The *umuD* gene products are molecular adaptors in the regulation of DNA damage tolerance

A dissertation presented

by

Jaylene N. Ollivierre

to

The Department of Chemistry & Chemical Biology, Northeastern University, Boston, MA 02115

in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

in the field of

**Chemistry**

Northeastern University
Boston, Massachusetts
March, 2012
The *umuD* gene products are molecular adaptors in the regulation of DNA damage tolerance

By

Jaylene Ollivierre

Abstract of Dissertation

Submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Chemistry in the department of Chemistry and Chemical Biology in the College of Science Northeastern University

Boston, MA
Abstract

The homodimeric $umuD$ gene products play key roles in regulating the cellular response to DNA damage in \textit{Escherichia coli}. $Umud_2$ is composed of 139-amino acid subunits and is upregulated as part of the SOS DNA damage response. Subsequently, damage-induced RecA:ssDNA nucleoprotein filaments mediate the slow autocleavage of the N-terminal 24-amino acid arms of $Umud_2$ yielding $Umud'_2$. It was previously proposed that $Umud$ cleaves only in the \textit{trans} conformation, in which the arm of one monomer utilizes that active site of the adjacent monomer for cleavage. Cleavage in \textit{trans} would therefore require dimerization. However, isoenergetic models of $Umud_2$ suggested that the arms may adopt \textit{cis} (intramolecular) or \textit{trans} (intermolecular) conformations, and may be unbound from or bound to the globular C-terminal domain. The dynamic nature of the N-terminal arms may explain how a number of distinct protein-protein contacts that prevent and facilitate mutagenic translesion synthesis (TLS) are made. The overall goal of my research is to determine the conformation and dynamics of the $Umud$ proteins in order to understand its regulatory role in response to DNA damage. Chapter 1 presents the relevant background and details of structure, function and interactions of $Umud$ with proteins involved in DNA replication and DNA damage repair.

In an effort to learn more about the structural dynamics and functions of $Umud$ proteins, we designed a $Umud$ protein variant that is defective in dimerization. Such a variant would not only answer the question as to whether $Umud_2$ is active in the \textit{cis} conformation, but also address the possibility that $Umud$ may be functionally active as a monomer. Although models of $Umud_2$ with the arms in the \textit{cis} conformation have been proposed, evidence that this conformation is physiologically relevant has been lacking. Wild-type $Umud_2$ and $Umud'_2$ form exceptionally
tight dimers in solution; however, in chapter 2 we show that the single amino-acid change N41D generates stable, active UmuD and UmuD' monomers that functionally mimic the dimeric wild-type proteins. The UmuD N41D monomer is proficient for cleavage and interacts physically with DNA polymerase IV (DinB). Furthermore, the N41D variants facilitate UV-induced mutagenesis and promote overall cell viability. Taken together, these observations show that a monomeric form of UmuD retains substantial function \textit{in vivo} and \textit{in vitro}.

UmuD$_2$ and UmuD'$_2$ can display differential interactions with their partner proteins which can lead to dramatically different cellular outcomes. These key differences may be due to the dynamics of the N-terminal arms of UmuD$_2$. Previous biochemical evidence supported a model in which the arms of UmuD are stably bound to its globular domain. However, recent experiments suggest that the N-terminal arms of UmuD are somewhat dynamic. Chapter 3 describes the use of Electron Paramagnetic Resonance (EPR) spectroscopy to probe the conformational dynamics of the N-terminal arms of the \textit{umuD} gene products and variants. We determined that the arms of UmuD$_2$ display a large degree of motion, are largely unbound from the globular C-terminal domain, and that the free energy of dissociation is +2.1 kJ/mol.

In chapter 4, we discuss the dimer exchange and conformation-dependent cleavage of the UmuD proteins. To further understand the dynamic regulatory roles of the \textit{umuD} gene products, we monitored the kinetics of exchange and cleavage of the UmuD$_2$ and UmuD'$_2$ homodimers as well as of the UmuDD' heterodimer under equilibrium conditions. We found that the heterodimer is the preferred but not exclusive dimeric protein form, and that both the heterodimer and homodimers exhibit slow exchange kinetics. In addition, the heterodimer efficiently cleaves to
form UmuD′. Together, this work reveals an intricate UmuD lifecycle that involves dimer exchange and cleavage in the regulation of the DNA damage response.

Chapter 5 discusses the use of UmuD N-terminal truncations, Δ1-7 (UmuD 8) and Δ1-17 (UmuD 18), in accessing the role of the arms in regulating protein-protein interactions. Extensive characterization reveals that the loss of even the N-terminal seven amino acids results in a notable change in domain conformation, binding affinity to DinB as detected by tryptophan fluorescence, and the facilitation of UV-induced mutagenesis and UV survival. Additionally, we have also discovered a smaller version of UmuD that is also UV-inducible. Given the information above, it is plausible that this smaller UmuD may be involved in yet another level of DNA damage and repair regulation.
Acknowledgements

I could have no better scientific influence than my graduate advisor Dr. Penny Beuning. Thank you for planting in me a passion for science and discovery through your own example, for taking the time to teach me things at the bench and other life lessons, for your patience, encouragement and support. I would also like to express my deepest gratitude to Dr. David Budil for being a wonderful mentor and for introducing me to the world of spectroscopy and protein dynamics. I can ask for no better scientific parents. I would also like to thank my thesis committee members Dr. Mary Jo Ondrechen and Dr. Carolyn Lee-Parsons for their career advice and guidance.

To my husband Jacques, thank you for your enduring love, support, and encouragement, and for providing an escape from the woes of grad school. You’ve seriously kept me sane. Of course, I would not be here without my parents Mr. and Mrs. Ollivierre, and my brother Kari. I miss you every day. Thank you for instilling in me a sense of conviction, ambition, and adaptability.

I would not be here without the guidance of many teachers over the years. Graduate school was not something I had given much thought before my undergraduate advisor Dr. Ewa Pater and my inorganic chemistry professor Dr. Ed Miller encouraged me to do so. Good teachers teach, great teachers inspire. College and grad school would also not have been the same without the safe haven provided by the Hollands. I know I always have a home on the Point.

To my lab mates, Michelle, Jason, Lisa, Ramya, Philip, Nick D and Jana, thank you for being great coworkers and my dysfunctional family. I wish nothing but the best for all of you in your future pursuits. I would also like to thank the undergraduates I’ve worked with over the years.
Your infectious enthusiasm and curiosity kept me motivated, and I’ve learnt so much from you. I would especially like to thank Jacque, Ali, Sarah, Kylie and Tyler for their hard work and for being so reliable. I could ask for no better team, or friends. I would like to thank members of the Engen lab, especially Jing, Thomas and Sean for their advice and assistance with everything mass spec. Finally, I would like to thank the College of Science and the Department of Chemistry and Chemical Biology for this exceptional educational opportunity.
Contents

Abstract .......................................................................................................................................... 3
Acknowledgements ....................................................................................................................... 7
LIST OF FIGURES AND TABLES .......................................................................................... 12
LIST OF ABBREVIATIONS .................................................................................................... 15
Chapter 1: Introduction ............................................................................................................. 19
  1.1 The *umuDC* gene products are an integral part of the SOS response to DNA damage. 20
  1.2 Structural dynamics of UmuD and UmuD′ ................................................................. 24
  1.3 The *umuD* gene products interact with the α, ε, β subunits of DNA Pol III.............. 27
  1.4 UmuD-Beta clamp interactions ................................................................................... 28
  1.5 Molecular interactions of UmuD and UmuC ............................................................... 31
  1.6 Molecular interactions of UmuD and DinB ................................................................. 32
  1.7 Molecular interactions of UmuD and UmuD′ with Lon and ClpXP proteases .......... 34
  1.8 CONCLUSIONS .......................................................................................................... 36
  1.9 REFERENCES .............................................................................................................. 36
Chapter 2: The Dimeric SOS Mutagenesis Protein UmuD is Active as a Monomer ............ 43
  2.1 INTRODUCTION ............................................................................................................ 44
  2.2 EXPERIMENTAL PROCEDURES ................................................................................. 48
    2.2.1 Bacteriological techniques ................................................................................... 48
    2.2.2 Proteins, strains, and plasmids ............................................................................. 48
    2.2.3 Determination of binding constant of UmuD N41D ............................................... 49
    2.2.4 Thermal shift assay .............................................................................................. 50
    2.2.5 UmuD *in vitro* cleavage assays ......................................................................... 50
    2.2.6 Immunoblotting .................................................................................................... 51
    2.2.7 UV survival and mutagenesis assays ................................................................. 52
    2.2.8 Heterodimerization of UmuD and UmuD′ .......................................................... 52
    2.2.9 Formaldehyde cross-linking of UmuD and β-clamp ............................................ 52
  2.3 RESULTS ......................................................................................................................... 53
    2.3.1 Design of UmuD variants ..................................................................................... 53
    2.3.2 UmuD N41D and UmuD′ N41D are monomers *in vitro* ....................................... 56
    2.3.3 N41D mutation significantly reduces the UmuD′ melting temperature .......... 58
    2.3.4 UmuD N41D monomer undergoes efficient cleavage ........................................ 60
    2.3.5 DinB inhibits UmuD N41D cleavage .................................................................... 62
    2.3.6 UmuD N41D and UmuD′ N41D are expressed as active monomers *in vivo* .... 62
Chapter 5: Characterization of newly identified forms of the polymerase manager protein UmuD ................................................................. 129

5.1 INTRODUCTION ............................................................................................................. 130

5.2 EXPERIMENTAL PROCEDURES .................................................................................. 133
5.2.1 Edman degradation of UmuD truncations ............................................................... 133
5.2.2 In vitro characterization of UmuD truncations ......................................................... 134
5.2.3 UV survival and mutagenesis assays .................................................................... 135
5.5.4 Immunoblotting.................................................................................................. 135
5.2.5 UmuD and DinB binding by Trp fluorescence ...................................................... 135

5.3 Results ......................................................................................................................... 136
5.3.1 Multiple UmuD protein bands observed during purification ....................... 136
5.3.2 Loss of N-terminal residues result in considerable change in UmuD arm characteristics ........................................................................ 137
5.3.3 UmuD 18 cleaves efficiently, UmuD 18 is not cleavable ................................ 139
5.3.4 UmuD 8 is proficient for UV-induced mutagenesis; UmuD 8 and UmuD 18 do not confer resistance to UV radiation ......................................................... 142
5.3.5 UmuD N-terminal arm conformation affects DinB interactions ................. 144

5.4 DISCUSSION ................................................................................................................. 148
5.5 REFERENCES ............................................................................................................... 150

Chapter 6: Future Directions ................................................................................................. 153

6.1 FURTHER ACKNOWLEDGEMENTS ..................................................................... 156

6.2 REFERENCES .............................................................................................................. 157
LIST OF FIGURES AND TABLES

Figure 1.1 Regulation of SOS-induced genes after DNA damage………………………….. 21
Figure 1.2 Model of full-length UmuD and crystal and NMR structures of UmuD’……….22
Figure 1.3 Polymerase switching in response to DNA damage…………………………… 23
Figure 1.4 Protein interaction sites on UmuD………………………………………………. 27
Figure 1.5 Residue substitutions in the β clamp that are implicated in interactions with UmuD, UmuD’ and α subunit of Pol III……………………………………………………………………………… 31
Figure 2.1 Homology models and structures of the UmuD proteins……………………. 55
Figure 2.2 UmuD N41D and UmuD’ N41D are monomers ................................. 57
Figure 2.3 Thermal shift assay result for wild-type UmuD and UmuD N41D…………… 60
Figure 2.4 UmuD Cleavage and DinB-dependent inhibition of cleavage of UmuD N41D in vitro is comparable to wild-type UmuD…………………………………………………………………… 61
Figure 2.5 UmuD expression levels and cleavage products in vivo. UmuD heterodimer formation under equilibrium conditions……………………………………………………………….. 64
Figure 2.6 UmuD and UmuD’ N41D are proficient for UV-induced mutagenesis and confer resistance to ultraviolet radiation............................................................... 66
Figure 2.7 UmuD N41D cross-links to the β-clamp as a monomer......................... 67
Figure. 3.1 UmuD and UmuD’ constructs......................................................... 80
Figure. 3.2 Least-squares fits to experimental spectra of different spin-labeled UmuD variants................................................................................................................................. 87
Figure 3.3 Spectrum of free IMSL in aqueous buffer at 297 K and least-squares fit to the lineshape............................................................... 89
Figure 3.4 A. Native gel showing IMSL-labeled wild-type UmuD, UmuD-S60A, UmuD3A and UmuD’-A31C, compared to their unlabeled counterparts, under non-denaturing conditions.

Figure 3.5 Left side: least-squares fits to spectra from spin-labeled UmuD-S60A at different temperatures.

Figure 3.6 Study of the temperature reversibility of the EPR spectrum of spin-labeled UmuD S60A.

Figure 3.7 A van’t Hoff plot for UmuD-S60A.

Figure 4.1 Homology models of UmuD and the structure of UmuD’ showing the location of amino acid substitutions constructed for FRET.

Figure 4.2 UmuD₂ (D) and UmuD’₂ (D’) homodimers exhibit slow exchange kinetics.

Figure 4.3 The heterodimer is the preferred UmuD protein conformer by FRET and native gel.

Figure 4.4 Melting of the UmuD protein homodimers and heterodimers.

Figure 4.5 Cleavage efficiency of the UmuD homodimer and the heterodimers.

Figure 4.6 Comparing cis and trans cleavage using UmuD active site and cleavage site variants.

Figure. 4.7 Model for the life cycle of the UmuD proteins.

Figure. 5.1 Two more UmuD derivatives have been identified: UmuD 8 and UmuD 18.

Figure 5.2 Characterization of UmuD 8 and UmuD 18 in vitro.

Figure 5.3 UmuD 8 cleaves as efficiently as wild-type UmuD, whereas UmuD 18 does not cleave to UmuD’.

Figure 5.4 UmuD 8 is proficient for UV-induced mutagenesis.
Figure 5.5 UmuD 8 and UmuD 18 do not confer resistance to UV radiation…………….. 144

Figure. 5.6 UmuD N-terminal arm conformation affects DinB interaction………………..145

Figure. 5.7 UmuD dimerization is not required for interaction with DinB………………… 147

Table 1.1 E. coli proteins that interact with the β-clamp via the β-binding pentapeptide motif QL[SD]LF or similar sequence……………………………………………………………………….. 29

Table 2.1 Strains and Plasmids……………………………………………………………………...... 47

Table 3.1 Dynamic parameters and scaling factors for the spin-labeled UmuD variants………………………………………………………………………………………………………….. 88

Table 3.2 Dynamic parameters and scaling factors for spin-labeled UmuD-S60A as a function of temperature…………………………………………………………………………………….. 93

Table 4.1 Strains and Plasmids……………………………………………………………………...... 105

Table 5.1 Strains and Plasmids……………………………………………………………………... 133

Table 5.2 Dissociation binding constants for UmuD proteins and DinB determined as by tryptophan fluorescence……………………………………………………………………... 148
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>Å</td>
<td>Angströms</td>
</tr>
<tr>
<td>η</td>
<td>Solvent viscosity</td>
</tr>
<tr>
<td>k</td>
<td>Boltzmann constant</td>
</tr>
<tr>
<td>τ</td>
<td>Protein rotational correlation time</td>
</tr>
<tr>
<td>Φ_{I,II,III}</td>
<td>Fraction of components</td>
</tr>
<tr>
<td></td>
<td>I (slow), II (medium), III (fast)</td>
</tr>
<tr>
<td>γ_e</td>
<td>Electronic gyromagnetic ratio</td>
</tr>
<tr>
<td>A_{x,y,z}</td>
<td>x,y,z spectral components of the A tensor</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine-5′-triphosphate</td>
</tr>
<tr>
<td>ASM</td>
<td>Active site mutant</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichoism</td>
</tr>
<tr>
<td>cP</td>
<td>Centipoise</td>
</tr>
<tr>
<td>CSM</td>
<td>Cleavage site mutant</td>
</tr>
<tr>
<td>CSU</td>
<td>Contacts of Structural Units</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DEAE</td>
<td>Diethylaminoethyl</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
</tbody>
</table>
\( E. \, coli \quad \text{Escherichia coli} \)

EDTA
Ethylene diaminetetraacetic acid

EPR
Electron Paramagnetic Resonance

FRET
Fluorescence Resonance Energy Transfer

\( \Delta G \)
Change in Gibbs free energy

\( G \)
Gauss

GHz
Gigahertz

\( g_{x,y,z} \)
x,y,z components of the g tensor

\( \Delta H_{I,II,III} \)
Gaussian inhomogeneous line-width parameters
  
  I (slow), II (medium), III (fast)

\( \Delta H \)
Change in enthalpy

Hepes
4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid

HX-MS
Hydrogen deuterium exchange mass spectrometry

IDP
Intrinsically Disordered Protein

IMSL
Iodomethyl spin label (3-iodomethyl-1-oxy-2,2,5,5-
tetramethylpyrroline)

J
Joule

K
Equilibrium constant

K
Kelvin (unit)

kDa
Kilo-Dalton

\( K_D \)
Equilibrium dissociation binding constant

\( k_{\text{obs}} \)
Association rate constant

kHz
Kilohertz
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>kJ</td>
<td>Kilojoule</td>
</tr>
<tr>
<td>min</td>
<td>Minutes</td>
</tr>
<tr>
<td>mol</td>
<td>Mole</td>
</tr>
<tr>
<td>mL</td>
<td>Milliliters</td>
</tr>
<tr>
<td>µM</td>
<td>Micromolar</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>mW</td>
<td>milliwatt</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular weight cutoff</td>
</tr>
<tr>
<td>nM</td>
<td>Nanomolar</td>
</tr>
<tr>
<td>ns</td>
<td>Nanoseconds</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>Pa</td>
<td>Pascal</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>pM</td>
<td>Picomolar</td>
</tr>
<tr>
<td>Pol</td>
<td>Polymerase</td>
</tr>
<tr>
<td>Oligo</td>
<td>Oligonucleotide</td>
</tr>
<tr>
<td>$R_{I, II, III}$</td>
<td>Isotropic rotational diffusion constants</td>
</tr>
<tr>
<td></td>
<td>I (slow), II (medium), III (fast)</td>
</tr>
<tr>
<td>ΔS</td>
<td>Change in entropy</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDSL</td>
<td>Site-directed spin-labeling</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single stranded DNA</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris[2-carboxyethyl] phosphine</td>
</tr>
<tr>
<td>TLS</td>
<td>Translesion synthesis</td>
</tr>
<tr>
<td>t&lt;sub&gt;1/2&lt;/sub&gt;</td>
<td>Half-time, half-life</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>T</td>
<td>Temperature</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VMD</td>
<td>Visual Molecular Dynamics</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Parts of this chapter have been published in:


1.1 The *umuDC* gene products are an integral part of the SOS response to DNA damage

Organisms are constantly bombarded by harmful DNA damaging agents that can lead to stalling of the replication machinery and cell death [1]. The tightly regulated bacterial SOS system is a stress-induced response to DNA damage and an integral part of UV-induced mutagenesis in *Escherichia coli* (*E. coli*) [1]. The first stage of this damage response involves relatively accurate DNA repair processes but as the response progresses, it shifts to a potentially mutagenic damage tolerance mode to ensure cell survival [1,2]. This switch from accurate DNA repair to mutagenic damage tolerance is regulated in part by the *umuD* gene products, UmuD\(_2\) and UmuD'\(_2\).

Together, UmuC and UmuD'\(_2\) form the Y family DNA polymerase V (UmuD'\(_2\)C), a low-fidelity DNA polymerase that has a specialized ability to copy damaged DNA in a process known as translesion DNA synthesis (TLS) [1,3,4].

Expression of the *umuDC* (“UV-mutable”) gene products is negatively regulated by the LexA repressor as part of the SOS transcriptional response. LexA binds to a sequence in the operator region of the genes [1,5,6]. Derepression of *umuDC* occurs when the RecA protein binds to single-stranded regions of DNA that develop at replication forks that are stalled by DNA damage [7]. The RecA/ssDNA nucleoprotein filament serves as a coprotease to facilitate cleavage of the LexA repressor (Figure 1.1). As the cellular concentration of LexA diminishes, the genes whose expression is normally repressed by LexA are transcribed[1].
Figure 1.1 Regulation of SOS-induced genes after DNA damage. The SOS response is induced by the formation of a RecA nucleoprotein filament on single stranded DNA (RecA*). This stimulates the auto-proteolysis of the LexA repressor which leads to the induction of at least 57 genes including umuDC, which encodes TLS DNA polymerase Pol V (UmuD′C). UmuD undergoes RecA* facilitated cleavage of its N-terminal 24 amino acids to yield UmuD′, the form that is active in translesion synthesis. Lon and ClpXP proteases play a role in regulating the levels of UmuD, UmuD′, and UmuC in the cell. ClpXP also specifically targets UmuD′ in UmuDD′ heterodimers (not shown, discussed later).

The umuDC genes are among the most tightly regulated SOS genes; the equilibrium dissociation constant (Kd) is 0.2 nM for LexA binding to the “SOS-box” in the promoter region [5]. In comparison, the Kd values for LexA binding to the “SOS-boxes” of the recA and lexA genes are estimated to be 2 nM and 20 nM, respectively [8]. Immunoblotting assays have shown the
cellular steady-state levels of UmuD₂ to be ~180 copies per uninduced cell and ~2400 copies per cell under SOS induction [9]. A single protein in a compartment with the volume of a typical *E. coli* cell is present at a concentration of ~1 nM. The level of UmuC is approximately 12-fold lower than UmuD₂ with about 15 molecules per cell in the absence of induction and ~200 molecules of UmuC per cell under SOS induced conditions [9]. Expression of the *umuDC* genes initially produces UmuD dimer (139 amino acids per monomer) which undergoes a RecA/ssDNA-stimulated autodigestion reaction after induction resulting in UmuD′₂ (115 amino acids per monomer) (Figure 1.1 and 1.2) [10,11]. UmuD₂ is the predominant species for the first approximately 20-40 min after SOS induction, after which UmuD′₂ is the predominant species [2]. UmuD proteins exist in solution as UmuD₂ and UmuD′₂ homodimers as well as the UmuD-UmuD′ heterodimer, which is more stable than either of the homodimers [12,13,14,15]. The *Kₜₐₙ* for UmuD₂ 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 [16]. From here on, UmuD₂ and UmuD′₂ will be simply referred to as UmuD and UmuD′.

*Figure 1.2* Model of full-length UmuD [17] and crystal [15] and NMR [14] structures of UmuD′. UmuD model shown in the *trans*, elbows down conformation (left). Crystal structure (middle) and NMR structure (right) of UmuD′ in *trans* conformation. The N-terminal arms of UmuD′ are cleaved between Cys24 and Gly25. Residues 1-24 are shown in magenta; residues 25 to 40 are shown in blue. Active site Ser60 and Lys97 are highlighted in red and green, respectively.
The two different forms of UmuD provide a temporal switch between accurate and mutagenic phases of the cellular response to DNA damage [2,18,19]. The combination of uncleaved UmuD and UmuC specifically decreases the rate of DNA replication and increases resistance of cells to killing by UV radiation [2,20]. Uncleaved UmuD$_2$C improves DNA damage survival by allowing time for error-free repair mechanisms to act before the combination of cleaved UmuD$'$$_2$ and UmuC (UmuD$'$$_2$C, Pol V) initiates potentially error-prone TLS (Figure 1.3)[2]. The combination of UmuC and noncleavable UmuD S60A significantly delay recovery of cell growth after UV radiation [2,19]. Therefore, a model for a umuDC-dependent DNA damage checkpoint in *E. coli* was proposed wherein a delay in DNA synthesis provides time for error-free nucleotide excision repair to remove DNA lesions. TLS is then enabled by the presence of UmuD$'$ [3,4,21]. This model suggests that the different *umuD* gene products, in combination with UmuC, are involved in distinct survival pathways after UV damage.

**Figure 1.3** Polymerase switching in response to DNA damage. Replicative DNA polymerases (purple) are generally unable to copy damaged DNA. A polymerase switch occurs allowing a TLS polymerase (green) access to DNA. The TLS polymerase synthesizes DNA opposite the lesion and far enough beyond it that the replicative polymerase can resume synthesis without disruption due to the lesion.

Increasing UmuD$'$$_2$C protein complex concentration was found to antagonize RecA-mediated recombination of a UV-damaged gene [22]; this effect was also observed *in vitro* [23]. In the
proposed model, high concentrations of the UmuD’2C proteins induce replisome switching from recombination to SOS mutagenesis. Indeed, the UmuD’2C complex has been shown to bind directly to the RecA/ssDNA filament and could disrupt the DNA pairing activity of RecA [24,25]. Notably, in the presence of homologous DNA sequences, homologous recombination repair is more prevalent than TLS in responding to DNA damage [26].

The mutagenic potential of Y family polymerases may be further regulated by preferential formation of heterodimers between UmuD and UmuD’, thereby depleting the cell of mutagenically active UmuD’ homodimers [12]. UmuD’ is degraded by the ATP-dependent protease ClpXP while in a heterodimeric complex with UmuD [27]. Formation of UmuDD’ heterodimers in preference to mutagenically active UmuD’ homodimers therefore specifically targets UmuD’ for proteolysis. UmuD also targets its UmuD homodimer partner for proteolytic degradation by ClpXP [28]. The ATP-dependent serine protease Lon is responsible for the degradation of both UmuD and UmuC proteins in vivo [29]. Targeted proteolysis of the umuD and umuC gene products is one mechanism for returning protein levels to their uninduced state.

1.2 Structural dynamics of UmuD and UmuD’

The umuD gene products interact with multiple replication factors such as DNA polymerases UmuC (as UmuD’2C, Pol V), DinB (Pol IV) and components of the Pol III holoenzyme. The latter include the polymerase subunit α, proofreading subunit ε, and the processivity clamp β [9,16,30,31,32]. These interactions are due in part to the relative flexibility of full-length UmuD and its UmuD’ cleavage product as shown biochemically and by X-ray crystallography, nuclear magnetic resonance spectroscopy (NMR), Electron Paramagnetic Resonance Spectroscopy
(EPR, discussed in Chapter 3), Hydrogen-deuterium exchange mass spectrometry (HX-MS) and circular dichroism (CD) [14,15,16,33,34]. The cleaved form UmuD’ contains disordered N-terminal arms that expose the C-terminal globular domain to solvent upon cleavage, while in full-length UmuD, the arms may be bound or unbound from the globular domain [14,33,34,35]. Therefore, UmuD and UmuD’ make specific contacts that facilitate a variety of protein-protein interactions [14,16,32,36].

The crystal structure of UmuD’ reveals 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 (Figure 1.2) [15]. Although two dimer interfaces (designated as molecular and filament) were observed in the crystal structure, NMR and cross-linking experiments support the conclusion that the so-called filament dimer interface is the form present in solution (Figure 1.2) [14,15,36,37,38,39,40]. However, there is evidence that the filament structure may be biologically relevant [15]. To date, there is no high resolution structure of the UmuD₂ homodimer. Cross-linking studies of a series of single-cysteine derivatives of UmuD are consistent with the UmuD₂ homodimer interface resembling the interface of the UmuD’₂ homodimer, involving contacts between the C-termini of the monomers and intermolecular interactions between Asn41 and Leu44 of α helix 1 [14,40]. NMR experiments also suggest that the UmuDD’ heterodimer most closely resembles the UmuD₂ homodimer [14].

Four models of the UmuD homodimer have been generated from NMR, Electron Paramagnetic Resonance (EPR), and cross-linking studies, and by homology to LexA [14,17,41,42]. 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 1.2). Each UmuD monomer cleaves the N-terminal arm of its partner at Cys24-Gly25 [17]. A trans, elbows up model 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 within the dimer[17]. Indeed, evidence that the UmuD N41D variant is active as a monomer supports a scenario in which cleavage in cis is possible and efficient (discussed in Chapter 2) [43]. Cross-linking and chemical modification experiments suggest that the trans, elbows down conformation of the N-terminal domain is the most prevalent in solution [17,41], however, given the dynamic nature of UmuD, all four conformations may be physiologically relevant [16,17,33,34,43].

Circular dichroism (CD) spectroscopy simulating physiological conditions detected random coil conformations for both UmuD dimers [16], rather than the β-sheet-rich structure determined by X-ray crystallography and NMR spectroscopy [14,35,36]. 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) [44]. Like their IDP counterparts, UmuD dimers are capable of making a remarkable number of specific protein-protein contacts. UmuD₂ and UmuD’₂ make a number of distinct protein-protein contacts with considerable functional implications (Figure 1.4).

Both UmuD₂ and UmuD’₂ interact with the RecA:ssDNA nucleoprotein filament, Y family DNA polymerases UmuC and DinB, the α, β, and ε subunits of the replicative DNA Pol III, and proteases Lon and ClpXP [1,11,17,29,31,32,41,45]. UmuD₂ strongly interacts with the β processivity clamp while UmuD’ preferentially interacts with the α-catalytic subunit [46].
UmuD₂ also prevents DinB-induced -1 frameshift mutations [47], whereas UmuD'₂ activates UmuC for TLS [1,48]. Degradation of UmuD₂ is carried out by the Lon protease [45]. Also, UmuD delivers either its UmuD or UmuD' partner to ClpXP for degradation [49]. The multiple interactions of UmuD₂ and UmuD'₂ are critical for regulating mutagenesis in *E. coli*.

**Figure 1.4** Protein interaction sites on UmuD. (a) The β-clamp interacts with residues 14-19, 24, 52 and 126 (blue) [50]. RecA interacts with residues 34, 81, 57, 67 and 112 (cyan) [39]. DinB interacts with residue 91 on UmuD (green) [31]. (b) ClpXP interacts with residues 9-12, 33-37, 41-51, 85-109 (red) [28]. Lon binds to regions close to the residues that are important for interaction with ClpXP, residues 15-19 (violet) [45].

**1.3 The *umuD* gene products interact with the α, ε, β subunits of DNA Pol III**

DNA replication requires the coordination of many different proteins to accomplish the goal of simultaneous replication of the two antiparallel stands of DNA. 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 *E. coli* cell cycle.

DNA replication is carried out by the dynamic Pol III holoenzyme comprised of at least seventeen subunits: two or three [51] αεθ polymerase cores, two or three β processivity clamps, and the δδ'γτ₃ψχ clamp loader complex [52]. Processive replication of the leading and lagging
strands is completed by the polymerase core, and requires interaction with the ring-shaped β clamp [53,54], which encircles DNA and decreases the dissociation of the polymerase core from the DNA template [55,56,57].

UmuD₂ and UmuD’₂ differentially interact with the α, β, and ε components of Pol III in the replisome [32]. Interestingly, UmuD₂ interacts more strongly with β than with α, whereas cleaved UmuD’₂ interacts more strongly with α than β [32]. UmuD₂ and UmuD’₂ also physically interact with the C-terminal domain of the 3’-5’ ε exonuclease subunit, with which the α subunit also interacts [32,58,59]. It has been suggested that these interactions facilitate higher-order regulation and that the *umuD* gene products are central to the role in coordinating the replication machinery in response to DNA damage [32].

DNA replication is highly efficient, with up to 1000 nucleotides added per second, and an error frequency of less than 1 in 100,000 additions in *E. coli* [60,61]. However, Pol III cannot replicate damaged DNA, which can often take the form of bulky lesions that may distort the double helix [1,62]. Specialized Y-family DNA polymerases with more open, flexible active sites are able to copy damaged DNA in a process known as Translesion Synthesis (TLS) [63]. These specialized polymerases, which include Pol V (UmuD’₂C) and Pol IV (DinB), are induced as part of the SOS DNA damage stress response [64,65].

### 1.4 UmuD-Beta clamp interactions

To date, 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) (Table 1.1) [66]. UmuD contains a $^{14}\text{TFPLF}^{18}$ sequence within its N-terminal arm [66]. Although the motif lies in a region of UmuD that is important for its interaction with the β clamp[50], the interaction does not depend on the sequence identity of the motif [30]. That is, a UmuD variant containing mutations in this motif binds to the β clamp with similar affinity as that of wild-type UmuD (see below) [17].

**Table 1.1 E. coli** proteins that interact with the β-clamp via the β-binding pentapeptide motif QL[SD]LF or similar sequence [66].

<table>
<thead>
<tr>
<th>β interacting proteins</th>
<th>β-binding sequence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>UmuD</td>
<td>$^{14}\text{TFPLF}^{18}$</td>
<td>[67,68]$^1$</td>
</tr>
<tr>
<td>DNA Pol V (UmuC)</td>
<td>$^{55}\text{QLNLF}^{56}$</td>
<td>[69,70]</td>
</tr>
<tr>
<td>DNA Pol IV (DinB)</td>
<td>$^{346}\text{QLVLGL}^{351}$</td>
<td>[69,71,72]</td>
</tr>
<tr>
<td>DNA Pol II (Pol B)</td>
<td>$^{79}\text{QLGLF}^{83}$</td>
<td>[69]</td>
</tr>
<tr>
<td>DNA Pol III (α-subunit)</td>
<td>$^{920}\text{QADMF}^{924}$</td>
<td>[73]</td>
</tr>
<tr>
<td>δ-subunit Clamp Loader</td>
<td>$^{70}\text{AMSLF}^{74}$</td>
<td>[74]</td>
</tr>
<tr>
<td>MutS</td>
<td>$^{812}\text{QMSLL}^{816}$</td>
<td>[75]</td>
</tr>
<tr>
<td>Hda</td>
<td>$^{6}\text{QLSLPL}^{11}$</td>
<td>[76]</td>
</tr>
</tbody>
</table>

$^1$Although these residues reside in an important region for interactions with the β clamp, their identity is not required for UmuD to interact with β.

UmuD and UmuD′ affinity chromatography and *in vitro* cross-linking studies confirm that the β clamp has a higher affinity for UmuD than UmuD′[50]. 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 [50]. UmuD lacking its N-terminal nine residues is proficient for interactions with the β clamp, while UmuD lacking the N-terminal 19 residues results in reduced formaldehyde
cross-linking to β [50]. The UmuD variant UmuD 3A (T14A L17A F18A), a non-cleavable variant with mutations of the most conserved residues of the TFPLF motif, possesses some of the biological functions of the cleaved form UmuD′. Although the $K_d$ values for interaction of the β clamp with UmuD (5.5 +/- 0.8 μM) and UmuD 3A (6.1 +/- 0.5 μM) are similar, their interactions with the β clamp may be different, as observed by intrinsic tryptophan fluorescence of the β clamp [17]. Tryptophan fluorescence is a relatively sensitive probe of the microenvironment. The single tryptophan of the β clamp is located on a flexible loop between domains I and II of β, and thus is a sensitive reporter of conformational changes [17,54]. The β clamp tryptophan fluorescence differed dramatically upon UmuD vs. UmuD 3A binding, suggesting they have different binding modes [17].

UmuD, UmuD′, the α catalytic subunit, UmuC, DinB and clamp loader all interact with the β clamp around the β clamp hydrophobic pocket, approximately defined by residues Leu177, Pro242, Val247, Val360, Met362 (Figure 1.5) [17,71,77,78,79]. UmuD, UmuD′ and the α subunit interact with overlapping regions of β, suggesting that there may be competition for binding (Figure 1.5) [32]. This implies that UmuD plays a regulatory role following the SOS response, interacting with components of Pol III, interfering with α binding to β, slowing replication and allowing time for error-free repair mechanisms to act [2,32]. It is also possible that α and UmuD or UmuD′ bind the homodimeric β clamp simultaneously. A model has been proposed in which cleavage of UmuD to form UmuD′ reduces binding to the β clamp, thereby releasing the DNA damage checkpoint and enabling translesion synthesis [80].
Figure 1.5 Residue substitutions in the β clamp that are implicated in interactions with UmuD (left), UmuD’ (middle) and α subunit of Pol III (right) [81]. 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. The hydrophobic channel is shown in brown (residues Leu177, Pro242, Val247, Val360, Met362) [78] [71], while the rim interaction residue Leu98 is shown in black. Structures were generated using VMD [82] and coordinates for β (2POL) from the PDB [54].

1.5 Molecular interactions of UmuD and UmuC

Due to the difficulty in acquiring large quantities of pure, active UmuC and Pol V, protein interaction studies have been somewhat limited, especially considering that the umuC gene was identified in the 1970s. However, the physical interaction between UmuD’ and UmuC was confirmed using immunoprecipitation, yeast two-hybrid assay and glycerol gradient analysis [13,83]. Additionally, the interaction between full-length UmuD and UmuC was verified by using affinity chromatography and velocity sedimentation in glycerol gradients, but not immunoprecipitation from cell extracts [13]. From this, it was concluded that UmuC associates strongly with UmuD’ in vivo, whereas in vitro, UmuC interacts efficiently with both forms of the umuD gene products [13]. The likely stoichiometry was determined to be one UmuC with either a dimeric UmuD or UmuD’ [13]. UmuD and UmuD’ appear to interact with the C-terminus of
Umuc, as a Umuc construct lacking its C-terminal 25 residues showed dramatically reduced binding to both UmuD and UmuD′ [84]. In addition to the UmuD and UmuD′ homodimers, Umuc also interacts with the UmuDD′ heterodimer, which acts to inhibit SOS mutagenesis, possibly by titrating out the dimeric UmuD′ species that is active in TLS [10,13,83,85].

1.6 Molecular interactions of UmuD and DinB

The dinB (“damage-inducible”) gene encoding DNA Pol IV (DinB) was discovered in a screen using reporter fusions to identify DNA damage-inducible genes [86]. DinB (Pol IV) is the other Y family lesion bypass polymerase in E. coli and is the only Y family polymerase that is conserved throughout all domains of life [87,88]. The expression level of chromosomal DinB under DNA damaging conditions is 6-12 times higher than that of Umuc or PolB (DNA Pol II) with about 2500 molecules of DinB in an SOS induced cell [89]. DinB is also found on the recombinant F′ plasmid that was constructed to determine mutation spectra of specific revertible lac′ alleles [89,90]. The expression level of DinB in an uninduced state from the F′ plasmid in E. coli strain CC108 is approximately 750 molecules, as compared to 250 molecules expressed from the chromosome in the absence of SOS induction [89]. DinB has a misincorporation error frequency of 10⁻³ – 10⁻⁵ [91]. Unlike UmuD′C, DinB elongates templates with bulged structures causing potentially deleterious -1 frameshift mutations [87,89].

Umud, Umud′ and RecA play important roles in the regulation of DinB and direct physical interactions between DinB and Umud, Umud′ and RecA have been detected [31]. Although this may have initially seemed surprising, the expression levels of Umud (180 molecules uninduced; 2400 molecules in SOS induced cells) and DinB (250 molecules uninduced; 2500 molecules in
SOS induced cells) before and after SOS induction align [9,89]. The stoichiometry of the complex was found to be one DinB molecule to one UmuD₂ dimer [31]. DinB and UmuD₂ bind with a $K_D$ of 0.62 µM [31]. It was also determined that DinB, RecA and UmuD₂ form a stable ternary complex under physiological conditions in vitro [31]. Genetic and biochemical analysis show that full-length UmuD as well as the non-cleavable UmuD variant UmuD S60A strongly inhibit the -1 frameshift mutator effect of DinB [31]. UmuD and UmuD’ also inhibit DinB activity in adaptive mutagenesis [31]. Pre-steady state kinetics experiments led to the proposal that DinB bound to DNA containing a repetitive sequence is in equilibrium between a template slipped conformation, which leads to frameshift mutagenesis, and a non-slipped conformation [92]. UmuD appears to prevent DinB-dependent frameshift mutagenesis by favoring the non-slipped conformation upon binding to DinB [92]. UmuD also modulates DinB function by facilitating efficient extension of correctly paired primer termini while blocking extension of mismatched termini [31,92].

Using peptide array mapping and structural homology models of both DinB and UmuD, it was proposed that UmuD interacts with several hydrophobic residues on the surface of DinB in the thumb and finger domains. DinB residue F172 in the thumb domain was identified as a likely site of interaction with UmuD. Indeed, DinB F172A has lower affinity for UmuD ($K_D$ reduced ~56-fold) and exhibits less UmuD-dependent -1 frameshift suppression in vivo and in vitro than wild-type DinB [31]. The DinB interacting surface on UmuD is a discontinuous surface when mapped onto a model of trans-UmuD [17,41]. Alternatively, isoenergetic models of UmuD in which the N-terminal arms are in a non-cleavable conformation provide alternative interacting surfaces across the side of UmuD [17]. UmuD D91, on the outer surface of UmuD, was proposed
as a likely residue to be important for interaction with DinB (Figure 1.3). UmuD D91A has reduced affinity for DinB ($K_D$ reduced by over 24-fold) and dramatically reduced suppression of -1 frameshift mutagenesis compared to wild-type UmuD [31]. This suggests that there may be multiple biologically relevant conformations of UmuD that can interact with DinB or other polymerases [16,17,34]. These interactions may aid in the suppression of frameshift mutagenesis by blocking the open active site that is needed to elongate bulged templates [31,93,94,95]. By creating a ternary complex model of DinB, UmuD$_2$ and RecA, it was suggested that UmuD$_2$ and RecA work together in restricting the open active site of DinB thereby preventing -1 frameshift mutagenesis on bulged templates [31,92]. Therefore, the presence of full-length UmuD actually enhances accurate TLS by DinB while suppressing extension of bulged templates that would cause frameshift mutagenesis.

### 1.7 Molecular interactions of UmuD and UmuD$'$ with Lon and ClpXP proteases

Regulation of UmuD protein levels by ClpXP and Lon proteases is an important part of the SOS response to DNA damage. Proteolytic degradation of the $umuD$ gene products is involved in cessation of SOS mutagenesis [1,29,96]. ClpXP is composed of the ATP-dependent unfoldase ClpX hexamer, and the double-ringed, 14-subunit serine protease, ClpP [97,98,99,100,101]. The domain structure of the Lon protease is quite similar in that it contains an ATPase domain, a sensor and substrate discrimination domain (SSD) and a protease domain [102]. The mechanism of degradation begins when ClpX unfolds the substrates using repeated cycles of ATP hydrolysis and translocates the unfolded peptide into the ClpP chamber where proteolysis occurs. Substrate recognition involves the N- or C-terminal regions of the target protein binding to the substrate-processing site on ClpX [103,104]. These signals may become apparent after cleavage, as in the
case of LexA, or upon a conformational change in the target protein [105,106]. However, the addition of an 11-amino acid (AANDENYALAA) ssrA tag to improperly translated nascent polypeptides will result in direct targeting to ClpXP for degradation [99,100,107,108,109,110]. This C-terminal ssrA tag is encoded by the ssrA transfer mRNA and is added cotranslationally to proteins translated without an in-frame stop codon [109,110]. In addition, substrate recognition by ClpXP involves the interaction of tethering sites with adaptor proteins. These adaptor proteins are not degraded themselves but work to enhance the degradation of the target protein [111,112]. One example is the SspB-mediated degradation of ssrA tagged protein. Here, one part of the target protein binds the tethering site on ClpX while the SspB protein interacts with the ssrA tag enabling efficient delivery to ClpXP for degradation [113,114].

Similar to SspB-facilitated degradation of ssrA-tagged target proteins, UmuD’ is a substrate for ClpXP but is only degraded when bound to full-length UmuD [27,28]. Therefore, the preferential formation of UmuDD’ heterodimer specifically leads to a decrease in the steady-state levels of UmuD’ in vivo [28]. Although the residues found within the N-terminal 24 amino acids of UmuD serve as the degradation signal for ClpXP degradation of UmuD’, UmuD serves as an adaptor and is not itself degraded [27]. UmuD also serves as an adaptor in the context of UmuD2 homodimers, leading to degradation of one UmuD in the dimer [28]. UmuD residues 9-12 are necessary for UmuD’ instability and therefore protease recognition (Figure 1.3) [27]. Amino acids 15-19 of UmuD are also implicated in the degradation of the UmuDD’ heterodimer by ClpXP (Figure 1.3) [27]. On the other hand, while residues 15-19 are also important for Lon-mediated degradation of UmuD, residues 9-12 are not involved in recognition by Lon[45]. ClpXP recognition sites can also be found on the surface of UmuD’, in particular, residues 33-37,
41-51 and 85-109 were found to interact robustly with ClpXP (Figure 1.3) [27]. The UmuD-facilitated degradation of UmuD’ can be impeded by the SspB-tethering peptide and the SspB-tethering motif is interchangeable with the sequence in UmuD. Because the N-terminal domain of ClpX mediates interactions with both SspB and UmuD, it was determined that UmuD acts as a ClpX delivery factor that is critical in tethering itself and UmuD’ to ClpX. This seems to be a primary mechanism for bringing SOS mutagenesis to an end [115].

1.8 CONCLUSIONS

The UmuD proteins play a central role in coordinating DNA replication and the DNA damage response. It is therefore important to understand how these proteins are involved in such a finely-tuned network of interactions with specific binding hierarchies. This can be accomplished in part by studying the structural and conformational dynamics, conformational-dependent cleavage activity, as well as dimer exchange kinetics of the UmuD proteins.

1.9 REFERENCES


764.
Microbiol 60: 231-253.
protein-protein interaction motif in the eubacterial DNA replication and repair systems.
variant that acts as a UmuD′ mimic. Journal of Biological Chemistry 281: 9633-9640.
68. Sutton MD, Narumi I, Walker GC (2002) Posttranslational modification of the umuD-
encoded subunit of Escherichia coli DNA polymerase V regulates its interactions with
synthesis and mutagenesis in E. coli cells. DNA Repair (Amst) 1: 703-708.
differentially alter the dual activities of UmuC. Molecular Microbiology 59: 460-474.
71. Bunting KA, Roe SM, Pearl LH (2003) Structural basis for recruitment of translesion DNA
polymerase pol IV/DinB to the beta-clamp. EMBO J 22: 5883-5892.
Biochemical Analysis of Sliding Clamp/Ligand Interactions Suggest a Competition
73. Dohrmann PR, McHenry CS (2005) A bipartite polymerase-processivity factor interaction:
only the internal β binding site of the α subunit is required for processive replication by
processivity clamp opening by the δ subunit wrench of the clamp loader complex of E.
75. Lopez de Saro FJ, O’Donnell M (2001) Interaction of the β sliding clamp with MutS, ligase,
77. Duzen JM, Walker GC, Sutton MD (2004) Identification of specific amino acid residues in
the E. coli β processivity clamp involved in interactions with DNA polymerase III,
UmuD and UmuD′. DNA Repair (Amst) 3: 301-312.
processivity clamp opening by the delta subunit wrench of the clamp loader complex of E.
synthesis and mutagenesis in E. coli cells. DNA Repair 1: 703-708.
80. Sutton MD, Duzen JM (2006) Specific amino acid residues in the beta sliding clamp
establish a DNA polymerase usage hierarchy in Escherichia coli. DNA Repair 5: 312-
323.


Chapter 2: The Dimeric SOS Mutagenesis Protein UmuD is Active as a Monomer

This research was originally published in the Journal of Biological Chemisty as:

2.1 INTRODUCTION

Organisms are constantly bombarded by harmful DNA damaging agents that can lead to stalling of the replication machinery and cell death [1]. The tightly regulated bacterial SOS system is a stress-induced response to DNA damage and is an integral part of UV-induced mutagenesis in Escherichia coli (E. coli) [1]. The first stage of this damage response involves relatively accurate DNA repair processes but as the response progresses, it shifts to a potentially mutagenic damage tolerance mode to ensure cell survival [1,2]. This switch from accurate DNA repair to mutagenic damage tolerance is regulated in part by the umuD gene products. The UmuD_2 homodimer is composed of 139-amino acid subunits and is expressed 20-30 minutes after the induction of the SOS response [1-3]. Interaction with the damage-induced RecA:ssDNA nucleoprotein filament facilitates the slow self-cleavage and removal of the N-terminal 24 amino acids of UmuD_2, yielding the C-terminal homodimer of 115 amino-acid subunits, UmuD'_2. Together, UmuC and UmuD'_2 form the Y family DNA polymerase V (UmuD'_2C), a low-fidelity DNA polymerase that has a specialized ability to copy damaged DNA in a process known as translesion DNA synthesis (TLS) [1,4,5].

UmuD_2 and UmuD'_2 make a number of distinct protein-protein contacts with considerable functional implications. Both UmuD_2 and UmuD'_2 interact with the RecA:ssDNA nucleoprotein filament, Y family DNA polymerases UmuC and DinB, the α, β, and ε subunits of the replicative DNA pol III, and proteases Lon and ClpXP [1,6-12]. UmuD_2 strongly interacts with the β processivity clamp while UmuD'_2 preferentially interacts with the α-catalytic subunit [12]. UmuD_2 also prevents DinB-induced -1 frameshift mutations [7], whereas UmuD'_2 activates UmuC for TLS [1,6]. Degradation of UmuD_2 is carried out by the Lon protease [10]. Also,
UmuD delivers either its UmuD or UmuD’ partner to ClpXP for degradation [13]. The multiple interactions of UmuD_2 and UmuD’_2 are critical for regulating mutagenesis in *E. coli*.

The structural flexibility of UmuD_2 and UmuD’_2 dimers permits a broad range of interactions [6,14-17]. For full-length UmuD_2, the N-terminal 39 amino acid arms are relatively stably bound to the globular C-terminal domain to produce a distinct binding surface [6,15]. Upon cleavage of the N-terminal 24 amino acids, the remaining 15 amino acids of the arm appear unbound from the C-terminal domain and are quite disordered [14,16]. This leaves the globular domain solvent exposed and available for interaction with a variety of proteins [6,15].

Whereas X-ray [16] and NMR [15] structures indicate that UmuD’_2 has an overall β-sheet fold, circular dichroism (CD) [17] experiments show that both the UmuD_2 and UmuD’_2 dimers resemble a random coil under physiological conditions. Additionally, the NMR and X-ray structures of UmuD’_2 are substantially different, with the active site of UmuD’_2 correctly formed and poised for catalysis only in the X-ray structure [15,16]. These structural differences highlight the plasticity of UmuD’_2, and have resulted in the classification of the umuD gene products as intrinsically disordered proteins (IDPs) [17]. Many IDPs play key roles in regulation despite their lack of a well-defined structure [18-20].

Although attempts at high-resolution structures of UmuD_2 have been unsuccessful, four isoenergetic models of full-length UmuD_2 have been proposed wherein the N-terminal arms are in the *trans* or *cis* conformations, with the elbows up or down [8]. In the *trans* elbows down version, the N-terminal arm of one monomer loops down across the globular C-terminal domain
of its partner where it crosses the catalytic site [8]. The model of UmuD₂ in the trans, elbows up conformation shows each arm bound to the outer edge of the C-terminal domain potentially allowing both regions of the protein to be solvent exposed [8]. For the UmuD₂ models in cis, the elbows up or elbows down conformations suggest that the arms can bind their own globular domains [8]. The proposed models are also consistent with cross-linking experiments completed at physiologically relevant concentrations [21,22]. Additionally, the relatively efficient cross-linking of mono-cysteine derivatives at position 37 and 38 of UmuD₂ implies that these residues are close to the dimer interface [21]. However, it is intriguing that the identical residues do not cross-link in the case of UmuD₂′, possibly suggesting a greater degree of structural flexibility relative to full-length UmuD [21]. Taken together, these results suggest that the UmuD proteins may adopt multiple conformations in solution.

In an effort to learn more about the structural dynamics and functions of UmuD proteins, we set out to create variants that were defective in dimerization. Such a variant would not only answer the question as to whether UmuD₂ is active in the cis conformation, but also address the possibility that UmuD may be functionally active as a monomer. Although models of UmuD₂ with the arms in the cis conformation have been proposed, evidence that this conformation is physiologically relevant has been lacking to date. We focused on the region near the α-helix composed of residues 39 and 44, as it has been proposed to be important for UmuD₂ dimerization [21-23]. We also considered the importance of α-helices in protein stability as in the case of leucine zipper dimerization, transmembrane helix interactions and numerous other examples [24,25]. With this in mind, we hypothesized that disrupting the contacts between the α-helical regions of UmuD₂ would result in a reduction in dimerization efficiency. We generated a
single point mutation, N41D, that is likely to not only alter the hydrogen bonding network, but also to produce repulsive interhelical electrostatic interactions. In this work, we have found that the single amino acid substitution N41D is sufficient to shift the dimer-monomer equilibrium of UmuD significantly to the monomer form. This single mutation also renders UmuD’ a monomer. We find that this variant of UmuD is a monomer under most conditions, is active in autocleavage, and is proficient for UV-induced mutagenesis.

Table 2.1 Strains and Plasmids

<table>
<thead>
<tr>
<th>Strains and Plasmids</th>
<th>Relevant Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td>GW8017</td>
<td>AB1157 ΔumuDC</td>
<td>[21]</td>
</tr>
<tr>
<td>PB103</td>
<td>AB1157 ΔumuDC ΔrecJ</td>
<td>P1 (JW2860)→ GW8017 [26]</td>
</tr>
<tr>
<td>BL21 DE3</td>
<td></td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGY9738</td>
<td>o1εumuD′C; pSC101-derived, SpecR</td>
<td>[27]</td>
</tr>
<tr>
<td>pGY9739</td>
<td>o1εumuD′C; pSC101-derived, SpecR</td>
<td>[27]</td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector; pSC101-derived, SpecR</td>
<td>[28]</td>
</tr>
<tr>
<td>pSG4</td>
<td></td>
<td>[29]</td>
</tr>
<tr>
<td>pSG5</td>
<td></td>
<td>[29]</td>
</tr>
</tbody>
</table>
2.2 EXPERIMENTAL PROCEDURES

2.2.1 Bacteriological techniques

The *E. coli* strains and plasmids used for this study are listed in Table 2.1. The operator sequences of pGY9739 and pGY9738 contain the *o1c* mutation where a single base substitution leads to modestly increased expression of *umuD* and *umuC* [27,30]. Strains were grown in Luria broth at 37 ºC supplemented with spectinomycin (60 µg/mL) or ampicillin (100 µg/mL).

Competent cells were prepared using the CaCl$_2$ method [31]. UmuD and UmuD’ N41D were constructed using a QuikChange kit (Stratagene). Mutations were confirmed by DNA sequence analysis (Massachusetts General Hospital Core Facility, Cambridge, MA). Mutagenic primer sequences are as follows:

Forward: 5′-GCGCATCGATCTGGATCAACTGTTGATCC
Reverse: 5′-GGATCAACAGTTGATCCAGATCGATGC

2.2.2 Proteins, strains, and plasmids

UmuD N41D and UmuD’ N41D expression plasmids were constructed in pSG5 and pSG4.

Expression of UmuD and UmuD’ proteins was accomplished as previously described [29]. Cells were harvested and UmuD and UmuD’ proteins were purified according to published methods [32]. The β-clamp was also purified using the method published for UmuD and UmuD’ [29]. DinB and RecA proteins were purified as described [29,33].

*Native PAGE*—Purified wild-type UmuD$_2$ and UmuD’$_2$ (10 µM) as well as UmuD N41D and UmuD’ N41D (10 µM) were each incubated in non-denaturing sample buffer (NDS, 62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 0.01% bromophenol blue) for 5 min, with or without the
addition of 10 mM dithiothreitol (DTT). Electrophoresis was carried out by 10-20% polyacrylamide gel electrophoresis (PAGE) (Lonza) using non-denaturing electrode buffer (NDE, 25 mM Tris-Base, 192 mM Glycine) at a constant 120 V at room temperature until proteins were resolved. The gel was then stained with Sypro Ruby (Molecular Probes), which has a detection limit of 0.25-1 ng for most proteins. Fluorescence was detected with a Storm 860 phosphoimager using excitation wavelength of 635 nm. UmuD monomer and dimer bands were quantified using ImageQuant (GE Healthcare).

2.2.3 Determination of binding constant of UmuD N41D
Concentrations of UmuD N41D and UmuD’ N41D ranging from 5-200 µM were combined with 10 mM DTT and NDS buffer in preparation for native PAGE. Wild-type UmuD2 and UmuD’2 at concentrations of 10, 100 and 200 µM were included as controls. Protein bands were detected with Sypro Ruby and quantified as described above. The calculations used to determine the $K_D$ for homodimerization (including equations and derivations) have been previously described [34].

The following equation was applied to the data set to obtain the $K_D$ for homodimerization:

$$\frac{[AA]}{[A]_T} = \frac{1}{2} + \frac{K_D - (K_D^2 + 8K_D[A]_T)^{0.5}}{(8[A]_T)}$$

(1)

In this equation, $[A]_T$ total concentration of A added to the reaction and $[AA]$ is the concentration of dimers. The $K_D$ was obtained by generating a best fit curve for the average of three trials using nonlinear regression curve fitting in GraphPad Prism.
2.2.4 Thermal shift assay

An optical 96-well reaction plate (Applied Biosystems) was used to analyze 16 µL reaction volumes. Each well contained UmuD $T_m$ buffer (50 mM Hepes, pH 7.5 and 100 mM NaCl), 45 µM UmuD protein and 25X Sypro Orange (Molecular Probes). The plates were sealed using optical adhesive film (Applied Biosystems). An iCycler iQ5 Real-Time PCR detection system (Bio-Rad) was used to heat the plate from 10 to 70 °C in 0.1 °C increments. Changes in the fluorescence were monitored concurrently with a charge-coupled device (CCD) camera. The temperature midpoint for the protein unfolding transition, $T_m$, was calculated by fitting the fluorescence data to a Boltzmann model using curve-fitting program Microsoft Excel XLfit5 add-on program (ID Business Solutions) [35]:

$$I = \left( A + \frac{(B - A)}{1 + e^{(T_m - T)/C}} \right)$$

In this equation, $I$ is the measured fluorescence intensity at temperature $T$. The pretransitional and posttransitional fluorescence intensities are denoted $A$ and $B$, respectively, and the slope factor is denoted $C$. Fitting did not include data points after the fluorescence intensity maximum.

2.2.5 UmuD in vitro cleavage assays

RecA:ssDNA nucleoprotein filament-dependent cleavage of wild-type UmuD$_2$ and UmuD N41D at concentrations of 0.5, 1, 2, 5 and 10 µM was assayed [21,29,36] for 60 min in LG buffer at 37 °C. The addition of sodium dodecyl sulfate (SDS)-PAGE loading buffer (25 mM Tris-HCl, 5% glycerol, 2% SDS, 1.25% β-mercaptoethanol, 0.1% bromophenol blue) was used to quench the
reaction before analysis by 4-20% SDS-PAGE (Pierce). Alkaline cleavage was carried out as previously reported [29]. Reactions were incubated at 37 ºC for 48 h. Cleavage products were analyzed by 4-20% SDS-PAGE (Pierce). Proteins were detected by staining with Sypro Ruby as described above.

To assay the inhibition of UmuD cleavage by DinB, wild-type UmuD2 and UmuD N41D at concentrations of 2 µM were incubated at room temperature for 1 h in the absence of wild-type DinB or in the presence of 2 µM or 4 µM wild-type DinB in LG buffer. In a separate reaction, 3.15 µM RecA, 0.35 µM 24-mer DNA oligo, and 0.68 µM ATPγS were incubated at room temperature for 1 h. The reactions were combined and incubated at 37 ºC for 1 h. The 20-µL reaction was then quenched with 5 µL of 4X SDS-PAGE loading buffer before analysis by 4-20% SDS-PAGE (Pierce). Sypro Ruby was used for protein detection as described above.

2.2.6 Immunoblotting

The level of UmuD and UmuD′ expression from the low-copy plasmids pGY9739 and pGY9738, respectively, in GW8017 was determined by immunoblotting. Induction of expression with UV-irradiation was accomplished as previously described [32]. For experiments using denaturing conditions, cell pellets were resuspended in 50 µL of 0.85% saline and 50 µL SDS-PAGE buffer. Cells were lysed by heating for 20 min at 95 ºC before loading 15 µL aliquots onto 4-20% SDS-PAGE gels (Pierce). Electrophoresis was carried out using Tris-HEPES-SDS running buffer (100 mM Tris base, 100 mM HEPES, 1% SDS) at 120 V for 1 h 15 min. These assays were carried out with loading amounts appropriate to give a signal within the linear range of quantification (Fig 2.5D).
For experiments carried out under non-denaturing conditions, a gentle lysis BugBuster protein extraction reagent (Novagen) was used per manufacturer directions. For native PAGE, 10 µL of supernatant was added to 20 µL of NDS buffer and half of the reaction was loaded onto 10-20% polyacrylamide gels (Lonza). Electrophoresis was accomplished using NDE buffer at 120 V for 3.5 h. Immunoblotting experiments were completed as previously described [32].

2.2.7 UV survival and mutagenesis assays
Survival and mutagenesis assays were performed as previously described [29,32]. The data represent an average of at least three trials and error bars represent the standard deviation.

2.2.8 Heterodimerization of UmuD and UmuD′
The heterodimerization experiment was carried out by incubating a mixture of UmuD and UmuD’ proteins at 10 µM in LG buffer and 10 mM DTT for 15 min. NDS buffer was added directly to the reaction for a total volume of 25 µL. Electrophoresis was carried out by 10-20% native PAGE (Lonza) using NDE buffer until proteins were resolved. Bands were detected and quantified as previously described.

2.2.9 Formaldehyde cross-linking of UmuD and β-clamp
Formaldehyde cross-linking reactions were completed by incubating 10 µM UmuD and Beta protein in X-link buffer (20 mM Hepes, pH 7.5, 50 mM NaCl, 0.1 mM EDTA) and 80 mM formaldehyde for a total reaction volume of 12 µL. The reactions were incubated for 45 min at room temperature before the addition of NDS buffer. Proteins were resolved by 10-20% native
PAGE (Lonza) for 5 h at 120 V. Sypro Ruby staining was then used to detect cross-linked proteins.

2.3 RESULTS

2.3.1 Design of UmuD variants

Previous investigations into the conformation of the N-terminal arms of UmuD$_2$ have concluded that while the arms are likely to be dynamic, the active form of this protein is *trans*, where the arm of one monomer is bound to the globular C-terminal domain of the adjacent monomer [8,13,17,37,38]. In creating monomeric UmuD proteins, we address two issues: 1) Can UmuD function as a monomer, that is, can it cleave itself to form UmuD$'$ and facilitate mutagenesis? 2) If UmuD is active as a monomer, this would suggest that the wild-type UmuD$_2$ dimer can also adopt an active *cis* conformation, where the N-terminal arm of the monomer binds to and is cleaved by its respective active site.

The strategy involved perturbation of the two areas of contact between subunits of the dimer; the N-terminal “neck” region between N41 and P48 and the C-terminal tail from V135 to R139 (Fig. 2.1A). We identified important candidate residues for mutation by using the evolutionary trace method for identifying active sites and functional interfaces, BLAST searches, and the multiple sequence alignment software CLUSTALW [39-41]. Results of this comprehensive search generated 15 orthologs of *E. coli* UmuD$'_2$ (PDB: 1I4V) [15] with residues N41 and P48 emerging as highly conserved (Fig. 2.1C). The matches include homologs of UmuD from a number of bacterial species: Signal Peptidase, RumA, MucA, ImpA, LexA, bacteriophage λ cI repressor and putative proteins suggested to be involved in the DNA damage response (Fig.
2.1C). Like the UmuD₂ protein, RumA, MucA, ImpA and LexA also undergo a RecA-facilitated self-cleavage reaction [42-46]. The chemical mechanism by which these proteins cleave is also similar in that the active sites contain a serine-lysine dyad that is located in the globular domain and the cleavage site is the dipeptide sequence (Ala/Cys)-Gly [42,47,48]. Many of these orthologs are also organized into an operon in which there is a umuD-like gene located upstream of a umuC-like gene. Some examples include mucAB, impAB and rumAB [42,44,49-53].
Figure 2.1 Homology models and structures of the UmuD proteins. A, Homology model of dimeric UmuD\( _2 \) in \textit{trans}, elbows down (left) (8). N41 (blue) is located in the interface of the “neck” region of UmuD. The active site residues S60 (red) and K97 (grey) are depicted. An arrow shows the location of the C-terminal “tails”. Chain A and chain B are shown in orange and light pink, respectively. Illustration was prepared using Chimera (69). Detail of “neck” region with residue N41 highlighted. B, Energy-minimized homology model of UmuD\( _2 \) in \textit{cis}, elbows down(8). The N-terminal arms of each monomer loop down and across its respective active site. C, Asparagine residues are conserved across UmuD-like proteins. Multiple sequence alignment of UmuD (PDB: 1I4V) and similar proteins. P, Plasmid-borne homologs; \( \text{Pv} \), \textit{Proteus vulgaris}; S, \textit{Salmonella typhi}; Sm, \textit{Serratia marcescens}; As, \textit{Acinetobacter sp}; Pa, \textit{Protochlamydia ameobophila}; Pm, \textit{Prochlorococcus marinus}; Dv, \textit{Desulfovibrio vulgaris}; St, \textit{Salmonella typhimurium}; Cv, \textit{Chromobacterium violaceum}; bacteriophage \( \lambda \) \textit{cI} repressor. UVPM, UV protection and mutation protein; PUVP, Putative UV protection protein; PSOSM, Putative SOS mutagenesis protein; SigP, Signal Peptidase. All proteins were aligned through a combination of BLAST and evolutionary trace searches and CLUSTALW multiple alignment (38-40). D, NMR (top) (16) and X-ray (bottom) (25) structures of UmuD\( _2' \). Chain A and B are shown in green and yellow, respectively.

We hypothesized that the UmuD N41D mutation within the “neck” region of UmuD\( _2 \) would disrupt the hydrogen bonding network that provides stability, while simultaneously generating destabilizing electrostatic and hydrophobic-hydrophilic interactions with the helix of the
opposing monomer (Fig. 2.1). We carried out the following analysis using the solution NMR structure of UmuD’ (Fig 2.1D, top) [15]. The crystal structure of UmuD’ is shown for comparison (Fig 2.1D, bottom) [23]. We used the contacts of structural units (CSU) program to analyze the effect of constructing the N41D mutation [55]. This software calculates the solvent accessibility of an atom, putative hydrogen bonds, and the various stabilizing and destabilizing interactions that occur between residues (hydrophobic-hydrophobic, aromatic-aromatic, hydrophobic-hydrophilic) [55]. Based on this calculation, N41 supports a hydrogen bonding network that involves residues I38, D39, Q42, Q46 and H47 within that chain. The substitution N41D supports hydrogen bond formation between Q42 and L40 within chain A, and R37 of chain B. However, destabilizing interactions were also identified between D41 of chain A and I38 and D41 of chain B.

As a result of these searches, we also constructed P48G, a triple mutant V135S K136A R139A, and combinations of mutations from both the N-terminal and C-terminal regions N41D K136A R139A and P48G K136A R139A. Expression of UmuD$_2$ P48G and V135S K136A R139A was substantially lower than that of wild-type UmuD$_2$. For variants N41D K136A R139A and P48G K136A R139A, expression could not be confirmed even by western blotting. Therefore, we focused on the UmuD N41D variant, which we obtained in high yield.

2.3.2 UmuD N41D and UmuD’ N41D are monomers in vitro

The dimeric and monomeric conformations of wild-type and N41D UmuD and UmuD’ were determined by native gel electrophoresis (Fig. 2.2A). UmuD$_2$ and UmuD’$_2$ dimers are clearly resolved. The predominant species of UmuD N41D migrates farther into the gel than either wild-
type UmuD₂ or UmuD'₂ dimers, consistent with UmuD N41D being a monomer. The presence of a small amount of UmuD N41D dimer is apparent. In the presence of DTT this species is eliminated and only the monomer form is observed. UmuD contains one cysteine per monomer, Cys24, so the addition of reducing agent DTT was necessary to prevent cross-linking of Cys24 on the N-terminal arms of UmuD₂. For UmuD' N41D, no covalent dimers are observed in the presence or absence of DTT as the N-terminal 24-amino acid arms, including the readily cross-linked C24, are not present in UmuD'₂.

**Figure 2.2** UmuD N41D and UmuD' N41D are monomers. A, Native gel showing the resolution of UmuD and UmuD' monomers (Mono) and dimers (Di). The plus (+) sign indicates the addition of DTT to the sample. B, Native gel showing monomer and dimer formation as a function of increasing concentrations of UmuD N41D as well as wild-type UmuD. C, Native gel showing monomer and dimer formation as a function of increasing concentrations of UmuD' N41D as well as wild-type UmuD'. D, Calculated dimer concentration (fraction bound) as a function of total concentration of UmuD N41D ([A]ₜₒₜ). Dimer concentrations obtained as a function of total concentration of UmuD N41D to give a $K_D$ of 52 ± 8.0 μM (s.e.m.) (solid line).
To determine the extent to which the N41D mutation alters the dimer-monomer equilibrium, the $K_D$ for dimerization was determined by analyzing a range of concentrations via native PAGE in the presence of DTT (Fig. 2.2B). We analyzed UmuD N41D at concentrations of 5 µM to 200 µM. At the lowest concentrations used, no dimer or only trace amounts of dimer are formed; at higher concentrations there is appreciable dimer present but the monomer is still the major species. The resulting curve was produced by calculating the fraction bound (observed dimer) as a function of the total concentration of UmuD N41D protein added to the reaction. The best fit curve generated a $K_D$ for dimerization for UmuD N41D of $52 \pm 8.0$ µM (s.e.m.) (Fig. 2.2D). This $K_D$ is over six orders of magnitude greater than the upper limit of the $K_D$ for wild-type UmuD, which is reported to be $\approx 10^{-11}$ M [17]. In the case of wild-type UmuD', no monomer form was observed in the native gel analysis (Fig. 2.2C). The $K_D$ for dimerization of wild-type UmuD' is less than 10 pM [17]. In the case of UmuD' N41D, no dimer could be detected at concentrations up to 200 µM by native PAGE. Therefore, this single amino acid change is sufficient to strongly favor the monomeric form of the protein.

### 2.3.3 N41D mutation significantly reduces the UmuD' melting temperature

In order to characterize the stability of UmuD N41D, we analyzed the melting profiles of the wild-type and variant UmuD proteins. Experiments were performed using purified protein and Sypro Orange fluorescent dye. Sypro Orange is weakly fluorescent in aqueous solution, but becomes highly fluorescent when in contact with non-polar environments such as the hydrophobic sites of a protein. Upon the melting of a protein, the hydrophobic regions are exposed and the dye binds. This results in an increase in fluorescence emission, followed by a
gradual decrease in the intensity that may be due to precipitation or aggregation of the complex of unfolded protein and the probe [56]. It was previously shown that wild-type UmuD$_2$ undergoes two melting transitions: one near 30 °C due to the dissociation of the N-terminal arms and one at 60 °C due to melting of the globular domain [35,56] (Fig. 2.3A). No clear transitions were observed for UmuD N41D; this may be because the N-terminal arms are only transiently bound to the globular domain resulting in exposure of the most hydrophobic regions of the protein (Fig. 2.3A). A range of concentrations were tested above and below the determined $K_D$ for UmuD N41D with no change in denaturation profile. The addition of stabilizing agents, such as polyethylene glycol, sucrose, or glycerol, also did not give a clear denaturation transition for UmuD N41D. Therefore, having the N-terminal arms bound to the dimeric globular domain may be a source of stability for the wild-type UmuD$_2$ protein and may explain why this is the only biochemically observed conformation to date [13,17,38].

In comparing the melting curves of wild-type UmuD’$_2$ and UmuD’ N41D at 45 µM, we determined a $\Delta T_m$ of 6.0 °C, with UmuD’ N41D destabilized relative to wild-type UmuD’$_2$ (Fig. 2.3B). This decrease in melting temperature is likely due to the loss of stability provided through dimerization. Under the conditions of these experiments, UmuD’$_2$ is dimeric while UmuD’ N41D is monomeric. This experiment was carried out using a range of concentrations, all of which gave consistent results.
Figure 2.3  A. Thermal shift assay result for wild-type UmuD [35] (green) and UmuD N41D (orange).  B. Thermal shift assay result for wild-type UmuD′ and UmuD′ N41D. Change in calculated melting temperature between wild-type UmuD′ [35] (navy blue) and UmuD′ N41D (magenta) variant protein at 45 μM is 6.0 ºC.

2.3.4 UmuD N41D monomer undergoes efficient cleavage

The cleavage of the N-terminal 24 amino acids of UmuD2 to yield the UmuD′2 homodimer is required to activate UmuC for its role in TLS [1]. UmuD binds to the RecA:ssDNA nucleoprotein filament which brings together the active site residues Ser60 and Lys97, facilitating deprotonation of Ser60 that cleaves the peptide bond between Cys24 and Gly25 [15]. The RecA:ssDNA-dependent cleavage of UmuD N41D was assayed in vitro to determine whether cleavage of the monomer was possible (Fig. 2.4A). The concentrations used in this assay were well below the calculated $K_D$ for dimerization for UmuD N41D. A range of concentrations was used to rule out the possibility that the cleavage proficiency is concentration dependent. UmuD N41D not only cleaves under these conditions, but the efficiency of cleavage was found to be independent of concentration and near that of wild-type UmuD2 (Fig. 2.4A).
To rule out the possibility that UmuD N41D was able to cleave itself because the RecA:ssDNA nucleoprotein filament facilitated dimerization, a cleavage assay under alkaline conditions was carried out (Fig. 2.4B). This reaction occurs in the absence of additional protein or DNA co-factors. At pH 10, Ser60 can be activated as a nucleophile without the addition of RecA:ssDNA, although cleavage efficiency is reduced. We found that the UmuD N41D monomer cleaves as efficiently as wild-type UmuD$_2$ under these conditions, with no dependence on concentration. Together, these observations indicate that the UmuD N41D monomer is active for cleavage. Furthermore, these observations provide strong evidence to support the model that wild-type UmuD$_2$ can cleave in the cis conformation.

**Figure 2.4** Cleavage and DinB-dependent inhibition of cleavage of UmuD N41D *in vitro* is comparable to wild-type UmuD. A, Cleavage products in the presence (+) and absence (-) of the RecA:ssDNA nucleoprotein filament are indicated. B, Cleavage under alkaline conditions (pH 10) is shown. The percentage of cleavage product was determined as the ratio of the density of D-N41D to D-WT. C, Alkaline cleavage assay

---

61
the UmuD' band to the total density of UmuD and UmuD' protein in each lane. C, DinB inhibits the cleavage of wild-type UmuD and the UmuD N41D monomer. DinB efficiently inhibits the cleavage of both N41D variant (dashed line) and wild-type (solid line) proteins to the same extent in vitro. Error bars representing standard deviation for three independent experiments are shown for each point.

2.3.5 DinB inhibits UmuD N41D cleavage

DinB, UmuD and RecA proteins form a ternary complex in solution, which suppresses the mutagenic -1 frameshift activity of DinB [7,57]. The physical interaction between UmuD and DinB inhibits the RecA-facilitated cleavage of UmuD in vitro [7]. Therefore, we tested the extent to which DinB inhibits cleavage of UmuD N41D compared to wild-type UmuD. UmuD protein concentrations used in this experiment were significantly below the $K_D$ for dimerization of UmuD N41D. UmuD N41D and DinB were incubated separately from the rest of the components of the cleavage reaction. This ensures that UmuD N41D can bind to DinB without competition from RecA:ssDNA. We found no difference in the capacity of DinB to inhibit RecA:ssDNA-mediated cleavage of wild-type UmuD or UmuD N41D (Fig 2.4C). The addition of 4 μM DinB to a cleavage reaction containing 2 μM wild-type UmuD and UmuD N41D resulted in an over 33% reduction in cleavage efficiency for both UmuD proteins. Therefore, the dimerization defect of UmuD N41D does not appear to affect the molecular interactions necessary for proper DinB-UmuD interactions, which suggests that the monomer is functionally active and capable of physical interactions with DinB in this context.

2.3.6 UmuD N41D and UmuD' N41D are expressed as active monomers in vivo

We next determined the proficiency of UmuD and UmuD' N41D variants for in vivo functions. The expression level and cleavage activity in vivo of wild-type UmuD and UmuD N41D
expressed from plasmids were determined by immunoblotting under native and denaturing conditions. We UV-irradiated cells harboring low-copy plasmids that expressed UmuD₂ or UmuD’₂ or the N41D variants and resolved the proteins by electrophoresis under denaturing conditions. The extent of cleavage observed for wild-type UmuD₂ and the UmuD N41D monomer was 39% and 41%, respectively (Fig. 2.5A). The expression levels for wild-type and variant proteins were similar (Fig. 2.5A). We also used gentle, non-denaturing extraction conditions and analyzed the cellular proteins by native gel electrophoresis. Under native conditions, we found that UmuD N41D and UmuD’ N41D are both resolved as monomers with comparable expression levels as seen under denaturing conditions. Full-length UmuD and cleaved UmuD’ N41D can also be seen on the blot thus confirming the cleavage proficiency of the monomer in vivo. However, for wild-type UmuD₂ under native conditions, only a single band was observed (Fig. 2.5B). This was intriguing as we would expect both full-length and cleaved UmuD₂ proteins to be present as their dimeric forms. Western blots of proteins resolved under denaturing conditions were performed on the identical samples and the presence of both UmuD and UmuD’ proteins in nearly equivalent amounts was observed (Fig. 2.5C). From this, we conclude that the UmuDD’ heterodimer is likely the most stable conformation in vivo. Although this phenomenon was previously observed in vitro by cross-linking [58], we believe this is the first report of heterodimer formation observed in vivo. These findings shed new light on the predominant conformation of wild-type UmuD after UV irradiation and highlight the possibility that the heterodimer may play a larger role in this regard.
Figure 2.5 UmuD expression levels and cleavage products \textit{in vivo}. A, Immunoblot of gel run under denaturing conditions showing the steady-state expression levels of UmuD from plasmids pGY9739 and pGY9738 in GW8017. Relative UmuD expression levels are shown below the blot, where UmuD N41D is normalized to wild-type UmuD and UmuD’ N41D is normalized to wild-type UmuD’. Percent cleavage for UmuD and UmuD N41D are also indicated. B, Immunoblot of gel under native conditions showing that UmuD and UmuD’ are expressed as dimers while UmuD N41D and UmuD’ N41D are expressed as monomers \textit{in vivo}. There is no cleavage product seen under these conditions for wild-type UmuD, however the cleavage product for UmuD N41D is evident. C, Immunoblot of identical samples from B, under denaturing conditions. In this blot, both full-length and cleaved products are present in almost equivalent amounts suggesting that a stable wild-type UmuDD’ heterodimer was observed. D, Western blots in A, B, and C were carried out to give a signal within the linear range of the assay. The intensity of the signal \textit{versus} the amount of crude lysate loaded is shown. The point on the y axis of the graph represents the band from gels in A–C with highest intensity, which is the heterodimer in B. E, Heterodimer formation under equilibrium conditions using purified protein \textit{in vitro}. Native gel showing preferential heterodimer formation (box) when wild-type UmuD and UmuD’ proteins are combined in equivalent amounts. UmuD N41D and UmuD’ N41D do not form heterodimers as expected. Combining N41D variant and wild-type protein also does not result in heterodimerization.
2.3.7 UmuD and UmuD' rapidly form heterodimers in vitro

It has been known for some time that UmuD₂ forms exchangeable dimers and that UmuD₂ and UmuD'₂ form heterodimers [8,15,58]. Glutaraldehyde cross-linking of equimolar amounts of UmuD₂ and UmuD'₂ was previously used to demonstrate the preferential formation of the UmuDD' heterodimer. In this case, no UmuD₂ or UmuD'₂ homodimers could be detected after 20 minutes, leading the authors to suggest that the heterodimer complex was favored [58]. Subsequently, it was shown that the arms of the active site variant UmuD₂ S60A could be cleaved by incubating with a non-cleavable variant in which the cleavage site is mutated but the active site is intact [37]. From this and similar experiments, it was concluded that the dimers must be exchangeable in order for cleavage to occur and that cleavage occurs in trans [8,37].

By mixing equal amounts of wild-type UmuD₂ and UmuD'₂ under equilibrium conditions, we have found that the heterodimer conformation is indeed the most stable conformation (Fig. 2.5E). This is consistent with our observations in vivo (Fig. 2.5B). We detect 71% heterodimer formation with 29% UmuD' present, most likely as a result of cleavage during incubation and/or electrophoresis (Fig. 2.6). We do not observe full-length UmuD₂ homodimer. Experiments were also performed to determine the ability of monomeric UmuD N41D and UmuD' N41D to form heterodimers or to form exchangeable dimers with wild-type UmuD and UmuD' (Fig. 2.5E). Heterodimerization did not occur with any combination of proteins. Notably, UmuD N41D or UmuD' N41D are unable to form dimers even with wild-type UmuD or UmuD' as partners.
2.3.8 UmuD N41D and UmuD’ N41D facilitate UV mutagenesis and survival

UmuD’2C is required for UV-induced mutagenesis in *E. coli* [1]. As part of its mutagenic signature, Pol V inserts guanine opposite the 3’-thymine of (6-4) T-T photoproducts that are a result of exposure to UV light [59]. This activity can be detected via the reversion of the *argE3* auxotrophic marker in the arginine biosynthetic pathway [29]. To assess the proficiency of UmuD N41D monomers in UV-induced mutagenesis, we compared the mutation frequency of Δ*umuDC* strains harboring plasmid-borne wild-type UmuD2 and UmuD’2 as well as UmuD N41D and UmuD’ N41D variants. We found that the mutation frequency of cells expressing UmuD N41D and UmuD’ N41D were similar to those expressing wild-type UmuD2 and UmuD’2, respectively (Fig. 2.6A). This suggests that monomeric UmuD’ is able to interact with the numerous protein partners required for mutagenesis, including UmuC, RecA, and the β clamp [1].

**Figure 2.6** UmuD and UmuD’ N41D are proficient for UV-induced mutagenesis and confer resistance to ultraviolet radiation. A, UV-induced mutation frequency for wild-type UmuD or UmuD’ and UmuD N41D or UmuD’ N41D in plasmids pGY9739 (*umuDC*)/pGY9738 (*umuD’C*) and empty vector pGB2 in strain GW8017. Mutation frequencies are reported relative to that of GW8017 pGY9739, which is set to 100%. B, Survival assays were carried out with pGY9739 (*umuDC*) and pGY9738 (*umuD’C*) plasmids in PB103. pGY9738 (*umuD’C-N41D; ■); pGY9738
AB1157 $\Delta umuDC \Delta recJ$ strains are hypersensitive to UV light [60], a phenotype that can be suppressed by complementation with low-copy plasmids bearing the $umuDC$ genes. We determined that $\Delta umuDC \Delta recJ$ strains harboring plasmids expressing UmuD N41D and UmuD' N41D were not sensitive to UV light (Fig. 2.6B). Even though UmuD N41D and UmuD' N41D are defective in dimerization, the level of resistance to UV light for strains harboring these variants surpassed that of strains expressing the wild-type UmuD and UmuD' proteins. As highlighted above, the expression levels of these variants are similar to those of wild-type UmuD and UmuD', and the cleavage efficiency of UmuD N41D is also comparable to that of wild-type UmuD$_2$ both in vivo and in vitro (Fig 2.4, 2.5). Notably, defects in UmuD$_2$ cleavage, as with the UmuD$_2$ S60A active site variant, result in hypersensitivity to killing by UV light, while defects in dimerization apparently confer UV-resistance [58,61]. These observations confirm that dimeric UmuD$_2$ and UmuD'$_2$ are not essential for UV survival or for overall cell viability.

**Figure 2.7** UmuD N41D cross-links to the $\beta$-clamp as a monomer. Formaldehyde was used as a cross-linker. Native gel showing wild-type UmuD or the UmuD N41D monomer cross-linked to the $\beta$ clamp. Reaction contained equimolar amounts (10 $\mu$M) of UmuD proteins and the $\beta$ clamp.
2.3.9 UmuD N41D interacts with the β clamp as a monomer

Interactions between the UmuD proteins and the β clamp have been well characterized using cross-linking methods [62] [63]. These experiments concluded that residues on both the N-terminal arms and C-terminal globular domain of UmuD_2 interact with the β clamp [63,64]. Truncations constructed in the N-terminal arms of UmuD_2 indicate that residues 9-19 are particularly important for binding, therefore the affinity of UmuD’_2 for the β clamp is weaker [63]. We utilized formaldehyde cross-linking to assess the relative binding affinity of wild-type UmuD_2 versus the UmuD N41D monomer for the β clamp (Fig 2.7). Results clearly demonstrate that there is a difference in the stoichiometry of binding although the resolution of the experiment does not allow for a quantitative assignment. In the case of cross-linked UmuD_2 and the β clamp, the multiple bands observed may indicate the binding of one or two UmuD dimers per dimeric β clamp. As this experiment was completed under non-denaturing conditions, wild-type UmuD_2 is most likely cross-linked to the β clamp as a dimer. Furthermore, we included the active site variant UmuD_2 S60A to confirm that UmuD’_2 is not one of the forms being detected. Conversely, cross-linking the UmuD N41D monomer to the β clamp produced one clean band. Considering the scenario described above, we conclude that the complex detected is composed of two UmuD N41D monomers in complex with the β clamp, with one UmuD monomer binding to each β clamp monomer.
2.4 DISCUSSION

In this, we have used numerous in vitro and in vivo methods to characterize UmuD N41D and UmuD’ N41D as active monomers that functionally mimic the wild-type dimers. It was found that alkaline and RecA-mediated cleavage for both wild-type UmuD₂ dimer and UmuD N41D monomer proteins occur with similar efficiency. Moreover, cleavage of the N-terminal arms of the UmuD N41D monomer shows that it is a stable and active species and also leads to the conclusion that the wild-type UmuD dimer may cleave in the cis conformation. The conformational dynamics of the UmuD and UmuD’ monomer and dimer proteins was further investigated using the thermofluor assay which clearly shows a difference in overall structural stability and hydrophobicity. It is also evident that the UmuD monomer interacts with partner proteins such as DinB and the β clamp in vitro and is active in facilitating mutagenesis in vivo.

We have estimated the $K_D$ for dimerization for monomeric UmuD N41D as approximately 54 μM. This is over six orders of magnitude greater than the $K_D$ determined for wild-type UmuD₂ [17]. As only a single mutation was made in the N-terminal “neck” region of the protein, this region appears to be critical for modulating the dimerization stability of wild-type UmuD₂ and UmuD’₂. Certainly, the charge-charge repulsion is most likely the key contributor to the dimerization defect as an N-terminal truncation of wild-type UmuD’ through residue 45 reportedly does not affect dimer formation [23].

Intriguingly, LexA is an extremely tight dimer, with a $K_D$ of less than 20 pM [65], but unlike wild-type UmuD, the LexA protein sequence contains residue D101 corresponding to UmuD N41 (Fig 2.1C). On further inspection, the LexA crystal structure [48] (PDB: 1JHH) shows the
carboxyl groups of the D101 residues of each monomer pointing away from the interface, and solvent exposed. Therefore, the destabilizing charge-charge repulsion that is attributed to the dimerization defect of UmuD N41D is averted in LexA.

The dynamic nature of UmuD proteins is also evident in the results of thermofluor assays on the UmuD N41D and UmuD’ N41D monomers. A sharp melting transition was not obtained for full-length UmuD N41D monomer even with the addition of a host of stabilizing agents. However, the melting profile for wild-type UmuD$_2$ has been reported as two distinct unfolding transitions that represent the dissociation of the N-terminal arms at approximately 30 °C, and the melting of the dimeric globular domain at 60 °C, respectively [35]. As previously observed, there is a single melting transition at 60 °C for wild-type UmuD’$_2$ [35]. The T$_m$ of 54 °C for the monomeric UmuD’ N41D most likely represents unfolding of the globular domain, and we attribute this reduction in melting temperature as compared to wild-type UmuD’$_2$ to the loss of the stability that is provided through dimerization. Moreover, this study demonstrates the significance of the N-terminal arms and their considerable contributions to stability. Although the contacts within the dimer interface of the C-terminal domains are important, the wrapping of the UmuD$_2$ arms across the globular domain also seems to provide additional structural support that is not present in the UmuD N41D variant.

UmuD$_2$ has been categorized as an intrinsically disordered protein (IDP) [17]. These proteins exhibit a high degree of flexibility that allows the accommodation of a large number of protein-protein contacts [66]. IDPs defy the traditional structure-function paradigm as it is not essential to fold into a stable three-dimensional structure for basic cellular function [66]. These proteins
function via molecular recognition where the structure of the IDP is altered upon binding to a partner protein or as effectors that modify the activity of a single binding partner or an assembly [67]. The dimeric UmuD proteins have been shown to interact with a growing list of partners including TLS polymerases UmuC and DinB, the α, β and ε subunits of the replicative polymerase pol III, RecA, and the Lon and ClpXP proteases [1,6-12]. Therefore, UmuD2 is reminiscent of a hub protein that may be stabilized by binding to multiple structured interacting partners which allow for the regulation of protein expression, multi-protein complex formation and degradation as required [17,66]. Many of these interactions may be transient but highly specific. Regardless, evolutionary selection of a dimeric over monomeric UmuD protein has advantages. Not only is dimeric UmuD2 structurally more stable, but it can also adopt more conformations and potentially bind twice as many interacting partners at any given time.

Monomeric UmuD N41D undergoes RecA:ssDNA-facilitated cleavage of its N-terminal arms as well as cleavage under alkaline conditions (Fig. 2.4A,B). Cleavage is independent of protein concentration indicating that the UmuD N41D monomer cleaves via an intramolecular reaction. Interestingly, the monomeric λ cI repressor protein also undergoes RecA:ssDNA-facilitated intramolecular self-cleavage [68]. It was previously shown that wild-type UmuD2 with the N-terminal arms in the trans conformation cleaves to form UmuD′2 by an intermolecular pathway [37]. The idea of trans or intermolecular cleavage was first confirmed by introducing plasmids carrying active site mutants (ASM: S60A and K97A) and cleavage site mutants (CSM: C24D and G25D) into E. coli ΔumuDC strains [37]. Cleavage was only observed when both ASM and CSM mutants were introduced together into cells suggesting that the reaction is intermolecular [37]. The UmuD2 ortholog MucA also undergoes intermolecular cleavage [37]. However,
isoenergetic models of UmuD with the N-terminal arms in the \textit{cis} conformation led us to consider intramolecular cleavage as a viable mechanism [8]. This is in agreement with the cleavage of other UmuD-like proteins such as LexA and λ cl that are primarily intramolecular in nature [47]. As monomeric UmuD N41D must cleave itself intramolecularly, it is plausible that wild-type UmuD cleaves itself both intermolecularly and intramolecularly.

Other evidence also suggests that the UmuD N41D monomer may be a competitive substrate for RecA-facilitated cleavage [36]. It was determined that cross-linking residues close to the dimer interface of UmuD reduced the cleavage efficiency by the RecA nucleoprotein filament substantially [36]. The monocysteine residues used in this study include V34C, I38C and L44C and are in the vicinity of the N-terminal “neck” region where the key interactions for dimerization occur [36]. Since UmuD N41D is proficient for cleavage, the essential interactions with RecA that are required for this process are still intact. Therefore, the RecA nucleoprotein filament may only be needed to activate the catalytic site of UmuD by inducing S60 and K97 to adopt the correct conformation for cleavage [15].

It has been generally accepted that a defect in dimerization of the UmuD proteins would render UmuD inactive in most, if not all, of its functions and make cells non-mutable [23,32,69,70]. Indeed, UmuD’ variants A50V, T51I, H82Y or G129S [69], all suggested to disrupt dimerization, result in reduced levels of spontaneous or induced mutagenesis [32,69]. In contrast, we have shown that the UmuD N41D and UmuD’ N41D monomers are active in facilitating mutagenesis and cell survival after treatment with DNA damaging UV light. Moreover, the UmuD N41D monomer has little or no defect in cleavage either \textit{in vitro} or \textit{in vivo}. Taken together, our
observations indicate that although UmuD₂ is a tight dimer, dimerization is not required for the cellular functions of UmuD in regulating mutagenesis.

2.5 REFERENCES

Chapter 3: Electron spin labeling reveals the highly dynamic N-terminal arms of the SOS mutagenesis protein UmuD

This chapter has been published as:

3.1 INTRODUCTION

The dimeric UmuD proteins are actively involved in regulating mutagenesis as part of the tightly controlled SOS response in *Escherichia coli* (*E. coli*). This stress-induced response occurs upon exposure to DNA damaging agents, such as toxic chemicals or ultraviolet radiation, which can cause mutations that may lead to uncontrollable cell proliferation or cell death [1]. Full-length UmuD₂, which prevents mutagenesis, is a homodimer of 139-amino acid subunits and is the initial umuD gene product that appears after induction of the SOS response[1]. Subsequently, the damage-induced RecA:ssDNA nucleoprotein filament facilitates autocleavage of the N-terminal 24-amino acids of UmuD₂ to yield UmuD′₂, the form that enables mutagenesis [1]. UmuD₂ and UmuD′₂ interact with potentially mutagenic translesion DNA synthesis (TLS) polymerases UmuC (polymerase subunit of pol V) and DinB (pol IV) as well as the α, β and ε subunits of the replicative polymerase, pol III.[1-4] UmuD₂ and UmuD′₂ can display differential interactions with their partner proteins which can lead to dramatically different cellular outcomes [1-4]. These key differences may be due to the dynamics of the N-terminal arms of UmuD₂ [1].

Previous biochemical evidence supported a model in which the arms of UmuD are stably bound to its globular domain [5,6]. However, recent experiments suggest that the N-terminal arms of UmuD are somewhat dynamic [7]. In order to address this apparent contradiction, we utilized site-directed spin-labeling (SDSL) with 3-iodomethyl-1-oxy-2,2,5,5-tetramethylpyrroline (iodomethyl spin label, IMSL) and electron paramagnetic resonance (EPR) spectroscopy to characterize the local motion of the N-terminal arms of UmuD₂ and find that they are indeed quite dynamic. The locations of the single Cys residues used for
labeling are shown as blue spheres in Figure. 3.1. Because UmuD$_2$ is a very tight dimer ($K_d < 10$ pM) [8], it is expected to be dimeric in experiments reported here and for simplicity, we will refer to dimeric UmuD$_2$ as UmuD.

The X-ray [9] and NMR [10] (Figure. 3.1B) structures of the cleaved protein UmuD' have been determined. These structures reveal substantial differences in the overall shape and compactness of UmuD', and show that the N-terminal arms remaining after cleavage (residues 25-40, pink in Figure. 3.1) are extended and not engaged in dimerization [9,10].

**Figure. 3.1** UmuD and UmuD' constructs. Active site residues K97 (purple) and S60 (red) are shown. Residues 25-40 are shown in pink. A. Homology model of UmuD with the arms in *trans*, elbows down. The UmuD variant with amino acid substitutions T14A, L17A and F18A (shown in orange, yellow, and green, respectively), which render UmuD non-cleavable, is referred to as UmuD3A.[11] The blue sphere represents the native residue C24, where the N-terminal arms of wild-type UmuD, active site variant UmuD-S60A and non-cleavable UmuD3A are spin labeled with IMSL. In wild-type UmuD, residues 1-24 (cyan) are cleaved off to yield UmuD'. B. The NMR structure of UmuD'[10] with residues 25-40 (pink) in an extended conformation is shown. A different spin labeling site of A31C (blue spheres) was constructed, as residue C24 is removed during cleavage and therefore not present in UmuD'.

80
The absence of high resolution structural information for UmuD has left many questions about the position of the arms. Hydrogen-deuterium exchange mass spectrometry was used to show that the N-terminal arms of UmuD are highly solvent-exposed [7]. As expected, the arms of UmuD' show more deuterium exchange than UmuD [7]. On the other hand, cross-linking studies reveal that the N-terminal arms of UmuD can be stably bound to the globular domains [5].

Isoenergetic homology models suggest that the N-terminal arms (residues 1-39) of UmuD can adopt either trans (Figure. 3.1A) or cis conformations, with the elbows up or down [11]. For UmuD in trans with the elbows down, the arm of each monomer loops across the active site of the partner monomer located in the globular C-terminal domain. For UmuD conformations in cis, each arm interacts with its respective C-terminal domain. For cleavage to occur in either cis or trans, the arms must be bound to the C-terminal domain in the elbows down conformation [5]. In the elbows up conformation, the arms bind to the outer edge of the globular C-terminal domain, away from the active site. These models suggest that the N-terminal arms of UmuD are dynamic and may adopt multiple conformations.

Cleavage of the arms is accomplished by the active site dyad of Ser60 and Lys97 [1]. While RecA:ssDNA is required for efficient UmuD cleavage, and was not included in our experiments, a small amount of cleavage product may be present in the samples. Therefore, the active site mutation UmuD-S60A was constructed to reduce cleavage as a potential complication. The triple mutant UmuD3A (T14A, L17A and F18A) is a non-cleavable variant of UmuD (Figure. 3.1A) that acts as a biological mimic of UmuD', even though the
UmuD3A arms are full length. Biochemical analysis of UmuD3A indicates that the three mutations, which are located in the N-terminal arms, perturb the interaction with the globular C-terminal domain, preventing cleavage,[2,7,11] a suggestion we test here. Therefore, the N-terminal arms of UmuD3A may be extended, much like the arms of UmuD’ (Figure. 3.1A).

We hypothesize that the N-terminal arms of UmuD are not stably bound to the globular C-terminal domain but rather are dynamic in solution. We utilized EPR spectroscopy to characterize the motion of the N-terminal arms of UmuD by comparing the spectra of wild-type UmuD, active-site variant UmuD-S60A, and the non-cleavable UmuD’ mimic UmuD3A, [11] to that of UmuD’-A31C, which is labeled on the unbound and solvent-exposed arms (Figure. 3.1B).

3.2 EXPERIMENTAL PROCEDURES

3.2.1 Protein purification, labeling and analysis

Wild-type UmuD, UmuD-S60A, UmuD3A, and UmuD’-A31C were purified following a reported procedure [12]. The A31C variant of UmuD’ was constructed using a QuikChange site-directed mutagenesis kit (Agilent). Protein with a monomer concentration of 300 µM was utilized for EPR. UmuD₂ protein samples (95% purity) were reacted with a thiol-reactive nitrooxide derivative, 3-iodomethyl-1-oxy-2,2,5,5-tetramethylpyrrole (iodomethyl spin label, IMSL, Toronto Research Chemicals) for site-directed spin labeling. This was accomplished by adding an appropriate amount of the 100 mM stock solution of spin label to achieve a molar ratio of 1:3 protein to spin label. The mixture was incubated at 37 °C for 4 h and then maintained at 4 °C overnight. The labeled protein was then dialyzed against exchange buffer [25 mM HEPES (pH 7.5), 150 mM NaCl, 0.1 mM EDTA, 5% glycerol] for at least 24 hours using a Slide-A-
Lyzer Dialysis Cassette (Thermo Scientific) with a 3500-Da cutoff membrane. Zeba 7K desalting spin columns (0.5 mL) (Thermo Scientific) were then used to remove residual DTT, and the labeled protein was subsequently stored at –80 °C. Protein samples were analyzed by using a Bruker EMX instrument outfitted with a high-sensitivity cylindrical cavity and variable-temperature module at 24 °C (297 K) unless noted otherwise, and the temperature was controlled to within 0.3 °C by a nitrogen stream. The spectra were obtained at a microwave frequency of 9.37 GHz, 6.0 mW microwave power and 0.5 G 100 kHz field modulation amplitude.

### 3.2.2 Lineshape fitting procedures and best-fit parameters

The LabView-based multicomponent lineshape fitting program of Altenbach [13] was used to analyze the spectra. For the fitting procedure, the magnetic parameters of the nitroxide label were set as follows: the electronic g tensor components were fixed at $g_x = 2.0081$, $g_y = 2.0059$, and $g_z = 2.0023$, and the components of the $^{14}$N hyperfine tensor were fixed at $A_x/\gamma_e = 6.6$ G, $A_y/\gamma_e = 6.9$ G, and $A_z/\gamma_e = 35.1$ G, where $\gamma_e$ is the electronic gyromagnetic ratio.[14] To model the magnetic properties of the remaining nuclei, the calculated spectra were convoluted with a superhyperfine pattern consisting of 12 equivalent $^1$H nuclei with isotropic coupling $a_{\text{H}}/\gamma_e = 0.25$ G, one $^1$H nucleus with isotropic coupling $a_{\text{H}}/\gamma_e = 0.40$ G, and a $^{13}$C nucleus with isotropic coupling $a_{\text{C}}/\gamma_e = 6.90$ G at one of the two carbon atoms alpha to the nitroxide at natural abundance (1.108%).

Except as noted below, each spectrum was fitted using three components, since three species with significantly different mobilities could be visually distinguished in most of the spectra, particularly at the lower temperatures studied. The scale factor for each component and the isotropic rotational diffusion constants $R_1$, $R_{\text{II}}$, and $R_{\text{III}}$, were varied, together with two Gaussian
inhomogeneous line-width parameters. The first line-width parameter, $\Delta H_I$, was applied to the slowest component, while the second ($\Delta H_{II,III}$) was applied to the components with the two highest $R$ values. The least-squares parameters for each protein variant studied are given in Table 3.1, for UmuD-S60A at each temperature in Table 3.2, and the fitting results are plotted in Figure 3.6 below and Figure 3.7.

For the UmuD3A and UmuD'-A31C variants, reliable parameters for a slow component could not be obtained under the assumption of a three-component spectrum. The slow component was therefore eliminated from the UmuD3A and UmuD'-A31C variants.

Parameter uncertainties were estimated from the covariance matrix of the parameters at the least-squares solution according to the procedure described by Budil et al [15]. The uncertainties were similar for a given parameter in different fits, so for clarity, uncertainties are reported as average values for each class of parameter.

### 3.2.3 Estimation of rotational correlation time for UmuD dimer

The expected order of magnitude for the rotational correlation time of a protein may be estimated from the Debye Stokes Einstein equation,

$$\tau = \frac{4 \pi r_e^3 \eta}{3 kT}$$

where $r_e$ is the effective hydrodynamic radius of the protein (including bound waters), $k$ is Boltzmann’s constant, and $\eta$ is the solvent viscosity, which for water is $8.9 \times 10^{-4}$ Pa s ($0.890$ cP) at $T=297$ K. Approximating the UmuD dimer as a sphere with radius 25 Å, the above equation...
gives an average rotational correlation time of 14 ns, consistent with the observed rotational correlation times of the slowest component in Table 3.1 and in Table 3.2 near room temperature.

3.3 RESULTS

EPR coupled with SDSL is a powerful tool for detecting conformational changes in proteins. In SDSL, a site of interest on the protein is modified with a paramagnetic spin label such as a stable nitroxide radical reagent. The spin label reacts specifically with cysteine residues in solvent-exposed regions resulting in chemical modification of the sulphydryl group [16]. Wild-type UmuD and the full-length variants UmuD-S60A and UmuD3A contain only one labeling site per monomer at the native C24 residue in the N-terminal arms (Figure. 3.1A). After cleavage to UmuD′, the N-terminal 24- amino acids, which include C24, are released. Therefore, we constructed A31C to enable labeling of UmuD′ (Figure. 3.1B). Because the N-terminal arms of UmuD′ are thought to be unbound and therefore conformationally unconstrained [9,10], this variant provides an important control for a high mobility component.

To determine the dynamics of the UmuD arms, we analyzed the influence of the nitroxide motion on the spectral lineshape. The change in the mobility of the probe may be due to tumbling of the entire protein, local backbone fluctuations, the dynamics of the spin-label, or interaction with other molecules [16]. EPR spectra were obtained at 24 °C for wild-type UmuD, UmuD-S60A, UmuD3A, and UmuD′-A31C with covalently bound IMSL spin-label. The spectra of wild-type UmuD and UmuD-S60A exhibit three distinct components reflecting the presence of three spin-probe populations with different mobilities (Figure.
3.2A). Features corresponding to each of the three components, slow (I), intermediate (II) and fast (III), are indicated. The spectrum of UmuD3A shows only a trace of the slow component whereas UmuD'-A31C exhibits only a fast component (Figure. 3.2A). The individual components are plotted separately for clarity in Figure. 3.2B. Dynamic parameters and scaling factors for each component in the spectra of the spin-labeled UmuD variants are presented in Table 3.1.
Figure 3.2 A. Least-squares fits (blue lines) to experimental spectra (black lines) of different spin-labeled UmuD variants. The EPR spectra of wild-type UmuD, UmuD-S60A, UmuD3A, and UmuD′-A31C were obtained at 24 °C. Wild-type UmuD, UmuD-S60A and UmuD3A are labeled at native C24, while UmuD′-A31C is labeled at A31C. The slow, intermediate and fast spectral components are labeled I, II, III, respectively. B. Calculated
spectra of slow (red lines), intermediate (green lines) and fast (blue lines) components from the fits shown in A. The relative populations of each component are shown on the spectra of the respective UmuD variant. The components for wild-type UmuD are scaled by a factor of 2.0 and those of UmuD-S60A by a factor of 1.6 to improve clarity. The fitted fast component spectra are also omitted for these two variants for the sake of clarity. Fitting parameters are given in Table 3.1 [14].

Table 3.1 Dynamic parameters and scaling factors for the spin-labeled UmuD variants.

<table>
<thead>
<tr>
<th>Variant</th>
<th>$\log_{10} R_{I_0}$</th>
<th>$\log_{10} R_{III_0}$</th>
<th>$\tau_{II}$ (ns)</th>
<th>$\tau_{III}$ (ns)</th>
<th>$\Delta H_I$ (G)</th>
<th>$\Delta H_{II,III}$ (G)</th>
<th>$\phi_I$</th>
<th>$\phi_{II}$</th>
<th>$\phi_{III}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT UmuD</td>
<td>6.99</td>
<td>7.01</td>
<td>17.1</td>
<td>1.2</td>
<td>4.01</td>
<td>0.45</td>
<td>0.47</td>
<td>0.30</td>
<td>0.22</td>
</tr>
<tr>
<td>UmuD-S60A</td>
<td>6.99</td>
<td>7.01</td>
<td>9.40</td>
<td>1.2</td>
<td>4.01</td>
<td>0.45</td>
<td>0.47</td>
<td>0.30</td>
<td>0.22</td>
</tr>
<tr>
<td>UmuD-3A</td>
<td>--</td>
<td>7.01</td>
<td>9.40</td>
<td>1.2</td>
<td>4.01</td>
<td>0.45</td>
<td>0.47</td>
<td>0.30</td>
<td>0.22</td>
</tr>
<tr>
<td>UmuD′-A31C</td>
<td>--</td>
<td>--</td>
<td>8.76</td>
<td>2.3</td>
<td>3.11</td>
<td>0.39</td>
<td>0.30</td>
<td>0.17</td>
<td>0.29</td>
</tr>
</tbody>
</table>

(a) $\log_{10}$ of isotropic diffusion constant $R$ for components I (slow), II (intermediate), and III (fast) expressed in units of $s^{-1}$. Average uncertainty in these parameters is 0.05.
(b) Average rotational correlation time for each component, calculated as $\tau = 1/6R$
(c) Inhomogeneous Gaussian derivative line peak-to-peak linewidth for the slowest component ($\Delta H_I$) and the fastest two components ($\Delta H_{II,III}$). Average uncertainty in these parameters is 0.1 G.
(d) Normalized fraction of components I (slow), II (intermediate), and III (fast). Average uncertainty in these parameters is 0.07.

The EPR spectra of spin-labeled wild-type UmuD and non-cleavable UmuD-S60A both show a low-mobility component with a rotational correlation time of $\tau \sim 17$ ns, suggesting that the spin label experiences similar local environments in both proteins. The observed $\tau$ is quite close to the expected value for the UmuD dimer based on its hydrodynamic radius (Table 3.1). This is consistent with the observation that the N-terminal arms of full-length UmuD fold across the globular C-terminal domain in an elbows down conformation, essentially immobilizing the spin-label on the protein [5] so that $\tau$ reflects rotation of the entire molecule. The slow-moving component is more apparent for UmuD-S60A as the arms may be in a conformation ready for cleavage, but cleavage is inhibited by mutation of the active
site residue S60 (Figure 3.1). In addition, the S60A mutation may stabilize the arm in a bound conformation due to a slight increase in the hydrophobic interactions between the arms and the globular domain. Both variants also exhibit a minor (~20%) component with high mobility. The τ values for this component are significantly longer than that of the free label (Table 3.1, Figure 3.3), suggesting that it arises instead from spin labels attached to the short peptide produced by cleavage of the N-terminal arms. This assignment is supported by the observation of nonspecific cleavage products in native gels of the sample (Figure 3.4).

**Figure 3.3** Spectrum of free IMSL in aqueous buffer at 297 K (solid line) and least-squares fit to the lineshape (dashed line). The best fit parameters were $a_0/\gamma_e = 16.24$ G, $\Delta H = 0.504$ G, and $\log_{10} R \cdot s = 9.84$. Superhyperfine interactions with $^{1}H$ nuclei and $^{13}C$ nucleus were included as described in the fitting procedure given above.

For UmuD’-A31C, no low- or intermediate-mobility component was observed. This is in agreement with both the X-ray [9] and NMR [10] structures of UmuD’ that show that the N-
terminal arms are unbound from the C-terminal domain and solvent exposed. Somewhat similarly, the spectrum of UmuD3A appears to have only the more mobile components, suggesting that the spin-label is solvent exposed and moving independently of the body of the protein, consistent with fast motional dynamics. This observation strongly supports the model that the N-terminal arms of UmuD3A are mostly in an extended conformation similar to those of UmuD’, and are largely unbound from the C-terminal globular domain [7, 11]. The faster of the two components has a $\tau$ value that is significantly slower than that of the component assigned to the cleavage product for the other variants; furthermore, cleavage of UmuD3A was not observed in the native gel (Figure 3.4). We hypothesize that the observed components reflect different conformations of the N-terminal arm in the UmuD3A variant, but this question warrants more detailed study.
Figure 3.4 A. Native gel showing IMSL-labeled wild-type UmuD, UmuD-S60A, UmuD3A and UmuD′-A31C, compared to their unlabeled counterparts, under non-denaturing conditions [17]. Wild-type UmuD sample contains residual UmuD′ as well as the UmuD/D′ heterodimer, whereas UmuD-S60A contains less of the lower molecular weight species. UmuD$_2$ multimers are also faintly detected. The relative proportions of the different species are given in the table below the gel. Numbers may not sum to 100% due to rounding. B. IMSL-labeled UmuD is cleavable upon incubation with RecA/ssDNA nucleoprotein filament [18], albeit with lower efficiency than unlabeled UmuD.

The intermediate mobility component in the wild-type and UmuD-S60A variants most likely represents a conformation in which the N-terminal arms are dissociated from the C-terminal domain of the protein. To investigate this possibility, a temperature study was conducted from 275 to 323 K using UmuD-S60A. Wild-type UmuD stocks contain small amounts of UmuD′, which results in formation of the heterodimer UmuD/D′ (Figure 3.4). Therefore, the active site variant UmuD-S60A, which contains significantly less cleavage product, was used for these temperature studies. Three-component fitting was used to estimate the contributions
of the fast, intermediate and slow components to the spectral intensity (Figure. 3.5).[19]

Dynamic parameters and scaling factors for each of the components as a function of
temperature are presented in Table 3.2. As temperature increases, the fraction of the slow
component decreases and that of the intermediate component increases. The spectral changes
are reversible with temperature (Figure. 3.6), strongly supporting the conclusion of
temperature-dependent interconversion of protein conformations.

**Figure 3.5** Left side: least-squares fits (green lines) to spectra (blue lines) from spin-labeled
UmuD-S60A at different temperatures. Right side: Calculated spectra of the slow (green
lines) and intermediate motion (blue lines) components from the fits shown on the left. As
the temperature increases from 275 K to 323 K, a conversion from the slow component to the
intermediate component is observed. Fitting parameters are given in Table 3.2.
Table 3.2 Dynamic parameters and scaling factors for spin-labeled UmuD-S60A as a function of temperature

<table>
<thead>
<tr>
<th>T (K)</th>
<th>$\log_{10} R_{SI}$</th>
<th>$\log_{10} R_{II}$</th>
<th>$\log_{10} R_{III}$</th>
<th>$\tau_{SI}$ (ns)</th>
<th>$\tau_{II}$ (ns)</th>
<th>$\tau_{III}$ (ns)</th>
<th>$\Delta H_I$ (G)</th>
<th>$\Delta H_{II,III}$ (G)</th>
<th>$\phi_I$</th>
<th>$\phi_{II}$</th>
<th>$\phi_{III}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>275</td>
<td>6.19</td>
<td>7.98</td>
<td>9.50</td>
<td>107.6</td>
<td>1.7</td>
<td>0.05</td>
<td>2.88</td>
<td>0.05</td>
<td>0.616</td>
<td>0.321</td>
<td>0.063</td>
</tr>
<tr>
<td>283</td>
<td>6.71</td>
<td>8.08</td>
<td>9.17</td>
<td>32.5</td>
<td>1.4</td>
<td>0.11</td>
<td>2.79</td>
<td>0.09</td>
<td>0.617</td>
<td>0.277</td>
<td>0.105</td>
</tr>
<tr>
<td>288</td>
<td>6.82</td>
<td>7.92</td>
<td>9.19</td>
<td>25.2</td>
<td>2.0</td>
<td>0.11</td>
<td>1.55</td>
<td>0.17</td>
<td>0.537</td>
<td>0.283</td>
<td>0.180</td>
</tr>
<tr>
<td>293</td>
<td>6.97</td>
<td>7.60</td>
<td>9.18</td>
<td>17.9</td>
<td>4.2</td>
<td>0.11</td>
<td>0.00</td>
<td>0.17</td>
<td>0.511</td>
<td>0.298</td>
<td>0.191</td>
</tr>
<tr>
<td>297</td>
<td>6.99</td>
<td>7.86</td>
<td>9.26</td>
<td>17.1</td>
<td>2.3</td>
<td>0.09</td>
<td>0.01</td>
<td>0.14</td>
<td>0.500</td>
<td>0.329</td>
<td>0.171</td>
</tr>
<tr>
<td>303</td>
<td>7.08</td>
<td>7.86</td>
<td>9.27</td>
<td>13.9</td>
<td>2.3</td>
<td>0.09</td>
<td>0.01</td>
<td>0.14</td>
<td>0.480</td>
<td>0.350</td>
<td>0.170</td>
</tr>
<tr>
<td>308</td>
<td>7.14</td>
<td>7.88</td>
<td>9.29</td>
<td>12.1</td>
<td>2.2</td>
<td>0.09</td>
<td>0.14</td>
<td>0.14</td>
<td>0.464</td>
<td>0.366</td>
<td>0.169</td>
</tr>
<tr>
<td>313</td>
<td>7.19</td>
<td>7.89</td>
<td>9.26</td>
<td>10.8</td>
<td>2.1</td>
<td>0.09</td>
<td>0.15</td>
<td>0.15</td>
<td>0.437</td>
<td>0.391</td>
<td>0.172</td>
</tr>
<tr>
<td>318</td>
<td>7.22</td>
<td>7.93</td>
<td>9.20</td>
<td>10.0</td>
<td>2.0</td>
<td>0.11</td>
<td>0.20</td>
<td>0.20</td>
<td>0.394</td>
<td>0.399</td>
<td>0.206</td>
</tr>
<tr>
<td>323</td>
<td>7.28</td>
<td>7.91</td>
<td>9.26</td>
<td>8.7</td>
<td>2.1</td>
<td>0.09</td>
<td>0.16</td>
<td>0.16</td>
<td>0.374</td>
<td>0.443</td>
<td>0.184</td>
</tr>
</tbody>
</table>

(a) $\log_{10}$ of isotropic diffusion constant $R$ for components I (slow), II (intermediate), and III (fast) expressed in units of s$^{-1}$. Average uncertainty in these parameters is 0.05.

(b) Average rotational correlation time for each component, calculated as $\tau = 1/6R$

(c) Inhomogeneous Gaussian derivative line peak-to-peak linewidth for the slowest component ($\Delta H_I$). and the fastest two components ($\Delta H_{II,III}$). Average uncertainty in these parameters is 0.1 G.

(d) Normalized fraction of components I (slow), II (intermediate), and III (fast). Average uncertainty in these parameters is 0.07. The individual uncertainties in $\phi_{II}$ and $\phi_{III}$ were propagated to estimate the uncertainties shown in the van’t Hoff plot (Figure 3.7).
Figure 3.6 Study of the temperature reversibility of the EPR spectrum of spin-labeled UmuD S60A. Shown are the initial spectrum at room temperature (top), exhibiting the broad line at low field corresponding to the low-mobility component, the spectrum at 323 K, the highest temperature studied (middle), and recovery of the spectrum upon returning to 297 K (bottom), with the initial spectrum taken at 297 K superimposed as a dashed line.

A van’t Hoff plot (Figure. 3.7) was used to quantify the changes in enthalpy and entropy for the dissociation of the arms from the globular domain. The change in enthalpy ($\Delta H^\circ = +17$ kJ/mol) reflects the energy cost of the N-terminal arms of UmuD unfolding and coming into contact with solvent. The change in entropy ($\Delta S^\circ = +50$ J/mol K) is also consistent with the arms dissociating from the C-terminal globular domain as temperature increases. The
relatively small free energy ($\Delta G^\circ = +2.1 \text{ kJ/mol}$) of dissociation of the arms suggests that once the arms are cleaved, the 24-amino acid peptide readily dissociates. The small free energy for the dissociation of the arms is comparable to that of formation of a single amide hydrogen bond in water, which is less than 1 kJ/mol [20]. In comparison, the favorable free energies of formation for bZip homodimers and heterodimer are 32-38 kJ/mol [21]. Therefore, the N-terminal arms of UmuD are highly dynamic and are at most only very weakly bound to the globular domain.

![Figure. 3.7](image_url)

**Figure. 3.7** A van’t Hoff plot, where $T$ is the temperature and $K$ is the equilibrium constant, obtained from relative populations of mobile (II) and immobile (I) UmuD-S60A components as a function of temperature ($K=\phi_{II}/\phi_I$). The best fit line provides an estimate for $\Delta H^\circ = +17 \text{ kJ/mol}$ and $\Delta S^\circ = +50 \text{ J/mol K}$.

### 3.4 DISCUSSION

Understanding the conformational dynamics of the N-terminal arms of UmuD is key in assessing the role of this protein in regulating the DNA damage response. In this work, we have shown that the N-terminal arms display a large degree of motion, and that they are mostly unbound from the globular domain and exposed to solvent. Most of the dynamics observed in the present study are in the fast-motion limit at the experimental frequency of 9
GHz. Much greater detail should be resolvable by measuring these spin-labeled variants at higher EPR frequencies such as 94 and 230 GHz.

The dissociation of the N-terminal arm from the body of UmuD allows for a large number of specific protein contacts to be made, both with the arms and with the globular domain.[6] As a protein with intrinsically disordered regions, the overall conformation of UmuD is greatly influenced by interactions with partner proteins and it is likely that the dynamics of the N-terminal arms will change under such conditions.[22] Many IDPs undergo a disorder-to-order transition when bound to their partners, which is generally characterized by decreased flexibility.[22] Indeed, when the N-terminal arms of UmuD$_2$ were cross-linked into the presumed elbows-down conformation shown in Figureure 3.1A, there was almost no reduction in the ability of UmuD$_2$ to interact with the β processivity clamp, an interaction thought to be central to the role of UmuD in regulating cellular responses to DNA damage [6].

Moreover, our study addresses the question of whether cleavage of the UmuD$_2$ arms in the cis conformation, in which one globular domain of the dimer cleaves its own arm, is possible.[11,17] The small free energy of dissociation and therefore highly dynamic nature of the arms suggest that the arms can readily alternate between cis and trans conformations. This is in agreement with the observation that a monomeric variant of UmuD is active and therefore must cleave in cis.[17] Overall, the dynamic N-terminal arms of UmuD allow it to adopt multiple conformations that are critical for its interactions with the DNA replication and damage response machinery.
In summary, EPR spin labeling has shown that the N-terminal arms of UmuD are highly dynamic and are mostly unbound from the globular domain. This conformation exposes binding surfaces that could facilitate interactions with multiple protein partners participating in both DNA replication and DNA damage responses, thereby regulating bacterial mutagenesis.

3.5 REFERENCES

Chapter 4: Dimer exchange and conformation-dependent cleavage of the SOS response protein UmuD
4.1 INTRODUCTION

The *umuD* gene products are critical components of the *Escherichia coli* (*E. coli*) SOS response, which occurs upon exposure to DNA damaging agents [1,2]. If DNA integrity is not maintained, the damaged DNA may lead to disruption of the replication fork and ultimately, cell death. The initial stage of the damage response pathway involves accurate repair processes, but subsequently shifts to a potentially mutagenic damage tolerance mode, which ensures continuation of replication and cell survival [2]. The damage tolerance pathway is regulated in part by UmuD2, its cleavage product UmuD′2, and their interactions with components of the cellular replication machinery.

Full-length UmuD2 is expressed at the onset of SOS; UmuD2, a dimer of 139-amino acid subunits, is the predominant form for the first 20-30 mins [2,3]. The presence of the damage-inducible RecA nucleoprotein filament (RecA:ssDNA) mediates the slow self-cleavage and removal of the N-terminal 24-amino acids of the UmuD2 homodimer, producing the UmuD′2 homodimer [4,5,6]. UmuD residues Ser60 and Lys97 form a Ser-Lys catalytic dyad, where lysine is proposed to deprotonate serine for subsequent nucleophilic attack on the peptide backbone [7]. This mechanism of cleavage also permits non-RecA-mediated cleavage at elevated pH [6,8].

UmuD2 cleavage was initially identified as a *trans* (domain-swapped, intermolecular, Figure. 4.1A) cleavage event, in which the active site of each C-terminal globular domain cleaves the arms (N-terminal 39 amino acids) between Cys24 and Gly25 of its adjacent partner [9,10], though it has also been shown that *cis* (intramolecular, Figure. 4.1B) cleavage is both possible
and efficient [11]. In cis cleavage, each N-terminal arm is cleaved by the active site in its respective C-terminal globular domain [11,12]. Furthermore, evidence that the monomeric UmuD protein variant retains substantial in vivo and in vitro function suggests that the wild-type UmuD₂ dimer can also cleave in the cis conformation [11]. Additionally, isoenergetic models of UmuD₂ show the N-terminal arms in trans (Figure 4.1A) and cis (Figure 4.1B) conformations with the elbows up or elbows down. In elbows up, each arm contacts the outer edge of the globular domain, permitting solvent accessibility of the active site region. In the elbows down conformation each N-terminal arm contacts the globular C-terminal domain and is positioned for cleavage. These models are supported by cross-linking experiments that were performed under physiologically relevant conditions [13,14].

The UmuD proteins have been shown to adopt multiple conformations in solution leading to the presentation by UmuD₂ and UmuD₂' of diverse binding surfaces for protein interactions [15,16,17,18,19,20]. X-ray [21] and NMR [18] structures show that the shorter arms of UmuD₂' are extended and disordered, and are unbound from the globular C-terminal domain. An analysis of the dynamics of the N-terminal arms of UmuD₂ as probed by Electron Paramagnetic Spectroscopy reveals that the arms are highly dynamic and are largely unbound from the globular domain [22]. High levels of deuterium uptake as determined by hydrogen exchange mass spectrometry also suggests that residues surrounding the arm and neck regions are solvent exposed and flexible [23].

There are also apparent structural differences in the dimer interfaces of the umuD gene products. For UmuD₂, efficient cross-linking of mono-cysteine derivatives at residues 37 and 38 was
observed, while the identical residues do not cross-link in UmuD’_2 [13]. For UmuD’_2 however, S57C cross-links robustly as compared to the same residue in UmuD_2 [16]. High-resolution structures of UmuD’_2 confirm that residues L40, N41, L43 and L44 of the α1 helix (neck region), as well as I73 and V135 are important for dimerization [18,19]. Intriguingly, the mutation N41D in the dimer interface renders UmuD’ a monomer, with no dimeric species detected at high concentrations. In the case of UmuD N41D, dimer was detected at higher protein concentrations, although the monomer was the predominant species [11]. This is despite the observation that UmuD lacking 45 N-terminal residues forms dimers as detected by cross-linking [19]. The NMR [18] structure of the UmuDD’ heterodimer, which was the only dimeric species detected upon cross-linking [24], has also been solved. It was determined that the dimerization interface of the heterodimer is similar to that of the UmuD’_2 homodimer[25]; however, in the case of UmuD_2 a region adjacent to the α1 helix also contributes to dimerization [17,18]. As in the UmuD’_2 homodimer, the secondary structure of the UmuDD’ heterodimer is composed mostly of β-sheets [17,18].

Differences in the conformation of the active site region are also evident in both high-resolution structures of UmuD’_2. In the X-ray structure, the active site residues are in the correct orientation for catalysis, while the NMR structure shows a more open conformation [18,21]. Additionally, circular dichroism (CD) experiments reveal that both UmuD_2 and UmuD’_2 proteins resemble random coils under physiological conditions including near-physiological concentration[20], leading to the suggestion that UmuD_2 and UmuD’_2 are intrinsically disordered proteins. However, the structure of full-length UmuD_2 may vary depending on the conformation of the N-terminal arms and interaction with partner proteins [16,20,22,26]. The umuD gene products have
been observed to adopt multiple conformations under different conditions, which are likely to play roles in modulating their multiple protein-protein interactions.

The *umuD* gene products perform distinct functions in preventing and facilitating mutagenesis. This entails a diverse array of specific interactions with the RecA:ssDNA nucleoprotein filament, translesion synthesis (TLS) DNA polymerases UmuC (polymerase subunit of DNA Pol V) and DinB (DNA Pol IV), the α, β and ε subunits of the replicative DNA polymerase III, as well as Lon and ClpXP proteases [3,5,12,16,27,28,29]. Affinity chromatography experiments suggest that UmuD′₂ interacts preferentially with the α catalytic subunit; while full-length UmuD₂ interacts more strongly with the β processivity clamp [3]. UmuD′₂ and UmuC (UmuD′₂C) form the active Y family DNA polymerase V (Pol V), which facilitates damage tolerance and mutagenesis [2,30]. Full-length UmuD₂ and UmuC (UmuD₂C) act non-catalytically in a DNA damage checkpoint [31,32], which slows DNA replication in response to DNA damage and allows time for more accurate repair mechanisms to be utilized [31]. While some evidence suggests that the UmuDD′ heterodimer also interacts with UmuC[25,33], there is no evidence that the heterodimer interacts with DinB [27]. UmuD prevents the -1 frameshift activity of DinB[27], whereas UmuD′₂C blocks accurate recombination repair by direct interaction with the RecA nucleoprotein filament [34,35].

UmuD₂ and UmuD′₂ are degraded by the Lon and ClpXP proteases at the termination of the SOS response [25,28,29]. Full-length UmuD is degraded by both Lon and ClpXP, while UmuD′ within the UmuDD′ heterodimer is specifically degraded by ClpXP [28]. In the heterodimer, ClpXP binds the UmuD arm at a specific recognition site (L⁹R¹⁰E¹¹I¹²), leading to the
degradation of the UmuD′ subunit [9,36]. This mechanism leads to low levels of UmuD′2 and the predominance of UmuDD′ in non-SOS cell [9,15,25]. It was therefore concluded that the degradation of the umuD gene products through this process occurs under both non-SOS and SOS conditions as a mechanism to keep the cell at or return to a non-mutable state, respectively [9,25].

It was previously demonstrated that UmuD dimers exchange readily [12], and that the subunits of UmuD2 and UmuD′2 exchange to form the UmuDD′ heterodimer [24]. This exchange occurs in solution despite equilibrium dissociation constants \( K_D \)s for dimerization for UmuD2 and UmuD′2 of <10\(^{-11}\) M.[20] It was proposed that the UmuDD′ heterodimer is the most stable and the predominant dimeric form through time-dependent formation of UmuDD′ heterodimers in which UmuD and UmuD′ subunits exchanged and were cross-linked using glutaraldehyde [24]. We subsequently isolated only UmuDD′ heterodimers from UV-irradiated cells under native conditions [11]. To probe the kinetics of exchange of the umuD gene products, we utilized Fluorescence Resonance Energy Transfer (FRET) to determine the relative stability and the rates of exchange of the UmuD2 and UmuD′2 homodimers, and the UmuDD′ heterodimer. We also determined the cleavage activity of heterodimer and homodimers of the umuD gene products in cis and trans. We find that dimers of the umuD gene products exchange relatively slowly and that the UmuDD′ heterodimer is the predominant, but not exclusive, dimeric form. For simplicity, we will refer to the UmuD protein dimers as UmuD and UmuD′.
Table 4.1 Strains and Plasmids

<table>
<thead>
<tr>
<th>Strains and Plasmids</th>
<th>Relevant Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td>PB103 AB1157</td>
<td>ΔumuDC ΔrecJ</td>
<td>P1 (JW2860) → GW8017 [37]</td>
</tr>
<tr>
<td>BL21 DE3</td>
<td></td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGY9739</td>
<td>α1umuD′C; pSC101-derived, SpecR</td>
<td>[38]</td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector; pSC101-derived, SpecR</td>
<td>[39]</td>
</tr>
<tr>
<td>pSG4</td>
<td>umuD', AmpR</td>
<td>[40]</td>
</tr>
<tr>
<td>pSG5</td>
<td>umuD, AmpR</td>
<td>[40]</td>
</tr>
</tbody>
</table>

4.2 EXPERIMENTAL PROCEDURES

4.2.1 Plasmids and proteins

*E. coli* plasmids and strains used in this work are listed in Table 4.1. Wild-type UmuD and UmuD′ proteins were purified from pSG5 and pSG4 plasmids, respectively, as previously described.[40,41] Variants of UmuD and UmuD′ were constructed in pSG5 and pSG4, respectively, and were purified as described.[41] The plasmid expressing RecA, pTXB-recA3, was a generous gift from Prof. Scott Singleton (University of North Carolina at Chapel Hill). The RecA protein used in this study was purified according to a published protocol.[42] For UV survival assays, the pGY9739 plasmid encoding *umuD* and *umuC* was used along with a derivative expressing UmuD G92C.[38] For these assays, pGB2 was used as the empty vector control.[38] UmuD and UmuD′ variants were constructed using a QuikChange kit (Agilent) and mutations were confirmed by DNA sequence analysis (Massachusetts General Hospital Core
Facility, Cambridge, MA). The primer sequence for UmuDG92C is as follows:

(5′ATCGCTGCTGTTGACTGCGAGTTTACGGTGAAAAAATTG)

4.2.2 UV survival assays

UV survival assays were performed as previously described in strain PB103.[40,41] The data represent an average of three trials and error bars represent the standard deviation.

4.2.3 Native gel and subunit exchange

To analyze the formation of the UmuDD′ heterodimer under non-denaturing conditions, equimolar amounts (5 µM) of UmuD and UmuD′ were combined and incubated for at least 15 min at room temperature. Electrophoresis was carried out as previously described using 10-20% native PAGE (Lonza) for 3.5 hr at 120 V.[11] Sypro Ruby gel stain (Invitrogen) was used to detect the different protein species. Fluorescence was detected by a Storm 860 phosphoimager using excitation wavelength of 635 nm. Protein bands were quantified using ImageQuant (GE Healthcare).

4.2.4 Protein Labeling with Alexa Dyes

Purified UmuD proteins at a concentration of 50 µM were labeled with thiol-reactive Alexa Fluor 488 maleimide (Invitrogen) or Alexa Fluor 647 maleimide (Invitrogen). For all FRET studies reported here, the UmuD or UmuD′ variants used for labeling were UmuD S60A C24A G92C or UmuD′ S60A G92C, respectively. Proteins were dialyzed against 50 mM K2HPO4 pH 7.4, at 4 °C using 0.1-0.5 mL capacity Slide-A-Lyzer dialysis cassettes (Pierce) with 7-kDa cutoff membrane. After dialysis, 50 mM Tris[2-carboxyethyl] phosphine (TCEP, Pierce) was
added to a final concentration of 1 mM. Labeling reactions were carried out with 8-10 equivalents fluorophore to cysteine. Both the protein solution and fluorophore dissolved in Dimethyl Sulfoxide (DMSO) were warmed separately at 37 °C for 5 min. Warmed contents were combined and mixed gently by shaking. The reaction was then wrapped in foil to protect from light and incubated at room temperature for 4 h, or overnight at 4 °C.

Sephadex G-25 columns (GE Healthcare) with dimensions of approximately 20 cm x 2 cm were pre-equilibrated with Qd buffer (20 mM Hepes, pH 7.5, 0.1 mM EDTA). Labeling reactions were then loaded onto prepared Sephadex G-25 columns and 1 mL fractions were eluted with Qd buffer. UmuD or UmuD’ elutes in the void volume. To confirm the presence of labeled UmuD or UmuD’, fractions were analyzed by 14% SDS-PAGE and labeled proteins were detected using a Storm phosphoimager at an excitation wavelength of 635 nm.

Fractions containing the most protein were subsequently pooled and loaded onto a 1 mL Hi-Trap Q column (GE Healthcare) pre-equilibrated with Qa buffer (20 mM Hepes, pH 7.5, 0.1 mM EDTA, 100 mM NaCl). A P-1 peristaltic pump with a flow rate of 1 mL/min was used to load the fractions containing labeled protein onto the Q column. The column was washed with Qa to remove any excess fluorophore. Protein was eluted with Qc (20 mM Hepes [pH 7.5], 0.1 mM EDTA, 500 mM NaCl) and 1 mL fractions were collected, up to 8 fractions. Fractions were again analyzed by 14% SDS-PAGE and labeled proteins were detected using a Storm phosphoimager. Protein was concentrated using spin concentrators with a 5-kDa-cutoff membrane (VivaSpin; GE Healthcare) and buffer was exchanged so that the protein was stored in Qe buffer (20 mM Hepes, pH 7.5, 100 mM NaCl, 5% glycerol) at -80 °C. Protein
concentration was determined by Bradford assay. Protein labeling efficiency in all cases was over 60%.

4.2.5 Fluorescence Resonance Energy Transfer (FRET) assays of subunit exchange

Fluorescence Resonance Energy Transfer (FRET) was used to monitor the rate of exchange of UmuD and UmuD’ subunits and the formation of the UmuDD’ heterodimer. Separate aliquots of UmuD C24A S60A G92C or UmuD’ S60A G92C were labeled with A488 and A647. Equimolar amounts of A488- and A647-labeled protein at 0.4 µM (except where noted) were combined in Qe buffer for FRET analysis using a Cary Eclipse spectrofluorometer (Varian). The excitation and emission wavelengths of the donor (A488) and acceptor (A647) are 495 nm and 667 nm, respectively.

Calculations used to determine the half-time ($t_{1/2}$) of the association reaction (including equations and derivations) have been previously described.[43] The following equation was applied to the data set to obtain the association rate constant ($k_{obs}$):

$$AB_t = AB_{max} (1-e^{-k_{obs}t})$$ (1)

In this equation, $AB_t$ is the amount of AB complex present at different time points, and $AB_{max}$ is the maximum of AB formed. The $k_{obs}$ was obtained by generating a best fit curve using GraphPad Prism. The corresponding $t_{1/2}$ for each data set was then calculated:

$$t_{1/2} = \frac{\ln2}{k_{obs}}$$ (2)
4.2.6 Homodimer and Heterodimer Cleavage assays

RecA: ssDNA-dependent and alkaline cleavage assays were carried out as previously described [11]. Total UmuD protein concentration used was 10 µM. Reactions were incubated in a 37 °C water bath and quenched at the time points indicated. Cleavage reactions were analyzed by 16% SDS-PAGE. Protein bands were detected and quantified as described above.

4.2.7 Thermal Shift Assay

The thermal shift assay was completed as previously described.[11] Heterodimer and homodimer reactions at a concentration of 20 µM were incubated for 30 min in Tm buffer (50 mM Hepes, pH 7.5, and 100 mM NaCl) before the addition of Sypro Orange (Molecular Probes). Detection, analysis and curve fitting was also completed as previously described [23].

4.3 RESULTS

4.3.1 Design of UmuD protein variants for FRET

We used FRET with two different pools of labeled UmuD to monitor subunit exchange. In order to carry out FRET experiments, it was necessary to design a single-cysteine variant of UmuD to provide a unique site for attachment of fluorescent labels. The UmuD or UmuD’ protein variants we used for labeling were UmuD S60A C24A G92C or UmuD’ S60A G92C, respectively. Non-cleavable S60A variants are routinely used to prevent cleavage of UmuD to UmuD’ [5,10,44]. UmuD contains only one cysteine (residue 24) per monomer at the cleavage site, which is between residues C24 and G25 [45]. Cleavage of UmuD to form UmuD’ removes the N-terminal 24 amino acids, which includes residue C24. In UmuD, the C24A mutation was constructed to prevent labeling on the N-terminal arm [44,46]. It has been previously determined that the N-
terminal arms of UmuD are quite dynamic and can adopt multiple conformations [11,12,22,26,47]. Based on the *cis* and *trans* models of UmuD and the NMR structure of UmuD′ (Figure 4.1. A,B,C), residue G92 is on the surface of the proteins and does not appear to interact with the N-terminal arms. The mutation G92C was constructed to provide a labeling site that is distant from both the active site and the arm binding site (Figure 4.1A,B). The UmuD G92C variant undergoes RecA:ssDNA-dependent cleavage *in vitro* (Figure 4.1D) and facilitates UV survival *in vivo* (Figure 4.1E) with an efficiency that is comparable to wild-type UmuD in both cases. UmuD G92C variant functionally mimics wild-type UmuD and can be used to monitor subunit exchange. Therefore, proteins were labeled with the Alexa Fluor dyes at position G92C on the globular C-terminal domain.
Figure 4.1 Homology models of UmuD [12] and the structure of UmuD′ [18] showing the location of amino acid substitutions constructed for FRET: A. The *trans* (elbows down) conformation of UmuD. B. The *cis* (elbows up) conformation of UmuD. C. NMR [18] structure of UmuD′. For clarity, chain A residues 1-24 (orange), 25-40 (red) and 41-139 (yellow) are depicted. Chain B residues 1-24 (purple), 25-40 (cyan) and 41-139 (blue) are also shown. The N-terminal 24 amino acids of UmuD are cleaved to form UmuD′. The remainder of the N-terminal arm is extended in UmuD′. UmuD and UmuD′ were labeled at site G92C (green sphere) on the globular C-terminal domain with Alexa dyes. In UmuD, the native residue C24 (gray) was changed to Alanine so that the protein is only labeled at the G92C position. The S60A (pink) active site mutation prevents cleavage of the N-terminal arms in *cis* or *trans*. D. Wild-type UmuD and UmuD G92C cleavage to UmuD′ in the presence of the RecA:ssDNA nucleoprotein filament after 1 hr at 37 °C. E. Survival assays carried out in strain PB103 (∆umuDCΔrecJ). pGY9739 (—, *umuDC*); pGY9739-G92C (▲, *umuDC*-G92C); pGB2 (●, empty vector). Error bars for three independent experiments are shown in D and E.
4.3.2 Dimers of the umuD gene products exhibit slow exchange kinetics

FRET was used to monitor the exchange of UmuD and UmuD' homodimers, as well as the UmuDD' heterodimers. Although it was previously demonstrated that the UmuD dimers exchange [10,11], the kinetics of this exchange has not been assessed in detail. UmuD dimer exchange was previously demonstrated by mixing the active site variant UmuD S60A with the UmuD C24D G25D cleavage site variant, which retains a functional active site.[10] Because cleavage was detected, it was concluded that the active site variant and cleavage site variant exchanged to allow cleavage in trans to occur [10,12].

To monitor subunit exchange of UmuD dimers over time, we labeled two separate pools of purified UmuD S60A C24A G92C with Alexa Fluor 488 (A488) and with Alexa Fluor 647 (A647). The A488 fluorophore served as the energy donor, and A647 as the energy acceptor. Upon mixing of the two different labeled proteins, we observed a gradual increase in the FRET signal as monitored at 667 nm (Figure. 4.2A). This indicates a slow time-dependent exchange of the subunits of the dimers as the A488 and A647 labels are brought closer together. There is little effect on the rate of exchange of the dimers when the acceptor concentration is either doubled or halved, suggesting that this process is concentration independent. This result is not surprising as these experiments were carried out above the reported equilibrium dissociation constant $K_D$ for dimerization for UmuD ($K_D < 10^{-11}$ M) [20]. Given the exceptionally low $K_D$ for homodimerization, UmuD and UmuD' are expected to be dimeric under most in vivo and in vitro conditions. The $t_{1/2}$ determined for the exchange of equimolar amounts of donor and acceptor labeled protein is 26 min (Figure. 4.2A).
Figure 4.2 UmuD$_2$ (D) and UmuD$'_2$ (D$'$) homodimers exhibit slow exchange kinetics. A. FRET of UmuD homodimers. B. FRET of UmuD$'$ homodimers. C. FRET of UmuD+D$'$ heterodimers. Concentrations analyzed are 1:2 (blue), 1:1 (green) and 1:0.5 (red), where 1:1 indicates mixing of 0.4 µM Alexa Fluor 488 (A488) labeled protein and 0.4 µM Alexa Fluor 647 labeled protein for FRET. The calculated half-time of the association ($t_{1/2}$) in minutes for each experiment are indicated on the graph. D-A488 or D-A647 represents UmuD S60A C24A G92C labeled with Alexa Fluor 488 or Alexa Fluor 647, respectively. D$'$-A488 or D$'$-A647 represents UmuD$'$S60A G92C labeled with Alexa Fluor 488 or Alexa Fluor 647, respectively.

FRET was also used to observe the rate of subunit exchange of UmuD$'$ dimers by using A488 and A647 labeled protein. There have been no previous reports of UmuD$'$ dimer exchange, presumably because exchange of UmuD$'$ dimers cannot be monitored by cleavage assays (see above). Subunit exchange of UmuD$'$ dimers displays sigmoidal kinetics with a sharp initial increase after approximately 10 minutes of mixing. In contrast to the UmuD dimers, the rate of exchange of subunits of UmuD$'$ dimers is modestly affected by a change in the concentration of
the acceptor-labeled protein (Figure. 4.2B). After mixing equimolar concentrations of donor and acceptor labeled protein, the observed $t_{1/2}$ is 21 minutes. When the concentration of acceptor labeled protein is either doubled or halved, the observed $t_{1/2}$ for the reactions are 24 and 15 minutes, respectively. This change in the rate suggests that the UmuD$'_2$ dimer subunits are more readily exchangeable than the full-length UmuD$_2$ dimer subunits.

We then mixed Alexa-647-labeled UmuD$_2$ and Alexa-488-labeled UmuD$'_2$ and monitored formation of UmuDD$'_h$ heterodimers (Figure. 4.2C). Under these equilibrium conditions, we observed that UmuD and UmuD$'$ dimers exchange to form stable heterodimers. We observe only slight changes in $t_{1/2}$ when the concentration of acceptor labeled UmuD was varied. The $t_{1/2}$ determined for the exchange of equimolar amounts of donor and acceptor labeled protein is 20 min (Figure. 4.2C).

### 4.3.3 The heterodimer is the predominant UmuD protein conformer

Previous work indicated that the heterodimer is the only dimeric species detected after 20 min when equimolar amounts of UmuD and UmuD$'$ are mixed and cross-linked with glutaraldehyde [24]. In this work, for heterodimer cross-linking to occur, UmuD and UmuD$'$ dimers must exchange to form the UmuDD$'_h$ heterodimer [18,24]. We wanted to determine whether the UmuDD$'_h$ heterodimer is preferred over the UmuD or UmuD$'$ homodimers under equilibrium conditions and without the possible confounding effect of potential differences in cross-linking efficiency. Therefore, we first mixed labeled UmuD and UmuD$'$ to form the UmuDD$'_h$ heterodimer and then added unlabeled UmuD or UmuD$'$ (Figure. 4.3). We found that the pre-formed heterodimer was not disrupted with the addition of either unlabeled UmuD (Figure.
4.3A) or UmuD’ (Figure. 4.3B), as there is no decrease in FRET intensity upon addition of unlabeled proteins. This indicates that once formed, the heterodimer does not readily exchange to form homodimers. We then carried out an analogous experiment in which UmuD’ dimers labeled with both A488 and A647 were initially formed as determined by FRET. Then, when unlabeled UmuD S60A C24A was added to this mixture of labeled A488 and A647 UmuD’, there was a marked decrease in FRET intensity (Figure. 4.3C) that began approximately 10 minutes after the addition of unlabeled UmuD protein. After the initial lag, the FRET intensity decreased steadily until about 40% of the FRET-detected UmuD’ dimers were disrupted. This is indicative of the preferential formation of the UmuDD’ heterodimer over UmuD’ homodimers (Figure. 4.3C).
**Figure 4.3** The heterodimer is the preferred UmuD protein conformer. A. Unlabeled UmuDS60A C24A at 0.8 µM is added to pre-formed heterodimer FRET reaction of 0.8 µM, total concentration. B. Unlabeled UmuD’ S60A at 0.8 µM is added to pre-formed heterodimer FRET reaction of 0.8 µM, total concentration. C. Unlabeled UmuD S60A C24A is added to the FRET reaction of labeled A488 and A647 UmuD’ homodimers resulting in a decrease in FRET intensity. No change was observed with the addition of unlabeled UmuD’ S60A to a FRET reaction of labeled A488 and A647 UmuD S60A C24A (data not shown). D. Preferential heterodimer formation under equilibrium conditions. UmuD and UmuD’ (wild-type heterodimer), and UmuD S60A and UmuD’ S60A (non-cleavable heterodimer) were combined in equivalent amounts and analyzed under non-denaturing conditions using native PAGE. The heterodimer is the predominant species for all combinations of wild-type and/or variant UmuD and UmuD’, with up to 70% detected for the non-cleavable heterodimer. Incubation times are indicated above the gel, however, it should be noted that under the native conditions in the gel, exchange is expected to take place.

We also used native polyacrylamide gel electrophoresis (PAGE) under equilibrium conditions to show that although the heterodimer is the preferred dimeric form, detectable amounts of UmuD and UmuD’ homodimers are present (Figure 4.3D). When we first mixed wild-type UmuD and UmuD’ dimers, we observed a disproportionate amount of UmuD’ homodimer, in addition to the
majority UmuDD' heterodimer detected by native PAGE. We hypothesized that even in the absence of the RecA:ssDNA nucleoprotein filament, this larger amount of UmuD' could result from cleavage of UmuD. However, on incubation of the heterodimer followed by analysis by SDS-PAGE, we did not observe cleavage, which suggests that cleavage was stimulated by the native PAGE conditions (data not shown). Therefore, we mixed non-cleavable variants of UmuD S60A and UmuD' S60A and observed predominantly UmuDD' heterodimer but still clearly detectable amounts of UmuD and UmuD' homodimers (Figure 4.3D). This is in contrast to previous work, in which the heterodimer was the only cross-linked species observed. Our observation that UmuD and UmuD' homodimers are also present along with the UmuDD' heterodimer suggests that the efficiency of cross-linking was a confounding factor in the previous work showing that the heterodimer was the only dimeric form [24].

A protein melting study was conducted to determine the stability of the UmuD heterodimer in comparison to the UmuD and UmuD' homodimers (Figure 4.4). These experiments utilize Sypro Orange which is weakly fluorescent in an aqueous environment. As the temperature increases and the protein unfolds, the dye binds to the hydrophobic regions of a protein and becomes highly fluorescent [48]. This is marked by an increase in fluorescence intensity, followed by a gradual decrease due to aggregation of the protein and the dye. The melting transitions observed for wild-type UmuD₂ and UmuD'₂ are consistent with previous reports. For UmuD, the first transition at 22 °C is due to the dissociation of the N-terminal arms from the body, and the second at 60 °C is due to the melting of the globular C-terminal domain[23]. The first and second melting transitions for UmuD S60A are at 23 °C and 59°C, respectively. UmuD' has only one melting transition at 61 °C, which corresponds to the melting of the globular domain, while
Umud’ S60A is modestly destabilized at 57 °C. For heterodimer analysis, equivalent amounts of purified Umud and Umud’ proteins were combined and preincubated to form the heterodimer as the predominant species (see Figure. 4.3D). The heterodimer melting profiles reveal two melting transitions, which also correspond to the dissociation of the N-terminal arms and of the globular domain. Melting temperatures for the heterodimers do not deviate significantly from the homodimers. For the wild-type heterodimer (Umud and Umud’), the first and second melting transitions are 22 °C and 61 °C, respectively.

**Figure. 4.4** Melting of the Umud protein homodimers and heterodimers. Melting profile of Umud S60A (green), Umud’ S60A (purple), and the non-cleavable heterodimer formed from mixing Umud S60A and Umud’ S60A (orange). Table summarizing melting temperatures is shown below.
4.3.4 Trans and cis heterodimers display different cleavage kinetics

Our observations above suggested that UmuD within the UmuDD’ heterodimer is cleavable. To characterize this further and to compare the cleavage kinetics in the different dimeric forms, RecA-mediated cleavage time-course assays were performed. When wild-type UmuD’ and UmuD S60A are mixed, cleavage can only occur upon exchange such that the active site of UmuD’ cleaves the arm of UmuD S60A, which is cleavage in trans. When wild-type UmuD and UmuD’ S60A are mixed, the UmuD arm can be cleaved by the active site of UmuD within either the heterodimer for cleavage in cis, or within the UmuD₂ homodimer for cleavage in cis or trans. Equivalent amounts of UmuD and UmuD’ were combined and preincubated before the cleavage reaction was initiated (Figure. 4.5). Cleavage reactions in the wild-type heterodimer (UmuD + UmuD’) and trans heterodimers (UmuD S60A + UmuD’) displayed similar kinetics to that of the UmuD₂ homodimer, with complete cleavage observed after 60 min (Figure. 4.5). Cleavage observed in the trans heterodimer (UmuD S60A + UmuD’) reaction must be a result of cleavage in the heterodimer conformation as the S60A substitution renders UmuD inactive. Therefore, the active site of the UmuD’ subunit is required for the cleavage reaction to occur. This observation suggests that RecA:ssDNA-mediated cleavage in the heterodimer conformation utilizes a similar mechanism and has equivalent efficiency to UmuD₂ cleavage. Though cleavage was observed for the cis heterodimer (UmuD + UmuD’ S60A) form, the reaction was significantly less efficient as compared to the UmuD₂ homodimer. This is somewhat surprising because in this reaction cleavage can also occur within the UmuD homodimer in the absence of subunit exchange. To determine whether the difference in cis and trans cleavage of the heterodimer was due to differential interactions with RecA:ssDNA, an alkaline cleavage assay was performed. This reaction occurs at pH 10 where S60 can be activated as a nucleophile in the absence of any
protein or DNA co-factors. We found that cleavage of the cis and trans heterodimer under these conditions were of similar efficiency, and comparable to that of the wild-type heterodimer. The cis conformation of the arms may be less effective in cleavage because it may not the preferred binding conformation for the RecA:ssDNA filament. Additionally, our observation that the cis reaction results in less cleavage overall than the trans reaction indicates that dimer exchange competes effectively with cleavage of wild-type UmuD.

Figure 4.5 Cleavage efficiency of the UmuD homodimer and the heterodimers. (A) UmuD₂ homodimer (blue), wild-type heterodimer (UmuD and UmuD’, brown), and trans heterodimer (UmuDS60A and UmuD’, green) exhibit similar RecA:ssDNA mediated cleavage kinetics where cleavage approaches nearly 90% after the reaction is carried out for 60 min. Rate of cleavage for the cis heterodimer (UmuD and UmuDS60A, purple) is significantly slower, with complete cleavage achieved after 90 min. Cartoons on the left depict experimental procedure, and correspond to the colors of the graph on the right. Red X shows position of the active site mutation (B) Cleavage under alkaline condition (pH 10) of the UmuD protein homodimer and heterodimers. RecA-mediated and alkaline cleavage reactions were resolved under denaturing conditions and UmuD bands were quantified to determine percent cleavage based on UmuD disappearance.
4.3.5 UmuD homodimer cleaves in cis

The wild-type UmuD homodimer potentially cleaves in both cis and trans conformations, while the monomeric UmuD N41D variant can only cleave in the cis conformation [11]. Both wild-type UmuD₂ and UmuD N41D display complete cleavage after 120 min, although the rate of cleavage is slower for UmuD N41D initially (Figure. 4.6). This result suggests that the wild-type UmuD homodimer cleaves in both cis and trans to achieve the efficiency that is observed. We also assayed cleavage in trans exclusively by combining the active site variant UmuD S60A and the cleavage site variant UmuD G25D [10] without preincubating the proteins (Figure. 4.6). For cleavage to occur, UmuD S60A and UmuD G25D dimers must first exchange in the presence of the RecA:ssDNA, and any cleavage detected results from cleavage in trans. Cleavage was less efficient in this context, indicating that the rate of cleavage in the trans dimers results from the time required for dimer exchange to first take place before cleavage can occur.
Figure 4.6 Comparing cis and trans cleavage using UmuD active site and cleavage site variants. RecA-mediated cleavage of wild-type UmuD (pink, cleaves in cis and trans), UmuD N41D (orange, cleaves in cis) and cleavage site variant UmuD G25D and active site variant UmuD S60A (gray, cleaves in trans with dimer exchange). Cartoons on the left depict experimental procedure, and correspond to the colors of the graph on the right. Red X on N-terminal arm and on C-terminal body shows the location of the G25D and S60A mutations, respectively. The percentage of cleavage product was determined as the ratio of the density of the UmuD′ band to the total density of UmuD and UmuD′ protein in each lane. For the trans with dimer exchange reaction, conversion of 50% of the reaction to UmuD′ is designated as full (100%) cleavage.

4.4 DISCUSSION

In this work we determined the kinetics of exchange of the UmuD₂ and UmuD′₂ homodimers, and UmuDD′ heterodimer. We also show that the UmuD′₂ dimers are exchangeable. Under equilibrium conditions, the UmuDD′ heterodimer is the predominant but not exclusive UmuD protein conformer, and once formed, is unlikely to re-exchange into homodimers. Moreover, the heterodimer can be cleaved in a RecA:ssDNA nucleoprotein-dependent manner to the UmuD′ homodimer, although cleavage within the heterodimer in the cis conformation was found to be less efficient than cleavage in trans in the presence of RecA:ssDNA.
The intrinsic error-prone activity of Pol V is tightly controlled on the transcriptional and post-translational levels [2]. At the onset of DNA damage, cellular levels of UmuD increase from approximately 180 to 2400 molecules, while expression of UmuC increases from approximately 15 to 200 molecules [49]. UmuD is later activated for its role in cellular mutagenesis by undergoing cleavage of the N-terminal 24 amino acids to form UmuD'. Previous work suggests that dimer exchange is utilized as a mechanism to decrease the number of active Pol V (UmuD'_2C) species as both UmuD_2C and UmuDD'_C were rendered insoluble and catalytically inactive [33]. This was accomplished by the addition of free UmuD_2 which resulted in the formation of increasing amounts UmuDD' over 1 h, and the release of insoluble UmuC [33]. However, our previous work also points to a role for the UmuDD' heterodimer during SOS mutagenesis, as the heterodimer was the only form of the protein observed shortly after exposure to ultraviolet radiation [11]. Although the experiments in the present work were performed in the absence of UmuC, we still observed slow exchange of UmuD_2 and UmuD'_2 to UmuDD' (Figure 4.2C), suggesting that the exchange property is independent of UmuC. Taken together, these observations suggest that the slow exchange of the active UmuD'_2C to form the insoluble UmuDD'_C complex allows sufficient time for TLS to occur, while gradually reducing the population of mutagenically active Pol V.

Formation and degradation of the UmuDD' heterodimer for ClpXP proteolysis allows for cells to return to the non-mutagenic state [24,25,29]. In the heterodimer, the UmuD subunit is tethered to ClpXP and delivers UmuD' for degradation. It was also determined that in the UmuD_2 homodimer, UmuD can deliver its partner UmuD for ClpXP degradation [29,36]. Therefore, exchange of UmuD or UmuD' subunits with the UmuD adaptor for ClpXP may be a mechanism for regulating degradation. Indeed, degradation of the UmuDD' heterodimer and UmuD
homodimer by ClpXP, along with degradation of UmuD homodimers by Lon provide additional levels of post-translational regulation at every stage of the SOS response [25,28].

We have also determined that UmuD within the UmuDD′ heterodimer is efficiently cleaved; specifically, the observed cleavage efficiency of the trans heterodimer is comparable to that of the UmuD₂ homodimer. For the heterodimer, trans cleavage was found to be more efficient than cis cleavage, though cis cleavage eventually reached the same high level as trans cleavage. RecA may have a binding preference for trans conformation over cis conformation of the N-terminal arms of the heterodimer since RecA cleavage was more efficient in the trans heterodimer than the cis heterodimer. This may be due to a preferential interaction of RecA with the UmuD subunit of the heterodimer. On the other hand, it does not appear that RecA discriminates between the UmuD N41D monomer, which cleaves in cis, or the wild-type UmuD dimer, which potentially cleaves in both cis and trans, as both were cleaved with similar efficiency.

UmuD₂ and UmuD′₂ homodimer exchange as well as the predominance of the UmuDD′ heterodimer regulate the SOS response in multiple ways (Figure 4.7). During the DNA damage response, the heterodimer can both up- and down-regulate the amount of UmuD′ in the cell by serving as a substrate for either RecA:ssDNA to allow cleavage to produce higher levels of UmuD′ or for ClpXP to facilitate degradation of UmuD′. This mechanism would create a highly dynamic environment throughout the SOS response in which UmuD′ levels are constantly changing in response to the level of DNA damage [50]. After TLS is completed, Pol V in the form of UmuD′₂C may exchange to form UmuDD′C which likely allows for dissociation of UmuC and subsequent degradation of the UmuDD′ heterodimer to terminate induced
mutagenesis. Since the heterodimer can serve as both the cleavage substrate of RecA:ssDNA and the degradation substrate for ClpXP, the heterodimer has the potential to both promote and suspend TLS.

![Diagram](image)

**Figure. 4.7** Model for the life cycle of the UmuD proteins. Formation of the RecA:ssDNA filament after DNA damage induces cleavage of the LexA repressor and expression of UmuD. RecA:ssDNA dependent cleavage of UmuD produces UmuD', the form that is active in TLS. UmuD and UmuD' irreversibly exchange to form the UmuDD' heterodimer. The heterodimer may be degraded by ClpXP or become a RecA:ssDNA substrate for cleavage to UmuD'. The UmuDD' heterodimer can be cleaved to UmuD' in the presence of RecA:ssDNA.

### 4.5 REFERENCES


Chapter 5: Characterization of newly identified forms of the polymerase manager protein UmuD
5.1 INTRODUCTION

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

Full-length UmuD is a homodimer of 139-amino acid subunits, and is expressed 20-30 minutes after the induction of the SOS response [1-3]. UmuD interacts with the RecA:ssDNA nucleoprotein filament to facilitate the slow autocleavage of the N-terminal 24-amino acids, forming UmuD′ [4-6]. The UmuD′ product is a homodimer of 115-amino acid subunits, and together with UmuC, forms the Y-family polymerase Pol V (UmuD′2C). This low fidelity polymerase copies damaged DNA, albeit possibly in an error-prone fashion, in a process known as translesion DNA synthesis (TLS) [1,3,7].

The N-terminal arms of UmuD are quite dynamic and adopt multiple conformations in regulating interactions with partner proteins [8-12]. UmuD cleaves in the *trans* (intermolecular) conformation, in which the arm of one monomer loops over the active site of the adjacent monomer [9,13]. Isoenergetic models of the UmuD dimer also propose that the *cis* conformation of the arms is also possible [9]. In *cis* (intramolecular), the arm binds its respective C-terminal domain [9]. Evidence that the monomeric variant UmuD N41D cleaves efficiently suggests that the *cis* conformation is likely an active conformation [14]. Additionally, the arms of UmuD may
be bound ("arms down") or unbound ("arms up") from the C-terminal domain which may significantly alter the interacting surface that is presented for binding [9].

The *umuD* gene products interact with multiple factors involved in DNA replication and the SOS damage response [15]. UmuD and UmuD’ interact specifically with Y-family polymerases UmuC and DinB. The noncatalytic UmuDC complex protects cells from the potentially harmful effects of error-prone DNA replication by delaying SOS mutagenesis [2,16]. This function is distinct from the role of UmuD’C in error-prone TLS [1,3,7]. Although the UmuDC and UmuD’C complexes have been studied for decades, difficulty purifying UmuC has left the localized interaction sites undetermined. Additionally, UmuD interacts with DinB at residue D91 on UmuD, resulting in an inhibition of the -1 frameshift activity of DinB [17]. Both UmuD and UmuD’ also interact differentially with the α, β, and ε components of the replicative polymerase DNA pol III [18].

The potentially deleterious effects of the *umuD* gene products call for tight regulation at the transcriptional and post-translational levels. The *umu* operon is repressed by LexA and is one of the most tightly controlled in the SOS regulon [1]. Cleavage of UmuD to UmuD’ not only activates UmuD’ for mutagenesis, but converts UmuD to a substrate that is insensitive to proteolysis by Lon [19]. UmuD and UmuD’ exist by themselves as homodimers, but can also exchange subunits to preferentially form the UmuDD’ heterodimer [9,12,20,21]. The UmuD’ subunit of the heterodimer is targeted for degradation by the ClpXP protease as a way of curbing mutagenesis [19,22,23]. A full-length UmuD subunit of the UmuD homodimer can also be degraded by ClpXP [22]. It was found that the N-terminal arm of UmuD carries the ClpX
recognition sequence, and that UmuD acts as the delivery factor for its bound UmuD′ or UmuD partner [23].

We identified N-terminally truncated versions of UmuD during purification of overexpressed full-length UmuD protein. These truncated proteins, UmuD 8 (UmuD Δ1-7) and UmuD 18 (UmuD Δ1-17), were used as tools to study the conformation of the N-terminal arms, the effect on cleavage, as well as the effect on protein-protein interactions. It is still unclear as to whether these truncated UmuD proteins are simply artifacts of overexpression, intermediates in the protein degradation process or cleavage products with physiological relevance. However, we considered that these truncations can be used to understand how potential binding partners interact specifically with the N-terminal arms of UmuD.

Interestingly, we found that the loss of just the N-terminal seven amino acids of UmuD results in significant changes in conformation of the N-terminal arms. Although UmuD 8 cleaves as efficiently as full-length UmuD in vitro and in vivo, UmuD 18 does not cleave to form UmuD′. We have also determined that UmuD 8 is proficient for UV mutagenesis, but oddly, does not confer resistance to UV radiation in a ΔrecJ strain. Additionally, by utilizing full-length UmuD, UmuD 8, the UmuD′ mimic UmuD 3A, UmuD′, and monomer variants UmuD N41D and UmuD′ N41D, we show that DinB affinity for the UmuD proteins is greatly impacted by the conformation of the N-terminal arms.
5.2 EXPERIMENTAL PROCEDURES

Table 5.1. Strains and Plasmids

<table>
<thead>
<tr>
<th>Strains and Plasmids</th>
<th>Relevant Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td>PB103 AB1157</td>
<td>∆umuDC ∆recJ</td>
<td>P1 (JW2860)→ GW8017 [24]</td>
</tr>
<tr>
<td>BL21 DE3</td>
<td></td>
<td>Laboratory Stock</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGY9739</td>
<td>αf umuD'; pSC101-derived, Spec&lt;sup&gt;R&lt;/sup&gt;</td>
<td>[25]</td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector; pSC101-derived, Spec&lt;sup&gt;R&lt;/sup&gt;</td>
<td>[26]</td>
</tr>
<tr>
<td>pSG4</td>
<td>umuD'</td>
<td>[27]</td>
</tr>
<tr>
<td>pSG5</td>
<td>umuD</td>
<td>[27]</td>
</tr>
</tbody>
</table>

5.2.1 Edman degradation of UmuD truncations

UmuD was purified as described [28]. UmuD proteins at concentrations of 20 µM were resolved by 16% SDS-PAGE. The proteins were transferred to a PVDF membrane in CAPS buffer (0.4% CAPS, 20% methanol, pH 10) for 2.5 hours at 50 V at 4 °C. The membrane was washed in distilled water several times, Coomassie Blue stained for 30 seconds, and destained in a solution of 40% methanol, 5% acetic acid for 1 min. The membrane was rinsed with distilled water for 5 min three times before it was air dried. For storage, the membrane was placed between two pieces of sterile membrane filter paper (Whatman), wrapped in foil, and placed in -20°C. The protein bands of interest were excised and sequenced (Tuft Core Facility, Boston, MA).

Construction of expression vectors for UmuD truncations— NdeI restriction sites were introduced into the pSG5 expression vector [27] at positions 8 for UmuD8, and 18 for UmuD18
using a QuikChange kit (Stratagene). There was already an NdeI site at the beginning of the umuD gene. The resulting plasmids were digested using NdeI (NEB), and re-ligated using T4 DNA Ligase (NEB). Mutations were confirmed by DNA sequence analysis (Massachusetts General Hospital Core Facility, Cambridge, MA). Mutagenic primer sequences are as follows:

umuDAsp8NdeI2 forward (5′-GTTGTTTATCAAGCATATGGATCTCCGCG)
umuDAsp8NdeI2 reverse (5′-CGCGGAGATCCATATGCTTGATAAAACAAC)
umuDPhe18NdeI2 forward (5′-GTGACTTTTCATATGTTTAGCGATCTTGTTCAGTG)
umuDPhe18NdeI2 reverse (5′-CACTGAACAAGATCGCTAAACATATGAAAAGTCAC).

UmuD8 and UmuD18 were constructed in pSG5, and expressed and purified as previously described [27,28].

KpnI restriction sites were introduced into pGY9739 at positions 1 and 8, and 1 and 18 to create the UmuD8 and UmuD18 truncations, respectively, using a QuikChange kit (Stratagene). The resulting plasmids were digested using KpnI (NEB), and re-ligated using T4 DNA Ligase (NEB). Mutations were confirmed by DNA sequence analysis (Massachusetts General Hospital Core Facility, Cambridge, MA). Mutagenic primers are as follows:

UmuDMet1KpnI19739 forward (5′-GGCAGGGTACCATGTTGTTTATCACC)
UmuDPhe18KpnI2 forward: (5′-GAAATTGTGACTGGTACCATGTTTAGGATC)
D9739Asp8KpnI forward: (5′-GATTATTATGTGTTTATCGGTACCATGGATCTCGCG)

5.2.2 In vitro characterization of UmuD truncations

Thermal shift assays of full-length UmuD, UmuD', UmuD8 and UmuD18 were completed as previously described [14]. UmuD N-terminal arms cross-linking with BMH at 10 μM was
completed essentially as previously described [29]. The RecA:ssDNA dependent and alkaline cleavage assays were also carried out as previously described [14,27].

5.2.3 UV survival and mutagenesis assays

Survival and mutagenesis assays were performed as previously described [27,28]. Data represent the average of at least three trials, and the error bars show the standard deviation.

5.5.4 Immunoblotting

Immunoblotting procedure was completed as previously described [14]. Where time points were taken, the amount of culture was normalized to the first time point at OD$_{600}$.

5.2.5 UmuD and DinB binding by Trp fluorescence

Determination of the binding constants of the UmuD proteins and DinB by Trp fluorescence was completed and analyzed as previously described [30]. DinB contains four tryptophans while UmuD contains none. To determine the fluorescence of DinB alone, 2 µM DinB in a buffer containing 50 mM Hepes (pH 7.5), 150 mM NaCl and 5% glycerol was excited at 278 nm. Fluorescence emission was monitored from 300 to 500 nm. To determine binding constants, UmuD proteins were then titrated into the sample containing DinB.
5.3 RESULTS

5.3.1 Multiple UmuD protein bands observed during purification

During the purification of wild-type or variant UmuD proteins, we often observed by SDS-PAGE distinct protein bands that were smaller than full-length UmuD, but larger than UmuD’. These bands were confirmed to be similar to UmuD by western blot, as they reacted with the α-UmuD/UmuD’ antibody (Fig. 5.1A). The purified protein samples were later sequenced by Edman degradation to determine the N-terminal sequence. We found that the first band below that of UmuD, as seen by SDS-PAGE, has the sequence of the N-terminal eight residues of D, L, R, E, I, V, T and F, which corresponds to UmuD residues 8-15 (Fig. 5.1B). This protein will be referred to as UmuD 8 from here on. Edman degradation of the protein corresponding to the second band reveals an N-terminal sequence that is consistent with UmuD residues 18-25, and will be referred to as UmuD 18 (Fig 5.1C). Residue C7 of UmuD 18 was not detected by Edman degradation, which is attributable to the high cross-linking efficiency of this residue [9,31] (see Fig 5.1C). We set out to identify the source of these UmuD truncations, but during characterization, we found interesting features that may be used to further study protein-protein interactions, in particular, UmuD and DinB.
Figure 5.1 Two more UmuD derivatives have been identified: UmuD 8 and UmuD 18 (A) Homology model of UmuD prepared in VMD (grey). Truncations that yield UmuD 8, UmuD 18 and UmuD' are highlighted in red. Full-length UmuD and UmuD' have been previously characterized (left). Immunoblot of purified UmuD protein under denaturing conditions showing four distinct UmuD protein bands. Relative amount of UmuD is indicated, with each band intensity normalized to that of full-length UmuD (right). (B) The sequence of the eight N-terminal residues of UmuD 8 and UmuD 18 were determined using Edman degradation. Residue candidates 1 (green) and residue candidates 2 (gray) are shown.

5.3.2 Loss of N-terminal residues result in considerable change in UmuD arm characteristics

Truncation of the N-terminal arms of UmuD to produce UmuD 8 and UmuD 18 result in a considerable change of the melting profile. It was previously shown that wild-type UmuD melts in two transitions. The first transition at approximately 30 °C is due to the melting of the N-terminal arms, while the second transition at approximately 60 °C is due to the melting of the
globular C-terminal domain [11]. As the arms of UmuD’ (residues 25-40) are not in contact with the C-terminal domain, only one melting transition is observed at approximately 60 °C [11]. The melting profile for UmuD 18 strongly resembles that of UmuD’ as there is only one transition observed at 62 °C. This suggests that the N-terminal arms of UmuD 18 are also not bound to the C-terminal domain. On the other hand, the N-terminal arms of UmuD 8 appear to be more stably bound to the C-terminal domain as there is a melting transition observed at 43 °C.

**Figure 5.2** Characterization of UmuD 8 and UmuD 18 *in vitro* (A) Thermal shift assays of UmuD proteins. Melting of the proteins as a function of temperature was monitored by changes in Sypro Orange fluorescence. Results for UmuD (purple), UmuD 8 (green), UmuD 18 (blue) and UmuD’ (red) are shown using 40 µM protein. (B) UmuD protein arms were cross-linked using BMH. Percent of cross-linked dimers are indicated under the lanes. Cross-linking reaction was carried out for 45 min at room temperature after the addition of BMH. Protein was visualized by Sypro Ruby staining. C. (left) Model of UmuD (arms-down) showing residues 1-7 (blue), 8-17 (purple) 18-24 (cyan), 25-40 (lime). UmuD 3A mutations: residue T14 (orange), L17 (tan), F18 (pink) and active site residue S60 (red) are also highlighted.
The conformation of the N-terminal arms of UmuD 8 and UmuD 18 was also assessed by BMH cross-linking (Fig. 5.2B). The homobifunctional BMH cross-linker is 13 Å in length and reacts with free cysteine thiols. The wild-type UmuD model with N-terminal arms in the “down” conformation, shows a distance of at least 20 Å separating the C24 residues within the dimer (Fig. 5.2C) [9]. Therefore, the arms must be unbound from the C-terminal domain for cross-linking to occur. As previously reported, the UmuD 3A variant is more readily cross-linked than wild-type UmuD. The UmuD 3A variant (T14A, L17A, F18A) is a biological UmuD’ mimic with full-length arms, in which the N-terminal arms are defective for binding to the C-terminal domain [9,11]. UmuD 8 and UmuD 18 are also readily cross-linked, suggesting that the arms are more likely to be unbound from the globular domain (Fig. 5.2B). Even a deletion of the first eight residues is sufficient to shift the N-terminal arms toward a conformation that is more readily cross-linked than wild-type UmuD.

5.3.3 UmuD 18 cleaves efficiently, UmuD 18 is not cleavable
Cleavage of UmuD to UmuD’ is required for the activation of Pol V (UmuD’2C) in translesion DNA synthesis (TLS) [1]. The removal of the N-terminal 24 amino acids is facilitated by binding of UmuD to the RecA:ssDNA filament, which positions the UmuD active site residues S60 and K97 in the correct orientation for the cleavage reaction [12]. Cleavage of UmuD 8 and UmuD 18 was assayed alongside full-length UmuD. We found cleavage of UmuD 8 to be quite efficient in a reaction carried out for 1h (Fig. 5.3A). Over time, cleavage of UmuD 8 mirrored that of wild-type UmuD (Fig 3C). We also assessed cleavage of UmuD 8 under alkaline conditions (pH 10) in the absence of the RecA:ssDNA filament and found that the cleavage was similar to that of full-length UmuD under these conditions as well. At pH 10, cleavage is less
efficient overall, but the active site S60 can be activated as a nucleophile without the addition of RecA:ssDNA. UmuD 18 does not undergo RecA:ssDNA facilitated cleavage (data not shown).

To rule out that the defect in UmuD 18 cleavage was due to a poorly formed active site, we performed a RecA:ssDNA dependent cleavage assay in which UmuD 18 and the active site variant UmuD S60A were mixed and allowed to form heterodimers. The N-terminal arms of UmuD S60A can then be cleaved in trans by the active site of UmuD 18 [9,13]. We found that cleavage in the context of UmuD S60A UmuD 18 dimers was indeed attainable (Fig. 5.3D). This confirms that UmuD 18 and UmuD S60A can form heterodimers and that the active site of UmuD 18 is competent for cleavage.
Figure 5.3 UmuD 8 cleaves as efficiently as wild-type UmuD, whereas UmuD 18 does not cleave to UmuD’. (A) Relative cleavage to UmuD’ in the presence of RecA:ssDNA nucleoprotein filament. (B) Results for cleavage to UmuD’ under alkaline conditions (pH) are also represented. Results are normalized to cleavage of wild-type UmuD to form UmuD’. (C) Comparison of the kinetics of cleavage of UmuD 8 and wild-type UmuD. Reactions were carried out over 6 h at 37 °C. Percent cleavage product was determined as a ratio of the density of the UmuD’ band to the total density of the uncleaved UmuD proteins and UmuD’ protein for each reaction. (D) Mixing equal amounts of UmuD 18 with the active site variant UmuD S60A results in cleavage. UmuD proteins at 10 μM were used and cleavage was carried out at 37 °C for 1 h.
5.3.4 UmuD 8 is proficient for UV-induced mutagenesis; UmuD 8 and UmuD 18 do not confer resistance to UV radiation

UmuD'C performs TLS in the face of UV damage and is required for UV-induced mutagenesis in *E. coli* [1]. To determine the proficiency of UmuD 8 and UmuD 18 in UV-induced mutagenesis, we compared the mutation frequency in Δ*umuDC* strains harboring plasmid-borne full-length UmuD, UmuD' and the truncated proteins UmuD 8 and UmuD 18 (Fig. 5.4A). We also compared the corresponding active site variant, S60A, of each protein. As Pol V (UmuD'_2C) inserts guanine opposite the 3'-thymine of (6-4) T-T photoproducts [32,33], polymerase activity can be detected via the reversion of the *argE3* auxotrophic marker in the *E. coli* biosynthetic pathway [27]. We found that the mutation frequency of cells expressing UmuD 8 is similar to that of full-length UmuD (Fig. 5.4A). Cleavage and expression level of UmuD 8 is also comparable to that of full-length UmuD *in vivo* (Fig. 5.4B). This suggests that UmuD 8 functions similarly to UmuD in this context, and is able to interact with protein partners that are required for mutagenesis, including UmuC, RecA and β [1]. However, UmuD 18, which appears to be similar to UmuD'_2C *in vitro*, is not proficient for UV-induced mutagenesis (Fig. 5.4A).
Figure 5.4 UmuD 8 is proficient for UV-induced mutagenesis. (A) Mutagenesis assays were performed with pGY9739 (umuDC), pGY9738 (umuD′C), pJNO8 (umuDC-Δ1-7) and pJNO18 (umuDC-Δ1-17) in strain GW8017. Immunoblot showing steady state expression levels of the UmuD proteins from plasmids pGY9739 (umuDC), pGY9738 (umuD′C), pJNO8 (umuDC-Δ1-7) and pJNO18 (umuDC-Δ1-17) in strain GW8017. A smaller and previously unidentified UmuD protein is also present in lanes 1 and 3. Work on identifying and characterizing this protein is currently in progress. Relative expression levels are shown below the blot, where uncleaved proteins are normalized to full-length UmuD WT (lane 1). Cleavage of UmuD WT (lane 1, 46%) and UmuD 8 (lane 6, 47%) are also indicated.

It was previously reported that ΔumuDC ΔrecJ strains are hypersensitive to UV radiation and that this phenotype can be suppressed by complementing with low-copy plasmids harboring the umuDC genes [14,34]. RecJ is an exonuclease that aids in DNA replication restart by degrading DNA at stalled replication forks [35,36]. In the absence of RecJ, replication restart is postponed
and DNA synthesis is carried out by the TLS polymerase Pol V [35]. We found that both UmuD 8 and UmuD 18 failed to confer resistance to UV light (Fig. 5.5B). In the case of UmuD 8, this observation was unexpected it is as proficient for UV-induced mutagenesis as full-length UmuD. This observation may suggest a direct interaction between UmuD and the exonuclease RecJ that facilities DNA repair after UV exposure.

**Figure 5.5** UmuD 8 and UmuD 18 do not confer resistance to UV radiation. (A) Survival assays were performed with pGY9739 (umuDC) and pGY9738 (umuD′C) in strain PB103. pGY9739 (▲, umuDC; pGY9739-S60A (●, umuDC-S60A); pGY9738 (■, umuD′C); pGY9738-S60A (♦, umuD′C-S60A); pGB2 (―, empty vector). (B) Survival assays with pJNO8 (umuDC-Δ1-7) and pJNO18 (umuDC-Δ1-17) in strain PB103. pJNO8 (♦, umuDC-Δ1-7); pJNO18 (●, umuDC-Δ1-17); pJNO8-S60A (■, umuDC-Δ1-7-S60A); pJNO18-S60A (▲, umuDC-Δ1-17-S60A). Error bars represent the standard deviation from three experiments.

### 5.3.5 UmuD N-terminal arm conformation affects DinB interactions

It was previously shown that DinB binds UmuD and RecA, an interaction that restricts the active site of DinB preventing -1 frameshift mutagenesis [17]. A $K_D$ for binding of 0.62 µM between UmuD and DinB was also determined with the primary interaction site at residue D91 on the C-terminal domain of UmuD [17]. It was also reported that an overproduction of either UmuD or UmuD’ reduced the -1 frameshift mutagenesis activity of DinB [17]. With this in mind, we
considered that there may be an interaction between UmuD′ and DinB that merits further exploration. In addition, we utilized in-house variants including the UmuD′ mimic with full-length arms UmuD 3A, the UmuD N41D and UmuD′ N41D monomers, as well as the N-terminally truncated UmuD 8 and UmuD 18 proteins, to further understand the effect of UmuD arm conformation on interactions with DinB.

**Figure 5.6** UmuD N-terminal arm conformation affects DinB interaction. Binding constants were determined by tryptophan fluorescence for A. UmuD 3A and DinB, B. UmuD 18 and DinB, C. UmuD′ and DinB, D. UmuD S60A and DinB, E. UmuD 8 and DinB. The curve represents the fraction of DinB fluorescence quenched by increasing concentrations of UmuD.
protein which produced the $K_D$ values shown for each figure. The data is normalized to $B_{\text{max}}$ and the error bars represent standard deviation from at least three independent experiments. F. (left) Model of UmuD (arms-down) showing residues 1-7 (blue), 8-17 (purple) 18-24 (cyan), 25-40 (lime). UmuD 3A mutations: residue T14 (orange), L17 (tan), F18 (pink) and active site residue S60 (red) are also highlighted. Residue implicated in DinB binding, D91, is also shown in brown. (right) NMR structure of UmuD’ with remaining arm residues 25-40 (cyan) shown. Residues S60 (red) and D91 (brown) are also highlighted.

By utilizing tryptophan fluorescence we have determined that the conformation of the N-terminal arms of the UmuD proteins greatly affects DinB binding to the globular C-terminal domain of UmuD. DinB displays increased affinity for UmuD 3A and UmuD 18 (Fig. 5.6AB, Table 1), proteins in which the N-terminal arms are unbound from the C-terminal domain. In comparison, DinB binding to UmuD S60A is somewhat weaker and may be a result of the N-terminal arms preventing DinB access to the area surrounding D91 on UmuD S60A (Fig. 5.6F). Moreover, the reduced arm flexibility of UmuD 8 as shown by thermofluor melting studies and BMH cross-linking (Fig. 5.2AB), also results in a significant weakening in the observed $K_D$ for binding between UmuD 8 and DinB. This suggests that DinB binds UmuD in the arms up position, where access to the residue D91 is uninhibited, and may provide further rationale as to how RecA:ssDNA mediated cleavage of UmuD is inhibited in the presence of DinB.
Figure 5.7 UmuD dimerization is not required for interaction with DinB. Binding constants were determined by tryptophan fluorescence for A. UmuD N41D monomer and DinB, B. UmuD′ N41D monomer and DinB. The curve represents the fraction of DinB fluorescence quenched by increasing concentrations of UmuD protein which produced the $K_D$ values shown for each figure. The data is normalized to $B_{max}$ and the error bars represent standard deviation from at least three independent experiments.

However, our results also suggest that UmuD arms in the up position are involved in binding to DinB. The $K_D$ for binding between UmuD′ and DinB is greatly reduced with the loss of the N-terminal 24 amino acids. Additionally, the monomers UmuD N41D and UmuD′ N41D also interact with DinB, and binding constants are within the range determined for the UmuD protein dimers. Like the UmuD′ dimer, the UmuD′ N41D monomer also displays weakened affinity for DinB as compared to the full-length UmuD N41D monomer. Taken together, this suggests that
UmuD binding to DinB does not require dimerization or a *trans* conformation of the N-terminal arms.

**Table 5.2** Dissociation binding constants for UmuD proteins and DinB determined as by tryptophan fluorescence.

<table>
<thead>
<tr>
<th>UmuD</th>
<th>$K_D$ for DinB</th>
</tr>
</thead>
<tbody>
<tr>
<td>UmuD S60A</td>
<td>7.4 ± 2.5 µM</td>
</tr>
<tr>
<td>UmuD 8</td>
<td>11.2 ± 2.8 µM</td>
</tr>
<tr>
<td>UmuD 3A</td>
<td>1.1 ± 0.3 µM</td>
</tr>
<tr>
<td>UmuD 18</td>
<td>1.3 ± 0.2 µM</td>
</tr>
<tr>
<td>UmuD'</td>
<td>11.4 ± 3.1 µM</td>
</tr>
<tr>
<td>UmuD N41D</td>
<td>4.3 ± 1.2 µM</td>
</tr>
<tr>
<td>UmuD' N41D</td>
<td>7.5 ± 2.0 µM</td>
</tr>
</tbody>
</table>

5.4 DISCUSSION

UmuD interacts with multiple proteins that are involved in DNA replication and repair. Some of the well characterized interactions include those with the β clamp, RecA, the TLS polymerase DinB, and the Lon and ClpXP proteases [10,17,37,38]. The primary interaction sites on UmuD for Lon are located on the N-terminal arms, while RecA and DinB predominantly interact with the C-terminal domain of UmuD. The β clamp and ClpXP protease also interacts with UmuD on both the N-terminal and C-terminal domains, although it was shown in both cases that the N-terminal arms are especially important for interaction [10,38]. We would expect a decrease in affinity between N-terminally truncated UmuD and protein partners that directly interact with the N-terminal arms, but how do the dynamic arms of UmuD, which have been shown to be either bound or unbound from the globular domain, impact the binding of proteins to the C-terminal domain of UmuD?
Although DinB interacts with residue D91 located on the C-terminal globular domain of UmuD, the conformation of the N-terminal arms appear to affect these interactions by either blocking or allowing access to C-terminal domain [17,37]. Indeed, our tryptophan fluorescence studies show that DinB has the highest affinity for the UmuD 3A and UmuD 18 variants, and this can be attributed to the arms up conformation of these UmuD variants [8,11]. On the other hand, the interactions between DinB and UmuD S60A or UmuD 8 are weakened as the arms are likely in the “down” conformation where access to the C-terminal domain of UmuD is somewhat reduced. Indeed, previous EPR studies of UmuD S60A show that the arms of UmuD are at most weakly bound to the globular domain in the absence of interacting partners [8]. Additionally, UmuD in which the arms at C24 were stably cross-linked to the body at F94C provides a surface that is required for stable interactions with the β clamp [10].

As previously discussed, the N-terminal arms of UmuD may also regulate interactions with DinB. Although residue D91 on UmuD is required for binding to DinB, it is possible that UmuD utilizes this residue for initial docking to DinB, and that this binding event induces conformational changes in the N-terminal region of UmuD that favors the arms up conformation. Our results also support the idea that UmuD and not UmuD’ suppresses the -1 frameshift activity of DinB as the N-terminal arms of UmuD are essential for interaction. Therefore, the DinB active site conformational change that prevents the accommodation of bulged templates in the presence of UmuD may indeed be induced by the N-terminal arms of UmuD.

UmuD protein interaction with the TLS polymerase DinB provides another example in which the N-terminal arms of UmuD mediate protein-protein interactions [15]. Moreover, this study
provides some of the first evidence for UmuD in the arms-up conformation binding to a partner protein. Thus far, studies of UmuD interactions with the β clamp and α subunit of the replicative polymerase Pol III all show that the UmuD arms down conformation is favored [10,30]. Perhaps, the arms up conformation of UmuD that is induced upon binding to DinB offers a mechanism by which proteins such as the β clamp and α subunit can be excluded from interactions with the DinB-UmuD complex.

5.5 REFERENCES


Chapter 6: Future Directions
The *umuD* gene products have been studied for over twenty years, resulting in a great deal of structural and functional information, as well as a growing protein-protein interaction network. The next challenge would be to put this information into context, and to develop a better understanding of the sequence of events. This is no trivial task, especially since many of the factors involved in DNA replication and the DNA damage response are not amenable to high resolution experimental methods. However, with the rapid development of single molecule and computational techniques, we can be optimistic that many answers are right around the corner.

As discussed in Chapter 2, the UmuD N41D and UmuD′ N41D monomers functionally mimic wild-type dimers, which raises the question as to whether dimerization is required *in vivo*. Therefore, we can use the UmuD monomer variants as tools in determining how UmuD structure and conformation affects interactions with DNA replication or DNA damage protein complexes. Do the UmuD proteins facilitate the switch between DNA replication and TLS? Is UmuD actively involved in the removal of the replicative polymerase from the replication fork, and does it play a role in regulating TLS polymerase access to the damage site? Designing a single molecule fluorescence assays in which the polymerase switch can be monitored may prove useful. For example, α can be used to extend a primer-template up to a lesion, followed by the addition of a TLS polymerase for lesion bypass. UmuD proteins may then be introduced into the experiment at various points in determining whether the UmuD proteins aid in the removal of α and or the αβ complex from DNA, or to further access the role of the UmuD proteins in polymerase switching.
The dynamics of the N-terminal arms of UmuD play a major role in the regulation of protein-protein interactions with the arms and the globular C-terminal domain (discussed in Chapter 3). It would be interesting to further access the dynamics of the arms in the presence of other proteins, either individually or in complex. Does binding of UmuD to a partner protein induce the arms “up” or “down” conformation? With this information, it can be speculated as to whether multiple binding events are feasible, based on the additional areas that may become exposed upon binding. A combination of EPR, molecular modeling and known binding constants may provide insights into the multiple binding events with the UmuD proteins.

How is the UmuD/D′ heterodimer the most stable conformer of the UmuD proteins, and what additional roles does it play in regulating mutagenesis? It was previously shown that the heterodimer is a substrate for the ClpXP protease[1,2], regulating the levels of the UmuD′, the form that promotes mutagenesis. However, recent results suggest that the heterodimer may play an active role in DNA repair as it was detected after treatment of cells with UV irradiation (discussed in Chapters 2 and 4). Does the heterodimer promote mutagenesis through interactions with UmuC? Furthermore, we have determined that both the UmuD and UmuD′ homodimers exchange. What is the biological relevance of homodimer exchange? Is exchange mediated by interaction with other proteins? From a structural perspective, the possible mechanisms of heterodimer formation are also quite intriguing. It is still not clear whether the homodimers dissociate and reassociate into heterodimers, or whether exchange occurs via the formation of a higher ordered structure.
The truncated UmuD proteins (discussed in Chapter 5), UmuD 8 and UmuD 18 may also be used to screen for partner proteins that interact specifically with the N-terminal arms, or the C-terminal domain where the arms may bind. We determined the $K_D$ for binding between DinB and full-length UmuD (previously determined), UmuD3A, UmuD', and UmuD 8. From this, it was concluded that the first seven amino acids of UmuD greatly influence that binding between the two proteins, although the primary interaction point on UmuD is located on the C-terminal domain at residue D91[3]. Similarly, we have found evidence for an interaction between UmuD and the RecJ exonuclease, based on in vivo UV survival and mutagenesis assays. This warrants additional study as it was previously thought that RecJ promotes the homologous recombination pathway. On the other hand, UmuD prevents mutagenesis while UmuD' promotes mutagenesis. It is plausible that the UmuD proteins may participate in the regulation of homologous recombination, or regulate whether DNA repair occurs via the homologous recombination or TLS pathways.

6.1 FURTHER ACKNOWLEDGEMENTS

Funding: We are pleased to acknowledge generous financial support from a New Faculty Award from the Camille & Henry Dreyfus Foundation (PJB), the NSF (CAREER Award, MCB-0845033 to PJB), a Cottrell Scholar award from the Research Corporation for Science Advancement (PJB), and the NU Office of the Provost.

Chapter 2: Acknowledgments—Jacquelyn Sikora performed and analyzed cleavage assays for Fig 2.4a and b. We thank April Gu in Civil and Environmental Engineering at Northeastern University for use of the real time PCR instrument. Srinivas Somarowthu and Alexandra Wallace
are kindly acknowledged for their help and technical assistance. We thank Česlovas Venclovas for careful reading of the manuscript.

Chapter 3: Acknowledgements—The EPR spectrometer is supported by NSF grant CHE-0443616. We also acknowledge technical assistance from Stefano Gulla, Mohammad Tehrani, Michael Ordazzo, Bilyana Koleva, and James Glick.

Chapter 4: Acknowledgements—We thank April Gu in Civil and Environmental Engineering at Northeastern University for use of the real time PCR instrument and Zhaohui Zhou in Chemistry and Chemical Biology at Northeastern University, for use of the fluorometer. We also thank Jana Sefcikova, Sarah Jensen, Tyler Cipriani, Michael Ordazzo, Bilyana Koleva, and Sharotka Simon (Brandeis University) for technical assistance.

Chapter 5: Acknowledgements—Qiuying Huang prepared samples for Edman degradation, Fig 1B,C, and performed cleavage assays and analysis for Fig. 3A,B. We thank April Gu in Civil and Environmental Engineering at Northeastern University for use of the real time PCR instrument and Zhaohui Zhou in Chemistry and Chemical Biology at Northeastern University, for use of the fluorometer. We also thank Michelle Silva, Tyler Cipriani and Philip Nevin for their technical assistance.

6.2 REFERENCES
