DNA Polymerase Exchange and Lesion Bypass in Escherichia Coli

DNA polymerase exchange and lesion bypass in *Escherichia coli*

A dissertation presented

by

James Evon Kath

to

The Committee on Higher Degrees in Biophysics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biophysics

Harvard University

Cambridge, Massachusetts

October 2015

Dissertation Advisor: Professor Joseph J. Loparo James Evon Kath

DNA polymerase exchange and lesion bypass in *Escherichia coli*

Abstract

Translesion synthesis (TLS) alleviates replication stalling at DNA lesions. Bypass of lesions by specialized translesion DNA polymerases involves exchange with high-fidelity replicative polymerases. As a consequence of their lesion bypass activity, TLS polymerases are mutagenic, requiring careful regulation of polymerase selection. In this dissertation, I describe a single-molecule reconstitution of polymerase exchange and lesion bypass. Using *Escherichia coli* polymerases as a model system, I have determined that the dimeric processivity clamp can simultaneously bind a replicative polymerase and a translesion polymerase, facilitating rapid exchange during synthesis and lesion bypass. Overlapping sets of polymerase-clamp interactions additionally allow the TLS polymerase Polymerase IV to displace the replicative polymerase Polymerase III. I finally describe the observation of single Polymerase IV molecules in living cells and initial efforts to determine their localization and dynamics during TLS.

Contents

Acknowledgments

First and foremost, I would like to thank my advisor, Joe Loparo. I first met Joe at the Biophysics BBQ in August 2010, when he was a new professor with more microscopes than lab members. Watching him establish and build his lab was as important a training experience as any of my research projects, and it was a pleasure to be a part of that. I deeply appreciate Joe's sense of perspective on how to move research forward, and how to work efficiently while maintaining priorities and balance. He also can align a laser faster than anyone I know.

Even if you have a great advisor, it's always a bit of a gamble to rotate in a new lab, since you don't really know what its culture will become. I am grateful for the other "Loparettes" who contribute the great science and good attitudes that make our lab a warm, collaborative place – Hyeongjun, Jacob, Linda, Seungwoo, Lizz, Sankalp, and the two honorary lab members, Wendy and Allen. I am especially indebted to Thomas Graham, who shares the distinction as Joe's "first student" with me. Thomas is a talented and humble scientist, and I am a stronger experimentalist from working with him. I also would be remiss to not mention all the Walter lab members who have welcomed us as their extended lab. They have taught me about the important of scotch, as well as the difference between "Sure" and "Walter Sure".

I have also received critical mentorship throughout my graduate studies from outside my lab. Graham Walker was kind enough to welcome me as a rotation student and has provided important feedback and encouragement through the years. His postdoc, Jamie Foti, was patient enough to work with me during that

rotation, even though I came in having never cloned a gene or purified a protein. Johannes Walter served on my two of my committees and has offered invaluable suggestions and advice on research, papers, and grants. I also would also like to thank Wesley Wong, Sam Reck-Peterson, Matt Waldor, and John Mekalanos for serving on my committees and providing helpful feedback.

A special thanks is reserved for the "heart and soul" of the Biophysics Program, Jim Hogle and Michele Jakoulov. They have been advocates on behalf of myself and many other students, and have gently offered correction when needed. I don't think I can say anything about them that hasn't been said by others, except: after my graduate school visits in 2009, I wrote a list of pros and cons comparing Harvard and an unnamed West Coast school and their names were weighty additions that helped tip the scales to my eventual choice.

Graduate school has been an incredible period in my life, and I am grateful for all the students I have shared it with: my Biophysics classmates and fellow participants in the weekly "George Xu lunch club;" members of the Cold Spring Harbor Advanced Bacterial Genetics course; and Perkins dormmates and fellow members of our weekly journal club, the Advanced Society for the Advancement of the Advanced Societal Advancements. There's nothing better than leaving after a long day at lab to go drink PBR and talk about science (or whatever).

Finally, I would like to thank my parents, Bonnie and Bill, and my sister, Karen, for their support. Most importantly, I want to thank my fiancée Kate for all her love. I eagerly look forward to our life together.

Chapter 1
Introduction

1.1 DNA replication in *Escherichia coli*

If the dream of every cell is to become two cells, as François Jacob quipped, then a completed duplication of its genome is the dream made reality. Although the means of regulating DNA replication varies across the domains of life, the core replication machinery, or replisome, shares remarkable structural and functional commonalities. A helicase hydrolyzes ATP to separate the DNA duplex into its two strands; a primase initializes the duplication of each strand by synthesizing a short primer, typically from ribonucleotides; and a DNA polymerase extends the primer by sequentially adding deoxyribonucleotides in an order complementary to the template sequence. As the two DNA strands run in antiparallel directions, copying of both strands is accomplished by the continuous synthesis of one, the leading strand, and discontinuous synthesis of the other, the lagging strand, in "Okazaki fragments."

Non-catalytic accessory proteins contribute to the stability and speed of the replisome; in most organisms the DNA polymerase's processivity – the number of consecutive nucleotide additions per binding event – is increased from single digits to thousands or more by its binding to a processivity factor, or sliding clamp. This ring-shaped protein is loaded by a multi-protein clamp loader complex onto DNA at the end of the primer, where it encircles the double helix. The processivity factor is therefore topologically constrained to slide along DNA until it is unloaded, and specific interactions with the DNA polymerase in turn increase its stability as it

synthesizes DNA at the primer terminus. Finally, single-stranded DNA binding proteins play a critical role by protecting the information-containing DNA bases and disrupting secondary structure in the (typically) brief time between strand separation and duplication.

Much of our knowledge of DNA replication comes from the study of the model bacterium *Escherichia coli,* the focus of this thesis. Although DNA replication in *E. coli* shares much in common with that in the cells of humans and other eukaryotes, there are several key differences. Most notably, two replisomes initiate synthesis at a single defined location, or origin, on the bacterium's circular chromosome, and travel in opposite directions as separate replication forks until they meet to terminate synthesis at the opposite point. In eukaryotes, in contrast, bidirectional replication initiates at multiple origins on each linear chromosome.

Within each fork during *E. coli* replication (see Figure 1.1 for a schematic), a single replicative DNA polymerase, Polymerase III, performs the majority of DNA synthesis, with copies on the leading and lagging strands that are physically coupled to one another by interactions with $\tau_3\delta\delta'\chi\psi$, the clamp loader complex, which is also responsible for placing copies of the bacterial processivity factor β at new primers. In eukaryotes, there is a division of labor between distinct polymerases on the leading and lagging strand, and no similar coupling has been determined (1). Finally, the *E. coli* helicase DnaB, which additionally binds the clamp loader complex, travels along the lagging, rather than leading, strand (2).

Despite these differences, structural and functional conservation of the replisome across the tree of life makes *E. coli* a valuable model system for studying

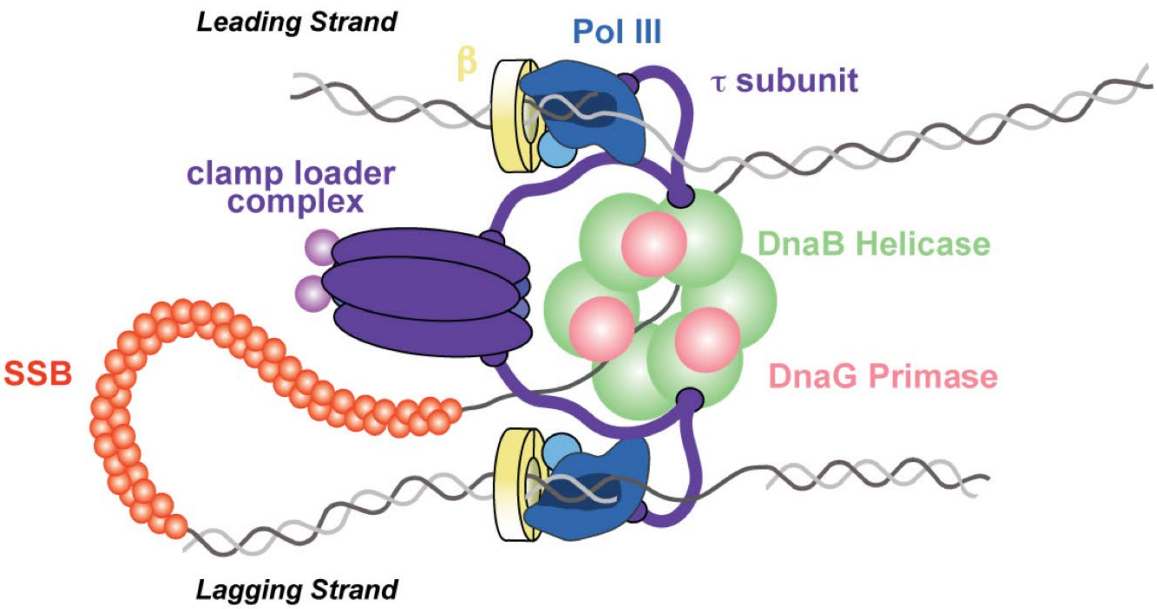

Figure 1.1 A schematic representation of coordinated leading and lagging strand synthesis by the *E. coli* replisome within a single replication fork. Image courtesy of Joseph Loparo.

DNA replication. In addition to its extensive characterization and a wealth of genetic tools and alleles, *E. coli* remains one of the few model organisms for which DNA replication can be fully reconstituted from initiation to coordinated leading and lagging strand replication – although recent efforts have brought the yeast replisome closer to this goal (3).

1.2 DNA damage, repair, and the SOS response

In addition to stability and speed, DNA replication exhibits a remarkable fidelity. *E. coli* Polymerase (Pol) III has an error rate of about 10^{-7}, resulting in it generating less than one error per genome duplication; mismatch repair further reduces the error rate by an additional 100-fold (4). Although the largest contribution to fidelity is at nucleotide binding and incorporation, the differences in free energy between proper and improper base pairing are not great enough to

achieve a sufficient fidelity. Therefore, replicative DNA polymerases use geometric selectivity that requires proper Watson-Crick base pairing before the active site residues are positioned for catalysis (5). Structural studies of Polymerase I, the polymerase involved in the completion of Okazaki fragments, suggest that improper pairings at the insertion site or even several nucleotides downstream within the DNA-binding footprint of a replicative polymerase disrupt the active site and can serve as a signal for the primer strand to be transferred to a exonuclease domain for proofreading (6).

Due to these strict geometric constraints, DNA damage that distorts base pairing or the structure of DNA serves as a potent block to replicative DNA polymerases, and can potentially lead to replication fork collapse and genomic instability (7). Damage can caused by exogenous sources, such as UV radiation, which crosslinks pyrimidine bases, and chemicals that form adducts with DNA, such as the carcinogen benzo[a]pyrene diol epoxide (8). However, DNA modifications can also be caused by endogenous sources of DNA damage, often the byproducts of metabolism. At any given time, the *E. coli* genome is estimated to contain about 100 DNA lesions caused by common oxidation and alkylation products (9).

Unsurprisingly, conserved pathways exist that reverse or repair DNA damage. These pathways were often discovered through mutations that sensitized cells to DNA damage – *e.g.*, (10, 11) – and include: base excision repair, in which specific, common lesions are recognized by glycosylases and removed, followed by a filling in of the gap by a polymerase; nucleotide excision repair, in which lesions are recognized more generally through DNA distortions and a larger patch containing

the DNA lesion is excised; and recombination, in which the damaged strand is repaired using a homologous sequence on a sister chromatid (9). Since in all three cases an undamaged DNA strand is used to template new synthesis, they are considered to be generally "error-free" forms of repair.

Although some repair factors are constitutively expressed, many are upregulated in response to DNA damage. A significant observation of this phenomenon, by Jean Weigle, was that UV-damaged phage have an increased survival in cells that were themselves irradiated with UV before infection (12). Painstaking genetics and biochemistry revealed that this and similar phenomena were due to activity of a damage-responsive transcriptional repressor, LexA, which controls the expression of over 50 genes (13). Named the SOS response by Miroslav Radman (after the international distress code), de-repression of these genes occurs after the accumulation of single-stranded DNA in cells caused by failed replication at sites of DNA damage, subsequent filamentation of the RecA recombinase at these single-stranded regions, and the autoproteolytic cleavage and inactivation of LexA, catalyzed by the interaction of the RecA filament with the repressor (14).

1.3 Translesion synthesis

As the *E. coli* DNA repair pathways and their constituent factors were characterized, additional genes were isolated that are independent of the canonical repair pathways, but either sensitized cells to DNA damage when mutated, or were themselves regulated by DNA damage. Curiously, the *dinB* gene (15) and the operon *umuDC* (16), both controlled by the SOS response, were found to induce

mutagenesis following UV irradiation – the former in the untargeted mutagenesis of λ phage transfected into irradiated cells (17), and the latter (requiring both *umuD* and *umuC* acting together) throughout the chromosome (16, 18). One model was that the UV mutagenesis phenotype was due to factors that reduced the fidelity of Pol III, allowing it to tolerate and replicate through DNA damage. Another model, proposed independently by Evelyn Witkin and Bryn Bridges, was that mutagenesis was due to the action of separate error-prone polymerases (14).

The latter proposal proved prescient, but it took another 30 years to demonstrate that *dinB* and *umuDC* were in fact error-prone DNA polymerases able to replicate through DNA lesions in a process called translesion synthesis (TLS). Together, they are the fourth and fifth *E. coli* DNA polymerases, with DinB as the single subunit Polymerase IV (19) and Polymerase V made up of a catalytic subunit, UmuC, and two cleaved subunits of UmuD, or $UmuD_2$' (20) (see Table 1.1 for a list of *E. coli* polymerases). Biochemical assays revealed that Pol IV can bypass alkylated DNA damage caused by methyl methanesulfonate (MMS) (21, 22) and nitroaromatic adducts formed by treatment with the antibiotic nitrofurazone (NFZ) (23) – and *dinB* mutants are significantly sensitive to these agents. Similarly, Pol V is able to bypass a large range of adducts, including UV products and abasic sites (24). Subsequently, a previously discovered and unrelated DNA polymerase, Pol II, encoded by the gene *polB* and also part of the SOS response, was discovered to also engage in TLS (25-27), although without as clear a preference or phenotype as Pol IV and Pol V (7).

Name	Pol family	Biological function	Regulated by the SOS response?
Pol I	A	DNA replication, Okazaki fragment maturation, DNA repair	No
Pol II	B	TLS	Yes
Pol III	C	DNA replication, DNA repair	No
Pol IV	Y	TLS (homolog of human Pol κ)	Yes
Pol V	Y	TLS	Yes

Table 1.1 DNA polymerases of *E. coli*

Studies of error-prone, or TLS, polymerases in other organisms, and the analysis of an increasing number of genomic sequences revealed that *E. coli* Pol IV and Pol V were in fact representatives of a large DNA polymerase family – called the Y family, following the previous naming convention of the A, B, C, D, and X polymerase families (28). Since these initial discoveries, Y family polymerases have been found in all three domains of life, with four in the human genome (29).

Structural studies of Y-family polymerases proved to be immensely informative of their function. Despite sharing little sequence similarity to replicative polymerases, Y family polymerases adopt a similar right-handed structure and use the same two-metal catalytic mechanism (30, 31). However, in contrast to the replicative polymerases of the other families, the active sites of Y family polymerases were generally found to make less specific interactions with the base pair formed by the incoming nucleotide, to be more solvent accessible, and to undergo less dramatic conformational changes upon nucleotide addition (31-33). These features explain both the TLS and mutagenesis activities of Y-family polymerases. By lacking the strict geometric selectivity of replicative polymerases, and having what is in effect a "pre-formed" active site, TLS polymerases are more

able to tolerate DNA-distorting lesions, at the expense of reduced specificity for a proper Watson-Crick base pair. Consistent with this, Pol IV and Pol V have error rates in the range of 10^{-3} to 10^{-4} (24) and lack exonuclease domains that are activated by DNA duplex distortions (34) and could therefore interfere with bypass.

The wide conservation of Y-family polymerases, and the presence of polymerases from other families with lesion bypass activity (such as the B-family *E. coli* Pol II and human Pol ζ), argues that TLS is an ancient and important pathway, complementary to DNA repair. TLS allows for greater tolerance of damage by alleviating lesion-induced replication stalling, letting normal replication resume without requiring the disassembly and reloading of the replisome, and leaves lesions for post-replicative repair. TLS is additionally advantageous when both strands are damaged, as in is the case of an interstrand crosslink – the initial excision of the crosslink leaves a gap across from an adduct that stalls replicative polymerases (35, 36).

Although the early focus on the induced mutagenesis phenotype has led to the label of these polymerases as "error-prone," many appear to be specialized to copy specific classes of lesions with a surprising degree of fidelity. *E. coli* Pol IV, for example, efficiently and accurately inserts dCTP across from a N^2-furfuryl-dG lesion (23), while human Pol η inserts dATP across from both bases of a thymine-thymine dimer (37), and makes specific contacts with a dimer-containing duplex that splints it into a B-form-like DNA structure (38).

As the primary function of TLS polymerases is damage tolerance, and lesion bypass activity is intrinsically connected to higher error rates, it is likely that

induced mutagenesis is an evolutionary side effect of TLS. Nevertheless, TLS mutagenesis has been implicated in several phenomena. In prokaryotes, stress-induced mutagenesis is the mechanistically controversial (39, 40) error-prone synthesis of double-strand break intermediates by Pol IV, which has been proposed to be a means of increasing the diversity of a population during periods of intense selection (41, 42). Pol IV has additionally been shown to be involved in the potentiation of antibiotic lethality by reactive oxygen species, through the incorporation of oxidized nucleotides into the genome (43), and increased mutagenesis induced by sublethal levels of β-lactam antibiotics (44). Finally, misregulation and overexpression of eukaryotic TLS polymerases have been connected to the mutator phenotype of several cancers (45).

1.4 Polymerase selection and exchange

The double-edged nature of TLS polymerases – the fitness benefits of lesion bypass and potential defects of increased mutagenesis – as well as the specificity of different TLS polymerases for different lesions, suggests that polymerase selection is carefully regulated. As discussed above, the levels of *E. coli* TLS polymerases are controlled by the SOS response. Following induction, the levels of Pol II and Pol IV are increased roughly 10-fold from their initial levels of 50 and 200 copies, respectively (46, 47). Expression of the more mutagenic Pol V, in contrast, is fully repressed during normal growth, and does not form the active $UmuD_2$'C polymerase complex until later in the SOS response (7).

Another focus of polymerase selection has been ring-shaped processivity clamps: the homodimeric bacterial clamp, β, and the homotrimeric human clamp, proliferating cell nuclear antigen (PCNA). All five *E. coli* polymerases bind a hydrophobic binding cleft on β via a conserved clamp-binding motif (CBM) (48, 49), a requirement for processive synthesis. Given each subunit of the β and PCNA has a binding cleft, Pagès and Fuchs proposed that the clamp may mediate polymerase exchange by simultaneously binding polymerases, much like a workman's toolbelt (50). In this "toolbelt model," they reasoned, the relative occupancies of different polymerases would be a function of their concentrations and affinities for the clamp. The switch from a replicative to TLS polymerase could be signaled by a distorting lesion, with the eventual release of the polymerase from the clamp and replacement with another if it failed to bypass the lesion – a "trial and error" approach.

As for DNA replication, *E. coli* has served as an important model system for the study of polymerase selection. A structure of the C-terminal β-binding (or "little finger") domain of Pol IV revealed an second, unexpected interface between internal Pol IV residues and a site near the clamp dimer interface the authors named the β "rim" (51). In support of the toolbelt model, when the full Pol IV structure is overlaid onto this domain, the full polymerase "hand" points well away from the DNA running through the clamp, suggesting an inactive binding mode that would allow Pol IV to occupy one subunit of the clamp during Pol III synthesis (Figure 1.2). Following the release of Pol III from DNA, perhaps following the encounter of a DNA lesion, Pol IV could break its interaction with the rim, exchanging it for interactions with the DNA duplex and initiate bypass.

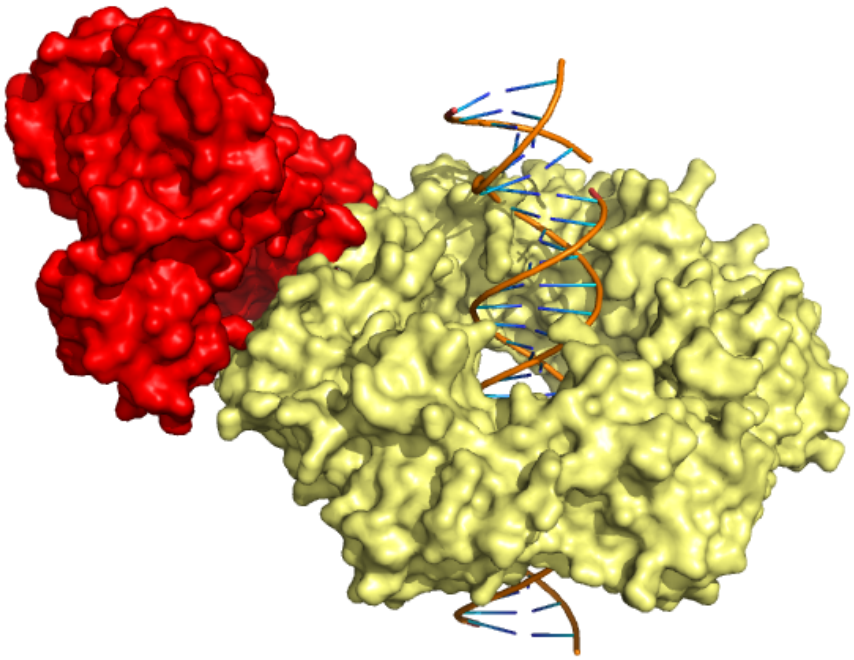

Figure 1.2 A proposed inactive β-binding mode for Pol IV. This model was constructed in PyMOL (Schrödinger) by docking the full structure of Pol IV (PDB 4IR9) (58) onto the co-crystal structure of the Pol IV little finger domain bound to β (PDB 1UNN) (51).

Although attractive, the toolbelt model suffers from several challenges. First, in the absence of a detailed Pol III-β co-structure, it is unclear if the large, heterotrimeric replicative polymerase could accommodate Pol IV, even in its proposed inactive binding mode. Second, Pol III was recently discovered to have a second, cryptic β-binding CBM on its ε exonuclease subunit (52, 53). Although it is not critical for processive DNA synthesis and weaker then the essential CBM within the polymerase subunit α, the interaction may stabilize the Pol III αεθ core, leading to the proposal that Pol III binds both subunits of β during active synthesis, occluding other factors from binding. Finally, although each clamp may be multivalent, the number of clamp-binding proteins – at least 10 in bacteria (54) and over 50 (55) in eukaryotes – far exceeds the number of binding sites.

Beyond the validity of the toolbelt model, another important question is under what contexts polymerase exchange and TLS occurs. One possibility is that the exchange between replicative and translesion polymerases typically occur within the intact replisome as it encounters and stalls at a DNA lesion. The second is that TLS occurs largely in single-stranded gaps left at lesion sites as the replisome translocates past and reinitiates synthesis downstream. This "lesion skipping" activity has been recently demonstrated by the Marians lab for an *in vitro* reconstitution of *E. coli* replication (56, 57).

This thesis seeks to clarify the mechanism of polymerase exchange, test the validity of the toolbelt model, and determine under what contexts TLS occurs using the *E. coli* replisome and TLS polymerases as a model system. A major focus will be on the development and application of single-molecule techniques, which have superior resolution and avoid ensemble averaging that makes it impossible to probe stochastic polymerase exchange events. Broadly speaking, the thesis will be divided in two parts. Chapters 2 and 3 will report an *in vitro* single-molecule reconstitution of polymerase exchange and TLS to determine the role of clamp-binding interactions and the validity of the toolbelt model. The second part, Chapter 4, will report the observation of individual DNA polymerases in living bacteria in an initial effort to determine where and when TLS occurs, and the role of different protein-protein interactions. Finally, I will discuss outstanding questions and potential future directions for this field.

References

1. Indiani C, O'Donnell M (2006) The replication clamp-loading machine at work in the three domains of life. *Nat Rev Mol Cell Biol* 7(10):751–761.

2. LeBowitz JH, McMacken R (1986) The *Escherichia coli* dnaB replication protein is a DNA helicase. *J Biol Chem* 261(10):4738–4748.

3. Yeeles JTP, Deegan TD, Janska A, Early A, Diffley JFX (2015) Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* 519(7544):431–435.

4. Schaaper RM (1993) Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem* 268(32):23762–23765.

5. Kunkel TA (2004) DNA Replication Fidelity. *J Biol Chem* 279(17):16895–16898.

6. Johnson SJ, Beese LS (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell* 116(6):803–816.

7. Fuchs RP, Fujii S (2013) Translesion DNA synthesis and mutagenesis in prokaryotes. *Cold Spring Harb Perspect Biol* 5(12):a012682.

8. Ling H, *et al*. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc Natl Acad Sci USA* 101(8):2265.

9. Strauss, BS (2005) Excision repair and bypass. *The Bacterial Chromosome*, ed Higgins NP (ASM Press, Washington, D.C.), pp. 431-447.

10. Clark AJ, Margulies AD (1965) Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12. *Proc Natl Acad Sci USA* 53:451–459.

11. Howard-Flanders P, Boyce RP, Theriot L (1966) Three loci in *Escherichia coli* K-12 that control the excision of pyrimidine dimers and certain other mutagen products from DNA. *Genetics* 53(6):1119–1136.

12. Weigle JJ (1953) Induction of Mutations in a Bacterial Virus. *Proc Natl Acad Sci USA* 39(7):628–636.

13. Simmons LA, Foti JJ, Cohen SE, Walker GC (2007) The SOS Regulatory Network. *EcoSal* 5.4.3:1–44.

14. Friedberg EC, *et al*. The SOS responses of prokaryotes to DNA damage. (2005) *DNA Repair and Mutagenesis*, 2nd. (ASM Press, Washington,D.C.), pp. 463-508.

15. Kenyon CJ, Walker GC (1980) DNA-damaging agents stimulate gene expression at specific loci in *Escherichia coli*. *Proc Natl Acad Sci USA* 77(5):2819–2823.

16. Kato T, Shinoura Y (1977) Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. *Mol Gen Genet* 156(2):121-131.

17. Brotcorne-Lannoye A, Maenhaut-Michel G (1986) Role of RecA protein in untargeted UV mutagenesis of bacteriophage lambda: evidence for the requirement for the *dinB* gene. *Proc Natl Acad Sci USA* 83(11):3904–3908.

18. Elledge SJ, Walker GC (1983) Proteins required for ultraviolet light and chemical mutagenesis. Identification of the products of the *umuC* locus of *Escherichia coli*. *J Mol Biol* 164(2):175–192.

19. Wagner J, *et al*. (1999) The *dinB* gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell* 4(2):281–286.

20. Tang M, *et al*. (1999) UmuD'(2)C is an error-prone DNA polymerase, Escherichia coli pol V. *Proc Natl Acad Sci USA* 96(16):8919–8924.

21. Bjedov I, *et al*. (2007) Involvement of *Escherichia coli* DNA polymerase IV in tolerance of cytotoxic alkylating DNA lesions *in vivo*. *Genetics* 176(3):1431–1440.

22. Cafarelli TM, *et al.* (2013) A Single residue unique to DinB-like proteins limits formation of the polymerase IV multiprotein complex in *Escherichia coli*. *J Bacteriol* 195(6):1179–1193.

23. Jarosz DF, Godoy VG, Delaney JC, Essigmann JM, Walker GC (2006) A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature* 439(7073):225–228.

24. Tang M, *et al*. (2000) Roles of E. coli DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* 404(6781):1014–1018.

25. Paz-Elizur T, Takeshita M, Goodman M, O'Donnell M, Livneh Z (1996) Mechanism of translesion DNA synthesis by DNA polymerase II. Comparison to DNA polymerases I and III core. *J Biol Chem* 271(40):24662–24669.

26. Napolitano R, Janel-Bintz R, Wagner J, Fuchs RP (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J* 19(22):6259–6265.

27. Becherel OJ, Fuchs RP (2001) Mechanism of DNA polymerase II-mediated frameshift mutagenesis. *Proc Natl Acad Sci USA* 98(15):8566–8571.

28. Ohmori H, *et al*. (2001) The Y-family of DNA polymerases. *Mol Cell* 8(1):7–8.

29. Foti JJ, Walker GC (2010) SnapShot: DNA polymerases II mammals. *Cell* 141(2):370

30. Zhou B, Pata J, Steitz T (2001) Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. *Mol Cell* 8(2):427–437.

31. Ling H, Boudsocq F, Woodgate R, Yang W (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell* 107(1):91–102.

32. Steitz TA, Yin YW (2004) Accuracy, lesion bypass, strand displacement and translocation by DNA polymerases. *Philos Trans R Soc Lond B Biol Sci* 359(1441):17–23.

33. Yang W (2005) Portraits of a Y-family DNA polymerase. *FEBS Lett* 579(4):868–872.

34. Beese L, Steitz T (1991) Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J* 10(1):25.

35. Kumari A, *et al*. (2008) Replication bypass of interstrand cross-link intermediates by *Escherichia coli* DNA polymerase IV. *J Biol Chem* 283(41):27433–27437.

36. Räschle M, *et al*. (2008) Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* 134(6):969–980.

37. Prakash S, Johnson RE, Prakash L (2005) Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem* 74:317–353.

38. Biertümpfel C, *et al*. (2010) Structure and mechanism of human DNA polymerase η. *Nature* 465(7301):1044–1048.

39. McKenzie GJ, Lee PL, Lombardo MJ, Hastings PJ, Rosenberg SM (2001) SOS mutator DNA polymerase IV functions in adaptive mutation and not adaptive amplification. *Mol Cell* 7(3):571–579.

40. Slechta ES, *et al*. (2003) Adaptive mutation: general mutagenesis is not a programmed response to stress but results from rare coamplification of *dinB* with *lac*. *Proc Natl Acad Sci USA* 100(22):12847–12852.

41. Ponder RG, Fonville NC, Rosenberg SM (2005) A switch from high-fidelity to

error-prone DNA double-strand break repair underlies stress-induced mutation. *Mol Cell* 19(6):791–804.

42. Pomerantz RT, Kurth I, Goodman MF, O'Donnell ME (2013) Preferential D-loop extension by a translesion DNA polymerase underlies error-prone recombination. *Nat Struct Mol Biol* 20(6):748–755.

43. Foti JJ, Devadoss B, Winkler JA, Collins JJ, Walker GC (2012) Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* 336(6079):315–319.

44. Gutierrez A, *et al*. (2014) β-lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nature Commun* 4:1610–9.

45. Hoffmann J-S, Cazaux C (2010) Aberrant expression of alternative DNA polymerases: a source of mutator phenotype as well as replicative stress in cancer. *Semin Cancer Biol* 20(5):312–319.

46. Courcelle J, Khodursky A, Peter B, Brown PO, Hanawalt PC (2001) Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* 158(1):41–64.

47. Sutton MD (2010) Coordinating DNA polymerase traffic during high and low fidelity synthesis. *Biochim Biophys Acta* 1804(5):1167–1179.

48. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc Natl Acad Sci USA* 98(20):11627–11632.

49. Wijffels G, *et al*. (2004) Inhibition of protein interactions with the β_2 sliding clamp of *Escherichia coli* DNA polymerase III by peptides from β_2-binding proteins. *Biochemistry* 43(19):5661–5671.

50. Pagès V, Fuchs RP (2002) How DNA lesions are turned into mutations within cells? *Oncogene* 21(58):8957–8966.

51. Bunting K, Roe S, Pearl L (2003) Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the β-clamp. *EMBO J* 22(21):5883–5892.

52. Jergic S, *et al*. (2013) A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J*:1–12.

53. Toste Rêgo A, Holding AN, Kent H, Lamers MH (2013) Architecture of the Pol III-clamp-exonuclease complex reveals key roles of the exonuclease subunit in processive DNA synthesis and repair. *EMBO J* 32(9):1334–1343.

54. Wijffels G, *et al*. (2004) Inhibition of protein interactions with the beta sliding clamp of Escherichia coli DNA polymerase III by peptides from beta-binding proteins. *Biochemistry* 43(19):5661–5671.

55. Moldovan G-L, Pfander B, Jentsch S (2007) PCNA, the maestro of the replication fork. *Cell* 129(4):665–679.

56. Heller RC, Marians KJ (2006) Replication fork reactivation downstream of a blocked nascent leading strand. *Nature* 439(7076):557–562.

57. Yeeles JTP, Marians KJ (2013) Dynamics of leading-strand lesion skipping by the replisome. *Mol Cell* 52(6):855-865.

58. Sharma A, Kottur J, Narayanan N, Nair DT (2013) A strategically located serine residue is critical for the mutator activity of DNA polymerase IV from *Escherichia coli*. *Nucleic Acids Res* 41(9):5104–5114.

Chapter 2

Polymerase exchange on single DNA molecules reveals processivity clamp control of Polymerase IV translesion synthesis

2.1 Introduction

The presence of multiple DNA polymerases in an organism, each with specialized functions, points to the existence of specific mechanisms of polymerase exchange. Some handoffs occur regularly. Within the *Saccharomyces cerevisiae* model eukaryotic replisome, Okazaki fragments are initiated by the primase activity of Polymerase (Pol) α, which hands the primer off to Pol δ for extension (1). In *Escherichia coli* replication, Okazaki fragment termination occurs once every few seconds by an exchange between the replicative polymerase Pol III, and Pol I, which removes RNA primers and fills in the resulting gaps (2). A second exchange then hands off the processed Okazaki fragment to DNA ligase.

Other types of exchange occur less frequently. Translesion DNA polymerases contribute to the tolerance of DNA damage, but their slower speeds and higher mutagenesis rates suggest that exchange with replicative polymerases is limited to particular contexts, such as sites of DNA damage. Improper exchange of either type can be deleterious, by slowing or stalling replication, and unnecessarily exposing

Material in this chapter was originally published as:
Kath et al. (2014) Polymerase exchange on single DNA molecules reveals processvity clamp control of translesion synthesis. *Proc Natl Acad Sci USA* 111:7646-7652.

single-stranded DNA intermediates that are sensitive to double-strand breaks and other forms of damage.

The development of paradigms for proper and improper DNA polymerase exchange requires appropriate model systems and experimental techniques. One such model system, the focus of this chapter, is the exchange between *E. coli* Pol III and Pol IV. As discussed in Chapter 1, Pol IV is a Y-family DNA polymerase capable of bypassing nitroaromatic and alkylation DNA damage; although not essential, deletion of the Pol IV gene *dinB* sensitizes cells to corresponding DNA damaging agents (3).

Previous efforts to reconstitute this model system for polymerase exchange were accomplished by stalling Pol III bound to the β processivity clamp at a primer terminus by nucleotide omission in order to synchronize the population of molecules and simulate a lesion-induced block (4, 5). Pol IV was then added, and, shortly after, DNA synthesis was initiated by adding the remaining nucleotides. The studies inferred exchange by observing DNA synthesis of the primed single-stranded circular DNA template over a short period in which Pol III would fully complete synthesis, but the slower Pol IV would not. The observation that Pol IV can displace and replace Pol III on a timescale faster than Pol III dissociates from the primer terminus suggests that exchange is not simply passive. Rapid exchange also requires both binding clefts of β, as well as the secondary rim-binding residues of Pol IV involved in its inactive binding mode, evidence that both polymerases bind to the clamp simultaneously during or prior to exchange.

Although informative, these studies have several limitations. First, it is unclear if stalling by nucleotide omission accurately reflects the behavior of Pol III as it approaches and stalls at a lesion. The alternative model that Pol III occludes Pol IV binding during active synthesis (6) cannot be tested using this approach. Second, although Pol III could be stably stalled on a DNA substrate, Pol IV could not, and these studies were therefore not able to accurately resolve an exchange back to Pol III following Pol IV synthesis to determine if Pol III might have an inactive binding mode. Finally, and perhaps most fundamentally, the use of a gel-based, ensemble technique means that potentially distinct subpopulations of exchange reactions, as well as rare, stochastic events, are averaged out.

2.2 Methods: Single molecule observation of DNA synthesis

An alternative approach is one in which polymerase exchange is monitored on individual DNA substrates. Single-molecule techniques have previously been used to monitor DNA synthesis, and can broadly be divided into two categories. The first involves the use of F*ö*rster Resonance Energy Transfer, or FRET, where an excited fluorophore can transfer energy to a separate, spectrally shifted fluorophore, if they are in close enough spatial proximity. The distance between the two fluorophores, or the FRET pair, can therefore be determined by measuring the relative fluorescence of a colocalized pair in a sparse field under a microscope. Previous studies have labeled a polymerase and an immobilized DNA substrate with fluorophores, and observed the change in fluorescence of the FRET pair as the polymerase moved toward or away from the labeled DNA site during synthesis (7,

8). Although this technique involves the direct observation of a polymerase and can have base pair (bp) resolution, it has a limited dynamic range – 50% FRET transfer typically occurs at ~5 nm, corresponding to less than 15 bp – and it requires polymerization to be slowed significantly to resolve synthesis.

The second approach involves monitoring the extension of linear, tethered DNA molecules during primer synthesis by exploiting the differential elasticity of single-stranded (ss) and double-stranded (ds) DNA. At a constant tension of a few piconewtons (pN), ssDNA is entropically coiled, while dsDNA is stretched to nearly its crystallographic length. Conversion of ssDNA to dsDNA by a polymerase therefore increases the extension of a tethered DNA molecule, which can be monitored individually.

This approach has been used to study primer extension by different polymerases: T7 DNA polymerase using an optical trap (9), the Klenow fragment of *E. coli* Pol I using magnetic tweezers (10), and *E. coli* Pol III in flow stretching experiments (11). These DNA stretching techniques have a much larger dynamic range (limited only by the length of the DNA substrate) and are able to measure rapid rates of synthesis. In contrast to fluorescence-based approaches, the presence of the polymerase is not directly observed – although a recent study combined the two to simultaneously measure DNA synthesis and the stoichiometry of T7 polymerases within a replication fork (12).

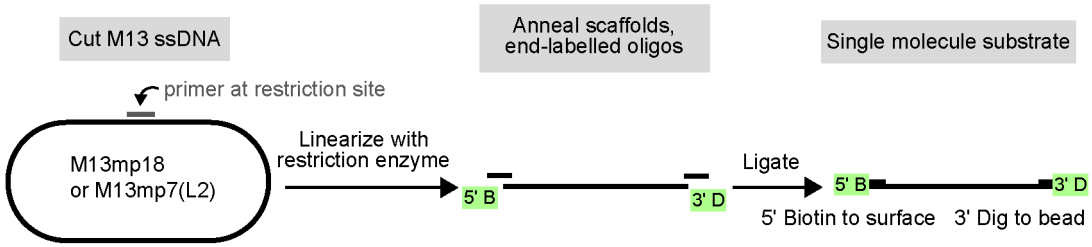

Figure 2.1. A schematic of the construction of single-molecule substrates from M13 phage genomes: the circular ssDNA is linearized with a restriction enzyme, and end-labeled using scaffold-mediated ligation.

Because the processivities of primer extension by Pol III and Pol IV range from hundreds of bp to over a kilobase pair (kb) (11, 13), I chose to use a DNA flow stretching technique to investigate polymerase exchange. This approach has previously been used to observe primer extension by Pol III (11), and has a much greater throughput than magnetic tweezers or optical tweezers (14).

2.2.1 A single-molecule DNA substrate

Previous single-molecule studies of primer synthesis have generated ssDNA substrates by separating two strands of a dsDNA duplex with an alkaline buffer using force, or by using an exonuclease to chew away one strand (9, 10, 15). To remove this separation step at the beginning of each experiment, I designed and constructed linear single-molecule DNA substrates using the circular, 7.2 kb genome of the phage M13. This substrate was labeled on one end with a biotin-containing oligonucleotide to couple it to a streptavidin surface of a flow cell, and on the other end with a digoxigenin-containing oligonucleotide to couple it to an anti-digoxigenin functionalized bead (Figure 2.1).

Names	**Sequences** (5' to 3')
mp18-SalI	CTGCAGGTCGACTCTAGA
mp18-scaffold-1	GCGGGCAATATGTACCTCTAGAGGATCCCC
mp18-scaffold-2	ATGCCTGCAGGTCGAACTATGCGACTGGAC
N2-dG-insert	CTACCT/*N*^{2}-fluoro-dG/TGGACGGCTGCGA
dG-control-insert	CTACCTGTGGACGGCTGCGA
N2-dG_minus3	TCGCAGCCGT
N2-dG-scaffold-1	AAAACGACGGCCAGTGAATTTCGCAGCCGTCC
N2-dG-scaffold-2	GGTAGACTGAATCATGGTCATAGC
mp7L2-AlwNI	AGCGCAGTCTCTGAATTTAC
mp7L2-scaffold-1	GCGGGCAATATGTACTCTCTGAATTTACCG
mp7L2-scaffold-2	GAATGGAAAGCGCAGACTATGCGACTGGAC
M13-3'-dig	GTACATATTGCCCGCAAAAAA-Dig
M13-5'-biotin	BioTEG-GTCCAGTCGCATAGT
dideoxy-M13-block	GCTAACGAGCGTCTTTCCAGAGCCTAATTT GCCAGTTA/ddC/

Table 2.1. Oligonucleotides used to construct single-molecule substrates. N2-dG-insert was purchased from Chemgenes; the rest were purchased from Integrated DNA Technologies.

The protocol is as follows: 16 μL of M13mp18 ssDNA (NEB, 250 ng μL^{-1}) were annealed in 20 μL with 1 μM of the oligonucleotide mp18-SalI (Table 2.1) by heating to 65˚C for 10 min and slowly cooling to room temperature. 10 μL of the annealed DNA were linearized at the double-stranded DNA region with 10 U SalI (NEB) and 1X Buffer 3 in 50 μL at 37˚C for 1 h. 40 μL of the restriction digest reaction were mixed, to a final volume of 55 μL, with 30 nM of the end-labeled oligonucleotides M13-5'-biotin and phosphorylated M13-3'-dig, and 30 nM of the scaffolding oligonucleotides mp18-scaffold-1 and mp18-scaffold-2 (Table 2.1). The scaffolds were annealed to the linearized phage DNA and the end-labeled oligonucleotides by heating to 65˚C for 20 min and cooling to room temperature, also inactivating SalI; scaffold-mediated ligation was subsequently performed overnight at 16˚C with 400 U DNA ligase (NEB). The reaction was stopped by heat-

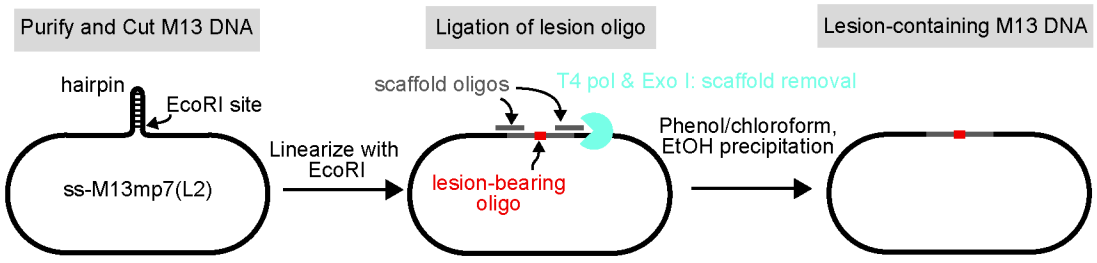

Figure 2.2. A schematic of the generation of ssDNA containing a site-specific lesion: an internal hairpin is cleaved with EcoRI, and a chemically synthesized, lesion-containing oligonucleotide is ligated in with scaffolds to re-circularize the molecule. Excess primers and scaffolds are removed using T4 DNA polymerase and exonuclease I.

inactivating ligase at 65˚C for 10 min and adding EDTA (20 mM final) to the cooled mixture. The stock solution (final substrate concentration ~5 nM) was stored at 4˚C.

2.2.2 A lesion-containing DNA substrate

An additional advantage in using M13 as a template for the single-molecule substrate is the existence of protocols to generate M13 genomes with site-specific DNA lesions. The protocol chosen for this study uses M13mp7(L2), a mutant phage that contains an EcoRI site within a stable hairpin in its genome (16) (Figure 2.2). Cleaving the EcoRI site allows a lesion-containing oligonucleotide to be inserted using scaffold-mediated ligation.

For this study, I chose to generate templates containing N^2-furfuryl-dG, a lesion efficiently bypassed by Pol IV (17). A 20-mer oligonucleotide containing the N^2-furfuryl-dG lesion was constructed as previously described (17). Briefly, a 20-mer oligonucleotide containing a fluoro substituent at the N^2 position of a single guanine base was purchased from Chemgenes, designated N2-dG-insert (Table 2.1).

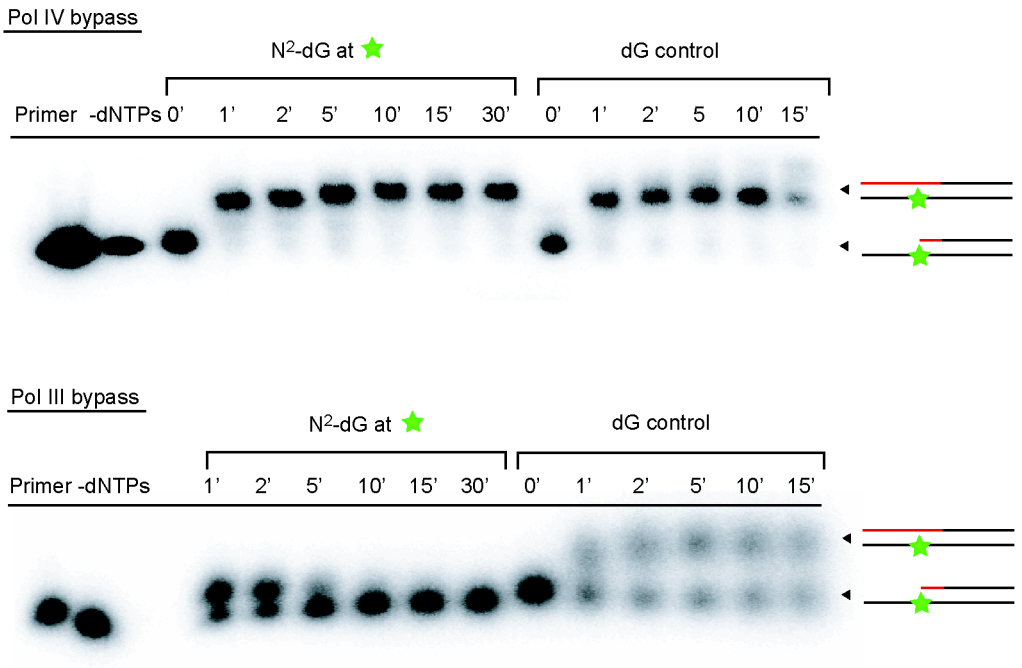

Figure 2.3. A time course for distributive synthesis (in the absence of β) by Pol III or Pol IV on an oligonucleotide template with a *N*²-furfuryl-dG (left) or control dG (right) base at the +3 position from the primer terminus. The lesion is a strong block for Pol III, but is rapidly bypassed by Pol IV. Arrows denote the location of the labeled 10-mer primer and the fully extended 20-mer product. Degradation of the primer by Pol III for the -dNTP control and on the lesion template is incomplete due to the short primer and the large footprint of the Pol III core.

The fluorine was displaced with a fufuryl group by treating the oligonucleotide with fufurylamine, followed by HPLC purification and validation with MALDI-TOF.

'Running start' bulk primer extension reactions were performed using the *N*²-furfuryl-dG-containing 20-mer annealed with the 5'-^{32}P-phosphorylated primer N2-dG_minus3 (Table 2.1), with either 50 nM Pol III or Pol IV in replication buffer (see Section 2.2.5). 30 μL reactions containing the primer-template (20 nM final) were initiated by adding dNTPs (250 μM) and incubating at 37°C. At the indicated times, 3 μL of the each reaction was added to 10 μL of stop buffer (50 mM Tris-HCl pH 7.5, 25 mM EDTA, 0.5% SDS). The control reaction lacking dNTPs was stopped

after 15 min. Samples were separated on a 12% urea-PAGE gel and the dried gel was exposed to a phosphor screen and imaged with a Personal Molecular Imager. As expected, the lesion was a strong block for Pol III, but was rapidly bypassed by Pol IV (Figure 2.3), in agreement with previous biochemical and genetic data for this lesion (17).

The protocol to purify M13mp7(L2) phage and ligate the lesion-containing oligonucleotide at the digested EcoRI site (16) was adapted for this study with the following modifications. Following PEG precipitation, DNA was extracted from the isolated phage pellet two to three times with 25:24:1 phenol/chloroform/isoamyl alcohol and once with pure chloroform. DNA in the final aqueous layer was ethanol precipitated and redissolved in 10 mM Tris pH 8.5 buffer to ~2 μg μL^{-1}. 100 μg of M13mp7(L2) DNA was linearized in a 100 μL digestion reaction with 40 U EcoRI-HF (New England Biolabs, NEB) and 1X Buffer 4 at 23˚C for 8 h and purified with sequential phenol/chloroform/isoamyl alcohol and chloroform extractions, and an ethanol precipitation, then dissolved in 100 μL Tris buffer. Purification of linear ssDNA prevented degradation in later steps.

30 pmol of the 5'-phosphorylated N^2-furfuryl-dG oligonucleotide insert (N2-dG-insert dG) or a control insert with dG at the equivalent position (dG-control-insert) were ligated at 16˚C overnight into 20 pmol of the purified, linear ssDNA using 25 pmol each of annealed scaffold oligonucleotides N2-dG-scaffold-1 and N2-dG-scaffold-2 (Table 2.1) and 800 U T4 DNA ligase (NEB) in a 60 μL reaction. To remove the scaffold oligonucleotides, unligated linear M13 DNA, and excess insert, the mixture was subsequently treated at 37˚C for 4 h with 18 U T4 DNA polymerase

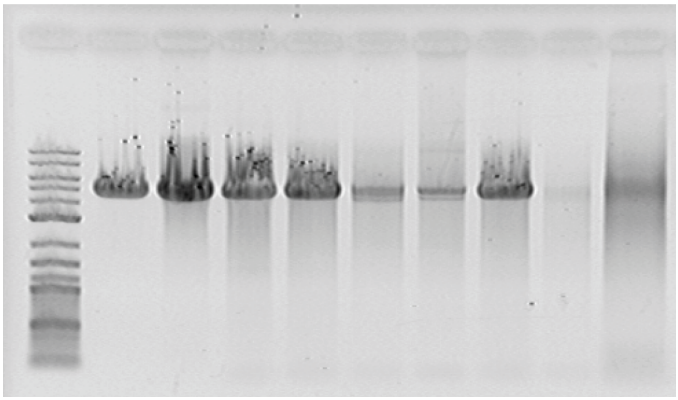

Figure 2.4. Successive steps of preparation of lesion-containing M13 ssDNA; each lane of the agarose gel contains 700 ng DNA, or the equivalent amount of the reaction, unless otherwise specified. Lanes numbered 1-10 from the left: lane 1, 2-log ladder (5 μL, NEB); lane 2, M13mp18 ssDNA (NEB); lane 3, purified M13mp7L2 ssDNA; lane 4, after EcoRI digestion and subsequent purification; lane 5, after scaffold-mediated ligation with lesion-containing insert; lane 6, after treatment with T4 DNA polymerase and exonuclease I; lane 7, purified N^2-furfuryl-dG M13mp7L2. Treatment of mock-ligated (using scaffolds, but no insert) substrate shows that digestion of the hairpin is nearly complete: lane 8, M13mp7L2, after the mock ligation; lane 9, after treatment with T4 polymerase and exonuclease I; lane 10, the same, but loading ten times the amount as lane 9.

and 80 U exonuclease I (NEB). The DNA was finally purified by sequential phenol/chloroform/isoamyl alcohol and chloroform extractions, ethanol precipitation, and was dissolved in 50 μL 10 mM Tris buffer to obtain the lesion-containing (or control) single-stranded phage DNA.

The progress of DNA construction was monitored by taking samples at each step and separating them on a 0.8% TAE agarose gel stained with ethidium bromide (Figure 2.4). As a control to guarantee that EcoRI completely linearized the hairpin-containing M13mp7(L2) ssDNA, which, if uncut, could contaminate lesion-containing DNA, a mock ligation reaction was performed with digested ssDNA and scaffolds, but without the insert. The mock reaction was then treated with T4 DNA polymerase and Exo I, which together degrade linear, but not circular, ssDNA. Degradation was nearly complete, demonstrating efficient digestion of the hairpin (Figure 2.4)

Single-molecule DNA substrates made from lesion-containing M13mp7(L2) were prepared similarly to those made from undamaged M13mp18. N^2-furfuryl-dG lesion or control dG-containing M13mp7(L2) was annealed with mp7L2-AlwNI (Table 2.1) and digested with 20 U AlwNI (NEB) in 1X Buffer 4 at 37˚C for 1 hr. Linearized DNA was ligated to M13-5'-biotin and phosphorylated M13-3'-dig using the scaffolding oligonucleotides mp7L2-scaffold-1 and mp7L2-scaffold-2 (Table 2.1). The annealed oligonucleotide mp7L2-scaffold-1, near the 3' terminus of the linear M13 template, served as the primer for DNA synthesis, ~3150 nucleotides from the N^2-furfuryl-dG site.

2.2.3 Measuring DNA extension by flow stretching

Single-molecule experiments were performed using custom microfluidic flow cells, constructed as previously described (18) with glass coverslips functionalized with a ratio of biotinylated polyethylene glycol succinimidyl valerate (PEG-SVA) and methyl-PEG-SVA (Laysan Bio) of 0.75%:15% (w/v) in 0.1 M $NaHCO_3$ pH 8.2. Prior to an experiment, the flow cell was incubated with 0.2 mg mL^{-1} streptavidin (Sigma) in PBS for 30 min, then washed and incubated with blocking buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 2 mM EDTA, 0.2 mg mL^{-1} BSA, and 0.005% Tween 20) for an additional 30 min. 2-4 μL of M13 substrate stock (~5 nM, see above) were diluted with 500 μL blocking buffer and drawn into the flow cell at 0.025 mL min^{-1} with a syringe pump (Harvard Apparatus 11 Plus), allowing binding of DNA by the 5'-biotinylated ends to immobilized streptavidin sites. A stock of α-digoxigenin-functionalized polystyrene beads (tosyl-activated, 2.8 μm diameter, Dynal) was

prepared as previously described (18); 2 μL of the bead stock were diluted with 500 μL blocking buffer and drawn into the flow cell at 0.025 mL min^{-1} to specifically bind the 3'-digoxigenin-labeled DNA substrates. Excess beads and DNA were removed from the flow cell by washing with 1 mL of blocking buffer (> 100 volumes) at 0.035 mL min^{-1}.

Immediately prior to the primer synthesis reactions, ~150 μL of replication buffer was introduced to exchange buffer. In experiments using SSB in Section 2.3.4, this buffer included SSB and further served as a pre-incubation step to coat ssDNA. A solution of proteins and nucleotides in 500 μL replication buffer (see Section 2.2.5) was added at 0.015 mL min^{-1}. A magnet exerting a weak force of ~1 pN was used to lift the tethered paramagnetic beads off the surface; laminar flow at this rate through the flow cell exerts a constant force of ~3 pN on the tether. After 2 min to allow the flow to stabilize, several hundred beads were visualized using dark-field microscopy through a 10X objective (Olympus) and imaged with a QIClick CCD camera (Q-Imaging). Data were recorded for 2750 frames at 2 Hz using the software package Micro-Manager (www.micro-manager.org). Primer extension on each individual molecule was observed by the motion of its bead in the direction of flow as coiled ssDNA was converted to extended dsDNA, as previously indicated. Synthesis was not observed when dNTPs were excluded.

During primer extension experiments, individual beads were fit to two-dimensional Gaussians and tracked with high accuracy (σ ~20 nm) using the software package DiaTrack (Semasopht); a bead nonspecifically stuck to the surface were used to subtract drift uniformly from all trajectories. Raw data for bead

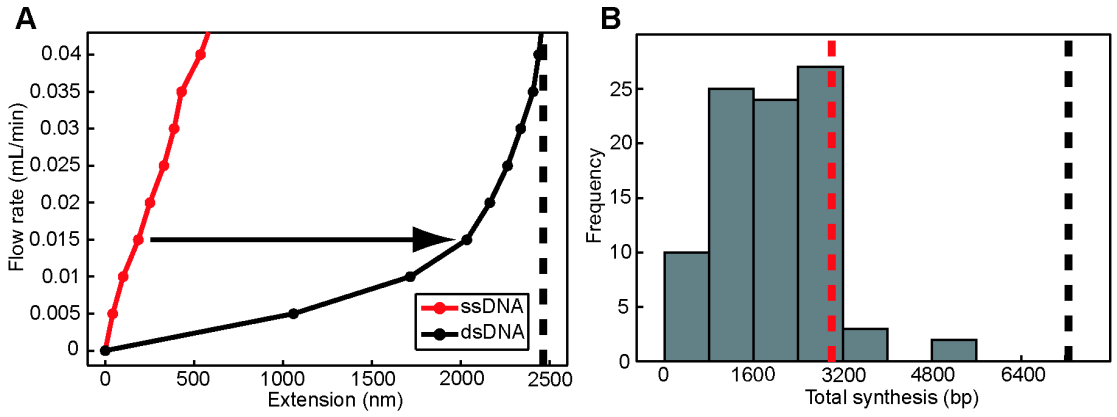

Figure 2.5 (A) The differential extensions of ssDNA and dsDNA at increasing flow rates. dsDNA was generated from the single-molecule substrate stock using ϕ29 DNA polymerase (New England Biolabs, NEB). The conversion factor (3.9 bp nm^{-1}), used to calculate DNA synthesis from bead displacement, was determined by dividing the total substrate length, 7249 bp, by the difference in the extension of dsDNA and ssDNA at 0.015 mL min^{-1}, 1848 nm. **(B)** Annealing a 3'-dideoxy-terminated oligonucleotide onto the ssDNA substrate at the +3000 bp position blocks synthesis by T7 DNA polymerase exo^- in 95% of trajectories (N = 91), confirming the conversion factor. The predicted location of the oligonucleotide block is marked with a dotted red line, and the full length of the substrate is marked with a dotted black line.

displacement in nanometers were converted into the number of base pairs synthesized using a calibration factor of 3.9 bp nm^{-1}, determined by dividing the substrate length by the differential extension of ssDNA and dsDNA at the flow rate used (Figure 2.5A).

To confirm the calibration factor, and demonstrate the ability to observe a site-specific block of replication in the single-molecule primer extension assay, the dideoxy chain-terminated oligonucleotide dideoxy-M13-block (Table 2.1) was annealed onto the M13mp18 single-molecule substrate, ~3000 bp from the primer terminus. Synthesis was performed with T7 DNA polymerase exo^-, a gift from Charles Richardson; primer extension terminated at the expected location (Figure 2.5B).

For experiments with SSB, which partially extends ssDNA, a different

calibration factor was used. This factor was determined by introducing a solution of SSB (1 μM, as a tetramer) in replication buffer into the flow cell and measuring the extension of the ssDNA substrates. Correcting for the new differential extension between the ssDNA-SSB filament and dsDNA, the new factor was calculated to be 5.6 bp nm^{-1}, in close agreement with the value of 5.7 bp nm^{-1} obtained by the van Oijen lab in a similar experimental set-up (data not published).

2.2.4 Single-molecule data analysis

Single-molecule trajectories were selected where the tethered DNA length increased in the direction of flow (*y*) but not in the transverse direction (*x*). Trajectories that had a rapid, simultaneous jump in both *x* and *y* represented sticking or unsticking of the bead to the surface of the flow cell and were excluded from analysis.

Synthesis trajectories were fit to segmented lines, with each segment corresponding to a Pol III event, a Pol IV event, or a pause using custom MATLAB code (Mathworks). Initial estimates for boundaries between segments were selected manually. The middle 80% of each region was then fit to a line, and new segment boundaries were determined from the intersection between adjacent segments. The processivity and rate for a segment are defined as the rise and slope, respectively.

Statistically significant synthesis events were defined as having a processivity greater than 3σ of the trajectory's noise (determined for individual trajectories, but generally ~200 bp); events were otherwise defined as pauses. For experiments with both Pol III and Pol IV present, a cutoff of 45 bp s^{-1} was used to

assign significant synthesis events to Pol III (faster) or Pol IV (slower). This cutoff captures 93% of Pol III events and 95% of Pol IV events in experiments with individual polymerases (see below).

Single molecule data were binned to generate distributions, and were normalized to integrated counts, generating probability densities to facilitate comparisons. Fits of normalized histograms to one-term exponentials of the form $A*\exp(-x/\lambda)$ for processivities and $A*\exp(-t/\tau)$ for pauses were determined using the MATLAB command 'fit,' which generated the exponential fit constant (τ or λ). Statistical comparisons were made between full datasets with the two-tailed Wilcoxon rank-sum test, using the MATLAB function 'ranksum,' and a significance level of $P < 0.05$, using the Bonferroni correction for multiple sample comparisons.

2.2.5 Proteins and buffers

Experiments were performed in replication buffer (50 mM HEPES-KOH pH 7.9, 12 mM $Mg(OAc)_2$, 80 mM KCl, and 0.1 mg mL^{-1} bovine serum albumin), with 5 mM DTT, 1 mM ATP, 760 μM dNTPs, 15 nM $\tau_3\delta\delta'\chi\psi$, 30 nM clamp (either β, β^R, or β^+/β^C, as dimers), and the indicated concentrations of Pol III and/or Pol IV. 60 μM dNTPs were used for single-molecule lesion bypass experiments. Single-stranded binding protein (SSB) was excluded from primer extension experiments unless otherwise noted, as SSB extends ssDNA at low force, reducing the contrast with dsDNA that is used to observe replication (19).

E. coli proteins were purified from overproducing strains as previously described and were untagged unless otherwise noted: Pol IV (20) and Pol IV^C (ΔC5)

(21); the Pol III holoenzyme subunits α, δ, and δ' (22); ε and θ (23); the wildtype clamp β (24), N-terminally his_6- and heart muscle kinase (HMK)-tagged β^R (E93K-L98K), and β^+/β^C, a stable dimer formed from N-terminally Myc-tagged β and his_6- and HMK-tagged β(ΔC5) (25); τ and refolded ψ within the χψ complex (11); and SSB (26). The Pol III αεθ core and clamp loader assembly with the stoichiometry $\tau_3\delta\delta'\chi\psi$ were each reconstituted and purified following reported protocols (11).

2.3 Results

2.3.1 Primer extension by individual polymerases

Using the flow-stretching single-molecule approach (Figure 2.6A), I characterized primer extension by Pol III and Pol IV, each individually. Synthesis by either polymerase occurred in discrete steps of processive synthesis interspersed by pauses (Figure 2.6B). Distributions for the processivity (Figure 2.6C-D) and rate (Figure 2.7A) of each synthesis step were generated from a large number of events; the data were in agreement with previous single-molecule experiments for Pol III (11) and bulk data for Pol IV (13). Pauses between synthesis steps were exponentially distributed, consistent with a single rate-limiting step, and I observed that increasing the concentration of Pol III from 5 nM to 30 nM reduced the pause length (Figure 2.8A-C, time constant τ decreases from 19.7 to 12.4 s). Given that biophysical and structural data suggest that only one Pol III binds the clamp dimer (5, 27-29), this argues that pauses observed during synthesis result from stochastic dissociation of Pol III from the clamp and the diffusion-limited recruitment of a new polymerase from solution (11).

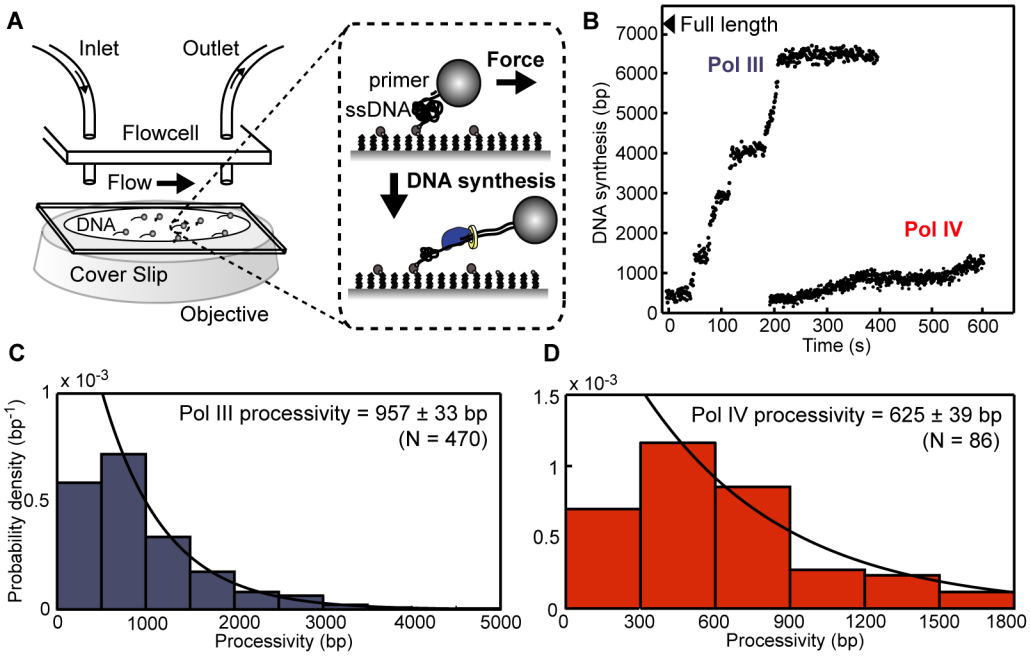

Figure 2.6 A single-molecule primer extension assay. **(A)** Under an applied force of ~3 pN, ssDNA is entropically collapsed, while dsDNA is extended to nearly its crystallographic length. Primer extension results in motion of tethered beads in the direction of flow. **(B)** Representative trajectories of synthesis by Pol III and Pol IV on individual DNA molecules. **(C, D)** Processivity distributions for Pol III (C, 5 nM) and Pol IV (D, 30 nM); values represent means +/- s.e.m.

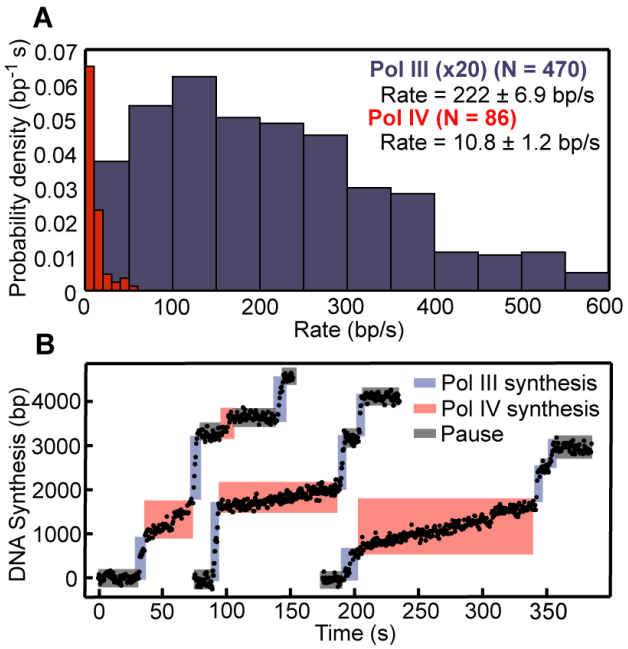

Figure 2.7 Observing exchange between Pol III and Pol IV. **(A)** Rate distributions for Pol III (blue) and Pol IV (red); values represent means +/- s.e.m. **(B)** Sample trajectories of rapid exchange between Pols III and IV.

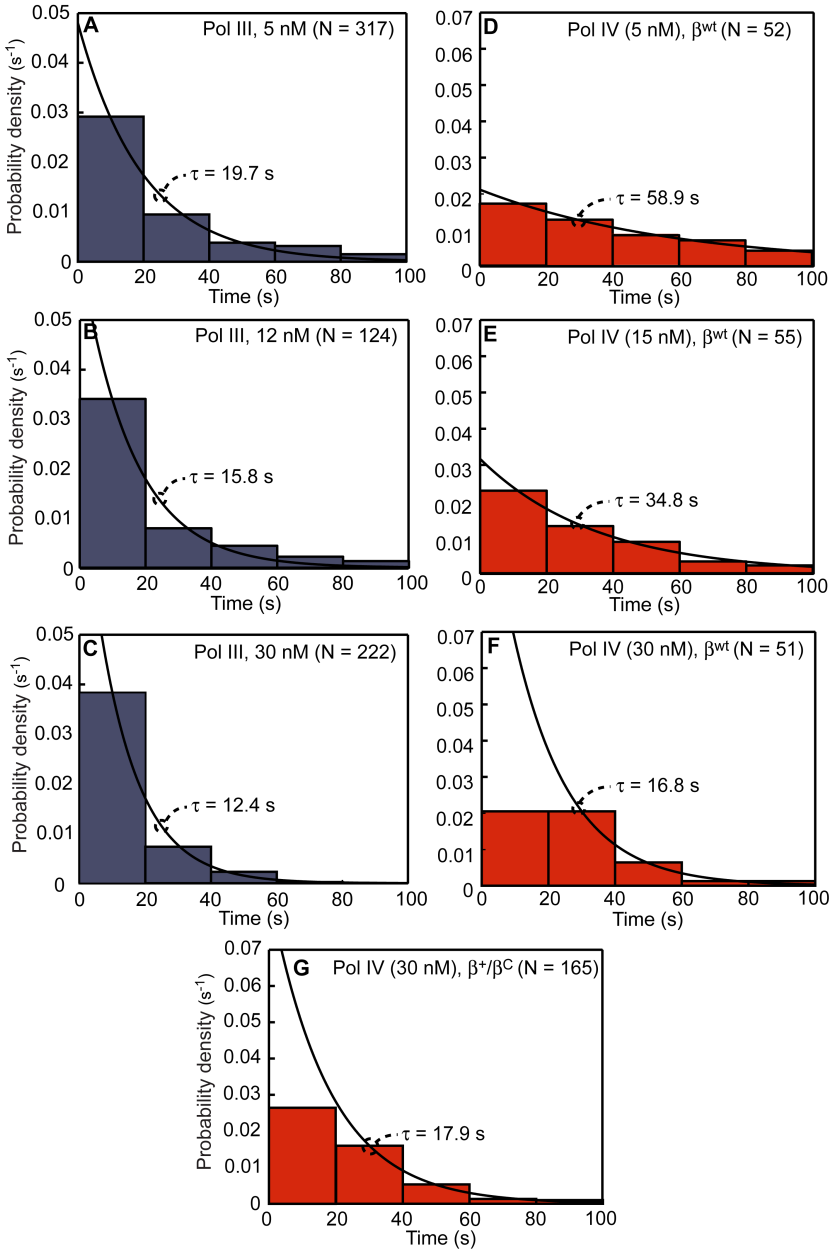

Figure 2.8 Effects of polymerase concentration and the number of clefts on pausing. Pauses between polymerase synthesis events in single-molecule trajectories are exponentially distributed, suggesting a single rate-limiting step, and inversely related with concentration; increasing Pol III from **(A)** 5 nM to **(B)** 12 nM and **(C)** 30 nM, or Pol IV from **(D)** 5 nM to **(E)** 15 nM and **(F)** 30 nM reduces these pause times. This demonstrates that pauses observed during synthesis by Pol III or Pol IV alone represent dissociation of a polymerase followed by the diffusion-limited recruitment of another from solution. Association times (pauses) of Pol III are shorter due to a greater k_a for clamp binding, measured in surface plasmon resonance experiments (14). Pauses observed in 30 nM Pol IV experiments with **(G)** the single-cleft clamp, β^{+}/β^{C}, are not significantly different from pauses observed in experiments with the wildtype clamp, β^{WT}. Observed pauses are therefore not due to switching between two Pol IV molecules potentially bound to the same β dimer. τ represents the exponential constant for fits to pause distributions; the first bins of D-G are under-sampled due to limits on the experimental resolution, and were therefore not used for fits.

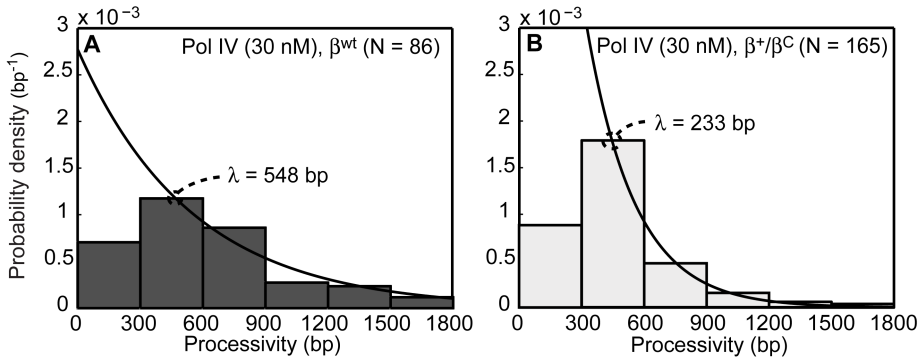

Figure 2.9 Effect of the number of clefts available on Pol IV processivity. The observed processivity of Pol IV is greater in experiments with **(A)** β^{WT} than with **(B)** β^{+}/β^{C} ($P < 10^{-3}$), supporting structural data that two Pol IV molecules can bind a single β (5) – the apparent processivity is artificially increased by rapid, unresolvable switches between two polymerases bound to the clamp. λ represents the exponential constant for fits to processivity distributions; the first bins are under-sampled due to limits on the experimental resolution, and were therefore not used for fits.

In contrast, results from structural and biophysical experiments suggest that two Pol IV molecules may simultaneously bind to the dimeric β (5, 30). To test this, I used a mutant clamp with a single binding cleft, β^{+}/β^{C} (25). While increasing the concentration of Pol IV from 5 nM to 30 nM also decreased pauses between Pol IV synthesis steps (Figure 2.8D-F, τ decreases from 58.9 to 16.8 s), pausing was not affected by the use of β^{+}/β^{C} (Figure 2.8G). The Pol IV processivity, however, dropped almost in half in experiments with β^{+}/β^{C} (Figure 2.9). Together, this implies that two Pol IV molecules can occupy β simultaneously, but that exchange between the two occurs on a timescale faster than our resolution, increasing the apparent processivity. Similar to Pol III, the concentration-dependent pauses observed result from recruitment of a Pol IV molecule to the clamp from solution.

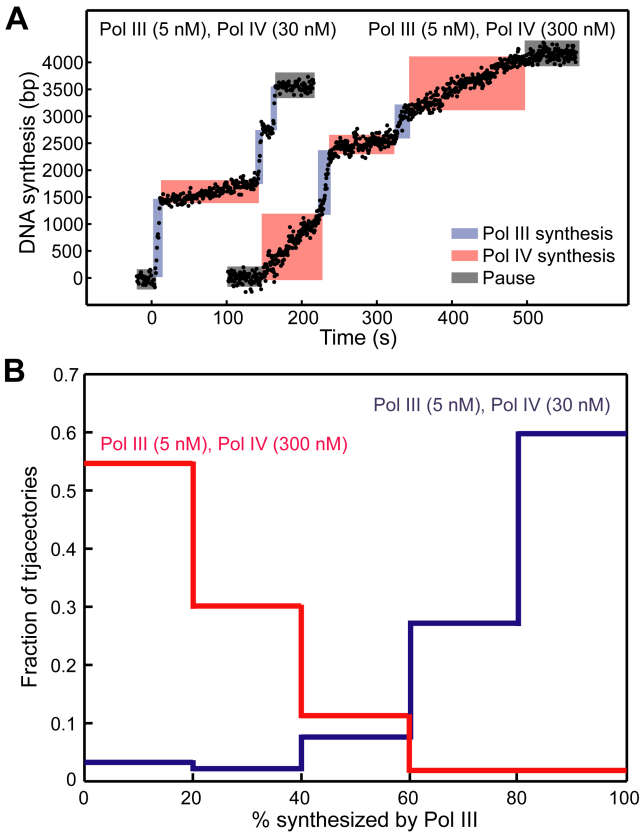

Figure 2.10 Effects of increasing concentrations of Pol IV on competition with Pol III. **(A)** Example trajectories at a low and high ratio of Pol IV to Pol III with assignments. At high concentrations of Pol IV, individual pauses are no longer observable. **(B)** Distributions for the percent of synthesis completed by Pol III in individual trajectories in the presence of low (30 nM Pol IV, average of 78% Pol III synthesis, N = 92 trajectories) and high (300 nM Pol IV, 22% Pol III synthesis, N = 53) levels of Pol IV.

2.3.2 Observation of Pol III-Pol IV exchange and lesion bypass

The dramatically different rates of the two polymerases (Figure 2.7A) enable assignment of synthesis events to either Pol III or Pol IV. I therefore performed primer extension with a mixture of Pol III (5 nM) and Pol IV (30 nM). This ratio was chosen to approximate that found in cells during exponential growth (31), with concentrations reduced (from about 20 nM for Pol III and 300 nM for Pol IV) so that distinct synthesis events could be resolved. If the fraction of active protein differs for each, then the molar ratio of active polymerases will be shifted by a constant

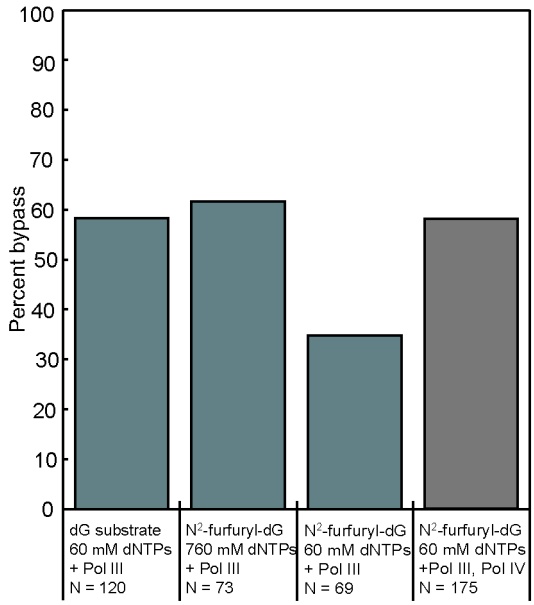

Figure 2.11 The percentage of trajectories that bypass the lesion site in single-molecule experiments; bypass is defined as synthesis past the site plus 3σ of the noise (3350 bp). Different conditions include lesion and control substrates, at low and high dNTP concentrations, and with or without Pol IV.

factor. Under these conditions, Pol III performed 78% of DNA synthesis (Figure 2.10), likely due to stronger interactions with β (31); however, one or more Pol IV events were also observed in 75% of trajectories (Figure 2.7B), exchanging with Pol III.

To observe polymerase exchange in the physiological context of a DNA lesion, I used the M13 substrate containing a site-specific N^2-furfuryl-dG adduct, a 'cognate lesion' for Pol IV. N^2-furfuryl-dG is a minor groove lesion that is efficiently and accurately bypassed by Pol IV, and an analog of the primary adduct formed by the antibiotic nitrofurazone, an agent to which Pol IV knockout strains are significantly sensitive (17).

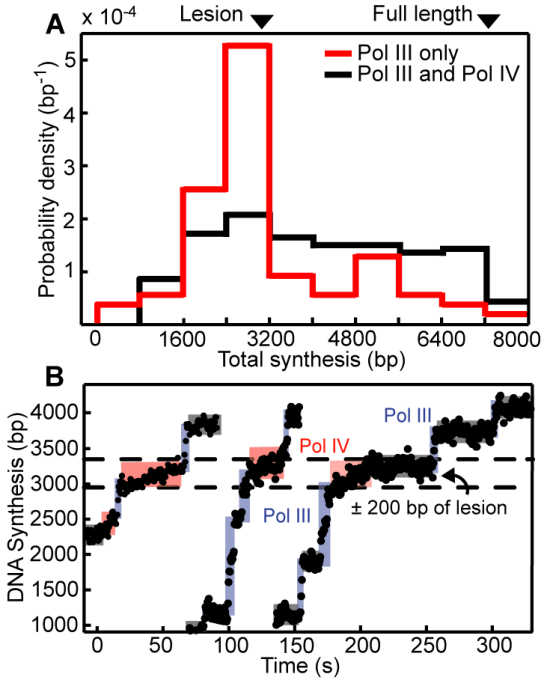

Figure 2.12 A single-molecule reconstitution of polymerase exchange and bypass at a DNA lesion. **(A)** An N^2-furfuryl-dG lesion, at ~3150 bp, blocks processive synthesis by Pol III (5 nM, N = 69), but is bypassed when Pol IV (30 nM) is added (N = 175). **(B)** Rapid exchange from Pol III to Pol IV and back is observed at the lesion site.

Although the lesion strongly blocked Pol III in ensemble synthesis in the absence of the clamp (Figure 2.3), I found that it blocked only 65% of trajectories in single-molecule experiments (Figure 2.11 and Figure 2.12A). Previous studies have shown Pol III lesion bypass efficiency is strongly promoted by β (32) and by increased dNTP levels, which biases polymerase over exonuclease activity (33); I indeed observed that higher dNTP levels increased bypass (Figure 2.11). The addition of both polymerases to the primer extension reaction alleviated the block at the N^2-furfuryl-dG position (Figure 2.12A) and revealed polymerase exchange at the lesion site and bypass by Pol IV (Figure 2.12B).

2.3.3 Kinetics of exchange support the toolbelt model

The observation of exchange from Pol III to Pol IV and back is uniquely accessible in this single-molecule reconstitution and permits the investigation of the role of β in polymerase trafficking. In the toolbelt model, polymerase exchange is only limited by the timescale of conformational changes of Pol III and Pol IV simultaneously bound to β. In an alternative model, in which steric effects prevent concurrent binding, exchange requires dissociation of the first polymerase, followed by recruitment of the second from solution, which would be sensitive to protein dilution. Quantifying exchange by measuring the time between the termination of synthesis by one polymerase and the subsequent initiation of synthesis by the other allows us to distinguish between these two models.

The time for exchange from Pol III to Pol IV on undamaged DNA (Figure 2.13C) was more rapid than the diffusion-limited recruitment time of Pol IV from solution (Figures 2.13A and 2.8), seen in exponential fits and a statistical comparison of the two data sets ($P < 10^{-5}$). Furthermore, reducing the concentration of Pol IV in exchange experiments from 30 to 15 nM did not affect the timescale of exchange (Figure 2.14), while the same dilution increased pause times in experiments with Pol IV alone (Figures 2.13A and 2.8E). This argues that exchange during active synthesis occurs between two polymerases bound to the clamp. Our observation of β-mediated exchange implies the second Pol III CBM, in the ε subunit, does not exclude Pol IV from binding the clamp in the absence of a lesion-induced stall, in contrast to a previous suggestion (6); rather, Pol IV can compete with ε for a cleft, allowing Pol IV to bind β while Pol III is synthesizing DNA.

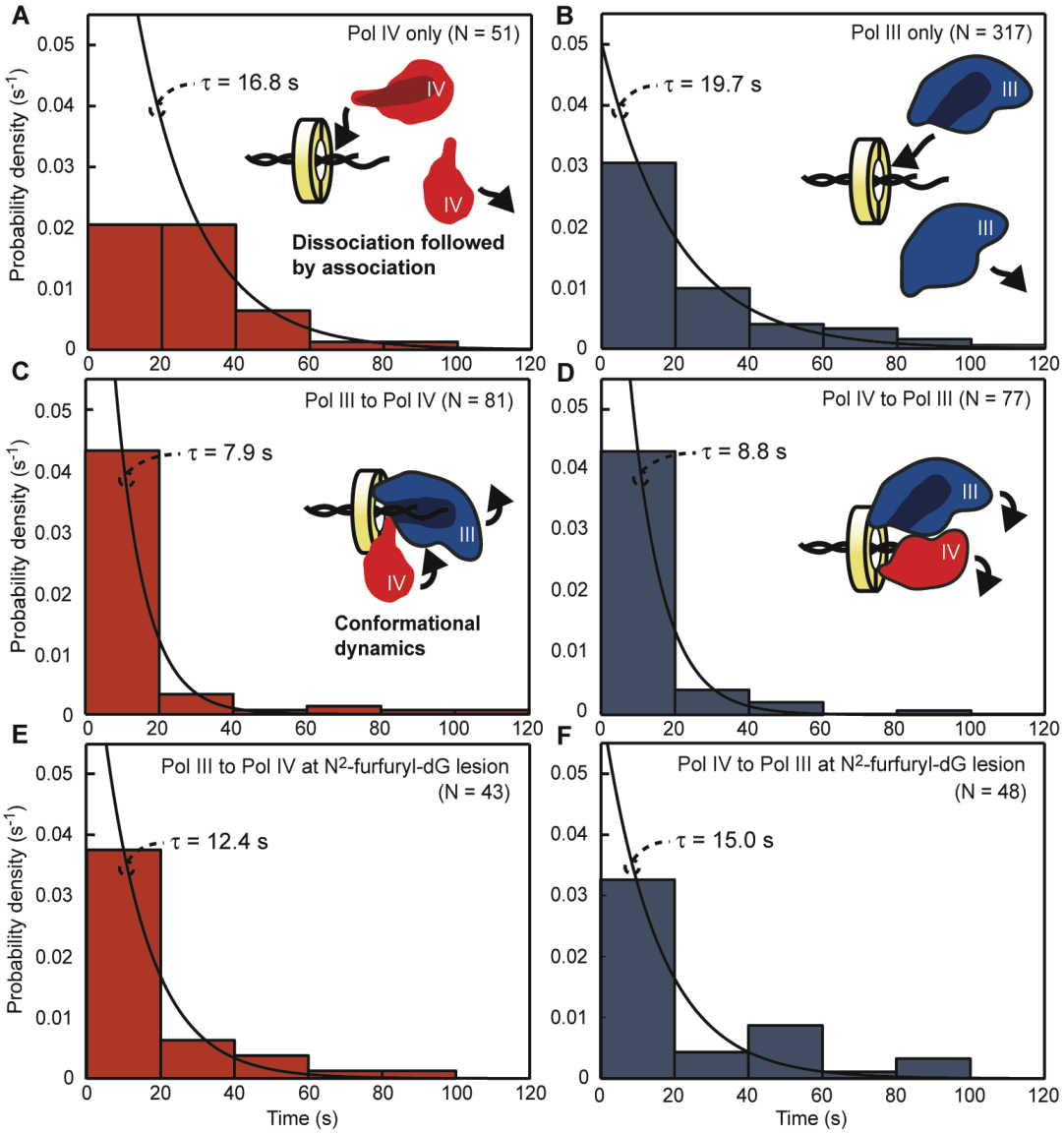

Figure 2.13 β acts as a molecular toolbelt for Pols III and IV to promote lesion bypass. Each distribution is fit to an exponential with associated time constant τ. Pauses represent association of a new polymerase from solution in experiments with Pol IV (**A**, 30 nM) or Pol III alone (**B**, 5 nM). **(C)** Exchange from Pol III (5 nM) to Pol IV (30 nM) ($P < 10^{-5}$ vs. A) and **(D)** back to Pol III ($P < 10^{-9}$ vs. B) is rapid due to simultaneous binding of both polymerases to the clamp. **(E)** Exchange to Pol IV at the N^2-furfuryl-dG site is also rapid ($P < 0.01$ vs. A, *NS* vs. C), indicating that lesion bypass is β-mediated. **(F)** Exchange back to Pol III is intermediate between B ($P = 0.04$) and D ($P = 0.02$), suggesting that both types of exchange occur.

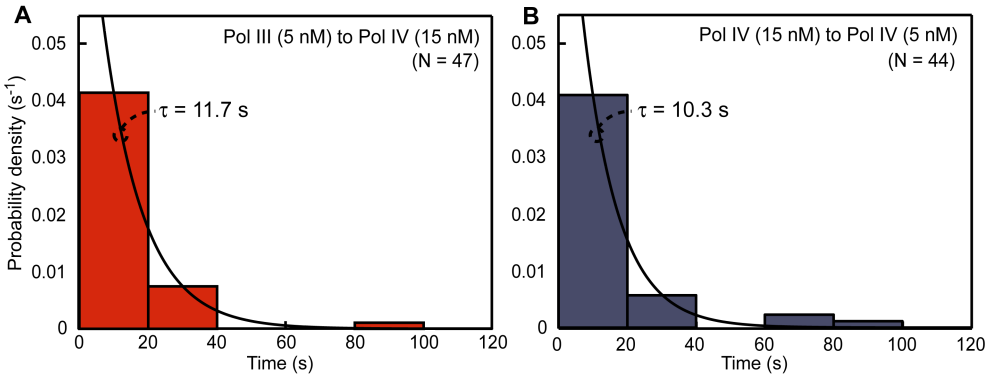

Figure 2.14 Reducing the Pol IV concentration from 30 nM to 15 nM results in a timescale of exchange with Pol III statistically indistinguishable from Figure 2.13C, despite the same dilution increasing the association time in experiments with Pol IV alone (Figure 2.8E,F). Exchange from **(A)** Pol III (5 nM) to Pol IV (15 nM) is more rapid than pauses in experiments with 15 nM Pol IV ($P < 10^{-5}$ vs. Figure 2.8E); exchange from **(B)** Pol IV to Pol III remains rapid ($P < 10^{-7}$ vs. Figure 2.13A).

Importantly, I observed that exchange back to Pol III following Pol IV synthesis was also more rapid than the recruitment time of Pol III from solution (Figure 2.13B,D, $P < 10^{-9}$). Furthermore, exchange from Pol IV to Pol III in the presence of the single-cleft β^{+}/β^{C} (Figure 2.15, τ = 27.3 s) was slower than in experiments with the wildtype clamp ($P < 10^{-3}$), closely matching the recruitment time of Pol III from solution (*NS* vs. Figure 2.13B). These data demonstrate that Pol III can bind the opposing β cleft in an inactive conformation while Pol IV is carrying out synthesis, and that eliminating the second cleft abolishes rapid exchange. The fact that the Pol IV processivity preceding exchange to Pol III matches that of Pol IV on β^{+}/β^{C} (Figures 2.9B and 2.15B) shows that a single Pol IV is bound to β in the presence of Pol III and strongly implies that Pol III does not displace Pol IV during the exchange back, but takes over after Pol IV stochastically releases DNA.

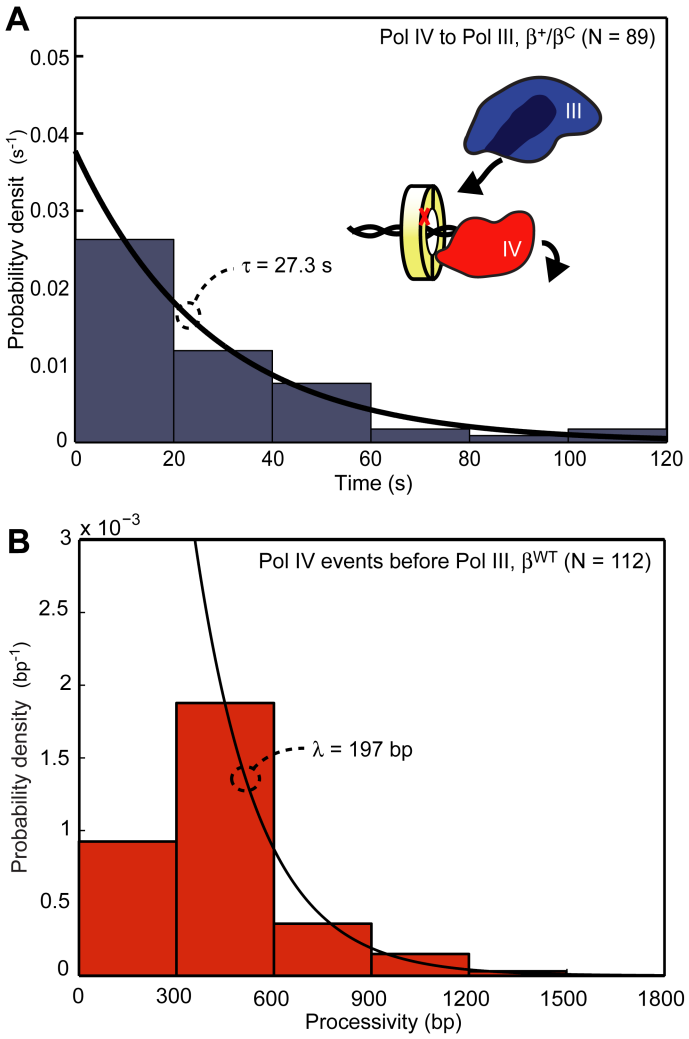

Figure 2.15 Stoichiometry of Pol III and Pol IV on the clamp during rapid exchange. **(A)** Exchange from Pol IV (5 nM) to Pol III (30 nM) is diffusion-limited (*NS* vs. Figure 2.13B; $P < 10^{-3}$ vs. Figure 2.13D) in experiments with the single-cleft clamp, β^{+}/β^{C}, further demonstrating that Pol III requires access to the free cleft to bind the clamp during Pol IV synthesis. **(B)** The processivity of Pol IV events preceding exchange to Pol III is less than the processivity of Pol IV alone (30 nM) with β ($P < 10^{-3}$ vs. Figure 2.9A) and not significantly different from Pol IV with β^{+}/β^{C} (Figure 2.9B), demonstrating that a single Pol IV, at low concentrations, binds the clamp during exchange with Pol III.

To test if polymerase exchange can occur in the physiological context of Pol III encountering a DNA lesion, I collected data for exchange from Pol III to Pol IV within the experimental resolution (±200 bp) of the N^2-furfuryl-dG position (Figure 2.12B). These exchange times (Figure 2.13E) matched those for exchange on undamaged DNA (*NS* vs. Figure 2.13C) and were more rapid than recruitment of Pol

IV from solution ($P < 0.01$ vs. Figure 2.13A), indicating that Pol IV was bound to β when Pol III encountered the lesion. The switch back to Pol III after the lesion (Figure 2.13F) was intermediate between rapid, β-mediated exchange and recruitment from solution, suggesting a mixture of these two types of exchange.

2.3.4 Binding of Pol IV at a secondary site on β reduces the Pol III processivity

These results support a model in which Pol IV, at a ratio to Pol III consistent with normal growth, can bind β in an inactive conformation, thereby promoting rapid bypass of DNA lesions encountered during synthesis. During the SOS DNA damage response, however, the cellular concentration of Pol IV increases roughly 10-fold, while Pol III levels remain constant (31). To test if an increased ratio of Pol IV to Pol III alters polymerase exchange, I performed primer extension experiments with 5 nM Pol III and 300 nM Pol IV. At these concentrations, Pol IV outcompeted Pol III, performing 78% of DNA synthesis (Figure 2.10). Although the average processivity of Pol III synthesis preceding exchange was only modestly affected by Pol IV under normal conditions, possibly due to disruption of the binding of the ε subunit of Pol III to β (Figure 2.16A), under SOS-like conditions it dropped almost in half (Figure 2.17A, $P < 10^{-5}$), reflecting a decrease of the lifetime of Pol III on the clamp (Figure 2.16B). This reduction in processivity was dose-dependent with the Pol IV concentration and β-mediated; Pol IV^{C}, a mutant that lacks its CBM, did not reduce the Pol III processivity (Figure 2.17A).

To further define which interactions with β mediate this activity, I tested the effects of mutant clamps under SOS-like conditions. β^{R}, a clamp mutant that

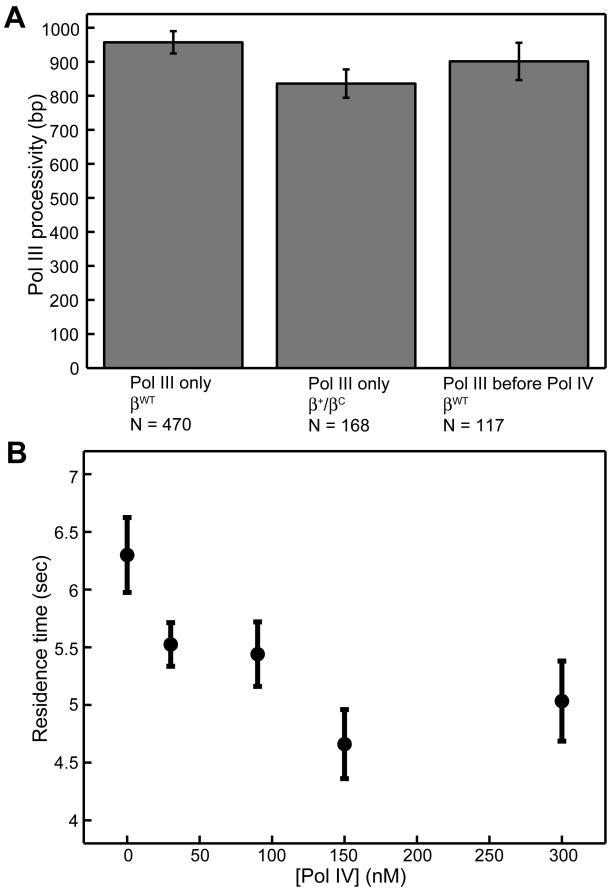

Figure 2.16 Effects of Pol IV at low and high concentrations on Pol III stability. **(A)** Pol III processivity is modestly reduced by the absence of the ε-β contact, and similarly by the presence of Pol IV on the clamp, suggesting Pol IV can compete with ε for a cleft. The average Pol III processivity in the following experiments (*left to right*): (1) Pol III only (5 nM) in experiments with the wild type clamp, β^{WT}; (2) Pol III only (5 nM) with the single-cleft clamp, β^{+}/β^{C}, which eliminates binding by the weaker ε exonuclease subunit CBM; (3) Pol III (5 nM) events that precede Pol IV (30 nM) events with the wild type cleft. A similar reduction in Pol III processivity was previously observed in single-molecule leading strand experiments when the ε CBM was weakened (27). **(B)** The average lifetime of Pol III on the clamp for individual synthesis events decreases with increasing concentrations of Pol IV. The lifetime for each event was determined by dividing the processivity by the rate.

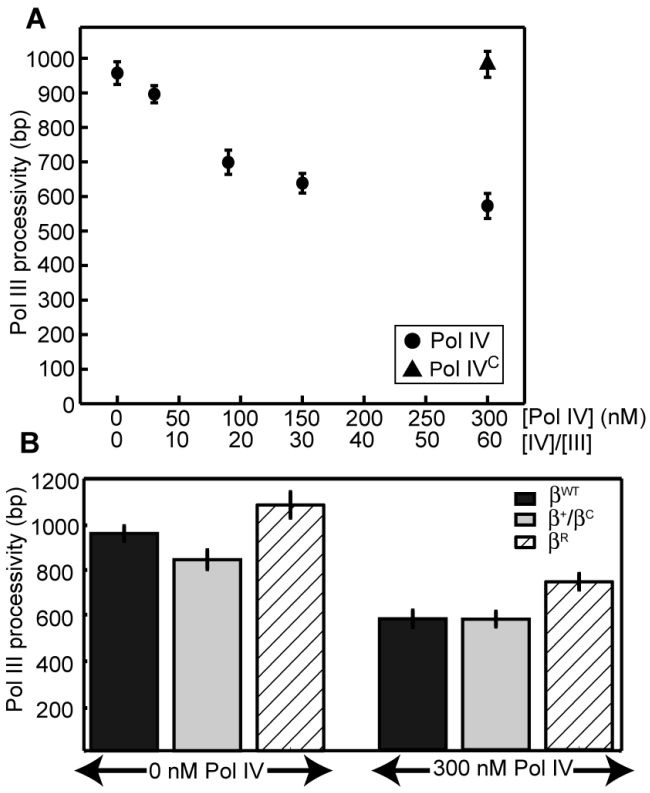

Figure 2.17 Binding of Pol IV at a secondary site on β reduces the Pol III processivity. **(A)** The Pol III processivity decreases with increasing concentrations of Pol IV (circles), but is unaffected by Pol IV^{C}, a mutant lacking its CBM (triangle). **(B)** Reduction in the Pol III processivity with β^{WT} (black) also occurs with the single-cleft mutant clamp, β^{+}/β^{C} (grey, *NS*), but is partially alleviated by β^{R}, a clamp with a weakened Pol IV-interacting rim interface (cross-hatched, $P < 0.01$). N ranges from 71 to 470 for A and B. Values represent means +/- s.e.m.

weakens the secondary Pol IV-β interaction at the 'rim' site (5), did not affect the synthesis of Pol III alone, but partially restored the Pol III processivity at a high Pol IV concentration (Figure 2.17B, $P < 0.01$). While the processivity of Pol III was reduced slightly with the single-cleft β^{+}/β^{C}, consistent with the model that the ε subunit stabilizes it on the clamp (6, 27), a high concentration of Pol IV with β^{+}/β^{C} reduced the Pol III processivity equivalently to the wildtype clamp condition (Figure 2.17B), further supporting a role for the non-cleft rim contact in Pol III displacement.

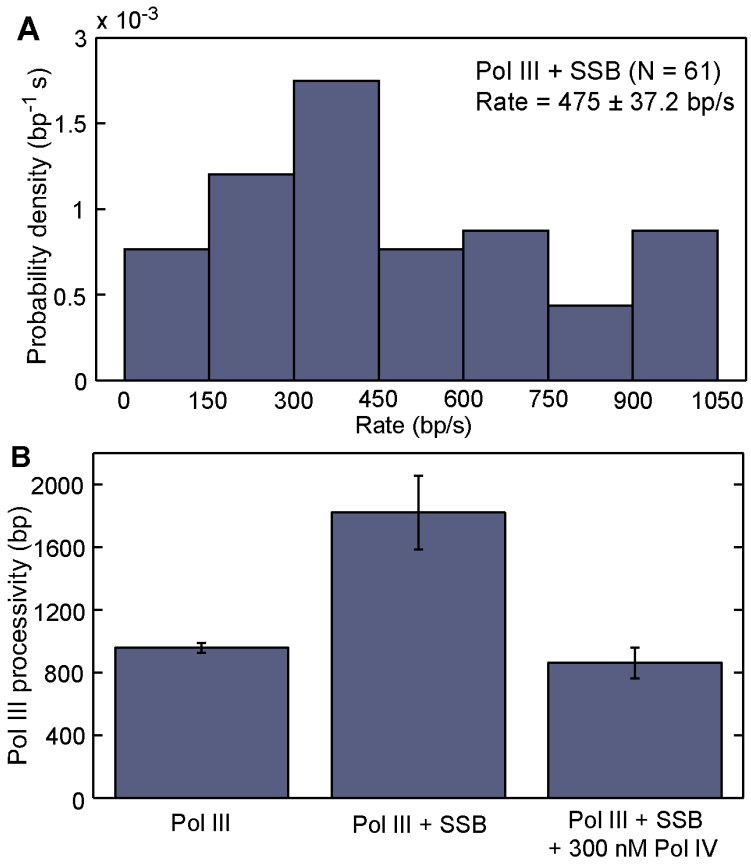

Figure 2.18 (A) SSB (1 μM, as tetramer) increases the rate of Pol III ($P < 0.001$ vs. Figure 2.7A) but not Pol IV (data not shown). **(B)** SSB increases the processivity of Pol III ($P < 0.001$), which 300 nM Pol IV reduces ($P < 0.001$). *N* is 470, 61, and 79, respectively. Values represent means ± SEMs.

Although SSB was excluded from the experiments above – it extends ssDNA and reduces the contrast with dsDNA used to observe primer extension – I tested the ability of Pol IV to displace Pol III to further verify that Pol IV was acting via its binding to β, and not to naked ssDNA. Including SSB (1 μM, as a tetramer) in primer exchange reactions increased the processivity and rate of Pol III by about two-fold (Figure 2.18), but did not affect Pol IV (data not shown). Importantly, 300 nM Pol IV significantly reduced the Pol III processivity in the presence of SSB (Figure 2.18B).

2.4 Discussion

By visualizing the TLS reaction in its entirety at the single-molecule level, these data provide a comprehensive view of how Pol IV access to the replication fork is regulated through interactions with β (Figure 2.19). Pol IV, at relatively low concentrations during normal growth, is able to associate with the rim site and compete with the weakly bound ε subunit of Pol III for its cleft (Figure 2.13). By occupying the rim and cleft sites of β in an inactive mode during normal growth conditions, Pol IV is available for rapid exchange and translesion synthesis when Pol III stalls upon encountering a lesion, as proposed in the toolbelt model (34).

We have also demonstrated a novel inactive binding mode for Pol III that allows it to remain bound to the cleft of one β protomer until the switch back (Figures 2.13 and 2.15); this is the other half of the polymerase exchange reaction that has been difficult to resolve by bulk biochemical studies (4, 5). The Pol IV processivity preceding a switch back to Pol III is not reduced from that of Pol IV alone (Figure 2.15B), suggesting that Pol III does not actively displace Pol IV during translesion synthesis, but relies on the lower processivity of Pol IV to minimize the mutagenic load.

At higher concentrations of Pol IV, corresponding to the SOS damage response, I observed a decrease in the Pol III processivity (Figure 2.17A). I propose that this is due to an increased occupancy of Pol IV at the low-affinity rim sites of β (Figure 2.19). Adjacent to the Pol III α subunit, Pol IV would be positioned to dynamically replace α's strongly bound CBM during a transient release from the cleft, as seen in a molecular model of both polymerases bound to β (Figure 2.20).

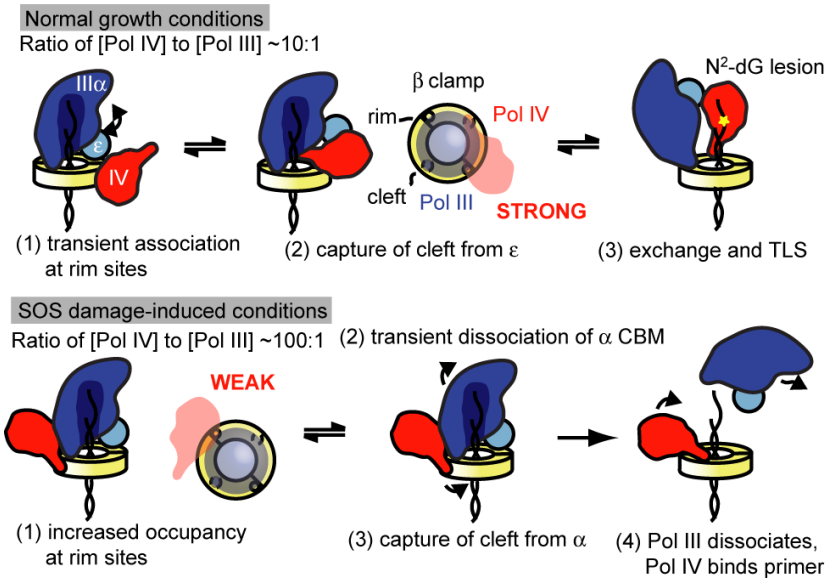

Figure 2.19 A model for how the Pol IV occupancy at rim sites and competition with Pol III subunits dictates polymerase exchange at different Pol IV concentrations. The small θ subunit of Pol III, which binds ε, is not represented for clarity.

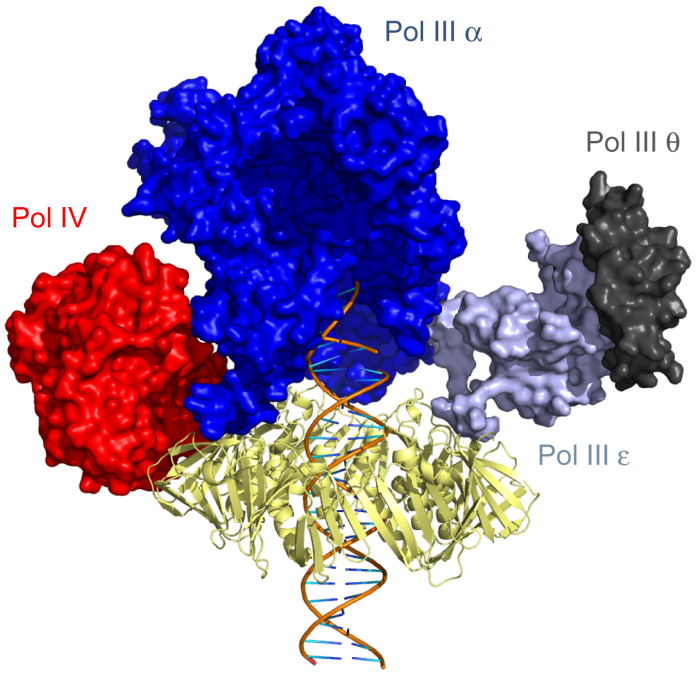

Figure 2.20 A model of Pol III and Pol IV bound to β (yellow), with Pol IV positioned at the rim site to capture the Pol III α CBM during a transient release from its cleft, showing that concurrent binding of the two polymerases is consistent with structural data for Pol III, Pol IV, and β. This model, constructed in PyMOL (Schrödinger), was based on a previously published model of the 'closed' Pol III αεθ complex bound to DNA and β (29). The structure of Pol IV (PDB 4IR9) (51) was docked onto the rim site of β using the co-crystal structure of the Pol IV little finger domain bound to β (PDB 1UNN) (30).

This role for the rim contact is consistent with biochemical results that β^{+}/β^{C} can support polymerase exchange to Pol IV when Pol III is stalled by nucleotide omission (5). Without the critical α-β contact following cleft capture by Pol IV, Pol III would dissociate from the primer terminus, allowing Pol IV to take its place. Unless the displaced Pol III could bind an adjacent, unoccupied protomer, or is stabilized by additional interactions, it would likely dissociate from the clamp entirely.

The requirement of the Pol IV CBM indicates that binding at rim sites is not sufficient for a reduction of the Pol III lifetime on DNA, and that Pol IV must also compete for the cleft bound by α. Shared contacts, such as a single cleft of β during competition between Pol III and Pol IV bound at additional sites on β, have been proposed to be important in facilitating dissociation and subunit exchange in multi-protein complexes upon transient contact release (35). This phenomenon has also been observed in the "dynamic processivity" of phage T4 and T7 replication, where additional polymerases are able to associate with moving replisomes and undergo exchange on a timescale faster than that of stochastic dissociation of the synthesizing polymerase (12, 36, 37), and in the facilitated dissociation of the *E. coli* DNA-binding protein Fis by nucleoid proteins (38). Secondary contacts, such as the rim site for Pol IV shown here, play an important role in orienting proteins to exploit transient changes in occupancy of these shared sites, resulting in binding partner exchange.

During coordinated leading and lagging strand replication, a displaced Pol III may remain associated with the replisome via additional contacts with the clamp loader complex. These contacts, however, do not appear to prevent Pol IV from

accessing β; previous biochemical experiments with a fully reconstituted replisome have demonstrated that Pol IV can replace Pol III (39). Furthermore, overexpression of Pol IV beyond SOS levels in cells has been shown to arrest replication and induce toxicity due to unregulated access of Pol IV to the replication fork (39-41). Removing the CBM residues alleviates Pol IV toxicity, while mutating the rim-contacting residues partially alleviates it (42). These data are explained by this model: contacts with the rim site and subsequently with the cleft provide a molecular path for Pol IV to displace Pol III from the primer terminus following SOS induction.

A putative interaction with the rim site could also position the other *E. coli* Y-family polymerase, Pol V, on the clamp, when it is expressed later in the SOS damage response (43). This binding activity would create a hierarchy for access to the primer terminus, a view of the toolbelt model in which clamp-polymerase interactions do more than merely increase the local polymerase concentration at the DNA template. During normal growth conditions, Pol III, which is preferentially loaded to the primer terminus (44), performs the majority of DNA synthesis. Pol IV is able to simultaneously bind β in an inactive mode, although it is currently unclear how other β-interacting proteins would influence its occupancy *in vivo*. Upon SOS induction, the rim site positions Pol IV to preferentially bind a cleft of β when it becomes available. Such a competitive advantage would ensure timely access of Y-family polymerases to the primer template and is likely to be important with several proteins competing for an open cleft.

Non-cleft contacts may play a similar role in regulating access to the DNA template in other domains of life. PCNA plays a key role in coordinating the hand-off

of DNA intermediates between a polymerase, flap endonuclease-1 (FEN1), and ligase during both Okazaki fragment maturation and long-patch base excision repair in eukaryotes (45). Structural studies have revealed that FEN1, Polymerase β, and Ligase I bind overlapping but distinct regions of DNA intermediates, which may facilitate displacement during handoff (46). In the archaeon *Sulfolobus solfataricus*, the three enzymes each bind distinct monomers of the heterotrimeric PCNA during Okazaki fragment maturation (47). However suggestive, it remains unknown if the homotrimeric eukaryotic PCNA can simultaneously bind any combination of these three proteins.

Eukaryotic TLS is regulated, in part, by ubiquitination of PCNA; all four human Y-family polymerases have ubiquitin-binding domains (48). Structural data of monoubiquitinated PCNA and its conformations support the model that the secondary ubiquitin site allows the TLS polymerase Pol η to bind an occupied clamp and positions it to compete for the cleft upon transient dissociation of the replicative polymerase (49, 50). In contrast to bacteria, where the occupancy of the rim site is controlled by polymerase concentration, analogous sites in eukaryotes are introduced by posttranslational modification. I anticipate that the single-molecule approaches described here will serve as powerful tools to elucidate the role of these interactions in translesion synthesis.

2.5 Contributions

Pol III replisome components were purified by Slobodan Jergic in the laboratory of Nicholas Dixon (University of Wollongong), while mutant clamp proteins were

provided by Mark Sutton (University of Buffalo). The N^2-furfuryl-dG-containing 20-mer oligonucleotide was constructed and purified by Deena Jacob in Graham Walker's laboratory at MIT, with help from Deyu Li. M13mp7(L2) phage stock was a gracious gift from John Essigman, also at MIT.

References

1. Clausen AR, *et al*. (2015) Tracking replication enzymology *in vivo* by genome-wide mapping of ribonucleotide incorporation. *Nat Struct Mol Biol* 22(3):185–191.

2. Kornberg A, Baker TA (1992) *DNA Replication*, 2nd (W.H. Freeman and Company, New York).

3. Benson RW, Norton MD, Lin I, Comb Du WS, Godoy VG (2011) An active site aromatic triad in *Escherichia coli* DNA Pol IV coordinates cell survival and mutagenesis in different DNA damaging agents. *PLoS ONE* 6(5):e19944.

4. Indiani C, McInerney P, Georgescu R, Goodman M, O'Donnell M (2005) A sliding-clamp toolbelt binds high-and low-fidelity DNA polymerases simultaneously. *Mol Cell* 19(6):805–815.

5. Heltzel J, Maul R, Scouten Ponticelli S, Sutton M (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc Natl Acad Sci USA* 106(31):12664.

6. Toste Rêgo A, Holding AN, Kent H, Lamers MH (2013) Architecture of the Pol III-clamp-exonuclease complex reveals key roles of the exonuclease subunit in processive DNA synthesis and repair. *EMBO J* 32(9):1334–1343.

7. Christian TD, Romano LJ, Rueda D (2009) Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc Natl Acad Sci USA* 106(50):21109–21114.

8. Shi X, *et al*. (2012) Quantitative fluorescence labeling of aldehyde-tagged proteins for single-molecule imaging. *Nat Meth* 9(5):499-503.

9. Wuite GJ, Smith SB, Young M, Keller D, Bustamante C (2000) Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* 404(6773):103–106.

10. Maier B, Bensimon D, Croquette V (2000) Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proc Natl Acad Sci USA* 97(22):12002–12007.

11. Tanner NA, *et al*. (2008) Single-molecule studies of fork dynamics in Escherichia coli DNA replication. *Nat Struct Mol Biol* 15(2):170–176.

12. Loparo JJ, Kulczyk AW, Richardson CC, van Oijen AM (2011) Simultaneous single-molecule measurements of phage T7 replisome composition and function reveal the mechanism of polymerase exchange. *Proc Natl Acad Sci USA* 108(9):3584-3589.

13. Wagner J, Fujii S, Gruz P, Nohmi T, Fuchs R (2000) The β clamp targets DNA polymerase IV to DNA and strongly increases its processivity. *EMBO Reports* 1(6):484–488.

14. van Oijen AM, Loparo JJ (2010) Single-molecule studies of the replisome. *Annu Rev Biophys* 39(1):429–448.

15. Tanner NA, *et al*. (2008) Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Res* 37(4):e27–e27.

16. Delaney JC, Essigmann JM (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions *in vivo*. *Methods Enzymol* 408:1–15.

17. Jarosz DF, Godoy VG, Delaney JC, Essigmann JM, Walker GC (2006) A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature* 439(7073):225–228.

18. Tanner NA, van Oijen AM (2010) Visualizing DNA replication at the single-molecule level. *Methods Enzymol* 475:259–278.

19. Fu H, Le S, Chen H, Muniyappa K, Yan J (2013) Force and ATP hydrolysis dependent regulation of RecA nucleoprotein filament by single-stranded DNA binding protein. *Nucleic Acids Res* 41(2):924–932.

20. Beuning PJ, Simon SM, Godoy VG, Jarosz DF, Walker GC (2006) Characterization of *Escherichia coli* translesion synthesis polymerases and their accessory factors. *Methods Enzymol* 408:318-340.

21. Maul RW, Scouten Ponticelli SK, Duzen JM, Sutton MD (2007) Differential binding of *Escherichia coli* DNA polymerases to the β-sliding clamp. *Mol Microbiol* 65(3):811–827.

22. Wijffels G, *et al*. (2004) Inhibition of protein interactions with the β sliding clamp of Escherichia coli DNA polymerase III by peptides from β-binding proteins. *Biochemistry* 43(19):5661–5671.

23. Hamdan S, *et al*. (2002) Hydrolysis of the 5'-p-nitrophenyl ester of TMP by the proofreading exonuclease (ε) subunit of Escherichia coli DNA polymerase III. *Biochemistry* 41(16):5266–5275.

24. Oakley AJ, *et al*. (2003) Flexibility revealed by the 1.85 Å crystal structure of the beta sliding-clamp subunit of Escherichia coli DNA polymerase III. *Acta Crystallogr D Biol Crystallogr* 59(7):1192–1199.

25. Scouten Ponticelli SK, Duzen JM, Sutton MD (2009) Contributions of the individual hydrophobic clefts of the *Escherichia coli* β sliding clamp to clamp

loading, DNA replication and clamp recycling. *Nucleic Acids Res* 37(9):2796–2809.

26. Mason CE, *et al*. (2013) *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(2):274–285.

27. Jergic S, *et al*. (2013) A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J* 32(9):1322-1333.

28. Lamers MH, Georgescu RE, Lee S-G, O'Donnell M, Kuriyan J (2006) Crystal structure of the catalytic α subunit of *E. coli* replicative DNA polymerase III. *Cell* 126(5):881–892.

29. Ozawa K, *et al*. (2013) Proofreading exonuclease on a tether: the complex between the *E. coli* DNA polymerase III subunits α, ε, θ and β reveals a highly flexible arrangement of the proofreading domain. *Nucleic Acids Res* 41(10):5354-5367.

30. Bunting K, Roe S, Pearl L (2003) Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the β-clamp. *EMBO J* 22(21):5883–5892.

31. Sutton MD (2010) Coordinating DNA polymerase traffic during high and low fidelity synthesis. *Biochim Biophys Acta* 1804(5):1167–1179.

32. Tomer G, Reuven NB, Livneh Z (1998) The β subunit sliding DNA clamp is responsible for unassisted mutagenic translesion replication by DNA polymerase III holoenzyme. *Proc Natl Acad Sci USA* 95(24):14106–14111.

33. Gon S, Napolitano R, Rocha W, Coulon S, Fuchs RP (2011) Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci USA* 108(48):19311–19316.

34. Pagès V, Fuchs RP (2002) How DNA lesions are turned into mutations within cells? *Oncogene* 21(58):8957–8966.

35. Ha T (2013) Single-Molecule Approaches Embrace Molecular Cohorts. *Cell* 154(4):723–726.

36. Yang J, Zhuang Z, Roccasecca RM, Trakselis MA, Benkovic SJ (2004) The dynamic processivity of the T4 DNA polymerase during replication. *Proc Natl Acad Sci USA* 101(22):8289–8294.

37. Hamdan SM, *et al*. (2007) Dynamic DNA helicase-DNA polymerase interactions assure processive replication fork movement. *Mol Cell*

27(4):539–549.

38. Graham JS, Johnson RC, Marko JF (2011) Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic Acids Res* 39(6):2249–2259.

39. Indiani C, Langston LD, Yurieva O, Goodman MF, O'Donnell M (2009) Translesion DNA polymerases remodel the replisome and alter the speed of the replicative helicase. *Proc Natl Acad Sci USA* 106(15):6031–6038.

40. Foti JJ, Devadoss B, Winkler JA, Collins JJ, Walker GC (2012) Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* 336(6079):315–319.

41. Uchida K, *et al*. (2008) Overproduction of *Escherichia coli* DNA polymerase DinB (Pol IV) inhibits replication fork progression and is lethal. *Mol Microbiol* 70(3):608–622.

42. Wagner J, Etienne H, Fuchs RP, Cordonnier A, Burnouf D (2009) Distinct β-clamp interactions govern the activities of the Y family Pol IV DNA polymerase. *Mol Microbiol* 74(5):1143–1151.

43. Beuning PJ, Sawicka D, Barsky D, Walker GC (2005) Two processivity clamp interactions differentially alter the dual activities of UmuC. *Mol Microbiol* 59(2):460–474.

44. Downey CD, McHenry C (2010) Chaperoning of a replicative polymerase onto a newly assembled DNA-bound sliding clamp by the clamp loader. *Mol Cell* 37(4):481–491.

45. Chapados BR, *et al*. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* 116(1):39–50.

46. Tsutakawa SE, *et al*. (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* 145(2):198–211.

47. Beattie TR, Bell SD (2012) Coordination of multiple enzyme activities by a single PCNA in archaeal Okazaki fragment maturation. *EMBO J* 31(6):1556–1567.

48. Sale JE, Lehmann AR, Woodgate R (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nature Rev Mol Cell Biol* 13(3):141–152.

49. Freudenthal BD, Gakhar L, Ramaswamy S, Washington MT (2010) Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA

polymerase exchange. *Nat Struct Mol Biol* 17(4):479–484.

50. Tsutakawa SE, *et al*. (2011) Solution X-ray scattering combined with computational modeling reveals multiple conformations of covalently bound ubiquitin on PCNA. *Proc Natl Acad Sci USA* 108(43):17672–17677.

51. Sharma A, Kottur J, Narayanan N, Nair DT (2013) A strategically located serine residue is critical for the mutator activity of DNA polymerase IV from *Escherichia coli*. *Nucleic Acids Res* 41(9):5104–5114.

Chapter 3

Exchange between *Escherichia coli* Polymerases II and III on a single processivity clamp

3.1 Introduction

Although Pol II was the second *E. coli* DNA polymerase to be discovered, its cellular role remains enigmatic (1). It is encoded by the gene *polB*, which is non-essential (2). Since Pol II is regulated by the SOS response and has some lesion bypass activity, it is considered to be a TLS polymerase (3-6). In contrast to Pols IV and V, however, it is a B-form polymerase with 3'-5' proofreading activity, two classifications that are shared by high fidelity replicative polymerases in other organisms (3, 7, 8). Additionally, lesion bypass by Pol II is generally slow (9, 10), and *polB* mutant cells have minor or negligible survival defects when treated with DNA damaging agents (11, 12). Other activities attributed to Pol II are replication restart following UV irradiation (13), stationary phase competition (14, 15), and proofreading misinsertion errors, especially on the lagging strand (16).

As with the other *E. coli* polymerases, processive DNA synthesis by Pol II requires an interaction with the β sliding clamp via a clamp-binding motif (CBM) at the polymerase's C-terminus (17, 18). β is a head-to-tail dimer with both monomers presenting a protein-binding cleft on the same face, inspiring the proposal that exchange occurs between a replicative polymerase and a translesion polymerase bound to the same clamp, with β serving as a molecular "toolbelt" (19). An alternative model is one where a single polymerase occludes binding of others,

requiring dissociation of one followed by association of another from solution (20).

In Chapter 2, I have demonstrated using a single-molecule reconstitution of primer exchange that Pol III and Pol IV can simultaneously bind β and exchange, in support of the toolbelt model. A previous study used a bulk biochemical reconstitution of exchange between Pol II and stalled Pol III to show that Pol II exchange was less efficient than Pol IV exchange (21). Here I report a detailed analysis of exchange between Pol II and Pol III at the single-molecule level, and a comparison of exchange involving Pol II and Pol IV within the fully reconstituted *E. coli* replisome. These results support the model that Pol II and Pol III can simultaneously bind β, but engage in a different mode of exchange that points to a role for Pol II outside of the replication fork.

3.2 Methods

Single-molecule primer extension reactions using the undamaged M13mp18 single molecule DNA template were performed and analyzed as described in Chapter 2. For primer extensions with both Pol II and Pol III, a cutoff of 45 bp s^{-1} was used in experiments to distinguish Pol II (slower) and Pol III (faster). This cutoff captured 93% of Pol III events and 94% of Pol II events in experiments with each polymerase alone.

Pol III core replisome proteins, Pol IV, and β clamp (wildtype and mutants) were also purified as in Chapter 2. Additional *E. coli* proteins were purified with published protocols and were expressed without affinity tags: Pol II (22), helicase DnaB (23), and primase DnaG (24).

A rolling-circle DNA template was prepared as previously described using T7 DNA polymerase (New England Biolabs) to extend a tailed oligonucleotide primer annealed to M13mp7(L2) single-stranded (ss) DNA, generating a fork structure (25). Substrates were purified with phenol/chloroform extraction. Rolling-circle replication reactions with the *E. coli* replisome were performed were performed by my collaborator, Seungwoo Chang, as previously described (26), with: 60 nM DnaB (hexameric), 180 nM DnaC (monomeric), 30 nM Pol III αεθ core, 15 nM $\tau_3\delta\delta'\chi\psi$, 30 nM β (dimeric), 300 nM DnaG, and 250 nM SSB (tetrameric); 60 μM dNTPs supplemented with α-^{32}P-labeled-dATP, 200 μM UTP, GTP, and CTP, and 1 mM ATP.

The Pol III replisome was loaded onto the fork structure by mixing αεθ, β, clamp loader, and helicase with dCTP, dGTP, ATP, and ~1.5 nM DNA substrate and incubating at 37 °C for ~5 minutes. Synthesis at 37 °C was initiated by adding the dATP and dTTP, SSB, and primase. 10 seconds after initiation, the indicated concentrations of Pol II or Pol IV were added. Reactions were quenched after 10 minutes by adding 25 mM EDTA and separated on a denaturing alkaline agarose gel (0.6%). The dried gel was exposed to a phosphor screen and imaged with a Personal Molecular Imager.

3.3 Results

3.3.1 Single-molecule analysis of Pol II synthesis

Bulk biochemical analysis of polymerase activity often requires the synchronization of a population of molecules. The measurement of a polymerase's processivity, for example, is often accomplished by stable loading of a polymerase

on a template in the presence of accessory proteins such as β, followed the initiation of primer extension of the labeled primer in the presence of an excess of unlabeled trap DNA. This approach can be challenging for translesion polymerases, for which a lower stability can limit pre-loading. Single-molecule measurements of polymerase activity, in contrast, do not share this limitation and allow for the measurement of a range of complex activities during unsynchronized, active synthesis.

Motivated to study the activity of Pol II polymerase exchange and avoid this pre-loading requirement, I chose to use the single-molecule flow-stretching assay that I previously used to study exchange between the Pol III and Pol IV (Chapter 2). In this assay, primed ssDNA templates are coupled to micron-scale beads within a microfluidic flow cell. Laminar flow is used to exert a constant, ~3 pN force on the bead which extends the ssDNA tether. The differential extension of ssDNA and dsDNA at this force results in lengthening of the tethers during primer extension, which is measured for individual molecules by observing bead displacement in dark-field microscopy.

Primer extension by Pol II or Pol III individually in the presence of β occurs in processive synthesis steps interspersed by pauses (Figure 3.1). The processivity (~300 bp) and rate (19.5 bp/s on average) of Pol II (Figure 3.2) and are in agreement with bulk biochemical measurements of primer extension (17). A direct comparison to an equivalent single-molecule analysis of Pol IV (27), and bulk experiments with Pol V (28), further reveals that Pol II is the fastest *E. coli* translesion polymerase, although all three are significantly slower and less processive than the replicative polymerase, Pol III (see Figure 2.7).

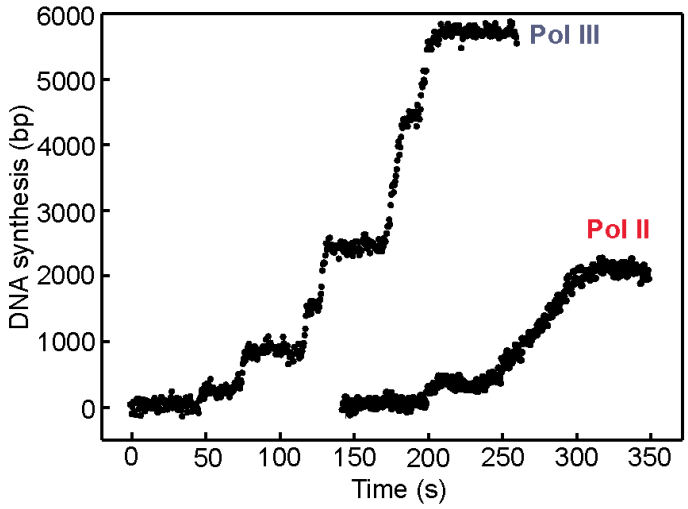

Figure 3.1 Synthesis by Pols II or Pol III on individual molecules occurs in processive steps interspersed by pauses.

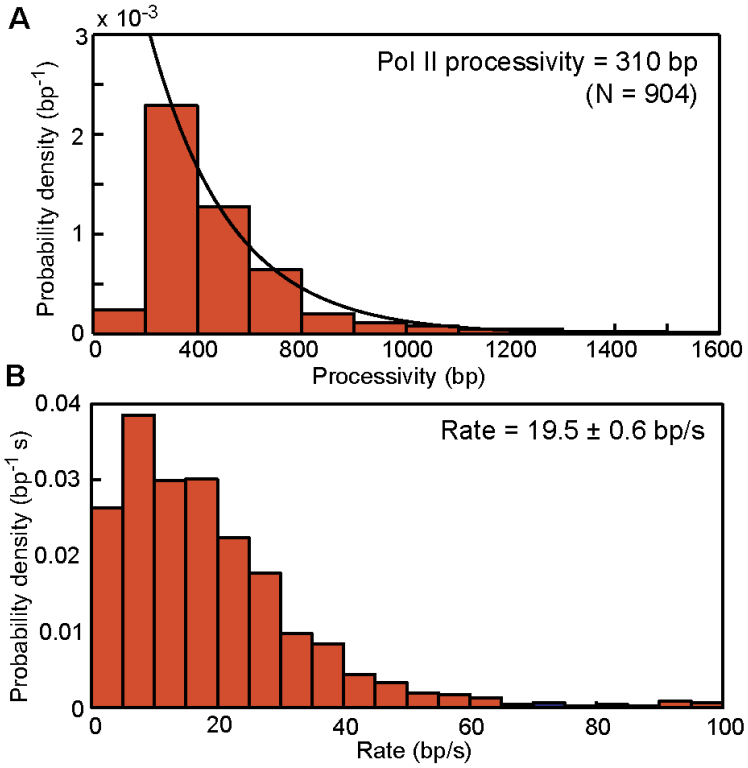

Figure 3.2 Primer synthesis by individual Pol II molecules. (A) Processivity distribution for synthesis steps, fit to a single exponential with a constant of 310 bp (95% confidence bounds: 250 bp, 410 bp). (B) Rate distribution for steps, with mean and SE.

3.3.2 Observing exchange between Pol II and Pol III

As I and others have previously shown, the pause times observed during single-molecule primer synthesis are inversely proportional to the concentration, and are therefore represent polymerase dissociation from β and the diffusion-limited recruitment of a new polymerase from solution (27, 29). The timescale of exchange between two different polymerases should also be diffusion-limited, unless the two polymerases can simultaneously bind β, as in the toolbelt model, or in some other complex at the primer terminus. In this scenario, the exchange timescale would likely be much faster, limited by conformational dynamics on the clamp.

I therefore performed primer experiments in the presence of both Pol II (15 nM) and Pol III (5 nM) and observed that exchange readily occurred between polymerases on individual DNA molecules (Figure 3.3). The ratio of polymerase concentrations were chosen to match the ratio in healthy, replicating cells (30), but reduced by roughly fivefold so the diffusion-limited association of a polymerase from solution could be clearly measured. For each polymerase at the indicated concentrations, the diffusion timescale was determined by measuring the pause lengths and fitting the distribution to an exponential (Figure 3.4A,B). That the pause times for Pol III are shorter despite a lower polymerase concentration likely reflects the greater k_a for clamp binding by Pol III, as previously measured in surface plasmon resonance experiments (31, 32).

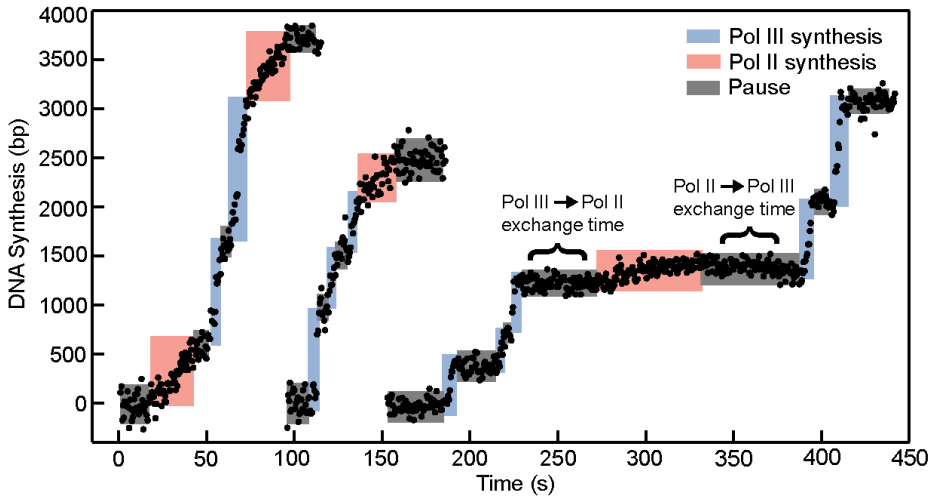

Figure 3.3 Exchange between Pol III (5 nM, blue) and Pol II (15 nM, red) observed during synthesis of individual DNA templates. Exchange times, between events by different polymerases, are highlighted.

In comparison to the diffusion timescale of Pol II (Figure 3.4A), the timescale of exchange from Pol III to Pol II (Figure 3.4C) was significantly faster ($P < 0.001$), with most events occurring within the time resolution of our assay. This result shows that for exchange on most molecules, Pol II is not being recruited from solution. Similarly, the reverse exchange reaction, from Pol II back to Pol III, was significantly faster than the diffusion timescale of Pol III alone ($P < 0.001$, Figure 3.4B vs. 3.4D), also reflecting a diffusion-independent mechanism.

To determine if this rapid exchange is due to each polymerase simultaneously binding a single cleft on the same clamp, I used a mutant clamp that contains a single binding cleft, β^{+}/β^{C}, purified with dual affinity tag chromatography from a mixture of wildtype β and β^{C}, which has a mutated binding cleft (31). For exchange with β^{+}/β^{C} from Pol III to Pol II, and from Pol III to Pol II, I found a significant increase in the timescale (Figure 3.5) compared to experiments with wildtype β ($P < 0.001$), supporting an important role for the homodimeric clamp in

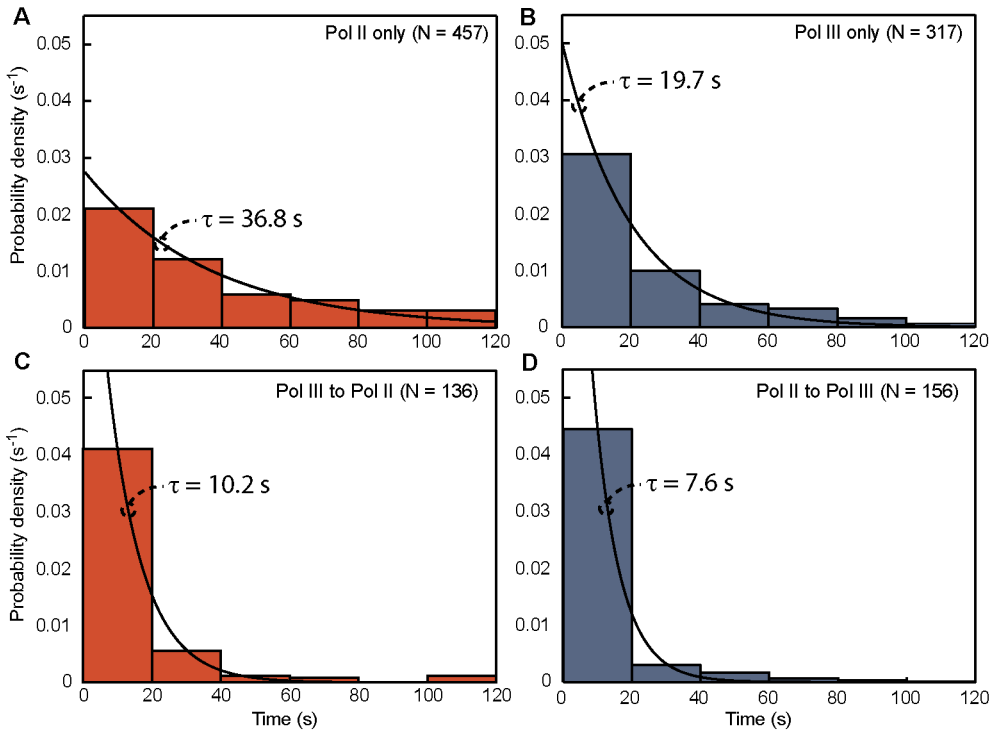

Figure 3.4 Quantification of exchange supports the toolbelt model for Pol II and Pol III. Exchange by (*A*) Pol II (15 nM) or (*B*) Pol III alone (5 nM) represents dissociation of a polymerase followed by the diffusion-limited association of a new polymerase. Exchange timescales from (*C*) Pol III to Pol II or (*D*) Pol II to Pol III at matched concentrations are more rapid ($P < 0.001$), indicating β–mediated exchange.

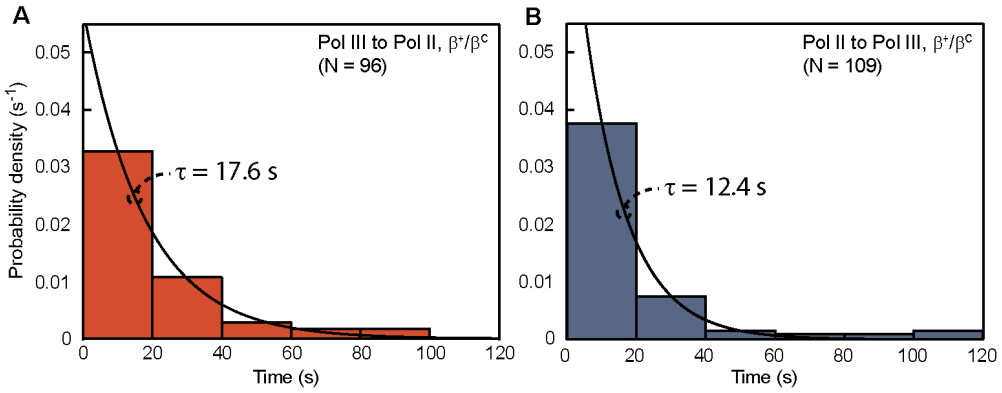

Figure 3.5 In experiments with the single cleft clamp β^{+}/β^{C}, exchange from (A) Pol III to Pol II and (B) Pol II to Pol III are intermediate between a diffusion-limited timescale (Fig. 3A,B) and the rapid, β-mediated exchange (Fig. 3C,D) ($P < 0.001$ for all comparisons).

coordinating exchange. Interestingly, exchange times with β^{+}/β^{C} remained faster than the diffusion-limited timescales. A possible explanation is that Pol II makes additional contacts with β or even Pol III, similar to Pol IV (33, 34). Nevertheless, the data from the single-cleft clamp β^{+}/β^{C} further support the toolbelt model for Pol III and Pol II simultaneously binding the clamp.

3.3.3 Inhibition of Pol III alone but not the full replisome by Pol II

Exchange between two polymerases bound to the same clamp, or within a multi-protein complex, can either occur after the first polymerase fully dissociates from a critical binding surface, terminating its synthesis, or through a more complex mechanism involving the exchange of a shared contact or contacts within a multivalent binding interface. In the latter case, the second polymerase gains a "foothold" by binding a secondary binding surface and can then capture a critical interaction site during a transient, partial dissociation of the first polymerase (35). In this scenario, the presence of the second polymerase can lead to the premature dissociation of the first, reducing its processivity. This "dynamic processivity" has been observed for polymerases bound to the same helicase within the T4 and T7 replisomes (36-38), and for Pol IV capturing a binding cleft from Pol III following its association at a unique binding surface on the β "rim" (27).

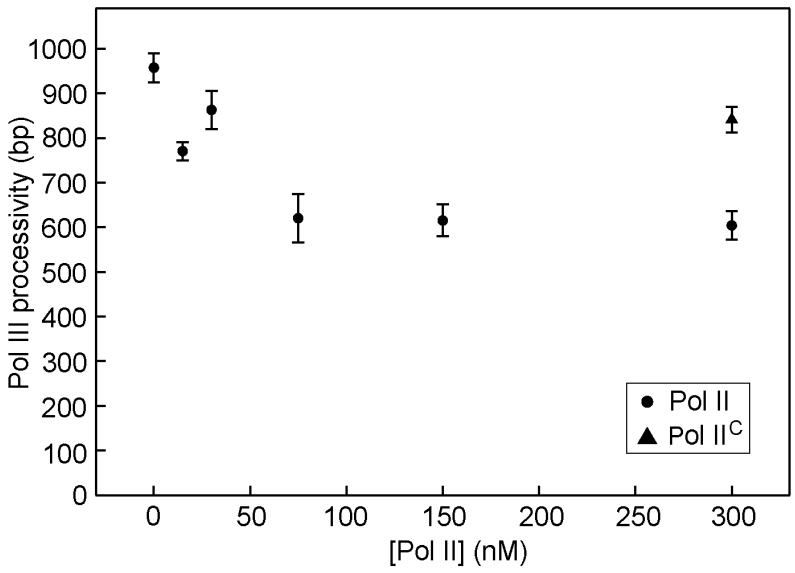

Figure 3.6 Increased concentrations of Pol II significantly reduce the Pol III processivity ($P < 0.01$ for 75-300 nM Pol II), indicative of dynamic processivity between the polymerases. Removing the Pol II cleft-binding motif rescues the effect (*NS* vs. Pol III alone). Values represent means with SEs.

To determine if dynamic processivity exists between Pol III and Pol II, I performed single-molecule experiments with increasing concentrations of Pol II, simulating SOS induction. The ability to assign individual synthesis events to each polymerase allows us to unambiguously determine the effect on Pol III, even as the relative contribution by each polymerase changes. As was previously shown for Pol IV, the presence of Pol II leads to the reduction of the Pol III processivity in a dose-dependent manner (Figure 3.6, $P < 0.01$). This reduction further depends on the capture of a β binding cleft from Pol III, as Pol II^{C}, a mutant lacking the clamp-binding residues, does not significantly affect the Pol III processivity distribution.

In the absence of detailed structural information of the Pol II-β interaction, it is difficult to determine if Pol III displacement involves Pol II binding first to a secondary site of β, independent of the cleft, or by a different mechanism; previous

data have shown that mutating the secondary Pol IV binding site on the 'rim' of β has no effect on exchange of Pol II with Pol III (21). We instead decided to test the ability of Pol II to displace Pol III within the context of the fully reconstituted replisome to determine if a potential Pol II binding site remains accessible. This assay involves pre-loading Pol III, β, the clamp loader complex $\tau_3\delta\delta'\chi\psi$, and the DnaB helicase on a rolling-circle M13 template. Synthesis is then initiated by adding nucleotides, including α-^{32}P-labeled-dATP, primase, and SSB; in the absence of TLS polymerases, Pol III rapidly makes several revolutions around the circular template, generating long leading-strand products that are visualized with alkaline agarose gel electrophoresis.

Adding Pol IV to the reaction after Pol III initiation inhibits synthesis of long products in a dose-dependent manner; in contrast, the Pol III replisome is largely resistant to inhibition by Pol II (Figure 3.7). A previous study using a minicircle template also showed that Pol IV is more efficient than Pol II at displacing Pol III (39), although the lack of inhibition by Pol II in this assay is more striking. These data are also consistent with the result that a roughly equivalent level of overexpression of Pol IV, but not Pol II, impedes growth in the strain lacking the Rep helicase, which makes cells more sensitive to Pol III stalling (21).

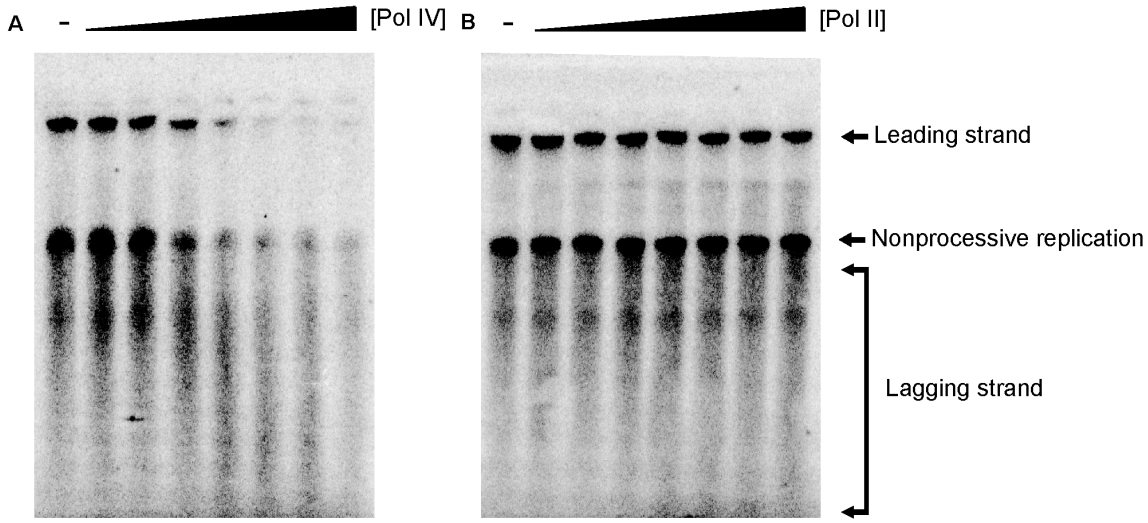

Figure 3.7 The fully reconstituted replisome is resistant to access by Pol II, but not Pol IV. Rolling-circle replication by the Pol III holoenzyme is initiated, the indicated TLS polymerase is added 10 seconds later, and reactions are quenched 10 minutes later. Concentrations of polymerase added (left to right) for (A) Pol IV are: 0, 39, 78, 156, 312, 625, 1250, and 2500 nM; and for Pol II: 0, 23, 47, 94, 188, 375, 750, and 1500 nM.

3.4 Discussion

At least 10 proteins contain clamp-binding motifs that interact with β (40). As the β dimer has only two binding clefts, this suggests that β occupancy is context dependent. In this chapter, I have shown that Pol II and Pol III can simultaneously bind the clamp and rapidly exchange. Since Pol III can accommodate either Pol II or Pol IV, this suggests that the Pol III-β interaction may be flexible enough to accommodate binding of a range of clamp-binding proteins without disrupting Pol III primer extension, giving other polymerases and DNA-modifying enzymes a means to rapidly take over from Pol III if and when it stalls.

Increasing concentrations of Pol II lead to the displacement of Pol III from primer extension reactions, suggesting that Pol II either binds to additional sites on the β, or perhaps to Pol III directly, similar to Pol IV (33, 34). Binding to such a site

can give Pol II a "foothold," which it can use to strip Pol III off the clamp through direct competition for common binding sites during transient dissociations. Further studies using surface plasmon resonance and mutagenesis are required to identify potential binding interacts.

In contrast, replication by Pol III within the full replisome effectively blocks up to a 50-fold excess of Pol II. This suggests that interactions within the full replisome either block secondary Pol II binding sites on β, or stabilize Pol III against the dynamic exchange observed in primer extension. In a striking contrast, Pol IV remains able to displace Pol III despite a ~15-fold lower affinity for β (32). Since displacement by Pol IV depends on its interaction with the rim site (Seungwoo Chang, unpublished data), this argues that this binding mode is unique and could give Pol IV priority over other TLS polymerases for access to the replisome following SOS induction.

Although Pol II may be excluded from the full replisome, there are other cellular contexts when it may exchange with Pol III. The Pol III primer extension reactions described here resemble Okazaki fragments that are prematurely released from the replisome. Premature loop release can either occur when the lagging strand Pol III encounters a lesion (41), or through stochastic loop release in the absence of a roadblock (42, 43). That Pol II can access released Okazaki fragments but not the replication fork is consistent with the observation that Pol II preferentially influences the fidelity on the lagging strand (16). In addition to structural studies, further clarification of the role of Pol II requires increasingly

complex reconstitutions of polymerase exchange, and single-molecule imaging of Pol II dynamics.

3.5 Contributions

As in Chapter 2, Pol III replisome components were purified by Slobodan Jergic in the laboratory of Nicholas Dixon (University of Wollongong). The heterodimeric mutant clamp and Pol II variants were provided by Mark Sutton (University of Buffalo). Seungwoo Chang (Harvard Medical School) performed rolling-circle replication reactions displayed in Figure 3.7.

References

1. Fuchs RP, Fujii S (2013) Translesion DNA synthesis and mutagenesis in prokaryotes. *Cold Spring Harb Perspect Biol* 5(12):a012682.

2. Campbell JL, Soll L, Richardson CC (1972) Isolation and partial characterization of a mutant of *Escherichia coli* deficient in DNA polymerase II. *Proc Natl Acad Sci USA* 69(8):2090–2094.

3. Bonner CA, Hays S, McEntee K, Goodman MF (1990) DNA polymerase II is encoded by the DNA damage-inducible *dinA* gene of *Escherichia coli*. *Proc Natl Acad Sci USA* 87(19):7663–7667.

4. Qiu Z, Goodman MF (1997) 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(13):8611–8617.

5. Napolitano R, Janel-Bintz R, Wagner J, Fuchs RP (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J* 19(22):6259–6265.

6. Becherel OJ, Fuchs RP (2001) Mechanism of DNA polymerase II-mediated frameshift mutagenesis. *Proc Natl Acad Sci USA* 98(15):8566–8571.

7. Iwasaki H, Ishino Y, Toh H, Nakata A, Shinagawa H (1991) *Escherichia coli* DNA polymerase II is homologous to alpha-like DNA polymerases. *Mol Gen Genet* 226(1-2):24–33.

8. Wang F, Yang W (2009) Structural Insight into Translesion Synthesis by DNA Pol II. *Cell* 139(7):1279–1289.

9. Paz-Elizur T, Takeshita M, Goodman M, O'Donnell M, Livneh Z (1996) Mechanism of translesion DNA synthesis by DNA polymerase II. Comparison to DNA polymerases I and III core. *J Biol Chem* 271(40):24662–24669.

10. Al Mamun AA, Humayun MZ (2006) *Escherichia coli DNA* polymerase II can efficiently bypass 3,N^4-ethenocytosine lesions in vitro and in vivo. *Mutat Res* 593(1-2):164–176.

11. Escarceller M, *et al*. (1994) Involvement of *Escherichia coli* DNA polymerase II in response to oxidative damage and adaptive mutation. *J Bacteriol* 176(20):6221–6228.

12. Berardini M, Foster PL, Loechler EL (1999) DNA polymerase II (*polB*) is involved in a new DNA repair pathway for DNA interstrand cross-links in *Escherichia coli*. *J Bacteriol* 181(9):2878–2882.

13. Rangarajan S, Woodgate R, Goodman MF (1999) A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated *Escherichia coli*. *Proc Natl Acad Sci USA* 96(16):9224–9229.

14. Yeiser B, Pepper E, Goodman M, Finkel S (2002) SOS-induced DNA polymerases enhance long-term survival and evolutionary fitness. *Proc Natl Acad Sci USA* 99(13):8737.

15. Corzett CH, Goodman MF, Finkel SE (2013) Competitive fitness during feast and famine: how SOS DNA polymerases influence physiology and evolution in *Escherichia coli*. *Genetics* 194(2):409-420.

16. Banach-Orlowska M, Fijalkowska IJ, Schaaper RM, Jonczyk P (2005) DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in *Escherichia coli*. *Mol Microbiol* 58(1):61–70.

17. Bonner CA, *et al*. (1992) Processive DNA synthesis by DNA polymerase II mediated by DNA polymerase III accessory proteins. *J Biol Chem* 267(16):11431–11438.

18. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc Natl Acad Sci USA* 98(20):11627–11632.

19. Pagès V, Fuchs RP (2002) How DNA lesions are turned into mutations within cells? *Oncogene* 21(58):8957–8966.

20. Toste Rêgo A, Holding AN, Kent H, Lamers MH (2013) Architecture of the Pol III-clamp-exonuclease complex reveals key roles of the exonuclease subunit in processive DNA synthesis and repair. *EMBO J* 32(9):1334–1343.

21. Heltzel JMH, Maul RW, Wolff DW, Sutton MD (2012) *Escherichia coli* DNA Polymerase IV, but not Pol II, dynamically switches with a stalled Pol III* replicase. *J Bacteriol* 194(14):3589-3600.

22. Maul RW, Scouten Ponticelli SK, Duzen JM, Sutton MD (2007) Differential binding of *Escherichia coli* DNA polymerases to the β-sliding clamp. *Mol Microbiol* 65(3):811–827.

23. San Martin MC, Stamford NP, Dammerova N, Dixon NE, Carazo JM (1995) A structural model for the *Escherichia coli* DnaB helicase based on electron microscopy data. *J Struct Biol* 114(3):167–176.

24. Stamford NP, Lilley PE, Dixon NE (1992) Enriched sources of *Escherichia coli* replication proteins. The dnaG primase is a zinc metalloprotein. *Biochim Biophys Acta* 1132(1):17–25.

25. Tabor S, Huber HE, Richardson CC (1987) *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J Biol Chem* 262(33):16212–16223.

26. Tanner NA, et al. (2011) *E. coli* DNA replication in the absence of free β clamps. *EMBO J* 30(9):1830–1840.

27. Kath JE, *et al.* (2014) Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. *Proc Natl Acad Sci USA* 111(21):7647-7652.

28. Karata K, Vaisman A, Goodman MF, Woodgate R (2012) Simple and efficient purification of *Escherichia coli* DNA polymerase V: cofactor requirements for optimal activity and processivity *in vitro*. *DNA Repair* 11(4):431–440.

29. Tanner NA, *et al*. (2008) Single-molecule studies of fork dynamics in Escherichia coli DNA replication. *Nat Struct Mol Biol* 15(2):170–176.

30. Sutton MD (2010) Coordinating DNA polymerase traffic during high and low fidelity synthesis. *Biochim Biophys Acta* 1804(5):1167–1179.

31. Scouten Ponticelli SK, Duzen JM, Sutton MD (2009) Contributions of the individual hydrophobic clefts of the *Escherichia coli* β sliding clamp to clamp loading, DNA replication and clamp recycling. *Nucleic Acids Res* 37(9):2796–2809.

32. Heltzel JMH, *et al*. (2009) Sliding clamp–DNA interactions are required for viability and contribute to DNA polymerase management in *Escherichia coli*. *J Mol Biol* 387(1):74–91.

33. Bunting K, Roe S, Pearl L (2003) Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the β-clamp. *EMBO J* 22(21):5883–5892.

34. Scotland MK, *et al*. (2015) 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(9):e1005507.

35. Ha T (2013) Single-molecule approaches embrace molecular cohorts. *Cell* 154(4):723–726.

36. Yang J, Zhuang Z, Roccasecca RM, Trakselis MA, Benkovic SJ (2004) The dynamic processivity of the T4 DNA polymerase during replication. *Proc Natl Acad Sci USA* 101(22):8289–8294.

37. Hamdan SM, *et al*. (2007) Dynamic DNA helicase-DNA polymerase interactions assure processive replication fork movement. *Mol Cell*

27(4):539–549.

38. Loparo JJ, Kulczyk AW, Richardson CC, van Oijen AM (2011) Simultaneous single-molecule measurements of phage T7 replisome composition and function reveal the mechanism of polymerase exchange. *Proc Natl Acad Sci USA* 108(9):3584-3589.

39. Indiani C, Langston LD, Yurieva O, Goodman MF, O'Donnell M (2009) Translesion DNA polymerases remodel the replisome and alter the speed of the replicative helicase. *Proc Natl Acad Sci USA* 106(15):6031–6038.

40. Wijffels G, *et al.* (2004) Inhibition of protein interactions with the β sliding clamp of *Escherichia coli* DNA polymerase III by peptides from β-binding proteins. *Biochemistry* 43(19):5661–5671.

41. Pagès V, Fuchs RP (2003) Uncoupling of leading- and lagging-strand DNA replication during lesion bypass *in vivo*. *Science* 300(5623):1300–1303.

42. Hamdan SM, Loparo JJ, Takahashi M, Richardson CC, van Oijen AM (2010) Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* 457(7226):336–339.

43. Kurth I, Georgescu RE, O'Donnell ME (2013) A solution to release twisted DNA during chromosome replication by coupled DNA polymerases. *Nature* 496(7443):119-122.

Chapter 4

Observing individual Polymerase IV molecules in living cells

4.1 Introduction

The ability to observe the localization and dynamics of proteins in living cells makes fluorescence microscopy a powerful method for studying DNA replication. A pioneering study by David Sherratt's lab at Oxford University used strains altered to express different replisome genes as fusions to the bright fluorescent protein Ypet, and imaged the position of the fusions throughout the cell cycle (1). Contrasting a previous proposal that both replication forks remain localized within a single "replication factory" throughout the cell cycle, these researchers observed that the two forks diverged in opposite directions from the midcell plane following initiation, moving toward the quarter cell position before replication termination and cell division.

A further improvement to live cell fluorescence imaging was the use of powerful lasers and sensitive EMCCD cameras to observe single fluorescent proteins, enabling quantitative studies of the composition of the replisome (2), and its dynamics (3). These single-molecule techniques have been further combined with fluorescent proteins that can be photoconverted by ultraviolet light from one form (typically a dark state, or a green fluorescent state) to another (often a red fluorescent state). In a technique called photoactivated localization microscopy (PALM), a low laser intensity is used to stochastically convert single fluorescent

proteins within the cell; the selective observation of activated fusion proteins can be used to determine localization patterns, diffusive behavior, and binding kinetics (4). This approach was recently used to observe the localization of Polymerase (Pol) I to sites of alkylation damage during base excision repair (5).

As I have previously focused on the ability of the TLS polymerase Pol IV to exchange with Pol III within a biochemical reconstitution (Chapter 2), a natural direction was the observation of Pol IV's localization and dynamics in living cells. Given the important role of the β clamp in facilitating polymerase exchange, and the requirement of the two sets of Pol IV clamp-binding residues for successful TLS (6), we predicted that Pol IV localizes with the replisome. It is unclear, however, if localization occurs during normal growth, or requires the presence of DNA damage and induction during the SOS response. Other factors that may influence Pol IV localization is its interaction with the RNA polymerase subunit NusA (7), which is involved in transcription-coupled repair, and an interaction with RecA (8) which may recruit Pol V to cell membranes during the early stages of the SOS response (9).

A recent study reported a Pol IV-EYFP fusion formed foci after treatment with nalidixic acid (10). Although this plasmid-expressed fusion was only partially functional, and cellular localization was not investigated, this nevertheless suggested that quantitative imaging of Pol IV in living cells was feasible. Here I report a functional, genomically expressed fusion of Pol IV to the photoactivable fluorescent protein PAmCherry, and initial efforts to observe its localization, dynamics, and replisome association in the presence of cognate lesions.

4.2 Methods

4.2.1 Strain construction

Fusions of fluorescent proteins to Pol IV and other replisome proteins were made using λ Red recombineering. In this genome editing system, the λ phage's recombinogenic Red operon is expressed prior to electroporating cells with linear DNA with ~50 bp homology arms; after recovery, cells containing the desired mutations are selected and verified. To express the Red operon, I used the Cm^R pSIM5 (11) plasmid transformed into the E. coli K-12 type strain, MG1655, following published protocols. After recombineering, the temperature-sensitive recombineering plasmids were cured by growth at 37 °C.

Linear DNA fragments used in recombineering were amplified by polymerase chain reaction (PCR) with Q5 Hot Start polymerase (New England Biolabs), or "homemade" Pfu polymerase. The plasmid pKD4 (12), modified by Gibson assembly (13) and/or site-directed mutagenesis, was used as a template. This plasmid contains the oriRγ replicon and thus requires the *pir* gene product for replication; linear PCR products amplified from pKD4 can therefore be PCR purified and electroporated into cells without obtaining undesired plasmid transformants. pKD4 contains a kanymycin resistance (Kan^R) cassette flanked by FRT sites. This cassette, included downstream of a fusion, is used to select for a genome modification, and can be removed using pCP20, a temperature-sensitive helper plasmid that expresses Flp recombinase (12). Removal of the cassette leaves behind a ~80 bp scar consisting of a single FRT sequence.

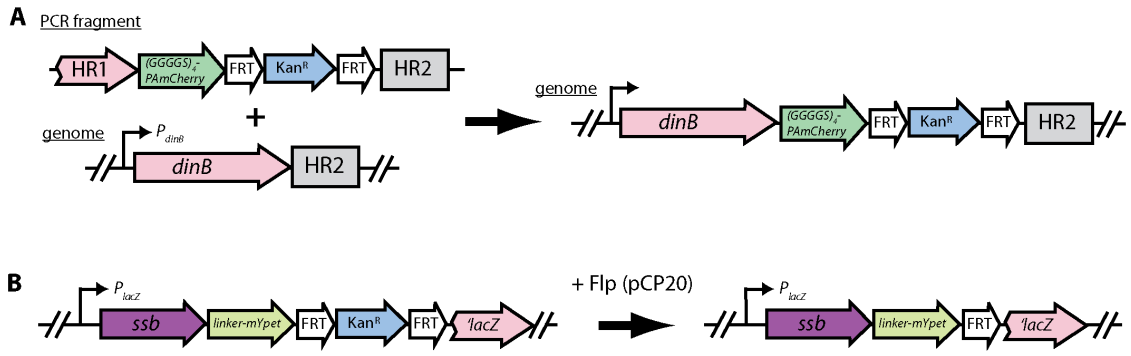

Figure 4.1 (A) Linear PCR fragments containing PAmCherry and a FRT-flanked Kan^R cassette were transformed into cells expressing the λ Red operon. Sequences homologous to the 3' end of the Pol IV gene *dinB* and the region immediately downstream target the fragment for recombination with the genome, creating an in-frame fusion. **(B)** Expression of the Flp recombinase from the plasmid pCP20 removes the Kan^R cassette from the *lacZ::ssb-mYpet* locus, leaving a scar.

To make a C-terminal fusion to the Pol IV gene *dinB*, the photoactivatible mCherry gene *pamcherry1* (14) was amplified from pBAD/HisB-PAmCherry1 (Addgene #31931) with oligonucleotides containing a Gly-Ser linker, pKD4_G4S_PAmCherry_for and pKD4_G4S_PAmCherry_rev (see Table 4.1 for a list of primers). Gibson assembly was used to stitch it together with a linear fragment of the pKD4 backbone, amplified by inverse PCR with pKD4_9_30_G4S_rev and pKD4_31_48_PAmCherry_for. The resulting plasmid, pKD4-G4S-PAmCherry, was used to amplify a linear fragment encoding the a 20 amino acid $(GGGGS)_4$ linker, PAmCherry, and the FRT-Kan^R-FRT cassette with oligonucleotides dinB-G4S-FP_KI_for and dinB-FP_KI_rev, which each contain ~50 nucleotide homology to the 3' region of the *dinB* gene. Recombineering with this fragment made an in-frame insertion of the linker and PAmCherry after the final amino acid of Pol IV, with the antibiotic resistance cassette following the fusion (Figure 4.1A).

A variant of this fusion, Pol IV-PAmCherry-FLAG, was made by introducing the residues GSDYKDDDDK to the C-terminus of PAmCherry within the plasmid

Names	**Sequences** (5' to 3')
pKD4_G4S_PAmCherry_for	GGTGGTGGTGGTTCTGGCGGCGGTGGCAGTGGTGGCGGTGGCAGTAT GGTTAGCAAGGGCGAGGAG
pKD4_G4S_PAmCherry_rev	ACTTCGAAGCAGCTCCAGCCTACACTTACTTGTACAGCTCGTCCATGC
pKD4_9_30_G4S_rev	GCCACCGCCGCCAGAACCACCACCACCAGAACCACCACCACCAATCGCT CAAGACGTGTAATGC
pKD4_31_48_PAmCherry_for	TACAAATAAGTGTAGGCTGGAGCTGCT
dinB-G4S-FP_KI_for	GTGACGTTGCTTGACCCGCAAATGGAAAGACAACTGGTGCTGGGATT AGGTGGTGGTGGTTCTGGT
dinB-FP_KI_rev	CAGTGATACCCTCATAATAATGCACACCAGAATATACATAATAGTAT ACATCTTATGAATATCCTCCTTAGTTCC
pKD4-PAm-FLAG-for	GGCAGCGACTACAAAGACGATGACGACAAGTAAGTGTAGGCTGGAGCT
pKD4-PAm-FLAG-rev	TGTAGTCGCTGCCCTTGTACAGCTCGTCCATG
pUC19-linker-for	TCGGCTGGCTCCGCTGCTGGTTCTGGCGAATTCATGAGTAAAGGAGAAGAA
pUC19-linker-rev	TCGGCTGGCTCCGCTGCTGGTTCTGGCGAATTCATGAGTAAAGGAGAAGAA
mYpet-ITA-for	ATTACACGTCTTGAGCGATTGAGCTCGGCTGGCTCCGCTGCTG
mYpet-ITA-rev	GAAGCAGCTCCAGCCTACACTTATTTGTAGAGTTCATC
pKD4-ITA-for	TACAAATAAGTGTAGGCTGGAGCTGCT
pKD4-ITA-rev	CAGCAGCGGAGCCAGCCGAGCTCAATCGCTCAAGACGTGTAAT
SSB-mYpet_for	TGCAGCATTACACGTCTTGAGCGATTGATGGCCAGCAGAGGCGTA
SSB-mYpet-rev	GCCAGAACCAGCAGCGGAGCCAGCCGAGCTGAACGGAATGTCATCATC
lacZ-SSB-mYpet-KI_for	TGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGG CCAGCAGAGGCGTAAA
lacZ-SSB-mYpet-KI_for	TCATCATATTTAATCAGCGACTGATCCACCCAGTCCCAGACGAAGATG AATATCCTCCTTAGTTCCTA

Table 4.1. Oligonucleotides (Integrated DNA Technologies) used for in strain construction.

pKD4-G4S-PAmCherry using the site-directed mutagenesis oligos pKD4-PAm-FLAG-for and pKD4-PAm-FLAG-rev. The C-terminal FLAG-tagged fusion was generated with recombineering as described above.

Cell viability requires SSB tetramers within the replisome contain at least two SSB monomers with free C-termini to interact with replication proteins (15); previous studies have therefore used a SSB-mYpet fusion expressed from a second copy at about one-fourth of the level of endogenous SSB (1, 2). To reproduce this fusion, I introduced a sequence encoding the linker SAGSAAGSGEF to the plasmid pUC19-mYpet (16) using the site-directed mutagenesis oligos pUC19-linker-for and pUC19-linker-rev. A fragment containing the linker-mYpet sequence was amplified with the oligos mYpet-ITA-for and mYpet-ITA-rev, and combined with a linear fragment of the pKD4 backbones amplified by inverse PCR with pKD4-ITA-for and pKD4-ITA-rev in a Gibson assembly reaction.

The resulting plasmid, pKD4-linker-mYpet, was then digested with SacI (New England Biolabs), which cuts at the unique restriction site at the 5' terminus of the linker sequence. The gel-purified linear DNA was combined with a fragment containing the SSB gene, amplified from an expression plasmid (17) with the primers SSB-mYpet_for and SSB-mYpet-rev. A recombineering fragment containing both the SSB-mYpet and FRT-Kan^R-FRT sequences was amplified with lacZ-SSB-mYpet-KI_for and lacZ-SSB-mYpet-KI_for, which contain homology to knock it into the *lacZ* gene, replacing the first ~1700 nucleotides. Expression of the resulting fusion within the *lac* operon can be induced by adding IPTG to growth media.

Strain designation	Relevant genotype	Origin or reference
MG1655 pSIM5	*E. coli* K-12 type strain with Cm^R recombineering plasmid	(11)
AB1157 polA-PAmCherry	*polA-pamcherry thr-1, araC14, leuB6(Am), Δ(gpt-proA)62, lacY1, tsx-33, qsr'-0, glnV44(AS), galK2(Oc), LAM-, Rac-0, hisG4(Oc), rfbC1, mgl-51, rpoS396(Am), rpsL31(strR), kdgK51, xylA5, mtl-1, argE3(Oc), thi-1*	(5)
RW118	*lexA+ rpsL31 xyl-5 mtl-1 galK2 lacY1 tsx-33 supE44 thi-1 hisG4[Oc] argE3[Oc] araD139 thr-1* Δ[*gpt-proA*]*62 sulA211*	(31)
RW542	*lexA51(Def) rpsL31 xyl-5 mtl-1 galK2 lacY1 tsx-33 supE44 thi-1 hisG4[Oc] argE3[Oc] araD139 thr-1 Δ[gpt-proA]62 sulA211*	(32)
JEK625	MG1655 *dinB-pamcherry*-FRT-Kan^R-FRT	λ Red: PAmCherry-FRT-Kan^R-FRT → MG1655 pSIM5 *dinB* locus
JEK772	MG1655 *dinB-pamcherry-FLAG*-FRT-Kan^R-FRT	λ Red: PAmCherry-FLAG-FRT-Kan^R-FRT → MG1655 pSIM5
JEK774	RW118 *dinB-pamcherry-FLAG*-FRT-Kan^R-FRT	P1*vir*: JEK772 → RW118
JEK776	RW542 *dinB-pamcherry-FLAG*-FRT-Kan^R-FRT	P1*vir*: JEK772 → RW542
JEK398	MG1655 *lacZ::ssb-mypet*-FRT-Kan^R-FRT	λ Red: *ssb-mypet*-FRT-Kan^R-FRT → MG1655 pSIM5 *lacZ* locus
JEK762	RW542 *lacZ::ssb-mypet*-FRT	(1) P1*vir*: JEK398 → RW542 (2) pCP20 Flp recombination
JEK766	RW542 *lacZ::ssb-mypet*-FRT *dinB-pamcherry*-FRT-Kan^R-FRT	P1*vir*: JEK625 → JEK762

Table 4.2. Strains used in this study.

All fusions were transduced into a clean MG1655 isolate or another genetic background using P1*vir* phage and were verified by PCR amplification of genomic DNA and sequencing (see Table 4.2 for strain list). Strains containing both the SSB and Pol IV fusions were made by using the Flp recombinase to remove the Kan^R cassette from the *lacZ::ssb-mypet*_locus (Figure 4.1B) and introducing the Pol IV fusions by transduction.

4.2.2 Growth conditions and strain validation

Nitrofurazone (NFZ) sensitivity was used to test the ability of genomic fusions of the Pol IV gene *dinB* to perform translesion synthesis. Overnight cultures in LB were diluted 1:1000 into 50 mL LB and grown to OD_{600} 1.0 at 37 °C. Samples of different strains were taken, serially diluted in 0.9% NaCl, and stamped to LB agar plates containing NFZ or solvent. Images of the plates were taken after 16 hr. growth at 37 °C.

Imaging experiments require the use of minimal media to reduce background and cellular fluorescence. Single colonies were inoculated into 3 mL rich LB media for ~8 hours growth at 37 °C, and then diluted 1:1000 into 3 mL freshly prepared M9 media supplemented with 0.4% glucose, 1 mM thiamine hydrochloride, 0.2% casamino acids, 2 mM Mg_2SO_4, and 0.1 mM $CaCl_2$ for overnight growth. Minimal media was supplemented with 0.5 mM IPTG for strains containing the *lacZ::ssb-mypet* allele to induce expression of the fusion. After ~16 hours, saturated cultures were diluted 1:1000 (for MG1655-based strains) or 1:200 (for strains containing the

lexA51^{+} mutation) into beveled flasks containing 50 mL freshly prepared, identically supplemented M9 media, for growth at 37 °C with aeration.

For imaging, 1 mL of a culture was isolated at early log phase (OD600 = 0.15 ± 0.05) and concentrated ~200-fold by centrifugation for 2 min. at 10,000xg. 10 minutes before sample collection, an agar pad was cast by depositing ~500 μL molten 3% GTG agarose (NuSieve) in M9 minimal media between two, cleaned 25x25 mm cover glass slides (VWR). IPTG was found to contribute significantly to background fluorescence and was therefore excluded from the agar pad. 1-2 μL of concentrated cells were deposited onto a small square of agarose and sandwiched between a 25x25 mm and a 24x60 mm No. 1.5 cover slide, with the cells on the side of the larger slide. To reduce background fluorescence, 24x60 mm slides were sonicated in ethanol and 1 M KOH for 30 minutes, repeating twice for each and washing with deionized water between steps. Clean glass was stored in water. Cell samples were imaged up to 45 minutes after sample collection.

For the DNA damage condition, nitrofurazone (NFZ, 100 mM in *N,N*-dimethylformamide [DMF], freshly prepared) was diluted into cultures at OD_{600} ~ 0.15 to a final concentration of 100 μM. After one hour of further growth, cells were isolated and imaged. To determine the fraction of cells that survived treatment with NFZ, cultures (six per condition) were serially diluted in 0.9% NaCl and spread onto LB plates both before and after treatment with NFZ or DMF solvent control. Colony-forming units (CFUs/mL) were counter after 16 hours growth at 37 °C.

For Western blotting, cells containing FLAG-tagged PAmCherry fusions were grown in supplemented M9 to OD_{600} ~ 0.15, and a cell pellet from 5 mL culture was

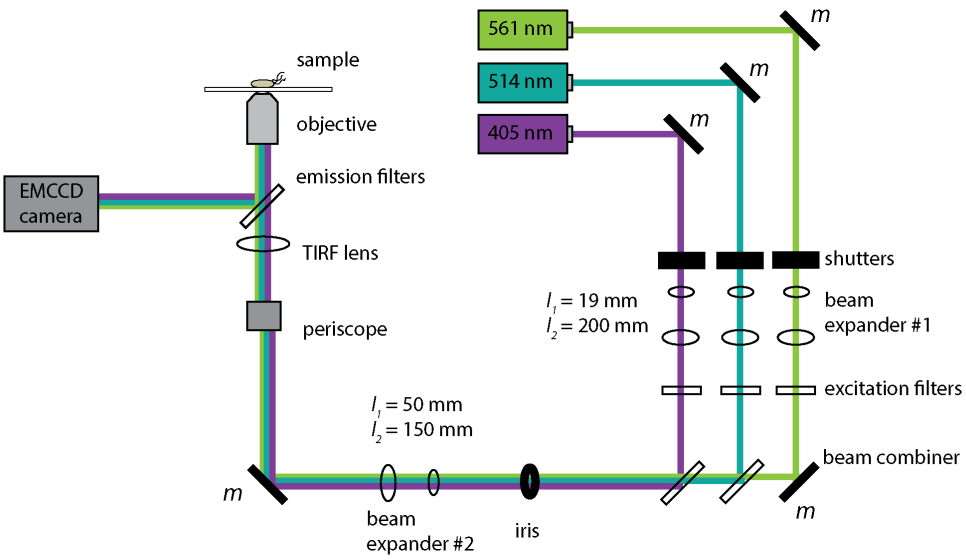

Figure 4.2 Schematic of the custom fluorescence microscope. Mirrors are denoted as *m*, lenses as *l*.

isolated by centrifugation. The pellet was lysed by freeze-thawing, followed by incubation for 10 minutes on ice in (100 x OD_{600}) μL lysis buffer, prepared as previously described (16). Cell lysates were separated by SDS-PAGE chromatography, and probed with 1:2000 goat anti-DDDDK (abcam #1257) and 1:2000 donkey anti-goat IgG-HRP (Santa Cruz Biotechnology).

4.2.3 Fluorescence microscopy

In order to image PAmCherry fusions and co-localize them with a second, mYpet-tagged protein, we built a custom two-color PALM fluorescence microscope (Figure 4.2). This microscope uses three laser lines, independently controlled with computer-controlled Uniblitz shutters: 405 nm to activate PAmCherry (Coherent OBIS 100 mW), 561 nm (Coherent Sapphire 200 mW) to image it, and a 514 nm (Coherent Sapphire 150 mW) to image mYpet.

The three laser lines were expanded with a telescope, cleaned up with excitation filters (ZET561/10X, ZET514/10X, and ZET405/20X, Chroma), and combined (561: mirror, 514: ZT514rdc dichroic, 405: ZT405rdc dichroic, Chroma). A second telescope further expanded the beam, which was focused to the sample with a Nikon 100X 1.49 NA TIRF objective, using a 400 mm lens to adjust the laser angle. Fluorescence signal was passed with a filter cube (dichroic: ZT405/514/561rpc, emission filter: ZET442/514/561m, long pass filter: ET5251p, Chroma) and directed to a Hamamatsu ImageEM C9100-13 camera.

Neutral density filters were used to independently adjust the photon fluence of each laser line to optimize signal-to-noise and bleaching, giving uniform imaging conditions for comparable strains at each exposure. A TIRF lens was used to direct excitation into a near-TIRF excitation, to maximize signal to noise. The TIRF angle, set with a micrometer, was found to be stable for long periods of time.

For PALM experiments, a pre-bleaching step of 561 nm excitation was used to bleach background fluorescence before simultaneous activation and imaging with 405 nm and 561 nm. For experiments at longer exposures (250 milliseconds, ms), this pre-bleaching would take prohibitively long at the lower 561 nm excitation, so an automated filter wheel (ThorLabs) was used to change the 561 nm excitation power by a factor of ten between the pre-bleaching and PALM steps. For imaging of strains with the SSB-mYpet fusion, 10 frames were taken under 514 nm excitation between the pre-bleaching and PALM imaging steps.

4.2.4 Data analysis

Live cell PALM imaging involves the analysis of large, complex data sets. To minimize potential sources of bias, we sought to automate as much of the analysis process as possible. Cell outlines were obtained using the MATLAB-based MicrobeTracker program, version 0.937 (18). A cell was eliminated from analysis by deleting its outline from the MicrobeTracker output only if it was partially obscured by the edge of the field, or if it contained a bright, background spot that did not bleach during the 561 nm pre-bleaching.

Portions of microscopy movies containing 405 nm + 561 nm PALM imaging (*i.e.*, excluding the pre-activation bleaching portion) were analyzed with the MATLAB-based u-track software (19) to locate and fit point-spread functions (PSFs) in each frame, and link PSFs together to form single-molecule tracks. The fit mode 'xyAsc' (variable sigma fits for the PSF position and amplitude) was used for all movies. For short exposure movies (13.3 ms), which are noisier, a significance criterion (α) of 0.01 was used to determine candidate localizations that were significantly distinguishable from background noise ($P < 0.01$). Single gaps in tracks were allowed, given the ability of fluorescent proteins to blink (20), but tracks were required to be at least two frames long to reduce the chance of spurious detections.

For movies with longer exposures (250 ms), a significance threshold of $P < 10^{-6}$ was used and no gaps were allowed in tracks one frame or longer. A maximum Brownian search radius of 3 (long exposures) or 5 pixels (short exposures) was used when linking PSFs in frames; a multiplier factor of 5 was used in the latter case. These parameters are used to link detections in sequential frames.

For movies containing SSB-mYpet imaging, those frames were extracted and the first five were averaged to produce a single SSB image. PSFs were located and fit using the xyAsc mode with a significance threshold of $P < 0.001$. For both PALM and SSB foci analysis, a custom MATLAB script, provided by the Harvard Medical School Image and Data Analysis Core, was used to define regions of interest in u-track corresponding to each cell outline determined in MicrobeTracker. Analyzing each cell in u-track individually prevented trajectories crossing from one cell to the other, improved tracking accuracy, and allowed fluorescence data to be more easily analyzed on a single cell basis.

Custom MATLAB scripts were written to take in u-track output files and generate arrays and structures containing the properties of each track or localization – positions, amplitudes, diffusion coefficients, etc. – tagged with movie number and cell number.

4.3 Results

4.3.1 A functional Pol IV-PAmCherry fusion

One major caveat to live cell imaging is that the addition at of a ~25 kDa fluorescent protein alters the behavior and activity of many proteins. In addition, many fluorescent proteins are able to form dimers (21), and even though dimerization surfaces have been the target of mutagenesis to generate monomeric variants, residual binding can alter cellular localization, especially for proteins that already participate in oligomerization (16, 20).

Since Pol IV is not essential for cell growth – the doubling rate of MG1655 *ΔdinB* in LB is 25.5 ± 0.8 min., compared to 24.9 ± 0.6 min. for MG1655 – viability of a Pol IV fusion strain does not necessarily indicate that the fusion is active for TLS. I therefore tested strains containing fusions of Pol IV to different fluorescent proteins for nitrofurazone (NFZ) resistance. NFZ is a DNA damaging agent that generates N^2-dG adducts that can be efficiently bypassed by Pol IV, and strains lacking Pol IV, or expressing variants that are impaired for catalysis, TLS, or clamp binding, are significantly sensitive to growth on plates containing 8 μM NFZ (6, 22, 23).

Using the λ Red recombineering system, I generated C-terminal fusions of the genomic Pol IV gene *dinB* to the photoactivatable proteins PAmCherry and mMaple3 using a 20-amino acid flexible Gly-Ser linker, and compared their NFZ resistance levels to the MG1655 parent strain and MG1655 Δ*dinB*. Although fusions of mMaple3 to other proteins are well behaved (20), the Pol IV-mMaple3 fusion was moderately sensitive to 8 μM NFZ, while the Pol IV-PAmCherry fusion exhibited normal levels of resistance (Figure 4.3). A partial inhibition of the Pol IV-PAmCherry strain growth at higher concentrations of NFZ (10.5 μM) reveals either a subtle TLS defect, or potentially reduced expression. Consistent with the previous report of a Pol IV-EYFP fusion, I found a Pol IV-mYpet fusion was even more sensitive to NFZ than the mMaple3 fusion, and that for both mYpet and PAmCherry, sensitivity increased with a reduced linker length (data not shown).

To verify that the full length Pol IV-PAmCherry was being expressed, and NFZ resistance was not due to proteolytic cleavage at the linker, I attempted to detect the

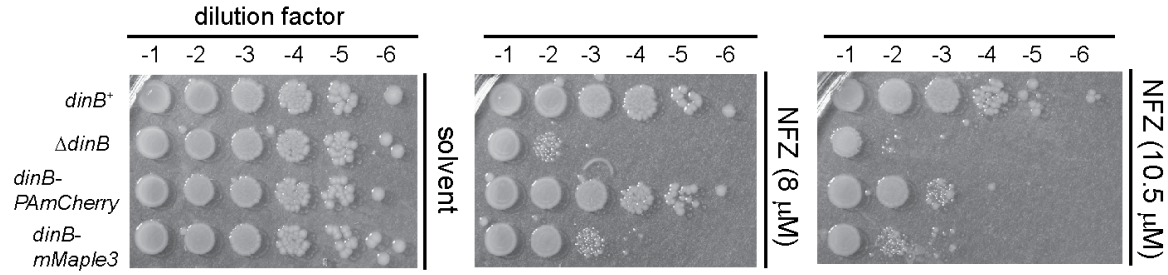

Figure 4.3 Growth of strains on plates containing 8 μM NFZ reveals a significant defect of a *ΔdinB* strain, a moderate defect of a strain containing a genomic fusion of Pol IV to mMaple3, and a near wildtype NFZ resistance of a Pol IV-PAmCherry strain. Growth on plates containing 10.5 μM NFZ reveals a subtle TLS defect of the Pol IV-PAmCherry strain. Images are representatives of three experiments.

fusion by Western blot analysis of whole cell lysates. Initial efforts to blot with antibodies against Pol IV (a gift of Veronica Godoy, Northeastern University) and mCherry (Adrian Salic, Harvard Medical School, and abcam #125096) were unsuccessful due to low antibody specificity in lysates. I therefore generated a chromosomal FLAG-tagged fusion, Pol IV-PAmCherry-FLAG, and transduced the allele into isogenic strains containing either the wildtype SOS repressor *lexA* (strain RW118), or *lexA51*, a Δ*lexA* frameshift mutation that generates constitutive SOS induction (strain RW542). Blotting with a specific anti-FLAG antibody revealed increased signal at the expected molecular length in the Δ*lexA* strain, evidence that the full length fusion is being produced by its native, LexA-regulated promoter (Figure 4.4). In the future, purified Pol IV-PAmCherry-FLAG will be used as a standard to estimate the cellular concentration and compare it to literature values.

I finally tested the NFZ resistance of strains containing Pol IV-PAmCherry in the isogenic *lexA*$^+$ and Δ*lexA* strains. Although the Δ*lexA* allele appears to (surprisingly) sensitize cells slightly to NFZ, no significant change in resistance was observed for the Δ*lexA dinB-PAmCherry* strain (Figure 4.5).

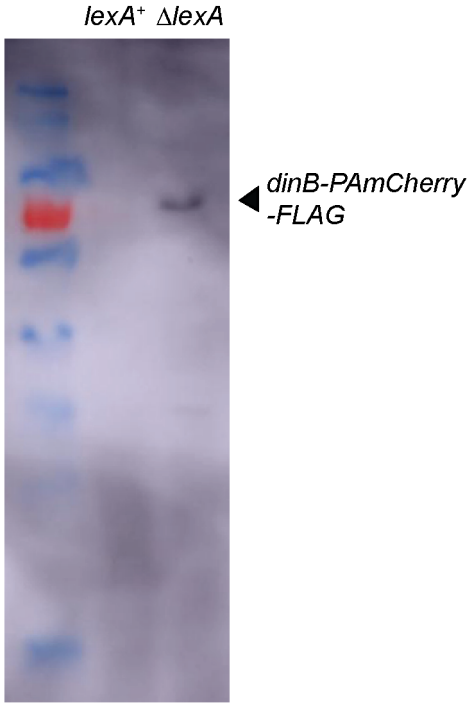

Figure 4.4 Western blot analysis of cell lysates expressing Pol IV-PAmCherry-FLAG reveals inefficient detection in a *lexA*⁺ strain (lane 2) and increased expression of the full-length Pol IV-PAmCherry-FLAG fusion (lane 3, expected length of 68.7 kDa). EZ-Run Prestained *Rec* Protein Ladder (Fisher Scientific), which contains a 72 kDa standard conjugated to an orange dye, was ran in lane 1.

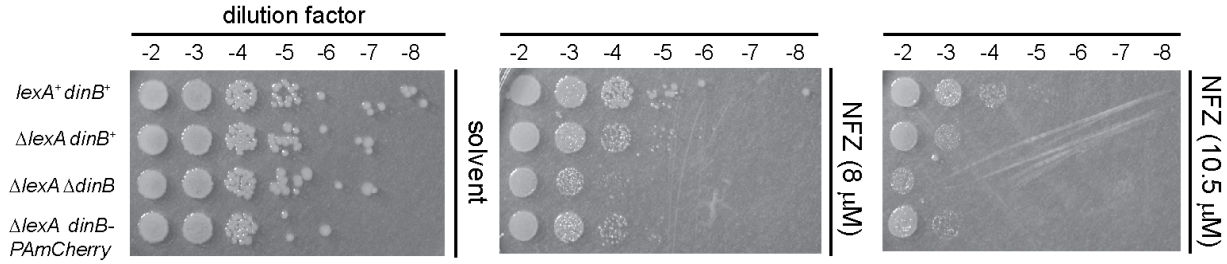

Figure 4.5 A Pol IV-PAmCherry strain containing the Δ*lexA* mutation has an equivalent NFZ resistance to its isogenic parent. Surprisingly, the Δ*lexA* strain, which expresses TLS polymerases to higher levels, is slightly more sensitive to NFZ than an otherwise equivalent *lexA*⁺ stain (note Δ*lexA* strains grow more slowly and therefore have smaller colonies).

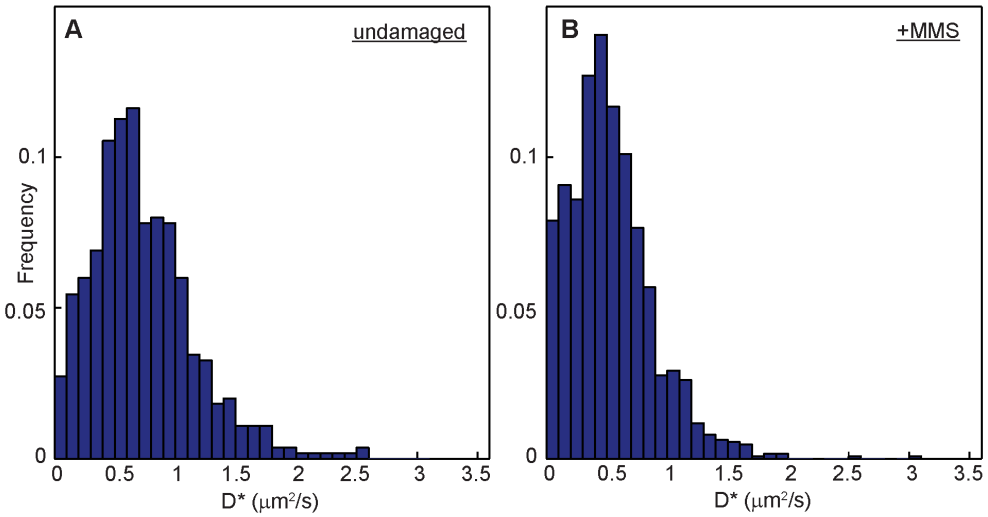

Figure 4.6 Distributions for the effective diffusion coefficient *D** of Pol I-PAmCherry: in normally growing cells (**A**, N = 550) molecules diffuse rapidly, while in MMS-treated cells (**B**, N = 1266), a fraction involved in DNA repair are statically bound to genomic sites. The distributions reported here are in good agreement with a previous measurements (5).

4.3.2 Imaging Pol IV-PAmCherry diffusion with short exposures

Before imaging the Pol IV-PAmCherry fusion, we first tested our ability to successfully perform PALM imaging in live cells using a previously described Pol I-PAmCherry fusion strain, a generous gift of Stephan Uphoff of the University of Oxford. Low intensity 405 nm laser illumination is used to stochastically activate individual PAmCherry molecules, and simultaneous high intensity 561 nm illumination is used to excite and image their localization within cells. A short exposure time (13.3 ms, the minimum for a 208x208 pixel ROI) was used to detect rapidly diffusing molecules, and for trajectories that lasted at least 5 frames before bleaching, we were able to calculate the effective (2D) diffusion coefficient (*D**) for molecules.

In good agreement with the reported description of this strain (5), we found that Pol I-PAmCherry rapidly diffuses throughout the cell during normal growth,

while pre-incubation of cells on an agar pad containing 100 mM methyl methanesulfonate (MMS) reveals an increased population of statically bound (low diffusion coefficient) molecules, indicative of excision repair (Figure 4.6).

In comparison to the Pol I-PAmCherry strain, where we were able to activate and image on average 53 molecules per cell, we were only able to image 12 molecules per cell in the MG1655 *dinB-PAmCherry* strain (Figure 4.7). Although the total number of fluorescent proteins in a cell cannot be accurately counted due to the presence of an immature, dark fraction, inefficient photoactivation (90-95% undetected for PAmCherry), and blinking (20), we had expected a greater number of Pol IV-PAmCherry molecules per cell, reflecting their reported cellular copy numbers in *lexA*$^{+}$ cells (~500 molecules per cell for Pol I and ~200 for Pol IV) (24).

This discrepancy could be explained by reduced expression of the Pol IV-PAmCherry fusion relative to endogenous Pol IV levels, or inefficient localization and tracking of the smaller, faster diffusing Pol IV fusion. Although it does not suggest a serious problem with the Pol IV-PAmCherry strain, one concern was that the low numbers of imaged molecules per cell might make imaging sensitive to presence of rare background fluorescence spots.

To increase our experimental throughput, and we switched to the Δ*lexA dinB-PAmCherry* strain described above, recognizing that deleting the *lexA* repressor changes the cellular context by increasing the copy number of not only the Pol IV-PAmCherry fusion, but also the other members of the SOS regulon (25). As expected, the number of localizations in the Δ*lexA* strain increased significantly, to an average of 32 molecules per cell (Figure 4.7). Analyzing the distribution of the effective

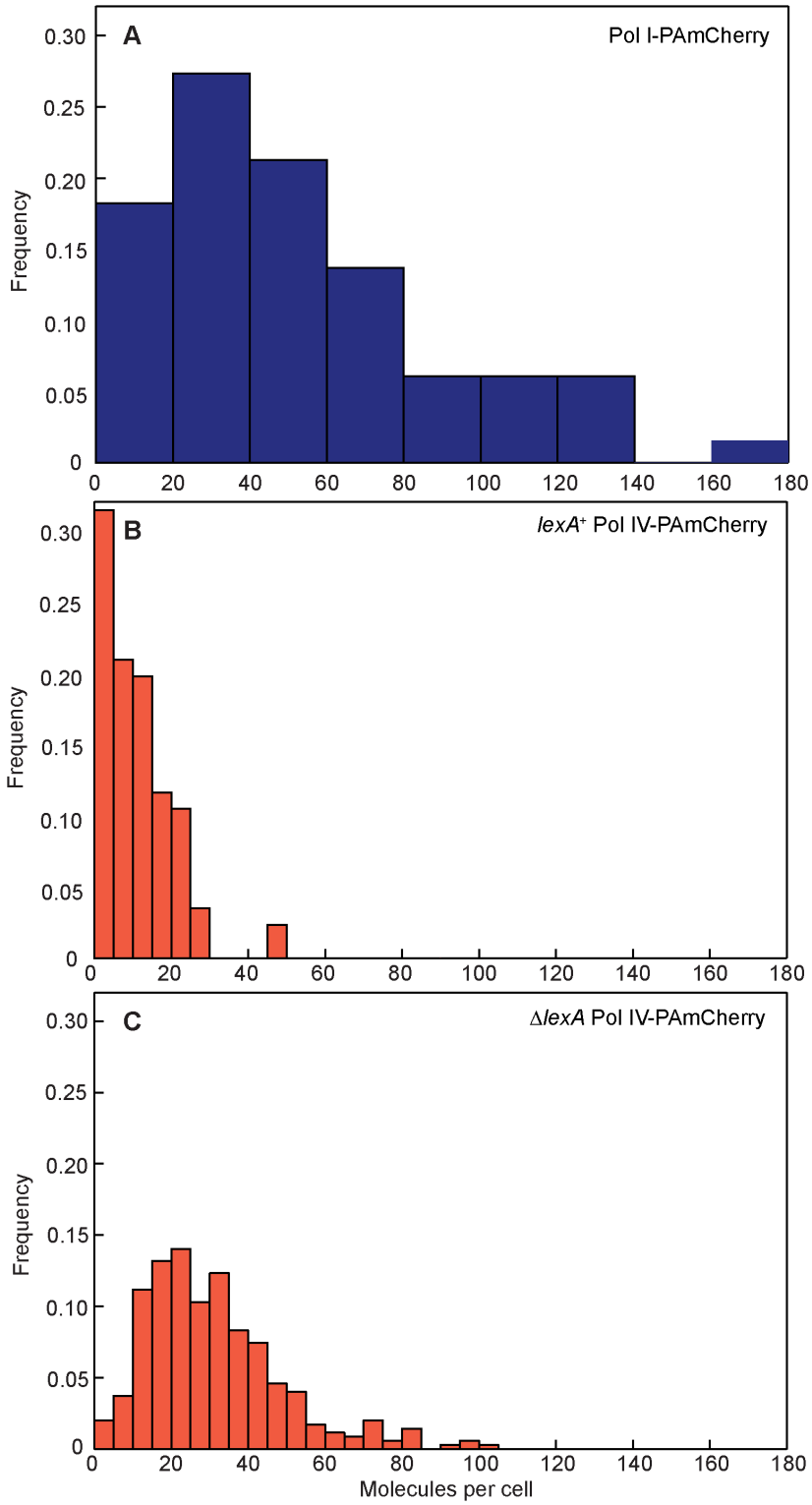

Figure 4.7 The number of molecules per cell for (**A**) AB1157 *polA-pamcherry* (*N* = 66 cells) (**B**) MG1655 *dinB-pamcherry* (*N* = 86), and (**C**) RW542 (Δ*lexA*) *dinB-pamcherry* (*N* = 349), in which the Pol IV-PAmCherry expression increases. To minimize the chance that spurious detections were scored as molecules at these short exposures, localizations were required to be observed in at least two successive frames.

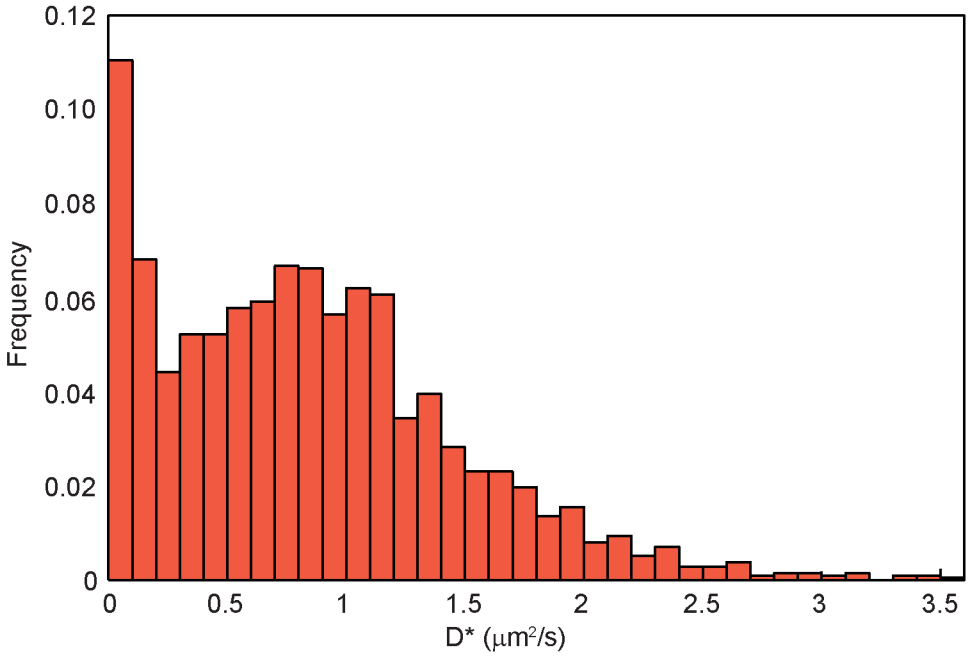

Figure 4.8 The distribution for the diffusion coefficient of Pol IV-PAmCherry molecules (N = 2,129, 5 frames or longer) in a Δ*lexA* background in the absence of DNA damage. A significant fraction of molecules is static.

diffusion coefficient (Figure 4.8) revealed that Pol IV-PAmCherry diffuses faster than Pol I-PAmCherry, as expected based on their molecular weights (68.7 kDa vs. 130.7 kDa); interestingly, we observed a sizeable fraction of statically bound Pol IV molecules (22.2% with $D^* < 0.15$ μm²/s).

4.3.3 Imaging Pol IV-PAmCherry localizations with long exposures

While short exposure times give a snapshot of both static and diffusing Pol IV-PAmCherry molecules, longer exposures average out the signal of the mobile fraction, enabling the selective observation of bound molecules. Using a 250 ms exposure with a concurrent reduction in excitation power, we observed an average of 11.2 static Pol IV-PAmCherry molecules per cell (Figure 4.9A). These localizations

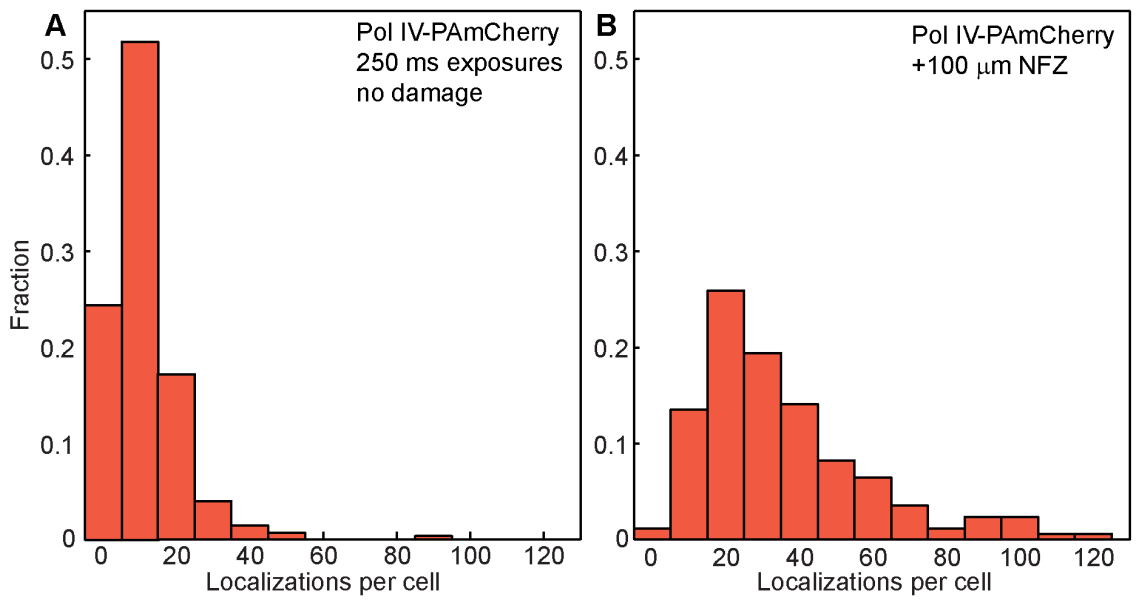

Figure 4.9 Distributions for the number of static localizations per cell for the Δ*lexA* Pol IV-PAmCherry strain at long exposures (**A**) during normal growth (N = 274 cells) and (**B**) after incubation with 100 μM NFZ for one hour (N = 170).

were distributed evenly throughout the cell, as seen by their distribution along the long and short cell axes (Figure 4.10A).

Treating cells in liquid culture for one hour with 100 μM NFZ significantly increased the number of localizations under equivalent imaging conditions, to 37.2 per cell (Figure 4.9B). The increased number of localizations was not due to increased autofluorescence due to damage, indicated by imaging of the NFZ-treated parent strain (data not shown). Although cell growth was reduced after treatment, cell viability was largely unchanged (fold change: 1.20 ± 0.99, mean and standard deviation), as determined by plating for colony-forming units (CFUs); during the same period, the number of CFUs for solvent-treated cells increased by a factor of 2.54 ± 0.99.

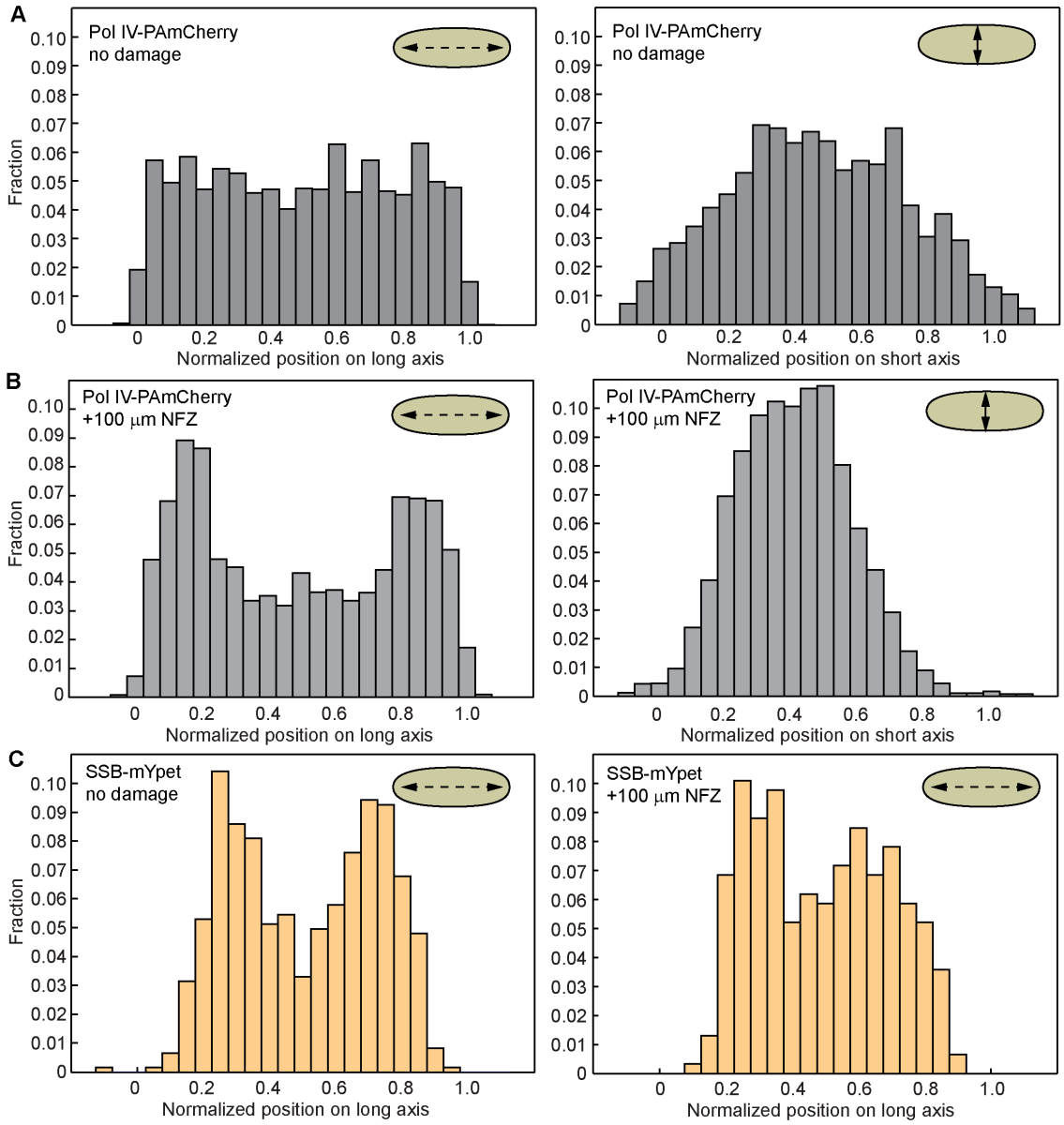

Figure 4.10 Positions of all Pol IV-PAmCherry localizations in cells (**A**) during normal growth (N = 3,078 localizations) and (**B**) after treatment with 100 μM NFZ for one hour (N = 6,317). Normalized positions along both the long and short cell axes are displayed. For comparison, the positions of SSB-mYpet along the long cell axis in the (**C**) absence (N = 605) and (**D**) presence (N = 307) of NFZ damage are displayed.

Interestingly, while localizations were distributed uniformly throughout the cell during normal growth, they appeared to preferentially localize at the ~20% and ~80% positions along the long cell axis (Figure 4.10B). In order to determine if this binding could correspond to replisomes at the quarter cell positions, we introduced a fusion of the Single-stranded DNA binding protein (SSB) to the bright, yellow fluorescent protein mYpet as a second copy within the chromosome's *lac* operon; including IPTG in the culture induces expression of SSB-mYpet, specifically labeling the replisome (1) without a reduction in cell viability (Table 4.3).

Using two-color PALM microscopy, we were able to image SSB-mYpet and Pol IV-PAmCherry orthogonally and observe its position along the cell axis (Figure 4.10C,D). Measuring the distance between Pol IV-PAmCherry molecules to the nearest SSB focus within each cell, we found Pol IV-PAmCherry localized moderately closer to the replisomes in the presence of DNA damage (Figure 4.11, localizations within 200 nm SSB increases from 11.1% to 14.6%).

Finally, we found that the timescale of polymerase binding, inferred from the number of frames localizations ere observed, was significantly longer in the presence of NFZ ($P < 0.001$). By binning trajectory lengths and fitting to an exponential distribution, we were able to estimate these time scales (Figure 4.12). For normal growth, binding occurred with a time constant of 151 ms. Cells treated with NFZ, however, were not accurately fit with a single exponential, so we used the sum of two exponentials, revealing two timescales: 146 ms (96.9%) and 824 ms (3.1%). Fitting the undamaged sample to two exponentials also revealed a longer time scale (1.13 s), but at a lower fraction (0.9%), suggesting that the treatment

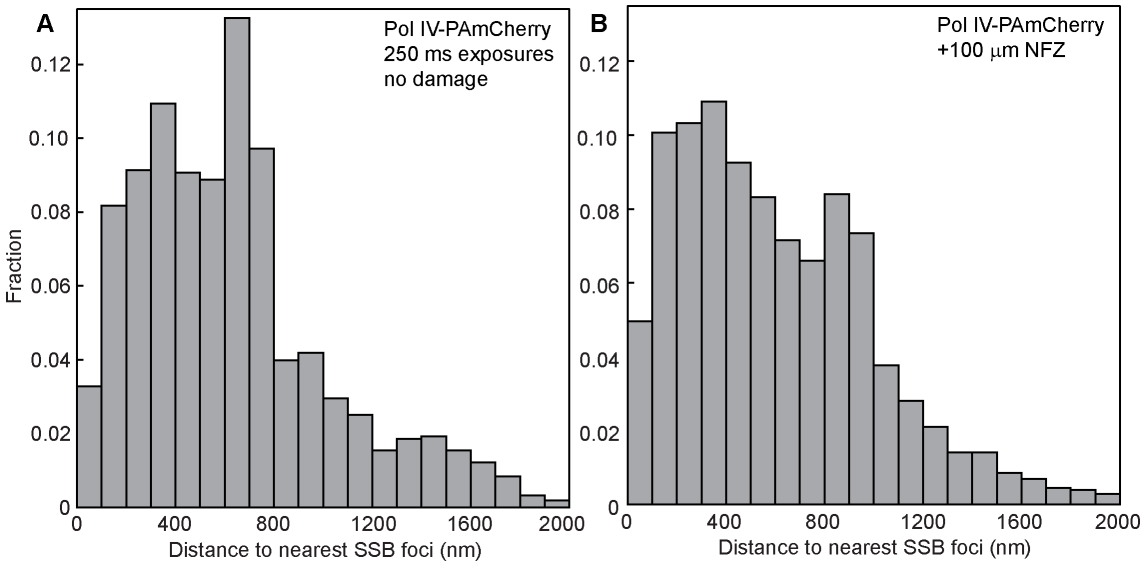

Figure 4.11 Distributions for the distance of each Pol IV-PAmCherry in cells to the nearest SSB focus in nanometers (nm), (**A**) during normal growth (N = 3,078) and (**B**) after treatment with 100 μM NFZ for one hour (N = 6,317).

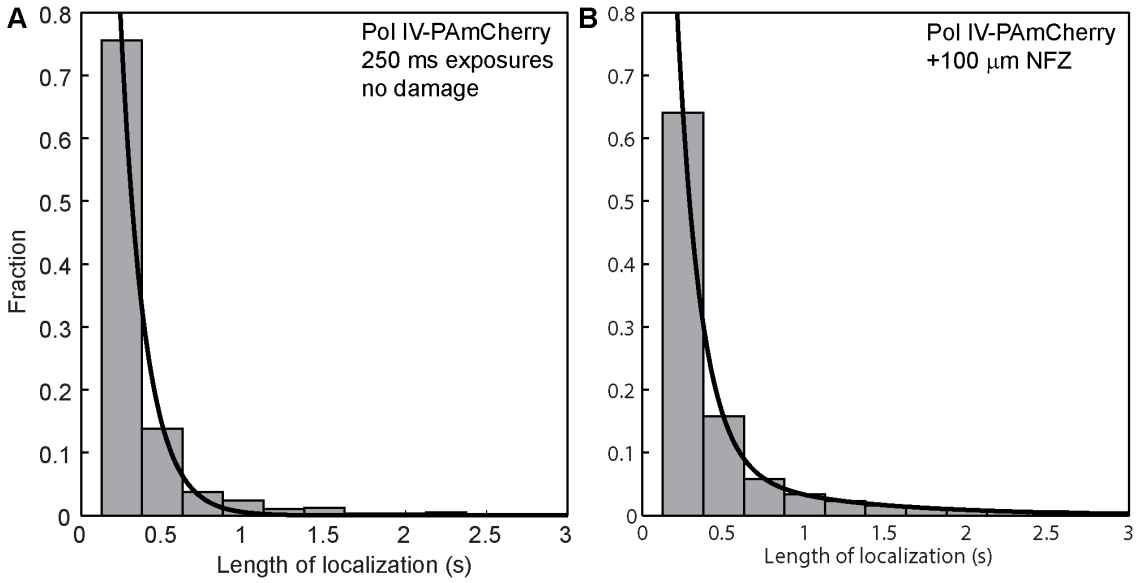

Figure 4.12 The length in seconds of Pol IV-PAmCherry localizations at long exposures, (**A**) during normal growth (N = 3,078) and (**B**) after treatment with 100 μM NFZ for one hour (N = 6,317). Displayed fits are to a single (A) or double (B) exponential function.

with NFZ either results in the appearance of or increase in a separate population of bound molecules.

4.4 Discussion and future directions

In this chapter, I report a novel fusion of Pol IV to the fluorescent protein PAmCherry, with near wildtype TLS activity. Using PALM imaging we are able to localize individual Pol IV-PAmCherry molecules and determine their binding times. Even in cells that are activated for the SOS response, Pol IV-PAmCherry binds throughout the cell, without any clear preference for replication forks. Although these data do not exclude the possibility of a low occupancy of Pol IV, they argue that Pol IV is not binding to a large fraction of the ~50 β dimers that localize at and behind the replisome. In contrast, we find that treatment of cells with NFZ, which produces Pol IV cognate lesions, increases the prevalence and timescale of Pol IV binding. Binding is biased near the one and three-quarters cell position, perhaps reflecting binding at stalled or broken replication forks.

Several next steps are needed to further characterize Pol IV binding in live cells. First, to selectively observe the longer binding timescale in NFZ-treated cells, we will image cells with longer exposures (500, 750, and 1000 ms). We will also image samples fixed with formaldehyde at each timescale to correct for photobleaching; after the correction, binding times should be equivalent across exposures (5). We will additionally treat cells with different concentrations of NFZ to determine how binding varies, and with other damaging agents that generate Pol IV cognate (4-NQO) or non-cognate (UV) DNA lesions.

A major question is what factors are recruiting Pol IV. The broad distribution of Pol IV molecules in the absence of damage could either represent non-specific binding to DNA, or binding to the RNA polymerase subunit NusA (7); as NusA is essential (26), we will use *nusA11,* a temperature-sensitive allele impaired for Pol IV-dependent TLS (27), to test this hypothesis. Since TLS by Pol IV (and NFZ resistance) depends on binding to β (6), we will additionally image damage-treated strains containing Pol IV fusions with its clamp-binding residues mutated, using different exposures. Pol IV was recently shown to interact directly with Pol III (28), and Pol IV(T120P), a mutant impaired for this interaction, can also be tested. Another possibility is that the localizations we observe are at double-strand breaks (DSBs) through an interaction with RecA (8); we will test this hypothesis by imaging NFZ-treated cells with a fluorescent fusion of the DSB-binding phage protein Gam (29), and by imaging Pol IV(C66A), a mutant that binds RecA more strongly (30).

4.5 Contributions

The imaging and analysis for this chapter was done in close collaboration with Elizabeth Thrall, a postdoctoral fellow in the Loparo lab. Dr. Thrall also designed the custom fluorescence microscope. This project benefited from helpful discussions with Stephan Uphoff (University of Oxford), who generously provided the *polA-pamcherry* strain. Initial imaging experiments were performed at the Harvard Medical School Nikon Imaging Center with help from Talley Lambert; assistance and custom modifications to the u-track software package were provided by Hunter Elliott and Joy Xu of the Harvard Medical School Image and Data Analysis Core.

References

1. Reyes-Lamothe R, Possoz C, Danilova O, Sherratt DJ (2008) Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* 133(1):90–102.

2. Reyes-Lamothe R, Sherratt DJ, Leake MC (2010) Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* 328(5977):498–501.

3. Moolman MC, *et al*. (2014) Slow unloading leads to DNA-bound β_2-sliding clamp accumulation in live *Escherichia coli* cells. *Nature Commun* 5:5820.

4. Uphoff S, Kapanidis AN (2014) Studying the organization of DNA repair by single-cell and single-molecule imaging. *DNA Repair* 20:32–40.

5. Uphoff S, Reyes-Lamothe R, Garza de Leon F, Sherratt DJ, Kapanidis AN (2013) Single-molecule DNA repair in live bacteria. *Proc Natl Acad Sci USA* 110(20): 8063–8068.

6. Heltzel JM, Maul RW, Scouten Ponticelli SK, Sutton MD (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc Natl Acad Sci USA* 106(31):12664-12669.

7. Cohen SE, Godoy VG, Walker GC (2009) Transcriptional modulator NusA interacts with translesion DNA polymerases in *Escherichia coli*. *J Bacteriol* 191(2):665–672.

8. Godoy V, *et al*. (2007) UmuD and RecA directly modulate the mutagenic potential of the Y family DNA polymerase DinB. *Mol Cell* 28(6):1058–1070.

9. Robinson A, *et al*. (2015) Regulation of Mutagenic DNA Polymerase V Activation in Space and Time. *PLoS Genet* 11(8):e1005482.

10. Mallik S, Popodi EM, Hanson AJ, Foster PL (2015) Interactions and localization of *Escherichia coli* error-prone DNA polymerase IV after DNA damage. *J Bacteriol* 197(17):2792–2809.

11. Sharan SK, Thomason LC, Kuznetsov SG, Court DL (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* 4(2):206–223.

12. Datsenko K, Wanner B (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97(12):6640-6645.

13. Gibson DG, *et al*. (2009) Enzymatic assembly of DNA molecules up to several

hundred kilobases. *Nat Meth* 6(5):343–345.

14. Subach FV, *et al*. (2009) Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat Meth* 6(2):153–159.

15. Antony E, *et al*. (2013) Multiple C-terminal tails within a single *E. coli* SSB homotetramer coordinate DNA replication and repair. *J Mol Biol* 425(23):4802–4819.

16. Landgraf D, Okumus B, Chien P, Baker TA, Paulsson J (2012) Segregation of molecules at cell division reveals native protein localization. *Nat Meth* 9(5):480–482.

17. Lohman TM, Green JM, Beyer RS (1986) Large-scale overproduction and rapid purification of the *Escherichia coli* ssb gene product. Expression of the ssb gene under λ P_L control. *Biochemistry* 25(1):21–25.

18. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C (2011) High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol Microbiol* 80(3):612–627.

19. Jaqaman K, *et al*. (2008) Robust single-particle tracking in live-cell time-lapse sequences. *Nat Meth* 5(8):695–702.

20. Wang S, Moffitt JR, Dempsey GT, Xie XS, Zhuang X (2014) Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. *Proc Natl Acad Sci USA* 111(23):8452–8457.

21. Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* 67:509–544.

22. Jarosz DF, Godoy VG, Delaney JC, Essigmann JM, Walker GC (2006) A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature* 439(7073):225–228.

23. Jarosz DF, Cohen SE, Delaney JC, Essigmann JM, Walker GC (2009) A DinB variant reveals diverse physiological consequences of incomplete TLS extension by a Y-family DNA polymerase. *Proc Natl Acad Sci USA* 106(50):21137-21142.

24. Sutton MD (2010) Coordinating DNA polymerase traffic during high and low fidelity synthesis. *Biochim Biophys Acta* 1804(5):1167–1179.

25. Courcelle J, Khodursky A, Peter B, Brown PO, Hanawalt PC (2001) Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* 158(1):41–64.

26. Cardinale CJ, *et al*. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* 320(5878):935–938.

27. Cohen SE, *et al*. (2010) Roles for the transcription elongation factor NusA in both DNA repair and damage tolerance pathways in *Escherichia coli*. *Proc Natl Acad Sci USA* 107(35):15517–15522.

28. Scotland MK, *et al*. (2015) 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(9):e1005507.

29. Shee C, *et al*. (2013) Engineered proteins detect spontaneous DNA breakage in human and bacterial cells. *Elife* 2:e01222.

30. Cafarelli TM, *et al*. (2013) A single residue unique to DinB-like proteins limits formation of the polymerase IV multiprotein complex in *Escherichia coli*. *J Bacteriol* 195(6):1179–1193.

31. Ho C, Kulaeva OI, Levine AS, Woodgate R (1993) A rapid method for cloning mutagenic DNA repair genes: isolation of *umu*-complementing genes from multidrug resistance plasmids R391, R446b, and R471a. *J Bacteriol* 175(17):5411–5419.

32. Fernández De Henestrosa AR, *et al*. (2000) Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol Microbiol* 35(6):1560–1572.

Chapter 5
Discussion and future directions

5.1 Conclusions

As outlined in Chapter 1, this dissertation sought to clarify the mechanism of polymerase exchange and its implications for translesion synthesis (TLS). In particular, one major question was the validity of the toolbelt model, a proposal that both a replicative polymerase and a translesion polymerase can simultaneously bind the same multimeric ring-shaped processivity clamp (1). This mechanism could facilitate polymerase exchange after the replicative polymerase encounters a DNA lesion roadblock, and the resumption of normal synthesis after translesion synthesis.

Previous studies were not able to observe polymerase exchange during DNA synthesis, and did not study exchange in the presence of a lesion (2, 3). I therefore developed a single-molecule reconstitution of the full polymerase exchange reaction, including bypass of a site-specific and chemically defined lesion. Using this approach, I have shown that the *Escherichia coli* replicative polymerase Pol III and two separate TLS polymerases, Pol IV (Chapter 2) and Pol II (Chapter 3), can simultaneously bind the processivity clamp β and rapidly exchange.

Another major question is under what contexts polymerase exchange and TLS occurs. Toolbelt binding of Pol IV to clamps during Pol III replication could make Pol IV a core component of the *E. coli* replisome. However; as the three other polymerases and at least three additional replication-associated proteins contain

clamp-binding motifs that interact with β (4), the proposed Pol III-Pol IV-β complex may only exist under some contexts, such as in the presence of DNA damage, or during the SOS DNA damage response, when the Pol IV copy number increases (5). Given recent reports that the reconstituted *E. coli* replisome can "skip" lesions on the leading strand, leaving behind single-stranded (ss) DNA gaps (6, 7), it is also possible that exchange between Pol III and Pol IV may primarily occur on clamps left behind in these gaps.

In an initial effort to answer these questions, discussed in Chapter 4, I developed a functional fusion of Pol IV to PAmCherry, which I used to observe the diffusion and binding of individual polymerases molecules in living cells. Even in cells constitutively induced for the SOS response, the majority of polymerases diffuse broadly throughout the cell. These results suggest that Pol IV does not associate with the replisome in the absence of damage, although it does not eliminate the possibility of low occupancy binding that would be difficult to detect. Pol IV localization increases dramatically after treatment with nitrofurazone, and we are currently investigating if this depends on interactions with β or other Pol IV-binding proteins.

5.2 Future directions

5.2.1 Reconstituting polymerase exchange within the full replisome

As the *in vitro* studies in this dissertation have largely focused on reconstituting polymerase exchange and TLS using a single-molecule primer extension assay, a natural next step would be to use a single-molecule rolling circle

replication assay previously used to study coordinated synthesis by the fully reconstituted T7 and Pol III replisomes (8). Instead of observing synthesis by motion of a tethered bead, long DNA products (up to hundreds of kilobase pairs, or kb) can be directly visualized using sparse labeling with an intercalating dye. As the DNA substrate uses M13 ssDNA as the template (leading) strand, it could be easily adapted to use the lesion-containing M13mp7(L2) substrate constructed for Chapter 2. The Pol III replisome would therefore encounter a lesion on the leading strand every ~7.2 kb.

One limitation of this assay is that since it does not use a large, bright bead to precisely determine DNA length and suppress its fluctuations, it would be unable to resolve short synthesis events by TLS polymerases, which would instead appear as pauses in rapid replisome progression. However, the absence of a large, bright bead makes the simultaneous imaging of two fluorescently labeled polymerases possible. Fluorescently labeled Single-stranded DNA-binding protein (SSB) on the lagging strand could also be used to track the replication fork. This could potentially be used to observe "lesion skipping" by Pol III, which involves the slow generation of a single-stranded gap past the leading-strand lesion (and a corresponding increase in SSB binding) before synthesis resumes (6). The ability to simultaenously observe the Pol III and Pol IV stoichiometry within the replication fork, and correlate it with replisome behavior, the kinetics of TLS, and even the probability of replisome collapse would be powerful.

This technique also allows the flow cell to be used like a affinity column, by pre-loading the replisome onto the substrates and rapidly removing free

components by flow (9). Such an approach could be used to definitely test if Pol III remains on the clamp during TLS before it resumes synthesis, and the number of times a back-and-forth exchange can occur in the absence of free polymerases in solution.

5.2.2 Imaging SSB to distinguish continuous TLS and lesion skipping

One major limitation of the live cell polymerase imaging approach described in Chapter 4 is its inability to observe the dynamics of DNA replication within a single cell – unlike *in vitro* assays where DNA substrates can be extended and synthesis spatially resolved, the relevant activity is generally constrained within a diffusion-limited spot.

One promising approach is to use the intensity of fluorescently labeled SSB within a replication fork to infer replisome dynamics, as discussed above. A constant (but perhaps fluctuating) amount of ssDNA, and therefore SSB, would reflect normal synthesis on the lagging strand. Rapid exchange to an associated TLS polymerase on the leading strand in the presence of DNA damage would likely result in a rapid reduction in the SSB signal as replisome progression temporarily slows and Pol III on the lagging strand completes its Okazaki fragment, while a slow increase in SSB, followed by the resumption of normal synthesis would correspond to lesion skipping. The choice between these two "pathways" could be correlated with the type of damage and the presence of labeled TLS polymerases at the replication fork. As increased concentrations of the primase DnaG increases "lesion skipping" activity *in vitro* (7), this could also be tested *in vivo*.

One challenge with developing such an assay is the lack of sufficiently photostable fluorescent proteins. Even though one study reported that fluctuations in the intensity of SSB within the replication fork could be correlated with lagging strand synthesis (10), my initial efforts to replicate it were unsuccessful due to rapid photobleaching, and this study was later retracted. An alternative approach would be to label SSB in cells with bright, photostable organic dyes using the SNAP-tag technology (11), or by labeling recombinant SSB *in vitro* (12) and electroporating it into cells (13).

5.2.3 Regulators of polymerase exchange

The development of single-molecule techniques to study TLS can also be used to study how other *Escherichia coli* factors influence polymerase access. The recombinase RecA and the Pol V subunit $UmuD_2$ have previously been shown to bind Pol IV *in vitro* and reduce its frameshifting mutagenesis (14, 15). RecA also alters competition between Pol III and Pol IV or Pol V *in vitro* (16). Single-molecule studies could clarify the mechanism by which these proteins influence polymerase competition.

Several additional candidates regulators were identified in a genetic selection for multicopy suppressors of Pol IV overproduction lethality, which I describe in the Appendix. Two interesting candidates are Rep, a replisome-associated helicase that interacts with Pol IV (17) and increases replisome stability (18), and RecB, a component of the RecBCD complex that processes double-strand breaks and loads RecA (19). If validated, these suppressors can be purified and

included in single-molecule assays, or overproduced in cells during live-cell imaging of Pol IV.

5.2.4 Observing exchange within the eukaryotic replisome

A major goal of this dissertation was the development of paradigms for polymerase exchange and the regulation of TLS. An important direction is to test if these principles extend to the eukaryotic replisome. A previous study observed rapid replacement of a stalled *Saccharomyces cerevisiae* replicative polymerase, Pol δ, and the translesion polymerase Pol η on PCNA (20). The single-molecule primer extension assay used to study exchange between Pol III and Pol II or Pol IV could be adapted to study exchange between *S. cerevisiae* or human polymerases. Interesting questions are the validity of the toolbelt model for the homotrimeric PCNA, the role of PCNA ubiquitination in exchange (21), and whether the human homolog of Pol IV, Pol κ, has an analogous inactive binding mode. Although the large size of eukaryotic cells makes fluorescence imaging challenging, recent advances in light-sheet microscopy (22, 23) will additionally allow for quantitative imaging of replisome dynamics and TLS within the nucleus of mammalian cells.

References

1. Pagès V, Fuchs RP (2002) How DNA lesions are turned into mutations within cells? *Oncogene* 21(58):8957–8966.

2. Indiani C, McInerney P, Georgescu R, Goodman M, O'Donnell M (2005) A sliding-clamp toolbelt binds high-and low-fidelity DNA polymerases simultaneously. *Mol Cell* 19(6):805–815.

3. Heltzel JM, Maul RW, Scouten Ponticelli SK, Sutton MD (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc Natl Acad Sci USA* 106(31):12664.

4. Wijffels G, *et al*. (2004) Inhibition of protein interactions with the β sliding clamp of *Escherichia coli* DNA polymerase III by peptides from β-binding proteins. *Biochemistry* 43(19):5661–5671.

5. Courcelle J, Khodursky A, Peter B, Brown PO, Hanawalt PC (2001) Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* 158(1):41–64.

6. Yeeles JT, Marians KJ (2011) The *Escherichia coli* replisome is inherently DNA damage tolerant. *Science* 334(6053):235–238.

7. Yeeles JT, Marians KJ (2013) Dynamics of leading-strand lesion skipping by the replisome. *Mol Cell* 52(6):855–865.

8. Tanner NA, *et al*. (2008) Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Res* 37(4):e27.

9. Tanner NA, *et al*. (2011) *E. coli* DNA replication in the absence of free β clamps. *EMBO J* 30(9):1830–1840.

10. Lia G, Michel B, Allemand J-F (2011) Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science* 335(6066):328-331.

11. Stöhr K, *et al*. (2010) Quenched substrates for live-cell labeling of SNAP-tagged fusion proteins with improved fluorescent background. *Anal Chem* 82(19):8186–8193.

12. Roy R, Kozlov AG, Lohman TM, Ha T (2009) SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* 461(7267):1092–1097.

13. Sustarsic M, *et al*. (2014) Optimized delivery of fluorescently labeled proteins in live bacteria using electroporation. *Histochem Cell Biol* 142(1):113–124.

14. Godoy V, *et al*. (2007) UmuD and RecA directly modulate the mutagenic potential of the Y family DNA polymerase DinB. *Mol Cell* 28(6):1058–1070.

15. Foti JJ, Delucia AM, Joyce CM, Walker GC (2010) $UmuD_2$ inhibits a non-covalent step during DinB-mediated template slippage on homopolymeric nucleotide runs. *J Biol Chem* 285(30):23086–23095.

16. Indiani C, Patel M, Goodman MF, O'Donnell ME (2013) RecA acts as a switch to regulate polymerase occupancy in a moving replication fork. *Proc Natl Acad Sci USA* 110(14):5410–5415.

17. Sladewski TE, Hetrick KM, Foster PL (2011) *Escherichia coli* Rep DNA helicase and error-prone DNA polymerase IV interact physically and functionally. *Mol Microbiol* 80(2):524–541.

18. Guy CP, *et al*. (2009) Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol Cell* 36(4):654–666.

19. Friedberg EC, *et al*. Recombinational repair, replication fork repair, and DNA damage tolerance. (2005) *DNA Repair and Mutagenesis*, 2nd. (ASM Press, Washington,D.C.), pp. 569-612.

20. Zhuang Z, *et al*. (2008) Regulation of polymerase exchange between Pol η and Pol δ by monoubiquitination of PCNA and the movement of DNA polymerase holoenzyme. *Proc Natl Acad Sci USA* 105(14):5361.

21. Freudenthal BD, Gakhar L, Ramaswamy S, Washington MT (2010) Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nat Struct Mol Biol* 17(4):479–484.

22. Zhao ZW, *et al*. (2014) Spatial organization of RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution microscopy. *Proc Natl Acad Sci USA* 111(2):681–686.

23. Chen BC, *et al*. (2014) Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science* 346(6208):1257998.

Appendix

Selecting for suppressors of Pol IV over-production lethality

A.1 Introduction

The majority of studies on polymerase exchange in TLS in *Escherichia coli* – including much of this thesis – have concentrated on the mechanism by which TLS polymerases gain access to the replication fork. However, given the slow rate and higher mutagenesis of Pol IV, it is clear that even though Pol IV is in excess of Pol III during normal growth, the fraction of genomic DNA it synthesizes must be low. This suggests several possibilities: (i) even if Pol IV associates with the replisome under normal conditions, exchange is inefficient in the absence of a lesion-induced stall and the DNA synthesized following exchange is limited; (ii) TLS largely occurs in gaps left behind the replisome as it "skips" over a lesion (1), and even in those gaps, the amount of DNA synthesized by a TLS polymerase is small; (iii) regulatory controls actively prevent unscheduled access by Pol IV.

These three possibilities are by no means mutually exclusive and relative contributions could depend on cellular context. However, a few gene products have been identified that may limit Pol IV access to the replication fork. These were identified in a screen by Graham Walker's lab at MIT to find Pol IV-binding proteins. Two hits, the recombinase RecA and the Pol V component $UmuD_2$, were shown to form a complex with Pol IV *in vitro* and reduce Pol IV frameshifting mutagenesis (2, 3). Another hit was the conserved RNA polymerase subunit NusA, an essential

protein involved in termination and transcription-coupled nucleotide excision repair. NusA was also shown to bind Pol IV *in vitro* (4), and the temperature-sensitive allele *nusA11* was found to be sensitive to damaging agents that produce Pol IV-specific cognate lesions (5).

A.2 Methods

In an effort to find additional genes that may limit unscheduled Pol IV access, I developed an unbiased screen for multicopy Pol IV suppressors. This approach took advantage of the fact that expressing Pol IV beyond its SOS-induced levels is lethal (6). For IPTG-induced expression from the medium copy vector pINIII, killing by Pol IV is dependent on its incorporation of 8-oxo-dGTP, followed by unchecked strand cleavage by base excision repair nucleases (7). Our lab has further shown using this system that lethality can be suppressed by mutating either of the two β–binding interactions, and reduced by increasing the Pol III exonuclease subunit's affinity for β, arguing that unscheduled access involves competition of Pol IV with Pol III for β (Seungwoo Chang, in preparation).

Although previous Pol IV killing assays have typically used arabinose- or IPTG-inducible expression of the Pol IV gene *dinB* from a plasmid, I instead chose to introduce a Tet-inducible *dinB* expression cassette to the *E. coli* genome. The Tet expression system is superior as it lacks the all-or-nothing induction of the P_{BAD} and P_{lac} operons (8); expression from the genome also simplifies screening and minimizes the chance of selecting for *dinB* suppressor mutants.

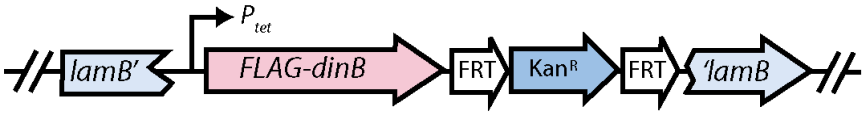

Figure A.1 The genomic *dinB* expression cassette was introduced to the genome, disrupting the *lamB* gene. Versions were made with and without the N-terminal FLAG tag residues MDYKDDDDKGS.

Using information from the iGEM parts registry, I constructed sequences that contain two Tet repressor-binding sequences, a strong ribosome binding site, and *dinB* (Figure A.1). Two versions of the *dinB* construct were made, one untagged, and one that contained a N-terminal FLAG tag. I then used the λ Red system to knock these two cassettes into the non-essential *lamB* gene of the strain MG1655 Z2, which constitutively expresses the Tet repressor from the genome (9).

For this screen, I obtained a multicopy plasmid library containing 1- to 4-kb random *E. coli* genomic DNA fragments cloned at the multiple cloning site of the p15A-based vector, pBAD33, a generous gift from of Thomas Bernhardt (Harvard Medical School). Plating viability of the Pol IV expression strain containing the empty pBAD33 vector was reduced by five orders of magnitude in the presence of 20 ng/mL anhydrotetracycline (aTc) inducer (Figure A.2), and was equivalent for strains expressing untagged and FLAG-tagged Pol IV (data not shown).

I next transformed the pBAD33 genomic library into an electrocompetent preparation of the P_{tet}-*FLAG-dinB* strain, selecting for plasmids in the absence of *dinB* induction. The transformation generated 9×10^6 colony-forming units (~8% of viable CFUs) after a minimal recovery period. Assuming (conservatively) an average insert size of 1500 bp, this represents ~3000X coverage of the *E. coli* genome, although coverage is likely lower due to redundancy, and the non-uniform

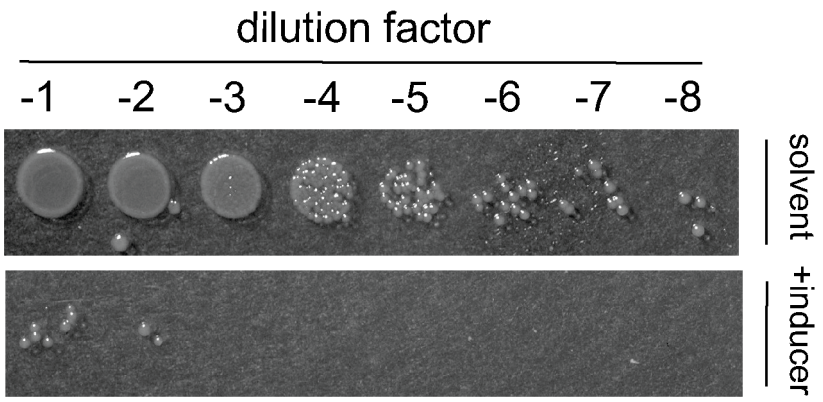

Figure A.2 The Pol IV expression strain MG1655 *lamB::*P_{tet}*-FLAG-dinB* containing the empty vector pBAD33 was serially diluted and spotted on LB agar supplemented with chloramphenicol (20 mg/mL) and either solvent or anhydrotetracycline (20 ng/mL) to induce Pol IV expression. This figure is representative of three isolates tested twice.

distribution of the genomic GATC sequences that were digested to create the library. I then pooled all plasmid-containing colonies in LB with added glycerol, plated an aliquot, and stored the remainder at -80 °C.

Instead of screening individual colonies, I chose to use high-throughput sequencing to determine which genes within the library were enriched after selection by *dinB* overexpression. I therefore plated the library-containing overexpression strain onto plates containing either solvent or 20 ng/mL aTc; ~0.1% of plasmid-containing colonies survived in the presence of inducer. Roughly 5000 colonies were isolated from each and pooled separately in LB for plasmid isolation. Pooled plasmid DNA from the "no inducer" and "aTc" samples were then submitted to the Dana-Farber Molecular Biology Core Facility for library prep with a Nextera XT kit and sequencing with an Illumina MiSeq (150 bp, paired ends).

With assistance from the Harvard School of Public Health Bioinformatics Core, the reads were trimmed of adaptor sequences using the program cutadapt

v1.4.2. Trimmed reads were mapped to either the MG1655 genome (NC_000913.3) or the pBAD33 plasmid sequence using STAR v2.4.1.

A.3 Preliminary results

Since the *E. coli* genome inserts were not isolated prior to sequencing, a large fraction of reads mapped to the plasmid backbone (74.0% for the no inducer sample, 74.3% for the aTc sample); nevertheless, a large number of reads still mapped to the *E. coli* genome (no inducer: 372,640; aTc: 553,817). Only 2.6% of the 4,497 *E. coli* gene features lacked mapped reads for the no inducer sample; this increased to 10.6% in the presence of aTc, evidence of selection due to *dinB* overexpression. Some reads in the presence of inducer may be the result of insufficient selection pressure, or cells containing escape mutants within the *dinB* expression cassette.

To determine which genes were enriched following *dinB* overexpression, I normalized the number of mapped reads per gene by the median non-zero read counts (45 for the no inducer sample, 31 for the aTc sample), which is less biased than normalizing by the total number of reads per sample (10). All gene features were then ranked by the increase in the number of normalized reads between the aTc and no inducer samples. *E. coli* genes that were also found in the pBAD33 plasmid backbone (*araC*, *insA*, *crl, yrhA*, and a 5S ribosome gene) were removed from the list of candidates; although they were all found to be enriched in the aTc sample, this was likely due to the large number of reads from the pBAD33 backbone, and the slight increase in the fraction of pBAD33 reads for the aTc sample.

The top 50 enriched genes are listed in Table A.1. As only one sample was submitted for each condition, enrichment cannot be tested for statistical significance; rather, this list suggests several candidates for further study. One notable candidate is *recB*, a component of the *E. coli* RecBCD complex that has both helicase and nuclease activities. RecBCD processes DNA double-strand breaks, and RecBC in the absence of RecD constitutively loads RecA onto recombination intermediates (11). It is possible that increased production of RecB suppresses Pol IV overprodution lethality by either acting upstream, shifting the normal balance between TLS and recombination (12) and decreasing Pol IV access, or downstream, by promoting increased homologous recombination to repair double-strand breaks caused by unscheduled Pol IV access (7).

Another interesting candidate connected to replication and repair is *rep*, a non-essential helicase that interacts with the main DnaB helicase, increasing replisome stability (13). As Rep and Pol IV physically interact (14), and a *rep* mutant strain is sensitive to moderate overexpression of Pol IV (15), it is an attractive hypothesis that increased Rep levels stabilize the Pol III replisome against Pol IV access.

Several genes involved in nucleotide metabolism are also enriched in cells overexpressing Pol IV: *amn* (AMP nucleosidase), *rihB* (ribonucleoside hydrolase 2), *nrdE* (ribonucleosidase-diphosphate reductase 2, alpha subunit), and *gsk* (inosine/guanosine kinase). Expression of these genes may alter the nucleotide pool composition and therefore synthesis by and access of the TLS and replicative polymerases. NrdE in particular is a component of the damage-inducible

ribonucleotide reductase, NrdEF, which increases dNTP levels and makes Pol III more permissive to DNA damage (16); increasing NrdE levels could potentially decrease Pol III stalling and, therefore, Pol IV access.

Follow-up studies are required to determine if candidates are true suppressors, by subcloning cleanly defined genes into the pBAD33 vector and verifying that the plasmids suppress lethality. Western blotting against the N-terminally FLAG-tagged *dinB* during overexpression will also determine whether or not multicopy supressors affect Pol IV levels. Mechanistic studies can then follow – for example, if expression of *rep* reduces Pol IV overexpression lethality, Rep can be purified and included in a reconstitution of the full replisome to see if it reduces Pol IV access (see Chapter 3.3.3).

Rank	Gene number	Gene description	Increase +aTc: Diff. of norm. reads (fold increase)
1	b1981	shiA; shikimate transporter; K08172 MFS transporter	351.5 reads (2259.4-fold)
2	b1982	amn; K01241 AMP nucleosidase	335.0 (3015.2)
3	b1943	fliK; flagellar K02414 flagellar hook-length control protein	102.8 (38.5)
4	b3851	rrsA; K01977 16S ribosomal RNA	82.8 (5.3)
5	b4007	rrsE; K01977 16S ribosomal RNA	78.4 (5.2)
6	b3968	rrsB; K01977 16S ribosomal RNA	73.8 (4.5)
7	b1603	pntA; K00324 NAD(P) transhydrogenase subunit alpha	70.0 (31.2)
8	b1919	dcyD; K05396 D-cysteine desulfhydrase	70.0 (70.0)
9	b3756	rrsC; K01977 16S ribosomal RNA	65.7 (4.5)
10	b2059	wcaA; putative glycosyl transferase	61.6 (75.0)
11	b3278	rrsD; K01977 16S ribosomal RNA	60.5 (4.1)
12	b2591	rrsG; K01977 16S ribosomal RNA	60.5 (4.5)
13	b1918	yecS; K10009 cystine transport system permease protein	59.7 (39.5)
14	b1604	ydgH; DUF1471 family periplasmic protein	57.7 (40.0)
15	b2057	wcaC; K13684 putative colanic acid biosynthesis glycosyltransferase	53.9 (67.4)
16	b3240	aaeB; K03468 p-hydroxybenzoic acid efflux pump subunit AaeB	52.5 (181.9)
17	b2820	recB; K03582 exodeoxyribonuclease V beta subunit	52.4 (7.2)
18	b0201	rrsH; K01977 16S ribosomal RNA	52.4 (4.3)
19	b2379	alaC; K14261 alanine-synthesizing transaminase	50.8 (31.7)
20	b2162	rihB; K10213 ribosylpyrimidine nucleosidase	47.5 (21.0)
21	b1065	mdtH; K08162 MFS transporter	46.4 (59.6)
22	b0204	rrlH; K01980 23S ribosomal RNA	45.6 (5.8)
23	b0925	ldtD; murein L,D-transpeptidase	45.0 (8.5)
24	b3854	rrlA; K01980 23S ribosomal RNA	44.9 (4.7)
25	b2721	hycE; K15830 formate hydrogenlyase subunit 5	44.8 (11.3)
26	b4009	rrlE; K01980 23S ribosomal RNA	43.8 (4.2)
27	b4492	ydbA; pseudogene	43.7 (2.3)
28	b4575	yjgX; pseudogene	43.3 (7.5)
29	b4257	yjgN; DUF898 family inner membrane protein	42.7 (960.7)
30	b4065	yjcE; K03316 monovalent cation:H+ antiporter, CPA1 family	42.3 (15.1)

Table A.1. Candidate multicopy suppressors of Pol IV overexpression lethality

31	b2589	rrlG; K01980 23S ribosomal RNA	40.8 (4.3)
32	b1917	yecC; K10010 cystine transport system ATP-binding protein	40.3 (82.5)
33	b0918	kdsB; K00979 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase)	40.1 (16.0)
34	b2984	yghR; putative ATP-binding protein	39.7 (66.2)
35	b0477	gsk; K00892 inosine kinase	39.1 (35.9)
36	b0950	pqiA; K03808 paraquat-inducible protein A	39.0 (16.4)
37	b3970	rrlB; K01980 23S ribosomal RNA	38.9 (3.9)
38	b2675	nrdE; K00525 ribonucleoside-diphosphate reductase alpha chain	38.7 (5.8)
39	b1944	fliL; K02415 flagellar FliL protein	37.6 (30.2)
40	b1256	ompW; outer membrane protein W; K07275 outer membrane protein	37.4 (80.1)
41	b1521	uxaB; K00041 tagaturonate reductase	37.0 (11.2)
42	b3275	rrlD; K01980 23S ribosomal RNA	37.0 (3.9)
43	b2722	hycD; K15829 formate hydrogenlyase subunit 4	36.2 (20.6)
44	b0776	bioF; K00652 8-amino-7-oxononanoate synthase	35.4 (7.9)
45	b1014	putA; K13821 RHH-type transcriptional regulator, proline utilization regulon repressor / proline dehydrogenase / delta 1-pyrroline-5-carboxylate dehydrogenase	35.3 (3.0)
46	b3212	gltB; K00265 glutamate synthase (NADPH/NADH)	33.3 (8.7)
47	b3778	rep; K03656 ATP-dependent DNA helicase Rep	33.0 (11.2)
48	b2347	yfdC; putative inner membrane protein	32.0 (16.2)
49	b0095	ftsZ; K03531 cell division protein FtsZ	31.2 (155.8)
50	b4623	insO; pseudogene	30.5 (10.7)

Table A.1. (Continued). Candidate multicopy suppressors of Pol IV overexpression lethality

References

1. Yeeles JT, Marians KJ (2011) The *Escherichia coli* replisome is inherently DNA damage tolerant. *Science* 334(6053):235–238.

2. Godoy V, *et al*. (2007) UmuD and RecA directly modulate the mutagenic potential of the Y family DNA polymerase DinB. *Mol Cell* 28(6):1058–1070.

3. Foti JJ, Delucia AM, Joyce CM, Walker GC (2010) $UmuD_2$ inhibits a non-covalent step during DinB-mediated template slippage on homopolymeric nucleotide runs. *J Biol Chem* 285(30):23086–23095.

4. Cohen SE, Godoy VG, Walker GC (2009) Transcriptional modulator NusA interacts with translesion DNA polymerases in *Escherichia coli*. *J Bacteriol* 191(2):665–672.

5. Cohen SE, *et al*. (2010) Roles for the transcription elongation factor NusA in both DNA repair and damage tolerance pathways in *Escherichia coli*. *Proc Natl Acad Sci USA* 107(35):15517–15522.

6. Uchida K, *et al*. (2008) Overproduction of *Escherichia coli* DNA polymerase DinB (Pol IV) inhibits replication fork progression and is lethal. *Mol Microbiol* 70(3):608–622.

7. Foti JJ, Devadoss B, Winkler JA, Collins JJ, Walker GC (2012) Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* 336(6079):315–319.

8. Morgan-Kiss RM, Wadler C, Cronan JE (2002) Long-term and homogeneous regulation of the *Escherichia coli araBAD* promoter by use of a lactose transporter of relaxed specificity. *Proc Natl Acad Sci USA* 99(11):7373–7377.

9. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Res* 25(6):1203–1210.

10. Dillies MA, *et al*. (2013) A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 14(6):671–683.

11. Friedberg EC, *et al*. Recombinational repair, replication fork repair, and DNA damage tolerance. (2005) *DNA Repair and Mutagenesis*, 2nd. (ASM Press, Washington,D.C.), pp. 569-612.

12. Pages V, Mazon G, Naiman K, Philippin G, Fuchs RP (2012) Monitoring bypass of single replication-blocking lesions by damage avoidance in the *Escherichia coli* chromosome. *Nucleic Acids Res* 40(18):9036–9043.

13. Guy CP, *et al*. (2009) Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol Cell* 36(4):654–666.

14. Sladewski TE, Hetrick KM, Foster PL (2011) *Escherichia coli* Rep DNA helicase and error-prone DNA polymerase IV interact physically and functionally. *Mol Microbiol* 80(2):524–541.

15. Heltzel JM, Maul RW, Wolff DW, Sutton MD (2012) *Escherichia coli* DNA polymerase IV, but not Pol II, dynamically switches with a stalled Pol III* Replicase. *J Bacteriol* 194(14):3589–3600.

16. Gon S, Napolitano R, Rocha W, Coulon S, Fuchs RP (2011) Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci USA* 108(48):19311–19316.