Novel Function and Regulation of Mutagenic DNA Polymerases in *Escherichia coli*

by

Daniel F. Jarosz

B.S. Chemistry and Biochemistry
Minor Mathematical Physics
University of Washington, 2001

Submitted to the Department of Chemistry in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN CHEMISTRY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2007

© 2007 Massachusetts Institute of Technology
All rights reserved

Signature of Author: ____________________________  Department of Chemistry  May 16, 2007

Certified by: ____________________________________  Graham C. Walker  Professor of Biology  Thesis Supervisor

Accepted by: ____________________________________  Robert W. Field  Professor of Chemistry  Chairman, Committee on Graduate Students
This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor John M. Essigmann

Professor Graham C. Walker

Professor Catherine L. Drennan

[Signatures]

Chairman

Thesis Supervisor
Novel Function and Regulation of Mutagenic DNA Polymerases in *Escherichia coli*

by

Daniel F. Jarosz

Submitted to the Department of Chemistry on May 16, 2007 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Chemistry

**ABSTRACT**

The observation that mutations in the *Escherichia coli* genes *umuC* and *umuD* abolish mutagenesis induced by UV-light strongly supported the counterintuitive notion that such mutagenesis is an active rather than passive process. Biochemical studies have revealed that *umuC* and its homolog *dinB* encode novel, low to moderate fidelity DNA polymerases with the ability to catalyze synthesis on imperfect DNA templates in a process termed translesion synthesis (TLS). Similar enzymes exist in nearly all organisms, constituting the Y-superfamily of DNA polymerases. Although DinB is the only Y-family DNA polymerase conserved among all domains of life, its precise function has remained elusive. Here we show that ΔdinB *E. coli* strains are sensitive to DNA damaging agents that form lesions at the N2 position of guanine. *In vitro* bypass studies of an N2-guanine adduct by DinB demonstrate considerable preference for correct nucleotide insertion and an increased catalytic proficiency on the lesion-bearing template relative to undamaged DNA. Moreover, DinB and its mammalian and archaeal orthologs possess similar substrate specificities. Mutation of a single residue in the active site of *E. coli* DinB suggests that its enhanced activity is coupled to lesion recognition and that its TLS function is required for resistance to DNA damaging agents *in vivo*.

Regulation of the mutagenic potential of DinB is critical for maintenance of genomic integrity. We present evidence indicating that abortive TLS products generated by a DinB variant are subject to the proofreading function of DNA polymerase III. Moreover, both the TLS activity and -1 frameshift mutator potential of DinB are modulated in a highly sophisticated manner by the DNA damage-inducible proteins RecA and UmuD2. These biochemical data, coupled with genetic analyses and molecular modeling, indicate that DinB is a specialized and remarkably controlled translesion DNA polymerase.

In addition, we present evidence that the *umuC* participates in several novel biological functions in addition to its established role in TLS. A novel *umuC* gain-of-function allele confers striking resistance to hydroxyurea and *umuC* mediates the expression of genes and physiological responses under conditions of SOS induction. Taken together, these observations hint at a largely uncharacterized function of Y-family polymerases in sculpting physiological responses, including active mechanisms of cell death, in response to environmental stress.

Thesis Supervisor: Graham C. Walker
Title: Professor of Biology
To my wife Mirna for her unwavering support and patience
# Table of Contents

Title Page 1  
Signature Page 3  
Abstract 5  
Dedication 7  
Acknowledgements 11  
Chapter 1: Introduction 13  
Chapter 2: A Single Amino Acid Governs Enhanced Activity of DinB DNA Polymerases on Damaged Templates 55  
Chapter 3: A Cluster of Aromatic Amino Acids Controls DinB Function in Translesion Synthesis 87  
Chapter 4: Conservation of Substrate Specificity Among DinB and its Archaeal and Mammalian Orthologs 121  
Chapter 5: UmuD and RecA Directly Modulate the Mutagenic Potential of the Y-family DNA Polymerase DinB 145  
Chapter 6: Y-family DNA Polymerases Respond to DNA Damage Independent Inhibition of Replication Fork Progression 189  
Chapter 7: umuC Mediates Diverse Physiological Outcomes During the SOS Response in *Escherichia coli* 227  
Chapter 8: Conclusions 255  
Appendix A: Genetic Evidence that DinB Acts on DNA Damage that is Recalcitrant to Excision Repair Mechanisms in *Escherichia coli* 267  
Appendix B: Bacterial Histone-like Proteins HupA and the Putative Histone Demethylase YcfD May Regulate DinB Function 275  
Appendix C: The *umuD* Gene Products Control DinB Function in Translesion Synthesis Under Conditions of SOS Induction 291  
Appendix D: Curriculum Vitae 303
Acknowledgements

I would like to sincerely thank my mentor Graham Walker for his invaluable guidance and support during my time in the lab. Graham is not only a singularly gifted scientist but also a truly exceptional human being. I am extremely grateful for the opportunity to pursue my dissertation research in his group. I would also like to thank our collaborators, John Essigmann, Jim Delaney, Will Neeley, Cathy Joyce, Angela DeLucia, Janice Pata, and Errol Friedberg, without whom much of the work presented in this thesis would have been impossible. In addition, I am indebted to John Essigmann and Cathy Drennan both for their advice along the way and for their experimental suggestions as members of my thesis committee.

The Walker lab has been an amazingly intellectually stimulating and perpetually amusing place to work. I want to thank Mary Ellen Wiltrout for tolerating the inevitable expansion of my experiments onto her bench and for being such a good-natured baymate, Sharotka (Xu) Simon for convincing me to exercise and determining the $K_d$ between DinB and UmuD$_2$, Kay Jones for always keeping me entertained, Rachel Woodruff for stimulating conversations about genetics, Jamie Foti for coffee breaks and agreeing to tackle DinB, Brenda Minesinger for her fascinating stories, Susan Cohen for always listening to my gripes and outrageous models, Sanjay D’Souza for his good humor, Bryan Davies for his unforgettable quotes, Lyle Simmons for his thoughtful advice, Michi Taga for ensuring I never lost perspective and maintaining the radioactive area, Hajime Kobayashi for his wit and maintenance of the coffee club, and Katherine Gibson for helpful experimental advice. I also want to thank former lab members Lise Barra for her spirit, Penny Beuning for lending a sympathetic ear and our many side-splitting conversations, Gail Ferguson for barely tolerating those conversations, Veronica Godoy for a truly synergistic working relationship, Jeelan Moghraby for always offering an eukaryotic perspective, and Laurie Waters for her friendship, love of Seattle, and opera parties. Thanks to Melanie Berkmen as well for thoughtful suggestions in group meeting and help with the fluorescence microscope. I am also grateful to Marianne White and Judy Carlin for keeping the lab running smoothly. My time at MIT was enriched by my interaction JoAnne Stubbe, who taught me to always think critically and that enzymes aren’t magic.

In addition, I would like to gratefully acknowledge the contribution of my mentors at the University of Washington, in particular my undergraduate advisor Rachel Klevit who taught me that being a successful researcher requires perpetual optimism. I would also like to thank the
faculty and staff of the Early Entrance Program at the Halbert and Nancy Robinson Center for Young Scholars (Kate Noble, Nancy Sisko, Maren Halvorsen, Bill Horder, Ron Geballe, and Nancy Robinson) for offering me the opportunity to pursue my undergraduate education in a unique and genuinely remarkable environment. My education has been the product of many exceptionally devoted, and exceptionally underpaid, public school teachers – in particular I would like to thank Jan Kragen.

I would not be where I am today without the steadfast support of my family. I thank my parents for teaching me the value of curiosity and inquiry from an early age and also my sister for always keeping me humble. Finally, I would like to thank my wife Mirna for her constant encouragement. Whether engaging in a lively discussion of an unexpected piece of data over dinner or forcing me to relax from time to time, this journey would not have been nearly as enjoyable without her by my side.
Chapter 1

Introduction

Introduction

DNA damage from both exogenous and endogenous sources is a difficulty with which all organisms must contend. Both the frequency and chemical diversity of such damage is considerable (1). For example, it is estimated that nearly 10,000 abasic sites are spontaneously generated in a single eukaryotic cell each day (2). Due to the universality of this problem, DNA repair pathways evolved early and have been conserved through evolution (Figure 1). In addition, several alternative approaches, collectively termed DNA damage tolerance pathways because they promote survival but do not remove DNA damage, have also been evolutionarily conserved (1). One such mechanism, translesion synthesis (TLS), plays a crucial role in the DNA damage response of organisms from bacteria to humans (3, 4). In TLS, specialized DNA polymerases incorporate a deoxyribonucleotide opposite to an otherwise replication-blocking DNA lesion and continue DNA synthesis past the site of damage (Figure 1). The majority of these enzymes belong to the Y-family of DNA polymerases, various subfamilies of which are found throughout evolution (5). The consequence of their broadened substrate specificity, however, is relaxed fidelity relative to replicative polymerases (4). Therefore regulation of Y-family polymerases is essential for the maintenance of genomic integrity. Indeed, Y-family polymerases are responsible for both spontaneous and induced mutagenesis in many organisms (1, 4, 6, 7).

SOS transcriptional regulation

The SOS response to DNA damage in *Escherichia coli* was the first inducible response to genotoxic stress to be characterized. Many molecular details of this response are now well understood (1). Transcription of genes that are induced as part of the SOS response is typically
repressed by the product of the $\text{lexA}^+$ gene (Figure 2). When replication is stalled by DNA
damage or another mechanism, the $\text{recA}^+$ gene product binds to single stranded DNA produced
at the replication fork, forming a nucleoprotein filament in the presence of nucleoside
triphosphates. This filament stimulates a latent autoproteolytic activity of LexA, thereby
inactivating LexA and allowing transcription of more than 40 genes. Both $\text{lexA}^+$ and $\text{recA}^+$ are
also SOS-regulated (1). Recent results have indicated that this simple view of the SOS-response
is far from complete, however (Figure 2). Agents that do not damage DNA, such as $\beta$-lactam
antibiotics, can induce the SOS response (8) through the two-component signal transduction
system $\text{dpiBA}$ and expression of the $\text{dinB}$ gene in particular through a $\text{lexA}$-independent
mechanism (9), presumably in an attempt to mitigate antimicrobial lethality by inhibiting cell
division. This observation raises the possibility that crosstalk between the SOS-response and
other cellular signaling pathways may be more extensive than previously realized. Maximal
transcription of $\text{dinB}$ in stationary phase requires a functional $\text{rpoS}$ gene, an effect that is also
$\text{lexA}$-independent (10). This may have particularly important implications for bacteria living
under conditions of nutrient starvation. The SOS-response also appears to be oscillatory at the
single-cell level, and this oscillation is dependent on the $\text{umuDC}$ genes (11). Finally, the SOS-
response is one component of a broader cellular response to DNA damage. Exposure of $E.\text{coli}$ to
the DNA damaging agent mitomycin C (MMC) results in expression changes of more than 1,000
genes (12).

A number of the genes regulated by the SOS-response were initially identified using
randomly generated Mu d1-generated transcriptional fusions to the $\text{lac}$ operon (13, 14). A
collection of $E.\text{coli}$ strains bearing these fusions was treated with MMC and examined for
expression of $\beta$-galactosidase. Certain of these fusions exhibited inducible expression of $\beta$-
galactosidase dependent on recA+ and lexA+ and were thus named din (for damage-inducible) (14). Many of these genes and their gene products have still not been characterized in detail. Although dinB, which encodes the TLS polymerase DNA pol IV was identified in this experiment, deletion of the gene did not initially show any marked phenotypes – in striking contrast to umuD and umuC (see below). Both umuD and umuC were subsequently shown to be transcriptionally induced as part of the SOS response (15).

Mutagenic function of umuD+-C+ and dinB+

Early studies of mutagenesis induced by UV-irradiation indicated that mutation of either the recA+ or lexA+ genes results in a nonmutable phenotype [reviewed in (1)]. A screen for additional nonmutable mutants identified the umuD+ and umuC+ genes (6). Loss of function mutants of each of these umu genes also show modest sensitivity to UV-irradiation (1). UmuD and LexA are structurally related to the lambda repressor, which undergoes RecA-nucleoprotein activated autocleavage, and to peptide hydrolases that employ a Ser-Lys catalytic diad in their mechanism (1). Both LexA and UmuD form homodimers in solution and similarly to LexA, interaction of UmuD2 with the RecA nucleoprotein filament induces a latent autoproteolytic activity causing UmuD2 to remove its N-terminal 24 amino acids to form UmuD'2 (16, 17). It is UmuD'2 that is active in mutagenesis induced by UV-irradiation and associates with UmuC to form DNA polymerase V (UmuD'2C) (17-20).

In contrast to the marked phenotypes displayed by mutants of umuD and umuC, mutants of dinB show more enigmatic phenotypes (21). Although deletion of the dinB+ gene has almost no discernable effect on spontaneous mutagenesis (22), the dinB+ gene is required for untargeted mutagenesis of λ phage, in which E. coli are UV-irradiated and transfected with unirradiated λ.
but UV-induced mutagenesis is seen in the \( \lambda \) DNA (7). The mutation spectrum observed is
distributed between base substitution mutations and -1 frameshift events with a strong preference
for mutation at G:C basepairs (23).

The \( \text{dinB}^+ \) gene is also important for the phenomenon of adaptive mutagenesis in \textit{E. coli} (24). The report of adaptive mutation under conditions of nonlethal selection challenged the
prevailing notion established by Luria and Delbruck that mutants arise spontaneously during
growth and are pre-existing at the time of selection (25). In the case of adaptive mutation,
stationary phase \textit{E. coli} that are unable to metabolize lactose by virtue of a +1 frameshift in an
episomal fused \textit{lacI-lacZ} allele are plated on medium with lactose as the sole carbon source (26).
New \textit{lac}^+ mutants appear over time despite the fact that the bacteria are not growing. The
molecular details of the mechanism responsible for this mutagenesis are quite controversial, but
the involvement of \textit{dinB} is well-established (27-35). Deletion of \textit{dinB} results in a 5-10-fold
reduction in the number of adaptive mutants that appear (24). Adaptive mutagenesis is also
regulated by a number of genes including \textit{rpoS} (10, 36), the chaperones \textit{groES} and \textit{groEL} (37),
and \textit{ppk} (38). Both \textit{groES} and \textit{groEL} mutants are also impaired for \textit{umuDC} dependent UV-
induced mutagenesis (39).

Overproduction of \textit{dinB} leads to an increase in the frequency of spontaneous and 4-NQO-
induced -1 frameshift and, to a lesser extent, base substitution mutagenesis (40, 41). Curiously, a
modest preference is observed for spontaneous mutagenesis on the lagging strand and this
appears to result from extension of terminal mismatches (42). Moreover, a considerable fraction
of the lagging strand directed mutator phenotype of a constitutively SOS-induced \textit{recA730} strain
requires \textit{dinB}^+ (43). Notwithstanding the relative ambiguity of these phenomena, they laid the
foundation, particularly in concert with the known mutagenic phenotypes of the \textit{dinB} homolog.
umuC, for a view of dinB as an agent of mutation. It has recently been shown, however, that 

\( \Delta \text{dinB} \) strains of \( E. \text{coli} \) display increased sensitivity to the DNA damaging agents nitrofurazone (NFZ) and 4-nitroquinoline-1-oxide (4-NQO) (44). Despite this marked sensitivity to both NFZ and 4-NQO, deletion of the \( \text{dinB}^+ \) gene does not reduce mutagenesis induced by either agent. These data suggest that the \( \text{dinB}^+ \) gene product is able to contend with DNA damage produced by at least some DNA damaging agents with comparable fidelity to other repair processes available to the \( E. \text{coli} \) cell.

Eukaryotic pol \( \kappa \) also participates in a variety of mutagenic phenomena (45). Although pol \( \kappa^{-/} \) mice are viable and fertile, they display a mutator phenotype (46). Pol \( \kappa^{-/} \) mouse embryonic fibroblasts are sensitive to benzo[a]pyrene (47) and are impaired in replicating a gapped plasmid containing a site-specific \( N^2-\text{B}[a]P-dG \) lesion (48). Pol \( \kappa^{-/} \) mice also show increased spontaneous mutagenesis (46), and pol \( \kappa \) appears to play a role in recovery from a \( \text{B}[a]P \)-induced S-phase checkpoint (49). Similarly to \( E. \text{coli} \) DinB, overproduction of pol \( \kappa \) not only increases mutation frequency but also promotes double strand breaks (DSB), increased homologous recombination and nonhomologous end joining, loss of heterozygosity, and aneuploidy (50). The striking panel of genetic abnormalities induced by improper expression of pol \( \kappa \) clearly indicates the importance of proper regulation of potentially mutagenic Y-family polymerases in multicellular organisms.

**Biochemical activities of DinB and UmuD\(_2\)C**

Although decades of genetic characterization had clearly established their roles in spontaneous and induced mutagenesis, the biochemical function of the \( \text{umuD}^+\text{-C}^+ \) and \( \text{dinB}^+ \) gene products remained elusive for many years. The discovery that \( \text{REV1} \) from \( \text{Saccharomyces} \)
*cerevisiae* encodes a dCMP transferase activity hinted at a possible direct role in DNA synthesis (51). Previous models had suggested that these gene products might interact with replicative polymerases, thereby modulating their fidelity (1). The subsequent findings that Rad30/XP-V, DinB, and UmuD'2C are also DNA polymerases led to the discovery of a fifth superfamily of DNA polymerases, the Y-family (19, 20, 52). These polymerases participate in TLS and are characterized by their weak activity and comparatively low fidelity on undamaged DNA (4).

The need for both Y-family and other DNA polymerases to be highly regulated is evident from their large numbers in many organisms – e.g. 5 in *E. coli*, 10 in *S. cerevisiae*, and 16 in *H. sapiens* – with error rates varying over many orders of magnitude (1).

Unlike DNA pol III, *E. coli*'s replicative DNA polymerase, DinB and UmuD'2C catalyze relatively distributive DNA synthesis that is modestly stimulated by the addition of the β processivity clamp subunit of DNA polymerase III (53-55). AP lyase activity has been demonstrated for both DinB and UmuD'2C, although genetic studies have not established a relevance for this function *in vivo* (56). The *in vitro* DNA polymerase activity of UmuD'2C and DinB on both damaged and undamaged DNA has been examined in some detail. Their specialized function comes with a mutagenic risk, as Y-family polymerases replicate DNA with lower fidelity than their replicative relatives. While UmuD'2C and DinB display poor activity and fidelity on undamaged DNA relative to replicative DNA polymerases, on certain types of damaged templates they compare far more favorably. UmuD'2C replicates undamaged templates with an error frequency of $10^{-3}$-$10^{-4}$ and has an error frequency of $10^{-2}$ for T^T cyclobutane dimers (57, 58). Like other Y-family polymerases, DinB and its orthologs appear to act with a range of error frequencies on damaged and undamaged templates. DinB from *E. coli* replicates undamaged and certain damaged templates with error frequencies between $10^{-3}$ and $10^{-5}$ (44, 59),
while its mammalian ortholog pol κ synthesizes undamaged DNA with a slightly higher error frequency of $10^{-2}$-$10^{-3}$ (60). Both enzymes also have the ability to produce -1 frameshift mutations during DNA synthesis at an appreciable frequency (52, 59, 60). The difference between the fidelity of these polymerases when replicating damaged substrates may correlate with the clear UV-induced mutagenic signature of *umuDC* in vivo and the comparative lack of *dinB*+-dependent mutagenesis induced by nitrofurazone or 4-NQO (6, 44). Furthermore DinB shows an increased catalytic proficiency on an $N^2$-dG damaged substrate relative to an undamaged control, which is dependent on a single active site residue (44). Certain data suggest that DinB/pol κ function may in fact be antimutagenic in vivo. Although deletion of either gene does not alter the frequency of spontaneous mutation (7, 45), it results in an increase in the frequency of mutagenesis induced by certain DNA damaging agents (44, 48). Such behavior is analogously observed with respect to UV-induced mutation in human cells deficient in pol η function, which replicates relatively accurately over thymine-thymine cyclobutane dimers (61, 62). Substrate specificity also appears to be conserved between DinB and pol κ. Both enzymes are able to bypass $N^2$-benzo[a]pyrene-adducted template G (48, 63, 64). Moreover, they each display a striking 10-15 fold increased activity on a template $N^2$-furfuryl-dG relative to undamaged DNA (44). Increased proficiency is also displayed by pol η replicating its cognate substrate T-T cyclobutane pyrimidine dimers (65). Curiously, both DinB and pol κ proficiently extend from a variety of lesions and mismatched primer ends in addition to their insertion specificities (44, 52, 66-70). It has been suggested, at least for pol κ, that this property reflects a separate role in the extension step of TLS (66).
Substrate specificity of Y-family DNA polymerases

Given the considerable chemical diversity of DNA modification, polymerase usage is a nontrivial problem. The number of DNA polymerases available to an organism is limited and the structural variation of DNA adducts likely precludes the existence of a single cognate lesion for each DNA polymerase. Both DinB and UmuD'2C show considerable preference for certain classes of substrates, however. DinB has been shown to catalyze DNA polymerization on substrates containing mismatched primer ends (52), covalent adducts at the $N^2$ position of a template G (44, 63, 71), and may cooperate with DNA polymerase II to mutagenically bypass oxidative lesions produced by methylene blue and UV-irradiation (71). Experiments described in chapter 4 of this thesis indicate that such substrate promiscuity is not derived merely from a relatively open active site, but rather from specialized activity. UmuD'2C appears to possess slightly broader range of substrate specificities. Abasic sites (19, 72), T^T cyclobutane pyrimide dimers and 6-4 (T-T) photoproducts (57), and several C8 covalent adducts of a template G (63, 71) are bypassed by UmuD'2C with varying efficiencies and fidelities. Computational modeling studies have bolstered the notion that the eukaryotic polymerases DNA pol κ and DNA pol η are functional orthologs of DinB and UmuD'2C, respectively (73). Pol κ appears to possess similar substrate specificities to those of DinB. It has been implicated in modest bypass of mismatched primer template ends (59, 67), thymine glycols (68), and proficient and accurate bypass of bulky covalent modifications at the $N^2$ position of a template G (44, 48, 64, 74). Whereas pol η appears to act primarily in preferential and relatively accurate bypass of T^T cyclobutane pyrimide dimers (62, 65, 75), it also participates in more modest bypass of thymine glycol (76), intrastrand GpG crosslinks induced by the chemotherapy agent cisplatin (77), and estrogen derived covalent
purine modifications (78, 79). Thus, while these polymerases each have relatively broad substrate specificities, they exhibit preferential function on a smaller subset.

Translesion synthesis can be divided into two steps – insertion an extension. During insertion, a nucleotide is added opposite a damaged template base, while during the extension step nucleotides are incorporated past the site of the lesion. In prokaryotic systems, an individual Y-family DNA polymerase participates in both steps. Indeed, when replicating an \( N^2 \)-furfuryl-dG damaged template \( E. \) \( \text{coli} \) DinB is more catalytically proficient at extension than insertion (44). UmuD\(_2\)C is also able to replicate opposite and past T^T cyclobutane pyrimidine dimers and 6-4 (T-T) photoproducts (57). In eukaryotes, which are comparatively polymerase rich, the insertion and extension phases of TLS are often performed by separate polymerases (80). For bypass of T^T dimers, pol \( \eta \) is able to insert nucleotides opposite to the lesion, but a eukaryotic X-family DNA polymerase, pol \( \zeta \), can catalyze extension (80). In the case of abasic sites in yeast, pol \( \delta \), pol \( \iota \), and Rev1 are all thought to contribute to the insertion step with varying efficiencies, while pol \( \zeta \) catalyzes extension (81). Moreover, for bypass of certain minor groove \( N^2 \)-dG adducts, pol \( \iota \) appears to catalyze insertion and pol \( \kappa \) promotes extension (66). One possible reason for this separation of TLS into insertion and extension steps is that it provides increased opportunity for regulation, as a mutation is not fixed in the genome until a mismatch is extended from.

**Loose grips and open active sites**

Although complete structures of the Y-family polymerases from \( E. \) \( \text{coli} \) have not yet been solved, structural analysis from \( S. \) \text{solfataricus} (Dpo4) and \( S. \) \text{acidocaldaricus} (Dbh) homologs have yielded profound insights into function (82, 83). While these enzymes share no clear
sequence homology with replicative polymerases, their structures reveal a similar right-hand fold consisting of a thumb, palm, and fingers domain. However, Y-family polymerases possess an additional little finger domain or PAD that appears to play an important role in both substrate specificity and processivity (84). Unlike the tight grip seen in active sites of canonical DNA polymerases (85), however, Y-family polymerases possess open active sites that are relatively solvent accessible (Figure 3). Moreover, an α-helix responsible for several orders of magnitude of fidelity in canonical DNA polymerases (the O-helix) (86) is entirely absent in Y-family polymerases, providing a structural rationale for their comparatively low fidelity when replicating undamaged DNA.

Structural insight into Y-family polymerases encountering their cognate substrates is considerably more limited. A study of Dpo4 encountering a cyclobutane pyrimidine dimer is the most definitive to date (87). Such UV-induced damage presents a particular problem for replicative polymerases, as their active sites can only accommodate one base at a time. The relative openness of the Dpo4 active site allows the enzyme to fit a covalently linked T^T cyclobutane pyrimidine dimer within its active site (87). While Dpo4 replicates past the 3’ T of the T^T with appreciable efficiency, it replicates past the second base with considerably higher activity and fidelity (87). This is particularly interesting given that structural analysis of the second addition reveals that the incipient base pair adopts a Hoogsteen conformation, rather than the more typical anti conformation. Hoogsteen basepairing occurs in the major groove and involve the N7 atom of purines in contrast to canonical Watson-Crick basepairing, which occurs in the minor groove (88). In the case of Dpo4, the conformation appears to be induced by the enzyme, as both bases in a T^T in isolated duplex DNA adopt a Watson-Crick conformation by
nuclear magnetic resonance (NMR) spectroscopy (87). The scope of Hoogsteen basepairing in TLS is a subject of intense investigation and has been observed crystallographically in pol ι (89).

An induced conformational change between an open and substrate-bound closed form is a hallmark of A and B family DNA polymerases (90). Indeed, this conformational change is believed to provide a major basis of replicative polymerases’ exquisite fidelity (86, 91). While not observed crystallographically for a Y-family polymerase, several pre-steady state kinetic studies have suggested that such a conformational change may play a critical role in translesion synthesis (44, 92, 93). These observations provide further evidence that Y-family polymerases catalyze translesion synthesis in an orchestrated fashion rather than exclusively by virtue of their promiscuous active sites.

**Sloppier copiers or specialized DNA polymerases?**

The notion that Y-family polymerases always act as low fidelity polymerases is inconsistent with the fact that XP-V patients, who bear mutant Rad30/pol η alleles, are prone to skin and other cancers suggesting that the function of pol η may be antimutagenic in humans (94, 95). XP-V cells in culture also show increased UV-induced mutagenesis (61). This is in striking contrast to other mammalian Y-family polymerases in mammals such as Rev1, whose action generates clear mutagenic signatures (51, 96). The recent discovery of the sensitivity of a ΔdinB strain to two DNA damaging agents that produce adducts at the N² position of dG has allowed us to determine whether dinB clearly falls into one of these categories (44). DinB catalytic function is clearly required for resistance to these agents, but deletion strains show either the same or increased induced mutation frequency relative to the wild-type when treated with either agent (44). These observations, taken together with the fact that a chromosomal
deletion does not alter either spontaneous or induced mutagenesis suggest that dinB is largely anti-mutagenic under many circumstances.

These in vivo observations stimulated in vitro studies which led to the discovery that E. coli DinB is able to bypass an adduct produced by NFZ, $N^2$-furfuryl-dG, with 15-fold greater efficiency than undamaged DNA. This property appears to be shared by DinB orthologs from archaea and mammals (44). When we generated a model of DinB encountering $N^2$-furfuryl-dG, we noted that a pocket in the enzyme is appropriately positioned to accommodate an $N^2$ modification on a template dG, bringing the adduct into contact with the ‘steric gate’ of DinB (F13). In DNA polymerases, ‘steric gates’ are responsible for the occlusion of improper rNTP substrates from the growing DNA chain, but we wondered whether this contact might also provide a basis for increased activity on an $N^2$-dG adducted template. Strikingly, mutation of F13 to V13 renders DinB virtually inactive for TLS but does not reduce its polymerase activity on undamaged templates (44). Moreover, the mutant is unable to complement NFZ and 4-NQO sensitivity, indicating that lesion bypass is required for resistance to these agents in vivo. It is provocative that throughout evolution no DinB orthologs possess a non-aromatic residue at this position.

Given the multiple and redundant high fidelity DNA repair and tolerance pathways available to the cell, why employ potentially mutagenic TLS? An emerging body of evidence suggests that certain types of DNA damage, particularly modification at the $N^2$ position of dG, may be particularly recalcitrant to such repair pathways. In mammalian cells damaged with acetylamino fluorine, the $N^2$-dG isomer is the most persistent lesion observed despite being the least commonly produced (74). Moreover, certain other $N^2$-dG adducts have been shown to be recalcitrant to repair by the E. coli nucleotide excision repair system in vitro (97). These
observations are particularly interesting given the recent finding that pol κ may play a role in nucleotide excision repair in mammalian cells (98).

**Modulation of function by protein-protein interactions**

Genetic characterization over nearly thirty years has underscored the importance of the recA and umuD gene products in regulation of umuC-dependent mutagenesis (1). Recent studies have recapitulated these results with purified components and identified the pivotal role of the β processivity clamp in dictating UmuD’2C function. Initial reports of UmuD’2C polymerase activity invoked a requirement for UmuD’2, RecA, SSB, and in one case various components of the polymerase III holoenzyme for UmuC activity (55). The demonstration of polymerase activity of UmuD’2C established UmuD’2 as a subunit of DNA polymerase V.

X-ray and NMR structures of the pol V subunit UmuD’2 have yielded considerable insight into its function (99, 100) (Figure 3). Additionally, EPR-derived distance constraints have been used to model the structure of the full-length UmuD2 (101). In the X-ray structure of UmuD’2, the catalytic serine and lysine required for autoproteolysis are located within hydrogen bonding distance of each other, while the N-terminus containing the scissile bond is located >50 Å from the active site (99). In contrast, the UmuD2 model suggests that the N-terminus of the molecule curls upon itself to bring the scissile bond in proximity to the active site (101). Such structural plasticity may be especially important given the relatively large number of proteins with which UmuD2 and UmuD’2 interact (102). A heterodimeric form of the umuD gene products, UmuD-D’, is the most thermodynamically stable form of the protein and targets it for ClpXP mediated proteolysis (1). A structural model of UmuD-D’ has been constructed based upon NMR analysis (100). Recent computational and biochemical studies have revealed that
UmuD₂ may possess remarkable structural plasticity, perhaps providing a mechanistic basis for its specific interactions with so many diverse proteins (103).

Aside from activating UmuD'₂C, RecA has numerous cellular roles. The recA gene is required not only for induction of the SOS-response but also for homologous recombination (1). Biochemical studies differ to some extent on the mode of RecA activation of UmuD'₂C as well as on the role of ATP in the process (55). Recent studies have suggested that RecA binds to UmuC as a subunit of the UmuD'₂C holoenzyme and that another molecule of ATP-associated RecA binds to UmuD', thereby stimulating the affinity of the holoenzyme for the primer terminus (104). It has been assumed that RecA is bound to the single-stranded DNA (ssDNA) template in this activating role, but it has now been proposed that stimulation of UmuD'₂C activity by the RecA-nucleoprotein filament occurs in trans (105). This has significant implications for models of UmuD'₂C action given that the most proficient transactivating RecA-nucleoprotein filament is one formed on gapped DNA. These observations foreshadow what appears to be remarkably complex regulation of Y-family polymerases via protein-protein interactions. Initial studies of UmuD'₂C activity also reported an enhancement of activity provided by single stranded DNA binding protein (SSB) (55). This effect, observed at substoichiometric quantities of SSB, has now been attributed to increased formation of dynamic RecA filaments on short ssDNA templates in the presence of DNA (55).

Protein regulators of DinB function have been comparatively less well-characterized. A recent report has implicated certain forms of the umuD gene products in regulation of a novel function of DinB (106) and the sigma factor RpoS (10) as well as the chaperone GroEL/GroES (37) affect DinB levels, perhaps indirectly. The recent identification of an additional phenotype
for ΔdinB E. coli strains (44) should promote investigation of DinB regulation considerably in the future.

Management role of replicative processivity clamps

Interactions with replicative processivity clamps are crucial for regulating Y-family polymerase activity and dictating their access to DNA. Although they are characterized by low processivity on undamaged DNA, Y-family polymerases exhibit an increased processivity in the presence of the β clamp. Indeed, DinB processivity is enhanced 300-fold by the β clamp (53), whereas that of UmuC is stimulated between 5- and 100-fold (54, 55). In either case, the processivity enhancement due to β is far less than that of pol III (~10^5-fold) (107). Mutation or deletion of the β interaction motif in either UmuC or DinB causes a loss of translesion synthesis in vivo (108). Most prokaryotic proteins that interact with the β processivity clamp do so via a conserved interaction motif: QL[S/D]LF (109), which bears similarity to the conserved eukaryotic PCNA interaction motif, QxxLxxFF (110).

Recent structural studies have shown that proteins as diverse as the δ subunit of the clamp loader and DinB, which interact with β via the conserved interaction motif, bind to the same hydrophobic channel on β at the interface between β Domains II and III (111-113). Thus, mutations in β near this hydrophobic channel can regulate specific DNA polymerase usage (114, 115). A co-crystal structure of the C-terminal little finger domain of DinB with the β clamp illustrates that in addition to the conserved β-binding motif interaction, DinB also interacts with β at its dimer interface via a hydrophobic loop in the little finger domain (112). When the structure of full-length S. solfataricus Dpo4 was superimposed on the DinB little finger in this structure, the active site of Dpo4 was surprisingly far from the DNA that is expected to be
running through the center of the β clamp, leading the authors to speculate that this orientation of DinB may represent a recruited-but-inactive state (Figure 4) (112).

What is the role of the β clamp in managing multiple DNA polymerases? Notably, all DNA polymerases in E. coli interact with β at the same site (116). The co-crystal structure of the DinB little finger and the β clamp suggests that it may be possible for β to bind two DNA polymerases simultaneously, with one polymerase in an inactive but still recruited conformation. Indeed, both DinB and the α catalytic subunit of pol III were found to bind to β simultaneously (117). Thus, switching polymerase access to the primer terminus may occur with two DNA polymerases bound to the β clamp (117). The hierarchy of affinities of DNA polymerases in E. coli for the processivity clamp has been investigated genetically (114, 118, 119). Whereas upon UV-irradiation the Pol III appears to possess the greatest affinity followed by Pol IV, Pol V, and Pol II (114), during conjugal replication the hierarchy appears to be first Pol III, then Pol II, Pol IV, and finally Pol V (119). Further work will be required to analyze competition among polymerases for access to the β clamp under various conditions.

The β clamp also interacts with UmuD2 and UmuD′2. Moreover, UmuD2 interacts with β more strongly than UmuD′2, possibly indicating a role in umuDC dependent replication pausing (102). UmuD binds β in the vicinity of the same hydrophobic channel where other β-binding proteins interact (120). Curiously, the N-terminal region of UmuD contains a cryptic β-binding motif (14TLPLF18) that by itself is insufficient to bind to β (109). UmuD variants containing mutations in this motif bind to β with essentially the same affinity as wild-type UmuD (103), but with a strikingly different tryptophan fluorescence emission spectrum of β (103), indicating that while this motif itself is not responsible for the strength of the interaction, it is important for determining the nature of the complex.
In eukaryotes it has recently been discovered that posttranslational modification of the PCNA processivity clamp by ubiquitination is required for TLS \textit{in vivo} (121-124). Furthermore, interaction with PCNA is a requirement for fluorescent focus formation of green fluorescent protein (GFP) tagged pol \( \kappa \) in response to DNA damage (125). The alternative processivity clamp Rad9-Rad1-Hus1 (9-1-1) has been also shown to be important for DinB function in \textit{Schizosaccharomyces pombe} (126). The role of ubiquitin in TLS is even more elaborate than simple PCNA modification. It has recently been shown that eukaryotic TLS polymerases bind ubiquitin and are themselves subject to ubiquitination (127, 128). Such modification appears to be important for their function in mutagenesis and in prevention of lethality induced by various DNA damaging agents. The regulation of TLS polymerases has clear implications for cancer, as many XP-V patients bear versions of Rad30 that are in principle competent for translesion synthesis but are lacking one or more putative regulatory domains (95).

The scope of posttranslational modification in regulating TLS is likely to be considerable. Of particular interest presently is how ubiquitination or other posttranslational modification dictates the site of TLS. It has long been assumed that TLS on the leading strand occurs at stalled or regressed replication forks (1, 129). However, early models of DNA replication proposed discontinuous DNA synthesis on both the lagging and leading strands following UV-irradiation (130). Such a model would allow for TLS to occur either at the replication fork or post-replicatively at such gaps, and evidence exists for both mechanisms (123, 131-133). In \textit{S. cerevisiae} the levels of Rev1, a Y-family DNA polymerase critical for mutagenesis, are highest not during S-phase when the majority of replication occurs, but rather later during G2/M (133). In \textit{S. pombe} PCNA is ubiquitinated upon DNA damage specifically in G2, suggesting a cellular signal for DNA damage outside of replication (123). Whether TLS occurs exclusively at gaps
left as a result of replication reinitiation is unclear. Indeed, measurements of the rate of nascent DNA synthesis following DNA damage would suggest that if this were the chief mechanism of TLS, repriming would be quite slow, occurring over tens of minutes (131). Curiously, measurements of the kinetics of TLS in vivo have suggested that it can take ca. 10 minutes to be completed (134). Thus, an attractive aspect of a post-replicative TLS model is that it allows for lesion bypass to occur in parallel with replication of the genome rather than in series, thereby allowing the cell to maintain a reasonable doubling time.

**Novel phenomena involving dinB and umuDC**

In addition to the well known function of Y-family polymerases in TLS, other functions of *umuDC* and *dinB* include UmuD$_2$C-dependent cold sensitivity, involvement in a primitive DNA damage checkpoint, enhanced survival in response to DNA damage independent replication stalling, and replication arrest-stimulated recombination (1, 106, 135-137).

Overexpression of the *umuDC* gene products leads to inhibition of growth at 30°C, known as *umuDC* mediated cold sensitivity. The *umuDC* genes are the only SOS regulated genes required for the manifestation of cold sensitivity and the degree of cold sensitivity is proportional to the amount of expression. This phenomenon is associated with the rapid and reversible inhibition of DNA synthesis as well as *sulA* independent filamentation (135). Strikingly, the genetic requirements for cold sensitivity are different from those needed for TLS (138). Namely, neither RecA nor the catalytic activity of UmuC is needed, and UmuD, but not UmuD', is required. Cold sensitivity appears to result from an exaggeration of a DNA damage induced checkpoint in which UmuD$_2$C delays the resumption of DNA synthesis after DNA damage, perhaps through interaction with the β clamp, to allow error free repair processes to
occur (136, 139). The response is temporally regulated by the cleavage of UmuD to UmuD'. Overproduction of dnaQ, which encodes the epsilon proofreading subunit of DNA polymerase III, has also been shown to suppress the phenomenon of umuDC dependent cold sensitivity (140).

Both E.coli Y-family polymerases have been implicated in enhancing cellular survival under conditions of depleted deoxyribonucleotide pools, such as occurs after the addition of hydroxyurea (HU). Strains carrying a umuC122::Tn5 allele, resulting in a truncated protein that retains an intact polymerase domain but is deficient for induced mutagenesis, are strikingly resistant to HU (106). Although seemingly unrelated, cold sensitivity and resistance to HU share a genetic requirement for umuD. HU resistance requires the catalytic activity of UmuC122 and DinB as well as certain forms of the umuD gene products. Moreover, this resistance may be due at least in part to failed communication with the toxin/antitoxin pairs MazEF and RelBE that would normally lead to cell death (106). The increased mutation frequency observed in a umuC122::Tn5 strain upon HU treatment may imply that under conditions of deoxyribonucleotide limitation, DinB and UmuD'2C take over a considerable fraction of DNA replication. Furthermore, recent studies have shown that Y-family DNA polymerases participate in oxidation induced mutagenesis by virtue of their ability to incorporate oxidized nucleotides during replication (141, 142). Taken together, these results suggest that Y-family polymerases may play a larger role in DNA replication when the deoxyribonucleotide pool is significantly perturbed such as under conditions of HU treatment or oxidative stress.

Interestingly, dinB has also been implicated in replication arrest-stimulated recombination (137). Deletions of tetA fragments that are set in tandem repeats are elevated at the permissive temperature in a strain background bearing a temperature sensitive mutant of the
replicative DNA helicase (*dnaB107*). This type of mutagenesis is reduced in a *dnaB107 dinB*, contributing to a model in which RadA, RecG, and RuvAB can stabilize a D-loop/recombination intermediate that allows DinB to extend the invading 3' strand and promote continued replication (137).

**Concluding perspectives and future directions**

Although initially considered simple agents of mutation, recent developments have led to the more nuanced view that the relative infidelity of Y-family polymerases results from their broadened substrate specificity. Indeed, the potential for base substitution or even frameshift mutagenesis may be far preferable to the risk of chromosomal rearrangements induced by stalled replication forks – particularly in higher eukaryotes where most of the genome is non-coding. However, several key questions regarding regulation remain. In the case of DinB/pol κ branch of the superfamily, TLS may be particularly important because its apparent cognate substrates appear to be recalcitrant to other repair mechanisms (74, 97). However, other TLS polymerases act on lesions that can be efficiently and accurately repaired by any number of other mechanisms. One possibility is that the intrinsic mutation frequency of TLS can be significantly modulated in certain cases by proofreading of mismatch repair *in trans* (143, 144). Further studies will be required to elucidate what dictates the choice of TLS over other DNA repair and damage tolerance mechanisms. It is still unclear what regulates access to the replication fork, presumably preventing these potentially mutagenic enzymes from inappropriately compromising genomic integrity. *E. coli* delays the mutagenic function of UmuD'2C by timing the cleavage of UmuD₂ to UmuD'₂. Moreover, whether selection of the proper polymerase occurs merely by stochastic competition or by more a more orchestrated mechanism is a topic of intense debate.
Another key question is whether TLS occurs solely at static stalled replication forks or perhaps additionally at gaps left after replication re-initiation. Recent results in *S. cerevisiae* have suggested that replication blocking lesions can induce both leading and lagging strand gaps (132), evoking early models of replication in *E. coli* (130). Although it is clear that regression of stalled replication forks also plays an important role in TLS (147), these observations have led to the speculation that TLS polymerases may act post-replicatively at such gaps (133). Such a model is particularly attractive given the weak activity of Y-family DNA polymerases, as it allows replication and TLS to be carried out simultaneously (134). Further research will give considerable insight into these and other problems. Indeed, the exquisite regulation of Y-family polymerases may be particularly important in eukaryotes which according to some estimates rely on translesion synthesis 50-fold more than prokaryotes (48).

**Acknowledgements:** We thank members of the Walker laboratory for thoughtful comments and also wish to thank Penny Beuning (Northeastern University) and Susan Cohen for contributing to the manuscript. G.C.W. is an American Cancer Society Research Professor. Funding was provided by National Institutes of Health grant CA021615 to G.C.W. and NIEHS grant P30 ES002109 to the MIT Center for Environmental Health Sciences.
References


97. Zewail-Foote, M., V. S. Li, H. Kohn, D. Bearss, M. Guzman, and L. H. Hurley. 2001. The inefficiency of incisions of ecteinascidin 743-DNA adducts by the UvrABC nuclease and the unique structural feature of the DNA adducts can be used to explain the repair-dependent toxicities of this antitumor agent. Chem. Biol. 8:1033-49.


Figure 1. Cellular responses to DNA damage, including DNA damage tolerance via translesion synthesis. If an unrepaired lesion (shown in red) remains in the chromosome during replication, it may be bypassed by translesion synthesis in which a specialized polymerase (shown in green) is recruited to the replication fork and copies over the lesion.
Figure 2. The inducing signal for the SOS response forms when RecA polymerizes on a region of ssDNA, which is formed due to the failure to replicate damaged DNA. The RecA/ssDNA nucleoprotein filament is referred to as RecA*. Binding to RecA* induces LexA to undergo autoproteolytic cleavage, which inactivates it as a repressor and leads to the induction of at least 40 genes, among which are the Y family DNA polymerases UmuD'2C and DinB. The cleavage of UmuD to UmuD' is also facilitated by the binding of UmuD2 to RecA*, which provides temporal regulation of the potentially mutagenic translesion synthesis activity of UmuC. Transcription of dinB is also regulated by rpoS, dpiBA, and β-lactam antibiotics. The chaperone GroEL/GroES is required for both DinB and UmuD'2C function. Extensions on UmuD2 and UmuD'2 represent the N-terminal arms; extensions from DinB and UmuC represent their C-termini including their β binding motifs.
Figure 3. X-ray and NMR structures reveal key mechanistic details of TLS. a) The structure of *Bacillus stearothermophilus* replicative DNA Polymerase I in a closed conformation (85) shows numerous close protein (shown in yellow) contacts with DNA (shown in red). An α-helix (orange) performs a geometric check to ensure the fidelity of the incipient basepair (cyan). b) In contrast, the Y-family polymerase Dpo4 from *Sulfolobus solfataricus* (82) shows a loose grip on the DNA, a relatively open active site, and has no α-helix to check the geometry of the incipient basepair. c) A model of UmuD₂ (101) and d) an NMR structure of UmuD'₂ (100) indicate the structural rearrangements that occur upon RecA-mediated autocleavage. The structural plasticity of these molecules is likely important for their ability to interact with various cellular factors.
Figure 4. A model for polymerase switching that may occur in the transition from a DNA damage checkpoint to translesion synthesis and replication. In a DNA damage checkpoint, UmuC acts with UmuD$_2$ to slow the rate of DNA synthesis. Autocleavage of UmuD$_2$, which removes its N-terminal 24 amino acids to form UmuD'$_2$, releases the checkpoint and is required for UmuC polymerase function. After UmuC polymerizes several base pairs past the lesion (21), the replicative polymerase DnaE (pol III $\alpha$ subunit) can resume DNA synthesis. Inset Crystal structure of the little finger domain of DinB (red) with beta (one monomer in blue, one monomer in green) (112).
Chapter 2

A Single Amino Acid Governs Enhanced Activity of DinB DNA Polymerases on Damaged Templates

Abstract and Introduction

Translesion synthesis (TLS) performed by Y-family DNA polymerases is a major mechanism of DNA damage tolerance (1). Such TLS can be quite accurate or error-prone, as in the cases of DNA pol η (XP-V/Rad30) bypass of a cyclobutane pyrimidine dimer or DNA pol V (UmuD'2C) bypass of a TT (6-4) photoproduct, respectively (2, 3). Although DinB is the only Y-family DNA polymerase conserved among all domains of life, the biological rationale for this striking conservation has remained enigmatic (4). Here we report that the E. coli dinB gene is required for resistance to certain DNA damaging agents that form adducts at the $N^2$-position of deoxyguanosine (dG). We demonstrate that DinB (DNA pol IV) catalyzes accurate TLS over one such $N^2$-dG adduct ($N^2$-furfuryl-dG), and that DinB and its mammalian ortholog, DNA pol κ, insert deoxycytidine (dC) opposite $N^2$-furfuryl-dG with 10-15 fold greater catalytic proficiency than opposite undamaged dG. We also show that mutating a single amino acid, the 'steric gate' residues of DinB (F13V) and of its archaeal homolog Dbh (F12A), separates their abilities to perform TLS over $N^2$-dG adducts from their abilities to replicate an undamaged template. These data lead us to propose that DinB and its orthologs are specialized to catalyze relatively accurate TLS over certain $N^2$-dG adducts that are ubiquitous in nature, that lesion bypass occurs more efficiently than synthesis on undamaged DNA, and that this specificity may be achieved at least in part through a lesion-induced conformational change.
Results and Discussion

Although DinB is strongly upregulated as part of the SOS DNA damage response and dinB+ function has been implicated in untargeted mutagenesis of λ phage, adaptive mutagenesis, and -1 frameshift mutagenesis when dinB+ is overexpressed in exponential phase (5-9), these phenotypes seemed inadequate to account for the strong conservation of the DinB subfamily of DNA polymerases during evolution. We therefore exposed an E. coli strain bearing a precise deletion of the dinB gene to a variety of DNA damaging agents to gain insights into DinB function in vivo. The ΔdinB strain displays a remarkable sensitivity to nitrofurazone (NFZ) (Figure 1) that can be complemented in trans by dinB+ under its native promoter on a low copy-number plasmid (Figure 2). The killing curve of a ΔumuC strain is indistinguishable from wild type (Figure 1), indicating that DinB is responsible for most TLS over potentially lethal NFZ-induced adducts. The ΔdinB mutant shows an increased sensitivity to killing by 4-nitroquinoline-1-oxide (4-NQO) as well (Figures 1-2), but in this case TLS by UmuD'2C also makes a contribution to survival in a dinB+ background (Figure 1). Deletion of polB, which encodes DNA polymerase II and is also induced by the SOS response (10), does not increase sensitivity to either agent (data not shown).

Before forming stable N2-dG adducts in vivo, nitrofurans such as NFZ must be reduced and acetylated (11). Likewise, at least half of the adducts that 4-NQO also produces are N2-dG adducts (11, 12). To address whether the NFZ resistance of a dinB+ strain arises from N2-dG lesion bypass, wild-type DinB was expressed and purified from E. coli (Figure 3) and oligonucleotide substrates were constructed that contained a site-specific N2-furfuryl-dG (Figure 4), a structural analog of the major N2-dG adduct formed by NFZ. While E. coli DNA polymerase I is strongly blocked by this lesion (Figure 5A), DinB has strikingly different properties. In the presence of all four deoxyribonucleotide triphosphates, DinB
displays an increased catalytic proficiency on the $N^2$-furfuryl-dG template relative to an undamaged template (Figure 5B). Standing-start (13) experiments (described in Figure 6) indicate that DinB is 15-fold more proficient at adding dC opposite $N^2$-furfuryl-dG than opposite undamaged dG (Figure 5C). DNA pol κ, the mammalian DinB ortholog, is also significantly more proficient at adding dC opposite $N^2$-furfuryl-dG than undamaged dG (Figure 5D) indicating that this remarkable specificity has been conserved in eukaryotes. Furthermore, DinB bypass of $N^2$-furfuryl-dG is not only proficient, but also quite accurate (Figure 5E). This is achieved in part from a preference for correct dC insertion and in part from a preference for elongating from dC correctly paired with $N^2$-furfuryl-dG (Table 1).

These observations suggest that a major physiological role of DinB and its orthologs is to catalyze accurate TLS over certain $N^2$-dG adducts. This hypothesis received strong support from our construction of a dinB mutant that virtually eliminates DinB’s ability to perform this type of TLS without impairing its ability to replicate undamaged DNA. We designed the dinB mutant after constructing a Sulfolobus solfataricus Dpo4-based (14, 15) homology model of DinB encountering an $N^2$-furfuryl-dG lesion (Figure 7). We noticed a pocket in the enzyme next to the template base that could potentially accommodate the $N^2$-furfuryl-dG adduct, bringing it into proximity with Phe13. This residue corresponds to Phe12 of the S. acidocaldarius DinB homolog (Dbh), the ‘steric gate’ that prevents the improper incorporation of rNTP substrates by that enzyme (16). Speculating that an active site rearrangement involving the $N^2$-furfuryl-dG adduct, the Phe13 steric gate residue, and the incoming nucleotide might favor catalysis, we mutated the planar hydrophobic Phe13 steric gate to a sterically different but still hydrophobic valine residue.

DinB(F13V), purifies to homogeneity indistinguishably from wild-type DinB (Figure 8), proved to be a most interesting mutant protein. Primer-extension assays revealed that DinB(F13V) is virtually unable to carry out TLS over $N^2$-furfuryl-dG, while its activity on
undamaged DNA is largely unaffected (Table 1, Figure 9A). The F13V mutation has a modest effect on DinB’s ability to discriminate against ribonucleotides, increasing the frequency of their misincorporation from $<10^{-5}$ (limit of detection) to ca. $10^{-3}$. Since the steric gates of all DinB orthologs are Phe or Tyr residues, we wondered whether the corresponding mutation in these enzymes would likewise separate their TLS activities from their ability to replicate undamaged templates. We therefore assayed the archaeal DinB ortholog Dbh and its steric gate mutant Dbh(F12A) (16) on $N_2$-furfuryl-dG and undamaged templates. While wild-type Dbh replicates both templates with comparable efficiencies at 37°C, the F12A derivative displays disproportionately reduced activity on the damaged template (Figure 9B-C).

To determine whether the F13V mutation specifically eliminates $N_2$-dG lesion bypass without affecting other properties of DinB, we examined bypass of two other well-studied lesions, (+)-trans-anti-benzo[a]pyrene-$N_2$-dG ($N_2$-B[a]P-dG) (Figure 4b) and a tetrahydrofuran abasic site analog (17, 18) (Figure 4c). Although DinB catalyzed bypass of the $N_2$-B[a]P-dG lesion is inefficient (17) compared to bypass of $N_2$-furfuryl-dG, the F13V mutation similarly eliminates its ability to perform this type of TLS (Figure 9D). Furthermore, just like the wild-type enzyme (18), DinB(F13V) is unable to efficiently bypass a tetrahydrofuran abasic site analog (Figure 10), indicating that the F13V mutation has not relaxed the specificity of DinB in vitro. While it is possible that the F13V mutation also affects DinB bypass of some other lesion, these data indicate that it specifically eliminates bypass of $N_2$-dG lesions.

To establish whether $N_2$-dG lesion bypass is required for dinB-dependent resistance to NFZ and 4-NQO, we examined the ability of a low copy number plasmid carrying the $dinB(F13V)$ allele under its own promoter to complement a $ΔdinB$ strain for NFZ and 4-NQO resistance (Figure 2A-B). While the mutant protein is expressed from this plasmid in vivo
(data not shown), pdinB(F13V) is unable to complement NFZ or 4-NQO resistance, an observation consistent with an \(N^2\)-dG adduct being responsible for NFZ lethality.

Furthermore, pdinB(F13V) exacerbates the sensitivity of the \(\Delta\text{dinB}\) strain to these agents, even to a greater degree than a plasmid encoding a catalytically inactive DinB(D103N) mutant protein (pdinB003) (19) (Figure 2A-B), suggesting that it is interfering with some cellular process that can otherwise contribute modestly to NFZ resistance. The plasmid-borne \(\text{dinB}(F13V)\) allele does not affect viability of the \(\text{dinB}^+\) strain, but it has a dominant negative effect on survival after treatment with either NFZ or 4-NQO (Figure 11). We conclude that this dominance is largely due to an impairment of TLS rather than ribonucleotide misincorporation into DNA by DinB(F13V) because dominance is still observed in an \(\text{rnhB}\) mutant (20) (Figure 11) and the mutant enzyme still favors dNTP incorporation \textit{in vitro} (Table 1). Taken together, our data indicate that the aromatic steric gate residue of DinB is required for TLS over \(N^2\)-dG adducts both \textit{in vivo} and \textit{in vitro}.

\textit{In vivo}, does DinB prevent mutagenesis caused by NFZ or 4-NQO, or does it promote mutagenesis, a behavior frequently attributed to Y-family DNA polymerases? Loss of \(\text{dinB}^+\) does not alter the frequency of NFZ-induced rifampicin resistant (Rif') mutations (29.6 \(\pm\) 9.7 \(\times\) \(10^{-9}\) for \(\text{dinB}^+\) vs. 31.4 \(\pm\) 11.8 \(\times\) \(10^{-9}\) for \(\Delta\text{dinB}\)). Given the drastically greater ability of a \(\text{dinB}^+\) strain to survive NFZ treatment compared to a \(\Delta\text{dinB}\) strain, this indicates that DinB is not a mutagenic polymerase when bypassing NFZ-induced lesions that are lethal in its absence. Loss of \(\text{dinB}^+\) function results in an increase in the frequency of Rif' mutants upon 4-NQO treatment (48.1 \(\pm\) 1.7 \(\times\) \(10^{-9}\) for \(\text{dinB}^+\) vs. 453 \(\pm\) 181 \(\times\) \(10^{-9}\) for \(\Delta\text{dinB}\)). This implies that DinB carries out accurate TLS over a class of lethal 4-NQO-induced lesions that are bypassed in a more error-prone fashion in its absence, most likely by UmuD'\(\_2\)C (21). These data indicate that the \(\text{dinB}\) gene product bypasses the lethal lesions generated by NFZ and 4-
NQO with unexpectedly high fidelity *in vivo*, thus resembling the behaviour of DNA pol η when bypassing cyclobutane pyrimidine dimers (2, 22).

The *in vitro* data discussed above indicate that the F13V mutation virtually eliminates DinB’s ability to bypass \( N^2 \)-furfuryl-dG and does not relax its specificity with respect to which lesions it can bypass, but does DinB(F13V) replicate undamaged DNA with reduced fidelity? Using a set of \( \Delta \)dinB strains carrying various plasmid-borne dinB alleles (Figure 2A-B), we examined the frequencies of spontaneous and NFZ-induced mutation to Rif\(^r\). We observed no increase in the frequency of spontaneous or NFZ-induced Rif\(^r\) mutations between the \( \text{dinB}^+ \) and \( \text{dinB(F13V)} \) alleles (Figure 2C), indicating that the F13V mutation does not result in DinB(F13V) becoming a mutator polymerase. We also compared the effect of \( \text{dinB(F13V)} \) on spontaneous mutation to that of \( \text{dinB}^+ \) using derivatives of the strain CC102 (23). This strain carries a lacZ allele that reverts by a GC→AT transition, the most frequent DinB error we detected *in vitro*. Here again, we detected no increase in Lac\(^+\) reversion between the \( \text{dinB(F13V)} \) derivative \((11 \pm 5 \times 10^{-9})\) and that of the \( \text{dinB}^+ \) strain \((8 \pm 5 \times 10^{-9})\), indicating that the F13V mutation does not decrease the fidelity of DinB.

DinB’s highly conserved steric gate residue, Phe13, clearly plays a critical role in bypass of \( N^2 \)-dG adducts, but further work will be required to establish whether it participates in an \( N^2 \)-dG lesion-induced conformational change that permits preferential replication of this type of damaged DNA template. Nevertheless, certain of our observations are consistent with there being such a lesion-induced conformational change (Table 1). These include: i) wild type DinB’s property of detectably incorporating low levels of rNTPs only when acting on the \( N^2 \)-furfuryl-dG bearing template, ii) no detectable increase in rNTP incorporation by DinB(F13V) on the \( N^2 \)-furfuryl-dG bearing template relative to an undamaged control, and iii) a lower apparent \( K_m \) for dCTP when DinB is bound to an \( N^2 \)-furfuryl-dG standing-start
primer template rather than to the corresponding dG template, coupled with a higher apparent
$V_{\text{max}}$ for the damaged primer/template itself.

DinB may have a role as a mutator polymerase under certain conditions of biological
stress, or in certain sequence contexts (7, 19). However, since other amino acids can serve as
steric gates in other DNA polymerases (24), the evolutionary conservation of this aromatic
steric gate residue with a critical role in TLS suggests that $N^2\text{-dG}$ adduct bypass is an
important and physiologically relevant property of the DinB subfamily of Y family DNA
polymerases. Present at levels comparable to 8-oxo-G in vivo, $N^2\text{-dG}$ adducts are formed
from byproducts of diverse cellular processes, including lipid peroxidation (25).
Furthermore, there is evidence that these minor groove adducts may be recalcitrant to
excision repair (26). In mammalian cells the $N^2\text{-dG}$ adduct of acetylaminofluorine persists
despite being the least common dG isomer formed by the carcinogen (27). Finally, our
results emphasize that Y-family DNA polymerases, while perhaps relatively error-prone
under certain conditions (7, 8, 10, 19), can also be specialized for proficient and accurate
replication of a particular class of damaged DNA.
Methods

Strains and plasmids

Strain KM1190 with the genotype AB1157 lexA51(Def) sulA11 was from K. McEntee (UCLA). The strain AB1157 lexA51(Def) sulA11 ΔdinBW2::cat bears a precise deletion of the dinB gene and was constructed by the method of Wanner et al (28) using the primers

\textit{dinBW2F}: \ 5'-ACTCGTTAAATGCTGAATCTTTACGCATTTCTCAAA \\
\textit{CCGTGTTAGCGCTGGAGCTGGAGCTGCTTC-3'} and \textit{dinBW2R}: \ 5'-GTGATATTGACCGGATTCTTTATACAGGAAAT-3'. A precise deletion of the rnhB gene was constructed similarly, using the primers \textit{rnhBDF}: \ 5'-GCTGCAATGCGGATGCAGGCGGCACAAGGCTGGAGGAGTTAGCAATGATTCTTGAGCTGCTTC \\
\textit{AGCGAGAATTCGATGCTATACCTCTTATAGGAGGAAT-3'}. The plasmids pWSK29 (vector), pGY768 (p\textit{dinB}'), and pGY782 (p\textit{loc,dinB}') have been described previously (8). To express DinB protein, the \textit{dinB} gene was amplified from pYG782 by PCR using the primers \textit{5'}- \texttt{GAGGTGACATATGCGTAAAATCATTCATGTGG-3'} and \textit{3'-CCTGGATCCCGTAACTCAGTGATATTGACC-3'}. The amplification fragment was digested with NdeI and BamHI and ligated into an expression vector, pET11T (29), that was subjected to the same enzymatic treatment to form pDFJ1. Derivatives of pDFJ1 and pYG768 expressing DinB(F13V) were made from the appropriate parent plasmid using a Quickchange kit from Stratagene according to the manufacturer's recommendations. Expression of soluble DinB from pYG768 and pYG768-F13V was verified by immunoblot (data not shown). Strains were grown at 37°C in Luria Bertani (LB) broth unless otherwise noted, and ampicillin (100 μg/mL) was added when required.
Synthesis of \( N^2 \)-furfuryl-dG containing template

Briefly, the 2-fluoro-\( O^\alpha \)-(trimethylsilylethyl)-2'-deoxyinosine phosphoramidite (ChemGenes, Wilmington, MA), was incorporated into the oligonucleotide 5'-
GAAGACCTXGGCGTCC-3' using phenoxyacetal protected phosphoramidites (30), after which the resin was treated with 0.1 N \( \text{NaOH} \) at 25°C for 8 h, and the pH was adjusted to 7 with dilute acetic acid prior to desalting (with a 10 mL water wash) by Sep-Pak (Millipore).

After lyophilization, 85 nmol of oligonucleotide was dissolved in 80 uL dimethyl sulfoxide, 40 uL \( N,N \)-diisopropylethylamine and 20 uL furfurylamine (Aldrich), which was mixed in an Eppendorf Thermomixer at 60°C for 12 h. The trimethylsilyl ethyl group was removed by adding 800 uL 5% acetic acid directly to the mixture, which was incubated at room temperature for 3.5 h, followed by addition of 3 mL water and neutralization with \( \text{NaOH} \) prior to desalting on a Sep-Pak. Matrix assisted laser desorption ionization-time of flight (MALDI-TOF) was performed as described (31), which revealed 20% of the oligonucleotides retained the trimethylsilyl ethyl group; therefore, complete removal of the group was achieved by treating the oligonucleotides with 0.25% acetic acid for 1 h at room temperature, and neutralization with \( \text{NaOH} \). Oligonucleotides were purified by reversed-phase HPLC using an analytical column (Microsorb-MV 100-5 C18 250 x 4.6 mm, Varian) at a flow of 1 mL/min and a gradient of 0 to 30% B over 60 min (A: 0.1 M triethylammonium acetate; B: 100% acetonitrile). The retention times were 29.1 min for the furfuryl and 28.5 min for the tetrahydrofurfuryl oligonucleotides.

Protein purification

DinB and DinB(Fl13V) were expressed and purified in the same fashion. BL21(DE3) pLysS cells harboring pDFJ1 were grown to an \( \text{OD}_{600} \) of 0.9 at 37 °C subsequent to induction with 1 mM IPTG at 30 °C for 3.5 h. Cells were harvested and lysed using lysozyme,
centrifuged for 45 min at 20,000 x g and the supernatant was treated with deoxyribonuclease I (Sigma). This was purified on a monoS column (Pharmacia) with a 0-1 M KCl gradient in 50 mM HEPES pH 7.5, 5 mM MgCl₂, 2 mM DTT, and 5% glycerol (buffer A) with DinB eluting at 300 mM KCl. DinB-containing fractions were identified by SDS-PAGE, diluted two-fold in buffer A with 1 M (NH₄)₂SO₄, and loaded onto a phenylsepharose column equilibrated in the same buffer. The column was washed and DinB was eluted with a 1-0 M gradient of (NH₄)₂SO₄ in buffer A, resulting in a DinB peak at 50 mM. The DinB-containing fractions were again identified by SDS-PAGE and dialyzed against buffer A with 100 mM KCl before storage at -80 °C.

Sensitivity and mutation frequency determination

From single colonies, strains were grown for 16 h to saturation, and then diluted 1:1000 into LB and grown to 5 x 10⁹ cfu/mL. From this freshly saturated culture, dilutions were plated on LB agar containing ampicillin and 0-15 μM NFZ or 4-NQO. NFZ and 4-NQO stock solutions were freshly prepared in N,N-dimethyl formamide. Approximately 10 of these colonies were suspended in M9 salts and deposited on plates containing rifampicin (100 μg/mL) to determine the number of Rif' mutants. This number was corrected for the number viable cells in each colony. GC to AT transitions were measured using a ΔdinB derivative of the CC102 episome (23).

Template synthesis and construction

The N²-furfuryl-dG adduct was made using a postsynthetic derivatization approach (30), described in detail in supplemental information. MALDI-TOF mass spectrometry of the purified oligonucleotide revealed a mass of 4986.26 (4986.27 calculated) for the single, negatively charged moiety (Figure 12). The 16-mer lesion-bearing oligo was ligated to 5'-
GGTTACTCAGATCAGGCCTGC-3’ at the 5’ end and 5’-
GGCTGCAGCTGTACTATCATATGC-3’ at the 3’ end using standard protocols and gel
purified to remove the ligation scaffolds 5’-AGGTCTTGCAGGCCTGAC-3’ and 5’-
CAGCTGCAGCCGGACGC-3’. The Benzo[a]pyrene lesion is in the sequence context 5’-
GACTACGTACTGTCACATXCCACACGCTATCTGGCCAGATCCGC-3’.

Primer extension assays

Assays were performed and quantitated using either standing or running start primers
of the sequences 5’-GCATATGATAGTACAGCTGCAGCCGGACGC C-3’ or 5’-
GCATATGATAGTACAGCTGCAGCCGGACGC-3’ respectively, for all templates except
the N2-B[a]P-dG bearing substrate (13). For that substrate, the primer 5’-
GCGGATCTGGCCAGATAGCTGT-3’ (running) was used. Briefly, assays were
conducted in a 10-μL volume with either 1, 10, or 50 nM DinB, 10 nM primer/template,
either 250 μM dNTPs or 0-2,000 μM dCTP, and in 50 mM Hepes pH 7.5, 100 mM KCl, 7.5
mM MgCl2, 5% glycerol, and 0.1% BSA. Reactions were initiated with dNTPs and
quenched after incubation for 15 min (or as noted in the figure legends) at 37 °C. Percent
extension is defined as the percent of primers that are extended past the lesion. The same
conditions were used to assay Dbh and pol κ, except using 10 nM and 2 nM enzyme,
respectively. Products were analyzed on a 12% denaturing polyacrylamide gel and
quantitated on a phosphorimager (Molecular Dynamics).

Acknowledgements

We wish to acknowledge C. Joyce and A. DeLucia (Yale) for the gift of Dbh and
Dbh(F12A) as well as E. Friedberg, J. Stirman, and P. Fischhaber (UT-Southwestern) for the
gift of M. musculus pol κ. T. Nohmi (NIHS, Tokyo) provided pYG768 and pYG782. We
also thank N. Geacintov (NYU) for providing benzo[a]pyrene damaged substrate and A.
Banerjee (Harvard) for assistance with Figure 7. J. Tuttle (Duke) and J. Bowers (MIT) offered advice with protein purification. D.F.J performed the protein purification, lesion bypass assays, and hypothesized the involvement of the steric gate residue in TLS. V.G.G. discovered the sensitivity of a ΔdinB strain to NFZ and 4-NQO and performed the mutagenesis experiments. J.C.D constructed and purified the \( N^2 \)-furfuryl-dG-containing oligonucleotide substrate. D.F.J., V.G.G., and G.C.W. wrote the manuscript. This work was supported by NIH grants to G.C.W. and J.M.E.
References


26. Zewail-Foote, M., V. S. Li, H. Kohn, D. Bearss, M. Guzman, and L. H. Hurley. 2001. The inefficiency of incisions of ecteinascidin 743-DNA adducts by the UvrABC nuclease and the unique structural feature of the DNA adducts can be used to explain the repair-dependent toxicities of this antitumor agent. Chem Biol 8:1033-49.


Table 1. Kinetic parameters for \(N\)-glycytyl-DC bypass mismatch extension and INTp insertion by wild-type DnaB and DnaB(F13E) proteins. A\(_{\text{max}}\) and \(K_m\) for the DNA substrates

<table>
<thead>
<tr>
<th>N-Glyc</th>
<th>INTp</th>
<th>N-Glyc</th>
<th>INTp</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 \times 10^7</td>
<td>2 \times 10^7</td>
<td>2 \times 10^7</td>
<td>2 \times 10^7</td>
</tr>
<tr>
<td>(K_m = 2 \times 10^7)</td>
<td>(K_m = 2 \times 10^7)</td>
<td>(K_m = 2 \times 10^7)</td>
<td>(K_m = 2 \times 10^7)</td>
</tr>
<tr>
<td>(A_{\text{max}} = 2 \times 10^7)</td>
<td>(A_{\text{max}} = 2 \times 10^7)</td>
<td>(A_{\text{max}} = 2 \times 10^7)</td>
<td>(A_{\text{max}} = 2 \times 10^7)</td>
</tr>
</tbody>
</table>

*Note: The data shows a 20% increase in activity for N-glycytyl-DC. *
Figure 1. Sensitivity of strains bearing a deletion of dinB to NFZ and 4-NQO. a, NFZ sensitivity. A lexA(Def) background was used to minimize differences in SOS-induction between each agent. b, 4-NQO sensitivity. Error bars represent the the standard deviation as determined from three experiments.
Figure 2. Importance of DinB F13 residue *in vivo*. a, pdinB(F13V) is unable to restore NFZ resistance in the ΔdinB strain, like pdinB+, but instead exacerbates the sensitivity, like pdinBOO3, which encodes a catalytically-inactive DinB(D103N) protein. Error bars represent one standard deviation determined from three experiments. b, pdinB(F13V) and pdinBOO3 also exacerbate the sensitivity of the ΔdinB strain to 4-NQO. Error bars represent one standard deviation. c, Spontaneous and induced mutation frequencies per 10⁹ bacteria to Rif⁺. Error is the standard deviation from 8 experiments.
Figure 3. SDS-PAGE of purified native DinB. Approximately 30 mg of enzyme are obtained per liter from the purification protocol. Lane 1 contains molecular weight markers, and lanes 2-6 contain 1.9, 9.5, 19, 38, 57 µg of purified DinB. The specific activity of the enzyme on undamaged template is ca. 1,000 pmol min⁻¹ mg⁻¹.
Figure 4. Structures of lesions used in this study. a, $N^2$-furfuryl-dG. b, $N^2$-[B[a]]P-dG. c, THF, an apurinic site analog. The proposed formula for the modified base containing the $N^2$-dG adduct produced by nitrofurazone is $C_{12}H_{14}N_{6}O_{3}$. 
Figure 5. Bypass of $N^2$-furfuryl-dG. a) Primer (lane 1) extension products of E. coli pol I (5 nM) on undamaged dG (lane 2, 95.3% extension) and $N^2$-furfuryl-dG-damaged templates (lane 3, 8.2% extension; see Supplementary Information). b) Running-start primer extension reactions with 1, 10, or 50 nM DinB protein and 250 μM dNTPs. Lanes 1-3, undamaged dG template (0.03%, 2.2%, 81.8% extension); lanes 4-6, $N^2$-furfuryl-dG-damaged template (0.05%, 65.5%, 91.1% extension). c) Plot of initial reaction velocity vs. initial [dCTP] in standing-start assays on undamaged dG (closed circles) and $N^2$-furfuryl-dG-damaged templates (open circles). Error bars represent one standard deviation determined from three reactions. d) As in c), but using mammalian DinB ortholog pol κ. e) Fidelity of DinB bypass of $N^2$-furfuryl-dG measured using standing-start incorporation and extension assays (13). The error of these measurements is ~20%.

<table>
<thead>
<tr>
<th>Nucleotide</th>
<th>$f_{inc} \times f_{ext}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>G$^*$-dCTP</td>
<td>1</td>
</tr>
<tr>
<td>G$^*$-dTTP</td>
<td>$1.6 \times 10^{-3}$</td>
</tr>
<tr>
<td>G$^*$-dATP</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>G$^*$-dGTP</td>
<td>$&lt;10^{-3}$</td>
</tr>
</tbody>
</table>
Primer Extension Experiments

Figure 6. Schematic of running start, standing start, and mismatch extension assays. Template is shown in blue and radiolabeled primer is shown in red. In running start assays, DinB encounters the lesion after initial polymerization on an undamaged portion of the template (1 base pair for all experiments in this publication). In contrast, in standing start experiments DinB initially polymerizes opposite the lesion. For fidelity measurements, standing start experiments were performed with a single dNTP at a time. Mismatch extension experiments are performed as the standing start experiments, but with primers that have different bases (X) at their 3' ends opposite to the lesion. The ability of DinB to extend from the correctly paired vs. mismatched base is then determined by assaying the addition of the next correct base (dA) on each substrate.
Figure 7. A homology model of DinB, constructed in part with SWISS-Model, reveals an $N^2$-dG lesion binding pocket and suggests a mechanism for rate enhancement on a damaged template. The templating base is shown in yellow, the incoming nucleotide in blue, the 3' base of the primer in green, and the 'steric gate’ residue in red. a) The conformation of the incoming nucleotide is not favorable to phosphodiester bond formation. b) A space-filling representation of the model viewed from the minor groove shows a pocket that is correctly positioned to accommodate an $N^2$-dG adduct and that the 'steric gate’ residue is at the back of this cavity. c) A schematic representation of a model for rate enhancement suggesting that binding of the $N^2$-furfuryl-dG adduct stimulates a conformational change in the active site, propagated through the 'steric gate’ residue, that positions the incoming nucleotide for efficient phosphodiester bond formation, providing rate enhancement on a damaged template. d) A space-filling representation of the lesion-bound model shows that the pocket is occupied by the furfuryl moiety.
Figure 8. SDS-PAGE of purified DinB(F13V). Approximately 27 mg of enzyme are obtained per liter in this purification. Lane 1 contains molecular weight markers, and lanes 2-8 contain 0.9, 2.3, 4.5, 9, 13.5, 18, 27 μg of purified DinB(F13V). The specific activity of the enzyme on undamaged template is ca. 1,150 pmol min⁻¹ mg⁻¹.
Figure 9. A single mutation in DinB or its archaeal ortholog Dbh separates their TLS and DNA polymerase activities. a) Running start primer extension reactions using 1, 10, and 50 nM DinB(F13V) on undamaged dG (lanes 1-3, 0%, 79.1%, 86.8% extension) and N2-furfuryl-dG-damaged templates (lanes 4-6, 0%, 7.5%, 16.3% extension). DinB(F13V) retains DNA polymerase activity but is compromised for TLS. b) Plot of relative initial velocity vs. initial [dCTP] for Dbh on undamaged dG (closed circles) and N2-furfuryl-dG (open circles). c) As in b) for Dbh(F12A); activity is disproportionately reduced on the N2-furfuryl-dG-damaged template. d) Running-start primer extension assays with 1, 10, and 50 nM DinB or DinB(F13V) on undamaged and N2-B[a]P-dG-damaged templates. Lanes 1-12 show 29.8%, 62%, 88.7%, 1.3%, 17.8%, 32.4%, 11.5%, 35.5%, 53.8%, 0%, 0%, and 0.2% extension.
Figure 10. Products of running start primer extension reactions (30 min) with 1, 10, and 50 nM DinB and DinB(F13V) on undamaged and tetrahydrofuran bearing templates reveal that the F13V mutation does not render DinB able to bypass this lesion. Primers are 14.8%, 70.1%, 73.2%, 3.9%, 6.2%, 6.9%, 7.5%, 24.5%, 57.8%, 2.6%, 6.2%, and 10.2% extended.
Figure 11. Dominant negativity of pdinB(F13V). The dominant negative effect of pdinB(F13V) on survival is not significantly altered in an rnhB mutant, indicating that the mechanism of the dominant negative effect is not merely through improper incorporation of rNTPs by DinB(F13V) in vivo.
Fraction of the sample much smaller than the bypass efficiency observed (ca. 5%).

The purified oligonucleotide observed a MALDI-TOF mass of 4.986.26, which is consistent with the sequence predicted from the observed MALDI-TOF spectra. The purified material was then subjected to HPLC purification on a Vydac C18 reversed-phase column. The purified material was then analyzed by MALDI-TOF mass spectrometry, which showed a peak corresponding to the sequence predicted from the sequence of the purified material.

Figure 12: MALDI-TOF mass spectra of the N-Furinyl-5'-adduct. Oligonucleotides were synthesized as described in the Supplementary Material and were purified by reversed-phase HPLC. The purified material was then subjected to MALDI-TOF mass spectrometry.
Chapter 3

A Cluster of Aromatic Amino Acids Controls DinB Function in Translesion Synthesis

Abstract

Y-family DNA polymerases catalyze the replication of imperfect DNA templates with low to moderate fidelity. Certain of these polymerases, including *Escherichia coli* DinB and its mammalian ortholog DNA polymerase κ, display increased activity on damaged templates relative to undamaged DNA. The detailed mechanisms that contribute to this remarkable property are poorly understood. Here we report that a cluster of aromatic hydrophobic amino acids surrounding the steric gate residue of *Escherichia coli* DinB dictates several features of its function. In particular, Tyr 79 plays a critical role in determining the length of products produced by DinB. Moreover, the DNA synthesized by DinB is subject to the proofreading exonuclease activity of the epsilon subunit of DNA polymerase III. Taken together, these data suggest both that residues within the active site of DinB collaborate to control detailed aspects of its polymerase function and that the length of products produced by DinB is tuned not only to limit potential introduction of mutations but also to avoid reversal by the proofreading activity of DNA polymerase III.
Introduction

Y-family DNA polymerases exist in virtually all organisms and possess the remarkable ability to duplicate imperfect templates that otherwise stall DNA replication (1, 2). Such broadened substrate specificity comes at a mutagenic penalty, however, as Y-family DNA polymerases display lower fidelities on undamaged DNA templates than their replicative counterparts (3). Organisms therefore place a premium on restricting the improper access of Y-family DNA polymerases to primer termini where translesion synthesis (TLS) is required (4, 5).

Studies of Y-family DNA polymerase function in both prokaryotes and eukaryotes have revealed diverse mechanisms that contribute to regulation of their activity (1). In *Escherichia coli*, interactions with the β processivity clamp subunit of DNA polymerase III (6-8) and the *umuD* gene products UmuD₂ and UmuD₂' (9-13) strongly influence the function of both DinB (DNA pol IV) and UmuD₂'C (DNA pol V). Other protein-protein interactions, notably with SSB (11) and RecFOR (14) also play critical roles in mediating TLS.

Several Y-family DNA polymerases also display a remarkable intrinsic preference for catalytic action on certain damaged substrates. *In vitro* studies reveal that *Saccharomyces cerevisiae* pol η acts preferentially on a cyclobutane pyrimidine dimer containing substrate (15, 16), whereas *E. coli* DinB and its mammalian ortholog *M. musculus* DNA polymerase κ (pol κ) each display a strikingly elevated catalytic proficiency on certain N²-dG damaged substrates (17).

At the molecular level, the structural features within Y-family polymerases that contribute to their remarkable catalytic capabilities are wide-ranging. Relatively accommodating active sites (18, 19) and divergent C-terminal ‘little finger’ domains (20) have been shown to facilitate TLS by several Y-family DNA polymerases. Specialized TLS function does not arise
merely from open active sites, however. Several recent studies have suggested that the active sites of both pol \( \kappa \) (21) and DinB (13) may be somewhat closed under many conditions, and that this may occur at least in part through interaction with other cellular proteins in some cases.

Although limited aspects of the genetic mechanisms governing access of certain Y-family polymerases to primer termini are understood, comparatively little is known about the intrinsic biochemical preference of Y-family polymerases for activity on damaged substrates. Certain amino acids in the active sites of both pol \( \eta \) and DinB are critical TLS function specifically. In *E. coli* DinB and *S. cerevisiae* pol \( \eta \), mutation of Phe 13 or Phe 34, respectively, results in a polymerase variant that is perfectly proficient for normal DNA synthesis but unable to efficiently catalyze TLS (17, 22). Other specific determinants of TLS function likely exist, some of which may be common to all Y-family DNA polymerases and others of which may be restricted to certain subfamilies of TLS polymerases.

In its most basic sense lesion bypass involves insertion of a nucleotide opposite an adducted base and extension from that nucleotide in a subsequent addition. Physiologically, however, these specialized polymerases synthesize DNA for several nucleotides beyond the site of DNA damage to complete the process of translesion synthesis and avoid reversal by the proofreading activity of the \( \varepsilon \) subunit of the replicative DNA polymerase (23). Therefore, regulating the length of DNA products during translesion synthesis is of critical importance to the cell, as those that are too small risk being removed by the \( \varepsilon \) proofreading subunit of pol III and those that are too long carry increased mutagenic potential. Here we report that a cluster of aromatic hydrophobic residues surrounding the 'steric gate' residue of DinB dictate crucial aspects of its TLS function. In particular, we demonstrate that Tyr 79, which is invariant among DinB orthologs, influences the extension steps of DinB-catalyzed TLS. Moreover, our data
indicate that the proofreading activity of the ε subunit of DNA polymerase III plays an important role in processing immature TLS intermediates produced by DinB. These data, taken together with an analysis of the range of DinB’s preferential activity on a damaged substrate, indicate that the size of DNA fragments synthesized by DinB is exquisitely tuned not only to evade reversal by proofreading function but also to avoid the excessive introduction of genomic instability during TLS.
Results

A cluster of aromatic hydrophobic residues surrounds the steric gate of DinB

We have previously reported that the steric gate residue of DinB is indispensable for its TLS function (17) in addition to its well established role in discrimination against improper rNTP incorporation (24), a result that arose from our analysis of a model of DinB encountering an \(N^2\)-furfuryl-dG lesion. We wondered whether additional residues might contribute to the preferential activity of DinB on a damaged substrate. When we examined our model in greater detail, we noted that the steric gate residue of DinB (Phe 13) is surrounded by a cluster of aromatic hydrophobic residues: Phe 12, Phe 76, and Tyr 79 (Figure 1A). In our homology model, Tyr 79 appears to stack with DinB's Phe 13 steric gate, whereas Phe 76 and Phe 12 to be oriented with a rotation of roughly 90 degrees relative to it. Provocatively, evolutionary conservation of Phe 12 is universal and this residue corresponds to Phe 34 of \(S. \textit{cerevisiae}\) pol \(\eta_1\), mutation of which eliminates its TLS function (22). Phe 76, in contrast, is more weakly conserved even among DinB orthologs. Strikingly, Tyr 79 is invariant among DinB orthologs from diverse organisms, while the structurally related but functionally distinct enzymes Rev1 and Rad30 possess a Phe residue at this position (Figure 1B). Curiously \(E. \textit{coli}\) UmuC also shares a tyrosine residue at this position.

Mutations in this pocket reveal its critical importance for DinB TLS function

To examine the effect of this cluster of aromatic hydrophobic residues on DinB function, we mutated each of Phe 12, Phe 76, and Tyr 79 to alanine, valine, and leucine and examined the ability of a plasmid encoding each variant to complement \(\text{dinB}^+\)-dependent sensitivity to the DNA damaging agent nitrofurazone (NFZ) in a \(\Delta\text{dinB}\) strain. Mutants of Phe 12 are unable to
fully complement NFZ sensitivity, but show a far less dramatic phenotype than the steric gate mutant dinB(F13V) (Figure 2). In contrast, mutation of Phe 76 has little discernable effect on complementation, suggesting a possible rationale for its comparative lack of evolutionary conservation.

Strikingly, mutation of Tyr 79 has a profound effect on sensitivity to both NFZ and the DNA damaging agent 4-nitroquinoline-1-oxide (4-NQO) (Figure 3A-B). A low copy number plasmid expressing DinB(Y79L) is unable to complement the NFZ sensitivity of a AdinB strain, instead conferring a further ca. 50-fold sensitivity beyond that of a AdinB strain transformed with an empty vector control. This exacerbated sensitivity is even greater than that conferred by expression of a steric gate mutant of DinB. Given that NFZ resistance requires DinB TLS function (17), these observations strongly suggest that Tyr 79 plays a crucial role in TLS over N²-dG adducts.

We wondered whether the role of Tyr 79 in mediating TLS substantially involved its hydroxyl group. Our initial studies indicated that a plasmid expressing a DinB(Y79F) variant is fully able to complement the NFZ sensitivity of a AdinB strain over the dose regime used in our experiments (Figure 3C). Curiously, when treated with ca. 10-fold higher doses of NFZ, we noted that the Y79F variant appears to confer remarkable resistance to the AdinB strain (Figure 3D). The reason for this effect is unclear but it may indicate that Y79 is either posttranslationally modified or that it participates in a specific hydrogen-bonding interaction under conditions of elevated NFZ exposure. Interestingly, deletion of dinB confers little further sensitivity in this dose regime, indicating that the Y79F variant likely interferes with an alternative cellular process that is less favorable with respect to cellular viability.
**Tyr 79 mutants catalyze lesion bypass**

We initially interpreted our observation that a DinB(Y79L) variant is unable to restore NFZ or 4-NQO resistance to a ΔdinB strain (Figure 3A-B) as reflecting the inability of that mutant protein to carry out TLS over the $N^2$-dG lesions produced by those agents. Unexpectedly, when we examined the ability of DinB(Y79L) to catalyze TLS over a site specific $N^2$-furfuryl-dG lesion *in vitro* we discovered that it is entirely able to bypass that lesion (Figure 4).

However, the lengths of the products synthesized by DinB(Y79L) are much shorter than those produced by wild-type DinB. Indeed, the vast majority of products formed by the action of DinB(Y79L) during TLS over a site specific $N^2$-furfuryl-dG adduct stall three nucleotides after the lesion. These data suggest that DinB(Y79L) represents a new class of mutant DinB proteins that are able to catalyze narrowly defined lesion bypass but are unable to finish the final primer extension steps that are required physiologically to complete TLS. Taken together with the profound NFZ sensitivity conferred to a ΔdinB strain by a plasmid expressing DinB(Y79L), these observations strongly indicate that the final extension step of TLS has at least as much physiological importance as lesion bypass itself.

In an effort to explain the unexpected NFZ resistance conferred by expression of DinB(Y79F) we examined the ability of this mutant protein to catalyze TLS *in vitro*. Although its expression results in a nearly 200-fold resistance to NFZ under certain conditions, we observed only a slight increase in its ability to bypass a site-specific $N^2$-furfuryl-dG lesion relative to wild-type DinB (Figure 4). This observation suggests that DinB(Y79F) does not confer NFZ resistance *in vivo* primarily due to an unforeseen unique aspect of its TLS activity, but rather that it likely does so via a different mechanism.
Dominance of DinB(Y79L) requires proofreading function

We wondered whether the profound NFZ sensitivity conferred by expression of DinB(Y79L) arises due to a futile cycle in which the TLS intermediates it produces are constantly reversed by the action of the ε proofreading subunit of DNA polymerase III. This appears to be the case, as we found that dominance conferred by DinB(Y79L) requires proofreading function. Expression of DinB(Y79L) in a proofreading proficient AB1157 strain of E. coli results in a striking sensitivity to NFZ that is abolished in a strain background in which the dnaQ gene, which encodes the ε proofreading subunit of DNA polymerase III, is deleted. These data indicate that proofreading function plays a critical role in mediating NFZ sensitivity conferred by expression of DinB(Y79L).

In principle, this phenomenon could arise as a result of dnaQ dependent dinB function, but deletion of dinB still confers sensitivity to NFZ in a ΔdnaQ background (Figure 5C), albeit of a slightly lesser magnitude than when it is deleted from a wild-type strain (17). These data suggest that the requirement of dnaQ+ for dinBY79L dominance arises not from loss of dinB+ function but rather from a novel feature of the DinB(Y79L) mutant protein, presumably arising from an altered interaction with proofreading function. We therefore examined the ability of a catalytically deficient DinB variant and the 'steric gate' DinB(F13V) variant to confer NFZ sensitivity in a ΔdnaQ strain. Because DinB(F13V) can catalyze normal DNA synthesis we anticipated that it could possibly require dnaQ+ to exert NFZ sensitivity, but that the dominance of a catalytically deficient DinB(D104N) variant with respect to NFZ sensitivity should not similarly require proofreading function. This is indeed the case as a plasmid bearing the catalytically deficient dinB003 allele [encoding DinB(D104N)] does still confer dominance in a
\texttt{dnaQ} strain but the same plasmid carrying the `steric gate' mutant \texttt{dinB}(F13V) does not (Figure 5D).

These observations suggest that \texttt{DinB}(Y79L) and \texttt{DinB}(F13V) may exert their dominance by inducing a futile cycle in which the immature TLS intermediates they produce are constantly destroyed by the action of the ɛ proofreading activity of DNA polymerase III.

Intriguingly, this model suggests that ɛ may have access to products synthesized by these DinB variants before wild-type DinB can be recruited to continue stalled DNA synthesis. Whether this is merely due to stochastic competition or rather due to an active mechanism of exclusion, it is perhaps provocative to note that the UmuD\textsubscript{2} protein has been shown to interact with both DinB (13) and the C-terminus of DnaQ (25).

**Defining the range of preferential DinB function**

These observations led us to examine DinB function in more detail. We performed nested primer extension assays (Figure 6) to determine the range of nucleotides around a site-specific \textit{N}^{2}\text{-furfuryl-dG} lesion over which DinB retains preferential activity relative to an undamaged control oligonucleotide of the same sequence. Individual standing start primer extension assays were performed using primers with 3’ termini starting 3 base pairs before the site of the lesion and ending 3 base pairs after the lesion. \(V_{\text{max}}\) and \(K_{m}\) values were measured for each addition (Table 1). Intriguingly, DinB appears to show two modes of preferential activity on the \textit{N}^{2}\text{-furfuryl-dG} bearing substrate. The catalytic proficiency of DinB (\(V_{\text{max}}/K_{m}\)) on the damaged template is comparable or slightly higher than it is on an undamaged control over the entire base pair range tested (Figure 7). Strikingly, the profoundly increased catalytic proficiency of DinB on the damaged template is restricted to nucleotide insertion opposite and
extension from the $N^2$-furfuryl-dG lesion. Although DinB is exceptionally proficient at catalyzing insertion opposite this lesion (Figure 7) (17), it is even more capable at extension from it, showing an astounding more than 25-fold increased catalytic proficiency on the damaged template for this reaction.

**DinB(Y79L) is impaired at extension from an elongated primer**

To uncover the basis for the inability of DinB(Y79L) to catalyze the final extension steps of TLS, we examined the activity of that enzyme in primer extension assays using a primer that is synthetically extended 3 base pairs beyond the site of the $N^2$-furfuryl-dG lesion in our oligonucleotide construct. If the inability of DinB(Y79L) variant to perform the extension steps of TLS arises merely from a general defect in processivity or requires lesion bypass to be revealed, it should behave similarly to the wild-type enzyme on this substrate. However, DinB(Y79L) is unable to efficiently extend from this synthetically elongated primer relative to wild-type DinB (Figure 8), suggesting that the defect of DinB(Y79L) arises largely from its specific inability to perform the extension steps of TLS.
Discussion

In this paper, we present evidence that the aromatic hydrophobic residues surrounding the lesion binding pocket within the active site of DinB play a key role in shaping TLS over $N^2$-dG adducts. Phe 12, Phe 76, and Tyr 79 are conserved to greater or lesser extents among evolutionarily diverse DinB orthologs and their mutation has differing effects on $\text{dinB}^+$-dependent NFZ resistance. Mutation of Phe 76 has no discernible effect on NFZ resistance, whereas mutation of Phe 12 results in a $\text{dinB}$ variant that is unable to fully complement the NFZ sensitivity of a $\Delta\text{dinB}$ strain. These data are consistent with the observations that F76 is modestly conserved among DinB orthologs and that F12 is universally conserved among Y-family DNA polymerases. Moreover, the equivalent residue of F12 in $S.\ cerevisiae$ pol $\eta$ has been shown to be a critical determinant for TLS over cyclobutane pyrimidine dimers catalyzed by that enzyme (22). Efforts are ongoing to determine whether this is similarly the case for DinB.

Strikingly, we discovered that mutation of Tyr 79 results in a DinB variant that is entirely unable to complement the NFZ sensitivity of a $\Delta\text{dinB}$ strain. In fact, expression of DinB(Y79L) from a low copy number plasmid under its native promoter exacerbates the sensitivity of a $\Delta\text{dinB}$ strain to NFZ by more than 50-fold. Based on our previous observation of similar behavior for the steric gate mutant of DinB, which is unable to catalyze efficient TLS but is perfectly active as a normal polymerase, we anticipated that DinB(Y79L) might behave similarly in a lesion bypass assay. To our complete surprise, the DinB(Y79L) variant is completely proficient at minimal lesion bypass, but appears unable to carry out the final extension steps of TLS.

In vivo, the extension steps of TLS are of critical importance as they prevent immature intermediates from degradation by the exonuclease activity of the $\varepsilon$ proofreading subunit of the
replicative DNA polymerase. Consistent with the notion that the striking NFZ sensitivity conferred by expression of DinB(Y79L) arises from its inability to form full-length TLS products, we observed that this variant does not confer NFZ sensitivity when expressed in a proofreading deficient strain background. Intriguingly, similar behavior was observed for the NFZ sensitivity conferred by the separation of function steric gate mutant DinB(F13V), whereas expression of a catalytically deficient DinB(D104N) variant still confers NFZ sensitivity in an AdnaQ strain background. These observations suggest that although DinB(Y79L) formally represents a new class of TLS deficient mutants, it shares similar features with the ‘steric gate’ DinB(F13V) mutant with respect to the mechanisms of its dominance in vivo.

Given its universal conservation among DinB orthologs, Tyr 79 may be particularly critical for DinB function. Expression of a DinB(Y79F) variant fully complements the NFZ sensitivity of an AdinB strain over most dose regimes, but confers approximately 200-fold greater resistance than expression of wild-type DinB when challenged with high doses of the drug. Moreover, DinB(Y79F) behaves as wild-type DinB with respect to its ability to catalyze TLS over a site specific $N^2$-furfuryl-dG lesion in vitro. These observations suggest that the hydroxyl group of Tyr 79 could participate in a critical hydrogen bonding interaction or even be posttranslationally modified under these high dose conditions in a manner that limits the ability of wild-type DinB to perform TLS.

Our results underscore the fact that physiologically, and even at the biochemical level, TLS is far more nuanced than mere insertion of a dNTP opposite to a damaged base. The fact that expression of DinB(Y79L) confers greater NFZ sensitivity than expression of either DinB(F13V) or a catalytically deficient DinB variant indicates that the final extension steps of
TLS are at least as physiologically important if not more so than insertion of a dNTP opposite to an adducted base.

Intriguingly, the striking intrinsic biochemical preference of DinB for action on a lesion-bearing template is most apparent for the insertion and primary extension steps of TLS. DinB is even more proficient at the first extension from $N^2$-furfuryl-dG than it is at insertion opposite to that lesion. The much more modest preferential activity of DinB on the lesion bearing template during the following extension steps of TLS may reflect a physiological premium placed on maintaining the integrity of the newly synthesized DNA. Indeed, the ability of DinB(Y79L) to confer dominance in a $\Delta$dinB+ wild-type E. coli strain but not in a $\Delta$dnaQ dinB+ strain indicates that the ε proofreading subunit of DNA polymerase III gains access to immature TLS intermediates before wild-type DinB can finish the extension steps of TLS.

Why does expression of DinB(F13V) or DinB(Y79L) result in such pronounced NFZ sensitivity relative to a $\Delta$dinB strain or even expression of a catalytically deficient DinB variant? Our data suggest that this hierarchy could be explained by the unique features of each mutant protein. In the absence of dinB+ function, a $\Delta$dinB strain must employ alternative mechanisms to contend with lethal DNA adducts produced by NFZ. A catalytically-deficient DinB variant can interfere with these other processes, but its defect in TLS is manifested far before dNTP insertion opposite a replication blocking lesion ever takes place. In contrast, both DinB(F13V) and DinB(Y79L) are perfectly able to participate in the early steps of TLS preceding actual lesion bypass. The relatively late manifestation of their TLS defects may restrict which alternative mechanisms the cell can employ. Whereas DinB(F13V) is unable to efficiently catalyze dNTP insertion opposite $N^2$-furfuryl-dG, DinB(Y79L) is able to efficiently bypass this lesion but cannot continue extension for more than three nucleotides. However, both of these mutants
require dnaQ+ to exert their dominance with respect to NFZ sensitivity. Our findings suggest that during the early phases of TLS, before bona fide lesion bypass has occurred, commitment to the process is relatively reversible and the cell can employ other DNA repair and damage tolerance mechanisms. In contrast, initiation of lesion bypass may commit the cell to resolving the replication blocking lesion by TLS. The proofreading function of DNA polymerase III normally ensures proper hand-off from a TLS polymerase to DNA polymerase III by reversing improperly short TLS products. Both DinB(F13V) and DinB(Y79L) behave as wild-type DinB prior to cellular commitment to TLS. The abortive TLS products these DinB variants produce, however, are perpetually reversed by the action of ε, locking the cell into a futile cycle of ineffective TLS. Taken together, these observations hint at the critically important interplay between TLS and the replicative DNA polymerase.
Materials and Methods

Bacterial strain and plasmid construction

The strains, plasmids, and primers used in this study are described in table 3. The plasmid borne \textit{dinB} variants were constructed from pYG768 (26) or pDFJ1 (17) using a Quickchange site directed mutagenesis kit (Stratagene) according to the manufacturer’s instructions. Transformations were performed according to standard procedures and plasmids were maintained with ampicillin (100 μg/mL) whenever necessary.

Sensitivity and mutation frequency measurements

Sensitivity to DNA damaging agents was determined essentially as described previously (17, 27). Briefly, \textit{E. coli} strains were grown to exponential phase in LB medium and plated on LB agar containing between 0-10 μg/mL NFZ or 4-NQO. A concentrated stock solution of each DNA damaging agent was first made in N,N-dimethylformamide and diluted appropriately for each experiment.

Oligonucleotide synthesis

The synthesis and characterization of oligonucleotides used as templates in this study has been described previously (17). The nested primers used to define the range of preferential DinB TLS activity are described in table 3.
Protein purification and primer extension assays

Purification of wild-type DinB and DinB(Y79L) was performed as previously described (17, 27). Briefly, each protein was expressed from the plasmid pDFJ1 or its DinB(Y79L) expressing variant in BL21(DE3) pLysS cells with 1 mM IPTG at 30°C. Four hours post induction, cells were harvested by centrifugation and subjected to lysozyme treatment. The resulting lysate was treated with DNAse and RNAse as described (17) and purified on monoS and phenylsepharose columns (GE Healthcare). DinB(Y79L) purified virtually indistinguishably from wild-type DinB, except for exhibiting a broader peak during elution from the phenylsepharose column. Similar specific activity is maintained through the entire peak.

Primer extension assays were performed as described previously (17), using equal concentrations of wild-type DinB and DinB(Y79L). Products were separated on a 16% denaturing polyacrylamide gel and quantified using a typhoon phosphorimager (GE Healthcare). In cases where nucleotide addition resulted in multiple products the sum of all products was used to calculate $V_{\text{max}}$ and $K_m$.

Acknowledgements

We would like to thank Eunha Kim for assistance in producing several mutants used in this study as well as members of the Walker laboratory for thoughtful comments on the manuscript and Steve Bell for use of the typhoon phosphorimager. This work was supported by NIH grants to GCW and JME. GCW is an American Cancer Society Research Professor.
References


Table 1. Kinetic parameters for nested lesion bypass primer extension assays.

<table>
<thead>
<tr>
<th>Primer</th>
<th>( V_{\text{max}, dG} )</th>
<th>( V_{\text{max}, N^2-dG} )</th>
<th>( K_m, dG )</th>
<th>( K_m, N^2-dG )</th>
<th>( V_{\text{max}}/K_m, dG )</th>
<th>( V_{\text{max}}/K_m, N^2-dG )</th>
<th>Fold Stimulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-5</td>
<td>557.8</td>
<td>668.9</td>
<td>83.2</td>
<td>27.9</td>
<td>( 6.7 \times 10^6 )</td>
<td>( 2.4 \times 10^7 )</td>
<td>3.6</td>
</tr>
<tr>
<td>G-4</td>
<td>755.9</td>
<td>2492.7</td>
<td>173.9</td>
<td>129.3</td>
<td>( 4.3 \times 10^6 )</td>
<td>( 1.9 \times 10^7 )</td>
<td>4.4</td>
</tr>
<tr>
<td>G-3</td>
<td>464.2</td>
<td>304.6</td>
<td>9.6</td>
<td>5.1</td>
<td>( 4.8 \times 10^7 )</td>
<td>( 6.0 \times 10^7 )</td>
<td>1.2</td>
</tr>
<tr>
<td>G-2</td>
<td>1081.7</td>
<td>1149.6</td>
<td>109</td>
<td>51.7</td>
<td>( 9.9 \times 10^6 )</td>
<td>( 2.2 \times 10^7 )</td>
<td>2.2</td>
</tr>
<tr>
<td>G-1</td>
<td>894.3</td>
<td>990.5</td>
<td>231.6</td>
<td>16</td>
<td>( 3.8 \times 10^7 )</td>
<td>( 6.2 \times 10^7 )</td>
<td>16.0</td>
</tr>
<tr>
<td>G</td>
<td>619.8</td>
<td>886.3</td>
<td>42.2</td>
<td>2.4</td>
<td>( 1.5 \times 10^7 )</td>
<td>( 3.8 \times 10^8 )</td>
<td>25.7</td>
</tr>
<tr>
<td>G+1</td>
<td>240.4</td>
<td>167.0</td>
<td>67.1</td>
<td>41.7</td>
<td>( 3.6 \times 10^6 )</td>
<td>( 4.0 \times 10^6 )</td>
<td>1.1</td>
</tr>
<tr>
<td>G+2</td>
<td>188.2</td>
<td>150.3</td>
<td>24.9</td>
<td>8.4</td>
<td>( 7.6 \times 10^6 )</td>
<td>( 1.8 \times 10^7 )</td>
<td>2.4</td>
</tr>
<tr>
<td>G+3</td>
<td>446.1</td>
<td>438.2</td>
<td>59.8</td>
<td>19.6</td>
<td>( 7.5 \times 10^7 )</td>
<td>( 2.2 \times 10^7 )</td>
<td>3.0</td>
</tr>
</tbody>
</table>

\( V_{\text{max}, dG} \) units are pmol min\(^{-1}\) mg\(^{-1}\)

\( K_m \) units are \( \mu \)M

\( V_{\text{max}}/K_m \) units are pmol min\(^{-1}\) mg\(^{-1}\) M\(^{-1}\).

Table 2. Kinetic parameters for DinB(Y79L) catalyzed bypass of a site-specific \( N^2 \)-furfuryl-dG lesion (F-dG). The DinB(Y79L) variant displays an approximately twofold lower catalytic proficiency on an undamaged template, but a striking 20-fold reduction in activity on the damaged template relative to wild-type DinB. \( K_m \) units are \( \mu \)M; \( V_{\text{max}} \) units are pmol min\(^{-1}\) mg\(^{-1}\); \( V_{\text{max}}/K_m \) units are pmol min\(^{-1}\) mg\(^{-1}\) M\(^{-1}\).

<table>
<thead>
<tr>
<th></th>
<th>( K_m )</th>
<th>( V_{\text{max}} )</th>
<th>( V_{\text{max}}/K_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-control</td>
<td>330</td>
<td>1072</td>
<td>( 3.2 \times 10^6 )</td>
</tr>
<tr>
<td>F-dG</td>
<td>1551</td>
<td>1416</td>
<td>( 0.9 \times 10^6 )</td>
</tr>
<tr>
<td>Strains</td>
<td>Genotype</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>BL21 (DE3)</td>
<td><em>E. coli</em> B. Standard strain used for protein overproduction</td>
<td>GCW lab stock</td>
<td></td>
</tr>
<tr>
<td>pLysS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td><em>thr-1 araC-14 leuB-6(Am) Δ(gpt-proA)62 lacY1 tsx-33 qsr'-0 glnV44(AS) galK2(Oc) LAM Rac-0 hisG4(Oc) rfbC1 mgl-51 rpoS396(Am) rpsL31 kdgK51 xylA5 mtl-1 argE3(Oc) thi-1</em></td>
<td>GCW lab stock</td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but Δ<em>adinB</em>::Cm</td>
<td>Jarosz et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>Δ<em>adinB</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but Δ<em>adnaQ</em>::Tet</td>
<td>Sutton et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>Δ<em>adnaQ</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>As above but Δ<em>adinB</em>::Cm</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>Δ<em>adnaQ</em> Δ<em>adinB</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDFJ1</td>
<td>DinB expression plasmid under T7 promoter in pET11T backbone; Ampicillin resistance</td>
<td>Jarosz et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>pDFJ1(Y79L)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDFJ1(Y79F)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB</td>
<td>pYG768. DinB under its native promoter in a low copy number pWSK29 plasmid</td>
<td>Kim et al. (1997)</td>
<td></td>
</tr>
<tr>
<td>pDinB(F12A)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(F12V)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(F12L)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(F76A)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(F76V)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(F76L)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(Y79A)</td>
<td>As pDFJ1 with indicated mutation introduced using quickchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>pDinB(Y79V)</td>
<td>As pDFJ1 with indicated mutation introduced using quikchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(Y79L)</td>
<td>As pDFJ1 with indicated mutation introduced using quikchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pDinB(Y79F)</td>
<td>As pDFJ1 with indicated mutation introduced using quikchange site-directed mutagenesis</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pWSK29</td>
<td>Empty vector for pDinB and derivatives</td>
<td>Kim et al. (1997)</td>
<td></td>
</tr>
</tbody>
</table>

**Primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-5</td>
<td>GCATATGATAGTACAGCTGCAGCCGGA</td>
</tr>
<tr>
<td>G-4</td>
<td>GCATATGATAGTACAGCTGCAGCCGGAC</td>
</tr>
<tr>
<td>G-3</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACG</td>
</tr>
<tr>
<td>G-2</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACGC</td>
</tr>
<tr>
<td>G-1</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACGCC</td>
</tr>
<tr>
<td>G</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACGCC</td>
</tr>
<tr>
<td>G+1</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACGCC</td>
</tr>
<tr>
<td>G+2</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACGCC</td>
</tr>
<tr>
<td>G+3</td>
<td>GCATATGATAGTACAGCTGCAGCCGGACGCC</td>
</tr>
</tbody>
</table>

*Table 3.* Strains, plasmids, and primers used in this study.
Figure 1. Cluster of aromatic hydrophobic residues surrounding the 'steric gate' residue of DinB, F13. A) F12, F13, F76, and Y79 form a group of hydrophobic residues surrounding a putative binding pocket for a template furfuryl modification on a template G. B) These residues display striking conservation among DinB orthologs through evolution. F76 is less conserved than other residues, whereas Y79 is invariant among DinB orthologs.
Figure 2. Aromatic hydrophobic residues surrounding F13 are important for DinB function. Percent survival of a ΔdinB strain bearing plasmids expressing DinB and variants when challenged with 2.0 μg/mL NFZ. Mutation of Phe 12 inhibits complementation, but mutation of Phe 76 has no statistically significant effect on complementation.
Figure 3. Tyr 79 is critical for DinB function. A-B) Mutation of Tyr 79 to Ala, Val, or Leu generates a DinB variant that, when expressed from a low copy number plasmid under its native promoter, is unable to complement the NFZ or 4-NQO sensitivity of a ΔdinB strain. Plots show percent survival vs. NFZ or 4-NQO concentration and error bars represent the standard deviation as determined from three independent transformants. C) Mutation of Tyr 79 to Phe generates a DinB variant (closed squares) that behaves as WT DinB (closed circles) with respect to its ability to complement NFZ sensitivity when expressed from a plasmid under its native promoter. D) Expression of DinB(Y79F) confers striking resistance during treatment with 40 μM NFZ, conditions under which killing by that agent does not depend on dinB."
Figure 4. Mutation of Tyr 79 alters the ability of DinB to perform TLS. Primer extension reactions containing 1, 5, and 50 nM WT DinB, DinB(Y79L), or DinB(Y79F) and 5 nM $^32$P-labeled $N^2$-furfuryl-dG containing primer/template separated on a 16% denaturing polyacrylamide gel. WT DinB and DinB(Y79F) fully extend the primer to the end of the template, whereas DinB(Y79L) stalls three nucleotide additions after the lesion.
Figure 5. Proofreading function is required for dominance of dinB(Y79L) and dinB(F13V). A) Percent survival vs. NFZ dose of AB1157 bearing an empty vector control (pWSK29, open circles); pDinB (closed circles); pDinB(Y79L) (closed triangles); pDinB(F13V) (closed squares). Error bars represent one standard deviation determined from three independent transformants. B) Percent survival vs. NFZ dose of AB1157 ΔdnaQ bearing the plasmids described in A. Error bars represent one standard deviation as determined from three independent experiments.
Denotes the site of the lesion

**Figure 6.** Schematic of DNA substrates used for nested primer extension assays to determine the range of DinB’s preferential activity surrounding a site-specific $N^2$-furfuryl-dG lesion.
Figure 7. Fold preferential activity of DinB surrounding a site-specific $N^2$-furfuryl-dG lesion. The ratio of catalytic proficiencies is plotted for the indicated insertion on a damaged vs. undamaged template. DinB displays at least comparable catalytic proficiency on the damaged template over the entire range tested, but shows striking preference for insertion opposite (G-1 primer) and immediate extension from (G primer) the $N^2$-furfuryl-dG lesion.
Figure 8. Proposed model for order of dominance of DinB variants employed in this study. DinB(Y79L) and DinB(F13V) behave normally during the early phases of TLS, manifesting their defects only at or after the point at which the cell has committed to lesion resolution by TLS. The abortive products they produce are then subject to reversal by the proofreading function of the ε subunit of Pol III. A catalytically deficient DinB variant manifests its phenotype early in TLS, thereby allowing other mechanisms for lesion resolution. Figure adapted from reference 23.
Chapter 4

Conservation of Substrate Specificity Among DinB and its Archaeal and Mammalian Orthologs

DF Jarosz, JC Delaney, AM DeLucia, CM Joyce, EC Friedberg, JM Essigmann, and GC Walker
in preparation.
Abstract

DNA damage tolerance mechanisms, including translesion synthesis (TLS), contend with chemically diverse substrates that have escaped DNA repair. We recently showed that the most evolutionarily conserved subfamily of TLS DNA polymerases, DinB and its orthologs, act preferentially on $N^2$-furfuryl-dG damaged templates. Here we report that this property arises from specialized function rather than simply from an accommodating active site. In addition, we demonstrate that the substrate specificities of *Escherichia coli* DinB, its archaeal ortholog *Sulfolobus acidocaldarius* Dbh, and its mammalian ortholog *M. musculus* DNA polymerase $\kappa$ are remarkably similar with respect to minor modifications at the $N^2$ position of a template dG. Efficient bypass was observed for each enzyme when bypassing both $N^2$-furfuryl-dG and the reduced $N^2$-tetrahydrofurfuryl-dG. Differences emerge with respect to bypass of bulky $N^2$-dG modifications with DNA polymerase $\kappa$ showing the greatest promiscuity. This result is unexpected, given the higher number of alternative DNA polymerases in eukaryotes, but offers insight into the similarities and differences among DinB and its orthologs from evolutionarily diverged organisms.
Introduction

Organisms employ numerous coordinated strategies in their responses to genotoxic stress that can be broadly placed into two classes: DNA repair and DNA damage tolerance (1). A novel damage tolerance mechanism, known as translesion synthesis (TLS) involves the DNA polymerase-catalyzed bypass of template base modifications (1-3). This process can occur with a range of fidelities (1, 3) and often involves the function of Y-family DNA polymerases (2). These DNA polymerases share general structural homology with replicative DNA polymerases (4-6), but they replicate undamaged DNA templates with substantially lower fidelity and also lack exonuclease activity (1, 3).

Despite their moderate fidelities, these enzymes play critically important physiological roles in both simple prokaryotic and complex multicellular eukaryotic organisms. Escherichia coli strains lacking either Y-family DNA polymerase possessed by that organism show a dramatic competitive disadvantage with wild-type (7) and mutation of RAD30, which encodes the Y-family DNA polymerase pol η, has been associated with the cancer-prone disease phenotype Xeroderma pigmentosum (8). Indeed, eukaryotes may employ TLS more commonly than prokaryotes to contend with replication-blocking DNA modification (9).

DinB is the only Y-family DNA polymerase that is present in all three domains of life (2). However, relatively little is known about its substrate specificity compared to other alternative DNA polymerases (1). We recently demonstrated that ΔdinB E. coli strains show striking sensitivity to the DNA damaging agent nitrofurazone (NFZ) and that DinB, as well as its archaeal and mammalian orthologs, can efficiently bypass a site-specific N²-furfuryl-dG adduct, a structural mimic of the major N²-dG lesion produced by NFZ (10). DNA polymerase κ (pol κ), the eukaryotic DinB ortholog, has also been shown to catalyze bypass of certain N²-dG adducts.
in several organisms with wide-ranging efficiencies (1, 11). The substrate specificities of archaeal DinB orthologs have been even less well characterized.

We wondered what mechanisms govern DinB’s specialized lesion bypass function and to what extent the substrate specificity of DinB overlaps with its archaeal and DinB orthologs. In this paper we present evidence that DinB’s remarkable catalytic preference for action on an $N^2$-furfuryl-dG bearing substrate does not arise merely from a generally accommodating active site. Moreover, we find that *E. coli* DinB, an isoform of *Mus musculus* DNA pol $\kappa$, and *Sulfolobus acidocaldarius* Dbh exhibit strikingly similar substrate specificities over a wide-range of template $N^2$-dG modifications. However, DNA pol $\kappa$ appears to be comparatively promiscuous with respect to its ability to catalyze TLS over bulky lesions. These findings extend our understanding of the driving force leading to the maintenance of DinB and its orthologs throughout evolution.
Results

Preferential DinB TLS function on $N^2$-dG adducts does not arise exclusively from its open active site

We initially wondered whether our finding that DinB can efficiently bypass $N^2$-furfuryl-dG lesions reflected the open active sites shared by Y-family DNA polymerases (4-6). To answer this question, we examined the behavior of DinB when bypassing a variety of previously characterized site-specific methylation lesions (12): 1-methyldeoxyadenosine (m1a), 3-methyldeoxycytosine (m3C), 3-ethyldeoxycytosine (e3C), and a synthetic tetrahydrofuran (THF) abasic site mimic (Figure 1). These lesions are structurally diverse, distinct from furfuryl modification of the $N^2$ position of dG, and are not very sterically bulky. Strikingly, even under conditions that greatly favor of DinB-mediated lesion bypass, we found that it was virtually unable to catalyze TLS over m1A, m3C, e3C, or THF containing substrates (Figure 2). This observation is particularly surprising given the relatively modest nature of these base modifications, particularly the THF abasic site analog, which at least in principle should not present a strong block to replication. DinB has the capability to bypass a THF abasic site mimic in the presence of the $\beta$ processivity clamp and $\gamma$ clamp loader subunits of DNA polymerase III, but our observations are consistent with previous reports that it can only do so extremely inefficiently in the absence of these factors (13). In vivo, however, the vast majority of TLS past abasic sites is mediated by the $umuDC$ gene products (13). Taken together, these data indicate that DinB function in TLS past $N^2$-dG adducts is specialized and that it is relatively intolerant of other even modestly modified DNA substrates.
**Generation of a collection of $N^2$-dG adducted substrates**

We have previously investigated mechanisms governing the enhanced activity of DinB when bypassing an $N^2$-furfuryl-dG lesion and found that it depends critically its steric gate residue (10), which forms the back of a putative binding pocket for the furfuryl moiety of the lesion (Figure 3). This putative lesion-binding pocket is roughly as long (ca. 7.5 Å) as it is wide (ca. 7.4 Å), is bounded below by Ser 44 (3.3 Å below $N^2$ of a template G), and is relatively open above. To gain further insight into the structural diversity of lesions that this pocket might accommodate, we synthesized a collection site-specific $N^2$-dG modified oligonucleotide substrates (Figure 1B) using a post-synthetic derivitization approach (10, 14). These modifications range from a simple methylation to the elaborate polycyclic benzo[a]pyrene lesion and constitute a wide-range molecular toolkit with which to probe the promiscuity of DinB and its orthologs with respect to substrate specificity.

**DinB orthologs from all domains of life display similar substrate specificities**

We characterized bypass of this collection of $N^2$-dG adducts by DinB orthologs from all three domains of life: *E. coli* DinB (15), *S. acidocaldarius* Dbh (16), and *M. musculus* pol $\kappa$ (17-20). DinB shares 25% identity with Dbh and 32% identity with pol $\kappa$, whereas Dbh shares only 24% identity with pol $\kappa$. These percentages are also reflected among the residues that comprise the putative lesion binding pocket of DinB.

These enzymes display striking similarity with respect to their abilities to bypass the collection of $N^2$-dG adducts that we examined. DinB, Dbh, and pol $\kappa$ each bypass $N^2$-methyl-dG and $N^2$-ethyl-dG with comparable efficiency to an undamaged template of the same sequence context (Figure 4A-C). Intriguingly, Dbh always showed a dramatically lower processivity than
either DinB or pol κ in our primer extension assays, adding only two nucleotides after the site of the lesion (Figure 4B). In comparison, even the distributive activity of DinB and pol κ appears relatively processive. Each enzyme also appears to bypass \(N^2\)-ethyl-dG with a slightly reduced efficiency relative to both \(N^2\)-methyl-dG and dG itself. Taken together, these data suggest that although these small \(N^2\)-dG modifications can be accommodated within the lesion binding pockets of all three DinB orthologs, they are not optimal for preferential lesion bypass by those enzymes.

**The double bonding character of \(N^2\)-furfuryl-dG is dispensable for efficient lesion bypass**

Our inability to observe preferential lesion bypass by DinB and its orthologs over minor \(N^2\)-dG modifications led us to investigate what aspects of \(N^2\)-furfuryl-dG contribute to its apparently unique properties. We therefore examined the ability of DinB, Dbh, and pol κ to bypass a sterically similar \(N^2\)-tetrahydrofurfuryl-dG lesion (Figure 1A) with different bonding character. Whereas both DinB and pol κ preferentially bypass this lesion, Dbh acts with lower fidelity on an \(N^2\)-tetrahydrofurfuryl-dG substrate than on an undamaged control (Figure 5A-C). These observations indicate that Dbh is likely more sensitive than either DinB or pol κ to the double bonding character of an \(N^2\)-dG adduct. Intriguingly, pol κ appears to catalyze slightly more proficient TLS over \(N^2\)-tetrahydrofurfuryl-dG than over \(N^2\)-furfuryl-dG (Figure 5C), perhaps reflecting a preference for the subtly altered geometry of this lesion. Taken together, these data suggest that both DinB and pol κ may employ a mechanism of substrate selection based largely on shape. In contrast, Dbh may be particularly sensitive to additional aspects of the damaged template. However, these effects are relatively subtle. Our results strongly suggest
that, at least with respect to $N^2$-methyl-dG, $N^2$-ethyl-dG, $N^2$-tetrahydrofurfuryl-dG, and $N^2$-furfuryl-dG, DinB, Dbh, and pol $\kappa$ display strikingly similar substrate specificities.

**Divergent behavior of DinB orthologs bypassing a bulky $N^2$-dG Adduct**

We wondered whether DinB, Dbh, and pol $\kappa$ would also behave similarly on a much larger substrate and therefore examined their abilities to bypass a site-specific $N^2$-beno[a]pyrene-dG lesion. DinB has been shown to weakly bypass this lesion (10, 21) and pol $\kappa$ from various organisms can catalyze bypass of various B[a]P-dG adducts with varying efficiencies (9). We found that DinB and its orthologs display strikingly different abilities to bypass this bulky $N^2$-dG adduct (Figure 6A-C). When present in high concentrations DinB can insert a nucleotide opposite this lesion with modest efficiency, but is strikingly deficient at extension from the damaged base (Figure 5A). In comparison, Dbh appears entirely unable to detectably insert a dNTP opposite to this lesion even in very high concentrations (Figure 5B). In striking contrast, pol $\kappa$ is able to bypass this lesion with modest efficiency (Figure 5B). These observations suggest that, at least with respect to bulky $N^2$-dG lesions, pol $\kappa$ is more promiscuous than its bacterial and archaeal orthologs.
Discussion

The chemical diversity of even naturally occurring DNA damage is remarkable (1) but presents a potential problem for polymerases that catalyze translesion synthesis given their limited numbers in all organisms. Here we report that the specialized ability of DinB to promote preferential TLS over a certain $N^2$-dG adducts arises not merely by virtue of its open active site, but rather due to a specialized function. A pocket in the active site of DinB is properly positioned and of the appropriate size to accommodate a modestly sized $N^2$-modification on a template dG. The TLS activity of DinB appears to be relatively restricted, as it is unable to efficiently bypass a variety of other lesions minorly perturbative we examined.

With respect to bypass of $N^2$-methyl-dG and $N^2$-ethyl-dG, DinB orthologs from archea (S. acidocaldarius Dbh) and mammals (M. musculus pol $\kappa$) possess strikingly similar substrate specificities to DinB given their evolutionary diversity. Although each enzyme bypasses both these minor modifications and an undamaged control with comparable efficiencies, none replicate them with preferential activity. This observation indicates that even though the methyl and ethyl modifications can be accommodated by the putative binding pocket, lesion recognition is not coupled to enhanced catalysis for these substrates.

In contrast, both DinB and pol $\kappa$ bypass $N^2$-furfuryl-dG and the sterically similar $N^2$-tetrahydrofurfuryl-dG with comparable preferential efficiencies, indicating that for these enzymes size matters more than bonding character for substrate selection. Dbh does show a measurably lower activity on $N^2$-THF-dG than on $N^2$-furfuryl-dG, indicating that the minor differences between these substrates including both bonding character and planarity of the lesion may be more important for this enzyme. Intriguingly, pol $\kappa$ appears to bypass $N^2$-THF-dG slightly better than $N^2$-furfuryl-dG. This may again reflect a subtle difference in the lesion
binding pocket between the two enzymes. Many of the residues that line the putative lesion
binding pocket are conserved among DinB, Dbh, and pol κ, but Dbh notably does not possess a
Phe 76 equivalent and instead bears a lysine residue at this position.

The striking similarity in substrate specificities of DinB orthologs does not extend to
bypass of bulky N²-dG lesions, however. Whereas pol κ can catalyze bypass of this lesion under
our assay conditions, DinB can only weakly insert a nucleotide opposite to such a damaged base
and appears unable to extend from it. Dbh is entirely unable to act on this substrate. These
observations suggest, somewhat unexpectedly, that pol κ may be the most promiscuous with
respect to substrate specificity than either DinB or Dbh. This is a somewhat counterintuitive as
the propensity of eukaryotes have far more alternative DNA polymerases than bacteria or
archaeal species might have translated into increased specialization. Instead, both our results and
those of others (11) suggest that eukaryotic pol κ is strikingly accommodating with respect to
bulky modification at the N²-position of a template dG.

Taken together, all of these data indicate that there is a unique feature of lesions similar
to N²-furfuryl-dG that makes them ideal substrates for DinB and its orthologs. Human pol κ has
been shown to preferentially catalyze the insertion of a nucleotide opposite to an N²-benzyl-dG
modification but not N²-napthyl-dG (11), indicating that the size of a five- or six-membered ring
may be preferred for this enzyme. DinB and its orthologs are thought to function in TLS by
bypassing endogenous DNA lesions relatively accurately (1, 22). Although furfuryl
modifications of dA and dG can be formed from reactive species generated by oxidation of
deoxyribose (23), it is unlikely that they are the sole reason for the striking conservation of DinB
orthologs through evolution. Rather, it seems plausible that DinB, Dbh, and pol κ each function
in bypass of endogenous substrates that share structural and steric similarity with these adducts.

The identity and source of such an adduct merits considerable further investigation.
Materials and Methods

Lesion synthesis and oligonucleotide construction

Oligonucleotide synthesis and construction was carried out as described previously in chapter 2 using a post-synthetic derivitization approach followed by scaffolded ligations (10, 14).

Primer extension assays

Primer extension assays were performed using 5'-32P-labeled primers and modified templates as previously described (10, 24). Reactions were initiated with the addition of 250 μM dNTPs, quenched after 20 min with 95% formamide, 25 mM EDTA, 0.5% bromophenol blue, and 0.5% xylene cyanol, and separated on a 16% denaturing polyacrylamide gel, which was quantified using a Typhoon phosphorimager (GE Healthcare).

Acknowledgements

We would like to thank members of the Walker and Essigmann laboratories, as well as Janice Pata (Wadsworth Center), for thoughtful discussions. DFJ performed the protein purification, lesion bypass assays, and wrote the manuscript. JCD synthesized the oligonucleotide substrates used in this study. AMD and CMJ provided purified Dbh and ECF provided mammalian pol κ.

This work was supported by NIH grants to GCW and JME. GCW is an American Cancer Society Research Professor.
References


Figure 1. Structures of DNA lesions employed in this study. A) Structures of a synthetic abasic analog, 1-methyldeoxyadenosines, and 3-alkyldeoxycytosines used to probe the structural breadth of DinB TLS function. B) Structures of the collection of $N^2$-dG adducts used in this study.
Figure 2. Primer Extension Assays Reveal DinB’s Preference for $N^2$-dG Adducted Substrates. DinB is able to efficiently bypass a site-specific $N^2$-furfuryl-dG lesion but unable to efficiently bypass THF, m1A, m1C, or e3C. Each of the three lanes represent independently constructed assays, conditions for which are described in materials and methods.
Figure 3. Putative binding pocket for $N^2\text{-dG}$ adducts in the active site of DinB. The template is colored in yellow, the incoming deoxynucleotide is colored in blue, the primer terminus is colored in green, and the ‘steric gate’ residue is colored in red.
Figure 4. DinB-, Dbh-, and pol κ-catalyzed bypass of site-specific $N^2$-methyl and $N^2$-ethyl lesions. Reactions contain increasing amounts of enzyme (1 nM, 5 nM, 50 nM). Processivity differences are observed among DinB and its orthologs, but each enzyme replicates these modified templates with an equal or lesser efficiency than an undamaged control template of the same sequence context.
Figure 5. DinB-, Dbh-, and pol κ-catalyzed bypass of site-specific $N^2$-tetrahydrofuranyl-dG and $N^2$-furanyl-dG lesions. Reactions contain increasing amounts of enzyme (1 nM, 5 nM, 50 nM). Whereas both DinB and pol κ act preferentially on each lesion, Dbh displays reduced activity on $N^2$-tetrahydrofuranyl-dG relative to dG alone. Reactions were performed as indicated in materials and methods.
Figure 6. Bypass of a bulky $N^2$-B[a]P-dG lesion by DinB and its orthologs Dbh and Pol κ.
Chapter 5

UmuD and RecA Directly Modulate the Mutagenic Potential of the Y-family DNA Polymerase DinB

This chapter has been submitted for publication as VG Godoy, DF Jarosz, SM Simon, A Abyzov, VA Ilyin, GC Walker “UmuD and RecA Directly Modulate the Mutagenic Potential of the Y-family DNA Polymerase DinB.” Submitted.
Abstract

DinB is the only translesion Y-family DNA polymerase conserved among bacteria, archaea, and eukaryotes. DinB and its orthologs possess a specialized lesion-bypass function but also display potentially deleterious -1 frameshift mutagenic phenotypes when overproduced. We show that the DNA damage inducible proteins UmuD2 and RecA act in concert to modulate its potentially mutagenic activity. Structural modeling suggests that the relatively open active site of DinB is enclosed by interaction with these proteins thereby preventing the template bulging responsible for -1 frameshift mutagenesis. Intriguingly, residues that define the UmuD2 interacting surface on DinB statistically co-vary throughout evolution, suggesting a driving force for the maintenance of a regulatory protein-protein interaction at this site. Together, these observations indicate that proteins like RecA and UmuD2 may be responsible for managing the mutagenic potential of DinB orthologs throughout evolution.
Introduction

Decades after their discovery, the dinB (1) and umuDC (2, 3) genes of Escherichia coli were shown to encode specialized Y-family DNA polymerases, DNA pol IV (4) and pol V (5, 6) respectively, that catalyze the insertion of deoxyribonucleoside triphosphates (dNTPs) opposite potentially lethal replication blocking lesions in a process termed translesion synthesis (TLS) (7, 8). TLS can proceed with a range of fidelities (7), but in all cases Y-family polymerases replicate undamaged DNA with a reduced fidelity relative to the enzymes that replicate the majority of the genome (9). Therefore, Y-family polymerases must be excluded from improper access to replication intermediates to maintain genomic integrity (10).

DinB is the only Y-family DNA polymerases conserved among all domains of life (8) and under conditions of DNA damage it is the most abundant DNA polymerase in E. coli (11). We recently showed that an important function of DinB and its orthologs is to carry out highly proficient and accurate TLS past a particular class of N²-deoxyguanosine adducts (12). On the surface, these observations seemed incompatible with DinB’s role in the elevated frequency of -1 frameshift mutations observed during λ untargeted mutagenesis (13) and adaptive mutagenesis (14, 15) or with the -1 frameshift mutator effect caused by DinB overproduction (16).

Both dinB and umuDC are regulated transcriptionally by the SOS regulatory network (17), their expression being induced when LexA undergoes facilitated autocleavage upon interaction with the RecA::ssDNA nucleoprotein filament formed after DNA damage. The subsequent control of UmuC function is remarkably complex and involves the gene products of the co-transcribed umuD gene. UmuD₂ undergoes an SOS-mediated posttranslational modification when it too is subject to facilitated autodigestion upon interaction with RecA::ssDNA nucleoprotein filaments to yield UmuD₂' (18, 19). The conversion of UmuD₂ to
UmuD'2 activates UmuC polymerase function (5, 6) and RecA::ssDNA filaments are required in trans for UmuD'2C-catalyzed translesion synthesis (20). In addition, RecF, RecO, and RecR cooperate to alleviate the inhibition of UmuD'2C mediated TLS that is brought about by DNA pol III (21). Finally, UmuD'2C must interact with the β processivity clamp of DNA pol III to function in vivo (22, 23).

Despite the remarkable evolutionary conservation of DinB, the details of its biochemical regulation are comparatively unknown (24). Previous studies of DinB and its orthologs, as well as of other Y-family DNA polymerases, have focused on the pivotal role of processivity clamps in regulation of TLS and mutagenesis (22, 25-30). However, given DinB’s documented potential to cause deleterious -1 frameshift mutations (11, 31), we thought it possible that DinB might be subject to control beyond simply transcriptional induction and interaction with the processivity clamp. We therefore searched for additional regulatory factors that might manage this potentially problematic function. Remarkably, we found that UmuD, UmuD', and RecA, previously known only to regulate UmuC function, regulate both the activity and mutagenic properties of DinB via protein-protein interactions that enclose its active site. Our findings suggest that mechanistic features of this regulation may be maintained in eukaryotes consistent with a common pattern of regulation for these DNA polymerases through evolution.
Results

DinB interacts with numerous cellular factors

To identify proteins that might regulate DinB function, we covalently coupled purified recombinant DinB to an affinity resin (Figure 1). Interacting proteins from lysates of constitutively SOS-induced *E. coli* were eluted and separated by SDS-PAGE. Those that bound in a DinB dependent fashion included the chaperones GroEL and DnaK and also the ribosomal protein L3 as well as lesser amounts of other ribosomal proteins. The presence of the chaperones was anticipated since GroEL has been shown to regulate the function of both DinB (32) and UmuD'2C (33) and DnaK regulates UmuC levels *in vivo* (34). The ribosomal proteins L3 and others seem less likely to be *bona fide* regulators of DinB function although this remains a formal possibility.

However, we also identified UmuD, UmuD', and RecA as DinB interacting proteins. This was a complete surprise because, despite their intensively studied, highly nuanced roles in regulating UmuC function (17, 31, 35), none of these factors had previously been implicated in regulating DinB function aside from the indirect role of RecA in mediating DinB induction via the SOS regulatory network. Intriguingly, the levels of UmuD *in vivo* (180 molecules in non-SOS induced cells; 2400 molecules in SOS induced cells) parallel those of DinB (250 molecules in non-SOS induced cells; 2500 molecules in SOS induced cells) and greatly exceed what is required to interact with UmuC (17 molecules in non-SOS induced cells; 200 molecules in SOS-induced cells) (16, 36). We therefore investigated the ability of these proteins to affect DinB function *in vivo* and *in vitro*. 
DinB forms a stable interaction with UmuD$_2$ and RecA

To ascertain whether the interactions we observed between DinB, RecA, UmuD, and UmuD$'$ were direct, we first performed a farwestern blot in which we probed membranes containing UmuD, UmuD$'$, and RecA with DinB. Each of the interactions appears to be direct in nature (Figure 2A). To analyze the stoichiometry of the DinB-UmuD interaction, we crosslinked DinB and UmuD with formaldehyde and analyzed the products by immunoblot using an antibody against DinB. The crosslinked species corresponds to the molecular weight of a DinB-UmuD$_2$ complex and the reaction appears to be inhibited by high concentrations of NaCl (Figure 2B), suggesting that the interface may partly involve ionic or polar interactions. The propensity of RecA to multimerize (37) made it difficult to establish the stoichiometry of the DinB-RecA interaction.

To test whether DinB, UmuD$_2$ and/or UmuD$'_2$, and RecA form a stable ternary complex in solution, a DinB variant with a hexahistidine affinity tag at its C-terminus was incubated with RecA – both alone and in combination with UmuD$_2$, UmuD$'_2$, and the heterodimeric species UmuDD$'$ (38). Complexes that formed with DinB were isolated using Ni$^{2+}$ affinity resin. Using physiologically relevant concentrations (2.5 $\mu$M) of each protein, we determined that DinB interacts stably with UmuD$_2$ and UmuD$'_2$ (Figure 2C). Curiously, we did not observe binding of the heterodimeric UmuDD$'$ to DinB. The small amount of UmuD that appears in the gel may arise either from a slight excess of UmuD$_2$ or from UmuDD$'$ rearranging to generate a small amount of UmuD$_2$. We also observed the formation of a stable stoichiometric complex between DinB and RecA (Figure 2C); fluorescence anisotropy does not reveal such an association (39). Moreover, it appears that RecA stimulates DinB association with UmuD, but not with UmuD$'$ or UmuDD$'$, by approximately two-fold. Taken together, these data indicate that DinB, RecA, and
UmuD\textsubscript{2} (and to a lesser extent UmuD\textsubscript{2}'\textsubscript{2}), can form ternary complexes under physiological conditions.

**UmuD suppresses DinB-dependent mutagenic phenomena in vivo**

Since prior genetic studies had not implicated either the *umuD* or *recA* gene products in *dinB\textsuperscript{+}-dependent phenomena, we investigated whether these interactions were important *in vivo*.

Because the cellular levels of UmuD mirror those of DinB and are much higher than those of UmuC, we wondered whether the -1 frameshift mutator effect associated with overexpression of DinB might be a consequence of the number of molecules of DinB exceeding those of UmuD in the cell. Strikingly, that appears to be the case as we found that co-overproduction of UmuD (and to a lesser extent UmuD\textsuperscript{2}') eliminates most of the -1 frameshift mutagenesis caused by DinB overproduction (Figure 3A). Co-overproduction of a noncleavable UmuD variant, UmuD(S60A), completely eliminates DinB dependent frameshift mutagenesis (Figure 3A), indicating that full-length UmuD is sufficient for maximal inhibition.

We then investigated the possibility that the *umuD* gene products might similarly modulate the phenomenon of adaptive mutagenesis in the widely studied *E. coli* strain FC40 (40), which is dependent on DinB-promoted -1 frameshifts (14, 15). Under the conditions required to observe such mutagenesis, DinB levels are elevated by approximately 2-4-fold (11, 41). Remarkably, overproduction of UmuD or UmuD\textsuperscript{2}' strikingly reduced adaptive mutagenesis (Figure 3B). The 5-fold reduction in the frequency of adaptive mutagenesis caused by UmuD overexpression is equivalent to the decrease caused by *dinB* inactivation (14, 15). These observations suggest that the *umuD* gene products are able to modulate the -1 frameshift activity
of DinB both in exponential phase and also under the conditions of an adaptive mutagenesis experiment.

Initially, we were surprised that deletion of *umuD* has no effect on *dinB*<sup>+</sup>-dependent NFZ resistance but this point is discussed further below. Although *ΔrecA* strains do exhibit sensitivity to this agent (data not shown), the multitudinous physiological roles of RecA (17, 37) complicate the interpretation of this result.

To investigate whether DinB reciprocally affects UmuD<sub>2</sub> function, we examined the effect of DinB overproduction on UV-induced mutagenesis, a phenomenon that is critically dependent on *umuD*<sup>+</sup> as well as on *umuC*<sup>+</sup> (2, 3). Expression of *dinB*<sup>+</sup> from a low copy number plasmid suppressed UV-induced mutagenesis by a factor of 8.9 (± 2.2). This effect, which is likely related to DinB dependent inhibition of RecA-mediated UmuD<sub>2</sub> autocleavage (Figure 9E), is also consistent with DinB and UmuD<sub>2</sub> interacting *in vivo*.

**Identification of the molecular interface between DinB and UmuD<sub>2</sub>**

Our observation that UmuD suppresses the -1 frameshift activity of DinB *in vivo* was especially intriguing in light of structural studies of archaeal DinB homologs, which have been shown to possess remarkably open active sites (42, 43). We therefore analyzed the interaction between DinB and UmuD<sub>2</sub> using cellulose filter peptide arrays (44). The membranes were probed with either DinB or UmuD<sub>2</sub>, and interacting peptides were identified and mapped onto structural models of DinB (12) or UmuD<sub>2</sub> (45, 46). Interestingly, the UmuD<sub>2</sub> interacting peptides on DinB localize to a single face of the protein (Figure 4A). Further, the presence of an extended interacting surface on DinB suggests that its interaction with UmuD<sub>2</sub> is qualitatively different from its interaction with the β processivity clamp, which depends on a highly conserved
peptide motif (28, 47). Most intriguingly, the interaction interface suggests that UmuD2 may suppress mutagenesis by helping to enclose the strikingly open active site of DinB, thereby preventing the DNA template bulging necessary for -1 frameshift mutagenesis (42, 48).

The DinB interacting interface forms a somewhat less contiguous surface when mapped onto a specific UmuD2 model we had proposed (45). However, the DinB binding interface forms a contiguous surface when mapped onto one of four isoenergetic models of UmuD2 (46), in which its N-terminus is raised to reveal an interacting surface across the side of the protein (Figure 4B). These observations may hint at a biological function for alternative UmuD2 conformers.

In an effort to design a DinB variant that is unable to interact with UmuD2, we identified a strongly interacting group of peptides from the DinB peptide array and examined conservation of this region in numerous umuD-containing organisms. Three residues, P166, F172, and L176 were strikingly conserved and we determined the effect of changing each residue to an alanine (Figure 4A). Although the mutant proteins DinB(P166A) and DinB(L176A) were insoluble (data not shown), we were able to express and purify DinB(F172A) in soluble form (Figure 6). Moreover, we found that the dinB allele encoding DinB(F172A) complements the NFZ sensitivity of a ΔdinB strain (Figure 4D), indicating that this mutant is proficient for TLS in vivo. A reciprocal approach was used to generate a UmuD2 mutant that might be impaired with respect to its ability to interact with DinB. The variant UmuD(D91A) is soluble and proficient for facilitated autoproteolysis (Figure 7).

We determined that the dissociation constant between DinB and UmuD2 is 0.62 μM using fluorescence spectroscopy (Figures 4C, 12). Provocatively, the levels of UmuD2 rise from ca.0.35 μM under non-SOS-induced conditions to ca. 4.5 μM under conditions of SOS-induction
indicating that DinB and UmuD$_2$ are capable of interaction within the range of physiologically relevant concentrations. Furthermore, the DinB(F172A) and UmuD(D91A)$_2$ proteins were each greatly impaired with respect to their ability to bind their partners (Figure 4C), indicating that the interfaces we identified by peptide array mapping are functionally relevant.

**Mutation of the interface between DinB and UmuD$_2$ impairs function in vivo**

To determine whether the physical interaction between DinB and UmuD$_2$ we observed and analyzed in vitro is important for modulation of DinB dependent frameshift mutagenesis in vivo, we examined whether the -1 frameshifts produced by DinB(F172A) could be inhibited by UmuD. Overproduction of DinB(F172A) results in an increase in -1 frameshift mutagenesis by approximately 6-fold (Figure 7A). However, co-overproduction of UmuD or UmuD' does not substantially reduce the -1 frameshift mutation frequency (Figure 7A). These data suggest that a direct interaction of UmuD$_2$ or UmuD'$_2$ at the interface we have identified on DinB is important for modulation of -1 frameshift mutagenesis in vivo.

Reciprocally we examined whether UmuD(D91A) could suppress the -1 frameshift mutagenesis promoted by overproduction of wild-type DinB. Overproduction of UmuD(D91A) only modestly suppresses -1 frameshift mutagenesis (ca. 1.5-fold vs. >25-fold for WT UmuD) (Figure 7B). Moreover, even significant overproduction of DinB(F172A) was insufficient to impair UV-induced mutagenesis (data not shown). These observations suggest that a direct interaction between DinB and UmuD$_2$ is crucial for the ability of each protein to modulate the function of the other in vivo.
UmuD₂ inhibits DinB-dependent -1 frameshift activity in vitro

To gain more detailed insights into the mechanisms governing modulation of DinB function by UmuD in vivo, we reconstituted DinB dependent -1 frameshift activity in vitro with a substrate containing a G:G mispair that can be extended with either i) dGTP to generate a full-length product or ii) dATP to generate a -1 frameshift product that is one nucleotide shorter than the template (Figure 9A) (4, 49). DinB is unable to extend from this mispair using dGTP under our experimental conditions (data not shown). In contrast, DinB can readily act on this substrate using dATP, albeit at a rate that is 10-fold lower than its ability to extend from a G:C basepair in the same sequence context (Figure 9B). We were initially surprised to find that addition of UmuD₂ alone did not alter the -1 frameshift activity of DinB (data not shown). However, we then discovered that, when RecA is added in a stoichiometric ratio with DinB, the addition of UmuD₂ nearly completely inhibits the reaction (Figure 9B). In striking contrast, addition of UmuD₂ and RecA results in a ca. 20-fold enhancement of the ability of DinB to extend from a correctly paired terminus in the same sequence context (Figure 9B). These observations indicate that UmuD₂ and RecA act in concert to modulate DinB function in a highly sophisticated manner, promoting its ability to extend a properly paired primer terminus while suppressing its ability to extend a mismatched terminus.

Our discovery that RecA is required for UmuD₂ to modulate DinB mutagenic function in vitro led us to examine whether recA⁺ is similarly required for UmuD-dependent suppression of -1 frameshift mutagenesis in vivo. It is, as co-overproduction of UmuD has almost no effect on the frequency of DinB-dependent -1 frameshift mutagenesis in a ΔrecA strain (Figure 9C). Taken together, these findings provide strong evidence that RecA is required for UmuD₂ dependent modulation of DinB function.
In vitro, UmuD₂ reduced the maximal -1 frameshift activity of wild-type DinB by one half at a concentration of 840 nM, while a concentration of 3.6 μM is needed to cause an equivalent effect on DinB(F172A) (Figure 9D). Relative to wild-type UmuD₂, UmuD(D91A)₂ also shows a marked 10-fold decrease in its ability to inhibit DinB dependent -1 frameshift activity in vitro. Additionally, we observed that DinB is able to inhibit the RecA-mediated autocleavage of UmuD₂ in vitro, and moreover that the DinB(F172A) variant was unable to do so efficiently (Figure 9E). All of these data underscore the notion that a physical interaction between DinB and UmuD₂ exists under physiological conditions and is required for UmuD₂ dependent modulation of DinB function.

A TLS-deficient DinB variant is proficient for -1 frameshift function

It has been suggested that the -1 frameshift mutator signature of DinB is a direct consequence of structural features that enable it to act as a TLS DNA polymerase (42, 48). To ascertain whether DinB’s -1 frameshift mutator activity is separable from its function in TLS, we examined the DinB(F13V) variant, which is able to catalyze DNA synthesis on undamaged DNA but is virtually unable to perform TLS on certain adducted templates (12). We found that when overexpressed, this mutant is able to promote -1 frameshift mutagenesis in vivo (Figure 11A), indicating that DinB’s -1 frameshift mutator activity can be genetically separated from its ability to carry out proficient and accurate TLS over certain N²-dG adducts.

Curiously, the mutation frequency induced by overexpression of DinB(F13V) is about 5-fold greater than that produced by overexpression of wild-type DinB (Figure 3A) even though the levels of each protein are comparable in vivo (data not shown). This observation suggests either that DinB(F13V) has an increased -1 frameshift mutator activity or that some other
mechanism is responsible for this phenomenon \textit{in vivo}. Although the first explanation is formally possible, the -1 frameshift activity of DinB(F13V) \textit{in vitro} is slightly reduced (ca. 3-fold) relative to wild-type DinB. This observation suggests that DinB(F13V) promotes increased frameshifts by virtue of either its recruitment to, or association with, frameshift intermediates. Curiously, we were unable to observe UmuD$_2$ dependent inhibition of DinB(F13V) -1 frameshift activity either \textit{in vivo} or \textit{in vitro} (Figures 11A-B). These observations may be in part due to the fact that DinB(F13V) has a reduced affinity for UmuD$_2$ relative to the wild-type enzyme (Figure 12).

\textbf{RecA and UmuD$_2$ may modulate DinB function by restricting its open active site}

In an effort to explain how UmuD$_2$ suppresses the intrinsic -1 frameshift mutator activity of DinB in a RecA-dependent manner, we generated a model of a ternary complex among DinB, RecA, and UmuD$_2$. The structure of RecA (50) and models of DinB (12) and UmuD$_2$ (46) were docked using several constraints. First, UmuD$_2$ was optimally positioned on DinB using our peptide array data (Figure 4A-B). We then used distance constraints between RecA and UmuD$_2$ from published monocysteine crosslinking studies (51) to orient RecA relative to UmuD$_2$. Finally, we analyzed the RecA binding interface on DinB with an additional peptide array experiment (Figure 13). Together, these data were used to generate the working model shown in Figure 14A-B, which suggests that RecA and UmuD$_2$ act in concert to enclose the relatively open active site of DinB, perhaps thereby physically preventing the template bulging necessary for -1 frameshift mutagenesis. It is also provocative that in our model, RecA is positioned appropriately to interact with the end of a RecA-nucleoprotein filament, suggesting that this
interaction may also play a pivotal role in targeting DinB to RecA-coated substrates, a concept that has been previously proposed for UmuD'2C (52, 53).

Although the UmuD protein is only conserved among certain bacteria, we wondered whether the UmuD-binding interface on DinB might be maintained throughout evolution to interact either with a highly diverged UmuD or with a different but functionally equivalent partner protein. Since overproduction of the eukaryotic ortholog of DinB, pol κ, similarly promotes -1 frameshift mutagenesis (54), it is possible that a eukaryotic regulatory protein might interact in a similar location on pol κ. Although residues on DinB that participate in the interface are not strongly conserved, we considered whether they would exhibit statistical covariance through evolution. We therefore assembled an alignment of numerous DinB and pol κ sequences from all domains of life and examined which positions showed statistically significant covariance (55) with at least two of three residues that comprise the UmuD2 interacting interface of DinB. Strikingly, these residues define an interface on pol κ that is similar to the one we identified experimentally on DinB (Figure 14C). These observations suggest that there may be an evolutionary driving force for the maintenance of this interface, perhaps as a site for regulatory protein-protein interactions.
Discussion

In this paper, we have used several *in vitro* methods to demonstrate completely unanticipated direct interactions among DinB, UmuD$_2$, UmuD'$_2$, and RecA (Figure 2A-C) and present evidence that these interactions are physiologically relevant. The dissociation constant between DinB and UmuD$_2$ is 620 nM while the uninduced cellular concentrations of DinB and UmuD are approximately 400 nM and 350 nM respectively. Thus, there are sufficient cellular quantities of UmuD to interact with nearly every molecule of DinB. The estimated K$_d$ between DinB and RecA is *ca.* 1 µM and the RecA concentration under normal conditions is 2.5 µM. The cellular levels of all of these proteins rise at least 10-fold upon SOS induction (11, 36, 56). Thus, DinB is likely to exist as at least a binary and perhaps ternary complex under many physiological conditions. Indeed, we have been able to isolate a ternary complex of these three proteins that is stable on the minute timescale (Figure 2C). Furthermore, these findings provide a potential rationale for the observation that the physiological levels of the *umuD* gene products greatly exceed those of UmuC.

We find that the -1 frameshift mutagenesis induced by overproduction of DinB can be suppressed by co-overproduction of UmuD (Figure 3A-B). Moreover, a noncleavable UmuD variant completely suppresses -1 frameshift mutagenesis, indicating that only the function of full-length UmuD is required to control DinB's mutator potential (Figure 3A). Reciprocally, the DinB-binding deficient UmuD(D91A) variant (Figure 4C), although perfectly proficient for RecA-mediated autoproteolysis (Figure 8) does not efficiently reduce DinB dependent -1 frameshift mutagenesis (Figure 7B). Intriguingly, when mapped onto a structural model of UmuD$_2$, the residue Asp91 in UmuD does not form a contiguous interface with the other residues identified as part of the DinB-binding interface as it is occluded by the curled N-terminus of
UmuD (Figure 4B). However, recent computational studies have suggested that isoenergetic conformations of UmuD₂ can be formed in which its N-terminus is raised (46), thereby potentially favoring interaction with DinB but also rendering it incompetent for autoproteolysis (Figure 5). Such conformational flexibility is completely consistent with the recent discovery that UmuD₂ and UmuD'₂ are intrinsically disordered proteins (Simon and Walker, in preparation). Our observation that DinB inhibits UmuD₂ autoproteolysis in vitro and UV-induced mutagenesis in vivo is also consistent with this notion. Overproduction of UmuD' can also suppress -1 frameshift mutagenesis, albeit to a lesser degree, indicating that the interface between DinB and UmuD₂ or UmuD'₂ does not critically depend on the UmuD N-terminus, in agreement with our observations using peptide arrays (Figure 4B). Overproduction of UmuD, and to a lesser extent UmuD', also inhibits adaptive mutagenesis (Figure 3B), which occurs via a -1 frameshift event (40). Irrespective of the precise molecular mechanisms of adaptive mutagenesis, it is clear that the umuD gene products can play a role in modulating it.

Our efforts to reconstitute the UmuD₂ modulation of DinB frameshift activity in vitro using a mismatched substrate (4, 49) revealed that RecA needs to be present in stoichiometric quantities with DinB for this to occur. Moreover, the addition of RecA and UmuD₂ to an assay in which DinB replicates a template with a properly paired terminus results in a remarkable increase in DinB catalytic proficiency (Figure 9B). Consistent with these observations, we found that recA⁺ is required for the suppression of DinB-dependent -1 frameshift mutator activity in vivo, as the co-overproduction of UmuD₂ does not suppress this mutagenesis in a recA strain (Figure 9C). A noncleavable UmuD(S60A) variant is fully proficient for suppression of -1 frameshift mutagenesis (Figure 3A), indicating that RecA's role in this phenomenon is distinct from its function in promoting UmuD autocleavage.
Taken together, these discoveries significantly expand our view of how the fidelity of Y-family DNA polymerases is regulated in response to DNA damage or environmental stress. Under normal conditions the levels of UmuD2 are relatively low but match those of DinB. Upon SOS induction, UmuD2 predominates for 30-40 minutes, a phase during which accurate repair and damage tolerance mechanisms operate (57). The subsequent phase, in which UmuD'2 predominates, is when potentially mutagenic TLS by UmuD'2 C takes place. Our results suggest that in uninduced cells, as well as during the first phase of SOS induction, UmuD2 and RecA would act in concert to restrict the -1 frameshift mutagenic potential of DinB by closing in its active site and simultaneously stimulating its activity on templates with properly paired termini. This control would then be relaxed during the second more mutagenic phase of SOS induction. Our results further suggest that DinB would be mutagenic under conditions of chronic SOS induction.

We propose, as have others (42, 48), that the propensity of DinB and its orthologs for -1 frameshift mutagenesis may arise as a result of their unique active sites that are specialized for TLS function. The equivalent levels of DinB and UmuD point to a possible mechanism through which the mutagenic potential of DinB is regulated, and perhaps exploited, by the cell. Elevation of DinB levels above those of its UmuD2 manager protein, either synthetically by DinB overproduction or apparently naturally during adaptive mutagenesis, renders the cell vulnerable to the full -1 frameshift mutator potential of DinB. Indeed, the precise tuning of DinB levels relative to those of UmuD2 under both basal and SOS induced conditions may facilitate the modulation of its function in response to environmental stress.

Our discovery that UmuD2 and RecA appear to modulate the mutagenic potential of DinB by enclosing its open active site may have implications for the control of DinB orthologs in
other organisms. Overproduction of mammalian pol κ has similarly deleterious mutagenic consequences to DinB overproduction (58), despite the fact that its active site appears to be comparatively closed (59). A comparison of the pol κ structure with our working model of the DinB-UmuD₂-RecA complex suggests that part of the pol κ structure may play a role equivalent to RecA, but that its function might be further regulated by a partner protein interacting in a manner analogous to UmuD interacting with DinB.

We had anticipated that deletion of either umuD or recA would affect DinB TLS function, and were therefore initially surprised that deletion of umuD had no effect on dinB⁺-dependent resistance to NFZ in a wild-type E. coli strain. Deletion of recA dramatically increased sensitivity to NFZ but, because of the central role of RecA in coordinating numerous aspects of the DNA damage response (17), we were unable to infer that this results from a loss of DinB mediated TLS. This led us to wonder whether the -1 frameshift mutator activity of DinB is genetically separable from its ability to proficiently and accurately bypass certain N²-dG adducts. We have previously characterized a separation of function mutant of DinB’s steric gate residue, DinB(F13V), which is active as a conventional DNA polymerase but is virtually unable to catalyze TLS on N²-dG adducted templates (12). In vivo, DinB(F13V) is exceptionally proficient at promoting -1 frameshift mutagenesis, although its levels are comparable to wild type by immunoblot, indicating that certain mechanistic attributes of DinB that are required for its proficient and accurate N²-dG lesion bypass ability are not required for its -1 frameshift mutator activity (Figure 11A).

Taken together our results indicate that DinB plays at least two separable roles in nature, one in which it promotes survival by accurately bypassing a common class of N²-dG adducts and another in which it promotes mutagenesis under stressful conditions. The additional levels of
DinB regulation we have described suggest ways in which the balance of these two diverse roles could be tuned to the physiological conditions being experienced by the organism.
Experimental Procedures

Protein expression and purification

DinB, UmuD, UmuD', and RecA were purified as described previously (12, 60, 61), but 50 mM Hepes pH 7.2 was used exclusively. Plasmids expressing DinB(F172A) and UmuD(D91A) were constructed from pDFJ1 and pSG5 (60) using a Quickchange kit (Stratagene). Both DinB(F172A) and UmuD(D91A) behaved as wild-type DinB and UmuD during purification. A plasmid encoding (His)₆HMK-DinB (62) was constructed in pET16B using standard cloning procedures. (His)₆HMK-DinB was purified using Ni²⁺-NTA affinity resin (Qiagen) following the manufacturer’s instructions.

Affinity chromatography

The details of the procedure are described in the figure legends.

Crosslinking and binding measurements

Crosslinking reactions were initiated as described (63) and allowed to proceed for 10 min before quenching with SDS-PAGE loading buffer containing 5% β-mercaptoethanol. Fluorescence spectroscopy was performed as described previously (60).

Farwestern assays and peptide array experiments

Farwestern blots were performed as previously described (63). Cellulose filter peptide arrays were synthesized with overlapping 12-mer peptides offset by two residues (MIT CCR Core Facility). The arrays were probed with 1 μM UmuD₂ or DinB, and washed and developed as described (64). Control arrays were performed using DinB or UmuD₂ antibodies alone.
Mutagenesis assays

DinB dependent -1 frameshift mutagenesis experiments were performed as described (16) except that IPTG was not added to the media. Briefly, P90C containing the CC108 episome and indicated plasmids were selected on minimal media, grown in L medium containing ampicillin, washed and plated on minimal lactose medium. Viable counts were determined by plating on minimal glucose medium in parallel. Ampicillin (100 μg/mL) and spectinomycin (60 μg/mL) were used as necessary for plasmid maintenance. Adaptive mutagenesis was performed as previously described (65). A table of the strains and plasmids used in this study is available as supplemental material (supplemental Table 1).

DinB and UmuD₂ activity assays

DinB was assayed as described previously (12) except 50 nM enzyme and 10 nM primer/template was used. The oligonucleotides 5'-
ATCCTAGTCCAGGCTGCTGACAACTCGGGAACGTGCTACATGAAT-3', 5'-
ATTCTAGCAGCGTTCCC-3', and 5'ATTCTAGCAGCGTTCCG-3' were designed based on those used previously (4). Reactions were initiated with the appropriate dNTP, quenched after 20 min, and separated on a 16% denaturing polyacrylamide gel, which was quantified using a Typhoon phosphorimager (GE Healthcare). UmuD₂ autocleavage reactions were performed as described previously (60).

Molecular modeling and statistical covariance

A model of the DinB-UmuD₂ complex was constructed using the application 3D-dock based on the following constraints: E168 < 6 Å from either UmuD chain, and L176, P177, K180,
and F172 < 8 Å from D91 of either UmuD chain. The UmuD2·RecA model was made using the following constraints: RecA S117 < 7 Å from either UmuD chain (51), UmuD residues L101, R102, V34, S81 are 6-12 Å from RecA, and UmuD residue E11 is 6-25 Å from RecA. Resulting complexes were filtered based the RecA residues T243 and R244 < 10 Å from DinB to generate a model of the DinB·RecA·UmuD2 ternary complex. We performed the same procedure by docking DinB to the UmuD2·RecA models using the constraints described above and obtained similar results. Statistical covariance was performed by aligning 84 DinB and pol κ sequences from diverse organisms using ClustalW and analyzing significant pairwise correlation of alignment positions with the CRASP algorithm (55).

Acknowledgements

We wish to thank Steve Bell for use of the phosphorimager. VGG performed the mutagenesis and peptide array experiments and identified the interaction between DinB and UmuD. DFJ performed the protein purification, in vitro experiments, and constructed the mutant plasmids used in this study. VGG, DFJ, and GCW wrote the manuscript. SMS measured the dissociation constant between DinB and the umuD gene products. AA and VAI generated the model of the DinB·RecA·UmuD2 complex. This work was supported by NIH grant CA021615 to GCW and NIEHS grant P30 ES002109 to the MIT Center for Environmental Health Sciences. GCW is an American Cancer Society Professor. SMS was a Cleo and Paul Schimmel fellow.
References


<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>E. coli B. Standard strain used for protein overproduction</td>
<td>GCW lab stock</td>
</tr>
<tr>
<td>BL21 (DE3)</td>
<td>thr-1 araC-14 leuB-6(Am) Δ(gpt-proA)62 lacY1 tsx-33 qsr'-0 gltV44(AS) galK2(Oc) LAM Rac-0 hisG4(Oc) rfbC1 mgl-51 rpoS396(Am) rpsL31 kgdK51 xylA5 mtl-1 argE3(Oc) thi-l</td>
<td>GCW lab stock</td>
</tr>
<tr>
<td>AB1157</td>
<td>Δ(lac-pro)xu ara gal</td>
<td>Cairns and Foster (1991)</td>
</tr>
<tr>
<td>P90C</td>
<td>As P90C but with recA gene deletion</td>
<td>This work</td>
</tr>
<tr>
<td>VG187</td>
<td>As P90C but with a precise deletion in dinB</td>
<td>This work and Jarosz et al. (2006)</td>
</tr>
<tr>
<td>VG347</td>
<td>As P90C but bearing an F'128 episome with the Lac- LacI33 allele revertible to Lac+</td>
<td>Cairns and Foster (1991)</td>
</tr>
<tr>
<td>FC40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC108</td>
<td>A derivative of the F'128 episome with a Lac-reversible allele</td>
<td>Kim et al. (1997)</td>
</tr>
<tr>
<td>pYG782</td>
<td>dinB gene cloned in the pWSK30 plasmid (AmpR) which is as pWSK29 but its MCS cloned in opposite orientation under Plac</td>
<td>Kim et al. (1997)</td>
</tr>
<tr>
<td>pYG768</td>
<td>Same as pYG768 but the dinB gene is cloned under its own promoter</td>
<td>Kim et al. (1997)</td>
</tr>
<tr>
<td>pYG768-F172A</td>
<td>Same as pYG768 but the dinB gene encodes the F172A mutation.</td>
<td>This work</td>
</tr>
<tr>
<td>pYG782-F172A</td>
<td>dinB gene encoding the F172A mutation cloned in pWSK30</td>
<td>This work</td>
</tr>
<tr>
<td>pGW60101</td>
<td>pBR322 (SpecR) vector</td>
<td>This work</td>
</tr>
<tr>
<td>pGW6030</td>
<td>umuD gene cloned in pGW60101 plasmid under the lac promoter</td>
<td>GCW collection</td>
</tr>
<tr>
<td>pGW6030-D91A</td>
<td>umuD gene encoding the D91A mutation cloned in pGW60101 under lac promoter</td>
<td>This work</td>
</tr>
<tr>
<td>pGW6030-S60A</td>
<td>umuD gene encoding the S60A mutation cloned in pGW60101 under lac promoter</td>
<td>This work</td>
</tr>
<tr>
<td>pGW6040</td>
<td>Same as pGW6030 but with the umuD' gene.</td>
<td>GCW collection</td>
</tr>
<tr>
<td>pSG5-6030-D91A</td>
<td>Same as pGW6030-D91A but in pET11T</td>
<td>Beuning et al. (2006).</td>
</tr>
<tr>
<td>pET11T-DinB-F172A</td>
<td>Same as pYT782-F172A but in pET11T</td>
<td>This work</td>
</tr>
</tbody>
</table>

Table 1. Strains and plasmids used in this study.
Figure 1. Diagram of the procedure followed to find DinB interacting proteins. DinB affinity resin was generated by coupling 3 mL of Affigel-10 N-hydroxysuccinimide derivatized agarose beads (Bio-RAD) with 30 mg of DinB at 4°C. The coupling was monitored by determination of the protein concentration remaining in solution, and quantitative conversion was observed after 4 h. The resin was blocked with excess 1 M Tris(hydroxymethyl) aminomethane pH 7.8 for 1 h and loaded into three 1 mL columns. A control resin was generated using 30 mg hen egg white lysozyme (Sigma) due to its elevated isoelectric point (pI 11.0) and was treated identically with respect to subsequent experimental steps. The DinB affinity columns and lysozyme controls were equilibrated in 50 mM Hepes pH 7.5, 100 mM KCl, 5% glycerol, 1 mM DTT (wash buffer). One liter of lexA(Def) E. coli cells were grown to fresh saturation in Luria Burtani (LB) broth and a lysate was generated using a French pressure cell. The lysate was cleared by ultracentrifugation at 100,000 x g and separated into three equal aliquots, one of which was left untreated and the other two were digested with either DNAse or RNAse. Each aliquot was divided in half and applied to both a DinB affinity column and a lysozyme control. Each column was washed with 50 mL of wash buffer and proteins remaining on the resin were identified by removing 100 μL of each resin and boiling with 100 μL SDS-PAGE loading buffer. The eluates were separated by SDS-PAGE and identified by Edman degradation.
Figure 2. DinB interacts directly with UmuD and RecA. A. Farwestern blot demonstrates that UmuD directly interacts with \(^{32}\)P-labeled (His)\(_6\)HMK-DinB. Either 50 or 100 pmols of UmuD or UmuD' were separated by 12% SDS-PAGE and transferred to a PDVF membrane. The HMK-DinB protein probe was radioactively labeled and incubated with the membrane for 15 min. after which the membrane was exposed to film. B. Crosslinking experiment suggests that DinB interacts with the UmuD\(_2\) homodimer. One hundred pmol of DinB and UmuD\(_2\) were mixed in a 10 \(\mu\)L volume in 50 mM Hepes pH 7.5, 25-500 mM NaCl, and 1 mM DTT and incubated for 10 min at 25\(^{\circ}\)C. C. DinB forms a stable binary and ternary complexes with RecA and UmuD\(_2\). DinB(His)\(_6\) pulls down UmuD, UmuD', and RecA on a Ni\(^{2+}\) affinity resin. The presence of RecA slightly increases the amount of UmuD that is recovered.
Figure 3. UmuD<sub>2</sub> regulates the -1 frameshift activity of DinB in vivo. A. Lac<sup>+</sup> reversion assay of the frameshift allele in CC108 demonstrates that UmuD<sub>2</sub> modulates DinB’s -1 frameshift function. B. UmuD<sub>2</sub> also affects the number of Lac<sup>+</sup> revertants in an adaptive mutagenesis experiment.
**Figure 4.** Molecular characterization of the interaction between DinB and UmuD₂. A. Peptide array mapping of the UmuD binding interface on DinB reveals a surface composed of the thumb and finger domains of the polymerase. Several hydrophobic residues in the most strongly interacting peptide are conserved among DinB orthologs from organisms containing umuD. B. Peptide array mapping of the DinB binding interface on UmuD₂ reveals a discontinuous interface on a structural model of trans-UmuD₂ that is rendered contiguous in an alternative isoenergetic trans-UmuD₂ conformer. C. Alanine mutants of DinB Phe172 or UmuD Asp 91 result in a weakened interaction determined by fluorescence spectroscopy. D. A low-copy number plasmid encoding DinB(F172A) is fully able to rescue the NFZ sensitivity of a ΔdinB E. coli strain.
Figure 5. DinB interacting interface mapped onto four isoenergetic models of UmuD₂. The N-terminal arms of each UmuD monomer can either interact with the other molecule of UmuD in the dimer in a \textit{trans} conformation, or with itself in a \textit{cis} conformation (46). Moreover, the arms themselves can either adopt a conformation that permits autocleavage (catalytically competent) or prevents it (catalytically incompetent) in these models. The DinB interacting peptides from cellulose filter peptide array analysis localize to a more contiguous surface in the catalytically incompetent conformations.
Figure 6. SDS-PAGE of DinB(F172A). DinB(F172A) was purified as described (12) and behaved as wild-type throughout the purification.
Figure 7. Single amino acid changes on the interface between DinB and UmuD perturb regulation of -1 frameshift activity. A. The DinB(F172A) variant has a lower affinity for UmuD2 and is not as responsive as wild type DinB to regulation by UmuD2. B. The UmuD(D91A) variant has a lower affinity for DinB and does not regulate -1 frameshift activity as well as wild type UmuD.
Figure 8. Autocleavage assay and SDS-PAGE of UmuD(D91A). UmuD(D91A) is proficient for RecA* mediated autocleavage (15 min timepoint) and can be purified to homogeneity just as wild-type UmuD.
Figure 9. UmuD₂ and RecA directly modulate DinB -1 frameshift function. A. Schematic of a mismatched DNA substrate that can be extended either by dGTP to generate a full length product or by dATP, thereby forming a dNTP stabilized misalignment and generating a -1 frameshift product. B. Plot of reaction velocity vs. dNTP substrate concentration for DinB alone and in combination with RecA and UmuD₂. Extension of the GG mismatch by DinB alone with dATP (open squares) is detectable but weaker than extension of a GC basepair in the same sequence context (open circles). The addition of RecA in stoichiometric ratios with DinB and saturating (10 μM) UmuD₂ profoundly inhibits DinB activity on a GG mismatch (closed squares) but stimulates DinB activity on a GC (closed circles) by more than 20-fold. C. recA⁺ is required for UmuD dependent inhibition of DinB promoted -1 frameshift mutagenesis in vivo. Overproduction of DinB promotes -1 frameshift mutagenesis in a ΔrecA background but the co-overproduction of UmuD₂ has little effect on mutation frequency. D. Plot of percent frameshift inhibition vs. UmuD variant concentration. The frameshift activity of DinB is efficiently inhibited by UmuD₂ (closed circles) but the frameshift activity of DinB(F172A) is more inert to UmuD₂ suppression (open circles). The UmuD(D91A) variant is also very inefficient at inhibiting the -1 frameshift activity of wild-type DinB (open triangles). DinB efficiently inhibits UmuD₂ autocleavage in vitro. Wild-type DinB (open circles) inhibits E. UmuD₂ autocleavage far better than DinB(F172A) (closed circles).
**Figure 10.** Fluorescence spectroscopy data for the interactions between UmuD₂ and DinB. Fraction DinB bound was determined as described in experimental procedures and is plotted against increasing concentrations of UmuD₂. Closed circles represent the interaction of UmuD₂ with wild type DinB and open circles represent the interaction with the DinB(F172A) variant. Data were fit as described in experimental procedures.
Figure 11. A TLS deficient variant of DinB is proficient for -1 frameshift function. A. The DinB(F13V) variant can promote -1 frameshift mutagenesis but is not controlled by co-overproduction of UmuD. B. The -1 frameshift activity of DinB(F13V) is poorly inhibited by UmuD in vitro. Plot of frameshift activity vs. UmuD concentration indicates that DinB(F13V) (open squares) retains much of its frameshift activity at concentrations of UmuD that inhibit virtually all DinB frameshift activity (closed circles). All reactions contain RecA in stoichiometric ratios with DinB.
Figure 12. Fluorescence spectroscopy data for interactions between UmuD$_2$ and DinB(F13V). Fraction DinB bound was determined as described in experimental procedures and is plotted against increasing concentrations of UmuD$_2$. Closed circles represent the interaction of UmuD$_2$ with DinB(F13V).

Figure 13. Putative RecA Interaction Site on DinB. DinB model on which RecA interacting residues (colored in red) are mapped based on peptide array analysis.
Figure 14. RecA and UmuD₂ enclose the open active site of DinB. A-B. *In silico* modeling of a ternary complex of the proteins. The surface representation of DinB is shown in blue, UmuD₂ in yellow, and RecA in orange. The DNA is relatively enclosed in the complex. C-D. Statistical covariance of DinB/polκ residues across evolution. Residues that display statistical covariance with the UmuD₂ binding interface on *E. coli* DinB define an interface in a similar position on polκ, suggesting a possible rationale for the maintenance of this interface as a site of regulatory protein-protein interactions.
Chapter 6

Y-family DNA Polymerases Respond to DNA Damage-independent Inhibition of Replication Fork Progression

This chapter was previously published as VG Godoy, DF Jarosz, FL Walker, LA Simmons, and GC Walker “Y-family DNA Polymerases Respond to DNA Damage-independent Inhibition of Replication Fork Progression.” (2006) *EMBO Journal* 25(4):868-79
Abstract

In *Escherichia coli*, the Y-family DNA polymerases Pol IV (DinB) and Pol V (UmuD'2C) enhance cell survival upon DNA damage by bypassing replication-blocking DNA lesions. We report a unique function for these polymerases when DNA replication fork progression is arrested not by exogenous DNA damage, but with hydroxyurea (HU), thereby inhibiting ribonucleotide reductase, and bringing about damage-independent DNA replication stalling. Remarkably, the *umuC122*~Tn5* allele of *umuC*, *dinB*, and certain forms of *umuD* gene products endow *E. coli* with the ability to withstand HU treatment (*HU^R*). The catalytic activities of the UmuC122 and DinB proteins are both required for *HU^R*. Moreover, the lethality brought about by such stalled replication forks in the wild-type derivatives appears to proceed through the toxin/antitoxin pairs *mazEF* and *relBE*. This novel function reveals a role for Y-family polymerases in enhancing cell survival under conditions of nucleotide starvation, in addition to their established functions in response to DNA damage.
Introduction

In both eukaryotes and prokaryotes (1), initiation of DNA replication is exquisitely regulated, and sophisticated systems have evolved to contend with the potentially lethal consequences of inhibition of replication fork progression (2). Depletion of dNTP pools leads to arrest of cell division in eukaryotes (3) and prokaryotes (4) until DNA replication is properly restored. Mutations in components of such checkpoints result in genomic instability and elevated mutation frequencies that may lead to cancer in higher organisms (5). Responses to arrest of fork progression include induction of DNA damage tolerance pathways. While the rationale for such a response is clear when stalling is brought about by exogenous DNA damage, it is more enigmatic (6) when replication fork progression is inhibited in a DNA damage-independent fashion.

Y-family polymerases possess properties that are advantageous for the resolution of replication forks stalled by DNA damage as they have the ability to insert nucleotides opposite DNA lesions that block replicative DNA polymerases, a process termed translesion synthesis (TLS) (7). TLS often ensues with comparatively low fidelity, meaning that bypass of DNA damage takes place at a potentially mutagenic cost (8). Notable exceptions exist, however, such as eukaryotic Pol η bypassing cyclobutane pyrimidine dimers (9).

The Y-family DNA polymerases are encoded in E. coli by the dinB and umuDC genes which are both regulated by the LexA transcriptional repressor as part of the SOS response to DNA damage (10). Initially, full length UmuD is expressed from the umuDC operon. The UmuD homodimer interacts with UmuC to effect a DNA damage checkpoint function (11), and cold sensitivity due to overproduction of umuDC (12) appears to result from an exaggeration of this function (13). UmuD thereafter undergoes removal of its first 24 amino-acids, dependent on the
RecA nucleoprotein filament (14), to form UmuD'. The UmuD' homodimer (UmuD'$_2$) is a positive effector of UmuC, the catalytic subunit of Pol V (15). Transcription of the dinB gene is weakly repressed by LexA, so that basal levels of DinB are high compared to those of UmuC (16). Indeed, upon SOS induction Pol IV is the most abundant DNA polymerase in the cell (16). Among Y-family polymerases, the DinB subfamily is strikingly conserved, and it is the only branch present in all domains of life (17).

Hydroxyurea (HU) has been widely used to investigate responses to DNA damage-independent replication arrest (18). HU inhibits class I ribonucleotide reductases (RNR), such as that of aerobically-grown E. coli (19), by scavenging a stable di-iron tyrosyl radical that is essential for catalysis. RNRs catalyze the conversion of ribonucleotides into deoxyribonucleotides – the rate-limiting step in DNA biosynthesis in most organisms (19). Levels of intracellular deoxyribonucleotide triphosphates (dNTPs) are thought to decline upon HU treatment such that DNA replication is arrested through substrate starvation (20).

We report that the E. coli Y-family polymerases Pol IV and Pol V play a role upon DNA damage-independent replication stalling. Strains bearing novel umuC alleles are unexpectedly HU$^R$, challenging the notion that replication inhibition by HU arises solely from dNTP starvation. Genetic analyses demonstrate that the dinB and umuD gene products also participate in the DNA damage-independent response to inhibition of replication fork progression. Together, these data suggest combined action of the UmuC derivatives together with the dinB and umuD gene products at these stalled replication forks. Moreover, we also find that the lethality of such replication fork arrest in wild type derivatives is alleviated independently by mutation of the mazEF and relBE toxin/antitoxin pairs, suggesting that the action of these Y-
family polymerases may prevent mazEF or relBE-mediated lethality under conditions of nucleotide starvation.
Results

_E. coli_ carrying the _umuC122::Tn5_ allele are unexpectedly resistant to HU

We were interested in whether the _umuC^+ _gene product might be part of the cellular response when replication fork progression is inhibited in a DNA damage-independent manner by dNTP depletion. We therefore examined a set of strains carrying null alleles of _umuC_ for their sensitivity to killing by HU (Figure 1A). A strain in which the _umuDC_ operon has been deleted is as sensitive to killing by HU as its _umuD^+C^+ _parent. Intriguingly, a strain carrying a precise _ΔumuC_ deletion that leaves the _umuD^+ _gene intact displays a modest level of resistance to killing (Figure 1A). Perhaps either or both of the _umuD^+ _gene products might contribute to HUR in the absence of UmuC (see below).

We also tested _umuC122::Tn5 (umuC122)_ , which is known to behave as a _umuC_ null allele with respect to induced mutagenesis caused by UV radiation and many chemicals (21). We found that strains carrying _umuC122_ are at least 100-fold more resistant to killing by HU than their _umuC^+ _parents and can in fact multiply during HU treatment (Figure 1A). We observed this HU^R _phenotype in all strain backgrounds tested, including AB1157 (Figure 1B) (22). These observations indicate that _umuC122_ is a gain-of-function _umuC _allele with regard to cell survival after HU treatment. This is plausible since the Tn5 insertion results in a missense mutation followed immediately by a termination codon giving rise to a predicted 32 kDa UmuC protein lacking its last 102 residues (23). The truncation occurs downstream of the conserved polymerase domain common to Y-family DNA polymerases (24). Immunoblotting confirmed that the _umuC122_ allele indeed encodes a UmuC derivative of this molecular weight (Figure 1C). We observed that the truncated UmuC122 protein appears to be expressed at higher levels than wild-type UmuC (data not shown), though this may be because one of the synthetic peptides
used to raise antibodies against UmuC lies immediately at the C-terminus of the UmuC122 protein, perhaps resulting in a more accessible epitope relative to full-length UmuC. The UmuC122 protein may also lack one or more C-terminal motifs that would normally target the protein for Lon-mediated proteolytic degradation (25). Overexpression of UmuC did not confer statistically significant HU\(^R\) (data not shown).

**The umuC122 allele alleviates the lethal effects of class I RNR inhibition by HU**

The observation of an HU\(^R\) phenotype as a consequence of a umuC mutation was unanticipated since most previously reported HU\(^R\) mutants affect RNR (20). By immunoblotting we showed that the levels of the small and large subunits of RNR are not affected in strains bearing the umuC122 allele during HU treatment (Figure 2A). Also, we found that the protective effect of umuC122 is observed with other RNR inhibitors such as guanazole (Figure 2B). We sought evidence that the umuC122 mutation helps cells recover from the lethal consequences of HU-mediated RNR inhibition instead of acting by some other mechanism. Therefore, we took advantage of the fact that anaerobically grown *E. coli* utilize an HU-insensitive class III RNR rather than the HU-sensitive class I RNR used during aerobic growth (26). As shown in Figure 2C, we found that the anaerobically grown HU-treated umuC\(^+\) and umuC122 strains were both insensitive to HU. These observations indicate that the umuC122 mutation alleviates the lethality caused by HU inhibition of the class I RNR in *E. coli* through a mechanism that does not involve alteration of RNR protein levels.
Resistance to HU requires the catalytic activity of the truncated UmuC122 protein

To facilitate further analysis of the genetic requirements for *umuC122*-mediated HUR, we tested whether a plasmid-borne *umuDumuC122* (pDC122), expressed in a Δ*umuDC* derivative, conferred HUR. This was indeed the case (Figure 3A). To determine whether this HUR requires the catalytic activity of UmuC122, we used the *umuC104* allele (D101N) (Figure 3D) (23), which alters a conserved catalytic residue common to all Y-family polymerases (24). The addition of pDC104 had little effect on resistance to killing by HU (Figure 3A). However, introduction of the D101N mutation into pDC122 eliminated HUR (Figure 3A), indicating that the UmuC122 protein must be catalytically active to observe this phenotype.

A unique *umuC* missense allele also confers resistance to HU

We also tested the response to HU of the *umuC125*, a *umuC* allele bearing a A39V mutation, which does not affect the ability of UmuC to function in UV mutagenesis but eliminates the cold-sensitivity observed when it is over-expressed together with UmuD (27). We found that Δ*umuDC* cells containing pDC125, the plasmid-borne version of *umuC125*, are also resistant to HU, although not to as high a level as observed with pDC122 (Figure 3A). This observation indicates that HUR is not a unique property of the *umuC122* allele, but can be mimicked, at least in part, by a simple missense mutation affecting the N-terminus of UmuC.

The *umuD* gene is required for HU resistance

The data presented in Figure 1A suggest that the *umuD* gene product(s) might contribute to HUR in the absence of UmuC. Furthermore, the *umuD* gene products influence the biological function of UmuC (15). We therefore assessed whether either form of the *umuD* gene product is
required for the high level of HU\(^R\) we observed in *umuC122* bearing strains. The *umuD*(S60A) mutation (23) eliminates the serine that serves as the nucleophile in RecA-mediated UmuD autocleavage, so that only full-length UmuD is produced \([pD(S60A)C]\). Alternatively, the DNA encoding the first 24 amino acids in the N-terminus of UmuD can be deleted so that UmuD' is synthesized directly \((pD'C)\) (15).

As shown in Figure 3B, Δ*umuDC* cells with a plasmid carrying *umuD*(S60A) *umuC122* \([pD(S60A)C122]\) exhibited a lower level of HU\(^R\) than the corresponding cells bearing the *umuC122* plasmid \((pDC122)\), but nevertheless were substantially HU\(^R\). Similarly, Δ*umuDC* cells bearing \(pD'C122\) exhibited a lower level of HU\(^R\) than the corresponding \(pDC122\) bearing strain, but were still HU\(^R\). These results suggest that the full degree of HU\(^R\) displayed by a *umuDumuC122* strain requires both forms of the *umuD\(^+\)* gene product. The two forms might act sequentially, first the UmuD\(_2\) homodimer and then the UmuD'\(_2\) homodimer. If so, it would appear that the component of HU\(^R\) requiring UmuD\(_2\) is more substantial than the component requiring UmuD'\(_2\). Another possibility is that a component of the HU\(^R\) requires the action of the UmuD-UmuD' heterodimer, which is known to be the most stable form *in vitro* (28).

We performed similar experiments with the *umuC125* plasmid-borne allele (Figure 3C, note y-axis scale) in which we examined HU\(^R\) when *umuD*(S60A) and *umuD'* were combined with *umuC125*. Interestingly, in contrast to the situation with *umuC122*, the strain bearing *umuC125* displayed comparable HU\(^R\) relative to the strain bearing \(pDC125\), whereas the strain bearing \(pD'C125\) showed substantially less HU\(^R\). These data, combined with the fact that the level of HU\(^R\) of a *umuC125* strain is less than that of a *umuC122* (Figure 3C, note y-axis scale), suggest that the UmuC125 protein is less proficient at the UmuD'-dependent component of HU\(^R\) than the UmuC122 protein.
The *dinB*\(^+\) gene is required for HU\(^R\)

The results presented to this point indicate that the high level resistance of certain *umuC* mutants to killing by HU also requires certain forms of the UmuD protein. Involvement of DinB in HU\(^R\) would be consistent with reports that DinB cooperates with UmuC in translesion synthesis past certain lesions (29). Furthermore, under both induced and uninduced conditions, the intracellular levels of the *umuD* gene products are much higher than the estimated intracellular concentrations of UmuC, but are approximately equal to those of DinB (30). Therefore, we constructed a strain with a precise deletion of the *dinB*\(^+\) gene in *umuC*\(^+\) and *umuC122* backgrounds. In a *umuC*\(^+\) strain, loss of *dinB*\(^+\) results in a slight sensitivity to HU (Figure 4A). However, introduction of the ∆*dinB* mutation into the strain carrying the *umuC122* allele eliminates the high level of HU\(^R\) observed in this strain (Figure 4A). Thus, the *dinB*\(^+\) gene product is essential for the HU\(^R\) exhibited in *umuC122* strains.

We asked whether HU\(^R\) could be restored in a *umuC122ΔdinB* mutant by introducing plasmids carrying the *dinB*\(^+\) gene. We were unable to complement HU\(^R\) *in trans* with low- or high-copy number plasmids bearing *dinB*\(^+\). However, by transducing the wild type copy of the *dinB*\(^+\) gene into the *umuC122ΔdinB* mutant, the HU\(^R\) phenotype was restored (Figure 4B). The possibility that the restoration is due to a closely linked locus rather than to *dinB*\(^+\) is inconsistent with the data presented in the following section. These observations suggest that the level of DinB expression or a *cis* regulatory element is critical for the ability of *dinB*\(^+\) to contribute to the HU\(^R\) of a *umuC122* strain. Perhaps DinB cannot contribute to HU\(^R\) if its levels do not correlate with those of the products of the *umuD*\(^+\) gene.
The catalytic activity of DinB is required for HU resistance

To test whether DinB must be catalytically active to contribute to HUR, we introduced the \textit{dinB003} mutation into the chromosome of a \textit{umuC122} strain. This mutation (D103N) alters a conserved aspartic acid residue required for phosphodiester bond formation (31). The large loss of HUR we observed (Figure 4B) suggests that DinB is indeed acting as a DNA polymerase as it contributes to HUR. Thus, it appears that HUR results from the combined action of two DNA polymerases, DinB and a mutant form of UmuC, acting together with UmuD and UmuD'.

DNA synthesis is slowed in both parental and \textit{umuC122} strains during HU challenge

To explain the observation that both Y-family polymerases are required for HUR, we asked whether HUR was simply due to an extensive alteration in the rate of DNA replication. We measured DNA synthesis by examining the ability of \textit{thyA}^- derivatives of wild type and \textit{umuC122} strains to incorporate thymidine (\(^{3}H\)-Thy) in 10 min pulses during HU treatment. We found that the amount of DNA synthesis is reduced during HU treatment in both wild-type and \textit{umuC122} strains compared to untreated controls (Figures 5A and 5B). Any minor changes that we observe in the ability to incorporate \(^{3}H\)-Thy into the DNA do not appear to account for the striking difference in viability, i.e. competence to develop colonies, between the HU-treated wild-type and \textit{umuC122} strains. This remarkable and unexpected result led us to examine the cells microscopically during HU treatment (see below).

A strain bearing a \textit{mazEF} or \textit{relBE} mutation is also resistant to HU

Although wild-type and \textit{umuC122} strains display comparable levels of bulk DNA synthesis during HU treatment, only in the \textit{umuC122} mutant is this activity beneficial for survival. It seemed possible that the wild-type strain loses viability not directly due to stalled
replication forks that arise during HU treatment, but instead due to events that occur downstream of such stalled forks. Examination by microscopy of an HU-treated parental culture revealed drastically fewer cells (>90% reduction at 5 h) than in the *umuC122* strain, most likely due to cell lysis. Hence, we considered the phenomenon of thymineless death, which is also thought to be the product of stalled replication forks formed by substrate starvation (32). In *E. coli* strain MC4100, thymineless death is mediated at least in part by the *mazEF* genes (33), which encode a toxin-antitoxin pair. We speculated that HUR and thymineless death may proceed through similar mechanisms.

Therefore, we examined the sensitivity to HU of an MC4100 derivative harboring a deletion of the *mazEF* genes (34). Not only does deletion of these genes protect cells from the lethal consequences of HU challenge (Figure 6A), but the mechanism of HUR is also likely to be related to that of the *umuC122* strain. Microscopical examination during HU treatment indicates that *umuC122* and *mazEF* strains appear quite similar at the single cell level (Figure 6B). No morphological difference is visible among the strains 1 h into HU treatment (panels A-D), but each HU<sup>S</sup>-parental strain had to be concentrated an additional 5-fold to analyze comparable numbers of cells relative to its HU<sup>R</sup> derivative. Finally, at 5 h (panels E-H), we observed similar responses in both HU<sup>S</sup>-parental strains (concentrated 15-fold relative to their HU<sup>R</sup> derivatives). In comparison, the *umuC122* and *mazEF* strains show extreme elongation and no dead cells, suggesting that HU<sup>R</sup> may arise through a similar mechanism in both strains. Therefore, it is plausible that the HU<sup>R</sup> phenotype of the *umuC122* mutant may be due to a failure to transduce a signal in a *mazEF*-dependent pathway leading to cell death and lysis (34).

We then tested whether a different TA pair may protect cells from the lethality caused by thymine starvation or HU challenge. Thus, we constructed P90C derivatives that harbored
deletions of either the mazEF or relBE genes (35). We also transduced the umuC122 allele into the E. coli strain HM21, the donor of the mazEF and relBE deletion alleles. We tested the mazEF and relBE strains in both backgrounds for HU^R and response to thymine starvation using trimethoprim (Tp) to inhibit thymidilate synthase (thyA). We found that the relBE deletion protects cells from inhibition of fork progression upon thymine starvation similarly to mazEF (Figure 6C), and that both strains showed comparable responses upon HU challenge (Figure 6D). Moreover, we found the umuC122 allele confers resistance to both Tp (Figure 6C) and HU (Figure 6D) in the HM21 strain background, though this HU^R is of a lower magnitude than that observed in the P90C strain. In contrast, the umuC122 allele does not confer resistance to Tp in the P90C background. This results suggest that there may be communication between pathways that couple HU- and Tp-induced stalled replication forks to cell death, and that a factor(s) involved in such communication is/are absent in the P90C strain, which bears a ~105 kb deletion on its chromosome. Moreover, both pathways appear to utilize the relBE and mazEF TA pairs as their ultimate executioners.

**HU-treated strains bearing the truncated UmuC protein have a high mutation frequency**

Our findings raise the possibility that the four proteins we have identified as being critical for HU^R - DinB, a UmuC derivative, UmuD, and UmuD' - enhance cell survival under conditions of low dNTP concentrations. They may even take over much of DNA replication, thereby helping cells to replicate even in the presence of HU (Figure 1A). If DNA replication upon HU challenge is DinB- and UmuC-dependent, one would expect such DNA synthesis on undamaged DNA to be less accurate than that carried out by the DNA Pol III holoenzyme. Therefore, we tested whether the mutation frequency to rifampicin resistance is changed before or after HU
treatment in a *umuC122* strain. We determined that untreated strains encoding UmuC122 protein have a spontaneous mutation frequency of $4\pm2\times10^{-7}$, identical to the mutation frequency of the untreated *umuC*+ parental strain ($4\pm3\times10^{-7}$). However, after HU treatment, the mutation frequency of the *umuC122* strain increases *ca.* 100-fold to $7\pm3\times10^{-5}$ while the mutation frequency of the *umuC*+ parental strain remains at *ca.* $10^{-7}$. These data suggest that it may be possible to explain the HU<sup>R</sup> phenotype of strains bearing the *umuC122* allele by a model in which one or both of the Y-family polymerases are responsible for a significantly greater proportion of DNA replication during HU treatment than under normal conditions.
Discussion

We examined the effect of inhibiting replication fork progression in a DNA damage-independent manner with HU in strains bearing different alleles of the umuC gene and found that cells bearing a carboxy-terminal truncation allele umuC122::Tn5 (21) are strikingly resistant to HU treatment (Figure 1A). Moreover, an unusual point mutation in UmuC (umuC125 allele, A39V) (27) displays a similar phenotype (Figures 3A and 3D). We have shown that umuC122 is a gain-of-function allele that mediates HU\textsuperscript{R} and encodes a gene product that could, in principle, perform DNA polymerization since its polymerase domain is intact (24). DNA polymerase activity in such a mutant protein is not unprecedented since truncations of the carboxy terminal domain of human Y-family polymerase η are TLS proficient \textit{in vitro} (36). XP-V patients (37) bearing these C-terminal truncations tend to have more tumors than those carrying other Pol η alleles (36). Indeed, we show that cells expressing a catalytically inactive UmuC122 protein are sensitive to HU (Figure 3A). We have also shown that the DinB protein (Figure 4A), and its catalytic activity (Figure 4B), is needed to observe the phenotype. In addition, we have learned that certain umuD gene products are required for the HU\textsuperscript{R} phenotype (Figures 3B and 3C).

Analysis of HU-treated cultures by microscopy (Figure 6B) revealed not only that the HU-treated wild-type cells die, but that many also disappear over the course of treatment, presumably through cell lysis. These data challenged our expectation that stalled replication forks would simply arrest cell division and prevent colony formation. We had not anticipated that they would bring about cell lysis in and of themselves.

We have shown that cells treated with HU are affected in a process downstream of RNR inhibition (Figures 2B and 2C). The current model for replication stalling elicited by dNTP
depletion is that substrate starvation brings about fork arrest and concomitant cell death (38). However, HU-treated *Saccharomyces cerevisiae* cells have been shown to exhibit both normal replication forks that can still sustain very slow DNA synthesis, as well as stalled replication forks (18). Moreover, HU-treated *S. cerevisiae* show a reduction in levels, but not an absence, of dNTPs (39). Hence, the dNTP starvation model may be too simplistic to account for all these observations.

Therefore, we considered whether the HUR mediated by these gain-of-function alleles of *umuC* is due to an abrogation in a pathway that would normally lead to cell death under conditions of dNTP starvation. We found that *E. coli* strains bearing a deletion of such a function (*mazEF::Kan*) (34) are also HUR (Figures 6A and 6B). We also found that deletion of *relBE* protects cells from both thymine starvation and HU challenge (Figures 6C and 6D). It is likely that the function of the *mazEF* and *relBE* gene products is to slow metabolism, thereby enabling stasis and resumption of balanced growth (35). However, when challenged with dNTP starvation, cells are unable to recover from this stasis and eventually perish. Based on these data, HU-induced death of *E. coli* may be brought about not by stalled replication forks directly, but rather through a series of downstream processes involving the TA pairs *mazEF* and *relBE*. The UmuC variants, acting in combination with the *dinB* and *umuD* gene products, may mitigate such *mazEF* or *relBE*-induced death, either directly or indirectly. Further studies will be needed to establish whether and to what extent replication fork collapse is required to signal such lethal pathways, as well as other factors that might be involved. It will be interesting to look for a function that would bestow TpR in the P90C *umuC122* derivative (Figure 6D). This strain harbors a large deletion (Δ(lac-pro), ca. 105Kb) compared to the HM21 background, where the *umuC122* derivative is TpR (Figure 6C).
In *E. coli*, intracellular dNTP pools are at least 10-fold lower (10 μM) in the presence of HU than in untreated cells (100 μM) (40). One explanation for HU-induced stalled replication forks is that the replicative DNA polymerase cannot catalyze efficient DNA synthesis as its $K_m$ for dNTPs (3-40 μM for DNA Pol III) (41) is higher than the concentrations of dNTPs present in the HU-treated cells. In comparison, the $K_m$ for dNTPs of Pol IV (0.12 μM for His-DinB with the processivity clamp) and Pol V (0.08 μM with RecA versus 1200 μM without) are much lower (42). Therefore, it appears the *E. coli* Y-family DNA polymerases have the potential to operate efficiently at low dNTP concentrations, conditions at which DNA Pol III would operate poorly. Furthermore, such capabilities seem to be dramatically regulated through protein-protein interactions.

All these data are consistent with the notion that DinB, UmuC, and the *umuD* gene products are recruited to stalled replication forks upon HU treatment. We propose that the UmuC derivatives alter the highly dynamic process of polymerase switching, so that Y-family polymerases are defective in the switch back to the replicative polymerase. Ordinarily, UmuC, UmuD, and DinB would be part of a transient complex relieving arrested replication forks, regardless of how they arise. Both Y-family polymerases would work together to enhance cell survival, perhaps with DinB extending primers that are misaligned on their templates (43) and UmuC continuing replication before hand off of the primer terminus to the replicative DNA polymerase. Such polymerase switching is regulated by numerous factors in *E. coli* including the *umuD* gene products (44). In contrast, the UmuC variants would be recruited to HU-induced stalled forks and would be proficient to catalyze DNA synthesis, but would be unable to sense the signal to hand off the primer terminus to the replicative DNA polymerase. Hence, these UmuC derivatives would retain access to the replication fork unlike the wild-type protein. The
unexpected finding (Figure 5A) that wild type cells still carry out DNA replication upon HU challenge may be explained by a futile cycling of Y-family polymerase recruitment and subsequent handoff to the replicative DNA polymerase which cannot function effectively at the low dNTP levels of the cell. Furthermore, although *umuC122* is non-mutable *in vivo* with respect to UV, its gene product may be able to catalyze DNA polymerization on undamaged templates. Under normal circumstances, such prolonged access to the fork would be detrimental, but during the unique stress of HU treatment (low dNTPs) it is advantageous for survival, albeit at a mutagenic penalty.

Why does this apparent failure to hand off to the replicative polymerase in the *umuC* mutants prevent HU-induced death? Although it is possible that UmuC communicates directly with either or both of the *mazEF* and *relBE* gene products, thereby signaling cell death in response to stalled replication forks, it is perhaps more likely that the prolonged action of the UmuC derivatives at the replication fork prevents the generation of an intermediate that would lead to the *mazEF*- and *relBE*-dependent process of cell death and lysis. We suggest a factor that responds to one of these intermediates that is specific to thymineless death is missing in the P90C strain, explaining why the *umuC122* derivative behaves as the wild type upon Tp challenge. The carboxy-terminus of UmuC harbors interaction sites for both UmuD$_2$ and UmuD’$_2$ (13), which are absent in the UmuC122 protein. Perhaps the lack of this domain alters the ability of the UmuC122 protein to return the primer terminus to the replicative DNA polymerase. Moreover, the data in figures 3B and 3C highlight the role of UmuD cleavage in HU$^R$. Alternatively, the truncated UmuC122 protein may remain at the replication fork due to altered interaction with the β subunit of Pol III since deletion of its C-terminus may modify the accessibility of its β-binding-motif (residues 357-361) (45). It is clear that *umuC122* and Δ*umuC*
are both loss of function alleles for UV- and chemical-induced mutagenesis in exponentially growing cells. However, phenomena tested using *umuC122* should be reevaluated. In comparison, the A39V mutation in the UmuC125 protein is in close proximity to the active site (ca. 6Å, Figure 3D). The phenotype conferred by the *umuC125* allele may be due to either disruption of regulatory protein-protein interactions with similar consequences to the *umuC122* mutation, or to alteration of the biochemical properties of the protein, such as a reduction in k\textsubscript{off} for the primer/template, K\textsubscript{m} for dNTP substrates, or both. In either case, the consequence is prolonged access to the replication fork under conditions of nucleotide starvation, resulting in survival during HU challenge.

If these polymerases replicate DNA in the presence of HU, mutability should be markedly higher in the mutant strains relative to the wild-type. Indeed, the *umuC122* bearing strain displays a 100-fold higher mutation frequency upon HU treatment than its untreated counterpart or the wild type strain. Intriguingly, prior to the discovery of Y-family polymerases, it has been reported that imbalances in dNTP pools increase mutagenesis, perhaps by decreasing the fidelity of DNA synthesis (46). This reduction in fidelity could perhaps now be attributed to the recruitment of such Y-family polymerases to the replication forks under conditions of nucleotide imbalance.
Materials and Methods

Strains and plasmids

We used different *Escherichia coli* K12 strains and their isogenic derivatives (Table I): P90C (47), AB1157 (Bachmann, 1987) and HM21 (48). A precise deletion of *dinB* was constructed using the method described by Wanner et al. (49) with primers FW2 (5'-acgcttaatgctgaatctttacgcatttctcaaacc-3') and RW2 (5'gtgatattgaccgatttttcagcgagaattcgatgcat3'). The deletion was transduced by P1 (50) into the appropriate strains from BW25113 (49). P1 transduction was also used to transfer the *umuC122* allele (21), a deletion of the *umuDC* operon (51), and a precise deletion of *umuC*. Wild-type and *umuC122 thyA* derivatives were constructed by P1 transduction from the strain EGSC#6827. The *dinB003* allele (43) was constructed on the chromosome of BW25113 using the plasmid-borne allele as a template. The *umuDC*-containing plasmids are derivatives of pGB2 (13). The non-cleavable *UmuD(S60A)* allele (27) was introduced by site directed mutagenesis using a Quikchange kit (Stratagene, La Jolla) with the following oligonucleotide: (5'-gcaagtgggtcatgatggctgtggatgggtgg-3') and its reverse complement. The *umuC122* allele was reconstructed in the same plasmid system using the primer (5'-cactcgagaggagagttacagatggcatctggtc-3') and its reverse complement.

Strains were grown routinely in liquid or solid media (LB) or in minimal M9 medium with the addition of hydroxyurea (HU; 30-100mM), ampicillin (Amp; 100 µg/ml), spectinomycin (Sp; 60µg/ml), chloramphenicol (Cm; 10-20 µg/ml), kanamycin (Km; 50 µg/mL), rifampicin (Rif; 100 µg/mL), trimethoprim (Tp; 3-7 µg/ml), diaminopimelic acid (DAP; 30 µg/ml) and thymine (50 µg/ml) whenever required. The *dinB*⁺ locus was reconstructed on the chromosome using the same approach as the *dinB003* construction in the Cm⁵ derivative of
The locus was transduced with P1 phage, and presence of the full-length dinB+ gene was verified by PCR with the primers dinBF (5'-atgcgtaaaatcattcatgtgga-3') and dinBR (5'-tcataatcccagcaccagggt-3').

**Hydroxyurea treatment**

Cultures were routinely treated in LB broth containing HU (Calbiochem) by diluting saturated cultures 1:1000. Treatment of ca. 10^6 bacteria/mL was for 6 h or as noted in the text or figure legends. Viability was checked throughout treatment. For anaerobic treatment with HU, cultures were treated as above for 6 h with 55 mM HU in an anaerobic chamber (Coy Laboratory Products) with a mixture of 5% Carbon Dioxide, 10% Hydrogen and 85% Nitrogen. Samples for western blotting were either TCA precipitated (20%) or concentrated 100-fold. The αUmuC antibody was used at a dilution of 1:20,000. The secondary antibody dilution and further detection were performed following manufacturer’s instructions (Pierce Biotechnology).

For the thymidine incorporation during HU treatment (100mM) we used a 1:1 mixture of M9 medium (50) with 0.3% casein to LB with 10 μg/ml of thymidine. The ^3^H-thymidine (PerkinElmer) incorporation was carried out in 10 min pulses after which the sample was immediately TCA precipitated (10% final).

**Acknowledgements:** We would like to thank Sue Lovett (Brandeis) for providing us with the ΔumuC mutant strain, Hannah Engelberg-Kulka (Hebrew University) for the mazEF strains, Kim Lewis (Northeastern University) for the HM21 mazEF and relBE strains, and the E. coli genetic stock center for the #6827 strain. We also thank JoAnne Stubbe (MIT) for the antibodies to RNR subunits, Alan Grossman (MIT) for use of the microscope, and Michael Malamy (Tufts Medical
School) for use of the anaerobic chamber. VGG discovered the resistance of the *umuC122* strain to HU and generated the bacterial strains used in this study. DFJ performed the immunoblots, determined the contribution of the *umuD* gene products, and generated the plasmids used in this study. VGG and DFJ performed the replication assays and wrote the manuscript along with GCW. LAS performed the microscopy. LAS was supported in part by a postdoctoral fellowship from NCI. This work was supported with the NIH Grant No. CA21615-27. GCW is an American Cancer Society Research Professor.
References


217
>(000
m~ x
S0)
00

N)

-- x

- -x

-A

-m

N

w(0 (D0(D(w(0(0w
0000000

000

0000000

ul u

CD
z

: a

C)

N)

CN) N)aa
N)

>+N K0
C:)
+ C)00

N)N).•

> > >»

> > > >

>
'0
-a
-a
"oC
V V V 0 -l

-a "o
'0 "0 -0 -0
r3 V3 V V Q)

0000000ow-R

mClWf +(~
IIT- >

W

~II
Q~

_•

~

...

mlW
fwWWmlWWe

I

fl

-uU

0
i ci
c-it0 rA
c000 ie0
.-,0000000000008
-

CD
N L'CCC

-A
r
r

r

;k

c.

k cr

C
o

CDCZD:

DmCD

CD
CD

j

>

W5CD
-•. 0
CD0
o

.
C

0( 0

0(0

-

0-:

D

=3

C
--

g•0)-•

(hO

w

w

..,C DJ
D,2r
r
VN(a

0)

CD2

C
M ca0-<

A ca

Cm

DCCZ CC i3 2 M

-

C

2,

2

5 02 cDO

C0

:3 CD
CD

:

D D
0.)

.D
CDDC
(C)) =r
C
:3.0
-,:3 no N(0

~-

r-=

ca

)

--- C

3

ncD

0

Dý3ý

C,
0
LO
:3=3C0
Z(E

-- (CD

0

2LC
CD
mC

C

-C

Z

CD

zC

c
CD

,
5-,

C o =0
3
CD 0•
CD
C
00
CD
CD
CD, CD
CDO0CDCD0
D
=
CD
CD CD-")
CD

0

d 3 L

)a)Q

o---••
co - _. CD=

C0D..
0
Ca CD
CDDV
CD
CD :C

CDCDn0n

CD 0)~~
:3
-C
c
CD

+.ý )
-

CDCD
CCD CD -C/
CD
0CA -•

=r
CD

CD
Cn

----

C

CD30
n
CD CD

X C

.*.i. -'---

,

C

z

q

-:
)
0CrD
CD(
CDM.•C

2.•,•h

NC
0
:C CD CD0),

0~m
:3

cn

2r0c.
=5

z3

0CL 0
D W M

-

rCDC
FC,-o
.

CD
0 0 a-.
__ . CD
.
DD
<
•)"
-"
0•0
-•.
-

CDO
cA =3J(0t @0.

0.0

--1 3

DC C

N) N) N) =3*5*5"*--,
CD c =3
c- cN

.•r••l."'•

0,
,, ,
Ca-i

=0)

:3
c

tt••

0) C0):3:3
a)C
a
ca0-1
,

~(

5 aoccC

CD
> o

C

0) CD
m"m-.
I•r0)

0
0a-00-•

0)
:3
. u
0

D N):C

CDCDC

O

"Q r r)OC
CD

lb 2)w
.=3 Ma'

,

(a <:
DI(
, CD

0

CD (0W

M. M.

C7)0

?DO0i
z 1-(0 c/)

c

k5~.

fl)CD

CD

z

M/D

0)~

-l.

:CD

L

CD
CD

00-)

-bU

CDCA0

T T

mm
CDCD CD
0)

0-

0-,

cr cr
0 0

00

CD CD CD

S-nI-Is --

CD CD
i0

000000

EY E

00 0)
7 7:3

a-o

CD CD
0)

0 00
X x x

0~

CT,
0

-n
0
C1)
CD


Figure 1. Bacterial cells bearing the *umuC122* allele are HU^R_. (A) Survival timecourse in hydroxyurea reveals HU^R_ of a strain bearing the *umuC122* allele (open circles). In comparison, a strain bearing a Δ*umuC* allele (closed triangles) is slightly HU^R_, while both parental (closed circles) and Δ*umuDC* strains (open triangles) are sensitive to the reagent. CFUs were determined by serial dilution. Error bars represent the standard deviation determined from at least five samples. (B) Comparison of survival in HU of both AB1157 and P90C backgrounds. Error bars represent the standard deviation determined from at least five samples. (C) The truncated UmuC122 protein is expressed *in vivo* as determined by immunoblotting. Lane 1 shows a cell-free extract from a Δ*umuDC* strain with vector only (pGB2), lane 2 shows the same strain but instead bearing the plasmid pDC, and lanes 3 and 4 show two independent isolates of the same strain bearing pDC122. Plasmid-borne copies were used to facilitate detection of UmuC in the absence of SOS induction.
Figure 2. HUR proceeds through RNR inhibition. (A) Immunoblot of large and small subunits of RNR shows no difference in levels between wild-type (lanes 1 and 2) and *umuC122* (lanes 3 and 4) strains during HU treatment. Lanes 1 and 3 contain twice as much total protein as lanes 2 and 4 (3.25 μg of total protein). (B) *umuC122* also alleviates cell death during challenge with other RNR inhibitors. The left panel shows results of treatment with 100 mM guanazole, while untreated results are shown on the right. Lane 1 shows the parental P90C strain, lane 2 shows the Δ*umuDC* strain, lane 3 shows *umuC122*, and lane 4 shows the Δ*umuC* strain. (C) Class I RNR is sensitive to HU while class III RNR, used exclusively in anaerobic growth, is indifferent to the reagent. The parental (P90C) and *umuC122* strains were treated with HU for 6 h with (+O₂) and without (-O₂) oxygen. CFUs reported are the average of four samples and error bars represent the standard deviation as determined from these samples.
Figure 3. *umuC* requirements for observation of HU\(^R\). (A) In a \(\DeltaumuDC\) strain, addition of the plasmid-borne *umuC* alleles pDC122 and pDC125 confer HU\(^R\). pDC122 carries the *umuD\(^+\)* gene but *umuC* has a stop codon at residue 322, thus reconstructing the truncated allele present on the chromosome by virtue of the Tn5 insertion. pDC104 encodes UmuC(D101N), rendering UmuC catalytically inactive, while pDC122C104 encodes UmuC122(D101IN). pDC125 encodes UmuC(A39V), an allele that separates the UV-induced mutagenesis and cold sensitivity phenotypes of *umuC*. CFUs were determined by serial dilution, and treatment was carried out with Sp (for plasmid maintenance) and 50mM HU. Values reported are the average of three experiments and error bars represent the standard deviation obtained from those values. (B) The *umuD* gene products are also required for HU\(^R\). The resistance conferred by a plasmid-borne *umuC122* allele depends upon the *umuD* gene products. pD(S60A)C122 is as pDC122 but encodes a UmuD protein with a mutation (S60A) rendering the protein unable to undergo autoproteolysis to become UmuD\(\prime\). The plasmid pDC122 is as pDC122 but encodes only UmuD\(\prime\) instead of the full-length protein. Reported values are the average of three experiments and error bars represent the standard deviation as determined from those experiments. (C) The *umuD* gene products are also required for the HU\(^R\) conferred by *umuC125*. Plasmids and data analysis are as in (B). (D) A structural representation of the UmuC active site reveals the proximity of A39 to residues essential for catalysis (D6, D101). The template is shown in red, and the primer in green. Model is courtesy of Dr. D. Barksy (LLNL, Livermore, CA).
Figure 4. The dinB gene and its catalytic activity are necessary to avert HU lethality. (A) The HU<sup>R</sup> of a umuC122 dinB<sup>+</sup> strain (open circles) is eliminated by deletion of the dinB gene (closed triangles). In contrast, deletion of the dinB gene has only a mild effect on the parental strain (open triangles, closed circles). (B) Reconstruction of the dinB<sup>+</sup> locus on the chromosome restores HU<sup>R</sup> to the umuC122 ΔdinB strain. However, transduction of the dinB003 allele, which encodes a catalytically inactive DinB(D103N), does not restore HU<sup>R</sup>, indicating that the catalytic activity of DinB is required. Treatment was for 6 h with 100 mM HU in rich medium. umuC122 dinB<sup>+</sup> refers to the reconstructed wild type gene with a linked cat gene upstream the dinB promoter. Reported values are the average of three experiments and error bars represent one standard deviation.
Figure 5. DNA synthesis is slowed in both wild type and umuC122 strains. Thymidine-requiring derivatives of both strains were used for the experiments shown. $^3$H-thy was added at 1μCi/ml for 10 min at each time point shown, after which cells were immediately precipitated with 10% TCA. For both the wild-type shown in (A) (circles) and umuC122 shown in (B) (squares) strains, bulk DNA replication is slowed during hydroxyurea treatment. The straight line represents the background cpm. Error bars represent the standard deviation of three samples.
Figure 6. Survival phenotypes under dNTP starvation. (A) Survival timecourse in 100 mM HU of the parental MC4100 derivative (closed circles) and the mazEF mutant strain (open circles) in LB. Error bars shown represent the standard deviation of two samples. (B) Strains bearing the indicated alleles and wild type control backgrounds were treated with 100 mM HU to determine cell morphology under HU treatment. Micrographs are presented for treated cells only because untreated samples of each strain showed indistinguishable morphologies over 5 h without HU. Panels A-D show representative images of cells after 1 h of HU treatment. (A) P90C wild type and umuC122 control DIC image, (C, D) MC4100 wild type and mazEF control DIC image. Images labeled 1 show DAPI staining, images labeled 2 show DEAD staining, and images labeled 3 show LIVE staining. Panels E-H are corresponding representative images of cells following treatment with HU for 5 h. Images were colorized using OpenLab software (Improvision) and were sized in Canvas (Deneba Systems). The white bar in (A) represents 2 μm. Exposure times for the images were as follows: DIC, 0.03s; DAPI, 0.13s; LIVE, 0.01s; DEAD, 0.13s. The LIVE/DEAD stain was used according to the manufacturer’s recommendations (Molecular Probes). (C) Mutation in the relBE gene-products protects cells from thymine starvation. Error bars represent the standard deviation of three samples. (D)
Deletion of the *relBE* genes also promotes HU<sup>R</sup>. Treatment with Tp in HM21 (4 µg/ml) and P90C (7 µg/ml) strains was performed in M9 minimal medium. CFUs were determined after 16 h incubation. HU challenge (100mM) was carried out in LB medium with DAP (30 µg/ml) in HM21. Error bars shown represent the standard deviation of three samples.
Chapter 7

*umuC* Mediates Diverse Physiological Outcomes During the SOS Response in *Escherichia coli*

DF Jarosz, MA Kohanski, LA Simmons, JJ Collins, and GC Walker. *In preparation.*
Abstract

The SOS response to DNA damage in *Escherichia coli* is mediated by the LexA transcriptional repressor. Its DNA damage-dependent interaction with the RecA nucleoprotein filament stimulates a latent autoproteolytic activity leading to its inactivation, thereby allowing the transcription of more than forty genes. Among these genes are *umuDC*, which encode the well-characterized alternative DNA polymerase UmuD'C and promote replication pausing in overproduction. Here we report that the *umuC* gene is also required for full induction of many SOS-regulated genes. Moreover we show that *umuC* mediates diverse physiological changes during the SOS response including cell division, motility, and resistance to acid stress. Taken together, these findings challenge the notion that UmuC acts exclusively at the terminus of the SOS response and suggest that, at least under conditions of SOS induction, the biological functions of the *umuC* gene products may be far more complex.
Introduction

In both prokaryotes and eukaryotes coordinated responses to genotoxic stress include induction of various genes involved in DNA repair, mutagenesis, and cell cycle control (1). The SOS response to DNA damage influences the expression of more than forty genes in *Escherichia coli* (2-5), whose gene products function in diverse aspects of the DNA damage response including DNA repair, transcriptional regulation, nucleoside metabolism, and cell division (5). In addition, the expression of numerous genes of unknown function is influenced by the SOS response, indicating that many aspects of this well-studied regulatory network have yet to be elucidated.

The SOS regulatory network is controlled by the LexA transcriptional repressor, which is inactivated in response to DNA damage (1). RecA plays a central role in controlling the SOS response in addition to its function in homologous recombination. Upon DNA damage RecA forms a nucleoprotein filament with ssDNA, which interacts with LexA and facilitates its inactivation via autoproteolysis (1). In addition to DNA damage, β-lactam antibiotics can induce the SOS response in *Escherichia coli* through the two component system *dpiAB* (6). This phenomenon has been interpreted as an effort to promote survival by inhibiting cell division. The β-lactam antibiotic ceftazidime has also been shown to induce at least one SOS-regulated gene, *dinB*, in a manner that is both *lexA* and *recA* independent (7).

In addition to its pivotal role in influencing DNA metabolism, the SOS response also regulates cell division. Expression of SulA, which inhibits FtsZ ring formation and hence cell division (8, 9), is controlled by LexA as part of the SOS regulatory network (10). Indeed, SulA has been suggested to function as part of a primitive DNA damage checkpoint in *Escherichia coli* (11-13). SulA dependent inhibition of cell division by DNA damage induces filamentation
of *Escherichia coli* cells, and constitutively active mutants of *lexA* must be constructed in a *sulA* background to allow for cell division (1). However, SulA-independent inhibition of cell division has also been observed under a variety of conditions including DNA damage (14), temperature stress (12), and high pressure growth (15). These observations suggest that the molecular connections between cell division and DNA replication are likely remarkably complex.

SulA independent filamentation induced by temperature stress depends on the SOS regulated *umuDC* gene products (12). During the SOS response, the effector protein UmuD initially exists as a homodimer (16), which promotes replication pausing in conjunction with UmuC (17, 18). UmuD₂ undergoes RecA::ssDNA nucleoprotein filament mediated autoproteolysis to remove its N-terminal 24 amino acids forming UmuD'₂ (19, 20), which associates with UmuC to form DNA polymerase V and is active in UV-induced mutagenesis (20). The most thermodynamically stable form of the *umuD* gene products (21), UmuDD', targets the molecule for proteolytic degradation by ClpXP (22, 23). The genetic underpinnings of other types of SulA independent filmentation are poorly understood, but are generally considered to be *lexA* dependent (14).

The SOS response and its physiological consequences are regulated at numerous levels in *Escherichia coli*. Here we report that *sulA* independent filamentation in a *lexA* deficient strain requires the function of the *umuC* gene product. Most unexpectedly, *umuC* appears to influence the induction of the SOS response. We also demonstrate that *umuC* deletion results in striking panel of phenotypes under conditions of SOS induction, including resistance to several agents that perturb DNA replication. Taken together, these observations point to a central role for *umuC* in coordinating the SOS response in *Escherichia coli*. 
Results

SOS-dependent colony and cellular morphologies are \textit{umuC}\textsuperscript{+}-dependent

When constructing a \textit{umuC} mutant in a \textit{lexA} deficient background, we no longer observed the small colonies characteristic of its constitutively SOS-induced parent strain (Figure 1A). This phenomenon requires constitutive SOS induction, as no alteration in colony morphology was observed when we constructed a \textit{umuC} mutant a \textit{lexA}\textsuperscript{+} otherwise isogenic background (Figure 1A). This difference in colony morphologies is not merely a result of differing growth rates. The doubling time of the \textit{lexA umuC} double mutant in rich medium at 37\textdegree C is slightly greater than its \textit{lexA} deficient parent (data not shown).

The altered colony morphology of the \textit{umuC} mutant is also reflected on the microscopic level. The SulA independent filamentation observed in a \textit{lexA} deficient background is suppressed by deletion of \textit{umuC} (Figure 1B). Strikingly, the length distribution of the \textit{lexA umuC} double mutant more closely resembles a \textit{lexA}\textsuperscript{+} strain than its \textit{lexA} deficient parent (Figure 1C). The effects of \textit{umuC} on both colony and cellular morphology are dependent on \textit{lexA}-, but not \textit{sulA}-, deficiency (data not shown). Surprisingly, deletion of both \textit{umuD} and \textit{umuC} does not affect colony or cellular morphology in a \textit{lexA}\textsuperscript{+} or \textit{lexA} deficient background (Figure 6). This unexpected observation led us to examine the extent of SOS induction in the \textit{lexA} deficient \textit{AumuC} mutant.

Full activation of the SOS response requires \textit{umuC} function

To ascertain the basis for the lack of filamentation in the \textit{lexA umuC} double mutant we examined the extent of SOS induction in that strain relative to its parent microscopically using a
reporter plasmid bearing green fluorescent protein (GFP) under the control of a synthetic strongly LexA-repressed promoter (24). In wild-type E. coli we observe a small subpopulation of cells that are SOS-induced, consistent with previous studies (25), whereas virtually all cells show GFP fluorescence in the lexA deficient strain (Figure 2A). Strikingly, deletion of umuC drastically reduces the percentage of cells that are SOS-induced (Fig. 4). This entirely unexpected result suggests that umuC may positively regulate SOS controlled genes.

This observation led us to examine the effect of umuC on the transcription of SOS regulated genes in more detail using the well-characterized series of din::Mud(Ap^R, lac) fusions (26). Two umuC deletion alleles, one marked with chloramphenicol resistance and the other with kanamycin resistance, were transduced into lexA deficient [lexA51(Def)] and lexA non-inducible (lexA3; constitutively non-SOS-induced) strains bearing Mud(Ap^R, lac) fusions to dinA (polB), dinB, dinD, and dinF. The β-galactosidase activity of each strain was measured to determine the steady state transcript levels of each fusion. Consistent with our previous observations, deletion of umuC in these strains suppresses the small colony morphology of these lexA deficient strains and drastically reduces the amount of β-galactosidase activity they manifest (Figure 2B-C). Taken together, these observations strongly indicate that umuC mediates the full induction of the SOS response.

umuC influences expression of genes involved in diverse cellular processes

To further elucidate the role of umuC in sculpting physiology during the SOS response, we examined the global effects of umuC on transcription using microarrays. To our surprise, we found that the transcript levels of numerous genes are controlled by umuC under conditions of SOS induction (Figure 3). As expected, the SOS response is profoundly downregulated in the
*lexA umuC* double mutant. Nearly 22% of the genes whose transcription is controlled by LexA (5), particularly those that are most strongly induced by the SOS response, are downregulated in the double mutant strain. However, many SOS regulated genes are also either unaffected (72%) or even upregulated (6%) upon deletion of *umuC*. Marked upregulation of the cell division machinery, the acid stress response, protein synthesis, and quorum sensing is also observed along with downregulation of motility (Figure 3).

To confirm the unexpected results of the microarray, we examined the motility and acid stress response of the *lexA umuC* double mutant. As predicted by the microarray, the *umuC* mutant strain exhibits a profound motility defect and has considerably less flagellin relative to its *lexA* deficient parent (Figure 4A). Moreover, the *umuC* mutant shows remarkably increased resistance to low pH, consistent with the elevated levels of *gadA* and *gadB* transcripts in the mutant (Figure 4B). Taken together, these observations indicate that the collection of transcriptional changes revealed in the microarray experiment are reflected at the physiological level in the *lexA umuC* double mutant strain.

**Inactivation of umuC promotes survival in response to agents that perturb DNA replication**

We exposed the *lexA umuC* double mutant strain to numerous DNA damaging agents in an effort to understand its physiology in greater detail. The strain is insensitive or only modestly sensitive to ultraviolet light, ionizing radiation, and nitrofurazone (data not shown). Although deletion of *umuC* does not affect the behavior of a wild-type strain when challenged with hydroxyurea (27) (Figure 5A), we observed striking resistance to that agent in the *lexA umuC* double mutant (Figure 5A). This hydroxyurea resistance arises due to induction of the SOS
response and not sulA deficiency as sulA single mutants are only modestly resistant to hydroxyurea (Figure 5D).

The resistance of the lexA umuC double mutant to hydroxyurea, an agent that perturbs DNA replication, motivated us to examine its behavior when challenged with another agent that elicits a similar response. Novobiocin inhibits DNA gyrase (28) thereby similarly perturbing DNA replication without directly damaging DNA. Remarkably, the lexA umuC mutant shows substantial resistance to novobiocin relative to its lexA deficient parent (Figure 5C). Notably, this resistance again requires that the strain be lexA deficient, as deletion of umuC has no effect on the novobiocin resistance of a lexA + strain (Figure 5C).

Different forms of the umuD gene products mediate these phenomena

We considered the possibility the umuD gene products, which regulate many functions of umuC, might also regulate these phenomena. To uncover the functions of the umuD gene products in both altered morphology and hydroxyurea resistance, we transformed plasmid-borne copies of UmuD, UmuD′, and a noncleavable UmuD variant (UmuD^{S60A}) into a lexA deficient ΔumuDC strain. The lexA deficient ΔumuDC strain does show slight but reproducible sensitivity to hydroxyurea (Figure 6A-B). These observations indicate that the umuD gene products potentiate the altered morphology and hydroxyurea resistance of the umuC mutant. Transformation with either UmuD or UmuD′ results in small colonies similar to the untransformed lexA ΔumuDC strain (Figure 6B). In contrast, transformation with the noncleavable UmuD^{S60A} variant results in large colonies similar to those of a lexA umuC double mutant (Figure 6B). Taken together, these observations indicate that the altered colony morphology of the lexA umuC double mutant requires the function of full-length UmuD rather
than UmuD'. On the cellular level, transformation of a \textit{lexA AumuDC} strain with UmuD does not alter the size distribution of the population (Figure 6A). Transformation with UmuD', however, results in substantially smaller cells (Figure 6A). Intriguingly, transformation with the UmuD\textsuperscript{S60A} variant results in a predominance of extremely elongated cells (Figure 6A). Hydroxyurea resistance can be achieved with either UmuD or UmuD', whereas the UmuD(S60A) variant exacerbates sensitivity to this agent (Figure 6C). Taken together, these data indicate that the different \textit{umuD} gene products influence these phenomena in multiple ways.
Discussion

Even in a simple organism such as *E. coli*, responses to DNA damage are remarkably complex. Some studies estimate that up to 1,000 genes, or 25% of the *E. coli* genome, are regulated by DNA damage (29). One of the best characterized responses to DNA damage in *E. coli* and numerous other bacteria, the SOS response, influences the expression of many genes including *umuD* and *umuC*. Several aspects of the SOS response remain to be elucidated, however. For example, it induces numerous genes of unknown function (5) and the response displays a largely unexplained oscillatory character that appears to depend upon the *umuDC* operon (30).

We find that deletion of *umuC* in a constitutively SOS induced background results in a panel of striking physiological abnormalities. *E. coli* strains deficient in both *lexA* and *umuC* display altered colony and cellular morphology that is dependent on different forms of the *umuD* gene products. UmuD' appears to promote reduced cell length in liquid medium, whereas full-length UmuD appears to promote enlarged colony size on solid medium. The apparent discrepancy between the enlarged colony morphology of the *lexA umuC* strain and its reduced cell length in liquid medium may reflect a corresponding difference in the status of UmuD cleavage between those growth conditions.

A *umuC* gene product also appears to potentiate *lexA*-dependent cell death in response to treatment with replication inhibitors. Deletion of *umuC* in a *lexA* deficient background provides striking resistance to the mechanistically distinct replication inhibitors hydroxyurea and novobiocin. There appears to be something unique about the type of replication stress exerted by these agents, as the *lexA umuC* strain is largely unaffected by diverse DNA damaging agents that
covalently modify DNA. We have previously characterized the role of toxin-antitoxin pairs, particularly *mazEF* and *relBE*, in promoting hydroxyurea-induced cell death (27). The transcript levels of both *relBE* and *mazEF* are essentially unchanged in our microarray experiment. These observations suggest that the hydroxyurea resistance we observe here likely arises via a different mechanism. Resistance to novobiocin in the *lexA umuC* strain may be explained at least in part by upregulation of cardiolipin synthase, as *cls* mutants are known to be sensitive to novobiocin (31). The molecular underpinnings of this *umuC* dependent HU resistance remain to be elucidated, however.

Surprisingly, a *umuC* gene product also appears to potentiate the full expression of certain, but not all, SOS-regulated genes. Indeed, some SOS-regulated genes are even induced in the *umuC* mutant. There is no clear pattern among those genes that are positively regulated by *umuC*, although they appear to be rather strongly regulated by LexA. Curiously, a *umuD* gene product appears to negatively regulate SOS induced genes in the absence of *umuC*, as the robust SOS induction in a *lexA umuDC* strain is far more similar to its *lexA* deficient parent than to the weak SOS induction in a *lexA umuC* strain. The mechanism of such regulation is not readily apparent. The *umuD* gene products interact with a host of cellular factors (32) by virtue of their remarkable biophysical properties (Simon and Walker, in preparation), but none of these are known transcriptional regulators.

We also find that *umuC* influences the transcription of numerous genes in addition to those in the SOS regulatory network. Flagellar genes as well as those involved in acid stress resistance are differentially regulated in the *lexA umuC* mutant. Indeed, numerous other genes including those involved in cell-cell communication, glycerol metabolism, and many of unknown function, are also differentially regulated in the double mutant strain. Consistent with
these observations, flagellar motility and acid resistance are dramatically altered in the mutant. All of these data are consistent with the notion that umuC and umuD perform critical functions in modulating cellular physiology during the SOS.

The umuDC genes play critical roles both in promoting replication pausing and translesion DNA synthesis after DNA damage (1, 33). Moreover, posttranslational modification of UmuD to form UmuD' provides a temporal switch during the SOS response separating accurate DNA repair and damage tolerance mechanisms from more error prone damage tolerance mechanisms. Deletion of umuC in a wild-type background has relatively minor consequences in the absence of exogenous treatment. In contrast, we observe that its deletion has profound consequences under conditions of SOS induction. Taken together, our findings suggest that during the SOS response umuC function mediates diverse physiological responses including filamentation, motility, cell death in response to replication inhibitors, acid resistance, and even the SOS response itself. In the absence of SOS induction, however, loss of umuC function does not appear to mediate such a wide array of physiological responses. The findings presented here, in conjunction with the role of a gain-of-function umuC allele in promoting hydroxyurea resistance (27), suggest that the biology of this Y-family DNA polymerase extends beyond simple TLS. Dissecting the genetic pathways involved in these phenomena will offer considerable insight into mechanisms governing responses to environmental stress.


Materials and Methods

Strains and plasmids

The strains and plasmids used in this study are described in Table 1. Strains were commonly grown in LB medium at 37°C, with the exception of the din::Mud(ApR, lac) strains, which were grown at 30°C. β-galactosidase assays were performed according to Miller, except that 1X M9 salts were used for an assay buffer. Conditions employed for fluorescence microscopy experiments are described in the figure legends.

Hydroxyurea and novobiocin treatment

Hydroxyurea treatment was performed by 1:1000 dilution of freshly-saturated cultures into LB containing hydroxyurea at the indicated concentration. Cultures were grown for 12 h and plated onto LB to determine remaining colony forming units. Percent survival is defined relative to the untreated culture. Novobiocin treatment was performed by plating serial dilutions of freshly saturated culture onto LB agar plates containing 400 μg/mL novobiocin.

Glutamate-dependent acid resistance assays

Glutamate-dependent acid resistance assays were performed essentially as described previously (34). Freshly saturated cultures were diluted 1:1000 into LB with 0.2% glucose at pH 5.5 and were pre-conditioned for 12 h at 37°C. These pre-conditioned cultures were diluted 1:1000 into minimal medium E salts at pH 2.5 with or without 1.5 mM glutamate. Treatment was performed for 4 h at 30°C, after which colony forming units were measured by serial dilution on LB agar plates.
Microarray experiments

RNA samples were prepared from freshly saturated cultures using an RNEasy kit (Qiagen) according to the manufacturer’s directions. Microarray data collection and analysis was performed as previously described (35).

Flagellin purification and motility assays

Flagellin was purified by vortexing freshly saturated cultures grown in LB (10 mL) at high speed for 45 s. Cells were pelleted by centrifugation at 5000 x g for 15 min and the supernatant was reserved. Ammonium sulfate was added to the supernatant to 50% saturation and the solutions were incubated on ice for 2 h followed by centrifugation at 25,000 x g for 90 min. The pellet formed was dissolved in 50 mM Hepes pH 7.5 with 100 mM KCl and analyzed by SDS-PAGE. Swarming assays were performed by spotting 5 μL of a freshly saturated culture onto 0.25% LB agar and measuring the swarm diameter at the time indicated.

Acknowledgements: We would like to thank S. Lovett (Brandeis) for the gift of the ΔumuC::Cm mutant and members of the Walker and Collins laboratories for thoughtful discussions. DFJ performed the physiological experiments, generated the strains used in this study, and wrote the manuscript along with GCW. MAK and JJC performed and analyzed the microarray experiments. This work was supported by NIH grants to GCW and JJC.
References


<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB1157</td>
<td>thr-1 araC-14 leuB-6(Am) Δ(gpt-proA)62 lacY1 tsx-33 qsr'·-0 glnV44(AS) galK2(Oc) LAM Rac-0 hisG4(Oc) rfbC1 mgl-51 rpoS396(Am) rpsL31 kdgK51 xylA5 mtl-1 argE3(Oc) thi-1</td>
<td>GCW lab stock</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but ΔumuC::Kan</td>
<td>This work.</td>
</tr>
<tr>
<td></td>
<td>ΔumuC::kan</td>
<td>Keio Collection.</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but ΔumuC::Cm</td>
<td>S. Lovett</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but ΔumuDC595::Cm</td>
<td>This work</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but ΔsulA::Tet</td>
<td>This work</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 but lexA300(Def) and ΔsulA::Tet</td>
<td>This work</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 lexA300(Def) but ΔumuC::Cm</td>
<td>This work.</td>
</tr>
<tr>
<td></td>
<td>ΔumuDC</td>
<td>S. Lovett</td>
</tr>
<tr>
<td>AB1157</td>
<td>As AB1157 lexA300(Def) but ΔumuDC595::Cm</td>
<td>GCW lab stock.</td>
</tr>
<tr>
<td>GW1000</td>
<td>F thr-1 leu-6 proA2 his-4 argE3 galK2 strA31 ilvR tif-1 sulA11 pro+ lacΔU169</td>
<td>GCW lab stock.</td>
</tr>
<tr>
<td>GW1012</td>
<td>As GW1000 but dinA::Mud(ApR, lac) lexA3(Ind-) malE::Tn5</td>
<td>Kenyon et al. 1980</td>
</tr>
<tr>
<td>GW1013</td>
<td>As GW1000 but dinA::Mud(ApR, lac) lexA51</td>
<td>Kenyon et al. 1980</td>
</tr>
<tr>
<td>GW1013</td>
<td>As GW1013 but ΔumuC::Cm</td>
<td>This work</td>
</tr>
<tr>
<td>GW1013</td>
<td>As GW1013 but ΔumuC::Kan</td>
<td>This work</td>
</tr>
<tr>
<td>GW1032</td>
<td>As GW1000 but dinB::Mud(ApR, lac) lexA3 malE::Tn5</td>
<td>Kenyon et al. 1980</td>
</tr>
<tr>
<td>GW1033</td>
<td>As GW1000 but dinB::Mud(ApR, lac) lexA51</td>
<td>Kenyon et al. 1980</td>
</tr>
<tr>
<td>GW1033</td>
<td>As GW1033 but ΔumuC::Cm</td>
<td>This work</td>
</tr>
<tr>
<td>Strain</td>
<td>Description</td>
<td>Reference/Source</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>GW1033</td>
<td>As GW1033 but <em>ΔumuC::Kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>ΔumuC::Kan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GW1042</td>
<td>As GW1000 but _dinD::Mud(Ap&lt;sup&gt;R&lt;/sup&gt;, lac) <em>lexA3 malE::Tn5</em></td>
<td>Kenyon <em>et al.</em> 1980</td>
</tr>
<tr>
<td>GW1043</td>
<td>As GW1000 but _dinD::Mud(Ap&lt;sup&gt;R&lt;/sup&gt;, lac) <em>lexA51</em></td>
<td>Kenyon <em>et al.</em> 1980</td>
</tr>
<tr>
<td>ΔumuC::Cm</td>
<td>As GW1043 but <em>ΔumuC::Cm</em></td>
<td>This work</td>
</tr>
<tr>
<td>ΔumuC::Kan</td>
<td>As GW1043 but <em>ΔumuC::Kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>GW1062</td>
<td>As GW1000 but _dinE::Mud(Ap&lt;sup&gt;R&lt;/sup&gt;, lac) <em>lexA3 malE::Tn5</em></td>
<td>Kenyon <em>et al.</em> 1980</td>
</tr>
<tr>
<td>GW1063</td>
<td>As GW1000 but _dinE::Mud(Ap&lt;sup&gt;R&lt;/sup&gt;, lac) <em>lexA51</em></td>
<td>Kenyon <em>et al.</em> 1980</td>
</tr>
<tr>
<td>ΔumuC::Cm</td>
<td>As GW1063 but <em>ΔumuC::Cm</em></td>
<td>This work</td>
</tr>
<tr>
<td>ΔumuC::Kan</td>
<td>As GW1063 but <em>ΔumuC::Kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>KM1190</td>
<td>As AB1157 but _lexA51(Def) _sulA11 ilv&lt;sup&gt;+&lt;/sup&gt; _his&lt;sup&gt;+&lt;/sup&gt; <em>srlC5</em></td>
<td>K. McEntee</td>
</tr>
<tr>
<td>ΔumuC::Cm</td>
<td>As KM1190 but <em>ΔumuC::Cm</em></td>
<td>This work</td>
</tr>
<tr>
<td>KM1190</td>
<td>As KM1190 but <em>ΔumuDC595::Cm</em></td>
<td>This work</td>
</tr>
<tr>
<td>ΔumuDC::Cm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGB2</td>
<td>Empty vector for plasmids expressing the <em>umuD</em> gene products (_pSC101-based; spec&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>R. Woodgate</td>
</tr>
<tr>
<td>pUmuD</td>
<td>As pGB2 with <em>umuD&lt;sup&gt;+&lt;/sup&gt;</em> in <em>hpal</em> fragment</td>
<td>M. Sutton 2001. Lab Stock</td>
</tr>
<tr>
<td>pUmuD'</td>
<td>As pGB2 with <em>umuD'&lt;sub&gt;+&lt;/sub&gt;</em> in <em>hpal</em> fragment</td>
<td>M. Sutton 2001. Lab stock</td>
</tr>
<tr>
<td>pUmuD(S60A)</td>
<td>As pUmuD with indicated mutation introduced by site-directed mutagenesis</td>
<td>This work</td>
</tr>
</tbody>
</table>

**Table 1.** Strains and plasmids used in this study.
**Figure 1.** *umuC* mediates the altered morphology of constitutively SOS-induced *E. coli* strains. 
A) The colony morphology of AB1157 is unaffected by deletion of *umuC*, but the small colonies characteristic of a *lexA*(Def) strain are suppressed by deletion of *umuC*. B) Cellular morphology is similarly affected by *umuC*. DIC micrographs indicated that the slight filamentation of a *lexA*(def) strain is suppressed by deletion of *umuC*. C) The cell lengths of a *lexA* *umuC* strain more closely resemble those of a *lexA*+ strain than its *lexA*(Def) parent.
Figure 2. *umuC*<sup>+</sup> mediates full induction of the SOS response in *E. coli*. A) Overlaid fluorescence micrographs of the indicated strains with a membrane stain (FM464; 50 ms exposure) and GFP fluorescence (300 ms exposure). The percentages indicate the number of green cells of the indicated numer scored. B) *dinD* fusion strains plated on LB with X-gal. GW1042 is *lexA*(Ind-); GW1043 is *lexA*(Def). Detailed strain genotypes are described in detail in Table 1. C) β-galactosidase activities of indicated strains bearing *dinA, dinB, dinD, and dinE* fusions.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold Change</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>recN</td>
<td>Recombination and repair</td>
<td>-22.1 ± 2.6</td>
<td>-5.3</td>
</tr>
<tr>
<td>dinD</td>
<td>Unknown</td>
<td>-13.6 ± 3.1</td>
<td>-6.3</td>
</tr>
<tr>
<td>dinI</td>
<td>Modulator of RecA function</td>
<td>-13.0 ± 1.9</td>
<td>-5.0</td>
</tr>
<tr>
<td>recA</td>
<td>Recombination and repair</td>
<td>-12.7 ± 2.0</td>
<td>-4.7</td>
</tr>
<tr>
<td>dinB</td>
<td>DNA polymerase IV</td>
<td>-2.0 ± 0.2</td>
<td>-3.1</td>
</tr>
<tr>
<td>uvrA</td>
<td>Nucleotide excision repair</td>
<td>-1.7 ± 0.2</td>
<td>-1.8</td>
</tr>
<tr>
<td>uvrB</td>
<td>Nucleotide excision repair</td>
<td>-2.6 ± 0.6</td>
<td>-3.4</td>
</tr>
<tr>
<td>polB</td>
<td>DNA polymerase II</td>
<td>-2.0 ± 0.3</td>
<td>-2.6</td>
</tr>
<tr>
<td>dinF</td>
<td>Unknown</td>
<td>+1.7 ± 0.1</td>
<td>+1.9</td>
</tr>
</tbody>
</table>

DNA Metabolism/Transcription Factors

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold Change</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>priB</td>
<td>Primosome</td>
<td>+3.6 ± 0.5</td>
<td>+3.5</td>
</tr>
<tr>
<td>dps</td>
<td>DNA protecting system</td>
<td>+3.3 ± 1.0</td>
<td>+2.3</td>
</tr>
<tr>
<td>ndk</td>
<td>Nucleoside diphosphate kinase</td>
<td>+2.9 ± 0.5</td>
<td>+2.7</td>
</tr>
<tr>
<td>rdgC</td>
<td>Modulator of RecA function</td>
<td>+1.7 ± 0.2</td>
<td>+4.4</td>
</tr>
<tr>
<td>fis</td>
<td>DNA-binding transcriptional regulator</td>
<td>+1.7 ± 0.2</td>
<td>+4.4</td>
</tr>
<tr>
<td>bipA</td>
<td>Global transcriptional regulator</td>
<td>+2.5 ± 1.2</td>
<td>+3.3</td>
</tr>
</tbody>
</table>

Cell Division Machinery

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold Change</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>minE</td>
<td>Cell division</td>
<td>+2.9 ± 0.4</td>
<td>+3.6</td>
</tr>
<tr>
<td>minC</td>
<td>Cell division inhibitor</td>
<td>+2.2 ± 0.2</td>
<td>+3.1</td>
</tr>
<tr>
<td>mreB</td>
<td>Cell wall; actin like</td>
<td>+1.8 ± 0.2</td>
<td>+2.6</td>
</tr>
<tr>
<td>ftsX</td>
<td>Cell division ATPase</td>
<td>+1.8 ± 0.3</td>
<td>+3.9</td>
</tr>
<tr>
<td>ftsY</td>
<td>Cell division</td>
<td>+1.7 ± 0.7</td>
<td>+2.7</td>
</tr>
</tbody>
</table>

Motility and Chemotaxis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold Change</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>fliC</td>
<td>Flagellin</td>
<td>-67.7 ± 1.2</td>
<td>-4.8</td>
</tr>
<tr>
<td>flgE</td>
<td>Flagellar hook protein</td>
<td>-34.2 ± 7.5</td>
<td>-4.8</td>
</tr>
<tr>
<td>flgD</td>
<td>Flagellar hook assembly protein</td>
<td>-23.3 ± 4.5</td>
<td>-4.8</td>
</tr>
<tr>
<td>flfD</td>
<td>Flagellar filament capping protein</td>
<td>-20.6 ± 2.1</td>
<td>-4.2</td>
</tr>
<tr>
<td>cheA</td>
<td>Chemotactic sensory histidine kinase</td>
<td>-12.6 ± 1.2</td>
<td>-4.0</td>
</tr>
<tr>
<td>cheW</td>
<td>Chemotaxis protein</td>
<td>-10.8 ± 0.8</td>
<td>-4.0</td>
</tr>
</tbody>
</table>

Metabolism and Acid Stress Response

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold Change</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>gadA</td>
<td>Glutamate decarboxylase; acid resistance</td>
<td>+10.6 ± 6.1</td>
<td>+2.3</td>
</tr>
<tr>
<td>gadB</td>
<td>Glutamate decarboxylase; acid resistance</td>
<td>+46.5 ± 3.5</td>
<td>+3.5</td>
</tr>
<tr>
<td>gadC</td>
<td>Predicted glutamate transporter</td>
<td>+11.4 ± 6.1</td>
<td>+3.3</td>
</tr>
<tr>
<td>gadE</td>
<td>Transcriptional activator</td>
<td>+3.3 ± 1.9</td>
<td>+1.4</td>
</tr>
<tr>
<td>gadX</td>
<td>Transcriptional dual regulator</td>
<td>+2.3 ± 0.9</td>
<td>+1.6</td>
</tr>
<tr>
<td>gatD</td>
<td>Galactitol-1-phosphate dehydrogenase</td>
<td>+3.8 ± 0.8</td>
<td>+3.2</td>
</tr>
<tr>
<td>aldB</td>
<td>Aldehyde dehydrogenase</td>
<td>+3.2 ± 0.4</td>
<td>+4.9</td>
</tr>
<tr>
<td>hdeD</td>
<td>Acid-resistance membrane protein</td>
<td>+3.2 ± 1.6</td>
<td>+1.7</td>
</tr>
<tr>
<td>cls</td>
<td>Cardiolipin synthase</td>
<td>+1.9 ± 0.3</td>
<td>+3.1</td>
</tr>
<tr>
<td>ahpC</td>
<td>Alkylhydroperoxide reductase</td>
<td>+2.4 ± 0.7</td>
<td>+3.8</td>
</tr>
</tbody>
</table>

Quorum Sensing

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold Change</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>lsrB</td>
<td>Autoinducer-2 transporter</td>
<td>+3.1 ± 0.8</td>
<td>+3.5</td>
</tr>
<tr>
<td>lsrK</td>
<td>lsr kinase</td>
<td>+1.8 ± 0.2</td>
<td>+3.0</td>
</tr>
<tr>
<td>lsrR</td>
<td>lsr regulator</td>
<td>+1.7 ± 0.3</td>
<td>+2.7</td>
</tr>
</tbody>
</table>

**Figure 3.** Selected transcripts that are differentially regulated in the *lexA umuC* mutant strain.
Figure 4. $umuC^+$ influences diverse phenotypes under conditions of SOS induction. 

A) Swarming assays reveal a severe defect in the $lexA$ $umuC$ double mutant and flagellin preparations show a drastic reduction in FliC, consistent with the results of the microarray, that is specific to conditions of SOS induction. 

B) Glutamate-dependent acid resistance assay for the indicated strains. Percent survival is plotted with (gray bars) and without (black bars) glutamate. Error bars represent the standard deviation determined from three independent transductants.
Figure 5. Deletion of *umuC* in a constitutively SOS-induced strain background confers profound resistance to agents that perturb DNA replication. A-B) Percent survival during HU challenge for the indicated strains in *lexA*+ and *lexA*(Def) backgrounds. Error bars represent one standard deviation determined from three experiments. C) Percent survival during novobiocin treatment for the indicated strains in *lexA*+ (black bars) and *lexA*(Def) (gray bars) backgrounds. Error bars represent one standard deviation determined from three experiments. D) Percent survival during HU challenge for a *sulA* strain (black bars) and a *sulA umuC* strain (gray bars) indicates that the majority of HU resistance conferred by *umuC* deletion occurs due to constitutive SOS induction.
Figure 6. The *umuD* gene products mediate altered cellular and colony morphologies as well as resistance to replication inhibitors in a *lexA*(Def) *ΔumuC* strain. A) Expression of UmuD or UmuD' does not affect the small colony morphology of a *lexA*(Def) *ΔumuDC* strain, whereas expression of UmuD(S60A) results in large colonies similar to a *lexA*(Def) *ΔumuC* strain. Comparatively, expression of UmuD results very similar cellular morphologies to a *lexA*(Def) *ΔumuDC* strain alone, expression of UmuD' results in shorter cells, and expression of UmuD(S60A) results in a distribution of very short cells and very long cells. B) Transformation of a *lexA*(Def) *ΔumuDC* strain with the same series of plasmid-borne UmuD variants indicates that UmuD indicates that resistance to 25 mM HU requires either UmuD' or UmuD, but that UmuD(DS60A) does not promote survival, exacerbating the sensitivity of the strain to HU.
Chapter 8

Conclusions
Why is DinB the Only Y-family DNA Polymerase Conserved in all Domains of Life?

The striking evolutionary conservation of DinB/pol κ (1) suggests that it performs a crucial biological function. However, despite this conservation, deletion of dinB does not affect growth rate or spontaneous mutagenesis (2). Evidence presented in chapters 2 and 4 of this thesis strongly suggest that a critical function for DinB and its orthologs is to bypass templates with modifications at the \( N^2 \) position of dG (3). The substrate specificities of DinB and its mammalian and archaean orthologs with respect to \( N^2 \)-dG modification is remarkably similar. Intriguingly, these enzymes seem to share a preference for action on \( N^2 \)-furfuryl-dG, a structural analog of the \( N^2 \)-dG lesion produced by the DNA damaging agent nitrofurazone (NFZ). Given their largely overlapping substrate specificities, it would be interesting to investigate whether heterologous expression of Dbh or pol κ can rescue the NFZ sensitivity of a \( \Delta \)dinB E. coli strain. Although \( N^2 \)-furfuryl-dG can be generated in vivo by oxidation of deoxyribose (4), it is likely a member of a class of ubiquitous replication blocking \( N^2 \)-dG lesions that DinB and its orthologs throughout evolution are specialized to bypass. Provocatively, a study of human pol κ activity on \( N^2 \)-dG modified templates in vitro revealed that it has a preference for activity on \( N^2 \)-benzyl-dG but not larger or smaller modifications (5), suggesting that the lesion binding pockets of DinB and its orthologs are specialized to accommodate substrates with structural similarity to both that adduct and \( N^2 \)-furfuryl-dG.

Identification of the endogenous adducts that DinB and its orthologs are specialized to bypass is of chief importance. Provocatively, deletion of dinB does confer a competitive disadvantage to E. coli strains (6). The importance of DinB TLS function for this effect could be addressed using the DinB(F13V) mutant characterized in chapter 2, which separates DinB function in TLS over \( N^2 \)-dG adducts from its ability to catalyze normal DNA synthesis.
Moreover, the effect of general catalytic activity, shorter patch synthesis, interaction with the \( \beta \) -processivity clamp, and UmuD binding could be addressed using other \( \text{dinB} \) alleles characterized in this thesis. Isolation of endogenous DinB substrates is a much more difficult proposition, but may be possible using an artful covalent crosslinking strategy followed by mass spectrometry.

Is the mutagenic potential of DinB important for its maintenance through evolution? Such an assertion is difficult to prove, but it is interesting that even eukaryotic pol \( \kappa \) can act with a relatively high error frequency (7). The results presented in chapter 5 suggest that the mutagenic function of DinB, which has been shown to be adaptive in certain circumstances (8, 9), can be exquisitely tuned by the cell during the course of the SOS response. These observations indicate that mutagenic TLS may not always be undesirable. The fact that \( \Delta \text{dinB} \) \( \text{E. coli} \) strains do not show an elevated frequency of NFZ- or 4-NQO-induced mutagenesis (3) indicates that DinB-catalyzed bypass of lesions produced by those agents is accurate relative to alternative mechanisms the cell can employ. \textit{In vitro} primer extension assays reveal, however, that DinB’s fidelity when replicating past \( N^2 \)-fufuryl-dG is moderate, with error frequencies between \( 10^{-3} \) and \( 10^{-5} \). This is still several orders of magnitude higher than the error frequencies of the replicative DNA polymerase III (2). It is possible that there simply is no selective pressure to generate and maintain a more faithful DinB enzyme. Alternatively, DinB’s modest and tunable mutagenic potential may be beneficial to the cell under certain circumstances (10).

\textbf{Determinants of DinB Function in Translesion Synthesis}

The results presented in chapters 2 and 3 indicate that the intrinsic preference of DinB for action on \( N^2 \)-dG adducted templates is determined by specific residues in its active site. The pocket of aromatic hydrophobic amino acids lining the putative lesion binding pocket of DinB
play critical roles in dictating diverse aspects of its TLS function including enhanced catalytic proficiency and product length. Other determinants of DinB TLS function likely exist. Indeed, a random mutagenesis screen using hydroxylamine has revealed several \( \text{dinB} \) alleles that behave as \( \text{dinB}(F13V) \) or \( \text{dinB}(Y79L) \) with respect to their ability to complement the NFZ sensitivity of a \( \Delta \text{dinB} \) strain. These mutants, and other residues that influence TLS function that are both general to Y-family DNA polymerases and specific to DinB and its orthologs are poorly understood and merit further study.

**How Elaborate is Regulation of DinB Function?**

DinB activity appears to be controlled by the cell at nearly every level. Its transcriptional regulation as part of the SOS response to DNA damage (11) and in response to treatment with \( \beta \)-lactam antibiotics (12, 13) is extremely intricate and well-characterized (2). The findings discussed in chapters 2-5 suggest that both intrinsic preference for catalysis on proper substrates and an elaborate network of protein-protein are critical for dictating DinB function. In addition interaction with to the \( \beta \) processivity clamp (14), interaction with both RecA and the \( \text{umuD} \) gene products appears to selectively modulate DinB function, especially under conditions of SOS induction. It is likely that these and additional protein modulators of DinB function may differ depending on the physiological state of the cell. Moreover, the products of DinB are subject to the action of the \( \varepsilon \) proofreading subunit of DNA polymerase III, which has been shown to interact with the C-terminus of UmuD (15).

Provocatively, the mutagenic potential of DinB appears to be tunable. Although overexpression of DinB causes a -1 frameshift mutator effect, data presented in chapter 5 suggest that co-overproduction of UmuD suppresses this phenomenon. This effect requires \( \text{recA}^+ \)
function and DinB, UmuD, and RecA form a ternary complex under physiological conditions. Moreover, UmuD and RecA modulate DinB function by inhibiting its activity on mismatched or $N^2$-dG adducted templates but promoting its activity on normal templates. These observations suggest that the umuD gene products, acting in conjunction with RecA, are able to modulate DinB function in a highly sophisticated manner. During the course of the SOS response, the forms of the umuD gene products change in a highly orchestrated fashion (16, 17), providing a mechanism through which DinB’s mutagenic potential might also be tuned through the response. Taken together, these data suggest that DinB function is generally accurate, but that its mutagenic potential can be unleashed by the cell under certain conditions.

The affinity chromatography experiment described in chapter 5 revealed numerous factors aside from UmuD, UmuD', and RecA that associate with DinB. For example, HupA, HupB, and YcfD were found to interact with DinB both biochemically and genetically in experiments described in appendix B. Their potential to regulate DinB function has not been addressed directly, but in Sulfolobus solfataricus the activity of a B-family polymerase has been shown to be specifically modulated by the function of small histone-like DNA binding proteins (18). The scope of chromatin structure in modulating TLS has not been investigated in any organism, but is likely to be important given the recent suggestions that much of TLS takes place not at stalled replication forks but rather at gaps left after downstream replication reinitiation (2, 19, 20).

**Alternative Biological Functions of Y-family DNA Polymerases**

Canonically, Y-family DNA polymerases are thought to function at the terminus of DNA damage and stress responses. The data presented in chapters 6 and 7, however, suggests that
under certain conditions *E. coli* UmuC may participate in several additional functions aside from its established roles in TLS (21, 22) and in effecting a primitive replication checkpoint (23-25). Unexpectedly, a gain-of-function allele of *umuC* confers dramatic hydroxyurea resistance (26), challenging the notion that killing by that agent proceeds exclusively through dNTP starvation. Killing by HU in the wild-type strain appears to be mediated by toxin-antitoxin pairs, many of which have overlapping function (27). In a distinct but equally provocative phenomenon described in chapter 7, *umuC*\(^+\) appears to mediate numerous physiological responses during the SOS response. Deletion of *umuC* in a constitutively SOS-induced strain mitigates the extent to which certain SOS genes are expressed and renders the strain comparatively immotile and strikingly resistant acid stress and agents that perturb DNA replication. In contrast, deletion of *umuC* in a wild-type background does not affect these phenomena. These observations suggest that, at least under conditions of SOS induction, *umuC* functions in genetic pathways that promote diverse facets of cellular physiology.

**Genetic Interactions Between *dinB* and Other DNA Repair Pathways**

The genetic interactions between *dinB* and certain DNA repair glycosylases, described in appendix A, are quite provocative. Moreover, pol \(\kappa\) has recently been shown to function in nucleotide excision repair in mouse cells (28). These observations hint at a potentially intricate coordination of error-free and error-prone aspects of the DNA damage response. The putative function of DinB and its orthologs in excision repair is unclear, but it may represent an effort to mitigate the tendency of \(N^2\)-dG lesions to evade such repair mechanisms (29, 30).
Conclusions

There are several broad implications of the work presented in this thesis. Principally, DinB and its orthologs do not function as ‘sloppier copiers’ of imperfect DNA templates (31), but rather are specialized DNA polymerases (10) with apparently tunable mutagenic potential (32), which can be adaptive in certain circumstances (33). The striking similarity in substrate specificities among DinB and its mammalian and archaeal orthologs suggests that proficient and accurate bypass of a class of $N^2$-dG lesions is a driving force for the evolutionary maintenance of DinB among all domains of life. DinB function in TLS, as well as its mutagenic potential, is elaboratey controlled both by intrinsic features (3, 34) and by its protein-protein interactions with diverse cellular factors (35), including RecA and the $umuD$ gene products.

In addition to its function in TLS as the catalytic subunit of DNA polymerase V (21, 22), the data presented in chapters 6 and 7 indicate that UmuC promotes cell death in response to agents that perturb DNA replication (26). Moreover, under conditions of SOS induction, $umuC^+$ function influences the expression of numerous SOS-regulated transcripts along with diverse physiological responses such as motility, quorum sensing, and the acid stress response. These observations hint at a largely uncharacterized function of Y-family polymerases in sculpting physiological responses, including active mechanisms of cell death, in response to environmental stress.
References


30. Zewail-Foote, M., V. S. Li, H. Kohn, D. Bearss, M. Guzman, and L. H. Hurley. 2001. The inefficiency of incisions of ecteinascidin 743-DNA adducts by the UvrABC nuclease and the unique structural feature of the DNA adducts can be used to explain the repair-dependent toxicities of this antitumor agent. Chem. Biol. **8**:1033-49.


Appendix A

Genetic Evidence that DinB Acts on DNA Damage that is Recalcitrant to Excision Repair Mechanisms in *Escherichia coli*
Introduction and Results

Given the multiple and redundant high fidelity DNA repair and tolerance pathways available to the cell, why employ potentially mutagenic translesion synthesis (TLS)? An emerging body of evidence suggests that certain types of DNA damage, particularly modification at the $N^2$ position of dG, may be recalcitrant to such repair pathways. For example, in mammalian cells damaged with acetylaminofluorine, the $N^2$-dG isomer is the most persistent lesion observed despite being the least commonly produced (1). Moreover, certain other $N^2$-dG adducts have been shown to be recalcitrant to repair by the *E. coli* nucleotide excision repair system (2).

Is TLS then necessary to contend with these persistent lesions? To answer this question, double deletion mutants of *dinB* and each of the 12 known DNA repair glycosylases were constructed in *E. coli* and their behavior when treated with the DNA damaging agents nitrofurazone (NFZ) and 4-nitroquinoline-1-oxide (4-NQO) was examined. Deletion of either *alkA, uvrA, nth* or *nei* confers considerable sensitivity to a wild-type strain. If DinB dependent TLS acts on replication blocking substrates that it shares with any of these enzymes a synergistic relationship should be observed with respect to survival. However, such a synergy was not observed with respect to survival for any of the glycosylases examined in combination with *dinB* (Table 1). Taken together, these observations indicated that DinB likely performs TLS across lesions that are persistent due to their inability to be resolved by other DNA repair and damage tolerance pathways. Considerable redundancy also exists among DNA repair glycosylases in *E. coli* (3). It is therefore a formal possibility that DinB could catalyze TLS opposite certain ubiquitous lesions that are repaired by redundant mechanisms. Nonetheless, our inability to observe a synergistic relationship in any of the double mutants indicates that the lethal NFZ-
induced lesions that DinB bypasses are not efficiently repaired by nucleotide excision repair (NER) or base excision repair (BER).

Quite provocatively \textit{dinB} appears to be epistastic to \textit{nth} and partially epistatic to \textit{uvrA}. The epistasis with \textit{nth} is remarkable in light of the lack thereof with \textit{nei}, which is considered to have somewhat overlapping substrate specificity (3). Moreover, the fact that \textit{dinB} is epistatic to \textit{nth} with respect NFZ sensitivity and not vice versa suggests that \textit{nth} function requires \textit{dinB} under these conditions. Alternatively put, this implies that during NFZ challenge \textit{dinB} function is counterproductive in the absence of \textit{nth}. One possible model for these data is that during NFZ treatment DinB generates a potentially toxic intermediate that can be repaired by Nth. Similarly, the observation that, at least for NFZ sensitivity, \textit{dinB} is partially epistatic to \textit{uvrA} suggests that DinB is required for at least some of UvrA’s function under these conditions. Indeed, DinB may generate a toxic intermediate that could be repaired by either Nth or UvrA. Examination of a triple mutant might be especially informative in that case. These observations are particularly interesting given the recent finding that pol \textit{k} may play a role in nucleotide excision repair in mammalian cells (4).

Despite its remarkable conservation throughout evolution, a genomic deletion of \textit{dinB} does not confer any dramatic phenotypes in \textit{E. coli} in the absence of external treatment (3, 5). Quite strikingly, \textit{ΔdinB} also strains do not show a significantly altered spontaneous mutation frequency to rifampicin resistance in a wild-type background and these observations are echoed in \textit{alkA}, \textit{alkB}, \textit{uvrA}, and \textit{mutM} backgrounds (data not shown). Provocatively, \textit{nth} mutants are known to display a weak spontaneous mutator effect (6). It would be interesting to determine what role, if any, \textit{dinB} function may play in such a phenomenon.
Materials and Methods

Construction of glycosylase mutants

Precise deletions of each individual glycosylase in this study were obtained from the Nara Institute (Japan) as part of the Keio collection (7) and transduced into AB1157 and its \textit{AdinB} derivatives (8) using standard methods. The genotype of the strains was confirmed by polymerase chain reaction (PCR).

Nitrofurazone sensitivity and mutagenesis assays

Nitrofurazone sensitivity assays were performed essentially as previously described (8, 9). Briefly, stains were grown to fresh saturation in liquid LB medium and plated at the appropriate dilution onto LB agar containing NFZ concentrations between 0 and 3 \( \mu \text{g/mL} \). In parallel, viable counts were measured on LB agar alone. Colonies were counted after 24 hours of growth and percent survival was scored.

2. Zewail-Foote, M., V. S. Li, H. Kohn, D. Bearss, M. Guzman, and L. H. Hurley. 2001. The inefficiency of incisions of ecteinascidin 743-DNA adducts by the UvrABC nuclease and the unique structural feature of the DNA adducts can be used to explain the repair-dependent toxicities of this antitumor agent. Chem. Biol. 8:1033-49.


<table>
<thead>
<tr>
<th>E. coli Gene</th>
<th>Substrate/Activity</th>
<th>Fold Killing in WT Strain</th>
<th>Fold Killing in AdiB Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>ung+</td>
<td>U</td>
<td>0.8 ± 0.1</td>
<td>0.9 ± 0.5</td>
</tr>
<tr>
<td>mug+</td>
<td>U, T, or εC opposite G</td>
<td>0.4 ± 0.2</td>
<td>0.6 ± 0.4</td>
</tr>
<tr>
<td>fpg+ (mutM+)</td>
<td>8-oxo-G opposite C</td>
<td>2.1 ± 0.2</td>
<td>0.2 ± 0.1</td>
</tr>
<tr>
<td>mutY+</td>
<td>A opposite 8-oxo-G</td>
<td>2.7 ± 1.5</td>
<td>1.4 ± 0.2</td>
</tr>
<tr>
<td>nth+</td>
<td>Ring-saturated or fragmented pyrimidines</td>
<td>1390 ± 70</td>
<td>0.2 ± 0.1</td>
</tr>
<tr>
<td>alkA+</td>
<td>3-meA, εA, hypoxanthine, broad substrate range</td>
<td>5.7 ± 0.9</td>
<td>1.7 ± 0.3</td>
</tr>
<tr>
<td>nei+</td>
<td>Thymine glycol</td>
<td>1.2 ± 0.2</td>
<td>0.4 ± 0.1</td>
</tr>
<tr>
<td>phrA+</td>
<td>CPD photolyase</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>ada+</td>
<td>N6-meG alkyltransferase</td>
<td>0.6 ± 0.1</td>
<td>0.1 ± 0.05</td>
</tr>
<tr>
<td>alkB+</td>
<td>1-meA, 3-meC demethylase</td>
<td>1.8 ± 0.9</td>
<td>2.8 ± 0.4</td>
</tr>
<tr>
<td>uvrA+</td>
<td>NER</td>
<td>16,800 ± 1800</td>
<td>2430 ± 100</td>
</tr>
</tbody>
</table>

Table 1. Survival of glycosylase mutant strains during NFZ challenge.
Appendix B

Bacterial Histone-like Proteins HupA and the Putative Histone Demethylase YcfD
May Regulate DinB Function
Introduction and Results

Recent studies have revealed that at least a large fraction of translesion synthesis (TLS) catalyzed by Y-family DNA polymerases may take place not at the replication fork, but rather at gaps that persist in the genome after replication has reinitiated downstream of a DNA lesion (1, 2). Depending on the extent of DNA condensation that occurs prior to TLS, an unforeseen layer of regulation may therefore be required to promote lesion bypass. The possible influence of chromatin structure on TLS has not been investigated in detail, but in human cells subjected to ultraviolet irradiation pol η, Rad18, Rad6, and Rev1 are present in the chromatin-associated fraction (3).

Our efforts to identify protein regulators of DinB function, described in chapter 5 of this thesis, revealed several interacting proteins aside from UmuD, UmuD', and RecA. We also identified the bacterial histone-like proteins HupA and HupB and the protein of unknown function YcfD by Edman degradation. In Escherichia coli HupA and HupB can form both heterodimeric and homodimeric species, the distribution of which changes depending on growth conditions (4). The presence of these proteins was particularly intriguing because they have been shown to tightly bind to gapped DNA substrates (5) and DNA recombination and repair intermediates (6). The hupA+ gene is also required for SOS induction in vivo (7). No function has been assigned to the ycfD gene product, but it shows considerable sequence homology to the cupin superfamily of proteins (8) and a recently-characterized family of histone demethylases that employ a novel mechanism requiring Fe^{2+} and 2-oxoglutarate (Figure 1A) (9, 10).

Within the genome of E. coli, the ycfD gene is located downstream of purB, phoP, and phoQ within the genome of E. coli in an apparent operon. Similar architecture is conserved among many bacterial species, as is the YcfD protein, and the phoPQ genes have been
implicated in regulation of adaptive mutation under conditions of nonlethal selection (11). Additionally, *hupA* and *hupB* mutants show altered frequencies of *dinB*<sup>+</sup>-dependent adaptive mutagenesis (PL Foster, personal communication). A homology model of YcfD reveals a classic cupin fold with properly positioned putative Fe<sup>2+</sup>-binding residues and a helix-hairpin-helix C-terminal domain (Figure 1B).

We recapitulated the interactions of DinB with C-terminally hexahistidine tagged variants of each of HupA, HupB, and YcfD *in vitro* using a Ni<sup>2+</sup> affinity resin (data not shown). We also introduced deletions of *hupA*, *hupB*, and *ycfD* into wild-type and *lexA* deficient *E. coli* backgrounds, and introduced a *dinB* deletion into each of these strains to generate a series of double mutants (Table 1). The *hupA* deletion strains are remarkably sensitive to the DNA damaging agent nitrofurazone (NFZ) (Figure 2A). Moreover, deletion of *dinB* in a Δ*hupA* strain background confers no further NFZ sensitivity, indicating that *hupA* is epistatic to *dinB* for this phenotype (Figure 2B). This relationship is preserved in both *lexA<sup>+</sup>* and *lexA* deficient backgrounds, indicating that the involvement of the *hupA* gene product in *dinB*<sup>+</sup>-dependent NFZ sensitivity does not arise entirely from its function in mediating SOS induction. Deletion of *hupB* does not appear to affect NFZ resistance regardless of *lexA* status, but a Δ*ycfD* strain does show slight sensitivity to NFZ in a *lexA* deficient background (Figure 2A). Deletion of *dinB* does not affect the NFZ sensitivity of this strain, an effect that merits further characterization (Figure 2B).

Deletion of *hupA*, *hupB*, or *ycfD* also results in a variety of other phenomena. In a *lexA* deficient background, deletion of *ycfD* causes a 4-fold spontaneous mutator effect and a larger colony morphology (data not shown). This does not appear to result from a change in cell shape, however, as this strain appears similar to its *lexA* deficient parent on the microscopic level (Figure 3). Deletion of either *hupA* or *hupB* also confers a substantial increase in doubling time.
in rich medium that is nearly restored to that of wild-type by deletion of \textit{dinB} (Table 3). Taken together, these observations suggest that both the \textit{hupA} and the \textit{ycfD} gene products regulate DinB function.

**Conclusions and Future Directions**

These findings suggest that \textit{hupA} is required for DinB function independently of its role in SOS induction (7). Moreover, \textit{ycfD} may regulate DinB function under conditions of SOS-induction. Certain of these effects are likely direct given our ability to observe stable interactions between DinB and each of these proteins.

Why would DinB interact with a putative histone demethylase? It is tempting to speculate the HupA, perhaps in a post-translationally modified form, may bind to aberrant DNA structures that are potential DinB substrates. Demethylation of HupA by YcfD might then be necessary for DinB to gain access to the primer terminus and catalyze translesion synthesis. Although there is no published evidence that either HupA or HupB are posttranslationally modified, we consistently observe a small amount of a higher molecular weight species by MALDI-TOF mass spectrometry in our purified fractions of each protein that disappears upon the addition of YcfD, 2-oxoglutarate, FeSO$_4$, and ascorbate. This effect, as well as any effects HupA, HupB, or YcfD have on DinB-dependent TLS \textit{in vitro} need to be examined in greater detail.

The \textit{dinB}+-dependent slowed growth rate of \textit{hupA} and \textit{ycfD} deletion strains underscores the notion that the control of DinB function provided by HupA and YcfD is crucial for cellular growth and division. It is possible that HupA and YcfD prevent DinB from catalyzing indiscriminate DNA synthesis and that such uncontrolled DinB function would slow growth. If
so, such DinB function is unlikely to be very mutagenic as the slight mutator phenotype of a \( \Delta ycfD \) strain is unaltered by deletion of \( \text{dinB} \). Complementation of this phenotype with plasmid-borne \( \text{dinB} \) variants might shed considerable light on the mechanisms governing this effect.

**Materials and Methods**

**Bacterial strains**

Bacterial strains used in this study are described in Table 1 and were generated by P1(vir) transduction of deletion mutants from strains in the Keio collection (12) into AB1157 and AB1157 \( \text{lexA}(\text{Def}) \) backgrounds. Strains were routinely grown in Luria-Burtani (LB) broth at 37°C.

**Sensitivity and mutagenesis assays**

NFZ sensitivity assays (13, 14) and spontaneous rifampicin resistance reversion assays (14) were performed as previously described.
References


### Table 1. Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (DE3) pLysS</td>
<td><em>E. coli</em> B. Standard strain used for protein overproduction</td>
<td>GCW lab stock</td>
</tr>
<tr>
<td>AB1157</td>
<td><em>thr</em>-1 <em>araC</em>-14 <em>leuB</em>-6*(Am) *Δ(gpt-proA)*62 <em>lacY1</em> <em>tsx</em>-33 <em>qsr</em>'-0</td>
<td>GCW lab stock</td>
</tr>
<tr>
<td></td>
<td><em>glnV44</em>(AS) <em>galK2</em>(Oc) <em>LAM Rac</em>-0 <em>hisG4</em>(Oc) <em>rfbC1</em> <em>mgl</em>-51 <em>rpoS396</em>(Am)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>rpsL31</em> <em>kgdK51</em> <em>xylA5</em> <em>mlt</em>-1 <em>argE3</em>(Oc) <em>thi</em>-1</td>
<td></td>
</tr>
<tr>
<td>DFJ135</td>
<td>As AB1157 but sulA lexA300(Def)</td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 Δ<em>hupA</em></td>
<td>As AB1157 but Δ<em>hupA::kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 Δ<em>hupA ΔdinB</em></td>
<td>As above but Δ<em>dinB</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 lexA(Def) Δ<em>hupA</em></td>
<td>As DFJ135 but Δ<em>hupA::kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 lexA(Def) Δ<em>hupA</em></td>
<td>As above but Δ<em>dinB</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 Δ<em>hupB</em></td>
<td>As AB1157 but Δ<em>hupB::kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 Δ<em>hupB ΔdinB</em></td>
<td>As above but Δ<em>dinB</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 lexA(Def) Δ<em>hupB</em></td>
<td>As DFJ135 but Δ<em>hupB::kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 lexA(Def) Δ<em>hupB</em></td>
<td>As above but Δ<em>dinB</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 Δ<em>ycfD</em></td>
<td>As AB1157 but Δ<em>ycfD::kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 Δ<em>ycfD ΔdinB</em></td>
<td>As above but Δ<em>dinB</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 lexA(Def) Δ<em>ycfD</em></td>
<td>As DFJ135 but Δ<em>ycfD::kan</em></td>
<td>This work</td>
</tr>
<tr>
<td>AB1157 lexA(Def) Δ<em>ycfD</em></td>
<td>As above but Δ<em>dinB</em></td>
<td>This work</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDFJ1</td>
<td>DinB expression plasmid under T7 promoter in pET11T backbone;</td>
<td>Jarosz et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Ampicillin resistance</td>
<td></td>
</tr>
<tr>
<td>pHupA</td>
<td>HupA(His)$_6$ expression plasmid under <em>lac</em> promoter; Chloramphenicol</td>
<td>Kitagawa et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>resistance</td>
<td></td>
</tr>
<tr>
<td>pHupB</td>
<td>HupB(His)$_6$ expression plasmid under <em>lac</em> promoter; Chloramphenicol</td>
<td>Kitagawa et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>resistance</td>
<td></td>
</tr>
<tr>
<td>pYcfD</td>
<td>YcfD(His)$_6$ expression plasmid under <em>lac</em> promoter; Chloramphenicol</td>
<td>Kitagawa et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>resistance</td>
<td></td>
</tr>
<tr>
<td>Strain</td>
<td>Reversion Frequency to Rif&lt;sup&gt;R&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------</td>
<td></td>
</tr>
<tr>
<td>AB1157</td>
<td>2.9 x 10&lt;sup&gt;-9&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>ΔycfD</td>
<td>4.3 x 10&lt;sup&gt;-9&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>ΔhupA</td>
<td>3.8 x 10&lt;sup&gt;-10&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>lexA300(Def)</td>
<td>4.8 x 10&lt;sup&gt;-9&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>lexA300 ΔycfD</td>
<td>1.5 x 10&lt;sup&gt;-8&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>lexA300 ΔhupA</td>
<td>1.4 x 10&lt;sup&gt;-9&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Spontaneous mutation frequencies to rifampicin resistance of ycfD and hupA deletions.
Table 3. Doubling times of $\Delta ycfD$ and $\Delta hupA$ strains and their $\Delta dinB$ derivatives in rich medium as determined by OD$_{600}$ measurements.
Figure 1. YcfD shows homology with the cupin superfamily of proteins and specifically with a novel class of histone demethylases. A) Alignment of YcfD homologs from diverse bacterial species. Putative Fe$^{2+}$ binding residues are indicated with arrows. B) A homology model of YcfD suggests that it adopts a typical cupin fold with the putative Fe$^{2+}$ binding residues in proximity to one another and a C-terminal helix-hairpin-helix motif.
Figure 2. Deletion of either hupA or hupB confers sensitivity to NFZ. A) Percent survival vs. NFZ dose in a lexA<sup>+</sup> background. Parental strain (closed circles); ΔhupA (closed squares); ΔhupB (closed triangles); ΔycfD (closed diamonds). B) Percent survival vs. NFZ dose in a lexA deficient background. Parental strain (open circles); ΔhupA (open squares); ΔhupB (open triangles); ΔycfD (open diamonds). C) hupA is epistatic to dinB. Strains as indicated in A-B and ΔdinB (closed crossed circles); lexA(Def) ΔdinB (open crossed circles); ΔdinB ΔhupA (closed crossed squares); lexA(Def) ΔdinB ΔhupA (open crossed squares). D) ycfD is epistatic to dinB in a lexA(Def) background. Strains as indicated in A-C and ΔdinB ΔycfD (closed crossed diamonds); lexA(Def) ΔdinB ΔycfD (open crossed diamonds).
Figure 3. Fluorescence micrographs of the *lexA*(Def) strain and its *ΔycfD* derivative indicate that the difference in colony morphologies is not dramatically manifested at the cellular level. Red membrane stain is FM464; Blue DNA stain is DAPI.
Appendix C

The *umuD* Gene Products Control DinB Function in Translesion Synthesis Under Conditions of SOS Induction
Results and Discussion

Our discovery that DinB interacts with UmuD, UmuD', and RecA led us to examine the effect of the umuD gene on nitrofurazone (NFZ) resistance in vivo. We were initially surprised that deletion of either umuDC or umuC did not seem to affect NFZ resistance (Figure 1). We then considered that the K_d for the interaction between DinB and UmuD lies between the non-SOS-induced and SOS-induced levels of UmuD (1, 2). When we examined the NFZ sensitivity of ΔumuDC strain in a constitutively SOS-induced lexA deficient background, we found that it is indeed sensitive to that agent whereas a lexA(Def) ΔumuC strain is not (Figure 1). These data indicate that NFZ resistance requires umuD+ under conditions of SOS induction.

We wondered whether this effect might arise due to the direct interaction between DinB and UmuD. Therefore, we examined the ability of DinB(F172A), a DinB variant that is impaired with respect to UmuD interaction (1), to complement the NFZ sensitivity of a lexA(Def) ΔdinB strain when expressed. Expression of this variant complements the NFZ sensitivity of a lexA+ ΔdinB strain, just as expression of wild-type DinB does (Figure 1). However, its expression is unable to confer full NFZ resistance to the lexA deficient ΔdinB strain (Figure 1). This result strongly suggests that the requirement of umuD+ for NFZ resistance under conditions of SOS induction arises from a direct interaction between DinB and UmuD.

The umuD gene products are subject to elaborate posttranslational modification. The initially formed UmuD2 protein undergoes facilitated autoproteolysis that depends on RecA:ssDNA filaments that are formed upon DNA damage. This process results in the smaller species UmuD', which is active in UV-induced mutagenesis (3, 4). This modification appears to function as a temporal switch in the DNA damage response in E. coli, dividing error-free DNA repair and damage tolerance mechanisms from mutagenic translesion synthesis (TLS). We
wondered which form(s) of the *umuD* gene products are required for resistance to NFZ during SOS induction. We therefore examined whether plasmids expressing UmuD, UmuD', or a UmuD(S60A) noncleavable variant could complement the NFZ sensitivity of a *lexA*(Def) Δ*umuDC* strain. Strikingly, expression of either UmuD or UmuD', but not UmuD(S60A) fully restores the NFZ resistance of the *lexA*(Def) Δ*umuDC* strain. These data indicate that the function of UmuD' is required for NFZ resistance.

Studies described in chapter 5 indicate that UmuD2 is able to modulate DinB’s ability to extend from a bulged -1 frameshift intermediate. We wondered whether UmuD2 might similarly modulate DinB catalyzed TLS over a site-specific *N*²-furfuryl-dG lesion. Addition of UmuD2 and RecA to primer extension assays does not affect DinB-catalyzed insertion opposite *N*²-furfuryl-dG, but has a profound effect on extension from that lesion (Figure 3). This effect was not recapitulated on an undamaged control oligonucleotide of the same sequence (data not shown), but UmuD2 and RecA promote rather than inhibit DinB activity on undamaged templates (1). Moreover, addition of UmuD2 alone had no effect on DinB-catalyzed insertion or extension on either substrate. Taken together, these observations suggest that the TLS function of DinB can be directly modulated by the *umuD* gene products, in conjunction with RecA, under conditions of SOS induction.

Why might the cell modulate DinB TLS function in such a nuanced manner? This is an especially interesting question given that eukaryotes often use two polymerases to perform translesion synthesis – one for insertion and another for extension (5-7). In *E. coli*, where DNA polymerases are comparatively limited in number, the *umuD* gene products may provide this additional level of control.
At first glance, the requirement of umuD\(^+\) for dinB\(^+\)-dependent NFZ resistance in vivo seems at odds with the observation that UmuD\(_2\) and RecA act in concert to modulate DinB function in vitro. However, these data may reflect the limitations of our in vitro assay. It is possible that DinB's activity is modulated by UmuD\(_2\) in vivo entirely analogously to what we observe in vitro. Loss of this regulatory function may result in a mutagenic risk, but should not negatively impact survival as DinB can still perform TLS. In contrast, UmuD'\(_2\) may interact with DinB in vivo in a manner that requires other cellular factors that are missing in our assay system. This interaction, in addition to interactions UmuD'\(_2\) makes with other proteins, appear to be critical for activation of DinB function in vivo.

Taken together, these observations are consistent with a model in which posttranslational modification of UmuD\(_2\) represents a temporal switch in the SOS response to DNA damage (Figure 4). Initially, UmuD\(_2\) predominates, and DinB may be recruited to an aberrant template and can even synthesize DNA up to an opposite a DNA lesion. Autocleavage of UmuD\(_2\) to form UmuD'\(_2\) appears to be required for the extension steps of TLS to be completed, however, thereby committing the cell to a potentially mutagenic damage tolerance process rather than a more accurate means to contend with genotoxic stress.

**Materials and Methods**

The strains and plasmids used in this appendix are described in tables accompanying chapters 5 and 7. NFZ sensitivity experiments and lesion bypass assays were performed as previously described (8, 9). RecA was added in stoichiometric quantities with DinB. UmuD, UmuD', and RecA were used at 10 \(\mu\)M.
References


Figure 1. The *umuD* gene is required for NFZ resistance in a *lexA* deficient background (gray bars), but not in a *lexA*+ background (black bars). One hundred percent survival is defined as that of the *lexA*+ or *lexA*(Def) parental strains, respectively. Expression of the UmuD-binding deficient DinB variant DinB(F172A) is unable to fully complement the NFZ sensitivity of the *lexA*(Def) ΔdinB strain. The NFZ dose used in this experiment is 7.6 μM; error bars represent the standard deviation determined from three independent transformants; pWSK29 is the empty vector for pDinB and pDinB-F172A.
Figure 2. Complementation of the NFZ sensitivity of a \textit{lexA(Def) \DeltaumuDC} strain. Percent survival is relative to the parental \textit{lexA(Def)} strain bearing an empty vector control. Expression of either UmuD or UmuD' complements sensitivity, but expression of the noncleavable UmuD(S60A) variant does not. Error bars represent the standard deviation determined from three independent transformants.
**Figure 3.** UmuD₂ and RecA specifically control the ability of DinB to extend from N²-furfuryl-dG lesion, but do not affect insertion opposite to it. Plots display initial reaction velocities vs. substrate concentrations. DinB + RecA (black); DinB + RecA + UmuD₂ (red); DinB + RecA + UmuD'₂ (blue).
Figure 4. Model for the function of the umuD gene products in a temporal switch between error-free and error-prone phases of the SOS response to DNA damage in *E. coli*. 
Curriculum Vitae
Daniel F. Jarosz

Education
Massachusetts Institute of Technology
Cambridge, MA
Ph.D. Candidate Department of Chemistry, degree expected spring 2007. Thesis work:
Novel Function and Regulation of Mutagenic DNA Polymerases in Escherichia coli

University of Washington
Seattle, WA
Bachelor of Science with distinction, Chemistry and Biochemistry, minor in Mathematical Physics, June 2001. Thesis: Structural Studies of Mg\(^{2+}\)-bound Calmodulin from S. cerevisiae.

Awards

Research Experience
MIT Department of Chemistry
Cambridge, MA
Advisor: Graham C. Walker, Department of Biology
Novel Function and Regulation of Mutagenic DNA Polymerases in Escherichia coli 2002-2007

MIT Department of Chemistry
Cambridge, MA
Advisor: JoAnne Stubbe
Characterization of Subunit Exchange in Ribonucleotide Reductase from S. cerevisiae 2001-2002

University of Washington Department of Biochemistry
Seattle, WA
Advisor: Rachel E. Klevit
Structural Studies of Mg\(^{2+}\)-bound Calmodulin from S. cerevisiae. 1999-2001

Boston University Department of Biology
Boston, MA
Advisor: Dean R. Tolan
Evolution of Aldolase in Primitive Fish 2000

Teaching Experience
MIT Department of Chemistry
Cambridge, MA
Advanced Biological Chemistry (5.08). Teaching Assistant, Spring 2002.
Advanced Chemical Experimentation and Instrumentation (5.33). Teaching Assistant, Fall 2001

MIT Department of Biology
Cambridge, MA
Member of Extended HHMI Education Group. 2003-2007

University of Washington Early Entrance Program
Seattle, WA
Physics and Mathematics Teaching Assistant, Fall 1998-Spring 1999
Publications


*Equally contributing author

Selected Oral Presentations

