INVESTIGATING THE ROLE OF UMUD IN DNA POLYMERASE REGULATION

A thesis presented

by

Michelle C. Silva

to

The Department of Chemistry and Chemical Biology

In partial fulfillment of the requirements for the degree of
Master of Science

In the field of
Chemistry

Northeastern University
Boston, Massachusetts

December 2009
INVESTIGATING THE ROLE OF UMUD IN DNA POLYMERASE REGULATION

by

Michelle C. Silva

ABSTRACT OF THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Chemistry in the Graduate School of Arts and Sciences of Northeastern University, December, 2009
ABSTRACT

DNA is constantly bombarded with both endogenous and exogenous damaging agents, such as benzo[a]pyrene, found in cigarette smoke and charred meat, and ultraviolet light. To overcome DNA lesions caused by these agents, repair pathways are employed by the cell. Unrepaired DNA damage is capable of disrupting normal DNA replication, which can be mutagenic or lethal to a cell. DNA polymerase III (pol III), the main replicative polymerase in *Escherichia coli*, cannot bypass a DNA lesion. When it encounters damage, DNA pol III stalls. To combat DNA damage, the cell initiates the SOS response which results in the transcription of at least 57 SOS genes. The products of these genes are involved in DNA damage tolerance mechanisms, including translesion synthesis (TLS), which is the process of copying damaged DNA. TLS is implemented by a group of potentially mutagenic DNA polymerases with the ability to bypass DNA lesions by inserting nucleotides opposite the lesion. One of these polymerases, pol V, is composed of UmuD’2 (the RecA/ssDNA-facilitated cleavage product of UmuD2) and UmuC which is the polymerase subunit. It has been suggested from genetic analysis that UmuD2 together with UmuC play a role in a primitive DNA damage checkpoint inhibiting replication to allow time for accurate DNA repair processes to act. Recently, it has been hypothesized that UmuD alone may be critical for regulating the cellular response to DNA damage. The current study will examine the role of UmuD in regulating polymerase access to the replication site.
AKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Penny Beuning for giving me the opportunity to work in her lab these past two years. Working with the proteins DNA polymerase III $\alpha$ and UmuD has allowed me to become a better chemist. I would also like to thank the members of my thesis committee, Dr. Mary Jo Ondrechen and Dr. David Budil, for supporting me throughout the thesis process.

I would also like to thank all past and present members of the Beuning lab for supporting me throughout this process. In particular, I would like to thank my bench-mate, Erin Ronayne for putting up with me and for her work in expressing the $\alpha$ variants and assembling various cleavage assays.

Finally, I would like to thank my family, especially my parents. Thank you for supporting me and allowing me the opportunity to have the best education possible. Without you, I would not have been who I am today. Eu te amo!!!!

Thank you all.
TABLE OF CONTENTS

ABSTRACT ............................................................................................................ 4

AKNOWLEDGEMENTS ...................................................................................... 5

TABLE OF CONTENTS ....................................................................................... 6

LIST OF FIGURES ................................................................................................ 9

LIST OF TABLES ................................................................................................ 10

LIST OF ABBREVIATIONS .............................................................................. 11

CHAPTER 1: Overview of DNA Replication in *Escherichia coli* ............... 15

1.1 *E. coli* DNA Polymerase III ............................................................................. 15

1.2 The Polymerase Core ........................................................................................... 22

*The Polymerase Subunit, α*........................................................................... 22

*The ε Subunit* ........................................................................................................ 25

*The θ Subunit* ....................................................................................................... 26

1.3 The Clamp Loader Complex and the β Clamp Subunit ............................. 27
CHAPTER 2: Analyzing the Interactions between DNA Pol III and UmuD. 78

2.1 Introduction ...........................................................................................................78

2.2 Materials and Methods .........................................................................................81

Proteins and Plasmids ...............................................................................................81

Primer-Template DNA Complex ..................................................................................82

Thermal-Shift Assays .................................................................................................83

UmuD in-vitro Cleavage Assays ..................................................................................84
Tryptophan Fluorescence Assay

His-tagged Protein Pull-Down Experiments

Gel Filtration

2.3 Results

The presence of \( \alpha \) induces a conformational change in the arm region of UmuD.

The presence of \( \alpha \) indirectly inhibits UmuD cleavage.

The determination of the binding constant between \( \alpha \) and UmuD.

A primer-template DNA complex may be needed for \( \alpha \) to bind UmuD.

2.4 Discussion

2.5 References

CHAPTER 3: Future Directions

3.1 Localize UmuD interaction site on DNA polymerase III \( \alpha \) subunit

3.2 Determining protein-protein interaction parameters using ITC

3.3 Using EPR to analyze conformational effects due to binding

3.4 References
LIST OF FIGURES

Figure 1.1 DNA polymerase III in the dimer form .......................................................18

Figure 1.2 The τ subunit ..............................................................................................32

Figure 1.3 Loading the β clamp onto the primer-template DNA duplex ......................33

Figure 1.4 The β clamp ..............................................................................................38

Figure 1.5 Model of full-length UmuD and NMR structure of UmuD' .......................48

Figure 2.1 The homology model of the α subunit of DNA pol III ...............................82

Figure 2.2 The melting curves of UmuD and α ..........................................................89

Figure 2.3 UmuD cleavage assays with wild-type α ..................................................92

Figure 2.4 UmuD cleavage assays with α variants .....................................................93

Figure 2.5 Tryptophan fluorescence assay ...............................................................96

Figure 2.6 The melting curves of α, UmuD, and primer-template complex ...............100
LIST OF TABLES

Table 1.1 The components of DNA polymerase III .......................................................19

Table 2.1 Calculated T_m of α and UmuD in various complexes.................................100
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>~</td>
<td>Approximately</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>Å</td>
<td>Ångströms</td>
</tr>
<tr>
<td>Ala</td>
<td>Alanine, A</td>
</tr>
<tr>
<td>AP</td>
<td>Abasic Sites</td>
</tr>
<tr>
<td>Arg</td>
<td>Arginine, R</td>
</tr>
<tr>
<td>Asn</td>
<td>Asparagine, N</td>
</tr>
<tr>
<td>Asp</td>
<td>Aspartic Acid, D</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine Triphosphate</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge-coupled Device</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>Cys</td>
<td>Cysteine, C</td>
</tr>
<tr>
<td>DEER</td>
<td>Double Electron Electron Resonance</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</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 II</td>
<td>DNA Polymerase II</td>
</tr>
<tr>
<td>DNA pol III</td>
<td>DNA Polymerase III</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>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>DNA pol β</td>
<td>DNA Polymerase β</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>EPR</td>
<td>Electron Paramagnetic Resonance</td>
</tr>
<tr>
<td>FRET</td>
<td>Fluorescence Resonance Energy Transfer</td>
</tr>
<tr>
<td>GF</td>
<td>Gel Filtration</td>
</tr>
<tr>
<td>Gly</td>
<td>Glycine, G</td>
</tr>
<tr>
<td>Hepes</td>
<td>4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid</td>
</tr>
<tr>
<td>His</td>
<td>Histidine, H</td>
</tr>
<tr>
<td>hrs</td>
<td>Hours</td>
</tr>
<tr>
<td>IDPs</td>
<td>Intrinsically Disordered Proteins</td>
</tr>
<tr>
<td>Ile</td>
<td>Isoleucine, I</td>
</tr>
<tr>
<td>IMSL</td>
<td>Iodomethyl Spin Label</td>
</tr>
<tr>
<td>ITC</td>
<td>Isothermal Titration Calorimetry</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>$K_d$</td>
<td>Dissociation Constant</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilo Dalton</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis Constant</td>
</tr>
<tr>
<td>Leu</td>
<td>Leucine, L</td>
</tr>
<tr>
<td>Lys</td>
<td>Lysine, K</td>
</tr>
<tr>
<td>Met</td>
<td>Methionine, M</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------</td>
</tr>
<tr>
<td>min</td>
<td>Minutes</td>
</tr>
<tr>
<td>mL</td>
<td>Milliliters</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>MTSL</td>
<td>Methane Thiosulfonyl Spin Label</td>
</tr>
<tr>
<td>n</td>
<td>Number of Binding Sites</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>nM</td>
<td>Nanomolar</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometers</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance Spectroscopy</td>
</tr>
<tr>
<td>nt/s</td>
<td>Nucleotides Per Second</td>
</tr>
<tr>
<td>NTA</td>
<td>Nitrilotriacetic Acid</td>
</tr>
<tr>
<td>OB</td>
<td>Oligonucleotide/Oligosaccharide Binding</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine, F</td>
</tr>
<tr>
<td>PHP</td>
<td>Polymerase and Histidinol Phosphatase</td>
</tr>
<tr>
<td>pM</td>
<td>Picomolar</td>
</tr>
<tr>
<td>PMT</td>
<td>Photomultiplier Tube</td>
</tr>
<tr>
<td>Pro</td>
<td>Proline, P</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Real Time Polymerase Chain Reaction</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>sec</td>
<td>Seconds</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine, S</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. aquaticus</td>
<td><em>Thermus aquaticus</em></td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine, T</td>
</tr>
<tr>
<td>TLS</td>
<td>Translesion Synthesis</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting Temperature</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine, Y</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>Val</td>
<td>Valine, V</td>
</tr>
<tr>
<td>WT</td>
<td>Wild-Type</td>
</tr>
<tr>
<td>ΔH</td>
<td>Enthalpy</td>
</tr>
<tr>
<td>ΔS</td>
<td>Entropy</td>
</tr>
<tr>
<td>μg</td>
<td>Micrograms</td>
</tr>
<tr>
<td>μL</td>
<td>Microliters</td>
</tr>
<tr>
<td>μM</td>
<td>Micromolar</td>
</tr>
</tbody>
</table>
CHAPTER 1

Overview of DNA Replication in *Escherichia coli*:

DNA Polymerase III, DNA Damage and the Y Family DNA Polymerases

1.1 *E. coli* DNA Polymerase III

DNA replication requires the coordination of many different proteins working together to accomplish the goal of replicating two antiparallel stands of DNA simultaneously. This process is tightly regulated so that DNA is replicated in a timely and accurate manner. Moreover, DNA replication is highly processive, which allows efficient replication of over three million base pairs in every cell cycle.

In *Escherichia coli*, DNA polymerase III (DNA pol III) is responsible for the majority of DNA replication (Figure 1.1). In each cell, there are approximately ten copies of DNA pol III present (Kornberg et al. 1992). Along with other auxiliary proteins, DNA pol III is able to semi-discontinuously replicate DNA at speeds of approximately 1 kilobase (kb) per second (Kelman et al. 1995), making only one error in approximately $10^5$ nucleotide additions (Bloom et al. 1997). In the presence of proofreading and mismatch repair, there is approximately one error per $10^{12}$ nucleotide additions (Topal et al. 1976). At each replication fork, DNA pol III acts as an asymmetric dimer (Johanson et al. 1984; Maki et al. 1988; O'Donnell et al. 1990; Yuzhakov et al. 1996); one monomer acts in the
continuous replication of the leading strand and another acts in the discontinuous replication of the lagging strand (Figure 1.1). Even though only two polymerases are needed to replicate DNA, it has been shown that the replisome may contain three DNA polymerases, in which the third polymerase can function on the lagging strand or serve as a spare polymerase, able to replace either polymerase when needed (McInerney et al. 2007).

DNA pol III consists of ten subunits (Maki et al. 1988; Kelman et al. 1995; Johnson et al. 2005) that can be classified into three subassemblies: the core, the clamp and the clamp loader complex (Onrust et al. 1995) (Table 1.1). The core is composed of three subunits: α, ε, and θ. The α subunit contains the polymerase activity and is responsible for synthesizing DNA. Alone, its processivity is low, as it typically copies 1-3 kb per binding event as opposed to >50 kb for that of the entire complex (Maki et al. 1985; Studwell et al. 1990). The ε subunit is a 3’ to 5’ exonuclease that is responsible for the proofreading capability of the core. In the absence of ε, the frequency of mutations due to misincorporations during replication increases by approximately forty-fold (Maki et al. 1987). When coupled, the exonuclease and polymerase activities of both α and ε subunits increase significantly (Maki et al. 1987). The third component of the core, θ, has no known effect on processivity (Studwell et al. 1990).
The β clamp, also known as the processivity clamp, is the major contributor to the processivity of DNA pol III (Kornberg et al. 1992). Processivity is a measurement of the number of nucleotides a polymerase incorporates into the nascent DNA in one association event. The β clamp encircles DNA and tethers the α subunit to its DNA substrate. As a result, the β clamp allows DNA pol III to replicate both strands with high processivity (Figure 1.1).

The clamp loader complex (or the γ complex) is composed of six subunits: τ, γ, δ, δ’, χ, ψ (Table 1.1). The τ subunit coordinates replication on both strands by coupling the polymerase cores to the clamp loader complex (McHenry 2003). When τ is coupled to the core, processivity increases approximately six-fold (Kornberg et al. 1992). The γ subunit is an ATPase. Along with δ and δ’, it is responsible for loading the clamp onto the DNA (Kelman et al. 1995; Johnson et al. 2005). The ψ and χ subunits bind to single stranded DNA binding protein (SSB) and help regulate replication on the lagging strand (Glover et al. 1998).
Figure 1.1 DNA polymerase III in the dimer form. The two polymerase cores (green), which are tethered to the β clamps (light blue), contain the three subunits α, ε, and θ. The γ complex (blue), assembled with two τ subunits, couples the polymerization sites of both the leading strand and the lagging strand. The single-stranded lagging strand is threaded through DnaB (orange) and is coated with SSB (gray). Also shown here is the primase (light green) which synthesizes the RNA primers (red) on the lagging strand. The ψ and χ subunits are not shown here but they connect the γ complex to SSB.
<table>
<thead>
<tr>
<th>Mass (kDa)</th>
<th>Gene</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>dnaE</td>
<td>polymerase</td>
<td>(McHenry 1982; Lopez de Saro et al. 2003b; Dohrmann et al. 2005; Bailey et al. 2006; Lamers et al. 2006; Wing et al. 2008)</td>
</tr>
<tr>
<td>27.5</td>
<td>dnaQ</td>
<td>3’-5’ exonuclease</td>
<td>(Maki et al. 1987; Taft-Benz et al. 1998; Perrino et al. 1999; Taft-Benz et al. 1999; Hamdan et al. 2002; Lehtinen et al. 2004; Kirby et al. 2006; Wieczorek et al. 2006; Ozawa et al. 2008)</td>
</tr>
<tr>
<td>10</td>
<td>holE</td>
<td>stabilizes the core</td>
<td>(Scheuermann et al. 1984; Studwell-Vaughan et al. 1993; DeRose et al. 2003; Lehtinen et al. 2004; Mueller et al. 2005; Kirby et al. 2006; Ozawa et al. 2008)</td>
</tr>
</tbody>
</table>
**Clamp Loader**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>dnaX</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>τ</td>
<td>71</td>
<td>coordinates</td>
<td>replication</td>
</tr>
<tr>
<td>γ</td>
<td>47.5</td>
<td>dnaX ATPase</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tsuchihashi et al. 1992; Hingorani et al. 1998; Ellison et al. 2001; Jeruzalmi et al. 2001a; Podobnik et al. 2003; Simonetta et al. 2009)</td>
<td></td>
</tr>
<tr>
<td>δ</td>
<td>35</td>
<td>holA “wrench”; opens clamp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Hingorani et al. 1998; Ellison et al. 2001; Jeruzalmi et al. 2001a; Jeruzalmi et al. 2001b; Burnouf et al. 2004; Simonetta et al. 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>$\delta'$</td>
<td>33</td>
<td>holB</td>
<td>mediator between $\gamma$ &amp; $\delta$</td>
</tr>
<tr>
<td>$\chi$</td>
<td>15</td>
<td>holC</td>
<td>binds SSB</td>
</tr>
<tr>
<td>$\psi$</td>
<td>12</td>
<td>holD</td>
<td>bridges $\delta$ &amp; $\chi$</td>
</tr>
</tbody>
</table>
1.2 The Polymerase Core

The DNA pol III core includes the polymerase and proofreading exonuclease activity of the replisome. Alone, the core can replicate DNA at a rate of approximately 20 nucleotides per second (nt/s) with a processivity of 11 kb, much slower and with lower processivity than that of the entire replisome (Kelman et al. 1995; Johnson et al. 2005). The following is an in-depth description of each of the core’s subunits.

The Polymerase Subunit, \( \alpha \)

The \( \alpha \) subunit, a product of the \textit{dnaE} gene and a member of the polymerase C family, is a protein of approximately 130 kDa, containing several distinct domains. Although this polymerase has been studied for decades, the crystal structure of the \( \alpha \) subunit of \textit{E. coli} was only solved recently (Lamers et al. 2006). With a resolution of approximately 2.3 Å, it confirms that like other polymerases, the structure resembles a right hand with three characteristic domains: the palm, the fingers, and the thumb domains (Figure 2.1) (Brautigam et al. 1998; Rothwell et al. 2005). The palm domain contains the active site of the polymerase consisting of three aspartic acid residues: Asp401, Asp403, and Asp555. This domain is similar to the palm domains of other polymerases in the X family, especially to that of pol \( \beta \) (Lamers et al. 2006). Modeling DNA onto the \textit{E. coli} DNA pol III \( \alpha \) structure using the human pol \( \beta \) co-crystal structure with DNA (Sawaya et al. 1997) as the template, showed that the DNA strand sterically clashes with a short \( \alpha \)-helix in the
palm domain (Lamers et al. 2006). This suggests that although DNA pol III and pol β have similar active sites, they do not bind to DNA in the same way.

The finger domain includes four sub-domains: the index finger, the middle finger, the ring finger, and the little finger. These sub-domains are responsible for binding the incoming nucleotide. The thumb domain guides the newly formed DNA duplex as it leaves the active site. The polymerase and histidinol phosphatase (PHP) domain is located where the “wrist” would be, relative to the polymerase domain. The exact role of the PHP domain is unknown but it was proposed to possess pyrophosphatase activity that would hydrolyze the pyrophosphate product produced during replication because of the domain’s sequence similarity to histidinol phosphatases (Aravind et al. 1998). The crystal structure of DNA pol III α suggests that this domain is unlikely to harbor such activity (Lamers et al. 2006). However, it has been demonstrated that the DNA pol III α subunit of Thermus thermophilus contains a Zn$^{2+}$-dependent 3’ to 5’ exonuclease activity (Stano et al. 2006).

The C-terminal domain is not present in the crystal structure of E. coli DNA pol III α. This domain, which is located C-terminal to the tip of the little finger domain, includes the oligonucleotide/oligosaccharide binding (OB) fold (Bailey et al. 2006; Lamers et al. 2006) and the binding sites for both the τ subunit (Gao et al. 2001; Jergic et al. 2007) and the β clamp (Lopez de Saro et al. 2003a; Dohrmann et al. 2005). The recently-solved
structure of DNA pol III α from eubacterial *Thermus aquaticus* (Bailey et al. 2006) includes this C-terminal domain, consisting of an αβ fold which interacts with the fingers domain. The OB fold domain consists of four β-strands connected with two α-helices arranged in a β barrel, similar to that of other OB folds (Theobald et al. 2003) and has been shown to bind single-stranded DNA (ssDNA) specifically (McCauley et al. 2008). Compared to other ssDNA binding proteins, the α OB-fold domain is somewhat unusual in that it does not actively melt DNA, but rather binds to ssDNA melted by other factors (McCauley et al. 2008). Along with the rest of the polymerase, this function provides insight into the regulation of DNA pol III α, discussed below.

Recently, a co-crystal structure of *T. aquaticus* DNA pol III α bound to primer-template DNA and an incoming deoxynucleoside 5′-triphosphate has been determined with a resolution of 4.6 Å (Wing et al. 2008). When compared to the structure of *T. aquaticus* DNA pol III α without DNA, it is possible to see significant movements of the thumb, finger, and β-binding domains. These movements position the protein onto the DNA substrate, allowing interaction with the DNA backbone at the minor groove. This structure also indicates that the DNA and incoming nucleotide bind in a similar fashion to that of pol β. The C-terminal domain undergoes an approximately 30° rotation putting the OB fold in position to bind single stranded DNA. The internal β-binding motif also seems to be correctly positioned in the structure with DNA in order to bind to the hydrophobic pocket on the β clamp (Dohrmann et al. 2005; Wing et al. 2008).
The structure of a ternary complex of pol C, the replicative polymerase of the gram-positive bacterium *Geobacillus kaustophilus*, with primed DNA and an incoming dideoxynucleoside substrate has been recently solved to 2.4 Å resolution (Evans et al. 2008). Unlike the DNA pol III α polymerase, pol C contains an intrinsic 3' to 5' exonuclease domain as an insert within the PHP domain. Instead of being located in the C-terminal domain, the OB fold of pol C is located N-terminal to the palm domain and is positioned so it could bind the template strand approximately 15-20 nucleotides away from the polymerase active site. The thumb domain contains a conserved β-strand that binds DNA in the minor groove, allowing for the detection of mismatched base pairs after incorporation (Evans et al. 2008). Flexibility in the palm domain suggests a large conformational change upon DNA binding (Evans et al. 2008; Lamers et al. 2008).

**The ε Subunit**

The ε subunit, a 27.5 kDa protein encoded by the dnaQ (also known as mutD) gene, is responsible for the 3' to 5' exonuclease activity of the polymerase core (Kornberg et al. 1992) and forms a tight 1:1 complex with DNA pol III α. Whereas in *E. coli* the polymerase and exonuclease reside on two different polypeptides, in other cases, the exonuclease activity and the polymerase activity are part of the same polypeptide, as in gram positive pol C and in pol I (Brautigam et al. 1998; Steitz 1999; Evans et al. 2008), another eubacterial polymerase. Such an interaction between α and ε enhances the overall
activity of each protein. In fact, it has been shown that $\alpha$ is three times more active in complex with $\varepsilon$ than alone and the $\varepsilon$ subunit’s exonuclease activity is nearly forty-fold greater within the complex than alone (Maki et al. 1987). The $\varepsilon$ subunit in complex with $\alpha$ also decreases the $K_M$ of $\alpha$ for DNA from 0.67 $\mu$M to 0.38 $\mu$M (Maki et al. 1987).

Although the structure of the N-terminal domain of $\varepsilon$ has been determined by NMR (Hamdan et al. 2002), no structures for the full-length $\varepsilon$ subunit have been determined due to the difficulty in purifying large amounts of protein. The NMR structure includes the 185 N-terminal residues of $\varepsilon$ responsible for its exonuclease activity. The 57 residues that are not present in the structure include the C-terminal domain, which contains a flexible linker that has been shown to bind to the $\alpha$ subunit (Taft-Benz et al. 1998; Perrino et al. 1999; Taft-Benz et al. 1999; Ozawa et al. 2008). The $\varepsilon$ subunit binds $\alpha$ at the PHP domain located in the $\alpha$ N-terminal domain (Wieczorek et al. 2006).

**The $\theta$ Subunit**

The third subunit of the polymerase core is $\theta$, the 10 kDa product of the *holE* gene (Kornberg et al. 1992). The $\theta$ subunit binds to $\varepsilon$ close to the $\varepsilon$ active site (Studwell-Vaughan et al. 1993; DeRose et al. 2003) and has not been shown to bind directly to $\alpha$ (Studwell-Vaughan et al. 1993; Jonczyk et al. 1998). Attempts at isolating $\theta$ by itself after over-expression have proven to be unsuccessful unless denaturation and refolding
protocols are used (Scheuermann et al. 1984). Although its function is somewhat enigmatic, the θ subunit may have a stabilizing effect on the \( \alpha:\epsilon \) complex (Mueller et al. 2005; Kirby et al. 2006). Such a function is further supported by a series of yeast two-hybrid experiments that show that the interaction between \( \alpha \) and \( \epsilon \) is strengthened in the presence of θ (Taft-Benz et al. 2004). The θ subunit also seems to stabilize the exonuclease activity of the \( \epsilon \) subunit by using site-directed mutagenesis. The \( \epsilon \) Ile170Thr/Val215Ala double mutant variant exhibited no exonuclease activity in the absence of θ (Lehtinen et al. 2004). Upon addition of θ, activity of the \( \epsilon \) I170T/V215A variant was partially restored.

1.3 The Clamp Loader Complex and the \( \beta \) Clamp Subunit

The clamp loader complex consists of at least two \( \tau \) subunits, up to three \( \gamma \) subunits, and one each of the \( \delta \), \( \delta' \), \( \chi \), and \( \psi \) subunits (Kelman et al. 1995; Johnson et al. 2005). As mentioned above, DNA pol III can be assembled as a trimer (McInerney et al. 2007). In this case, three \( \tau \) subunits are present in the absence of \( \gamma \) subunits. Although not involved with the clamp loading process, the \( \tau \) subunits play the critical role of managing replication on both strands (McHenry 2003). The \( \gamma \) subunits, closely related to \( \tau \), bind ATP, and facilitate loading the \( \beta \) clamp onto DNA (Ellison et al. 2001). \( \delta \) and \( \delta' \) are directly involved with the loading of the clamp, \( \delta \) as the “wrench” and \( \delta' \) as the mediator between \( \tau \) and \( \delta \) (Ellison et al. 2001).
The γ Complex: δ:γ1:γ2:γ3:δ; Loading the Clamp

The dnaX gene encodes both τ and γ. τ is the full-length product of the gene and γ is transcribed due to a –1 frameshift, which causes a stop codon to be inserted prematurely, forming the shorter product, γ (Blinkowa et al. 1990; Flower et al. 1990; Tsuchihashi et al. 1992). This frameshift is caused by two factors: a downstream pseudoknot and a heptanucleotide sequence known for causing frameshifts (Figure 1.2A). Therefore, γ consists of only the first three of the five τ domains (Figure 1.2B). These three domains contain the ATPase site and so both τ and γ are ATPases. In the crystal structure of the γ complex, containing γ3δδ′ (Jeruzalmi et al. 2001a), it is possible to distinguish these domains. Domains I and II contain the nucleotide binding site in which ATP binds at the interface between the subunits and Domain III forms a circular collar with the other subunits of the γ complex.

A crystal structure of the first 243 residues of γ (Domains I and II) solved with and without nucleotides (Podobnik et al. 2003), shows a conformational change when ATP binds, suggesting a mechanism for binding the clamp and ATP hydrolysis (Figure 1.3). In the crystal structure of the γ complex (Jeruzalmi et al. 2001a), the γ subunits, together with δ and δ′, are arranged in a heptameric complex resembling an opened ring in the order δ′:γ1:γ2:γ3:δ where δ′ and δ do not bind each other (Figure 1.3). At each γ interface there is an ATP binding site (Jeruzalmi et al. 2001a). Such a configuration suggests that upon ATP binding, the γ complex undergoes a conformational change from a closed state
(without ATP) to an open state (with ATP) (Podobnik et al. 2003). The δ subunit is then free to interact with the β clamp.

The δ′ subunit, the 33 kDa member of the γ complex encoded by the holB gene, acts as a mediator between γ and δ (Kornberg et al. 1992). Although sequence and structure alignments of δ′ and γ suggest that these two subunits are homologous, δ′ does not have a functional nucleotide binding domain, as shown in the crystal structure of the δ′ subunit (Guenther et al. 1997). This crystal structure shows that δ′, like γ, consists of three consecutive domains organized in a C-shaped architecture. The first domain consists of a β sheet with five parallel strands surrounded by six α-helices similar to the RecA nucleotide binding domain, although it is non-functional (Story et al. 1992; Guenther et al. 1997). This δ′ domain also contains a zinc-binding module whose function is unknown. But because it is found on what resembles a phosphate binding loop, it may help in coupling DNA binding with ATP hydrolysis.

The δ subunit, a 35 kDa product of the holA gene, is considered the “wrench” of the clamp loader complex because when it binds to the β clamp, it causes a spring-like conformational change that allows the dimeric ring of the β clamp to transition from its default closed state, where both dimeric interfaces are intact, to an open state, where only one dimeric interface exists (Figure 1.3) (Jeruzalmi et al. 2001b; Millar et al. 2004). This spring-like opening mechanism facilitates loading of the β clamp onto the primer-
template DNA duplex (Kelman et al. 1995; Jeruzalmi et al. 2001a; Johnson et al. 2005). The wrench-like interaction between δ and the β clamp, is shown in a crystal structure involving the N-terminal 140 residues of δ and a mutant form of the β clamp (Jeruzalmi et al. 2001b), which cannot dimerize. Attempts to crystallize δ with the full-length β clamp dimer have been unsuccessful (Jeruzalmi et al. 2001b). In fact, δ binds to the monomer form approximately fifty-fold more tightly than to the dimer. The binding interface between these two subunits is contained in the hydrophobic tip of δ and a hydrophobic pocket on the surface of the β clamp that is also known for binding other components of the replisome (Jeruzalmi et al. 2001b; Lopez de Saro et al. 2003a) and DNA polymerases (Becherel et al. 2002; Sutton et al. 2005; Beuning et al. 2006a; Scouten Ponticelli et al. 2009; Burnouf et al. 2004). When the structure of the monomer form of the β clamp without δ was compared to that with δ, a 10-14 Å distortion of the curvature of the β clamp can be seen. Such an opening is enough to allow ssDNA into the ring (Ellison et al. 2001).

The crystal structures of the clamp loader subunits suggest a detailed process for loading the β clamp onto a primed DNA strand (Ellison et al. 2001), powered by the binding of two ATP molecules (Hingorani et al. 1998) (Figure 1.3). When ATP binds to the closed ring-shaped γ complex, a conformational change takes place disrupting the interaction between δ and δ'. The δ subunit is then free to bind to the hydrophobic pocket of the β clamp dimer. This binding event distorts the β dimer interface allowing DNA to enter the
clamp creating an opened ring-like structure consisting of the opened γ complex and clamp threaded onto the DNA substrate. Site-directed mutagenesis (Goedken et al. 2005) and electron microscopy (Miyata et al. 2005) have shown that DNA can bind to the center chamber of the complex. Once ATP is hydrolyzed, the two dimer interfaces of the β clamp are restored allowing the γ complex to relax back to its closed state. This causes the δ subunit to release its hold on the β clamp, the rate-limiting step of the clamp loading process (Anderson et al. 2009). Once this occurs, the β clamp and primer-template DNA duplex is assembled and is ready for the loading of the polymerase (O'Donnell 1987).

A new crystal structure, showing the clamp loader complex coupled to DNA (Simonetta et al. 2009) supports the notched screw-cap model for clamp loading as described above. The structure shows that the complex forms a right-handed, spiral-like structure and is loaded onto dsDNA like a cap, allowing ssDNA to exit through a slit formed by the complex. In this structure, the γ complex does not recognize DNA via the primer as previously thought (Miyata et al. 2005). Rather, recognition occurs on the phosphate backbone of the template strand alone. This allows for both DNA and RNA primers to be recognized, the mechanism of which was previously unclear (Lahue et al. 1989).
Figure 1.2 (A) The segment of the *dnaX* gene, responsible for coding the γ subunit. The two factors that cause the -1 frameshift are a heptanucleotide sequence (bold) and a downstream pseudoknot. As a result, the γ subunit consists of the first 430 residues of the *dnaX* gene product shared with the τ subunit, followed by a glutamic acid residue and a stop codon. (B) A side-by-side comparison of the domains belonging to both the τ and γ subunits. The γ subunit, the shorter protein, contains only the first three domains of the entire protein. These domains include the ATPase active site and the collar domain. Domains IV and V of the τ subunit contain the binding sites for both the pol III α subunit and DnaB.
Figure 1.3 The process of loading the β clamp onto the primer-template DNA duplex. ATP binds the γ complex, causing a conformational change from a closed state to an opened state which allows the N-terminal domain of the δ subunit to bind to the β clamp. This interaction disrupts the dimer interface of the β clamp, creating an opening small enough for the primer-template DNA duplex to enter. The bound ATP is then hydrolyzed, causing the γ complex to relax back to its closed state. The δ subunit then releases the β clamp, reestablishing the dimer interface.

The χ and ψ Subunits
The χ and ψ subunits, products of the holC and holD genes respectively, are subunits of the clamp loader complex (Kelman et al. 1995; Johnson et al. 2005). These two subunits function during replication on the lagging strand (Glover et al. 1998). Because DNA polymerases replicate DNA only in the 5' to 3' direction, the lagging strand must be replicated in the opposite direction to which the replication fork is moving. This is
accomplished by replicating DNA in ~1 kb fragments, called Okazaki fragments (Okazaki et al. 1968). As a result, an abundance of ssDNA is present and coated with SSB. The subunits χ and ψ help to coordinate replication on the lagging strand by connecting the γ complex to SSB (Glover et al. 1998). The ψ subunit acts as a mediator between the γ complex and χ by binding the collar domains (Domain III) of the τ and γ subunits (Olson et al. 1995). The χ subunit binds to the C-terminal domain of SSB, thereby coupling it to the replisome and allowing the clamp loader complex to be in close proximity to the primer-template DNA on the lagging strand (Witte et al. 2003). Together, the χ and ψ subunits constitute a tightly held complex that increases the affinity of τ and γ for δ and δ’ (Olson et al. 1995).

The β Clamp Subunit

Once it is loaded onto the primer-template DNA duplex, the β clamp (dnaN) has two specific roles: to tether the polymerase to the DNA and to contribute to the mobility of the polymerase on the DNA strand. The β clamp, a ring-shaped homodimer (Figure 1.4) (Kong et al. 1992), is the major contributor to processivity (Kornberg et al. 1992), which allows the polymerase to maintain close contact with the DNA so that if α does dissociate, it can quickly re-associate with the DNA, thus providing high processivity to DNA pol III. In order to facilitate processive DNA synthesis, the clamp must remain bound to the DNA with or without the polymerase present. This has been shown (Stukenberg et al. 1991). Other studies observed that the clamp can remain on a circular
plasmid two to three times longer than the time it takes for cells to divide (Yao et al. 1996). With the use of single molecule fluorescence spectroscopy, it has been suggested that the motion of the clamp is due to the attractive interactions of positively charged residues on the inside of the clamp that come in contact with the negatively charged phosphate backbone of the DNA (Laurence et al. 2008). Therefore, the clamp is able to remain bound to DNA and move with the polymerase.

Binding experiments of the clamp with DNA pol III α and other components suggest that the same hydrophobic region on the surface of β to which δ binds, is also responsible for binding DNA pol III α and other DNA polymerases (Wagner et al. 2000; Becherel et al. 2002; Lopez de Saro et al. 2003a; Sutton et al. 2005; Beuning et al. 2006a; Scouten Ponticelli et al. 2009). It has also been suggested that the DNA pol III α C-terminus binds to the β clamp (see below) and τ. This suggests that the polymerases, δ, and τ compete with one another for binding, creating a mechanism for polymerase loading and switching on the β clamp (Leu et al. 2003; Lopez de Saro et al. 2003a; Lopez de Saro et al. 2003b; Georgescu et al. 2009).

According to Dohrmann and McHenry, an internal binding site (α residues 920–924) instead of the twenty C-terminal residues was shown to be responsible for the interaction between the β clamp and α (Dohrmann et al. 2005). Replacement of all residues in this internal binding site eliminated binding to the β clamp altogether, but binding between
the τ subunit and α was not affected. When an analogous set of mutations was made in the C-terminal binding site for α, the β clamp still showed affinity of the α subunit (Dohrmann et al. 2005). In fact, α was still able to participate in processive replication when the entire C-terminal binding site was removed. The absence of the C-terminal peptide, however, was found only to have an effect on τ’s interaction with α. These findings suggest that the clamp does not bind to the α C-terminus, but instead to an internal binding site. One possible reason for this discrepancy is that the previous study used a α variant with a relatively large C-terminal truncation (Dohrmann et al. 2005), rather than using site-directed mutant variants or more modest deletions (Lopez de Saro et al. 2003a). The truncations that were used, may not be stable.

There is also evidence that two different polymerases can bind to the clamp at the same time. β, a homodimer, has two hydrophobic pockets per functional protein, thus allowing it to bind two DNA polymerases simultaneously (Fujii et al. 2004; Johnson et al. 2005). Such a situation allows replication to alternate between the two DNA polymerases without the need for them to dissociate from and re-associate with the clamp (the “toolbelt” theory) resulting in high processivity even under conditions where multiple polymerases are used. The ability of the clamp to bind both DNA pol III and pol IV (a Y family polymerase) was investigated using fluorescence resonance energy transfer (FRET). FRET was detected between the pol III alpha subunit and pol IV, in a β-clamp-dependent manner, providing experimental evidence that two different DNA polymerase protein molecules can simultaneously bind to the β clamp (Indiani et al. 2005). The
crystal structure of β with the C-terminal little finger domain of pol IV (Bunting et al. 2003) showed that pol IV is angled off to the side, providing enough room for another polymerase, like DNA pol III, to bind. Modeling the full length structure of Dpo4 (Ling et al. 2001) (a pol IV homolog from *Sulfolobus solfataricus*) with a primer-template duplex and incoming nucleotide onto that of the little finger structure of the *E. coli* pol IV with the β clamp, showed that when bound in this position, it is likely that the polymerase does not have access to the DNA strand. But when the polymerase is needed, the little finger domain can then undergo a conformational change (Bunting et al. 2003), positioning the polymerase onto the DNA substrate.

Another model for polymerase switching is of dynamic processivity (Yang et al. 2004) hypothesizing that during replication, rather than only one polymerase replicating the entire DNA fragment, the polymerase is constantly being replaced by another polymerase without affecting overall processivity. Such a scheme was observed during bacteriophage T4 DNA replication (Yang et al. 2004). The addition of a catalytically inactive variant Asp408Asn of gp43, the T4 DNA polymerase, to an active replication fork, arrested replication while still retaining wild-type-like affinity for DNA and the clamp (gp45), suggesting that the active polymerase is quickly (<1 min) replaced by the inactive variant (Yang et al. 2004). This dynamic processivity of polymerases implies that multiple replicative polymerases may be required to replicate normal, undamaged DNA.
Figure 1.4 Residue substitutions in the β clamp that affect interaction with UmuD (left), UmuD’ (middle) and α subunit of pol III (right). Positions in green are important to the interaction between the β clamp and all three proteins listed above. Positions in purple exhibit only a modest effect. Substitutions that result in an increase or decrease in the affinity of UmuD and UmuD’ for the β clamp by formaldehyde or glutaraldehyde cross-linking are shown in red. Residue Lys74 shown in grey (left) cross-links to UmuD using formaldehyde. Structures were generated using the coordinates for β (2POL) from the protein data bank (PDB) (Kong et al. 1992; Duzen et al. 2004).

The τ Subunit

Although similar to the γ subunit, the τ subunit’s role in the replisome is distinct from the rest of the γ complex. As a central component of the DNA pol III replisome, τ coordinates replication on both strands by connecting the subassemblies of the core and β clamp at the two replication forks by binding to α (Gao et al. 2001), the clamp loader complex (Olson et al. 1995), and the helicase DnaB (Dallmann et al. 2000). Numerous
distinct roles have been assigned to τ, which are summarized in a review by McHenry (McHenry 2003). The roles of the τ subunit in replication are to: 1) coordinate replication on both strands; 2) bind to DnaB; 3) prevent premature removal of the β clamp; 4) function in the processivity switch.

When the polymerase core was first isolated (McHenry 1982), it was observed that two polymerases were coupled by two τ subunits, suggesting that τ may be a key factor in coordinating replication (Figure 1.1). This hypothesis was tested by varying the concentration of τ (Kim et al. 1996b). At low concentrations of τ, shorter DNA fragments were observed when the diluted replication reaction samples were analyzed by agarose gel electrophoresis (Kim et al. 1996b). These results suggest that the two polymerase cores need to be coupled through τ in order to effectively replicate both the lagging and leading strands.

DNA pol III α binds τ through the α C-terminal domain with a $K_d$ of 4 nM (Gao et al. 2001). The α binding site on τ was determined to be at the τ C-terminal domain (Dallmann et al. 2000), which is not part of γ (Figure 1.2B). Along with determining the NMR solution structure for this domain (Su et al. 2007), combinatorial binding studies were also conducted to determine exactly which residues bind to α (Jergic et al. 2007). It was concluded that the 18 C-terminal residues of τ were needed to bind α (Jergic et al. 2007).
Four different reconstituted clamp loader complexes, $\gamma_3$-$\delta\delta'\chi\psi$, $\tau_1\gamma_2$-$\delta\delta'\chi\psi$, $\tau_2\gamma_1$-$\delta\delta'\chi\psi$, and $\tau_3$-$\delta\delta'\chi\psi$ have similar rates of loading the $\beta$ clamp onto DNA (McInerney et al. 2007). The complex containing $\tau_3$ may coordinate three DNA polymerases at the replication fork. A model has been proposed in which two of the pol III cores function on the lagging strand to synthesize Okazaki fragments. Alternatively, one pol III core is utilized on the lagging strand with the third pol III core held “in reserve” off of the DNA (McInerney et al. 2007). The latter model suggests a possible switching mechanism between high and low fidelity DNA polymerases (McInerney et al. 2007).

The affinity of the $\tau$ subunit for DnaB (Dallmann et al. 2000) also seems to affect the rate at which the replication fork proceeds (Kim et al. 1996a; Yuzhakov et al. 1996). DnaB, a hexameric helicase, unwinds DNA ahead of the fork while encircling only the lagging strand (Figure 1.1). Without the replisome, the helicase unwinds DNA at a rate of approximately 35 nucleotides per second (nt/s), similar to that in the presence of the complex without $\tau$. When $\tau$ is added, the rate increases to 500-700 nt/s, approaching that of DNA pol III (Dallmann et al. 2000).

The $\tau$ subunit may indirectly prevent the premature removal of the $\beta$ clamp by the $\gamma$ complex, thereby maintaining processivity. The length of DNA produced is directly proportional to the concentration of $\beta$ and inversely proportional to the concentration of
the γ complex (Kim et al. 1996c). This suggests that the clamp removal function is inhibited during normal replication in the presence of the γ complex, allowing the β clamp to stay on the DNA.

As mentioned, the lagging strand is discontinuously replicated. In order to complete replication of the lagging strand, DNA pol III α must constantly dissociate and re-associate with different β clamps. Such a cycle is known as the processivity switch and is thought to involve τ (Leu et al. 2003). It has been shown that the very last nucleotide must be incorporated by the polymerase core/τ complex to the newly formed strand, leaving only a nick, before the switch can be activated (Leu et al. 2003). In such a scenario, τ loses affinity for α when primed DNA is present. Then when replication is completed, τ gains affinity for α, releasing α from the clamp (Leu et al. 2003).

1.4 DNA Damage Disrupts DNA Replication

DNA damage is ubiquitous, arising from numerous exogenous and endogenous sources. For example, it is estimated that 10,000 abasic sites (AP) are formed per human cell per day (Friedberg et al. 2006). DNA replication is highly efficient, but replicative DNA polymerases generally cannot copy damaged DNA templates. The outcome of the replisome encountering DNA damage in the template may depend on whether the damage is encountered during leading strand or lagging strand synthesis, but the rate of progress of replication is typically decreased (Rupp 1996; Rudolph et al. 2007). A lesion
in the leading strand has been observed to slow progression of the replication fork, but was not observed to interfere with lagging strand replication (Pages et al. 2003). On the other hand, a lesion in the lagging strand does not block overall progression of the replication fork, instead, the lagging strand DNA polymerase appears to re-initiate downstream of the lesion at the next Okazaki fragment, leaving a gap (Higuchi et al. 2003; Pages et al. 2003; McInerney et al. 2004). Indeed, it has been shown that replication can restart downstream of obstacles, even in the case of leading strand synthesis (Heller et al. 2006).

1.5 Specialized DNA Polymerases Facilitate DNA Damage Tolerance

In *E. coli* and some other bacteria, DNA damage and other stresses lead to induction of the SOS response (Friedberg et al. 2006). Stalling of DNA replication at damaged sites results in the accumulation of single stranded DNA (ssDNA). RecA polymerizes on the ssDNA, forming a nucleoprotein filament that serves as the inducing signal for the SOS response. As the result, the expression of at least 57 genes is induced (Simmons et al. 2008; Foti et al. 2009). SOS-regulated genes code for proteins involved in regulation of cell division, nonmutagenic repair of chemically modified DNA or in damage tolerance mechanisms, which can be mutagenic or error prone (Jarosz et al. 2007; Schlacher et al. 2007; Simmons et al. 2008).
Three of the five known *E. coli* DNA polymerases belong to the group of SOS inducible proteins (Nohmi 2006; Jarosz et al. 2007). These SOS inducible DNA polymerases are pol II (*polB*), pol IV (*dinB*), and pol V (*umuDC*; UmuD\_2C), the latter two belonging to the Y family of DNA polymerases, which are characterized by their ability to perform translesion synthesis (TLS) on damaged DNA templates as well as their relatively low fidelity on undamaged DNA (Yang 2003). Y family DNA polymerases also lack intrinsic 3'-to-5' exonucleolytic proofreading and exhibit low processivity (Beard et al. 2001; Yang 2003; Jarosz et al. 2007).

Although high-resolution structures of *E. coli* Y family polymerases have not yet been experimentally determined, the structures of Y family polymerases from *Sulfolobus solfataricus* and *Sulfolobus acidocaldaricus* homologs provide insights into their function (Ling et al. 2001; Silvian et al. 2001; Trincao et al. 2001). While there is no obvious sequence homology between replicative and TLS polymerases, the crystal structures of Y family DNA polymerases reveal a similar right-hand structure of the catalytic domain consisting of thumb, palm, and finger domains, common to other DNA polymerases. Another domain, the little finger domain, is present in the Y family polymerases providing additional DNA binding contacts in the major groove (Ling et al. 2001; Yang 2003). It also seems to be responsible for both substrate specificity and processivity (Boudsocq et al. 2004). Moreover, the O helix that provides high fidelity in the replicative polymerases (Bell et al. 1997; Beard et al. 2003; Zhang et al. 2006) is not present in Y family polymerases, suggesting a structural basis for the low fidelity of Y
family DNA polymerases while replicating undamaged DNA. The specialized ability of TLS polymerases to replicate damaged DNA has been attributed to their loose, flexible active sites that accommodate aberrant DNA structures and to having fewer contacts with their DNA substrates than replicative DNA polymerases (Ling et al. 2001; Zhou et al. 2001). Although the crystal structures show that the catalytic domains of Y-family polymerases have similar overall folds, these polymerases exert different efficiencies and fidelity in bypassing various DNA lesions (Boudsocq et al. 2002; Yang 2003).

Regulation of Y Family DNA Polymerases

Expression of the *umuDC* and *dinB* gene products is negatively regulated as part of the SOS response at the transcriptional level by the LexA repressor, which binds to a sequence in the operator region of the genes (Kitagawa et al. 1985; Courcelle et al. 2001; Friedberg et al. 2006). Derepression of the *umuDC* and *dinB* operons occurs when the RecA protein binds to single-stranded regions of DNA that develop at replication forks stalled by DNA damage. The RecA/ssDNA nucleoprotein filament serves as a coprotease to facilitate auto-cleavage of the LexA repressor. As the cellular concentration of LexA diminishes, the genes whose expression is normally repressed by LexA are transcribed (Friedberg et al. 2006).

The *umuDC* genes are some of the most tightly regulated SOS genes; the $K_d$ is 0.2 nM for LexA binding to the “SOS-box” in the promoter region (Kitagawa et al. 1985).
Immunoblotting assays have been used to directly measure the cellular steady-state levels of UmuD to be ~180 copies per uninduced cell and ~2400 copies per cell under SOS induction (Woodgate et al. 1991). The level of UmuC is approximately 12-fold lower than UmuD with ~200 molecules of UmuC per cell under SOS induced conditions and about 15 molecules per cell in the absence of induction (Woodgate et al. 1991).

Expression of the \textit{umuDC} genes initially produces UmuD (139 amino acids) which undergoes a RecA/ssDNA-stimulated autodigestion reaction after induction resulting in UmuD' (115 amino acids) (Figure 1.5) (Nohmi et al. 1988; Rajagopalan et al. 1992). UmuD is the predominant species for the first ~20-40 min after SOS induction, after which UmuD' is the predominant species. UmuD proteins exist in solution as UmuD$_2$ and UmuD'$_2$ homodimers as well as the UmuD-UmuD' heterodimer, which is more stable than either of the homodimers (Woodgate et al. 1989; Battista et al. 1990; Peat et al. 1996a; Ferentz et al. 2001). The $K_d$ for UmuD$_2$ dimerization is estimated to be in the low-pM range, so UmuD is likely to be present in the cell as a dimer under most conditions (Simon et al. 2008).

SOS-induced mutagenesis is also regulated at the post-translational level. UmuD is functionally inactive for facilitating TLS until it undergoes RecA/ssDNA-mediated cleavage to generate UmuD' (Burckhardt et al. 1988; Nohmi et al. 1988; Shinagawa et al. 1988). Full-length UmuD also inhibits -1 frameshifting mutagenesis by DinB (Godoy et al. 2007). Efficient cleavage of UmuD \textit{in vitro} and \textit{in vivo} was observed at elevated levels of activated RecA, suggesting that TLS likely occurs when cells are under more severe conditions.
environmental stress (Burckhardt et al. 1988; Woodgate et al. 1991). The removal of the UmuD N-terminal 24 amino acids through the cleavage of the Cys24-Gly25 bond occurs via an intermolecular pathway, that is, one protomer of the dimer acts as an enzyme, while the other is the substrate (McDonald et al. 1998). The crystal structure of UmuD' revealed that the active site, consisting of conserved serine (Ser60) and lysine (Lys97) residues, is found at the end of a cleft within the C-terminal globular domain of the protein and these residues are poised for cleavage (Figure 1.5) (Peat et al. 1996a). In the NMR structure of UmuD' Ser60 and Lys97 are further apart and not correctly oriented for catalysis. It has been suggested that the crystal structure of UmuD2 mimics the effect of the RecA-ssDNA nucleoprotein filament as it serves to realign these residues, thereby activating UmuD2 for self-cleavage (Ferentz et al. 2001; Sutton et al. 2001a).

The mutagenic potential of these proteins may be further regulated by the preferential formation of heterodimers between UmuD and UmuD', thereby depleting the cell of mutagenically active UmuD' homodimers (Battista et al. 1990). UmuD' is degraded by the ATP-dependent protease, ClpXP while in a heterodimeric complex with UmuD (Gonzalez et al. 1998; Gonzalez et al. 2000). The preferred formation of UmuDD' heterodimers rather than mutagenically active UmuD' homodimers specifically targets UmuD' for proteolysis. UmuD also targets its UmuD partner in the homodimer for proteolytic degradation by ClpXP (Neher et al. 2003). The ATP-dependent serine protease Lon is responsible for the degradation of both UmuD and UmuC proteins in vivo (Frank et al. 1996).
The appearance of the two different forms of UmuD provides a temporal switch between accurate and mutagenic phases of the cellular response to DNA damage (Opperman et al. 1996; Opperman et al. 1999; Sutton et al. 2001b). Uncleaved UmuD together with UmuC was found to specifically decrease the rate of DNA replication and to increase resistance of cells to killing by UV radiation (Marsh et al. 1985; Opperman et al. 1999). Uncleaved UmuD$_2$C (inactive DNA pol V) helps cells survive DNA damage by allowing time for error-free repair mechanisms to act before cleaved UmuD'$_2$C (DNA pol V) initiates potentially error-prone TLS (Opperman et al. 1999). Noncleavable UmuD(S60A) together with UmuC significantly delayed the recovery of cell growth after UV radiation (Opperman et al. 1996; Opperman et al. 1999). Therefore, a model for a $umuDC$-dependent DNA damage checkpoint in *E. coli* was proposed wherein a delay in DNA synthesis could provide the cells time to initiate error-free nucleotide excision repair for removal of DNA lesions. Then, TLS is enabled by the presence of UmuD' (Tang et al. 1998; Reuven et al. 1999; Tang et al. 1999). This discovery suggested that the different $umuD$ gene products in combination with UmuC are involved in distinct survival pathways after cells suffer DNA damage caused by UV radiation.

The UmuD'$_2$C protein complex was found to be an antagonist of RecA-mediated recombination as the recombination of a UV-damaged gene was reduced by increasing the UmuD'$_2$C concentration (Sommer et al. 1993). In the proposed model, the UmuD'$_2$C proteins at a high enough concentration provide a factor for the replisome to switch from
recombination to SOS mutagenesis. Notably, in the presence of homologous DNA sequences, homologous recombination repair is more prevalent than TLS in responding to DNA damage (Berdichevsky et al. 2002).

**Figure 1.5** Model of full-length UmuD and NMR structure of UmuD'. UmuD model shown in the trans, elbows up conformation (left). Solution structure of UmuD' also in trans conformation (right). The N-terminal arms of UmuD' have been cleaved between Cys24 and Ala25 (light blue). Active site Ser60 and Lys97 are highlighted in red and green, respectively (Ferentz et al. 1997; Ferentz et al. 2001; Beuning et al. 2006b).

**Structural dynamics of UmuD and UmuD'**

The *umuD* gene products interact with multiple factors involved in replication such as polymerases UmuC (pol V as UmuD₂C), DinB (pol IV) and components of the pol III holoenzyme including the polymerase subunit α, proofreading subunit ε, and the processivity clamp β (Woodgate et al. 1991; Sutton et al. 2002b; Beuning et al. 2006b; Godoy et al. 2007). These interactions are due in part to the relative flexibility of full-
length UmuD and its cleavage product UmuD’ as shown biochemically and by X-ray crystallography, nuclear magnetic resonance spectroscopy (NMR) and circular dichroism (CD) (Peat et al. 1996a; Ferentz et al. 2001; Simon et al. 2008). The cleaved form UmuD contains disordered N-terminal arms that allow the C-terminal globular domain to become solvent exposed upon cleavage, while in full-length UmuD, the arms are more stably bound to the globular domain (Peat et al. 1996a; Ferentz et al. 2001). Therefore, UmuD and UmuD’ are able to make highly specific contacts that can facilitate a variety of protein-protein interactions (Ferentz et al. 1997; Sutton et al. 1999; Ferentz et al. 2001; Simon et al. 2008).

The crystal structure of UmuD’ revealed extended N-terminal arms (residues 25-39) and a globular C-terminal body (residues 40 to 139) that contains the catalytic dyad Ser60 and Lys97 (Peat et al. 1996a). Although two dimer interfaces were observed in the crystal structure, designated as molecular and filament interfaces, NMR and cross-linking experiments support the conclusion that the so-called filament dimer interface is present in solution (Lee et al. 1994; Guzzo et al. 1996; Lee et al. 1996; Peat et al. 1996a; Peat et al. 1996b; Ferentz et al. 1997; Ohta et al. 1999; Ferentz et al. 2001). However, there is evidence that the filament structure may be biologically relevant (Peat et al. 1996b). While structural information is not yet available for the UmuD2 homodimer, cross-linking studies of monocysteine derivates of UmuD are consistent with the UmuD2 homodimer interface resembling the interface of the UmuD’ homodimer involving contacts between the C-termini as well as interactions involving Asn41 and Leu44 (Lee et al. 1994).
UmuD' variants that resulted in decreased UV-induced mutagenesis also have severe deficiencies in their abilities to form homodimers in vivo (McLenigan et al. 1998; Ohta et al. 1999; Beuning et al. 2009).

Four models of the UmuD homodimer have been generated by utilizing data from NMR, electron paramagnetic resonance (EPR), and cross-linking studies, and by homology to LexA (Ferentz et al. 2001; Luo et al. 2001; Sutton et al. 2002a). One model shows UmuD with the N-terminal arms in trans with the elbows down, where the N-terminal arm of one monomer folds down across the C-terminal body of the adjacent monomer and crosses the catalytic site (Figure 1.5). Each UmuD monomer cleaves its partner’s N-terminal arm at Cys24-Ala25 (Beuning et al. 2006b). A trans, elbows up version positions the arms along the outer edge of the globular domains. Two cis versions with elbows up and elbows down suggest that each N-terminal arm could bind over its own globular domain (Beuning et al. 2006b). The N-terminal region (residues 1-14) is likely to be in a random extended conformation. Cross-linking and chemical modification experiments suggest that the trans, elbows down conformation of the N-terminal domain is the most prevalent in solution (Sutton et al. 2002a; Beuning et al. 2006b). However, given the dynamic nature of UmuD, all four conformations may be physiologically relevant (Beuning et al. 2006b; Simon et al. 2008).
CD spectroscopy showed that under physiological conditions the conformations of both UmuD dimers resemble a random coil (Simon et al. 2008) rather than the β-sheet-rich structure determined by X-ray crystallography and NMR spectroscopy (Peat et al. 1996a; Ferentz et al. 1997; Ferentz et al. 2001). At higher salt concentration both UmuD and UmuD' dimers have more typical β-sheet appearance. Thus, the umuD gene products belong to the group of intrinsically disordered proteins (IDPs) (Radivojac et al. 2007). Like their IDP counterparts, UmuD dimers are capable of making a remarkable number of specific protein-protein contacts.

It has been shown that UmuD and UmuD' interact with the α, β, and ε subunits of DNA pol III (Sutton et al. 1999). Uncleaved UmuD interacts less strongly with α subunit in vitro than cleaved UmuD' (Sutton et al. 1999). Moreover, uncleaved UmuD interacts more strongly with β processivity clamp than cleaved UmuD' as observed by affinity chromatography (Sutton et al. 1999). Different interactions of UmuD and UmuD' with the β processivity clamp suggest that these interactions regulate how the umuD gene products access the replication fork. Overexpression of the α subunit or the β clamp inhibits cleavage of UmuD to UmuD' in vivo, further supporting the specific interactions between the pairs of proteins (Sutton et al. 1999). It has also been found that overexpression of the β clamp reduces UV-induced mutagenesis (Tadmor et al. 1992).
**UmuD-β clamp interactions**

Interactions with the β processivity clamp are important for the lesion-bypass activity of UmuD’2C (Becherel et al. 2002; Sutton et al. 2005; Beuning et al. 2006a). The β clamp has been shown to increase the processivity of all five *E. coli* DNA polymerases (Kornberg et al. 1992; Wagner et al. 2000; Maor-Shoshani et al. 2002). Both UmuD and UmuD’ interact with the β clamp and the α and ε subunits of pol III (Sutton et al. 1999). To date, the interactions between the *umuD* gene products and the β clamp have been studied in much more detail than other interactions involving the *umuD* gene products. Proteins that interact with the β clamp, with the exception of UmuD and UmuD’, contain the eubacterial clamp-binding motif (QL[S/D]LF) (Dalrymple et al. 2001). UmuD contains a 14TFPLF18 motif within its N-terminal arm (Dalrymple et al. 2001). Although the motif lies in a region of UmuD that is important for interaction with the β clamp (Sutton et al. 2002b), the interaction does not depend on the sequence identity of the motif (Beuning et al. 2006b).

UmuD and UmuD’ affinity chromatography and *in vitro* cross-linking studies confirm that the β clamp has a higher affinity for UmuD than UmuD’ (Sutton et al. 2002b). However, it has been shown that both the N-terminal arms and C-terminal globular domains of UmuD are important for interaction with the β clamp (Sutton et al. 2002b). UmuD lacking its N-terminal nine residues is proficient for interactions with the β clamp, while UmuD lacking the N-terminal 19 residues resulted in a reduction of the efficiency of formaldehyde cross-linking to β (Sutton et al. 2002b). The non-cleavable UmuD-3A
(Thr14Ala/Leu17Ala/Phe18Ala) variant possesses some of the biological functions of the cleaved form UmuD’. Although the $K_d$s for interaction of the $\beta$ clamp with UmuD (5.5 +/- 0.8 $\mu$M) and UmuD-3A (6.1 +/- 0.5 $\mu$M) were found to be quite similar, the manner in which these proteins interact with the $\beta$ clamp may be different, as observed by using tryptophan fluorescence (Beuning et al. 2006b). Tryptophan fluorescence is a relatively sensitive probe of the environment; the single tryptophan of the $\beta$ clamp is located on a flexible loop between domains I and II of $\beta$, and thus is a sensitive reporter of possible conformational changes (Kong et al. 1992; Beuning et al. 2006b).

Studies have shown that UmuD, UmuD’, the $\alpha$ catalytic subunit, UmuC and clamp loader all interact with the $\beta$ clamp around the $\beta$ clamp hydrophobic pocket approximately defined by residues Leu177, Pro242, Val247, Val360, Met362 (Figure 1.4) (Jeruzalmi et al. 2001a; Becherel et al. 2002; Duzen et al. 2004; Beuning et al. 2006a). UmuD, UmuD’ and the $\alpha$ subunit interact with overlapping regions of $\beta$, suggesting that there may be competition for binding to the clamp (Figure 1.4) (Sutton et al. 1999). This implies that UmuD plays a regulatory role following the SOS response by interacting with components of pol III, slowing replication and allowing time for error-free repair mechanisms to act (Opperman et al. 1999; Sutton et al. 1999). On the other hand, there is also the possibility that $\alpha$, UmuD or UmuD’ may bind the homodimeric $\beta$ clamp simultaneously. A model has been proposed in which cleavage of UmuD to form UmuD’ reduces binding to the $\beta$ clamp, thereby releasing the DNA damage checkpoint (Sutton 2006).
UmuD interacts with multiple components of the replisome. These interactions govern replication in response to DNA damage, including translesion DNA synthesis. In order to understand how UmuD regulates replication, the following study will probe the interactions between UmuD and the α subunit of DNA pol III.

1.6 References


Crowther, G. Otting and N. E. Dixon (2007). "The unstructured C-terminus of the
tau subunit of *Escherichia coli* DNA polymerase III holoenzyme is the site of

processivity clamp loader gamma complex of *E. coli* DNA polymerase III." *Cell*

Jeruzalmi, D., O. Yurieva, Y. Zhao, M. Young, J. Stewart, M. Hingorani, M. O'Donnell
and J. Kuriyan (2001b). "Mechanism of processivity clamp opening by the delta
subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III." *Cell*

support the formation of an initiation complex between the DNA polymerase III


protein interactions within the *Escherichia coli* DNA polymerase III core." *J


by increasing the affinity of tau and gamma for delta,delta' to a physiologically relevant range." *J Biol Chem* **270**(49): 29570-7.


CHAPTER 2

Analyzing the Interactions between DNA Pol III α and UmuD

2.1 Introduction

DNA polymerase III (DNA pol III) is the main polymerase in *E. coli* and is responsible for replicating the entire genome. It contains ten subunits that allow for accurate and processive replication to occur (Maki et al. 1988; Kelman et al. 1995; Johnson et al. 2005). The ten subunits are organized into three subassemblies: the polymerase core; the β clamp; and the clamp loader complex (Onrust et al. 1995). The polymerase core consists of the α polymerase, and two other subunits: the ε proofreading subunit and the θ subunit. The clamp loader loads the processivity clamp onto the replication site consisting of a primer and a template strand (Ellison et al. 2001). Once the β clamp is positioned, the clamp loader complex regulates the loading of the polymerase core onto the β clamp (Gao et al. 2001). Such an ensemble of proteins allows for continuous replication to occur on the leading strand and discontinuous replication to occur on the lagging strand (O'Donnell 1987). In this way, DNA pol III efficiently replicates DNA.

DNA is constantly bombarded with both endogenous and exogenous damaging agents, such as benzo[a]pyrene, which is found in cigarette smoke and charred meat, and ultraviolet light (Lindahl et al. 1999). To overcome DNA lesions caused by these agents,
repair pathways are employed by the cell (Nohmi 2006). Unrepaired DNA damage is capable of disrupting normal DNA replication, which can be mutagenic or lethal to a cell. DNA pol III cannot normally bypass a DNA lesion. When it encounters damage, DNA pol III stalls, continuously adding and removing a nucleotide opposite the lesion (Fujii et al. 2007). To deal with DNA damage, *E. coli* initiates the SOS response which results in the activation of various correctional procedures.

The initial molecular trigger for the SOS response is an abundance of ssDNA present at the replication fork caused by the stalled polymerase (Sassanfar et al. 1990). RecA coats the ssDNA forming a nucleoprotein filament that facilitates the auto-cleavage of the LexA repressor and up-regulates the transcription of at least 57 SOS genes (Simmons et al. 2008). The products of these genes are involved in DNA damage tolerance mechanisms, such as recombination, DNA repair and translesion synthesis (TLS), which is the process of copying damaged DNA. TLS is implemented by a group of potentially mutagenic DNA polymerases that can replace the stalled pol III (Jarosz et al. 2007).

The Y family is a group of specialized DNA polymerases with the ability to bypass DNA lesions by inserting nucleotides opposite the lesion (Cordonnier et al. 1999; Sutton et al. 2001; Goodman 2002; Pages et al. 2002). Of the five DNA polymerases present in *E. coli*, two are members of this family: DNA pol IV (DinB) and DNA pol V (UmuD′2C) (Napolitano et al. 2000; Wagner et al. 2002). Pol V is composed of the UmuD′2 dimer
(the RecA/ssDNA facilitated cleavage product of UmuD$_2$) and the UmuC monomer which includes the polymerase activity (Reuven et al. 1999; Tang et al. 1999). UmuD together with UmuC serves as a primitive DNA damage checkpoint, inhibiting replication to allow time for accurate DNA repair processes to act.

Past studies have used affinity chromatography to show that $\alpha$ and other components of DNA pol III interact with UmuD (Sutton et al. 1999). In the same study, it was determined that overproduction of $\alpha$ in vivo inhibited UmuD auto-cleavage (Sutton et al. 1999). The $\alpha$ subunit and pol IV can also interact with the $\beta$ clamp simultaneously (Ling et al. 2001; Bunting et al. 2003; Fujii et al. 2004; Indiani et al. 2005). This data suggests that the $\beta$ clamp acts as a tool belt, providing access to the necessary polymerase. When a DNA lesion is encountered, $\alpha$ stalls and a Y family polymerase quickly takes over replication (Fujii et al. 2004). Once the DNA lesion has been passed by, $\alpha$ once again takes over.

The following study will examine the role of UmuD in regulating which polymerase has access to the replication fork. A thermal-shift assay was employed to look at the thermal stability of UmuD in the presence of $\alpha$. Additionally, to determine the effect of $\alpha$ on the auto-cleavage activity of UmuD in-vitro, cleavage was assayed. Although previous research has shown that $\alpha$ and UmuD bind, no binding constant has been determined. Numerous attempts have proven to be unsuccessful.
2.2 Materials and Methods

Proteins and Plasmids

UmuD was expressed from the pSG5 plasmid as previously described (Beuning et al. 2006a). Plasmids that express His-tagged wild-type α and the truncations α270 and α280 were provided by Dr. Meindert Lamers and Prof. John Kuriyan from UC Berkeley (Lamers et al. 2006). Previous studies showed the possibility of two different binding sites, one in the N-terminal PHP domain and one in the little finger domain (Figure 2.1). Within these two binding sites, specific residues were chosen as possible binding factors. To determine the necessity of each site, site-directed mutagenesis (using mutagenic primers ordered from Operon and the QuikChange kit from Stratagene) was used to introduce a mutation to attempt to eliminate binding (αLeu77Gly). The mutations αHis824Tyr and αArg895Cys, associated with temperature sensitive alleles (Vandewiele et al. 2002), were also constructed. All α variants and truncations were purified using the established protocol (Lamers et al. 2006).
Figure 2.1 A homology model of the $\alpha$ subunit of DNA pol III is shown with the domains highlighted (Sharma, unpublished). The bar below shows the positions of all $\alpha$ mutated variants and truncations used in the current study.

Primer-Template DNA Complex

A primer-template complex was designed as the substrate for replication. It is composed of two ssDNA fragments (Operon), the primer (CAGGTCTTTCGCCAGCCTGATCTGAG) and the template (GGTTACTCAGATCAGCCTGCGAAGACCTGGGCGTCCGGCTGCAGCTGTACTATC). Samples containing 50 $\mu$M each DNA strand were annealed in a solution
consisting of 50 mM Hepes, pH 7.5, and 100 mM NaCl. The annealing process involved heating the solution at 95 °C for 2 min, cooling the solution to 54 °C holding for 30 min, and then cooling it to 25 °C and holding for 20 min.

**Thermal-Shift Assays**

Reactions were assembled in 96-well PCR plates (Applied Biosystems) in which each sample (16 μL total volume) consisted of thermal-shift assay buffer (50 mM Hepes), and 25x Sypro Orange (Invitrogen) (concentration not disclosed) (excited at 490 nm). In experiments in which the melting transitions of UmuD were observed in the presence of α, each sample contained 45 μM UmuD (monomer concentration used throughout Chapter 1) and 1 μM α. Samples were incubated for 2 hrs at room temperature before detection. In experiments to determine if DNA is a cofactor in the interaction between α and UmuD, 30 μM UmuD, 15 μM α and 25 μM DNA (annealed primer-template) were used. α and the primer-template complex were allowed to bind by incubating for 1 hr at room temperature before UmuD was added. The samples were then incubated for another hour before detection. Appropriate controls were as discussed. Once the plate was sealed with an optical adhesive film (Applied Biosystems), the iCycler iQ5 Real Time PCR (Bio-Rad) was used to increase the temperature from 25 °C to 75 °C with an increment of 0.1 °C and a 10 sec dwell time per temperature increment. The fluorescence intensities emitted at 575 nm and detected by the built-in CCD camera were plotted versus temperature. In order to calculate the melting temperature (T_m), XLfit 5, a Microsoft
Excel add-on program, was used to fit the data to the sigmoidal Boltzmann model represented by the equation:

\[
Intensity = \left( A + \frac{B-A}{1+e^{\frac{C-Temperature}{D}}} \right)
\]

(Equation 1)

where C is the Tm, A and B are pre- and post-transitional fluorescence intensities, and D is a slope factor. The assays were repeated several times with similar observations and consistent melting temperatures.

*Umud in vitro Cleavage Assays*

RecA/ssDNA nucleoprotein filament-facilitated UmuD cleavage reactions were carried out as previously described in LG Buffer (Beuning et al. 2006a). Samples containing α, UmuD and buffer were incubated for 2 hr at room temperature before the addition of RecA/ssDNA. Once RecA/ssDNA was added, reactions were carried out at either 37 °C or 30 °C. The reaction temperature was adjusted to 30 °C wherever possible because of the relatively low melting temperature of α (approximately 37 °C to 40 °C). Alkaline cleavage of UmuD was also carried out as previously described (Beuning et al. 2006a). The glycine containing buffer (pH 10) was added after 30 min incubation at room temperature. Again the reactions were carried out at either 37 °C or 30 °C. Reactions of both cleavage assays were analyzed using 18% SDS-PAGE.
Tryptophan Fluorescence Assay

Tryptophan is a fluorescent amino acid that is excited at 278 nm and can be used to assay binding (Helene et al. 1971; Helene et al. 1972; Beuning et al. 2006a). This is accomplished by monitoring the change in fluorescence of the tryptophan residues while varying the concentration of the binding partner. To determine the equilibrium constant for α binding to UmuD, this assay is attractive because UmuD has no tryptophan residues and α has eight tryptophans. As a result, the tryptophan environment of α can be monitored while titrating in UmuD. Throughout the assay, the α concentration was kept constant at 200 nM while the UmuD concentration was varied from 0 to 15 μM. The emission spectrum of UmuD alone at each concentration was subtracted from those of α and UmuD. The center of spectral mass was calculated as previously described (Beuning et al. 2006a). Graphing the center of spectral mass versus concentration of UmuD produced binding curves.

His-tagged Protein Pull-Down Experiments

Two separate conditions (with and without primer-template DNA complex) were tested using a pull-down method designed to use a His-tagged protein as “bait” in order to determine its binding partner. In both instances, a 1:2 α:UmuD monomer mole ratio was used. In the presence of DNA, the mole ratio was slightly altered to 3:6:5 α:UmuD:DNA. This assay required 50 μL of a 50% slurry consisting of nickel-NTA agarose (Qiagen) to be added to an 800 μL centrifuge spin column (Pierce). To equilibrate the column, 350 μL of wash buffer (50 mM Hepes, pH 7.5; 10% Glycerol; 2 mM β-mercaptoethanol) was
added to the resin and eluted by centrifugation (1300 x g for 30 sec). The “bait” protein, α, was added to the column and incubated at 4°C for 1 hr with a gentle rocking motion. In the assays containing DNA, α was pre-incubated with DNA for 2 hr at room temperature before loading onto the column. After the flow-through from the load was eluted by centrifugation (1300 x g for 30 sec), the column was washed five times with 350 μL wash buffer. The appropriate amount of UmuD was added to the column and was allowed to equilibrate with the nickel resin at 4°C for 1 hr with a gentle rocking motion. Once the flow-through from the UmuD load was eluted by centrifugation (1300 x g for 30 sec), 250 μL of elution buffer (50 mM Hepes, pH 7.5, 10% Glycerol, 2 mM β-mercaptoethanol, 0.5 M imidazole, pH 7.5) was added to the column and was incubated for another 5 min. Once this was accomplished, α and bound proteins were eluted from the column by centrifugation (1300 x g for 30 sec). All fractions were analyzed by 14% SDS-PAGE.

**Gel Filtration**

In an attempt to find the binding constant for the interaction between α and UmuD, a Superdex 200 (GE Healthcare) column (25 mL void volume) was employed. Initially the column was equilibrated at 4°C with GF buffer (20 mM Hepes, pH 7.5, 0.1 mM EDTA, 1 mM dithiothreitol) at a flow-rate of 0.5 mL/min for 1 hr before the protein was loaded onto the column.Samples were prepared using 100 μg of both α and UmuD (approximately a 10-fold molar excess of UmuD) and were pre-incubated for 2 hr at room temperature. Once loaded, the protein samples were eluted at a flow-rate of 0.5
mL/min with GF buffer. Fractions were analyzed using 14% SDS-PAGE, and protein was visualized by using Coomassie brilliant blue or silver staining.

2.3 Results

The presence of α may induce a conformational change in the arm region of UmuD.

A thermal-shift assay, often used in drug discovery screening (Pantoliano et al. 2001) and to optimize conditions for protein crystallography (Ericsson et al. 2006), was used to determine whether the presence of α had a structural impact on UmuD. This assay involves the use of an environmentally-sensitive fluorescent dye where the fluorescence intensity is proportional to the hydrophobicity of its environment. When a folded protein is present, the dye is in an aqueous environment and so the fluorescence of the dye is quenched. As the temperature reaches the melting point of the protein, the dye binds to the exposed hydrophobic interior of the unfolding protein, increasing the fluorescence intensity which is detected by a Real Time PCR (RT-PCR) instrument. By observing the fluorescence intensity as a function of temperature, the folded and unfolded states of the protein can be distinguished. The temperature at the midpoint of the transition between these two states is known as the melting temperature, or T_m. Because the T_m is related to the stability of the protein, a protein with a higher T_m is more stable compared with a protein with a lower T_m.
Previously, it has been shown that UmuD melts in two transitions: one at approximately 28 °C and another at approximately 60 °C. The absence of the first transition with UmuD' suggests that the N-terminal arms of UmuD dissociate from the body at a lower temperature than the entire protein melts (Fang et al.). In order to observe the effect of α on these two melting transitions, the thermal-shift assay was conducted again in the presence of α (Figure 2.2). By comparing the melting curve of UmuD with the melting curve of UmuD with α, it is clear that upon addition of α, the first melting transition disappears. This suggests that α binds UmuD in such a way that the arms cannot dissociate from the body or that the arms are not bound to the body. More research is needed to differentiate these possibilities (See chapter 3). The melting temperatures recorded for UmuD are very similar to the ones previously recorded, with a T_m of 31.79 °C for the first transition and a T_m of 60.41 °C for the second transition (Fang et al.). In the presence of α, UmuD has a T_m of 60.12 °C. α, on the other hand, has a T_m of 32.26 °C. Because of its low stability, precautions must be taken when working with α to prevent it from melting.
**Figure 2.2** Melting curves of 45 μM UmuD (blue), 45 μM UmuD with 1 μM α (green), and 1 μM α (red). The blue curve shows the melting transitions for UmuD at 31.79 °C and 60.41 °C. The disappearance of the first melting transition suggests that α may bind UmuD in the “elbows unbound” or locked in the “elbows down” conformation. The insert includes the zoomed in melting curve of α with a calculated Tm of 32.2 °C. Compared to UmuD, α is significantly less stable.

The presence of α indirectly inhibits UmuD cleavage.

RecA/ssDNA cleavage assays were employed to determine if the presence of α would have an impact on the cleavage activity of UmuD. To do this, the reaction was conducted after the addition of 10 to 25 μM α which resulted in an approximately 50% cleavage inhibition at 37 °C (Figure 2.3A). Because the melting point of α is approximately 37 °C, the assay was conducted at 30 °C to minimize the amount of denatured α present in the
reaction sample. This resulted in a slightly higher inhibition (approximately 60%). To
determine if the inhibition of cleavage was directly due to the binding between \(\alpha\) and
UmuD, the assay was conducted under alkaline conditions without the RecA/ssDNA
nucleoprotein filament (Figure 2.3B). Instead of inhibition of cleavage activity, an
increase in cleavage activity was seen with increasing amount of \(\alpha\) present. Two
conclusions can be made about the interaction between \(\alpha\) and UmuD. The observed
increase in cleavage under alkaline conditions suggests that \(\alpha\) binds UmuD in such a way
that allows the cleavage site on the arms to stay in close contact with the active site, in an
“elbows down” conformation, thus increasing the cleavage activity. Also, this suggests
that \(\alpha\) indirectly inhibits UmuD because of the competitive binding between \(\alpha\), UmuD,
and the RecA/ssDNA filament. To confirm the conformation of the arms of UmuD in the
presence of \(\alpha\), additional experiments must be conducted.

The cleavage assay was also used to probe the site on \(\alpha\) to which UmuD binds.
RecA/ssDNA facilitated cleavage was assayed in the presence of \(\alpha\) variants and \(\alpha\)
truncations (Figure 2.4A). Several \(\alpha\) residues were selected representing possible UmuD
binding sites and were changed to another amino acid that is predicted to disrupt binding.
The variant \(\alpha\)Leu77Gly gave a decrease in UmuD cleavage compared to that with wild-
type \(\alpha\) (Figure 2.4A). Variants \(\alpha\)His824Tyr and \(\alpha\)Arg895Cys, associated with known
temperature sensitive alleles (Vandewiele et al. 2002) were also tested. Like \(\alpha\)Leu77Gly,
these variants reduced UmuD cleavage compared to wild-type \(\alpha\), suggesting that these
variants bind UmuD more strongly than wild-type \(\alpha\).
The α truncations α270 and α280 include the N-terminal PHP domain (Figure 2.1). The presence of α280 and α270 had little to no effect on cleavage (Figure 2.4A). In another assay, various concentrations of α280 were tested (10 to 40 µM) (Figure 2.4B). Approximately four times more α280 than wild-type α were required to observe the same magnitude of cleavage inhibition. These results suggest that either the N-terminal domain of α does not contain a UmuD binding site, or that this particular binding site has less affinity for UmuD than full-length α. Other truncations are currently being analyzed.
Figure 2.3 After each cleavage assay, samples containing UmuD and UmuD’ were resolved using 18% SDS-PAGE. (A) RecA/ssDNA nucleoprotein filament-facilitated UmuD cleavage assays were performed in the presence of 10 and 25 µM α wild type and at two temperatures (37 °C and 30 °C). (B) Cleavage assays were conducted under alkaline conditions. Percentage of cleavage was calculated by comparing the density of cleaved product to the total amount of protein present.
Figure 2.4 UmuD cleavage assays were performed in the presence of α at 30 °C. (A) SDS-PAGE analysis of UmuD cleavage in the presence of α variants and truncations, indicated above the gel. (B) UmuD cleavage determined as a function of α280 concentration. Percentage of cleavage, shown below each gel, was calculated by comparing the density of cleaved product to the total amount of protein present.
The determination of the binding constant between α and UmuD was unsuccessful using various methods.

To determine the binding constant that governs the interaction between α and UmuD, various methods were employed with little success. A tryptophan fluorescence assay seemed to be an ideal method to use as the assay conditions can be considered to be near physiological because they do not require immobilization. The signal would be easy to interpret because UmuD has no tryptophans and α has eight tryptophans. Also, in this case an intrinsic fluorescence signal is measured as opposed to an extrinsic signal. In the case of extrinsic fluorescence experiments, a fluorescent dye is used to monitor certain physical aspects of the binding pair which would produce a change in the emitted fluorescent signal (Hawe et al. 2008). Again, by adding such a fluorescent dye, unnatural conditions would be introduced to the binding pair that could affect the affinity. The decision to use the tryptophan fluorescence assay was further supported because of its successful use in finding the binding constant between the β clamp and UmuD (Beuning et al. 2006b).

Although it was initially thought to be an ideal technique, it proved to be difficult to implement when used to determine binding between α and UmuD. Some preliminary experiments seemed promising, but were not reproducible from one trial to the next. To troubleshoot this assay, many experimental variables were changed. This included instrumental variables such as changing the slit-width, varying the PMT (photomultiplier tube) detector voltage, and others. Sample conditions were also tested by changing the
protein concentrations, varying the buffer composition and the use of additives such as glycerol to prevent aggregation. Occasionally, data produced a trend that resembled a binding curve with two binding sites (Figure 2.5). Attempts to reproduce this trend proved to be unsuccessful. Future plans involve using a technique that would provide an accurate number of binding sites and binding constants for the interactions (Chapter 3).
Figure 2.5 These two charts show preliminary binding curves collected using the intrinsic tryptophan fluorescence of α. (Excitation at 278 nm; entrance and exit slits at 5 nm) (A) An initial potential binding site was observed when 200 nM α was titrated with 0 to 2 μM UmuD ($K_D$ of approximately 13 nM). (B) A second potential binding site was observed when α was titrated with 0 to 15 μM UmuD ($K_D$ of approximately 7 μM).
The His-tagged protein pull down assay was employed because it provides a concrete way to verify the interaction between α and UmuD and to differentiate between specific and non-specific binding using α variants. Because α is expressed with a His-tag, α can be captured by a spin column containing nickel-NTA agarose resin. UmuD does not bind to this resin so any UmuD present in the final elutions would be due to binding to α. Unfortunately, no binding was observed. Several concentrations of both proteins were used with no success. We suspect that α binding to DNA might be required for its interaction with UmuD, so DNA was added to these experiments. The presence of DNA only decreased the affinity of the α His-tag for the resin.

To attempt to determine a binding constant, gel filtration chromatography was employed. With this method, the biggest concern was selectivity of the column, because α is much larger than UmuD (~130 kDa compared to ~15 kDa monomer molecular weight). The column resin selected for this experiment needed to be able to resolve α alone versus α in complex with UmuD. Superdex 200 provided the desired resolution with a 10 kDa to 600 kDa optimum separation range. Unfortunately no interaction was observed. In order to determine if DNA is needed for α to bind to UmuD, α was pre-incubated with primer-template complex before loading. Fractions collected during this experiment showed no detectable α after staining with either Coomassie dye or silver staining. This may be a direct result of the conformational changes α undergoes when binding to DNA (Wing et
al. 2008), which could result in poor resolution that would prevent sufficient separation of the proteins in complex.

*A primer-template DNA complex may be needed for α to bind UmuD.*

As mentioned above, various techniques have been employed in an attempt to determine the binding constant between α and UmuD with no success to date. A thermal-shift assay was used to determine if other cofactors are required for the interaction between α and UmuD. As previously discussed in Chapter 1, α undergoes numerous conformational changes when in complex with DNA (Wing et al. 2008). This suggests that the addition of a primer-template DNA complex may stabilize α so that it is in a favorable conformation to bind to UmuD.

A thermal-shift assay was conducted to test this hypothesis (Figure 2.6 and Table 2.1). The melting temperatures were similar to those previously recorded for both α and UmuD (38.28 °C and 60.05 °C respectively). When a primer-template complex was pre-incubated with α, the T<sub>m</sub> was increased by 1.10 °C. The melting curve for UmuD did not change in the presence of the primer-template complex, which is expected because UmuD does not bind to DNA. When α and UmuD were allowed to bind, no significant change in the T<sub>m</sub> of α was observed and a slight destabilizing effect was seen with UmuD. To determine whether DNA has any effect on the UmuD-α interaction, UmuD was allowed to bind to α that was previously incubated with the primer-template
complex. In this experiment, the $T_m$ of $\alpha$ was increased by 1.37 °C and the $T_m$ of UmuD was decreased by 1.75 °C. Assuming that the interaction between $\alpha$ and UmuD would only produce small changes in melting temperatures, it can be concluded that $\alpha$ is most stable when in complex with both the primer-template complex and UmuD. By examining the $T_m$ of $\alpha$ in various complexes, it is possible to conclude that $\alpha$ may need DNA in order to bind to UmuD. Attempts at verifying this using gel filtration and pull-downs proved to be unsuccessful. Other techniques will be employed (see Chapter 3).
Figure 2.6 The melting curves of $\alpha$ (15 $\mu$M), UmuD (30 $\mu$M), and primer-template complex (25 $\mu$M) in multiple combinations. The shift in the melting transition of $\alpha$ suggests that the primer-template complex stabilizes the $\alpha$ subunit so that it may bind to UmuD.

Table 2.1 Calculated $T_m$ of $\alpha$ and UmuD in various complexes.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>UmuD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>38.28 °C</td>
<td>---</td>
</tr>
<tr>
<td>UmuD</td>
<td>---</td>
<td>60.05 °C</td>
</tr>
<tr>
<td>$\alpha$ + DNA</td>
<td>39.38 °C</td>
<td>---</td>
</tr>
<tr>
<td>UmuD + DNA</td>
<td>---</td>
<td>59.00 °C</td>
</tr>
<tr>
<td>$\alpha$ + UmuD</td>
<td>38.25 °C</td>
<td>59.01 °C</td>
</tr>
<tr>
<td>$\alpha$ + UmuD + DNA</td>
<td>39.64 °C</td>
<td>58.30 °C</td>
</tr>
</tbody>
</table>
2.4 Discussion

DNA pol III α is the *E. coli* polymerase responsible for replicating the entire genome. When a DNA lesion is encountered, α stalls and replication is disrupted (Pages et al. 2003; Fujii et al. 2007). To rescue the stalled replication fork, Y family polymerases are employed to replicate past the lesion. Once the lesion has been bypassed, α regains access to the replication fork and continues to replicate (Fujii et al. 2007). Although it is known that pol III and Y family polymerases alternate in the presence of DNA lesions, the exact mechanism for this switch is unknown. Developing evidence suggests that UmuD plays a major role in regulating which DNA polymerase has access to the replication fork, including evidence that UmuD binds to α (Sutton et al. 1999). Recently in our lab, it has been observed that the presence of UmuD inhibits the ability of α to replicate undamaged DNA (Sefcikova, unpublished).

The research described here investigates the mechanism by which UmuD regulates the pol III α subunit. Thermal-shift assays were employed to observe the thermal stability of UmuD in the presence of α. These experiments showed that instead of melting in two transitions, UmuD melts in one transition in the presence of α. In particular, the UmuD melting transition that is not observed in the presence of α is the transition that we have shown is due to the N-terminal arm unbinding from the globular domain (Fang, unpublished).
To test whether $\alpha$ interacts with the arms of UmuD, RecA/ssDNA facilitated UmuD cleavage assays were carried out in the presence of $\alpha$. This showed that $\alpha$ inhibits cleavage by approximately 50%. RecA/ssDNA facilitates cleavage most likely by inducing a conformational change in UmuD that helps form the active site (Jarosz et al. 2007). To determine if inhibition of cleavage is directly related to the interaction between $\alpha$ and UmuD, the cleavage assay was conducted under alkaline conditions without RecA or ssDNA. This showed that the presence of $\alpha$ increases the cleavage of UmuD, suggesting that $\alpha$ binds UmuD in such a way that allows the cleavage site more access to the UmuD active site. It can also be concluded that by binding to UmuD, $\alpha$ prevents UmuD from binding the RecA/ssDNA filament, inhibiting cleavage.

All attempts to date to determine the binding constant for the interaction between $\alpha$ and UmuD have been unsuccessful. The tryptophan fluorescence assay seemed promising but probably because of the $\alpha$ subunit’s large number of tryptophan residues, the changes in the tryptophan environment were minimal and inconsistent. The little data gathered from this method suggest that there are two UmuD binding sites on $\alpha$. To determine the position of these binding sites, the thermal-shift assay was used with a number of $\alpha$ variants and truncations. The N-terminal PHP domain may be a weak binding site for UmuD. Thermal-shift experiments conducted in the presence of a primer-template complex showed evidence that DNA may be needed for $\alpha$ to bind to UmuD.
From these data, a preliminary mechanism for the complementary regulation of pol III α and UmuD can be established. Full-length UmuD binds α and inhibits replication. UmuD binding to α also interferes with the ability of UmuD to undergo RecA/ssDNA-facilitated cleavage. However, this inhibition of UmuD cleavage is not complete, and UmuD' is still formed after DNA damage. UmuD' can then form a complex with UmuC which facilitates TLS. Once the lesion has been bypassed, pol III can once again regain access to the replication site (Fujii et al. 2007). More research needs to be conducted to test this hypothesis.

2.5 References


sensitive mutator alleles of $dnaE$, encoding modified alpha-catalytic subunits of


CHAPTER 3

Future Directions

Data have been presented that begin to elucidate the interactions between the DNA polymerase III α subunit and UmuD, a DNA damage-inducible protein. Genetic evidence had suggested that UmuD acts together with UmuC to inhibit DNA replication (Marsh et al. 1985). Our laboratory has recently discovered that only UmuD is required to inhibit the DNA polymerase activity of α. Furthermore, this thesis presents the observation that α inhibits RecA/ssDNA-dependent cleavage of UmuD in vitro, but not RecA/ssDNA-independent alkaline cleavage of UmuD. Therefore, evidence is emerging to support the model that there is an intricate network of protein-protein interactions that exquisitely regulate the cellular response to DNA damage. In order to further develop this model, several future investigations are proposed.

3.1 Localize UmuD interaction site on DNA polymerase III α subunit

Preliminary data presented here and by others suggest that UmuD plays a significant role in regulating both replicative and TLS polymerases. In order to verify and extend these preliminary results, other α variants and truncations will be analyzed for their effect on UmuD cleavage. By doing so, it may be possible to pinpoint the UmuD binding site on the pol III α subunit. Structural data provided by the cleavage assays, together with a
calculated binding constant would illustrate the mechanism for polymerase regulation by UmuD.

3.2 Determining protein-protein interaction parameters using isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a powerful technique used to study both protein-ligand and protein-protein interactions (Doyle 1997; Jelesarov et al. 1999; Leavitt et al. 2001). In a single experiment, ITC can directly measure multiple thermodynamic parameters for the interacting pair. When two molecules bind, heat is either generated or absorbed. An isothermal titration calorimeter measures this difference in heat, which allows for accurate determination of binding constants ($K_D$), the number of binding sites (n), and enthalpy ($\Delta H$) and entropy ($\Delta S$) of binding (Pierce et al. 1999). Other than making sure that the protein samples are in a buffer system with low ionization enthalpy and that the protein concentration is accurately calculated, sample preparation is minimal for ITC (Pierce et al. 1999; Liang 2008). No labeling or immobilization is required, so the proteins can be assayed under physiological conditions (Pierce et al. 1999). This allows for the determination of an accurate binding constant and the possibility of artifacts due to labeling or other unnatural conditions is minimized. Thermodynamic data collected using ITC can also be used to distinguish different proteins when the $K_D$ values are similar (Liang 2008).
The relative ease of use and the large amount of data that can be collected with this method makes ITC a suitable candidate for the analysis of UmuD-α binding. As discussed in chapter 2, various attempts have been made to determine the binding constants between α and UmuD. In doing so, preliminary data were gathered showing the possibility that UmuD may bind α at two sites. Along with determining the binding constant, ITC can also be used to determine the number of binding sites. Determining the other thermodynamic parameters can also result in a better understanding of the structural dynamics involved.

3.3 Using electron paramagnetic resonance to analyze conformational effects due to binding

Electron paramagnetic resonance (EPR) is a technique used to study basic molecular mechanisms of proteins with the use of nitroxide spin labels (Borbat et al. 2001). As the name implies, the technique is used to look at the resonance of unpaired electrons. But because most biological molecules lack unpaired electrons, nitroxide spin labels are used (Klare et al. 2009). Attachment of a spin label covalently to a cysteine residue results in the ability to conduct targeted site-directed dynamic studies (Borbat et al. 2001).

Molecular distances can be determined by using two spatially separated spin labels utilizing an experimental method called double electron electron resonance (DEER) (Jeschke et al. 2007). Time resolution can also be used to look at local motion by the direct measurement of spin relaxation times. As the tumbling motion of the spin label
slows down due to factors such as binding, a drastic change can be observed in the EPR spectra (Borbat et al. 2001).

As suggested by the thermal-shift and cleavage assays, α binds UmuD in such a way that the arms cannot dissociate from the body allowing the cleavage site more access to the active site. EPR in conjunction with site-directed spin labeling can be used to verify this. Incidentally, UmuD contains only one cysteine residue at position 24 in the elbow region of the arms. If this cysteine were labeled with an iodomethyl spin label (IMSL) or methane thiosulfonyl spin label (MTSL), the dynamic motion of the arms can be monitored as UmuD interacts with α. This would help us to determine whether α binds UmuD in the predicted manner. Alternatively, single cysteine α variants can be created and the UmuD binding sites can be localized.

This work is the first to demonstrate that the α subunit of DNA polymerase III inhibits RecA/ssDNA-dependent UmuD cleavage in vitro. Furthermore, α actually increases UmuD cleavage under alkaline conditions, suggesting that α facilitates binding of the N-terminal arms of UmuD over its body so they can access the cleavage site. The extension of this work to probe the molecular details of these interactions will shed light on the regulation of DNA damage responses.
3.4 References


