Elucidating the role of the *Escherichia coli* RarA protein at the interface of DNA replication and DNA repair

by

Tyler H. Stanage

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Degree of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: December 12, 2017

The dissertation is approved by the following members of the Final Oral Committee:

Michael M. Cox, Professor, Biochemistry
James L. Keck, Professor, Biomolecular Chemistry
M. Thomas Record Jr., Professor, Biochemistry
Jue D. Wang, Professor, Bacteriology
Aaron A. Hoskins, Assistant Professor, Biochemistry

To my teachers

Acknowledgements

For the past four and a half years, I have had the privilege of calling Madison my home. I cannot think of a better place to have spent these critical years of my scientific career.

I want to extend my deepest gratitude to my PhD advisor, Dr. Michael Cox, for making my graduate school experience fulfilling, productive, and unforgettable. His constant support and care for both his lab members and the undergraduates he teaches are admirable qualities I strive to replicate in my own career. Mike gave me incredible amounts of independence in my project, but was always happy to discuss results and problems alike for what seemed like unlimited amounts of time. Professionally, Mike has always been a strong advocate for preparing his students for life after graduate school. Thanks to Mike, I was able to travel to conferences across the United States and around the world. In addition, we traveled to Australia to jump start a collaboration that has been instrumental in progressing my project. Moving forward, I have so much to thank Mike for, including securement of my postdoctoral researcher position in Simon Boulton's laboratory at the Francis Crick Institute in London. I will be forever thankful to Mike for giving me the tools necessary to be the best scientist I can be.

I have been incredibly lucky to have such a supportive and encouraging thesis committee. Dr. James Keck has been a *de facto* second PhD advisor throughout my time at UW-Madison. His constant upbeat attitude, combined with useful insight and helpful suggestions for both my laboratory and professional lives have been invaluable. Dr. Jade Wang has pushed me to be a better scientist and presenter through her constant advice in our group meetings. I will never forget to include a model figure in any presentation or writing thanks to her! Dr. Tom Record has been a phenomenal source of knowledge and teaching. His course on biophysical chemistry gave me a greater appreciation for understanding the intricate details of biochemistry that many often

overlook. Dr. Aaron Hoskins is one of the most talented young scientists we have in our program—I am in constant awe in the work that his laboratory is producing, He was a great pleasure to teach for in our introductory biochemistry course, and I learned a lot about teaching style from him. Each of the professors on my committee has contributed an integral part to my graduate school career. I constantly tell others how important it is to have a committee that is as helpful and encouraging as mine.

My labmates have made going to work every day a great pleasure. I've most likely annoyed them throughout the years by asking so many questions, but they've always taken them in stride. Our lab has attracted a diverse set of people over the years, with wide-reaching talents and interests. I think this is partially due to Mike's ability to create a fostering and collaborative laboratory environment that is attractive to many people. The biggest thanks goes to Liz Wood, who is single-handedly responsible for ensuring our students maintain an efficient time to graduation. Her constant presence in the lab for the past 30+ years offers sustainability that is unparalleled. I've thoroughly enjoyed getting to know Liz over these past few years—from exchanging recipes to always being able to make me laugh—I will miss her dearly. Sindhu Chitteni-Pattu has always been a smiling face in lab—her generosity and supportiveness is infectious. I will always carry her positivity with me wherever I go. Carol Pfeffer has been a positive force of organization and support ever since I joined the lab. From helping me with travel and conference registrations to providing fantastic travel recommendations, I have always enjoyed our conversations and all of the advice she has provided me.

Early on in the lab, graduate students Erin Ronayne, Angie Gruber, and TaeJin Kim provided guidance through my project and helped me rapidly acclimate to graduate school life. Stefanie Chen, a postdoc in the lab, was also instrumental in helping me troubleshoot my project

and train me in numerous biochemical techniques. I am glad that all of these former lab members are continuing the great tradition of Cox lab alumni finding success in their scientific endeavors. Steven Bruckbauer and I joined the lab at the same time in early 2014. He has put up with my inexperience in genetics and microbiology with patience and humility. I would never have been able to do as many *in vivo* experiments without his constant guidance and troubleshooting ability. He is a great teacher and mentor, as evidenced by his successful undergraduates he has mentored in the lab. I wish him all of the best in the future. New graduate students Zach Romero and Miguel Osorio Garcia show tremendous promise in lab, despite their inability to win bets they made with me. Camille Henry has been the best support in lab while writing this thesis. Camille was always ready to drop whatever she was doing to help me with any questions I've had or to just give me chocolate. Her unique skillset will be a tremendous benefit to the Cox lab now and in the future. Finally, I'd like to thank all of our talented undergraduates, especially the ones I have mentored. Laura Sowin, Matthew Ritger, and Kamya Gopal have been phenomenal undergrads to mentor, and I know they'll all find success in the next steps of their careers.

I've been incredibly lucky to have phenomenal collaborators around the world. The two weeks we spent with the van Oijen laboratory at the University of Wollongong were some of the most scientifically stimulating moments of my graduate school career. Megan, Jacob, Lisanne, Slobodan, Andrew, Nick, and Antoine—thank you all for taking such an interest in my project and transforming it into a well-rounded story. Locally, I would be remiss not to thank the Keck and Wang laboratories for their constant outpouring of support and ideas throughout my graduate school career. In addition, I appreciate all of the work that Dan Stevens and Darrell McCaslin in the Biophysical Instrumentation Facility have done to help me with my project. Lastly, my work would be a collection of words and poorly constructed images/posters without the constant help

of Robin Davies and Laura Vanderploeg in our MediaLab. The MediaLab is a unique resource that many of us take for granted—I know I will sorely miss you both when I leave.

Despite oftentimes being described as "curmudgeonly," I have found a great group of friends that put up with me. After living alone for the first two years in Madison, I thought I would never want to have roommates again. I was very wrong. My roommates, Karl Wetterhorn and Markus Nevil, have always been there for me to vent about my project or reviewer #3. I am so incredibly grateful to have lived with these two unique individuals for the past two and a half years. I can truly say that living with them has made me a better person, even though I'll never be able to truly appreciate the "how many grains of sand make a pile"-like conversations they always seem to be having. Thank you Christine, for being a voice of reason and a source of confidence ever since I met you. To my fellow fifth year main crew—Karl, Markus, Alex, Haley, Megan, Stephanie, Ana—I've enjoyed all of the time we've spent together these past four years. So, so many memories were made that I will cherish forever. All of my other peers—thank you for being such a large, wonderful, diverse group of people. We constantly lean on each other to be successful in this graduate program, which is precisely what makes this program so highly regarded.

Throughout the ever-changing life of a graduate student, my family has remained my constant source of unadulterated support. My parents are the hardest working people I know. I try every day to give my work a fraction of the dedication that my parents give theirs, and I know that I have been successful in large part due to that effort. My brother Broc and sister-in-law Megan are two more great examples of hard-working parents. I thank them both for taking such a large interest in my work. My stepfamily is also full of hard-working individuals who have colored my life in ways I never would have imagined when we first met—I also thank them for being

completely supportive of me. Finally, I'd like to thank my extended family for providing a loving atmosphere for me to grow up and learn in.

Without any of the people I've mentioned here, this project would not have evolved or progressed. I am truly grateful for everyone providing me with advice, support, love, and mentorship. I stand on all of your shoulders in everything that I do.

Abstract

In each cell cycle, the replisome encounters replication blocks that include DNA lesions, tightly bound DNA-protein complexes, and breaks in the DNA. In *Escherichia coli*, many pathways have evolved to recognize, respond to, and repair these replication blocks so that replication can be completed in a timely and accurate manner. The highly conserved RarA (Replication associated recombination protein A) protein family has been implicated in the faithful resolution of these replication conflicts. Despite high conservation and extensive work conducted on this protein family, the overall function of RarA remains enigmatic.

Here, we carefully describe two novel DNA-dependent functions for the *E. coli* RarA protein. First, we demonstrated that RarA binds to and hydrolyzes ATP at double-stranded DNA ends in order to generate single-stranded DNA flaps. Second, using a single-molecule reconstituted DNA replication assay, we show that RarA generated single-stranded DNA gaps behind active replisomes in an ATP-dependent manner. This activity is achieved by RarA disengaging DNA polymerase III cores from their β_2 processivity clamps.

In vivo, we demonstrate that RarA—presumably using its gap creation activity—generates suitable substrates for post-replication repair pathways including RecFOR-mediated homologous recombination and translesion DNA synthesis. Loss of this function results in impaired growth and varied DNA replication phenotypes. Overexpression of *rarA* is toxic due to its ATPase activity and the resulting overuse of mutagenic TLS polymerases. Together, these results further elucidate the role of the RarA protein family at the interface of DNA replication and DNA repair.

Table of Contents

Table of Figures

* Work conducted in collaboration with Antoine van Oijen's lab at the University of Wollongong in Wollongong, Australia.

***** Work conducted in collaboration with Antoine van Oijen's lab at the University of Wollongong in Wollongong, Australia.

Table of Tables

Abbreviations

6-FAM	6-carboxyfluorescein
AAA+	ATPases associated with various cellular activities
ADP	adenosine-5'-diphosphate
ATP	adenosine-5'-triphosphate
ATPγS	adenosine-5'-(gamma-thiotriphosphate)
bp	base pair(s)
cssDNA	circular single-stranded DNA
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DSB	double-stranded break
dsDNA	double-stranded DNA
ldsDNA	linear double-stranded DNA
NFZ	nitrofurazone
nt	nucleotide(s)
NTP	nucleotide triphosphate
SSB	single-stranded DNA binding protein
ssDNA	single-stranded DNA
SSG	single-stranded gap
TLS	translesion DNA synthesis
UV	ultraviolet light

Chapter 1: Introduction

DNA Replication in Escherichia coli

In all organisms, accurate replication of the genome is essential for the faithful distribution of genetic material to daughter cells and the propagation of life. As such, DNA replication is one of the most highly conserved cellular processes from bacteria to humans. To ensure that this process is accurate and efficient, the cell has evolved an enzymatic complex, known as the replisome, to replicate a DNA template that is in constant flux. In the bacterium *Escherichia coli*, the replisome is comprised of four major protein complexes (**Figure 1-1**): (i) replicative helicase DnaB, (ii) single-stranded DNA binding protein (SSB), (iii) primase DnaG, and (iv) DNA polymerase III holoenzyme. Each DNA polymerase III holoenzyme is comprised of three polymerase cores (□□□) which carry out DNA synthesis (*1*), a toroidal clamp (□$_2$) that makes DNA synthesis highly processive (*2*), and a clamp loader complex (□□□□) which loads clamps onto the DNA in an ATP-dependent mechanism (*3*).

All DNA polymerases synthesize nascent DNA using a single-stranded DNA template. As such, the establishment of an active replication fork requires the double-stranded genome to be unwound into its constitutive single strands. At the beginning of each DNA replication cycle, this DNA melting occurs at a specific site known as the origin of replication (*oriC* in *E. coli*) (*4*, *5*). The highly-conserved initiator protein DnaA is able to recognize *oriC* and unwind the duplex in an ATP-dependent manner (*6*). The replicative helicase DnaB is then able to load onto this substrate and is primed for further unwinding of the DNA duplex as part of the replication fork (*7*). Once DnaB is loaded, a short RNA primer is synthesized by the primase DnaG (*8–10*). This functionalized RNA-DNA hybrid substrate, known as a primer-template junction, is now able to be recognized by the clamp loader complex (*11*, *12*). In an ATP-dependent manner, clamp loader opens a processivity clamp (□$_2$) and loads it onto the primer-template junction (*3*). The final step

in initiating DNA replication involves loading DNA polymerase III core onto the primed □$_2$-DNA complex, after which DNA replication can proceed (*13*).

The requirement of a free 3' hydroxyl group for DNA polymerase III to form a phosphodiester bond to an additional nucleotide necessitates that DNA synthesis occurs in a 5' to 3' orientation along the DNA template (*14*). This inherent polarity, combined with the polar nature of the template DNA, constrains the ability of DNA polymerase III to synthesize DNA continuously on both template strands. Thus, two modes of DNA synthesis occur: the first is termed leading strand DNA synthesis, while the second is known as lagging strand DNA synthesis (**Figure 1-2**). Leading strand DNA synthesis occurs continuously as DnaB unwinds additional template DNA. Lagging strand synthesis, however, occurs discontinuously in short (1-2 kilobases) fragments known as Okazaki fragments (*15*, *16*). These fragments are subsequently ligated together by the enzyme DNA ligase (*ligA/B*) (*17*). An additional side effect of discontinuous DNA synthesis on the lagging strand is the accumulation of large amounts of single-stranded DNA. This single-stranded DNA is bound by the single-stranded DNA binding protein (SSB). SSB not only protects labile single-stranded DNA from nucleolytic attack or modification, it also serves as an interacting partner for numerous proteins involved in DNA replication, recombination, and repair (*18*, *19*).

Despite the different modes of DNA synthesis on the leading and lagging strand, DNA replication is an incredibly efficient and processive process. In order for DNA polymerase III holoenzyme to be a processive enzyme, it must be bound to a □$_2$ processivity clamp (*20*). Without a beta clamp, DNA polymerase III can only synthesize 8-14 nucleotides before dissociating with the DNA template (*20*). However, when bound to a beta clamp, DNA polymerase III can synthesize thousands to tens of thousands of nucleotides prior to dissociation (*20*). While other

subunits of the DNA polymerase III holoenzyme contribute to processivity, beta clamp is the major contributor (*21*).

Even though DNA polymerase III is highly efficient and processive enzyme when replicating on an ideal template *in vitro*, the *E. coli* genome is an inherently imperfect template. The genome is a substrate for many major cellular processes including DNA replication, transcription, and DNA repair. As such, DNA replication is competing with many other essential processes in the cell for the same template. Despite its high processivity, DNA polymerase III is largely intolerant of imperfections in the DNA template, including modified DNA bases, DNA-protein complexes, and single-stranded DNA gaps (**Figure 1-3**) (*22*). Cells have evolved many pathways to allow for DNA replication to continue unimpeded, even in the presence of these imperfections. Largely, these pathways are focused on recognizing, responding to, and repairing these imperfections in the DNA template so that DNA replication can complete in a timely manner. Collectively, I will refer to these pathways as the replication-associated DNA repair pathways.

Replication-associated DNA repair pathways in Escherichia coli

When a replication fork encounters a barrier in the DNA template, it can stall or collapse (*23–30*). These barriers may include damaged/modified DNA bases, double-stranded DNA breaks, actively transcribing RNA polymerases, or protein-DNA complexes (*31–33*). If left unrepaired, collapsed replication forks can lead to genome instability, increased mutation rate, and even cell death. Estimates of replication fork collapse rates vary, but the growing consensus is that forks can collapse as often as once per cell cycle, even in the absence of exogenous genotoxic agents (*34*). Due to this frequency of replication fork collapse, all cells have evolved mechanisms to combat

the deleterious effects of these events. In *E. coli*, there are two major pathways that have evolved to facilitate the repair and/or tolerance of DNA damage encountered by the replication fork. The first pathway, known as translesion DNA synthesis (TLS), utilizes specialized DNA polymerases to bypass the site of DNA damage, rather than directly repairing it (*35*). The second pathway, known as recombinational DNA repair, utilizes a host of proteins that utilize an undamaged DNA template to directly repair the site of DNA damage in a nonmutagenic manner (*24*, *25*, *30*, *36*, *37*). I note that a third (albeit minor) pathway does exist for repairing DNA damage at a replication fork known as non-homologous end joining (NHEJ). A possible mechanism for NHEJ in *E. coli* has recently been proposed (*38*). In this work, I will mainly focus on the two major pathways listed above as they relate directly to the work described in later chapters.

Recombinational DNA Repair

An error-free alternative pathway to direct translesion DNA synthesis utilizes a diverse collection of proteins to catalyze homologous recombination at sites of DNA damage (*27*). When a replication fork stalls at a site of DNA damage, it can collapse to form a double-stranded DNA end. Alternatively, the replicative helicase can continue unwinding DNA ahead of a stalled replisome, generating a large stretch of single-stranded DNA that is coated by the single-stranded DNA binding protein SSB. In either situation, the replisome can disengage from the DNA template, necessitating *oriC*-independent replication initiation, known as replication restart (*39*). Prior to restarting replication, the damaged DNA substrates must be repaired and the replisome must be reloaded and primed. The repair of these damaged substrates is the primary function of homologous recombination in prokaryotes (*36*).

The primary driver of homologous recombination in most organisms is the RecA protein family (*40*). RecA is a highly conserved bacterial recombinase that catalyzes homologous

recombination via DNA strand exchange (**Figure 1-4**) (*41*). To accomplish this goal, RecA forms a nucleoprotein filament on single-stranded DNA termed the RecA filament or RecA* (*42*, *43*). RecA* serves both as a template for homologous recombination at sites of DNA damage and as a signal for induction of the SOS response by cleaving the transcriptional repressor protein *lexA* (*44*, *45*). There are three main steps in repairing broken replication forks via homologous recombination: 1) generation of single-stranded DNA and formation of the RecA filament; 2) catalysis of strand exchange; and 3) resolution of strand exchange products into a functionalized replication fork.

Since RecA generally forms a filament on single-stranded DNA, a double-stranded DNA end must be processed prior to loading RecA. Processing of a double-stranded end is catalyzed by the RecBCD helicase-nuclease complex (*46*). RecBCD binds tightly to DNA ends and proceeds to unwind the substrate in an incredibly fast (2,000 bp/s) and processive (30,000 bp) manner (*47*). As the duplex DNA is unwound, RecBCD degrades the 5' strand, creating a single-stranded DNA substrate containing a free 3' end which is subsequently bound by SSB (*48*). This endonucleolytic action continues until the RecBCD holoenzyme reaches a specific DNA sequence called the *chi* site found throughout the *E. coli* genome (*48–50*). RecBCD is then able to load RecA directly onto the single-stranded DNA substrate to form a RecA filament (*46*). In the case of DNA damage leading to a single-stranded DNA gap, the RecFOR pathway facilitates RecA loading and filament formation on this substrate (*51*). In fact, RecA is only able to effectively filament on SSB-coated DNA in the presence of the RecO or RecBCD protein (*51–53*). Loss of either the *recB* or *recO* pathways results in the inability to repair double-stranded DNA ends or single-stranded DNA gaps, respectively (*54*, *55*). These observations, combined with the *in vitro* results detailed above, result

in a model where RecBCD or RecFOR act upstream of RecA and facilitate RecA filament formation on single-stranded DNA at sites of DNA damage (*23*).

Once RecA has loaded onto DNA as a nucleoprotein filament, it is primed to catalyze the major reaction of homologous recombination termed DNA strand exchange. This reaction requires a RecA filament, a double-stranded DNA substrate that is homologous to the filamented single-stranded DNA, ATP, and SSB (*56*, *57*). Under these conditions, the RecA filament is able to perform a homology search on a sister chromosome by invading the duplex DNA (*43*). At a minimum, eight homologous base pairs are required for RecA to catalyze strand exchange, while heterologous DNA is strongly disfavored for this reaction to complete (*58*). Many accessory protein factors modulate the ability of RecA to perform strand invasion and eventual homologous recombination. Most of these factors stabilize or destabilize the RecA filament directly, often by binding RecA monomer-monomer interfaces (*59*). These factors are in constant flux, modulating the RecA filament to respond to the type and extent of DNA damage being repaired. Once strand exchange has occurred, recombination is completed by resolving the intermediate DNA substrates formed during this process. This process, which includes the resolution of a four-stranded recombination intermediate known as the Holliday Junction (*60*, *61*), is carried out by the RuvABC proteins (*62–66*). Upon resolution of these recombination intermediates, a replication fork has been reestablished in a nonmutagenic manner. This replication fork can then serve as a substrate for reloading a functional replisome using the replication restart pathway. Replication restart is primarily carried out by the primosome which is comprised of PriABC, DnaT, primase DnaG, and replicative helicase DnaB (*39*). The end result of replication restart is the reestablishment of a functional replisome, anchored by the reloading of DnaB, which then signals for the recruitment of the entire Pol III holoenzyme (*39*).

Translesion DNA Synthesis (TLS)

It has long been observed that when *E. coli* cells are stressed using ultraviolet (UV) irradiation, mutation rates increased in daughter cells' genomes (*67*). Several studies discovered the genes required for these increased mutation rates, which included the bacterial recombinase *recA*, the transcription factor *lexA*, and genes of unknown function (at the time) *umuC* and *umuD* (*68*, *69*). Indeed, Miroslav Radman had already hypothesized the existence of a DNA damage response pathway—the SOS response—that involved these same genes (*70*). However, while several models were proposed for these genes' involvement in damage-induced mutagenesis, it wasn't until the discovery and characterization of the three specialized DNA polymerases—Pol II, Pol IV, and Pol V—that the translesion DNA synthesis (TLS) model was fully formulated (*71*). In brief, the model for TLS involves the recruitment of specialized DNA polymerases to sites of DNA damage, whereby these polymerases incorporate nucleotides across from a damaged DNA template (*35*). The currently understood properties of each of the known TLS polymerases, their recruitment to sites of DNA damage, and phenotypes associated with their deletion will be explored below.

The first TLS polymerase, DNA polymerase II, encoded by the *polB* gene, was first described by Rolf Knippers, Thomas Kornberg and Malcolm Gefter in the early 1970's (*72*, *73*). When cells are stressed with DNA damaging agents (thereby inducing the SOS response), the expression of Pol II increases approximately sevenfold, from 50 to approximately 350 molecules (*74*). Pol II is the only TLS polymerase known to possess an exonuclease proofreading (3' to 5') function (*71*). Pol II is capable of bypassing two known types of DNA damage: apurinic (AP) sites (*75*) and some bulky guanine base adducts (*76*). However, bypass of these lesions by Pol II has been shown to increase the frequency of -2 frameshift mutations (*76*, *77*). Recruitment of Pol II to

sites of DNA damage and stimulation of its polymerase activity is mediated by an interaction with the beta processivity clamp (*78–80*). Despite its role in the SOS response, cells lacking *polB* do not show exhibit a significant phenotype when stressed by various DNA damaging agents. The only exception is when Δ*polB* cells are exposed to the mutagen 4-nitroquinoline-1-oxide (4NQO), which is known to cause bulky guanine base adducts (*81*), and hydrogen peroxide, which increases cellular concentrations of reactive oxygen species (ROS) (*82*). DNA polymerase II remains the most poorly understood of the three TLS polymerases.

DNA polymerase IV, encoded by the *dinB* gene, was discovered by Kenyon and Walker in 1980 after observing its cellular levels increased tenfold following induction of the SOS response (*83*). However, unlike Pol II or Pol V, Pol IV is present at rather high levels (approximately 250 copies/cell) in the absence of DNA damage (*84*). Pol IV is the most conserved of the *E. coli* TLS polymerases; as such, it has been studied extensively and compared to its eukaryotic counterparts (*35*). Evidence from *in vitro* and *in vivo* studies suggests Pol IV is able to bypass both alkyl and bulky guanine base adducts (*84*). However, with no proofreading capability, Pol IV activity results in an increase in -1 frameshift mutations (*84*). Indeed, overexpression of *dinB* has been shown to dramatically increase the frequency of -1G frameshift mutations, the accumulation of which eventually results in cell death (*85*). Like Pol II, Pol IV is recruited to sites of DNA damage through an interaction with the beta processivity clamp (*86*). This well-studied interaction between Pol IV and the beta clamp has led to a "tool belt" model for polymerase switching, whereby replicative Pol III and TLS polymerases bind to a single beta clamp simultaneously and can switch to an active position when damaged DNA is encountered by the replisome (*87–89*). Physiologically, Pol IV function is not required for normal growth. However, cells lacking *dinB* are sensitive to various DNA damaging agents, including alkylating agents such

as methyl methanesulfonate (MMS), or drugs that introduce bulky guanine base adducts like nitrofurazone (NFZ) or 4-nitroquinoline-1-oxide (4NQO) (*81*, *90*, *91*). The sensitization of Δ*dinB* cells to these genotoxic agents has been attributed to simply lacking the only TLS polymerase that is able to bypass these lesions *in vitro* and *in vivo.* Pol IV continues to be an active area of investigation, especially in the context of its abnormally high expression rate in the absence of the SOS response.

The final known TLS polymerase in *E. coli* is DNA polymerase V, encoded by the *umuC* and *umuD* genes. The catalytic subunit of DNA polymerase V is encoded by the *umuC* gene, while the *umuD* gene product serves in multiple roles as a regulatory subunit of Pol V activity (*92*, *93*). Unlike TLS Pols II and IV, expression of Pol V is completely regulated by the SOS response, with no detectable Pol V molecules present until very late in the SOS response. Indeed, Pol V is the most stringently regulated TLS polymerase, with expression only occurring 30-50 minutes after induction of the SOS response (*94*). This transcriptional and temporal regulation, along with a post-translational RecA-mediated proteolytic regulation of the regulatory *umuD* subunit, showcase the evolved constraints on usage of the Pol V enzyme (*95*, *96*). Correspondingly, Pol V is both the slowest and most error-prone of all *E. coli* DNA polymerases (*97*); thus, its activity is only used in the most severe of DNA damage events. Nonetheless, *in vitro* studies have documents Pol V's ability to bypass a diverse collection of damaged DNA substrates, making it an ideal last-ditch effort to thwart the propagation of DNA damage to daughter cells. Due to these observations, many groups have termed Pol V as the primary or major TLS polymerase in *E. coli*. Like the other TLS polymerases, Pol V is targeted to sites of DNA damage and its polymerase activity is stimulated by an interaction with the beta processivity clamp (*98*). Unlike other TLS Pols, Pol V is strongly activated by the presence of RecA nucleoprotein filaments, suggesting an additional mechanism

of action at long-lived sites of DNA damage (*99*). Interestingly, cells lacking Pol V are modestly sensitized to moderate levels of ultraviolet light, unlike cells lacking the other two TLS polymerases (*100*). This finding underlines the hypothesis that Pol V activity is required when cells are subjected to high amounts of genotoxic stress. DNA polymerase V is perhaps the most well-characterized TLS polymerase due to its early association with the SOS response, its complex regulatory mechanisms, and its relationship with the well-studied bacterial recombinase RecA. Its *in vitro* and *in* vivo mechanisms continue to be an active area of research today (*95*, *101*).

Lesion Skipping as a Mechanism for DNA Repair Pathway Choice

Despite decades of work on characterizing the two major replication-associated DNA repair pathways in *Escherichia coli*, much less is known about how the cellular machinery chooses which pathway to utilize when faced with diverse types of replisome blockages. Many have concluded that a repair pathway is chosen based on the type of substrate needing repair. This is despite the fact that both translesion DNA synthesis and homologous recombination are able to repair single-stranded DNA gaps left behind when the replicative helicase continues ahead of a stalled replisome. These gaps were described as early as 50 years ago, when Rupp and Howard-Flanders observed single-stranded DNA gaps formed following UV irradiation of cells lacking the nucleotide excision repair pathway (direct excision and repair of damaged nucleotide bases by the *uvrABC* proteins) (*102*). Recent work by Robert Fuchs and colleagues has illuminated the competition between TLS and homologous recombination to repair these gaps; however, the exact mechanism of the gap formation remains to be elucidated (*103*).

Ken Marians and colleagues have demonstrated the ability of DNA polymerase III, in the presence of the DnaG primase, to bypass leading-strand DNA lesions *in vitro* (*104*, *105*). However, this process occurred on the timescale of minutes, and was specific to only one type of DNA lesion—a UV-induced pyrimidine dimer. In addition to their work on leading-strand DNA lesions, Marians has also examined the trigger of lagging strand DNA synthesis to move to the next Okazaki fragment *in vitro*. In this work, Marians proposes two possible models for DNA polymerase recycling during lagging strand synthesis: one involving a signal derived from DnaG-dependent primer formation and another involving a direct collision of Pol III with an adjacent Okazaki fragment (*106*). McHenry and colleagues have ruled out the collision model in favor of the signaling model (*107*). While the source of the signal to recycle Pol III is still unknown, McHenry hypothesized a possible role for the clamp loader complex DnaX (*107*). This recycling signal could theoretically be harnessed as a mechanism to bypass sites of DNA damage on the lagging strand during DNA synthesis. In addition, exploitation of this signal for bypassing sites of DNA damage could explain the presence of gaps at the replication fork following UV irradiation as observed by Rupp and Howard-Flanders. In this work, we hypothesize the function of a highly-conserved AAA+ ATPase, RarA, is to catalyze this lesion skipping *in vivo*.

The Escherichia coli RarA protein

The *E. coli* RarA (Replication associated recombination A) protein was first described by David Sherratt and colleagues in 2001, after demonstrating its high conservation across all domains of life and its relationship to known DNA metabolism proteins DnaX and RuvB (*108*). Indeed, RarA shares roughly 40% amino acid identity with its yeast (Mgs1) and human (WRNIP1) homologues and 25% amino acid identity with the *E. coli* Holliday Junction helicase RuvB and clamp loader DnaX (*108*). Disruption of the *rarA* gene (using a chloramphenicol cassette insertion

at the RarA N-terminus) alone did not cause any significant phenotype with respect to growth, DNA repair capacity, or mutagenesis rate (*108*). However, when the *rarA* disruption was combined with DNA recombination mutants, growth rate and recombination frequencies were decreased, suggesting a role for RarA in DNA replication and recombination (*108*). Fluorescently-tagged RarA co-localizes with active replication forks, again suggesting a role for its function at the interface of DNA replication and recombination (*19*, *109*, *110*). Work by Shinagawa and colleagues demonstrated a synergistic effect on growth rate upon loss of both *rarA* and the bacterial recombinase *recA* (*111*). In addition, overexpression of *rarA* resulted in a 500-fold increase in mutagenesis, specifically the frequency of -1G frameshift mutations; overexpression of *rarA* was found to be lethal in the absence of *recA* (*111*). As noted previously, an increase in -1G frameshift mutations was also seen when *dinB* (Pol IV) was overexpressed (*85*); thus, it is hypothesized that the overexpression of *rarA* leads to increased Pol IV activity. The relationship between *rarA* function and DNA replication has been further underlined by observations that loss of *rarA* function alleviates the temperature sensitivity of strains carrying alleles that encode unstable DNA polymerase III holoenzyme (*111*, *112*). These findings suggested a destabilization role for RarA function on the replisome.

Biochemical characterization of the RarA protein has been a particular interest of our lab in the past ten years. A former graduate student in our laboratory, Asher Page, in collaboration with Nick George and James Keck's lab, solved the X-ray crystal structure of *apo* RarA (**Figure 1-5**) (*113*). The structure placed RarA in the AAA+ ATPase superfamily, specifically in the clamp-loader clade (*113*, *114*). Indeed, RarA shares extensive structural homology with the *E. coli* and *S. cerevisiae* clamp loaders DnaX and replication factor C, respectively. RarA was shown to be a DNA-dependent ATPase, with its activity stimulated by M13 bacteriophage genomic DNA (*113*).

In addition, RarA interacts with the single-stranded DNA binding protein (SSB) with the second highest affinity of all of its interacting partners (*113*). This interaction was found to be necessary for localization of RarA to replication forks *in vivo* (*19*). Despite the numerous studies published on the RarA protein, its overall function remains highly enigmatic. However, some additional insights have been achieved through study of its yeast and human homologues.

Eukaryotic RarA homologues

As previously mentioned, the RarA protein family is one of the most highly conserved DNA metabolism protein families. The yeast homologue of RarA, called Mgs1 (Maintenance of genome stability protein 1), was first described by Shinagawa and colleagues at the same time as Sherratt first described *E. coli* RarA (*115*). Like RarA, overproduction of ATPase-competent Mgs1 resulted in severe genome instability and in hypersensitization to DNA damaging agents (*115*). Like loss of *rarA,* loss of *mgs1* function did not confer a growth phenotype, but did modestly increase basal levels of homologous recombination. Combination of an *mgs1* deletion with mutants in the DNA damage response (specifically the *rad6* pathway) was found to be synthetically lethal, suggesting a role for *mgs1* in an alternative DNA damage pathway (*116*). Like in *E. coli*, a connection between *mgs1* and DNA replication was observed when deletion of *mgs1* rescued the inviability of DNA polymerase δ mutants (*117*). It should be noted that in yeast and most eukaryotes, two primary DNA polymerases exist: Pol ε, which synthesizes the leading strand, and Pol δ, which synthesizes the lagging strand (*118*). Mgs1 localizes to the replication fork in the absence and presence of DNA damage (*119*). Localization is dependent on a conserved interaction between Mgs1 and PCNA, the DNA replication processivity clamp in eukaryotes (*119*, *120*). Similarly, the human homologue of Mgs1, called WRNIP1 (Werner's helicase interacting protein

1), also localizes to the replication fork through an interaction with PCNA (*121*, *122*). Interestingly, localization of Mgs1/WRNIP1 to the replication fork is independent of its ATPase activity, thereby suggesting a role for this activity after it is localized to the replication fork (*121*). A possible role for WRNIP1 function at replication forks includes the observation that it functions upstream of the TLS polymerase Pol η in human cells treated with ultraviolet light (*123*). Despite these diverse observations for WRNIP1 function *in vivo*, its precise cellular function remains unknown. Many groups have chosen instead to investigate the biochemical properties of Mgs1/WRNIP1 *in vitro* to understand its DNA binding and ATPase activities.

Mgs1/WRNIP1 are DNA-dependent ATPases, with both single-stranded and double-stranded DNA being reported as substrates for this reaction (*115*, *124*, *125*). Unsurprisingly, these proteins also bind a variety of DNA substrates, including double-stranded DNA, single-stranded DNA, and replication fork mimics (*124*, *126*). Besides an ATPase-independent DNA annealing activity, no known DNA-dependent activity has been attributed to RarA family proteins. Mgs1/WRNIP1 interacts with PCNA, the eukaryotic DNA replication processivity clamp. Upon polyubiquination, PCNA acts as a scaffold for many proteins, including Mgs1/WRNIP1, for localization to the replication fork (*119*, *120*, *127*, *128*). In addition, as its name suggests, WRNIP1 has been shown to interact with Werner's helicase (WRN), a member of the conserved RecQ family of helicases that are involved in many aspects of DNA/RNA metabolism (*129–132*). From a replication perspective, Mgs1/WRNIP1 binds to and stimulates the activity of DNA polymerase δ, the polymerase responsible for replicating the lagging strand in eukaryotes (*125*). Further implicating Mgs1/WRNIP1 in the maintenance of lagging strand DNA synthesis, Mgs1 stimulated the activity of flap structure-specific endonuclease (Fen1) *in vitro*, even though a protein-protein interaction between these two proteins was not observed (*124*). Fen1 is responsible for removing

single-stranded DNA flaps created by DNA Pol δ upon overreplication at sites of newly synthesized Okazaki fragments (*133*).

Unlike *E. coli* RarA, eukaryotic Mgs1/WRNIP1 have been shown to possess numerous interacting protein partners *in vitro*. However, like RarA, there have been no descriptions of the function of Mgs1/WRNIP1 ATPase or DNA binding function *in vitro*. Similarly, despite the wealth of observations published on the RarA family in the past two decades, its underlying function in all cells remains highly enigmatic. The most recent model for RarA protein family function involves its ability to bind polyubiqutinated PCNA, which could serve as a mechanism to facilitate a DNA polymerase switch between normal (Pol δ) and translesion (Pol η) DNA synthesis when DNA damage is encountered by the replication fork (*134*). However, this hypothesis remains to be tested directly.

In this work, I use the *E. coli* RarA protein as a model for the RarA protein family. In chapter 2, I will describe our work on RarA *in vitro*. This includes two novel functions for the RarA protein family: DNA flap formation and single-stranded DNA gap formation. In chapter 3, I will describe the consequences of these novel activities *in vivo*. Finally, in chapter 4, I will put these observations in context with the current literature and provide some future directions.

Figures

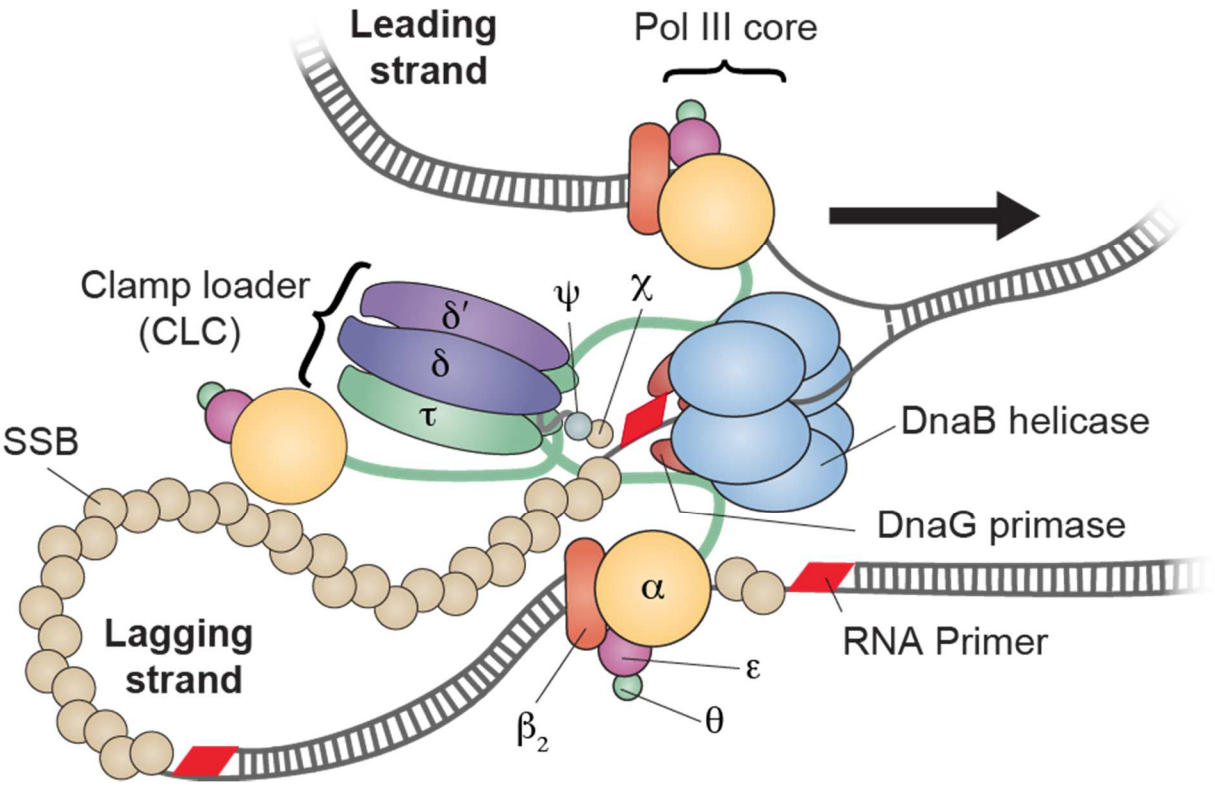

Leading strand
Pol III core
Clamp loader (CLC)
δ′
δ
τ
ψ
χ
SSB
DnaB helicase
DnaG primase
α
Lagging strand
RNA Primer
ε
θ
β2

Figure 1-1: Organization of the *E. coli* replisome

The replicative helicase DnaB translocates along the lagging strand, separating parental duplex DNA into its constitutive template strands. This action allows for DnaG primase to synthesize an RNA primer. On the leading strand, DNA synthesis occurs continuously from this RNA primer; however, on the lagging strand, DNA synthesis occurs in 1-2 kilobase segments called Okazaki fragments. As single-stranded DNA is exposed on the lagging strand, it is bound and protected by the single-stranded DNA binding protein SSB. Using this unwound single-stranded DNA template, Pol III holoenzyme synthesizes double-stranded DNA product in a coupled and processive manner. This processivity is largely conferred by the beta clamp (β_2) through its association with interaction with the Pol III core subunits. Beta clamp is initially loaded onto primer-template junctions by the clamp-loader complex (CLC). Pol III core subunits are tethered together (up to three at a time) through interactions with the CLC τ subunits, which also interact with DnaB helicases to scaffold the entire replisome together as replication progresses. Figure was adapted from (*135*).

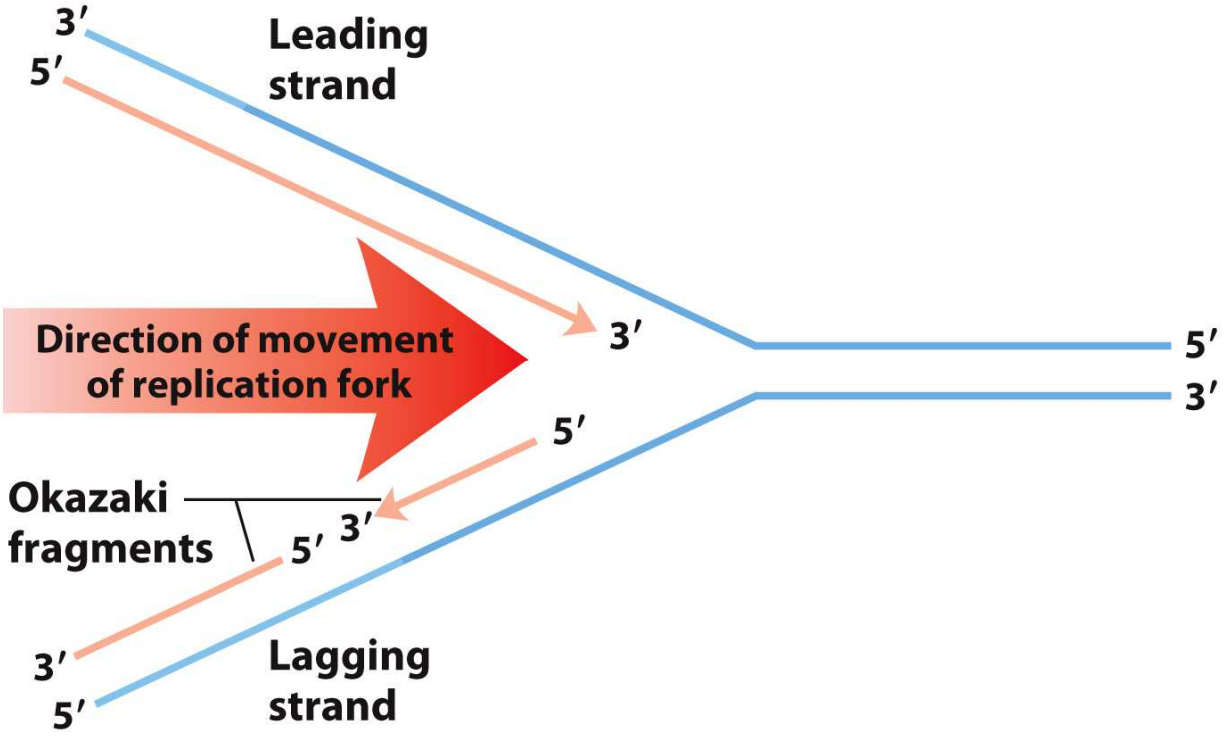

3′
5′
Leading strand
Direction of movement of replication fork
3′
5′
3′
5′
Okazaki fragments
5′
3′
3′
5′
Lagging strand

Figure 1-2: DNA replication occurs continuously on the leading strand and discontinuously on the lagging strand

A model showing an idealized DNA replication fork. As parental duplex DNA is unwound, DNA synthesis occurs continuously on the leading strand and discontinuously on the lagging strand. Leading strand synthesis occurs in the direction of overall replication fork movement, while lagging strand synthesis occurs in the opposite direction of overall replication fork movement. As a consequence, lagging strand synthesis utilizes short 1-2 kilobase-sized fragments called Okazaki fragments. Figure is adapted from Lehninger's Principles of Biochemistry (*136*).

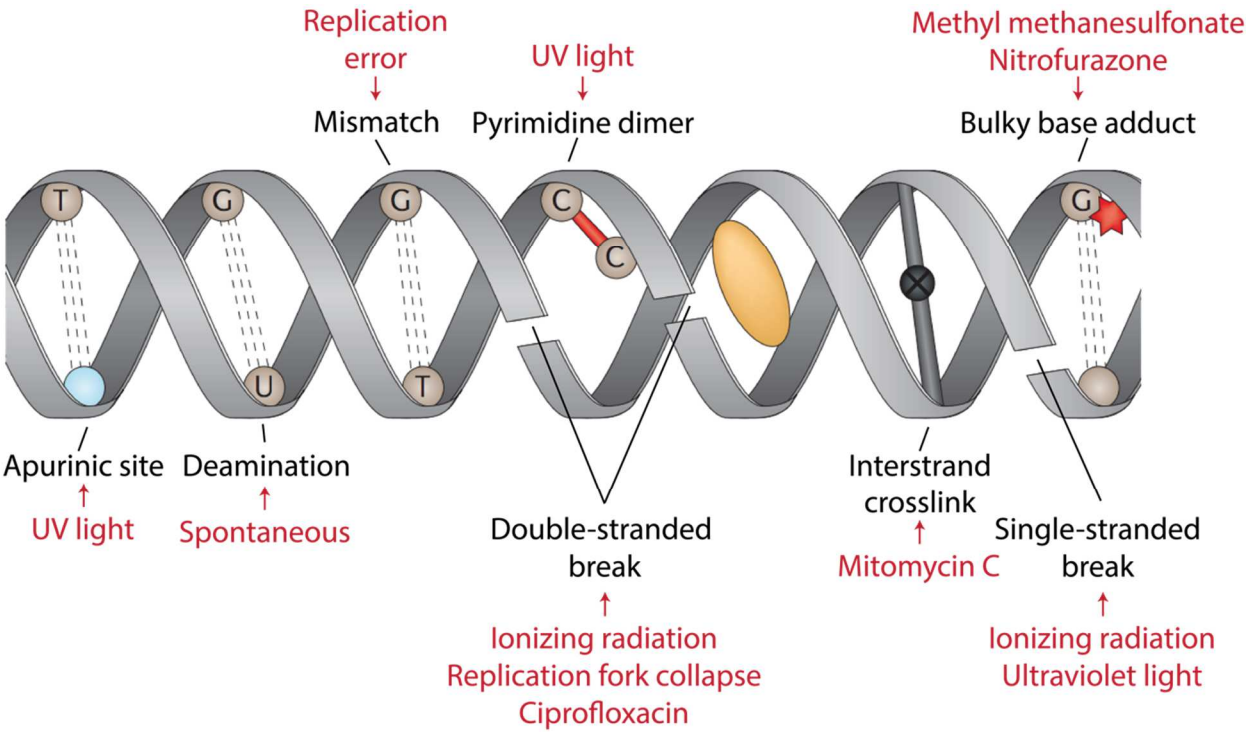

Replication error
Mismatch
UV light
Pyrimidine dimer
Methyl methanesulfonate
Nitrofurazone
Bulky base adduct
T
G
G
C
C
G
U
T
Apurinic site
UV light
Deamination
Spontaneous
Double-stranded break
Ionizing radiation
Replication fork collapse
Ciprofloxacin
Interstrand crosslink
Mitomycin C
Single-stranded break
Ionizing radiation
Ultraviolet light

Figure 1-3: Types of DNA damage encountered by replication forks

A schematic showing some of the most common types of DNA damage encountered by the replication fork in *E. coli*. Genotoxic agents and/or radiation sources that are known to cause the indicated DNA damage are shown in red. Damage ranges in severity from mismatched DNA bases to highly toxic double-stranded DNA breaks. *E. coli* has evolved several pathways to repair or tolerate this diverse set of DNA damage. Figure is adapted from Helleday *et al.* (*137*).

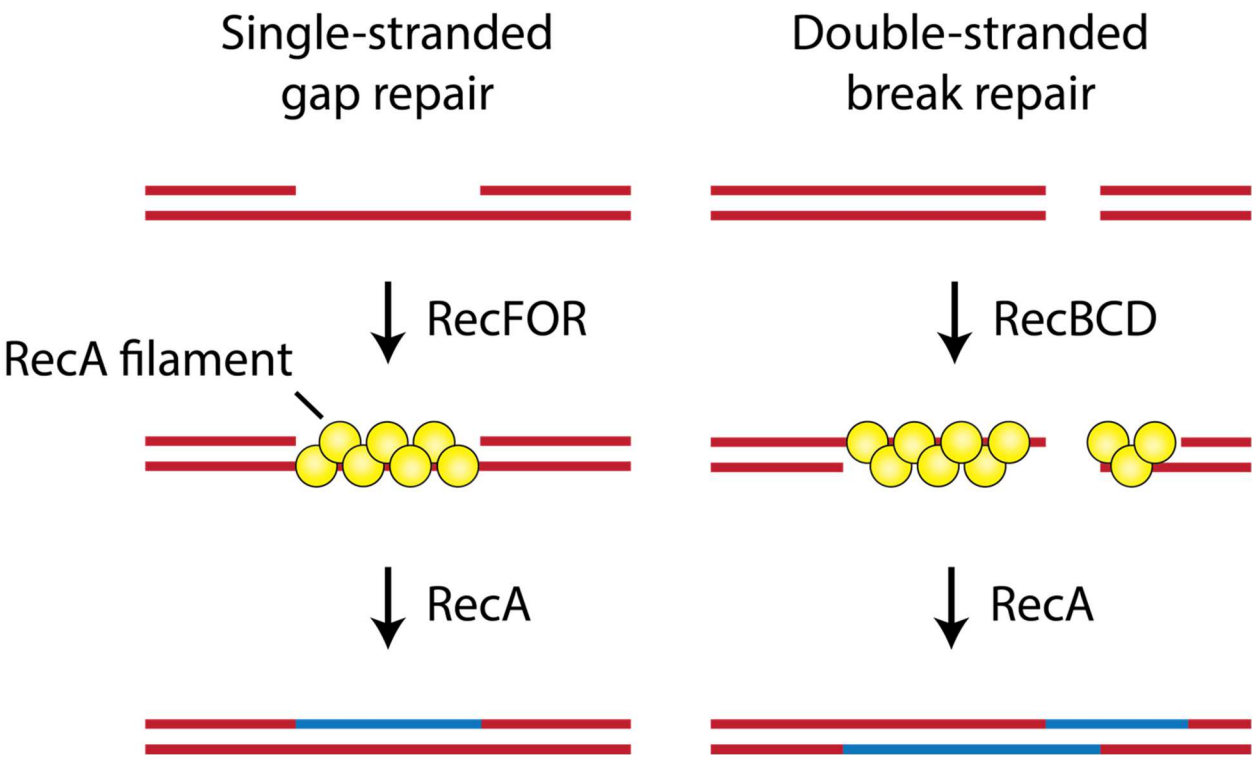

Single-stranded
gap repair
Double-stranded
break repair
RecFOR
RecBCD
RecA filament
RecA
RecA

Figure 1-4: Recombinational DNA repair is catalyzed by the RecA protein

Recombinational DNA repair is the major pathway in *E. coli* for repairing single-stranded DNA gaps and double-stranded DNA breaks. The bacterial recombinase RecA is loaded onto SSB-coated single-stranded DNA by the RecFOR proteins (left). Double-stranded DNA ends must first be processed by the RecBCD helicase-nuclease complex to generate single-stranded DNA (right). In both cases, RecA catalyzes strand exchange between the damaged chromosome and an undamaged sister chromosome. This process generates recombination intermediates called Holliday Junctions (not pictured) that are then resolved to complete the repair process.

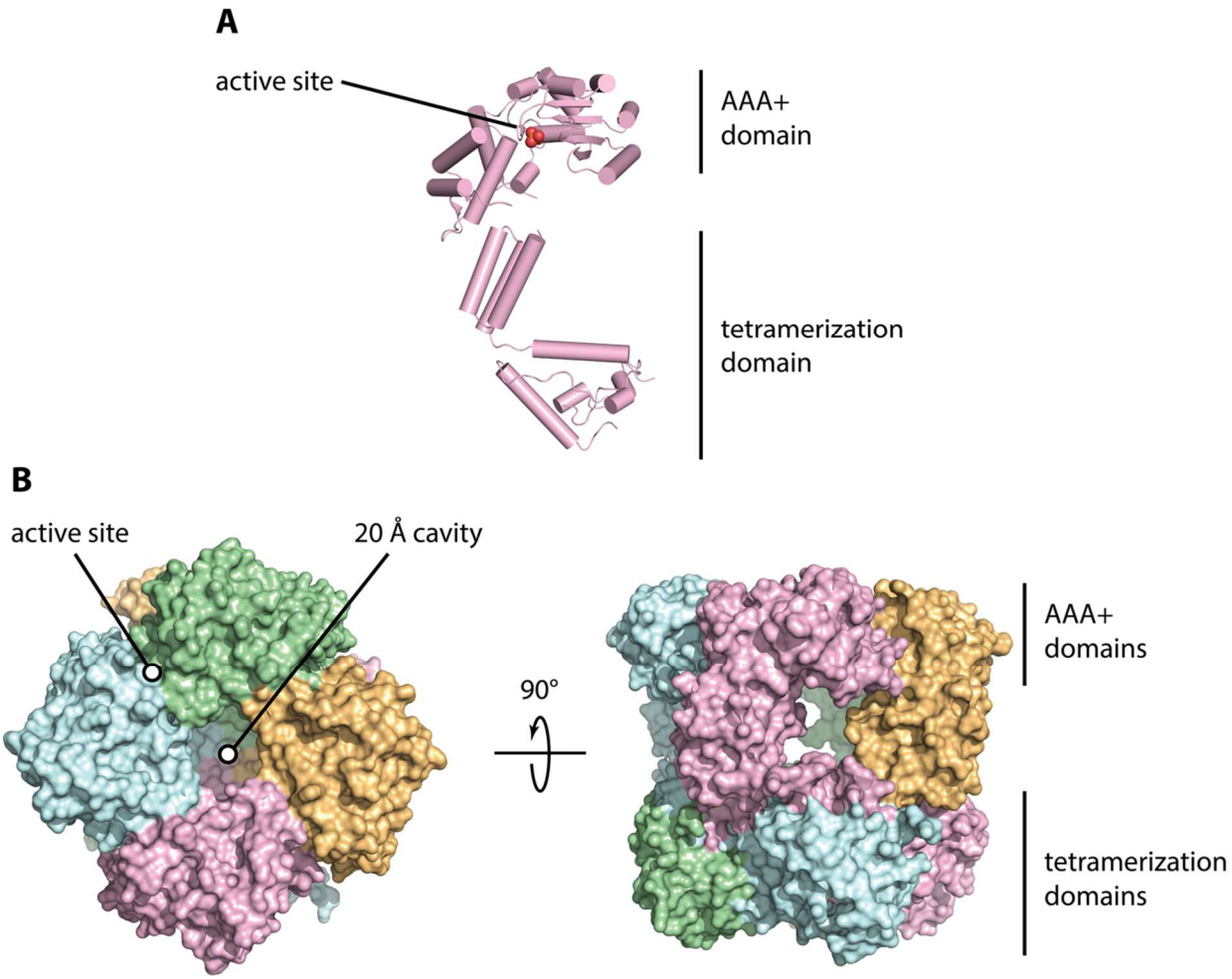

A
active site
AAA+
domain
tetramerization
domain
B
active site
20 Å cavity
90°
AAA+
domains
tetramerization
domains

Figure 1-5: X-ray crystal structure of the *E. coli* RarA protein

(**A**) X-ray crystal structure of the *E. coli* RarA monomer. RarA is comprised of two major domains: a AAA+ ATPase domain, and a tetramerization domain. A phosphate ion is bound in the active site of the AAA+ domain (orange and red spheres). The extensive alpha-helical structure of the tetramerization domain facilitates oligomerization of the enzyme. (**B**) Homotetrameric structure of *E. coli* RarA protein. A top-down view (left) reveals a 20 Å cavity formed by the AAA+ domains upon tetramerization. A hallmark of the AAA+ superfamily, the active site of ATP hydrolysis is found at the interface between two oligomers. A side view of the crystal structure reveals the extensive contacts between oligomers facilitated by the tetramerization domains. This crystal structure has been published previously (*113*). PDB: 3PVS.

References

1. R. Reyes-Lamothe *et al.*, Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science*. **328**, 498–501 (2010).

2. R. J. LaDuca, J. J. Crute, C. S. McHenry, R. A. Bambara, The beta subunit of the *Escherichia coli* DNA polymerase III holoenzyme interacts functionally with the catalytic core in the absence of other subunits. *J. Biol. Chem.* **261**, 7550–7557 (1986).

3. B. A. Kelch, D. L. Makino, M. O'Donnell, J. Kuriyan, How a DNA Polymerase Clamp Loader Opens a Sliding Clamp. *Science.* **334**, 1675–1681 (2011).

4. A. C. Leonard, J. E. Grimwade, The orisome: Structure and function. *Front. Microbiol.* **6**, 1–13 (2015).

5. M. L. Mott, J. M. Berger, DNA replication initiation: mechanisms and regulation in bacteria. *Nat. Rev. Microbiol.* **5**, 343–354 (2007).

6. K. Sekimizu, D. Bramhill, A. Kornberg, ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell*. **50**, 259–265 (1987).

7. J. H. LeBowitz, R. McMacken, The *Escherichia coli* dnaB replication protein is a DNA helicase. *J. Biol. Chem.* **261**, 4738–48 (1986).

8. J.-P. Bouché, L. Rowen, A. Kornberg, The RNA Primer Synthesized by Primase to Initiate Phage G4 DNA Replication*. **253** (1978).

9. J. P. Bouché, K. Zechel, A. Kornberg, dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form. *J. Biol. Chem.* **250**, 5995–6001 (1975).

10. L. Rowen, A. Kornberg, Primase, the dnaG Protein of *Escherichia coli*. *J. Biol. Chem.* **253**, 758–764 (1978).

11. G. D. Bowman, M. O'Donnell, J. Kuriyan, Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature*. **429**, 724–730 (2004).

12. K. R. Simonetta *et al.*, The Mechanism of ATP-Dependent Primer-Template Recognition by a Clamp Loader Complex. *Cell*. **137**, 659–671 (2009).

13. K. O. Johanson, C. S. McHenry, The beta subunit of the DNA polymerase III holoenzyme becomes inaccessible to antibody after formation of an initiation complex with primed DNA. *J. Biol. Chem.* **257**, 12310–12315 (1982).

14. J. Josse, A. D. Kaiser, A. Kornberg, Enzymatic Synthesis of Deoxyribonucleic Acid: VIII. Frequencies of Nearest Neighbor Base Sequences in Deoxyribonucleic Acid. *J. Biol. Chem.* **237**, 864–875 (1962).

15. R. Okazaki, T. Okazaki, K. Sakabe, K. Sugimoto, A. Sugino, Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc. Natl. Acad. Sci. U. S. A.* **59**, 598–605 (1968).

16. K. Sakabe, R. Okazaki, A unique property of the replicating region of chromosomal DNA. *Biochim. Biophys. Acta - Nucleic Acids Protein Synth.* **129**, 651–654 (1966).

17. M. M. Gottesman, M. L. Hicks, M. Gellert, Genetics and function of DNA ligase in *Escherichia coli*. *J. Mol. Biol.* **77**, 531–547 (1973).

18. R. D. Shereda, A. G. Kozlov, T. M. Lohman, M. M. Cox, J. L. Keck, SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* **43**, 289–318 (2008).

19. A. Costes, F. Lecointe, S. McGovern, S. Quevillon-Cheruel, P. Polard, The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS Genet.* **6**, 1–15 (2010).

20. P. J. Fay, K. O. Johanson, C. S. McHenry, R. A. Bambara, Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of *Escherichia coli*. *J. Biol. Chem.* **256**, 976–983 (1981).

21. C. S. McHenry, DNA replicases from a bacterial perspective. *Annu. Rev. Biochem.* **80**, 403–436 (2011).

22. E. V Mirkin, S. M. Mirkin, Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.* **71**, 13–35 (2007).

23. M. M. Cox *et al.*, The importance of repairing stalled replication forks. *Nature*. **404**, 37–41 (2000).

24. M. M. Cox, Historical overview: searching for replication help in all of the rec places. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8173–8180 (2001).

25. S. Lusetti, M. M. Cox, The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Ann. Rev. Biochem*. **71**, 71–100 (2002).

26. A. Kuzminov, Collapse and repair of replication forks in *Escherichia coli*. *Mol. Microbiol.* **16**, 373–384 (1995).

27. A. Kuzminov, Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* **63**, 751–813 (1999).

28. J. T. P. Yeeles, K. J. Marians, The *Escherichia coli* replisome is inherently DNA damage

tolerant. *Science*. **334**, 235–8 (2011).

29. J. T. P. Yeeles, J. Poli, K. J. Marians, P. Pasero, Rescuing stalled or damaged replication forks. *Cold Spring Harb. Perspect. Biol.* **5**, 1–16 (2013).

30. B. Michel *et al.*, Rescue of arrested replication forks by homologous recombination. *Proc. Natl. Acad. Sci.* **98**, 8181–8188 (2001).

31. B. J. Brewer, When polymerases collide: Replication and the transcriptional organization of the *E. coli* chromosome. *Cell*. **53**, 679–686 (1988).

32. B. W. Trautinger, R. P. Jaktaji, E. Rusakova, R. G. Lloyd, RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol. Cell*. **19**, 247–258 (2005).

33. B. M. Alberts *et al.*, Studies on DNA Replication in the Bacteriophage T4 In Vitro System. *Cold Spring Harb. Symp. Quant. Biol.* **47**, 655–667 (1983).

34. T. Lindahl, Instability and decay of the primary structure of DNA. *Nature*. **362**, 709–715 (1993).

35. R. P. Fuchs, S. Fujii, Translesion DNA Synthesis and Mutagenesis in Prokaryotes. *Cold Spring Harb. Lab. Press*. **5**, 1–22 (2013).

36. M. M. Cox, The nonmutagenic repair of broken replication forks via recombination. *Mutat. Res. Mol. Mech. Mutagen.* **510**, 107–120 (2002).

37. B. Michel, G. Grompone, M.-J. Florès, V. Bidnenko, Multiple pathways process stalled replication forks. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12783–8 (2004).

38. R. Chayot, B. Montagne, D. Mazel, M. Ricchetti, An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **107**, 2141–2146 (2010).

39. B. Michel, S. J. Sandler, Replication restart in bacteria. *J. Bacteriol.* **199**, 1–13 (2017).

40. J. C. Bell, S. C. Kowalczykowski, RecA: Regulation and Mechanism of a Molecular Search Engine. *Trends Biochem. Sci.* **41**, 491–507 (2016).

41. E. P. C. Rocha, E. Cornet, B. Michel, Comparative and Evolutionary Analysis of the Bacterial Homologous Recombination Systems. *PLoS Genet.* **1**, e15 (2005).

42. Z. Chen, H. Yang, N. P. Pavletich, Mechanism of homologous recombination from the RecA–ssDNA/dsDNA structures. *Nature*. **453**, 489–484 (2008).

43. M. Prentiss, C. Prévost, C. Danilowicz, Structure/function relationships in RecA protein-mediated homology recognition and strand exchange. *Crit. Rev. Biochem. Mol. Biol.* **50**, 453–476 (2015).

44. K. McEntee, Protein X is the product of the recA gene of *Escherichia coli*. *Proc. Natl.*

Acad. Sci. U. S. A. **74**, 5275–5279 (1977).

45. M. Meghna Patel, Q. Jiang, R. Woodgate, M. Cox, M. Goodman, A New Model for SOS-induced Mutagenesis: How RecA Protein Activates DNA Polymerase V. *Crit Rev Biochem Mol Biol*. **45**, 171–184 (2010).

46. J. J. Churchill, D. G. Anderson, S. C. Kowalczykowski, The RecBC enzyme loads recA protein onto ssDNA asymmetrically and independently of chi, resulting in constitutive recombination activation. *Genes Dev.* **13**, 901–911 (1999).

47. P. R. Bianco *et al.*, Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature*. **409**, 374–8 (2001).

48. D. A. Dixon, S. C. Kowalczykowski, The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell*. **73**, 87–96 (1993).

49. D. G. Anderson, S. C. Kowalczykowski, The recombination hot spot chi is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme. *Genes Dev.* **11**, 571–581 (1997).

50. P. R. Bianco, S. C. Kowalczykowski, The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as the single strand of DNA containing the sequence 5'-GCTGGTGG-3'. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6706–11 (1997).

51. J. C. Bell, J. L. Plank, C. C. Dombrowski, S. C. Kowalczykowski, Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA. *Nature*. **491**, 274–278 (2012).

52. M. Ryzhikov, O. Koroleva, D. Postnov, A. Tran, S. Korolev, Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein. *Nucleic Acids Res.* **39**, 6305–6314 (2011).

53. M. S. Dillingham, S. C. Kowalczykowski, RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.* **72**, 642–71, Table of Contents (2008).

54. S. T. Lovett, C. Luisi-DeLuca, R. D. Kolodner, The genetic dependence of recombination in recD mutants of *Escherichia coli*. *Genetics*. **120**, 37–45 (1988).

55. K. C. Smith, T. C. V. Wang, R. C. Sharma, recA-Dependent DNA repair in UV-irradiated *Escherichia coli*. *J. Photochem. Photobiol. B Biol.* **1**, 1–11 (1987).

56. M. M. Cox, I. R. Lehman, recA Protein-promoted DNA Strand Exchange. *J. Biol. Chem.* **257**, 8523–8532 (1982).

57. M. M. Cox, I. R. Lehman, Directionality and polarity in recA protein-promoted branch migration. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6018–22 (1981).

58. C. Danilowicz, D. Yang, C. Kelley, C. Prévost, M. Prentiss, The poor homology stringency in the heteroduplex allows strand exchange to incorporate desirable mismatches without sacrificing recognition in vivo. *Nucleic Acids Res.* **43**, 6473–6485 (2015).

59. M. M. Cox, Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* **42**, 41–63 (2007).

60. R. Holliday, A mechanism for gene conversion in fungi. *Genet. Res.* **5**, 282–304 (1964).

61. Y. Liu, S. C. West, Happy Hollidays: 40th anniversary of the Holliday junction. *Nat. Rev. Mol. Cell Biol.* **5**, 937–946 (2004).

62. S. C. West, The RuvABC proteins and Holliday junction processing in *Escherichia coli*. *J. Bacteriol.* **178**, 1237–1241 (1996).

63. S. C. West, Processing of Recombination Intermediates By the RuvABC Proteins. *Annu. Rev. Genet.* **31**, 213–244 (1997).

64. G. J. Sharples, S. N. Chan, a a Mahdi, M. C. Whitby, R. G. Lloyd, Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. *EMBO J.* **13**, 6133–6142 (1994).

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

66. B. Müller, I. R. Tsaneva, S. C. West, Branch migration of holliday junctions promoted by the *Escherichia coli* RuvA and RuvB proteins: I. Comparison of RuvAB- and RuvB-mediated reactions. *J. Biol. Chem.* **268**, 17179–17184 (1993).

67. E. M. Witkin, The radiation sensitivity of *Escherichia coli* B: a hypothesis relating filament formation and prophage induction. *Proc. Natl. Acad. Sci. U. S. A.* **57**, 1275–1279 (1967).

68. T. Kato, Y. Shinoura, Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. *Mol. Gen. Genet.* **156**, 121–131 (1977).

69. G. Steinborn, Uvm mutants of *Escherichia coli* K 12 deficient in UV mutagenesis - I. Isolation of uvm mutants and their phenotypical characterization in DNA repair and mutagenesis. *MGG Mol. Gen. Genet.* **165**, 87–93 (1978).

70. M. Radman, SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sci*, 355–67 (1975).

71. M. F. Goodman, R. Woodgate, Translesion DNA polymerases. *Cold Spring Harb. Perspect. Biol.* **5**, 247–250 (2013).

72. T. Kornberg, M. L. Gefter, Purification and DNA synthesis in cell-free extracts: properties of DNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* **68**, 761–764 (1971).

73. R. Knippers, DNA Polymerase II. *Nature*. **227**, 815–817 (1970).

74. C. A. Bonner *et al.*, Purification and characterization of an inducible *Escherichia coli* DNA polymerase capable of insertion and bypass at abasic lesions in DNA. *J. Biol. Chem.* **263**, 18946–18952 (1988).

75. T. Paz-Elizur, M. Takeshita, M. Goodman, M. O'Donnell, Z. Livneh, Mechanism of translesion DNA synthesis by DNA polymerase II. Comparison to DNA polymerases I and III core. *J Biol Chem*. **271**, 24662–24669 (1996).

76. R. Napolitano, R. Janel-Bintz, J. Wagner, R. P. Fuchs, All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J.* **19**, 6259–65 (2000).

77. O. J. Becherel, R. P. Fuchs, Mechanism of DNA polymerase II-mediated frameshift mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8566–71 (2001).

78. O. J. Becherel, R. P. P. Fuchs, J. Wagner, Pivotal role of the beta-clamp in translesion DNA synthesis and mutagenesis in *E. coli* cells. *DNA Repair (Amst).* **1**, 703–708 (2002).

79. C. A. Bonner *et al.*, Processive DNA synthesis by DNA polymerase II mediated by DNA polymerase III accessory proteins. *J. Biol. Chem.* **267**, 11431–11438 (1992).

80. J. E. Kath *et al.*, Exchange between *Escherichia coli* polymerases II and III on a processivity clamp. *Nucleic Acids Res.* **44**, 1681–1690 (2015).

81. A. B. Williams, K. M. Hetrick, P. L. Foster, Interplay of DNA repair, homologous recombination, and DNA polymerases in resistance to the DNA damaging agent 4-nitroquinoline-1-oxide in *Escherichia coli*. *DNA Repair (Amst).* **9**, 1090–1097 (2010).

82. M. Escarceller *et al.*, Involvement of *Escherichia coli* DNA Polymerase II in Response to Oxidative Damage and Adaptive Mutation. *J. Bacteriol.* **10**, 6221–6228 (1994).

83. C. J. Kenyon, G. C. Walker, DNA-damaging agents stimulate gene expression at specific loci in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **77**, 2819–2823 (1980).

84. S. R. Kim, K. Matsui, M. Yamada, P. Gruz, T. Nohmi, Roles of chromosomal and episomal dinB genes encoding DNA pol IV in targeted and untargeted mutagenesis in *Escherichia coli*. *Mol. Genet. Genomics*. **266**, 207–215 (2001).

85. K. Uchida *et al.*, Overproduction of *Escherichia coli* DNA polymerase DinB (Pol IV) inhibits replication fork progression and is lethal. *Mol. Microbiol.* **70**, 608–622 (2008).

86. J. Wagner, S. Fujii, P. Gruz, T. Nohmi, R. P. Fuchs, The beta clamp targets DNA polymerase IV to DNA and strongly increases its processivity. *EMBO Rep.* **1**, 484–488

(2000).

87. J. M. H. Heltzel, R. W. Maul, S. K. Scouten Ponticelli, M. D. Sutton, A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc. Natl. Acad. Sci.* **106**, 12664–12669 (2009).

88. J. E. Kath *et al.*, Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. *Proc. Natl. Acad. Sci.* **111**, 7647–7652 (2014).

89. M. K. Scotland *et al.*, A Genetic Selection for dinB Mutants Reveals an Interaction between DNA Polymerase IV and the Replicative Polymerase That Is Required for Translesion Synthesis. *PLoS Genet.* **11**, 1–29 (2015).

90. A. Gutierrez, M. Elez, O. Clermont, E. Denamur, I. Matic, *Escherichia coli* YafP protein modulates DNA damaging property of the nitroaromatic compounds. *Nucleic Acids Res.* **39**, 4192–4201 (2011).

91. K. R. Ona, C. T. Courcelle, J. Courcelle, Nucleotide excision repair is a predominant mechanism for processing nitrofurazone-induced DNA damage in *Escherichia coli*. *J. Bacteriol.* **191**, 4959–4965 (2009).

92. M. Tang *et al.*, UmuD'(2)C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8919–24 (1999).

93. N. B. Reuven, G. Arad, A. Maor-Shoshani, Z. Livneh, The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J. Biol. Chem.* **274**, 31763–31766 (1999).

94. S. Sommer, A. Bailone, R. Devoret, The appearance of the UmuD'C protein complex in *Escherichia coli* switches repair from homologous recombination to SOS mutagenesis. *Mol. Microbiol.* **10**, 963–971 (1993).

95. A. Robinson *et al.*, Regulation of Mutagenic DNA Polymerase V Activation in Space and Time. *PLoS Genet.* **11**, 1–30 (2015).

96. K. Schlacher, M. M. Cox, R. Woodgate, M. F. Goodman, RecA acts in trans to allow replication of damaged DNA by DNA polymerase V. *Nature*. **442**, 883–887 (2006).

97. R. P. Fuchs, S. Fujii, J. Wagner, Properties and functions of *Escherichia coli*: Pol IV and Pol V. *Adv. Protein Chem.* **69**, 229–264 (2004).

98. S. Fujii, V. Gasser, R. P. Fuchs, The biochemical requirements of DNA polymerase V-mediated translesion synthesis revisited. *J. Mol. Biol.* **341**, 405–417 (2004).

99. M. Tang *et al.*, Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature*. **404**, 1014–1018 (2000).

100. C. T. Courcelle, J. J. Belle, J. Courcelle, Nucleotide excision repair or polymerase V-

mediated lesion bypass can act to restore UV-arrested replication forks in *Escherichia coli*. *J. Bacteriol.* **187**, 6953–6961 (2005).

101. A. J. Gruber *et al.*, A RecA Protein Surface Required for Activation of DNA Polymerase V. *PLoS Genet.* **11**, 1–37 (2015).

102. W. D. Rupp, P. Howard-Flanders, Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J. Mol. Biol.* **31**, 291–304 (1968).

103. R. P. Fuchs, Tolerance of lesions in *E. coli*: Chronological competition between Translesion Synthesis and Damage Avoidance. *DNA Repair (Amst).* **44**, 51–58 (2016).

104. J. T. P. Yeeles, K. J. Marians, Dynamics of leading-strand lesion skipping by the replisome. *Mol. Cell*. **52**, 855–865 (2013).

105. C. B. Gabbai, J. T. P. Yeeles, K. J. Marians, Replisome-mediated translesion synthesis and leading strand template lesion skipping are competing bypass mechanisms. *J. Biol. Chem.* **289**, 32811–32823 (2014).

106. X. Li, K. J. Marians, Two distinct triggers for cycling of the lagging strand polymerase at the replication fork. *J. Biol. Chem.* **275**, 34757–34765 (2000).

107. Q. Yuan, C. S. McHenry, Cycling of the *E. coli* lagging strand polymerase is triggered exclusively by the availability of a new primer at the replication fork. *Nucleic Acids Res.* **42**, 1747–1756 (2014).

108. F.-X. Barre *et al.*, Circles: the replication-recombination-chromosome segregation connection. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8189–8195 (2001).

109. I. F. Lau *et al.*, Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* **49**, 731–743 (2003).

110. D. J. Sherratt *et al.*, Recombination and chromosome segregation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**, 61–69 (2004).

111. T. Shibata *et al.*, Functional overlap between RecA and MgsA (RarA) in the rescue of stalled replication forks in *Escherichia coli*. *Genes to Cells*. **10**, 181–191 (2005).

112. B. Michel, A. K. Sinha, The inactivation of *rfaP* , *rarA* or *sspA* gene improves the viability of the *Escherichia coli* DNA polymerase III *holD* mutant. *Mol. Microbiol.* **104**, 1008–1026 (2017).

113. A. N. Page, N. P. George, A. H. Marceau, M. M. Cox, J. L. Keck, Structure and biochemical activities of *Escherichia coli* MgsA. *J. Biol. Chem.* **286**, 12075–12085 (2011).

114. J. P. Erzberger, J. M. Berger, Evolutionary Relationships and Structural Mechanisms of

Aaa+ Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).

115. T. Hishida, H. Iwasaki, T. Ohno, T. Morishita, H. Shinagawa, A yeast gene, MGS1, encoding a DNA-dependent AAA(+) ATPase is required to maintain genome stability. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8283–9 (2001).

116. T. Hishida, T. Ohno, *Saccharomyces cerevisiae* MGS1 is essential in strains deficient in the RAD6 -dependent DNA damage tolerance pathway. *EMBO J.* **21**, 2019–2029 (2002).

117. D. Branzei, M. Seki, F. Onoda, T. Enomoto, The product of *Saccharomyces cerevisiae* WHIP/MGS1, a gene related to replication factor C genes, interacts functionally with DNA polymerase Delta. *Mol. Genet. Genomics*. **268**, 371–386 (2002).

118. N. Yao, M. O'Donnell, Bacterial and Eukaryotic Replisome Machines. *JSM Biochem. Mol. Biol.* **3**, 1–15 (2016).

119. I. Saugar, J. L. Parker, S. Zhao, H. D. Ulrich, The genome maintenance factor Mgs1 is targeted to sites of replication stress by ubiquitylated PCNA. *Nucleic Acids Res.* **40**, 245–257 (2012).

120. T. Hishida, T. Ohya, Y. Kubota, Y. Kamada, H. Shinagawa, Functional and physical interaction of yeast Mgs1 with PCNA: impact on RAD6-dependent DNA damage tolerance. *Mol. Cell. Biol.* **26**, 5509–17 (2006).

121. H. Nomura *et al.*, WRNIP1 accumulates at laser light irradiated sites rapidly via its ubiquitin-binding zinc finger domain and independently from its ATPase domain. *Biochem. Biophys. Res. Commun.* **417**, 1145–1150 (2012).

122. N. Crosetto *et al.*, Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J. Biol. Chem.* **283**, 35173–35185 (2008).

123. A. Yoshimura, Y. Kobayashi, S. Tada, M. Seki, T. Enomoto, WRNIP1 functions upstream of DNA polymerase eta in the UV-induced DNA damage response. *Biochem. Biophys. Res. Commun.* **452**, 48–52 (2014).

124. J. H. Kim *et al.*, In vivo and in vitro studies of Mgs1 suggest a link between genome instability and Okazaki fragment processing. *Nucleic Acids Res.* **33**, 6137–6150 (2005).

125. T. Tsurimoto, A. Shinozaki, M. Yano, M. Seki, T. Enomoto, Human Werner helicase interacting protein 1 (WRNIP1) functions as novel modulator for DNA polymerase delta. *Genes to Cells*. **10**, 13–22 (2005).

126. A. Yoshimura *et al.*, Physical and functional interaction between WRNIP1 and RAD18. *Genes Genet. Syst.* **84**, 171–178 (2009).

127. N. Acharya *et al.*, Roles of PCNA-binding and ubiquitin-binding domains in human DNA polymerase eta in translesion DNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17724–17729 (2008).

128. C. Hoege, B. Pfander, G.-L. Moldovan, G. Pyrowolakis, S. Jentsch, RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature*. **419**, 135–141 (2002).

129. Y. I. Kawabe *et al.*, A Novel Protein Interacts with the Werner's Syndrome Gene Product Physically and Functionally. *J. Biol. Chem.* **276**, 20364–20369 (2001).

130. W. K. Chu, I. D. Hickson, RecQ helicases: multifunctional genome caretakers. *Nat. Rev. Cancer*. **9**, 644–654 (2009).

131. V. A. Bohr, Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. *Trends Biochem. Sci.* **33**, 609–620 (2008).

132. D. L. Croteau, V. Popuri, P. L. Opresko, V. A. Bohr, Human RecQ Helicases in DNA Repair, Recombination, and Replication. *Annu. Rev. Biochem.* **83**, 519–552 (2014).

133. M. Walker, J. G. Kublin, J. R. Zunt, Flap Endonuclease 1. *Annu. Rev. Biochem.* **82**, 119–138 (2013).

134. A. Yoshimura, M. Seki, T. Enomoto, The role of WRNIP1 in genome maintenance. *Cell Cycle*. **16**, 515–521 (2017).

135. J. S. Lewis *et al.*, Single-molecule visualization of fast polymerase turnover in the bacterial replisome. *Elife*. **6**, 1–17 (2017).

136. D. L. Nelson, M. M. Cox, *Lehninger Principles of Biochemistry 6th ed.* (2013).

137. T. Helleday, S. Eshtad, S. Nik-Zainal, Mechanisms underlying mutational signatures in human cancers. *Nat. Rev. Genet.* **15**, 585–598 (2014).

Chapter 2: Biochemical characterization of the *E. coli* RarA protein

Portions of this chapter have been published previously:

T.H. Stanage, A.N. Page, M.M. Cox, DNA flap creation by the MgsA/RarA protein of *Escherichia coli*. *Nucleic Acids Res.* **45**, 2724-2735 (2017).

Introduction

When a cellular replication fork encounters damage or barriers in the DNA template, it can stall or collapse (*1–11*). The damage/barriers may be DNA lesions, template strand breaks, actively transcribing RNA polymerases, or other tightly bound protein-DNA complexes (*12*, *13*). Failure to recognize, respond to, and resolve these replication fork encounters in a timely manner can threaten genome integrity and cell viability. As such, all cells have evolved diverse pathways to circumvent or directly repair replication fork impediments of all types. In *Escherichia coli*, two major pathways exist to facilitate the repair or tolerance of DNA damage encountered by the replication fork. The first pathway is translesion DNA synthesis, which utilizes specialized DNA polymerases to bypass sites of DNA damage not normally tolerated by DNA polymerase III (*14–16*). The second pathway is homologous recombination, which utilizes the bacterial recombinase RecA to repair double-stranded breaks (DSBs) and single-stranded DNA gaps (SSGs) using a non-mutagenic mechanism (*1*, *2*, *12*, *17*). For decades, work has focused on the repair of DNA damage and the biochemical characterization of the proteins involved in these pathways. However, how these substrates, specifically SSGs, are formed during replication remains highly enigmatic, despite evidence of their formation *in vivo* nearly 50 years ago (*18*). I present biochemical data in this chapter that supports the hypothesis that SSG formation at the replisome is an active process catalyzed by the highly conserved *E. coli* RarA protein.

The RarA (Replication associated recombination protein A) is part of what may be the most highly conserved family of DNA metabolism proteins, sharing 40% amino acid identity and 56-58% similarity with Mgs1 in yeast and WRNIP1 in mammals (*19*, *20*). The protein has also been called MgsA (Maintenance of genome stability A) to better identify it with its yeast homolog (*21*). The RarA protein is most closely related to two proteins in *E. coli*: the DNA polymerase III clamp

loader protein encoded by the *dnaX* gene and the Holliday Junction helicase, RuvB (*19*). Indeed, the proteins in the RarA family are members of the clamp loader clade of the AAA+ (ATPases associated with various cellular activities) superfamily, but adopt tetrameric complexes as opposed to the more common pentameric conformation (*22*). Members of the AAA+ superfamily are oligomeric enzymes that convert the chemical energy of ATP binding and hydrolysis to mechanical energy used to carry out various cellular processes, including functions associated with DNA replication (*23–26*).

Considerable work has been conducted with RarA/MgsA family proteins both *in vivo* and *in vitro*, but their function remains highly enigmatic. *In vivo*, all of the proteins co-localize with the replication fork throughout the cell cycle, both in the absence (*19*, *27–30*) and presence (*31–33*) of exogenous DNA damage. The localization is mediated by an interaction with single-stranded DNA binding protein (SSB) in bacteria (*22*, *27*) and ubiquitylated PCNA in eukaryotes (*31*, *32*, *34*, *35*). Knockout mutations tend to exhibit a modest phenotype, at most a slight growth defect and/or a tendency for hyperrecombination (*19*, *36*, *37*). However, combining a deletion of a *rarA* family gene with a defect in another gene involved in DNA repair or replication often results in an additively or synergistically enhanced phenotype (*19–21*, *29*, *30*, *37–44*). Overexpression of any RarA family protein results in severe genome instability and is lethal in *E. coli* strains lacking *recA* (*21*, *29*, *37*). Despite a wealth of *in vivo* experiments, the precise function of RarA family proteins remains unknown.

In vitro, numerous studies have shed limited light on the RarA protein. The *apo* X-ray crystal structure of the *E. coli* RarA protein revealed marked similarity between RarA and clamp loader proteins from various organisms (*22*). However, RarA adopts a unique homotetrameric configuration rather than the heteropentameric form seen amongst clamp loaders and other AAA+

proteins (*22*, *45*, *46*). The RarA protein family exhibits a DNA-dependent ATPase activity that is stimulated by circular single-stranded DNA (*22*, *37*, *40*), and in once case (WRNIP1) by double-stranded DNA ends (*47*). Both Mgs1 and WRNIP1 bind ssDNA, ssDNA-dsDNA junctions, and forked DNA in the presence or absence of ATP (*30*, *40*). RarA homologs physically interact with the processivity clamp PCNA in yeast, and the lagging strand DNA polymerase δ in both yeast and humans (*34*, *47*, *48*). These interactions have been hypothesized to both localize Mgs1/WRNIP1 to sites of replication fork stress and to directly stimulate DNA polymerase δ activity at the replication fork. RarA also tightly binds the conserved C-terminus of the single-stranded DNA binding protein (SSB) in both *E. coli* and *B. subtilis* (*22*, *27*). This interaction is required for RarA to localize to the replication fork, which is a characteristic with other SSB-interacting proteins, including the branch migration helicase RecG, the replication restart helicase PriA, and over a dozen additional proteins (*27*, *49*, *50*).

In this study, we further characterize the biochemical properties of the *E. coli* RarA protein using a combination of ATPase assays, DNA binding studies, fluorescence spectroscopy, and single-molecule DNA replication assays. We reveal an ATPase-dependent helicase activity operating at double-stranded DNA ends. In addition, we observe that RarA catalyzes single-stranded DNA gap formation behind actively replicating DNA polymerases *in vitro*. We hypothesize that RarA ensures optimal replisome progression by catalyzing DNA lesion bypass.

Results

RarA ATPase activity is stimulated by double-stranded DNA ends

Multiple previous reports on RarA and yeast Mgs1 indicated that the ATPase activity of proteins in this family was stimulated by single-stranded DNA (ssDNA) (*22*, *37*, *40*). One report on WRNIP1 indicated that the ATPase was stimulated by DNA ends (*47*). This could represent a distinction between family members. However, we note that in all previous reports of RarA family protein stimulation by ssDNA, the ssDNA employed was a derivative of M13 phage DNA that has a considerable degree of secondary structure (*51*, *52*). We considered the possibility that the secondary structure may have led to a misinterpretation of results.

We examined the response of the RarA DNA-dependent ATPase activity to a series of DNA substrates with different structures. RarA exhibits a weak DNA-independent ATPase activity (0.8 ± 0.20 μM min^{-1} under the conditions of this experiment with 0.5 μM RarA monomers; **Figure 2-1A**). A 16-nucleotide single-stranded DNA substrate (ssDNA; 0.5 μM molecules) with no secondary structure did not increase the ATPase rate (0.7 ± 0.10 μM min^{-1}) when compared to DNA-independent reactions. However, when a 16-base pair duplex (dsDNA; 0.5 μM molecules) was incubated in the reaction, ATPase activity was strongly stimulated (68.0 ± 2.0 μM min^{-1}) (**Figure 2-1A**). These results double-stranded DNA was the actual primary substrate for RarA ATPase activity.

To determine whether the interior or the ends of the dsDNA molecule were responsible for stimulating RarA ATPase activity, three DNA substrates were included in different ATPase reactions: circular double-stranded pUC19 DNA (cds DNA; 2686 bp), linear double-stranded M13mp8 DNA (lds DNA; 7229 bp), and a 48-base pair duplex. All reactions contained equimolar

concentrations of DNA as measured in nucleotides (48 μM), but greatly differing concentrations of dsDNA ends. Circular double-stranded DNA (no dsDNA ends) did not significantly stimulate ATPase activity compared to a reaction with no DNA present (**Figure 2-1B**). Similarly, linear double-stranded DNA (0.018 μM dsDNA ends) only marginally increased RarA ATPase rate (**Figure 2-1B**). However, a reaction containing a 48-base pair duplex (1 μM dsDNA ends) strongly stimulated RarA ATPase activity when compared to all other substrates (**Figure 2-1B**). Figure 2-1AB shows that double-stranded DNA ends maximally stimulated RarA ATPase activity, while closed-circular dsDNA and short ssDNA did not increase the rate of ATP hydrolysis relative to reactions lacking DNA.

RarA ATPase activity is affected by DNA end structure

To determine the effects of dsDNA end structure in stimulating RarA ATPase activity, we designed a DNA duplex hairpin substrate containing a single dsDNA end. This 36 nucleotide hairpin substrate consisted of a 16 base pair duplex with a four nucleotide 5'-GTAA-3' hairpin. This hairpin sequence was selected due to its propensity to form a hairpin and its thermostability (*53*). We altered this dsDNA end structure with single-stranded DNA overhangs and examined their effects on RarA ATPase activity. In all cases, the concentration of DNA molecules, and thus free ends, was kept constant. A blunt end duplex maximally stimulated RarA ATPase activity (26.21 ± 1.80 μM min^{-1}), followed by a dsDNA end with an 18-nucleotide 5' ssDNA overhang (19.79 ± 0.81 μM min^{-1}) (**Figure 2-1C**). A duplex end modified with an 18-nucleotide 3' ssDNA overhang modestly stimulated RarA ATPase activity (6.71 ± 0.50 μM min^{-1}) while a structure lacking exposed dsDNA ends (both ends capped in a hairpin structure) did not significantly stimulate ATPase activity when compared to a blunt end or 5' overhang substrate (**Figure 2-1C**). We note that both substrate specificity and ATPase rates are maintained in a reaction buffer with

physiological ionic strength (**Figure 2-2**). From these data, we identified both a blunt end duplex and a duplex modified with a short 5' ssDNA overhang as the most effective substrates for stimulating RarA ATPase activity, although all three substrates were stimulatory.

Single-stranded DNA gaps stimulate RarA ATPase activity

Single-stranded DNA gaps are common structures found at stalled replication forks *in vivo*. As such, we designed DNA substrates containing one single-stranded DNA gap of varying size and tested their ability to stimulate RarA ATPase activity. As previously noted, a closed-circle DNA substrate did not significantly stimulate RarA ATPase activity. This situation is reprised here with a 72 nucleotide substrate containing 32 base pairs of duplex DNA, two four-nucleotide 5'-GTAA-3' hairpins at either end, and no strand breaks (**Figure 2-3A**). However, when a single nick is present in the same substrate (in the center), ATPase activity is notably increased (9.91 ± 0.47 $\mu M\ min^{-1}$) (**Figure 2-3A**). The rate of RarA ATP hydrolysis increases further with the size of the ssDNA gap, up to eight nucleotides (20.79 ± 1.65 $\mu M\ min^{-1}$) (Figure 2-3A). Substrates containing a gap larger than eight nucleotides did not further increase the rates of ATP hydrolysis (**Figure 2-3A**). It should be noted that the stimulation of RarA ATPase activity by ssDNA gaps did not depend on the presence of a 5' phosphate or 3' hydroxyl group, as DNA substrates lacking either of these features still stimulated activity (**Figure 2-3B**).

RarA ATPase activity is equivalent at dsDNA ends and ssDNA gaps

To compare the stimulatory effect of a dsDNA end and ssDNA gap on RarA ATPase activity, a Michaelis-Menten kinetic analysis of reactions containing these substrates was conducted. ATP concentrations were held constant and in excess, while DNA substrate concentration was varied and ATPase rates measured. In effect, the results measure RarA binding to DNA ends, as seen indirectly by ATPase stimulation. RarA ATPase activity in the presence of

a blunt-end dsDNA substrate had a V_{max} of 59.76 ± 2.04 μM min^{-1} and an apparent K_m of 48.71 ± 6.21 nM (**Figure 2-4A**). Similarly, RarA ATPase activity in the presence of an 18-nucleotide ssDNA gap had a V_{max} of 56.45 ± 2.64 μM min^{-1} and an apparent K_m of 51.40 ± 8.03 nM (**Figure 2-4B**).

Finally, the effect of ATP concentration on RarA ATPase activity was investigated by conducting a Michaelis-Menten kinetic analysis of reactions containing saturating amounts of DNA and increasing amounts of ATP. The measured V_{max} for this reaction was 70.98 ± 2.08 μM min^{-1}, the measured K_m for the ATP substrate was 88.03 ± 7.15 μM, and the calculated k_{cat} was 141.96 min^{-1} (**Figure 2-4C**).

RarA binds to double-stranded DNA in the presence of ATP, and to single-stranded DNA in the absence of nucleotide cofactor

We next examined DNA binding directly, and also investigated the effect of nucleotide cofactor on RarA binding to double-stranded and single-stranded DNA. Using a fluorescently-labeled 16 base pair duplex or 16 nucleotide single-stranded DNA substrate, equilibrium fluorescence polarization binding studies were conducted.

RarA exhibited a very weak affinity for dsDNA in the absence of nucleotide cofactor or in the presence of ADP, with an apparent dissociation constant ($K_{d,app}$) or > 10 μM (**Figure 2-5A**). However, when the minimally-hydrolyzable ATP analog ATPγS was included in binding reactions, RarA possessed a much greater affinity for dsDNA, with an apparent dissociation constant of 58 ± 3 nM (**Figure 2-5A**). Note that the K_d value of RarA binding dsDNA in the presence of ATPγS compares very well with the apparent K_m derived from ATPase values determined as DNA end concentration was varied (**Figures 2-4AB**). Conversely, RarA exhibited a very weak affinity for ssDNA in the presence of ATPγS or ADP ($K_{d,app}$ = > 10000 nM and 9000

± 1000 nM, respectively), but a significant affinity in the absence of nucleotide cofactor ($K_{d,app}$ = 890 ± 28 nM) (**Figure 2-5B**).

RarA ATPase activity is stimulated by AT-rich duplexes

We next sought to determine if the nucleotide sequence present at DNA ends affected the RarA ATPase activity. Using the previously described hairpin duplex substrate as a template, six substrates were designed. A duplex comprised entirely of G-C base pairs was designed and then modified to contain an increasing number of A-T base pairs from the dsDNA end inward to the hairpin. The resulting five DNA substrates, as well as a duplex comprised entirely of A-T base pairs, were included in ATPase reactions with RarA protein.

The GC-rich duplex simulated RarA ATPase activity at a rate lower than the blunt end duplex examined in Figure 2-1C (20.01 ± 1.80 μM min^{-1}) (**Figure 2-6**). Conversion of the terminal-most base pair from G-C to A-T resulted in a significant increase in RarA ATPase activity (**Figure 2-6**). As more G-C base pairs were substituted with A-T base pairs in the substrate, the rate of ATP hydrolysis increased further. A duplex comprised entirely of A-T base pairs stimulated RarA ATPase activity maximally (35.24 ± 2.31 μM min^{-1}) (**Figure 2-6**). This result suggested that RarA might engage in the catalysis of strand separation at DNA ends.

RarA protein separates the strands of double-stranded DNA ends

To investigate the effect of RarA ATP hydrolysis on dsDNA ends, we used two assays. The first is an ensemble fluorescence intensity assay. We designed a hairpin duplex substrate (molecular beacon) modified with a 5' 6-carboxyfluorescein (6-FAM) fluorophore and a 3' 3-dabcyl quencher (**Figures 2-7AB**). When the DNA is folded, the fluorescence of the 6-FAM fluorophore is quenched by the 3-dabcyl quencher. When the strands of the DNA substrate are separated, the fluorescence intensity increases. This DNA substrate was included in ATPase

reactions with RarA protein while the fluorescence intensity of fluorescein was measured over time.

Addition of RarA protein to reactions containing the labeled DNA substrate resulted in an increase in fluorescence intensity over time (**Figure 2-7A**). Increasing the amount of RarA included in these reactions resulted in an increase in overall fluorescence intensity in a protein concentration-dependent manner (**Figure 2-7A**).

RarA R156A is a variant protein lacking the arginine finger residue required by AAA+ proteins to hydrolyze ATP. When this variant protein was included in a reaction containing the 6-FAM/dabcyl labeled DNA substrate, the increase in fluorescence intensity was dramatically less than that observed in a reaction containing wild-type protein (**Figure 2-7B**). These results suggest that ATP binding or hydrolysis is required for RarA to separate the strands of dsDNA ends. We note that the labels on this DNA substrate may interfere with RarA binding to the ends to some degree.

To better characterize this strand separation activity, and to control for DNA end structure, we set up a second assay using a quite different DNA substrate and stopped-flow spectrophotometry. For these experiments, we designed another hairpin duplex substrate containing 21 base pairs. This substrate contained a 2-aminopurine at the penultimate position on the 3' end of the strand (**Figure 2-7CD**). 2-aminopurine is a fluorescent base analog that mimics adenine both in its structure and its ability to base-pair with thymine (*54–57*). Additionally, its fluorescence intensity is naturally quenched when it is found in a duplex and unquenched when single-stranded (*54–57*). These properties make 2-aminopurine well suited for probing the interactions between RarA protein and dsDNA ends.

The fluorescence intensity of 2-aminopurine was monitored over a 30 second time period in reactions containing RarA protein and DNA upon addition of indicated nucleotide cofactor (**Figure 2-7C**). In a reaction lacking nucleotide cofactor, 2-aminopurine fluorescence marginally increased (**Figure 2-7C**). When ATP was included in the reaction, a modest sharp increase in fluorescence was observed in the first 150 ms, after which the fluorescence reached its peak (**Figure 2-7C**). Interestingly, when the minimally-hydrolyzable ATP analog ATPγS was included in the reaction, a large and steady increase in fluorescence was observed in the first 15 seconds (a shorter time scale than those seen in the assay presented in panels A and B), after which the fluorescence reached its peak (**Figure 2-7C**). However, when the ATPase-dead RarA R156A mutant was included in this reaction, fluorescence intensity only modestly increased (**Figure 2-7C**). Finally, we investigated whether this activity was RarA concentration-dependent. We included the indicated amounts of RarA protein in the presence of ATPγS and observed the fluorescence intensity over time. We observed an increase in both the rate of increase and overall fluorescence intensity when higher concentrations of RarA were included in these reactions (**Figure 2-7D**).

We note that the order of component addition was different between the method used in Figure 2-7A and B and the method used in Figure 2-7C and D. In the fluorimeter method (**Figure 2-7AB**), reactions were initiated when RarA was added to a buffered solution of ATP and DNA substrate, while in the stopped-flow method (**Figure 2-7CD**) RarA was pre-incubated with DNA and reactions were initiated upon addition of nucleotide cofactor. To ensure that the order of component addition did not significantly affect the resulted observed with the stopped-flow method, we performed a variation of the stopped-flow experiment where two different orders of addition were directly compared. In this experiment (**Figure 2-7E**), the addition of DNA and

ATPγS to RarA yielded a higher final signal, although the rate of approach to the endpoint was similar in both cases. This rate was in all cases faster than that observed in the experiments employing the 6-FAM/dabcyl-labeled DNA substrate. Thus, while the order of component addition does not significantly affect the overall activity, the DNA end modification present in the experiments in Figure 2-7AB may have affected the kinetics of this reaction.

RarA creates ssDNA gaps behind replicating Pol III replisomes in vitro

To observe the effects of RarA at replication forks, we utilized a single-molecule DNA replication assay (*58–60*). This work was done in collaboration with Antoine van Oijen's laboratory at the University of Wollongong in Wollongong, Australia. All single-molecule DNA replication experiments were conducted by Jacob Lewis, Lisanne Spenkelink, and Slobodan Jergic. This assay employs a rolling-circle DNA amplification scheme, allowing observation of processive DNA synthesis by the *E. coli* replisome in real time (**Figure 2-8A**). A double-stranded (ds) circular DNA is anchored to the surface of a microfluidic flow cell through a biotinylated 5'-flap. This flap also facilitates loading of the DnaB helicase. Replication is then initiated by introducing a laminar flow of buffer with the components required for DNA synthesis. As replication proceeds, the newly synthesized leading strand becomes part of the circle and later acts as a template for lagging-strand synthesis. With the leading strand attached to the surface and the continuously growing DNA product stretched in the buffer flow, the dsDNA circle moves away from the anchor point. Replication is visualized in real-time near-TIRF fluorescence imaging of stained dsDNA (**Figure 2-8B**). This stratagem allows quantification of the instantaneous rates of individual replisomes and their processivities.

In the experiments documented in Figure 2-8B, the replisome was pre-assembled onto the rolling-circle template in solution. Subsequently, the template was attached to the surface of a flow

cell. The flow cell was then washed to remove all unbound proteins. Replication is initiated by introduction of a replication solution that omits Pol III* and helicase, but includes SSB, DnaN (the β_2 clamp), DnaG primase, rNTPs, and dNTPs. The absence of free Pol III* in solution makes polymerase exchange impossible (*60*). Nonetheless, these conditions support highly processive DNA replication with synthesis rates and processivities identical to a situation with Pol III* in solution and are consistent with values reported previously (*58*, *60–63*).

When RarA was included at a concentration of 300 nM (tetramer) in the replication reaction solution, numerous gaps appeared in the rolling circle products synthesized by individual replisomes (**Figure 2-8B**, middle; **Figure 2-8C**, bottom). The stain used to visualize the duplex DNA binds poorly to single-stranded (ss) DNA. Thus, any gaps in the product can be visualized as breaks in the fluorescence signal along the growing DNA molecules. Introducing the same concentration of an ATPase defective RarA protein (RarA K63R) did not produce gaps and had no evident effect on replisome progress (**Figure 2-8B**, bottom). Thus, the appearance of gaps is dependent upon both the presence of RarA and its ATPase activity. It should be noted that RarA and RarA K63R bind to the SSB C-terminus with very similar affinities (**Figure 2-9A**). The absence of gaps when the ATPase mutant is added provides evidence that gap formation is not an artifact of strong binding of RarA to SSB. In addition, these gaps are not seen if RarA is added post-replicatively to these single-molecule DNA replication assays (**Figure 2-9B**). This indicates that this activity is not due to either a non-replication associated function nor a nuclease contamination of the RarA protein preparation.

Addition of RarA or RarA K63R did not significantly alter the rates or processivities of DNA synthesis in the rolling-circle assay. The rates were calculated to be 662 ± 72 bp/s (no RarA), 658 ± 33 bp/s (300 nM RarA added), and 685 ± 70 bp/s (300 nM RarA K63R added) (**Figure 2-**

10A). The processivities were calculated to be 68 ± 7 kb (no RarA), 70 ± 6 kb (300 nM RarA added), and 68 ± 8 kb (300 nM RarA K63R added) (**Figure 2-10A**).

On average, RarA-induced gaps were 2.2 μm long (SEM = 0.5 μm; *n* = 109 molecules) and appeared at an average frequency of once per 78 nm (**Figure 2-8B**; **Figure 2-10B**). Under these experimental conditions, dsDNA has a length of 3 nt/nm (measured by visualization of tethered 20-kb linear dsDNA under the same conditions). In comparison, ssDNA is much more compact. Based on previous measurements of SSB-coated DNA, we estimate that the ssDNA within gaps has a length of approximately 12 nt/nm (*60*). Applying these length conversions, RarA-induced gaps had an average length of 26 knt (SEM = 7 knt) and appeared at a frequency of once every 37 kb (**Figure 2-8B**; **Figure 2-10B**). These values imply that under these conditions leading- and lagging-strand synthesis become uncoupled for very long periods of time.

We next examined the dependence of these parameters on the RarA concentration. When RarA was included in the reaction mixture at 100 nM (tetramer), gaps appeared less frequently (once per 100 kb). However, the lengths of the gaps were unaffected (mean = 1.8 μm, SEM = 0.7 μm; *n* = 37 molecules) (**Figure 2-10B**). At 30 nM RarA, few gaps were observed. Thus the gap frequency is dependent on RarA concentration, whereas the gap length is independent. When RarA was included at 300 nM, but a five-fold higher concentration of β_2 was used, the length of the gaps reduced to 0.95 μm (SEM = 0.19 μm; *n* = 354 molecules) (**Figure 2-10B**). Based on these observations (together with results described below), we attribute the very long gaps to slow restart of Okazaki-fragment synthesis as a result of slow loading of β_2 clamps from solution under these experimental conditions.

We closely examined the relationship between RarA-induced gap formation and the β_2 clamp in these reactions. RarA could act to disengage the Pol III core and its associated β_2 clamp

from the DNA template, or it could act to separate that Pol III core from the β_2 clamp. If RarA disengages both the Pol III core and associated clamp, and both are re-used when lagging strand DNA synthesis re-initiates, no β_2 clamps would be left behind the fork at RarA-mediated gaps. However, if RarA separates Pol III core from its clamp, one would expect that clamps would be left behind at gaps. To distinguish these possibilities, we repeated the rolling-circle assays using fluorescently labeled β_2 clamp. As expected, no gaps were observed in the absence of RarA (**Figure 2-8D**, top). Fluorescently labeled β_2 clamps were only visible at the tips of DNA molecules, corresponding to the position of the replication fork. This indicates under the conditions of the assay (20 nM β_2 clamp is provided in the replication solution), clamps are predominantly recycled by Pol III* during synthesis of each new Okazaki fragment (*64*). Introduction of 300 nM RarA to the reaction led to products containing numerous gaps, as expected. However, multiple β_2 clamps were now visible on each DNA product (**Figure 2-8D**, bottom). This indicates that in the rolling-circle assay RarA action disengages Pol III core from the β_2 clamp, which remains associated with the dsDNA upstream of each gap. Many of the abandoned clamps were seen near gaps. It is possible that all of them were thus associated, since gaps of fewer than 500 nt would not be observed in this experiment.

An important caveat of these experiments is that only gaps on the lagging strand are detected in intact DNA molecules visualized. This is due to the configuration of our assay and rolling-circle replication, where the leading strand is used as a template for continued lagging strand synthesis (**Figure 2-8A**). Thus, any gaps present in the leading strand would result in early termination of rolling-circle replication, which would materialize as truncated products (and a lower processivity) in our assay. However, since no significant change in processivity was seen

upon addition of RarA, we conclude that these gaps are forming on the lagging strand during DNA synthesis (**Figure 2-10A**).

We note that in all single-molecule experiments, no DNA lesions have been purposefully introduced into these DNA substrates. RarA protein is acting to create gaps behind the replication fork. The concentration of RarA employed, and perhaps the absence of other factors that may affect RarA function in the cell, may amplify an activity that normally addresses replisome stress with more targeted precision *in vivo*.

Discussion

The work described in this chapter provides four new insights into the highly conserved yet enigmatic RarA protein family. First, *E. coli* RarA is a DNA-dependent ATPase specifically stimulated by double-stranded (ds) DNA ends and single-stranded DNA gaps. The effects of DNA ends on RarA ATPase activity are documented in Figures 2-1 – 2-4. This is likely to be a general feature of this protein family. Previous reports of RarA and yeast Mgs1 ATPase stimulation by single-stranded (ss) DNA likely reflects an interaction with the ends of secondary structure that is prevalent in the M13-based DNA substrates employed in all cases (*22*, *37*, *40*), although we did not directly test Mgs1. ATPase stimulate by dsDNA ends has been demonstrated for WRNIP1 (*47*).

Second, we characterized the RarA-DNA complex in various nucleotide-bound states. These experiments demonstrated that ATPγS-bound RarA possesses a high affinity for dsDNA, while *apo* RarA possesses a substantial affinity for ssDNA. These results begin to define an ATP hydrolytic cycle possessing multiple intermediates with different DNA affinities.

Third, we showed that RarA separates the strands at DNA duplex termini in the presence of ATPγS or ATP to create ssDNA flaps. In addition to a direct demonstration of this activity in Figure 2-7, the effects of AT base pair content at the DNA end documented in Figure 2-6 also contributes to this conclusion. The inability for an ATPase-deficient RarA protein to unwind dsDNA demonstrates the importance of ATP in this process. This helicase-like activity is the first DNA substrate-based activity detected to date for this protein family. We note that RarA is not the only protein, even in *E. coli*, which can produce a fraying or strand separation at the ends of the duplex. For example, the RecBCD helicase can produce this effect at DNA ends, albeit presumably as a prelude to complete DNA unwinding (*65*, *66*).

Finally, we described a gap-creation activity, catalyzed by RarA, that creates ssDNA gaps behind actively replicating DNA polymerases *in vitro*. To our knowledge, this is the first description of such an activity being catalyzed in an active manner *in vitro*. The experiments described in Figures 2-8 – 2-10 suggest RarA accomplishes this activity by outcompeting Pol III core for binding the β_2 processivity clamp during lagging strand DNA synthesis. This forces Pol III cores to disengage from the DNA template and search for an additional β_2 binding partner on a downstream RNA primer. The result – a ssDNA gap – is generated behind the replication fork. We showed that the number of these gaps increases with RarA concentration, suggesting an active mechanism for Pol III disengagement from the β_2 clamps. Surprisingly, the average gap size did not change when RarA concentration increased. This is most likely due to a constant β_2 clamp concentration. We hypothesize that as β_2 clamp concentration increases, more clamps would bind to downstream RNA primers, creating additional binding partners for Pol III cores upon upstream disengagement. Thus, the overall gap size would decrease, since the probability of Pol III core finding a closer primed β_2 clamp is higher when β_2 clamp concentration increases.

The mechanism by which RarA disengages Pol III core from its associated β_2 during lagging-strand DNA synthesis is most likely through a protein-protein interaction between RarA and the β_2 clamp (**Figure 2-11**). As mentioned previously, RarA shares both amino acid and structural homology with clamp-loader proteins from all organisms (*19*, *22*). In addition, RarA is present at the replication fork and has been shown to be responsible for the instability of at least two replisome protein subunits *in vivo* (*19*, *21*, *67*). All of these observations are consistent with the gap creation activity we have described here. Although more work is needed to explore mechanistic details, the simplest model for this activity would involve direct disengagement of Pol III cores from their associated β_2 clamp. Consistent with this model, β_2 clamps are left behind in

gaps created by RarA and replisomes continue following disengagement with the β_2 clamp (**Figure 2-10D**). Thus, replisomes are not completely disassembled at these gaps. Indeed, the N-terminus of RarA contains a hexapeptide motif (NH_2-S**NL**S**L**D**F**) that is consistent with other proteins that bind the β_2 clamp in eubacteria (i.e. **QL**S**L**P**L** in the *E. coli* Had protein; bold characters indicate residues that come in intimate contact with a hydrophobic pocket on the β_2 clamp) (*68*, *69*). Further experiments should be undertaken to see if RarA-mediated gap formation is dependent on this hexapeptide sequence (and thus β_2 clamp binding).

The DNA substrates used in the single-molecule rolling-circle DNA replication assays did not contain any DNA lesions. Thus, from a purely *in vitro* standpoint, a connection between DNA replication and repair bridged by this RarA function is not readily apparent. However, one can hypothesize the usefulness of creating ssDNA gaps behind the replisome when faced with DNA damage. This mechanism would provide the replisome with an active mechanism of lesion skipping, a process that has been observed in *in vitro* DNA replication assays conducted by Marians and colleagues (*70*, *71*). Marians observed that DNA polymerase III is able to bypass a site-specific DNA lesion on the leading strand by disengaging from the DNA template and re-engaging downstream of the DNA block. *In vitro*, this process occurred on the order of minutes, and was stimulated by DnaG primase, suggesting the requirement of a downstream RNA primer for this activity to occur. It is quite possible that RarA catalyzes this process *in vivo*, since halting DNA replication for minutes at a time in *E. coli* would not be beneficial for overall cell survival. RarA-mediated gap formation could thus serve as a mechanism to allow for the replisome to bypass DNA lesions and rely on post-replication repair pathways like homologous recombination and translesion DNA synthesis to repair the damaged DNA. This would be advantageous to maintain the high speed of DNA replication required by all cells efficiently divide and propagate.

Indeed, gaps left behind the replisome are ideal substrates for both RecFOR-mediated homologous recombination or translesion DNA synthesis (TLS) (*72*). Evidence that RarA creates substrates for both of these processes *in vivo* will be presented in Chapter 3.

Despite the discovery of two novel activities for the highly conserved RarA protein family, many questions remain. With the data presented in this chapter, it is still unclear if there is a connection between the DNA flap formation and gap creation activities of RarA. One major connection between both activities is that they both require a competent RarA ATPase activity. This observation, combined with the ATPase and DNA binding studies presented here, suggest that DNA binding and possibly DNA flap formation are required for gap formation. Further investigation using the single-molecule rolling-circle assay is required to answer these questions. Finally, a direct demonstration of this activity *in vivo* still remains to be seen. While DNA gaps have appeared behind the replisome in UV-irradiated cells, it is unclear whether this is in any part due to the action of RarA. Building upon the data we present in chapter 3 for the *in vivo* consequences of this gap creation activity is still required in future work.

Experimental Procedures

Protein expression and purification

E. coli replication proteins were produced as described previously: the β_2 sliding clamp (*73*), SSB (*74*), the $DnaB_6(DnaC)_6$ helicase-loader complex (*75*), DnaG primase (*76*), the Pol III $\tau_3\delta\delta'\chi\psi$ clamp loader (*59*), Pol III $\alpha\varepsilon\theta$ core (*59*), and wild-type, K63R mutant, and R156A mutant RarA protein were purified as described previously (*22*). All proteins were carefully tested for endo- and exonuclease contamination using gel-based DNA degradation assays utilizing supercoiled and linear dsDNAs and circular ssDNAs. No contaminating endo- or exonucleases were detected. Aliquots of purified proteins were thawed fresh from -80°C stocks prior to each experiment. RarA protein concentration was determined using the native extinction coefficient $\varepsilon = 5.44 \times 10^4\ M^{-1}\ cm^{-1}$.

DNA substrates

All DNA substrates used in this study were purchased from Integrated DNA Technologies (Coralville, IA) (**Table 2-1**). DNA substrates were annealed by heating equimolar amounts of each oligonucleotide (or a single oligonucleotide) in annealing buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 50 mM NaCl) to 95°C for five minutes and cooling to room temperature over 2 h. DNA concentrations were calculated using A_{260} values measured using a Cary 300 UV-Vis spectrophotometer and extinction coefficients provided by IDT. All DNA substrates were stored at 4°C prior to usage in experiments. DNA concentrations are in all cases reported both in terms of total nucleotides and total molecules.

All fluorescently labeled DNA substrates were HPLC-purified by IDT. Molecular beacon substrates (labeled with 6-carboxyfluorescein (FAM)) used in steady state fluorescence assays

were constructed using a 5' 6-FAM fluorophore and a 3' 3-Dabcyl quencher. 2-Aminopurine (2-AP) labeled substrates were constructed by substituting a single 2-AP base for adenine. All sequences of the DNA oligonucleotides used in this study are shown in Table 2-1.

ATPase activity assay

A coupled spectrophotometric assay was used to measure ATPase activity by RarA protein as described previously (*77*, *78*). All assays were carried out using a Varian Cary 300 dual-beam spectrophotometer equipped with a temperature-controlled 12-cell changer. The cell path length is 1 cm, and the bandwidth was 2 nm. All reactions were carried out at 37°C in 1x reaction buffer (25 mM Tris-acetate pH 7.5, 1 mM DTT, 3 mM potassium glutamate, 10 mM $Mg(OAc)_2$, 5% (w/v) glycerol), an ATP regeneration system (10 U/mL pyruvate kinase, 3 mM phosphoenolpyruvate), a coupled detection system (10 U/mL lactate dehydrogenase, 3 mM NADH), 3 mM ATP, and the indicated amounts of RarA protein and DNA substrate. In some experiments, the potassium glutamate concentration was increased to 100 mM as noted. V_{max} and apparent K_m values were calculated by constructing Michaelis-Menten plots in Graphpad Prism software.

Fluorescence polarization assay

RarA protein was incubated with 1 nM oANP031_FAM or annealed oANP031_FAM and oANP032, or 5 nM fluorescein-labeled *E. coli* SSB C-terminal tail peptide (Fluor-Trp-Met-Asp-Phe-Asp-Asp-Ile-Pro-Phe) in 1x reaction buffer supplemented with indicated amounts of nucleotide cofactor at room temperature for 30 min. Fluorescence polarization was measured at 25°C using a Beacon 2000 fluorescence polarization system. The polarization values of experiments reactions were background corrected by subtracting the average polarization value of reactions containing only labeled DNA (for ssDNA, 36 mP; for dsDNA, 62.5 mP) or labeled

peptide (44 mP) from experimental polarization values. For all binding experiments, data were fit to a simple one-site specific interaction model and apparent K_d values were determined using Graphpad Prism software.

Steady-state fluorescence assays

Molecular beacon DNA substrate (50 nM) was incubated in 1x reaction buffer supplemented with 3 mM ATP and an ATP regeneration system (10 U/mL pyruvate kinase, 3 mM phosphoenolpyruvate) at 37°C for 5 min. Indicated amounts of wild-type or R156A RarA protein were then added to the reaction mixture. Fluorescence intensities were measured by exciting the reaction mixture at 494 nm and measuring emitted light at 521 nm in a QuantaMaster Model C-60/2000 spectrofluorimeter (Photon Technologies International). Measurements were taken at 37°C for a total of 500 s. Fluorescence intensity values were normalized by subtracting values from no protein control reactions. All experiments were completed at least three times and the fluorescence values average for each condition.

Stopped flow spectrofluorimetry

Stopped flow fluorescence experiments were conducted using a Kintek SF-300X stopped flow instrument. 2-aminopurine substituted DNA substrates were pre-incubated with indicated amounts of RarA protein for at least 10 min prior to loading into mixing syringes. Syringe A was loaded with DNA and RarA protein or no protein in stopped flow buffer (50 mM Tris-acetate pH 7.5, 6 mM potassium glutamate, 20 mM $Mg(OAc)_2$, 10% (w/v) glycerol). Syringe B was loaded with indicated nucleotide cofactor (or no nucleotide) in stopped flow buffer. Mixing in the flow cell resulted in a final concentration of DNA substrate (100 nM), RarA protein (as indicated), and NTP (500 μM). 2-aminopurine was excited at 319 nm and emitted fluorescence was measured using a 341 nm long pass filter (Edmund Optics). Measurements were taken at 37°C for a total of

30 s. Fluorescence intensity values were background corrected by subtracting no protein control reactions from experimental values. All experiments were repeated at least three times and the fluorescence values were averaged for each condition.

Labeling of β_2 with AF647

β_2 labeling reactions were carried out at a protein concentration of 140 μM (as dimer) at room temperature in 500 μL of labeling buffer (50 mM Tris-HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 20% (v/v) glycerol). A 4-fold molar excess of Alexa Fluor 647 carboxylic acid succinimidyl ester (Invitrogen) dissolved in anhydrous DMSO was added to the protein solution and allowed to react for 1.5 h in the dark, yielding Fraction I. Fraction I was centrifuged (21,000 x *g*; 15 min) at 6°C and the supernatant carefully removed to yield Fraction II. Fraction II was applied at 1 mL/min flow to a column (1.5 x 10 cm) of Superdex G-25 resin (GE-Healthcare) equilibrated with gel filtration buffer (50 mM Tris-HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 5% (v/v) glycerol) to remove unreacted fluorophores. Fractions containing the labeled β_2 were pooled and dialyzed into storage buffer (50 mM Tris-HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 20% (v/v) glycerol). The degree of labeling was determined by UV/vis spectroscopy to be ~1 fluorophore per β dimer.

In vitro single-molecule rolling-circle DNA replication assay

Microfluidic flow cells were prepared as described (*79*). Briefly, a PDMS flow chamber was placed on top of a PEG-biotin-functionalized microscope coverslip. To help prevent non-specific interactions of proteins and DNA with the surface, the chamber was blocked with blocking agent (NEB, Ipwich, MA). The chamber was placed on an inverted microscope (Nikon Eclipse Ti-E) with a CFI Apo TIRF 100x oil-immersion TIRF objective (NA 1.49, Nikon, Japan) and connected to a syringe pump (Adelab Scientific, Australia) for flow of buffer. Reactions were

carried out at 31°C, maintained by an electrically heated chamber (Okolab, Burlingame, CA). Double-stranded DNA was visualized in real time by staining it with 150 nM SYTOX orange (Invitrogen) excited by a 568 nm laser (Coherent, Santa Clara, CA; Sapphire 568-200 CW) at 150 mW/cm^2. The red labeled β_2 was excited at 700 mW/cm^2 with a 647 nm (Coherent, Obis 647-100 CW) laser. For dual-color imaging the signals were separated via dichroic mirrors and appropriate filter sets (Chroma, Bellows Falls, VT). Imaging was done with an EMCCD camera (Photometrics, Tucson, AZ; Evolve 512 Delta).

Conditions for the pre-assembly replication reactions were adapted from published methods (*64*, *80*, *81*). Solution 1 was prepared as 30 nM $DnaB_6(DnaC)_6$, 1.5 nM biotinylated circular 2 kb dsDNA substrate and 1 mM ATP in replication buffer. This was incubated at 37°C for 3 min. Solution 2 contained 60 μM dCTP and dGTP, 3.3 nM Pol III* (assembled in situ by incubating $\tau_3\delta\delta'\chi\psi$ (410 nM) and Pol III cores $\alpha\varepsilon\theta$ (1.2 μM) in replication buffer at 37°C for 90 s), and 74 nM β_2 in replication buffer (without dATP and dTTP). Solution 2 was added to an equal volume of solution 1 and incubated for 6 min at 37°C. This was then loaded onto the flow cell at 100 μL/min for 1 min and then 10 μL/min for 10 min. The flow cell was washed with replication buffer containing 60 μM dCTP and dGTP. An imaging buffer was made with 1 mM UV-aged Trolox, 0.8% (w/v) glucose, 0.12 mg/mL glucose oxidase, and 0.012 mg/mL catalase (to increase the lifetime of the fluorophores and reduce blinking), 1 mM ATP, 250 μM CTP, GTP, and UTP, and 50 μM dCTP, dGTP, dATP, and dTTP in replication buffer. Replication was finally initiated by flowing in the imaging buffer containing 20 nM β_2, 75 nM DnaG, 250 nM SSB_4, and RarA when indicated, at 10 μL/min. All in vitro single-molecule experiments were performed at least four times. The analysis was done with ImageJ using in-house built plugins. The rate of replication

of a single molecule was obtained from its trajectory and calculated for each segment that had a constant slope.

Name	Sequence
oANP031	5'--ACACAGAAGGAGACGG--3'
oANP031_F	5'--/56-FAM/ACACAGAAGGAGACGG--3'
oANP032	5'--CCGTCTCCTTCTGTGT--3'
THS7	5'--TTGGAGTTTGCTTCCGGTAACGGAAGCAAACTCCAA--3'
THS7_FAM_Quencher	5'--/56-FAM/TTGGAGTTTGCTTCCGGTAACGGAAGCAAACTCCAA/3Dab/--3'
THS7 18nt 3' OH	5'--TTGGAGTTTGCTTCCGGTAACGGAAGCAAACTCCAATTTTTTTTTTTTTTTTT--3'
THS7 18nt 5' OH	5'--TTTTTTTTTTTTTTTTTTTGGAGTTTGCTTCCGGTAACGGAAGCAAACTCCAA--3'
THS9 0nt gap P	5'--/5Phos/GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS9 3nt gap	5'--GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS9 8nt gap	5'--GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTTTTTTCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS9 13nt gap	5'--GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTTTTTTTTTTTCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS9 18nt gap	5'--GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTTTTTTTTTTTTTTTCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS9 18nt gap P	5'--/5Phos/GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTTTTTTTTTTTTTTTCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS9 18nt gap dd	5'--GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTTTTTTTTTTTTTTTTGGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCC/3ddC/--3'
THS9 23nt gap	5'--GCGGAGTTACGTAACGGTAACGTTACGTAACTCCGCTTTTTTTTTTTTTTTTTTTTTCGGAAGCATCGGCTTAGTAATAAGCCGATGCTTCCG--3'
THS19	5'--TTGGAGTTTGCTTCTCAGCCGGTAACGGCTGAGAAGCAAACTCC/i2AmPr/A--3'
THS27	5'--CCGGCGCCGGCGGCCGGTAACGGCCGCCGGCGCCGG--3'
THS28	5'--AATTATAATTATTAATGTAAATTAATAATTATAATT--3'
THS29	5'--ACGGCGCCGGCGGCCGGTAACGGCCGCCGGCGCCGT--3'
THS30	5'--AAGGCGCCGGCGGCCGGTAACGGCCGCCGGCGCCTT--3'
THS31	5'--AATGCGCCGGCGGCCGGTAACGGCCGCCGGCGCATT--3'
THS32	5'--AATTCGCCGGCGGCCGGTAACGGCCGCCGGCGAATT--3'

Table 2-1: A list of oligonucleotides utilized in this chapter.

A list of oligonucleotides used to make DNA substrates in this chapter are shown above.

Figures

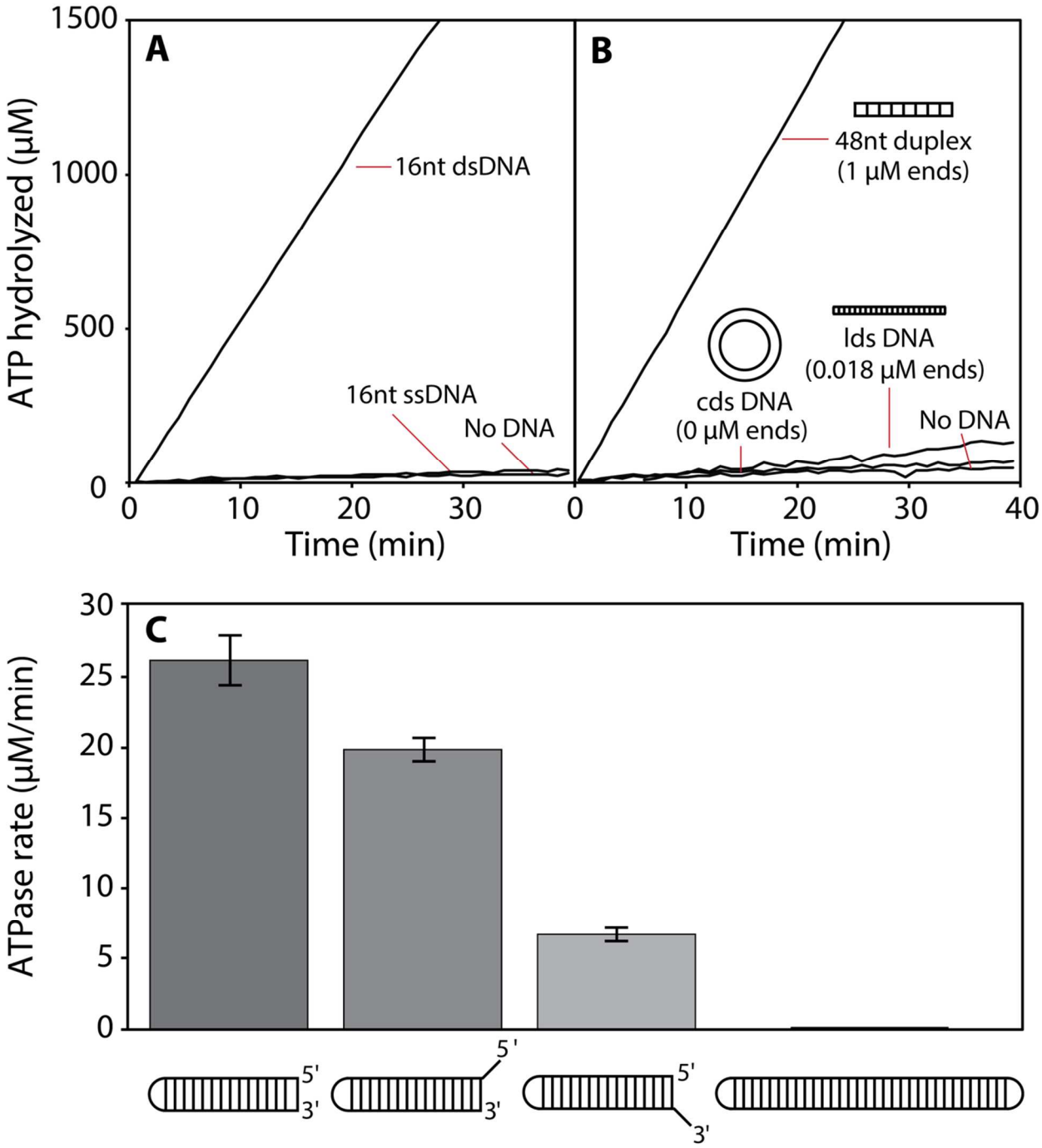

A
B
C
ATP hydrolyzed (μM)
Time (min)
16nt dsDNA
16nt ssDNA
No DNA
48nt duplex
(1 μM ends)
lds DNA
(0.018 μM ends)
cds DNA
(0 μM ends)
ATPase rate (μM/min)
5'
3'

Figure 2-1: Double-stranded DNA ends stimulate RarA ATPase activity

(A) The ATPase activity of RarA (0.5 μM) was measured in the presence of short single-stranded or double-stranded DNAs (0.5 μM total molecules; 8 or 16 μM in nucleotides, respectively). (**B**) The ATPase activity of RarA (0.5 μM) was measured in the presence of various DNA substrates (48 μM nucleotides) including a 48-base pair duplex, M13mp9 linear double-stranded DNA (ldsDNA), or pUC19 circular double-stranded DNA (cdsDNA). Each DNA was included at indicated concentrations of dsDNA ends. (**C**) The rate of ATP hydrolysis mediated by RarA (0.5 μM) was measured in the presence of various duplex substrates (THS7; THS7 5' OH; THS 18 nt 3' OH) containing different DNA end structures (0.1 μM molecules; 3.6-5.4 μM total nucleotides). Error bars represent one standard deviation from the mean.

* This work was conducted by Asher Page, a former graduate student in our laboratory.

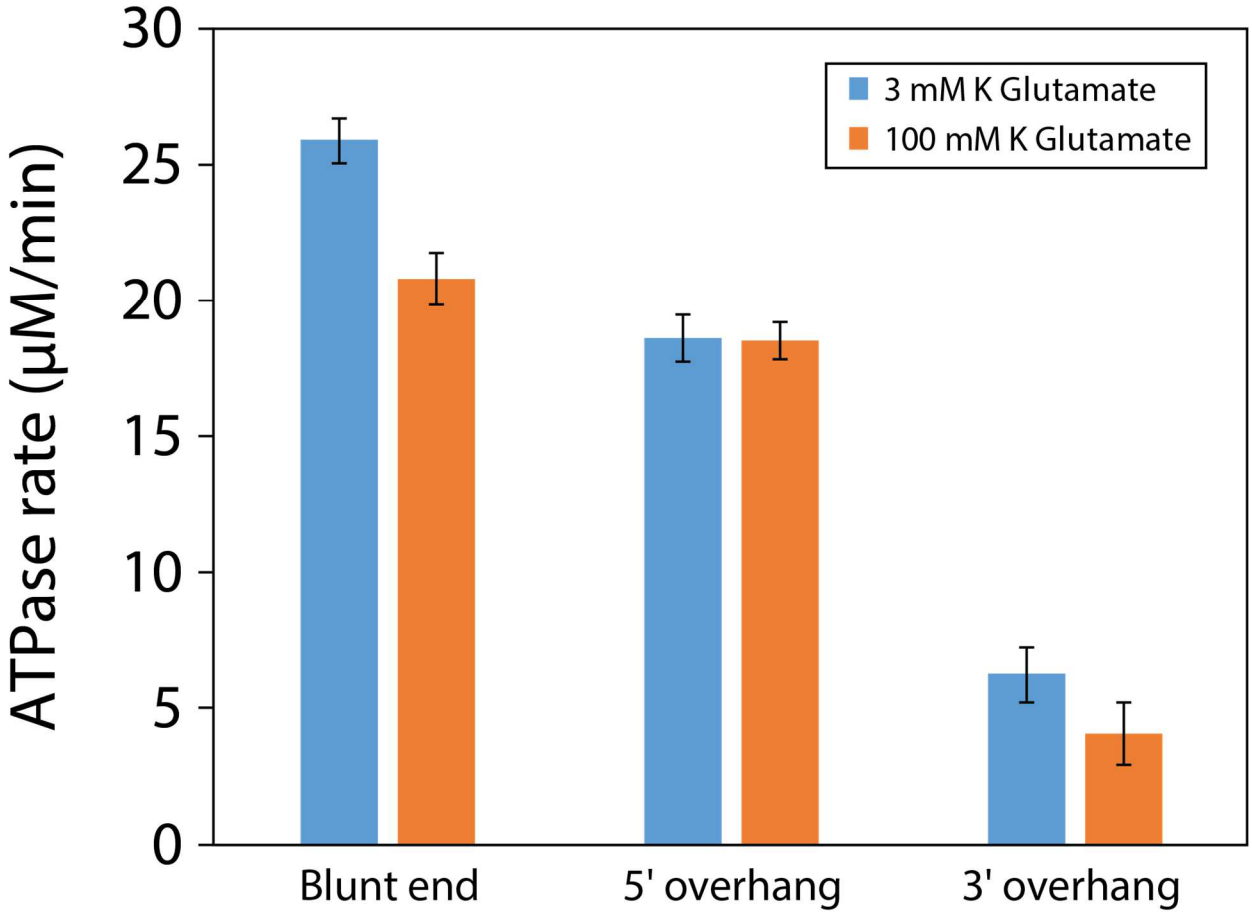
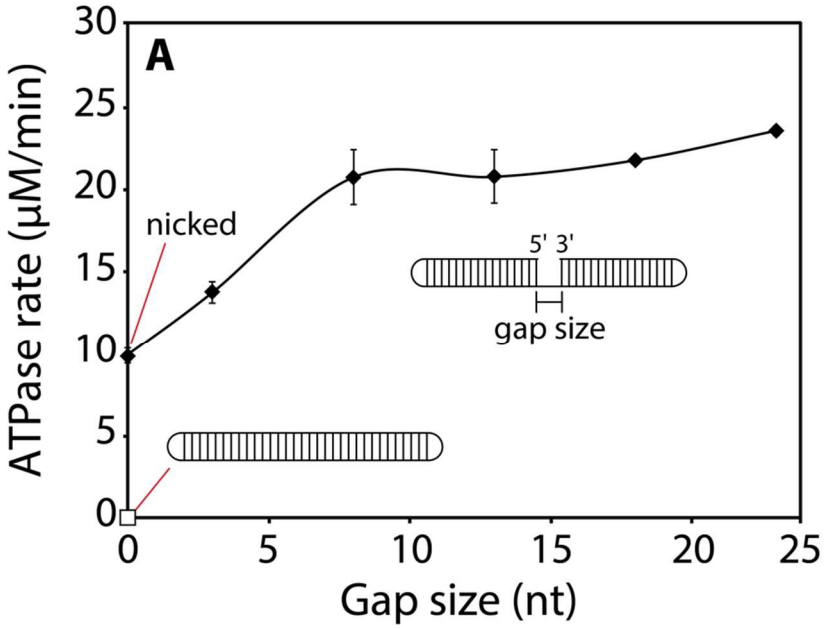
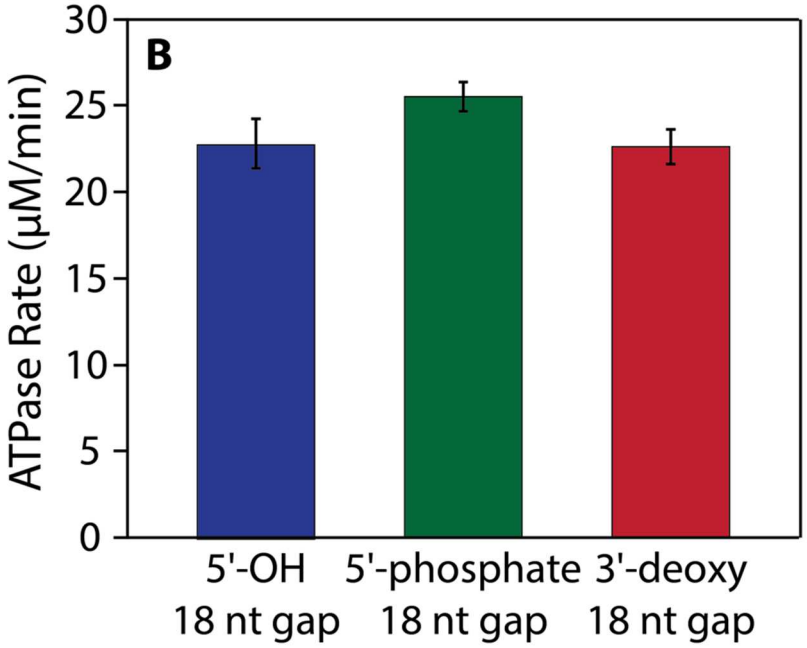

ATPase rate (μM/min)
30
25
20
15
10
5
0
3 mM K Glutamate
100 mM K Glutamate
Blunt end
5' overhang
3' overhang

Figure 2-2: RarA ATPase activity is not affected by higher ionic strength buffers.

RarA protein (500 nM) was incubated with blunt end (THS7), 5' overhang (THS7 18nt 5' OH), or 3' overhang (THS7 18nt 3'OH) DNA substrates (50 nM) in 1x reaction buffer (blue bars) or 1x reaction buffer with 100 mM potassium glutamate (orange bars). With the exception of a slight decrease in ATPase rate on a blunt end DNA substrate, RarA ATPase activity and substrate specificity were not significantly affected by additional potassium glutamate in the buffer.


A
ATPase rate (μM/min)
Gap size (nt)
nicked
5' 3'
gap size
0
5
10
15
20
25
30


B
ATPase Rate (μM/min)
5'-OH
18 nt gap
5'-phosphate
18 nt gap
3'-deoxy
18 nt gap
0
5
10
15
20
25
30

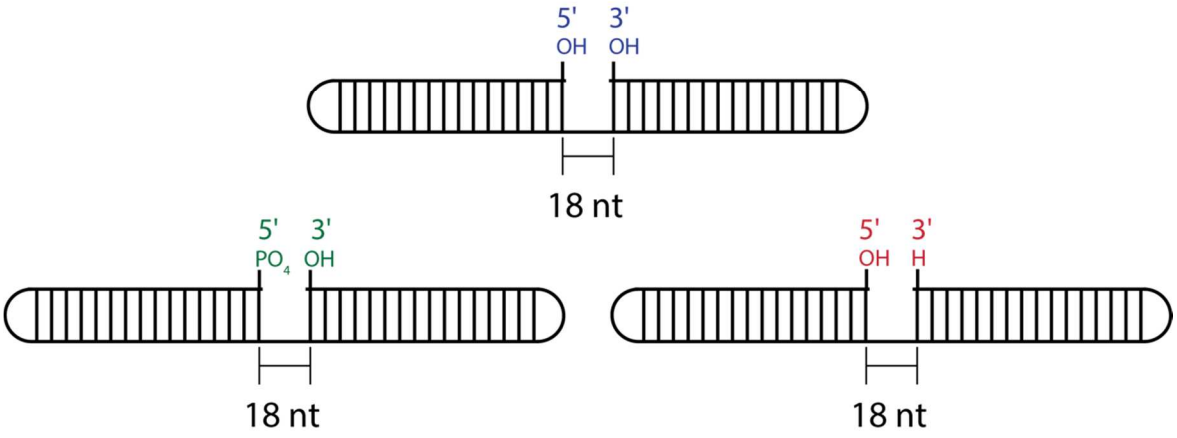

5' OH
3' OH
18 nt
5' PO4
3' OH
18 nt
5' OH
3' H
18 nt

Figure 2-3: Single-stranded DNA gaps stimulate RarA ATPase activity

(**A**) The rate of ATP hydrolysis was measured in the presence of RarA (0.5 μM) and gapped DNA hairpin substrates containing two 16 base pair duplexes separated by poly dT gaps of the indicated size and capped by two four-nucleotide hairpins (0.05 μM molecules, 3.6-4.75 μM total nucleotides). The white square represents a ligated closed-circle substrate while black diamonds represent substrates containing a gap ranging from a single nick to 23 nucleotides. Error bars represent one standard deviation from the mean. (**B**) ATPase rates were measured in the presence of RarA (0.5 μM) and gapped DNA substrates (0.05 μM molecules, 4.5 μM total nucleotides) containing various 5' or 3' modifications. A substrate containing no 5' phosphate (blue), with a 5' phosphate (green), and without a 3' hydroxyl group (red) were included in ATPase reactions with RarA protein. Error bars represent one standard deviation from the mean value.

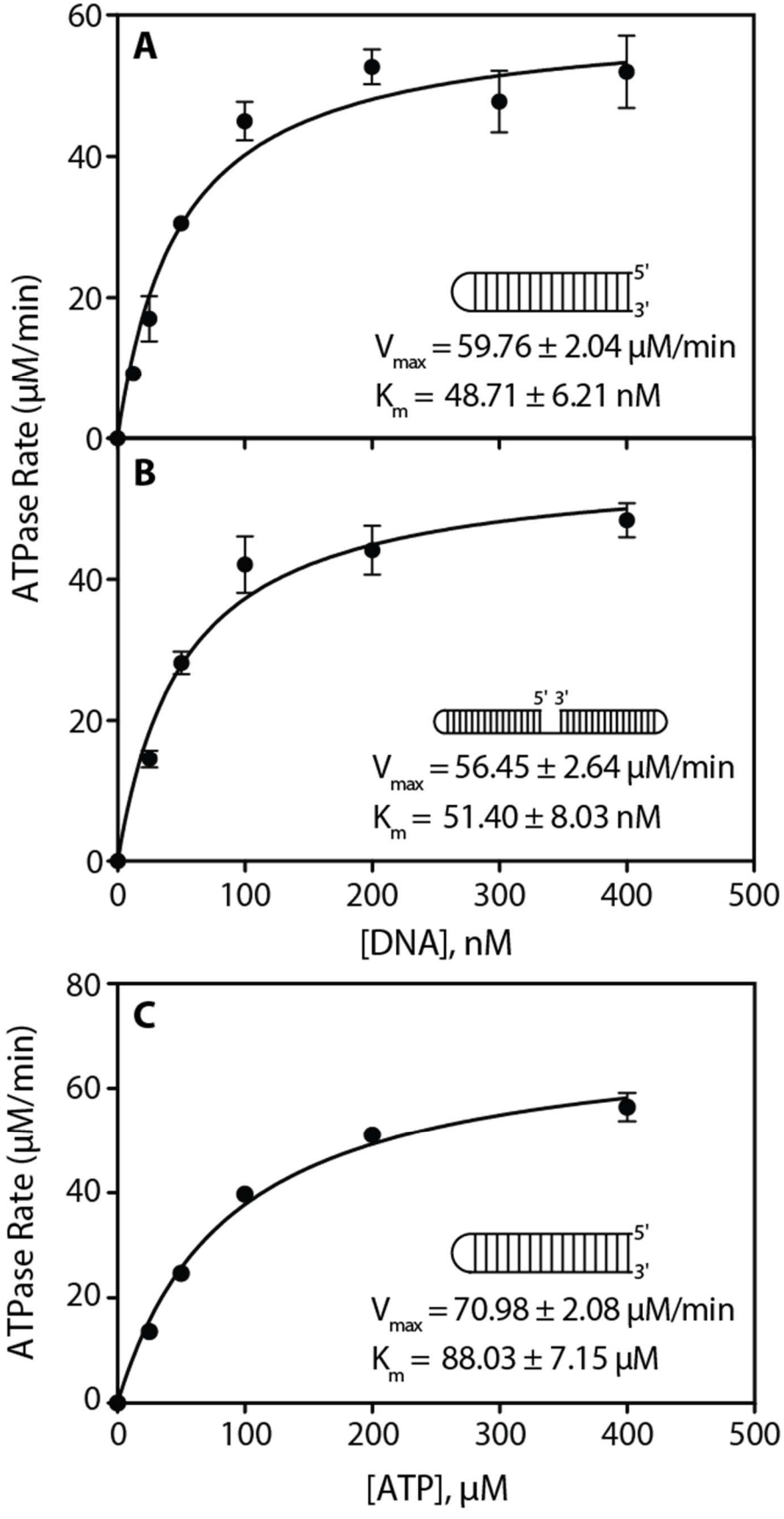
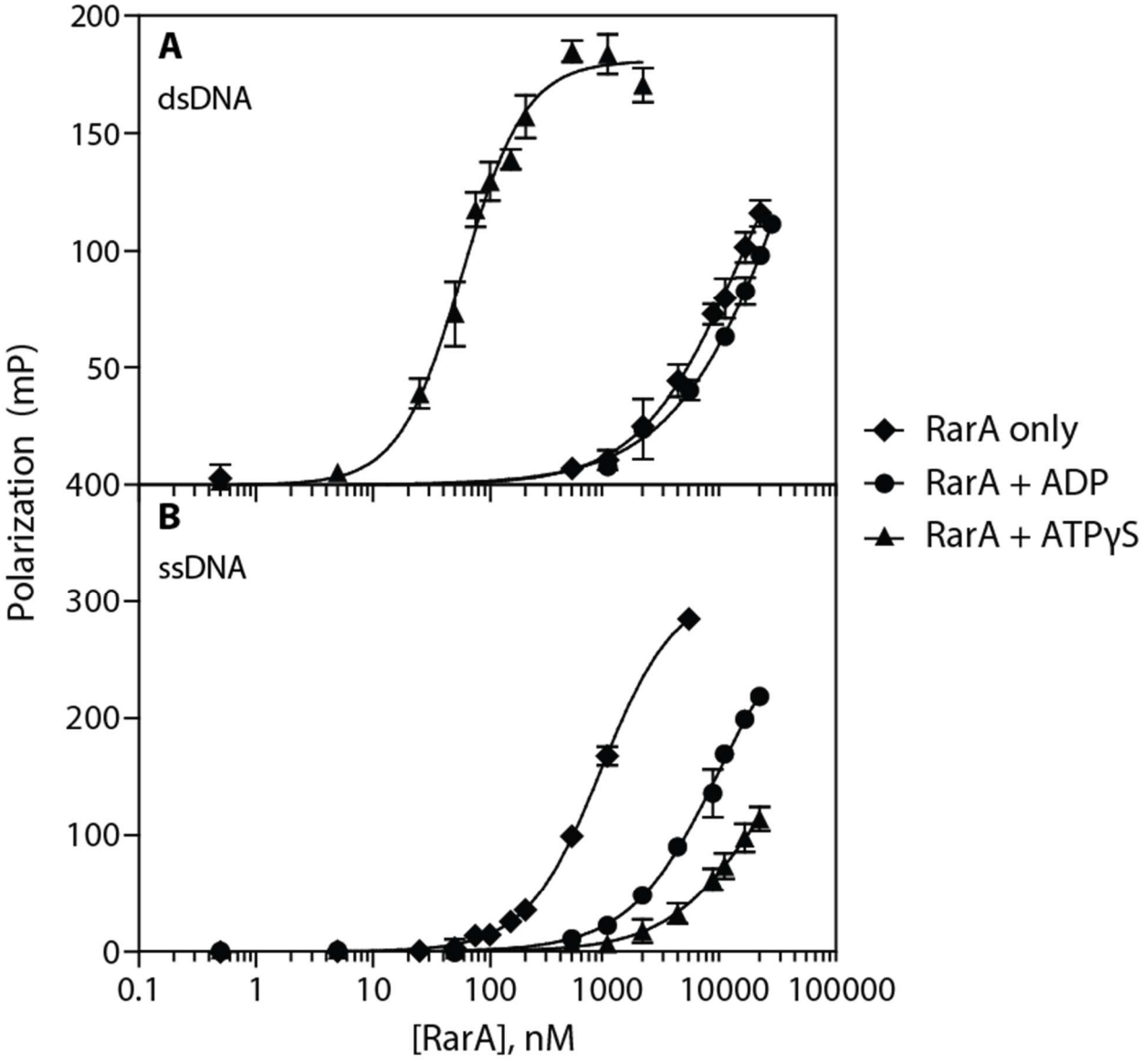

A
B
C
ATPase Rate (μM/min)
[DNA], nM
[ATP], μM
5'
3'
V_{max} = 59.76 ± 2.04 μM/min
K_m = 48.71 ± 6.21 nM
V_{max} = 56.45 ± 2.64 μM/min
K_m = 51.40 ± 8.03 nM
V_{max} = 70.98 ± 2.08 μM/min
K_m = 88.03 ± 7.15 μM

Figure 2-4: Michaelis-Menten kinetic analyses of RarA ATPase activity

The rate of ATP hydrolysis in the presence of different substrate concentrations were measured in the presence of 0.4 μM RarA protein. (**A**) ATP hydrolysis reactions were conducted in the presence of increasing concentrations of a 16 base pair blunt end DNA duplex hairpin substrate. (**B**) ATP hydrolysis reactions were conducted in the presence of increasing concentrations of a 90 nucleotide gapped DNA hairpin substrate containing two 16 base pair duplexes separated by an 18 nucleotide poly dT gap and capped with two four-nucleotide hairpins. (**C**) ATP hydrolysis reactions were conducted in the presence of blunt end duplex DNA substrate (0.4 μM molecules; 14.4 μM nucleotides) and increasing amounts of ATP. Each point is an average of at least three replicate experiments, with error bars representing one standard deviation from the mean ATPase rate.



C

Substrate	Condition	$K_{d,app}$ (nM)
dsDNA	RarA only	> 10000
	RarA + ADP	> 10000
	RarA + ATPγS	58 ± 3
ssDNA	RarA only	890 ± 28
	RarA + ADP	9000 ± 1300
	RarA + ATPγS	> 10000

Figure 2-5: RarA binds double-stranded DNA in the presence of ATPγS and single-stranded DNA in the absence of nucleotide cofactor

RarA binding affinity for dsDNA (**A**) and ssDNA (**B**) was measured using fluorescence polarization in the presence of various nucleotide cofactors. A table of calculated dissociation constants for each reaction condition is provided in (**C**). Error bars represent one standard deviation from the mean polarization value.

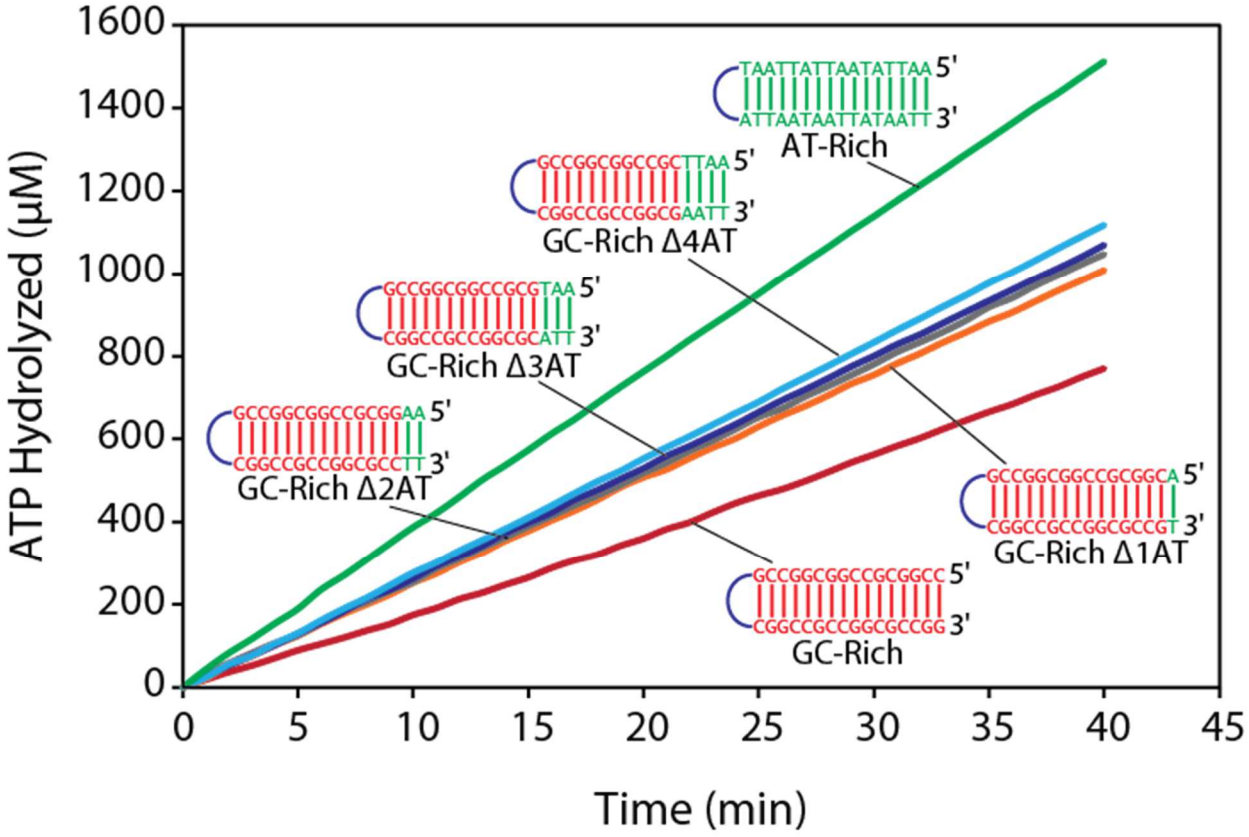

ATP Hydrolyzed (μM)
Time (min)
0
200
400
600
800
1000
1200
1400
1600
0
5
10
15
20
25
30
35
40
45
TAATTATTAATATTAA 5'
ATTAATAATTATAATT 3'
AT-Rich
GCCGGCGGCCGCTTAA 5'
CGGCCGCCGGCGAATT 3'
GC-Rich Δ4AT
GCCGGCGGCCGCGTAA 5'
CGGCCGCCGGCGCATT 3'
GC-Rich Δ3AT
GCCGGCGGCCGCGGAA 5'
CGGCCGCCGGCGCCTT 3'
GC-Rich Δ2AT
GCCGGCGGCCGCGGCA 5'
CGGCCGCCGGCGCCGT 3'
GC-Rich Δ1AT
GCCGGCGGCCGCGGCC 5'
CGGCCGCCGGCGCCGG 3'
GC-Rich

Figure 2-6: RarA ATPase rate is dependent on the GC-content of a DNA duplex

Six different DNA duplex substrates were designed with varying GC-content and included in RarA ATPase reactions. The rate of ATP hydrolysis was measured in the presence of these DNA substrates (0.1 μM molecules; 3.6 μM nucleotides) and RarA (0.4 μM). Reactions were repeated at least three times. Representative data are shown for each DNA substrate.

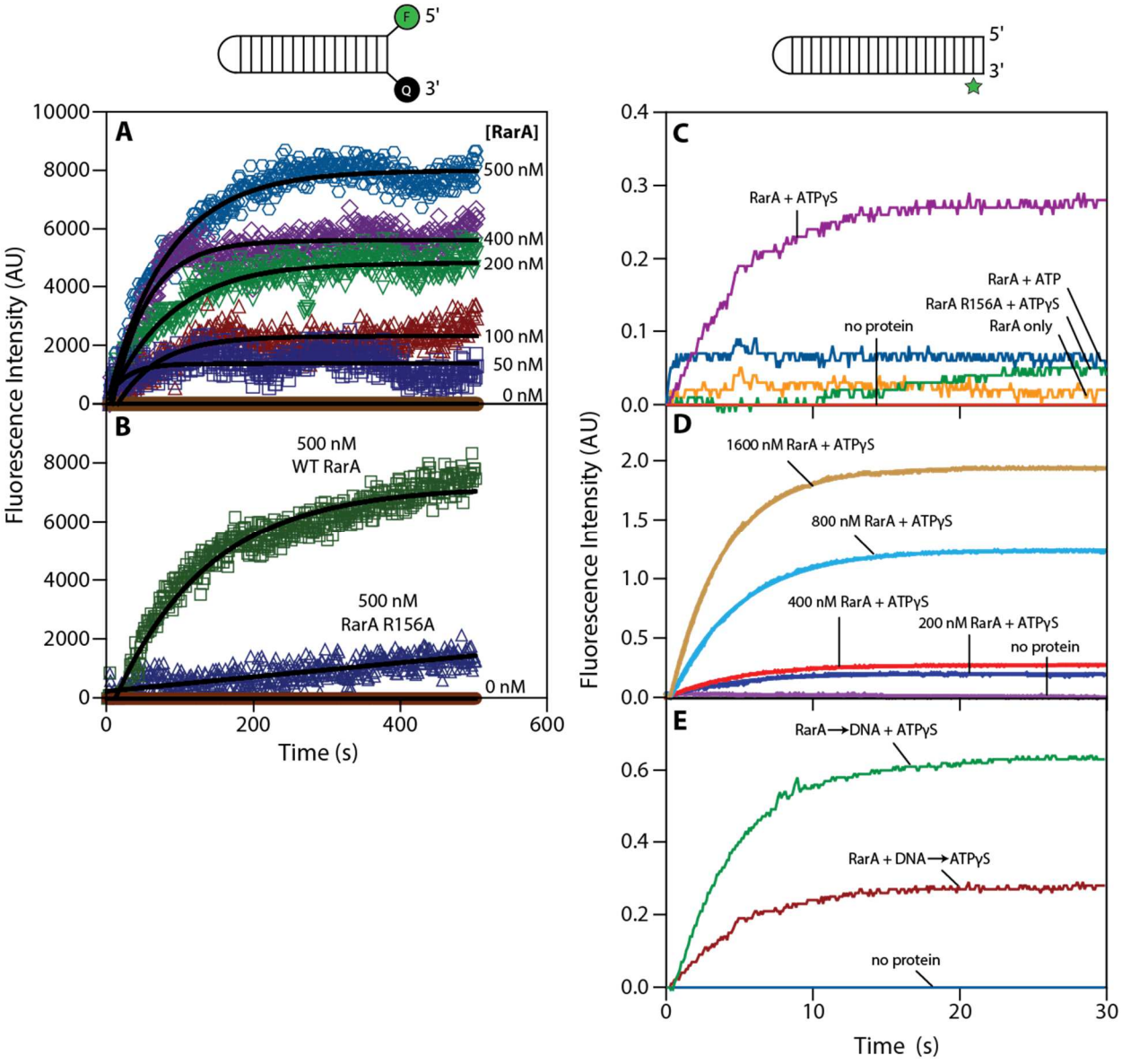
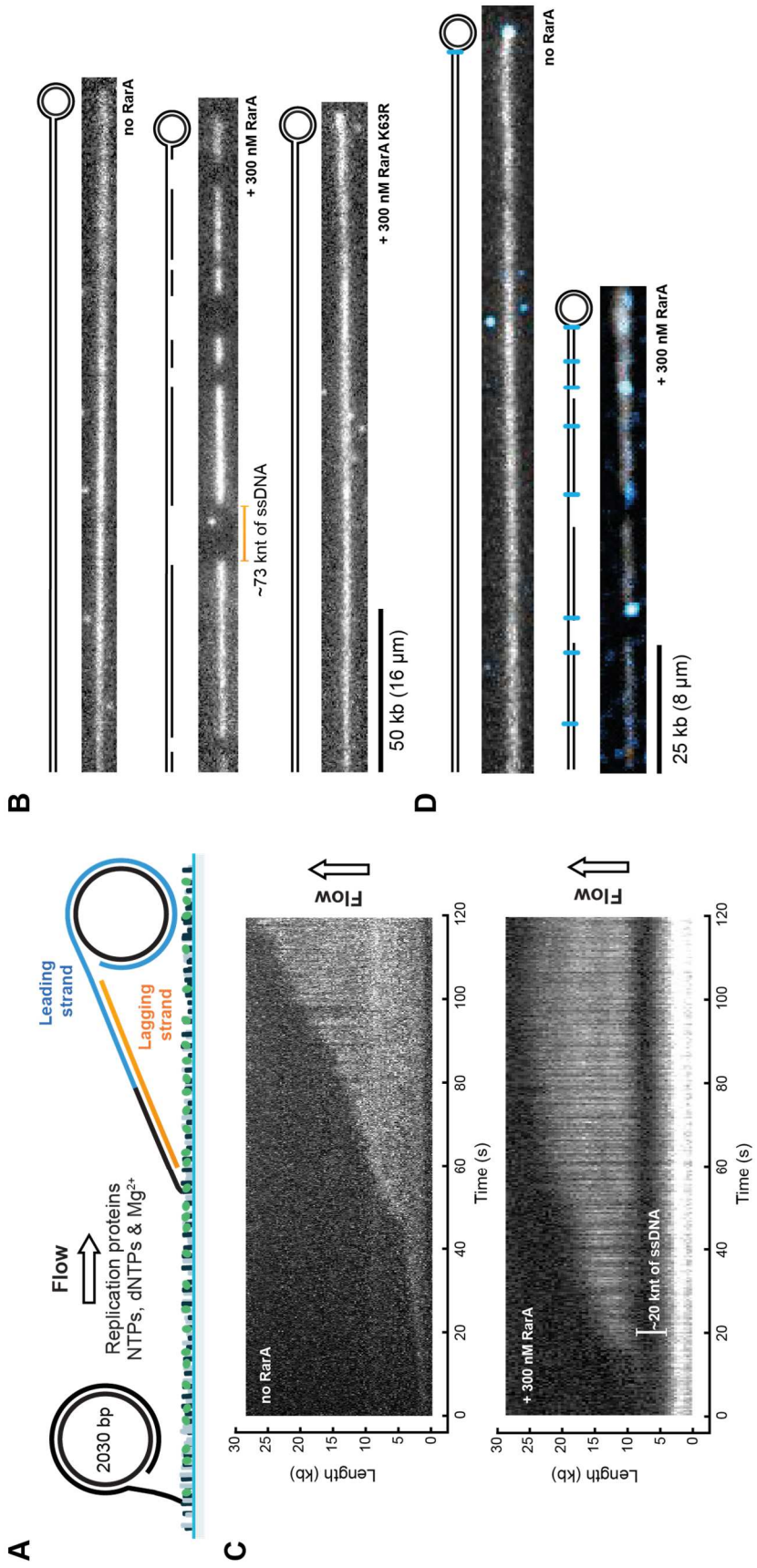

F
5'
Q
3'
A
[RarA]
500 nM
400 nM
200 nM
100 nM
50 nM
0 nM
B
500 nM
WT RarA
500 nM
RarA R156A
0 nM
Fluorescence Intensity (AU)
Time (s)
5'
3'
C
RarA + ATPγS
RarA + ATP
RarA R156A + ATPγS
RarA only
no protein
D
1600 nM RarA + ATPγS
800 nM RarA + ATPγS
400 nM RarA + ATPγS
200 nM RarA + ATPγS
no protein
E
RarA→DNA + ATPγS
RarA + DNA→ATPγS
no protein
Fluorescence Intensity (AU)
Time (s)

Figure 2-7: RarA separates the strands of a double-stranded DNA duplex

(**A**) The fluorescence intensity of a 6-FAM/3-Dabcyl labeled DNA duplex was measured for 500 s after the addition of RarA protein in the presence of ATP. (**B**) The fluorescence intensity of the same DNA substrate was measured over the course of 500 s after addition of either wild-type RarA or RarA R156A variant (0.5 μM) in the presence of ATP. (**C**) RarA (0.4 μM) was pre-incubated with a 2-aminopurine labeled DNA substrate (0.1 μM molecules; 4.6 μM nucleotides). Fluorescence intensity of 2-aminopurine was measured for 30 seconds following addition of indicated nucleotide cofactors. (**D**) Varying concentrations of RarA were pre-incubated with a 2-aminopurine labeled DNA substrate (0.1 μM molecules; 4.6 μM nucleotides). Fluorescence intensity of 2-aminopurine was measured for 30 seconds following addition of ATPγS. (**E**) RarA (0.4 μM) was added to a solution containing 2-aminopurine labeled DNA substrate (0.1 μM molecules; 4.6 μM nucleotides) and ATPγS. In a previous experiment, ATPγS was added to a solution containing RarA (0.4 μM) and 2-aminopurine labeled DNA substrate (0.1 μM molecules; 4.6 μM nucleotides). Fluorescence intensity of 2-aminopurine was measured over the course of 30 seconds in both experiments. All values represent the average of at least three experiments.



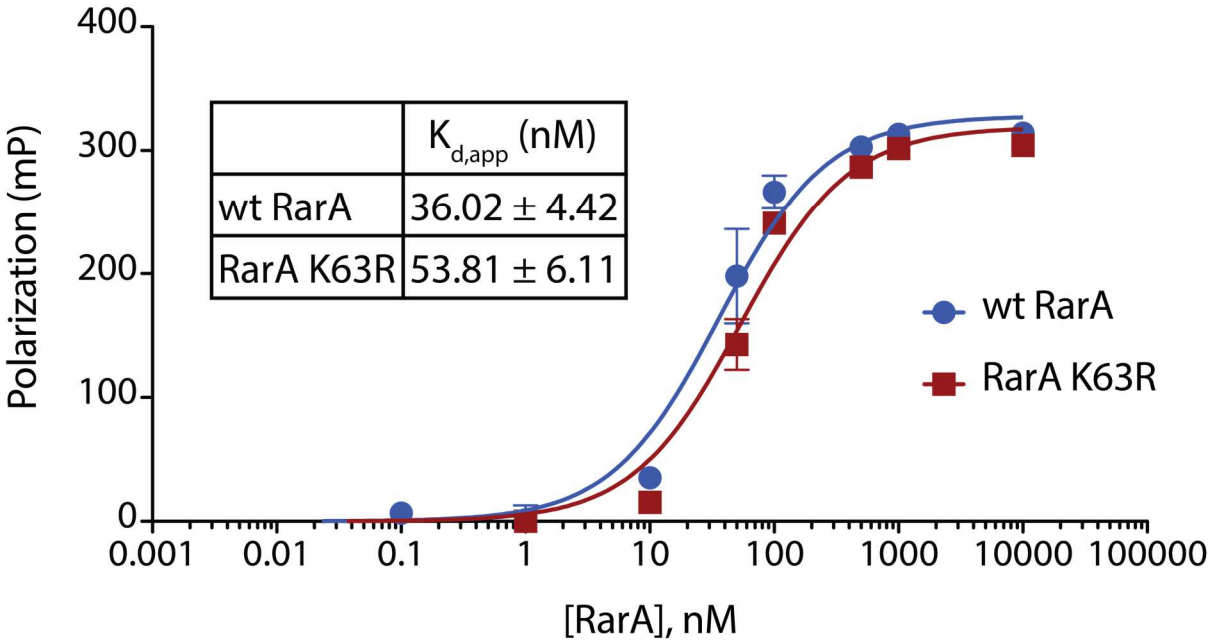
* This work was conducted in collaboration with J. Lewis, L. Spenkelink, S. Jergic, A. Robinson, and A. van Oijen at the University of Wollongong, Australia.

Figure 2-8*: RarA induces the formation of ssDNA gaps on the lagging strand in a reconstituted DNA replication assay *in vitro*.

(**A**) Schematic representation of the experimental design. 5'-biotinylated DNA is coupled to the passivated surface of a microfluidic flow cell through a streptavidin linkage. Addition of the *E. coli* replication proteins and nucleotides initiates DNA synthesis. The DNA products are elongated hydrodynamically by flow, labeled with intercalating DNA stain, and visualized using fluorescence microscopy. (**B**) Examples of individual DNA molecules produced by rolling circle replication in the absence of RarA, or in the presence of 300 nM RarA or its ATPase-dead mutant RarA K63R. The gray scale indicates the fluorescence intensity of stained DNA. (**C**) Kymographs of individual DNA molecules undergoing rolling circle replication in the absence or presence of 300 nM RarA. (**D**) Examples of individual DNA molecules produced by rolling circle replication in the presence and absence of RarA in which the β_2 clamp was fluorescently labeled with Alexa Fluor 647.

* This work was conducted in collaboration with J. Lewis, L. Spenkelink, S. Jergic, A. Robinson, and A. van Oijen at the University of Wollongong, Australia.

A



	$K_{d,app}$ (nM)
wt RarA	36.02 ± 4.42
RarA K63R	53.81 ± 6.11

B

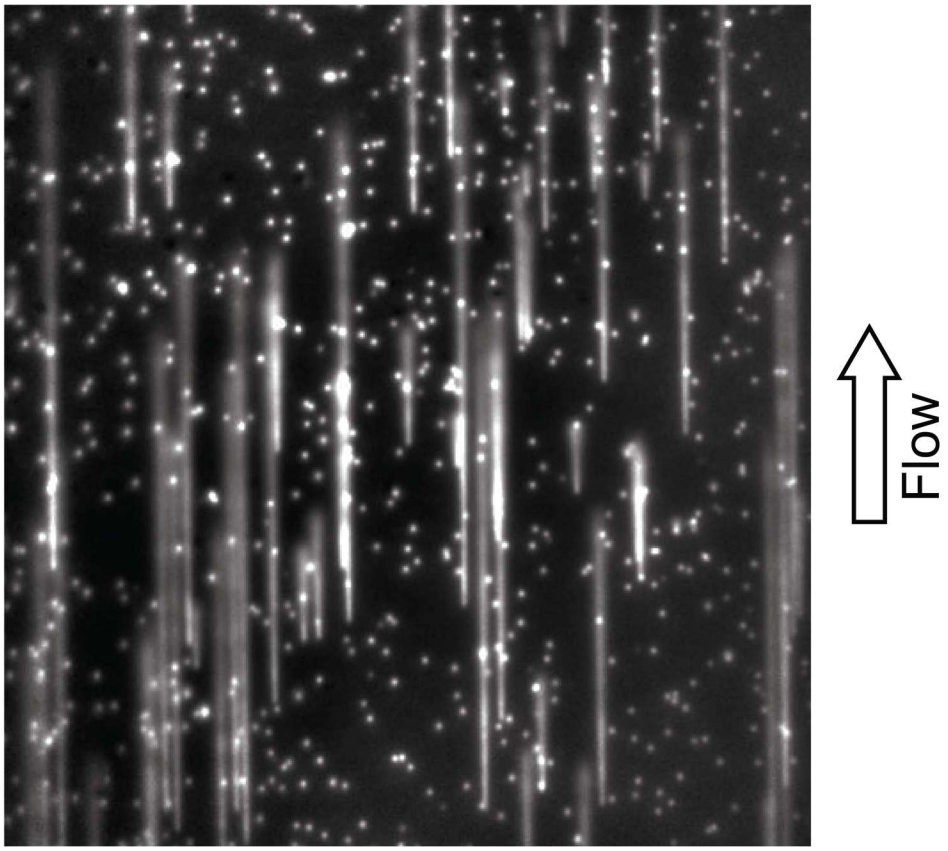
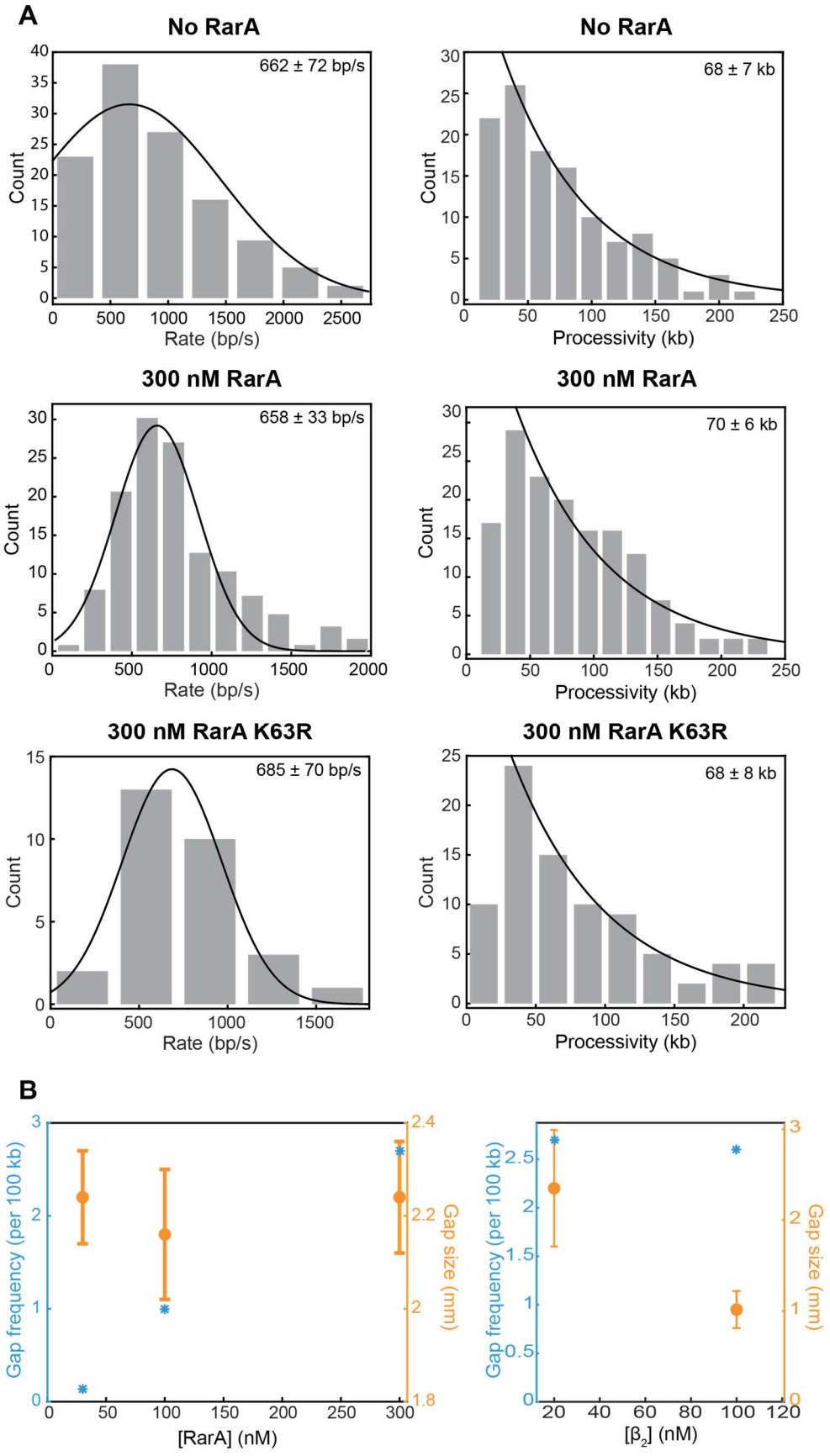


Figure 2-9: Controls showing that gap formation is not due to an increased affinity of RarA for SSB or a nuclease contamination

(**A**) RarA K63R binds the SSB C-terminal tail peptide with a similar affinity to wild type RarA. Wild type RarA or RarA K63R protein were incubated with fluorescently labeled *E. coli* SSB peptide at room temperature. Following incubation for 30 minutes, fluorescence polarization values were measured. Each point represents an average polarization value for a reaction containing the indicated concentration of RarA protein, while error bars represent one standard deviation from the average polarization value. Data were fit to a simple single-site binding curve and apparent dissociation constants ($K_{d,app}$) were calculated. (**B***) Post-replicative addition of RarA does not lead to gap formation. Representative field of view showing DNA when RarA was loaded after the replication reaction had stopped. RarA does not create gaps on the DNA outside the context of the replisome.

* This work was conducted in collaboration with J. Lewis, L. Spenkelink, S. Jergic, A. Robinson, and A. van Oijen at the University of Wollongong, Australia.



* This work was conducted in collaboration with J. Lewis, L. Spenkelink, S. Jergic, A. Robinson, and A. van Oijen at the University of Wollongong, Australia.

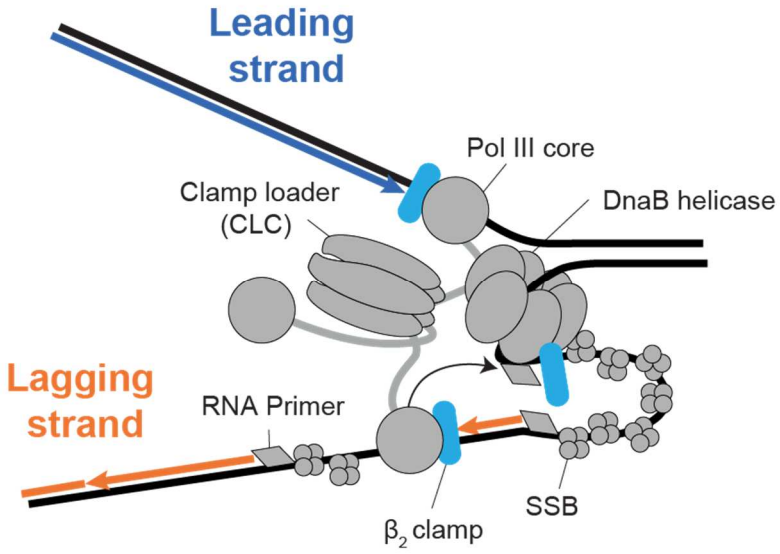
Figure 2-10*: Histograms of rates, processivities, and concentration effects in single-molecule DNA replication assays

(**A**) The rate histograms were fit with a Gaussian distribution (black line) to obtain the mean rate. The rates are 662 ± 72 bp/s without RarA (top, left), 658 ± 33 bp/s with 300 nM RarA (middle, left), and 685 ± 70 bp/s with 300 nM RarA K63R (bottom, left). Processivity distributions were fit with a single-exponential decay function (black line). The processivities are 68 ± 7 kb without RarA (top, right), 70 ± 6 kb in the presence of 300 nM RarA (middle, right), and 68 ± 8 kb with 300 nM RarA K63R (bottom, right). These data show that RarA does not affect the rate of replication and processivity under these conditions. The error bars represent the standard error of the mean.

(**B**) Comparison of gap sizes and gap frequencies for increasing RarA or beta clamp concentrations. Gap frequency (blue) and gap size (orange) as a function of RarA concentration at a constant β_2 concentration (left). The gap frequency increases with increasing concentrations of RarA. The gap size remains constant within this range of RarA concentrations. At a constant concentration of RarA (300 nM) and two different concentrations of β_2 (right), gap size decreases as β_2 concentration increases. Error bars represent the standard error from the mean.

* This work was conducted in collaboration with J. Lewis, L. Spenkelink, S. Jergic, A. Robinson, and A. van Oijen at the University of Wollongong, Australia.

RarA mediated disengagement of the lagging-strand polymerase results in loop release



re-assembly at the now distant 3' terminus is not possible, thus a ssDNA gap is left in the wake of the replisome

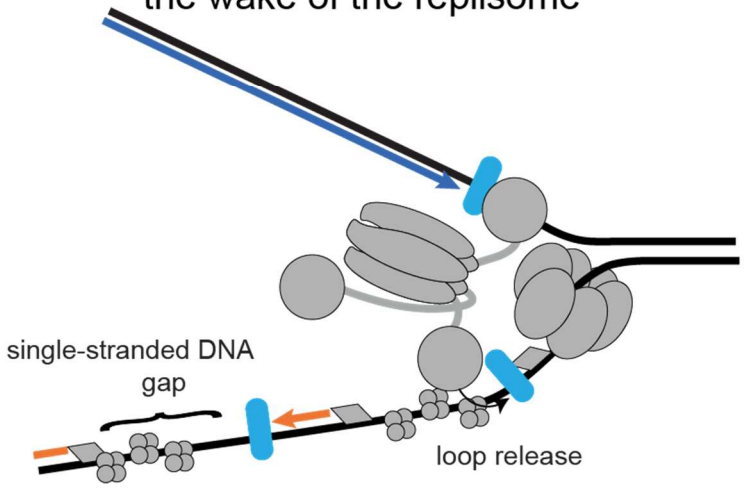


Figure 2-11: A model for RarA-mediated gap formation *in vitro*

A model depicting the hypothesized mechanism of action for RarA-mediated gap formation behind DNA replication forks *in vitro*. During lagging strand DNA synthesis, RarA disengages a Pol III core (circle) from its associated β_2 processivity clamp (blue oval) by binding to β_2 (black arrow). The SSB-bound single-stranded DNA loop downstream of the site of disengagement is released. The disengaged Pol III core is forced to search for a β_2 clamp loaded at an upstream Okazaki fragment. During this process, the released DNA loop becomes a single-stranded DNA gap that can serve as a template for downstream DNA metabolism.

References

1. M. M. Cox *et al.*, The importance of repairing stalled replication forks. *Nature*. **404**, 37–41 (2000).

2. M. M. Cox, Historical overview: searching for replication help in all of the rec places. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8173–8180 (2001).

3. S. C. Kowalczykowski, Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem. Sci.* **25** (2000), pp. 156–165.

4. A. Kuzminov, Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* **63**, 751–813 (1999).

5. A. Kuzminov, DNA replication meets genetic exchange: Chromosomal damage and its repair by homologous recombination. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8461–8468 (2001).

6. B. Michel, Replication fork arrest and DNA recombination. *Trends Biochem. Sci.* **25** (2000), pp. 173–178.

7. B. Michel, H. Boubakri, Z. Baharoglu, M. LeMasson, R. Lestini, Recombination proteins and rescue of arrested replication forks. *DNA Repair (Amst).* **6**, 967–980 (2007).

8. R. C. Heller, K. J. Marians, Replisome assembly and the direct restart of stalled replication forks. *Nat. Rev. Mol. Cell Biol.* **7**, 932–43 (2006).

9. H. L. Klein, K. N. Kreuzer, Replication, Recombination, and Repair: Going for the Gold. *Mol. Cell*. **9**, 471–480 (2002).

10. M. Lopes *et al.*, The DNA replication checkpoint response stabilizes stalled replication forks. *Nature*. **412**, 557–561 (2001).

11. H. Merrikh, Y. Zhang, A. D. Grossman, J. D. Wang, Replication–transcription conflicts in bacteria. *Nat. Rev. Microbiol.* **10**, 449–458 (2012).

12. M. M. Cox, The nonmutagenic repair of broken replication forks via recombination. *Mutat. Res. Mol. Mech. Mutagen.* **510**, 107–120 (2002).

13. E. V Mirkin, S. M. Mirkin, Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.* **71**, 13–35 (2007).

14. R. P. Fuchs, S. Fujii, Translesion DNA Synthesis and Mutagenesis in Prokaryotes. *Cold Spring Harb. Lab. Press*. **5**, 1–22 (2013).

15. R. P. Fuchs, Tolerance of lesions in *E. coli*: Chronological competition between Translesion Synthesis and Damage Avoidance. *DNA Repair (Amst).* **44**, 51–58 (2016).

16. M. F. Goodman, R. Woodgate, Translesion DNA polymerases. *Cold Spring Harb. Perspect. Biol.* **5**, 247–250 (2013).

17. S. Lusetti, M. M. Cox, The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Ann. Rev. Biochem*. **71**, 71–100 (2002).

18. W. D. Rupp, P. Howard-Flanders, Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J. Mol. Biol.* **31**, 291–304 (1968).

19. F.-X. Barre *et al.*, Circles: the replication-recombination-chromosome segregation connection. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8189–8195 (2001).

20. D. J. Sherratt *et al.*, Recombination and chromosome segregation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**, 61–69 (2004).

21. T. Shibata *et al.*, Functional overlap between RecA and MgsA (RarA) in the rescue of stalled replication forks in *Escherichia coli*. *Genes to Cells*. **10**, 181–191 (2005).

22. A. N. Page, N. P. George, A. H. Marceau, M. M. Cox, J. L. Keck, Structure and biochemical activities of *Escherichia coli* MgsA. *J. Biol. Chem.* **286**, 12075–12085 (2011).

23. M. J. Davey, D. Jeruzalmi, J. Kuriyan, M. O'Donnell, Motors and switches: AAA+ machines within the replisome. *Nat. Rev. Mol. Cell Biol.* **3**, 826–835 (2002).

24. J. P. Erzberger, J. M. Berger, Evolutionary Relationships and Structural Mechanisms of Aaa+ Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).

25. L. M. Iyer, D. D. Leipe, E. V. Koonin, L. Aravind, Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* **146**, 11–31 (2004).

26. T. Ogura, A. J. Wilkinson, AAA+ superfamily ATPases: Common structure-diverse function. *Genes to Cells*. **6**, 575–597 (2001).

27. A. Costes, F. Lecointe, S. McGovern, S. Quevillon-Cheruel, P. Polard, The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS Genet.* **6**, 1–15 (2010).

28. I. F. Lau *et al.*, Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* **49**, 731–743 (2003).

29. D. Branzei, M. Seki, F. Onoda, T. Enomoto, The product of Saccharomyces cerevisiae WHIP/MGS1, a gene related to replication factor C genes, interacts functionally with DNA polymerase Delta. *Mol. Genet. Genomics*. **268**, 371–386 (2002).

30. A. Yoshimura *et al.*, Physical and functional interaction between WRNIP1 and RAD18. *Genes Genet. Syst.* **84**, 171–178 (2009).

31. N. Crosetto *et al.*, Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J. Biol. Chem.* **283**, 35173–35185 (2008).

32. I. Saugar, J. L. Parker, S. Zhao, H. D. Ulrich, The genome maintenance factor Mgs1 is targeted to sites of replication stress by ubiquitylated PCNA. *Nucleic Acids Res.* **40**, 245–257 (2012).

33. H. Nomura *et al.*, WRNIP1 accumulates at laser light irradiated sites rapidly via its ubiquitin-binding zinc finger domain and independently from its ATPase domain. *Biochem. Biophys. Res. Commun.* **417**, 1145–1150 (2012).

34. T. Hishida, T. Ohya, Y. Kubota, Y. Kamada, H. Shinagawa, Functional and physical interaction of yeast Mgs1 with PCNA: impact on RAD6-dependent DNA damage tolerance. *Mol. Cell. Biol.* **26**, 5509–17 (2006).

35. R. A. Bish, M. P. Myers, Werner helicase-interacting protein 1 binds polyubiquitin via its zinc finger domain. *J. Biol. Chem.* **282**, 23184–23193 (2007).

36. D. Branzei *et al.*, Characterization of the slow-growth phenotype of S. cerevisiae whip/mgs1 sgs1 double deletion mutants. *DNA Repair (Amst).* **1**, 671–682 (2002).

37. T. Hishida, H. Iwasaki, T. Ohno, T. Morishita, H. Shinagawa, A yeast gene, MGS1, encoding a DNA-dependent AAA(+) ATPase is required to maintain genome stability. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8283–9 (2001).

38. T. Hishida, T. Ohno, Saccharomyces cerevisiae MGS1 is essential in strains deficient in the RAD6 -dependent DNA damage tolerance pathway. *EMBO J.* **21**, 2019–2029 (2002).

39. R. Lestini, B. Michel, UvrD controls the access of recombination proteins to blocked replication forks. *EMBO J.* **26**, 3804–14 (2007).

40. J. H. Kim *et al.*, In vivo and in vitro studies of Mgs1 suggest a link between genome instability and Okazaki fragment processing. *Nucleic Acids Res.* **33**, 6137–6150 (2005).

41. T. Hayashi *et al.*, Vertebrate WRNIP1 and BLM are required for efficient maintenance of genome stability. *Genes Genet. Syst.* **83**, 95–100 (2008).

42. Y. ichi Kawabe *et al.*, Analyses of the interaction of WRNIP1 with Werner syndrome protein (WRN) in vitro and in the cell. *DNA Repair (Amst).* **5**, 816–828 (2006).

43. A. Yoshimura, Y. Kobayashi, S. Tada, M. Seki, T. Enomoto, WRNIP1 functions upstream of DNA polymerase eta in the UV-induced DNA damage response. *Biochem. Biophys. Res. Commun.* **452**, 48–52 (2014).

44. A. Yoshimura *et al.*, Functional relationships between Rad18 and WRNIP1 in vertebrate

cells. *Biol. Pharm. Bull.* **29**, 2192–6 (2006).

45. G. D. Bowman, M. O'Donnell, J. Kuriyan, Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature*. **429**, 724–730 (2004).

46. D. Jeruzalmi, M. O'Donnell, J. Kuriyan, Crystal structure of the processivity clamp loader gamma (??) complex of *E. coli* DNA polymerase III. *Cell*. **106**, 429–441 (2001).

47. T. Tsurimoto, A. Shinozaki, M. Yano, M. Seki, T. Enomoto, Human Werner helicase interacting protein 1 (WRNIP1) functions as novel modulator for DNA polymerase delta. *Genes to Cells*. **10**, 13–22 (2005).

48. N. D. Vijeh Motlagh, M. Seki, D. Branzei, T. Enomoto, Mgs1 and Rad18/Rad5/Mms2 are required for survival of Saccharomyces cerevisiae mutants with novel temperature/cold sensitive alleles of the DNA polymerase delta subunit, Pol31. *DNA Repair (Amst).* **5**, 1459–1474 (2006).

49. R. D. Shereda, A. G. Kozlov, T. M. Lohman, M. M. Cox, J. L. Keck, SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* **43**, 289–318 (2008).

50. F. Lecointe *et al.*, Anticipating chromosomal replication fork arrest: SSB targets repair DNA helicases to active forks. *EMBO J.* **26**, 4239–4251 (2007).

51. H. B. Gamper, G. D. Cimino, J. E. Hearst, Solution hybridization of crosslinkable DNA oligonucleotides to bacteriophage M13 DNA. Effect of secondary structure on hybridization kinetics and equilibria. *J. Mol. Biol.* **197**, 349–362 (1987).

52. B. Reckmann *et al.*, Analysis of secondary structures in M13mp8 (+) single-stranded DNA by the pausing of D N A polymerase a. *Eur. J. Biochem.* **643**, 633–643 (1985).

53. V. P. Antao, S. Y. Lai, I. Tinoco, A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* **19**, 5901–5905 (1991).

54. S. M. Law, R. Eritja, M. F. Goodman, K. J. Breslauer, Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. *Biochemistry*. **35**, 12329–12337 (1996).

55. K. D. Raney, L. C. Sowers, D. P. Millar, S. J. Benkovic, A fluorescence-based assay for monitoring helicase activity. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6644–6648 (1994).

56. A. Ronen, 2-aminopurine. *Mutat. Res.* **75**, 1–47 (1979).

57. R. Eritja *et al.*, Synthesis and properties of defined DNA oligomers containing base mispairs involving 2-aminopurine. *Nucleic Acids Res.* **14**, 5869–5884 (1986).

58. N. A. Tanner *et al.*, Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Res.* **37**, 2–7 (2009).

59. N. A. Tanner *et al.*, Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat. Struct. Mol. Biol.* **15**, 170–176 (2008).

60. J. S. Lewis *et al.*, Single-molecule visualization of fast polymerase turnover in the bacterial replisome. *Elife*. **6**, 1–17 (2017).

61. M. Mok, K. J. Marians, The *Escherichia coli* preprimosome and DNA B helicase can form replication forks that move at the same rate. *J. Biol. Chem.* **262**, 16644–16654 (1987).

62. N. Y. Yao, R. E. Georgescu, J. Finkelstein, M. E. O'Donnell, Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13236–41 (2009).

63. P. McInerney, A. Johnson, F. Katz, M. O'Donnell, Characterization of a Triple DNA Polymerase Replisome. *Mol. Cell*. **27**, 527–538 (2007).

64. N. A. Tanner *et al.*, *E. coli* DNA replication in the absence of free beta clamps. *EMBO J.* **30**, 1830–1840 (2011).

65. J. a Farah, G. R. Smith, The RecBCD enzyme initiation complex for DNA unwinding: enzyme positioning and DNA opening. *J. Mol. Biol.* **272**, 699–715 (1997).

66. C. J. Wong, A. L. Lucius, T. M. Lohman, Energetics of DNA end binding by *E. coli* RecBC and RecBCD helicases indicate loop formation in the 3???-single-stranded DNA tail. *J. Mol. Biol.* **352**, 765–782 (2005).

67. B. Michel, A. K. Sinha, The inactivation of *rfaP* , *rarA* or *sspA* gene improves the viability of the *Escherichia coli* DNA polymerase III *holD* mutant. *Mol. Microbiol.* **104**, 1008–1026 (2017).

68. B. P. Dalrymple, K. Kongsuwan, G. Wijffels, N. E. Dixon, P. A. Jennings, A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc. Natl. Acad. Sci.* **98**, 11627–11632 (2001).

69. G. Wijffels *et al.*, Inhibition of Protein Interactions with the ??2 Sliding Clamp of *Escherichia coli* DNA Polymerase III by Peptides from ?? 2-Binding Proteins. *Biochemistry*. **43**, 5661–5671 (2004).

70. J. T. P. Yeeles, K. J. Marians, Dynamics of leading-strand lesion skipping by the replisome. *Mol. Cell*. **52**, 855–865 (2013).

71. R. C. Heller, K. J. Marians, Replication fork reactivation downstream of a blocked nascent leading strand. *Nature*. **439**, 557–562 (2006).

72. S. T. Lovett, Template-switching during replication fork repair in bacteria. *DNA Repair (Amst).* **56**, 118–128 (2017).

73. A. J. Oakley *et al.*, Flexibility revealed by the 1.85 angstrom crystal structure of the beta sliding-clamp subunit of *Escherichia coli* DNA polymerase III. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **59**, 1192–1199 (2003).

74. C. E. Mason *et al.*, *Escherichia coli* single-stranded DNA-binding protein: NanoESI-MS studies of salt-modulated subunit exchange and DNA binding transactions. *J. Am. Soc. Mass Spectrom.* **24**, 274–285 (2013).

75. S. Jergic *et al.*, A direct proofreader–clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J.* **32**, 1322–1333 (2013).

76. N. P. J. Stamford, J. Stamford, P. E. Lilley, N. E. Dixon, Enriched sources of *Escherichia coli* replication proteins. The dnaG primase is a zinc metalloprotein. *BBA - Gene Struct. Expr.* **1132**, 17–25 (1992).

77. J. E. Lindsley, M. M. Cox, Assembly and Disassembly of RecA Protein Filaments Opposite Filament Ends Occur at. *J. Biol. Chem.* **265**, 9043–9054 (1990).

78. S. W. Morrical, J. Lee, M. M. Cox, Continuous Association of *Escherichia coli* Single-Stranded DNA Binding Protein with Stable Complexes of recA Protein and Single-Stranded DNA. *Biochemistry*. **25**, 1482–1494 (1986).

79. H. J. Geertsema *et al.*, Single-molecule imaging at high fluorophore concentrations by local activation of dye. *Biophys. J.* **108**, 949–956 (2015).

80. J. S. Lewis *et al.*, Single-molecule visualization of *Saccharomyces cerevisiae* leading-strand synthesis reveals dynamic interaction between MTC and the replisome. *Proc. Natl. Acad. Sci.* **114**, 10630–10635 (2017).

81. R. E. Georgescu, I. Kurth, M. E. O'Donnell, Single-molecule studies reveal the function of a third polymerase in the replisome. *Nat. Struct. Mol. Biol.* **19**, 113–116 (2011).

Chapter 3: *In vivo* consequences of *rarA* loss of function and overexpression

Introduction

Replication forks can stall when encountering roadblocks, such as DNA lesions, template-strand breaks or DNA-bound proteins. The outcomes of fork stalling may include replication fork collapse or disengagement of the DNA polymerase III replisome (*1–11*). If these outcomes are not avoided or repaired, they can have catastrophic consequences on genome integrity and cell viability. The growing consensus is that these events aren't uncommon, with recent evidence suggesting that replication fork collapse occurs at least once per cell cycle under normal growth conditions (*1*, *2*, *12–18*). The majority of these conflicts are resolved using highly conserved nonmutagenic pathways utilizing homologous recombination (*1–11*, *19–21*).

When bacterial cells are stressed by conditions that inflict long-lasting or severe DNA damage, the transcriptional cascade termed the SOS response is induced (*22*, *23*). In the early stages of SOS, nonmutagenic DNA repair still dominates; however, in later stages of the SOS response, a different pathway to tolerate the DNA damage prevails. These pathway utilize specialized DNA polymerases to bypass sites of DNA damage using a process called translesion DNA synthesis (TLS) (*24–37*). In *E. coli*, TLS is carried out by DNA polymerase II (*polB*), DNA polymerase IV (*dinB*), and DNA polymerase V (*umuDC*). Despite its ability to be error-prone, translesion DNA synthesis becomes the predominant pathway in DNA damage tolerance when homologous recombination is unable to function (*38*).

When the SOS response is induced, levels of all TLS DNA polymerases increase. However, unlike DNA polymerase V, which is exclusively expressed in SOS conditions, DNA polymerases II and IV are present under normal growth condition at 30-50 and 250 molecules per cell, respectively (*39–44*). The reason for the constitutive expression of DNA polymerases II and IV remains enigmatic. Cells lacking all three TLS polymerases do no exhibit a growth defect, nor

are they highly sensitive to ultraviolet light (*45*), underlining the cell's preference for nonmutagenic homologous recombination as a DNA repair pathway choice. However, when cells lacking TLS polymerases are stressed with certain mutagens, they exhibit a decrease in viability, suggesting a role for this pathway under certain conditions (*40*, *46–52*). While many mechanisms have been presented for how *E. coli* cells choose which DNA repair pathway to utilize upon replication fork encounter with diverse DNA lesions, whether this mechanism requires an active component remains to be seen. In this work, we propose that the highly conserved RarA protein actively dictates DNA repair pathway choice in *E. coli*.

The *Escherichia coli* RarA protein is a AAA+ ATPase belonging to a family containing highly conserved homologous in yeast (Mgs1) and mammals (WRNIP1). Sequence conservation between RarA and its eukaryotic homologues is extensive, with RarA sharing 40% amino acid identity and 56-58% similarity with its *S. cerevisiae* and *H. sapiens* homologues (*53*). This high amount of sequence conservation suggests a conserved function across all domains of life. In *E. coli*, RarA is most closely related to the clamp loader complex encoded by the *dnaX* gene, which is responsible for loading the β_2 processivity clamp onto primer-template DNA (*53*, *54*). Despite its high sequence conservation and its implicated function in the maintenance of stalled replication forks, the precise function of RarA family proteins remains highly enigmatic.

Several dozen studies have now been published on the RarA/Mgs1/WRNIP1 protein family. Although these have yielded a complex, and sometimes contradictory plethora of observations, several themes are evident. First, RarA family members localize to the replication fork through interactions with either the single-stranded DNA binding protein, SSB (RarA), or ubiquitylated processivity clamp PCNA (Mgs1 and WRNIP1) (*53*, *55–61*). Second, the sequence and structure of RarA (and by extension other family members) place it in the clamp-loader clade

of AAA+ ATPases (*56*, *62*). Interestingly, *E. coli* RarA adopts a tetrameric conformation, unlike all other known members of the clamp-loader clade (*56*, *62*). Third, RarA has an effect on replisome stability and somehow promotes TLS (*63–67*). Fourth, RarA, Mgs1, and WRNIP1 all exhibit *in vitro* DNA-dependent ATPase activity specifically stimulated by double-stranded DNA ends and single-stranded DNA gaps (*56*, *64*, *68–72*). A recent observation that RarA creates single-stranded DNA gaps behind the replisome was also discussed in Chapter 2. Fifth, RarA function appears to complement a range of DNA damage tolerance pathways (*58*, *64*, *68*, *69*, *73–78*). These genetic results suggest that RarA does not belong to any currently defined pathway.

Utilizing a combination of growth assays, fluorescence microscopy, and DNA damage sensitivity assays, we provide evidence that RarA acts upstream of single-stranded DNA gap repair and translesion DNA synthesis to influence DNA repair pathway choice *in vivo*. This upstream action is likely a result of the *in vitro* RarA-mediated gap creation activity we demonstrated in Chapter 3. We provide a model in which RarA creates single-stranded DNA gaps behind the replisome in response to DNA damage encountered by the replication fork. We hypothesize that this allows for the replication fork to actively bypass sites of DNA damage and allow for post-replication repair of the lesions via homologous recombination or TLS.

Results

Loss of rarA function results in a growth defect in E. coli

We hypothesize that the main function of RarA-mediated gap formation is to promote optimal progression of the replication fork when lesions or other barriers are encountered. Whereas truly stalled forks may occur at most a few times per cell cycle, there may be many more circumstances where lesions or replication pause sites are simply bypasses, leaving daughter strand gaps behind the fork. Loss of RarA function may thus result in a growth defect. No growth or viability phenotype has previously been ascribed to strains with an interrupted *rarA* gene (*53*, *58*, *64–66*). Previous work has focused on a modified *rarA* gene in which a chloramphenicol-resistance cassette replaced either the first 600 nucleotides of the gene (*53*, *58*, *65*, *66*) or codons 113-349 (*64*) in an the AB1157 substrain of *E. coli*. As most of our constructs are based on *E. coli* substrain MG1655, we constructed a complete *rarA* deletion in the MG1655 background and characterized the knockout strain using various techniques.

First, we compared the growth of the Δ*rarA* strain to wild-type cells in rich medium (LB). The Δ*rarA* cells grew much more slowly than wild type cells, exhibiting a doubling time of 42 versus 29 min for wild type cells (**Figure 3-1A**). To determine the relative fitness cost of a *rarA* deletion, we carried out direct competition assays between the wild type strain and the Δ*rarA* strain using an approach developed by Lenski and colleagues (**Figure 3-1B**) (*79*). Wild type or mutant cells were modified to carry a neutral Δ*araBAD* allele which confers a red colony color when plated on tetrazolium arabinose (TA) agar plates. Overnight cultures of Δ*rarA* cells were mixed in a 50/50 ratio with isogenic wild type cells carrying the Δ*araBAD* allele. The mixed culture was then diluted and grown up again on successive days, with plating on TA plates to score for the

percentage of red colonies each day. Earlier work demonstrated that the Δ*araBAD* mutation does not affect growth rates by itself (*79*, *80*); nonetheless, we carried out the growth competitions twice with the Δ*araBAD* allele in either the wild type or Δ*rarA* cells. In both experiments, the wild type cells outcompeted Δ*rarA* cells and dominated the mixed cultures almost completely within 48 hours (**Figure 3-1B**). Base on the 24h time point, we calculated that the Δ*rarA* strain had a relative fitness $w = 0.5$, indicating a significant loss of fitness relative to wild type cells (*81*).

Loss of rarA results in decreased cell size and impaired DNA replication

To investigate why Δ*rarA* cells grow more slowly than wild type cells, we carried out single-molecule single-cell microscopy in collaboration with Megan Cherry and Andrew Robinson at the University of Wollongong in Wollongong, Australia. To see if there were any replication defects, we made use of strains expressing the ε subunit of DNA polymerase III subunit fused with the YPet fluorescence protein (*dnaQ*-YPet) from its native locus on the chromosome as described previously (*26*). The effect this allele had on growth rate, DNA content, and replication dynamics were previously analyzed and confirmed to be wild type in nature by several research groups (*82*, *83*).

We observed that Δ*rarA* cells were substantially smaller than *rarA*+ cells (3.7 [SEM = 0.03] versus 5.6 [SEM = 0.04] μm in length) and divided less frequently within the flow cell environment used for imaging (division time = 53 [SEM = 2] versus 26 [SEM = 1] min; $n = 20$ cells each) (**Figure 3-1CD**). Additionally, *dnaQ*-YPet *rarA*+ cells had between 0 and 10 replication foci (mean = 2.8; SEM = 0.03; $n = 2371$ cells) (**Figure 3-1D**), consistent with multi-fork DNA replication due to growth in rich imaging medium. Cells carrying the *rarA* deletion (*dnaQ*-YPet Δ*rarA*) had fewer foci (mean = 1.2; SEM = 0.03; $n = 1116$ cells), consistent with their slower growth rate. Approximately 7% of cells contained more than 2 replication foci, indicating

that Δ*rarA* cells are capable of multi-fork replication, but grow slow enough that this mode of replication is not usually necessary.

We also examined cell size and DNA content using flow cytometry, with results that were consistent with those found using single-cell fluorescence microscopy. Wild type or Δ*rarA* cells were grown to early exponential phase (OD_{600} = 0.2) in rich medium, washed in phosphate buffered saline, and diluted to an appropriate concentration. Both forward area light scattering and fluorescence from the DNA-specific dye Vybrant Dyecycle Green were measured in a flow cytometer. Data obtained from the flow cytometer were gated to exclude cell debris and doublet cells (**Figure 3-2AB**). Cell size was significantly reduced in cells lacking *rarA*, as measured by forward scattering area (**Figure 3-2CE**). DNA content was also significantly reduced and exhibited a bimodal distribution in Δ*rarA* cells as opposed to a unimodal distribution seen in wild type cells (**Figure 3-2DF**). These results further demonstrate the effects of deleting *rarA* on growth rate and DNA replication.

RarA creates substrates for RecFOR-mediated homologous recombination in response to UV damage

The RecF, RecO, and RecR proteins have been implicated in repair of daughter-strand gaps. These proteins all have a role in loading RecA onto SSB-coated single-stranded DNA at ssDNA gaps and in some cases dsDNA ends. Loss of function of any of the RecFOR proteins results in sensitivity to UV irradiation (*7*, *84–87*). If a significant proportion of the gaps that act as RecFOR substrates in UV-irradiated cells are created by RarA action, then loss of RarA function could decrease the numbers of UV-associated gaps. Fork stalling would be more likely and UV lesion repair would be channeled into pathways other than RecFOR-mediated daughter-strand gap repair. In this case, a *rarA* deletion could suppress or partially suppress *recFOR* mutations, depending upon how much RecFOR function was focused on gap repair.

As such, we tested the sensitivities of strains harboring Δ*recO* or Δ*recF* alleles in conjunction with Δ*rarA* to UV irradiation. When Δ*recO* cells were exposed to UV irradiation at an incident dose of 15 J/m^2, cell survival declined by approximately 2 orders of magnitude relative to wild type cells (**Figure 3-3A**). Cells lacking the *rarA* gene exhibited no decline in survival, despite growing more slowly (**Figure 3-3A**). When the Δ*recO* allele was combined with the Δ*rarA* allele, cells no longer were sensitive to UV irradiation at 15 J/m^2 compared to a single Δ*recO* knockout strain (**Figure 3-3A**). This Δ*recO* cells are not sensitive to UV light at low (15 J/m^2) doses solely due to the lack of the RecFOR pathway; instead, *rarA* action is somehow sensitizing these cells to this type of damage. Deletion of *rarA* also suppressed the UV sensitivity of Δ*recF* cells at a dose of 15 J/m^2 (**Figure 3-3B**), even though Δ*recF* cells are not as sensitive to UV-induced damage than Δ*recO* cells.

When the UV dose was increased to 30 J/m^2, a similar result was obtained (**Figure 3-3**) Survival by the cells lacking *recO* or *recF* declined further at this higher dose of UV light, as expected. Suppression of the UV sensitivity of Δ*recO* and Δ*recF* cells by deleting *rarA* was only partial when compared to the lower UV dose of 15 J/m^2, yet was still very clear. This is most likely due to an increased number of RarA-independent gaps (and therefore, RecFOR substrates) created at the higher dose of UV light.

RarA creates substrates for translesion DNA synthesis in an ATPase-dependent manner

Bacterial cells lacking the translesion DNA synthesis DNA polymerase IV (Δ*dinB*) function are highly sensitive to the agents methyl methanesulfonate (MMS), nitrofurazone (NFZ), and 4-nitroquinoline-1-oxide (4-NQO) during exponential growth in rich medium (*46–50*). DNA polymerase IV can bypass lesions at guanine-N^2 resulting from treatment with these agents (*47*,

88, *89*). The sensitivity of *dinB* mutants to these agents has been interpreted as reflecting the absence of the only capable TLS polymerase able to bypass this lesion. If RarA is creating suitable substrates for TLS action upon treatment of cells with these genotoxic agents, eliminating these RarA-dependent substrates should eliminate the need for TLS. This in turn would result in a suppression of the sensitivity of Δ*dinB* cells to these genotoxic agents when *rarA* was also deleted.

To test if this hypothesis was correct, we utilized a similar approach to that used with the suppression of Δ*recO or* Δ*recF* cells' sensitivity to ultraviolet light. Instead, we incorporated nitrofurazone (NFZ) into agar plates and tested the sensitivity of cells with a Δ*dinB* allele to grow on this media. Bacterial cultures were plated on either LB agar or LB agar containing NFZ upon reaching early exponential growth (OD_{600} = 0.2). As previously reported, Δ*dinB* cells were sensitized to NFZ by 2-3 orders of magnitude when compared to a wild type control (**Figure 3-4A**). Cells lacking *rarA* alone were not sensitized at the indicated concentration of NFZ compared to wild type cells (**Figure 3-4A**). However, when a Δ*rarA* allele was included with a Δ*dinB* allele, the previously seen sensitization to NFZ was completely suppressed (**Figure 3-4A**). Thus, the combination of *rarA* action and the absence of *dinB* in Δ*dinB* cells is responsible for the sensitization of these cells to NFZ. To test whether this suppression was dependent on RarA ATPase activity, we constructed a strain where *rarA* was replaced with a mutant *rarA* K63R allele which lacks competent ATPase activity. When *rarA* K63R was combined with a Δ*dinB* allele, these cells were also no longer sensitive to NFZ-induced DNA damage (**Figure 3-4A**). These data suggest that *rarA* action not only acts upstream of Pol IV by creating suitable substrates for TLS, but also excludes alternative repair pathways at the same time, leading to a sensitization of NFZ-induced DNA lesions.

The suppression effect of a *rarA* mutant is not specific to cells lacking DNA polymerase IV. Mutants of DNA polymerase II (Δ*polB*) are also sensitive to NFZ-induced DNA damage (*51*). Mutants of DNA polymerase V (Δ*umuDC*) are sensitive to high doses of UV irradiation (*45*). The sensitivity of these strains to NFZ or UV are suppressed when either a Δ*rarA* or *rarA* K63R allele are included in combination with either of the TLS polymerase knockouts (**Figure 3-4BC**). These data suggest that the mechanism by which RarA action leads to reliance on TLS pathways is not specific to any one TLS polymerase.

Overexpression of RarA leads to cell death in an ATPase-dependent manner

After documenting some phenotypes associated with loss of *rarA* function, we wanted to further characterize an earlier observation by Shibata and colleagues that overexpression of *rarA* was toxic (*64*). Shibata found that overexpression of *rarA* in cells lacking the bacterial recombinase *recA* was lethal, while *rarA* overexpression in *recA*+ cells lead to a slow growth phenotype. Thus, whatever was causing cells to growth slowly upon *rarA* overexpression was somehow suppressed by *recA* function. To better understand this phenotype, we repeated these experiments by constructing a pBAD-based plasmid harboring *rarA*, *rarA* mutants, or other proteins and tested the effects of overexpression in this system. Overexpression in all of these experiments was accomplished by plating exponential growth phase bacterial cultures on LB agar plates containing 0.2% L-arabinose. Expression was suppressed by plating on LB agar plates containing 0.2% glucose.

In agreement with Shibata and colleagues, overexpression of *rarA* in a Δ*recA* background was extremely toxic, killing most cells in a spot plate assay (**Figure 3-5A**). Surprisingly, overexpression of *rarA* in a *recA*+ background was also toxic, albeit less so than when in a Δ*recA* background (**Figure 3-5A**). To explore if this phenotype was dependent on the ATPase activity of

RarA, we overexpressed RarA K63R, a Walker A mutant lacking ATPase activity, in a *recA*+ background. As expected, overexpression of *rarA* K63R was not lethal, but did impart a growth defect in these cells, as evidenced by reduced colony size when grown on an agar plate containing L-arabinose (**Figure 3-5B**). To explore the cause of slow growth phenotype associated with overexpression of *rarA* K63R, we constructed a pBAD plasmid harboring an exonuclease-dead mutant of the *sbcB* gene (*90*). The *sbcB* gene encodes the exonuclease I protein, which is known to also bind the single-stranded DNA binding protein (SSB) with a similar affinity to RarA (*91*). Overexpression of this exonuclease I mutant phenocopied the slow growth phenotype associated with overexpression of *rarA* K63R. Thus, we hypothesize that sequestration of SSB binding sites by *rarA* K63R upon its overexpression is the cause of the slow growth phenotype. Thus, we wanted to test a mutant of *rarA* that is unable to bind SSB or hydrolyze ATP to ensure these were both responsible for the slow growth and lethality (respectively) upon overexpression. As expected, overexpression of *rarA* K430E, a mutant variant that is unable to tetramerize (and thus perform any RarA function), was not deleterious at all when compared to an empty vector control (**Figure 3-5B**).

In order to eliminate any side effects of monitoring cell viability after *rarA* overexpression using spot plate assays, we chose to further characterize these phenotypes in shaking cultures. This experiment involved the same strains used in the spot plate assays; however, overexpression of *rarA* was induced in shaking culture and the cell viability of those cultures was tracked over time by counting colony forming units (CFUs). As expected, overexpression of wild type *rarA* was highly toxic, reducing CFUs by roughly 4 orders of magnitude over the 180 min time course of the experiment when compared to an empty vector control (**Figure 3-6A**). As expected from the spot plating results, overexpression of the non-functional *rarA* K430E mutant did not have a

significant impact on cell viability (**Figure 3-6A**). Overexpression of the ATPase-dead RarA K63R mutant resulted in 2 orders of magnitude loss in cell viability when compared to an empty vector control. Thus, overexpression of RarA K63R was still somewhat toxic, albeit not nearly as much as a fully function RarA protein. To ensure that these mutant proteins were all expressed to a similar level to wild type, we utilized Western Blots to assess relative protein amounts at each time point assayed in these experiments. No significant differences between the expression levels of wild type RarA or the two RarA mutants were apparent in these Western Blots (**Figure 3-6B**). We repeated the overexpression experiments in a Δ*recA* background, and found no significant differences than in a *recA*+ background, albeit the cells were naturally less viable in the Δ*recA* strains (**Figure 3-6A**).

Overexpression of RarA delocalizes DNA polymerase III in an ATPase-independent mechanism

Since overexpression of *rarA* was observed to be toxic in an ATPase-dependent manner, we hypothesized that there may be a connection to our *in vitro* results presented in Chapter 2 showing that RarA creates single-stranded DNA gaps behind the replisome. Overexpression of wild type RarA could lead to destabilization of the replisome due to its predicted interaction with the β_2 clamp. To characterize the effect of *rarA* overexpression on replisome stability *in vivo*, we monitored cells containing fluorescently labeled DNA polymerase III as *rarA* was overexpressed. In collaboration with Megan Cherry and Andrew Robinson at the University of Wollongong in Australia.

We carried out single-cell time-lapse fluorescence microscopy using cells containing fluorescently labeled DNA polymerase III (*dnaQ*-YPet), as described in Figure 3-1CD. In this case, we utilized Δ*rarA* cells containing a pBAD plasmid harboring either wild type *rarA* or *rarA*

K63R. Cells were initially grown in media containing 0.2% glucose to inhibit expression from the pBAD plasmid. At t = 0 min, cells were transferred to media containing 0.2% L-arabinose to induce overexpression of *rarA* or *rarA* K63R. These cells were then imaged over the course of 3 hours. Three hours after inducing *rarA* overexpression, cells severely filamented into large, elongated structures (**Figure 3-7AB**). In addition, DNA polymerase III foci that were present prior to *rarA* overexpression almost completely disappeared, suggesting they were disengaged from the DNA template (**Figure 3-7AB**). At the end of the 3-hour experiment, nearly 80% of all cells lacked a DNA polymerase III focus (**Figure 3-7B**). However, these observations were not dependent on the ATPase activity of RarA. These experiments were repeated by overexpressing the ATPase-dead RarA K63R mutant protein, with largely similar results, albeit with slightly milder effects on cell length and replisome foci (**Figure 3-7CD**). These data, combined with the *in vitro* data presented in Chapter 3, suggest that these phenotypes are most likely (at least in part) due to strong binding of RarA or RarA K63R to either SSB or β_2 clamp, both of which would outcompete DNA polymerase III cores from engaging with the DNA template.

Loss of TLS polymerase function partially suppresses the toxic phenotype of rarA overexpression

Previous work by Shibata and colleagues showed that *rarA* overexpression drastically increased genome instability, specifically the incidence of -1G frameshift mutations and base substitutions (*64*). Coincidentally, overexpression of *dinB*, the gene that encodes the translesion DNA synthesis polymerase IV, also increases the frequency of -1G frameshift mutations (*92*, *93*). Additionally, Pol V has been shown to be largely responsible for an increased incidence of base substitution mutations *in* vivo (*37*). Thus, taking these observations together with our own connections between TLS polymerases and RarA, we hypothesized that the source of mutagenesis upon *rarA* overexpression was due to TLS Pols being preferentially used for DNA synthesis over

Pol III. To test this hypothesis, we deleted Pol IV and Pol V and repeated the overexpression spot plate assays described in Figure 3-4. If overuse of TLS Pols was responsible for the toxic phenotype associated with *rarA* overexpression, we hypothesized that deleting these TLS Pols would suppress this phenotype.

We constructed strains lacking either Pol IV (*dinB*), Pol V (*umuDC*), or both (*dinB/umuDC*), and repeated *rarA* overexpression experiments again utilizing pBAD-based vectors as described above (**Figure 3-4**). As expected, deletion of either *umuDC* or *dinB* suppressed the toxic phenotype of *rarA* overexpression by one or three orders of magnitude, respectively (Figure 3-8). These suppression effects were additive when both Pol IV and Pol V were deleted, again underlining the hypothesis that *rarA* acts to create substrates for TLS Pols, and is not specific to any one polymerase (**Figure 3-8**). It is unclear as to why the heterologous colony sizes are seen upon *rarA* overexpression in cells lacking TLS Pols. In all cases, agar plates contained both L-arabinose to overexpress *rarA* and ampicillin to maintain the pBAD plasmid, suggesting that the different colony sizes were not solely due to loss of the toxic *rarA* overexpression plasmid. We hypothesize that this effect can be attributed to the slow growth phenotypes associated with *rarA* overexpression and the previous observation that overexpression from pBAD-based vectors can result in heterologous expression levels (*94*). Regardless, it is clear that deletion of TLS Pols at least partially suppresses the toxic phenotype associated with *rarA* overexpression.

Discussion

We conclude that the function of RarA is to create single-stranded DNA gaps behind the replication fork upon encountering DNA damage. *In vivo*, this activity is utilized to (a) facilitate normal growth and replisome progression and (b) create DNA substrates that are able to be repaired using daughter-strand gap repair or tolerated using translesion DNA synthesis (**Figure 3-9**). Loss of *rarA* function leads to a growth defect, reduced cell size, and reduced DNA content when compared to wild type cells (**Figures 3-1**; **3-2**). In addition, loss of *rarA* suppresses the sensitivity of cells lacking daughter-strand gap repair (RecFOR) or translesion DNA synthesis (Pols II, IV, or V) to DNA damaging agents (**Figures 3-3**; **3-4**). Overexpression of *rarA* is lethal due to its ATPase activity (**Figures 3-5**; **3-6**). However, we also showed that *rarA* overexpression leads to cell filamentation and delocalization of replisomes in an ATPase-independent manner (**Figure 3-7**). Finally, we showed that the lethality associated with *rarA* overexpression can be partially suppressed upon deletion of TLS polymerase IV and V (**Figure 3-8**). Taken together, it is clear that a precise concentration of *rarA* is required in cells, as both deletion and overexpression are deleterious.

We have placed RarA function at the interface between DNA replication and repair. The most recent model for RarA function suggested it promoted a switch between DNA replication and translesion DNA synthesis by directly loading TLS Pols onto broken replication forks (*66*). Instead, we hypothesize that RarA facilitates a switch between DNA repair ahead of a replication fork to post-replication repair behind the replication fork. At sites of specific DNA damage on the lagging strand, RarA disengages a DNA polymerase from its processivity clamp, thereby forcing it to find the next available clamp. This process indirectly creates a large single-stranded DNA gap, as we have shown *in vitro* in Chapter 2. This function appears to be critical for normal cell

growth, as deletion of *rarA* resulted in slow growth, smaller cell size, and lower DNA content than wild type cells (**Figure 3-1**; **3-2**). From these phenotypes, we conclude that *rarA* function is required for optimal progression of the replication fork *in vivo*. If *rarA* is not present, more time-consuming repair pathways, like replication fork regression, would need to be employed to repair DNA lesions ahead of the fork. However, RarA-mediated gap formation behind the replication fork creates idealized substrates for both daughter-strand gap repair via homologous recombination or DNA damage tolerance via translesion DNA synthesis. Indeed, if processivity clamps are left behind in these gaps (as shown *in vitro* in Chapter 2), they serve as ideal scaffolds for direct TLS (*30*, *32*, *33*, *95*, *96*). Thus, RarA action may be part of the reason why Pol II and Pol IV are present during normal cell growth even in the absence of the SOS response (*41–43*).

From data presented in this chapter, we conclude that overproduction of *rarA* is lethal due to continuous disengagement of DNA polymerase III from SSB and/or β_2 processivity clamp at the replication fork. This would allow for TLS polymerases to preferentially fill the void left behind on β_2. Overuse of DNA polymerase IV has been shown to be lethal *in vivo*, supporting this hypothesis (*93*). In an active manner, RarA may also create too many gaps behind the replication fork, which could also overuse TLS. This also explains the previous observation that overexpression of *rarA* is more toxic in cells lacking homologous recombination, as TLS would be the only option to fill in the excess gaps (*64*). In addition, this further explains the increase in mutagenesis that is also seen upon *rarA* overproduction, specifically the mutations associated with Pol IV and Pol V use (*64*). Taken together with previous data, our data clearly show that *rarA* overproduction is lethal due to overuse of translesion DNA polymerases, as a consequence of RarA's gap creation activity.

It should be noted that many groups have observed single-stranded DNA gaps forming behind the replication fork when cells are treated with DNA damaging agents. Rupp and Howard-Flanders first documented the presence of single-stranded DNA behind active replication forks in 1968, after irradiating *E. coli* cells and isolating genomic DNA using a sucrose gradient (*97*). More recently, Páges and Fuchs observed single-stranded DNA gaps behind the replication fork following encounter with a site-specific DNA lesion (*98*). This process does not seem to be specific to bacteria, as recent work by Lopes and colleagues demonstrated the presence of single-stranded DNA gaps behind replication forks following UV irradiation in *S. cerevisiae* (*99*). It remains to be determined if these gaps are due to RarA action. Further work should be carried out to determine if these gaps decrease in frequency in the absence of *rarA*.

To our knowledge, this is the first suggestion that DNA lesion skipping is not an inherent property of DNA polymerase III. Rather, we present data that support the hypothesis that RarA directly catalyzes this process *in vivo* as a mechanism to bypass DNA damage. Marians, O'Donnell, and colleagues have shown that DNA polymerase III is able to directly bypass site-specific DNA lesions on both the leading and lagging strand template *in vitro* (*100–103*). However, these reconstituted DNA replication assays required several minutes to achieve lesion bypass. On the timescale of DNA replication *in vivo*, it is clear that this process cannot proceed on the order of minutes in the cell. Thus, RarA could be responsible for catalyzing this reaction *in vivo*.

Experimental procedures

Strain construction

All strains were *E. coli* MG1655 derivatives and are listen in Table 1. All parent strains were constructed using Lambda Red recombination as described by Datsenko and Wanner (*104*). All chromosomal mutations were confirmed using Sanger sequencing. When required, antibiotic resistance of a given strain was eliminated using FLP recombinase encoded by the pLH29 plasmid as described previously (*105*). Strains were transformed to harbor indicated plasmids using conventional methods.

Growth curves – plate reader

Overnight cultures of indicated strains were diluted 1:100 in LB. Three biological replicates were prepared in a clear bottom 96 well plate (Corning, Corning, NY). Cultures were grown at 37°C with continuous orbital shaking in a BioTek Synergy 2 plate reader (BioTek Instruments Inc., Winooski, VT). OD_{600} absorbance readings were measured every 10 min for over the course of the experiment and normalized by subtracting out the OD_{600} value of a blank sample containing only LB. Growth curves represent the average of at least three biological replicates, with error bars representing one S.D. from the average value.

Growth competition assays

Growth competition assays were conducted as previously described (*80*) using a method originally described by Lenski and colleagues (*79*). The Δ*araBAD* ΔP_{araB} marker was included on either the wild type (MG1655) or Δ*rarA* strains in separate experiments to control for any effect the marker may have had on cell fitness. These experiments are the result of three biological replicates, with error bars representing one S.D. from the average value.

Single-molecule time-lapse imaging and analysis

Cells were grown at 37°C in EZ rich defined medium (Teknova, Hollister, CA) that included 0.2% (*w*/*v*) glycerol. Flow cells were constructed by affixing a quartz piece, embedded with inlet and outlet tubing, to a (3-aminopropyl)triethoxysilane (APTES, Alfa Aesar, Haverhill, MA) functionalized glass coverslip according to the design specifications and procedure detailed previously (*26*).

For time-lapse imaging, flow cells were mounted to the microscope where temperature was maintained at 37°C by a combination of stage and objective lens heating. Cells were grown in shaking culture until reaching mid exponential growth phase, then loaded into mounted flow cells by pulling through the outlet with a syringe. The inlet was then placed into fresh, constantly aerated, EZ medium. Aerated medium was then pulled through the flow cell using a syringe pump, at a rate of 50 μL/min. To induce overexpression of *rarA*, 0.2% L-arabinose (Sigma) was added to the medium at $t = 0$ min.

A microscope constructed to the specifications described previously was used for these experiments (*26*). All *dnaQ* replisome labeled strains were excited at 18 W/cm^{-2} with a 514 nm laser light for 500 ms. Imaging was done using a 512 x 512 pixel EM-CCD camera (C9100-13, Hamamatsu). Image processing was performed using custom plugins. All cell outlining was manually performed in the open source MicrobeTracker Suite in MATLAB R2013b (Mathworks).

Single-molecule fluorescence imaging of cells grown in shaking culture

For comparison of *rarA*+ and Δ*rarA* cell morphologies and effects on labeled replisomes, indicated strains were grown overnight at 37°C in EZ glycerol, diluted 1000-fold in fresh medium, then grown an additional 3 hrs prior to imaging.

For imaging, glass coverslips were either functionalized with APTES (Alfa Aesar) according to the procedure detailed previously (*26*) or cleaned by sonicating 5M KOH for 25 min, rinsing with fresh MilliQ water, and drying under N_2 gas. Coverslip sandwiches were prepared by flattening 20 μL of cell suspension between an APTES-functionalized (bottom) and KOH-cleaned (top) cover slide. Cells were then imaged in both brightfield and fluorescence channels on the microscope as described above. Image processing was again performed using custom ImageJ plugins after manually outlining cells in the MicrobeTracker Suite in MATLAB R2013b (Mathworks).

Flow cytometry

Overnight cultures of indicated strains were diluted 1:100 in fresh LB medium in biological triplicate from three separate bacterial colonies. Cultures were grown at 37°C with shaking and aeration until the OD_{600} measured 0.2 (exponential growth phase). Aliquots (500 μL) were taken from each culture and placed on ice. Cells were collected via centrifugation and washed with 1 mL of phosphate buffered saline (PBS) three times. These aliquots were diluted 1:10 in fresh PBS, then either mock stained or stained with the DNA-specific dye Vybrant Dyecycle Green (Thermo Fisher Scientific, Waltham, MA) according to manufacturer's instructions (final dye concentration = 10 μM).

Samples were measured using a BD Accuri Flow Cytometer (BD Biosciences, San Jose, CA) equipped with a 488 nm wavelength excitation laser and 533 nm wavelength emission filter with a 30 nm bandwidth. A sample threshold of 10,000 FSC-A was set. At least 50,000 cells were measured for each biological replicate, with three total biological replicates measured per strain. Light scattering and fluorescence data were imported into FlowJo (FlowJo LLC, Ashland, OR) and gated as indicated to exclude cell debris and doublet cells.

Spot plate drug/UV/overexpression sensitivity assays

Overnight cultures of indicated strains were diluted 1:100 in fresh LB medium. Cultures were grown in biological triplicate at 37°C with aeration and shaking until the OD_{600} reached 0.2 (exponential growth phase). Aliquots (1 mL) were taken from each culture, serially diluted in LB to 10^{-6}, then 10 μL were spot plated on agar plates containing LB medium supplemented with indicated concentrations of nitrofurazone, glucose, or L-arabinose. When UV irradiation was used, plates were exposed to indicated doses of UV irradiation (254 nm) using a Spectrolinker XL-1500 UV crosslinker (Spectronics Corporation, Westbury, NY). Plates were incubated overnight at 37°C and imaged the next day using an ImageQuant LAS 4010 imaging system (GE Healthcare, Marlborough, MA).

Overexpression toxicity time-lapse cell viability curves

Overnight cultures of indicated strains were diluted 1:100 in fresh LB medium. Cultures were grown with aeration and shaking at 37°C until the OD_{600} measured 0.2 (exponential growth phase). L-arabinose was then added to the culture to a final concentration of 0.2%. 1 mL aliquots were taken from each culture at 0, 45, 90, and 180 minutes. The OD_{600} values were measured, then the aliquots were normalized to the OD_{600} value of the culture prior to addition of arabinose. After normalization, the cultures were serially diluted in phosphate buffered saline (PBS) to an appropriate concentration and 100 μL was plated on LB agar plates containing 100 μg/mL ampicillin and 0.2% glucose. The plates were incubated overnight at 37°C. Colonies were counted and colony forming units (CFUs) were calculated the following day.

Protein purification and western blot

Wild type RarA, RarA K63R, and RarA K430E mutant proteins were purified as described previously (*56*). All proteins were carefully tested for endo- and exonuclease contamination using

gel-based DNA degradation assays utilized supercoiled and linear double-stranded DNA and circular single-stranded DNA. No contaminating endo- or exonucleases were detected. Aliquots of purified proteins were thawed fresh from -80°C stock solutions prior to each experiment. RarA protein concentration was determined using the native extinction coefficient $\varepsilon = 5.44 \times 10^4\ M^{-1}\ cm^{-1}$ (*56*).

Cultures of indicated strains were grown as described in the time-lapse cell viability curve section above. After normalization, the cultures were centrifuged to pellet cellular material. Pellets were resuspended in 20 µL 2x Lamelli sample buffer (250 mM Tris-HCl pH 6.8, 4% SDS, 20% (*v*/*v*) glycerol, 1.43M betamercaptoethanol, 1 mg/mL bromophenol blue). Samples were boiled at 95°C for 5 minutes. Purified protein samples were made by mixing a protein solution of known concentration with 2x Lamelli sample buffer.

Samples (15 µL) were applied to a 4-20% gradient tris-glycine SDS-PAGE gel (Bio-Rad) and electrophoresed at 100 V for 30 min, then 200 V until completed. Gels were loaded into a western blot sandwich consisting of sponge, filter paper, gel, 0.45 µm PVDF (polyvinylidene fluoride) membrane (GE Healthcare), filter paper, and sponge. Samples were transferred from the SDS-PAGE gel to the PVDF membrane at 300 mA for one hour.

PVDF membranes were blocked using a solution of powdered milk in PBS-Tween 20 (1x phosphate buffered saline, 0.5% Tween-20, 2.5% (*w*/*v*) powdered milk) for one hour. Primary chicken antibody raised against *rarA* (Gene-Tel; Madison, WI) was diluted 1:5000 in blocking solution and applied to the PVDF membrane for one hour. Membranes were rinsed three times in PBS-T, then washed twice for 5 min, then once for 15 min. Secondary antibody (HRP-conjugated rabbit anti-chicken; Sigma) was diluted 1:10000 in PBS-Tween 20 and applied to the membrane for one hour. The membrane was then washed as before. SuperSignal West Pico chemiluminescent

substrate (ThermoFisher) was applied to the membrane and incubated at room temperature for 5 min. The membrane was then imaged using an ImageQuant 4010 Instrument (GE Healthcare) with an exposure time of 10 seconds.

Strain	Relevant Genotype	Parent Strain	Source/Technique
MG1655	*rarA+ recA+ recO+ recF+ dinB+ polB+ umuDC+*	-	George Weinstock
EAW98	*ΔrarA::Kan*	MG1655	Lambda RED recombination
EAW214	Δ(*araBAD ParaB*)::Kan	MG1655	Lambda RED recombination
THS100	Δ*rarA* ::FRT* Δ(*araBAD ParaB*)::Kan	EAW98	Transduction of EAW98 with P1 grown on EAW214
EAW170	*dnaQ::dnaQ-YPet Kan*	MG1655	Lambda RED recombination
THS04	Δ*rarA* ::FRT* *dnaQ* ::*dnaQ* -YPet Kan	EAW98	Transduction of EAW98 with P1 grown on EAW170
JJC5945	*dnaX* ::*dnaX* -YPet Kan	MG1655	Bénédicte Michel
EAW85	*recO-::* Kan	MG1655	Lambda RED recombination
EAW573	Δ*rarA* ::FRT* *recO-* ::Kan	EAW98	Transduction of EAW98 with P1 grown on EAW85
EAW90	*recF* -::Kan	MG1655	Lambda RED recombination
THS130	Δ*rarA* ::FRT* *recF-::* Kan	EAW98	Transduction of EAW98 with P1 grown on EAW90
EAW18	Δ*dinB* ::Kan	MG1655	Lambda RED recombination
THS13	Δ*rarA* ::FRT* Δ*dinB* ::Kan	EAW98	Transduction of EAW98 with P1 grown on EAW18
EAW644	*rarA::rarA* (K63R) Kan	MG1655	Lambda RED recombination
THS122	*rarA::rarA* (K63R) FRT Δ*dinB* ::Kan	EAW644	Transduction of EAW644 with P1 grown on EAW18
EAW21	Δ*polB* ::Kan	MG1655	Lambda RED recombination
THS124	*rarA::* FRT* Δ*polB* ::Kan	EAW98	Transduction of EAW98 with P1 grown on THS124
THS125	*rarA::rarA* (K63R) FRT* *ΔpolB::Kan*	EAW644	Transduction of EAW644 with P1 grown on EAW21
EAW296	*umuDC::* Kan	MG1655	Lambda RED recombination
SF2018	ZK126 *polB* ::Spc *dinB* ::Kan *umuDC* ::Cam	ZK126	Myron Goodman
THS10	Δ*rarA* ::Kan *umuDC* ::Cam	EAW98	Transduction of EAW98 with P1 grown on SF2018
THS126	*umuDC* ::FRT* *rarA::rarA* (K63R) Kan	EAW296	Transduction of EAW296 with P1 grown on EAW644
THS26	Δ*rarA* ::Kan + pBAD/myc-His A Nde	EAW98	Transformation of EAW98 with pBAD/myc-His A Nde plasmid
THS27	Δ*rarA* ::Kan + pBAD/myc-His A Nde(*rarA*)	EAW98	Transformation of EAW98 with pTHS02 plasmid
THS35	Δ*rarA* ::FRT* Δ*recA* ::Kan + pBAD/myc-His A Nde	EAW572	Transformation of EAW572 with pBAD/myc-His A Nde plasmid
THS36	Δ*rarA* ::FRT* Δ*recA* ::Kan + pBAD/myc-His A Nde(*rarA*)	EAW572	Transformation of EAW572 with pTHS02 plasmid
THS28	Δ*rarA* ::Kan + pBAD/myc-His A Nde(*rarA* K63R)	EAW98	Transformation of EAW98 with pTHS03 plasmid
THS37	Δ*rarA* ::FRT* Δ*recA* ::Kan + pBAD/myc-His A Nde(*rarA* K63R)	EAW572	Transformation of EAW572 with pTHS03 plasmid
THS32	Δ*rarA* ::Kan + pBAD/myc-His A Nde(*rarA* K430E)	EAW98	Transformation of EAW98 with pTHS08 plasmid
THS42	Δ*rarA* ::FRT* Δ*recA* ::Kan + pBAD/myc-His A Nde(*rarA* K430E)	EAW572	Transformation of EAW572 with pTHS08 plasmid
THS119	Δ*rarA* ::Kan + pBAD/myc-His A Nde(*SbcB15*)	EAW98	Transformation of EAW98 with pTHS10 plasmid
THS55	Δ*rarA* ::FRT* *dnaQ::dnaQ* -YPet / pBAD	EAW98	Transformation of EAW98 with pBAD/myc-HisA plasmid
THS56	Δ*rarA* ::FRT* *dnaQ::dnaQ* -YPet / pBAD(*wt rarA*)	EAW98	Transformation of EAW98 with pTHS02 plasmid
THS68	Δ*rarA* ::Kan Δ*umuDC* ::Cam + pBAD/myc-His A Nde(*rarA*)	THS10	Transformation of THS10 with pTHS02 plasmid
THS71	Δ*rarA* ::FRT* Δ*dinB* ::Kan + pBAD/myc-His A Nde(*rarA*)	THS13	Transformation of THS13 with pTHS02
THS79	Δ*rarA* ::FRT* Δ*dinB* ::Kan Δ*umuDC* ::Cam + pBAD/myc-His A Nde(*rarA*)	THS14	Transformation of THS14 with pTHS02

Table 3-1: A list of strains used in this study

A list of strains used in this chapter is shown above. All strains were constructed as described in the experimental procedures section.

Figures

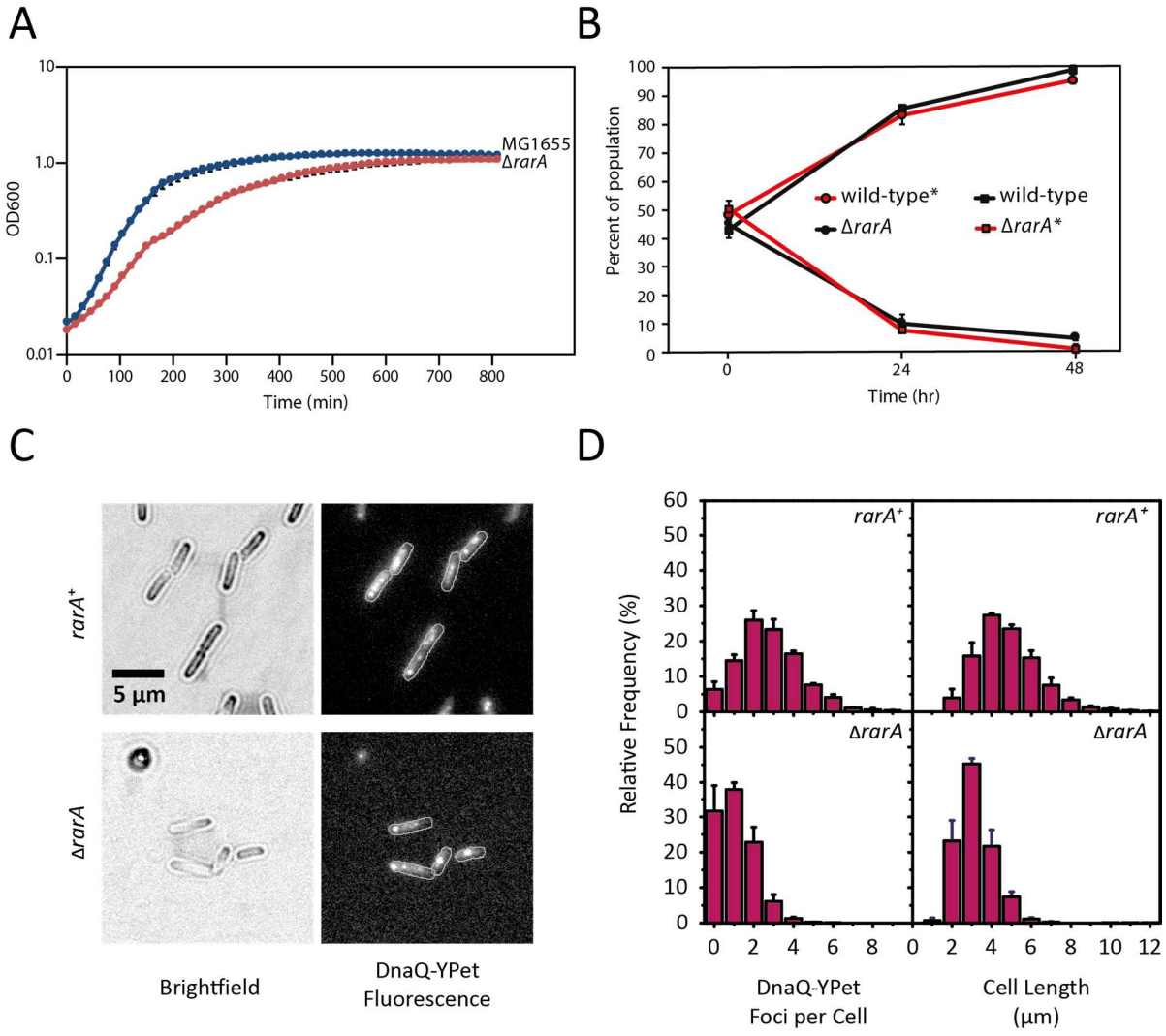

A
OD600
Time (min)
MG1655
ΔrarA
B
Percent of population
Time (hr)
wild-type*
ΔrarA
wild-type
ΔrarA*
C
rarA+
ΔrarA
5 µm
Brightfield
DnaQ-YPet
Fluorescence
D
Relative Frequency (%)
rarA+
ΔrarA
DnaQ-YPet
Foci per Cell
Cell Length
(µm)

Figure 3-1: Strains lacking *rarA* exhibit a growth defect compared to wild type MG1655 cells, exhibit smaller cell size and contain a reduced number of replisome foci.

(**A**) Overnight cultures were grown in LB medium, diluted 100-fold and allowed to grow for 1000 min. OD_{600} values were recorded over this time period. Traces represent OD_{600} values averaged over a minimum of three biological replicates. Doubling times for the wild type and Δ*rarA* strains were 29.2 ± 0.7 and 42.3 ± 0.5 min, respectively. (**B**) Using a growth competition assay, equal amounts of wild type MG1655 and Δ*rarA* strains are incubated together at $t = 0$ h. This co-culture was allowed to grow for 48 h, with samples taken at 0, 24, and 48 h. These samples were serially diluted and plated onto tetrazolium agar plates. Deletion of the *araBAD* operon acts as a marker (*) and is able to be differentiated from *araBAD*+ cells. Colonies representing each genotypic population were counted and divided by the total number of colonies to determine the percentage of the total population each strain inhabited. These experiments were conducted in triplicate, with error bars representing one S.D. from the mean. (**C***) Single-molecule fluorescence imaging of *rarA*+ (top) and Δ*rarA* (bottoms) strains containing DnaQ-YPet labeled DNA polymerase III. Cells were grown at 37°C in flow cells and imaged every 5 min for 180 min. (**D***) Histograms of DnaQ-YPet replisome foci per cell (left) and cell length (right) measurements for the *rarA*+ (top) and Δ*rarA* (bottom) strains. Error bars represent the SEM for each bin across at least two biological replicates. DnaQ-YPet foci per cell for *rarA*+ cells: mean = 2.8; SEM = 0.03; $n = 2738$ cells. DnaQ-YPet foci per cell for Δ*rarA* cells: mean = 1.3; SEM = 0.03; $n = 1424$ cells. Cell size for *rarA*+ cells: mean = 5.5; SEM = 0.03 μm in length; $n = 2892$ cells. Cell size for Δ*rarA* cells: mean = 3.7; SEM = 0.03 μm in length; $n = 1660$ cells.

* This work was conducted in collaboration with M. Cherry, A. Robinson, and A. van Oijen at the University of Wollongong in Wollongong, Australia.

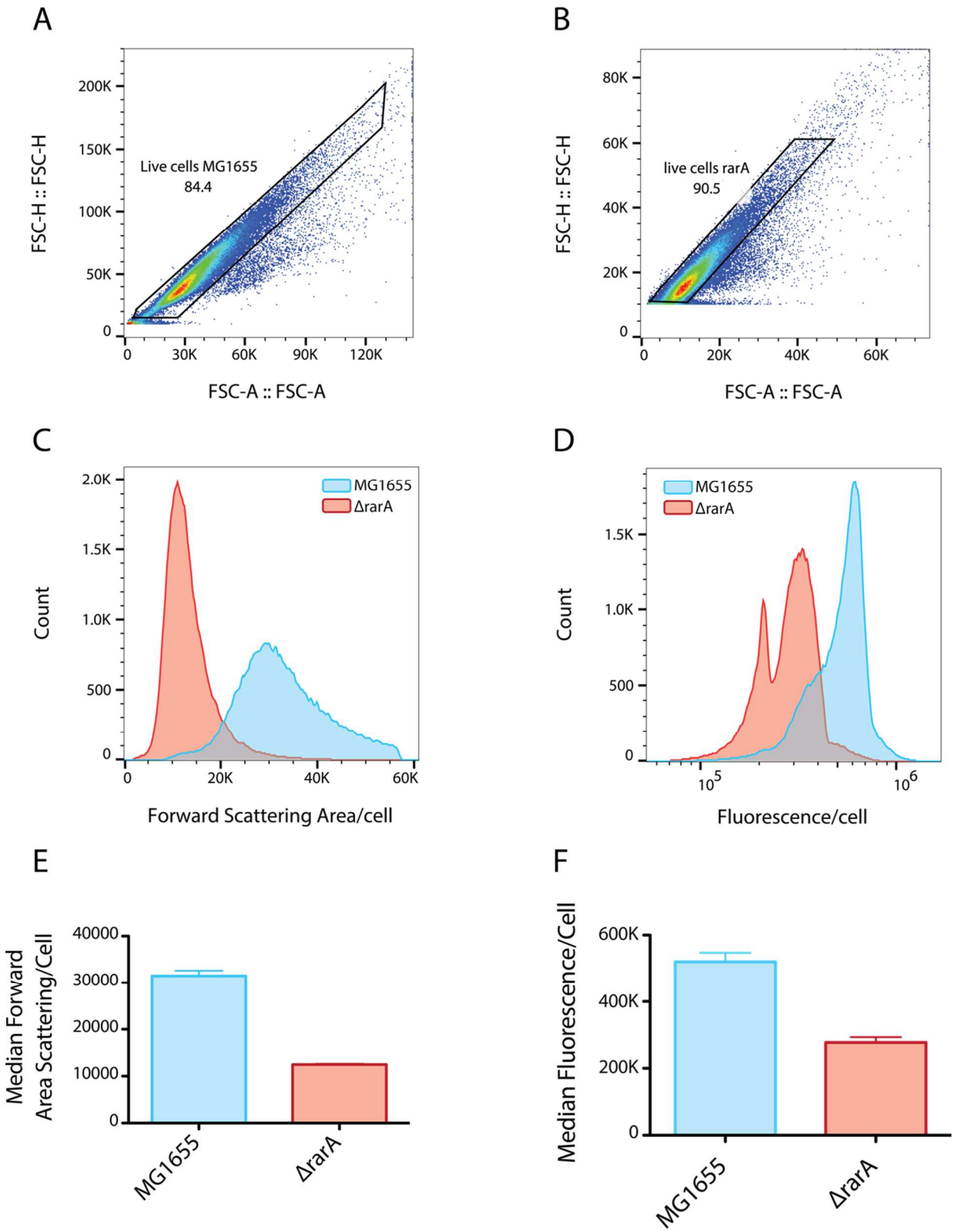

A
FSC-H :: FSC-H
FSC-A :: FSC-A
Live cells MG1655
84.4
200K
150K
100K
50K
0
0
30K
60K
90K
120K
B
FSC-H :: FSC-H
FSC-A :: FSC-A
live cells rarA
90.5
80K
60K
40K
20K
0
0
20K
40K
60K
C
Count
Forward Scattering Area/cell
MG1655
ΔrarA
2.0K
1.5K
1.0K
500
0
0
20K
40K
60K
D
Count
Fluorescence/cell
MG1655
ΔrarA
1.5K
1.0K
500
0
10^5
10^6
E
Median Forward Area Scattering/Cell
40000
30000
20000
10000
0
MG1655
ΔrarA
F
Median Fluorescence/Cell
600K
400K
200K
0
MG1655
ΔrarA

Figure 3-2: Δ*rarA* cells are smaller and contain less DNA than wild type cells during exponential growth phase

Wild type (MG1655) or Δ*rarA* cells were grown to early exponential growth phase (OD_{600} = 0.2), then diluted to an appropriate concentration (~1 x 10^6 cells/mL) and washed with PBS. Cells were not synchronized. DNA content was assessed using Vybrant Dyecycle Green as a stain. Each strain was grown up and assayed in biological triplicate, with > 50,000 cells being counted for each replicate. Wild type (**A**) and Δ*rarA* (**B**) cells were gated as indicated to exclude cell debris and doublet cells. 84.4% (wild type) to 90.5% (Δ*rarA*) of cells were included in further analysis. (**C**) Forward scattering area histograms of representative MG1655 (blue) and Δ*rarA* (red) samples are shown. (**D**) Fluorescence signal representing DNA content for representative MG1655 (blue) and Δ*rarA* (red) samples are shown. (**E**) Quantification of the forward scattering area data graphed in (C) is shown for MG1655 (blue) and Δ*rarA* (red) samples. Bar graphs represent the average of the media forward scattering area values across biological replicates, while error bars represent one S.D. from the averaged median value. (**F**) Quantification of the fluorescence data graphed in (d) is shown for MG1655 (blue) and Δ*rarA* (red) samples. Bar graphs represent the average of the median fluorescence values across biological replicates, while error bars represent one S.D. from the averaged median value.

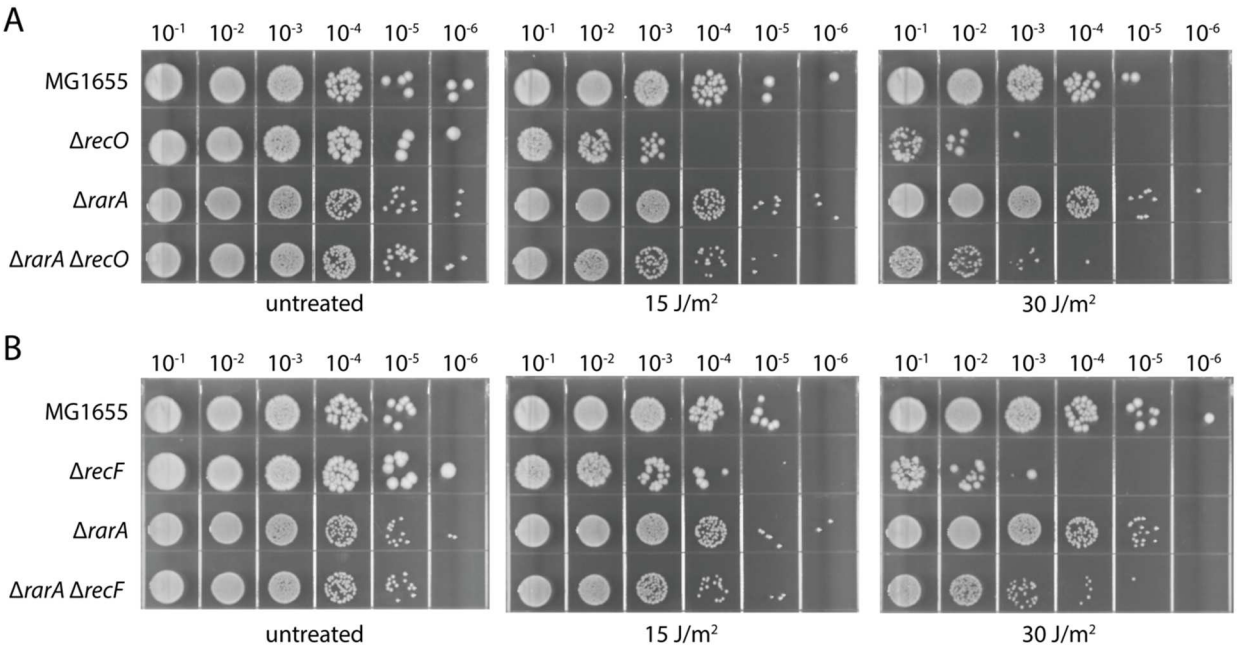

A
10-1 10-2 10-3 10-4 10-5 10-6
MG1655
ΔrecO
ΔrarA
ΔrarA ΔrecO
untreated
15 J/m2
30 J/m2
B
10-1 10-2 10-3 10-4 10-5 10-6
MG1655
ΔrecF
ΔrarA
ΔrarA ΔrecF
untreated
15 J/m2
30 J/m2

Figure 3-3: Deletion of *rarA* suppresses the sensitivity of Δ*recF* and Δ*recO* stains to UV light

Indicated strains were grown to exponential growth phase (OD_{600} = 0.2), serially diluted, spot plated onto LB agar plates, and irradiated at a dose of 15 or 30 J/m^2. (**A**) Δ*recO* cells are sensitized to low levels of UV light by 2 orders of magnitude compared to wild type cells. Deletion of *rarA* results in no observable decrease in cell viability and completely (at 15 J/m^2) or partially (at 30 J/m^2) restores cell viability to wild type levels in a Δ*recO* background. (**B**) Δ*recF* cells are sensitized to low levels of UV light by 1-2 orders of magnitude compared to wild type cells. Deletion of *rarA* in a Δ*recF* background completely (at 15 J/m^2) or partially (at 30 J/m^2) restores cell viability to wild type levels. All experiments were repeated at least three times using separate biological replicates, yielding consistent results.

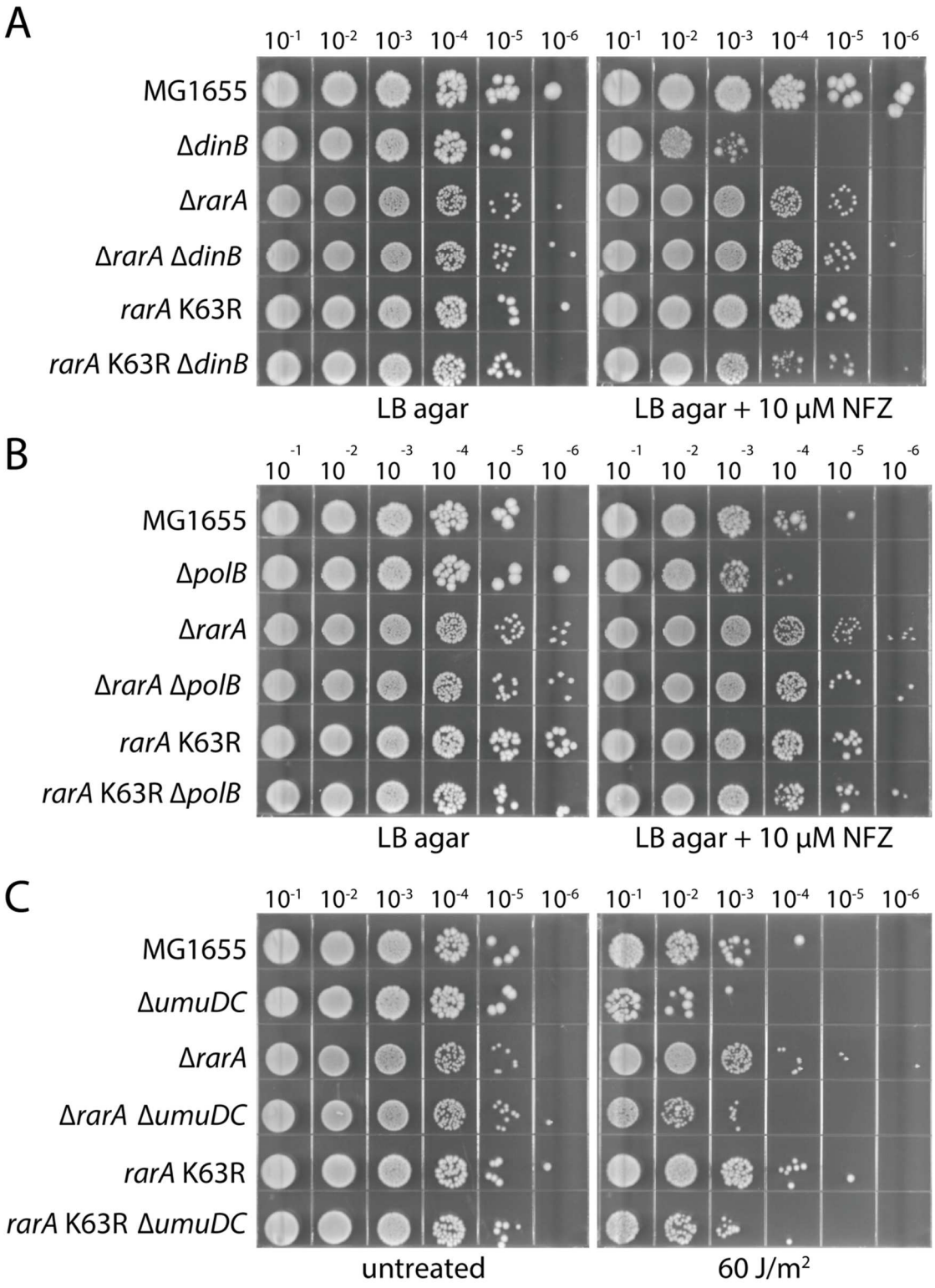

A
10-1 10-2 10-3 10-4 10-5 10-6
MG1655
ΔdinB
ΔrarA
ΔrarA ΔdinB
rarA K63R
rarA K63R ΔdinB
LB agar
LB agar + 10 μM NFZ
B
10-1 10-2 10-3 10-4 10-5 10-6
MG1655
ΔpolB
ΔrarA
ΔrarA ΔpolB
rarA K63R
rarA K63R ΔpolB
LB agar
LB agar + 10 μM NFZ
C
10-1 10-2 10-3 10-4 10-5 10-6
MG1655
ΔumuDC
ΔrarA
ΔrarA ΔumuDC
rarA K63R
rarA K63R ΔumuDC
untreated
60 J/m2

Figure 3-4: Loss of *rarA* ATPase activity suppresses the sensitivity of TLS polymerase knockout strains Δ*polB*, Δ*dinB*, and Δ*umuDC* to DNA damage

Indicated strains were grown to exponential phase (OD_{600} = 0.2), serially diluted and spot plated onto LB agar plates containing indicated supplements. (**A**) Loss of RarA ATPase activity suppresses the sensitivity of Δ*dinB* (Pol IV) cells to the mutagen nitrofurazone (NFZ). Cells lacking the *dinB* gene are sensitized to NFZ by 2-3 orders of magnitude compared to *dinB*+ cells. Loss of *rarA* or its ATPase activity (K63R) confers no sensitivity to NFZ. In a Δ*dinB* background, loss of *rarA* or its ATPase activity restores cell viability to wild type levels. (**B**) Loss of RarA ATPase activity suppresses the sensitivity of Δ*pol*B (Pol II) cells to NFZ. Cells lacking *polB* are mildly sensitized to NFZ by 1-2 orders of magnitude compared to wild type cells. In a Δ*polB* background, loss of *rarA* or its ATPase activity restores cell viability to wild type levels. (**C**) Loss of RarA ATPase activity partially suppresses the sensitivity of Δ*umuDC* (Pol V) mutants to ultraviolet light. Cells lacking Pol V are sensitized to high levels (60 J/m^2) of UV light compared to wild type cells. Deletion of *rarA* or inactivation of its ATPase activity (K63R) results in a 1 order of magnitude increase in resistance of UV light at this dose when compared to wild type cells. Deletion of *rarA* or inactivation of its ATPase activity in a Δ*umuDC* background restores cell viability to wild type levels at this incident dose of UV light. All experiments were repeated at least three times with consistent results.

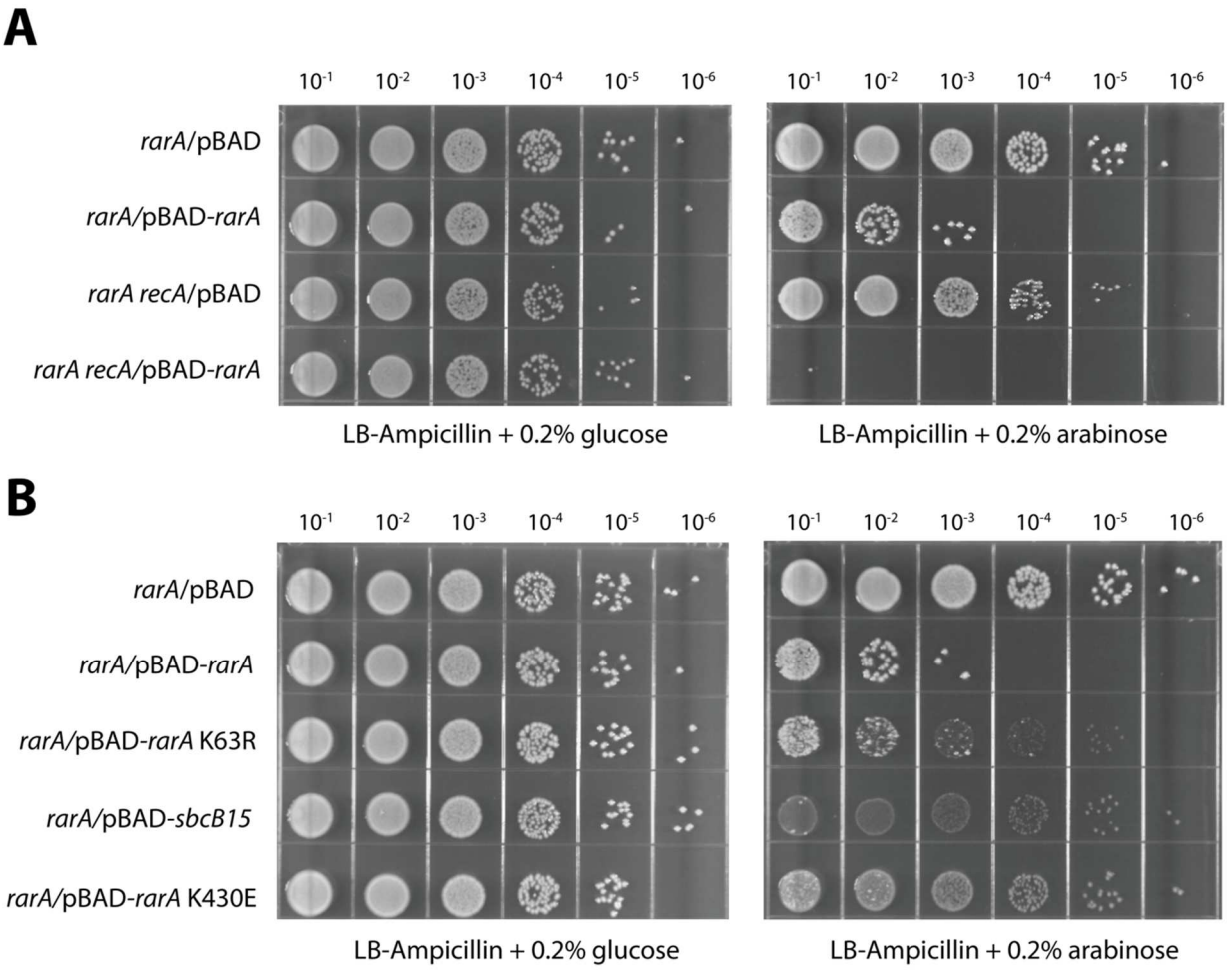
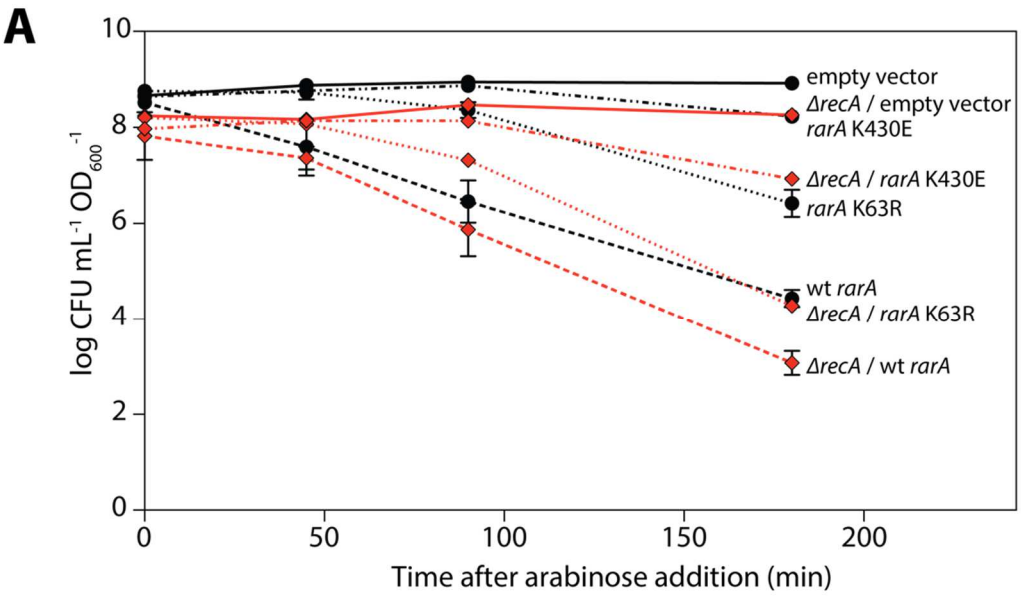

A
10-1 10-2 10-3 10-4 10-5 10-6
rarA/pBAD
rarA/pBAD-rarA
rarA recA/pBAD
rarA recA/pBAD-rarA
LB-Ampicillin + 0.2% glucose
LB-Ampicillin + 0.2% arabinose
B
10-1 10-2 10-3 10-4 10-5 10-6
rarA/pBAD
rarA/pBAD-rarA
rarA/pBAD-rarA K63R
rarA/pBAD-sbcB15
rarA/pBAD-rarA K430E
LB-Ampicillin + 0.2% glucose
LB-Ampicillin + 0.2% arabinose

Figure 3-5: Overexpression of *rarA* is lethal in both the presence and absence of *recA*

(**A**) All strains were Δ*rarA* while only those indicated were also Δ*recA*. Overexpression of *rarA* was induced in exponential growth phase cells (OD_{600} = 0.2) upon plating on LB agar containing L-arabinose (0.2%), while expression was repressed upon plating on LB agar containing 0.2% glucose. Plates are representative images of at least three separate experiments. (**B**) Δ*rarA* cells harboring the indicated gene on a pBAD plasmid were grown and spot plated as in (A). RarA K63R is a Walker A mutant that is incapable of hydrolyzing ATP, but can still bind SSB. SbCB15 is a nuclease-dead mutant of exonuclease I that is still capable of binding SSB. RarA K430E is a tetramerization-deficient mutant that is incapable of hydrolyzing ATP or binding SSB. Plates are representative images of at least three separate experiments.


A
10
8
6
4
2
0
log CFU mL-1 OD600-1
0
50
100
150
200
Time after arabinose addition (min)
empty vector
ΔrecA / empty vector
rarA K430E
ΔrecA / rarA K430E
rarA K63R
wt rarA
ΔrecA / rarA K63R
ΔrecA / wt rarA

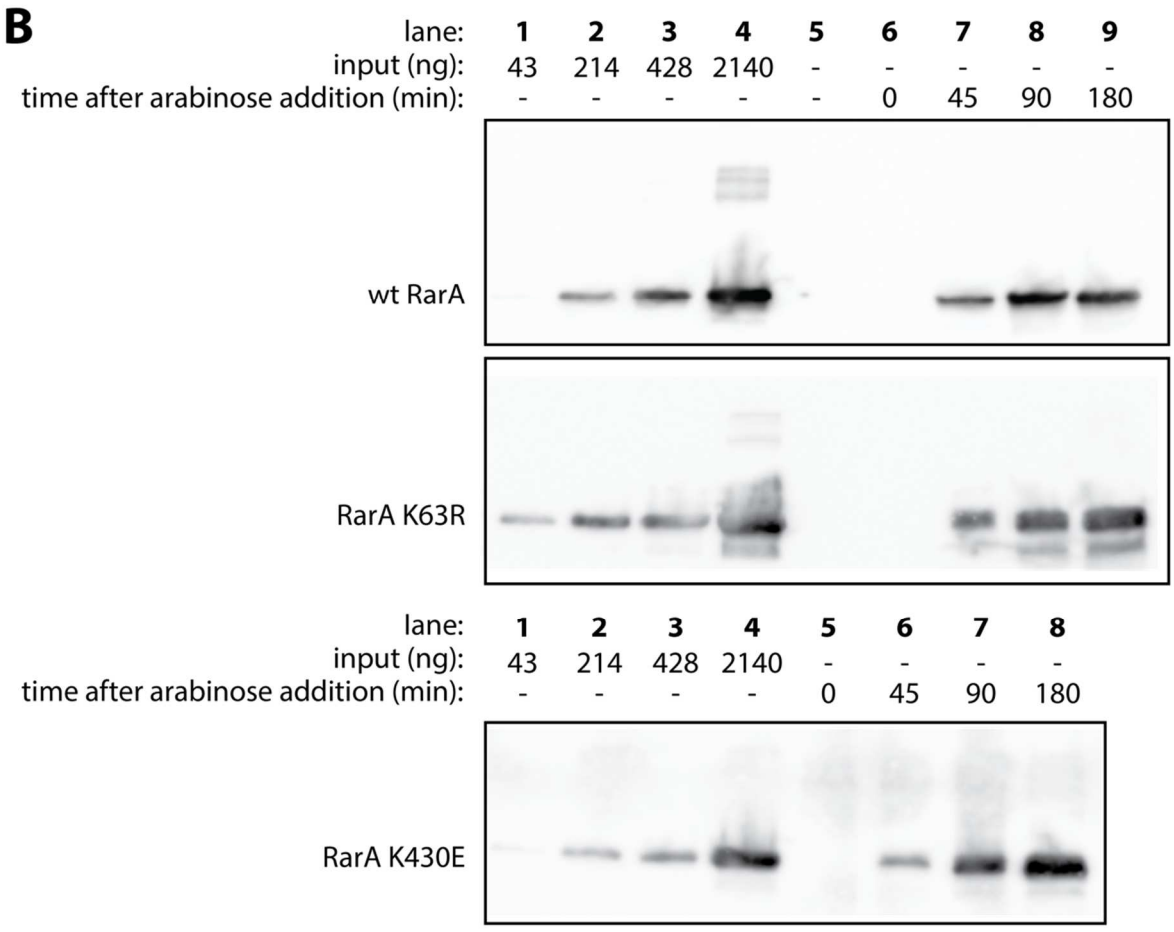
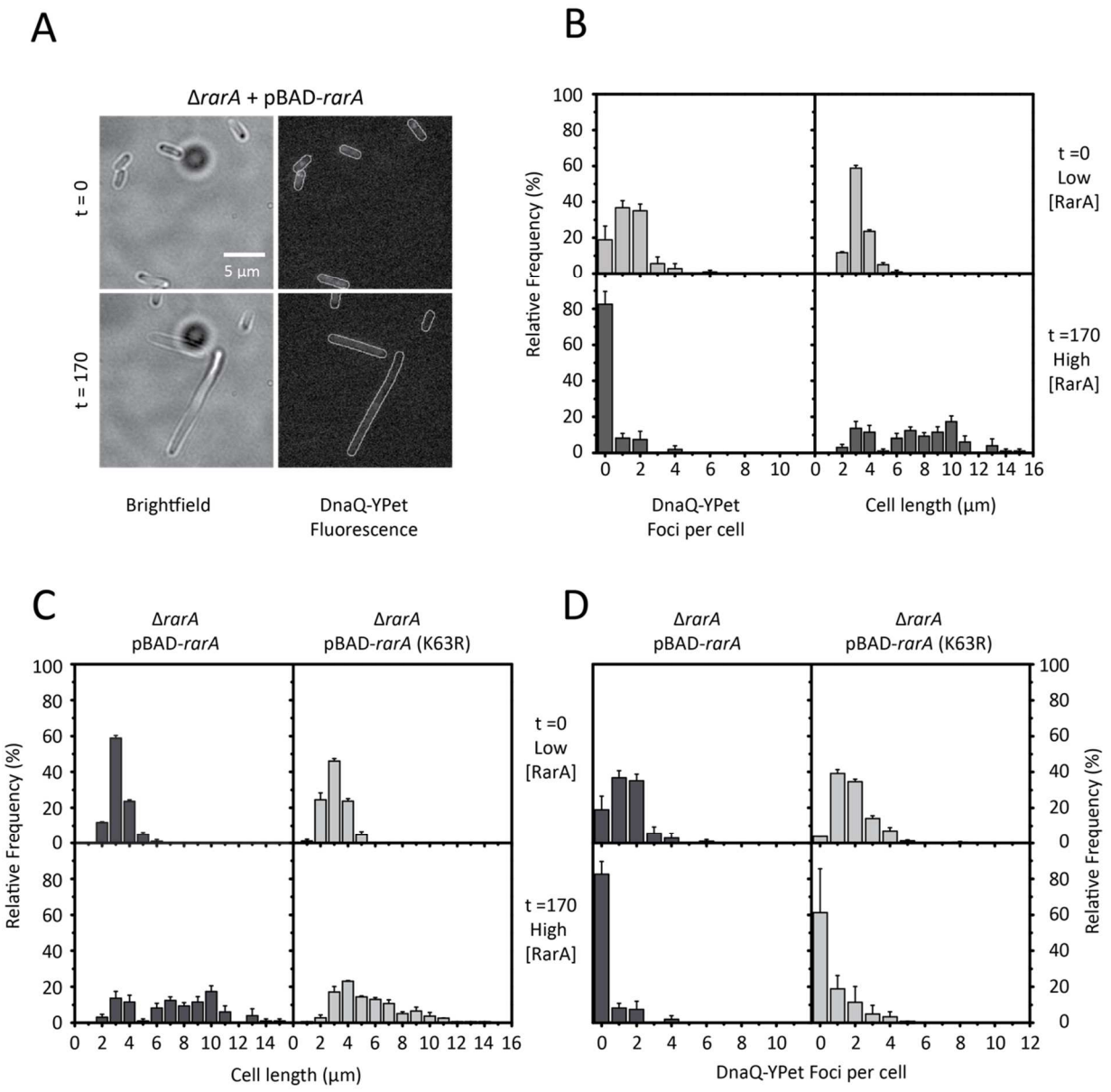

B
lane: 1 2 3 4 5 6 7 8 9
input (ng): 43 214 428 2140 - - - - -
time after arabinose addition (min): - - - - - 0 45 90 180
wt RarA
RarA K63R
lane: 1 2 3 4 5 6 7 8
input (ng): 43 214 428 2140 - - - -
time after arabinose addition (min): - - - - 0 45 90 180
RarA K430E

Figure 3-6: Overexpression of *rarA* is toxic in both *recA+ and* Δ*recA* cells after 180 minutes

(**A**) Δ*rarA* cells (black lines) or Δ*rarA* Δ*recA* cells (red lines) harboring the indicated gene on a pBAD expression vector were grown at 37°C in LB medium to exponential phase, after which arabinose was added to a final concentration of 0.2%. Samples were then collected at indicated time points and plated on LB medium + 0.2% glucose agar to calculate colony forming units (CFUs). (**B**) Relative expression levels were determined for wild type RarA, RarA K63R, and RarA K430E in Δ*rarA* strains after addition of L-arabinose in cells grown up in (A). Samples containing known concentrations of purified RarA or RarA mutants were loaded alongside experimental samples containing cell lysate with overexpressed RarA or RarA mutants. Experiments were performed at least three times, with representative western blots for each mutant shown.



* This work was conducted in collaboration with M. Cherry, A. Robinson, and A. van Oijen at the University of Wollongong in Wollongong, Australia.

Figure 3-7: Overexpression of *rarA* results in cell filamentation and delocalization of the replisome in an ATPase-independent manner

(**A**) Time-lapse imaging of a Δ*rarA* pBAD-*rarA* strain containing labeled DNA polymerase III (DnaQ-YPet). Cells were grown at 37°C in flow cells and imaged every 5 min for 180 mins. Overexpression was initially suppressed using 0.2% glucose then induced with 0.2% L-arabinose at $t = 0$ min. Representative brightfield (left) and fluorescence (right) micrographs are shown at $t = 0$ min (top) and $t = 180$ min (bottom) following induction of *rarA* overexpression. (**B**) Histograms of DnaQ-YPet replisome foci per cell (left) and cell length (right) are shown for the indicated time points. Error bars represent the standard error of the mean values for each bin across at least two biological replicates. DnaQ-YPet foci per cell at $t = 0$ min: mean = 1.4; SEM = 0.1; $n = 103$ cells. Cell size at $t = 0$ min: mean = 3.2; SEM = 0.07 μm; $n = 102$ cells. DnaQ-YPet foci per cell at $t = 170$ min: mean = 0.3; SEM = 0.1; $n = 80$ cells. Cell size at $t = 170$ min: mean = 7.2; SEM = 0.34 μm; $n = 80$ cells. (**C, D**) Overexpression of ATPase-deficient RarA K63R induces similar effects on cell filamentation and replisome focus loss compared to wild type RarA. Cells were grown as in (A), except that pBAD-*rarA* K63R was overexpressed instead of wild type RarA. Histograms of cell size (left) and DnaQ-YPet foci per cell (right) are shown comparing overexpression of wild type RarA (dark gray) and RarA K63R (light gray). Cell size for Δ*rarA* pBAD-*rarA* K63R at $t = 0$ min: mean: 3.0; SEM = 0.04 μm; $n = 432$ cells. DnaQ-YPet foci per cell at $t = 0$ min: mean: 1.8; SEM = 0.1; $n = 433$ cells. Cell size at $t = 170$ min: mean = 5.6; SEM = 0.12 μm; $n = 357$ cells. DnaQ-YPet foci per cell at $t = 170$ min: mean = 0.5; SEM = 0.1; $n = 357$ cells.

* This work was conducted in collaboration with M. Cherry, A. Robinson, and A. van Oijen at the University of Wollongong in Wollongong, Australia.

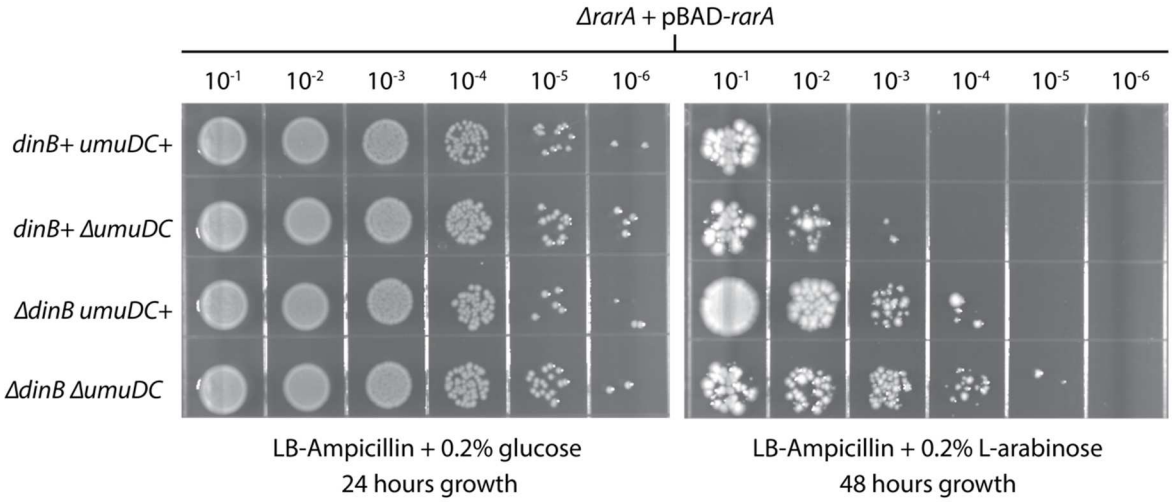

ΔrarA + pBAD-rarA
10-1 10-2 10-3 10-4 10-5 10-6
10-1 10-2 10-3 10-4 10-5 10-6
dinB+ umuDC+
dinB+ ΔumuDC
ΔdinB umuDC+
ΔdinB ΔumuDC
LB-Ampicillin + 0.2% glucose
24 hours growth
LB-Ampicillin + 0.2% L-arabinose
48 hours growth

Figure 3-8: Deletion of TLS polymerases Pol IV (Δ*dinB*) and Pol V (Δ*umuDC*) suppress the lethality of *rarA* overexpression

Cultures of Δ*rarA* cells containing indicated TLS polymerase mutations and harboring pBAD-*rarA* plasmids were growth to early exponential growth phase (OD_{600} = 0.2), serial diluted, and spot plated onto LB agar containing either 0.2% glucose or 0.2% L-arabinose. As before, overexpression was induced using 0.2% L-arabinose, while expression was repressed in the presence of 0.2% glucose. Plates are representative images from at least three independent biological replicate experiments.

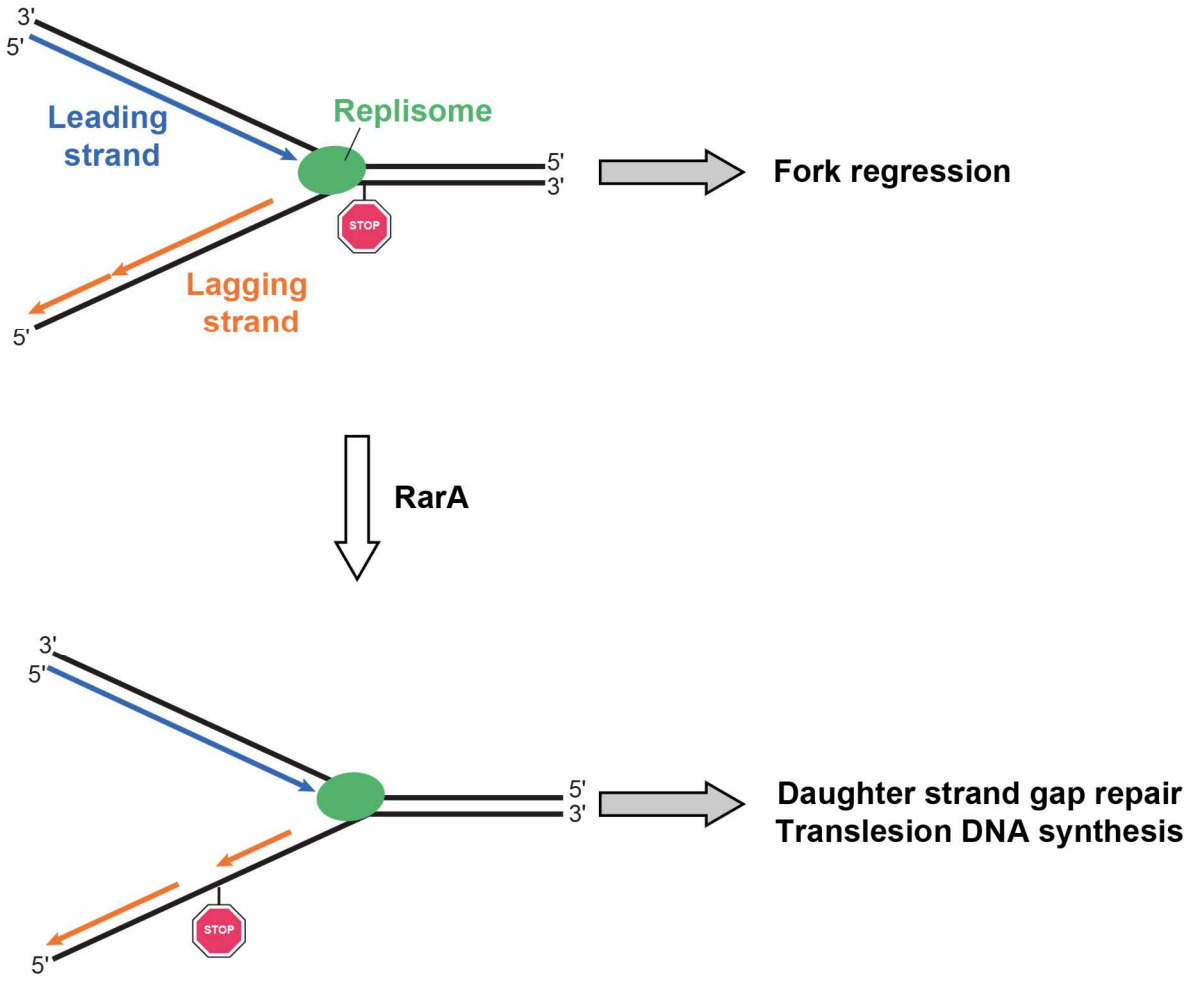

3'
5'
Leading strand
Replisome
5'
3'
STOP
Fork regression
Lagging strand
5'
RarA
3'
5'
5'
3'
Daughter strand gap repair
Translesion DNA synthesis
STOP
5'

Figure 3-9: A model for RarA-mediated gap creation behind replication forks *in vivo*

A simplified model depicting a replication fork encountering a DNA lesion on the lagging strand. RarA would displace the active Pol III core from its processivity clamp to generate a single-stranded DNA gap behind the replication fork. This single-stranded DNA gap can serve as a template for RecFOR-mediated daughter-strand gap repair or translesion DNA synthesis. In the absence of RarA, replication forks can be remodeled and the damage repaired directly using time-intensive pathways such as fork regression.

References

1. M. M. Cox, Historical overview: searching for replication help in all of the rec places. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8173–8180 (2001).

2. M. M. Cox *et al.*, The importance of repairing stalled replication forks. *Nature*. **404**, 37–41 (2000).

3. S. C. Kowalczykowski, Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem. Sci.* **25** (2000), pp. 156–165.

4. A. Kuzminov, Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* **63**, 751–813 (1999).

5. A. Kuzminov, DNA replication meets genetic exchange: Chromosomal damage and its repair by homologous recombination. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8461–8468 (2001).

6. B. Michel, Replication fork arrest and DNA recombination. *Trends Biochem. Sci.* **25** (2000), pp. 173–178.

7. B. Michel, H. Boubakri, Z. Baharoglu, M. LeMasson, R. Lestini, Recombination proteins and rescue of arrested replication forks. *DNA Repair (Amst).* **6**, 967–980 (2007).

8. R. C. Heller, K. J. Marians, Replisome assembly and the direct restart of stalled replication forks. *Nat. Rev. Mol. Cell Biol.* **7**, 932–43 (2006).

9. H. L. Klein, K. N. Kreuzer, Replication, Recombination, and Repair: Going for the Gold. *Mol. Cell*. **9**, 471–480 (2002).

10. M. Lopes *et al.*, The DNA replication checkpoint response stabilizes stalled replication forks. *Nature*. **412**, 557–561 (2001).

11. H. Merrikh, Y. Zhang, A. D. Grossman, J. D. Wang, Replication–transcription conflicts in bacteria. *Nat. Rev. Microbiol.* **10**, 449–458 (2012).

12. A. Kuzminov, Collapse and repair of replication forks in *Escherichia coli*. *Mol. Microbiol.* **16**, 373–384 (1995).

13. J. D. McCool, C. C. Ford, S. J. Sandler, A dnaT mutant with phenotypes similar to those of a priA2::kan mutant in *Escherichia coli* K-12. *Genetics*. **167**, 569–578 (2004).

14. B. Michel *et al.*, Rescue of arrested replication forks by homologous recombination. *Proc. Natl. Acad. Sci.* **98**, 8181–8188 (2001).

15. B. Michel, G. Grompone, M.-J. Florès, V. Bidnenko, Multiple pathways process stalled replication forks. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12783–8 (2004).

16. A. H. Syeda, M. Hawkins, P. McGlynn, Recombination and replication. *Cold Spring Harb. Perspect. Biol.* **6**, a016550 (2014).

17. S. M. Mangiameli, C. N. Merrikh, P. A. Wiggins, H. Merrikh, Transcription leads to pervasive replisome instability in bacteria. *Elife*. **6**, 1–27 (2017).

18. J. Courcelle, B. M. Wendel, D. D. Livingstone, C. T. Courcelle, RecBCD is required to complete chromosomal replication: Implications for double-strand break frequencies and repair mechanisms. *DNA Repair (Amst).* **32**, 86–95 (2015).

19. S. Lusetti, M. M. Cox, The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Ann. Rev. Biochem*. **71**, 71–100 (2002).

20. E. V Mirkin, S. M. Mirkin, Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.* **71**, 13–35 (2007).

21. A. Aguilera, T. García-Muse, Causes of Genome Instability. *Annu. Rev. Genet.* **47**, 1–32 (2013).

22. B. Michel, After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol.* **3**, 1174–1176 (2005).

23. M. Radman, SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sci*, 355–67 (1975).

24. A. J. Gruber *et al.*, A RecA Protein Surface Required for Activation of DNA Polymerase V. *PLoS Genet.* **11**, 1–37 (2015).

25. Q. Jiang, K. Karata, R. Woodgate, M. M. Cox, M. F. Goodman, The active form of DNA polymerase V is UmuD′2C–RecA–ATP. *Nature*. **460**, 359–363 (2009).

26. A. Robinson *et al.*, Regulation of Mutagenic DNA Polymerase V Activation in Space and Time. *PLoS Genet.* **11**, 1–30 (2015).

27. R. P. Fuchs, Tolerance of lesions in *E. coli*: Chronological competition between Translesion Synthesis and Damage Avoidance. *DNA Repair (Amst).* **44**, 51–58 (2016).

28. R. P. Fuchs, S. Fujii, Translesion DNA Synthesis and Mutagenesis in Prokaryotes. *Cold Spring Harb. Lab. Press*. **5**, 1–22 (2013).

29. C. B. Gabbai, J. T. P. Yeeles, K. J. Marians, Replisome-mediated translesion synthesis and leading strand template lesion skipping are competing bypass mechanisms. *J. Biol. Chem.* **289**, 32811–32823 (2014).

30. M. Ikeda *et al.*, DNA polymerase IV mediates efficient and quick recovery of replication forks stalled at N2-dG adducts. *Nucleic Acids Res.* **42**, 8461–8472 (2014).

31. H. A. Jeiranian, B. J. Schalow, C. T. Courcelle, J. Courcelle, Fate of the replisome

following arrest by UV-induced DNA damage in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **110**, 11421–11426 (2013).

32. J. E. Kath *et al.*, Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. *Proc. Natl. Acad. Sci.* **111**, 7647–7652 (2014).

33. J. E. Kath *et al.*, Exchange between *Escherichia coli* polymerases II and III on a processivity clamp. *Nucleic Acids Res.* **44**, 1681–1690 (2015).

34. L. M. Margara, M. M. Fernández, E. L. Malchiodi, C. E. Argaraña, M. R. Monti, MutS regulates access of the error-prone DNA polymerase Pol IV to replication sites: A novel mechanism for maintaining replication fidelity. *Nucleic Acids Res.* **44**, 7700–7713 (2016).

35. M. Jaszczur *et al.*, Mutations for Worse or Better: Low-Fidelity DNA Synthesis by SOS DNA Polymerase V Is a Tightly Regulated Double-Edged Sword. *Biochemistry*. **55**, 2309–2318 (2016).

36. M. F. Goodman, Better living with hyper-mutation. *Environ. Mol. Mutagen.* **57**, 421–434 (2016).

37. M. F. Goodman, R. Woodgate, Translesion DNA polymerases. *Cold Spring Harb. Perspect. Biol.* **5**, 247–250 (2013).

38. K. Naiman, V. Pagès, R. P. Fuchs, A defect in homologous recombination leads to increased translesion synthesis in *E. coli*. *Nucleic Acids Res.* **44**, 7691–7699 (2016).

39. C. A. Bonner, S. Hays, K. McEntee, M. F. Goodman, DNA polymerase II is encoded by the DNA damage-inducible dinA gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7663–7 (1990).

40. M. Escarceller *et al.*, Involvement of *Escherichia coli* DNA Polymerase II in Response to Oxidative Damage and Adaptive Mutation. *J. Bacteriol.* **10**, 6221–6228 (1994).

41. H. Iwasaki, A. Nakata, G. C. Walker, H. Shinagawa, The *Escherichia coli polB* gene, which encodes DNA polymerase II, is regulated by the SOS system. *J. Bacteriol.* **172**, 6268–73 (1990).

42. Z. Qiu, M. F. Goodman, The *Escherichia coli* polB locus is identical to dinA, the structural gene for DNA polymerase II: Characterization of pol II purified from A polB mutant. *J. Biol. Chem.* **272**, 8611–8617 (1997).

43. S. R. Kim, K. Matsui, M. Yamada, P. Gruz, T. Nohmi, Roles of chromosomal and episomal dinB genes encoding DNA pol IV in targeted and untargeted mutagenesis in *Escherichia coli*. *Mol. Genet. Genomics*. **266**, 207–215 (2001).

44. R. P. Fuchs, S. Fujii, J. Wagner, Properties and functions of *Escherichia coli*: Pol IV and Pol V. *Adv. Protein Chem.* **69**, 229–264 (2004).

45. C. T. Courcelle, J. J. Belle, J. Courcelle, Nucleotide excision repair or polymerase V-mediated lesion bypass can act to restore UV-arrested replication forks in *Escherichia coli*. *J. Bacteriol.* **187**, 6953–6961 (2005).

46. I. Bjedov *et al.*, Involvement of *Escherichia coli* DNA polymerase IV in tolerance of cytotoxic alkylating DNA lesions in vivo. *Genetics*. **176**, 1431–1440 (2007).

47. D. F. Jarosz, V. G. Godoy, J. C. Delaney, J. M. Essigmann, G. C. Walker, A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature*. **439**, 225–228 (2006).

48. D. F. Jarosz, V. G. Godoy, G. C. Walker, Proficient and accurate bypass of persistent DNA lesions by DinB DNA polymerases. *Cell Cycle*. **6**, 817–822 (2007).

49. D. F. Jarosz, S. E. Cohen, J. C. Delaney, J. M. Essigmann, G. C. Walker, A DinB variant reveals diverse physiological consequences of incomplete TLS extension by a Y-family DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21137–42 (2009).

50. K. R. Ona, C. T. Courcelle, J. Courcelle, Nucleotide excision repair is a predominant mechanism for processing nitrofurazone-induced DNA damage in *Escherichia coli*. *J. Bacteriol.* **191**, 4959–4965 (2009).

51. A. B. Williams, K. M. Hetrick, P. L. Foster, Interplay of DNA repair, homologous recombination, and DNA polymerases in resistance to the DNA damaging agent 4-nitroquinoline-1-oxide in *Escherichia coli*. *DNA Repair (Amst).* **9**, 1090–1097 (2010).

52. A. Gutierrez, M. Elez, O. Clermont, E. Denamur, I. Matic, *Escherichia coli* YafP protein modulates DNA damaging property of the nitroaromatic compounds. *Nucleic Acids Res.* **39**, 4192–4201 (2011).

53. F.-X. Barre *et al.*, Circles: the replication-recombination-chromosome segregation connection. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8189–8195 (2001).

54. A. Johnson, M. O'Donnell, CELLULAR DNA REPLICASES: Components and Dynamics at the Replication Fork. *Annu. Rev. Biochem.* **74**, 283–315 (2005).

55. A. Costes, F. Lecointe, S. McGovern, S. Quevillon-Cheruel, P. Polard, The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS Genet.* **6**, 1–15 (2010).

56. A. N. Page, N. P. George, A. H. Marceau, M. M. Cox, J. L. Keck, Structure and biochemical activities of *Escherichia coli* MgsA. *J. Biol. Chem.* **286**, 12075–12085 (2011).

57. I. F. Lau *et al.*, Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* **49**, 731–743 (2003).

58. D. J. Sherratt *et al.*, Recombination and chromosome segregation. *Philos. Trans. R. Soc.*

Lond. B. Biol. Sci. **359**, 61–69 (2004).

59. R. A. Bish, M. P. Myers, Werner helicase-interacting protein 1 binds polyubiquitin via its zinc finger domain. *J. Biol. Chem.* **282**, 23184–23193 (2007).

60. N. Crosetto *et al.*, Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J. Biol. Chem.* **283**, 35173–35185 (2008).

61. I. Saugar, J. L. Parker, S. Zhao, H. D. Ulrich, The genome maintenance factor Mgs1 is targeted to sites of replication stress by ubiquitylated PCNA. *Nucleic Acids Res.* **40**, 245–257 (2012).

62. J. P. Erzberger, J. M. Berger, Evolutionary Relationships and Structural Mechanisms of Aaa+ Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).

63. D. Vandewiele, A. R. Fernández de Henestrosa, A. R. Timms, B. A. Bridges, R. Woodgate, Sequence analysis and phenotypes of five temperature sensitive mutator alleles of dnaE, encoding modified α-catalytic subunits of *Escherichia coli* DNA polymerase III holoenzyme. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **499**, 85–95 (2002).

64. T. Shibata *et al.*, Functional overlap between RecA and MgsA (RarA) in the rescue of stalled replication forks in *Escherichia coli*. *Genes to Cells*. **10**, 181–191 (2005).

65. R. Lestini, B. Michel, UvrD controls the access of recombination proteins to blocked replication forks. *EMBO J.* **26**, 3804–14 (2007).

66. B. Michel, A. K. Sinha, The inactivation of *rfaP* , *rarA* or *sspA* gene improves the viability of the *Escherichia coli* DNA polymerase III *holD* mutant. *Mol. Microbiol.* **104**, 1008–1026 (2017).

67. A. Yoshimura, M. Seki, T. Enomoto, The role of WRNIP1 in genome maintenance. *Cell Cycle*. **16**, 515–521 (2017).

68. D. Branzei, M. Seki, F. Onoda, T. Enomoto, The product of *Saccharomyces cerevisiae* WHIP/MGS1, a gene related to replication factor C genes, interacts functionally with DNA polymerase Delta. *Mol. Genet. Genomics*. **268**, 371–386 (2002).

69. T. Hishida, H. Iwasaki, T. Ohno, T. Morishita, H. Shinagawa, A yeast gene, MGS1, encoding a DNA-dependent AAA(+) ATPase is required to maintain genome stability. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8283–9 (2001).

70. J. H. Kim *et al.*, In vivo and in vitro studies of Mgs1 suggest a link between genome instability and Okazaki fragment processing. *Nucleic Acids Res.* **33**, 6137–6150 (2005).

71. T. Tsurimoto, A. Shinozaki, M. Yano, M. Seki, T. Enomoto, Human Werner helicase interacting protein 1 (WRNIP1) functions as novel modulator for DNA polymerase delta. *Genes to Cells*. **10**, 13–22 (2005).

72. T. H. Stanage, A. N. Page, M. M. Cox, DNA flap creation by the RarA/MgsA protein of *Escherichia coli*. *Nucleic Acids Res.* **45**, 2724–2735 (2017).

73. D. Branzei *et al.*, Characterization of the slow-growth phenotype of *S. cerevisiae* whip/mgs1 sgs1 double deletion mutants. *DNA Repair (Amst).* **1**, 671–682 (2002).

74. T. Hayashi *et al.*, Vertebrate WRNIP1 and BLM are required for efficient maintenance of genome stability. *Genes Genet. Syst.* **83**, 95–100 (2008).

75. T. Hishida, T. Ohno, *Saccharomyces cerevisiae* MGS1 is essential in strains deficient in the RAD6 -dependent DNA damage tolerance pathway. *EMBO J.* **21**, 2019–2029 (2002).

76. N. D. Vijeh Motlagh, M. Seki, D. Branzei, T. Enomoto, Mgs1 and Rad18/Rad5/Mms2 are required for survival of *Saccharomyces cerevisiae* mutants with novel temperature/cold sensitive alleles of the DNA polymerase delta subunit, Pol31. *DNA Repair (Amst).* **5**, 1459–1474 (2006).

77. A. Yoshimura *et al.*, Functional relationships between Rad18 and WRNIP1 in vertebrate cells. *Biol. Pharm. Bull.* **29**, 2192–6 (2006).

78. A. Yoshimura *et al.*, Physical and functional interaction between WRNIP1 and RAD18. *Genes Genet. Syst.* **84**, 171–178 (2009).

79. R. E. Lenski, M. R. Rose, S. C. Simpson, S. C. Tadler, Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *Am. Nat.* **138**, 1315–1341 (1991).

80. R. T. Byrne *et al.*, Evolution of extreme resistance to ionizing radiation via genetic adaptation of DNA repair. *Elife*. **2014**, 1–18 (2014).

81. M. J. Wiser, R. E. Lenski, A comparison of methods to measure fitness in *Escherichia coli*. *PLoS One*, 1–11 (2015).

82. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell*. **133**, 90–102 (2008).

83. J. S. Lewis *et al.*, Single-molecule visualization of fast polymerase turnover in the bacterial replisome. *Elife*. **6**, 1–17 (2017).

84. K. C. Smith, T. C. V. Wang, R. C. Sharma, recA-Dependent DNA repair in UV-irradiated *Escherichia coli*. *J. Photochem. Photobiol. B Biol.* **1**, 1–11 (1987).

85. Yu-Chin Tseng, Jai-Li Hung, T. C. V Wang, Involvement of RecF pathway recombination genes in postreplication repair in UV-irradiated *Escherichia coli* cells. *Mutat. Res. Repair*. **315**, 1–9 (1994).

86. T.-C. Wang, K. C. Smith, Mechanisms for recF-dependent and recB-dependent pathways of postreplication repair in UV-irradiated *Escherichia coli* uvrB. *J. Bacteriol.* **156**, 1093–

1098 (1983).

87. M. M. Cox, Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* **42**, 41–63 (2007).

88. L. H. Sanders, A. Rockel, H. Lu, D. J. Wozniak, M. D. Sutton, Role of Pseudomonas aeruginosa dinB-encoded DNA polymerase IV in mutagenesis. *J. Bacteriol.* **188**, 8573–8585 (2006).

89. B. Yuan, H. Cao, Y. Jiang, H. Hong, Y. Wang, Efficient and accurate bypass of N2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase in vitro and in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8679–84 (2008).

90. S. R. Kushner, H. Nagaishi, a Templin, a J. Clark, Genetic recombination in *Escherichia coli*: the role of exonuclease I. *Proc. Natl. Acad. Sci. U. S. A.* **68**, 824–827 (1971).

91. D. Lu, J. L. Keck, Structural basis of *Escherichia coli* single-stranded DNA-binding protein stimulation of exonuclease I. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9169–74 (2008).

92. J. Wagner *et al.*, The dinB gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol. Cell*. **4**, 281–286 (1999).

93. K. Uchida *et al.*, Overproduction of *Escherichia coli* DNA polymerase DinB (Pol IV) inhibits replication fork progression and is lethal. *Mol. Microbiol.* **70**, 608–622 (2008).

94. D. Széliová, J. Krahulec, M. Šafránek, V. Lišková, J. Turňa, Modulation of heterologous expression from PBAD promoter in *Escherichia coli* production strains. *J. Biotechnol.* **236**, 1–9 (2016).

95. O. J. Becherel, R. P. P. Fuchs, J. Wagner, Pivotal role of the beta-clamp in translesion DNA synthesis and mutagenesis in *E. coli* cells. *DNA Repair (Amst).* **1**, 703–708 (2002).

96. S. Fujii, V. Gasser, R. P. Fuchs, The biochemical requirements of DNA polymerase V-mediated translesion synthesis revisited. *J. Mol. Biol.* **341**, 405–417 (2004).

97. W. D. Rupp, P. Howard-Flanders, Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J. Mol. Biol.* **31**, 291–304 (1968).

98. V. Pages, R. P. Fuchs, Uncoupling of Leading- and Lagging-Strand DNA Replication During Lesion Bypass in Vivo. *Science (80-.).* **300**, 1300–1303 (2003).

99. M. Lopes, M. Foiani, J. M. Sogo, Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell*. **21**, 15–27 (2006).

100. R. C. Heller, K. J. Marians, Replication fork reactivation downstream of a blocked nascent leading strand. *Nature*. **439**, 557–562 (2006).

101. J. T. P. Yeeles, K. J. Marians, Dynamics of leading-strand lesion skipping by the replisome. *Mol. Cell*. **52**, 855–865 (2013).

102. K. Higuchi *et al.*, Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication in vitro. *Genes to Cells*. **8**, 437–449 (2003).

103. P. McInerney, M. O'Donnell, Functional uncoupling of twin polymerases: Mechanism of polymerase dissociation from a lagging-strand block. *J. Biol. Chem.* **279**, 21543–21551 (2004).

104. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–5 (2000).

105. L. C. Huang, E. A. Wood, M. M. Cox, Convenient and reversible site-specific targeting of exogenous DNA into a bacterial chromosome by use of the FLP recombinase: The FLIRT system. *J. Bacteriol.* **179**, 6076–6083 (1997).

Chapter 4: Conclusion

Conclusion

The primary goal of this thesis was to further characterize the highly enigmatic RarA protein family using the *E. coli* RarA protein as a model. Previous work over the past 16 years by over a dozen research groups placed the function of the RarA protein family at the interface between DNA replication and repair. In this thesis, we utilized a host of biochemical, biophysical, and genetic assays to deepen our understanding not only of the RarA protein family, but of DNA replication as a whole. Our main conclusion was that RarA ensures optimal replication fork progression upon encountering DNA damage by catalyzing a process called lesion skipping.

Previous work defined two major pathways for repairing or tolerating DNA damage encountered by active replication forks in *E. coli*. For the past three decades, work in our laboratory has focused on one of these pathways—recombinational DNA repair at the fork—while just recently beginning to focus more on the second of these pathways—translesion DNA synthesis (TLS) and daughter-strand gap repair. While both of these pathways are well-characterized, a major gap in knowledge remains: how does the cell choose which repair pathway to utilize? Conventional knowledge pointed to the SOS response, which contains an inherent temporal regulation of genes involved in both pathways. Transcription of genes involved in homologous recombination and nonmutagenic repair of stalled replication forks were derepressed prior to genes involved in mutagenic TLS. Induction of the SOS response does explain the recruitment of regulated proteins to sites of DNA damage preventing replication fork progression (damage at or ahead of the fork). However, the SOS response fails to explain why certain TLS Pols—Pol II and Pol IV—are present throughout normal cell growth at significant concentrations (50-250 molecules per cell) even in the absence of DNA damage. This observation would suggest that TLS Pols have a role during a normal cycle of DNA replication. In addition, single-stranded DNA

gaps—substrates repaired/tolerated by TLS and daughter-strand gap repair—are formed following UV irradiation despite their source not being completely understood. We hypothesize that these gaps form in an active process to prevent time-intensive repair of lesions at or ahead of the fork. Instead, replisome progression continues while the gaps are simultaneously repaired by post-replication repair pathways. We have presented data here that point to the highly conserved RarA protein family as a probable candidate for catalyzing such an activity in cells.

In chapter 2, we detailed the ATPase and DNA binding characteristics of the *E. coli* RarA protein. Our work clearly shows that RarA recognizes damaged DNA substrates—including double-stranded DNA ends and single-stranded DNA gaps—in an ATP-dependent manner. In accordance with its sequence similarity to the *E. coli* Holliday Junction helicase RuvB, we also demonstrated RarA possesses a flap-formation activity that is also ATP-dependent. We have yet to determine an *in vivo* role for this flap-formation activity. Finally, we demonstrated that RarA creates single-stranded DNA gaps—the substrate required for both daughter-strand gap repair and TLS—behind active replication forks *in vitro*. We hypothesized that RarA-mediated gap formation proceeds by outcompeting active DNA polymerase III cores with their associated processivity clamp during lagging strand DNA synthesis. We directly observed fluorescently-labeled processivity clamps behind left behind in these single-stranded gaps as would be expected in our model. These are ideal substrates for both daughter-strand gap repair and TLS.

Despite these novel findings *in vitro*, many questions remain to directly link them to an *in vivo* pathway or function. First, our single-molecule rolling-circle assay was conducted in the absence of any DNA damage or replication stress. Thus, RarA action in this assay could be the result of an excess concentration compared to the single molecules of actively replicating DNA. Despite our efforts to match the concentration of RarA in our assays with those measured *in vivo*,

it is unclear how the setup of our assay may require an excess of RarA molecules to visualize an effect in the absence of DNA damage. Further work should be conducted to repeat these experiments in the presence of site-specific DNA damage in the template DNA. Observing RarA preferentially catalyze gap formation in the presence of DNA damage would begin to link these results to those described later *in vivo*. Second, a direct interaction between RarA and beta processivity clamp has yet to be demonstrated either *in vitro* or *in vivo*. Our model for gap creation *in vitro* relies entirely on a hypothesized protein-protein interaction between RarA and beta clamp; thus, this should be tested directly. Finally, the flap-forming activity of RarA has yet to be connected to any process *in vivo* or the gap creation process. Our model relies on RarA disrupting a protein-protein interaction between DNA polymerase III and its processivity clamp to generate a single-stranded DNA gap *in vitro*. If this is true, then no DNA-dependent ATPase or flap creation activity would be required for gap creation. However, since these gaps only form in the presence of ATPase-competent RarA, the two observations must be connected. Further investigation of the connection between RarA ATPase and gap creation activity is required.

While no direct connection to DNA damage was established in our observation of RarA-mediated gap creation *in vitro*, several insights into a role for gap creation in DNA repair pathway choice *in vivo* were detailed in chapter 3. If we view the observations in chapter 3 through the lens of the *in vitro* gap creation activity described in chapter 2, a clear role for RarA in the cell is established. Since single-stranded DNA gaps are shared DNA substrates between daughter-strand gap repair and translesion DNA synthesis, our hypothesis was that RarA creates substrates for these processes *in vivo*. Indeed, cells lacking either RecFOR-mediated daughter-strand gap repair or TLS were sensitized to DNA damage in a RarA-dependent manner. If RarA action (or simply ATPase activity) was eliminated in these same strains, these cells were no longer sensitive to these

genotoxic agents. The simplest interpretation of these data is that RarA action creates substrates that are processed by either daughter-strand gap repair or TLS, depending on the type of DNA damage encountered by the replication fork. However, if the substrate is created in a strain lacking the downstream repair pathways, the cells fail to survive. In light of the *in vitro* gap creation activity, these two observations do not appear to be a coincidence. One of the most important experiments required in the future is to directly observe these gaps form *in vivo* in a RarA-dependent manner.

If our model is correct, loss of *rarA* function would result in a loss of commitment to daughter-strand gap repair/TLS in favor of repair processes acting at or ahead of the fork. In these cells, DNA replication and repair would no longer be able to be decoupled; thus, all damage encountered by the fork would require repair prior to restarting replication. This could explain the growth, cell size, and DNA content defects we observed in strains lacking *rarA* documented in chapter 3. Conversely, we showed that overexpression of *rarA* resulted in cell filamentation, toxicity and ejection of replisomes from the genome. This complex phenotype was largely independent of RarA ATPase activity, suggesting a larger role for protein-protein interactions as causative agents. However, since overexpression of wild type *rarA* caused the most severe effect, a link between its toxicity and excess gap creation cannot be ruled out. This hypothesis is corroborated by our data linking *rarA* overexpression toxicity with the action of TLS Pols.

Overall, the major contribution of the work detailed in this thesis is the formulation of a new model of DNA repair pathway choice during DNA replication. Like all models, it requires extensive further investigation to disprove the aspects that are not completely correct. In this particular case, proposing a novel model arises with both the raw characteristics lacking detail and the high expectations of proof. This is the major limitation of this thesis—through this work, we

have attempted to answer a singular question with a singular answer, only to generate further questions with each additional answer. In a sense, this limitation is also a testament to the longevity of this project. This is the burden I must leave with future researchers with fresh perspectives and novel approaches.