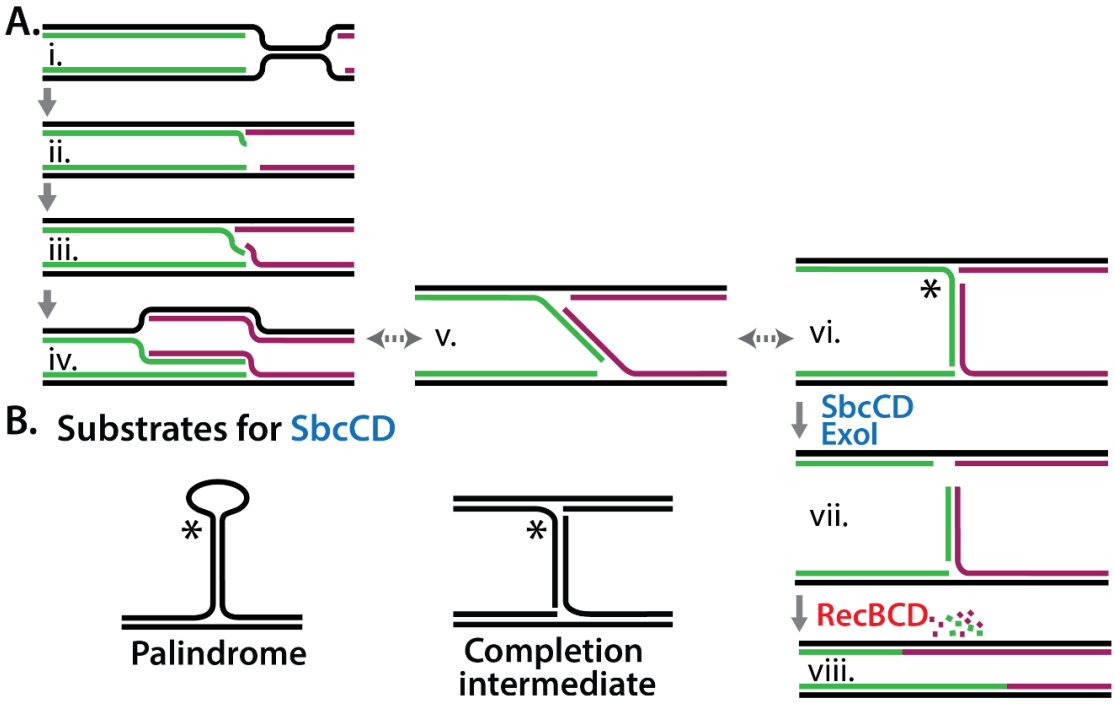

Figure 3.7. Model for completing replication. A) i-vi) Replication forks continue past their meeting point creating a palindromic substrate. vi) SbcCD-ExoI cleave and process the over- replicated intermediate. vii) RecBCD-mediated resection and joining of the DNA ends completes replication. B) In vitro and in vivo substrates for SbcCD, respectively.

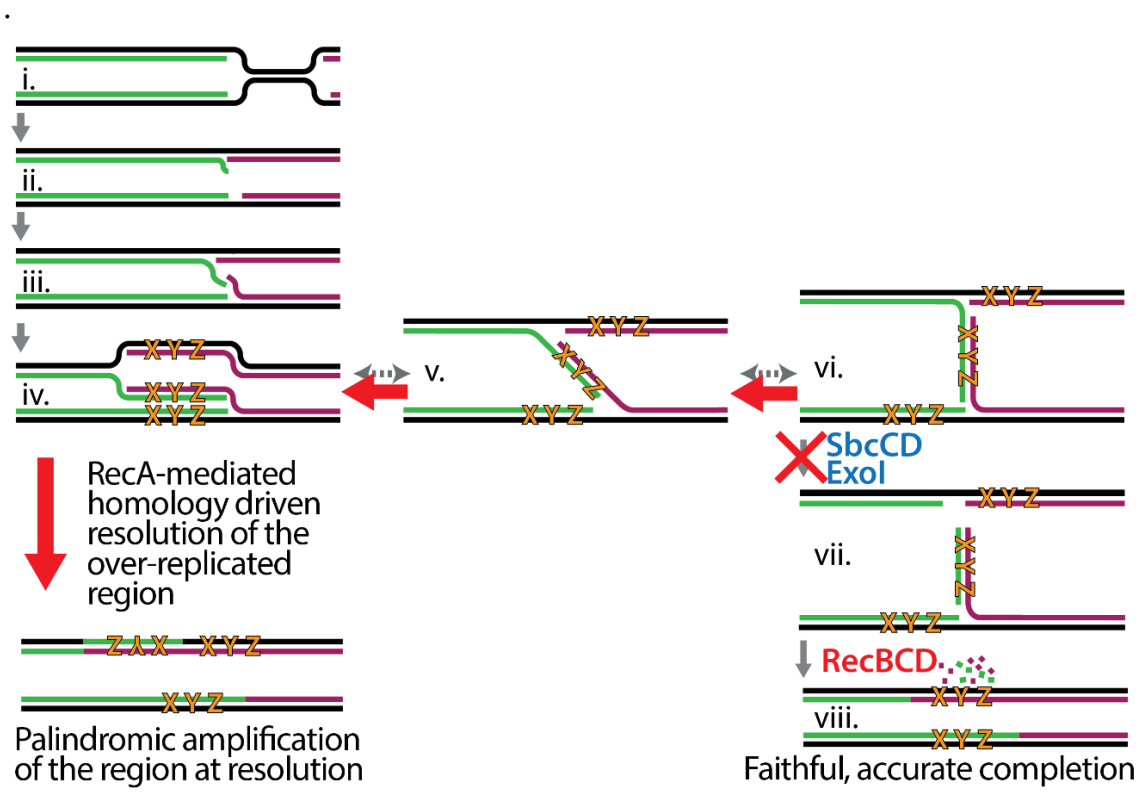

Figure 3.8. When the normal mechanism of completing replication is prevented, maintaining the region where forks converge becomes dependent on RecA and recombination. The model is shown as in *figure 5*. When processing of the over-replicated regions where forks converge is impaired, such as occurs in the absence of SbcCD ExoI, resolution of the joint molecules occurs through an aberrant form of RecA-mediated recombination. Resolution in this manner is termed aberrant because it produces chromosome amplifications at loci where forks converge. In addition, when the normal completion reaction cannot occur, cells fail to grow and these loci are degraded in the absence of RecA.

References

1. Costa A, Hood IV, Berger JM (2013) Mechanisms for Initiating Cellular DNA Replication. *Annu Rev Biochem* 82(1):25–54.
2. O'Donnell M, Langston L, Stillman B (2013) Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya. *Cold Spring Harb Perspect Biol* 5(7):a010108.
3. Wendel BM, Courcelle CT, Courcelle J (2014) Completion of DNA replication in *Escherichia coli*. *Proc Natl Acad Sci*:201415025.
4. Courcelle J, Wendel BM, Livingstone DD, Courcelle CT (2015) RecBCD is required to complete chromosomal replication: Implications for double-strand break frequencies and repair mechanisms. *DNA Repair* 32:86–95.
5. Heichinger C, Penkett CJ, Bähler J, Nurse P (2006) Genome-wide characterization of fission yeast DNA replication origins. *EMBO J* 25(21):5171–5179.
6. Wu P-YJ, Nurse P (2009) Establishing the Program of Origin Firing during S Phase in Fission Yeast. *Cell* 136(5):852–864.
7. Hill TM (1992) Arrest of bacterial DNA replication. *Annu Rev Microbiol* 46(1):603–633.
8. Kobayashi T, Hidaka M, Horiuchi T (1989) Evidence of a *ter* specific binding protein essential for the termination reaction of DNA replication in *Escherichia coli*. *EMBO J* 8(8):2435–2441.
9. Duggin IG, Wake RG, Bell SD, Hill TM (2008) The replication fork trap and termination of chromosome replication. *Mol Microbiol* 70(6):1323–1333.
10. Roecklein B, Pelletier A, Kuempel P (1991) The tus gene of *Escherichia coli*:

autoregulation, analysis of flanking sequences and identification of a complementary system in Salmonella typhimurium. *Res Microbiol* 142(2–3):169–175.

11. Taylor AF, Smith GR (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* 423(6942):889–893.
12. Taylor AF, Smith GR (1985) Substrate specificity of the DNA unwinding activity of the RecBC enzyme of *Escherichia coli*. *J Mol Biol* 185(2):431–443.
13. Taylor A, Smith GR (1980) Unwinding and rewinding of DNA by the RecBC enzyme. *Cell* 22(2):447–457.
14. Bianco PR, Kowalczykowski SC (2000) Translocation step size and mechanism of the RecBC DNA helicase. *Nature* 405(6784):368–372.
15. Dillingham MS, Webb MR, Kowalczykowski SC (2005) Bipolar DNA Translocation Contributes to Highly Processive DNA Unwinding by RecBCD Enzyme. *J Biol Chem* 280(44):37069–37077.
16. Chaudhury AM, Smith GR (1984) A new class of *Escherichia coli* recBC mutants: implications for the role of RecBC enzyme in homologous recombination. *Proc Natl Acad Sci U S A* 81(24):7850–7854.
17. Amundsen SK, Taylor AF, Chaudhury AM, Smith GR (1986) recD: the gene for an essential third subunit of exonuclease V. *Proc Natl Acad Sci* 83(15):5558–5562.
18. Barbour SD, Clark AJ (1970) Biochemical and genetic studies of recombination proficiency in *Escherichia coli*, I. Enzymatic activity associated with recB+ and recC+ genes. *Proc Natl Acad Sci* 65(4):955–961.
19. Rudolph CJ, Upton AL, Stockum A, Nieduszynski CA, Lloyd RG (2013) Avoiding

chromosome pathology when replication forks collide. *Nature* 500(7464):608–611.

20. Sharples GJ, Leach DRF (1995) Structural and functional similarities between the SbcCD proteins of *Escherichia coli* and the RAD50 and MRE11 (RAD32) recombination and repair proteins of yeast. *Mol Microbiol* 17(6):1215–1217.
21. Connelly JC, Leach DRF (2002) Tethering on the brink: the evolutionarily conserved Mre11-Rad50 complex. *Trends Biochem Sci* 27(8):410–418.
22. Hopfner KP, et al. (2001) Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell* 105(4):473–485.
23. Hopfner K-P, et al. (2002) The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* 418(6897):562–566.
24. Connelly JC, de Leau ES, Leach DR (1999) DNA cleavage and degradation by the SbcCD protein complex from *Escherichia coli*. *Nucleic Acids Res* 27(4):1039–1046.
25. Deng SK, Yin Y, Petes TD, Symington LS (2015) Mre11-Sae2 and RPA Collaborate to Prevent Palindromic Gene Amplification. *Mol Cell* 60(3):500–508.
26. Darmon E, et al. (2010) *E. coli* SbcCD and RecA control chromosomal rearrangement induced by an interrupted palindrome. *Mol Cell* 39(1):59–70.
27. Narayanan V, Mieczkowski PA, Kim H-M, Petes TD, Lobachev KS (2006) The Pattern of Gene Amplification Is Determined by the Chromosomal Location of Hairpin-Capped Breaks. *Cell* 125(7):1283–1296.
28. Xiao Y, Weaver DT (1997) Conditional gene targeted deletion by Cre recombinase demonstrates the requirement for the double-strand break repair Mre11 protein in murine embryonic stem cells. *Nucleic Acids Res* 25(15):2985–2991.

29. Luo G, et al. (1999) Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. *Proc Natl Acad Sci* 96(13):7376–7381.

30. Kang J, Bronson RT, Xu Y (2002) Targeted disruption of NBS1 reveals its roles in mouse development and DNA repair. *EMBO J* 21(6):1447–1455.

31. Zhu J, Petersen S, Tessarollo L, Nussenzweig A (2001) Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice. *Curr Biol* 11(2):105–109.

32. D'Amours D (2001) The yeast Xrs2 complex functions in S phase checkpoint regulation. *Genes Dev* 15(17):2238–2249.

33. Bruhn C, Zhou Z-W, Ai H, Wang Z-Q (2014) The Essential Function of the MRN Complex in the Resolution of Endogenous Replication Intermediates. *Cell Rep* 6(1):182–195.

34. Ajimura M, Leem SH, Ogawa H (1993) Identification of new genes required for meiotic recombination in *Saccharomyces cerevisiae*. *Genetics* 133(1):51–66.

35. Yeeles JTP, Dillingham MS (2010) The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes. *DNA Repair* 9(3):276–285.

36. Symington LS (2016) Mechanism and regulation of DNA end resection in eukaryotes. *Crit Rev Biochem Mol Biol* 51(3):195–212.

37. Hoa NN, et al. (2015) Relative contribution of four nucleases, CtIP, Dna2, Exo1 and Mre11, to the initial step of DNA double-strand break repair by homologous recombination in both the chicken DT40 and human TK6 cell lines. *Genes Cells*

Devoted Mol Cell Mech 20(12):1059–1076.

38. Yamamoto K, et al. (2016) Kinase-dead ATM protein is highly oncogenic and can be preferentially targeted by Topo-isomerase I inhibitors. *eLife* 5. doi:10.7554/eLife.14709.
39. Takeda S, Hoa NN, Sasanuma H (2016) The role of the Mre11-Rad50-Nbs1 complex in double-strand break repair-facts and myths. *J Radiat Res (Tokyo)* 57 Suppl 1:i25–i32.
40. Connelly JC, Kirkham LA, Leach DR (1998) The SbcCD nuclease of *Escherichia coli* is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proc Natl Acad Sci* 95(14):7969–7974.
41. Lengsfeld BM, Rattray AJ, Bhaskara V, Ghirlando R, Paull TT (2007) Sae2 Is an Endonuclease that Processes Hairpin DNA Cooperatively with the Mre11/Rad50/Xrs2 Complex. *Mol Cell* 28(4):638–651.
42. Paull TT, Gellert M (1998) The 3′ to 5′ Exonuclease Activity of Mre11 Facilitates Repair of DNA Double-Strand Breaks. *Mol Cell* 1(7):969–979.
43. Rattray AJ, Shafer BK, Neelam B, Strathern JN (2005) A mechanism of palindromic gene amplification in *Saccharomyces cerevisiae*. *Genes Dev* 19(11):1390–1399.
44. Chalker AF, Leach DR, Lloyd RG (1988) *Escherichia coli* sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71(1):201–205.
45. Gibson FP, Leach DR, Lloyd RG (1992) Identification of sbcD mutations as cosuppressors of recBC that allow propagation of DNA palindromes in *Escherichia*

coli K-12. *J Bacteriol* 174(4):1222–1228.

46. Barrick JE, et al. (2014) Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. *BMC Genomics* 15:1039.
47. Langmead B (2010) Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinforma Ed Board Andreas Baxevanis Al* CHAPTER:Unit-11.7.
48. Phillips GJ, Prasher DC, Kushner SR (1988) Physical and biochemical characterization of cloned sbcB and xonA mutations from *Escherichia coli* K-12. *J Bacteriol* 170(5):2089–2094.
49. Lehman IR, Nussbaum AL (1964) The deoxyribonucleases of *Escherichia coli*. V. On the specificity of Exonuclease I (phosphodiesterase). *J Biol Chem* 239:2628–2636.
50. Lloyd RG, Buckman C (1985) Identification and genetic analysis of sbcC mutations in commonly used recBC sbcB strains of *Escherichia coli* K-12. *J Bacteriol* 164(2):836–844.
51. Barbour SD, Nagaishi H, Templin A, Clark AJ (1970) Biochemical and genetic studies of recombination proficiency in *Escherichia coli*. II. Rec+ revertants caused by indirect suppression of rec- mutations. *Proc Natl Acad Sci U S A* 67(1):128–135.
52. Biek DP, Cohen SN (1986) Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in *Escherichia coli*. *J Bacteriol* 167(2):594–603.
53. Bassett CL, Kushner SR (1984) Exonucleases I, III, and V are required for stability of ColE1-related plasmids in *Escherichia coli*. *J Bacteriol* 157(2):661–664.
54. Hill TM, Marians KJ (1990) *Escherichia coli* Tus protein acts to arrest the

progression of DNA replication forks in vitro. *Proc Natl Acad Sci U S A* 87(7):2481–2485.

55. Paull TT, Gellert M (1999) Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev* 13(10):1276–1288.
56. Kitada K, Yamasaki T (2007) The MDR1/ABCB1 regional amplification in large inverted repeats with asymmetric sequences and microhomologies at the junction sites. *Cancer Genet Cytogenet* 178(2):120–127.
57. Kitada K, Yamasaki T, Aikawa S (2009) Amplification of the ABCB1 region accompanied by a short sequence of 200bp from chromosome 2 in lung cancer cells. *Cancer Genet Cytogenet* 194(1):4–11.
58. Tanaka H, Bergstrom DA, Yao M-C, Tapscott SJ (2005) Widespread and nonrandom distribution of DNA palindromes in cancer cells provides a structural platform for subsequent gene amplification. *Nat Genet* 37(3):320–327.
59. Tanaka H, Yao M-C (2009) Palindromic gene amplification — an evolutionarily conserved role for DNA inverted repeats in the genome. *Nat Rev Cancer* 9(3):216–224.
60. Marotta M, et al. (2017) Palindromic amplification of the ERBB2 oncogene in primary HER2-positive breast tumors. *Sci Rep* 7:41921.
61. Slamon DJ, et al. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235(4785):177–182.
62. Waddell N, et al. (2015) Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* 518(7540):495–501.

63. Dunham MA, Neumann AA, Fasching CL, Reddel RR (2000) Telomere maintenance by recombination in human cells. *Nat Genet* 26(4):447–450.

64. Zhong Z-H, et al. (2007) Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. *J Biol Chem* 282(40):29314–29322.

65. Lombard DB, Guarente L (2000) Nijmegen breakage syndrome disease protein and MRE11 at PML nuclear bodies and meiotic telomeres. *Cancer Res* 60(9):2331–2334.

66. Uziel T, et al. (2003) Requirement of the MRN complex for ATM activation by DNA damage. *EMBO J* 22(20):5612–5621.

67. Mellon I, Hanawalt PC (1989) Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 342(6245):95–98.

68. Courcelle J, Carswell-Crumpton C, Hanawalt PC (1997) recF and recR are required for the resumption of replication at DNA replication forks in Escherichia coli. *Proc Natl Acad Sci U S A* 94(8):3714–3719.

69. Chow K-H, Courcelle J (2007) RecBCD and RecJ/RecQ Initiate DNA Degradation on Distinct Substrates in UV-Irradiated *Escherichia coli*. *Radiat Res* 168(4):499–506.

70. Maloy SR, Nunn WD (1981) Selection for loss of tetracycline resistance by *Escherichia coli*. *J Bacteriol* 145(2):1110–1111.

71. Courcelle J, Hanawalt PC (1999) RecQ and RecJ process blocked replication forks prior to the resumption of replication in UV-irradiated *Escherichia coli*. *Mol Gen Genet MGG* 262(3):543–551.

72. Poteete A, Rosadini C, St. Pierre C (2006) Gentamicin and other cassettes for chromosomal gene replacement in *Escherichia coli*. *BioTechniques* 41(3):261–264.

Chapter 4

RecBCD and SbcCD-ExoI are required to complete convergent replication forks through a mechanism independent of recombination or double-strand breaks

Abstract

Completion of DNA replication on the chromosome requires RecBCD and the structure specific SbcCD-ExoI nucleases. However, many aspects of the substrates and mechanism by which these enzymes catalyze this reaction remain unknown. Here we use plasmid constructs to show that RecBCD and SbcCD-ExoI are required to complete replication in the presence of two converging replisomes but not one replisome. Completion on plasmids containing two converging replisomes occurs normally in the absence of *dif* sequences or RecA, indicating that the reaction does not involve double-strand breaks associated with fork collapse or aberrant segregation. Similar to the chromosome, in the absence of SbcCD-ExoI incisions, RecA becomes essential and maintains two-replisome plasmids through an aberrant form of recombination that is able to resolve the convergent replication forks but results in amplifications, rearrangements, and genomic instabilities. RecA was not essential to maintain plasmids containing one replisome under these conditions. The observations demonstrate that the substrate SbcCD-ExoI and RecBCD acts upon *in vivo* is created specifically when two replisomes converge and argues that the defect in *recBC* mutants arises from a failure to join the convergent strands of the replication forks.

Introduction

Before any cell can divide, its DNA must first be duplicated. To maintain the integrity of its genome during replication, a cell must effectively complete replication by resolving convergent replication forks and their newly synthesized DNA into discreet molecules while preserving the exact DNA sequence and copy number of the parent chromosome. Segregation in the absence of effective completion can result in chromosomal instabilities and large-scale mutations, yet the process by which DNA replication is completed has not been well understood (1–3). We have studied this process on the chromosome and have recently identified several gene products required for processing convergent replication forks at the end stage of replication.

Completion of DNA replication in *Escherichia coli* requires the RecBCD helicase-nuclease, the structure-specific nuclease SbcCD, and the single-stranded exonuclease ExoI (4, 5). SbcCD and ExoI act before RecBCD to process the DNA intermediates resulting from convergent replication forks without homologous recombination or RecA (4). While their involvement has been established, the exact role in processing intermediate DNA structures that arise during completion remains to be elucidated. In the absence of these enzymes responsible for the normal reaction, recombination acts to facilitate completion. When cells complete DNA replication through this aberrant pathway, completion is less efficient and contributes increased amplifications, rearrangements, and chromosomal instability (5). Chromosomal segregation has been proposed as source of the chromosomal instability observed in the absence of the normal completion pathway requiring RecBCD (6), and here we use plasmid minichromosomes as a model to test this concept.

Plasmid minichromosomes have long been used to understand the efficiency and control of plasmid replication, and as a result, many diverse mechanisms have been discovered through which plasmids initiate replication and self-regulate copy number (7–10). While much of this work has focused on the initiation of plasmid DNA replication and differences from chromosomal initiation (11, 12), once replication is initiated most plasmids utilize the normal cellular machinery of the host for replication (13). Furthermore, some plasmids replicate in a manner highly similar to that of the chromosome using two replication forks emanating from a single origin of replication which proceed in opposite directions around a circular molecule (14–16). Due to the mechanistic similarities of the converging replication forks and the utilization of the host machinery during the DNA replication of these plasmids, they likely complete replication in a manner analogous with what occurs on the chromosome, making them excellent tools for study of the completion process.

The completion of plasmid replication can be studied by directly observing plasmid species and by studying how stable or unstable they are in cells. Whether plasmids replicate unidirectionally or bidirectionally, instability of plasmids *in vivo* can arise from three different sources: (i) stochastic plasmid loss without active partitioning, (ii) decreased fitness of plasmid-containing cells, and (iii) direct defects in plasmid replication (17–21). Plasmids can be lost randomly in the absence of an active partitioning or plasmid addiction mechanism, a problem that would be exacerbated in low-copy plasmids but lessened in high-copy ones (19). Additionally, plasmid-containing cells could exhibit decreased viability due to the metabolic burden of maintaining the plasmid and could confer a fitness advantage to plasmid-free cells in culture (22). These

first two sources of instability can be circumvented by evaluating a high-copy plasmid over few generations. Direct defects or inefficiencies in plasmid replication could generate plasmid-free cells after cell division even under these conditions. This third source of instability would be expected to arise in bacterial strains defective in the completion of DNA replication, particularly in plasmids with two replisomes.

In some bacteria, the completion reaction is known to occur within a defined region of the chromosome known as the terminus. This region contains *ter* sequences which when bound by the protein Tus, block replication forks in an orientation specific manner (23, 24). The layout of these *ter* sequences serve to contain the reaction to the terminus region by forming what has been termed a "replication fork trap" (as reviewed in (25)). Interestingly, the mechanisms containing this reaction are not required for accurate completion, as cells lacking either *ter* sites or Tus replicate normally (26, 27). The overall structure of this replication fork trap suggests that it functions to increase the efficiency of the completion reaction by either limiting the progression of converged replication forks, or limiting the degradation emanating from the point of fork convergence. The role that this structure plays in the completion reaction remains to be examined and is easily studied using a plasmid, or minichromosome, model system.

By evaluating plasmid stability over short duration, two-replisome plasmids that are maintained at a high copy and lack an active segregation mechanism can be used as a model to study the completion of DNA replication on the chromosome. Performing this type of analysis in strains previously identified to be deficient in completing chromosome replication we can further understand the role those gene products play in the process. Furthermore, by evaluating how a replication fork trap impacts the stability of two-

replisome plasmids, we can better understand the function of the replication fork trap during completion of DNA replication on the chromosome. Here we show that two replisome plasmids complete replication in a manner analogous to, and require the same gene products required for, the completion of DNA replication on the chromosome. We also demonstrate that the efficiency of the completion reaction is enhanced by the presence of a replication fork trap.

Materials and methods

Strains and Plasmids. All strains used in this work are derived from SR108, a *thyA deoC* derivative of W3110 and listed in table 4.1. Plasmid constructions were performed according to published protocols for *in vivo* recombineering (28), construction by amplification (29), and Gibson assembly (30). pCLZ was constructed via construction by amplification method using the primers 5'-GTCGGTTCAGGGCAGGGTCGTGGAT CCCGCGGAACCCCTATTTGTTT-3' and 5'-GGCGGTTTGCGTATTGGGCGCGGTC TGACAGTTACCAATGC-3' to amplify the Ampicillin resistance cassette and provide homology to parent plasmid pCB104 kindly provided by Dr. Wegryzn. pCTZ was constructed using the primers 5'- GGCGGTTTGCGTATTGGGCGCATATTAGTTACA ACATCCTATATGGTCTGACAGTTACCAATGC-3' and 5'- GTCGGTTCAGGGCAG GGTCGTGGATCCACTTTAGTTACAACATACTTATTCGCGGAACCCCTATTTGT TT-3' to amplify the Ampicillin resistance cassette, insert a replication fork trap and provide homology to parent plasmid pCB104 kindly provided by Dr. Grzegorz Węgrzyn (31). Primers 5'-ATTGCTGATAAATCTGGA-3', 5'-CTTTGGAATCCAGTCCCTCTT CCTCCTGCTGATCTGCGACTTATCAAC-3', 5'-TCCAGATTTATCAGCAAT-3',

and 5’-GTTGATAAGTCGCAGATCAGCAGGAGGAAGAGGGACTGGATTCCAAA G-3’ were used to remove a Chi sequence from both plasmids via Gibson assembly.

Growth Rates. Cells were grown in 0.1-mL cultures in LB medium supplemented with 10µg/mL thymine (LBthy), at 37°C with agitation in a 96-well microtiter dish. Absorbance at 490 nm or 630 nm was measured over time using a BIO-Whittaker ELx808 plate reader 10 Ten-microliter aliquots of serial 10-fold dilutions were then spotted on LBthy plates to ensure equal numbers of viable cells were used. (4)

Transformation Efficiency. Fresh overnight cultures of each strain were diluted 1:100 in 10 ml of LBthy and grown to an OD630 of 0.4. Cells were pelleted, washed with 30 mL of DI H_20, pelleted and washed in 30mL 10% glycerol, and finally pelleted and resuspended in 200 µL of 10% glycerol. Cells were stored at -80ºC until needed. 40 µL of competent cells were mixed with 50ng of purified plasmid and resuspended in 1 mL SOC media and allowed to recover at 37°C for 30 minutes. 50 or 200µL of cells were spread on LBthy plates with 50 or 100 µg/mL ampicillin. Ten-microliter aliquots of serial 10-fold dilutions were then spotted on LBthy plates in the presence or absence of 50 or 100 µg/mL ampicillin. Transformants and viable colonies were counted following overnight incubation at 37°C and compared to establish the percent of transformants per viable cells. The same preparation of competent cells for each strain was used for all plasmids.

Plasmid Stability. Overnight cultures of bacterial strains containing plasmids grown with selection. Cultures containing the various plasmids were grown for 15 or 30 generations by repeatedly subculturing 1:50 or 1:1000 respectively in LBthy medium at 37°C with aeration. Ten-microliter aliquots of serial 10-fold dilutions were then spotted on LBthy plates in the presence or absence of 50 or 100 µg/mL ampicillin. Viable colonies were counted following overnight incubation at 37°C and compared to establish the percent of plasmid-containing cells before and after each of three passages. (4)

Total Genomic DNA Extraction. 0.75 mL cultures grown in LBthy medium supplemented with 100 µg/ml ampicillin were taken and placed into 0.75 Ml of cold 2× NET (100 mM NaCl, 10 mM Tris, pH 8.0, 10 mM EDTA). Each sample was pelleted, resuspended in 140 µl of 1.5 mg/mL lysozyme and 0.2 mg/mL RNaseA in TE (10 mM Tris, pH 8.0, 1 mM EDTA), and lysed at 37 °C for 30 min. At this time, proteinase K (10 µl, 10 mg/ml) and Sarkosyl [10 µl, 20% (wt/wt)] were added and incubated at 37 °C for 30 min. Samples were then extracted with 4 vol of phenol/chloroform (1/1) and dialyzed for 1 h on 47 mm Whatman 0.05-µm pore disks (Whatman #VMWP04700) floating on a 250-mL beaker of TE (1 mM Tris, pH 8.0, 1 mM EDTA).(4)

Southern Analysis of Plasmid Replication Intermediates. Total genomic DNA samples of mutant strains containing different plasmids were digested with SacII or EcoRV (New England Biolabs), all of the plasmids used lack SacII restriction sites expect pCLZ which contains a single site and contain a single EcoRV restriction site, and extracted with chloroform. Equal volumes were loaded onto a 0.5% and 1.0% agarose gel

containing 0.5× TBE (220 mM Tris, 180 mM Borate, 5 mM EDTA, pH 8.3) and electrophoresed at 1 V/cm respectively. Gels were transferred to Hybond N+ nylon membranes and probed with appropriate plasmid that had been labeled with 32P by nick translation according to the protocol supplied by Roche using [α-32P] dCTP (PerkinElmer). Radioactivity was visualized using a Storm 840 and its associated ImageQuant Software (Molecular Dynamics).(4)

Results

Construction and characterization of plasmids replicated by two replisomes.

To further characterize the involvement of chromosomal segregation and double-strand breaks as replisomes converge during the completion of replication, we engineered a plasmid that contains a bidirectional origin of replication that would allow us to characterize the genes, sequences, and intermediates associated with maintaining a minichromosome that contains convergent replication forks (Figure 4.1a). The two-replisome plasmid contained the bacteriophage lambda origin of replication, which loads dual helicases and utilizes the host's replication proteins (32–35). For the purposes of control and for comparison, we compared the replication and stability of these two-replisome plasmids to pBR322, a well-characterized plasmid that maintains a moderate copy number and utilizes the host's replication machinery, but replicates unidirectionally, and therefore only contains a single replisome.

The two-replisome plasmids could be propagated and were stable in the presence of selection, however, they were modestly less stable than the one-replisome plasmid

when grown without selection. Whereas the unidirectionally-replicating pBR322 was stably maintained without loss over 30 generations, only 0.1-1.0% of cells maintained the bidirectionally-replicating plasmids over the same span (Figure 4.1B). Despite the reduced stability of the two-replisome plasmids, they propagated at a rate generating more than 50 copies per chromosome, whereas the single replisome plasmid contained only ~15 copies per chromosome consistent with previously reported copy numbers for these plasmids (Figure 4.1c and (10, 31, 36)). Additionally, a Southern analysis of DNA purified from these cultures showed that replication of the two-replisome plasmids could be resolved as supercoiled monomeric replicons (Figure 4.1d), and did not impair the growth of cells in culture (Figure 4.1e). Thus, the two-replisome plasmid appears to initiate replication at modest frequencies, can replicate normally, and is capable of completing and resolving replication into circular monomeric units. Additionally, the reduced stability of the two-replisome plasmid relative to the one-replisome plasmid is not simply due to a reduced initiation rates, lower copy numbers, or altering the growth rate of the bacterial cultures.

Although greater than 90**%** of the one-replisome plasmid is resolved and maintained as circular monomeric units, the two-replisome plasmids contained a greater proportion of molecules as multimeric-linear and higher order species (Figure 4.2a). A caveat to this observation is the discovery of a SacII restriction site in pCLZ after these experiments were conducted. Given the low level of SacII used to lower the viscosity of the DNA preparation, its unlikely that the *in vivo* plasmid species were measurably altered. Furthermore, the presence of increased intermediates in both pCLZ and pCTZ

which does not contain a SacII site suggests that little if any alteration of *in vivo* plasmid species occurred as a result of preparation for Southern analysis.

We hypothesized that these multimeric plasmid species could account for the instability oberserved in the two-replisome plasmid. If these species replicate and segregate normally but fail to complete or be resolved into circular species, then whenever replication ceases, they would remain susceptible to exonucleoytic degradation and be prone to eventual loss. Consistent with this, we found that loss of the two-replisome plasmids did not correlate with the number of generations the plasmid has replicated (Figure 4.2b). Instead, the rate of plasmid loss correlated closely with the number of times the cultures reached stationary phase and ceased replicating (passages) (Figure 4.2c). Similarly, when we examined the ability of the aberrant species to persist in the absence of replication, the higher order multimeric species were degraded more rapidly than the closed monomeric species, consistent with the idea that these uncompleted intermediates are susceptible to degradation and loss (Figure 4.2d).

The requirement for RecBC to complete replication arises in the presence of two replisomes, but not one, and does not involve a double strand break intermediate.

Having established that the bidirectional-replicating plasmid, pCLZ can propagate as unit length circular monomers, we next asked whether the genetic requirement for completing replication on the two-replisome plasmids mirrored the requirements to complete replication on the chromosome. Completion of replication on the chromosome requires RecBC and -D to complete resection of the over-replicated region where forks

converge and join the nascent ends. In the absence of RecBC, joining does not occur and exonucleases continue to degrade these strands, resulting in a marked loss and inability to maintain the DNA in these regions (4, 5).

To examine whether the two-replisome plasmid similarly depended upon recBC, we compared the ability of the two-replisome plasmid to stably transform cells, relative to the one-replisome plasmid. As shown in *figure 4.3a* and *4.3b*, the unidirectional plasmid transformed and grew in all mutants examined. Similarly, wild type as well as several mutants that fail to degrade over-replicated sequences on the chromosome, including *recD*, *recG*, and *sbcCD xonA*, were all able to be stably transformed by the bidirectional-replicating plasmid. In contrast, *recBC* mutants could not be stably transformed with the two-replisome plasmid (Figure 4.3a). Following some transformation attempts, micro-colonies of *recBC* mutants could be observed on selective plates following extended (~48 hours) incubation of the transformation mixture (Figure 4.4a). However, these micro-colonies were unable to grow further when inoculated on plates or in culture. Further, when DNA was purified and examined directly from micro-colonies scraped from these plates, the plasmid DNA was found to migrate entirely as high molecular weight, linear species and lacked any closed circular monomers (Figure 4.4b). The ability to transform *recBC* mutants with the one-replisome plasmid, but not the two-replisome plasmid strongly suggests that the substrate processed by RecBCD is specific to a structure created by the presence of two-replisomes. Further, the presence of linear plasmid intermediates in the *recBC* microcolonies implies that the plasmid DNA enters and replicates successfully, but that it fails to join replicated ends to form circular

monomeric molecules, leading to eventual degradation and loss, similar to that seen on the chromosome.

Similar to two-replisome plasmids, *recBC* mutants also fail to maintain the region where replisomes converge on the chromosome (4, 5). If the failure to maintain regions where replisomes converge were due to defective repair of double strand breaks that arise at these sites, then one would expect mutants lacking RecA, which is essential for all double strand break repair, to exhibit a similar phenotype. However, RecA is not required to maintain this region on the chromosome (4), and as shown in *figure 4.3b*, it also does not affect the ability to stably transform or maintain plasmids containing two-replisomes. The ability of recA mutants to transform and maintain plasmids containing two-replisomes demonstrates that the RecBC-mediated completion reaction can occur independently of RecA, and that inability of recBC mutants to maintain plasmids containing two-replisomes is not associated with a failure to repair double strand breaks.

When the normal completion reaction is impaired, the ability to transform and maintain plasmids with two replisomes becomes dependent on the aberrant recombinational pathway to complete replication.

On the chromosome, the completion reaction is initiated by the SbcCD and ExoI nucleases which are required to incise and/or resect the DNA structure created at sites where convergent replication forks meet, before RecBCD can facilitate joining (5). When both nucleases are inactivated, the over-replicated structure at sites where replication forks converge persists, preventing replication from completing normally.

Under these conditions the ability to maintain these regions and continue to grow becomes dependent on an aberrant recombinational process that results in genomic instabilities and amplifications at sites where forks converge (5).

To determine if the two-replisome plasmids similarly rely upon these enzymes, we compared the ability to transform and maintain the plasmid in *sbcCD xonA* mutants, both in the presence and absence of RecA. We found that strains lacking the SbcCD-ExoI nucleases could be transformed and stably maintain the two-replisome plasmid (Figure 4.3). However, in the absence of these gene products, maintaining the two-replisome plasmid depends on the presence of RecA. This argues that similar to the chromosome, when the normal mechanism of completion is impaired, the reaction is shunted through an aberrant recombinational mechanism. For comparison, we also examined mutants lacking RecG. *recG* mutants also exhibit a chromosomal over-replication phenotype, similar to sbcCD xonA (4). However, the over-replication is thought to result from the promiscuous initiation of replication from R- and D-loops (37–39), and RecG does not appear to be directly involved in the completion reaction (4). We found that the two-replisome plasmid could be stably transformed in recG mutants both in the presence or absence of RecA (Figure 4.3a-b). The observation suggests that dependence on RecA to maintain the bi-directional plasmid in sbcCD xonA mutants is specific to an impaired ability to complete replication. The RecA-dependence of sbcCD xonA mutants was also specific to plasmid containing two replisomes, as the single replisome pBR322 could successfully transform sbcCD xonA mutants in the presence or absence of RecA (Figure 4.3a-b). The observation argues that the substrate these enzymes act upon is specifically created when two replisomes converge.

Similar to the chromosome, the effect of the *sbcCD* and *xonA* mutations were additive, as the absence of either gene product alone did not prevent the ability to maintain the two-replisome plasmid in the absence of RecA. Additionally, inactivation of both SbcCD and ExoI restored the ability of recBC mutants to transform and maintain the bidirectional plasmid, through a mechanism that depended upon RecA (Figure 4.3). The observation is consistent with what is observed on the chromosome, where loss of SbcCD ExoI blocks the completion reaction prior to RecBCD function and shunts the reaction through the aberrant recombinational pathway. Unexpectedly, in the absence of RecD, the sbcCD xonA mutations prevented transformation of the bidirectional plasmid, even in the presence of RecA (Figure 4.3). Further, these mutations did not prevent the transformation of plasmids replicated using only one replisome. On the chromosome, the over-replicated region where forks converge persists, leading to amplification of this region. However, the mutations do not impair the growth or viability of these cells. Thus, the inability to transform or stably establish two-replisome plasmids in mutants lacking the RecD SbcCD and ExoI nucleases is not clear but may relate to a severely impaired ability to resect or process the excess DNA sequences generated where forks converge.

The aberrant recombinational pathway is responsible for most of the instability of plasmids with two replisomes.

The results above show that plasmids containing two-replisomes are more unstable than those containing one replisome, and that the ability to transform plasmids with two replisomes depends on many of the same factors that are required to complete

replication on the chromosome. On the chromosome, completion can occur either through the normal SbcCD-ExoI-RecBCD-mediated pathway or through an aberrant recombinational mechanism when the normal pathway is impaired or prevented (5). Completion through the aberrant recombinational pathway is associated with genomic instabilities and amplifications on the chromosome.

By evaluating plasmid stability in high copy plasmid with no active segregation mechanism over a short duration without selection, we can evaluate the role of normal and aberrant pathways on plasmid and ultimately chromosomal replication. We observe that unidirectionally-replicating plasmid is maintained efficiently in nearly all the strains evaluated (Figure 4.5a-c). *recBC* and *recD* mutants both demonstrate significant instability of the unidirectional plasmids over 30 generations, with the instability observed in *recD* mutants being the most severe (Figure 4.5b and (4, 40)). Interestingly, the deficiencies in stability of plasmids in *recBC* and *recD* mutants are rescued by deficiencies in SbcCD and ExoI (Figure 4.5b and (5)).

To examine what the causes instability on the two-replisome plasmids, we examined how mutations in genes associated with the normal and recombinational pathways affected plasmid loss over time. To this end, we compared ability of plasmids to propagate in wild type and mutants in the absence of selection. We found that, with one exception (see Exo I below), inactivation of genes associated with the normal completion reaction, RecD, SbcCD, or the *sbcCD xonA* double mutant, did not alter the rate that plasmid loss occurred (Figure 4.5a). *recG* mutants also did not affect the plasmid loss rate. By contrast, loss of recA, which inactivates the aberrant recombinational mode of completing replication, significantly increased the stability of

the plasmid (Figure 4.5a). Inactivation of recA-dependent completion pathway similarly increased the stability of other mutants, including *recG, sbcCD,* and *xonA* (Figure 4.5c). By contrast, in the *sbcCD xonA recBC* strain that depends entirely on the recombination pathway, the stability of the plasmid was lower than wild type cells, and in the absence of RecA these strains could not be transformed (Figure 4.5c). Taken together, these observations are consistent with the idea that the RecA-mediated recombinational pathway is responsible for most of the instabilities observed on plasmids containing two replisomes, and even more so if the normal pathway of completion is impeded.

Unexpectedly, we also observed that inactivation of ExoI, increased the stability of the plasmid similar to recA. While the reason for this is not clear, early studies noted that ExoI associates strongly with RecA during purification (41, 42), suggesting a potential functional interaction between these proteins in vivo.

To further characterize the instability of two replisome plasmids, we examined the form of these plasmids in cells. To this end, total genomic DNA was purified from cultures containing the plasmid, electrophoresed and examined by southern analysis using 32P-labelled plasmid as the probe. One-replisome plasmids were maintained primarily in monomeric forms, with multimeric forms observable in strains known to be defective in completing DNA replication (Figure 4.6a and (4, 5)). The introduction of the *recA* mutation reduced the aberrant products occurring in the single-replisome plasmids significantly (Figure 4.6b). Of those strains that are able to be transformed by the two-replisome plasmid, each contained significant amount of abnormal higher order and linear products (Figure 4.6c). In each case however, inactivation of RecA led to reduced abnormal multimeric species and increased monomeric products (Figure 4.6d). Taken

together, the higher order plasmid products appear to correlate with the instability of the plasmids and suggest that the aberrant recombinational pathway to complete replication is responsible for the instability observed on plasmids, similar to that seen on the chromosome.

In mutant strains not as severely impaired in the ability to complete replication, a loss of RecA appears to have an inverse effect on plasmid stability. Mutant strains deficient in RecA that are transformable with a two-replisome plasmid demonstrate a plasmid stability greater than in the presence of RecA, and in most cases than that of wild type cells. The mutant strain that maintains the two-replisome plasmid the most stably is the *xonA recA* double mutant (Figure 4.5c, suggesting that the role ExoI is playing in the completion reaction is at least in part involved in preventing recombination. This is consistent with the idea that RecA is only involved in completion in the absence of the normal pathway. This suggests that in strains with an intact completion pathway, homologous recombination and RecA only serve to drive instability by inducing recombinationally-driven multimerization. Observations of reduced multimerization of one- and two-replisome plasmids in strains deficient in RecA support this idea (Figure 4.5).

The presence of a replication fork trap increases the efficiency of both the normal and aberrant completion reaction

The results presented above indicate that an impaired ability to complete replication through the normal recBCD-mediated pathway, or the persistence of the over-

replicated intermediates created by the collision of two replication forks leads to genetic instabilities and plasmid loss. On the *E. coli* chromosome, *ter* sequences serve to limit the amount of over-replication that can occur from colliding replication forks, and perhaps limit runaway replication and the aberrant process. To examine how the *ter* sequences affect the genetic stability and the ability to accurately complete replication, we constructed a two-replisome plasmid identical to the one previously used with an added replication fork trap, or *ter* trap, directly opposite the bidirectionally-replicating plasmid origin in proportion with the structures found on the *E. coli* chromosome (Figure 4.7a).

Cells deficient in SbcCD alone, exhibit a greatly reduced stability of two-replisome plasmids containing a *ter* trap, with all the cells in population being plasmid-free after only 15 generations (Figure 4.8A). This contrasts strongly with the observation that bidirectional plasmids without a *ter* trap are maintained at wild type levels (Figure 4.8a). This susceptibility of *sbcCD* mutants to loss of bidirectional plasmids with a *ter* trap suggests a fundamental replication intermediate persists in absence of processing by SbcCD and is contributing to the instability.

sbcCD xonA recD mutants have a greatly increased transformation efficiency with a two-replisome plasmid containing a *ter* trap as stable transformants cannot be obtained by transformation of a two-replisome plasmid without one (Figure 4.7b). Interestingly, the addition of a *ter* trap rescues the transformation of *sbcCD xonA* mutants in the absence of RecA (Figure 4.7c). This effect is all the more striking considering that stable transformants cannot be obtained from *sbcCD xonA recBC* and *sbcCD xonA recD* mutant strains without RecA (Figure 4.7c). Bidirectional plasmids in *recD* and *sbcCD xonA recD*

mutants demonstrate severe instability, but the addition of a *ter* trap restores it to wild type levels (Figure 4.8b). The *recD* mutants lack a functional RecD subunit of the RecBCD enzyme, resulting in an attenuated helicase action and loss of nuclease activity (43–45). Interestingly, there is no significant difference in stability of two-replisome plasmids with and without a replication fork trap in strains lacking *xonA, recG, xonA sbcCD,* and *sbcCD xonA recBC* mutants as well as those same strains lacking RecA (Figure 4.8a and 4.8c).

Strains containing RecA, maintain two-replisome plasmids containing a *ter* trap with increased levels of multimerization relative to those without a *ter* trap (Figure 4.9a, 4.9b). Two-replisome plasmids with and without a *ter* trap are maintained as largely monomeric forms in mutant strains without RecA, as compared to the same mutants with RecA and wild type cells (Figure 4.9c, 4.9d). This is further evidence suggesting that in strains with an intact completion pathway, homologous recombination and RecA only serve to drive instability. This also supports the idea that RecA only serve to drive instability by inducing recombinationally-driven multimerization in transformable mutants, or mutants that have an intact normal completion pathway.

Discussion

The data shown here illustrate that the completion DNA replication of plasmids that have two converging replisomes requires the same gene products required for completion of chromosomal DNA replication. This further suggests that the replication intermediates formed as replication fork converge on the chromosome are highly similar to those formed as replication forks converge on bidirectionally-replicating plasmids. We

also demonstrate that the *ter* trap serves to enhance the efficiency of the completion reaction by preventing runaway replication or degradation of the DNA. Additionally, when replisomes converge within a *ter* trap, SbcCD is required to process the DNA intermediate. Furthermore, when the normal mechanistic processing of the convergent replication forks does not occur, an aberrant recombinationally-mediated process facilitates completion inefficiently resulting in accumulation of over-replicated DNA intermediates. This is highlighted by the observation that the presence of RecA in strains deficient in *sbcCD xonA recBC* do not facilitate completion efficiently enough to rescue transformation efficiency or to maintain stability at wild type levels and that without RecA, the phenotype is even more severe.

The aberrant recombinational mechanism of completion appears to contribute to the plasmid instability even in wild type cells, as evidenced by the higher rate of plasmid loss of one-replisome vs two-replisome plasmids, the production of large multimeric species and plasmid loss, and the increased stability and reduction of those multimeric species in when the aberrant pathway is inactivated. Replication profiles of the chromosome in *E.coli* suggest that recombination may also contribute to some instabilities associated with the completion of replication in wild type cells as they contain more over-replicated regions at sites where replication completes than cells lacking the ability to recombine (4, 5). Based on the high copy number of these two-replisome plasmids, the number of completion events to be processed by RecBCD far exceeds that of the chromosome and may tax the capacity of the normal RecBCD-mediated completion reaction, leading to the high level of instability observed on the two-replisome plasmid even in wild type cells.

Two recent studies proposed that the inability to maintain the region where replisomes converge in *recBC* mutants is due to the formation of a double strand break caused during segregation of the chromosomes (6, 46) They proposed that *recA* cells may not display this phenotype on the chromosome because the broken chromosomes are rapidly degraded in these mutants. Several observations made in this work run counter to this hypothesis. Firstly, we have shown that genes previously shown to be required for completion of DNA replication on the chromosome are required for the maintenance high copy two replisome plasmids. The lack of an active partitioning mechanism suggests the instability of these plasmids in mutants defective in chromosomal completion is due to an inability to complete plasmid replication and not aberrant segregation generating a double-strand break. Secondly, we have shown that stable transformants of two-replisome plasmids into *recBC* mutants cannot be obtained, whether RecA is present or not. If double-strand breaks were being generated in *recBC* mutants, RecA would be required to repair them and rescue the plasmid stability and this effect would be lost in the absence of RecA. The observation that stable two-replisome plasmid transformants cannot be obtained in either *recBC* or *recBC recA* mutants suggests that the plasmid instability in these mutants stems from a deficiency in replication rather than repair.

Interestingly, some of the observations illustrated above appear to conflict with current understanding of roles of these gene products in the completion reaction, but upon closer examination could provide further insight into the mechanism by which replication is completed. Deficiencies in ExoI confer greatly increased stability of bidirectionally-replicating plasmids overtime when compared to wild type cells. This is an interesting result given the necessity of mutation in both ExoI and SbcCD to restore the growth

defects and lack of terminus DNA observed in replicating *recBC* mutants (5, 47). ExoI has been shown *in vitro* to be highly-processive 3'-5' single-stranded exonuclease, yet stops degrading abruptly when encountering double-strand DNA (48). This suggests that ExoI normally might play a role in preventing the aberrant recombinationally-mediated completion process, by degrading the highly recombinogenic 3'-ends of replication intermediates formed as replication forks converge both on the chromosome, and on bidirectional plasmids. This is consistent with the co-requirement of Exo1 and SbcCD in suppressing the completion and viability defects in *recBC* mutants. In the absence of ExoI, the aberrant recombinationally-mediated pathway is more efficient if not less mutagenic due to the persistence of these 3'-ends, a phenotype only observable in this model for studying completion. This is supported by the observation that the *xonA recA* double mutant is more stable than either single mutant, supporting the idea that homologous recombination is not an efficient method of completing replication and only acts to facilitate it in the absence of the normal process.

Several other interesting observations give some insight into the mechanistic completion of DNA replication by the role that the replication fork trap play in the stability of these bidirectional plasmids. Firstly, bidirectional plasmids in *sbcCD* mutants are unstable only if the plasmid contains a replication fork trap. This suggests that the replication intermediates targeted by SbcCD are present in bidirectional plasmids where replication forks are contained with a replication fork trap. The absence of the action of SbcCD results in decreased stability. This is consistent with a model previously proposed, in which one replication fork dissociates the other and replicates on to the opposing

newly replicated daughter strand generating a palindrome-like intermediate recognized and cleaved by SbcCD (5).

Lastly, the observation that bidirectionally-replicating plasmids with a replication fork trap are maintained at wild type levels in *recD* and *sbcCD xonA recD* mutants and yet bidirectional plasmids without one are barely maintained at extremely low levels in *recD* mutants and not at all in *sbcCD xonA recD* mutants suggests a role for RecBCD at the point of fork convergence. This idea is supported by recent work (49). The only difference between *recBC* and *recD* mutants or *sbcCD xonA recBC* and *sbcCD xonA recD* mutants is the presence in the *recD* mutant of a RecBC protein with attenuated helicase and lack of nuclease activity. The differential plasmid stability in the presence of this attenuated RecBC protein suggests a non-nucleolytic action for the enzyme during completion. Furthermore, this suggests that the helicase activity may be the important function to the completion reaction as near wild type plasmid stabilities are observed in *sbcCD xonA recD mutants* when there is a replication fork trap to contain the substrates of the reaction preventing the replication forks from outpacing the attenuated RecBC.

Acknowledgements

This work was supported by the National Science Foundation (Grant MCB0130486) and National Institute of Environmental Health Sciences (Grant R15ES025953). We are especially grateful to Dr. Grzegorz Węgrzyn and Dr. Stuart Austin for the generous gift of plasmids.

Table 4.1 Strains and plasmids used in this study

Strain/Plasmid	Relevant Genotype	Source, Reference, and/or construction
SR108 parental	λ⁻ *thyA deo IN(rrnD-rrnE)*	(1)
HL922	SR108 *recB21C22 argA81::*Tn*10*	(2)
CL851	SR108 *recB21C22 argA81::*Tn*10 recA::cam*	(3)
HL923	SR108 *recD1011 argA81::*Tn*10*	(2)
CL726	SR108 *recD1011 argA81::*Tn*10 recA::cam*	(3)
CL542	SR108 *recA::cam*	(3)
CL039	SR108 *xonA::cam*	(4)
CL718	SR108 *xonA::Cat300 D(srlR-recA)306::*Tn*10*	(5)
CL2344	SR108 *sbcCD::Gm*	(5)
CL3535	SR108 *sbcCD::Gm D(srlR-recA)306::*Tn*10*	(5)
CL2357	SR108 *xonA::Cat300 sbcCD::Gm*	(4)
CL3539	SR108 *xonA::Cat300 sbcCD::Gm D(srlR-recA)306::*Tn*10*	(5)
CL2542	SR108 *xonA::Cat300 sbcCD::Gm recB21C22 argA81::*Tn*10*	(5)
CL2575	SR108 *xonA::Cat300 sbcCD::Gm recB21C22 argA81 D(srlR-recA)306::*Tn*10*	(5)
CL2539	SR108 *xonA::Cat300 sbcCD::Gm recD1011 argA81::*Tn*10*	(5)
CL2577	SR108 *xonA::Cat300 sbcCD::Gm recD1011 argA81 D(srlR-recA)306::*Tn*10*	(5)
CL2456	SR108 *recG6200::tet857*	P1 transduction of *recG6200::tet857* from TP538 (6) into SR108
CL2579	SR108 *recG6200::tet857 recA::cam*	P1 transduction of *recA::cam* from CL542 (3) into CL2456

pBR322	Ampicillin, Tetracyclin resistance, pMB1 origin	(7)
pCLZ	Ampicillin, Chloramphenicol resistance, λ bacteriophage origin	Added AmpR cassette to terminus region of parent plasmid pCB104, a generous gift from Dr. Grzegorz Węgrzyn (8)
pCTZ	Ampicillin, Chloramphenicol resistance, λ bacteriophage origin, replication fork trap	Added AmpR cassette and surrounding replication fork trap using sites *terB* and *terC* in opposition to the terminus region of parent plasmid pCB104, a generous gift from Dr. Grzegorz Węgrzyn (8)

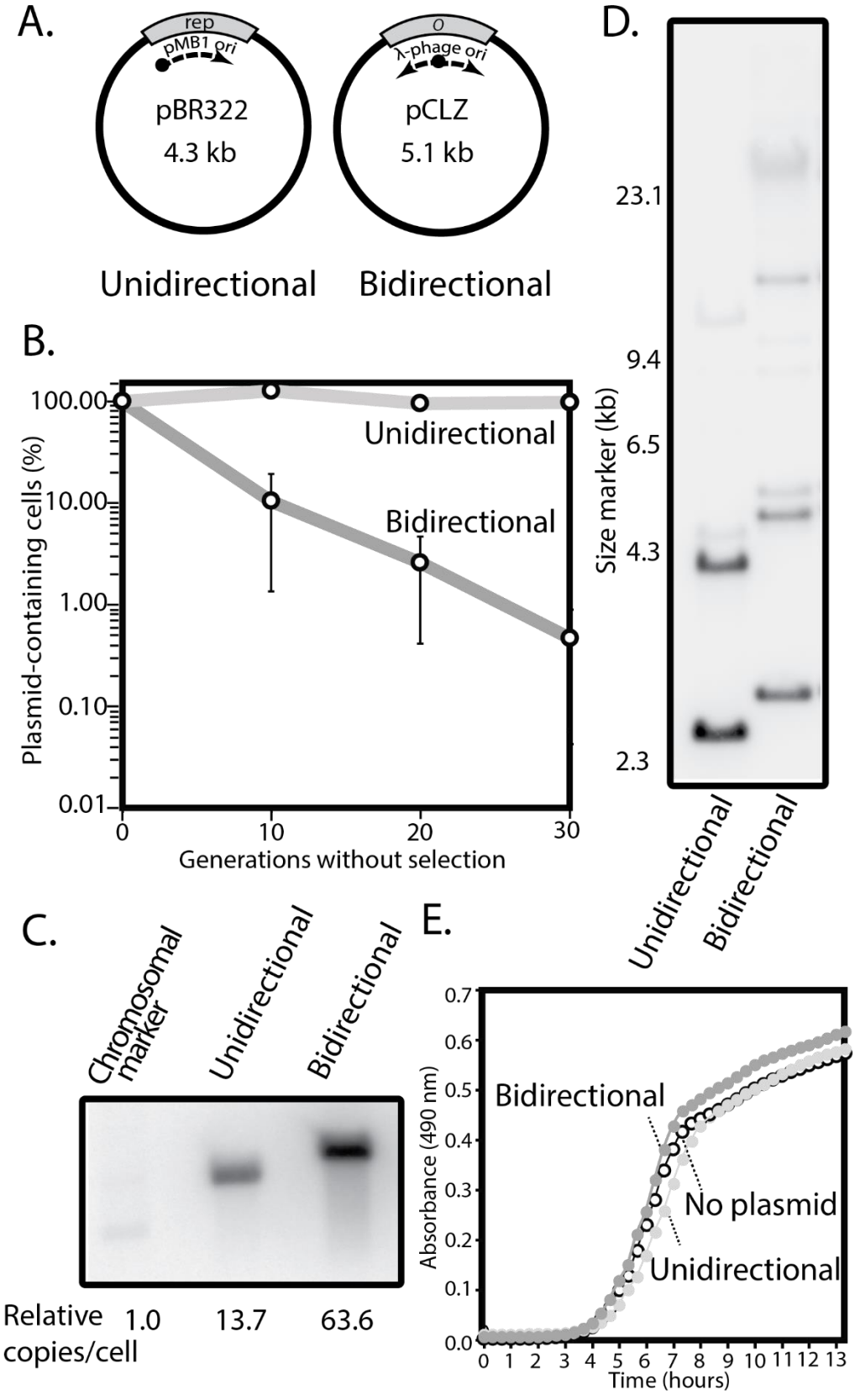
A.
rep
pMB1 ori
pBR322
4.3 kb
o
λ-phage ori
pCLZ
5.1 kb
Unidirectional
Bidirectional
B.
100.00
10.00
1.00
0.10
0.01
Plasmid-containing cells (%)
Unidirectional
Bidirectional
0
10
20
30
Generations without selection
D.
23.1
9.4
6.5
4.3
2.3
Size marker (kb)
Unidirectional
Bidirectional
C.
Chromosomal marker
Unidirectional
Bidirectional
Relative copies/cell
1.0
13.7
63.6
E.
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Absorbance (490 nm)
Bidirectional
No plasmid
Unidirectional
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Time (hours)

Figure 4.1. Two-replisome plasmids are a good model for studying completion. A) Diagram of the general structure of plasmids used in this study. B) Bidirectionally-replicating plasmids are less stable than a unidirectional plasmid of a similar size. Overnight cultures were grown with selection and subcultured without selection to evaluate plasmid stability over subsequent generations. Values are normalized to starting plasmid-containing cells. Data shown is the average of at least 3 independent experiments C). One and two replisome plasmids can be maintained at copy numbers similar to those previously reported. Whole genomic DNA samples were digested with EcoRV, electrophoresed at 1 V/cm, analyzed by Southern blotting and labelled with a P32-labelled PCR fragment of the ampicillin resistance cassette. D) Bidirectionally-replicating plasmids are maintained in highly multimeric forms as compared to unidirectional plasmids. This multimerization is enhanced by the presence of a replication fork trap. Total DNA extractions containing were digested with SacII which has no restriction sites on either plasmid, were electrophoresed at 1 V/cm, and analyzed by Southern blotting. A representative blot is shown. E) Cells containing these plasmids maintain a growth rate similar to that observed in plasmid-free cells. Absorbance (490nm) of cultures grown at 37°C is plotted over time. 10µL drops from 10-fold serial dilutions of overnight cultures were plated to ensure the same starting viable cell count. Colonies were observed following overnight 37C incubation. A representative plot is shown.

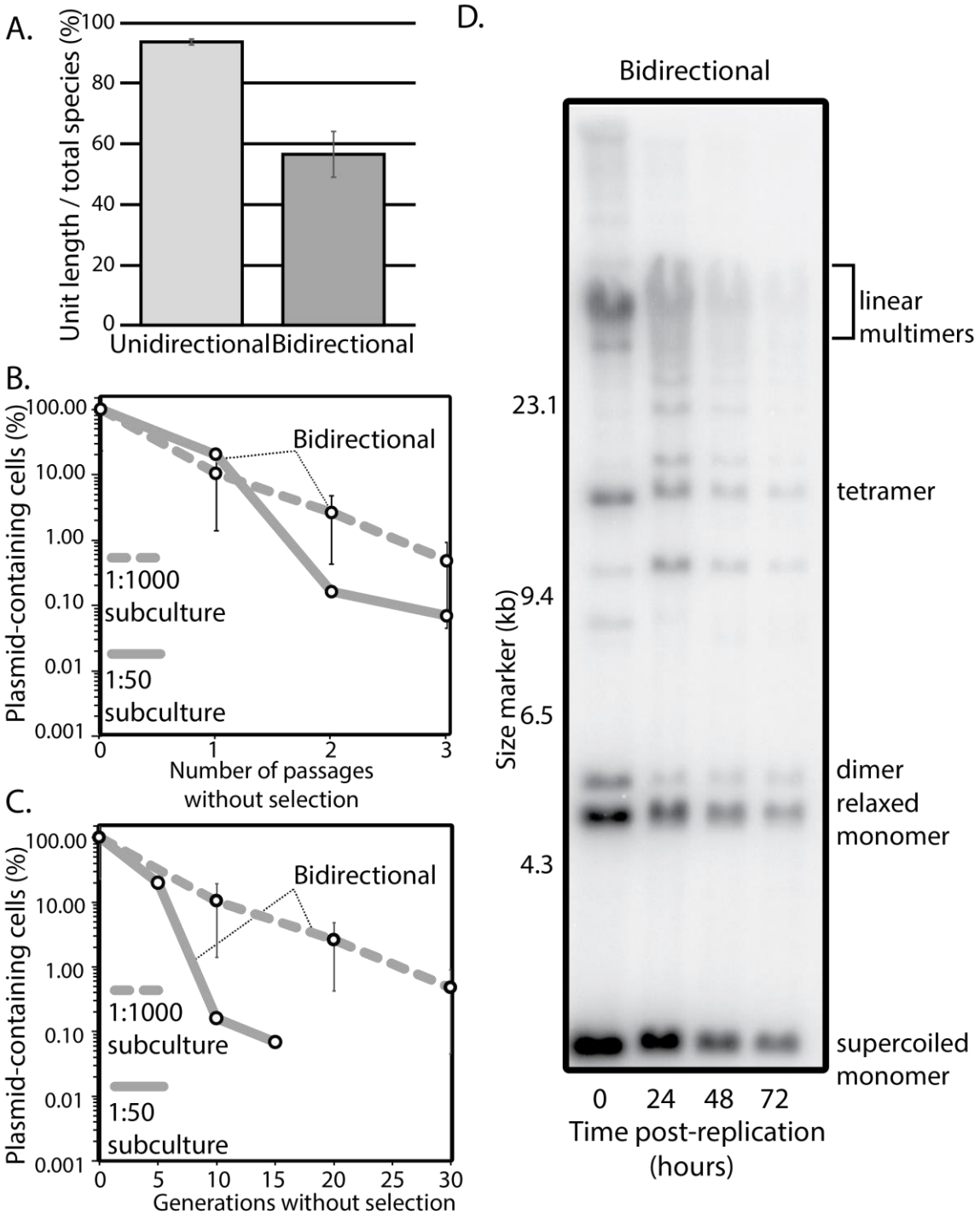
A.
Unit length / total species (%)
100
80
60
40
20
0
Unidirectional
Bidirectional
B.
Plasmid-containing cells (%)
100.00
10.00
1.00
0.10
0.01
0.001
Bidirectional
1:1000
subculture
1:50
subculture
0
1
2
3
Number of passages
without selection
C.
Plasmid-containing cells (%)
100.00
10.00
1.00
0.10
0.01
0.001
Bidirectional
1:1000
subculture
1:50
subculture
0
5
10
15
20
25
30
Generations without selection
D.
Bidirectional
linear
multimers
23.1
tetramer
9.4
6.5
dimer
relaxed
monomer
4.3
supercoiled
monomer
Size marker (kb)
0
24
48
72
Time post-replication
(hours)

Figure 4.2. Multimeric species contribute to instability when plasmids are not replicating. A) Unit length plasmid species represent a larger fraction of total plasmid species in unidirectional versus bidirectional plasmids. Signal representing unit length plasmid species was compared to total plasmid signal. Data shown is the average and standard error of three independent experiments. B) Bidirectionally-replicating plasmids are lost at a similar rate per passage. C) Yet when the same data is evaluated by the number of generations to reach stationary phase appear to differentially lose the plasmid. Overnight cultures were grown with selection and subcultured without selection to evaluate plasmid stability over subsequent generations. 10µL drops from 10-fold serial dilutions of the resulting overnight cultures were plated with and without selection and enumerated following overnight 37C incubation. Values are normalized to starting plasmid-containing cells. Data shown is the average of at least 3 independent experiments. D) Multimeric species of bidirectional plasmids are degraded *in vivo* post-replication. Samples of cultures containing each plasmid were taken every 24 hours and Southern analysis identified persistence of plasmid species.

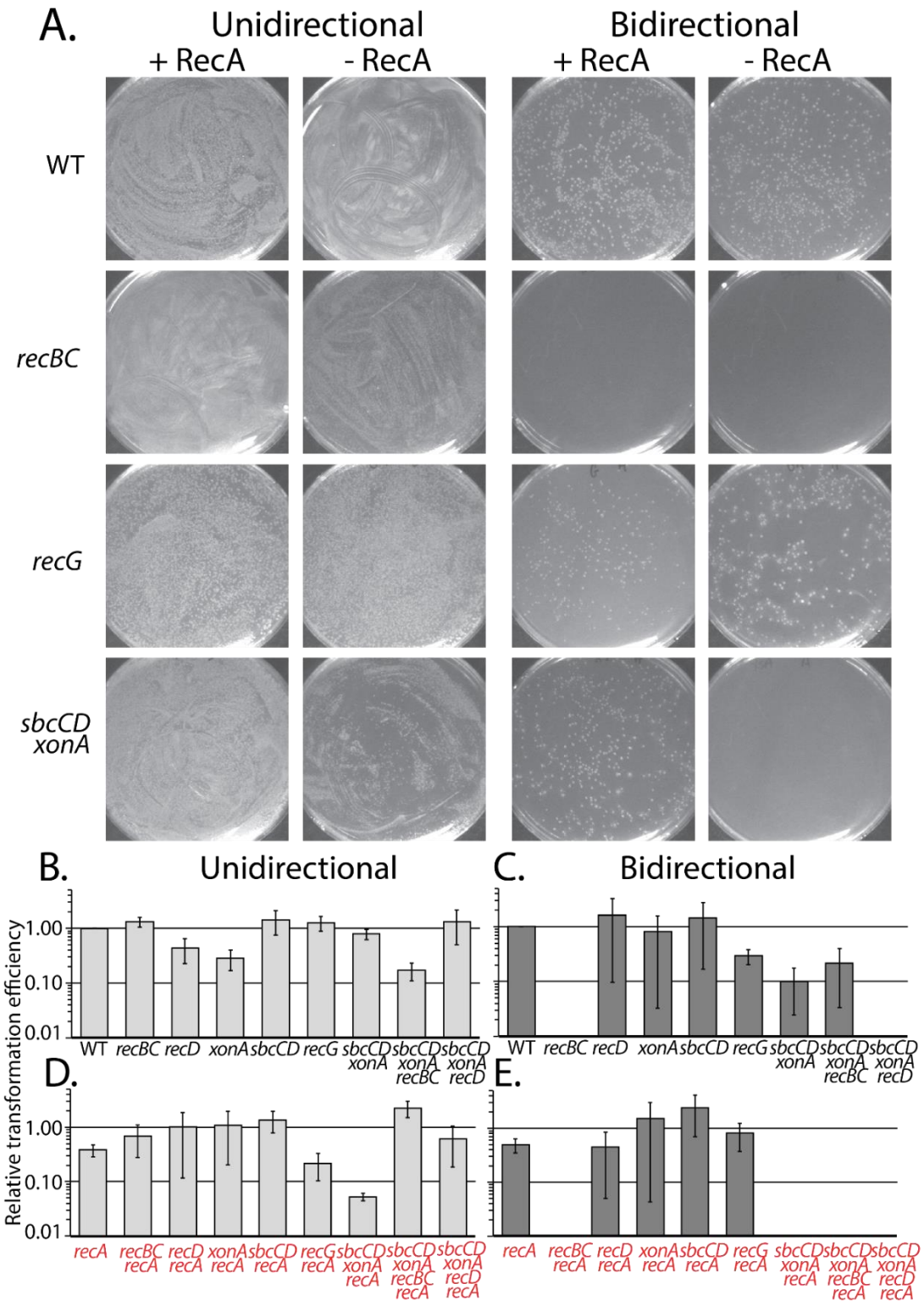
A.
Unidirectional
Bidirectional
+ RecA
- RecA
+ RecA
- RecA
WT
recBC
recG
sbcCD xonA
B.
Unidirectional
C.
Bidirectional
Relative transformation efficiency
1.00
0.10
0.01
WT
recBC
recD
xonA
sbcCD
recG
sbcCD xonA
sbcCD xonA recBC
sbcCD xonA recD
D.
E.
recA
recBC recA
recD recA
xonA recA
sbcCD recA
recG recA
sbcCD xonA recA
sbcCD xonA recBC recA
sbcCD xonA recD recA

Figure 4.3. Transformation of two replisome plasmids into strains lacking RecBCD and SbcCD-ExoI requires RecA. A). Strains lacking RecBCD are not efficiently transformed with bidirectionally replicating plasmids with or without RecA. Strains lacking SbcCD-ExoI require RecA for efficient transformation. Wild type cells or *recG* mutant cells transform efficiently with or without RecA. Photographs of plates containing representative transformants for each strain and plasmid are shown. B) All strains tested are effectively transformed with unidirectional plasmid. C) *RecBC* and s*bcCD xonA recD* mutants cannot be transformed with two-replisome plasmids. D) Strains lacking RecBCD and SbcCD-ExoI do not require RecA to be transformed with unidirectional plasmid, E) but do require it to be transformed with bidirectional plasmid. Percentages of transformants to viable cells are shown relative to that of wild type cells for each plasmid. The averages and standard error for data from two independent experiments are shown. A lack of relative transformation efficiency measurement indicates less than 5 viable transformants per transformation. The same preparation of competent cells for each strain was used for all plasmids.

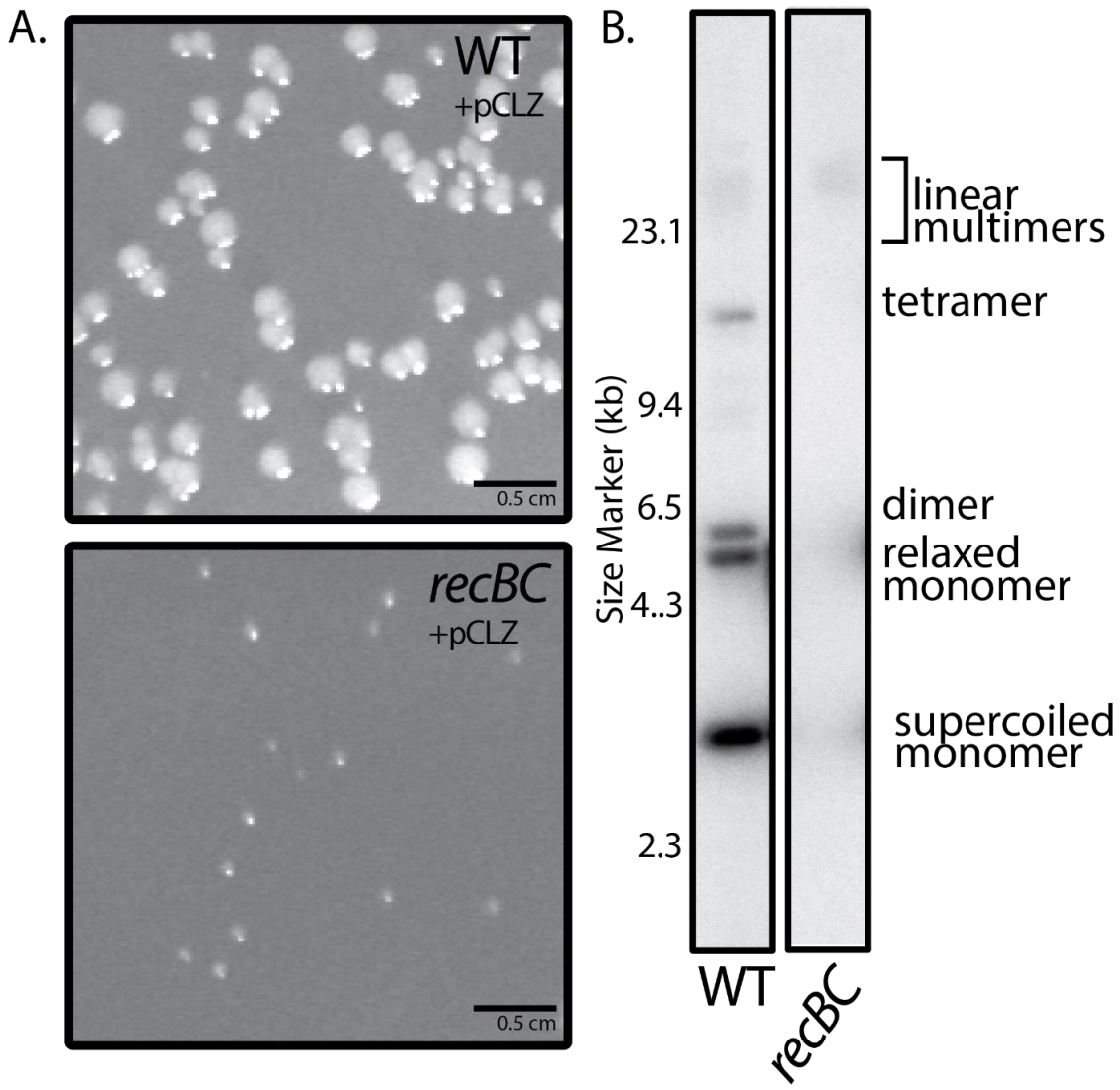

Figure 4.4. *recBC* mutants form microcolonies after transformation with two-replisome plasmids that cannot be propagated. A) Wild type cells transformed with pCLZ, a two-replisome plasmid form robust colonies capable of further culture, while strains lacking RecBCD form microcolonies after 48 hours of incubation at 37°C that cannot be propagated. B) DNA purified directly from the scraped microcolonies exhibited only linear multimers and lacked unit length species as determined by Southern analysis.

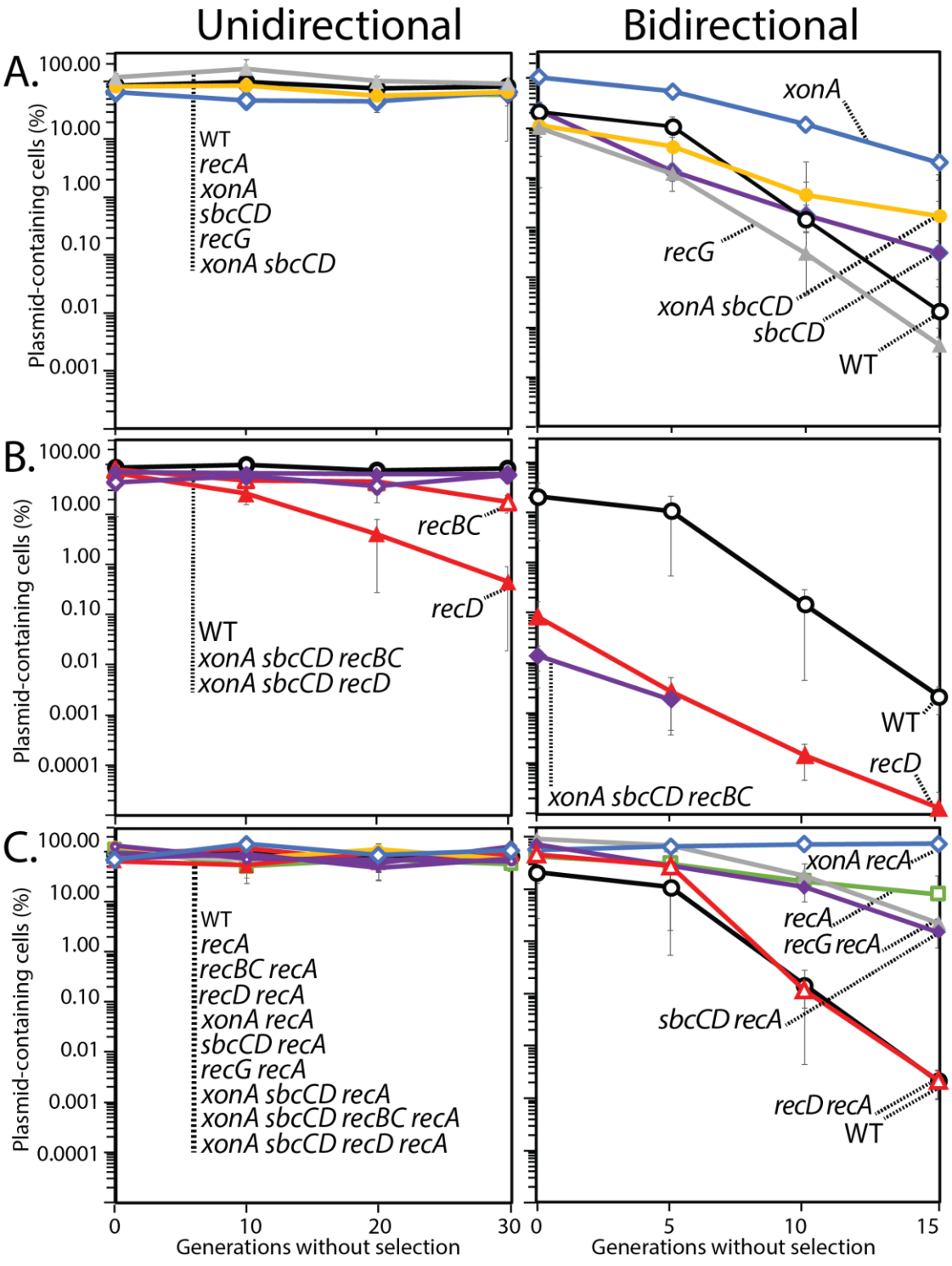
Unidirectional
Bidirectional
A.
B.
C.
Plasmid-containing cells (%)
100.00
10.00
1.00
0.10
0.01
0.001
0.0001
WT
recA
xonA
sbcCD
recG
xonA sbcCD
recBC
recD
xonA sbcCD recBC
xonA sbcCD recD
xonA recA
recA
recG recA
sbcCD recA
recD recA
xonA sbcCD recA
xonA sbcCD recBC recA
xonA sbcCD recD recA
0
10
20
30
5
15
Generations without selection

Figure 4.5. Mutants with defects in completing replication have greater instability. A) Mutants deficient in ExoI, SbcCD, and recG maintain one- and two-replisome plasmids at wild type rates. B) Strains lacking RecBCD and SbcCD-ExoI exhibit improved stability in one-replisome plasmids but not in two-replisome plasmids. C) Loss of RecA has no effect on the stability of one-replisome plasmids, but in two-replisome plasmids, strains lacking RecA, or homologous recombination are all far more stable than wildtype cells. This suggests that RecA is responsible for the instability in wild type cells. Overnight cultures were grown with selection and subcultured without selection to evaluate plasmid stability over subsequent generations. 10µL drops from 10-fold serial dilutions of the resulting overnight cultures were plated with and without selection and enumerated following overnight 37C incubation. Values are normalized to starting plasmid-containing cells. Data shown is the average of at least 3 independent experiments. Error bars depicted represent the standard error of the mean.

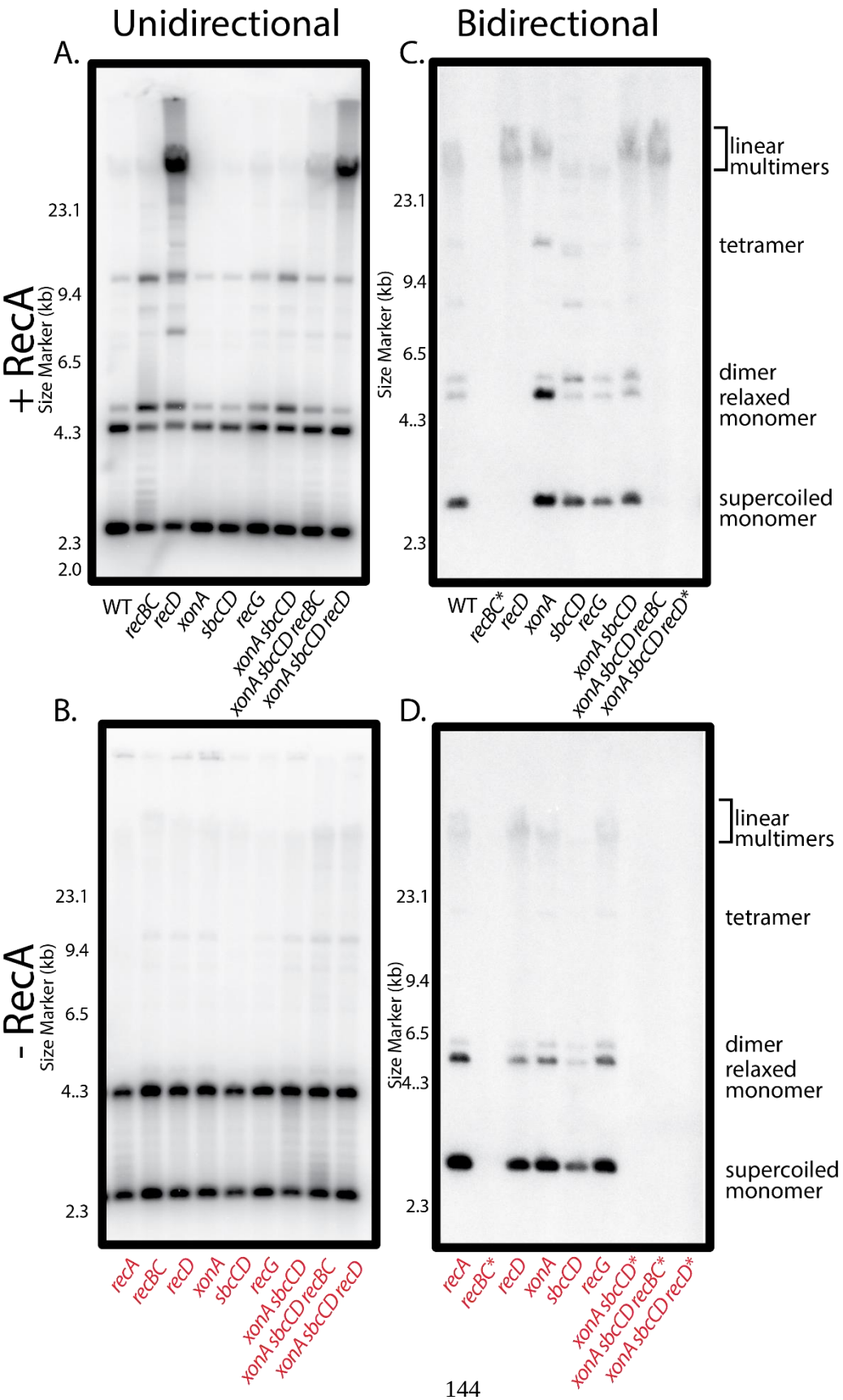
Unidirectional
Bidirectional
A.
C.
B.
D.
+ RecA
- RecA
Size Marker (kb)
23.1
9.4
6.5
4..3
2.3
2.0
WT
recBC
recD
xonA
sbcCD
recG
xonA sbcCD
xonA sbcCD recBC
xonA sbcCD recD
recBC*
recD*
recA
linear multimers
tetramer
dimer relaxed monomer
supercoiled monomer

Figure 4.6. Multimeric plasmid species are more prevalent in bidirectionally-replicating plasmids in mutants defective for completion. A) Unidirectional plasmids are maintained stably in most strains and multimerization is dependent upon RecA. B) Strains lacking RecA exhibit nearly all monomeric species. C) Two-replisome plasmids exhibit higher degrees of multimerization, that unidirectional and require gene products required for chromosomal completion. D) Without RecA, completion deficient strains cannot maintain the plasmids and strains that can exhibit mostly monomeric species. Strains marked with an asterisk (*) were unable to be transformed and therefore no sample was run but a lane was left open for clarity. Total DNA extractions containing uncut plasmid DNA were electrophoresed at 1 v/cm and probed with the respective ^{32}P-labelled plasmid. Representative blots are shown.

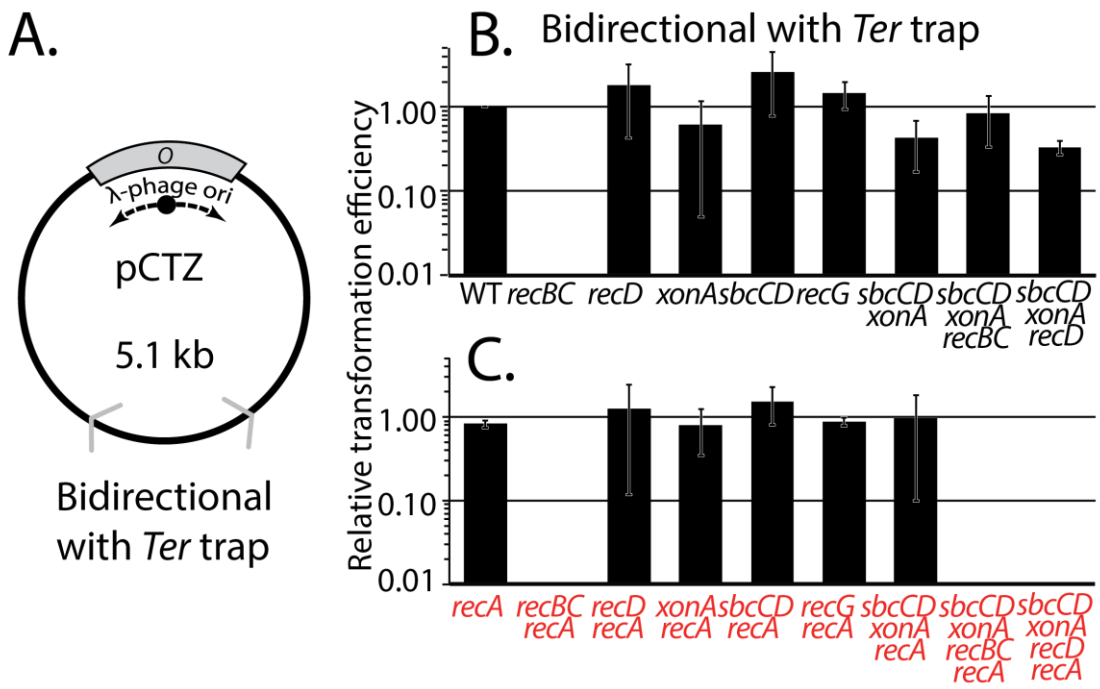

Figure 4.7. Transformation of two replisome plasmids containing a *ter* trap is more efficient in strains lacking RecBCD and SbcCD-ExoI and still requires RecA. A) Diagram of the general structure of the plasmid containing a fork trap used in this study. B) Strains lacking RecD and SbcCD-ExoI and RecD are more efficiently transformed with a two-replisome plasmid containing a *ter* trap. C) Transformation of the bidirectional plasmid with a *ter* trap in *sbcCD xonA recBC* and *sbcCD xonA recD* mutants still requires RecA. Experiment was performed as in Figure 3. The averages and standard error for data from two independent experiments are shown.

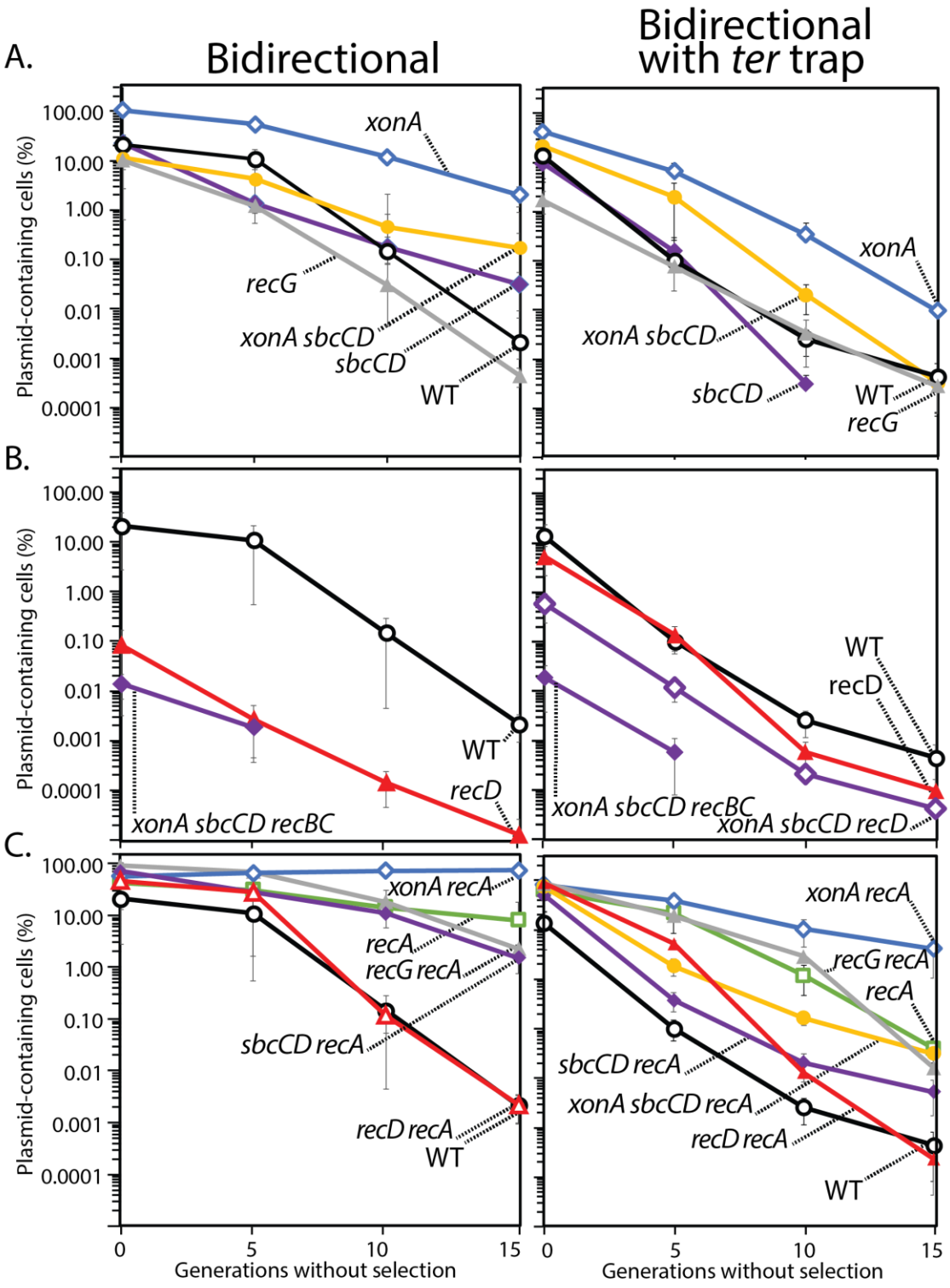
A.
Bidirectional
Bidirectional with ter trap
Plasmid-containing cells (%)
100.00
10.00
1.00
0.10
0.01
0.001
0.0001
xonA
recG
xonA sbcCD
sbcCD
WT
B.
xonA sbcCD recBC
recD
xonA sbcCD recD
C.
xonA recA
recA
recG recA
sbcCD recA
recD recA
0
5
10
15
Generations without selection

Figure 4.8. A replication fork trap increases the stability of two-replisome plasmids in strains lacking RecBCD and SbcCD-ExoI. A) Two-replisome plasmids are less stable when they contain a *ter* trap B) Strains lacking RecBCD and SbcCD-ExoI exhibit improved stability with a replication fork trap. C) Strains lacking RecA, or homologous recombination are all more stable than wild type cells with and without a *ter* trap. This experiment was performed as in Figure 4.

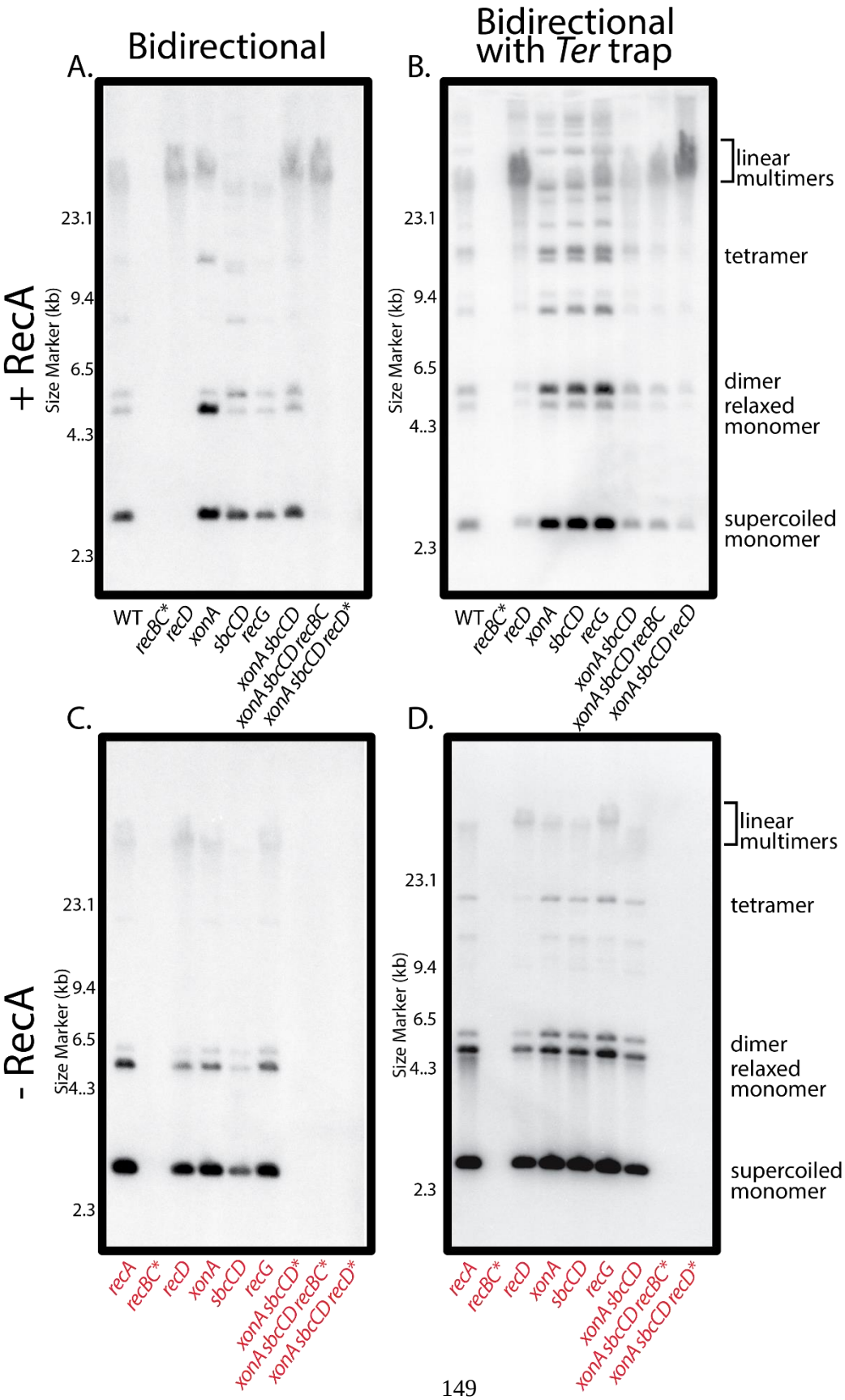
A.
Bidirectional
B.
Bidirectional
with Ter trap
+ RecA
- RecA
Size Marker (kb)
23.1
9.4
6.5
4..3
2.3
linear
multimers
tetramer
dimer
relaxed
monomer
supercoiled
monomer
WT
recBC*
recD
xonA
sbcCD
recG
xonA sbcCD
xonA sbcCD recBC
xonA sbcCD recD
recA
recBC*
recD
xonA
sbcCD
recG
xonA sbcCD
xonA sbcCD recBC*
xonA sbcCD recD*
C.
D.

Figure 4.9. Multimeric plasmid species are more prevalent in two-replisome plasmids with a *ter* trap. A) B) Blots of the two-replisome plasmid are show for comparison. C) The presence of a *ter* trap within a bidirectional plasmid stabilizes the plasmid in many of the mutants tested despite the increased incidence of multimerization. D) This multimerization is dependent upon RecA, and the stabilizing effect of the *ter* trap is still observable in completion mutants lacking RecA. Experiment was performed as in Figure 5. Representative blots are shown.

References

1. Burrell RA, et al. (2013) Replication stress links structural and numerical cancer chromosomal instability. *Nature* 494(7438):492–496.
2. Chan KL, Palmai-Pallag T, Ying S, Hickson ID (2009) Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol* 11(6):753–760.
3. Courcelle J, Wendel BM, Livingstone DD, Courcelle CT (2015) RecBCD is required to complete chromosomal replication: Implications for double-strand break frequencies and repair mechanisms. *DNA Repair* 32:86–95.
4. Wendel BM, Courcelle CT, Courcelle J (2014) Completion of DNA replication in *Escherichia coli*. *Proc Natl Acad Sci*:201415025.
5. Wendel BM, Cole JM, Courcelle CT, Courcelle J (2017) SbcC-SbcD and ExoI process convergent forks to complete chromosome replication. *Proc Natl Acad Sci*:201715960.
6. Sinha AK, et al. (2017) Division-induced DNA double strand breaks in the chromosome terminus region of *Escherichia coli* lacking RecBCD DNA repair enzyme. *PLoS Genet* 13(10):e1006895.
7. Fang FC, Helinski DR (1991) Broad-host-range properties of plasmid RK2: importance of overlapping genes encoding the plasmid replication initiation protein TrfA. *J Bacteriol* 173(18):5861–5868.
8. Aizenman E, Engelberg-Kulka H, Glaser G (1996) An *Escherichia coli* chromosomal “addiction module” regulated by guanosine [corrected] 3’,5’-bispyrophosphate: a model for programmed bacterial cell death. *Proc Natl Acad Sci U S A* 93(12):6059–6063.

9. Smith ASG, Rawlings DE (1998) Autoregulation of the pTF-FC2 Proteic Poison-Antidote Plasmid Addiction System (pas) Is Essential for Plasmid Stabilization. *J Bacteriol* 180(20):5463–5465.
10. Atlung T, Christensen BB, Hansen FG (1999) Role of the Rom Protein in Copy Number Control of Plasmid pBR322 at Different Growth Rates in *Escherichia coli* K-12. *Plasmid* 41(2):110–119.
11. Inoue N, Uchida H (1991) Transcription and initiation of ColE1 DNA replication in *Escherichia coli* K-12. *J Bacteriol* 173(3):1208–1214.
12. Veltkamp E, Nijkamp HJJ (1974) Replication of the Bacteriocinogenic Plasmid Clo DF13 in Thermosensitive *Escherichia coli* Mutants Defective in Initiation or Elongation of Deoxyribonucleic Acid Replication. *J Bacteriol* 120(3):1227–1237.
13. Kingsbury DT, Sieckmann DG, Helinski DR (1973) Temperature-Sensitive Mutants for the Replication of Plasmids in *ESCHERICHIA COLI* II. Properties of Host and Plasmid Mutations. *Genetics* 74(1):1–16.
14. Lovett MA, Sparks RB, Helinski DR (1975) Bidirectional replication of plasmid R6K DNA in *Escherichia coli*; correspondence between origin of replication and position of single-strand break in relaxed complex. *Proc Natl Acad Sci* 72(8):2905–2909.
15. Eichenlaub R, Figurski D, Helinski DR (1977) Bidirection replication from a unique origin in a mini-F plasmid. *Proc Natl Acad Sci* 74(3):1138–1141.
16. Matsubara K (1974) Preparation of Plasmid λdv from Bacteriophage λ: Role of Promoter-Operator in the Plasmid Replicon. *J Virol* 13(3):596–602.
17. Cooper NS, Brown ME, Caulcott CA (1987) A Mathematical Method for Analysing

Plasmid Stability in Micro-organisms. *Microbiology* 133(7):1871–1880.

18. Ayala-Sanmartin J, Gómez-Eichelmann MC (1989) Stability of ColE1-like and pBR322-like plasmids in *Escherichia coli*. *Mol Microbiol* 3(12):1745–1752.
19. Summers DK (1991) The kinetics of plasmid loss. *Trends Biotechnol* 9(8):273–278.
20. Summers D (1998) Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol Microbiol* 29(5):1137–1145.
21. Werbowy O, Werbowy S, Kaczorowski T (2017) Plasmid stability analysis based on a new theoretical model employing stochastic simulations. *PLOS ONE* 12(8):e0183512.
22. Bentley WE, Mirjalili N, Andersen DC, Davis RH, Kompala DS (1990) Plasmid-encoded protein: The principal factor in the "metabolic burden" associated with recombinant bacteria. *Biotechnol Bioeng* 35(7):668–681.
23. Hill TM, Tecklenburg ML, Pelletier AJ, Kuempel PL (1989) tus, the trans-acting gene required for termination of DNA replication in *Escherichia coli*, encodes a DNA-binding protein. *Proc Natl Acad Sci* 86(5):1593–1597.
24. Coskun-Ari FF, Hill TM (1997) Sequence-specific interactions in the Tus-Ter complex and the effect of base pair substitutions on arrest of DNA replication in *Escherichia coli*. *J Biol Chem* 272(42):26448–26456.
25. Duggin IG, Bell SD (2009) Termination Structures in the *Escherichia coli* Chromosome Replication Fork Trap. *J Mol Biol* 387(3):532–539.
26. Roecklein B, Pelletier A, Kuempel P (1991) The tus gene of *Escherichia coli*: autoregulation, analysis of flanking sequences and identification of a complementary system in Salmonella typhimurium. *Res Microbiol* 142(2–3):169–

175.

27. Iismaa TP, Wake RG (1987) The normal replication terminus of the *Bacillus subtilis* chromosome, terC, is dispensable for vegetative growth and sporulation. *J Mol Biol* 195(2):299–310.
28. Sawitzke JA, et al. (2012) The segregation of *Escherichia coli* minichromosomes constructed in vivo by recombineering. *Plasmid* 67(2):148–154.
29. Casini A, et al. (2014) One-pot DNA construction for synthetic biology: the Modular Overlap-Directed Assembly with Linkers (MODAL) strategy. *Nucleic Acids Res* 42(1):e7–e7.
30. Gibson DG, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343–345.
31. Słomińska M, Węgrzyn A, Konopa G, Skarstad K, Węgrzyn G (2001) SeqA, the *Escherichia coli* origin sequestration protein, is also a specific transcription factor. *Mol Microbiol* 40(6):1371–1379.
32. Moore DD, Denniston-Thompson K, Furth ME, Williams BG, Blattner FR (1977) Construction of Chimeric Phages and Plasmids Containing the Origin of Replication of Bacteriophage Lambda. *Science* 198(4321):1041–1046.
33. Anderl A, Klein A (1982) Replication of lambda dv DNA in vitro. *Nucleic Acids Res* 10(5):1733–1740.
34. Taylor K, Wegrzyn G (1995) Replication of coliphage lambda DNA. *FEMS Microbiol Rev* 17(1–2):109–119.
35. Potrykus K, Baranska S, Wegrzyn A, Węgrzyn G (2002) Composition of the lambda plasmid heritable replication complex. *Biochem J* 364(Pt 3):857–862.

36. Bolivar F, et al. (1977) Construction and characterization of new cloning vehicle. II. A multipurpose cloning system. *Gene* 2(2):95–113.

37. Rudolph CJ, Upton AL, Stockum A, Nieduszynski CA, Lloyd RG (2013) Avoiding chromosome pathology when replication forks collide. *Nature* 500(7464):608–611.

38. Azeroglu B, et al. (2016) RecG Directs DNA Synthesis during Double-Strand Break Repair. *PLOS Genet* 12(2):e1005799.

39. Dimude JU, Midgley-Smith SL, Stein M, Rudolph CJ (2016) Replication Termination: Containing Fork Fusion-Mediated Pathologies in *Escherichia coli*. *Genes* 7(8):40.

40. Biek DP, Cohen SN (1986) Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in *Escherichia coli*. *J Bacteriol* 167(2):594–603.

41. Bedale WA, Inman RB, Cox MM (1993) A reverse DNA strand exchange mediated by recA protein and exonuclease I. The generation of apparent DNA strand breaks by recA protein is explained. *J Biol Chem* 268(20):15004–15016.

42. Butland G, et al. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 433(7025):531–537.

43. Roman LJ, Kowalczykowski SC (1989) Characterization of the helicase activity of the *Escherichia coli* RecBCD enzyme using a novel helicase assay. *Biochemistry (Mosc)* 28(7):2863–2873.

44. Taylor AF, Smith GR (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* 423(6942):889–893.

45. Dillingham MS, Spies M, Kowalczykowski SC (2003) RecBCD enzyme is a bipolar

DNA helicase. *Nature* 423(6942):893–897.

46. Sinha AK, et al. (2018) Broken replication forks trigger heritable DNA breaks in the terminus of a circular chromosome. *PLoS Genet* 14(3):e1007256.
47. Gibson FP, Leach DR, Lloyd RG (1992) Identification of sbcD mutations as cosuppressors of recBC that allow propagation of DNA palindromes in *Escherichia coli* K-12. *J Bacteriol* 174(4):1222–1228.
48. Richardson CC, Kornberg A (1964) A deoxyribonucleic acid phosphatase-exonuclease from *Escherichia coli* I. Purification of the enzyme and characterization of the phosphatase activity. *J Biol Chem* 239(1):242–250.
49. Laureti L, Lee L, Philippin G, Pagès V (2017) A non-catalytic role of RecBCD in homology directed gap repair and translesion synthesis. *Nucleic Acids Res* 45(10):5877–5886.

Chapter 5

Concluding remarks

The completion of DNA replication is a fundamental process essential to genomic stability during DNA replication. In the previous chapters I have shown that the completion process involves the complex helicase/nuclease RecBCD, the structure-specific endonuclease SbcCD, and the single-stranded exonuclease ExoI (1, 2). These enzymes process convergent replication forks independently of homologous recombination and without RecA (1). In the absence of the gene products, the reaction is shunted to an aberrant pathway requiring homologous recombination and RecA (2). This aberrant pathway can result in genomic instability, insertions, deletions and rearrangements as observed by chromosomal and plasmid instabilities.

Based primarily on biochemical evidence, RecBCD has been proposed to act during double-strand break repair via homologous recombination to degrade the DNA from the double-strand DNA end, generating a 3’ end while actively loading RecA onto the single-strand end (3–6). As previously discussed, several early studies on the role of RecBCD proposed a role for RecBCD late in the resolution of recombinant DNA (7, 8). This was proposed based on the observation that recombinant molecules form in conjugating recBC mutants at similar rates to wild type cells, but over time the recBC mutants lost viability. By contrast, no recombinant molecules were formed or detected in recA mutants, suggesting that RecBCD acts late, after RecA, to finish the recombination reaction. This supports the idea that the function of RecBCD in double-strand break

repair via homologous recombination may not be in degrading from double-strand ends, but that RecBCD is acting later to complete the repair event.

Interestingly, the differential plasmid stabilities observed in bidirectionally-replicating plasmids with and without a replication fork trap in *recD* and *sbcCD xonA recD* supports a role for RecBCD at the point of fork convergence and suggests a possible non-nucleolytic action for the enzyme during completion. Recent work has demonstrated RecBCD playing a fundamental yet non-catalytic role in a repair event, and it could be inferred to be playing a similar role here (9). The only difference between *recBC* and *recD* mutants or *sbcCD xonA recBC* and *sbcCD xonA recD* mutants is the presence in the *recD* mutant of a RecBC protein with attenuated helicase and lack of nuclease activity and could support the idea of a non-nucleolytic action for the enzyme during completion. This combined with the observation that the *recD* mutant is recombination proficient despite lacking nuclease activity (10–12), suggests that the function of RecBCD in double-strand break repair via homologous recombination may not be degrading from double-strand ends, and that RecBCD is acting later to complete the repair event. An alternative explanation could simply be that the helicase activity may be the important function to the completion reaction as near wild-type plasmid stabilities are observed in *sbcCD xonA recD mutants* when there is a replication fork trap to contain the substrates of the reaction preventing the replication forks from outpacing the attenuated RecBC.

Together, the replication profiles and plasmid model studies demonstrate a role for RecBCD in completing DNA replication distinct from double-strand break repair via homologous recombination. Yet the identification of this new role during DNA

replication could have fundamental implications for its role in double-strand break repair. Many of the biochemical activities and mutant characteristics of this enzyme support the idea that RecBCD is acting late in the process to facilitate the completion of normal chromosomal replication. This leads to the idea that RecBCD could be acting to complete replication resulting from double-strand break repair in the same manner as it does during normal chromosomal replication. The data presented in this work here clearly show that RecBCD acts to complete replication and that this model of action agrees with some of the first interpretations of RecBCD function (7). This strongly supports the idea that RecBCD is fundamentally important for both replication and recombination and could be processing the same substrates in either event.

Striking similarities in structure between DNA intermediates associated with two converging replication forks and those that appear double-strand break repair homologous recombination which suggest a similar mechanism for processing (Figure 5.1). During homologous recombination, RecA is essential and facilitates strand invasion and homology recognition (13–15). Current models propose that after the broken molecule is paired with homologous sequence, the replisome is loaded and the DNA is resynthesized. Both ends of the broken chromosome replicate towards each other ultimately converging and completing resynthesis. The converging replication forks occurring during resynthesis are would generate the same structure as that arising from the converging replication forks resulting from normal chromosomal replication. Given the similarities between these two structures, it is possible that RecBCD is processing the two different occurrences of converging replication forks in an equivalent manner. This would suggest that during double-strand break repair, RecBCD is acting to resolve

converging replication forks resulting from resynthesis of the broken DNA, rather than initiating the process.

The rescue of the completion event by homologous recombination poses some interesting question for evolution and fitness of all cells. If recombination is required to resolve convergent replication forks during completion in the absence of the normal pathway involving RecBCD, SbcCD, and ExoI, then it would be advantageous to the cell to increase the aberrant pathway to promote resolution. Yet completion occurring via the aberrant pathway can frequently result in duplications, deletions and chromosomal rearrangements. This increased chromosomal instability occurs in cells capable of the aberrant pathway even in cells with an intact normal completion pathway. This highlights the ultimate bargain undertaken by a cell in controlling the completion of DNA replication. The cell must be balancing the upregulation of the aberrant pathway to ensure the completion of DNA replication in the event the normal pathway is not functional and suppressing the aberrant pathway to avoid the resulting genomic instabilities driven by the recombinationally-mediated mechanism even when the normal pathway is intact.

The striking fundamental nature of this process for completing DNA replication makes it tempting to speculate on the level of conservation of the process across diverse taxa. Mechanistically, converging replication forks would be even more fundamental to organisms replicating their chromosomes with multiple origins. (16–18) By looking at the protein players we have identified in the process, one could infer no high degree of conservation in the process when looking at RecBCD. RecBCD is not even conserved across bacteria (19). Yet the other players in completion that we identified suggest a much higher level of conservation

SbcCD are the bacterial orthologs of the double-strand break proteins Mre11-Rad50. Mre11-Rad50 are conserved across a wide range of evolutionarily divergent organisms as determined by similarities in structure and biochemical activities. SbcCD is a structure-specific endonuclease that targets DNA hairpins, or secondary structure arising at palindromic sequences. This high degree of conservation of SbcCD suggests that the mechanism of completion may be equally conserved across a wide range of organisms.

Future directions

While using yeast as a model organism is typically the next logical step in examining the level of conservation of this process, there are many difficulties with using this technique in a yeast system. The linear structure of the chromosomes and number of origins in *Saccharomyces cerevisiae* coupled with the variability in origin usage makes identifying the region where forks are converging difficult. Additionally, there is no known replication fork trap, like that found in bacteria, to constrain this reaction to a particular region in all cells in the population. This along with a well-defined cell cycle in yeast provides a unique challenge for study.

Bacillus subtilis, are a species of Gram-positive bacteria that maintains a replication fork trap similar to *E. coli*. This species is particularly interesting because it does not have all the same genes identified to play a role completion in *E. coli*, and those that are conserved have different mutant phenotypes and binding-partners. RecBCD is not conserved, but a protein AddAB is thought to be functionally analogous and would

warrant further study (19, 20). SbcCD is conserved but appears to have different *in vivo* activity and a third player SbcE may alternatively play a role (21, 22). Not only does the chromosomal structure and the interesting completion protein conservation make this an excellent model for this type of study, but the ease of genetic manipulation and availability of mutant libraries makes it highly amenable to these types of studies.

Another species that would make an excellent model for this type of replication study is the archaeon *Sulfolobus solfataricus*. DNA replication has been well-studied in *Sulfolobus*, compared to other archaea, and has some very unique properties. In general, the replication machinery in archaea more closely matches that in eukaryotes, than bacteria (21, 22). Secondly, *S. solfataricus* has a unique circular chromosome with three bidirectional origins of replication unevenly distributed across the chromosome (25). This model system provides a unique opportunity to understand the level of conservation as well as how multiple origins affect completion, without the complexities associated with eukaryotes. While not as thoroughly manipulated as some model bacteria, there are genetic tools to engineer mutations, making this a tempting model for study.

By understanding the role that SbcCD is playing in the completion DNA replication in *E. coli*, valuable insight can be obtained into the potential roles of Mre11-Rad50 in the completion of DNA replication in other organisms. Not only would this help us identify a fundamental process by which all converging replication forks are resolved across evolutionarily distant organisms, but by furthering understanding the novel and fundamental process by which convergent replisomes are resolved across diverse taxa, we can better understand the mechanism by which the normal process becomes perturbed in diseases as diverse as cancer.

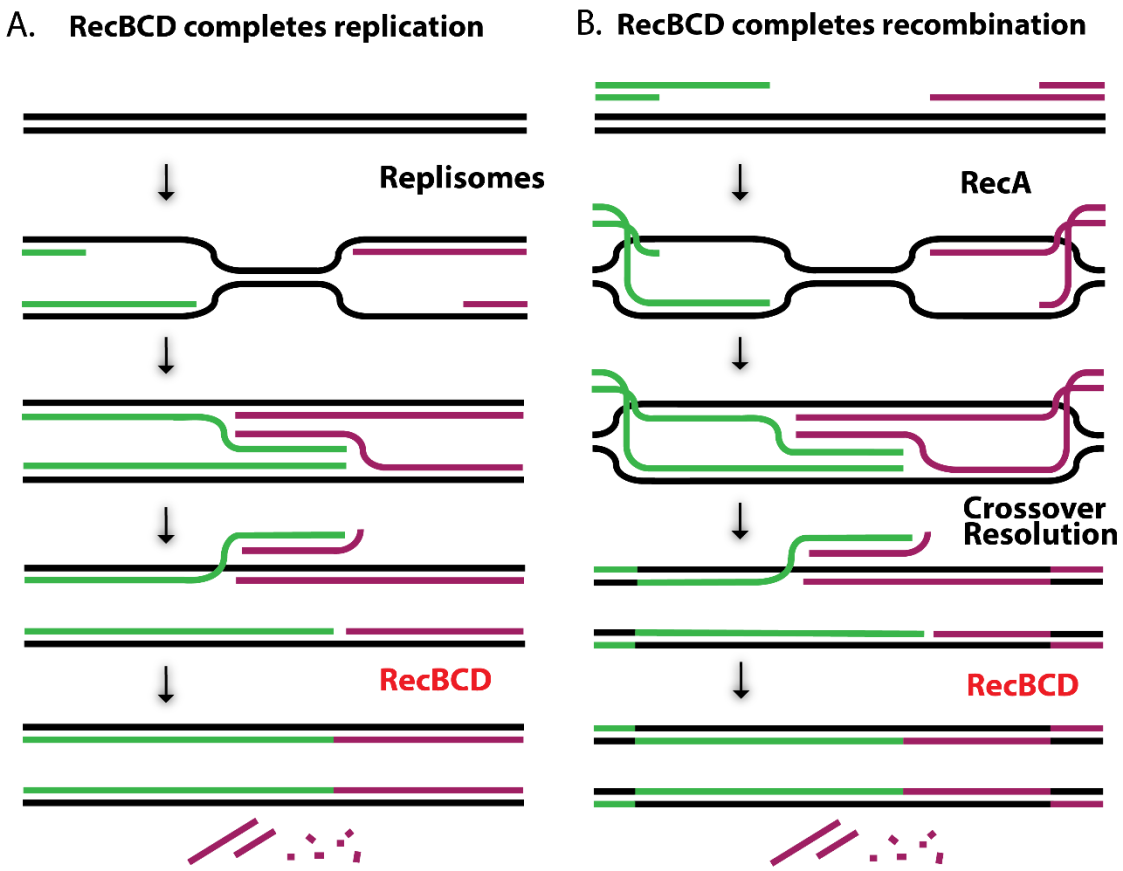

Figure 5.1. Model for the completion of DNA replication and the similarites to double-strand break repair via homologous recombination. A) As replication forks converge, one replication fork displaces the other and replicates on to the daughter strand via template-strand switching. The over-replicated region is processed by exo-nucleases generating recognizable substrates for RecBCD. B) Double-strand break repair via homologous recombination begins with degradation from ds-DNA ends generating 3' ss-DNA ends. RecA facilitates homology recognition and strand invasion. The newly invaded strands replicate and converge in the same way as duplicitous replication forks. Completion of the resulting replication follows a highly similar pathway. Figure adapted from (26)

References

1. Wendel BM, Courcelle CT, Courcelle J (2014) Completion of DNA replication in Escherichia coli. *Proc Natl Acad Sci*:201415025.

2. Wendel BM, Cole JM, Courcelle CT, Courcelle J (2017) SbcC-SbcD and ExoI process convergent forks to complete chromosome replication. *Proc Natl Acad Sci*:201715960.

3. Churchill JJ, Kowalczykowski SC (2000) Identification of the RecA protein-loading domain of RecBCD enzyme1. *J Mol Biol* 297(3):537–542.

4. Anderson DG, Kowalczykowski SC (1997) The Translocating RecBCD Enzyme Stimulates Recombination by Directing RecA Protein onto ssDNA in a χ-Regulated Manner. *Cell* 90(1):77–86.

5. Roman LJ, Dixon DA, Kowalczykowski SC (1991) RecBCD-dependent joint molecule formation promoted by the Escherichia coli RecA and SSB proteins. *Proc Natl Acad Sci* 88(8):3367–3371.

6. Jockovich M. E., Myers R. S. (2002) Nuclease activity is essential for RecBCD recombination in Escherichia coli. *Mol Microbiol* 41(4):949–962.

7. Birge EA, Low KB (1974) Detection of transcribable recombination products following conjugation in Rec+, RecB− and RecC− strains of Escherichia coli K12. *J Mol Biol* 83(4):447–457.

8. Hall JD, Howard-Flanders P (1972) Recombinant F′ factors from Escherichia coli K-12 strains carrying recB or recC. *J Bacteriol* 110(2):578–584.

9. Laureti L, Lee L, Philippin G, Pagès V (2017) A non-catalytic role of RecBCD in homology directed gap repair and translesion synthesis. *Nucleic Acids Res* 45(10):5877–5886.

10. Biek DP, Cohen SN (1986) Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in Escherichia coli. *J Bacteriol* 167(2):594–603.

11. Lovett ST, Luisi-DeLuca C, Kolodner RD (1988) The genetic dependence of recombination in recD mutants of Escherichia coli. *Genetics* 120(1):37–45.

12. Ivančić-Baće I, Salaj-Šmic E, Brčić-Kostić K (2005) Effects of recJ, recQ, and recFOR Mutations on Recombination in Nuclease-Deficient recB recD Double Mutants of Escherichia coli. *J Bacteriol* 187(4):1350–1356.

13. Chow SA, Chiu SK, Wong BC (1992) RecA protein-promoted homologous pairing and strand exchange between intact and partially single-stranded duplex DNA. *J Mol Biol* 223(1):79–93.

14. Hou R, Hill TM (2002) Loss of RecA function affects the ability of Escherichia coli to maintain recombinant plasmids containing a *Ter* site. *Plasmid* 47(1):36–50.

15. Galletto R, Amitani I, Baskin RJ, Kowalczykowski SC (2006) Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* 443(7113):875–878.

16. Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R (2004) Three replication origins in Sulfolobus species: Synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci U S A* 101(18):7046–7051.

17. Beach D, Piper M, Shall S (1980) Isolation of chromosomal origins of replication in yeast. *Nature* 284(5752):185–187.

18. Hause RJ, Shendure J (2014) Genetic Variation Meets Replication Origins. *Cell* 159(5):973–974.

19. Cromie GA (2009) Phylogenetic Ubiquity and Shuffling of the Bacterial RecBCD and AddAB Recombination Complexes. *J Bacteriol* 191(16):5076–5084.

20. Alonso JC, Stiege AC, Lüder G (1993) Genetic recombination in *Bacillus subtilis* 168: effect of recN, recF, recH and addAB mutations on DNA repair and recombination. *Mol Gen Genet MGG* 239(1–2):129–136.

21. Mascarenhas J, et al. (2006) *Bacillus subtilis* SbcC protein plays an important role in DNA inter-strand cross-link repair. *BMC Mol Biol* 7:20.

22. Lenhart JS, Schroeder JW, Walsh BW, Simmons LA (2012) DNA Repair and Genome Maintenance in *Bacillus subtilis*. *Microbiol Mol Biol Rev MMBR* 76(3):530–564.

23. Duggin IG, Bell SD (2006) The Chromosome Replication Machinery of the Archaeon Sulfolobus solfataricus. *J Biol Chem* 281(22):15029–15032.

24. Bell SP (2014) Terminating the replisome. *Science* 346(6208):418–419.

25. Duggin IG, Dubarry N, Bell SD (2011) Replication termination and chromosome dimer resolution in the archaeon Sulfolobus solfataricus. *EMBO J* 30(1):145–153.

26. Courcelle J, Wendel BM, Livingstone DD, Courcelle CT (2015) RecBCD is required to complete chromosomal replication: Implications for double-strand break frequencies and repair mechanisms. *DNA Repair* 32:86–95.