Specificity and Regulation of *Escherichia coli* Y Family DNA Polymerase V

A dissertation presented

by

Lisa A. Hawver

to

The Department of Chemistry and Chemical Biology
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the field of Chemistry

Northeastern University
Boston, Massachusetts

July 15, 2013
Specificity and Regulation of *Escherichia coli* Y Family DNA Polymerase V

by

Lisa A. Hawver

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Chemical Biology in the College of Science of Northeastern University
July 15, 2013
Abstract

Y family DNA polymerases are found in all domains of life and play important roles in many aspects of disease including bacterial virulence and cancer. In bacteria, these polymerases are regulated by the SOS response to DNA damage. The SOS response controls the expression of at least 57 genes, including *umuDC*, which encode UmuD’2C, or polymerase (pol) V. Pol V is further regulated, as it is only active in replication of damaged DNA once UmuD2 undergoes self-cleavage to UmuD’2. UmuD’2C’s cognate lesions are abasic sites as well as the UV photoproducts cyclobutane pyrimidine dimers and 6-4 thymine-thymine photoproducts, while DinB, the other Y family DNA polymerase in *E. coli*, bypasses N²-dG adducts such as N²-furfuryl-dG accurately and efficiently.

Y family polymerases have open active sites and lack specific protein features in the active site that ensure proper Watson-Crick base pairing. Y family polymerases possess two structural loops near the active site that we predict are important for binding the nascent base pair. We have defined these loops in *E. coli* DNA polymerase V as 3¹SNNDGCVi³⁸ (loop 1) and 5⁰KMGDP⁵⁴ (loop 2). I have shown that residues N32, N33, and D34 of loop 1 are important for pol V’s function as a DNA polymerase that bypasses thymine-thymine dimers caused by UV radiation. Additionally, I have shown that mutation of these residues reduce the ability of pol V to inhibit RecA-mediated homologous recombination, which suggests a role in the interaction between UmuC and the RecA protein.

After defining the important residues for pol V function, we wanted to understand the inherent basis of specificity by attempting to convert the specificity of pol V to that of DinB. Using hydroxylamine, we generated a library of UmuC mutations from which we then selected those that had the ability to confer resistance to nitrofurazone, which is believed to cause
predominantly $N_2$-furfuryl-dG lesions. I tested the ability of these selected UmuC variants to confer cellular survival after UV irradiation as well as their ability to bypass $N_2$-furfuryl-dG lesions in biochemical assays. All of the variants conferred resistance to nitrofurazone, but some did not show bypass of $N_2$-furfuryl-dG. Rather, they seem to disrupt important protein-protein interactions between UmuC and UmuD or UmuD', suggesting that protein interactions are important for specific lesion bypass and DNA damage tolerance. UmuC variants H282P and T412I specifically alter the cleavage of UmuD to UmuD', causing more cleavage than in the presence of wild-type UmuC. These two variants are also readily degraded by Lon protease. UmuC variant A9V shows decreased UV mutagenesis, depending on the length of the arm of UmuD'. The location of A9 in the active site is predicted to be adjacent to the incoming nucleotide, suggesting this residue may play an important role in the orientation of the incoming nucleotide.

This research has shown that not only is the active site important for the ability of pol V to bypass lesions, but protein-protein interactions also play a role in specificity of those lesions bypassed, as well as the overall ability of pol V to participate in TLS. Knowing which residues are important for specific protein-protein interactions have helped us to learn more about this specialized polymerase and how to alter its functionality.
Acknowledgements

I would first like to thank my advisor Dr. Penny J. Beuning. I was the last student that you formally trained in the lab, and I have always been proud of that. I know I could not be writing this today without your help and endless hours of editing and advising. Thank you for always listening to my ideas, and for pushing me in the right direction and for making me believe that I could get there on my own.

I would like to thank my dissertation committee members: Dr. Mary Jo Ondrechen, Dr. David Budil, Dr. Carla Mattos, and Dr. Samer Lone. Thank you for always being there for advice and for helping me through this big milestone.

I would like to thank the Department of Chemistry and Chemical Biology at Northeastern University.

I would also like to thank my amazing professors at RIT, who inspired me to become who I am today. Dr. Tubbs and Dr. Rosenberg – you have been the most amazing role models I could have asked for. Thank you.

I would like to thank all of my fellow lab mates in the Beuning Lab. I have always considered you all to be my (dysfunctional) family away from home. Dr. Jana Sefcikova, Dr. Jason Walsh, Dr. Jaylene Ollivierre, Dr. Michelle Silva, Ramya Parasuram, Nicholas DeLateur, Philip Nevin – we have grown up together as scientists and as friends and I couldn’t have asked for better friends. Also thanks to the new kids in the lab: Dr. Ingrid Salgado, Dave Murison, Nicole Antczak, Lisa Ngu, Caitlin Mills. Also thank you to the hard working undergrads who have
grown up and moved on to fantastic things – Danielle Kania, Stephanie Muser, Caitlin Gillooly, Megan Troy, Jennifer Endress – my time with you was wonderful and I wish you all the best in what your future holds – I am really proud of your accomplishments. I would also like to thank Mohammad Tehrani – without you, I wouldn’t have half of my thesis! Thank you for working really hard on the UmuC NFZ selection assay, and for giving me years of puzzles to solve and turn into an amazing degree. Also thanks to Anke Dierckx – it was a blast having you here, and you saved my sanity more times than I can count. Finally, thanks to the visiting students from Bridgewater State – Jim McIssac, Jeremy Mare, and Natasha Mariano – you are part of our lab, and I thank you for the endless laughs and for your support.

To my family: Mom – you have always been my number 1 fan, and you have pushed me to get this far. I wouldn’t be writing this without your help as my shoulder to cry or without you cheering me on. Thanks mom for always being there and finding a way to make it work – “it will get done – it always does.” To my sister Jamie – you have always been there to cheer me on as well, and I am really thankful that we grew up as friends. Thank you to my cute niece Gianna and my handsome nephew Michael – your aunt Lisa is a doctor now! I hope that I can someday show you this, and coach you through studying for your first big chemistry test. Also, thank you to Steve for listening to me complain through the years – it means a lot. Last, but not least, thank you to Joel, for being so patient through my heartache when experiments didn’t work right, for understanding that a grad student is really poor, and for saving the cork when we toasted to my first paper. Here’s to a few more corks!

I love you all.
Table of Contents

Abstract ............................................................................................................................................3

Acknowledgements ..........................................................................................................................6

Table of Contents .............................................................................................................................8

List of Figures ................................................................................................................................11

List of Tables ................................................................................................................................13

List of Abbreviations .....................................................................................................................14

Chapter 1: Background ..................................................................................................................19
  1.1 Introduction ............................................................................................................................19
  1.2 Transcriptional and Post-Translational Regulation of Y Family DNA Polymerases ..........21
  1.3 DNA polymerase V: UmuD'C ...............................................................................................24
  1.4 Discovery of pol V and SOS Mutagenesis ...........................................................................24
  1.5 Roles of pol V in Responding to Replication Stress ...........................................................26
  1.6 Specificity of pol V ...............................................................................................................28
  1.7 Notable UmuC Residues .......................................................................................................29
  1.8 Cellular Interactions of UmuC .............................................................................................36
  1.9 Specificity and Regulation of Polymerase IV, DinB .............................................................42
  1.10 Polymerase Switching .........................................................................................................45
  1.11 Summary .............................................................................................................................46
  1.12 References .........................................................................................................................47

Chapter 2: Characterization of Escherichia coli UmuC Active-Site Loops Identifies Variants that Confer UV Hypersensitivity .................................................................58
  2.1 Introduction ............................................................................................................................58
  2.2 Materials and Methods ........................................................................................................58
    2.2.1 Strains and Plasmids .......................................................................................................61
    2.2.2 UV Survival and Mutagenesis Assays ...........................................................................62
    2.2.3 Nitrofurazone (NFZ) and 4-Nitroquinoline-N-oxide (4-NQO) Survival Assays ..........62
    2.2.4 Immunoblotting ...............................................................................................................63
    2.2.5 Genetic Transduction .....................................................................................................63
  2.3 Results .....................................................................................................................................64
    2.3.1 Identification of UmuC active site loops .......................................................................64
    2.3.2 Variants in the N-terminal region of Loop 1 cause hypersensitivity to UV radiation 66
    2.3.3 N-terminal loop 1 variants confer sensitivity to 4-NQO and N33A causes sensitivity to NFZ .................................................................72
    2.3.4 Deletion of dnaQ or modification of the β-binding motifs of UmuC does not alter UV-hypersensitivity of N-terminal Loop variants ..........73
2.3.5 Variants in the N-terminal region of Loop 1 partially suppress inhibition of RecA-mediated homologous recombination ................................................................. 76
2.3.6 C-terminal region of Loop 1 does not play a significant role in UmuC function... 78
2.3.7 Variants in UmuC Loop 2 contribute little to UV survival .................................. 80
2.4. Discussion .................................................................................................................. 82
2.5 References .................................................................................................................. 88

Chapter 3: Point mutations in *Escherichia coli* DNA pol V that confer resistance to non-cognate DNA damage and enhance UmuD cleavage ........................................... 94
3.1 Introduction .................................................................................................................. 94
3.2 Materials and Methods ............................................................................................... 97
  3.2.1 Plasmids and strains .............................................................................................. 97
  3.2.2 Hydroxylamine mutagenesis and genetic selection of UmuC variants ............... 99
  3.2.3 NFZ survival assays ............................................................................................. 101
  3.2.4 UV survival and mutagenesis assays ................................................................... 101
  3.2.5 Immunoblotting .................................................................................................. 102
  3.2.6 Cold sensitivity ................................................................................................... 102
  3.2.7 Genetic transduction ............................................................................................ 103
  3.2.8 Pol V Purification ................................................................................................ 103
  3.2.9 Pol V in vitro lesion bypass assays ...................................................................... 104
3.3 Results ........................................................................................................................ 105
  3.3.1 Identification of UmuC variants that confer resistance to nitrofurazone *in vivo* ...... 105
  3.3.2 UmuC variants H282P and T412I weakly bypass $N^2$-furfuryl-dG *in vitro* ........ 108
  3.3.3 UmuC variants H282P and T412I enhance cleavage of UmuD to UmuD' *in vivo*.... 109
  3.3.4 UmuC H282P and T412I steady state levels are diminished in the presence of Lon protease .............................................................................................................. 111
  3.3.5 UmuC H282P and T412I fail to confer the *umuDC*-specific cold sensitive phenotype ................................................................................................................................. 112
  3.3.6 RecA-mediated homologous recombination is inhibited by UmuC H282P and T412I ............................................................................................................................ 113
  3.3.7 UmuC variants H282P and T412I retain the ability to confer UV Survival and UV-induced mutagenesis .............................................................................................. 115
3.4 Discussion .................................................................................................................. 118
3.5 References .................................................................................................................. 122

Chapter 4: Point mutation in *Escherichia coli* DNA pol V active site that confers resistance to non-cognate DNA damage has altered requirements for cognate damage ............................ 129
4.1 Introduction .................................................................................................................. 129
4.2 Materials and Methods ............................................................................................... 132
  4.2.1 Plasmids and strains ........................................................................................... 132
4.3 Results ........................................................................................................................ 135
  4.3.1 Identification of UmuC variant A9V that confers resistance to nitrofurazone *in vivo* .......................................................... 135
  4.3.2 UmuC variant A9V does not bypass $N^2$-furfuryl-dG *in vitro* ............................ 137
  4.3.3 UmuC variant A9V confers resistance to UV radiation yet cells harboring this variant exhibit low UV-induced mutation frequency .............................. 138
4.3.4 UmuC A9V fails to confer the umuDC-specific cold sensitive phenotype ............ 143
4.3.5 RecA-mediated homologous recombination is not inhibited by UmuC A9V .......... 144
4.4 Discussion ............................................................................................................... 146
4.5 References ............................................................................................................. 149

Chapter 5: Conclusions and Future Directions ......................................................... 155
5.1 References ............................................................................................................. 159
List of Figures

Figure 1.1 The domains and relative sizes of selected Y family polymerases .......................................................... 19
Figure 1.2. DNA damaging agents lead to the formation of lesions in DNA that disrupt replication and induce the SOS response .............................................................................................................................................. 22
Figure 1.3. DNA damage bypassed by DNA pol V ......................................................................................................................... 28
Figure 1.4. Model of notable residues of UmuC, the DNA polymerase subunit of UmuD′2C ................................................................. 30
Figure 1.5. Model of UmuC, the DNA polymerase subunit of UmuD′2C (Beuning 2006) .......................................................... 34
Figure 1.6. A comparison of biochemical requirements for TLS by pol V .......................................................... 37
Figure 1.7. Adducts of deoxyguanine bypassed by E. coli pol IV ......................................................................................................................... 42
Figure 1.8. E. coli UmuC is a Y family DNA polymerase that shares little homology with other members of the Y family .............................................................................................................................................. 65
Figure 2.1. N-terminal loop 1 variants 32, 33, and 34 cause hypersensitivity to UV radiation in a strain that is not normally sensitive to UV .............................................................................................................................................. 67
Figure 2.2. N-terminal loop 1 variants 32, 33, and 34 cause hypersensitivity to UV radiation in a wild-type strain .......................................................................................................................................................... 69
Figure 2.3. UmuC N-terminal loop 1 variants S31, N32, N33, and D34 confer hypersensitivity to UV radiation in a ΔumuC ΔrecJ strain .............................................................................................................................................. 71
Figure 2.4. UmuC N-terminal loop 1 variants S31, N32, N33, and D34 confer hypersensitivity to UV radiation in a ΔumuC ΔrecJ strain .............................................................................................................................................. 71
Figure 2.5. Cells harboring UmuC variants N32A, N33A, and D34A are sensitive to 4-nitroquinoline-N-oxide (4-NQO) .............................................................................................................................................. 73
Figure 2.6. Deletion of the proofreading subunit of pol III (dnaQ) or mutation of β clamp binding sites of UmuC do not suppress sensitivity to UV radiation caused by mutation of N32, N33, D34 .............................................................................................................................................. 75
Figure 2.7. UmuD′2C inhibits homologous recombination facilitated by RecA .............................................................................................................................................. 77
Figure 2.8. UmuC C-terminal loop 1 variants do not confer sensitivity to UV radiation and are proficient for mutagenesis .............................................................................................................................................. 79
Figure 2.9. UmuC loop 2 variants do not confer sensitivity to UV radiation and are proficient for mutagenesis .............................................................................................................................................. 81
Figure 3.1. E. coli Y family DNA polymerases UmuC (the polymerase component of pol V UmuD′2C) and pol IV, DinB, have specificity for distinct types of DNA lesions .............................................................................................................................................. 106
Figure 3.2. Cells harboring UmuC variants H282P and T412I confer resistance to NFZ in vivo .............................................................................................................................................. 107
Figure 3.3. Primer extension assay on undamaged DNA and DNA containing N2-furfuryl-dG with 32P-labeled standing start primer .............................................................................................................................................. 109
Figure 3.4. Immunoblot analysis of UmuC and UmuD′ steady-state levels show that UmuC H282P and UmuC T412I alter the extent of cleavage of UmuD to UmuD′ .............................................................................................................................................. 111
Figure 3.5. UmuC variants H282P and T412I do not confer a cold sensitive phenotype .............................................................................................................................................. 113
Figure 3.6. UmuD′2C inhibits RecA-mediated homologous recombination .............................................................................................................................................. 114
Figure 3.7. UmuC variants H282P and T412I confer resistance to UV radiation .............................................................................................................................................. 116
Figure 3.8 UmuC variants H282P and T412I confer high mutation frequencies.........................................................117
Figure 3.9 Immunoblot of purified wild-type pol V and variants H282P and T412I .................................................117
Figure 3.10 Summary findings from *in vivo* experiments of UmuC H282P and UmuC T412I and highlighted
UmuC residues that have been predicted to interact with UmuD′ and are important for protein
stability...........................................................................................................................................................120
Figure 4.1 *E. coli* UmuC homology model, based on the structure of Dpo4, depicting UmuC variants A9V and
A39V ..............................................................................................................................................................136
Figure 4.2 Cells harboring UmuC variant A9V confer resistance to NFZ *in vivo*. ..................................................137
Figure 4.3 Primer extension assay with a DNA template containing N²-furfuryl-dG, as well as an undamaged
DNA template................................................................................................................................................138
Figure 4.4 UmuC variant A9V confers resistance to UV radiation..........................................................................139
Figure 4.5. UmuC variant A9V when expressed from the *umuD′C* plasmid confers a very low mutation
frequency. ..........................................................................................................................................................141
Figure 4.6 UmuC variant A9V when expressed from the synthetically cleaved *umuD′C* plasmid has a very low
mutation frequency..........................................................................................................................................142
Figure 4.7 UmuC variant A9V fails to complement the cold sensitivity phenotype.................................................144
Figure 4.8 UmuD′₂C (pol V) inhibits RecA-mediated homologous recombination..........................145
List of Tables

Table 2.1  Strains and Plasmids ...............................................................61
Table 3.1  Strains and Plasmids ...............................................................98
Table 3.2  Nucleotide sequence analysis of umuD’C alleles containing missense mutations .................................101
Table 4.1  Strains and Plasmids ...............................................................133
Table 4.2  Nucleotide sequence analysis of umuD’C alleles containing missense mutations ........................................134
List of Abbreviations

% percent
' prime
~ approximately
α alpha
β beta
γ gamma
Δ delta, deletion
η eta
ι iota
κ kappa
ζ zeta
γ-HOPdG γ-hydroxypropano-deoxyguanosine
2-AAF N-2-Acetylaminofluorene
2-AF 2-aminofluorene
4NQO 4-nitroquinoline-1-oxide
6-4 T-T 6-4 thymine-thymine photoproduct
8-oxo-G 8-oxo-guanine
$^{32}$P Phosphorus-32
Å angstrom
A adenosine
A Alanine
Ala Alanine
Asn Asparagine
AP abasic (apurinic/apyrimidinic)
Arg Arginine
Asp Aspartic Acid
ATP adenosine triphosphate
ATPγS adenosine triphosphate gamma sulfur
BPDE Benzo[a]pyrene diolepoxide
BSA bovine serum albumin
C  carbon
C  cytosine
CEdG  $N^2$-(1-carboxylethyl)-2′-deoxyguanosine
cfu  colony forming units
Cys  Cysteine
CPD  cyclobutane pyrimidine dimer
D  Aspartic acid
da  deoxyadenine
dC  deoxycytosine
dG  deoxyguanosine
DNA  deoxyribonucleic acid
dNTP  deoxynucleotide triphosphate
dT  deoxythymine
E  Glutamic Acid
E. coli  Escherichia coli
EDTA  ethylenediaminetetraacetic acid
F  Phenylalanine
G  Glycine
G  guanine
Gln  Glutamine
Glu  Glutamic Acid
Gly  Glycine
g  g-force
GE  General Electric
H  Histidine
h  hour
H  hydrogen
H$_3$NO  Hydroxylamine
HEPES  4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
I  isoleucine
Ile  isoleucine
K    Lysine
kDa  kilodalton
Leu  Leucine
L    Leucine
L    liter
LB   Luria broth
Lys  Lysine
M    methionine
M    molar
mer  unit length of oligonucleotide
Met  methionine
mg   milligram
Mg   magnesium
µM   micromolar
min  minute
mL   milliliter
mM   millimolar
MW   molecular weight
N    Asparagine
N    nitrogen
\textsuperscript{2}-ff-dG \textsuperscript{2}-furfuryl-deoxyguanosine
\textsuperscript{6}-ff-dA \textsuperscript{6}-furfuryl-deoxyadenosine
NaCl sodium chloride
NFZ  nitrofurazone
nM   nanomolar
nm   nanometers
nt   nucleotide
OAc  acetate
O    oxygen
P    phosphorous
P    primer
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>Proline</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>pol</td>
<td>polymerase</td>
</tr>
<tr>
<td>pols</td>
<td>polymerases</td>
</tr>
<tr>
<td>Pro</td>
<td>Proline</td>
</tr>
<tr>
<td>Q</td>
<td>Glutamine</td>
</tr>
<tr>
<td>R</td>
<td>Arginine</td>
</tr>
<tr>
<td>RecA*</td>
<td>RecA nucleoprotein filament</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleotide</td>
</tr>
<tr>
<td>rNTP</td>
<td>ribonucleotide triphosphate</td>
</tr>
<tr>
<td>rNMP</td>
<td>ribonucleotide monophosphate</td>
</tr>
<tr>
<td>S</td>
<td>Serine</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>SOS</td>
<td>save our souls</td>
</tr>
<tr>
<td>SSB</td>
<td>single stranded DNA binding protein</td>
</tr>
<tr>
<td>T</td>
<td>Threonine</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>T</td>
<td>template; length of fully extended primer</td>
</tr>
<tr>
<td>T-T</td>
<td>thymine-thymine dimer</td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>TLS</td>
<td>translesion synthesis</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>Valine</td>
</tr>
<tr>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>W</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
<tr>
<td>XPV</td>
<td>Xeroderma pigmentosum variant</td>
</tr>
<tr>
<td>Y</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>
Chapter 1: Background

1.1 Introduction

The process by which DNA polymerases replicate damaged DNA is known as translesion synthesis (TLS) and was first described 35 years ago (Radman 1974). It was observed that DNA damage induced the *E. coli* SOS response, which is accompanied by mutagenesis of the DNA (Radman 1974). Originally DNA damage-induced mutagenesis was thought to result from the modification of replicative DNA polymerases, which allowed them to bypass DNA damage, albeit sometimes mutagenically (Woodgate 1989). However it was later discovered that the UmuC/UmuD’ complex (UmuD’$_2$C, pol V) and DinB (pol IV) are Y family DNA polymerases that have the specialized ability to carry out potentially mutagenic TLS (Yang 2005).

![Figure 1.1 The domains and relative sizes of selected Y family polymerases](Lone 2007). The DinB ortholog human DNA polymerase kappa is represented as hPolk. Dpo4 represents DNA polymerase IV from *Sulfolobus solfataricus*.

Y family DNA polymerases (Ohmori 2001), found throughout all domains of life, have five conserved sequence motifs but the overall size of the proteins can vary considerably (Figure 1.1) (Ohmori 1995; Boudsocq 2002; Yang 2005; Friedberg 2006; Yang 2007; Pata 2010). In addition to *E. coli* pol IV and pol V, the eukaryotic members of the family include Rev1, pol eta, pol kappa, and pol iota (Ohmori 2001). The domains of both replicative and Y family DNA polymerases are named for their general
resemblance to the parts of a right hand, including thumb, palm, and finger domains. Y family polymerases also possess a domain known as the ‘little finger’ domain unique to the Y family (Ling 2001). The Y family polymerases are characterized by small finger and thumb domains relative to replicative DNA polymerases, which result in an open, solvent-accessible active site in the palm domain of Y family members (Yang 2005; Chandani 2010). The active site of replicative polymerases contains an alpha helix, the role of which is to act as a steric check on fidelity and allow only a correct base pair to be formed (Kaushik 1996; Ogawa 2001; Beard 2003). Y family polymerases lack this alpha helix, contributing to their more open and flexible active sites and allowing them to accommodate lesions on the DNA template (Ling 2001; Jarosz 2007). The available crystal structures of Y family DNA polymerases (Pata 2010) tend to support the model of an open active site, as seen in the structure of Sulfolobus solfataricus Dpo4 in complex with DNA containing a thymine-thymine (T-T) cyclobutane pyrimidine dimer (Ling 2003). This structure demonstrates that both thymines are accommodated in the active site simultaneously (Ling 2003). Structures of other Y family DNA polymerases with or without DNA also generally show that these proteins have small finger domains and open, solvent-accessible active sites, suggesting a structural basis both for their ability to accommodate DNA lesions and for their low fidelity when copying undamaged DNA (Ling 2001; Silvian 2001; Zhou 2001; Yang 2005; Chandani 2009; Pata 2010; Sharma 2013).
1.2 Transcriptional and Post-Translational Regulation of Y Family DNA Polymerases

In *E. coli*, expression of the Y family polymerases along with other genes is induced via the SOS response to damaged DNA. This cellular response was named the SOS response by Miroslav Radman because there is a “danger signal which induces SOS repair” (Radman 1974). That “danger signal” is usually considered to be a DNA lesion that disrupts normal DNA replication (Radman 1974; Friedberg 2006). Evelyn Witkin suggested that there was a pathway in *E. coli* that is controlled by a repressor whose function is inactivated when DNA damage occurs and that again becomes activated as a repressor once the repair of DNA damage is complete (Witkin 1967). This repressor was discovered to be the LexA protein, the repressor of the SOS genes. The SOS response is initiated when a lesion in the DNA template prevents replicative polymerases from continuing with efficient replication, causing a region of single-stranded DNA (ssDNA) to develop (Figure 1.2). RecA is activated upon binding to ssDNA, forming a RecA/ssDNA nucleoprotein filament (Friedberg 2006). LexA then binds the RecA/ssDNA nucleoprotein filament, inducing LexA to cleave itself at its Ala$^{84}$-Gly$^{85}$ bond, approximately in the middle of the protein (Horii 1981). Once LexA is cleaved it no longer represses the SOS genes, allowing at least 57 genes, including *umuC*, *umuD*, and *dinB*, to be expressed during the SOS response (Friedberg 2006; Simmons 2008). In addition to its role in initiating the SOS response, RecA also plays more direct roles in the ability of Pol V to bypass lesions (see Section 1.8).
Figure 1.2. DNA damaging agents lead to the formation of lesions in DNA that disrupt replication and induce the SOS response. Single-stranded DNA (ssDNA) develops and becomes coated with RecA creating a RecA/ssDNA nucleoprotein filament, which signals the initiation of the SOS response. At least 57 genes are regulated by the LexA repressor, which represses the SOS genes by binding to consensus sequences (“SOS boxes”) in the promoter regions. LexA cleaves itself upon its interaction with the RecA/ssDNA nucleoprotein filament (Friedberg 2006; Simmons 2008). The cleavage of LexA ablates its repressor function and allows for the expression of the $umuDC$ and $dinB$ genes, among others. UmuD$_2$ also undergoes a cleavage reaction facilitated by the RecA/ssDNA nucleoprotein filament to form UmuD$'_2$, the active form in SOS mutagenesis (Burckhardt 1988; Nohmi 1988; Shinagawa 1988). The dashed curve in the UmuD$_2$ cartoon signifies that the arm is behind the globular domain, as the monomers are related to each other by a $C_2$ axis of symmetry. Both UmuD$_2$C (pol V) and DinB (pol IV) perform translesion synthesis (TLS) to bypass DNA adducts. It should be noted that it is not yet known exactly how UmuC and UmuD$'_2$ interact and the cartoon merely indicates that they form a complex.
The *umuD* gene products contribute an additional level of regulation of Y family DNA polymerases in *E. coli* (Friedberg 2006). Upon expression, UmuD₂ binds to the RecA/ssDNA nucleoprotein filament, stimulating the ability of UmuD to cleave itself at its Cys²⁴-Gly²⁵ bond and removing its N-terminal 24-amino acids to form UmuD'₂ by a mechanism similar to that of LexA (Burckhardt 1988; Nohmi 1988; Shinagawa 1988). The full-length UmuD₂ protein is the dominant form for approximately 20-40 min after expression is induced, after which time the cleaved form UmuD'₂ becomes the predominant form (Figure 1.2) (Opperman 1999). Full-length UmuD₂ and cleaved UmuD'₂ play distinct roles in the cellular response to DNA damage; UmuD₂ contributes to accurate DNA replication and repair while UmuD'₂ facilitates mutagenesis (Marsh 1985; Opperman 1999; Reuven 1999; Tang 1999; Sutton 2001; Godoy 2007; Ollivierre 2010). Therefore, this lag in the appearance of UmuD'₂ delays the use of a potentially mutagenic pathway, in part via direct interactions between the *umuD* gene products and Y family DNA polymerases.

Furthermore, the levels of UmuD, UmuD', and UmuC are regulated by ClpXP and Lon proteases (Frank 1996; Gonzalez 1998; Gonzalez 2000; Neher 2003). Lon protease recognizes two sites on UmuD, with the primary site ₁⁵FPLF₁⁸ and a secondary site ₂⁶FPSP₂⁹ (Gonzalez 1998). UmuD' is rapidly degraded by ClpXP when it is in the UmuDD' heterodimeric form (Frank 1996). Lon protease degrades UmuD and UmuC rapidly (Frank 1996).
1.3 DNA polymerase V: UmuD'2C

Y family DNA polymerase pol V, discovered in an experiment to identify nonmutable mutants of *E. coli*, is encoded by the *umuDC* genes and induced by the SOS response (Kato 1977; Bagg 1981). This specialized polymerase consists of one molecule of UmuC which possesses the polymerase activity, and one UmuD'2 dimer, therefore pol V is also referred to as UmuD'2C. Pol V is responsible for the majority of UV-induced mutagenesis in *E. coli* and it has been shown to bypass common lesions from UV radiation.

1.4 Discovery of pol V and SOS Mutagenesis

The *umuC* gene was discovered by characterizing mutants of *E. coli* that were deficient for UV-induced mutagenesis but were still viable (Kato 1977). Kato and Shinoura screened for UV-nonmutable (Umu) mutants by using the mutagen 4-nitroquinoline-1-oxide (4NQO) and then using UV irradiation in a second screen (Kato 1977). The *umuC* gene appeared to encode a protein that participates in “mutagenic repair” as well as reactivation by UV irradiation (Kato 1977). Steinborn independently discovered *uvm* mutants that are deficient in UV mutability using a similar method to Kato and Shinoura (Steinborn 1978). It was thought that perhaps the *uvm* gene was related to *umuC* in that they displayed similar nonmutable characteristics when *uvm* mutants were exposed to UV light, a phenotype of UmuC that is now well known (Steinborn 1978). Characterization of the *umuC* gene and its role in SOS mutagenesis was carried out long before the biochemical function of its gene product in translesion synthesis was determined.
Polymerase V is an error-prone DNA polymerase that is responsible for most SOS mutagenesis. It was first suggested that the UmuC and UmuD proteins were mediators that enabled DNA polymerase III holoenzyme (pol III HE), which typically stalls at sites of damage, to bypass DNA lesions (Bridges 1985a; Rajagopalan 1992). This stalling may occur for one of two reasons: either the polymerase cannot recognize the lesion as an instructional base, or the exonucleolytic proofreading subunit of pol III HE recognizes any base insertion as incorrect and hydrolyzes the newly incorporated nucleotide (Walker 1984; Bridges 1985b; Woodgate 1989; Fujii 2004a; Friedberg 2006). It was thought that UmuC-UmuD’ allows pol III HE to successfully replicate past a DNA lesion but with low fidelity (Woodgate 1989; Rajagopalan 1992). A two-step model for UV mutagenesis was proposed. In the model, the first step involves a RecA-mediated misincorporation event opposite the lesion. In the second step, umuC extends the primer from the misincorporated nucleotide, allowing replication to continue beyond this point (Bridges 1985a; Bridges 1985b). At the same time, it was discovered that the umuD gene product participates along with UmuC in this two-step process (Bridges 1985a).

Tang et al. showed that UmuD’2C in the presence of RecA, beta clamp, gamma clamp loader, SSB, and either polymerase III or II facilitates the bypass of an abasic lesion, at which time it was speculated that UmuD’2C had polymerase activity (Tang 1998). By 1999, two separate groups reported purifying different forms of the UmuC protein. The Livneh group purified a soluble form of UmuC, an N-terminal fusion with maltose binding protein (MBP) (Reuven 1999). The Woodgate and Goodman groups collaborated to purify a soluble complex UmuD’2C (Bruck 1996; Tang 1998; Tang 1999). Both groups determined that UmuD’2C is in fact a DNA polymerase (Reuven 1999; Tang
1999). UmuC had weak DNA polymerase activity, but with the addition of cofactors such as UmuD′, RecA, and SSB, this activity increased despite the absence of Pol III HE (Reuven 1999). The UmuD′2C complex also inhibits homologous recombination mediated by RecA, suggesting that the appearance of pol V actively prevents the relatively accurate recombination repair pathway while enabling SOS mutagenesis (Sommer 1993a; Szpilewska 1995; Rehrauer 1998).

A simple method of purification for pol V has been recently published, in which UmuD′ is expressed from an inducible promoter, while UmuC is expressed at basal levels (Karata 2012). Pol V had strong polymerase activity on a single stranded circular DNA template coated in SSB in the presence of the RecA nucleoprotein filament (RecA*) and the β clamp and γ clamp loader complex (Karata 2012). Purified pol V was able to weakly insert nucleotides opposite a thymine-thymine dimer even in the presence of only RecA* and SSB, though the addition of the β clamp and γ clamp loader leads to an increase in processivity (Karata 2012) (see section 1.8).

1.5 Roles of pol V in Responding to Replication Stress

Upon DNA damage, replication forks undergo regression; these regressed forks are stabilized by RecA and RecF (Courcelle 2003a; Courcelle 2003b). RecJ and RecQ partially degrade nascent DNA at stalled replication forks, while preventing TLS from occurring (Courcelle 2001). Recovery of DNA synthesis typically occurs once the generally accurate nucleotide excision repair process removes the lesion (Courcelle 2005). If the capacity of nucleotide excision repair is exhausted, pol V specifically allows DNA replication to recover (Courcelle 2009). In the absence of umuDC, recovery of DNA synthesis is modestly delayed (Courcelle 2005). However, in the absence of recJ,
replication restart is significantly delayed and in the absence of both recJ and umuDC, replication essentially does not recover. Without RecJ present to process the stalled replication fork, TLS by pol V is required for survival (Courcelle 2005; Courcelle 2006).

When UmuC and UmuD are overexpressed in *E. coli*, strains are cold sensitive, meaning the cells exhibit extremely slow growth at 30 °C without a growth defect at 42 °C. The cold sensitivity phenotype correlates with a specific decrease in the rate of DNA replication (Marsh 1985). This decrease in the rate of replication likely delays restart of replication in response to DNA damage to allow time for accurate methods of DNA repair such as nucleotide excision repair to operate and therefore may serve as a primitive DNA damage checkpoint (Opperman 1999). In this model, cleavage of UmuD to form UmuD’ releases the checkpoint, in part because UmuD’ has lower affinity for the beta clamp than does UmuD (Sutton 1999; Sutton 2002), and allows TLS to occur (Opperman 1999; Sutton 2001; Sutton 2006a). The cold sensitivity phenotype conferred by UmuD and UmuC is independent of their roles in TLS (Sutton 2001). Taken together, these observations suggest that inappropriate levels of the *umuDC* gene products, whether because they are deleted or overexpressed, disrupt the cellular responses to DNA damage or replication stress.
Figure 1.3. DNA damage bypassed by DNA pol V. *E. coli* DNA polymerase V is known to bypass common lesions that occur from UV radiation, such as thymine-thymine (T-T) *cis-syn* cyclobutane pyrimidine dimers (CPD) and T-T (6-4) photoproducts, as well as abasic sites and the C<sup>8</sup>-dG adduct formed from *N*-2-acetylaminofluorine (C<sup>8</sup>-AAF) (Becherel 1999; Tang 2000; Fujii 2004a; Fujii 2004b). The table shows the identity of the nucleotide inserted by pol V opposite each lesion.

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Pol V inserts</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-T CPD</td>
<td>A-A</td>
</tr>
<tr>
<td>T-T (6-4) dimer</td>
<td>A-G</td>
</tr>
<tr>
<td>Abasic site</td>
<td>A</td>
</tr>
<tr>
<td>C&lt;sup&gt;8&lt;/sup&gt;-AAF-dG</td>
<td>C</td>
</tr>
<tr>
<td>N&lt;sup&gt;2&lt;/sup&gt;-B[a]P-dG</td>
<td>A/C</td>
</tr>
</tbody>
</table>

1.6 Specificity of pol V

Polymerase V bypasses common lesions that result from UV radiation, such as thymine-thymine (T-T) *cis-syn* cyclobutane pyrimidine dimers (CPD) and T-T (6-4) photoproducts, as well as abasic sites and the C<sup>8</sup>-dG adduct formed from *N*-2-acetylaminofluorine (C<sup>8</sup>-AAF) (Figure 1.3) (Becherel 1999; Tang 2000; Pham 2001; Fujii 2004a; Fujii 2004b). Pol V efficiently bypasses UV photoproducts as well as abasic sites when in the presence of the beta clamp, RecA/ssDNA, and SSB (Tang 2000; Karata 2012). Pol V inserts G six-fold more frequently than it inserts A opposite the 3′-T of T-T (6-4) photoproducts, consistent with its mutagenic signature *in vivo* (Becherel 1999; Tang 2000). The CPD UV photoproduct is bypassed in an error-free manner (Becherel 1999; Tang 2000). *N*-2-acetylaminofluorine forms adducts at the C<sup>8</sup> position of the guanine
base (Fujii 2004b), giving C^8-AAF-dG, which is bypassed in an error-free manner by pol V when the lesion occurs in the context of (5'-GGCG^AAFFCC-) (Fujii 2004b). These lesions also commonly cause -2 frameshift mutations in continuous G sequences, including the NarI sequence, as well as -1 frameshift mutations in continuous three or four G sequences (Friedberg 2006). Pol V efficiently bypasses N^6-benzo[a]pyrene-dA and is implicated in bypass of some N^2-benzo[a]pyrene-dG adducts (Figure 1.7), as well as some oxidized bases (Lenne-Samuel 2000; Shen 2002; Seo 2006; Neeley 2007). Pol V displays error rates of 10^-3 – 10^-4 when copying undamaged DNA, therefore it is even less accurate than DinB (Tang 2000).

1.7 Notable UmuC Residues

There is currently no crystal structure of UmuC, so interpretation of experiments with UmuC variants must rely on models based on homology to Dpo4 and other Y family DNA polymerases (Lee 2005; Beuning 2006). UmuC variants with mutations at several residues have been characterized (Figure 1.4). Only recently was pol V efficiently purified (Karata 2012). Due to this, most characterization of variants has been carried out in vivo using complementation assays.

Residues D101, E102, and D6 (Figure 1.4A) are strictly conserved catalytic residues for metal ion binding, with D101 and E102 part of the conserved S[LI]DE motif (Steinborn 1978; Ling 2001). The UmuC D101N (umuC104) variant made by Steinborn cannot carry out UV-induced mutagenesis and polymerase activity is severely diminished (Steinborn 1978; Ling 2001).
Figure 1.4. Model of notable residues of UmuC, the DNA polymerase subunit of UmuD′C (Beuning 2006). Residues are highlighted as follows: (A) D6, D101, and E102 (silver) are conserved catalytic residues (Steinborn 1978; Ling 2001; Boudsocq 2002); (B) F10 and Y11 steric gate (orange) (Shurtleff 2009), A9 (blue) and A39 (cyan) are predicted to be located adjacent to the incoming nucleotide in the active site adjacent to the steric gate and are important for mutagenesis, UV resistance, the cold sensitive phenotype, and alleviation of the pol V inhibition of the RecA-mediated homologous recombination phenotype (Chapter 4, this dissertation) (Marsh 1991); (C) 31SNND34 (orange) are proposed to control the size of the active site (124) and 32NND34 (orange) are important for UV survival and UV mutagenesis (Hawver 2011); V29, L30, and V37 (pink) are thought to anchor an active site loop; I38 (pink) is likely to be adjacent to the incoming nucleotide (Chandani 2009), A39 (cyan) is important for mutagenesis, UV resistance, and the cold sensitive phenotype (Marsh 1991), UmuC loop 1 31SNNDGCV138 is highlighted in red, and UmuC loop 2 32KMGDP34 is highlighted in blue (Hawver 2011); (D) 32NND34 (orange), 59KDLFR64 (green), K25, C65, G66, W55, F56 (all in black) are thought to be important for RecA interaction (Hawver 2011; Chandani 2013); (E) 313LTP315 (violet) is the site of interaction with the dimer interface of the beta clamp (Beuning 2006); Q353 (black) is the C-terminal residue of the model. This model of UmuC is a truncation, as the C-terminal domain of UmuC lacks homology to proteins of known structure. The UmuC beta binding motif 35QLNL361 would be located just after Q353 (black) (Dalrymple 2001; Beuning 2006).
The steric gate residue of UmuC has been identified as Y11 (Figure 1.4B) (Shurtleff 2009). F10 is the residue N-terminal to the steric gate residue and was chosen for analysis based on analogy to the F13 steric gate residue of DinB (Shurtleff 2009). Mutating either UmuC F10 or Y11 to alanine caused cells harboring these variants to be sensitive to UV light (Shurtleff 2009; Kuban 2012), which was alleviated by combining either of these mutations with ablation of the beta-binding motif in UmuC (Shurtleff 2009). This observation suggests that the toxic effect of the F10 or Y11 variants is conferred via their access to replication forks (Shurtleff 2009). The Y11A variant was purified, and in the presence of co-factors SSB, RecA*, and the beta and gamma complex, this pol V variant enhanced the incorporation of ribonucleotides into DNA as expected, while the F10L variant increases sugar selectivity and fidelity (Vaisman 2012). It is important to note that under these conditions, pol V wild-type also incorporates ribonucleotides readily into DNA, however, Y11A incorporated ribonucleotides three times more efficiently than wild-type pol V (Vaisman 2012). Molecular modeling suggests that the F10L variant constricts the movement of neighboring Y11, closing the steric gate (Vaisman 2012). Recent in vivo studies suggest that because Y11A efficiently incorporates rNTPs into DNA, that the observed phenotypes of conferring UV sensitivity and low UV mutation frequency could be due to repair mechanisms in vivo, such as removal by Ribonuclease HII of ribonucleotide monophosphates from the newly generated DNA strand or by removal by nucleotide excision repair (NER) of UV-induced lesions from the template strand of DNA being copied, leaving double strand breaks that reduce cell viability (McDonald 2012). In the absence of either pathway, resistance to UV radiation and UV mutagenesis conferred by UmuC variant Y11A are significantly
increased (McDonald 2012). Also, pol IV inhibits pol V bypass of UV lesions, therefore reducing the access of pol V to the replication fork, and reducing the number of rNTPs incorporated by the Y11A variant, minimizing the incorporation of rNTPs (McDonald 2012).

Cells harboring the UmuC variant A39V were very sensitive to UV radiation and had decreased UV-induced mutagenesis (Figure 1.4B). Also, the A39V mutation fails to complement the $umuDC$-dependent cold sensitivity phenotype (Marsh 1991). Even though pol V contributes substantially to UV-induced mutagenesis, it contributes only modestly to survival after exposure to UV (Friedberg 2006). However, mutations at F10, Y11, and A39 are examples of point mutations in UmuC that confer dramatically increased sensitivity to UV on cells that harbor them. Similarly, A9V, discussed in Chapter 4 of this dissertation, is a mutation discovered from a random mutagenesis screen. A9V confers resistance to nitrofurazone (NFZ), which is thought to cause predominantly $N^2$-furfuryl-dG lesions, a cognate lesion of DinB (Jarosz 2006). A9 is predicted to be located in the active site adjacent to the incoming nucleotide, near the steric gate (Figure 1.4B). A9V has very low mutation frequency depending on the length of the UmuD’ arm. Synthetically cleaved UmuD’ has one extra residue for expression, and when the length of the arm is shortened to the naturally occurring length, the mutation frequency of A9V is restored to wild-type levels. This phenomenon does not affect UV survival, however.

Several residues, later defined to be part of UmuC active site loop 1, $^{31}$SNNDGCVI$^{38}$, were proposed to be important for UmuC activity based on analysis of the active site region thought to accommodate bulky lesions through analysis of
molecular dynamics runs using CHARMM 30 (Figure 1.4C) (Lee 2005; Brooks 2009; Chandani 2009; Seo 2009; Hawver 2011). Residues 31SNN33 are metaphorically referred to as the “flue” of a “chimney,” meaning these residues perhaps control the size of the active site opening of UmuC, which in turn would control the size of the adducts that can be bypassed (Figure 1.4C) (Chandani 2009). S31 was shown experimentally to be important for UV-induced mutagenesis, while I38 and A39 contribute to bypass of benzo[a]pyrene (Beuning 2009; Seo 2009). N32 theoretically plugs the chimney hole, controlling the size of the opening (Figure 1.4C) (Chandani 2009). V29, L30, V37, and I38 are thought to anchor an active site loop, referred to as loop 1, of UmuC (Chandani 2009; Hawver 2011). Furthermore, I38 is identified as the “roof amino acid” which influences dNTP insertion due to its location directly adjacent to the incoming nucleotide. L30 is a “flue handle” that controls the flue opening of the chimney (Chandani 2009; Seo 2009). In a later in-depth characterization, described in Chapter 2 of this dissertation, it was found that residues 32 through 34 of UmuC active site loop 1 are important for UV survival and UV-induced mutagenesis (Hawver 2011). When these three residues are mutated to alanine, cells harboring them confer hypersensitivity to UV radiation and exhibit extremely low mutation frequencies (Hawver 2011). UmuC variants N32A, N33A, and D34A also diminish pol V-dependent inhibition of RecA-mediated homologous recombination, suggesting that these three residues could be involved in an interaction between UmuC and RecA (Hawver 2011). UmuC active site loop 2, 50KMGDP54, in an alanine-scanning mutagenesis screen, was found to not have an impact on UV survival or UV-induced mutagenesis (Figure 1.4C) (Hawver 2011).
Strains harboring the \textit{umuC36} allele (E75K) were rendered nonmutable (Koch 1992). E75 is potentially a site of interaction between UmuC and UmuD', because in strains with UmuD' present at elevated levels, the non-mutable phenotype was suppressed (Figure 1.4D) (Bates 1991; Koch 1992). UmuC variants H282P and T412I, discussed in Chapter 3 of this dissertation, were found in a hydroxylamine mutagenesis selection to confer resistance to NFZ. These variants in UmuC caused UmuD to hypercleave to UmuD', suggesting they also may play a role in the interaction between UmuC and UmuD'. UmuC H282P and T412I are also readily degraded by Lon protease. The location of H282 is predicted to be in close proximity to E75 (Figure 1.4D). The T290K mutation (\textit{umuC25}) causes strains harboring it to be non-mutable (Koch 1992). By analogy to Dpo4, T290 is predicted to be in the hydrophobic core of the little finger and may contribute to protein stability (Figure 1.4D) (Boudsocq 2002).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure15.png}
\caption{Model of UmuC, the DNA polymerase subunit of UmuD',C (Beuning 2006). Highlighted residues are thought to be important for interaction with UmuD': E75 (orange) (Bates 1991; Koch 1992); T290 (blue) is in the hydrophobic core of the little finger and may also contribute to its stability (Koch 1992); H282P (green) discussed in Chapter 3 of this dissertation causes UmuD to become hypercleavable; E90, N82, 126TVLQRTH132 (magenta) and E89, P93, R94, E239 (black) have been found in a molecular modeling study to interact with UmuD' (Chandani 2013).}
\end{figure}
In a recent docking study, predictions were made for the sites of interaction between UmuC and its binding partners UmuD’ and RecA using ZDOCK and ClusPro (Chen 2003; Comeau 2004; Chandani 2013). The proteins being docked are treated as rigid, native structures, and results are evaluated by an energy scoring function that considers electrostatics, desolvation, and change in shape (Chen 2003; Comeau 2004; Chandani 2013). Two sites on UmuC were predicted to interact with UmuD’ (Chandani 2013). The UmuD’ closest to RecA is predicted to interact with UmuC residues 82, 90, and 126-132, while the UmuD’ further from RecA is predicted to interact with UmuD residues 89, 93, 94, and 239 (Figure 1.5) (Chandani 2013). Similarly, two regions of UmuC were proposed as binding sites for RecA, which are UmuC residues 32 through 34, and also residues 59 through 64 (Figure 1.4D) (Chandani 2013). Residue K59 is predicted to form hydrogen bonds with RecA residues E96, A147, A148; residue R63 is predicted to form hydrogen bonds with RecA residues P119, D120; and residue R64 is predicted to form hydrogen bonds with RecA residues S117, D120, and N124 (Chandani 2013). UmuC residues N32, N33, and D34, discussed above and also in Chapter 2 of this dissertation, have also been predicted to interact with RecA in a molecular modeling study (Chandani 2013). UmuC residue N32 is predicted to interact with RecA residue E68; N33 is predicted to interact with RecA residues E96 and Q194; and UmuC residue D34 is predicted to interact with RecA residue R196 (Chandani 2013). Interestingly, RecA residue E96 is predicted to act as a bridge between the two potential binding sites, as it interacts with UmuC residue K59 and also N33 (Chandani 2013). Both interaction sites are predicted to be located on the surface of the UmuC protein, making them likely candidates for protein-protein interactions (Figure 1.4D) (Lee 2005; Chandani 2013).
Other UmuC residues predicted to bind to RecA include K25 with RecA D100, C65 with RecA Q118, G66 with RecA D100, W55 with RecA E96, and F56 with RecA A148 and R196 (Chandani 2013).

The model of UmuC shown in Figure 4 is a truncation, with the last residue in the model indicated at Q353, as the C-terminal domain of UmuC lacks homology to proteins of known structure (Figure 1.4E). The UmuC C-terminal domain is important for UV-induced mutagenesis as well as for interactions with UmuD and UmuD’ (Sutton 2001). The UmuC beta-binding motif \(^{357}QLNL^{361}\) would be located just C-terminal to the end of the model, Q353 (Dalrymple 2001; Becherel 2002; Beuning 2006). Mutations in this motif cause almost a complete loss in UmuC-dependent UV-induced mutagenesis (Becherel 2002; Beuning 2006). The second area of interaction between UmuC and the beta clamp, \(^{313}LTP^{315}\) (Figure 1.4E), did not cause a loss in UV-induced mutagenesis when mutated. However, the UmuC-dependent cold-sensitivity phenotype was suppressed when either of these sites was mutated (Beuning 2006).

1.8 Cellular Interactions of UmuC

Initial characterization through a series of co-immunoprecipitation experiments showed that UmuC interacts with UmuD\(_2\) and interacts more strongly with UmuD’\(_2\) (Woodgate 1989). UmuD’\(_2\) is the form of the \(umuD\) gene products that is specifically required to activate UmuC for mutagenesis (Nohmi 1988). Moreover, induction of \(umuDC\) is all that is required for SOS mutagenesis (Sommer 1993b). It was not until approximately twenty years after the discovery of the \(umuDC\) genes and their roles in SOS mutagenesis that UmuD’\(_2\)C was found to be a DNA polymerase capable of copying damaged DNA (Reuven 1999; Tang 1999).
Table 1.6. A comparison of biochemical requirements for TLS by Pol V. (A) UmuC was purified by using a maltose binding protein tag (Reuven 1999). The DNA substrate used was a circular plasmid with approximately 339 nt of ssDNA. It was found that with SSB, ATP, and a RecA/ssDNA nucleoprotein filament, the beta clamp stimulates nucleotide incorporation by three- to five-fold (Reuven 1999; Reuven 2001; Maor-Shoshani 2002; Arad 2008). (B) The DNA substrate used was a single strand of linear DNA with the lesion 50-nt away from the 5′ end of the DNA. SSB and the beta clamp were required when ATPγS was present. A RecA nucleoprotein filament in trans or only a single RecA monomer is needed for Pol V activation (Tang 1999; Tang 2000; Pham 2001; Pham 2002; Schlacher 2005; Schlacher 2006a; Schlacher 2006b; Schlacher 2006c; Jiang 2009). (C) A modification of the Goodman and Woodgate requirements from section B: the circular DNA substrate used was approximately 3000-nt long. SSB enhances TLS in the presence of trans RecA nucleoprotein filament; the beta clamp stimulates TLS by eight to 33-fold. The RecA nucleoprotein filament is in trans and is formed with ATPγS prior to addition to reactions (Karata 2012; Kuban 2012; Vaisman 2012). (D) The circular DNA substrate used was approximately 2700-nt long. SSB was not required when ATP was present; the beta clamp stimulates TLS by 100-fold. The RecA nucleoprotein filament is in cis (Fujii 2004a; Fujii 2004b; Fujii 2009). The processivity of Pol V is increased three- to five-fold when ATP is replaced with ATPγS (Pham 2001).
Several additional proteins facilitate TLS by UmuD′2C. These proteins include RecA, the beta clamp and gamma clamp loader, and SSB. Activated RecA is strictly required for pol V-dependent TLS, although the exact mechanism by which activated RecA facilitates TLS is still in question (see below). The use of different experimental systems by three separate groups to study the biochemical properties of pol V has contributed to different conclusions about the roles and requirements of these accessory proteins (summarized in Figure 1.6). First, different forms of UmuC have been used, with one group using a maltose binding protein tag (MBP) to purify UmuC (Reuven 1999) while others purified a native UmuC-UmuD′ complex (Tang 1999; Fujii 2004b; Karata 2012). Another major experimental difference among the three groups is the DNA substrate used (Figure 1.6). The Livneh group used a gapped plasmid with an ssDNA region of approximately 339 nucleotides (Reuven 1999; Reuven 2001; Maor-Shoshani 2002; Arad 2008). The Goodman and Woodgate groups originally used linear ssDNA with the lesion located 50 nucleotides from the 5′ end (Tang 1999; Tang 2000; Pham 2001); in their more recent work, the length of the DNA varies (Pham 2002; Schlacher 2005; Schlacher 2006a; Jiang 2009). The Goodman and Woodgate collaboration recently discovered a simple purification method for pol V, and in these assays, they utilized a single-stranded circular DNA template that is 3,000 nucleotides long (Karata 2012; Kuban 2012; McDonald 2012; Vaisman 2012). Lastly, the Fuchs group also used circular ssDNA that is about 2700 nucleotides in length (Fujii 2004a; Fujii 2004b; Fujii 2009). These differences in experimental design possibly set the stage for the discrepancies in biochemical requirements for pol V as summarized in Figure 1.6 and discussed below.
RecA is a 38-kDa protein that is the product of the recA gene (Clark 1965). In addition to the significant roles RecA is known to play in responses to DNA damage, including homologous recombination, induction of the SOS response by serving as a co-protease in the autoproteolytic cleavage of LexA, and the regulation of SOS mutagenesis by facilitating cleavage of UmuD\(_2\) to UmuD\(_2^\prime\), RecA also has a direct role in TLS (Schlacher 2006a; Patel 2010). Furthermore, it was proposed that RecA has two distinct roles in pol V-mediated TLS (Pham 2002). First, RecA at the 3′ end of a primer stimulates pol V for TLS activity. Second, RecA bound to the template strand mediates extension past the lesion (Pham 2002). The RecA nucleoprotein filament may provide an “activated” RecA monomer for TLS, but the filament itself may or may not participate in the actual TLS reaction. A six-nucleotide overhang can only bind two RecA monomers, and TLS still occurs, an observation which argues against the need for a RecA filament (Pham 2002). Bypass of an abasic site in a three-nucleotide gap in the DNA to which RecA can still bind was successful as well; however, bypass of a lesion in a two-nucleotide gap was not (Pham 2002).

One viewpoint is that a “minimal mutasome,” which includes pol V, ATP, and two RecA molecules, one bound to UmuC and one bound to UmuD\(_2^\prime\), is the active form in TLS (Schlacher 2005; Sweasy 2005). Furthermore, in this model, RecA acts in \textit{trans}, such that the RecA/ssDNA filament is formed on a DNA strand not actively being copied by pol V (Schlacher 2006a). In other words, the \textit{trans} RecA/ssDNA transfers a RecA monomer as well as a molecule of ATP from its 3′ end to a molecule of pol V, thereby activating pol V for TLS (Jiang 2009). The newly activated pol V then performs a single round of TLS, and upon dissociation from the DNA is inactivated and must once again be
activated by RecA and ATP (Jiang 2009). On the other hand, there is also evidence that the RecA/ssDNA nucleoprotein filament primarily acts in cis for TLS by forming a filament on the single-stranded DNA downstream from the lesion on the primer strand, thus facilitating TLS by pol V (Fujii 2009). It has been shown that of the co-factors that pol V has been known to utilize, that the addition of the trans RecA nucleoprotein filament is a major contributor to the processivity of pol V (Karata 2012).

The beta clamp significantly increases the processivity of the replicative DNA polymerase pol III (Kornberg 1992). This enhancement of processivity extends to other polymerases as well, including pol V. The beta clamp and gamma clamp loader provide additional stability to the complex and may help pol V remain tethered to DNA (Pham 2001). However, the beta clamp only increases processivity of pol V modestly, with determinations ranging from three- to five-fold to ~100-fold (Maor-Shoshani 2002; Fujii 2004b; Karata 2012). All three groups determined that the beta clamp stimulates Pol V, but to varying extents and with different requirements for co-factors (Figure 1.6) (Tang 1998; Reuven 1999; Tang 1999; Tang 2000; Pham 2001; Maor-Shoshani 2002; Fujii 2004a; Fujii 2004b; Karata 2012). The Livneh and Fuchs groups determined that with native ATP, the presence of the beta clamp increased processivity of Pol V (Maor-Shoshani 2002; Fujii 2004b), whereas the Goodman/Woodgate group observed a three- to five-fold increase in processivity with the beta clamp present but only with ATP-gamma-S and SSB also present (Pham 2001). The Goodman/Woodgate group later saw an increase of eight to 33-fold in processivity of pol V in the presence of trans RecA nucleoprotein filament, SSB, and the beta clamp on circular ssDNA (Karata 2012). The Fuchs group suggested that SSB must be present when ATP-gamma-S is used, perhaps to
destabilize the highly stabilized RecA filament that is formed in the presence of ATP-

gamma-S (Fuchs 2004).

A direct physical interaction has been detected between UmuC and single-
stranded DNA binding protein (SSB) (Arad 2008). SSB coats single-stranded DNA and
helps to prevent dissociation of RecA from the ssDNA formed after a replicative
polymerase stalls at a lesion (Cox 1983). SSB also stimulates the formation of the RecA
filament by over 50-fold (Flory 1982). Lesion bypass using MBP-UmuC (Reuven 1999)
was at its most efficient when the concentration of SSB in the reaction was 50 nM and
ATP was present (Reuven 2001). Using a purified UmuD′2C complex, it was determined
that lesion bypass was optimal at 60 nM SSB and that SSB stimulated TLS by 1,040-fold
when ATP-gamma-S was present (Pham 2001). Similarly, SSB likely helps to form the
RecA nucleoprotein filament in the presence of ATP (Reuven 2001). However, in the
presence of ATPγS, SSB may hinder the formation of the RecA-nucleoprotein filament
when the filament is formed in the presence of SSB, by blocking access of RecA to the
DNA (Karata 2012). In a lesion bypass experiment in the absence of SSB, in which a
functional RecA/ssDNA nucleoprotein filament was formed, no bypass was observed,
leading to the conclusion that SSB may have a second function in TLS besides
stimulation of RecA nucleoprotein filament formation (Reuven 2001). On the other hand,
it has been suggested that SSB is not absolutely required for Pol V TLS in the presence of
ATP, but SSB may be needed to disrupt the more stable RecA filament formed in the
presence of ATPγS (Fujii 2004b). Furthermore, it is thought that SSB may play a direct
role in recruiting pol V to the 3′ end of the primer coated with RecA via a direct
interaction with UmuC (Arad 2008), as SSB interacts with pol V at the C-terminus of
SSB (Arad 2008; Karata 2012). It is important again to consider the differences in the form of pol V, as well as the DNA substrate and the method of formation of the RecA nucleoprotein filament, used for these experiments, which is summarized in Figure 1.6.

1.9 Specificity and Regulation of Polymerase IV, DinB

DinB is the other Y family polymerase found in *E. coli* (Wagner 1999; Ohmori 2001). The *dinB* (*damage-inducible*) gene was identified as being induced upon treatment with DNA damaging agents (Kenyon 1980; Ohmori 1995). Subsequently, the *dinB* gene product was demonstrated to be a DNA polymerase at around the same time as pol V(Reuven 1999; Tang 1999; Wagner 1999).

![Figure 1.7. Adducts of deoxyguanine bypassed by *E. coli* pol IV:](image)

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Pol IV inserts</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N^2)-ff-dG</td>
<td>C</td>
</tr>
<tr>
<td>(N^2)-B[a]P-dG</td>
<td>C</td>
</tr>
<tr>
<td>γHOP-dG</td>
<td>C (A/T)</td>
</tr>
<tr>
<td>(N^2)-CE-dG</td>
<td>C</td>
</tr>
</tbody>
</table>

*Figure 1.7. Adducts of deoxyguanine bypassed by *E. coli* pol IV: \(N^2\)-dG-\(γ\)-hydroxypropano-dG (Minko 2008), \(N^2\)-(1-carboxylethyl)-2′-dG (\(N^2\)-CE-dG) (Yuan 2008), \(N^2\)-furfuryl-dG (Jarosz 2006) \(N^2\)-+[trans, anti]-benzo[a]pyrrol-dG (Shen 2002; Fuchs 2004; Seo 2006). Unmodified dG is shown for comparison. The table shows the identity of the nucleotide inserted by pol IV opposite each lesion.*
Much like pol V, DinB has a preference for certain DNA lesions. Most notably, DinB possesses a 15-fold preference to insert C opposite $N^2$-furfuryl-dG and a 25-fold preference to extend from a base pair with $N^2$-furfuryl-dG in comparison to undamaged dG, meaning DinB is more active on DNA containing this lesion than it is on undamaged DNA, supporting the idea that this lesion is the cognate lesion of DinB (Figure 1.7) (Jarosz 2006; Jarosz 2009). The external source of $N^2$-furfuryl-dG has not yet been found, however, by analogy to the formation of kinetin ($N^6$-furfuryl-dA), it could be a result of ribose oxidation (Barciszewski 1999). Moreover, strains in which $dinB$ has been deleted are sensitive to nitrofurazone and 4-nitroquinoline-1-oxide (4NQO), both of which are thought to form $N^2$-dG adducts (Friedberg 2006; Jarosz 2006).

Continuing with the trend of accurate bypass of $N^2$-dG adducts, DinB also efficiently and accurately bypasses adducts that are known to be biologically important. DinB bypasses $N^2$-(1-carboxyethyl)-2'-deoxyguanine ($N^2$-CEdG), which was detected in 1 in $10^7$ bases in melanoma cells, and is formed as an adduct of methylglyoxal, a byproduct of glycolysis (Figure 1.7) (Yuan 2008). DinB also bypasses the $N^2$-dG adduct of benzo[a]pyrene (B[a]P), a carcinogen found in coal tars and cigarette smoke (Figure 1.7) (Phillips 1983; Lenne-Samuel 2000; Shen 2002; Fuchs 2004; Seo 2006). However, efficient bypass of some isomers of B[a]P also requires pol V, suggesting even further that polymerases possess specificity for certain lesions (Seo 2006). DinB also accurately inserts dCTP across from $\gamma$-hydroxypropano-deoxyguanosine ($\gamma$-HOPdG), the $N^2$-dG adduct of acrolein, which is a toxin and is a byproduct of fatty acid metabolism (Figure 1.7) (Minko 2008).
The presence of accessory proteins, specifically the beta processivity clamp and the gamma clamp loader, greatly increase the efficiency of bypass (Maor-Shoshani 2002) and decrease frameshift mutagenesis. While there are several mechanisms for frameshift mutations, the likely scenario of how frameshift mutations occur with DinB is thought to be a mix of Streisinger slippage and dNTP-stabilized misalignment (Tippin 2004). A nucleotide downstream from the primer terminus is flipped out of the DNA helix, and is therefore skipped over by DinB, which in turn generates a -1 frameshift mutation, as the flipped nucleotide is not copied by DinB, while dNTP-stabilized misalignment allows for a correct nucleotide to be added to the primer across from the next template base, which is adjacent to the flipped nucleotide (Streisinger 1966; Kobayashi 2002; Tippin 2004; Foti 2010). However, addition of RecA and UmuD2 alter the -1 frameshift mutator activity of DinB, which is a result of the elevated levels of DinB in a cell such that there is insufficient UmuD2 present to regulate DinB (Godoy 2007). Co-upregulation of both UmuD2 and DinB suppresses the -1 frameshift mutation activity (Godoy 2007; Foti 2010). Molecular modeling offers the explanation that RecA and UmuD2 may suppress the -1 frameshift activity of DinB by closing in the DinB active site (Godoy 2007). Deletion of *umuD* did not affect DinB-dependent resistance to nitrofurazone, suggesting that the -1 frameshift mutator activity may not be related to DinB TLS function (Godoy 2007). DinB residue F172 has been shown to interact with UmuD (Godoy 2007). DinB interacts with the beta processivity clamp, with similar sites of interaction as UmuC (Dalrymple 2001; Becherel 2002; Bunting 2003; Beuning 2006). The two sites of interaction on DinB for the beta clamp are \(^{346}\text{QLVLGL}^{351}\) (Lenne-Samuel 2000; Becherel 2002), which is at the C-terminus of DinB, and \(^{303}\text{VWP}^{305}\) (Bunting 2003).
1.10 Polymerase Switching

There are a variety of models used to describe how the multiple DNA polymerases in *E. coli* are utilized. The beta clamp facilitates polymerase switching that must take place in order for pol V to replace pol III on the DNA template (Burnouf 2004; Fujii 2004a; Fujii 2004b; Fujii 2009). In *E. coli*, it was observed that the beta clamp is donated by pol III to recruit pol V to the DNA template near the lesion (Fujii 2004a). The “dynamic processivity” model was suggested from experiments using a catalytically inactive mutant of T4 gp43 that was exchanged in less than a minute to replace the replicative polymerase in T4 (Yang 2004). Thus, this model suggests that DNA polymerases may exchange in a stochastic and rapid manner. The observation that elevated levels of DinB inhibit pol III suggests a model of rapid replacement of pols at the replication fork in *E. coli* as well (Furukohri 2008; Uchida 2008). Another model of polymerase switching is the “tool belt” model, which holds that multiple DNA polymerases are tethered to the beta clamp at the same time, allowing several DNA polymerases to be present at the replication fork and used when needed (Pages 2002; Friedberg 2006). However, the observation that only one binding site on the beta clamp is used to facilitate a switch between pol III and pol IV is inconsistent with the tool-belt model (Heltzel 2009). Finally, the gap filling model suggests that gaps are left at lesion sites when a replicative polymerase cannot bypass the lesion. Replication is initiated downstream of the lesion and the resulting gap is subsequently filled in by Y family polymerases performing TLS (Pages 2003; Heller 2006; Lehmann 2006; Lopes 2006). These models are not mutually exclusive.
Although pol V is very poorly processive depending on the accessory co-factors used, the processivity offered by the beta clamp to pol V allows the polymerase to synthesize an appropriate-length “TLS patch.” This patch of at least six nucleotides allows pol V to bypass the lesion and extend past it far enough for pol III to resume DNA synthesis without triggering proofreading at the newly bypassed lesion (Fujii 2004a). Pol III can detect distortions in DNA caused by inserting a nucleotide across a lesion even when pol III is recruited back to the replication fork four to five nucleotides after the lesion so a TLS polymerase must extend at least beyond that point.

Other work suggests that the situation is more complicated than a simple exchange between two DNA polymerases, as apparently several polymerases can compete or cooperate to bypass specific lesions (Napolitano 2000; Delmas 2006; Fuchs 2007; Neeley 2007; Courcelle 2009; Curti 2009). The presence of pol III, pol II, pol IV and pol V influence one another’s ability to access the primer terminus of the replication fork (Sutton 2006b). The two Y family DNA polymerases contribute to bypass of benzo[a]pyrene, for example (Lenne-Samuel 2000; Seo 2006). Additionally, using the characteristic mutagenic signature of each DNA polymerase, it was found that the polymerases in *E. coli* compete for access to DNA and that both pol IV and pol V contribute to spontaneous mutagenesis (Curti 2009).

1.11 Summary

Y family DNA polymerases are important for tolerating DNA damage. These specialized polymerases are regulated in different ways, and they possess specificity for certain DNA lesions. *E. coli* pol V is specific for UV-induced lesions, such as thymine-
thymine cyclobutane pyrimidine dimers and 6-4 thymine-thymine photoproducts, while DinB accurately bypasses \(^{N2}\)-dG lesions, such as \(^{N2}\)-fururyl-dG.

Pol V is regulated very tightly by the cleavage of UmuD to UmuD\(^{\prime}\), as pol V is not active for TLS until UmuC is paired with UmuD\(^{\prime}\)C. Also, it has been shown that several other co-factors, such as the beta clamp, gamma clamp loader, SSB, and the RecA-nucleoprotein filament, are important for pol V TLS under different conditions. Knowing which residues of a polymerase interact with which binding partners could be important for developing new approaches to controlling or changing the regulation and specificity of these polymerases. Y family polymerases play an integral role in antibiotic resistance and cancer, and discovering how these polymerases function is important for future directions in these important areas of medicine.

1.12 References


This chapter is reproduced with permission from the following


Authors in ASM journals retain the right to republish discrete portions of his/her article in any other publication (including print, CD-ROM, and other electronic formats) of which he or she is author or editor, provided that proper credit is given to the original ASM publication. ASM authors also retain the right to reuse the full article in his/her dissertation or thesis. For a full list of author rights, please see: http://journals.asm.org/site/misc/ASM_Author_Statement.xhtml

Penny J. Beuning conceived this study. Caitlin Gillooly created variants in plasmids and performed some of the mutagenesis experiments for the original alanine-scanning mutagenesis of loops 1 and 2. Lisa A. Hawver collected all other experimental data and prepared the manuscript draft, tables and figures.
Chapter 2: Characterization of *Escherichia coli* UmuC Active-Site Loops Identifies Variants that Confer UV Hypersensitivity

2.1 Introduction

*E. coli* has five DNA polymerases that replicate DNA under different circumstances (Friedberg 2006). The replicative polymerase in *E. coli* is DNA polymerase III, a member of the C family. DNA polymerases IV and V are members of the Y family, which specialize in copying damaged DNA in a process known as translesion synthesis (TLS) (Ohmori 2001; Friedberg 2006). Y family DNA polymerases also copy undamaged DNA in an error-prone manner, possibly subjecting DNA to untargeted mutagenesis and potentially leading to antibiotic resistance or cancer (Pages 2002; Cirz 2005; Friedberg 2006; Cirz 2007).

*E. coli* DNA polymerases IV (DinB) and V (UmuD′C) are the products of the *dinB* and *umuDC* genes, respectively. Due to their potentially mutagenic nature, these proteins are highly regulated in a specific cellular response to DNA damage and other stresses called the SOS response (Radman 1974; Friedberg 2006). The SOS response is initiated when single-stranded DNA (ssDNA) forms downstream from a lesion in DNA due to the inability of the replicative DNA polymerase to copy damaged DNA. RecA then coats the ssDNA to form a RecA/ssDNA nucleoprotein filament, which is the inducing signal for the SOS response. The RecA/ssDNA filament facilitates the autocleavage of LexA, the repressor of the SOS genes, and allows for the expression of at least 57 genes (Friedberg 2006; Simmons 2008). While many genes are induced in the SOS response, expression of only the *umuDC* genes is required for SOS mutagenesis (Sommer 1993b).
The *umuDC* genes in *E. coli* encode UmuD, a polymerase manager protein, and UmuC, the DNA polymerase subunit. The RecA/ssDNA nucleoprotein filament plays another role in regulation of SOS mutagenesis by facilitating the cleavage of the UmuD protein to form UmuD′ (Burckhardt 1988; Nohmi 1988; Shinagawa 1988). Approximately 20 to 40 min after SOS induction, UmuD cleaves its N-terminal arms between C24 and G25 to form UmuD′2, the form that interacts with UmuC to form pol V (UmuD′2C), which is active in SOS mutagenesis (Woodgate 1989; Reuven 1999; Tang 1999). UmuD′2 is required for UmuC to be active in translesion synthesis (Reuven 1999; Tang 1999).

Other protein interactions have also been found to be important for pol V activity. The β clamp substantially increases processivity in both pol III and pol IV in *E. coli* but only increases the processivity of pol V three- to five-fold (Kornberg 1992; Wagner 2000; Maor-Shoshani 2002; Fujii 2004b). The β clamp interacts directly with UmuC, utilizing a canonical (Dalrymple 2001; Becherel 2002; Lopez de Saro 2003; Sutton 2005; Beuning 2006a) β binding motif, an interaction that is critical for UmuC to participate in TLS (Becherel 2002; Sutton 2005; Beuning 2006a).

In addition to the roles of RecA in the induction of the SOS response and the cleavage of UmuD, RecA also plays a role in the activation of pol V for TLS (Reuven 2001; Schlacher 2005; Schlacher 2006; Fujii 2009; Jiang 2009). While the exact mechanism of pol V stimulation by RecA remains to be determined, it is understood that RecA is required for pol V mutagenesis. It is thought that pol V has a preference for binding to the end of RecA nucleoprotein filaments in order to target the pol V complex to DNA lesions, as well as binding deep in the groove of the RecA/ssDNA filament (Frank 2000).
Pol V bypasses the major lesions produced by UV radiation, which are thymine-thymine cyclobutane pyrimidine dimers (T-T CPD) and thymine-thymine (6-4) photoproducts (Tang 2000). Pol V generally accurately bypasses T-T CPDs by inserting two dA nucleotides opposite the lesion (Szekeres 1996; Tang 2000). Pol V bypasses T-T (6-4) photoproducts inaccurately by adding dA across from the 5′ T and dG across from the 3′ T (Becherel 1999; Tang 2000). Pol V also bypasses abasic sites by inserting dA, following the “A” rule (Lawrence 1990; Strauss 1991; Reuven 1999; Tang 1999; Tang 2000). Opposite C\(^8\)-dG-acetyaminofluorine, pol V inserts dA (Becherel 1999; Fujii 2004a; Fujii 2009); pol V bypasses N\(^2\)-benzo[a]pyrene-dG with relative inaccuracy, while it bypasses N\(^6\)-benzo[a]pyrene-dA fairly accurately (Lenne-Samuel 2000; Shen 2002; Yin 2004).

In this work, we sought to identify residues of UmuC that contribute to lesion bypass, and more specifically, to UV-induced mutagenesis. We hypothesized that two loops of UmuC interact with template DNA as well as with the incoming nucleotide and therefore are likely to contribute to mutagenesis. We show here that the N-terminal residues of loop 1, 31-34, play a significant role in cell survival in response to UV radiation. We also show that loop 2 does not play as significant a role in conferring survival or in UV-induced mutagenesis. Furthermore, we show that the mutation of each of the three UmuC loop 1 residues 32-34 to alanine results in less inhibition of recombination than wild-type UmuC, which could indicate the disruption of an interaction between pol V and RecA and therefore could explain the ability of these variants to confer UV hypersensitivity.
2.2 Materials and Methods

2.2.1 Strains and Plasmids

Low-copy plasmid pGY9738 carries the synthetic $umuD' C$ operon and encodes resistance to spectinomycin (60 μg/mL). Strains used (Table 2.1) were grown in Luria Broth at 37 °C unless otherwise noted. Competent cells were made by using the CaCl$_2$ method (Sambrook 1989).

Table 2.1 Strains and plasmids

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant genotype or description</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial Strain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>$argE3 umuDC^+$</td>
<td>Laboratory stock</td>
</tr>
<tr>
<td>GW8017</td>
<td>AB1157 $\Delta umuDC$</td>
<td>(Guzzo 1996)</td>
</tr>
<tr>
<td>PB102</td>
<td>AB1157 $\Delta umuC \Delta recJ$</td>
<td>P1(JW2860) → AB1157</td>
</tr>
<tr>
<td></td>
<td>$\Delta umuC$ (Baba 2006)</td>
<td></td>
</tr>
<tr>
<td>AB1157 $\Delta umuC$</td>
<td></td>
<td>(Jarosz 2006)</td>
</tr>
<tr>
<td>GW2771</td>
<td>$umuDC^+$</td>
<td>(Opperman 1996)</td>
</tr>
<tr>
<td>GW2771 $spq-2$</td>
<td>GW2771 $spq-2$</td>
<td>(Opperman 1996)</td>
</tr>
<tr>
<td>GW2771 $spq-2$</td>
<td>$dnaQ903::tet$</td>
<td>(Slater 1994; Shurtleff 2009)</td>
</tr>
<tr>
<td>$dnaQ903$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector; pSC101 derived, Spec$^r$</td>
<td>(Churchward 1984)</td>
</tr>
<tr>
<td>pGY9738</td>
<td>$o_1^C umuD' C$; pSC101 derived</td>
<td>(Sommer 1998)</td>
</tr>
</tbody>
</table>
Transformations were performed as described previously (Beuning 2006b). Variants of pGY9738 were made with Quikchange or Quikchange Lightning site-directed mutagenesis kits (Agilent Technologies). Beta binding sites are defined as $\beta_2$ ($^{313}\text{LTP}^{315}$) and $\beta_1$ ($^{357}\text{QLNF}^{361}$). Constructs with mutated beta binding sites are designated $\beta_1+2$, which indicates the presence of the mutations $^{313}\text{LTP}^{315} \to ^{313}\text{AAA}^{315}$ and $^{357}\text{QLNF}^{361} \to ^{357}\text{AAAAA}^{361}$ (Beuning 2006a). The presence of the mutations was confirmed by DNA sequencing (Massachusetts General Hospital DNA Core Facility, Cambridge, MA).

2.2.2 UV Survival and Mutagenesis Assays

UV survival and mutagenesis assays were performed as described previously (Beuning 2006b; Shurtleff 2009). Variants were exposed to 25 J/m$^2$ 254 nm UV radiation unless otherwise noted for mutagenesis assays. Each point shown represents the average of at least three trials and the error bars indicate the standard deviation.

2.2.3 Nitrofurazone (NFZ) and 4-Nitroquinoline-N-oxide (4-NQO) Survival Assays

Strains harboring alanine variants of the N-terminal loop 1 residues (32-34) were exposed to increasing concentrations of NFZ (5-nitro-2-furaldehyde semicarbazone, TCI America) and 4-NQO (Acros Organics), as described previously (Beuning 2006b; Shurtleff 2009). Stock solutions (10 mg/mL) were freshly prepared in $N,N$-dimethyl formamide (Fisher Scientific) and protected from light. Serial dilutions of overnight cultures were plated on LB-agar plates containing 60 $\mu$g/mL spectinomycin and the indicated amount of either NFZ or 4-NQO and were incubated at 37 °C for 20 to 24 h.
Each point shown represents the average of at least three trials and the error bars indicate the standard deviation.

### 2.2.4 Immunoblotting

Western blots were performed as described previously (Shurtleff 2009). Strains harboring variants of the N-terminal loop 1 residues 31-34 (alanine or conservative mutations) were exposed to 10 J/m² UV radiation, while all others were exposed to 25 J/m² UV radiation. Proteins were resolved by 14% SDS-PAGE and transferred to PVDF membrane. The membrane was blocked overnight in 5% milk in 1X TBS-Tween (20 mM Tris-HCl, pH 7.6, and 137 mM NaCl, 1% Tween 20). The membrane was then probed with anti-UmuC (Beuning 2006a) in 2.5% milk with 0.5X TBS-Tween and washed for 2 min and then 3x10 min each with 1X TBS-Tween buffer. The membrane was probed with horseradish peroxidase-conjugated goat anti-rabbit antibody (Pierce) in 2.5% milk with 0.5X TBS-Tween, developed with SuperSignal chemiluminescence reagent (Pierce), and exposed to film, which was subsequently developed with a Kodak photoprocessor.

### 2.2.5 Genetic Transduction

Strains used (Table 2.1) were grown in Luria Broth at 37 °C overnight. Cultures were centrifuged at 1500 x g for 10 min and resuspended in 2.5 mL of a solution containing 10 mM MgSO₄ and 5 mM CaCl₂ and transferred to glass test tubes. Recipient cells (100 µL) harboring each variant were incubated at 30°C for 30 min with 50 µL P1vir ΔyeaB (KanR). As controls, 100 µL of recipient cells were incubated without P1vir ΔyeaB and 100 µL of P1vir ΔyeaB were incubated without recipient cells. After incubation, 100 µL of sodium citrate and 400 µL of Luria Broth were added to each test
tube and cells recovered for 2 h at 30°C with no shaking. Reactions were centrifuged at 3800 x g for 5 min and resuspended in 200 µL of Luria Broth. A 100-µL aliquot of each reaction was plated on selective medium (spectinomycin, 60 µg/mL and kanamycin, 30 µg/mL) and incubated at 30°C for 20-22 h. Phage titer was determined as described (Miller 1992), except that 10-µL aliquots of serial dilutions of phage P1 were plated onto bacteria in top agar.

### 2.3 Results

#### 2.3.1 Identification of UmuC active site loops

We identified UmuC loops 1 and 2 through modeling and homology searches (Figure 2.1). The homology model of UmuC is based upon the crystal structure of Dpo4 since there is currently no crystal structure of UmuC (Beuning 2006a). UmuC loops 1 and 2 are predicted to be near the active site of the polymerase and are hypothesized to play a role in DNA template binding and alignment, as well as in interactions with the incoming nucleotide. We defined loop 1 as residues 31-38 and loop 2 as residues 50-54. For comparison, the sequences of some other Y family DNA polymerases are also shown, including *E. coli* DinB, *S. solfataricus* Dpo4, human Pol κ, human pol ι, and *S. cerevisiae* Rev1 (Figure 2.1). Loop 1 is highly conserved among UmuC sequences but not within the entire Y family of polymerases, while loop 2 is more variable.
Figure 2.1. *E. coli* UmuC is a Y family DNA polymerase that shares little homology with other members of the Y family. (A) Homology model of UmuC (Beuning 2006a). UmuC Loop 1 (31-38) is shown in red, and Loop 2 (50-54) is shown in blue. UmuC residues N32 (black), N33 (green), and D34 (purple) are predicted to determine the “gap” size that dictates which lesions can fit into the active site (Chandani 2009; Seo 2009; Chandani 2010). The backbone of UmuC is shown in yellow. DNA is rendered as sticks and colored by atom identity. The illustration was prepared using VMD (Humphrey 1996). (B) Amino acid sequences of representative Y-family polymerases showing conserved residues aligned with loop 1 (31-38, red box) and loop 2 (50-54, blue box). Secondary structure based on the crystal structure of Dpo4 is shown above the alignment (Ling 2001; Boudsocq 2002). Ec, *Escherichia coli*; St, *Salmonella typhimurium*; Hs, *Homo sapiens*; Ss, *Sulfolobus solfataricus*; Sc, *Saccharomyces cerevisiae*. The first five sequences above the black line are UmuC and its homologs which share almost 100% homology in loop 1.
2.3.2 Variants in the N-terminal region of Loop 1 cause hypersensitivity to UV radiation

Alanine-scanning mutagenesis was used to determine the extent to which mutating the residues of loop 1 (S31 – I38) contribute to the ability of UmuC to facilitate UV-induced mutagenesis \textit{in vivo}. We first characterized the alanine variants of UmuC N32, N33, and D34 for UV-induced mutagenesis, which is a prominent phenotype of UmuC. Variants of UmuC expressed in the GW8017 (ΔumuDC) strain do not normally cause sensitivity to UV. Moreover, cells in which the \textit{umuDC} genes have been deleted are only modestly sensitive to UV (Opperman 1999; Beuning 2006c). However, UmuC with alanine mutations at residues 32-34 cause significantly greater sensitivity to UV radiation than wild-type UmuC (Figure 2.2A). Strains harboring either UmuC N33A or D34A variants are substantially more sensitive to UV than strains without UmuC.
Figure 2.2. N-terminal loop 1 variants 32, 33, and 34 cause hypersensitivity to UV radiation in a strain that is not normally sensitive to UV. (A) Assays were performed with the pGY9738 plasmid and the following derivatives in GW8017: pGY9738 (umuDC, WT; ■); pGB2 (empty vector, EV; ♦); pGY9738-N32A (umuDC N32A; ×); pGY9738-N33A (umuDC N33A; ◊); pGY9738-D34A (umuDC D34A; ◌). (B) Conservative mutations of N32, N33, and D34 conferred hypersensitivity to UV radiation. Assays were performed with the pGY9738 plasmid and the following derivatives in GW8017: pGY9738 (umuDC, WT; ■); pGB2 (empty vector, EV; ♦); pGY9738-D34N (umuDC D34N; ▲); pGY9738-D34E (umuDC D34E; ×); pGY9738-N32Q (umuDC N32Q; ●); pGY9738-N33Q (umuDC N33Q; Δ); pGY9738-N33D (umuDC N33D; ±); pGY9738-N32D (umuDC N32D; ◊). (C) Immunoblot showing steady-state levels of UmuC variants expressed from plasmids encoding the umuDC genes in GW8017. The wild-type plasmid was pGY9738, and the empty vector was pGB2.
We made conservative substitutions of N32, N33 and D34 by changing N32 or N33 to Gln or D34 to Glu. These substitutions increase the length of the side chain at these positions by one methylene (CH₂) group while maintaining similar chemical characteristics of the side chain functional groups. We also mutated N32 or N33 to Asp or D34 to Asn to test the effects of altering the charge of the side chains on the ability of UmuC to contribute to UV survival. Each conservative substitution caused UV sensitivity in a ΔumuDC strain compared to wild-type UmuC (Figure 2.2B). Therefore, these residues may play an important role in the ability of UmuC to bypass UV-induced lesions in DNA, supporting the conjecture that these residues are involved in binding the DNA substrate, since altering these residues even in a conservative manner causes extreme sensitivity of the cells harboring these mutations to UV light.

We determined the steady-state expression levels of the UmuC variants to rule out the possibility that the decrease in survival of cells expressing these variants was due to a change in protein levels. Western blots showed that steady-state expression levels for each variant were similar to that of wild-type UmuC (Figure 2.2C). Therefore, the effects noted above are not due to altered expression levels of the UmuC variants. It should be noted that LexA is able to repress umuDC expression in this context; SOS induction causes an approximate two-fold increase in umuDC expression compared to levels in uninduced cells (Sommer 1993b).

Given the extreme sensitivity to UV light conferred by UmuC variants with substitutions at 32, 33, and 34, we examined the effect of UV radiation on cells expressing these UmuC variants in the AB1157 strain (umuDC⁺) to determine whether these variants caused sensitivity to UV radiation even in a strain with a chromosomal
copy of *umuDC*. UmuC variants with either alanine or more conservative substitutions at positions 32, 33, and 34 conferred extreme sensitivity to UV radiation in comparison to wild-type UmuC (Figure 2.3A and B).

The *argE3* reversion frequency of wild-type *umuDC* strains expressing UmuC variants containing single alanine substitutions at positions 32, 33, and 34 was determined by
exposure to 5 J/m² UV radiation. Cells expressing the UmuC variants showed a significant decrease in mutation frequency compared to cells expressing wild-type UmuC (Figure 2.3C). We previously observed that plasmid-borne umuC can suppress the extreme sensitivity to UV that is characteristic of strains in which both umuC and recJ have been deleted (Courcelle 2006; Ollivierre 2011). RecJ helps to partially degrade DNA at replication forks that are stalled due to lesions to help aid in re-start of DNA replication. Without RecJ, replication is delayed, and recovery of DNA synthesis relies on TLS specifically by pol V to bypass UV lesions (Courcelle 2006). However, changing UmuC N32, N33, and D34 to alanine caused the cells harboring plasmids expressing these variants to confer sensitivity to UV radiation and therefore fail to complement the ΔumuC ΔrecJ strain in comparison to cells expressing wild-type UmuC (Figure 2.4 A-C). Conservative mutations of N32, N33, and D34 also conferred hypersensitivity to UV radiation in this strain (Figure 2.4 A-C). Therefore, these UmuC variants fail to fulfill various cellular functions of UmuC, including induced mutagenesis and complementation of a ΔumuC ΔrecJ strain, in addition to conferring hypersensitivity on a wild-type strain.
Figure 2.4. UmuC N-terminal loop 1 variants S31, N32, N33, and D34 confer hypersensitivity to UV radiation in a ΔumuC ΔrecJ strain. (A) Mutation of N-terminal loop 1 residue N32 causes sensitivity to UV radiation. Assays were performed with pGY9738 plasmid and the following derivatives in PB102 (AB1157 ΔumuC ΔrecJ; pGY9738 (umuD'C, WT; ■); pGB2 (empty vector, EV; ●); pGY9738-N32A (umuD'C N32A; ▼); pGY9738-N32Q (umuD'C N32Q; □); and pGY9738-N32D (umuD'C N32D; □). (B) Mutation of N-terminal loop 1 residue N33 causes hypersensitivity to UV radiation. Assays were performed with pGY9738 plasmid and the following derivatives in PB102: pGY9738 (umuD'C, WT; ■); pGB2 (empty vector, EV; ♦); pGY9738-N33Q (umuD'C N33Q; Δ); pGY9738-N33D (umuD'C N33D; +); and pGY9738-N33A (umuD'C N33A; ◊). (C) Mutation of N-terminal loop 1 residue D34 causes hypersensitivity to UV radiation. Assays were performed with pGY9738 plasmid and the following derivatives in PB102: pGY9738 (umuD'C, WT; ■); pGB2 (empty vector, EV; ●); pGY9738-D34E (umuD'C D34E; ×); pGY9738-D34A (umuD'C D34A; ◢); and pGY9738-D34N (umuD'C D34N; ▲). (D) N-terminal loop 1 variant S31A causes hypersensitivity to UV radiation and confers upon strains a growth defect (G), a non-growth defect (N), or an intermediate growth defect (I). Assays were performed with the pGY9738 plasmid and the following derivatives in PB102: pGY9738-S31A (umumD'C S31A (N); ▲); pGY9738 (umuD'C, WT; ■); pGY9738-S31A (I) (umuD'C S31A (I); ×); pGB2 (empty vector, EV; ●); and pGY9738-S31A (G) (umuD'C S31A (G); Δ).

Cells expressing the S31A variant of UmuC exhibit a growth defect, observed in AB1157, GW8017, and PB102 strains, that was not observed with other variants (data not shown). Cells harboring S31A are extremely sensitive to UV radiation in the PB102 strain (Figure 2.4D). However, we sometimes observed these cells to be resistant to UV,
which we attribute to the possible acquisition of suppressor mutations. Because of the apparent instability of strains harboring the S31A variant, we decided not to pursue further characterization of this variant.

2.3.3 N-terminal loop 1 variants confer sensitivity to 4-NQO and N33A causes sensitivity to NFZ

We wanted to investigate whether the N-terminal loop 1 variants contribute to survival in the presence of other DNA damaging agents. We assayed survival of the AB1157 (umuDC\(^+\)) strain harboring low-copy plasmids expressing the UmuC N32A, N33A, and D34A variants in the presence of NFZ and 4-NQO. NFZ is thought to cause \(N^2\)-furfuryl-dG adducts, and strains in which \(dinB\), the gene encoding DNA polymerase IV, is deleted are very sensitive to NFZ (Jarosz 2006). The major adduct formed from 4-NQO is at the \(N^2\) position of guanine; 4-NQO also forms adducts at the \(C^8\) position of dG as well as at the \(N^6\) position of dA (Friedberg 2006). UmuC variants N32A, N33A, and D34A conferred sensitivity to 4-NQO in the \(umuDC^+\) strain (Figure 2.5A). UmuC variant N33A conferred sensitivity to NFZ, but N32A and D34A did not (Figure 2.5B). It has been observed that NFZ is a relatively weak DNA damaging agent (Ona 2009), which could account for these differences. Nonetheless, mutation of the N-terminal loop 1 residues causes sensitivity to 4-NQO as well as to some extent to NFZ.
2.3.4 Deletion of dnaQ or modification of the β-binding motifs of UmuC does not alter UV-hypersensitivity of N-terminal Loop variants

Disrupting dnaQ, which encodes for the ε subunit of pol III, disrupts the proofreading function of pol III and has been shown to allow pol III to bypass damaged DNA under some circumstances (Livneh 1986; Vandewiele 1998; Pages 2005). To determine whether the proofreading subunit of DNA polymerase III plays a role in the hypersensitivity to UV conferred by N32, N33, and D34 variants, we expressed UmuC N32A, N33A, and D34A in GW2771, GW2771 spq-2, and GW2771 spq-2 dnaQ903 strains (Table 2.1). We compared the effects of the UmuC N-terminal loop variants in GW2771, a wild-type strain, to GW2771 spq-2, where spq-2 is an antimutator allele of dnaE (amino acid substitution V832G in DnaE) (Lifsics 1992; Slater 1994) that suppresses the mutator effect of the disruption of dnaQ, and the isogenic strain in which dnaQ has been disrupted. Whereas N32A caused modest sensitivity to UV in these strains as in other strains (Figure 2.6A), variants N33A and D34A caused hypersensitivity to UV.
in all three strains (Figure 2.6 B-C). Therefore, deletion of \textit{dnaQ} did not suppress the UV hypersensitivity caused by UmuC variants N33A and D34A.

Next, we tested the role of binding to the \(\beta\) processivity clamp in the ability of UmuC mutations N32A, N33A and D34A to confer UV hypersensitivity. We combined the N32A, N33A and D34A individual mutations with mutation of both of the known \(\beta\) binding sites of UmuC to alanine. The mutations in the beta binding sites of UmuC are designated as \(\beta_{1\&2}\) indicating the multiple mutations \(313\text{LTP}^{315} \rightarrow 313\text{AAA}^{315}\) and \(357\text{QLNLF}^{361} \rightarrow 357\text{AAAAA}^{361}\). These mutations disrupt the binding of UmuC to the \(\beta\) clamp, and therefore the resultant UmuC should not be recruited to the replication fork (Beuning 2006a; Shurtleff 2009). N32A, N33A and D34A conferred hypersensitivity to UV radiation when combined with mutations in the \(\beta\)-binding motifs in the GW8017 strain (Figure 2.6D), indicating that the UV-hypersensitivity is caused by a mechanism that is at least partially independent of binding to the \(\beta\) clamp. When wild-type UmuC is present on the chromosome, in the case of AB1157, the mutation of the \(\beta\)-binding motifs partially suppresses the hypersensitivity conferred by N33A and D34A and almost fully suppresses the moderate sensitivity conferred by N32A (Figure 2.6E). Therefore, disruption of the \(\beta\)-binding motifs in this context appears to allow wild-type UmuC to confer resistance to DNA damage, presumably by allowing access of wild-type UmuC to DNA.
Figure 2.6. Deletion of the proofreading subunit of pol III (dnaQ) or mutation of β clamp binding sites of UmuC do not suppress sensitivity to UV radiation caused by mutation of N32, N33, D34. In A-C, the respective plasmids were assayed in the following strains GW2771 (×); GW2771 spq-2 (▲); GW2771 spq-2 dnaQ903 (●). N-terminal loop 1 variants (A) N32A, (B) N33A, or (C) D34A confer sensitivity to UV radiation despite the deletion of dnaQ. Assays were performed with derivatives of pGY9738 in the listed strains. (D) The hypersensitivity to UV radiation conferred by variants N32A, N33A and D34A in GW8017 was not suppressed by the mutation of the β binding motifs of UmuC. (E) The hypersensitivity to UV radiation conferred by UmuC variants N32A, N33A, and D34A in AB1157 was modestly suppressed by the mutation of the β binding motifs of UmuC. In D-E assays were performed with plasmid pGY9738 and derivatives in GW8017 (D) and AB1157 (E): pGB2 (empty vector, EV; ♦); pGY9738 (umuD/C, WT; ■); pGY9738-N32A (umuD/C N32A; ●); pGY9738-N32Aβ₁+₂ (umuD/C N32A β₁+₂; ○); pGY9738-N33A (umuD/C N33A; ◊); pGY9738-N33A β₁+₂ (umuD/C N33A β₁+₂; Δ); pGY9738-D34A (umuD/C D34A; ◇); and pGY9738-D34A β₁+₂ (umuD/C D34A β₁+₂; ×).
2.3.5 Variants in the N-terminal region of Loop 1 partially suppress inhibition of RecA-mediated homologous recombination

UmuD' and UmuC inhibit RecA-mediated homologous recombination when present at elevated levels such as in the SOS response, which may be important for regulating RecA and its involvement in cellular processes (Sommer 1993a; Szpilewska 1995; Boudsocq 1997; Rehrauer 1998; Sommer 2000; Sutton 2001). UmuC variants N32A, N33A, and D34A were expressed from the respective derivatives of plasmid pGY9738 in a ΔumuDC strain. Cells were exposed to bacteriophage P1vir ΔyeaB and plated on selective media to measure transductants that result from RecA-mediated homologous recombination. We observed that the N32A, N33A, and D34A variants confer intermediate inhibition of homologous recombination as compared to wild-type UmuD'C, which inhibits recombination (Figure 2.7A). This suggests that these mutations may disrupt a direct interaction between RecA and UmuC, although other interpretations are possible, such as a disruption of an interaction between pol V and RecA that is mediated by UmuD'2. Two residues, K342 and Y270 were shown previously (Sommer 2000) to enhance the inhibition of homologous recombination when mutated to glutamine and cysteine, respectively. These two residues are predicted to be in close proximity to N32, N33, and D34, offering further evidence that this region is important for interactions that modulate recombination (Figure 2.7B).
Figure 2.7. UmuD′:C inhibits homologous recombination facilitated by RecA. (A) Variants in plasmid pGY9738 (umuD′C, WT) were expressed in strain GW8017 (ΔumuDC). Cells harboring the variants N32A, N33A, and D34A show intermediate inhibition of RecA-mediated homologous recombination, whereas cells harboring other variants (Y11A, G52A, P54A, to right of line) inhibit homologous recombination to a similar extent as WT umuD′C. Transduction efficiency is measured in cfu/pfu as a percent of EV (1.31 x10⁻⁵ cfu/pfu, normalized to 100.0%): WT: 2.09%; N32A: 58.1%; N33A: 48.4%; D34A: 38.1%; Y11A: 0.358%; G52A: 0.956%; P54A: 0.000%. (B) Homology model of UmuC (6). The backbone of UmuC is shown in yellow. UmuC Loop 1 (31-38) is shown in red, and Loop 2 (50-54) is shown in blue. UmuC residues N32, N33, and D34 (green) are near previously studied residues K342 and Y270 (purple) in the UmuC model. Previously studied residue Y11 as well as residues G52 and P54 (orange) are shown for comparison. Cells harboring variants Y270C and K342Q show a significant decrease in RecA-mediated homologous recombination (Humphrey 1996).

In order to assess the specificity of this effect, we determined the ability of UmuC variants with other mutations to inhibit homologous recombination. The UmuD loop 2 mutants G52A and P54A exhibited an inhibition of homologous recombination that is similar to that of wild-type UmuC (Figure 2.7A). Moreover, the previously studied variant Y11A confers hypersensitivity to UV radiation (Shurtleff 2009), but we find that that the UmuC Y11A variant inhibits homologous recombination to a similar extent as wild-type UmuC. G52, P54, and Y11 are predicted to be located in the interior of the protein, and so are unlikely to be available for protein-protein interactions (Figure 2.7B). These observations suggest that N32, N33, and D34 have a specific role in UmuC
function in both UV resistance presumably via DNA binding and TLS as well as in modulating homologous recombination.

### 2.3.6 C-terminal region of Loop 1 does not play a significant role in UmuC function

Cells expressing UmuC variants at the C-terminus of loop 1 (G35A, C36A, V37A, or I38A) have a mutation frequency similar to or greater than that of cells expressing wild-type UmuC, therefore these variants are proficient for UV-induced mutagenesis (Figure 2.8A). Additionally, UmuC variants G35A, C36A, V37A, and I38A each complement the $\Delta$umuC $\Delta$recJ strain as well as wild-type UmuC for UV resistance (Figure 2.8B). For the cells harboring these variants, there is good correlation between UV-induced mutagenesis and the UV-resistance in the $\Delta$umuC $\Delta$recJ strain.

The steady-state levels of the loop 1 C-terminal UmuC variants G35A, C36A, V37A, and I38A after exposure to UV radiation were determined by using Western blotting (Figure 2.8C). Each of these variants is present at levels similar to wild-type UmuC. Therefore, the expression levels of UmuC are not altered by the presence of these loop 1 mutations.
Figure 2.8. UmuC C-terminal loop 1 variants do not confer sensitivity to UV radiation and are proficient for mutagenesis. (A) UV-induced mutation frequency of selected variants in plasmid pGY9738 (umuDC) in strain GW8017 (ΔumuDC). The wild-type plasmid was pGY9738 and the empty vector was pGB2. UV-induced arg+ revertants of GW8017 harboring the plasmids indicated were irradiated at 25 J/m². EV: induced mutants (IM) 7.66 x 10⁻⁷, spontaneous mutants (SM) 3.36 x 10⁻⁷, mutation frequency (MF) 4.31 x 10⁻⁷; WT: (IM) 1.85 x 10⁻⁵, (SM) 6.48 x 10⁻⁶, (MF) 1.21 x 10⁻⁶; G35A: (IM) 1.73 x 10⁻⁵, (SM) 7.58 x 10⁻⁶, (MF) 9.68 x 10⁻⁶; C36A: (IM) 3.25 x 10⁻⁵, (SM) 2.30 x 10⁻⁶, (MF) 3.03 x 10⁻⁶; V37A: (IM) 2.24 x 10⁻⁵, (SM) 5.41 x 10⁻⁶, (MF) 1.70 x 10⁻⁵; I38A: (IM) 2.80 x 10⁻⁵, (SM) 2.73 x 10⁻⁶, (MF) 2.53 x 10⁻⁵. (B) C-terminal loop 1 variants do not confer sensitivity to UV radiation. Assays were performed with pGY9738 plasmid and the following derivatives in PB102: pGY9738-C36A (umuDC C36A; ×); pGY9738 (umuDC, WT; □); pGY9738-I38A (umuDC I38A; Δ); pGY9738-V37A (umuDC V37A; ●); pGY9738-G35A (umuDC G35A; ○); pGB2 (empty vector, EV; ●). (C) Immunoblot showing steady-state levels of UmuC expressed from variant umuDC plasmids (GW8017). The wild-type plasmid was pGY9738 and the empty vector was pGB2.
2.3.7 Variants in UmuC Loop 2 contribute little to UV survival

We also used alanine-scanning mutagenesis to determine the importance of residues in loop 2 (K50-P54) in UV-induced mutagenesis. Cells expressing these UmuC variants have similar mutation frequencies as cells expressing wild-type UmuC (Figure 2.9A). Each of the loop 2 variants confers UV resistance in a ΔumuC ΔrecJ strain that is similar to that of cells harboring wild-type UmuC (Figure 2.9B). Of these variants, only cells expressing UmuC D53A are significantly more sensitive to UV radiation than cells expressing wild-type UmuC (Figure 2.9B) and exhibit a lower mutation frequency than cells expressing wild-type UmuC (Figure 2.9A). With the exception of D53A, mutating each residue of loop 2 to alanine did not have a significant effect on UmuC-dependent UV-induced mutagenesis or UV survival. A western blot (Figure 2.9C) shows that the steady-state expression level of each UmuC variant is similar to that of wild-type UmuC, indicating that the mutations constructed in UmuC do not alter its expression levels.
Figure 2.9. UmuC loop 2 variants do not confer sensitivity to UV radiation and are proficient for mutagenesis. (A) UV-induced mutation frequency of selected variants in plasmid pGY9738 (umuD'C) in strain GW8017 (ΔumuDC). The wild-type plasmid was pGY9738 and the empty vector was pGB2. UV-induced $arg^+$ revertants of GW8017 harboring the plasmids indicated were irradiated at 25 J/m². EV: induced mutants (IM) $7.66 \times 10^{-7}$, spontaneous mutants (SM) $3.36 \times 10^{-7}$, mutation frequency (MF) $4.31 \times 10^{-5}$; WT: (IM) $1.85 \times 10^{-5}$, (SM) $6.48 \times 10^{-6}$, (MF) $1.21 \times 10^{-5}$; K50A: (IM) $3.28 \times 10^{-5}$, (SM) $2.96 \times 10^{-6}$, (MF) $2.98 \times 10^{-5}$; M51A: (IM) $1.71 \times 10^{-5}$, (SM) $3.00 \times 10^{-6}$, (MF) $1.41 \times 10^{-5}$; G52A: (IM) $3.38 \times 10^{-5}$, (SM) $6.36 \times 10^{-6}$, (MF) $2.74 \times 10^{-6}$; D53A: (IM) $1.16 \times 10^{-5}$, (SM) $3.32 \times 10^{-6}$, (MF) $8.29 \times 10^{-6}$; P54A: (IM) $3.77 \times 10^{-5}$, (SM) $7.22 \times 10^{-6}$, (MF) $3.05 \times 10^{-5}$. (B) C-terminal loop 1 variants do not confer sensitivity to UV radiation. Assays were performed with pGY9738 plasmid and the following derivatives in PB102: pGY9738-K50A (umuD'C K50; ×); pGY9738 (umuD'C, WT; ■); pGY9738-G52A (umuD'C G52A; ○); pGY9738-M51A (umuD'C M51A; □); pGY9738-P54A (umuD'C P54A; ◊); pGY9738-D53A (umuD'C D53A; ●); pGB2 (empty vector, EV; ●). (C) Immunoblot showing steady-state levels of UmuC expressed from variant umuDC plasmids (GW8017). The wild-type plasmid was pGY9738 and the empty vector was pGB2.
2.4. Discussion

We identified loops 1 and 2 as likely to harbor residues important for the function of UmuC in mutagenesis. Loop 1, consisting of residues 31-38, is located just after predicted β sheet 2 in the fingers domain. Loop 2, consisting of residues 50-54, is located between predicted α helices B and C and is also in the fingers domain (Boudsocq 2002). Residue P54 of loop 2 was noted by Boudsocq et. al to be an important conserved residue for hydrophobic core formation (Boudsocq 2002). The residues in loop 1, however, are not conserved among the Y family polymerases (Boudsocq 2002), but are conserved among UmuC homologs, including those found on plasmids (Figure 2.1) (Woodgate 1992). Considering these residues are conserved among UmuC homologs, but are not conserved among Y family polymerases, we hypothesized that they may contribute to the specificity of UmuC in TLS. Loop 2 residues are more conserved among Y family polymerases and less conserved among UmuC homologs (Figure 2.1) suggesting loop 2 may not play as important a role in lesion bypass specificity. We observed that the N-terminal part of loop 1, residues 32-34, plays a significant role in UmuC function. When each of these residues is mutated to alanine, cells harboring the variants are non-mutable (Figure 2.3C). The ability of pol V to inhibit homologous recombination is also disrupted by mutating residues 32-34 (Figure 2.7A), so perturbation of the interaction between pol V and RecA may partially explain the extreme sensitivity to UV radiation conferred by these variants. On the other hand, mutations of the residues in the C-terminal region of loop 1 and loop 2 do not have as significant an impact on UmuC-dependent mutagenesis (Figure 2.8 and 2.9).
To date, there is no crystal structure of UmuC, the polymerase subunit of Pol V, so we rely on molecular modeling based on homology to other Y family polymerases as well as comparisons to polymerases in other families. Y family polymerases are notable for their lack of a specific α helix, the O helix, found in the active sites of A family DNA polymerases that strongly contributes to fidelity (Jarosz 2007). This α helix is positioned such that it interacts with the incoming nucleotide as well as with the template strand base and contributes to the high fidelity of replicative DNA polymerases (Patel 2001; Beard 2003; Johnson 2003). Loops 1 and 2 in UmuC as defined here are in a similar location in UmuC as the O helix is in the A family of DNA polymerases, which suggests a role in fidelity or specificity (Figure 2.1) (Jarosz 2007). Chandani and Loechler describe the opening next to the active site of Y family polymerases as a “chimney” (Chandani 2009; Chandani 2010), with a cluster of amino acids (UmuC S31, N32, N33) that control the size of the opening and potentially dictate which adducts are bypassed by UmuC. We observed that UmuC variants S31A, N32A, N33A, and D34A fail to complement the UV sensitivity of the ΔumuC ΔrecJ strain as well as the less sensitive ΔumuDC strain. Conservative mutations of these residues also failed to complement these strains, suggesting that they are extremely important for UmuC function, perhaps in determining which lesions UmuC bypasses by controlling the size of the active site opening. However, our observations suggest that the UV sensitivity conferred by mutations of the loop 1 N-terminal residues may not be entirely due to action at the replication fork, as concomitant mutation of the β binding motifs suppress their UV hypersensitivity only in the presence of wild-type UmuC and not in its absence.
UmuC I38 is a conserved residue that is located above the base of the incoming nucleotides, is next to the opening of the active site, and has been shown to be important for efficient bypass of $N^2$-benzo[a]pyrene-dG (Chandani 2009; Seo 2009). The residue neighboring I38, UmuC M51, also contacts the incoming nucleotide in the model, giving rise to the suggestion that these residues may play a significant role in lesion bypass, or more specifically, dNTP insertion. We find that cells (ΔumuC ΔrecJ) harboring plasmids expressing UmuC M51A were only modestly sensitive to UV radiation and were fully proficient for UV-induced mutagenesis (Figure 2.9B), suggesting that M51 may not have a significant role in UV mutagenesis, but may still have a role in dNTP insertion (Chandani 2009; Seo 2009). In our experiments, the I38A variant complemented the ΔumuC ΔrecJ strain and was fully proficient in UV-induced mutagenesis. The differences in these observations could be due to the fact that other studies probed UmuC bypass of adducts of a common metabolite of the carcinogen benzo-[a]-pyrene (Chandani 2009; Seo 2009), while here we are concerned mainly with photoproducts of UV radiation.

Mutating S31 to L results in a UmuC variant that is unable to complement a ΔumuDC strain in UV mutagenesis (Beuning 2009). S31 is located in the “flue” region of the UmuC protein (Chandani 2009) and predicted to be within 5 Å of the deoxyribose of the template base (Beuning 2009). We observed that strains expressing UmuC S31A had a growth defect and subsequently identified populations of strains expressing UmuC S31A with severe, intermediate, or no growth defect, indicating that this variant acquired suppressor mutations. Strains expressing UmuC-S31L did not have this growth defect (Beuning 2009). The location and behavior of S31A as well as S31L suggests that this
residue plays an important role in UmuC function and perhaps more specifically in template DNA alignment in the active site.

Human DNA pol η, a functional homolog of UmuC, is encoded by the *hRAD30A* gene and without it, humans develop the genetic disorder xeroderma pigmentosum variant (XPV), making them 100 times more susceptible to UV radiation-induced skin cancer from exposure to sunlight (Cleaver 1999; Johnson 1999; Masutani 1999; Glick 2001). Human pol η bypasses T-T CPDs formed from exposure to UV radiation accurately by inserting primarily adenine opposite each T of the CPD (Johnson 2000; Masutani 2000). Motif II of human pol η (residues 52-73) aligns with the loop 2 region in UmuC (Figure 2.1). The S62G variant is more efficient in bypassing T-T CPDs, 8-oxo-G, \( O^4 \)-MeT, \( O^6 \)-MeG, abasic sites, and ethenoA lesions, as well as in copying undamaged DNA, than is wild-type pol η (Glick 2001). A crystal structure of human pol η depicts S62 interacting with the 5′ base of a T-T CPD (Glick 2001; Biertumpfel 2010). Human pol η R61A, also in loop 2, exhibits decreased efficiency but improved fidelity opposite the TT CPD lesion (Biertumpfel 2010). We observed that several variants of UmuC loop 2 (residues 50-54) conferred increased mutagenesis on strains harboring those mutations, although these increases were not statistically significant compared to wild-type UmuC.

Human pol η loop 1 residue Q38 hydrogen bonds to both thymines of a T-T CPD lesion (Biertumpfel 2010). Mutating Gln to Ala decreases the efficiency of pol η, causing stalling after the T-T CPD, perhaps due to incorrect base stacking and therefore misalignment with the incoming nucleotide (Biertumpfel 2010). This interaction with the template base supports the idea that loop 1 residues interact with the nascent base pair in the active site (Chandani 2009; Chandani 2010).
There are very few examples of variants of UmuC that confer hypersensitivity to UV radiation as seen here with the N-terminal loop 1 residues 32-34. A variant of UmuC with a mutation of the steric gate residue Y11 failed to complement a \( \Delta umuDC \) strain for UV-induced mutagenesis, rendered these cells hypersensitive to UV radiation and was dominant negative (Shurtleff 2009). Though the Y11A variant confers similar hypersensitivity to UV radiation as the mutations N32A, N33A, and D34A, its apparent mechanism of action is different than that of the variants reported here. Disruption of the \( dnaQ \) gene, which encodes the proofreading subunit of DNA pol III, almost completely suppressed the UV hypersensitivity of cells expressing UmuC-Y11A, while with the variants reported here, disruption of \( dnaQ \) had almost no effect on UV sensitivity. Moreover, mutating the \( \beta \) clamp binding sites of UmuC did not suppress UV hypersensitivity conferred by N33A and D34A in a \( \Delta umuDC \) strain, in striking contrast to our observations with Y11A. By analogy to Dpo4, we hypothesized that the UmuC N-terminal loop 1 residues 31-34 interact with the DNA template, and mutating these residues possibly disrupts this interaction. This perturbed interaction with the DNA substrate may cause cells to confer hypersensitivity to UV radiation if the UmuC variants cannot bypass lesions caused by exposure to UV radiation. Because mutating the \( \beta \) clamp binding sites of UmuC along with mutations at residues 32-34 did not suppress their hypersensitivity to UV radiation in a \( \Delta umuDC \) strain, unlike the previously studied Y11A variant (Shurtleff 2009), there is likely an additional factor responsible for the observed UV hypersensitivity.

Even more intriguing is the possibility that a RecA interaction is interrupted by the mutation of N32, N33, or D34 to alanine. UmuC requires RecA for TLS, but the site
of interaction on UmuC is not completely elucidated (Schlacher 2005; Schlacher 2006; Jiang 2009). Our data indicate that mutating N32, N33, or D34 to alanine confers intermediate inhibition of RecA-mediated homologous recombination compared to that of the empty vector and wild-type UmuC (Figure 2.7A). A selection experiment for UmuD′ and UmuC variants that increased the inhibition of RecA-mediated homologous recombination identified (Sommer 1998) seven UmuC variants: F10L, Y270C, K277E, F287L, F287S, K342Q, F351I. We mapped these residues on a model of UmuC (Figure 2.7B) and observed that N32, N33, and D34 are in close proximity to Y270 and K342, two residues that when mutated were found to enhance the inhibition of RecA-mediated homologous recombination (Sommer 1998). All five residues (N32, N33, D34, Y270, K342) are located near the incoming template strand of ssDNA, consistent with the idea that inhibition of RecA-mediated homologous recombination by UmuD′2C occurs via a physical interaction of pol V with RecA (Sommer 1993a; Boudsocq 1997; Rehrauer 1998; Sommer 1998; Sommer 2000). Taken together, these observations suggest that a site of interaction between UmuC and RecA is the location on UmuC where the single-stranded DNA template enters UmuC. Our observations cannot rule out alternative models, however, including a scenario in which UmuD′2 mediates interactions with UmuC and RecA or in which competition for DNA substrates forms the basis of pol V-induced inhibition of homologous recombination. Disruption of the interaction of pol V with RecA could lead to the UV sensitivity that we observed, as that interaction is critical for TLS (Reuven 2001; Schlacher 2005; Schlacher 2006; Fujii 2009; Jiang 2009). A recent molecular modeling supports the possibility that 32NND34 could be a site of interaction for RecA (Chandani 2013). In sum, we show here that the N-terminus of loop
plays an essential role in cellular resistance to DNA damaging agents and facilitates possible interactions with RecA that could be critical for pol V-mediated UV mutagenesis.

2.5 References


Chapter 3: Point mutations in *Escherichia coli* DNA pol V that confer resistance to non-cognate DNA damage and enhance UmuD cleavage

3.1 Introduction

Y family DNA polymerases were classified as a distinct DNA polymerase family in 2001 (Ohmori 2001). DNA polymerases that are members of the Y family have a similar N-terminal palm domain, which possesses the active site, to DNA polymerases of other families, but with smaller fingers and thumb domains (Zhou 2001; Yang 2005; Pata 2010). Unlike replicative DNA polymerases, such as *E. coli* pol III of the C family, Y family polymerases do not possess 3′-5′ proofreading activity (Friedberg 2006). Their unique C-terminal domains, called the little finger (Ling 2001; Silvian 2001; Bunting 2003) or the polymerase-associated domain (PAD) (Trincao 2001), share little sequence similarity among the members of the family but are structurally similar (Ling 2001; Silvian 2001; Trincao 2001; Bunting 2003). The DNA substrate is bound by the thumb and fingers domains, whose smaller sizes allow for a more open, solvent-accessible active site that can accommodate lesions and unusual structures in DNA (Yang 2005; Chandani 2010).

*Escherichia coli* (*E. coli*) has five DNA polymerases, two of which (pols IV and V) are members of the Y family. *E. coli* DNA polymerase DinB (pol IV) is the product of the *dinB* gene, while UmuD’2C (pol V) is a product of the *umuDC* genes, all of which are upregulated in the SOS response (Friedberg 2006). When replication is disrupted by a lesion in the DNA template, the ssDNA downstream from the lesion becomes coated with RecA, forming a RecA-ssDNA nucleoprotein filament, which in turn initiates the SOS
response by facilitating the self-cleavage of the LexA repressor protein, leading to the expression of at least 57 genes, including *dinB* and *umuDC* (Radman 1974; Friedberg 2006; Simmons 2008). Once UmuD is expressed, it undergoes a self-cleavage reaction facilitated by the RecA-ssDNA nucleoprotein filament (Burckhardt 1988; Nohmi 1988; Shinagawa 1988). About 20 to 40 minutes after SOS induction, the first 24 amino acids of UmuD are cleaved to form UmuD′, which combines with UmuC to form pol V, which is active in translesion synthesis (TLS) and SOS mutagenesis (Woodgate 1989; Reuven 1999; Tang 1999; Friedberg 2006).

Several protein-protein interactions are important for the function of UmuD′C. While the RecA-ssDNA nucleoprotein filament is important for UmuD cleavage to UmuD′, RecA is also required for pol V-dependent SOS mutagenesis, though the exact mechanism of how RecA facilitates SOS mutagenesis is still unclear (Reuven 2001; Schlacher 2005; Schlacher 2006; Fujii 2009; Jiang 2009; Karata 2012). Cryo-electron microscopy experiments showed that pol V binds to the end of the RecA-ssDNA nucleoprotein filament in such a way that pol V is recruited to the lesion site; pol V was also observed to bind to the grooves of the filament (Frank 2000). Our previous work suggested that a potential RecA binding site on UmuC includes residues 32 through 34 on the active site loop 1 (Hawver 2011), which is supported by a recent protein docking study using ZDOCK (Chen 2003) and ClusPro (Comeau 2004) (Chandani 2013).

Interaction with the β clamp is also important for pol V functionality in TLS (Dalrymple 2001; Becherel 2002; Sutton 2005). Pol V is thought to interact with the β clamp at two distinct binding sites: at UmuC residues $\text{LTP}^{315}$ and $\text{QLNLF}^{361}$ (Beuning 2006a). The β clamp increases the processivity of DinB dramatically (Wagner...
2000), while interaction with UmuD′2C increases its processivity by anywhere from three
to 100 fold (Maor-Shoshani 2002; Fujii 2004; Karata 2012). Processivity and activity of
pol V are increased in the presence of various co-factors under differing conditions,
including SSB, RecA-ssDNA nucleoprotein filament, and the γ clamp loader (Reuven
1999; Reuven 2001; Maor-Shoshani 2002; Fujii 2004; Fujii 2009; Karata 2012; Vaisman
2012).

There are other regulatory and chaperone proteins that allow for cellular control
of pol V. Two proteases, Lon and ClpXP, help regulate the levels of pol V. The protein
partner of UmuD, whether UmuD′ in the heterodimer UmuDD′, or UmuD in the
homodimer UmuD2, is degraded by ClpXP (Gonzalez 2000; Neher 2003), as UmuD but
not UmuD′ contains the “tethering” sequence necessary to interact with ClpX (Neher
2003). The UmuD2 homodimer (Frank 1996a; Gonzalez 1998) and UmuC are also
degraded by Lon protease (Frank 1996a).

The heat shock and chaperone proteins GroEL, which interact directly with
UmuC, and GroES are required for UV mutagenesis, and mutations in these genes
decrease the stability of UmuC (Donnelly 1989; Liu 1990; Donnelly 1992). Expression of
umuD′ instead of umuD suppresses the nonmutability conferred by mutations in groE,
suggesting that UmuD′ stabilizes an otherwise unstable UmuC (Donnelly 1992). Thus,
protein interactions that stabilize UmuC are important for regulating pol V activity.

Whereas pol V specializes in bypass of lesions produced by UV radiation,
including thymine-thymine cyclobutane pyrimidine dimers (T-T CPDs) and thymine-
thymine 6-4 photoproducts (Tang 2000), DinB specializes in faithfully bypassing \(N^2\)-dG
lesions, including \(N^2\)-furfuryl-dG (Jarosz 2006; Jarosz 2009), \(N^2\)-benzo[a]pyrene-dG
(Shen 2002; Seo 2009), and $N^2$-(1-carboxylethyl)-2′-dG (Yuan 2008) by adding dC across from the lesion. While Pol V faithfully bypasses T-T CPDs by inserting two dA nucleotides opposite the lesion (Szekeres 1996; Tang 2000), it accurately adds dA opposite the 5′ thymine and mutagenically adds dG opposite the 3′ thymine of the 6-4 T-T photoproduct (Becherel 1999; Tang 2000).

Given that pol V and DinB bypass chemically distinct types of DNA damage, we wanted to determine whether we could identify mutations in pol V that would allow it to confer resistance to the DinB-specific damage caused by nitrofurazone (NFZ). We used random mutagenesis and genetic selection to discover UmuC variants (H282P and T412I) that allow pol V to confer resistance to NFZ. Characterization of these variants has led us to the surprising observation that the UmuC variants T412I and H282P cause UmuD to be hypercleavable relative to wild-type UmuC. We also found that UmuC H282P is readily degraded by Lon protease. We conclude that specificity may not be only a matter of protein structure, but also a matter of protein-protein interactions.

3.2 Materials and Methods

3.2.1 Plasmids and strains

Plasmids and strains are listed in Table 3.1. UmuC variants were constructed in the pGY9738 and pGY9739 plasmids, which both encode resistance to spectinomycin (60 µg/mL), for in vivo assays (Sommer 1998).
**Table 3.1 Strains and Plasmids**

<table>
<thead>
<tr>
<th>Bacterial Strain</th>
<th>Relevant Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory stock</td>
</tr>
<tr>
<td>YG7207 AB1157</td>
<td>ΔdinB</td>
<td>(Kim 1997)</td>
</tr>
<tr>
<td>GW8017 AB1157</td>
<td>ΔumuDC</td>
<td>(Guzzo 1996)</td>
</tr>
<tr>
<td>PB107 GW8017</td>
<td>Δlon</td>
<td>This work</td>
</tr>
<tr>
<td>PB103 AB1157</td>
<td>ΔumuDC ΔrecJ</td>
<td>(Ollivierre 2013)</td>
</tr>
<tr>
<td>RW644 BL21(λDE3)</td>
<td>ΔpolB::Ωspec,ΔdinB61::ble,ΔumuDC596::ermGT</td>
<td>(Karata 2012)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pGY9738</td>
<td>$o_1^C$ umuD'C; pSC101-derived</td>
</tr>
<tr>
<td>pGY9739</td>
<td>$o_1^C$ umuDC; pSC101-derived</td>
</tr>
<tr>
<td>pGYDΔC</td>
<td>$o_1^C$ umuD',ΔumuC; pSC101-derived</td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector, pSC101-derived</td>
</tr>
<tr>
<td>pYG768</td>
<td>dinB</td>
</tr>
<tr>
<td>pWSK30</td>
<td>Vector</td>
</tr>
<tr>
<td>pHUC25</td>
<td>pSC101 ori, laqF*, His-umuC, umuD'</td>
</tr>
<tr>
<td>pARAD2</td>
<td>pSC101 ori, umuD'</td>
</tr>
</tbody>
</table>

The pGY9738 plasmid contains a synthetic $umuD'C$ operon; whereas the pGY9739 plasmid contains the $umuDC$ operon. The pHUC25 plasmid for purification of pol V encodes resistance to kanamycin (50 μg/mL) (Karata 2012). Strains used (Table 3.1) were grown in liquid Luria Broth at 37 °C unless otherwise noted. Competent cells
were made using the CaCl₂ method (Sambrook 1989) and transformations were performed as described previously (Beuning 2006b). UmuC variants H282P and T412I were made in the pGY9738, pGY9739, and pHUC25 plasmids using Quikchange or Quikchange Lightning site-directed mutagenesis kits (Agilent Technologies). The presence of mutations was confirmed by DNA sequencing (Massachusetts General Hospital DNA Core Facility, Cambridge, MA; Macrogen, Cambridge, MA; or Eton Bioscience, Charlestown, MA).

3.2.2 Hydroxylamine mutagenesis and genetic selection of UmuC variants

Hydroxylamine (H₃NO) reacts with cytosine specifically, and causes cytosine-guanine to adenine-thymine transition mutations (Miller 1992). Two methods were used to introduce random mutations into plasmids of interest via hydroxylamine. Method One: We treated 10 µg of low-copy number plasmids pGY9738 and pGY9739 in 5 mM hydroxylamine hydrochloride, 5 M NaOH, pH 6.7 for 20 h at 37 °C in a sealed microcentrifuge tube (Rose 1987). DNA was then purified by ethanol precipitation. Method Two was carried out as described (Rose 1987; Miller 1992; Beuning 2009): To a 1 M solution of hydroxylamine in 1.5 M sodium phosphate buffer, pH 6.5, 7.5 µg of DNA was added in a 1:3 (W/V) ratio of hydroxylamine solution, and samples were incubated at 65 °C in sealed microcentrifuge tubes for 1 to 2 h. DNA was then purified (GeneJET PCR Purification Kit, Fermentas). AB1157 ΔdinB competent cells that were made using the CaCl₂ method (Sambrook 1989) were transformed with treated, purified plasmid DNA as described previously (Beuning 2006b). The amount of DNA used for the transformation was normalized by estimation of the concentration of each plasmid via electrophoresis on a 1.1% agarose gel stained with ethidium bromide. The transformation
mixtures (100 µL) were plated on LB-agar plates containing spectinomycin (60 µg/mL) and increasing concentrations from 0 to 6 µg/mL of NFZ (from a 1 mg/mL stock solution of 5-nitro-2-furaldehyde semicarbazone, TCI America, in N,N-dimethyl formamide, Fisher Scientific) and were grown at 37 °C for 24 h. The number of colony forming units (cfu) of treated plasmid was compared to the cfu of untreated plasmid for each concentration of NFZ. Positive hits were considered to be transformed treated DNA with higher cfu than the untreated DNA on a given concentration of NFZ.

Two rounds of validation were conducted before positive hits were analyzed by sequencing. Variants of interest were analyzed further by plating AB1157 ΔdinB strains harboring the selected plasmids which were cultured on plates containing 0.0, 2.1, 2.4, and 2.8 µg/mL NFZ. Colonies were picked from plates containing 2.4 and 2.8 µg/mL NFZ and the plasmid DNA was isolated. These plasmids were then used to transform AB1157 ΔdinB competent cells, which were then assayed for growth once again on agar plates with higher NFZ concentrations (0, 2.6, 3, 3.5, 4, 4.7, 6 µg/mL NFZ). From this process we obtained eight clones, which were then sequenced (Massachusetts General Hospital DNA Core Facility, Cambridge, MA). Only six hits were verified as variants of UmuC, two of which are the subject of this work (Table 3.2); two of the clones did not contain mutations in the umuC gene. These six variants were then reconstructed in both pGY9738 and pGY9739, which were the constructs used in subsequent experiments. In this work, we focus on H282P and T412I; characterization of another variant is reported in Chapter 4.
3.2.3 NFZ survival assays

Strains (AB1157 ΔdinB) harboring UmuC variants identified in the selection process (Table 3.2) were exposed to increasing concentrations of NFZ as described previously (Beuning 2006b; Shurtleff 2009; Hawver 2011). Serial dilutions (20-fold) of overnight cultures were plated on LB agar plates containing 60 µg/mL spectinomycin and the indicated amount of NFZ. The plates were incubated at 37 °C for 20 to 24 h. Error bars represent the standard deviation of the average of at least three trials.

Table 3.2 Nucleotide sequence analysis of umuDC alleles containing missense mutations

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Nucleotide Substitution</th>
<th>Deduced Amino Acid Substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGY9738</td>
<td>A^{845} → C^{845}</td>
<td>His^{282} → Pro^{282}</td>
</tr>
<tr>
<td>pGY9739</td>
<td>C^{1235} → T^{1235}</td>
<td>Thr^{412} → Ile^{412}</td>
</tr>
</tbody>
</table>

¹Nucleotide numbering is based on using the first nucleotide of UmuC as 1. ²Amino acid numbering begins with 1 for the first amino acid of UmuC.

3.2.4 UV survival and mutagenesis assays

UV survival and mutagenesis assays were performed as described previously (Beuning 2006b; Shurtleff 2009; Hawver 2011). GW8017 harboring plasmids expressing the selected variants were exposed to 25 J/m² 254-nm UV radiation for mutagenesis assays. UV survival assays were carried out in PB103 (AB1157 ΔumuDC ΔrecJ). Each
point shown represents the average of at least three trials, and the error bars indicate the standard deviation.

### 3.2.5 Immunoblotting

Western blotting was performed as described previously (Shurtleff 2009; Hawver 2011). Strains harboring the UmuC variants H282P and T412I in the pGY9738 plasmid and the pGY9739 plasmid were exposed to 25 J/m² UV radiation. Proteins were resolved by 14% SDS-PAGE (for UmuC) and 18% SDS-PAGE (for UmuD and UmuD′) and transferred to a polyvinylidene difluoride membrane (Millipore). After blocking in milk, the membrane was then probed with either anti-UmuC (Beuning 2006a) or anti-UmuD/UmuD′ (Beuning 2006c) in 2.5% milk with 0.5X TBS-Tween buffer and washed for 2 min and then 3 times for 10 min each time (for UmuC blots) or rinsed quickly and then washed for three times for 10 min each time (for UmuD/UmuD′ blots) with 1X TBS-Tween buffer. The membrane was probed with horseradish peroxidase-conjugated goat anti-rabbit antibody (Pierce) in 2.5% milk with 0.5X TBS-Tween buffer. The blots were then developed with SuperSignal chemiluminescence reagent (Pierce) and exposed to film (Kodak) which was developed with a Kodak photoprocessor. Blots were analyzed using ImageQuant TL 1D Gel Analysis software (GE Healthcare).

### 3.2.6 Cold sensitivity

Assays were carried out as described previously in the AB1157 strain with variants in the pGY9739 plasmid (umuDC) (Beuning 2006b). Equal volumes of transformation mixtures were plated for growth on LB-agar plates supplemented with...
spectinomycin (60 µg/mL) at permissive (37 °C) and non-permissive (30 °C) temperatures.

3.2.7 Genetic transduction

Transductions were performed as described previously (Hawver 2011) using P1
vir ΔyeaB phage. The phage was titered using the previously described method (Miller

3.2.8 Pol V Purification

Purification was carried out with similar conditions as those described previously
(Karata 2012). Strain RW644 harboring pHUC25 expressing wild-type pol V or UmuC
variants H282P or T412I along with pARAD2 was grown at 37 °C in 40 mL of LB broth
containing 25 µg/mL kanamycin and 25 µg/mL ampicillin overnight. Cultures were then
transferred to 2 L of fresh LB broth containing 50 µg/mL kanamycin and 50 µg/mL
ampicillin and vigorously shaken at 37 °C until the culture reached an OD_{600} of 0.4, at
which time L-arabinose (Acros Organics) was added to a concentration of 4 mg/mL and
cultures were shaken an additional 5 h. Cells were harvested by centrifugation at 3800 x g
for 15 min and then resuspended in 30 mL Buffer A (Karata 2012) and frozen at -80 °C.
Cells were lysed by sonication and centrifuged at 10,000 x g for 60 min. The supernatant
was purified in three steps using a 1 mL HisTrap FF column (GE Healthcare), then a
HiLoad 26/60 Superdex 200 (GE Healthcare), and finally a Bioscale Mini Cartridge CHT
Type I Hydroxyapatite column (Bio-Rad) as described (Karata 2012). Purified proteins
were analyzed by 16% SDS-PAGE, and concentrations of purified proteins were
determined using the Bradford assay (Bradford 1976).
3.2.9 Pol V in vitro lesion bypass assays

Pol V activity was analyzed in vitro by using a 5′-32P labeled 31-mer standing start DNA primer (5′ GCATATGATAGTACAGCTGCAGCGAGCCGACC 3′) and a 61-mer DNA template (3′ CGTATACTATCATGTGACGGTGCCTGCGGXTCCTAGAAGCGTCCGGACTAGACTCATTTGG 5′) in which the template base is X, which is an undamaged guanine, or N2-furfuryl-dG, a damaged base believed to result from treatment with NFZ, N2-furfuryl-dG (Jarosz 2006). The DNA template and 5′-32P labeled primer are combined in a 1:1 ratio of 500 nM each and are annealed in 1X annealing buffer (20 mM HEPES, pH 7.5, 5 mM Mg(OAc)2) by heating at 95 °C for 2 min, and then at 50 °C for 60 min. Reactions were carried out in 1X pol V reaction buffer (Karata 2012; Vaisman 2012) with 100 nM DNA, with the following final concentrations of proteins: 400 nM Pol V (WT or H282P or T412I variants), 2.1 nM RecA nucleoprotein filament (RecA*), 8.4 μM ATPγS, and 0.84 nM SSB (Affymetrix). Reactions were incubated at 37 °C for 10 min, and then initiated with 200 μM dNTPs. Samples were taken at 0, 20, and 60 min. RecA* (2.1 nM) was made as previously described (Vaisman 2012) by mixing 33.7 nM RecA (Affymetrix), 8.4 μM ATPγS, 2.1 nM 48-mer ssDNA oligo in 1X pol V reaction buffer. For reactions containing DinB, no cofactor proteins were used. DinB reactions contained 100 nM DNA and 10 nM wild-type DinB in 1X DinB reaction buffer (30 mM HEPES pH 7.5, 20 mM NaCl, 7.5 mM MgSO4, 2 mM β-mercaptoethanol, 1% BSA, 4% glycerol) and were initiated with 200 μM dNTPs. Reactions were quenched in 22 μL loading buffer (85% formamide, 50 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue). Time points were then analyzed on denaturing 8 M urea 16% polyacrylamide gels, and
were imaged on a Molecular Dynamics 860 storage phosphorimager. Data were analyzed using ImageQuant TL 1D Gel Analysis software (GE Healthcare).

3.3 Results

3.3.1 Identification of UmuC variants that confer resistance to nitrofurazone in vivo

We identified variants of the polymerase subunit of *E. coli* DNA polymerase V, UmuC, that conferred resistance to NFZ, which is considered to be a cognate lesion of DinB and a non-cognate lesion of UmuC (Jarosz 2006; Walsh 2011) (Figure 3.1A). By using hydroxylamine, we created random mutations in the *umuC* gene and selected colonies of AB1157 ΔdinB harboring the library of pol V variants that survived increasing concentrations of NFZ. From our selection, we found individual point mutations in UmuC that confer resistance to NFZ on a ΔdinB strain (Table 3.2).
Figure 3.1 *E. coli* Y family DNA polymerases UmuC (the polymerase component of pol V UmuD’C) and pol IV, DinB, have specificity for distinct types of DNA lesions. (A) UmuC’s cognate lesions are the UV photoproducts thymine-thymine cyclobutane pyrimidine dimers and 6-4 thymine-thymine dimers. DinB’s cognate lesions are N²-dG adducts, including N²-furfuryl-dG. Our goal was to change the specificity of UmuC to that of DinB as depicted by the dotted arrow. (B) Homology model of UmuC (Beuning 2006a) depicting one of the residues we discovered in our selection assay. Due to the lack of homology of the UmuC C-terminal domain (grey) to other proteins of known structure, residues 354-422 could not be modeled. The C-terminus of the model, residue Q353, is highlighted in green. Residues H282 (shown in purple) and T412 (not shown on model), found to be mutated to H282P and T412I in our selection experiment, enabled UmuC to confer resistance to NFZ, which is believed to cause predominantly N²-furfuryl-dG. The backbone of UmuC is shown in yellow, and DNA is rendered as sticks and colored by atom identity. This image was rendered by using VMD (Humphrey 1996).
We confirmed the NFZ phenotype by reconstructing the UmuC variants in plasmids harboring the *umuD'C* genes by site-directed mutagenesis and testing for NFZ resistance of strains harboring these plasmids. Cells harboring UmuC variants H282P and T412I confer resistance to NFZ in a *dinB*-deficient background (AB1157 *ΔdinB*) to a similar extent as wild-type DinB at a concentration of 1.5 µg/mL NFZ in this strain (Figure 3.2).

Since there is currently no crystal structure available for UmuC, we used a homology model based on the structure of Dpo4 to analyze the locations of the selected variants (Beuning 2006a). H282 is predicted to be located in the little finger domain on a short loop between and α-helix L and β-strand 10 and somewhat near the linker that joins the polymerase domain to the little finger domain, which is suggested to play a major role in specificity (Boudsocq 2002; Wilson 2013). T412I is omitted from the model because...
the ~70-residue C-terminal domain of UmuC where T412 is located has no homology to proteins of known structure and therefore cannot be modeled (Beuning 2006a).

3.3.2 UmuC variants H282P and T412I weakly bypass $N^2$-furfuryl-dG in vitro

The lesion $N^2$-furfuryl-dG is considered to be the cognate lesion of DinB, as DinB bypasses it accurately and efficiently in vitro (Jarosz 2006; Jarosz 2009; Walsh 2012). It has been suggested that NFZ causes $N^2$-furfuryl-dG lesions; therefore, since we selected for UmuC variants H282P and T412I that confer resistance to NFZ, we tested those variants for their ability to bypass $N^2$-furfuryl-dG in vitro by utilizing a primer extension assay (Figure 3.3). We discovered that both UmuC H282P and T412I bypass $N^2$-furfuryl-dG, though they do so weakly. However, both UmuC H282P and T412I were able to further extend a primer annealed to undamaged DNA (Figure 3.3). UmuC H282P extended the primer almost to the end of the undamaged template, while T412I extended undamaged DNA weakly. As a control, we included wild-type pol V, which does not bypass $N^2$-furfuryl-dG, but is active on undamaged DNA with full extension of primers to the end of the template (Figure 3.3). Wild-type DinB bypasses $N^2$-furfuryl-dG and extends primers to the end of the template, while showing weaker activity on undamaged DNA, which is consistent with previous work (Jarosz 2006) .
3.3.3 UmuC variants H282P and T412I enhance cleavage of UmuD to UmuD’ in vivo

In order for UmuC to become active for TLS, it must combine with the cleaved form of UmuD2, which is UmuD’2, to form pol V (Woodgate 1989; Reuven 1999; Tang 1999). A western blot of UmuD and UmuD’ expressed from the plasmid pGY9739 expressing *umuDC*, which harbored the UmuC variants, allows for the measurement of the extent of cleavage of UmuD to UmuD’ in the presence of UmuC variants *in vivo*. We
determined the levels of UmuD and UmuD' in the presence of UmuC variants H282P and T412I (Figure 3.4A). Over a time period of 120 minutes, the percentage of UmuD' relative to the total amount of umuD gene products (uncleaved UmuD and cleaved UmuD') was nearly twice as high in the presence of UmuC variants H282P or T412I in comparison to wild-type UmuC (Figure 3.4A). For comparison, we also probed the amount of UmuD' relative to UmuD when UmuC was absent (pGYDAC), and observed that the relative percentage of UmuD' to total UmuD and UmuD' was over twice as high in comparison to the that in the presence of wild-type UmuC, suggesting that UmuD undergoes greater self-cleavage in the absence of binding partner UmuC. We speculate that in the absence of binding partner UmuC, more UmuD is available to interact with the RecA nucleoprotein filament, leading to an increase in UmuD cleavage. We believe that this is the first example of mutations in UmuC that substantially affect the cleavage efficiency of UmuD to UmuD'.

Figure 3.4 Immunoblot analysis of UmuC and UmuD/D’ steady-state levels show that UmuC H282P and T412I are degraded by Lon protease (B,C). (A) UmuD was expressed from pGY9739 plasmids encoding *umuDC* genes in GW8017. The wild-type plasmid was pGY9739, and the empty vector was pGB2. Each percentage of cleavage is a measurement of the amount of UmuD’ in comparison to the total amount of UmuD and UmuD’. (B,C) UmuC was expressed from plasmids encoding *umuDC* genes in GW8017 (B) and GW8017 Δlon (C). The intensity of each UmuC band is normalized to wild-type UmuC from the respective strain, expressed as a percentage. The wild-type plasmid was pGY9739, and the empty vector was pGB2.

3.3.4 UmuC H282P and T412I steady state levels are diminished in the presence of Lon protease

We suspected that the enhanced cleavage of UmuD we observed in the presence of UmuC variants H282P and T412I (Figure 3.4A) may be due to a decrease in affinity between UmuD or UmuD’ and UmuC, which we predict may also result in a decrease in the protein levels of UmuC. Therefore, we probed the steady-state expression levels of UmuC protein for UmuC variants H282P and T412I in the presence (Figure 3.4B) and absence (Figure 3.4C) of Lon protease. A strain deficient in *umuDC* expressing the
selected UmuC variants was exposed to 25 J/m² UV radiation to initiate the SOS response and induce expression of the *umuDC* gene products, and samples were taken at 30, 60, and 120 min recovery after UV treatment. The steady state levels of UmuC H282P were dramatically decreased even at 30 minutes after exposure to UV radiation, while UmuC T412I steady state levels were modestly decreased, relative to the levels of wild-type UmuC (Figure 3.4B). Lon protease degrades UmuC at the end of the SOS response, and could be responsible for decreasing levels of UmuC, and therefore indirectly increasing the extent of cleavage of UmuD to UmuD′ (Donnelly 1992; Frank 1996a; Frank 1996b). We therefore constructed a derivative of GW8017 that lacks *lon* and determined the steady-state levels of UmuC H282P and T412I in this strain. We found that in the absence of *lon*, the levels of these UmuC variants were substantially higher (Figure 3.4C).

### 3.3.5 UmuC H282P and T412I fail to confer the *umuDC*-specific cold sensitive phenotype

Overexpression of the *umuDC* gene products leads to a cold sensitive phenotype, as cells harboring elevated levels of these gene products cannot grow or grow slowly at the non-permissive temperature of 30 °C (Marsh 1985; Opperman 1999; Sutton 1999; Sutton 2001b). This phenotype is thought to be a manifestation of a primitive DNA damage checkpoint, allowing time for error-free mechanisms of DNA repair such as RecA-mediated homologous recombination or nucleotide excision repair to take place (Opperman 1999; Sutton 2001b). The pGY9738 plasmid, expressing the cleaved form UmuD′, does not cause a cold sensitive phenotype. Plasmid pGY9739 expressing the UmuC variants H282P and T412I failed to confer the cold sensitive phenotype (Figure 3.
5), which could be because specific interactions between UmuD and UmuC that are required for the cold sensitive phenotype are disrupted by the H282P or T412I mutations or because there may be more UmuD′ than UmuD in the cell in the presence of these UmuC variants.

Figure 3.5 UmuC variants H282P and T412I do not confer a cold sensitive phenotype. When *umuDC* is overexpressed, it induces a cold sensitive phenotype (Sutton 2001b) for growth at 30°C. AB1157 was transformed with the respective plasmids and equal volumes of the transformation mixtures were grown on LB agar with appropriate antibiotics at the non-permissive (30°C) and permissive (37°C) temperatures.

3.3.6 RecA-mediated homologous recombination is inhibited by UmuC H282P and T412I

TLS by pol V is in competition with homologous recombination, an error-free repair pathway, for the RecA nucleoprotein filament, as RecA is necessary for pol V-mediated TLS, facilitates the cleavage of UmuD to UmuD′, and is required for repair by homologous recombination. The inhibition of RecA-mediated homologous recombination
by elevated levels of pol V is considered to be a mechanism of regulation of RecA (Sommer 1993; Szpilewska 1995; Boudsocq 1997; Rehrauer 1998; Sutton 2001a). Since we observed that the two selected UmuC variants seem to disrupt interactions with UmuD or UmuD', and since it is known that pol V inhibits RecA-mediated homologous recombination, we wanted to test if the selected UmuC variants also affect homologous recombination. UmuC variants H282P and T412I were expressed from plasmid pGY9738 in a ΔumuDC strain. These cells were exposed to bacteriophage P1vir ΔyeaB and plated on selective medium to measure the transductants that result from RecA-mediated homologous recombination. We observed that UmuC variants H282P and T412I inhibit RecA-mediated homologous recombination to a similar extent as wild-type pol V (Figure 3.6), suggesting that interactions required for pol V inhibition of homologous recombination are intact in these variants.

![Figure 3.6 UmuD'C inhibits RecA-mediated homologous recombination.](image)

**Figure 3.6 UmuD'C inhibits RecA-mediated homologous recombination.** Empty vector (pGB2), wild-type (pGY9738), and variants H282P and T412I expressed from plasmid pGY9738 (umuDC', wild-type) were used in strain GW8017. Transduction efficiency is measured in CFU/PFU as a percentage of that fraction for empty vector (9.26 X 10⁶ CFU/PFU, normalized to 100%): wild-type pol V, 5.09%; UmuC H282P, 7.97%; UmuC T412I, 20.4%.
3.3.7 UmuC variants H282P and T412I retain the ability to confer UV Survival and UV-induced mutagenesis

We have determined that both UmuC H282P and T412I alter the cleavage of UmuD to UmuD'. Since UmuD'$_2$ is essential for TLS by pol V, we wanted to test if these variants significantly alter the ability of pol V to confer resistance to UV light and to carry out UV-induced mutagenesis. We transformed a ΔumuDC ΔrecJ strain with plasmids expressing the UmuC variants H282P and T412I and exposed these cells to increasing amounts of UV radiation. It has been shown that ΔumuC ΔrecJ or ΔumuDC ΔrecJ strains are sensitive to UV radiation (Courcelle 2006; Shurtleff 2009; Hawver 2011; Ollivierre 2011; Ollivierre 2013); ΔumuDC ΔrecJ was used here. RecJ is an endonuclease that degrades replication forks in which replication is disrupted due to DNA damage (Courcelle 2006). Without RecJ, recovery of synthesis after UV-induced damage relies on TLS by pol V (Courcelle 2006). The UmuC H282P and T412I variants expressed from a plasmid also expressing synthetically cleaved UmuD' (pGY9738) (Figure 3.7A) or full-length UmuD (pGY9739) (Figure 3.7B) confer resistance to UV radiation.
We also wanted to determine if UmuC H282P and T412I retain the ability to perform UV-induced mutagenesis, the predominant phenotype of pol V. The *argE3* reversion frequency of Δ*umuDC* strains harboring plasmids expressing UmuC H282P or UmuC T412I was determined by exposure to 25 J/m² UV radiation. Cells expressing UmuC H282P and UmuC T412I showed a moderate to high mutation frequency, respectively (Figure 3.8). The lower mutation frequency of cells expressing UmuC H282P may be due to the severe degradation of this UmuC variant by Lon or because it may interact with UmuD' with less affinity. The high mutation frequency of UmuC T412I suggests that even though this variant may disrupt the interaction between UmuC and UmuD', which is required for UV-induced mutagenesis, the enhanced cleavage of UmuD to form UmuD' observed with this variant may actually compensate for this by increasing the cellular concentration of UmuD'. Though these UmuC variants seem to decrease the interaction between UmuC and UmuD', they must still interact somewhat due to the
observation that they are proficient for UV survival and UV mutagenesis (Figures 3.7 and 3.8) as well as the observation that UmuD’ still co-purifies with UmuC (Figure 3.9).

**Figure 3.8 UmuC variants H282P and T412I confer high mutation frequencies.** The UV-induced (25 J/m²) mutation frequency of GW8017 (ΔumuDC) expressing wild-type UmuC or variants UmuC-H282P or UmuC-T412I was determined compared to wild-type umuD’C (pGY9738) and the empty vector (pGB2). The mutation frequency is normalized to that conferred by wild-type pol V (100% mutation frequency). The mutation frequencies found were as follows: empty vector, 0.29%; wild-type (umuD’C) 100%; pGY9738-H282P, 60.3%; and pGY9738-T412I, 129%.

**Figure 3.9 Immunoblot of purified wild-type pol V and variants H282P and T412I.** Each lane contains 0.8 µg of total purified protein.
3.4 Discussion

We have shown that UmuC variants H282P and T412I confer resistance to nitrofurazone in vivo, and they do bypass the lesion believed to be caused by nitrofurazone, N²-furfuryl-dG, in vitro, though weakly. We wanted to characterize these variants further, so we tested the selected variants for several known UmuC and UmuD functions, and discovered that both UmuC H282P and T412I cause enhanced cleavage of UmuD to UmuD’ (Figures 3.4A and 3.10A). It is known that UmuD’ is required for UmuC to be active for TLS, however, the binding site on UmuC for UmuD’ has not yet been defined. Currently, there is no crystal structure for UmuC, so we must rely on a homology model to define the location of specific residues (Beuning 2006a). Moreover, the C-terminal domain of UmuC is unique to this protein and its orthologs and has no homology to proteins of known structure, so the C-terminal domain (residues 354-422) could not be modeled.

Several laboratories have identified possible UmuD’ binding sites on UmuC, often through the use of truncations of UmuC (Woodgate 1989; Jonczyk 1996; Sutton 2001b; Chandani 2013). Jonczyk and Nowicka determined by using truncations of UmuC in a yeast-two-hybrid experiment that (Jonczyk 1996) deletion of UmuC residues 1-13 or 397-422 eliminated the interaction between UmuC and UmuD’ (Jonczyk 1996). Sutton, et al. also determined by using truncations of UmuC that residues 397-422 are important for interaction with UmuD’ (Sutton 2001b). We have shown that the mutation T412I causes UmuD to be hypercleavable to UmuD’, as evidenced by immunoblotting. It seems likely that the UmuC T412I mutation weakens the interaction between UmuC and UmuD’; however, since UmuC T412I is still competent for UV-induced mutagenesis, which
requires UmuD′, and UmuD′ co-purifies with UmuC-T412I, clearly mutation of this one residue is not sufficient to completely eliminate the interaction. It has been recently suggested from results of docking studies using ZDOCK (Chen 2003) and ClusPro (Comeau 2004) that one monomer of UmuD′ interacts with UmuC residues 82, 90, 126-132, and the other UmuD′ monomer interacts with UmuC residues 89, 93, 94, and 239 (Figure 3.10B) (Chandani 2013). It is important to note that this docking study utilized the UmuC homology model, which lacks the C-terminal domain (residues 354-422), including T412.

The *umuC36* allele encoding the E75K mutation in UmuC greatly reduces UV-induced mutagenesis, which can be suppressed by overproduction of UmuD′ (Kato 1977; Bates 1991; Koch 1992; Woodgate 1994). It has been suggested that residues surrounding E75 may also be involved in the interaction of UmuC with UmuD or UmuD′, as the increasing the levels of UmuD′ restored mutagenesis (Woodgate 1989; Koch 1992). Figure 3.10B highlights residues of UmuC that are predicted to facilitate interaction with UmuD, showing that residue H282 is in close proximity to E75, as well as residues predicted to interact with UmuD′ from a recent protein docking study using ZDOCK (Chen 2003) and ClusPro (Comeau 2004; Chandani 2013).
Figure 3.10 Summary of findings from in vivo experiments of UmuC H282P and UmuC T412I and highlighted UmuC residues that have been predicted to interact with UmuD' and are important for protein stability. (A) A potential loss of interaction between UmuC H282P or UmuC T412I and UmuD causes UmuD to be hypercleavable, and UmuC to be degraded by Lon protease. (B) Homology model of UmuC (Beuning 2006a) depicting our variant H282P (purple) as well as several other residues predicted to interact with UmuD'. Due to the lack of homology of the UmuC C-terminal domain to other proteins of known structure, residues 354-422 could not be modeled. Residue H282 (purple) was shown to cause hypercleavability of UmuD to UmuD' in this work. Residues 82, 90, 126-132 (blue), and 89, 93, 94, and 239 (black) have been shown to interact with UmuD' in a recent molecular modeling study (Chandani 2013). Residue E75K (orange) is also thought to be important for UmuD' interaction (Koch 1992). Residue T290K (green) is thought to be important for UmuC stability (Woodgate 1994; Boudsocq 2002). The backbone of UmuC is shown in yellow, and DNA is rendered as sticks and colored by atom identity. This image was rendered by using VMD (Humphrey 1996).
UmuC H282P was shown here (Figure 3.4C) to be degraded by Lon protease more readily than wild-type UmuC (Figure 3.10A). It is known that Lon protease degrades UmuC if UmuC is not stabilized by UmuD’ or UmuD (Frank 1996a). Lon protease also degrades full-length UmuD (Frank 1996a). There have been two regions identified in UmuD that are important for Lon proteolysis, which are \(^{15}\text{FPLF}\) and \(^{26}\text{FPSP}\) (Gonzalez 1998). There are no similar sequences in UmuC; however, the mutation of H282 to proline causes significant degradation by Lon protease, and proline is present in both UmuD sequences required for Lon proteolysis. The sequence context of H282 in UmuC is \(^{280}\text{SEH(P)QYC}\), which is a poor match to the recognition sequences in UmuD. We speculate that mutating H282P could cause UmuC to become unstable, making this variant a better substrate for degradation, or that UmuC H282P could be targeted by Lon due to its weaker interactions with UmuD. Also, without UmuC present in adequate levels such as in the case of UmuC H282P, UmuD is free to interact with other binding partners such as the RecA nucleoprotein filament, explaining the appearance of more UmuD’ than is observed in the presence of wild-type UmuC (Figure 3.4A).

The \textit{umuC25} allele encodes the point mutation T290K in UmuC, and renders cells non-mutable (Steinborn 1978; Koch 1992). By comparison to the Dpo4 structure, it is predicted that T290 is located in a hydrophobic part of the UmuC little finger domain (Figure 3.10B) (Boudsocq 2002), and this mutation is expected to cause UmuC to become unstable (Boudsocq 2002), which likely would lead to degradation of the UmuC protein (Woodgate 1994). While our variants are not predicted to be in a similar
hydrophobic region, disruption of the little finger domain, such as the mutations H282P and T412I, could be enough to cause instability of UmuC.

UmuC is tightly regulated by the cleavage of UmuD, as well as by Lon protease, and by access to the RecA nucleoprotein filament. The UmuC mutations H282P and T412I identified in this work alter the protein interactions of UmuC. It is possible that disrupting important protein-protein interactions can alter the specificity of a polymerase, especially if changing these interactions affects the size or flexibility of the active site. Our work shows that specific protein interactions can be important for the specificity of cellular resistance to particular types of DNA damage.

3.5 References


Chapter 4: Point mutation in *Escherichia coli* DNA pol V active site that confers resistance to non-cognate DNA damage has altered requirements for cognate damage

4.1 Introduction

*E. coli* has five DNA polymerases, two of which are from the Y family of DNA polymerases (Ohmori 2001; Friedberg 2006). These specialized polymerases, found in all domains of life, have larger active sites than typical replicative polymerases such as polymerase III in *E. coli*. Due to their less constrained active sites, they can accommodate bulky DNA lesions that would otherwise disrupt replication. Thus, Y family polymerases participate in translesion synthesis (TLS), which is the process of bypassing DNA lesions.

The two Y family DNA polymerases found in *E. coli* are DinB (pol IV), which is a product of the *dinB* gene, and UmuD′_2C (pol V), which is a product of the *umuDC* genes. These genes are regulated by the SOS response to DNA damage, along with over 50 other genes (Friedberg 2006; Simmons 2008). The SOS response is initiated when single stranded DNA (ssDNA) downstream from a disrupted replication fork becomes coated with the RecA protein, forming a RecA-ssDNA nucleoprotein filament. The nucleoprotein filament is involved in many processes, one of which is facilitating the autoproteolytic cleavage of LexA, the suppressor of the SOS response, which in turn signals the expression of the SOS genes (Radman 1974; Friedberg 2006; Simmons 2008).

Since Y family polymerases are error-prone, they are highly regulated in the cell; in particular, Pol V is regulated in several different ways. One of the main modes of regulation is the cleavage of UmuC’s binding partner, UmuD, to UmuD′. Pol V is not
active for TLS until UmuC combines with UmuD′ to form UmuD′C (Reuven 1999; Tang 1999; Friedberg 2006). About 20 to 40 minutes after SOS induction, UmuD undergoes a self-cleavage reaction facilitated by the RecA-ssDNA nucleoprotein filament, in which the first 24 amino acids of UmuD are cleaved to form UmuD′ (Burckhardt 1988; Nohmi 1988; Shinagawa 1988). Furthermore, overexpression of umuDC inhibits DNA synthesis at the non-permissive temperature (30 °C), allowing for more accurate pathways of DNA repair to function prior to error-prone translesion synthesis (Marsh 1985; Sutton 2001b). Once UmuD is cleaved to UmuD′, pol V is active for TLS, a potentially error-prone pathway for DNA damage tolerance. Finally, the cellular levels of UmuC, UmuD and UmuD′ are managed by two proteases, Lon and ClpXP. The Lon protease is involved in degradation of UmuC and UmuD (Gonzalez 1998), whereas ClpXP degrades UmuD or UmuD′ bound to UmuD (Frank 1996; Gonzalez 2000; Neher 2003).

Along with UmuD and UmuD′, RecA is an important UmuC binding partner in forming the pol V mutasome (Schlacher 2005). In addition to the role of the RecA-ssDNA nucleoprotein filament in the autoproteolytic cleavage of UmuD to UmuD′, RecA also interacts with UmuC and facilitates translesion synthesis, though the exact mechanism is still not fully understood (Reuven 2001; Schlacher 2005; Schlacher 2006; Fujii 2009; Jiang 2009; Karata 2012). Cryo-electron microscopy shows that pol V binds to the end as well as within the grooves of the RecA-ssDNA nucleoprotein filament (Frank 2000). Our previous work suggested that a possible binding site for RecA on UmuC could be residues 32 through 34 (Hawver 2011), which was subsequently
predicted via a protein docking study using ZDOCK (Chen 2003) and ClusPro (Comeau 2004; Chandani 2013).

UmuD'\textsubscript{2}C bypasses thymine-thymine cyclobutane pyrimidine dimers (T-T CPDs) and 6-4 T-T photoproducts produced by exposure of DNA to UV radiation (Tang 2000). UmuD'\textsubscript{2}C bypasses T-T CPDs accurately by inserting two dA nucleotides opposite the lesion (Szekeres 1996; Tang 2000). UmuD'\textsubscript{2}C has a mutagenic signature, however, accurately inserting dA opposite the 5' thymine but mutagenically inserting dG opposite the 3' thymine of the 6-4 T-T photoproduct (Becherel 1999; Tang 2000). Another lesion that pol V bypasses inaccurately is C\textsuperscript{8}-dG-acetylaminofluorine, by adding dA opposite the lesion instead of dC (Becherel 1999; Fujii 2004; Fujii 2009). N\textsuperscript{2}-furfuryl-dG is considered to be the cognate lesion of DinB (Jarosz 2006; Jarosz 2009), as DinB bypasses this lesion, as well as other N\textsuperscript{2}-dG adducts such as N\textsuperscript{2}-benzo[a]pyrene-dG (Shen 2002; Seo 2009) and N\textsuperscript{2}-(1-carboxylethyl)-2'-dG (Yuan 2008) accurately by adding dC opposite the lesion.

Through a hydroxylamine mutagenesis screen, we discovered the UmuC point mutation A9V that allowed cells harboring pol V in the absence of pol IV to be resistant to NFZ. Surprisingly, we discovered that this mutation does not give pol V the ability to bypass N\textsuperscript{2}-furfuryl-dG, the predominant lesion believed to be caused by NFZ (Jarosz 2006) in biochemical assays. In an effort to understand the basis of resistance to NFZ of cells expressing UmuC A9V, we found evidence that this residue also modulates the activity of pol V via dependence of the length of the arm of UmuD.
4.2 Materials and Methods

See Chapter 3, Section 3.2 for detailed materials and methods.

4.2.1 Plasmids and strains

Plasmids and strains are listed in Table 4.1. UmuC variants were constructed in the pGY9738 and pGY9739 plasmids, which both encode resistance to spectinomycin (used at (60 µg/mL), for *in vivo* assays (Sommer 1998). The pGY9738 plasmid contains a synthetic *umuD′C* operon; whereas the pGY9739 plasmid contains the *umuDC* operon. The pGY9738-ΔG25 plasmid is a derivative of pGY9738, where the G25 residue of UmuD′ is deleted. This deletion was made using Quickchange or Quickchange Lightning site-directed mutagenesis kits (Agilent Technologies). The pHUC25 plasmid for purification of pol V encodes resistance to kanamycin (50 µg/mL) (Karata 2012). Strains used (Table 4.1) were grown in liquid Luria Broth at 37 °C unless otherwise noted. Competent cells were made using the CaCl₂ method (Sambrook 1989) and transformations were performed as described previously (Beuning 2006b). UmuC variant A9V was made in the pGY9738, pGY9738-ΔG25, pGY9739, and pHUC25 plasmids using Quickchange or Quickchange Lightning site-directed mutagenesis kits (Agilent Technologies). The presence of mutations was confirmed by DNA sequencing (Massachusetts General Hospital DNA Core Facility, Cambridge, MA; Macrogen, Cambridge, MA; or Eton Bioscience, Charlestown, MA).
Table 4.1. Strains and Plasmids

<table>
<thead>
<tr>
<th>Bacterial Strain</th>
<th>Relevant Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB1157</td>
<td>argE3</td>
<td>Laboratory stock</td>
</tr>
<tr>
<td>YG7207</td>
<td>AB1157 ΔdinB</td>
<td>(Kim 1997)</td>
</tr>
<tr>
<td>GW8017</td>
<td>AB1157 ΔumuDC</td>
<td>(Guzzo 1996)</td>
</tr>
<tr>
<td>PB103</td>
<td>AB1157 ΔumuDC ΔrecJ</td>
<td>(Ollivierre 2013)</td>
</tr>
<tr>
<td>RW644</td>
<td>BL21(λDE3), ΔpolB::Ωspec,ΔdinB61::ble,ΔumuDC596::ermGT</td>
<td>(Karata 2012)</td>
</tr>
</tbody>
</table>

Plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Genotype Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGY9738</td>
<td>o₁^{C}umuDC; pSC101-derived</td>
<td>(Sommer 1998)</td>
</tr>
<tr>
<td>pGY9738</td>
<td>o₁^{C}umuDC; pSC101-derived, umuD'ΔG25</td>
<td>This work</td>
</tr>
<tr>
<td>UmuD'ΔG25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGY9739</td>
<td>o₁^{C}umuDC; pSC101-derived</td>
<td>(Sommer 1998)</td>
</tr>
<tr>
<td>pGYDΔC</td>
<td>o₁^{C}umuD^{+},umuC; pSC101-derived</td>
<td>(Sutton 2001b)</td>
</tr>
<tr>
<td>pGB2</td>
<td>Vector, pSC101-derived</td>
<td>(Churchward 1984)</td>
</tr>
<tr>
<td>pYG768</td>
<td>dinB</td>
<td>(Kim 1997; Kim 2001)</td>
</tr>
<tr>
<td>pWSK30</td>
<td>Vector</td>
<td>(Kim 1997; Kim 2001)</td>
</tr>
<tr>
<td>pHUC25</td>
<td>pSC101 ori, laqF, His-umuC, umuD'</td>
<td>(Karata 2012)</td>
</tr>
<tr>
<td>pARAD2</td>
<td>pSC101 ori, umuD'</td>
<td>(Karata 2012)</td>
</tr>
</tbody>
</table>

4.2.2 Hydroxylamine mutagenesis and genetic selection of UmuC variants.

Hydroxylamine mutagenesis was carried out as previously described (see Chapter 3, Section 3.2.2). Two rounds of validation were conducted before positive hits were analyzed by sequencing. Variants of interest were analyzed further by plating AB1157 ΔdinB strains harboring the selected plasmids which were cultured on plates containing
0.0, 2.1, 2.4, and 2.8 µg/mL NFZ. Colonies were picked from plates containing 2.4 and
2.8 µg/mL NFZ and the plasmid DNA was isolated. These plasmids were then used to
transform AB1157 ΔdinB competent cells, which were then assayed for growth once
again on agar plates with higher NFZ concentrations (0, 2.6, 3, 3.5, 4, 4.7, 6 µg/mL
NFZ). From this process we obtained eight clones, which were then sequenced
(Massachusetts General Hospital DNA Core Facility, Cambridge, MA). Only six hits
were verified as variants of UmuC, one of which is the subject of this work (Table 4.2);
two of the clones did not contain mutations in the *umuC* gene. These six variants were
then reconstructed in both pGY9738 and pGY9739, which were the constructs used in
subsequent experiments. In this work, we focus on A9V; characterization of other
variants is reported in Chapter 3.

### Table 4.2. Nucleotide sequence analysis of *umuD'C* alleles containing missense
mutations

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Nucleotide Substitution(^1)</th>
<th>Deduced Amino Acid Substitution(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGY9739</td>
<td>C(^26) → T(^26)</td>
<td>Ala(^9) → Val(^9)</td>
</tr>
</tbody>
</table>

\(^1\)Nucleotide numbering is based on using the first nucleotide of UmuC as 1. \(^2\)Amino acid
numbering begins with 1 for the first amino acid of UmuC.
4.3 Results

4.3.1 Identification of UmuC variant A9V that confers resistance to nitrofurazone in vivo

Using hydroxylamine mutagenesis, we identified variants of *E. coli* UmuC that conferred resistance to nitrofurazone (NFZ), which predominantly causes the $N^2$-furfuryl-dG lesion, a cognate lesion of DinB (Jarosz 2006) but a non-cognate lesion of UmuC. From our selection, we found a mutation in UmuC, A9V, that confers resistance to NFZ on cells lacking the *dinB* gene (Figure 4.2). This mutation was reconstructed in a fresh plasmid (pGY9738), and it was confirmed that cells harboring UmuC variant A9V are resistant to nitrofurazone at a concentration of 1.5 µg/mL in a *dinB* deficient background (AB1157 Δ*dinB*). This residue is predicted to be located in the active site of UmuC, adjacent to the 5’ phosphates of the incoming nucleotide, as shown in the UmuC homology model (Figure 4.1A) (Beuning 2006a).
Figure 4.1 *E. coli* UmuC homology model, based on the structure of Dpo4, depicting UmuC variants A9V and A39V (Beuning 2006a). (A) UmuC homology model depicting residue A9 (purple) in the active site, just above the 3’ phosphates of the incoming nucleotide. The backbone of UmuC is shown in yellow, and DNA is rendered as sticks and colored by atom identity. This image was prepared using the VMD software (Humphrey 1996). (B) UmuC homology model with UmuC A9 (purple) and A39 (green) located near the incoming nucleotide.
Figure 4.2 Cells harboring UmuC variant A9V confer resistance to NFZ in vivo. Assays were performed with the pWSK30 plasmid (DinB EV), pGY768 plasmid (DinB WT), and pGY9738 plasmid with A9V derivative in the strain AB1157 ΔdinB. Cells were exposed to 1.5 µg/mL NFZ in LB agar plates with appropriate antibiotics. Shown is the structure of $N^2$-furfuryl-dG, the predominate lesion believed to be caused by NFZ (Jarosz 2006).

4.3.2 UmuC variant A9V does not bypass $N^2$-furfuryl-dG in vitro

$N^2$-furfuryl-dG is the cognate lesion of DinB, which DinB bypasses accurately and efficiently in vitro (Jarosz 2006; Jarosz 2009; Walsh 2012). We tested the bypass of $N^2$-furfuryl-dG by pol V with the UmuC variant A9V in a primer extension assay. Pol V containing UmuC A9V was unable to bypass $N^2$-furfuryl-dG in vitro, but it was able to extend undamaged DNA, though not to the end of the template (Figure 4.3). Wild-type pol V does not bypass $N^2$-furfuryl-dG, but it is active on undamaged DNA with full extension to the end of the template (Figure 4.3). Wild-type DinB bypasses $N^2$-furfuryl-dG and weakly extends undamaged DNA, which has been shown previously (Walsh 2012). Thus, the UmuC A9V variant confers resistance to NFZ in vivo, but does not bypass $N^2$-furfuryl-dG in vitro.
Figure 4.3 Primer extension assay with a DNA template containing $N^2$-furfuryl-dG, as well as an undamaged DNA template. Assays were carried out with $^{32}$P-labeled standing start primer, and extension was visualized by phosphorimaging. Pol V reactions contain 400 nM enzyme, cofactors 2.1 nM RecA*, 8.4 µM ATPγS, and 0.84 nM SSB, as well as 100 nM DNA, and 200 µM dNTPs. DinB reactions contain 10 nM enzyme, 100 nM DNA, and 200 µM dNTPs. Time points were 0, 20, 60 min for each reaction. All reactions were analyzed by 16% polyacrylamide gel electrophoresis.

4.3.3 UmuC variant A9V confers resistance to UV radiation yet cells harboring this variant exhibit low UV-induced mutation frequency

We tested UmuC A9V for the ability to confer resistance to UV radiation by exposing a ΔumuDC ΔrecJ strain harboring our variant expressed from pGY9739 (umuDC) and pGY9738 (umuD'C) to increasing amounts of UV radiation. The ΔumuC ΔrecJ strain is sensitive to UV radiation (Courcelle 2006; Shurtleff 2009; Hawver 2011; Ollivierre 2011), because the deletion of recJ results in dependence pol V and TLS to
facilitate recovery of synthesis at a disrupted replication fork (Courcelle 2006). The umuD gene was also deleted from the chromosome for our assays as we suspect that UmuD, or more specifically the cleaved form UmuD', may play a role in the differences in phenotypes that we have observed. UmuC A9V confers resistance to UV radiation to a similar extent as wild-type pol V when expressed from pGY9738 (umuD'C) (Figure 4.4A) or pGY9739 (umuDC) (Figure 4.4B).

**Figure 4.4 UmuC variant A9V confers resistance to UV radiation.** (A) Assays were carried out in the AB1157 ΔumuDC ΔrecJ strain with the pGY9738 (umuD'C) plasmid and the following derivatives: pGY9738 (umuD'C wild-type; ■), pGB2 (empty vector; ♦), and pGY9738-A9V (umuD'C A9V; ●). (B) Assays were carried out in the AB1157 ΔumuDC ΔrecJ strain with the pGY9739 (umuDC) plasmid and the following derivatives: pGY9739 (umuDC wild-type; ▲), pGB2 (empty vector; ♦), and pGY9739-A9V (umuDC A9V; ●).

Similarly, we used an argE3 reversion assay to determine the UV-induced mutability of ΔumuDC cells harboring UmuC A9V expressed from pGY9738 or pGY9739. The mutation frequency of cells expressing UmuC A9V from pGY9739 (umuDC) was about two-fold lower than that of cells expressing wild-type UmuC (Figure 4.5A). Surprisingly, the mutation frequency of cells expressing UmuC A9V from pGY9738 (umuD'C) was very low (Figure 4.5A). This was unexpected as pGY9739 expressed full-length UmuD initially, which is then cleaved to form UmuD'; whereas pGY9738 expresses cleaved UmuD' directly. We would expect a defect in UV-induced
mutagenesis in the context of non-cleavable UmuD, but not synthetically cleaved UmuD'.

The pGY9738 (umuD'C) plasmid expresses a synthetically cleaved UmuD', with one extra methionine residue (M24) present to allow gene expression (Figure 4.5B). To determine if this M24-containing UmuD' arm interfered with pol V-dependent mutagenesis, we deleted UmuD' residue G25 in order to shorten the length of the arm to its natural length while maintaining an N-terminal Met residue (Figure 4.5B).

Remarkably, with the shortened UmuD' arm, the mutation frequency of UmuC A9V increased approximately 4.5-fold, to a level similar to that of its corresponding “wild-type” (Figure 4.5A). We then determined the steady state levels for wild-type UmuC and the A9V variant expressed from pGY9738 (umuD'C) as well as from pGY9738-ΔG25, which is defined as the synthetically cleaved umuD'C with the corrected UmuD' arm length (umuD'C UmuD' ΔG25). The immunoblot showed that over a time period of two hours, umuD'C (pGY9738)-A9V UmuC steady state levels decrease significantly. However, the steady state levels of A9V expressed from umuD'C UmuD' ΔG25 remain relatively high (Figure 4.5C).
Figure 4.5. UmuC variant A9V when expressed from the umuD'C plasmid confers a very low mutation frequency. (A) The UV-induced (25 J/m²) mutation frequencies in strain GW8017 (ΔumuDC) expressing wild-type UmuC and variant A9V were determined. The empty vector was pGB2. The mutation frequencies were normalized to wild-type (umuDC′) and determined as follows: empty vector, 0.38%; wild-type (umuD'C) 100%; pGY9738-A9V 1.89%; wild-type (umuDC) 110%; pGY9739-A9V 58.5%; pGY9738 UmuD' ΔG25 wild-type 70.4%; UmuD' ΔG25 A9V 45.7%. (B) Diagram of the natural cleavage of UmuD to UmuD' (plasmid pGY9739) and the synthetically cleaved UmuD' (plasmid pGY9738). We deleted residue G25 from UmuD' in the pGY9738 plasmid in order to make the synthetically cleaved UmuD' arm length the correct number of amino acids. (C) Immunoblot analysis of the steady-state levels of UmuC expressed from different plasmids over a time course up to 120 minutes. UmuC was expressed from plasmids pGB2 (EV), pGY9738 (umuD'C), and pGY9738 UmuD' ΔG25 in GW8017 (ΔumuDC).

Due to the observation that cells harboring pGY9738 (umuD'C)-A9V as well as pGY9739 (umuDC)-A9V survived UV radiation in a strain that is ΔumuDC ΔrecJ, but ΔumuDC cells harboring the pGY9738-UmuC A9V plasmid were non-mutable, we wondered if the presence of recJ plays a role in the decrease of the mutation frequency of cells harboring plasmids expressing UmuC A9V. We utilized the argE3 reversion assay, exposing a ΔumuDC ΔrecJ strain harboring plasmids expressing UmuC A9V to 25 J/m² UV radiation to determine the mutation frequency. We observed that UmuC A9V, when
expressed from pGY9738 (umuD'C), still maintains a low mutation frequency (Figure 4.6A). The mutation frequency of A9V expressed from pGY9738-ΔG25 remains the same at around 50%, while the mutation frequency of pGY9738-ΔG25 but with wild-type UmuC has decreased two-fold (Figure 4.6A). It appears that deletion of recJ did not substantially alter the mutation frequency of cells expressing UmuC A9V, however, it decreased the mutation frequency of cells harboring pGY9738-ΔG25 alone. Steady-state levels of UmuC A9V expressed from pGY9738 increased in this strain, but still remained low in comparison to steady state levels of wild-type UmuC (Figure 4.6B).

Figure 4.6 UmuC variant A9V when expressed from the umuD'C plasmid has a very low mutation frequency. (A) The UV-induced (25 J/m2) mutation frequencies of strain PB103 (ΔumuDC ΔrecJ) expressing wild-type UmuC and variant A9V were found. The mutation frequencies found were as follows: empty vector (pGB2), 3.1%; wild-type pGY9738 (umuD'C) 100%; pGY9738-A9V 15.8%; wild-type pGY9739 (umuDC) 132.5%; pGY9739-A9V 151.6%; pGY9738-UmuD' ΔG25 41.4%; pGY9738-UmuD' ΔG25 A9V 53.3%. (B) Immunoblot analysis of the steady-state levels of UmuC in various plasmids over a time course up to 120 minutes. UmuC was expressed from plasmids pGY9738 (umuD'C), and pGY9738 UmuD' ΔG25 in GW8017 (ΔumuDC); the empty vector was pGB2 (EV).
When DNA is exposed to UV light, approximately 75% of the resulting lesions are thymine-thymine cyclobuane pyrimidine dimers (CPDs), while about 25% of the lesions are 6-4 photoproducts (Friedberg 2006). Pol V bypasses CPDs accurately, while it bypasses 6-4 photoproducts mutagenically (Szekeres 1996; Tang 2000). Our observations indicate that the length of the UmuD′ arm may control the ability of UmuC A9V to participate in mutagenic translesion synthesis, while it may not affect the ability to bypass lesions accurately. The exact sites of interaction between UmuC and UmuD′ is not yet known. In the context of wild-type UmuC, deletion of UmuD′ G25 has no effect on mutagenesis. However, in the context of UmuC A9V, deletion of UmuD′ G25 has a dramatic effect on mutagenesis, suggesting that in the context of UmuC with a modified active site, the length of the UmuD′ arm is critical.

4.3.4 UmuC A9V fails to confer the umuDC-specific cold sensitive phenotype

Overexpression of the umuDC gene products produces a cold sensitive phenotype for growth, growing poorly at 30 °C, while overexpression of umuD′C does not (Marsh 1985; Opperman 1996; Sutton 2001b). Full-length UmuD combined with UmuC regulates DNA synthesis in response to DNA damage in a primitive DNA damage checkpoint (Opperman 1999; Sutton 2001b), allowing more accurate pathways of DNA repair such as homologous recombination and nucleotide excision repair to remedy the damage (Marsh 1985; Murli 2000). Once the cleavage of UmuD to UmuD′ occurs, pol V (UmuD′C) is active for translesion synthesis, which is a more error prone method of DNA damage tolerance (Reuven 1999; Tang 1999). The pGY9738 (umuD′C) plasmid, with artificially cleaved UmuD′ (Figure 4.4B), does not confer a cold sensitive phenotype, as UmuD is already in its cleaved form, UmuD′ (Figure 4.7). We tested the
AB1157 strain harboring *umuDC* (pGY9739) -A9V and discovered that this variant fails to confer the cold sensitive phenotype (Figure 4.7), indicating that the presence of this mutation changes the regulation of UmuC. As expected, cells harboring pGY9738-ΔG25 did not display the cold sensitive phenotype. Likewise, cells expressing UmuC A9V from pGY9738-ΔG25 also fail to exhibit the cold sensitive phenotype (Figure 4.7).

![Figure 4.7 UmuC variant A9V fails to complement the cold sensitivity phenotype.](image)

**Figure 4.7 UmuC variant A9V fails to complement the cold sensitivity phenotype.** Elevated levels of UmuDC cause a cold sensitive phenotype for growth at 30°C. The ratio of colony forming units (cfu) at 37 °C versus 30 °C for AB1157 harboring each of the selected plasmids is shown.

### 4.3.5 RecA-mediated homologous recombination is not inhibited by UmuC A9V

UmuD'C inhibits RecA-mediated homologous recombination by competing with the error-free repair pathway, homologous recombination, for access to the RecA nucleoprotein filament (Sommer 1993; Szpilewska 1995; Boudsocq 1997; Rehrauer 1998; Sommer 2000; Sutton 2001a). We wanted to determine whether UmuC A9V could inhibit RecA-mediated homologous recombination, particularly because UmuC F10L, UmuC N32A, UmuC N33A, and UmuC D34A do, and they are located in close
proximity to A9. GW8017 cells (ΔumuDC) expressing UmuC variant A9V from pGY9738 (umuDC') were exposed to bacteriophage P1vir ΔyeaB and plated on selective medium. We then measured the transductants that result from RecA-mediated homologous recombination. UmuC A9V does not inhibit RecA-mediated homologous recombination, and is within standard deviation of empty vector (pGB2), which contains no umuDC' (Figure 4.8). The significant decrease in UV mutagenesis in cells harboring the UmuC A9V variant we observe could be because homologous recombination could repair many of the lesions resulting from UV damage in this context. By analogy, in the context of UmuC A9V, lesions caused by NFZ could be repaired by homologous recombination, which could explain the differences in in vivo NFZ resistance conferred by UmuC A9V versus the lack of bypass of N²-furfuryl-dG by pol V A9V.

Figure 4.8 UmuD'2C (pol V) inhibits RecA-mediated homologous recombination. Strain GW8017 (ΔumuDC) harbored the empty vector (pGB2), wild-type (pGY9738), or indicated plasmids expressing the UmuC variant A9V. Transduction efficiency is measured in CFU/PFU as a percentage of that fraction for empty vector (1.63 X 10⁻⁴ CFU/PFU, normalized to 100%); wild-type pol V (D'C), 8.82%; D'C A9V, 121%; pGY9738 UmuD'ΔG25, 8.91%; pGY9738 UmuD'ΔG25-A9V, 46.0 %.
4.4 Discussion

We have shown that cells harboring UmuC variant A9V confer resistance to nitrofurazone \textit{in vivo}, but fail to bypass $N^2$-fururyl-dG \textit{in vitro}. In order to resolve this apparent discrepancy, we tested several known phenotypes of UmuDC and pol V. Surprisingly, we observed that the correct length but not the correct sequence of the UmuD’ arm is critically important for UV-induced mutagenesis in cells harboring UmuC A9V, but the one extra residue in the synthetic UmuD’ arm makes no difference in cellular UV resistance. When we changed the length of the UmuD’ arm to the naturally occurring length, the mutation frequency rose six-fold, and the steady-state levels of UmuC A9V increased. We also determined that the deletion of \textit{recJ}, which is used in UV survival experiments, made no difference in the mutation frequency of cells expressing UmuC A9V nor the steady-state levels of A9V when expressed from pGY9738 (\textit{umuD’C}).

Several other pol V-specific phenotypes rely on the particular form of UmuD, including cold sensitivity. We observe here that the UmuC A9V mutation fails to confer the cold sensitivity phenotype, perhaps by altering the way UmuD interacts with UmuC. In a previous hydroxylamine mutagenesis selection Marsh \textit{et. al} discovered the mutation A39V (\textit{umuC125}) in UmuC, which failed to confer the cold sensitivity phenotype. The A39V variant also failed to complement the UV mutagenesis phenotype and failed to confer resistance to UV radiation (Marsh 1991; Sutton 2001b). The A39 residue is predicted to be in a similar location to A9 according to the UmuC homology model (Figure 4.1B). Both A9V and A39V are alanine to valine mutations, and are predicted to be located adjacent to the incoming nucleotide (Figure 4.1B). While the mutation
described here, A9V, produced a low mutation frequency in the context of a synthetic
UmuD′ construct, we determined that restoring the length of the UmuD′ arm to its natural
length restored the mutation frequency to a level that is comparable to that of wild-type
UmuDC. The location of these variants within the protein suggests that both A39 and A9
participate in shaping the active site and may play a role in positioning the incoming
nucleotide. Both of these residues appear important for mutagenesis by pol V, but
mutagenesis by the A9V variant is dependent on the correct length of the UmuD′ arm.
This raises the possibility that the arm of UmuD′ modulated the UmuC active site.

The *umuC36* allele encodes the mutation E75K in UmuC (Koch 1992). This
mutation also causes cells to be non-mutable, unless *umuD′* is overexpressed, leading to a
partial restoration of mutagenesis (Bates 1991). It has been suggested that the E75K
mutation disrupts the interaction between UmuC and UmuD′, perhaps due to the
predicted location of this residue on the surface of the protein (Bates 1991; Boudsocq
2002). Mutations in *groE* decrease the mutation frequency of a strain expressing wild-
type *umuDC* (Donnelly 1989), unless *umuD′* but not *umuD* is coexpressed with *umuC*
(Donnelly 1992). GroEL is a chaperone protein that interacts directly with UmuC, and is
said to be required for UV mutagenesis and for stability, including protein folding
(Donnelly 1992).

Other variants made in the active site of pol V fail to complement the UV
mutagenesis phenotype. Residues F10 and the steric gate Y11 are adjacent to residue A9.
When mutated to alanine, UmuC variants F10A and Y11A were defective in UV
mutagenesis. Moreover, these two variants conferred hypersensitivity to UV radiation
(Shurtleff 2009), whereas cells expressing UmuC A9V survived UV radiation. UV
hypersensitivity was alleviated when the beta clamp binding sites on UmuC were mutated to alanine, suggesting that the effects these residues have on UV survival is due to recruitment to the replication fork (Shurtleff 2009). This further suggests that the role of F10 and Y11 in tolerating UV-induced lesions is fundamentally different than the role of A9.

Mutations of residues N32, N33, and D34, located in active site loop 1 in UmuC, also conferred hypersensitivity to UV radiation (Hawver 2011), similar to Y11A and F10A, and cells expressing UmuC variants N32A, N33A, or D34A also showed a significant decrease in mutation frequency (Hawver 2011). Interestingly, cells harboring N32A and D34A conferred resistance to NFZ at a concentration of 1.5 µg/mL (Hawver 2011). Variants N32A, N33A, and D34A alleviated the inhibition of RecA-mediated homologous recombination, suggesting that these residues may not be directly involved with specificity, however, they may be involved in the interaction of UmuC with RecA (Hawver 2011). Sommer et al. discovered several UmuC mutants in a genetic selection, including F10L which is adjacent to A9 in the active site, that actually recovers the ability to inhibit RecA-mediated homologous recombination (Sommer 2000). The A9V variant also alleviates the inhibition of RecA-mediated homologous recombination, though its location is predicted to be buried in the active site and not on the surface of UmuC making A9 an unlikely candidate for a direct protein-protein interaction with a globular protein. F10 is also predicted to be buried in the active site, but it has been predicted to be important not for a protein-protein interaction, but for DNA binding (Boudsocq 2002). It is also important to note that the effects that F10A, Y11A, N32A, N33A, and D34A had on pol V phenotypes were not due to any decrease in steady state
levels (Shurtleff 2009; Hawver 2011). The steady-state levels of UmuC variant A9V decrease rapidly when expressed from the pGY9738 (umuD’C) plasmid that also expresses synthetically cleaved UmuD’, while steady-state levels of A9V when expressed from umuD’C UmuD’ ΔG25 are higher overall. More importantly, the effect that A9V has on UV mutagenesis is related to the length of the UmuD’ arm, suggesting that this variant, and its role in UV mutagenesis, could be regulated by UmuD and its cleavage to UmuD’.

This work reports our characterization of a point mutation in UmuC that was found in a genetic selection to identify variants of UmuC that confer resistance to NFZ, which is considered to be a cognate lesion of DinB. Surprisingly, we found that although UmuC A9V conferred resistance to NFZ, pol V A9V was unable to bypass N²-fufuryl-dG. We determined the ability of UmuC A9V to carry out UV-induced mutagenesis and found that this activity depends almost entirely on the correct length, but not necessarily the correct sequence, of UmuD’. Our findings support the idea that specificity and function of pol V does not rely only on the active site itself, but also on specific and intact protein-protein interactions.

4.5 References


Chapter 5: Conclusions and Future Directions

Polymerase V is an *E. coli* Y family DNA polymerase that specializes in bypassing UV lesions such as T-T cyclobutane pyrimidine dimers and 6-4 photoproducts (Tang 2000). In my dissertation research, I have uncovered many important residues for pol V function and specificity. Based on homology modeling (Beuning 2006) and homology searches, I focused on two loops in the active site of UmuC, $^{31}$SNNDGCVI$^{38}$ and $^{50}$KMGDP$^{54}$, that were predicted to be important for DNA binding and interaction with the incoming nucleotide (Chapter 2) (Hawver 2011). Loop 1 residues N32, N33, and D34, when mutated to alanine, confer UV hypersensitivity as well as low mutation frequencies, even in a wild-type background (Hawver 2011). I also learned that the UV hypersensitivity conferred by N32A, N33A, and D34A is not solely related to the recruitment of pol V to the replication fork, as mutating the $\beta$ binding sites on UmuC in association with these point mutations did not alleviate the UV hypersensitivity. In a genetic transduction experiment, I tested for changes in the pol V phenotype of inhibition of RecA-mediated homologous recombination (Sommer 1993; Szpilewska 1995; Boudsocq 1997; Rehrauer 1998) and discovered that N32A, N33A, and D34A exhibit intermediate inhibition of RecA-mediated homologous recombination, suggesting that these residues may play a role in interacting with RecA, which has been supported by a recent molecular modeling study (Hawver 2011; Chandani 2013).

I then characterized three UmuC variants found to confer resistance to NFZ; H282P, T412I (Chapter 3), and A9V (Chapter 4). Pol V does not normally confer resistance to NFZ; NFZ is thought to cause predominantly $N^2$-furfuryl-dG lesions, the cognate lesion of DinB (Jarosz 2006). I showed that H282P and T412I bypass $N^2$-
furfuryl-dG weakly, while A9V does not bypass $N^2$-furfuryl-dG at all. All three variants are active on undamaged DNA. In further biochemical characterization, I discovered that UmuC variants H282P and T412I facilitate the hypercleavability of UmuD to UmuD$'$. The steady-state levels of these variants are also diminished in the presence of Lon protease, in comparison to wild-type levels. UmuC variant A9V has altered requirements for cognate lesions, such as UV-induced lesions. I found that the length of the UmuD$'$ arm is critical for UV-induced mutagenesis in the context of the A9V mutation. Also, UmuC A9V does not inhibit RecA-mediated homologous recombination, indicating that perhaps RecA-mediated homologous recombination can repair lesions caused by NFZ in this context, explaining our observation that UmuC A9V confers resistance to NFZ but pol V containing UmuC A9V does not bypass $N^2$-furfuryl-dG lesions.

I have used many \textit{in vivo} assays to characterize several important UmuC variants. More in-depth \textit{in vitro} characterization would be a significant benefit in the further understanding of these variants and how they alter protein-protein interactions. An efficient protocol for purifying pol V was recently published (Karata 2012). Until then, it was nearly impossible to purify pol V without either growing excessive volumes of culture only to yield little product (Bruck 1996; Tang 1998; Tang 1999), or tagging pol V with an MBP tag that is almost as large as UmuC itself (Reuven 1999). In order to better characterize UmuC, it would be helpful to be able to purify UmuC alone, which has not yet been done without significant drawbacks. Once UmuC is purified in sufficient yield, it could be utilized in assays such as FRET to determine how other proteins such as UmuD, UmuD$'$, and RecA interact with UmuC. It is possible currently to try to localize a binding site using truncations of UmuC in a bacterial 2-hybrid assay, which could also be
done in order to further narrow down binding sites with the aforementioned binding partners.

Also, similar to our NFZ selection assay, it would be interesting to design a selection assay to uncover pol V variants that alter interactions with RecA. I have discovered several variants that I suspect could be important for RecA interaction, such as N32, N33, and D34. It would be interesting to design a selection to determine if these variants arise when selecting for UV sensitivity and then pol V-mediated homologous recombination. A study has been carried out by Sommer et al. where they identified umuD′ and umuC variants that restore mitomycin C (a DNA crosslinker) sensitivity in a strain that suppresses the inhibition of RecA-mediated homologous recombination caused by umuD′C (Sommer 2000). By using replica plating, one can deduce which colonies survived or did not survive, and further analysis can be done on the parent colony. A similar system, using UV radiation, where one replicated plate is exposed to UV and one is not, could be used for a selection to discover variants that confer hypersensitivity to UV radiation. Resulting colonies would then be tested for inhibition of RecA-mediated homologous recombination using the transduction assay.

Another useful selection would be to repeat the NFZ selection, but in reverse, to find DinB variants that confer resistance to UV radiation. After treating plasmids with hydroxylamine, the library can then be exposed to increasing amounts of UV radiation. Then, the colonies that confer resistance to UV radiation at higher dosages than cells harboring wild-type DinB can be selected and characterized further.

Another interesting direction for this work would be to try the same NFZ selection but in different strains. The selection reported in Chapters 3 and 4 was done in
an AB1157 ΔdinB strain. In order to narrow down our in vivo results, the selection experiment should be repeated in an AB1157 ΔdinB ΔrecA strain to rule out any possibilities of RecA-mediated homologous recombination repairing the DNA damage caused by NFZ. Also, it would be interesting to try a selection in an AB1157 ΔdinB ΔuvrA strain as well, as uvrA is involved in nucleotide excision repair (NER), a mechanism of DNA repair, and it is also upregulated by the SOS response. This may help to gain more confidence that the variants selected are in fact related to the conferral of resistance to NFZ, and that the resulting DNA damage is not being repaired via these other mechanisms.

Lastly, it is important to determine the specific DNA lesion that results from exposure to NFZ. To do this, one could expose a strain, preferably with repair mechanisms such as NER or RecA-mediated homologous recombination deleted, to NFZ, and then isolate the genomic DNA. After specific digestion of the DNA, the resulting nucleotides would be analyzed by mass spectrometry to determine the mass of the lesion. Another less direct way to do this could be to expose cells harboring a generic plasmid, such as pBR322, to NFZ, and harvest the plasmid DNA and digest this DNA with specific enzymes for specific types of suspected damage, like hOGG1, which is an 8-oxoguanine-specific DNA glycosylase that will leave an abasic site after removing the damage, which could then be cleaved with AP endonuclease.

The work in this dissertation has elucidated features of UmuC that are important for conferring resistance to both cognate and non-cognate DNA damage. Notably, the protein active site and protein-protein interaction sites appear to contribute to these
activities. Future work will further define regions of UmuC that are important for resistance to a range of DNA damage.

5.1 References


