Mechanisms of *E. coli* DNA polymerase III and polymerase management by UmuD during DNA replication

by David A. Murison
B.S. in Chemistry, Trinity College

to

The Faculty of
The College of Science of
Northeastern University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

March 24, 2017

Dissertation directed by
Penny J. Beuning
Associate Professor of Chemistry and Chemical Biology
Acknowledgements

I would like to thank my graduate advisor, Dr. Penny J. Beuning, for allowing me to work in her lab. I have learned nearly everything that I know about molecular biology and cell culture while under her guidance, and I thank her tremendously for her patience and support while learning many new techniques. I could not ask for a better scientific role model than Penny, and her dedication to sound-science will serve as a paragon in my future work.

To my mother and father, Susan C. Wagner and David R. Murison, I cannot put into words my gratitude and appreciation for all of the love, support, and encouragement you’ve given me as your son. Thank you for allowing me the opportunity to get the best education possible at every stage of my life. I love you both dearly, and I feel blessed to be your son.

I would like to thank my thesis committee members: Dr. David E. Budil, Dr. Mary Jo Ondrechen, and Dr. Mark C. Williams. Thank you for accepting my invitation to be on my committee, and I am grateful for the advisory and guidance each have provided during committee meetings and in review of my dissertation.

I would like to thank my collaborators: Dr. David E. Budil, Dr. Jaylene N. Ollivierre, Dr. Mark C. Williams, Mohammed Nabuan Naufer, Dr. George O’Doherty, Dr. Michael Cuccarese, and Dr. Pei Shi. Each of you have contributed to the expansion of my scientific understanding, and I thank you for sharing your time and expertise.
Thank you to all of the Beuning lab members, both past and present, for supporting me as friends, mentors, and colleagues throughout this process. I would especially like to thank Dr. Michelle C. Silva, Dr. Philip Nevin, and Nicholas DeLateur for training me when I initially joined the lab and knew almost nothing about molecular biology. Thank you to my fellow graduate students: Dr. Jason Walsh, Dr. Lisa Hawver, Dr. Ingrid Salgado, Dr. Ramya Parasuram, Dr. Mark Meunter, Nicole Antczak, Lisa Ngu, and Caitlyn Mills, Timothy Coulther, Lydia Ruffner, Caitlin Kramer, Bilyana Koleva, Hannah Stern, and Amy Hotchkiss. I wish you all the best, and hope our paths cross again in the future.

I would also like to thank the undergraduates I have worked with: Angelina Choy, Rebecca Timson, Deb Milton, and Ke Ji. Thank you for your hard work, enthusiasm, and contributions to numerous projects. I would like to thank Dr. Sunny Zhou for use of the fluorimeter that many of my projects relied on. Mark Sutton and Borja Ibarra are gratefully acknowledged for providing constructs used in the work in Chapter 4. Finally, I would like to thank the College of Science and the Department of Chemistry and Chemical Biology for providing me the opportunity to pursue my Ph.D. at Northeastern University.
DNA polymerase III (DNA pol III) is a ten-subunit complex that replicates the *Escherichia coli* genome with incredible speed and accuracy; however, DNA pol III is unable to replicate damaged DNA. Interruptions in replication fork progression caused by DNA damage invoke the SOS response that involves expression of the *umuD* gene products, along with more than 50 other genes. Full-length UmuD is expressed as a 139-amino-acid protein, which eventually cleaves its N-terminal 24 amino acids to form UmuD'. The N-terminal arms of UmuD are dynamic and contain recognition sites for multiple partner proteins. Cleavage of UmuD to UmuD' dramatically affects the function of the protein and activates UmuC for translesion synthesis (TLS) by forming DNA Polymerase V.

We constructed the truncations UmuD 8 (UmuD Δ1-7) and UmuD 18 (UmuD Δ1-17) to probe the roles of the N-terminal arms in the cellular functions of *umuD* gene products. We found that the loss of just the N-terminal seven (7) amino acids of UmuD results in changes in conformation of the N-terminal arms, as determined by electron paramagnetic resonance spectroscopy with site-directed spin labeling. UmuD 8 is cleaved as efficiently as full-length UmuD in vitro and in vivo, but expression of a plasmid-borne non-cleavable variant of UmuD 8 causes hypersensitivity to UV irradiation, which we determined is the result of a copy-number effect. UmuD 18 does not cleave to form UmuD', but confers resistance to UV radiation. We find that UmuD 8 mimics full-length UmuD in many respects, whereas UmuD 18 lacks a number of functions characteristic of UmuD.
Dimerization of the *umuD* gene products is important for a number of cellular functions. UmuD\(_2\) and UmuD'\(_2\) exist as homodimers, but their subunits can readily exchange to form UmuDD' heterodimers preferentially. Heterodimer formation is an essential step in the degradation pathway for the pro-mutagenic UmuD'. To better understand the mechanism by which UmuD subunits exchange, we measured the kinetics of exchange for a number of single-cysteine UmuD variants. By conjugating fluorescent molecules at different positions on the protein we were able to observe their effect on exchange kinetics. We found that labeling positions at the outer edge of the globular domain exhibited slow exchange kinetics, which suggests that this is a surface implicated in the subunit exchange mechanism.

DNA pol III core is the three-subunit subassembly of DNA pol III. It contains the catalytic polymerase subunit α, the 3’→5’ proofreading exonuclease ε, and a subunit of unknown function, θ. We employed optical tweezers to characterize DNA pol III core activity on a single DNA substrate. We showed that polymerization and exonucleolysis by DNA pol III core are force-dependent and occur as a series of short bursts separated by pauses. The measured force and concentration dependence of exonucleolysis initiation fits well to a two-step reaction scheme in which DNA pol III core binds bimolecularly to the primer-template junction, then converts into an exo-competent conformation. Exo initiation requires fluctuational opening of two base pairs at the primer-template junction, in agreement with temperature- and mismatch-dependent bulk biochemical assays. Taken together, our results support a model in which the pol and exo activities of DNA pol III core are effectively independent, and in which recognition of the 3’ end of the primer by either α or ε is governed by the primer stability. Thus, binding to an unstable primer is the primary mechanism for mismatch recognition during proofreading.
Table of Contents

Acknowledgements ........................................................................................................ ii

Abstract of Dissertation ............................................................................................... iv

Table of Contents .......................................................................................................... vi

Table of Figures ........................................................................................................... xi

List of Abbreviations .................................................................................................... xiv

Chapter 1: Introduction to DNA Replication in E. coli .................................................. 1

1.1 DNA polymerase III holoenzyme ......................................................................... 1

    1.1.1 DNA pol III core ......................................................................................... 2

    1.1.1a DNA pol III α Subunit ............................................................................... 3

    1.1.1b DNA pol III ε Subunit ............................................................................... 13

    1.1.1c DNA pol III θ Subunit ............................................................................... 16

    1.1.2 DNA pol III β Clamp and Clamp Loader Complex ..................................... 17

    1.1.2a The γ Complex ......................................................................................... 18

    1.1.2b The τ Subunit ........................................................................................... 21

    1.1.2c The β Clamp ............................................................................................ 24

1.2 DNA Damage Disrupts Replication by Replicative DNA pols ............................ 25

    1.2.1 The SOS response in E. coli ...................................................................... 26
1.2.2 The *umuD* gene products .......................................................................................................................... 28

1.2.2a Structure and dynamics of UmuD and UmuD’ ................................................................................................. 28

1.2.2b UmuD subunit exchange and heterodimer formation ......................................................................................... 31

1.2.2c UmuD is a polymerase manager protein ........................................................................................................... 31

1.3 Conclusions .......................................................................................................................................................... 32

1.4 References.......................................................................................................................................................... 33

Chapter 2: Altering the N-terminal arms of the polymerase manager protein UmuD modulates protein interactions ........................................................................................................................................... 51

2.1 Introduction .......................................................................................................................................................... 52

2.2 Materials and Methods ........................................................................................................................................ 54

2.2.1 *In vitro* characterization of truncated UmuD variants ........................................................................................ 56

2.2.2 UV survival and mutagenesis assays and inhibition of homologous recombination ................................. 57

2.2.3 Immunoblotting ................................................................................................................................................. 57

2.2.4 UmuD and DNA polymerase III α subunit binding by tryptophan fluorescence assay and FRET .......................................................... 57

2.3 Results ................................................................................................................................................................. 59

2.3.1 Loss of N-terminal residues changes UmuD arm characteristics ................................................................. 59

2.3.2 UmuD 8 cleaves efficiently, UmuD 18 is not cleavable ...................................................................................... 63
2.3.3 UmuD arm length attenuates affinity for binding sites on DNA pol III α subunit................. 65

2.3.4 UmuD 8 is proficient for UV-induced mutagenesis; UmuD 8 and UmuD 18 do not confer resistance to UV radiation .................................................................................................................. 67

2.4 Discussion ........................................................................................................................................................................ 74

2.5 References........................................................................................................................................................................... 79

Chapter 3: UmuD dimer exchange................................................................................................................................. 84

3.1 Introduction ....................................................................................................................................................................... 85

3.2 Materials and Methods ....................................................................................................................................................... 88

3.2.1 Strains, plasmids, and proteins ................................................................................................................................. 88

3.2.2 Protein labeling with fluorescent dyes ............................................................................................................................. 90

3.2.3 Fluorescence Resonance Energy Transfer (FRET) assay of subunit exchange ................................................. 91

3.2.4 BMH Cross-linking and RecA:ssDNA-dependent cleavage assays ................................................................. 93

3.2.5 UV survival and Immunoblotting ................................................................................................................................. 93

3.3 Results............................................................................................................................................................................... 93

3.3.1 Design of UmuD variants ........................................................................................................................................... 93

3.3.2 BMH cross-linking ....................................................................................................................................................... 97

3.3.3 UmuD variants exhibit different homodimer exchange kinetics ........................................................................... 98
3.3.4 Heterodimer Exchange: UmuDD’ heterodimer is preferred; kinetics of exchange mirror homodimer ................................................................. 100

3.3.5 UV survival: Single-cysteine UmuD variants confer resistance to UV .................................. 103

3.3.6 RecA:ssDNA Cleavage: UmuD single-cysteine variants retain cleavage activity \textit{in vivo} and \textit{in vitro} ................................................................. 103

3.4 Discussion .......................................................................................................................... 106

3.5 References .......................................................................................................................... 111

Chapter 4: Single-molecule mechanochemical characterization of \textit{E. coli} pol III core catalytic activity ............................................................................................................. 116

4.1 Introduction .......................................................................................................................... 117

4.2 Materials and Methods ...................................................................................................... 121

4.2.1 DNA pol III core expression, purification, and biochemical analysis.......................... 121

4.2.2 Single molecule DNA constructs .................................................................................. 122

4.2.3 Single molecule optical tweezers experiments .............................................................. 123

4.2.4 Single molecule data analysis ....................................................................................... 124

4.3 Results .............................................................................................................................. 124

4.3.1 DNA pol III core activity at constant force ................................................................... 124

4.3.2 Force-dependent polymerization and exonucleolysis velocities ................................. 127
4.3.3 Concentration-dependent pause times during exonucleolysis .......................................................... 133
4.3.4 Force-dependent pause times during exonucleolysis ............................................................................. 136
4.3.5 Pauses during polymerization .............................................................................................................. 137
4.3.6 Dwell times for polymerization and exonucleolysis ........................................................................... 139
4.3.7 Number of processively catalyzed base pairs during polymerization or exonucleolysis .......... 139

4.4 Discussion .............................................................................................................................................. 140
4.4.1 Force-dependent instantaneous velocity of DNA pol III core exonucleolysis ............................. 143
4.4.2 Force-dependent instantaneous velocity and pausing of DNA pol III core polymerization ........ 144
4.4.3 Two-step exonucleolysis initiation from concentration-dependent measurements ................... 144
4.4.4 Exonucleolysis initiation is determined by primer-template junction stability ............................. 145

4.5 Conclusions ........................................................................................................................................... 147

4.6 References ............................................................................................................................................ 149

Chapter 5: Conclusions and future work ....................................................................................................... 154

5.1 Do the umuD gene products modulate the DNA pol III processivity switch? ............................... 154

5.2 References ............................................................................................................................................ 158
Table of Figures

Figure 1.1 – DNA pol III α subunit domains ................................................................. 4

Table 2.1 Strains and Plasmids .................................................................................. 54

Figure 2.1 – Characterization of UmuD 8 and UmuD 18 in vitro. ................................ 58

Figure 2.2 – Continuous wave EPR spectra of UmuD variants ..................................... 61

Figure 2.3 – UmuD 8 cleaves as efficiently as wild-type UmuD; UmuD 18 does not cleave to
UmuD’ ......................................................................................................................... 64

Table 2.2. Equilibrium dissociation constants\(^1\) ........................................................ 66

Figure 2.4 – UmuD 8 S60A is able to disrupt the interaction between the α polymerase subunit
and β processivity clamp while UmuD 18 S60A cannot .................................................. 67

Figure 2.5 – UmuD 8 is proficient for UV-induced mutagenesis ..................................... 68

Figure 2.6 – Steady state expression levels results of the UmuD proteins from plasmids in strain
GW8017 ....................................................................................................................... 69

Figure 2.7 – UmuD 8 S60A confers sensitivity to UV light ............................................. 71

Figure 2.8 – Sensitivity induced by UmuD 8 S60A is alleviated by tighter control over
expression. ................................................................................................................... 72

Figure 2.9 – UV Sensitivity caused by UmuD 8 is not due to a deficient pol V interaction .... 73
Figure 2.10 – UmuD 8 and UmuD 18 do not inhibit homologous recombination to the same extent as UmuD'C. ................................................................. 74

Table 3.1 – Strains and Plasmids ................................................................. 88

Figure 3.1 – Design of UmuD mono-cysteine variants and BMH cross-linking ......................... 92

Figure 3.2 – Kinetics of UmuD homodimer subunit exchange .............................................. 96

Figure 3.3 – Kinetics of UmuDD' heterodimer formation ..................................................... 101

Figure 3.4 – In vivo characterization of mono-cysteine UmuD variants in PB103. ................. 102

Figure 3.5 – Cleavage efficiency of UmuD mono-cysteine variants ...................................... 104

Figure 4.1 – Schematic depiction of the single molecule experimental procedure using optical tweezers (not to scale) ................................................................. 120

Figure 4.2 – DNA pol III core force-extension curves and temporal trajectories ...................... 126

Figure 4.3 – Representative bimodal Gaussian distribution of DNA pol III core ................. 129

Figure 4.4 – Force dependence of the DNA pol III core instantaneous velocities .................. 131

Figure 4.5 – Concentration dependence of DNA pol III exonucleolysis ............................ 132

Figure 4.6 – Force-depends of DNA pol III core exonucleolysis ....................................... 135

Figure 4.7 – Force-dependence of DNA pol III core polymerization and processivity ........... 138

Figure 4.8 – Temperature-dependence of DNA pol III core activity .................................. 140
Figure 5.1 – UmuD proteins change FRET of DNA pol III α and τ......................................................... 155
# List of Abbreviations

~ Approximately
° Degrees
A Adenosine
A Alanine
Å Angstrom
A488 Alexa Fluor 488 dye
A647 Alexa Fluor 647 dye
Ala Alanine
Asn Asparagine
Asp Aspartic Acid
ATP Adenosine triphosphate
ATPase Enzyme that catalyzes the hydrolysis of ATP
βNT DNA pol β-like nucleotidyltransferase superfamily
BMH Bis-maleimidohexane
BSA Bovine serum albumin
C Cytidine
CaCl₂ Calcium Chloride
CBM Clamp Binding Motif
CD Circular Dichroism
CTD C-terminal Domain
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>Column volume</td>
</tr>
<tr>
<td>Cys</td>
<td>Cysteine</td>
</tr>
<tr>
<td>D</td>
<td>Aspartic Acid</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytidine triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNA pol</td>
<td>DNA polymerase</td>
</tr>
<tr>
<td>DNA pol C</td>
<td>DNA polymerase C</td>
</tr>
<tr>
<td>DNA pol I</td>
<td>DNA polymerase I</td>
</tr>
<tr>
<td>DNA pol III</td>
<td>DNA polymerase III</td>
</tr>
<tr>
<td>DNA pol III'</td>
<td>DNA pol III core + τ subunit</td>
</tr>
<tr>
<td>DNA pol III core</td>
<td>DNA polymerase III catalytic core (αεθ)</td>
</tr>
<tr>
<td>DNA pol IV</td>
<td>DNA polymerase IV</td>
</tr>
<tr>
<td>DNA pol V</td>
<td>DNA polymerase V</td>
</tr>
<tr>
<td>DNA pol β</td>
<td>DNA polymerase β</td>
</tr>
<tr>
<td>DNA pol λ</td>
<td>DNA polymerase λ</td>
</tr>
<tr>
<td>DnaB</td>
<td>E. coli helicase protein</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
</tbody>
</table>
dTTP  Deoxythymidine triphosphate

*E. coli*  *Escherichia coli*

EDTA  Ethylenediaminetetraacetic acid

EPR  Electron paramagnetic resonance

F  Phenylalanine

FRET  Fluorescence resonance energy transfer

G  Glycine

G  Guanosine

Glu  Glutamine

Gly  Glycine

h  Hour

HEPES  4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HhH  Helix-hairpin-helix

His  Histidine

I  Isoleucine

$K_a$  Equilibrium association constant

kb  Kilobase

$K_D$  Equilibrium dissociation constant

kDa  Kilodaltons

KF  Klenow Fragment (DNA pol I)

L  Leucine
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu</td>
<td>Leucine</td>
</tr>
<tr>
<td>Lys</td>
<td>Lysine</td>
</tr>
<tr>
<td>µL</td>
<td>Microliter</td>
</tr>
<tr>
<td>µM</td>
<td>Micromolar</td>
</tr>
<tr>
<td>Met</td>
<td>Methionine</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>mL</td>
<td>Milliliter</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>N</td>
<td>Asparagine</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometer</td>
</tr>
<tr>
<td>nM</td>
<td>Nanomolar</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>nt/s</td>
<td>Nucleotides per second</td>
</tr>
<tr>
<td>OB</td>
<td>Oligonucleotide/oligosaccharide binding</td>
</tr>
<tr>
<td>–OH</td>
<td>Hydroxyl group</td>
</tr>
<tr>
<td>P</td>
<td>Proline</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>PHP</td>
<td>Polymerase and histidinol phosphatase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>pM</td>
<td>Picomolar</td>
</tr>
<tr>
<td>pN</td>
<td>Piconewtons</td>
</tr>
<tr>
<td>pol</td>
<td>Polymerase</td>
</tr>
<tr>
<td>Pro</td>
<td>Proline</td>
</tr>
<tr>
<td>Q</td>
<td>Glutamine</td>
</tr>
<tr>
<td>RecA*</td>
<td>RecA:ssDNA nucleoprotein filament</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>rNTP</td>
<td>Ribonucleoside triphosphate</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>SSB</td>
<td>Single-stranded DNA binding protein</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>T</td>
<td>Thymidine</td>
</tr>
<tr>
<td>Taq</td>
<td><em>Thermus aquaticus</em></td>
</tr>
<tr>
<td>TLS</td>
<td>Trans-lesion Synthesis</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>Ts</td>
<td>Temperature Sensitive</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>W</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>-----</td>
<td>------------</td>
</tr>
<tr>
<td>Y</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction to DNA Replication in \textit{E. coli}

When cells divide, each daughter cell must possess its own copy of the chromosome, a string-like macromolecule made up of nucleic acids whose sequence holds instructions for life. In \textit{E. coli} and other prokaryotes, the chromosome is a single, circular molecule that is packaged and maintained within the cell nucleoid. DNA is copied by molecular machines called DNA polymerases, which are a specialized class of proteins that copy DNA. Under ideal conditions, replicative polymerases, such as DNA pol III, efficiently replicate DNA. However, DNA damage is a constant threat to all organisms and it interferes with the efficiency and accuracy of replicative DNA pols. The Y family of DNA pols are characterized by their ability to replicate damaged DNA, albeit in a potentially error-prone manner. Consequently, their expression and activity are tightly regulated.

In \textit{E. coli} and some other bacteria, DNA damage leads to expression of a specialized set of genes as part of the SOS response, which is a coordinated response to replication stress. The two Y family DNA pols in \textit{E. coli} are under SOS inducible regulation and may rescue the replisome if progression is blocked by DNA damage. This introduction first aims to familiarize the reader with the subunits that comprise the DNA pol III holoenzyme and later discusses the activities of the \textit{umuD} gene products during the SOS response to DNA damage.

1.1 DNA polymerase III holoenzyme

DNA polymerase III (DNA pol III) performs the bulk of DNA replication in \textit{E. coli}. To copy the over-three-million base pairs that compose the \textit{E. coli} genome, the replication process must be extremely fast, accurate, and processive. When the entire DNA pol III holoenzyme is assembled,
DNA replication can proceed at speeds reaching 1 kb per second (1) with fewer than one error in every 100,000 additions (2, 3); with proofreading and mismatch repair, error rates can decrease to $10^{-10}$ per genome (2, 4). Replication is a semi-discontinuous process in which DNA pol III is an asymmetric dimer (5-9) copying leading and lagging strands simultaneously. Many proteins are employed to carry out this process in every cell cycle. From here, the complex, multisubunit machine responsible for coordinating the highly-efficient and faithful replication process in *E. coli* will be discussed.

The DNA pol III holoenzyme consists of ten discrete subunits (1, 6, 9, 10) encoded by nine different genes that can be classified into three subassemblies: the catalytic core, the processivity clamp, and the clamp loader complex (9, 11). The catalytic core (DNA pol III core) subassembly contains three subunits: $\alpha$, $\varepsilon$ and $\theta$. The three proteins are referred to as the “core polymerase” because they were among the first proteins of the holoenzyme to be isolated (12) and cannot be isolated from one another under native conditions. Initial study of DNA pol III core was difficult due to its low abundance of approximately 40 molecules per cell (only half of which are assembled into holoenzyme) (13). Early attempts to isolate the holoenzyme required a 7400-fold enrichment from 2-3 kg of cells that yielded only 1 mg of purified polymerase (14).

1.1.1 DNA pol III core

The core complex possesses the polymerase and proofreading activities of the DNA pol III holoenzyme, but these activities are limited without accessory proteins. By itself, core replicates DNA at a rate of approximately 20 nucleotides per second (nt/s) with a processivity of 11 nucleotides per binding event (13, 15). These observations describe a specific activity similar to
that of DNA polymerase I. In fact, on singly-primed M13 ssDNA, DNA pol III core cannot replicate a full circle no matter how much polymerase is added or how long the reaction proceeds (16, 17). DNA pol III core is among the slowest polymerases known, but when associated with accessory proteins it becomes one of the fastest (1).

1.1.1a DNA pol III α Subunit

The α subunit of DNA pol III possesses the polymerase activity of DNA pol III core and the holoenzyme (13). It is the largest single protein in DNA pol III consisting of 1160 amino acids (130 kDa) whose sequence is encoded by the dnaE gene. Like many polymerases, the tertiary structure of α resembles a right hand with distinctive palm, thumb and fingers domains (Fig. 1.1) (18-20). Conservation of specific architecture and acidic side chains in the palm domain allow for the classification of all six polymerase families into two superfamilies (21). Members of the A, B, and Y families are defined by palm domains with a four-stranded antiparallel β-sheet (22-25) while C and X families have palm domains formed by a five-stranded mixed β-sheet (26-28) that belongs to the DNA pol β-like nucleotidyltransferase (βNT) superfamily (29). The α subunit of DNA pol III belongs to the C family of DNA polymerases (30, 31); members of this family are only found in bacteria.
Figure 1.1 – DNA pol III α subunit domains
The crystal structure of DNA pol III α (26) with characteristic Palm (red), Thumb (blue), Fingers (orange), and PHP (yellow) domains indicated.

There has been debate about the relationship between eukaryotic DNA pol β and the DNA pol III α subunit (21, 26). The implication of defining similarity between DNA pol β and DNA pol III α is that if their palm domains are homologous then they likely bind DNA and the incoming nucleotide in the same way (21). Structural comparison of the DNA pol III α and DNA pol β palm domains using the DALI database (32) returns DNA pol β as the closest match for DNA pol III α. However, when the first crystal structure of *E. coli* DNA pol III α was solved (26), Lamers, *et al.* concluded that α could not bind DNA in a DNA pol β-like manner; superposition of the ternary structure of DNA-bound DNA pol β (33) onto the structure of DNA pol III α yielded clashes between an α-helix adjacent to the palm and thumb domains and the modeled DNA substrate (26). In contrast, Wing, *et al.* determined that the palm domains of DNA pol III α
and DNA pol β are homologous based on the ternary structure of another C family polymerase, DNA pol III α from *Thermus aquaticus* (*T. aquaticus*) (21). The debate will likely continue until a ternary structure of DNA-bound *E. coli* DNA pol III α is solved at sufficient resolution.

The palm domain of DNA pol III α (residues 271-432 and 511-560) harbors three highly-conserved aspartic acid residues essential for catalysis: Asp401, Asp403, and Asp555 (26, 34). The carboxylates of these sidechains coordinate two Mg$^{2+}$ ions, and it has been proposed that the phosphoryl transfer reaction of all polymerases proceeds through a similar mechanism involving two divalent metal ions (20, 34, 35). In this canonical mechanism, metal ion 1 activates the 3ʹ-OH of the primer for attack of the α-phosphate on the incoming dNTP. Metal ion 2 interacts with β- and γ-phosphates of the incoming dNTP to stabilize both the transition state and pyrophosphate leaving group (20).

**DNA pol III α Fidelity and Nucleotide Selectivity**

Fidelity of DNA pols is measured by their ability to replicate a template accurately. The striking specificity exhibited by high fidelity replicative DNA pols can be attributed to two factors: nucleotide selection and proofreading. Accurate nucleotide selection by replicative DNA pols is an organism’s first line of defense against genome instability. The mechanisms employed for accurate nucleotide selection cannot be accounted for solely by the thermodynamics of base-pairing (36); structural and kinetic data indicate that conformational changes within DNA pols serve as kinetic checkpoints to select complementary bases (37, 38). Johnson, *et al.* describe this selection mechanism as “induced fit” in which correct Watson-Crick base pairing between incoming dNTP and the templating base places the 3ʹ-hydroxyl of the primer terminus, α-
phosphate of the incoming nucleotide, and catalytic metal ions in an orientation advantageous for phosphodiester bond formation (39).

Nucleotide selection is the rate-limiting step of the polymerization reaction (39). When the incorrect nucleotide is paired with the templating base in the active site, it does not induce a conformational change and therefore is released quickly. Conversely, binding of the correct nucleotide induces a conformational change that is slow relative to the rate of the chemical reaction. Structural studies have shown that nucleotide binding and discrimination dynamics occur within the palm and fingers domains of replicative DNA pols (40-44).

As previously mentioned, the *E. coli* DNA pol III holoenzyme is an extremely accurate, high fidelity polymerase that makes less than one error for every $10^5$ additions before proofreading (1, 3). Interestingly, the error rate for DNA pol III α subunit is similar with small variation depending on the identity of the incorrect base pair (3). This result indicates that the nucleotide selection of DNA pol III holoenzyme occurs within the polymerase subunit α and is not affected by auxiliary proteins. Unfortunately, the nucleotide selection mechanism of DNA pol III α has remained elusive to date as neither a binary nor ternary structure of isolated DNA pol III α has been solved at sufficient resolution. Structurally, dNTPs and rNTPs are similar: they share four nitrogenous bases, except for the presence of U rather than T in rNTPs, and their sugar structures only differ by a single hydroxyl (–OH) that is present at the 2’ position in rNTPs. DNA pols select correct Watson-Crick base pairs by conformational changes before the chemistry step, but addition of a single –OH is undetected by this mechanism. Sugar discrimination in most pols
occurs via steric clash between the sidechain or peptide backbone of an active site residue and the 2’-OH group on the ribose of an incoming rNTP (45).

High fidelity DNA pols routinely discriminate against incorporation of rNTPs even though the intracellular concentration of rNTPs exceeds those of dNTPs by 10-2000-fold (46-48). Therefore, it is imperative that DNA pols have a mechanism for discrimination against rNTPs. DNA pols generally use a specific “steric gate” residue or structural feature attributed to sugar discrimination that has not been identified in E. coli DNA pol III α. Further, a canonical gating residue has not yet been identified for the C-family polymerases to date. The crystal structure of C-family member Geobacillus kaustophilus DNA pol C has been solved in ternary complex with DNA and dGTP (49). The authors propose that ribose misincorporation is prevented by His1275 (49), but current literature does not provide biochemical characterization of this prediction. Crystallization and structural studies of the C family polymerases is hampered by their large size and conformational flexibility.

Incorporation of rNTPs by DNA pols can lead to replication stress and genomic instability (50); therefore, DNA pols have a method for selection of the correct sugar during DNA synthesis. High and low fidelity DNA pols differ in dNTP specificity and misincorporation rate (36), yet most are able to discriminate efficiently against rNTPs (45, 51). Structural features of DNA pols from many families have been identified and the identity of the steric gate amino acid is conserved within a given family. The steric gate for A family polymerases is Glu while B, X, Y, and RT families use Tyr or Phe (45). As a general rule, bulkier steric gate residues decrease the tolerance for rNTPs in the active site. For example, X family members DNA pol β and DNA pol
λ utilize a segment of peptide backbone to exclude rNTPs (52-54) and are known to incorporate rNTPs at higher rates. While the active site architecture of E. coli DNA pol III α has been shown to resemble that of X family polymerases (21, 26), specifically human DNA pol β, it is likely that the C family polymerases utilize a different sugar discrimination mechanism based on their low ribose misincorporation frequencies (1, 3).

**DNA pol III α Polymerase and Histidinol Phosphatase Domain**

N-terminal residues 1-280 of DNA pol III α form the polymerase and histidinol phosphatase (PHP) domain. The PHP domain has a TIM barrel-like fold comprising a seven-stranded β-barrel encircled by seven α-helices (26). PHP domains are found N-terminal to the palm domain of all C family polymerases (2, 55), yet play different enzymatic or structural roles. For example, DNA pol III α from *Thermus thermophilus* (*T. thermophilus*) and *Thermus aquaticus* (*T. aquaticus*, *Taq*) both contain PHP domains that possess Zn$^{2+}$-dependent exonuclease activity (21, 56), whereas the PHP domains from *E. coli* DNA pol III α and its ortholog DNA pol C from *G. kaustophilus* lack exonuclease activity (26, 49). Many conserved metal-binding sidechains are not present in the *E. coli* DNA pol III α PHP domain (55) which is consistent with its lack of exonuclease activity. Indeed, restoring metal-binding sidechains at these conserved positions was able to confer metal binding, but did not result in detectable exonuclease activity (57).

Many have speculated that the DNA pol III α PHP domain possess pyrophosphatase activity that would remove the byproduct of DNA synthesis and drive the polymerization reaction forward (26, 55). When the crystal structure of DNA pol III α was first solved, Lamers, *et al.* noted that there was a phosphate ion present in the putative PHP domain active site (26). The presence of a
shallow channel running from the palm domain to the PHP domain was noted. Recent work has shown that the DNA pol III α PHP domain possesses pyrophosphatase activity (58). Lapenta et al. identified three residues that participate in the hydrolysis reaction (H12, D19 and D210), and showed that the rate of DNA synthesis is linked to pyrophosphate hydrolysis.

The PHP domain of DNA pol III α also plays a non-enzymatic role in the context of the holoenzyme. The ε subunit of DNA pol III, a small 27-kDa protein that possesses the 3’→5’ proofreading exonuclease activity (59), tethers itself to the polymerase by wrapping its C-terminal linker sequence around the PHP domain of DNA pol III α (60). Therefore, the PHP domain provides a scaffold for association of the DNA pol III polymerase and exonuclease subunits. Interestingly, it has been noted that loss of PHP domain metal-coordinating residues within a C family polymerase coincides with the presence of an auxiliary proofreading exonuclease protein homologous to DNA pol III ε subunit (57).

**DNA pol III α Fingers Domain**

*E. coli* DNA pol III α has an unusually large fingers domain (residues 561-911) relative to other polymerases (26). It has been hypothesized that this structural feature shared among C family members may be the source of their increased catalytic rate (61). The fingers domain is further characterized by finger subdomains: Index (residues 641-756), middle (residues 561-640), ring (residues 779-838), and little finger (839-911) (26). As illustrated by a binary Cryo-EM structure of DNA-bound DNA pol III core assembly (αεθ) in complex with the β processivity clamp (62), the fingers domain of DNA pol III α wraps around double-stranded DNA. This structure also
shows that the majority of contacts between fingers domain and DNA duplex are via the peptide backbone and the positive charge of α-helix dipoles.

DNA pol III α binds dsDNA using tandem helix-hairpin-helix (HhH) motifs located at the C-terminus of the fingers domain (residues 835-912) (63). The presence of HhH motifs in DNA pol III α was initially predicted by sequence analysis (64) and later confirmed by solving the first crystal structure (26). The tandem, consecutive HhH motifs are considered part of a (HhH)_2 domain (65) capable of binding dsDNA in a symmetric, non-sequence-specific fashion (64, 65).

Biochemical and biophysical experiments have demonstrated the weak affinity of DNA pol III α for duplex DNA (62, 63). Recent structural investigations show that DNA pol III α does indeed make weak electrostatic contact with the DNA duplex which is consistent with the low processivity of the isolated DNA pol III α subunit (3, 13, 62). To compensate for its weak binding to DNA, it is thought that DNA pol III α has evolved a larger fingers domain to interact with a larger segment of DNA (62). Strong interactions between polymerase and DNA would likely slow synthesis by creating drag.

The high processivity of DNA pol III is derived from interaction of DNA pol III core complex with the β clamp. Biochemical and structural studies show that DNA pol III α subunit interacts with the β clamp at two α sites: an internal site at the end of the little finger (residues 920-924) and a site at the extreme C-terminus (residues 1154-1159) (62, 66). Only the internal site is required for processive DNA synthesis in vivo and in vitro (66).
**DNA pol III α Thumb Domain**

The polymerase thumb is one of the three characteristic domains that give DNA pols their distinctive right-handed structure (18-20). The canonical function of the polymerase thumb is to guide emerging DNA duplex as it exits the active site. Structural analysis of apo and DNA-bound *Taq* polymerases show that the thumb domain rotates inward upon substrate binding to interact with the minor groove of DNA duplex (20, 21, 67). The thumb domain in DNA pol III α (residues 433-510) is located on the side of the palm opposite the fingers domain and adjacent to the PHP domain (26).

The thumb domain structure is not conserved across polymerase families (20). It has been noted that the thumb domains of C family polymerases make a more invasive contact with the minor groove of nascent DNA compared with other, lower-fidelity polymerases (21, 49, 62). For example, the thumb domains of *E. coli* polymerases I, II, and IV each make hydrogen bonding and van der Waals contacts with the phosphodiester backbone of DNA, but do so only via the peptide backbone (62).

A recent structure of DNA-bound DNA pol III α, ε, and β clamp in synthesis mode shows that the ε subunit is wedged between the thumb domain of α and the β clamp (62). Further, the structure shows that a loop from the thumb domain (residues 464-470) makes the polymerase’s only non-backbone contact with DNA. A follow-up structure of the same DNA-bound DNA pol III complex, including the θ subunit, was captured in editing mode (68). Here, the thumb domain is seen acting as a wedge between primer and template strands of DNA. The thumb appears to redirect a mismatch at the 3’ primer terminus toward the active site of ε which is sandwiched
between the thumb and β clamp. The aromatic sidechain of residue Tyr453 appears to stack with the last base pair of the primer terminus which likely helps stabilize the mismatch. Interestingly, aromatic residues are conserved at this position in replicative C family polymerases (68).

While this is the first time a thumb domain has been shown to participate in the fidelity mechanism of a replicative polymerase, the concept has been hypothesized for quite some time. A structure of DNA-bound DNA pol C from G. kaustophilus, a fellow C family member, had previously led to the prediction that the thumb may sense and redirect mispairs to the exonuclease site (49). The role of the polymerase thumb in sensing mismatches is unique to C family polymerases and may be key to maintaining faithful replication.

**DNA pol III α C-terminal Domain**

Residues 917-1160 form the C-terminal domain (CTD) of DNA pol III α. This segment of the protein is flexible and disordered, but gains structure upon interaction with other DNA pol III subunits (62, 68). This domain is often referred to as the regulatory hub of the protein and contains both β binding motifs, a τ subunit interaction site (69), the UmuD binding site (70, 71), and an oligonucleotide binding (OB-fold) domain (72) responsible for ssDNA binding (63).

The OB-fold domain (residues 960-1071) is the second site of DNA binding in DNA pol III α which was predicted based on sequence and is absent from the E. coli crystal structure. The C-terminal tail region of DNA pol III α where the OB-fold is located has been shown to undergo significant conformational change when interacting with the β clamp and further upon DNA binding (62, 68, 73). While biochemical experiments have indicated there is an interaction between the OB-fold and ssDNA (63, 73), which has implications for the polymerase during
lagging strand synthesis and polymerase recycling, structural investigations have yet to show binding (62, 68).

OB domains bind a diverse array of substrates including oligonucleotides, oligosaccharides, or metal ions (74, 75). Further, OB domains are also known to mediate protein-protein interactions as has been shown for DNA pol III α (residues 956-975) and the polymerase manager protein, UmuD (71). UmuD inhibits DNA pol III α binding with ssDNA which is hypothesized to play a role in the DNA damage tolerance mechanism and polymerase switching.

The τ subunit of DNA pol III also interacts with the CTD of DNA pol III α. The affinity of the τ subunit for DNA pol III α is strong with equilibrium dissociation constants in the picomolar range (76, 77). Further, the site of interaction was localized to the extreme C-terminus of DNA pol III α between residues 1112-1160 (77). The τ subunit is part of the clamp loader complex that is responsible for assembling the β clamp on DNA (76, 78, 79). While the presence and function of τ is essential to *E. coli* (80), τ is not required for clamp loading. Its most vital role is in coordinating leading and lagging strand DNA synthesis of two DNA pol III core assemblies via the aforementioned interaction with DNA pol III α (11, 81). The τ subunit mediates the affinity of DNA pol III α for the β clamp upon sensing dsDNA as part of the “processivity switch” mechanism (82). This mechanism will be discussed in more depth to follow and highlights the importance of the α-τ interaction.

1.1.1b DNA pol III ε Subunit

The ε subunit of DNA pol III is a 27-kDa protein that contains the 3′→5′ exonuclease activity integral to the proofreading mechanism of DNA pol III core (2, 83). The ε protein is encoded by
the *dnaQ* gene (also known as *mutD*) which was first identified as an editing exonuclease (59) by investigation of mutator alleles *dnaQ49* (84) and *mutD5* (85). Many replicative polymerases, such as fellow C family members DNA pol C from *G. kaustophilus* and DNA pol III α from *T. aquaticus*, have polymerase and exonuclease domains within the same polypeptide (27, 49). However, DNA pol III from *E. coli* represents a large portion of DnaE1-type polymerases that utilize a separate exonuclease protein attached to the N-terminus of the polymerase subunit by a flexible peptide linker (31). While the stimulus driving this divergence in polymerase architecture is not well understood, the presence of an exonuclease auxiliary to the polymerase protein seems advantageous from a regulatory perspective. For example, the processivity of the holoenzyme is reduced from >50 kb to 1.5 kb in the absence of ε (86) which all but ensures that the exonuclease is present during replication.

In the context of DNA pol III core, ε is flanked by α and θ subunits, and together the three form a tight 1:1:1 complex (87, 88). The N-terminal portion of ε associates with θ and contains the exonuclease active site within its first 185 residues (89). The C-terminal segment of ε is responsible for binding DNA pol III α (90, 91) via a 57-residue portion of the ε C-terminus termed “the Q-linker.” The Q-linker contains a flexible 17-residue linker sequence (92) and a 40-residue segment that wraps around the PHP domain of α (90). Subunits α and ε increase the activity of each other (83, 93) and association of ε and θ has a slight stimulatory effect on the exonuclease activity of ε (89). The θ subunit makes contact with ε near the exonuclease active site and is thought to stabilize ε (94), thereby modulating the polymerase and exonuclease activities of DNA pol III core (95).
The natural substrate for 3’→5’ exonuclease activity of ε subunit is ssDNA at least three nucleotides in length (96). Kinetic analysis of the exonucleolysis reaction by ε shows that degradation of DNA occurs through a two-step mechanism in which product release is not rate-limiting (96). Rather, DNA melting or fraying at the 3’ terminus of the primer strand is the rate-limiting step in the exonucleolytic DNA editing reaction followed by product release.

The active site architecture of the ε subunit includes two divalent metals (Mn$^{2+}$ or Mg$^{2+}$) coordinated by the carboxylate side chains of Asp12, Glu14, and Asp167 (97). Structural data (98, 99) suggest that one metal ion is pentacoordinated, a feature that is supported by in silico investigation of the ε reaction mechanism (100). This configuration facilitates proton transfer to His162 from a metal-coordinating water, followed by nucleophilic attack on the phosphate backbone.

Three non-enzymatic roles have been established for the ε subunit at the replication fork (101, 102). First, the ε subunit has been shown to bind the β clamp which stimulates the rate and processivity of DNA synthesis by DNA pol III (86, 88). This additional point of contact between polymerase and clamp stabilizes the association of DNA pol III core and promotes polymerization mode of the complex during replication (103). Although the interaction between ε and β is weak (103), its transient nature is well-suited for periodic disruption such as is necessary for polymerase recycling during lagging-strand synthesis.

The exonuclease binds the clamp using a canonical clamp binding motif (Qxx(L/M)xF, residues $^{182}$QTSMAF$^{187}$) located near its catalytic domain (101). The processivity clamp is a homodimer ($^{\beta2}$), and each crescent-shaped monomer possesses a binding pocket near the dimer interface. As
previously noted, DNA pol III α is known to occupy the binding pocket of one β monomer using either its internal or C-terminal clamp binding site (66, 104, 105). Therefore, ε associates with the corresponding binding pocket of the second β monomer. Consequently, the exonuclease-clamp interaction has a second, allosteric role: To prevent error-prone polymerases, such as DNA pol IV (DinB), from accessing the replication fork during normal replication (101).

The third non-enzymatic function of ε is in the strand-displacement activity of DNA pol III holoenzyme (102). In fact, this activity does not require the exonuclease domain (residues 1-180) of ε at all, but does depend on the clamp binding motif ε C-terminal linker (residues 180-243) that binds the PHP domain of DNA pol III α (103). The strand-displacement activity is also contingent on other DNA pol III holoenzyme cofactors, and the function of ε as a scaffold to join the N-terminus of DNA pol III α and the clamp is similar to the role the τ subunit plays in coordinating the matrix of the C-terminal domain of DNA pol III α, the β clamp, SSB, and the χψ subunits of DNA pol III.

1.1.1c DNA pol III θ Subunit

The θ subunit of DNA pol III is the 8.8-kDa product of the hole gene (2). DNA pol III θ is the third protein comprising DNA pol III core in which it binds the N-terminal region of ε near the exonuclease active site (89, 106). NMR studies have shown that the θ fold consists of three α-helices connected by segments of flexible polypeptide chain (107, 108). Like DNA pol III ε, θ is intrinsically unstable in isolation, but formation of the ε:θ complex enhances stability of both proteins (106, 109). The contact between ε and θ is stabilized by extensive hydrophobic interactions (108, 110). Exposure of this surface to solvent likely contributes to destabilization of
the free subunits. The θ subunit enhances the 3′→5′ exonuclease activity of DNA pol III ε two-to fourfold (89), and deletion of θ slightly increases spontaneous mutation frequency (89, 97). A series of two-hybrid experiments have shown that the interaction between DNA pol III α and ε is enhanced by θ (94), but θ does not affect DNA synthesis by α or α:ε (86).

The θ subunit does not bind DNA and has not been assigned an enzymatic function. In fact, deletion of holE from the chromosome does not affect cell viability or morphology (111). While the data suggest that DNA pol III θ is not essential, the θ subunit is found in other gram negative bacteria that contain an ε-like proofreading exonuclease that is not part of the same polypeptide as the polymerase (94). Proposed non-enzymatic functions of DNA pol III θ include protection of ε from degradation (94) and serving as a checkpoint that modulates DNA pol III core activity during holoenzyme assembly.

1.1.2 DNA pol III β Clamp and Clamp Loader Complex

The DNA pol III holoenzyme relies on the β clamp for processive DNA synthesis. DNA pol III β2 is a homodimeric protein comprising two crescent-shaped subunits that together form a ring-shaped quaternary structure. Once loaded, the β subunit encircles dsDNA and slides freely along the double-helix. The β subunit promotes processive DNA synthesis by anchoring DNA pols to the dsDNA substrate.

The clamp loader complex comprises DNA pol III subunits τ, γ, δ, δ′, χ, and ψ (1, 9, 10). As the name implies, the clamp loader complex is responsible for loading DNA pol III β onto dsDNA. The stoichiometry of clamp loader subunits is variable for clamp loading activity, but during chromosomal replication the clamp loader complex is also responsible for coordination of
leading and lagging strand DNA synthesis, which requires at least two τ subunits (5, 6, 11). The γ subunit, which is expressed from the same gene as τ, binds the ATP required for clamp loading (10, 112). During the loading process, δ and δ' subunits directly interact with the β clamp with δ functioning as a “wrench” and δ' modulating the contact between δ and β (112-114).

1.1.2a The γ Complex

The clamp loader complex is powered by the ATPase activity of τ and γ subunits. Both τ and γ are expressed from the dnaX gene where τ is the full-length product (643 residues) and γ is produced by a -1 frameshift during translation that creates a premature stop codon (430 residues) (115-117). In fact, the frameshift producing γ is so efficient that the τ/γ ratio in vivo is close to 1 (115-118). The structure of τ is composed of five domains, two of which are unique to τ. N-terminal Domains I, II and III are shared by τ and γ subunits. The nucleotide binding site and ATPase activity are located in Domains I and II while Domain III is responsible for associating with other clamp loader subunits to form a circular collar upon assembly of the γ complex. The ATP binding sites are located at the interface of Domains I and II within a single subunit, and at the interface between two adjacent subunits within the complex (119). Crystal structures of γ subunit Domains I and II (residues 1-243) with and without ATPγS shows that nucleotide binding induces a large conformational change, which is integral to the β clamp opening process (119).

The stoichiometry of τ and γ subunits within the clamp loader can vary while maintaining clamp loading function (120-122). The “minimal” version of the clamp loader is the γ complex: a heptameric assembly comprising three γ subunits, a δ and δ' in the order δ':γ1:γ2:γ3:δ. The γ
complex is the simplest assembly that retains clamp loading function (123). Crystal structures of nucleotide-free (124) and ATPγS-bound (125) versions of the γ complex have been solved. Comparison of the open (ATP-bound) and closed (nucleotide-free) γ complex structures shows little difference in conformation. This may be due to the fact that only two of the three ATP binding sites are occupied in the open structure (125). Conformational change leading to ATP hydrolysis and clamp opening may require ATP-binding at all three sites. Other possible factors that may contribute to ATP-binding include: (1) interaction with the β clamp stimulates ATP-binding at the third site or (2) χ and ψ subunits, which are not present in the minimal γ complex, modulate ATP affinity through an interaction with Domain III of γ which forms the collar (126-128).

The closed, nucleotide-free clamp loader has low affinity for β, but upon ATP binding, the complex gains affinity for the clamp (129). The increased affinity is likely due to a conformational change that restores accessibility of the δ subunit, which is sequestered in the closed conformation (124, 130). Clamp binding by the δ subunit induces opening of one β2 interface by which the clamp and loader complex can be threaded onto DNA (131). The clamp-clamp loader complex has strong affinity for DNA, specifically a primer-template junction (78, 130), and threading of DNA through the assembly stimulates hydrolysis of ATP by the γ and/or τ motor proteins (132). ATP hydrolysis induces a clamp loader conformational change that releases it from β and, in doing so, closes the β2 ring around DNA (113, 132).

The δ subunit has a high affinity for the β clamp and, upon binding, triggers a spring-like conformational change at one interface of β2 (114). For this reason, δ is considered the “wrench”
of the clamp loader complex. The δ subunit is the 35-kDa product of holA and its structure comprises three flexible domains that form a C-shape similar to γ. Domains I and II of δ adopt the same fold associated with AAA+ family of ATPases (133, 134), but δ does not contain an ATP binding site and its sequence shares little homology with other AAA+ family members (135). The clamp opening activity of δ is not limited to its activity within the clamp loader complex. The δ subunit alone has been shown to stimulate dissociation of the clamp from circular DNA (113, 136). The interaction is illustrated by a crystal structure of the N-terminal 140 residues of δ bound to a β variant that cannot dimerize (114). The hydrophobic tip of δ’s N-terminus (Domain I) binds β at a hydrophobic pocket that is also utilized by other replication machinery including DNA pols I – V (114, 137-145). Interestingly, δ binds the monomeric form of β 50-times more strongly than the β2 dimer and crystallization of δ with wild-type β was unsuccessful (114). The δ-β interaction induces a conformational change within the β monomer which results in a steric clash at the β2 dimer interface. The open conformation of the clamp presents a 15 Å gap large enough for passage of ssDNA into the clamp pore (112, 114).

The δ’ subunit is the 37-kDa product of the holB gene and has been termed the “stator” owing to its structural rigidity relative to other clamp loader subunits (146) and its role as the fixed point around which other subunits move within the γ complex (10). The δ and δ’ subunits were originally thought to be genetically related (147), but instead are encoded by separate genes (148). The crystal structure of isolated δ’ subunit reveals the structural similarities it shares with other clamp loader subunits: three domains arranged in a C-shape architecture (146). Therefore, it is not surprising that the δ’ subunit is also a member of the AAA+ family. While the amino
acid sequence of δ’ shows homology to fellow AAA+ family members γ and τ (148-150), it does not contain a competent nucleotide binding domain or ATPase activity.

In the context of the γ complex, δ’ caps one flank of the three γ subunits opposite that of δ. The δ and δ’ subunits are in close proximity when part of the γ complex assembly, but only their C-terminal domains (Domain III) interact to form the collar (124). The gap between N-terminal Domains I and II of δ and δ’ does not change appreciably in the clamp-clamp loader complex (151) and δ’ has been shown to modulate binding of δ to the clamp (113).

1.1.2b The τ Subunit

The τ and γ subunits of DNA pol III are closely related: [1] they are both expressed from the dnaX gene; [2] they share the same N-terminal 430 residues, three-domain architecture, and crescent shape; and [3] they are both ATPases capable of powering the clamp loader complex. However, only the τ subunit is essential to cell viability (80). The C-terminal 213 residues that are unique to τ are assembled into two domains that coordinate replication machinery and facilitate processive DNA synthesis.

The isolation of DNA pol III’ (DNA pol III core + τ) from cell extracts (152) led to the observation that two DNA pol III core assemblies were coupled by two τ subunits. The τ subunit binds the C-terminus of DNA pol III α (residues 1113-1160) (77) via Domain V of its own C-terminus (residues 622-643) (69) with an equilibrium dissociation constant $K_d$ of 4 nM (76). The τ-α interaction is implicated in multiple replication process that include: [1] coordination of leading and lagging strand polymerases during replication (153, 154); [2] participation in the
processivity switch and polymerase recycling during lagging strand synthesis (73, 82); and [3] prevent premature removal from the β clamp (155).

Because DNA replication is unidirectional 5’→3’, the complementary strands of template DNA must be copied in opposite directions (2). The leading strand is replicated by nearly continuous synthesis while the lagging strand is discontinuously replicated in discrete Okazaki fragments (5). Therefore, the DNA pol III holoenzyme particle is an asymmetric dimer with leading and lagging strand polymerases. The τ subunit coordinates leading and lagging strand polymerases as part of the clamp loader assembly.

At least two τ subunits are required for coupling the two DNA pol III cores as part of the clamp loader complex (9). The clamp loader complex requires three copies of dnaX protein (156), and four different stoichiometries are possible: δγ3χψδ’, δγ2τχψδ’, δγτ2χψδ’, and δτ3χψδ’. Each of these reconstituted assemblies is able to load the β clamp onto DNA and does so at a similar rate (116, 117, 120). Because at least two polymerases are required for concurrent duplication of leading and lagging strands, δγ3χψδ’ (γ complex) and δγτ2χψδ’ assemblies are not relevant here. There has been some debate about the significance of a triple polymerase holoenzyme coordinated by a clamp loader with three τ subunits (δτ3χψδ’) (120). The “triple-polymerase” model proposes that two DNA pol III cores share the responsibility of synthesizing Okazaki fragments on the lagging strand. However, significant data supports a model where the DNA pol III holoenzyme contains two τ and one γ (152, 156-158). Interestingly, a strain with mutant dnaX expressing only τ displayed sensitivity to UV exposure and display impaired DNA pol IV-dependent mutagenesis associated with double-strand break repair (158).
Discontinuous DNA synthesis on the lagging strand requires polymerase recycling upon completion of each Okazaki fragment. The cycle is known as the “processivity switch” and requires τ (9, 73, 82). The “processivity switch” refers to a mechanism that triggers dissociation of the polymerase from its clamp in response to specific DNA structures (82). Early work using an in vitro rolling circle replication reaction indicated that lagging strand synthesis was dependent on a sufficient concentration of τ, and replicase assembled with γ complex is not able to recycle after synthesizing an Okazaki fragment (9, 154). Activation of the processivity switch requires completion of an Okazaki fragment by DNA pol III α such that only a nick remains at the 3’ terminus (73). The nick is first sensed by the OB domain of α which lowers the affinity of polymerase for the β clamp in response (21). The τ subunit has also been shown to bind ssDNA (82, 105) and, upon sensing the nick, τ induces a conformational change in the polymerase that disengages α from the clamp and increases its affinity for primed DNA (73, 82).

The τ subunit promotes DNA pol III processivity by preventing premature removal of the β clamp by the γ complex (155). The length of synthesized DNA produced by DNA pol III is directly proportional to τ concentration and inversely proportional to the concentration of γ complex. The γ complex is capable of binding an active replication complex and displacing β from DNA (113), but the presence of τ inhibits clamp unloading thereby promoting processive DNA synthesis (155).

The rate of synthesis by DNA pol III is regulated by the affinity of τ for DnaB (153, 159). DnaB leads the replication fork by unwinding double-stranded template DNA while encircling only the lagging strand (8). The C-terminal Domain IV of τ interacts directly with DnaB and increases its
rate of unwinding to at least 400 nt/s (76, 153). Without τ coupling DnaB to the replisome, the helicase slows to approximately 35 nt/s (153). This difference in DnaB activity suggests a mechanism in which the rate of ssDNA generated by DnaB is moderated by replication fork progression and coordination by τ.

### 1.1.2c The β Clamp

The β subunit is expressed from *dnaN* as a 41-kDa crescent-shaped monomer that preferentially forms ring-shaped homodimers in solution. Each β monomer comprises three domains that are structurally similar but share little sequence homology (160). Domains are connected by peptide loops along the outer rim, and the inner pore is lined with six α-helices per monomer. Bacterial processivity clamps, such as *E. coli* DNA pol III β₂, are homodimers (160, 161). Assembly of the two protomers into a ring shape occurs head-to-toe through non-covalent interactions at two identical interfaces. The inner pore of β₂ is approximately 35 Å in diameter, which permits passage of duplex DNA through the center. The clamp slides freely along dsDNA due to a layer or two of water molecules mediating interaction between the clamp and phosphate backbone (160).

DNA pol III β₂ is the primary contributor to processivity of the holoenzyme (2). The β clamp provides a scaffold for polymerases to maintain close contact with DNA during replication. Once loaded, the clamp remains stably bound to DNA with or without polymerase present with a half-life on the order of 2 h (162-164). DNA pols and other clamp interaction partners primarily bind β using conserved peptide motif, referred to as the clamp binding motif (CBM) (104). Processivity clamps have one protein-binding site per monomer which is located on the C-
terminal face of DNA pol III β (66, 124, 165). DNA pol III core interacts with both protein-binding sites on β₂ via one of two sites at the C-terminus of the α polymerase subunit (66, 104, 166) and a second CBM in the exonuclease subunit ε (103). Occupation of both protein-binding sites on β₂ has implications for different conformational states of DNA pol III during replication (62, 68).

### 1.2 DNA Damage Disrupts Replication by Replicative DNA Polys

DNA pol III is well-suited to perform the bulk of DNA replication in an efficient and faithful manner. However, the holoenzyme is unable to replicate efficiently noncanonical DNA structures that arise from DNA damage. DNA damage is a constant threat to all organisms as it can arise from exogenous sources, such as UV light, and endogenous sources, such as reactive oxygen species created during metabolism (167). For example, it is estimated that a single day in the sun can cause up to 100,000 UV photoproducts in a single mammalian skin cell (168). DNA damage caused by UV light generally affects pyrimidine bases and manifests in two different structures: cyclobutane thymine dimer (CPD) or (6-4) photoproduct. These lesions are covalent linkages between adjacent pyrimidine bases that cause prominent distortion in the helical structure of DNA (167).

The consequence of the replisome encountering DNA damage in the template depends on whether the damage is encountered by the leading-strand or lagging-strand polymerase (169, 170). Replication on the lagging-strand is discontinuous by nature. Therefore, when the lagging-strand DNA pol encounters a block in the template it is able to reassemble at a downstream primer and continue replicating while the leading-strand DNA pol is unaffected (170-173).
However, when the leading-strand DNA pol encounters a block it must be dealt with before leading-strand replication can continue. Meanwhile the lagging-strand DNA pol is unaffected and continues replicating ahead of the leading-strand (170). The rate of replication progress is decreased in both scenarios, but blockage on the leading strand appears more disruptive.

1.2.1 The SOS response in E. coli

DNA damage and other cellular stresses may induce the SOS response in E. coli (167). Interruptions in replication as a result of DNA pol III stalling at unrepaired DNA damage lead to an accumulation of ssDNA. RecA protein polymerizes on ssDNA in an ATP-dependent process, forming a RecA:ssDNA nucleoprotein filament (RecA*) (174). RecA is a 38-kDa protein comprising 352 amino acids encoded by the recA gene. One RecA monomer has a footprint of three nucleotides per monomer, and nucleation of five monomers is required for filamentation (175, 176). RecA protein is present in unstressed cells at fewer than 10,000 copies, which perform functions such as homologous recombination (177). However, upon induction of SOS, intracellular RecA increases to over 70,000 copies per cell (178).

Under normal conditions, the SOS response is suppressed by the LexA repressor protein. LexA binds the operator sequence of genes involved with SOS to prevent their expression. However, the binding affinity of LexA for a given operator depends on the nucleotide sequence (179-181). For example, LexA binds weakly to the recA operator so that expression is readily induced upon sensing stress. Conversely, other SOS genes, such as umuDC, are under tighter control with LexA binding the operator sequence with an equilibrium dissociation constant $K_d$ of 0.2 nM (182). RecA* facilitates the autocatalytic cleavage of the LexA repressor protein into two
inactive polypeptides (183). As the intracellular concentration of LexA decreases, dissociation of LexA protein from the operator regions leads to induction of SOS regulated genes. Therefore, genes with the highest affinity LexA binding are induced last.

The SOS response induces expression of at least 57 genes (184, 185) that participate in diverse functions ranging from regulation of cell division to mutagenic DNA damage tolerance (167, 184-187). Three of the five DNA pols in *E. coli* are under SOS inducible regulation (186, 188) including the Y-family polymerases DNA pol IV (*dinB*) and DNA pol V (*umuDC*). Members of the Y family of DNA polymerases are characterized by their ability to perform TLS on damaged DNA templates and non-canonical DNA structures (189, 190). As a tradeoff, Y-family DNA polymerases exhibit low fidelity on undamaged DNA and lack 3′→5′ exonucleolytic proofreading activity (186, 191, 192). Therefore, the activity of DNA pols IV and V are tightly controlled at the pre- and post-translational level.

The SOS response to DNA damage is temporally separated into early and late stages (193-195). During the early stages of SOS, the cell utilizes error-free repair pathways such as nucleotide excision repair and homologous recombination (193). The SOS-regulated genes associated with error-free DNA repair are weakly regulated by LexA. In the event that DNA damage persists, potentially-mutagenic pathways are employed; specifically, TLS by Y family DNA pol V. DNA pol V (*UmuD'C*) is capable of replicating past a variety of DNA lesions, albeit in a typically error-prone manner (145, 196, 197). Therefore, activation of DNA pol V is a cell’s last resort to promote survival through mutagenesis.
1.2.2 The *umuD* gene products

The *umuDC* operon encodes two proteins: UmuD and UmuC. Expression of *umuDC* is negatively regulated by the LexA repressor protein as part of the SOS response to DNA damage. UmuD is expressed as a 139 amino acid monomer that preferentially forms UmuD2 dimers in solution (198). Approximately 180 copies of UmuD are present in non-SOS-induced cells, but intracellular UmuD is increased to approximately 2400 copies following induction of the SOS response. The activities of the *umuD* gene products are functionally distinct and temporally separated. During the early stages of SOS, UmuD and UmuC specifically inhibit DNA replication and promote faithful DNA damage repair pathways. It has been suggested that UmuD2C acts in a primitive DNA damage checkpoint by controlling cell cycle progression in favor of maintaining genomic integrity. Full-length UmuD2 is the predominant form for approximately 20-40 minutes after initiation of SOS (193). However, if DNA damage persists, UmuD will undergo RecA-facilitated cleavage forming the mutagenically-active species, UmuD'. Together with UmuC, UmuD' forms the Y family polymerase DNA pol V. DNA pol V is capable of performing TLS across a variety of damaged DNA bases, albeit at low fidelity. Cleavage of UmuD to UmuD' marks a transition from accurate DNA repair to DNA damage tolerance.

1.2.2a Structure and dynamics of UmuD and UmuD'

UmuD is an intrinsically disordered protein (198, 199) comprising a C-terminal globular domain often referred to as the “body” and an unstructured N-terminus often referred to as the “arm.” To date, no high-resolution structure exists for full-length UmuD2, but NMR solution structures (200, 201) and an X-ray crystal structure (202) of UmuD'2 have been solved. UmuD' is formed
by the autocleavage of 24 amino acids from the N-terminal arm of UmuD. The N-termini of UmuD’ are thought to be largely disordered as the first seven residues (Gly25-Asp32) do not appear in the X-ray structure and the remainder of the arm is poorly ordered (203). Indeed, HXMS investigation indicated that the N-terminal arms of both UmuD and UmuD’ are dynamic (204). However, biochemical evidence supports a model where the longer N-terminal arms of full-length UmuD₂ may stably bind to the globular domain while the shortened N-termini of UmuD’ cannot (205, 206). Analysis of the N-terminal arm dynamics of UmuD and UmuD’ by EPR showed that a low-mobility component exists for full-length UmuD while only fast-component is detected for the N-termini of UmuD’ (207). Considered together, these data indicate that the N-termini of UmuD and its cleavage product UmuD’ exhibit different arm dynamics. Moreover, the ability of full-length UmuD N-termini to stably bind the globular domain presents a unique protein-protein interaction surface that is not present in UmuD’ and may modulate affinity for binding partners during the progression of SOS (201, 208).

The differences between the X-ray and NMR structures of UmuD’₂ are substantial. In fact, the RMSD of C-terminal peptide backbone atoms (residues 40-139) between the two structures is 4.59 Å (200). Notably, the difference in positioning of active site residues Ser60 and Lys97 within the two structures may indicate how conformation and dynamics affect cleavage. In the crystal structure, Ser60 and Lys97 sidechains are separated by 3 Å within each protomer, whereas they are over 7 Å apart in the NMR structure. Therefore, it has been suggested that the conformation of UmuD’₂ in the X-ray crystal structure mimics that of full-length UmuD₂ bound to the RecA filament where RecA advantageously positions the UmuD active site to facilitate cleavage (200, 209).
The crystal structure of UmuD'2 presented two different dimer interfaces, designated molecular and filament. The umuD gene products likely exist as dimers given that the equilibrium dissociation constant $K_d$ for dimerization is $< 10$ pM (198). NMR and cross-linking experiments support the conclusion that the filament dimer interface is prominent in solution (200-203, 210-213). However, evidence suggests that the molecular interface may be of biological significance (203, 214). Dimerization of the UmuD2 and UmuD'2 occurs through shared structural features, but the dimer interfaces of UmuD2 and UmuD'2 are distinct. High resolution structures of UmuD'2 indicate that the sidechains of residues Leu20, Asn41, Leu43, and Leu44 within the neck region participate in intermolecular interactions that hold the dimer together (200, 213). However, differences in this region of UmuD2 and UmuD'2 dimers are illustrated by cross-linking of mono-cysteine derivatives at positions 37 and 38 (211). For UmuD2, cross-linking was observed at these positions, while identical positions in UmuD'2 did not cross-link. Introduction of a single mutation, N41D, produces stable, monomeric UmuD and UmuD' proteins that functionally mimic the wild-type dimeric proteins (215). The results of this study show that the monomeric form of UmuD is proficient for cleavage, facilitates UV-induced mutagenesis, and is able to maintain physical interactions with known partner-proteins DinB and DNA pol III $\beta_2$ (215). Indeed, dimerization does not appear necessary for a variety of UmuD activities. However, the umuD gene products are known to preferentially form UmuDD' heterodimers when a mixture of UmuD2 and UmuD'2 homodimers is present (214, 216). Formation of UmuDD' heterodimers is required for the role of UmuD in the DNA damage checkpoint, and suppression of the SOS response following resumption of DNA replication.
1.2.2b UmuD subunit exchange and heterodimer formation

The *umuD* gene products exist in multiple dimeric forms. It has been demonstrated that UmuD₂ and UmuD’₂ homodimers exchange subunits (214, 217, 218), and that a mixture of UmuD₂ and UmuD’₂ leads to formation of UmuDD’ heterodimers preferentially (214, 216). UmuDD’ heterodimers are stable and resistant to subsequent homodimer formation, however, evidence suggests that subpopulations of UmuD₂ and UmuD’₂ homodimers remain (214).

The levels of UmuD and UmuD’ are strictly controlled at the transcriptional level by LexA, and post-translationally by proteolytic degradation. UmuD and UmuD’ are degraded by the Lon and ClpXP proteases at the end of the SOS response (219-221). Full-length UmuD₂ is degraded by both Lon and ClpXP, while UmuD’ is degraded by ClpXP in the context of the UmuDD’ heterodimer (220). The signal sequences required for recognition by both Lon and ClpXP proteases are located in the UmuD N-terminal arm, which is removed upon cleavage to UmuD’ (219, 220). In the heterodimer, UmuD is able to mark UmuD’ for degradation by ClpXP. ClpXP binds the UmuD arm at its recognition sequence leading to degradation of the adjacent subunit (220, 222, 223). This mechanism allows the cell to manage levels of mutagenic UmuD’ as it accumulates under normal or SOS-induced conditions.

1.2.2c UmuD is a polymerase manager protein

The combination of dynamic arms and intrinsically disordered body structure of the *umuD* gene products facilitate the use of unique contacts to modulate a variety of protein-protein interactions (70, 198, 200, 201, 208, 224). Full-length UmuD preferentially interacts with the C-terminal
domain of α and competes with the β clamp and with ssDNA for binding to α (70, 71). UmuD also interacts with the Y-family DNA polymerase DinB and suppresses its mutagenic -1 frameshifting activity (225, 226). UmuD’ has stronger affinity for UmuC than UmuD does, and the UmuD’2C complex is the active translesion DNA polymerase DNA pol V (197, 227, 228).

1.3 Conclusions

The dynamics and interactions of the umuD gene products are of critical importance to regulating DNA replication in response to DNA damage. The dynamics of UmuD N-terminal arm truncations are characterized in Chapter 2. The dynamics of UmuD subunit exchange are elucidated in Chapter 3. Finally, characterization of proofreading and polymerase activity of pol III core is described in Chapter 4. Chapter 5 discusses preliminary data and future directions for probing potential interactions between the umuD gene products and the DNA pol III α -τ complex.
1.4 References


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 Research. 2013.


151. Goedken ER, Levitus M, Johnson A, Bustamante C, O'Donnell M, Kuriyan J. Fluorescence measurements on the *E.coli* DNA polymerase clamp loader: implications


157. Hawker JR, Jr., McHenry CS. Monoclonal antibodies specific for the tau subunit of the DNA polymerase III holoenzyme of Escherichia coli. Use to demonstrate that tau is the product of the dnaZX gene and that both it and gamma, the dnaZ gene product, are integral components of the same enzyme assembly. J Biol Chem. 1987;262(26):12722-7.


Chapter 2: Altering the N-terminal arms of the polymerase manager protein UmuD modulates protein interactions

This chapter was recently published as:


David A. Murison constructed, expressed, and purified UmuD variants, performed EPR, tryptophan fluorescence quenching, UV survival, FRET, and Western blotting assays and wrote the manuscript; Jaylene N. Ollivierre constructed, expressed, and purified UmuD variants, performed thermofluor, BMH cross-linking, UmuD cleavage, mutagenesis, homologous recombination, and UV survival assays and wrote the manuscript; Qiuying Huang performed UmuD cleavage assays; David E. Budil designed EPR experiments and wrote the manuscript; Penny J. Beuning designed experiments and wrote the manuscript.
2.1 Introduction

*Escherichia coli* cells that are exposed to exogenous or endogenous DNA damaging agents invoke the SOS response that involves the induction of at least 57 genes (1, 2). The SOS response is temporally divided into two phases: the initial phase that allows time for error-free pathways to act, and a potentially mutagenic damage tolerance phase that may ensure survival (1, 3). Key participants in the later stage of the damage response, which is often referred to as SOS mutagenesis, include the *umuDC* and *dinB* gene products.

Full-length UmuD is a homodimer of 139-amino acid subunits, and is expressed 20-30 minutes after the induction of the SOS response (1, 3, 4). UmuD interacts with the RecA:ssDNA nucleoprotein filament to facilitate the autocatalytic cleavage of the N-terminal 24-amino acids, forming UmuD’ (5-7). The UmuD’ cleavage product is a homodimer of 115-amino acid subunits, and together with UmuC, forms the Y-family polymerase DNA pol V (UmuD’2C). This specialized DNA polymerase copies damaged DNA, albeit in a potentially error-prone fashion, in a process known as translesion DNA synthesis (TLS) (1, 4, 8, 9).

The N-terminal arms of UmuD are quite dynamic and can adopt multiple conformations, which regulate interactions with partner proteins (10-14). UmuD can cleave in the *trans* (intermolecular) conformation, in which the arm of one monomer loops over and is cleaved by the active site of the adjacent monomer (11, 15). Isoenergetic models of the UmuD dimer also suggest that the *cis* (intramolecular) conformation of the arms, in which each arm binds and is cleaved by its respective C-terminal domain, is possible (11). The monomeric variant UmuD N41D also cleaves efficiently, which suggests that the *cis* conformation is likely an active
conformation (16). Additionally, the arms of UmuD may be bound (“arms down”) or unbound (“arms up”) from the C-terminal domain which may significantly alter the interacting surface that is presented for binding (10, 11).

The umuD gene products interact with multiple factors involved in DNA replication and the SOS damage response (17). UmuD and UmuD’ interact specifically with Y-family polymerases UmuC and DinB (1, 8, 9, 18). The noncatalytic UmuDC complex protects cells from the potentially harmful effects of error-prone DNA replication by delaying SOS mutagenesis (3, 19). This function is distinct from the role of UmuD’2C in error-prone TLS (1, 4, 8). Additionally, both UmuD and UmuD’ interact differentially with the α polymerase, β processivity, and ε proofreading subunits of the replicative polymerase DNA pol III (20-22).

The umuD gene products are regulated at the transcriptional and post-translational levels. The umu operon is repressed by LexA and is one of the most tightly controlled in the SOS regulon (1). Cleavage of UmuD to UmuD’ activates UmuC for TLS, and also removes the degradation signal for Lon protease (23). UmuD and UmuD’ exist by themselves as homodimers, but can also exchange subunits to preferentially form the UmuDD’ heterodimer (11, 14, 24-26). Both the UmuD’ subunit of the heterodimer and one full-length UmuD subunit of the UmuD homodimer are targeted for degradation by the ClpXP protease as a way of attenuating mutagenesis (23, 27, 28). The N-terminal arm of UmuD harbors the ClpX recognition sequence, and thus UmuD acts as the delivery factor for its bound UmuD’ or UmuD partner (27, 28).
We previously showed that, even in full-length UmuD, the N-terminal arms are only loosely bound to the globular domain (10, 13). In order to probe this further, in this work we constructed and characterized variants of UmuD possessing N-terminal truncations. These truncated proteins, UmuD 8 (UmuD Δ1-7) and UmuD 18 (UmuD Δ1-17), were used to study the conformation of the N-terminal arms, their effects on cleavage, and other cellular functions of the umuD gene products, as well as their effects on protein-protein interactions. We found that the loss of just the N-terminal seven amino acids of UmuD results in changes in conformation of the N-terminal arms, but this truncated UmuD maintains interactions with the α polymerase subunit of DNA polymerase III. Although UmuD 8 is cleaved as efficiently as full-length UmuD in vitro and in vivo, UmuD 18 is not cleaved to form UmuD’. We have also determined that UmuD 8 is proficient for UV mutagenesis, but intriguingly, sensitizes cells to UV radiation.

### 2.2 Materials and Methods

#### Strains, plasmids, and proteins

Strains and plasmids used in this work are listed in Table 2.1.

<table>
<thead>
<tr>
<th>Strains and Plasmids</th>
<th>Relevant Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td>GW8017</td>
<td>AB1157 ΔumuDC</td>
<td>(29)</td>
</tr>
<tr>
<td>PB103</td>
<td>AB1157 ΔumuDC ΔrecJ</td>
<td>P1 (JW2860)→ GW8017 (24, 30)</td>
</tr>
</tbody>
</table>
For protein expression and purification, NdeI restriction sites were introduced into the pSG5 expression vector (33) at positions 8 for UmuD 8, and 18 for UmuD 18 using a QuikChange kit (Agilent). There was already an NdeI site at the beginning of the umuD gene. The resulting plasmids were digested using NdeI (NEB), and re-ligated using T4 DNA Ligase (NEB). Mutations were confirmed by DNA sequence analysis (Massachusetts General Hospital Core Facility, Cambridge, MA). Mutagenic primer sequences are as follows: UmuDAsp8NdeI2 forward (5′-GTTGTTTATCAAGCATATGGATCTCCGCG), UmuDPhe18NdeI2 forward (5′-GTGACTTTTCATATGTTTAGCGATCTTGTTCAGTG), and their respective reverse complementary sequences. UmuD 8 and UmuD 18 were constructed in pSG5, and expressed and purified as previously described (33, 34). In general, and unless otherwise noted, the biochemical experiments reported here used non-cleavable S60A variants to avoid complications due to the possibility of spontaneous cleavage.
For bacterial experiments, KpnI restriction sites were introduced into pGY9739 or the S60A derivative at positions 1 and 8, and 1 and 18 to create the UmuD 8 and UmuD 18 truncations, respectively, using a QuikChange kit (Agilent). The resulting plasmids were digested using KpnI (NEB), and re-ligated using T4 DNA Ligase (NEB). Mutations were confirmed by DNA sequence analysis (Massachusetts General Hospital Core Facility, Cambridge, MA). Mutagenic primers are as follows: UmuDMet1KpnI19739 forward (5′-CTTCAGGCAGGGTACCATGTTGTTTATCAAGCCTG), D_9739Asp8KpnI forward (5′-GTTGTTTATCGGTACCATGGATCTCCCGAAATTGTGAC), UmuDPhe18KpnI2 forward (5′-CGCGAAATTGTGACTGGTACCATGTTTAGCGATCTTGTTC), and their respective reverse complementary sequences.

2.2.1 *In vitro* characterization of truncated UmuD variants

Thermal shift assays of full-length UmuD, UmuD′, UmuD 8 and UmuD 18 were completed as previously described using a Bio-Rad CFX 96 Real-Time system (16). Cross-linking of the UmuD N-terminal arms with 10 mM bismaleimidohexane (BMH, Pierce) at 40 µM protein was completed as previously described (11, 35), except proteins were visualized by Coomassie-stained SDS-PAGE. The RecA:ssDNA-dependent and alkaline cleavage assays were also carried out as previously described (16, 33).

Site-directed spin-labeling of purified UmuD variants for electron paramagnetic resonance experiments were carried out using the thiol-reactive nitroxide derivative, 3-iodomethyl-1-oxy-2,2,5,5-tetramethylpyrroline (Toronto Research Chemicals). Labeling chemistry was carried out as previously described (10). Continuous wave experiments were performed at room temperature
using a Bruker EMX instrument equipped with a high-sensitivity cylindrical cavity. Spectra were obtained using a 9.37 GHz microwave frequency, 6.0 mW microwave power, and 1.0 G 100 kHz field modulation amplitude. Spectra were aligned and scaled using MatLab (MathWorks) to illustrate differences in line shape as a function of nitroxide probe motion.

2.2.2 **UV survival and mutagenesis assays and inhibition of homologous recombination**

Survival and mutagenesis assays were performed as previously described (33, 34). Genetic transduction was carried out as previously described using P1vir ΔyeaB (Kan') (36). Values represent the average of at least three trials, and the error bars show the standard deviation.

2.2.3 **Immunoblotting**

Immunoblotting procedure was completed as previously described (16), with rabbit polyclonal anti-UmuD/UmuD’ antibodies (11). Band densities were determined using ImageQuantTL software (GE).

2.2.4 **UmuD and DNA polymerase III α subunit binding by tryptophan fluorescence assay and FRET**

Equilibrium dissociation binding constants $K_d$ for the interaction between UmuD proteins and DNA polymerase III α subunit (pol III α) were determined by tryptophan fluorescence as previously described (21, 33). Pol III α truncations α1-280 and α917-1160 were described previously (21) and were used here to localize UmuD 8 and UmuD 18 binding sites on α. UmuD was titrated into a solution of α.
Figure 2.1 – Characterization of UmuD 8 and UmuD 18 in vitro.

(A) Model of UmuD (with arms down) showing residues 1-7 (purple), 8-17 (green) 18-24 (blue), 25-40 (red). UmuD 3A mutations T14 (orange), L17 (gray), F18 (cyan) and active site residue S60 (yellow) are also highlighted. (B) Thermal shift assays of UmuD proteins. Melting of the proteins as a function of temperature was monitored by changes in Sypro Orange fluorescence. Results for UmuD (purple), UmuD 8 (green), UmuD 18 (blue) and UmuD’ (red) are shown using 40 μM protein. (C) UmuD protein arms were cross-linked using BMH. Percent of cross-linked dimers are indicated below the lanes. The cross-linking reaction was carried out for 10 min at room temperature after the addition of BMH. Protein was visualized by Coomassie stain. (D) Continuous wave EPR spectra of UmuD variants. Arrow I indicates line shape from a partially-immobilized species; arrow II shows line shape characteristic of elevated mobility.
Protein labeling with Alexa488 and Alexa647 (Life Technologies) and FRET assays were performed as previously described (21). Purified UmuD variants were added to a final concentration of 40 μM, and incubated with fluorescent α and β proteins prior to analysis. FRET efficiency was calculated as previously described (21).

2.3 Results

2.3.1 Loss of N-terminal residues changes UmuD arm characteristics

In order to probe the dynamics and functions of the N-terminal arms of UmuD, we generated truncations lacking the N-terminal 7 or 17 residues, denoted as UmuD 8 and UmuD 18, respectively (Fig. 2.1A). Truncating the N-terminal arms of UmuD to create UmuD 8 and UmuD 18 changed the melting profile relative to wild-type UmuD (Fig. 2.1B). It was previously shown that wild-type UmuD melts in two transitions (13). The transition at approximately 30 °C is attributed to release of the N-terminal arms from the globular C-terminal domain, and the second transition at approximately 60 °C is associated with melting of the globular domain (13). The arms of UmuD' (residues 25-40) are not in contact with the C-terminal domain; therefore, only one melting transition is observed at approximately 60 °C for UmuD' (13). The melting profile for UmuD 18 resembles that of UmuD' with a single transition observed at 62 °C (Fig. 2.1B). This suggests that the N-terminal arms of UmuD 18 are also dissociated from the C-terminal domain. Conversely, the melting profile of UmuD 8 resembles that of wild-type UmuD in which there is an initial melting transition at 43 °C for UmuD 8. This observation is consistent with a model in which the first transition is due to dissociation of the N-terminal arms from the globular domain. Thus, the arms of UmuD 8 are apparently of sufficient length to interact with the
globular domain, whereas the arms of UmuD 18 are presumably too short to form a stable interaction.

The conformation of the N-terminal arms of UmuD 8 and UmuD 18 was also assessed by cross-linking with the homobifunctional cross-linker bismaleimidohexane (BMH). The BMH cross-linker is 13 Å in length and reacts with free cysteine thiols. The model of wild-type UmuD with N-terminal arms in the “down” conformation shows that the C24 residues within the dimer are separated by a distance of 20 Å (Fig. 2.1A) (11). Therefore, the arms must be unbound from the C-terminal domain for cross-linking of the single-Cys residues to occur. UmuD S60A C24A A7C was used as a control to represent maximal cross-linking as the first few residues of the UmuD arm are highly dynamic (13). As expected, the UmuD S60A C24A A7C variant was cross-linked the most readily (71%) due to the position of the cysteine near the end of the arm. UmuD 8 S60A (32%) and UmuD S60A (31%) exhibited similar cross-linking efficiencies, which were slightly lower than those of UmuD 18 S60A (39%) and UmuD 3A (41%) (Fig 2.1C). UmuD' was not used in this assay because full-length UmuD contains a single cysteine residue C24 which is at the cleavage site and is removed upon cleavage. The UmuD 3A variant possesses three alanine mutations (T14A L17A F18A) and is considered a UmuD’ mimic because the arms weakly interact with the C-terminal globular domain (11, 13). Together, these results show that deletion of the first eight residues does not change the cross-linking efficiency compared to UmuD S60A, but deletion of the first 18 residues causes increased cross-linking efficiency. We believe the increase is due to reduced interaction between the arms and the globular domain in the case of UmuD18.
Figure 2.2 – Continuous wave EPR spectra of UmuD variants.

(A) Overlaid spectra of UmuD’ A31C and the truncation variants: The N-terminal arms of UmuD exhibit increased mobility when shortened. (B) Overlaid spectra of truncation variants with UmuD 3A demonstrate greater intermediate and slow-motional contributions to the spectral line shape of the truncation variants.

Site directed spin labeling (SDSL) allows for detection of conformational changes as well as local dynamics in a protein by electron paramagnetic resonance (EPR) spectroscopy. Three UmuD variants were modified with the paramagnetic spin label 3-iodomethyl-1-oxy-2,2,5,5-
tetramethylpyrroline (iodomethyl spin label, IMSL) which specifically reacts with the sulphydryl group of cysteine residues. UmuD truncation variants UmuD 8 S60A and UmuD 18 S60A were labeled at the natural C24 position, and UmuD S60A C24A A7C was labeled at residue 7 near the end of the full-length arm. The spectra appear to be the superposition of spectra from at least two subpopulations of the nitroxide spin label: one displaying the characteristic three-line spectrum of a nitroxide undergoing fast motion, and additional components with broader lines indicating varying degrees of slower motion. Fig. 2.1D compares spectra from the previously characterized UmuD S60A variant (10) with the truncated variants UmuD 8 and UmuD 18. The spectrum of UmuD S60A (purple) exhibits peaks from a relatively immobilized species (arrow, I) and from a more mobile population (arrow, II) in addition to the characteristic three-line spectrum. We previously demonstrated a temperature-dependent equilibrium between these two components (10). When the arm is truncated (green spectrum, UmuD 8 S60A), the spectrum reflects an increase in the relative amount of the more mobile component, which becomes more pronounced as the truncation is increased in the UmuD 18 S60A construct (Fig. 2.1D, blue). This is consistent with the conclusion that the arm becomes more mobile as it is shortened. UmuD’ A31C, which was previously shown to exhibit only a fast-motional component (10), is compared with the two truncation variants of the present study in Fig. 2.2A. This comparison clearly highlights the immobile component that is present upon partial truncation of the arms. As expected, the immobile component becomes more prominent as the length of the arm is increased. Similarly, comparison of spin-labeled UmuD 3A, which shows only intermediate and fast-motional components (10), with UmuD 8 S60A and UmuD 18 S60A reveals more intermediate and slow-motional components than for UmuD 3A (Fig. 2.2B). We constructed UmuD A7C C24A S60A to place a spin label near the end of the N-terminal arm in the
expectation that this would exhibit only a fast motional component, similar to UmuD' A31C; however, to our surprise this construct exhibits some less-mobile component, which could indicate some structure in the extreme N-terminal ends of the UmuD arms.

### 2.3.2 UmuD 8 cleaves efficiently, UmuD 18 is not cleavable

Cleavage of UmuD to UmuD' is required for the activation of pol V (UmuD'C) in translesion DNA synthesis (1). The removal of the N-terminal 24 amino acids is facilitated by binding of UmuD to the RecA:ssDNA nucleoprotein filament, which positions the UmuD active site residues S60 and K97 in the correct orientation for the cleavage reaction (14). Cleavage of UmuD 8 and UmuD 18 was assayed alongside full-length UmuD. Cleavage of UmuD 8 was nearly as efficient as that of wild-type UmuD, whereas that of UmuD 18 was dramatically reduced (Fig. 2.3A and 2.3B). We also assessed cleavage of UmuD 8 under alkaline conditions (pH 10) in the absence of the RecA:ssDNA filament; at pH 10, cleavage is less efficient overall, but the active site serine can be activated as a nucleophile without the addition of RecA:ssDNA (5). Under alkaline conditions, similar to our observations with RecA-facilitated cleavage, the cleavage of UmuD 8 was similar to that of full-length UmuD (Fig. 2.3C). UmuD 18 does not undergo RecA:ssDNA-facilitated cleavage appreciably (Fig. 2.3ABD). To determine whether UmuD 18 has a functional active site, we performed a RecA:ssDNA-dependent cleavage assay in which UmuD 18 and the active site variant UmuD S60A were mixed and allowed to form heterodimers. The N-terminal arms of UmuD S60A can then be cleaved in trans by the active site of UmuD 18 (11, 15). We found that cleavage in the context of UmuD S60A/UmuD 18 heterodimers was indeed efficient (Fig. 2.3D). This confirms that UmuD 18 and UmuD S60A can form heterodimers and that the active site of UmuD 18 is competent for cleavage, suggesting
that the cleavage defect of UmuD 18 is due to decreased binding of its arm to its globular domain and presumably not due to a defect in its ability to interact with the RecA/ssDNA filament.

Figure 2.3 – UmuD 8 cleaves as efficiently as wild-type UmuD; UmuD 18 does not cleave to UmuD’. (A) Relative cleavage to UmuD’ in the presence of RecA:ssDNA nucleoprotein filament. UmuD proteins at a concentration of 10 μM were used and the reaction was carried out for 1 h at 37 °C. Percent cleavage product was determined as a ratio of the density of the UmuD’ band to the total density of the uncleaved UmuD proteins and UmuD’ protein for each reaction. (B) Comparison of the kinetics of cleavage of UmuD 8, UmuD 18, and wild-type UmuD. Reactions were carried out over 6 h. (C) Results for cleavage to UmuD’ under alkaline conditions (pH 10) for 48 h are also represented. Results are normalized to cleavage of wild-type UmuD to form UmuD’. (D) Mixing equal amounts of UmuD 18 with the active site variant UmuD S60A results in cleavage. UmuD proteins at 10 μM were used and cleavage was carried out at 37 °C for 1 h.
2.3.3 UmuD arm length attenuates affinity for binding sites on DNA pol III α subunit

The *umuD* gene products interact with the DNA pol III α polymerase subunit at two locations: at the N-terminal domain between residues 1-280 and at the C-terminal region between residues 956-975 (21). Since UmuD contains no tryptophans and α contains eight, we measured the intrinsic fluorescence of α in the presence of increasing amounts of UmuD to determine equilibrium dissociation constants $K_d$ for the UmuD truncation variants binding to α. We probed the affinity of UmuD 8 S60A and UmuD 18 S60A for three forms of the α subunit: full-length α, α1-280, and α917-1160. Our observations indicate that UmuD 8 S60A exhibits a strong affinity for the α917-1160 C-terminal fragment ($K_d = 0.2 \pm 0.4 \mu M$) which is similar to the equilibrium dissociation constant determined for UmuD S60A binding to the same α fragment ($K_d = 0.7 \pm 0.3 \mu M$) (21). On the other hand, UmuD 18 S60A displayed a weaker affinity for the C-terminal fragment ($K_d = 3.6 \pm 0.4 \mu M$), and the calculated equilibrium dissociation constant closely resembles the values determined for UmuD’ and UmuD3A ($K_d = 3.8 \pm 0.9 \mu M$ and $3.4 \pm 1.0 \mu M$, respectively) (Table 2.2) (21). These observations further support the idea that UmuD 8 mimics UmuD whereas UmuD 18 is similar to UmuD’.

Previous experiments have shown that full-length UmuD is capable of disrupting the interaction between the DNA polymerase III polymerase subunit α and the processivity clamp β (21). When purified α and β subunits are labeled with acceptor and donor fluorophores, FRET is observed when donor-labeled β clamp is in the presence of acceptor-labeled α subunit. As expected and shown previously, FRET efficiency was significantly decreased in the presence of wild-type UmuD, but negligibly affected by the presence of UmuD’ (Fig. 2.4) (21). When the same
experiment was performed using the UmuD arm truncation variants, we observed that UmuD 8 S60A was able to decrease FRET between α and β whereas in the presence of UmuD 18 S60A FRET efficiency was unchanged (Fig. 2.4). This result is consistent with our observation that UmuD 8 S60A has a higher affinity for the C-terminal region of α similar to that of full-length UmuD and thus can compete for binding to the β clamp. Our observations show a correlation between longer N-terminal arm length, stronger affinity for the C-terminus of α, and the ability to disrupt the α-β complex.

Table 2.2. Equilibrium dissociation constants

<table>
<thead>
<tr>
<th>UmuD variants</th>
<th>Full-length α (μM)</th>
<th>α1-280 (μM)</th>
<th>α917-1160 (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT UmuD</td>
<td>1.1 ± 0.6</td>
<td>7.3 ± 1.4</td>
<td>13.9 ± 5.1</td>
</tr>
<tr>
<td>UmuD S60A</td>
<td>10.6 ± 2.9</td>
<td>10.3 ± 4.3</td>
<td>0.7 ± 0.3</td>
</tr>
<tr>
<td>UmuD'</td>
<td>10.9 ± 0.6</td>
<td>8.6 ± 1.0</td>
<td>3.8 ± 0.9</td>
</tr>
<tr>
<td>UmuD 8 S60A</td>
<td>8.1 ± 0.9</td>
<td>3.5 ± 0.8</td>
<td>0.2 ± 0.4</td>
</tr>
<tr>
<td>UmuD 18 S60A</td>
<td>12.2 ± 0.2</td>
<td>5.9 ± 0.9</td>
<td>3.6 ± 0.4</td>
</tr>
<tr>
<td>UmuD 3A</td>
<td>8.7 ± 1.5</td>
<td>9.3 ± 3.2</td>
<td>3.4 ± 1.0</td>
</tr>
</tbody>
</table>

1Equilibrium dissociation constants \(K_d\) (μM) for WT UmuD, UmuD S60A, UmuD', and UmuD 3A were previously determined (21) and are reported here for ease of comparison.
Figure 2.4 – UmuD 8 S60A is able to disrupt the interaction between the α polymerase subunit and β processivity clamp while UmuD 18 S60A cannot. FRET was monitored between α labeled with Alexa Fluor 647 C2-maleimide and β labeled with Alexa Fluor 488 C5-maleimide. The bar graph shows FRET efficiency calculated in the presence and absence of purified UmuD proteins at 40 μM. Error bars represent standard deviation of three or more independent replicates.

2.3.4 UmuD 8 is proficient for UV-induced mutagenesis; UmuD 8 and UmuD 18 do not confer resistance to UV radiation

UmuD’2C performs TLS on UV-induced DNA damage and is required for UV-induced mutagenesis in *E. coli* (1). As Pol V (UmuD’2C) inserts guanine opposite the 3’-thymine of (6-4) T-T photoproducts (37, 38), polymerase activity can be detected via the reversion of the *argE3* auxotrophic marker in the *E. coli* arginine biosynthetic pathway (33). To determine the proficiency of UmuD 8 and UmuD 18 in UV-induced mutagenesis, we compared the mutation frequency of ΔumuDC strains harboring plasmid-borne full-length UmuD, UmuD’ and the truncated versions UmuD 8 and UmuD 18 (Fig. 2.5). We also compared the corresponding active
site variant, S60A, of each protein. We found that the mutation frequency of cells expressing UmuD 8 is similar to that of full-length UmuD, and as expected the non-cleavable UmuD 8 S60A shows greatly reduced UV-induced mutagenesis (Fig. 2.5). The cleavage efficiency and expression level of UmuD 8 are also comparable to that of full-length UmuD in vivo (Fig. 2.6). This suggests that UmuD 8 functions similarly to UmuD in this context, and is able to interact with protein partners that are required for UV-induced mutagenesis, including UmuC and RecA (1). However, UmuD 18 shows reduced UV-induced mutagenesis (Fig. 2.5). Although we did not detect cleavage of UmuD 18 in vitro or in vivo (Figs. 2.3 and 2.6), cells expressing UmuD 18 had low but detectable mutagenesis. The greater UV-induced mutagenesis with this non-cleavable UmuD 18 truncation than other non-cleavable UmuD proteins can likely be attributed to the more dynamic N-terminal arms of UmuD 18, which would allow it to partially mimic UmuD'. In addition, the lower mutation frequency could be explained if accurate TLS is promoted by altered interactions of UmuD 18.

![Graph showing mutation frequency](image)

**Figure 2.5 – UmuD 8 is proficient for UV-induced mutagenesis.**
Mutagenesis assays were performed in strain GW8017.
It was previously reported that Δ*umuDC ΔrecJ* strains are hypersensitive to UV radiation and that this phenotype can be suppressed by complementing with low-copy plasmids harboring the *umuDC* genes (16, 39). RecJ is an exonuclease that aids in DNA replication restart by degrading DNA at stalled replication forks (39, 40). In the absence of RecJ, replication restart is postponed and DNA synthesis is carried out by the TLS polymerase Pol V (39). We found that cells expressing UmuD 8 or UmuD 18 display a similar level of sensitivity to UV light as cells harboring empty vector (Fig. 2.7), and that UmuD 8 S60A surprisingly sensitized cells to UV, to
an even greater extent than UmuD S60A. This extreme sensitivity to UV light conferred by UmuD 8 S60A was also observed for a strain with recJ (GW8017, Fig. 2.7) and even in the context of AB1157 cells that harbor wild-type *umuD* on the chromosome (Fig. 2.8). This phenotype was unexpected as UmuD 8 is proficient for UV-induced mutagenesis and presumably the UmuD 8 S60A variant could be cleaved by the chromosomally-encoded *umuD* to form UmuD’. We constructed corresponding plasmids expressing catalytically-deficient UmuC (D101N) or lacking *umuC* altogether and found that in both contexts, UmuD 8 S60A conferred UV hypersensitivity (Fig. 2.9). The *umuDC*-encoding plasmid we typically use for these experiments harbors a promoter mutation resulting in higher-than-normal expression levels (31, 41). We therefore constructed an o⁺ version of the plasmid expressing UmuD 8 S60A and observed that this failed to confer UV hypersensitivity (Fig. 2.8), thus indicating that the extreme sensitivity to UV light caused by UmuD 8 S60A is due to a copy-number effect. Our observation that cells harboring the *oC₁* version of UmuD 8 S60A are extremely sensitive to UV light suggests that elevated levels of UmuD 8 S60A can be harmful to cells. We next examined another characteristic phenotype of *umuDC*, specifically the inhibition of RecA-mediated homologous recombination by elevated levels of UmuD’C (42-47). UmuD 8 and UmuD 18 show similar levels of inhibition of RecA-mediated homologous recombination as full-length UmuD, again indicating that the truncated proteins appear to be proficient for interaction with RecA. UmuD 18 S60A inhibits RecA-mediated homologous recombination, suggesting it partially mimics UmuD’ in its interactions with RecA-coated DNA (Fig. 2.10).
Figure 2.7 – UmuD 8 S60A confers sensitivity to UV light.
(A) UV Survival in strain GW8017 ΔrecJ which lacks chromosomal umuDC and recJ. (B) UV survival in GW8017. Plasmids encode umuDC that vary only in the umuD construct.
Figure 2.8 – Sensitivity induced by UmuD 8 S60A is alleviated by tighter control over expression.
(A) UV Survival in strain AB1157 which has chromosomal *umuDC*. UmuD 8 and UmuD 8 S60A confer UV sensitivity. (B) UV Survival of pUmuDC variants with wild-type promoter (*o*<sup>+</sup>) in GW8017.
Figure 2.9 – UV Sensitivity caused by UmuD is not due to a deficient pol V interaction. (A) UV Survival of plasmids possessing umuD only. The umuC sequence was removed entirely from umuDC operon in pGY9739 (B) UV survival of plasmids harboring umuC104 allele (D101N) in GW8017. The point mutation D101N in UmuC inactivates the protein as a polymerase.
UmuD proteins from plasmids pGY9739 (umuDC), pGY9738 (umuD’C), and those harboring umuD variants UmuD 8 and UmuD 18 in strain GW8017 were expressed to determine the extent of inhibition of RecA-facilitated homologous recombination. Plasmids encode umuDC that vary only in the umuD construct.

Figure 2.10 – UmuD 8 and UmuD 18 do not inhibit homologous recombination to the same extent as UmuD’C.

2.4 Discussion

The goal of this study was to characterize two truncation variants of the E. coli polymerase manager protein UmuD. We used several in vivo and in vitro techniques to investigate the effect of N-terminal arm length on protein conformation and activity. Surprisingly, we discovered that the variant UmuD 8 caused UV sensitivity in cells when expressed from a plasmid. We then attempted to characterize the cause of this sensitivity. Given that UmuD 8 is cleavable and viable for UV-induced mutagenesis, we were surprised to find that UmuD 8 was unable to confer resistance to UV in a ΔumuDC ΔrecJ strain by complementation (Fig. 2.7). Moreover, UV survival of a strain possessing a chromosomal copy of umuDC (AB1157) and harboring a low-copy plasmid bearing UmuD 8 was decreased relative to this strain harboring an empty vector.
(Fig. 2.8). Cells are even more sensitive to UV when harboring a plasmid encoding the non-cleavable UmuD\(^{8}\) S60A. The observed UV hypersensitivity phenotype was independent of UmuC catalytic activity in the context of Pol V, as introduction of the \textit{umuC}104 allele (D101N), which renders UmuC catalytically inactive (48), into the respective plasmid constructs did not confer resistance to UV. Complementation by plasmid-borne UmuD\(^{18}\), on the other hand, promoted survival relative to empty vector in the case of both strains. This result was also surprising as UmuD\(^{18}\) is non-cleavable and renders cells only weakly mutable (Figs. 2.3 and 2.5). Plasmids used for complementation contain the \(o^{C1}\) point mutation in the \textit{umuDC} operator which decreases the regulation of plasmid-borne \textit{umuDC} gene product expression by preventing LexA binding (31). When reverted to the wild-type \(o^{+}\) operator sequence, the UV hypersensitivity phenotype observed for plasmid-borne UmuD\(^{8}\) was alleviated. We therefore attribute this phenotype to a copy-number effect.

X-ray and NMR structures of UmuD\(^{'}\) (14, 49) show that the N-terminal arms are dissociated from the globular domain and are predominantly unstructured. While a high resolution structure of full-length UmuD has not been solved, homology models (Fig. 2.1A) and observations from biophysical experiments indicate that the full-length N-terminal arms are dynamic, but also can stably bind to the globular domain (10, 11, 13, 50). Indeed, interaction between arm residues Cys24 and Gly25 and globular domain active site residues Ser60/Lys97 is required for cleavage to occur. Both \textit{in vivo} and \textit{in vitro} cleavage experiments show that UmuD\(^{8}\) cleaves as efficiently as wild-type UmuD whereas UmuD\(^{18}\) is not cleavable (Figs. 2.3 and 2.6). The single melting transition observed in thermofluor (Fig. 2.1B) and elevated BMH cross-linking efficiency (39%, relative to 25% for WT UmuD and 41% for UmuD3A; Fig. 2.1C) suggest that the conformation
of UmuD 18 is more similar to that of UmuD’ in which the truncated arms weakly associate with the globular domain. The triple mutant UmuD 3A (T14A, L17A and F18A) possesses full-length arms, but is non-cleavable (10, 11, 13). The three point mutations in UmuD 3A prevent interaction between the arms and globular domain. We believe that, like UmuD 3A, UmuD 18 is structurally similar to UmuD’, and its arms are not cleavable because they interact with the globular domain far more weakly than those of full-length UmuD.

The umuD gene products have been shown to interact with an increasingly large number of partner proteins. The UmuD interactome includes translesion DNA polymerases DinB and UmuC, RecA, subunits α, β, and ε of replicative DNA polymerase III, as well as proteases Lon and ClpXP (1, 5-7, 12, 21, 23, 27, 33, 51, 52). Many of these interactions demonstrate preference for either UmuD or UmuD’. In addition, because UmuD variants lacking the N-terminal seven or eight residues maintain their interactions with both α (Table 2.2) and β (12), the disruption of α-β binding by UmuD 8 may be due to competitive interactions of UmuD 8 with both α and β. On the other hand, both UmuD 8 and UmuD 18 show reduced inhibition of RecA-mediated homologous recombination (Fig 2.10). Previous work identified several variants of UmuD’ localized to the N-terminal arms that enhanced the inhibition of RecA-mediated recombination, specifically G25D, S28T, P29L, E35K, as well as T95R, suggesting an important role for the N-terminal arms region of UmuD’ in modulating recombination (45). In addition, UmuD single-cysteine derivatives that cross-linked most efficiently to RecA are at UmuD positions 34 and 81 (53). These residues are present in both UmuD and UmuD’, and thus are also present in UmuD 8 and UmuD 18. Previous work showed that amino acid positions 19 and 24 are not implicated in
interaction with RecA (53), so it is not unexpected that UmuD 8 and 18 inhibit homologous recombination to a similar extent.

Previous work from our lab has shown that UmuD interacts with two regions of α (21). The first was localized to N-terminal residues 1-280 which make up the polymerase and histidinol phosphatase (PHP) domain (54-56), and the second was localized to the C-terminal region between residues 956-975. The C-terminal region of α binds more strongly to full-length UmuD-S60A relative to UmuD’ and UmuD 3A. UmuD’ and UmuD 3A share similar affinity for the α C-terminal region ($K_d = 3.8 \pm 0.9$ and $3.4 \pm 1.0 \mu M$ respectively), but differ in arm length. Considered together, these observations suggest that the interaction between the C-terminal region of α and UmuD requires that UmuD adopt an “arms-down” conformation in which the N-terminal arms of UmuD associate with its globular domain and create a specific binding site. UmuD 8 S60A exhibits similar affinity for the C-terminal region of α ($K_d = 0.2 \pm 0.4 \mu M$) compared to UmuD S60A ($K_d = 0.7 \pm 0.3 \mu M$ (21)). It has previously been shown that UmuD S60A interacts with the α917-1160 fragment in an arms-down fashion (21). Therefore, the N-terminal arms of UmuD 8 are also likely capable of associating with the C-terminal globular domains to achieve the preferred “arms-down” conformation. The binding constant determined for the interaction of UmuD 18 and α917-1160 ($K_d = 3.6 \pm 0.4 \mu M$) mirrors the values calculated for UmuD 3A and UmuD’. Like the UmuD 3A and UmuD’ arms, the arms of UmuD 18 are also likely free in solution given that this variant is non-cleavable and exhibits relatively high levels of BMH cross-linking, and thus shows a weaker interaction with the C-terminal region of α.
In a previous study, UmuD arm variants similar to UmuD 8 and UmuD 18 were used to show that the β processivity clamp of DNA polymerase III has greater affinity for full-length UmuD and this affinity is somewhat reduced as residues are removed from the N-terminal arms, although β also binds to UmuD’ (12). This preference was attributed to the presence of an interface created by the contact between the N-terminal arms and the C-terminal globular domain in full-length UmuD. The *umuD* gene products have been shown to inhibit DNA replication, which is presumably accomplished by their specific interactions with α, β, and likely other subunits of the replisome (3, 20, 21, 51, 57). Indeed, in a FRET assay, energy transfer between fluorescently-labeled α and β subunits was decreased in the presence of wild-type UmuD, UmuD-S60A, or UmuD 8, but no change in FRET was observed in the presence of UmuD’ or UmuD 18 (21). Taken together, this and previous work support a model of specific interactions of the *umuD* gene products that dictate protein interactions important for regulating DNA replication.
2.5 References


Chapter 3: UmuD dimer exchange

David A. Murison constructed, expressed, and purified UmuD variants, fluorescently labeled and purified proteins, performed experiments, analyzed data, wrote the manuscript; Rebecca Timson, fluorescently labeled and purified proteins; Penny J. Beuning designed experiments and wrote the manuscript.
3.1 Introduction

In the presence of DNA damage, specialized genes are upregulated as part of the SOS response (1). In Escherichia coli, the induced damage response pathway is temporally separated into two stages. The initial stage provides an opportunity for accurate repair pathways, such as homologous recombination and nucleotide excision repair, to take place. The later stages involve damage tolerance in which potentially mutagenic mechanisms are utilized to promote cell survival.

In Escherichia coli, the umuD gene products are involved in both early and late stages of DNA damage tolerance. Full-length UmuD₂ is initially expressed as a homodimer of, 139-amino acid subunits with dynamic N-terminal arms (2-4). The full-length protein exists as the predominant form for the first 20-30 minutes after SOS induction (1, 5), and its presence (together with UmuC) has been shown to inhibit the resumption of DNA replication and increase cell survival after UV irradiation (5). UmuD₂ interacts with multiple subunits of DNA polymerase III, the primary replicative polymerase in E. coli (6-9). Specifically, UmuD₂ has been shown to disrupt interaction between the β processivity clamp and α polymerase subunit (7). These activities are consistent with a model in which UmuD₂ provides time for the cell to repair damage accurately.

During the SOS response, the RecA:ssDNA nucleoprotein filament facilitates the autocleavage of UmuD₂ homodimers to UmuD'₂ homodimers. Formation of UmuD'₂ is achieved by removal of the N-terminal 24-amino acids from each arm. UmuD residues Ser60 and Lys97 form a catalytic dyad, in which lysine activates serine for nucleophilic attack on the peptide backbone between residues Cys24 and Gly25 of the arm (10). Cleavage to form UmuD'₂ activates the protein as a
participant in association with UmuC to form DNA polymerase V (UmuD′C, DNA pol V), one of two Y-family polymerases in *E. coli*. Cleavage of UmuD2 to UmuD′2 marks a “temporal switch” in which the cell transitions from faithful DNA damage repair mechanisms to DNA damage tolerance mode (11-13).

The chronological activity of *umuD* gene products in the SOS response is often described as a primitive DNA damage checkpoint: DNA synthesis is delayed so that error-free repair may take place, but later translesion synthesis (TLS) is activated by accumulation of UmuD′2 (11-13). DNA pol V is an error-prone TLS polymerase capable of replicating damaged DNA bases in a potentially mutagenic fashion (14, 15). TLS is a mode of DNA damage tolerance used to promote cell survival. A characteristic of TLS polymerases is a large, open active site capable of accommodating bulky, damaged DNA bases. As a tradeoff, the looser active site makes these polymerases relatively inaccurate when copying undamaged DNA. Pol V, for instance, exhibits error frequencies between $10^{-3}$ and $10^{-4}$ (16). Therefore, the activity of DNA pol V is tightly regulated at the pre- and post-translational levels.

As mentioned earlier, the activation of DNA pol V as a polymerase is managed by the availability of UmuD′2 homodimers after induction of SOS. Once DNA damage has been bypassed, the cell must also have a mechanism to reduce cellular concentrations of this error-prone polymerase, which is in part accomplished by proteolysis as UmuD2 and UmuD′2 are degraded by the ClpXP and Lon proteases (17, 18). Full-length UmuD2 is recognized by ClpXP using a specific sequence (L$^9$R$^{10}$E$^{11}$I$^{12}$) located on the N-terminal arm (17, 19, 20). However, in the context of UmuD′2, this recognition sequence is absent. UmuD2 and UmuD′2 homodimers
exchange to form UmuDD' heterodimers preferentially when both species are present. In the context of the UmuDD' heterodimer, uncleaved UmuD shuttles the UmuD' subunit to ClpXP for degradation. This mechanism allows the cell to manage levels of UmuD' during the later stages of SOS.

UmuD is an intrinsically disordered protein (IDP) and adopts multiple conformations in solution (21, 22). Like other IDPs, the dynamic nature of UmuD provides diverse binding surfaces, and thus, the ability to interact with many protein partners (3, 4, 21). Electron Paramagnetic Resonance (EPR) and hydrogen-deuterium exchange mass spectrometry (HXMS) studies of full-length UmuD2 show that the N-terminal arms in particular are highly dynamic (2, 3, 23, 24). The truncated arms of UmuD'2 are also dynamic (2, 3). Two dimer interfaces (termed molecular and filament) were observed in the crystal structure of UmuD'2, but NMR and cross-linking experiments suggest that the filament interface is the one found in solution (10, 19, 25-28).

Previous investigation by our lab and others examined the kinetics of exchange between UmuD2 and UmuD'2 homodimers as well as the preferential formation of UmuDD' heterodimers (29, 30). Our experiments led to an unexpected result when heterodimer formation was not observed when UmuD'2 was added to a mixture of UmuD2 fluorescently labeled at position G92C. This surprising observation was attributed to the location of fluorescent label interfering with the subunit exchange mechanism. In order to further probe the mechanism of UmuD dimer exchange, we investigate here the effect of different fluorescent probe locations on UmuD subunit exchange kinetics using novel UmuD mono-cysteine variants in our previously developed Förster Resonance Energy Transfer (FRET) assay (29). It has been demonstrated that
UmuD\textsubscript{2} and UmuD'\textsubscript{2} homodimers exchange readily (31, 32), and kinetic analysis indicates that the exchange occurs on a timescale relevant to the SOS response (11, 29). Indeed, we found that the rate of homodimer exchange between separate pools of fluorescently-labeled UmuD\textsubscript{2} varied with position of the fluorescent label. Further, all variants preferentially formed UmuD'D' heterodimers upon addition of UmuD'\textsubscript{2} and the kinetics of heterodimer formation are also affected by position of the fluorescent label.

3.2 Materials and Methods

3.2.1 Strains, plasmids, and proteins

Table 3.1 – Strains and Plasmids

<table>
<thead>
<tr>
<th>Strains and Plasmids</th>
<th>Relevant Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td>GW8017</td>
<td>AB1157 (\Delta)umuDC</td>
<td>(19)</td>
</tr>
<tr>
<td>PB103</td>
<td>AB1157 (\Delta)umuDC (\Delta)recJ</td>
<td>(29, 33)</td>
</tr>
<tr>
<td>BL21 DE3</td>
<td></td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGY9739</td>
<td>(umuD')C; pSC101-derived, Spec\textsuperscript{R}</td>
<td>(34)</td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector; pSC101-derived, Spec\textsuperscript{R}</td>
<td>(35)</td>
</tr>
<tr>
<td>pSG4</td>
<td>(umuD'), Amp\textsuperscript{R}</td>
<td>(36)</td>
</tr>
<tr>
<td>pSG5</td>
<td>(umuD), Amp\textsuperscript{R}</td>
<td>(36)</td>
</tr>
</tbody>
</table>
Plasmids and *E. coli* strains used in this study are listed in Table 3.1. For expression and purification of UmuD and UmuD' used for *in vitro* characterization, proteins were purified from pSG5 and pSG4 plasmids, respectively, as previously described (36, 37). Variants of UmuD and UmuD’ were constructed by site-directed mutagenesis using a QuikChange kit (Agilent). Mutations were confirmed by DNA sequencing analysis (Eton Biosciences, Charlestown, MA). Mutagenic primer sequences are as follows:

UmuD-I78C forward (5’- CGATAGCGCTTTGTACCGCCAGCCATGG), UmuD-T79C forward (5’- CGATAGCGCTATTTGCGCCAGCCATGG), UmuD-D91C forward (5’- GATATTGTCATCGCTGTGTTTGCGGCGAGTTTACGGTGAAAAAA), UmuD-E93C forward (5’- CTGTTGACGCTGTTTACGGTGAAAAATTGC), UmuD-K98C forward (5’- CGAGTTTACGGTGAAATGTTTGCAACTACGCCGACGGTA), UmuD-I108C forward (5’- CCGACGCTACAGCTTCTCCCTCCTGAACAGCGCTACGCTTCGCCCATT), UmuD-M110C forward (5’- CTTCATTCCATGAAAACGCGCTGCTCCCGCCCATTACCGATCGCTGCCTG), UmuD-T118C forward (5’- CGCGTACTCGGCCATTTCGATCTCGTACTGCTGCTG), UmuD-E122C forward (5’- GCCCATACATCAGTAGTGGTGCGGACAGCTG), UmuD-D123C forward (5’- GCCCATACATCAGTAGTGGTGAGAGTAGACG) and their respective reverse complement sequences. Biochemical experiments reported here used non-cleavable S60A variants to avoid complications arising from spontaneous cleavage. Biochemical experiments using mono-cysteine variants used mutation C24A to avoid secondary fluorescent
labeling of the N-terminal arms. Thus, all purified UmuD variants contain the C24A S60A mutations in addition to the single-Cys mutations noted.

For UV survival and immunoblotting assays, plasmid pGY9739 encoding umuD and umuC genes was used. Cysteine substitutions were introduced into pGY9739 by site-directed mutagenesis using a QuikChange kit (Agilent). Mutations S60A and C24A were not introduced into these constructs. Mutations were confirmed by DNA sequencing analysis (Eton Biosciences, Charlestown, MA).

3.2.2 Protein labeling with fluorescent dyes

Purified UmuD proteins were labeled with thiol-reactive Alexa Fluor 488 maleimide (Invitrogen) or Alexa Fluor 647 maleimide (Invitrogen). UmuD protein (20 nmols) was diluted to 100 µL using Qa buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 0.1 mM EDTA). To ensure reduction of cysteine sidechains, 0.5 M TCEP was added to 2 mM and reduction was allowed to proceed for 30 min at room temperature (approximately 22 °C). Buffer exchange into Ql buffer (100 mM HEPES pH 7.25, 100 mM NaCl, 0.1 mM EDTA) was performed using 2 mL Zeba Spin desalting columns (ThermoFisher Scientific) equilibrated with Ql buffer. The volume of reduced protein was increased to 500 µL using Qa buffer before loading onto the column. Eluted protein was concentrated to ≤ 100 µL using Amicon Ultra 0.5 mL MWCO 10 kDa concentrators (Millipore) and transferred directly to 50 nmol of Alexa Fluor fluorophore. Conjugation reactions were allowed to proceed for 4 h at room temperature while protected from light. Reactions were quenched by addition of β-mercaptoethanol to 2 mM.
UmuD protein was separated from unconjugated fluorophore by gel filtration using a 1.5 x 20 cm off-line column (BioRad) packed with Sephadex G-50 resin (GE Healthcare) and equilibrated with Qₐ buffer + 1 mM dithiothreitol. Labeling reactions were loaded directly onto the resin bed, and the eluate was collected in 1 mL fractions. Fractions were analyzed by 16% SDS-PAGE. Labeled protein and free label were detected with a Storm phosphoimager (GE Healthcare) using the 635 nm excitation for Alexa Fluor 647 and the 450 nm excitation for Alexa Fluor 488. Gels were subsequently stained with Coomassie to confirm the presence of protein.

Fractions containing clean, labeled protein were concentrated using a 10 kDa MWCO VivaSpin6 (GE Healthcare) spin column. The same spin column was also used to exchange protein into Qₑ buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 10% glycerol). Protein concentration was determined by Bradford assay, and labeled protein was stored at -80 °C. Labeling efficiency was determined by measuring the concentration of fluorophore by UV-Vis spectrophotometry and comparing with the protein concentration determined by Bradford assay. Protein labeling efficiency was above 75% for all variants.

3.2.3 Fluorescence Resonance Energy Transfer (FRET) assay of subunit exchange

The rate of UmuD₂ homodimer exchange was monitored by FRET as previously described (29) using equimolar amounts of A488- and A647-labeled UmuD at 1 µM in FRET buffer (20 mM HEPES pH 7.5, 100 mM NaCl) using λₑₓ = 493 nm and detecting fluorescence emission at 516 nm and 671 nm. To measure the rate of UmuDD’ heterodimer formation, an equimolar amount of UmuD’ S60A was added upon completion of the homodimer exchange reaction and
fluorescence emission was again measured at both 516 and 671 nm. The half-time ($t_{1/2}$) of subunit exchange was determined as previously described (29).

Figure 3.1 – Design of UmuD mono-cysteine variants and BMH cross-linking
(A) Model of UmuD (with arms down) showing arm residues 1-24 (red) removed upon cleavage to UmuD’. UmuD 3A mutations T14, L17, F18 (cyan), natural cysteine position C24 (magenta), and active site residue S60 (blue) are also highlighted. Locations of cysteine substitutions shown in green and yellow for contrast. (B) Distances between mono-cysteine positions measured in the model of full-length UmuD. (C) Cross-linking of UmuD variants using BMH. Reactions were resolved on 16% SDS-PAGE.

3.2.4 BMH Cross-linking and RecA:ssDNA-dependent cleavage assays

Cross-linking of the UmuD mono-cysteine variants with bis(maleimido)hexane (BMH, Pierce) at 40 µM protein and 2 mM BMH was carried out as previously described (28, 32), except proteins were visualized by Coomassie-stained SDS-PAGE. The RecA:ssDNA-dependent cleavage assays were also carried out as previously described (36, 38).

3.2.5 UV survival and Immunoblotting

Survival assays were performed as previously described (36, 37). Values represent the average of at least three trials, and the error bars show the standard deviation. Immunoblotting procedure was completed as previously described (38), with rabbit polyclonal anti-UmuD/UmuD’ antibodies (32). Band densities were determined using ImageQuantTL software (GE).

3.3 Results

3.3.1 Design of UmuD variants

Our previous work used a FRET assay to monitor UmuD exchange and showed that the UmuD₂ and UmuD’₂ homodimers undergo relatively slow, time-dependent exchange, occurring with a t₁/₂ of approximately 26 min (29). We found that the subunits of UmuD’₂ homodimers are slightly more readily exchangeable than the full-length subunits in UmuD₂ (29) and that the rate
of UmuDD’ heterodimer formation was similar to the rates of the homodimer exchange (29). In our previous investigation, variants UmuD S60A G92C and UmuD’ S60A G92C were fluorescently labeled at G92C. Conjugation of the fluorescent probe at the G92C position led to the intriguing observation that UmuD₂ dimers were resistant to heterodimer formation when mixed with unlabeled UmuD’ (29). Because the equilibrium dissociation constant for dimerization of UmuD₂ and UmuD’₂ is in the low-picomolar range (22), it is unlikely that UmuD and UmuD’ subunits exchange via dissociation into monomers. We hypothesize that subunit exchange proceeds through formation of higher-order structures, and that conjugation of the fluorescent dye molecules at the G92 position may interfere with a protein-protein interaction surface utilized in the subunit exchange mechanism. In this work, we further probed the mechanism of exchange by carrying out the FRET assay with a number of different labeled mono-cysteine UmuD variants. By monitoring changes in the rate of subunit exchange, we are able to identify positions involved in the subunit exchange mechanism.

Variants used for in vitro experiments including FRET, BMH cross-linking, and cleavage contain S60A and C24A mutations (Fig. 3.1A) in addition to their unique mono-cysteine substitution. Because cleavage can occur spontaneously during expression or purification of UmuD variants, non-cleavable S60A variants are routinely used to prevent cleavage. UmuD contains one natural cysteine at the 24 position of its N-terminal arm. Consequently, we also introduced the C24A mutation to prevent secondary labeling of the UmuD N-terminal arms. Previous work shows that this mutation does not significantly affect UmuD function (28, 39).
The positions tested here span various positions in the globular domain. UmuD residues D91 and G92 were predicted by molecular modeling to be involved in the interaction between UmuD and the α subunit of DNA polymerase III (pol III α) (6). These positions are located along the outer surface of the C-terminal globular domain (Fig. 1), and the sidechain of D91 is predicted to form a salt-bridge with R1068 of pol III α (6). By analogy, we predicted that D91 and neighboring acidic position E93 could be implicated in an interaction between UmuD subunits. Therefore, we designed single-cysteine variants at these positions.

A second pair of Asp/Glu residues are located in a loop neighboring D91/E93 (Fig. 3.1). The homology model of full-length UmuD₂ (32) and NMR solution structure of UmuD’₂ (26) show that E122 and D123 are found at the outer-most edge of the C-terminal globular domain. Therefore, these positions could make an ideal contact point for protein-protein interactions.

Additional cysteine substitutions were constructed at surface positions on different faces of the globular domain. Positions K98, I108, M110, Y114, and T118 line one edge of each monomer while I78 and T79 are part of a loop opposite the face that binds the N-terminal arm. Positions I78 and T79 were of particular interest based on previous investigation of UmuD dynamics by hydrogen-deuterium exchange (HDX) mass spectrometry. Residues I78 and T79 are part of a peptide (residues 73-88) that showed decreased deuterium uptake relative to the rest of the protein (3). This result was unexpected as most peptides showing protection are located at the dimer interface. Based on available structural data, residues 73-88 should be solvent-exposed. Thus, the observation that this region is protected from HDX could indicated a surface of UmuD involved in mediating higher-order structure.
Figure 3.2 – Kinetics of UmuD homodimer subunit exchange
Heat map of homodimer exchange kinetics plotted on (A) the homology model of full-length UmuD, (B) the NMR solution structure of UmuD’, and (C) the molecular dimer interface of UmuD’ in the X-ray crystal structure. Mono-cysteine positions used for conjugation of fluorescent dyes are shown by spheres. Coloring corresponds to relative speed of exchange kinetics where warm colors indicate fast exchange and cold colors indicate slow exchange. (D) Bar graph illustrating assignment of heat map color based on calculated $t_{1/2}$ of exchange for each UmuD variant. (E) UmuD mono-cysteine constructs and corresponding $t_{1/2}$ of homodimer exchange.
3.3.2 BMH cross-linking

We tested the efficiency of cross-linking each mono-cysteine variant using the homobifunctional cross-linker bismaleimidohexane (BMH). The thiol-reactive maleimides create a covalent linkage between reduced cysteine sidechains. The BMH molecule is 13 Å in length with maleimide functional groups at either end and therefore, for cross-linking to occur, cysteine thiols must be within 13 Å of one another. The distance between cysteine thiols of each mono-cysteine UmuD variant tested here ≥ 20 Å (Fig. 3.1B). Therefore, only interdimer cross-linking of UmuD subunits is expected to be detected by this assay. Cross-linking of C24 or A31C in the N-terminal arms can occur within the dimer if the arms are unbound from the globular domain.

The variant UmuD3A (39.6 ± 3.3%) is considered a UmuD’ mimic, but possesses full-length N-terminal arms. Three alanine mutation (T14A, L17A, F18A, Fig. 3.1A) prevent the arms from stably binding to the globular domain. Therefore, the arms of UmuD3A are highly dynamic (3) and cross-linking at the C24 position occurs readily. Because UmuD3A has the ability to cross-link both inter- and intra-molecularly with highly-dynamic arms, it represents maximal cross-linking in this assay. Surprisingly, mono-cysteine variants UmuD C24A S60A T79C (41.0 ± 3.0%) and UmuD C24A S60A E122C (33.2 ± 4.7%) cross-link as efficiently as UmuD3A, yet are expected only to be able to cross-link intermolecularly.

Cross-linking efficiency can also provide information about the accessibility of a cysteine thiol for modification. If a residue is buried, chemical conjugation is hindered. We believe that this is the case for UmuD C24A S60A K98C (6.9 ± 2.4%) which was unable to cross-link efficiently, and was difficult to label with fluorescent dyes. In fact, fluorescent labeling of UmuD C24A
S60A K98C was so inefficient that it was not used in the FRET exchange assay. While cross-linking efficiency is not necessarily a determinant of solvent accessibility, these observations indicate that K98C is sequestered within the globular domain.

There was no correlation between intramolecular distance and cross-linking efficiency, which is not surprising given that most likely cross-linking occurs in this case intermolecularly. Surprisingly, UmuD’ A31C exhibited the least efficient cross-linking (5.9 ± 2.1%). Indeed, investigation of this cysteine variant by EPR suggests that its N-terminal arms are sufficiently dynamic to allow intramolecular cross-linking (2). Moreover, positions 25-32 were not modeled in the X-ray crystal structure due to poor ordering in this region (25, 40).

3.3.3 UmuD variants exhibit different homodimer exchange kinetics

UmuD subunit exchange was measured using a FRET assay previously developed in our lab (29). Exchange was monitored by mixing UmuD labeled with Alexa Fluor 488 (A488) with UmuD labeled with Alexa Fluor 647 (A647). In this experiment, UmuD labeled with A488 acts as the donor fluorophore, and UmuD labeled with A647 is the acceptor. Emission by A647 depends on excitation by A488, and energy transfer only occurs when the fluorophores are in reasonably close proximity. Subunit exchange was monitored by both decrease in A488 and increase in A647 emission, which will be anti-correlated if FRET occurs.

Our lab previously showed that UmuD2 subunits undergo a slow, time-dependent exchange. The measured $t_{1/2}$ for UmuD C24A S60A G92C homodimer exchange was 26 min (29). In this work, UmuD homodimer exchange was observed for all mono-cysteine variants tested and the kinetics
of exchange were influenced by position of the fluorophore. All variants tested here except T79C exhibited faster exchange than the previously-characterized G92C.

We classified the rates of exchange into four categories. Three variants displayed fast exchange kinetics relative to other variants tested in this study (Fig. 3.2, red). UmuD C24A S60A I78C \( (t_{1/2} = 1.7 \pm 0.1 \text{ min}) \), UmuD C24A S60A M110C \( (t_{1/2} = 5.6 \pm 0.1 \text{ min}) \), and UmuD C24A S60A Y114C \( (t_{1/2} = 4.7 \pm 0.1 \text{ min}) \) are each located toward the interior of the UmuD\( _2 \) dimer, and the latter two positions are at one edge of the dimer interface, at the bottom in the perspective shown in Fig. 3.2.

Labeling positions near the interior of the C-terminal globular domain yielded intermediate-fast (Fig. 3.2, pink) exchange kinetics. Variants producing intermediate-fast exchange were UmuD C24A S60A E93C \( (t_{1/2} = 10.1 \pm 0.1 \text{ min}) \) and UmuD C24A S60A I108C \( (t_{1/2} = 9.5 \pm 0.1 \text{ min}) \). Labeling positions that exhibited intermediate-slow exchange kinetics (Fig. 3.2, light blue) include labeling positions D91C \( (t_{1/2} = 16.7 \pm 0.4 \text{ min}) \), T118C \( (t_{1/2} = 18.6 \pm 0.3 \text{ min}) \), E122C \( (t_{1/2} = 22.2 \pm 0.8 \text{ min}) \), and D123C \( (t_{1/2} = 14.6 \pm 0.1 \text{ min}) \). All of these positions are located on an outer edge of the C-terminal globular domain (Fig. 3.2).

Variant UmuD C24A S60A T79C exhibited the slowest exchange kinetics \( (t_{1/2} = 38.5 \pm 1.1 \text{ min}) \), which was also slower than the exchange of the previously-characterized G92C variant. Labeling at this position dramatically affected the rate of subunit exchange, and the reaction reached equilibrium only after 90 minutes. Intriguingly, the fastest and slowest exchange kinetics resulted from labeling at adjacent positions I78 and T79 (Fig. 3.2), respectively.
In general, labeling positions near the dimer interface exchange subunits more quickly, which could indicate that disruption of the dimer interface by the presence of the label facilitates exchange. Conversely, labeling positions toward the outside edge of each C-terminal domain experience slower exchange kinetics. Slower exchange kinetics are presumably a result of the fluorescent label obstructing a subunit association interaction surface utilized in the subunit exchange mechanism.

3.3.4 Heterodimer Exchange: UmuDD’ heterodimer is preferred; kinetics of exchange mirror homodimer

Previous work by us and others has shown that the UmuDD’ heterodimer is the predominant dimeric species after 20 min when equimolar amounts of purified UmuD₂ and UmuD’₂ are mixed (29, 30). Heterodimer formation is integral to UmuD regulation during the latter stages of the DNA damage response, as UmuD’ is degraded by ClpXP in the context of UmuDD’ heterodimer (41). Therefore, heterodimer formation allows ClpXP to specifically target UmuD’, and, in doing so, reduce levels of mutagenically-active UmuD’₂ (30).

Each mono-cysteine UmuD was tested for its ability to form a heterodimer with UmuD’. Upon completion of the homodimer exchange assay, UmuD₂ dimers were formed as determined by FRET. Next, an equimolar amount of unlabeled UmuD’ S60A was added to the reaction and heterodimer formation was observed as a decrease in FRET. The decrease in FRET intensity is the result of heterodimer formation as the UmuDD’ heterodimers contain only one fluorophore.
Our observations reveal that all the variants were able to form heterodimers preferentially, and the relative kinetics of heterodimer formation are similar to those observed during homodimer formation. UmuD variants labeled at the I78 and T79 positions displayed the fastest and slowest rates of heterodimer formation, respectively, with heterodimer formation of the UmuD C24A S60A T79C variant significantly slower than all other variants (Fig. 3.3).

**Figure 3.3 – Kinetics of UmuDD' heterodimer formation.**
(A) Half-lives (t_{1/2}) of exchange and corresponding heat map colors. (B) Graphical representation of heterodimer formation kinetics (t_{1/2}) illustrating choice of heat map color for each UmuD mono-cysteine variant.
Figure 3.4 – *In vivo* characterization of mono-cysteine UmuD variants in PB103. (A) and (B) UV survival of mono-cysteine UmuD variants. Error bars represent standard deviation of three independent experiments. (C) Western blot relative expression levels normalized to maximum expression measured (M110C).
3.3.5 UV survival: Single-cysteine UmuD variants confer resistance to UV

To test the effect of the single-cysteine variants on cell survival, we first performed site-directed mutagenesis on low-copy plasmids harboring the *umuDC* operon with a UV-inducible promoter to recreate the variants in suitable constructs for *in vivo* characterization. The plasmid constructs do not contain the S60A or C24A mutations used for *in vitro* characterization. The *E. coli* strain PB103 is hypersensitive to UV irradiation due to deletion of *umuDC* and *recJ* genes from the chromosome, but wild-type UmuD$_2$C from a plasmid is able to confer resistance to UV by complementation (Figs. 3.3A and 3.3B) (38, 42). As expected, plasmid-borne wild-type UmuDC (pGY9739) fully complemented, and all single-cysteine variants increased survival relative to empty vector (pGB2). Only three variants, D91C, Y114C, and D123C, produced a statistically-significant difference in survival at the 100 J/m$^2$ UV exposure compared to wild-type UmuDC and even these exhibited substantially increased survival relative to empty vector. Thus, all variants are able to complement the UV sensitivity of the Δ*umuDC* Δ*recJ* strain.

3.3.6 RecA:ssDNA Cleavage: UmuD single-cysteine variants retain cleavage activity *in vivo* and *in vitro*

Full-length UmuD is expressed as a 139-amino acid protein whose intracellular concentration is upregulated approximately 20 min after initiation of the SOS response. Initially, the full-length product of *umuD* is the predominant form during which time UmuD acts as a DNA polymerase manager allowing time for accurate DNA repair pathways to act (11-13). Upon interacting with the RecA:ssDNA nucleoprotein filament, UmuD undergoes an autocatalytic self-cleavage to remove its N-terminal 24-amino acids and forms UmuD' (43-45).
Figure 3.5 – Cleavage efficiency of UmuD mono-cysteine variants.

(A) RecA-mediated cleavage of unlabeled UmuD mono-cysteine variants. Reaction resolved by SDS-PAGE and UmuD bands detected by Coomassie stain. 

(B) RecA-mediated cleavage of UmuD mono-cysteine variants fluorescently labeled with A647. Purified UmuD mono-cysteine variants contain S60A and C24A mutations in addition to their specific cysteine substitution. Cleavage was performed by mixing with UmuD3A which has a competent active site, but uncleavable N-terminus. Reactions were resolved by SDS-PAGE and UmuD bands were detected using Storm phosphoimager at 635 nm excitation wavelength.

All single-cysteine variants were cleavable in vivo after UV irradiation, but with a wide range of cleavage efficiencies. UmuD Y114C showed the least cleavage (14%) while E122C cleaved the most (100%) with no detectable full-length protein 60 minutes after UV treatment. Cleavage and relative expression levels of UmuD I78C could not be determined because the protein runs high on SDS-PAGE which causes interference of full-length UmuD by the cross-reacting band. 

Steady-state expression levels of plasmid-borne UmuD variants in strain PB103 were measured.
by immunoblot assay (Fig. 3.4), and range between two-fold above to two-fold below that of wild-type UmuD.

Purified single-cysteine variants, which also harbor the C24A S60A mutations, were assayed for efficiency of RecA:ssDNA-mediated cleavage in vitro. In order to assay cleavage of the arms of the mono-cysteine variants, which also harbor the S60A mutation and are therefore non-cleavable, we needed to exchange subunits with a UmuD construct possessing a competent active site. The UmuD3A variant has three mutations in the N-terminal arm, T14A, L17A, F18A (Fig. 3.1A), that render the arms non-cleavable while maintaining an active site that is proficient for cleavage (32). Previous work has shown that the umuD gene products are able to undergo cleavage in both cis and trans orientations (31, 32, 38). Therefore, by first mixing each variant 1:1 with UmuD3A, the subunits exchange and we are able to assay for cleavage of mono-cysteine variants upon addition of the RecA:ssDNA nucleoprotein filament.

UmuD3A was able to cleave each of the mono-cysteine variants to a similar extent in trans (Fig. 3.5). We assayed purified, unlabeled mono-cysteine variants for cleavage (Fig. 3.5A) in addition to the same variants conjugated with A647 dye (Fig. 3.5B). Variants labeled with the A647 dye were chosen because of the increased steric bulk of A647 compared to A488; if dye molecules at a particular site were to interfere with cleavage, A647 would likely have a greater effect than A488. As was the case in vivo (Fig. 3.4C) and in vitro with unlabeled proteins (Fig. 3.5A), all of the variants were cleavable. Interestingly, the cleavage efficiencies of mono-cysteine variants were more uniform in vitro than in vivo. Therefore, our cysteine substitutions have little effect on the RecA:ssDNA-mediated cleavage reaction in vitro, but appear to be differentially affected by
intracellular factors that influence cleavage. In addition, steady-state expression levels of the variants vary by about five-fold, which could affect cleavage efficiency.

It should be noted that the difference in in vitro cleavage efficiencies of labeled and unlabeled mono-cysteine UmuD variants could be attributed to the method of detection. Cleavage reactions using A647-labeled UmuD variants were quantified by fluorescent detection. Therefore, only the labeled mono-cysteine variant is detected and the calculated percent cleaved is representative of only the mono-cysteine variant. The calculated percent cleaved for each unlabeled mono-cysteine UmuD is lower than its corresponding reaction with A647-labeled protein because in the former case the cleaved fraction is determined relative to total protein, including the variant being tested and UmuD3A, which are both detected by Coomassie staining. Additionally, the cleavage efficiency of wild-type UmuD alone (Figs. 3.5A and 3.5B) is greater than in all other reactions because each UmuD2 dimer contains two cleavable arms and two competent active sites as compared with reactions containing UmuD3A where only one arm is cleavable and only one active site is competent.

3.4 Discussion

In this study, we set out to expand our understanding of UmuD dimer exchange. We were able to demonstrate that conjugation of a fluorescent molecule at different positions on UmuD affects the kinetics of UmuD2 homodimer subunit exchange. In addition, we demonstrate that our variants preferentially form UmuDD’ heterodimers, and location of the dye molecule also influences the kinetics of heterodimer formation. Further, we demonstrate that each variant
retains the activities integral to UmuD function during the SOS response to DNA damage including RecA:ssDNA-mediated cleavage and resistance to UV irradiation.

Mono-cysteine derivatives of UmuD have previously been used to localize its dimer interface by cross-linking (28, 46). Cross-linking was performed using BMH, which has a 13-Å span and is functionalized with thiol-reactive maleimides at either end. The efficiency of cross-linking in our assay is primarily dependent on the following factors: (i) the reactivity of each cysteine position as a function of solvent-accessibility; (ii) the positioning of cysteine side chains on the UmuD structure which determines if UmuD protomers can cross-link intramolecularly, intermolecularly, or both; and (iii) the frequency that cysteine side chains are within 13 Å. To determine which variants may undergo intramolecular cross-linking by BMH, we calculated the distance between cysteine positions of each UmuD protomer (Fig. 3.1B) using the homology model of full-length UmuD (32). While none of the variant cysteine positions are within 13 Å of each other, previous work from our lab and others has shown that intramolecular cross-linking at Cys24 occurs readily using BMH due to the dynamic nature of the N-terminal arms (2, 3, 7, 28). Therefore, wild-type UmuD, UmuD3A, UmuD S60A, and UmuD' A31C, which each contain a single-cysteine in the arms, should be able to cross-link intramolecularly within the dimer and intermolecularly with protomers from other dimers. Moreover, the arms of UmuD3A are unable to associate with the C-terminal globular domain (2, 3, 32), which increases the probability of Cys24 residues being cross-linked by BMH. Surprisingly, the T79C variant (41 ± 3.0%), which is unable to cross-link intramolecularly, cross-linked as efficiently as UmuD3A (40 ± 3.3%). The relatively high cross-linking efficiency of E122C (33 ± 4.7%), which has the largest distance between cysteines (70 Å), was also unexpected. The E122C variant cross-linked more efficiently
than UmuD S60A (29 ± 2.9%), which is able to cross-link intramolecularly like UmuD3A. These observations suggest that the T79 and E122 positions of UmuD are frequently within 13 Å, which likely occurs through association of separate dimers or in an alternative dimer conformation.

To determine the solvent-accessibility and reactivity of each mono-cysteine variant, we measured the efficiency of A647 and A488 labeling reactions (Fig. 3.1B). Dye conjugation was performed using the same maleimide coupling reaction that is required for cross-linking with BMH. Therefore, fluorescent labeling efficiency is reliant on accessibility of each cysteine position. With the exception of K98C, all of the mono-cysteine variants were labeled at > 70% indicating that the cysteine side-chains are reactive and accessible. UmuD K98C was not amenable to fluorescent dye conjugation (percent labeled = 20% and 26% for A488 and A647, respectively), even though we are using a highly optimized labeling protocol for maximum labeling efficiency, nor did it cross-link well (percent cross-linked = 6.9 ± 2.4%). Together, these observations suggest that the K98C sulfhydryl group is not accessible for modification.

The *umuDC* operon encodes two proteins that together form the Y family polymerase DNA pol V (UmuD'zC) (15). Y family polymerases are characterized by their ability to perform TLS on damaged DNA templates, but exhibit low fidelity on undamaged DNA (47, 48). The activity of DNA pol V is tightly controlled at the transcriptional level by LexA and post-transcriptionally as part of a primitive DNA damage checkpoint (1, 11-13). Induction of the SOS response leads to an increase in UmuD levels from approximately 180 to 2400 copies per cell, while UmuC levels are approximately 12-fold lower (49). Previous work suggests that UmuD subunit exchange is a
mechanism cells use to manage DNA pol V activation through the formation of catalytically-inactive complexes, UmuD\textsubscript{2}C and UmuDD\textsubscript{C} (50). Indeed, LexA regulation of the \textit{umuDC} operon is not fully derepressed until approximately 15 min after DNA damage (51), and activated DNA pol V does not accumulate until \sim45 min after DNA damage (18).

In the present work, we expand upon our knowledge of UmuD\textsubscript{2} subunit exchange kinetics using a series of novel mono-cysteine UmuD variants (Fig. 3.1A). Variants T79C, E122C, T118C, D91C, and D123C exhibited intermediate-to-slow exchange kinetics (Fig. 3.2, light blue). The spatial positions of these variants are concentrated in a region of the C-terminal domain distal to the dimer interface. We also monitored the formation of heterodimer and found that the same five variants exhibited similar slow exchange kinetics. Indeed, each labeling position affected the kinetics of homodimer exchange and heterodimer formation similarly. This suggests that subunit exchange mechanism is mediated by the same interface for both species. However, for the T79C variant, the kinetics of heterodimer formation are faster than homodimer exchange and may indicate that this position is more important for the homodimer exchange mechanism.

Analysis of the X-ray crystal structure of UmuD\textsubscript{2} led to the identification of two potential dimer interfaces (25, 40). NMR and cross-linking experiments support the conclusion that the so-called filament dimer interface is prevalent in solution (10, 19, 25-28, 40, 52). However, evidence suggests that the molecular dimer interface may be implicated in the formation of higher-order oligomers (40). Indeed, the low-pM \( K_d \) of UmuD dimerization implies that subunit exchange does not occur via dissociation into monomers at physiologically-relevant concentrations (22).
has been suggested that subunit exchange proceeds through assembly of transient, higher-order structures in which the molecular interface mediates quaternary structure (29, 40).

Approximately 550 Å² of solvent-accessible surface area is buried per monomer in formation of the molecular dimer interface (40). The core of the interface is primarily stabilized by hydrophobic interactions with the exception of a pair of salt bridges between E93 and K55 residues. If the molecular interface is also the interface mediating subunit exchange, one would expect the variant E93C to exhibit slow exchange. Conversely, our observations show that labeling at this position results in intermediate-to-fast exchange. Furthermore, the five cysteine positions that exhibited slow exchange kinetics lie outside of the buried surface area. Considered together, the observations reported here do not support an exchange mechanism mediated by the molecular dimer interface; however, the molecular dimer interface may play a role facilitating DNA pol V association with the cell membrane prior to activation (53-55).
3.5 References


Chapter 4: Single-molecule mechanochemical characterization of *E. coli* pol III core catalytic activity

This chapter was recently published as:

"Single-molecule mechanochemical characterization of *E. coli* pol III core catalytic activity"

M. Nabuan Naufer performed the single-molecule DNA stretching experiments, analyzed the data, modeled the force dependence of pol and exo activities, and wrote the manuscript; David A. Murison synthesized the DNA constructs used for single-molecule experiments, purified the DNA pol III core protein complex, and performed primer-extension assays; Ioulia Rouzina carried out mathematical modeling; Penny J. Beuning and Mark C. Williams designed experiments and wrote the manuscript.
4.1 Introduction

Replicative DNA polymerases are responsible for duplicating chromosomal DNA, which carries a large amount of information during cell division, and hence must be replicated with high accuracy to sustain life. They catalyze the addition of deoxynucleoside triphosphate (dNTP) units to the DNA backbone in DNA replication. The addition of the dNTPs occurs directly on the DNA template strand, and the base of the new dNTP is complementary to the base on the template strand. Since bases are added to the 3′ end of the nascent strand, the polymerization reaction must proceed in the 5′ to 3′ direction. The tertiary structure of DNA polymerase is such that the enzyme fits over the previously formed base pairs (1, 2). These bases must be paired correctly for the polymerase to adopt its functional conformation (3-5).

DNA polymerase III is the replicative DNA polymerase in *E. coli* (6-11). It is an asymmetric dimer or trimer that synthesizes the leading and lagging strands simultaneously at the replication fork (6, 7, 12). A helicase unwinds the double-stranded DNA (dsDNA) into two anti-parallel template strands. After primase synthesizes the RNA primer, DNA pol III replicates the leading strand continuously and the lagging strand in Okazaki fragments in the 5′ to 3′ direction (1, 6-8).

Pol III is composed of 10 subunits that coordinate leading and lagging strand synthesis. The core of the polymerase (DNA pol III core) contains the polymerase subunit α, the proofreading exonuclease ε, and a subunit of unknown function θ (13, 14). The exact role of the θ subunit is still to be determined, but its presence increases the accuracy of DNA pol III and it has been suggested to stabilize the interaction between α and ε (15-17). Although the dimer is asymmetric (7, 18) to allow simultaneous polymerization of both template strands, the core of the each branch consists of the same α ε θ complex (13). The polymerase subunit α is a C-family DNA
polymerase, a family that is found only in prokaryotes (6). DNA pol III has an extremely high catalytic rate, at 10^3 bases/s (13, 19), and high fidelity, with error frequencies of approximately 10^{-5}/bp without proofreading (13, 20, 21) and 10^{-8}/bp with proofreading (13). In the presence of the β clamp, which tethers the core protein assembly to its DNA substrate, α exhibits very high processivity (13, 22, 23).

During replicative polymerization, tight coordination between the polymerization and exonucleolysis cycles is expected to exist, to permit efficient and faithful replication. It has been shown that mutations that lead to a loss in fidelity during E. coli replication are found in the dnaQ gene, which encodes the ε subunit (24-26). However, the molecular mechanism of the switching between the polymerase and exonuclease subunits is poorly understood. Elucidating the structure, function, and catalytic activity of these molecular motors is essential to understand the complex mechanisms of DNA replication. Here we report a single-molecule approach to manipulate these molecules and characterize the dynamics of DNA pol III core polymerization and exonucleolysis.

We observe that the mechanical tension applied to the substrate DNA promotes the switching between exonucleolysis and polymerization functions, which agrees with previous single molecule studies on DNA polymerases Klenow Fragment, T7 gp5, and φ29 DNA polymerase (27-29). The force dependence of T7 polymerization velocity is modeled as a function of the free energy change involved in ssDNA-dsDNA conversion. The kinetic scheme proposed for φ29 DNA polymerase describes the intra-molecular primer transfer as a consequence of a conformational change in the φ29 pol - DNA assembly induced by the applied tension on the
DNA template (28). The key difference between DNA pol III core and these polymerases is that the editing and polymerization activities of DNA pol III core are carried out by distinct subunits, ε and α, respectively. Hence the primer transfer between the catalytic exo and pol domains occurs intermolecularly. In addition, the exo activity of isolated ε is similar to that of DNA pol III core and ε is considered to be a highly efficient 3′-5′ exonuclease, capable of functioning independently of α (30, 31). The source of this exonucleolytic editing specificity is found to be the greater melting capacity of a mispaired 3′ terminus for both isolated ε and DNA pol III core (32, 33). According to previous bulk biochemical assays, the exonuclease activity of both ε and DNA pol III core is more efficient with ssDNA (32). Furthermore, a two-step kinetic scheme for the exonuclease reaction of isolated ε subunit suggests that the physiologically relevant substrate for the ε subunit within the holoenzyme complex is ssDNA at least three nucleotides in length (33).

Here, we investigate the force dependence of DNA pol III core polymerization and exonucleolysis. We are able for the first time to characterize these individual catalytic events on a single primer-template DNA substrate. We propose a two-state reaction scheme to describe the rate of force-induced exo initiation. According to our model, DNA pol III core bimolecularly binds at the primer-template junction and subsequently transforms to an exo-active conformation that is strongly affected by the applied template force. We show that this result is in quantitative agreement with the previously measured temperature-dependence of exo-activity. This analysis shows that the intermolecular switching of the primer between the polymerase and exonuclease subunits is a thermally driven process governed by destabilization of the primer-template junction, rendering it more susceptible to exonuclease binding.
Figure 4.1 – Schematic depiction of the single molecule experimental procedure using optical tweezers (not to scale).

(A) Linearized pBACgus11 is ligated with a DIG-dsDNA handle and a biotinylated oligonucleotide at its free ends, providing a single primer-template junction for DNA pol III core to bind. At forces above 30 pN, exonucleolysis is observed. Conversion from dsDNA to ssDNA upon the excision of nucleotides is registered as an increase in extension. Similarly, at forces below 30 pN incorporation of nucleotides due to polymerization is registered as a decrease in extension. (B) Force-extension curves of dsDNA (blue) and ssDNA (red). Stretch and release are represented as solid and empty circles, respectively. Solid lines represent the theoretical polymer models: extensible wormlike chain (34) for dsDNA and extensible freely jointed chain (35) for ssDNA. The arrows show the direction of extension change of exonucleolysis and polymerization at constant force experiments.
4.2 Materials and Methods

4.2.1 DNA pol III core expression, purification, and biochemical analysis

Wild-type DNA pol III core was expressed from the plasmid pET16b-dnaE-holEH dnaQ, which features a His-tag on the 0 subunit (a generous gift from Mark Sutton, Univ. at Buffalo), as described (36). Core was purified from a cell pellet harvested from 1 L of culture and stored at -80 °C. The cells were thawed on ice and lysed by sonication. Clarification was carried out by centrifugation at 12000 x g for 1 h at 4 °C. The supernatant containing soluble proteins was passed through a 0.45 μm filter before loading on a 5-mL His-Trap HP column (GE Healthcare) equilibrated with buffer HisA [20 mM HEPES; 500 mM NaCl; 50 mM imidazole; 10% glycerol, pH 7.5]. Bound protein complex was eluted using buffer HisB [20 mM HEPES; 500 mM NaCl; 300 mM imidazole; 10% glycerol, pH 7.5]. Fractions containing the desired protein were pooled and diluted 10-fold using buffer HeparinA [50 mM HEPES (pH 7.5); 0.1 mM EDTA; 10% glycerol; 1mM DTT] before loading onto a 5-mL Hi-Trap Heparin HP column (GE Healthcare) equilibrated with HeparinA. Bound proteins were eluted by addition of buffer HeparinA + 1 M NaCl in a linear gradient. Fractions containing intact core complex were pooled and diluted 6-fold with buffer HydroxyA [50 mM HEPES (pH 7.5); 150 mM NaCl; 1 mM DTT; 10% glycerol] and loaded onto a 5-mL hydroxyapatite column (BioRad Bioscale Mini CHT Type 1, 5 ml, 40 mm cartridge) equilibrated with buffer HydroxyA. Bound protein was eluted with buffer HydroxyB [200 mM sodium phosphate (pH 6.5); 150 mM NaCl; 1 mM DTT; 10% glycerol] in a step gradient. Fractions containing protein complex were pooled and dialyzed overnight at 4 °C against 2 L of storage buffer [30 mM HEPES (pH 7.5); 100 mM NaCl; 0.5 mM EDTA; 2 mM
DTT; 20% glycerol]. Protein purity was determined by SDS-PAGE, proteins were quantified by Bradford assay, and purified complex was stored at -80 °C.

Primer extension assays were carried out as described previously (37) using 32P-labeled primers annealed to 61-mer template. Reactions contained a final concentration of 25 nM DNA polymerase, 100 nM primer/template DNA, 100 µM dNTPs, 7.5 mM MgSO4, 30 mM HEPES (pH 7.5), 20 mM NaCl, 2 mM DTT, 1% (w/v) bovine serum albumin, and 4% glycerol. Reaction products were separated by denaturing 16% polyacrylamide gel electrophoresis and analyzed by phosphorimaging. The template sequence is 5′-ggttactcagatcaggcctgcgaagacctgggcgtccggctgcagctgtactatcatatgc; the primer sequences are Match: 5′- gcatatgatagtacagctgcagccggacgcc; MmT: 5′- gcatatgatagtacagctgcagccggacgcct; MmTC: 5′- gcatatgatagtacagctgcagccggacgcctc; MnTT: 5′- gcatatgatagtacagctgcagcggacgcctt; MmTTT: 5′-gcatatgatagtacagctgcagccggacgccttt.

4.2.2 Single molecule DNA constructs
Either a 38.5-kbp λ DNA or an 8.1 kbp pBacgus11 DNA were used in the single molecule stretching experiments. The 48.5 kbp linear λ DNA (Roche) with 12-nt 5′ overhangs at both the termini was digested with ApaI (New England Biolabs, NEB). The 5′ overhang of the resultant 38.5 kbp substrate was filled-in with KF in the presence of dGTP, dATP, biotin-14-dATP, and biotin-14-dCTP (NEB). At the opposite end a biotinylated oligonucleotide (5′-CTCTCTCTCTCTCTCTCTCTCTCTTTGGCC-3′, Integrated DNA Technologies, IDT) consisting a 3′ end complementary sequence to the ApaI-digested site was ligated with T4 DNA ligase. The 8.1 kbp construct was created by first linearizing the pBacgus11 (a gift from Borja
Ibarra) dsDNA vector (8041 bp) with BamHI and SacI (NEB) through dual digestion. Digoxigenin (DIG) labeled dsDNA handles with a complementary sticky end to the BamHI sequence were generated as described (28). A biotinylated oligonucleotide (5′-CTCTCTCTCTCTCTCTCTCTCTGGCCAGCT-3′, IDT) with a 3′ end complementary to SacI sequence, and the dsDNA DIG-handles were then ligated to their complementary positions at the linearized pBacgus11 DNA using T4 DNA ligase.

4.2.3 Single molecule optical tweezers experiments

We used optical tweezers to induce tension in single DNA molecules and thereby facilitate the DNA pol III core activity. Here, a single DNA molecule was attached by its labeled ends to derivatized polystyrene spheres. The 38.5 kbp λ DNA biotinylated at both the termini or 8.1 kbp pBACgus11 DNA, with biotinylated and DIG handles ligated at its respective termini, were tethered at the ends with streptavidin or streptavidin and anti-digoxigenin coated beads. One bead was immobilized by a glass micropipette attached to a flow cell while the other was held in a dual beam optical trap. By moving the glass micropipette attached to the flow cell, the DNA molecule was stretched and the force required to extend the DNA molecule was measured. The solution surrounding a single DNA molecule was replaced with DNA pol III core (20, 50, 100, 200 and 500 nM) diluted in the reaction buffer, 50 mM Hepes at pH 7.5, 25 mM Na+, 10 mM MgCl₂, 5 mM DTT and 1% BSA. In addition, 0.3 mM (each) dNTP were added to the experiments with dNTP at 0.2 μM DNA pol III core. For some experiments, the exo activity of T7 DNA polymerase (NEB) was used to initially create a partial ssDNA substrate, then exchanged for DNA pol III core. Data were collected at constant forces at 25 Hz, in which a detected change in the tension of the DNA substrate is compensated with a change in extension.
via a force feedback loop. The conversion between dsDNA-ssDNA upon exonucleolysis or polymerization, at constant DNA tensions, is registered as a change in extension as a function of time.

4.2.4 Single molecule data analysis

The extension-time trajectory was filtered with a moving average window of 8 Hz. The change in extension was converted to number of replicated or excised nucleotides by dividing the observed distance change by the expected change in extension at a given force accompanying the event of a single nucleotide incorporation. Theoretical polymer models, extensible worm like chain (34) for dsDNA, and extensible freely jointed chain (35) for ssDNA, were used to calculate the expected change in extension at a given force. Polymerization or exonucleolysis rate distributions were obtained from the moving average trace and fit to a bimodal Gaussian to find the instantaneous catalytic rate. Pauses were captured in the filtered trace by setting the consecutive events less than a cutoff of \( \frac{5}{\sqrt{w}} \) nt, to their mean values using a custom MATLAB code. Here 5 nt is the experimental noise and w is the number of data points included within the chosen window. We tested our algorithm with simulated data and the accuracy of recovered results were >90%, when the random noise level was set as 5 nt.

4.3 Results

4.3.1 DNA pol III core activity at constant force

We used optical tweezers to characterize the dynamics of DNA pol III core activity at the single molecule level. Both polymerization and exonuclease activity were measured at constant applied tensions on a single DNA substrate. To do this, a single dsDNA molecule with a 3’ recessed end
(~30 nt) was attached by its covalently-labeled free ends to polystyrene spheres, one held in an optical trap and the other immobilized on the end of a glass micropipette (Fig. 4.1A). By gradually moving the micropipette, the applied mechanical tension and the extension of a single DNA molecule was measured. In the absence of protein, an approximately constant force phase transition, referred to as DNA overstretching, is observed (38). At the low salt conditions used in these experiments, this transition occurs at about 62 pN and represents a conversion of DNA from dsDNA to ssDNA as the DNA is destabilized by force and primarily peels from its free end (39-44). The force-extension profiles of ssDNA and dsDNA cross at ~6 pN (Fig. 4.1B). DNA pol III core activity was measured at forces greater than this crossover force, at which ssDNA is longer than dsDNA. Therefore, at constant forces below the melting transition, conversion between ssDNA and dsDNA is registered as an increase in extension due to DNA pol III exonucleolysis (exo) and a decrease in extension due to DNA pol III core polymerization (pol) activity (Fig 4.1B). To measure DNA pol III core activity, we introduced purified DNA pol III core at fixed concentration to the flow cell containing a single DNA molecule captured between the beads as shown (Fig. 4.1A). The DNA was maintained at a constant force through a feedback loop and the change in DNA length at constant force was determined as a function of time. The number of nucleotides incorporated or excised as a function of time was then obtained by dividing the observed change in extension by the expected change in extension at a given force accompanying the conversion of one single-stranded nucleotide into its double-stranded counterpart, as described in Materials and Methods.
Figure 4.2 – DNA pol III core force-extension curves and temporal trajectories

(A) Force-extension curves of a bare dsDNA molecule (black) and a partially single-stranded dsDNA molecule (red) in the presence of 0.2 μM DNA pol III core. Changes in extension due to dsDNA-ssDNA conversion at different constant forces are shown in the multicolor profile. The arrows represent the direction of the pol III core velocity, indicating exonucleolysis or polymerization. (B) Temporal trajectories of the change in extension corresponding to the shown colors in (A) at the given constant forces (red).

As shown in Fig. 4.2, polymerization is observed at forces below 30 pN and exonuclease activity is observed at forces higher than 30 pN. From the measured extension vs. time trajectories, we obtained a distribution of instantaneous velocities for both exo and pol activity at several forces by applying a moving average filter as described in Materials and Methods. An example of a velocity distribution for exo activity at 50 pN is shown in Fig. 4.3. Each distribution was fit as a
sum of Gaussian functions, one at zero velocity and the other at finite velocity. The fit at finite velocity represents the distribution of instantaneous velocities characteristic of enzyme catalytic activity, while the zero-velocity distribution represents instrument noise as well as protein fluctuations, and these data are not included in the velocity analysis. The average instantaneous velocities given by the mean of the Gaussian were obtained for each trajectory and averaged over at least three trajectories for all forces, and these averages for both exo and pol activity are shown in Fig. 4.4. The instantaneous velocities are sensitive to the template tension, suggesting that the rate-limiting step is force-dependent. In addition, the sharp transition from pol to exo activity as a function of force also suggests that switching between these functions in the absence of force is expected to be thermally driven and the initiation of an exo event is facilitated by force.

4.3.2 Force-dependent polymerization and exonucleolysis velocities

The instantaneous velocities of T7 and Klenow Fragment DNA polymerases were previously modeled (27, 29) primarily attributing the force dependence to the activation enthalpy of converting n bases from the ss to the ds geometry. In that model, the projections of DNA segments along the direction of applied external force were determined “globally” by the change in the extension of ds and ss DNA extension as a function of force, averaged over thousands of bases. Therefore, that model does not distinguish the local orientations in the DNA segments inside the active site from the DNA segments further away from the active site. Consequently, Goel et al. (45, 46) suggested a “local” model attributing the force dependence of polymerization velocity to the orientation change of two DNA segments neighboring the active site. It was also shown that large conformational changes in the pol-template complex and a conserved active-site
geometry that induce a sharp kink at the 5′ end of the template during a catalytic pol event is a universal property shared by three families of polymerases. This model was further modified in Andricioaei et al. (47) by restricting the orientations of these local DNA segments (Restricted-Cone Local Model, RCLM) due to steric effects that were determined using molecular dynamic simulations on the *Thermus aquaticus* (Taq) DNA polymerase I complex. In this model, the instantaneous polymerization velocity ($v(F)$) is given by

$$v(F) = v_0 e^{-\Delta G(F)/k_BT},$$  \hspace{1cm} (1)

where $\Delta G$ is the force-dependent free energy contribution determined by the additional enthalpy and change in entropy associated with converting the two DNA segments from their “open” to “closed” forms, in the presence of force. For a given DNA segment $d$, the free energy contribution $\Delta G_d$ is given by

$$\Delta G_d(F) = -\int_0^F \delta\langle \cos \theta \rangle dF'.$$ \hspace{1cm} (2)

Here $\langle \cos \theta \rangle$ is the average angular orientation of a given DNA segment along the direction of force and is given by

$$\langle \cos \theta \rangle = \frac{\cos \theta_m e^{\xi \cos \theta_m} - \cos \theta M e^{\xi \cos \theta_M}}{e^{\xi \cos \theta_m} - e^{\xi \cos \theta_M}} - \frac{1}{\xi}.$$ \hspace{1cm} (3)
Figure 4.3 – Representative bimodal Gaussian distribution of DNA pol III core
Representative bimodal Gaussian distribution of DNA pol III core (0.2 μM) instantaneous velocities at 50 pN template tension determined from the moving average filter (yellow) shown in the inset. Zero (dashed blue) and nonzero (dashed red)-peaked Gaussians represent the paused states and the moving states of DNA pol III core, respectively. The peak position of the non-zero distribution represents the DNA pol III core mean instantaneous velocity at 50 pN. Inset: Representative temporal trajectory (magnified) of the moving and paused states of DNA pol III core. The pause detection (black, see Materials and Methods) is done using the moving average (yellow) of the extension data (blue).

The subscripts m and M refer to the minimum and maximum values used in RCLM, δ is the length of a given DNA segment, and $\xi = F\delta/k_BT$. Here $v_0$ is the zero-force instantaneous velocity, $kB$ is the Boltzmann constant, and $T$ is the absolute temperature. The reported $\theta$ and $\delta$ values for these two segments in the closed and open conformation in Andricioaei et al. (47) were used in the force ranges less than and greater than 9 pN. Although this model used Taq
DNA pol I, which is an A family polymerase, and *E. coli* DNA pol III alpha is a C family polymerase, another C family polymerase Taq DNA pol III alpha was also shown to exhibit a bend of the template at the active site (48), suggesting that the basic features of this model are applicable in the present case.

The best fit of our data to the model of Andricioaei *et al.* (47) yields $v_{0,pol} = 84.8 \pm 5.7$ nt/s (Fig. 4.4). DNA pol III core has a high velocity in the context of the complete replisome (13); however, the rate of replication of DNA pol III core alone, which is the relevant comparison here, was measured in bulk biochemical experiments to be 20 nt/s (49). This significant difference in the velocities with and without the β clamp is likely due to the weaker association of DNA pol III core with the substrate DNA in the absence of the β clamp. In contrast to single molecule assays, in bulk biochemical assays the catalytic rates are averaged over the paused states as well. Hence it is not surprising that the zero-force velocity ($v_{0,pol} = 84.8 \pm 5.7$ nt/s) in our measurements is higher than the 20 nt/s rate observed in bulk biochemical experiments.

In the case of exo activity, there is not a similar previously applied model, so we will initially assume a simple exponential dependence on force given by

$$v(F) = v_0 e^{-Fd/k_BT}.$$  \hspace{1cm} (4)

Here $v_0$ is the zero force exo velocity and $d$ is the force-independent length change required for each exonucleolysis event. The best fit yields, $v_0 = -20 \pm 5$ nt/s, $d = 0.11 \pm 0.02$ nm with dNTPs and $v_0 = -15 \pm 10$ nt/s, $d = 0.14 \pm 0.05$ nm in the absence of dNTPs. The value of $d$ reflects an elongation of DNA that occurs during each exo event, which is likely the slightly extended state.
of the terminal base pair when it is positioned for cleavage during the exo rate-limiting step. The fact that $d$ (in both cases) is slightly less than the total change in DNA length (~0.22 nm/bp) during an exo event supports this hypothesis.

**Figure 4.4 – Force dependence of the DNA pol III core instantaneous velocities.** Positive velocities represent polymerization and negative velocities represent exonucleolysis in the presence (red) and absence (blue) of dNTPs. The force dependence of polymerization is fitted to the RCLM (47) (black line) as described in the text, yielding $84.8 \pm 5.7$ nt/s as the zero-force velocity ($v_{0,\text{pol}}$). The exo force dependence is modeled as a simple exponential function of force, which yields $v_{0,\text{exo}} = -20 \pm 5$ nt/s, $d = 0.11 \pm 0.02$ with dNTPs (negative red line) and $v_{0,\text{exo}} = -15 \pm 10$ nt/s, $d = 0.14 \pm 0.05$ nm without dNTPs (blue line), where $d$ is the force-independent length change required for each exonucleolysis event. Error bars are standard errors of at least three independent measurements and uncertainties in the fitting parameters are from the standard deviation of the $\chi^2$-minimized fit.
Figure 4.5 – Concentration dependence of DNA pol III exonucleolysis. (A) Concentration dependence of pause (blue) and dwell (green) times during exo. Average pause time is significantly decreased with concentration. Dwell times do not exhibit a significant dependence on concentration. (B) Exo initiation rates, \( k_{\text{init}}(C) \) as a function of DNA pol III core concentration. \( k_{\text{init}}(C) \) is the reciprocal of the pause times shown in (A) and fitted to the two-state model as described in the text (Eq. 6). The best fit yields the dissociation constant \( K_d = 0.13 \pm 0.07 \mu \text{M} \) and transition rate to exo-active state \( k_2 = 11.3 \pm 3.1 \text{ s}^{-1} \). Error bars are standard errors of at least three independent measurements and uncertainties in the fitting parameters are from the standard deviation of the \( \chi^2 \)-minimized fit.

Other alternative models, such as the model that describes \( \phi 29 \) pol and exo activity (28), are not consistent with our data. For \( \phi 29 \), the total velocity was presented as a sum of exo and pol activity, such that fluctuations between the states determined the total velocity. However, we do not observe a significant change in exo velocity in the presence and absence of dNTP, showing that pol activity is not present at forces that primarily induce exo activity. Also, as is shown below, in contrast to \( \phi 29 \), the time that DNA pol III core spends on the DNA is much shorter than the switching time between pol and exo activities, such that in each processive event only one process is observed. Because pol and exo activity are accomplished by two separate proteins...
for DNA pol III core, and ε by itself is an independent ssDNA nuclease (33), it is not surprising that these activities are so well separated.

### 4.3.3 Concentration-dependent pause times during exonucleolysis

The time spent during an exo event is measured as the dwell time ($\tau_d$). The time spent between two consecutive exo events is measured as the pause time, determined from the pause-detecting trajectory shown in the inset of Fig. 4.3. The reciprocal of the average $\tau_p$ is the exo initiation rate ($k_{\text{init}}$) at a given force and concentration. The resulting concentration dependence of the pause and dwell times is given in Fig. 4.5A. Although $k_{\text{init}}$ (Fig. 4.5B) initially increases with concentration, this rate saturates at high concentrations. To describe this concentration dependence, we propose the following kinetic scheme.

$$
\frac{E+\text{DNA}_n}{k_{-1}} \xrightarrow{k_1} \text{E} \text{•DNA}_n \xrightarrow{k_2} \text{E} \text{•DNA}_n^* \xrightarrow{k_{\text{exo}}} \text{E} \text{•DNA}_{n-1}^*
$$

(5)

Here E is DNA pol III core in the solution, C is the DNA pol III core concentration, and DNA$_n$ is the substrate that is the primer-template terminus. The first step is the bimolecular DNA pol III core binding, which is in pre-equilibrium to the subsequent exo-active E•DNA$^*_n$ state. The rate $k_2$ is the rate of DNA pol III core transition to its exo-active state, and $k_{-2}$ is the rate of exo activity termination, due to dissociation that is determined by the dwell time measured below. As suggested by our measured force-dependence of that rate (also discussed below) as well as by its strong temperature dependence measured in the previous studies (32, 33), exo initiation is rate-limited by the requirement for destabilization of 2-3 bp at the primer terminus. Based on the DNA pol III core structure (50), such destabilization is most likely followed by strand transfer of
the 3' end of the primer from the pol to exo catalytic site. This step is not accompanied by a net change in DNA construct length, and is therefore not directly observed in our experiment. However, conventional biochemical measurements of the exo initiation rates (31-33) also suggest that DNA duplex destabilization is the rate-limiting step of the process, supporting the notion that the subsequent strand transfer between the pol and exo sites occurs rapidly.
Figure 4.6 – Force-dependence of DNA pol III core exonucleolysis

(A) Force-dependence of average pause times (blue) and dwell times (green) during exo at 0.2 μM DNA pol III core. Average pause times decrease with force. Average dwell times (green) exhibit insignificant force dependence. (B) Exo initiation rates, $k_{\text{init}}(F)$ as a function of force. The rate $k_{\text{init}}(F)$ is the reciprocal of the average pause times shown in (A) and is fitted to an exponential function of force as described in the text (Eq. 8), which yields $n_{\text{exo,init}}$ to be $1.84 \pm 0.20$ bp and $k_F$ to be $17.3 \pm 1.3$ s$^{-1}$. Here $n_{\text{exo,init}}$ is the minimum number of base pairs required to melt at the primer-template junction in order for the bound-DNA pol III core to transform to the exo-competent conformation. $k_F$ is the force-independent attempt rate at the given concentration. Error bars are standard errors of at least three independent measurements and uncertainties in the fitting parameters are from the standard deviation of the $\chi^2$-minimized fit. (C) Exonuclease activity of DNA pol III core complex is stimulated by base-pair mismatches at the primer-template junction. Fully-extended polymerization products also decrease when 1, 2, or 3 non-complementary bases are present at the junction. Reactions were quenched after 0, 1, 2.5, 5, and 10 min and analyzed by 12% denaturing polyacrylamide gel electrophoresis.

The next step is the catalytic exo activity that transforms $E\cdot\text{DNA}_n^*$ to $E\cdot\text{DNA}_{n-1}^*$. This process ($k_{\text{exo}}$) is fast, and is not observed in the kinetics of exo-initiation. This is because our analysis decouples the paused states from the moving states that are measured as a velocity ($v_{\text{exo}}$), which ranges from 20 to 100 nt/s, depending on the force, as described in the previous section. For the proposed reaction scheme (Eq. 5), the predicted exo initiation rate ($k_{\text{init}}$) is given by,

$$k_{\text{init}}(C) = \frac{k_C}{k_C + k_i}k_2 = \frac{1}{1 + K_d/C}k_2,$$

(6)

where,

$$K_d = \frac{k_i}{k_C}$$

(7)
is the equilibrium dissociation constant for DNA pol III core binding to the substrate. Note that because $k_{\text{init}}$ is the reciprocal of the average pause time ($1/\tau_p$) for a single DNA pol III core molecule to rebind to the substrate, this explicitly represents the on rate of an exo-competent state and therefore, $k_{-2}$ is disregarded in the equation. The best fit to the observed dependence of $k_{\text{init}}$ on concentration (Fig. 5B), yields $K_d$ to be $0.13 \pm 0.07$ µM and $k_2$ to be $11.3 \pm 3.1$ s$^{-1}$.

Eq. 6 assumes that the first step of DNA pol III core binding to DNA occurs in pre-equilibrium to the slower catalysis initiation step $k_2$, $(k_{-1} \gg k_2)$. According to our fitted values of $k_2$ ($11.3 \pm 3.1$ s$^{-1}$) and $K_d$ ($0.13 \pm 0.07$ µM), the bimolecular association rate $k_1 (k_{-1}/K_d)$ is much higher than $k_2$ ($k_1 > 10^8$ M$^{-1}$ s$^{-1}$), which is on the order of the diffusion rate ($10^9$ M$^{-1}$ s$^{-1}$). Thus, initial DNA pol III core binding to the primer-template junction is a nonspecific diffusion-limited process, leading to the slower step of catalysis initiation, which subsequently results in either pol or exo activity, depending on the stability of the primer-template junction.

### 4.3.4 Force-dependent pause times during exonucleolysis

We observe a significant increase in the exo initiation rate $k_{\text{init}}$ with increasing template tension (Fig. 4.6). The data in Fig. 4.6 is measured at 0.2 µM DNA pol III core concentration and therefore reflect the protein-saturated value of $k_{\text{init}}$ that is $\sim k_2$. Thus, the observed force-dependence of $k_{\text{init}}$ primarily corresponds to the rate at which the bound protein-DNA complex transforms into the exo-active state. It has been shown that destabilization of the primer-template junction increases the susceptibility of DNA to the exonuclease activity of DNA pol III core and its exo subunit ε (30, 32, 33). Because template tension uniformly destabilizes all base pairs of stretched DNA (51), the force dependence of $k_{\text{init}}$ is likely due to the increased probability of the duplex fraying, which can be described as (51)
\[ k_{\text{init}} = k_F e^{\text{exo,init} \left( \Delta G^\ast(F) - \Delta G^\ast_0 \right) / k_B T} \]  

(8)

Where,

\[ \Delta G^\ast(F) = \int_0^F x_{ss}(F')dF - x_{ds}(F')dF'. \]  

(9)

Here \( x_{ss} \) and \( x_{ds} \) are the ss and ds DNA extensions respectively (Fig. 4.1). The rate \( k_F \) is the maximum exo rate on completely destabilized dsDNA or on ssDNA. This maximum exo rate is expected to be reached at the melting force \( F_m \) (62 pN), at which the work performed by force to destabilize dsDNA, \( \Delta G^\ast(F) \), is equal to the free energy of bp melting in the absence of force \( \Delta G^\ast_0 \), . 2.23 \( k_B T \) (52). Best fit yields \( n_{\text{exo,init}} \) to be 1.84 ± 0.20 bp and \( k_F \) to be 17.3 ± 1.3 s\(^{-1}\) \( \mu \text{M}^{-1} \). This value for \( n_{\text{exo,init}} \) is very similar to the number of mismatches required for optimal exo activity by the subunit \( \varepsilon \) observed in our bulk primer extension assay for DNA pol III core, presented in Fig. 4.6, and previous biochemical assays using isolated \( \varepsilon \) (33).

### 4.3.5 Pauses during polymerization

The observed pauses between consecutive catalytic pol bursts appear to be unaffected by the applied force within the accuracy of our measurement (Fig. 4.7). The weighted average of the pause time over all forces during pol activity is 0.90 ± 0.22 s. This suggests that, in contrast to that observed for exo, initiation of a pol event is force-independent.
Figure 4.7 – Force-dependence of DNA pol III core polymerization and processivity

(A) Force-dependence of average pause times (red) and dwell times (blue) during pol at 0.2 μM pol III core. Average pause and dwell times are independent of force during polymerization. The weighted average over all the forces of dwell time (dashed blue) is 0.20 ± 0.05 s and pause time (dashed red) is 0.90 ± 0.22 s shown by the dashed line. Error bars are standard errors of at least three independent measurements.

(B) Average number of nucleotides (N) polymerized (red) or excised (blue) per single burst of catalytic activity of DNA pol III core as a function of force. N was calculated as a product of measured velocities and dwell times of pol and exo activities, respectively. The red triangle is the N determined from the zero-force velocity predicted from the RCLM (Fig. 4) and the red square is the zero-force measurement reported in a previous biochemical study (49).
4.3.6 Dwell times for polymerization and exonucleolysis

The observed catalytic bursts during both pol (Fig. 4.7) and exo (Figs. 4.5-6) occur at short time scales ($\tau_{d,\text{pol}} = 0.21 \pm 0.05 \text{ s}$, $\tau_{d,\text{exo}} = 0.15 \pm 0.05 \text{ s}$) that are independent of applied force or protein concentration (data not shown for pol). The reciprocal of the measured dwell time represents the termination rate of pol or exo activity ($k_{\text{off}} = 1/\tau_d = 5-7 \text{ s}^{-1}$) that is orders of magnitude higher than that observed for the DNA pol III core complex in the presence of the $\beta$ clamp (13, 22, 23). This significant difference in termination rates may account for the much weaker association of DNA pol III core with the primer-template junction in the absence of the $\beta$ clamp. Because the observed $k_{\text{init}}$ (C) (Fig. 4.5) for exo increases and saturates at high concentrations, this termination of catalysis primarily represents protein dissociation from the active conformation rather than intrinsic pausing during catalysis. Overall, this suggests that each exo event is associated with bimolecular protein binding to the primer-template junction, which is followed by a complex transition to its exo-active conformation.

4.3.7 Number of processively catalyzed base pairs during polymerization or exonucleolysis

Fig. 4.7B shows the number of nucleotides polymerized or excised per pol or exo activity burst, obtained as a product of the velocity (Fig. 4.4) and dwell time (Figs. 4.5, 4.7A). The number of catalyzed nucleotides vanishes at ~25 pN for both exo and pol, indicating a transition from pol to exo. Moreover, from the predicted zero-force pol velocity (Fig. 4.4) and the measured dwell time (Fig. 4.7), we can estimate the zero-force value of the processivity to be ~18 nt. This is consistent with previously measured bulk studies, in which DNA pol III core was shown to incorporate ~20 nucleotides before dissociating from the primer-template junction.
Figure 4.8 – Temperature-dependence of DNA pol III core activity

(A) Temperature-dependence of the exo initiation rate as reported in Brenowitz et al (32). Red circles are the reported exo initiation rates on a G-C paired primer-template junction, scaled by a factor of 0.02 to match the experimental conditions in this study (see discussion). The reported data is fit to an Arrhenius function as described in Eq. 11. The best fit (red line) yields \( n \Delta H_0 = 47 \) \( k_B T \). The dashed blue line at 8.1 s\(^{-1}\) is the maximum rate reported, which is observed on ssDNA at 310 K. (B) \( k_{\text{init}} \) determined from force- (blue) and temperature- (red) (32) dependent measurements as a function of total free energy \( \Delta G_{\text{tot}}(F,T) \) required to destabilize terminal dsDNA. The global fit is from the expression \( k_0 \exp[-n_{\text{exo,init}} \Delta G_{\text{tot}} / k_B T] / (1 + \exp[-n_{\text{exo,init}} \Delta G_{\text{tot}} / k_B T] \), that describes the probability of destabilizing \( n_{\text{exo,init}} \) terminal base pairs as a function of \( \Delta G_{\text{tot}}(F,T) \). Here \( \Delta G_{\text{tot}}(T) = \Delta H_0 - T \Delta S_0 \), where \( \Delta H_0 = 17 \) \( k_B T \) and \( T \Delta S_0 = 0.0471 T \) \( k_B T \) as described in the text. \( \Delta G_{\text{tot}}(F) \) is given by Eq. 8. The best fit yields \( n_{\text{init,exo}} = 2.16 \pm 0.13 \) bp and \( k_0 = 27.9 \pm 1.9 \) s\(^{-1}\).

4.4 Discussion

The force dependence of the catalytic functions of polymerases from bacteriophage \( \phi 29 \) and T7 as well as the \( E. \ coli \) polymerase Klenow Fragment (KF) has been previously studied using single-molecule stretching experiments (27-29). Our single molecule results for the force dependence of DNA pol III core catalytic activity qualitatively agree with the previously studied polymerases to the extent that force inhibits or facilitates polymerization and exonuclease
activity, depending on the force relative to the 6 pN crossover point. However, DNA pol III core is a weakly processive polymerase that only incorporates ~20 nucleotides before dissociating from the primer-template junction (53). This weak processivity of DNA pol III core imposes an additional challenge in studying its function and characterizing its activity, especially at the single-molecule level, demanding high resolution data acquisition. In this study we have successfully probed the force dependence of both the polymerase and exonuclease functions of DNA pol III core, and by modeling the force dependence of the observed pol and exo velocity and pausing, we obtain significant new insights into how these processes are regulated.

One major difference between DNA pol III core and the previously studied T7 and φ29 polymerases is that DNA pol III core is a multienzyme assembly in which the pol and exo domains are different subunits, α and ε, which can function independently even when not part of core (33). In fact, a recent study estimated that the distance between the polymerase and exonuclease active sites in DNA pol III core is greater than 7 nm (50). Therefore, for the exo activity to initiate after DNA pol III binding to duplex DNA, the 3′ end of the primer strand has to move from the pol catalytic site in α into the exo catalytic site in ε. The length change associated with the force dependence of the exo initiation rate $k_{\text{init}}(F)$ is only ~0.44 nm (2 bp × ~0.22 nm.bp, where ~0.22 nm.bp is the extension change associated with melting one DNA bp at $F > 30$ pN, Fig 4.1B). Because this length change is much smaller than the distance between the active sites of α and ε, α binding to the template must be disrupted to allow ε to bind to the frayed strand.
The structural autonomy of the two catalytic domains of DNA pol III core may result in more independent functions between the two proteins. In the case of T7 and φ29 DNA polymerases, duplex DNA binds to the polymerase active site and exonucleolysis is facilitated via an intramolecular transfer through several intermediate steps (28, 54, 55). Although the applied force favors exo and suppresses the pol activity of φ29, the underlying mechanochemistry is significantly different from DNA pol III core. Specifically, in contrast to pol III core, the dwell time of φ29 is much longer than its time of switching between pol and exo activities. The average velocity of φ29 catalysis appears to be a continuous function of the force with several fine features suggesting intermediate steps in the pol to exo switching process (28). In contrast, ε by itself has been shown to be an ssDNA exonuclease, and the catalytic activity of ε is similar on ssDNA and mispaired primer-extension termini. The ε subunit preferentially binds ssDNA, whereas α binds both ssDNA and dsDNA and prefers a primer-template junction (30, 56). In addition, a recent NMR study showed that a primer destabilized due to mismatches increases the propensity of the mismatch to reach the ε subunit, enabling ε to correct for the mismatches in a passive manner (57). Thus, a simple model for the regulation of exo and pol activity is based primarily on the preferential binding of each protein for specific DNA substrates. Because α binds strongly to a stable primer-template junction, while ε binds strongly to ssDNA free ends, the switch between pol and exo is determined by the stability of the primer-template junction, which is reduced upon the application of large forces or at high temperatures. These conditions, as well as mismatches at the primer-template terminus, induce a shift from pol to exo activity, and these activities will be considered independently below.
4.4.1 Force-dependent instantaneous velocity of DNA pol III core exonucleolysis

We have modeled the force-dependence of the pol and exo instantaneous velocities of DNA pol III core as independent processes. Here, for the first time we were able to directly measure the instantaneous exo velocity on a properly paired dsDNA substrate uniformly destabilized by a stretching force parallel to the DNA axis. Interestingly, this velocity ranges between 40 and ~110 nt/s as the stretching force increases from 30 to 55 pN (Fig. 4.4). Because at 55 pN the dsDNA is very close to its melting force of 62 pN, we approximate the latter measurement as the maximum catalytic exo rate of DNA pol III core. The observed velocities at higher forces (~100 nt/s) are independent of the DNA pol III core concentration (data not shown) and are also not strongly facilitated by the force. The exponential dependence of the observed exo velocities yields $v_{0,exo} = 20 \pm 5$ nt/s and $d = 0.11 \pm 0.02$ nm with dNTPs and $v_{0,exo} = 15 \pm 10$ nt/s and $d = 0.14 \pm 0.04$ nm without dNTPs. The independence of both parameters within uncertainty to the presence of dNTPs suggests that $\epsilon$ acts independently from the polymerase $\alpha$, even in the context of DNA pol III core. The length change required during a rate-limiting step of processive catalytic excision is significantly smaller than that of the length change required for exo initiation (2 bp ~ 0.4nm, Fig. 4.6B), which is the rate-limiting step for the exo process. Thus, while exo initiation is relatively slow and strongly dependent on the primer-template terminus stability, the instantaneous exo velocity is about ~10-100 fold faster and depends weakly on the base pair stability. This result indicates that the catalytic excision by itself is likely not strongly affected by the presence of mismatches or the DNA sequence at the junction.
4.4.2 Force-dependent instantaneous velocity and pausing of DNA pol III core polymerization

The instantaneous pol velocity is strongly affected by the applied force (Fig. 4.4). However, the dwell times during pol are approximately force-independent (Fig. 4.7A). The zero-force pol velocity ($v_{0,pol}$) was found to be $84.8 \pm 5.7$ nt/s nt/s based on the model of Eqs. 1 and 2. Therefore, the predicted number of nucleotides synthesized by DNA pol III core at zero force during an average dwell time of 0.21 s is $\sim$18 nt (Fig. 4.7B). Pauses during consecutive pol bursts are relatively longer than in exo and are poorly or not at all affected by the applied force. We find the average pause over all forces to be $0.90 \pm 0.22$ s at 0.2 μM DNA pol III core. Thus, in strong contrast to the exo activity, the instantaneous catalytic velocity of the pol activity is strongly affected by force, while the initiation of a pol event is insensitive to the applied force. This result is consistent with the fact that pol binds strongly to a stable primer-template junction, the presence of which does not depend strongly on force at pol-competent forces. In contrast, exo requires a highly force-dependent destabilized primer-template junction for initiation, as discussed below. However, once initiation of exo occurs, exo activity can proceed without further requirements for base pair destabilization.

4.4.3 Two-step exonucleolysis initiation from concentration-dependent measurements

We model the observed concentration dependence of exo-initiation rates using the proposed reaction scheme shown in Eq. 6. The DNA pol III core-saturated value for $k_{\text{init}} \approx k_2$ ($11.3 \pm 3.1$ s$^{-1}$ at 55 pN), the rate at which an exo-active state is achieved, can be compared to the rate of a single exo-cut by DNA pol III core measured previously using bulk biochemical assays (32). After correcting the results in Brenowitz et al (32) (as discussed below) to match our
experimental conditions, we find that $k_{\text{init}}$ for ssDNA at 310 K is $\sim 8.1 \text{ s}^{-1}$ (dashed line, Fig. 4.8A). This represents the maximum value that characterizes $k_{\text{init}}$ for a completely destabilized DNA substrate, obtained by increasing force or temperature. The similarity of this result to our single molecule measurements at high forces (55 pN) approaching the DNA melting force (62 pN) brings confidence that we are measuring the same process (Fig. 4.6). Furthermore, we measured $K_d$ to be $0.13 \pm 0.07 \mu M$, which agrees with the previously measured values of 0.14-0.46 $\mu M$ (30, 32) for comparable solution conditions.

### 4.4.4 Exonucleolysis initiation is determined by primer-template junction stability

We model the force dependence of the exo-initiation rate as a function of the complete work done by force to transform DNA from its double- to single-stranded form as described in Eqs. 8-9 (51). We find that at least two base pairs ($n_{\text{exo,init}}=1.84 \pm 0.20 \text{ bp}$) are required to be destabilized through thermodynamic fluctuations to enter the exo-active state of DNA pol III core. As shown in in our primer-extension assay with DNA pol III core (Fig. 4.6C) and as demonstrated by Miller et al (33) for isolated ε, there is a significant increase in exo activity when the number of mismatches at the primer-template terminus is altered from one to two, in remarkable agreement with our single-molecule results.

At concentrations above the DNA pol III core-DNA binding $K_d$, the force dependence measured is determined primarily by $k_2$. Thus, simple diffusion-limited bimolecular binding of the enzyme to its DNA substrate occurs in pre-equilibrium to exo initiation. The exo initiation rate is strongly affected by dsDNA stability, leading to a strong force and temperature dependence of $k_{\text{init}}$. If DNA pol III core exo activity is primarily rate-limited by the destabilization of the primer-
template terminus, our force-dependent measurements should be comparable to the previously measured temperature dependence of its activity (32). The force-dependence of exo initiation $k_{\text{init}}(F)$ in our study is analogous to the rate of single nucleotide excision on matched GC-terminated and single-stranded DNA reported in Brenowitz et al (32). To compare these values to those measured here, we scale the reported parameters of the Michaelis-Menten excision reaction ($V_{\text{max}}/K_m$) to determine $k_{\text{init}}(T)$, where,

$$k_{\text{init}}(T) = \left( \frac{V_{\text{max}}}{K_m} \right) \times \left( \frac{K_m}{C} \right). \quad (10)$$

Here we took into account that the $K_m$ is ~400 nM as measured in the same work at 2 nM DNA pol III core concentration (C), and shown in Fig. 4.8A (32). Furthermore, we modeled $k_{\text{init}}(T)$ as an Arrhenius function analogous to the Eq. 8 where,

$$k_{\text{init}}(T) = k_1 e^\frac{-n_{\text{exo,init}} \Delta G_0(T)}{k_BT} = k_1 e^\frac{-n_{\text{exo,init}} \Delta H_0 - T \Delta S_0}{k_BT} \quad (11)$$

Here $\Delta G_0(T)$, $\Delta H_0$ and $\Delta S_0$ are the free energy, enthalpy and entropy of a single bp melting at a reference temperature $T$, and $n_{\text{exo,init}}$ is the number of base pairs melted during the rate-limiting step of exo-initiation. The best fit yields $n_{\text{exo,init}} \Delta H_0$ to be 47 $k_BT$. Because the enthalpy of a single base pair melting, $\Delta H_0$, is ~17 $k_BT$ (58), $n_{\text{exo,init}}$, the minimum number of destabilized base pairs required in the rate-limiting step for exo, can be estimated to be ~2.5 bp. Furthermore, a global fit (Fig. 4.8B) to both the force and temperature dependence as a function of the free energy required to destabilize dsDNA, determined using the same values, yields $n_{\text{init},\text{exo}} = 2.16 \pm$
0.13 bp. This agrees remarkably well with the value determined for $n_{\text{exo,init}}$ (1.8 ±0.2 nt) obtained only from our force-dependent measurements. The excellent compatibility between the force and temperature dependence of the DNA pol III core exo activity confirms that the rate limiting step $k_2$ is primarily dependent on the stability of the primer-template terminus. Furthermore, the exo catalytic functions of DNA pol III core and ε are shown to be similar once scaled with their appropriate $K_d$ values (32). This suggests that the requirement for destabilization at the primer-template terminus for the onset of exo activity by either ε or DNA pol III core are very similar. However, DNA pol III core alters the geometry of bound ε, in which the 3′ end of the primer strand is required to be displaced a significant distance to the catalytic site of ε to trigger the onset of exo activity.

4.5 Conclusions

Taken together, our results support a model in which the pol and exo activities of DNA pol III core are effectively independent and the stability of the primer-template junction determines the selection between α and ε binding to the 3′ end of the primer strand. Once DNA pol III binds in a pol- or exo-competent conformation, enzymatic activity proceeds with relatively high velocity of 10-100 nt/s (Fig. 4.4). Despite these high catalytic rates for the pol- and exo- domains, the processivity of DNA pol III core remains low at 10-20 nt (Fig. 4.7B), as both catalytic events occur via short ~0.2 s bursts (Fig. 4.7A). These bursts are interrupted by DNA pol III core dissociation from the primer-template junction. Re-initiation of exo activity requires re-binding of DNA pol III to the primer-template junction from solution, followed by the slower melting of ~2 bp at the primer terminus. This is then followed by the much faster transfer of the 3′ end of the primer strand to the DNA pol III exo site. The force dependence we measure for exo
initiation closely matches the previously measured temperature dependence of the same process (32). This result supports a model in which mismatch recognition during proofreading is determined by primer template duplex end stability, rather than a model of duplex defect recognition (32, 33). Overall, the independent nature of the pol and exo states as well as frequent dissociation of DNA pol III from its substrate are expected to allow stronger regulation of these processes by other factors involved in *E. coli* replication.
4.6 References


Chapter 5: Conclusions and future work

5.1 Do the umuD gene products modulate the DNA pol III processivity switch?

The τ subunit of DNA pol III coordinates the subassemblies that comprise the replisome. The τ and γ subunits are encoded by dnaX from which τ is the full-length product and γ is created by a programmed ribosomal -1 frameshift during translation (1-3). Therefore, τ possesses two unique C-terminal domains relative to γ that make τ essential for cell viability (4, 5). The C-terminal domain of τ binds the C-terminal domain of DNA pol III α with an equilibrium dissociation constant $K_d$ of 4 nM (6-9). The τ-α interaction facilitates dimerization of DNA pol III cores that semi-discontinuously replicate the leading and lagging strands of template DNA (10).

Discontinuous replication on the lagging strand requires DNA pol III α to dissociate from DNA upon completion of each Okazaki fragment and recycle at a β clamp downstream. This polymerase recycling is known as the processivity switch and relies on the τ-α interaction (4, 11).

The C-terminus of DNA pol III α contains both β binding motifs (8, 12), the τ subunit interaction site (6), a UmuD binding site (13, 14), and an oligonucleotide binding (OB-fold) domain (15) responsible for ssDNA binding (16). Previous work from the Beuning Lab showed that UmuD is able to displace α from the β clamp (13) and inhibit α binding to ssDNA (14). These data suggest that such an interaction is associated with the primitive DNA damage checkpoint in which full-length UmuD specifically inhibits replication during the early stages of the SOS response to DNA damage (17-19). Interestingly, a similar change in affinity of DNA pol III α for ssDNA is observed during the processivity switch. Upon completion of an Okazaki fragment, the OB-fold of α no longer has ssDNA to bind, triggering a conformational change that decreases the
polymerase’s affinity for the clamp (20). Therefore, one can imagine a model in which UmuD binding the CTD of α induces a similar conformational change that inhibits ssDNA binding and, in doing so, displaces α from β.

Figure 5.1 – UmuD proteins change FRET of DNA pol III α and τ
(A) Linear representation of the DNA pol III α sequence showing N-terminal PHP domain, β binding sites, UmuD binding site, OB-fold domain, and τ binding site. (B) Changes in FRET intensity and (C) FRET efficiency of α647 and τ488 in the presence of different UmuD proteins.

The τ subunit also binds ssDNA (7, 11) and regulates affinity of the polymerase for primed DNA (21), which facilitates recycling. Therefore, both τ and UmuD have the ability to modulate the affinity of α for DNA by interacting at unique sites oriented on either side of the OB-domain.

These observations suggest that an interaction between UmuD and τ may exist. Such an
interaction would likely affect replication processivity, which is consistent with the role of UmuD in the DNA damage checkpoint.

Preliminary results generated using a FRET assay in which DNA pol III α is labeled with A647 (α^{647}) and τ is labeled with A488 (τ^{488}) suggest that addition of the UmuD proteins induces a change in the α^{647}-τ^{488} interaction (Figure 5.1). Surprisingly, the data suggest that addition of UmuD' induces a change in the α^{647}-τ^{488} interaction detected by an increase in FRET (Figure 5.1, C). Disruption of the α^{647}-β^{488} interaction was detected by a decrease in FRET upon addition of UmuD and UmuD S60A (13). A potential explanation for increased FRET efficiency between α and τ is that UmuD induces a conformational change such that τ is drawn closer to α. Recently, cryo-EM structures of the DNA pol III β_{2-α-ε-τ_c} complex both DNA-bound and DNA-free were published (22). A significant conformational change is illustrated upon comparing the DNA-bound and DNA-free structures. The positioning of the C-terminal “tail” of DNA pol III α is adjusted by 30 Å upon DNA-binding, and the authors note that τ_c appears to sequester the polymerase tail away from β_{2} in the DNA-free structure (22). Such a conformational change may be induced by interaction of UmuD with the α-τ complex as indicated by the observed increase in FRET. To validate the observed increase in α^{647}-τ^{488} FRET upon addition of UmuD', the experiment should be repeated using α^{488} and τ^{647}.

This work provides new insights into how the protein dynamics of the umuD gene products regulate their cellular activities during the SOS response to DNA damage. We have also characterized the mechanochemistry of E. coli DNA pol III core and suggest a model where polymerization and exonucleolysis activities are governed by primer stability at the 3’ terminus.
Future work will further probe the interactions of UmuD with the *E. coli* replisome to provide additional insights into the regulation of DNA replication in response to DNA damage.
5.2 References


