Specificity of DNA Damage Inducible DNA Polymerase IV from *Escherichia coli*

A thesis presented

by

Jason M. Walsh

to

The Department of Chemistry and Chemical Biology

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

in the field of Chemistry

Northeastern University
Boston, Massachusetts

December 2009
© 2009, Jason M Walsh
All rights reserved.
Specificity of DNA Damage Inducible DNA Polymerase IV from Escherichia coli

by

Jason M. Walsh

ABSTRACT OF THESIS

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

DNA polymerases are responsible for DNA replication during cell division. There are multiple families of polymerases (A, B, C, D, X) responsible for copying DNA during replication and repair. There is also a class of polymerases conserved throughout evolution, known as the Y family polymerases, that have reduced replication fidelity on undamaged DNA (Tang et al. 2000). However Y family DNA polymerases have the specialized property of replicating DNA by copying damaged DNA, a process known as translesion synthesis (TLS). Structural differences between Y family and replicative polymerases may account for the difference in enzymatic activity. However we demonstrate that the Klenow fragment (A family) can bypass a fluorescent cytosine analog known as 1, 3-diaza-2-oxophenothiazine (tC), that DinB, a Y family polymerase, cannot bypass. We show that DinB inserts dGTP faithfully, but cannot extend the DNA primer beyond that. Verifying which amino acid residues are responsible for both function and specificity of the Y family polymerases is accomplished by assessing the kinetic data of nucleotide incorporation events of DinB variants as compared to wild-type DinB. In order to determine candidates for residues to alter, we use a sequence alignment based approach as well as by using THEMATICS, a computational methodology developed by the Ondrechen group at Northeastern University that identifies ionizable active site residues. The variants are then tested for activity, lesion specificity and replication fidelity in vitro and in vivo.
ACKNOWLEDGEMENTS

First and foremost I would like to thank Dr. Penny Beuning for her support, her unending thirst for science, and her tireless efforts in doing everything for us in and even out of the lab, I think I can safely speak for everyone when I say that we notice it, we appreciate it immensely, and its rubbing off on us, both as people and scientists, thank you Penny for fighting for me. Secondly I would like to thank Dr. Mary Jo Ondrechen, and Dr. David Budil for their support through this process and for taking their time to read this thesis and give me feedback and constructive criticism to help make me a better scientist. I would also like to thank our Post-Doc: Dr. Jana Sefkicova who routinely puts up with my messes and looks the other way when I “borrow” her supplies and buffers and pipettes, and still takes the time out to teach me anything and everything about the lab. I would also like to thank the Beuning Lab members, especially (in no particular order) Brenna, Michelle, Jaylene, Lisa, Rajal, Mo, Susan, Phillip, Caitlin, Ali, Sri, Kelly, Ramya, Tara, Stephanie, Alicia, Susan, Kylie and especially to both: Khadijah Balfour-Jeffery whose in vivo work I used in this thesis and to Kataryna (Kasia) Wilk who has done many a DinB purification for me. Lastly I would like to thank my family: Jacqueline, Raymond, Carolyn, Jared, Jacob, and Trigger for supporting me through graduate school. I could not have done this without all of you.
# TABLE OF CONTENTS

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

ACKNOWLEDGEMENTS .................................................................................................... 5

List of Figures ...................................................................................................................... 8

List of Tables ..................................................................................................................... 10

List of Abbreviations ....................................................................................................... 11

Chapter 1: Introduction .................................................................................................. 15
  I. DNA Replication ........................................................................................................... 16
  II. DNA Damage and Repair .......................................................................................... 18
  III. Y Family Polymerases and the SOS Response ......................................................... 20
  IV. Y family Polymerase Structure and Function .......................................................... 23
  V. DinB Specificity ........................................................................................................... 24
  VI. Cellular Interactions of DinB ..................................................................................... 26
  VII. Significance ............................................................................................................... 27
  VIII. Summary .................................................................................................................. 27
  IX. References .................................................................................................................. 28

Chapter 2: Determination of DinB Specificity ................................................................ 32
  I. Introduction .................................................................................................................. 32
  II. Materials and Methods .............................................................................................. 35
      a. Expression and Purification of Wild Type DinB and Variants ............................... 35
      b. DNA Synthesis and Purification .......................................................................... 37
      c. Primer Extension Assays ....................................................................................... 40
      d. Nitrofurazone Sensitivity Assays ......................................................................... 41
III. Results .................................................................................................................................41
   a. Activity of Wild Type DinB ..........................................................................................42
   b. Activity and Kinetics of Ser55Thr and Cys66Ser ......................................................44
   c. Leu102lle Kinetics ........................................................................................................46
   d. Ser42Ala Kinetics ..........................................................................................................47
   e. Val40Ala Kinetics .........................................................................................................49
   f. Met57Ala .......................................................................................................................50
IV. Discussion .........................................................................................................................53
V. References ..........................................................................................................................55

Chapter 3: Fluorescent Cytosine Analog Bypass .................................................................57
I. Introduction .........................................................................................................................57
II. Materials and Methods ....................................................................................................60
III. Results ..............................................................................................................................61
   a. Activity of DinB on Template Containing tC Modification .........................................61
   b. Kinetics .........................................................................................................................65
IV. Discussion ........................................................................................................................66
V. References .........................................................................................................................67

Chapter 4: Future Consideration .........................................................................................69
I. Introduction ........................................................................................................................69
II. DinB Variants ....................................................................................................................70
III. Fluorescent Cytosine Analog ........................................................................................70
IV. References .......................................................................................................................71
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>DNA Repair Mechanisms</td>
<td>19</td>
</tr>
<tr>
<td>1.2</td>
<td>Some common lesions observed in DNA</td>
<td>20</td>
</tr>
<tr>
<td>2.1</td>
<td>Homology model of DinB</td>
<td>32</td>
</tr>
<tr>
<td>2.2</td>
<td>Conservative amino acid residue mutations</td>
<td>35</td>
</tr>
<tr>
<td>2.3</td>
<td>Solid phase DNA synthesis</td>
<td>38</td>
</tr>
<tr>
<td>2.4</td>
<td>Addition of the furfuryl lesion to deoxyguanosine triphosphate.</td>
<td>39</td>
</tr>
<tr>
<td>2.5</td>
<td>Activity of DinB on lesion containing DNA at various [dNTP]</td>
<td>42</td>
</tr>
<tr>
<td>2.6</td>
<td>Initial rate of dCTP incorporation by WT DinB vs. concentration of dCTP</td>
<td>43</td>
</tr>
<tr>
<td>2.7</td>
<td>Double reciprocal plot of dCTP incorporation by wild type DinB across from $N^2$-furfuryl dG.</td>
<td>43</td>
</tr>
<tr>
<td>2.8</td>
<td>Activity assessments of Ser55Thr and Cys66Ser with WT DinB as a control.</td>
<td>44</td>
</tr>
<tr>
<td>2.9</td>
<td>Initial rates of incorporation of dCTP by Ser55Thr across from $N^2$-furfuryl-dG</td>
<td>45</td>
</tr>
<tr>
<td>2.10</td>
<td>Double reciprocal plot of dCTP incorporation by Ser55Thr across from $N^2$-furfuryl dG.</td>
<td>45</td>
</tr>
<tr>
<td>2.11</td>
<td>Initial rates of incorporation of dCTP by Leu102Ile across from $N^2$-furfuryl-dG plotted versus concentration of dCTP.</td>
<td>46</td>
</tr>
<tr>
<td>2.12</td>
<td>Double reciprocal plot of dCTP incorporation by Leu102Ile across from $N^2$-furfuryl dG.</td>
<td>47</td>
</tr>
<tr>
<td>2.13</td>
<td>Initial rates versus dCTP concentration of insertion of dCTP by Ser42Ala across from $N^2$-furfuryl dG.</td>
<td>48</td>
</tr>
</tbody>
</table>
Figure 2.14 Double reciprocal plot of initial rate of insertion versus substrate concentration of insertion of dCTP across from $N^2$-furfuryl dG by Ser42Ala. 48

Figure 2.15 Initial rates versus dCTP concentration of insertion of dCTP by Val40Ala across from $N^2$-furfuryl dG. .................................................................49

Figure 2.16 Double reciprocal plot of initial rate of insertion versus substrate concentration of insertion of dCTP across from $N^2$-furfuryl dG by Val40Ala.50

Figure 2.17 Survival assay of E. coli grown with various mutations in DinB. ..............51

Figure 2.18 Met57Ala and WT enzyme activity assay via primer extension with dNTP mix. ..................................................................................................................52

Figure 2.19 Close up of the DinB active site with the position of the residues changed...52

Figure 3.1 Hydrogen bonding between (A) canonical guanosine and cytosine and (B) guanosine and the fluorescent cytosine analog. (Stengel et al. 2007). ..........58

Figure 3.2 Wild-type extension on tCDNA/running start .........................................................61

Figure 3.3 Increase [DinB] using only dGTP.................................................................62

Figure 3.4 Primer extension assay by DinB wild type on MismatchExtensionA,C,G,T...63

Figure 3.5 Wild type DinB extension from Primer22; on tC DNA.. .................................64

Figure 3.6 Plot of initial velocities incorporation versus dGTP concentration.............65

Figure 3.7 Double reciprocal plot of concentration of substrate versus initial velocity of the enzymatic reaction of insertion of dGTP across from tC fluorescent cytosine analog...............................................................66
List of Tables

Table 2.1  Kinetic parameter results calculated for wild type DinB and variants. ..........54

Table 3.1  Kinetic values for the insertion of dGTP across from DNA template
containing tC modified cytosine. .................................................................66
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>percent</td>
</tr>
<tr>
<td>2-AAF</td>
<td>N-2-acetylaminofluorene</td>
</tr>
<tr>
<td>2-AF</td>
<td>aminofluorene</td>
</tr>
<tr>
<td>Ala</td>
<td>Alanine</td>
</tr>
<tr>
<td>Asn</td>
<td>Asparagine</td>
</tr>
<tr>
<td>AP</td>
<td>abasic</td>
</tr>
<tr>
<td>Arg</td>
<td>Arginine</td>
</tr>
<tr>
<td>Asp</td>
<td>Aspartic Acid</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BL21</td>
<td>competent <em>E. coli</em> cells</td>
</tr>
<tr>
<td>BPDE</td>
<td>Benzo[a]pyrene diolepoxide</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>CEdG</td>
<td>$N^2$-(1-carboxylethyl)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>CPG</td>
<td>Controlled pore glass</td>
</tr>
<tr>
<td>Cys</td>
<td>Cysteine</td>
</tr>
<tr>
<td>dA</td>
<td>deoxyadenine</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenine triphosphate</td>
</tr>
<tr>
<td>dC</td>
<td>deoxycytosine</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytosine triphosphate</td>
</tr>
<tr>
<td>dG</td>
<td>deoxyguanosine</td>
</tr>
<tr>
<td>dGTP</td>
<td>deoxyguanosine triphosphate</td>
</tr>
</tbody>
</table>
DMT  dimethoxytrityl
DNA  deoxyribonucleic acid
dNTP deoxynucelotide triphosphate
dT  deoxythymine
dtCTP deoxy1,3-diaza-2-oxophenothiazine triphosphate
dTTP deoxythymine triphosphate

E. coli  Escherichia coli
EDTA  ethylenediaminetetraacetic acid
F  Fluorine
FPLC  fast protein liquid chromatography
G  guanine
Gln  Glutamine
Glu  Glutamic acid
Gly  Glycine
g  g-force
GE  General Electric
HEPES  4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
I  isoleucine
I₂  molecular iodine
Ile  Isoleucine
IPTG  Isopropyl β-D-1-thiogalactopyranoside

\( k_{\text{cat}} \)  catalytic constant; turnover number

\( K_D \)  dissociation constant
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>KF</td>
<td>Klenow fragment</td>
</tr>
<tr>
<td>$K_m$</td>
<td>Michaelis-Menten constant</td>
</tr>
<tr>
<td>Leu</td>
<td>leucine</td>
</tr>
<tr>
<td>L</td>
<td>liter</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>mer</td>
<td>unit length of nucleic acid</td>
</tr>
<tr>
<td>Met</td>
<td>methionine</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>Mg</td>
<td>Magnesium</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mL</td>
<td>milliliter</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NFZ</td>
<td>nitrofurazone</td>
</tr>
<tr>
<td>NH₄</td>
<td>ammonium</td>
</tr>
<tr>
<td>nM</td>
<td>nanomolar</td>
</tr>
<tr>
<td>nm</td>
<td>nanometers</td>
</tr>
<tr>
<td>OAc</td>
<td>acetate</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenyl methyl sulfonyl fluoride</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
</tbody>
</table>
pol  polymerase
Q  Glutamine
RNA  ribonucleic acid
S  Serine
SDS-PAGE  sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser  Serine
SO$_4$  sulfate
SSB  single stranded DNA binding protein
T  threonine
tC  1,3-diaza-2-oxophenothiazine
TCA  trichloroacetic acid
Thr  Threonine
TLS  translesion synthesis
UV  ultraviolet
V  Valine
Val  Valine
$V_{\text{max}}$  maximum velocity
$V_0$  initial velocity
wt  wild-type
$\mu$M  micromolar
Chapter 1: Introduction

Virtually all organisms have at least one member of the superfamily of DNA polymerases known as the Y family polymerases. The Y family polymerases are important to DNA damage tolerance mechanisms because they facilitate replication of damaged DNA, rescuing stalled replication forks in order to transmit genetic information to the next generation. The goal of this research is to assess both the replication fidelity of *E. coli* Y family polymerase DinB (DNA polymerase IV) and the range of DNA damage that DinB bypasses, in the case of both wild-type DinB and variants. This work will help determine the amino acid residues that are essential to conferring the ability of DinB to bypass specific DNA lesions. The active site of DinB, and of other related Y family DNA polymerases, is considered to be more flexible and solvent accessible in order to accommodate either lesions which may be large and bulky or to accommodate more than one nucleotide in the case of thymine-thymine cyclobutane dimers. The crystal structures of Y family polymerases that have been determined demonstrate that they adopt a right hand shape that is similar overall to that of other DNA polymerases. However, the primary structure shows little conservation between Y family DNA polymerases and those in other families (Ling et al. 2001). We are using a dual approach to identify candidate amino acid residues to change in order to test their function. First, the THEMATICS (theoretical microscopic titration curves (Ondrechen et al. 2001)) program developed by Ondrechen et al. was used to identify active site and other residues that contribute to translesion synthesis activity. THEMATICS is a computational method that calculates theoretical titration curves of ionizable residues. Residues that display non-
Henderson-Hasselbalch behavior are likely to be important for activity or specificity (Ondrechen et al. 2001). Therefore we have chosen to mutate these amino acids in order to elucidate residues that are directly responsible for activity. Second, we are using multiple sequence alignments to identify residues that are conserved throughout evolution and therefore are likely to be important for activity or specificity. Both of these methods have been used to identify amino acid residues to change in order to test their role in DinB activity.

I. DNA Replication

DNA replication is carried out by a complex molecular machine which includes a variety of DNA polymerases, subdivided into families that have specific features necessary for DNA replication under a variety of conditions. In addition to DNA polymerase, there are other enzymes involved in replication such as: helicase, which unwinds a portion of the DNA double helix; RNA Primase, which synthesizes RNA primers on the template strands; DNA polymerase I, which removes the RNA primers and replaces them with DNA; and DNA ligase, which forms the phosphodiester bond at nicks in the backbone. The proofreading nuclease removes incorrectly placed nucleotides from the daughter strand.

The *E. coli* replisome is the protein complex that replicates DNA as the cell prepares for division. The DNA synthesis activity of the replisome is provided by the DNA polymerase III (pol III) core, which consists of α, ε, and θ subunits (total MW 166 kDa). Due to the requirement to replicate both strands of DNA simultaneously, two α polymerase subunits are required, one for the leading strand and one for the lagging strand. The ε subunit of the core possesses 3’-5’ exonuclease activity (Pomerantz et al.
The two catalytic cores of pol III are connected by the ATPase clamp loader, which loads the β processivity clamp onto DNA. The polymerase itself is only weakly processive, catalyzing the incorporation of approximately 100-200 nucleotides per binding event (Pomerantz et al. 2007). A processivity factor known as the β sliding clamp, a homodimer of MW 40.6 kDa subunits, increases the processivity of pol III to 1000-2000 bases (Pomerantz et al. 2007). The five-subunit clamp loader complex assembles sliding clamps onto DNA (Pomerantz et al. 2007). The clamp loader also links the leading and lagging strand polymerases, acting as a bridge between the two. The clamp loader binds to the helicase in order to couple double stranded DNA unwinding with DNA synthesis (Pomerantz et al. 2007). Single stranded binding protein (SSB) prevents the formation of secondary structures that could disrupt DNA replication (Pomerantz et al. 2007).

Pre-steady state chemical kinetics have established that the actual chemistry of nucleotide addition to a growing DNA strand occurs in a series of steps: [1] initial binding of the polymerase to DNA of length n nucleotides; [2] binding of the incoming nucleotide to form the DNA-polymerase-dNTP complex; [3] isomerization of the ternary complex from inactive to active; [4] phosphoryl group transfer and nucleophilic attack of the α-phosphate group on the strand by the 3′ hydroxyl group of the primer strand (Wong et al. 1991); [5] a second isomerization step from inactive to active form; [6] pyrophosphate release; and finally, [7] release of the DNA_{n+1} final product (Stengel et al. 2007). There is also sufficient data to suggest that the other families of DNA polymerases carry out their reactions in a very similar scheme, although the rates of each step can vary significantly (Ahn et al. 1997; Washington et al. 2001; Fiala et al. 2004; Fiala et al. 2007).
In general, DNA replication by replicative DNA polymerases occurs with very high fidelity (1 error in approximately $10^7$ bases replicated not including the contributions of proofreading and mismatch repair mechanisms).

The Y family DNA polymerases, notably *E. coli* pol IV and pol V, have the specialized ability to copy damaged DNA, albeit with a high error rate on undamaged DNA. Pol V has an error rate of approximately 1 in $10^3$ to $10^4$ base pairs, with pol IV (DinB) being about 5-10 fold more accurate (Tang et al. 2000) when copying undamaged templates. The misincorporation and mismatch extension efficiencies for pol IV were determined to range from $10^{-3}$ to $10^{-5}$ (Kobayashi et al. 2002). Gap-filling analyses show that small frameshifts account for a large percentage of pol IV errors, but the remainder of errors are divided among other base substitutions with the most common being T $\rightarrow$ C, and either T $\rightarrow$ G or T $\rightarrow$ A (Kobayashi et al. 2002). Pol IV nucleotide misincorporation accuracies are larger at pyrimidine template sites with purine template sites being about 5-10 fold worse (Kobayashi et al. 2002).

II. DNA Damage and Repair

Cells have evolved multiple mechanisms to deal with DNA that has lost its integrity due to exposure to damaging agents. The focus here is on the SOS response and translesion synthesis via the Y family of DNA polymerases. However, the canonical DNA repair systems will be briefly discussed. In base excision repair, a single base is removed by a DNA glycosylase, and a short patch of new DNA is synthesized (Fortini et al. 2007). In nucleotide excision repair, a group of nucleotides are removed, whereas in mismatch repair a mismatched, but usually undamaged, base pair is removed (Friedberg 2006). Homologous recombination involves the exchange of DNA between two
homologous DNA molecules (Kowalczykowski et al. 1994). These are systems in which the damage is directly removed from the DNA, by excising the damaged nucleotide or fully restoring the area including the damaged base, and replacing it with the correct complementary DNA (Friedberg 2006). These are very important to maintaining DNA integrity, but DNA damage tolerance pathways also play an important role in cell survival (Figure 1.1). It has been suggested that the DNA damage tolerance mechanisms may have fueled selection pressure for sexual transmission of genetic material (Friedberg 2006).

**Figure 1.1** DNA Repair Mechanisms. Excision repair, in which the incorrect nucleotide is removed, homologous recombination repair, and DNA damage tolerance via translesion synthesis.
DNA damage can come from a variety of sources such as changes in the nitrogenous bases that occur spontaneously, including deamination (Figure 1.2), or the environment (toxins, UV light from the sun, ingested foods and preservatives); metabolism of said toxins; general metabolism; as well as ubiquitous oxygen free radical species.

Figure 1.2 Some common lesions observed in DNA. (A) Deamination of nitrogenous bases in DNA (B) DNA adducts from environment or metabolism (Friedberg 2006).

III. Y Family Polymerases and the SOS Response

DNA replication is a key biochemical process for the maintenance and transmission of genetic information. Passing genetic material to the next generation is the desired outcome of virtually all Darwinian selection pressure. DNA is constantly subject to damage that requires DNA repair in order to restore its information content. However,
mutagenesis of the DNA due to the lowered fidelity of Y family polymerases may also play an important role in evolution.

The Y family polymerases are a superfamily of polymerases that are found in all domains of life (Ohmori et al. 2001). The DinB subfamily is conserved throughout evolution. They have an important function in rescuing stalled replication forks, so that DNA replication can continue even in the presence of distorting lesions. These properties demonstrate the importance of the Y family DNA polymerases to the cell.

The process of TLS was first described roughly 30 years ago (Radman 1975). It was observed that DNA damage induces the SOS response with accompanying mutagenesis of the DNA (Radman 1975). DNA damage-induced mutagenesis was thought to result from the modification of normal DNA polymerases, which allowed them to bypass the DNA damage. It was also observed that under these conditions, DNA is synthesized with lower fidelity due to either a modified conformation of replicative DNA polymerase or some other unknown mechanism (Woodgate et al. 1989). By the late 1990s there was substantial evidence that Rev1, the UmuC/UmuD′ complex, and DinB were DNA polymerases that could carry out SOS-induced TLS (Yang 2005).

Many components of this DNA damage induced SOS response are well understood (Friedberg 2006; Jarosz et al. 2007; Schlacher et al. 2007). Polymerase III cannot bypass damaged nucleotides and fails to continue replicating the DNA, resulting in the appearance of a long region of single-stranded DNA. This is likely due to the helicase being uncoupled from the replisome as it continues to unwind the DNA downstream of the stalled polymerase. RecA then polymerizes on the exposed single stranded DNA to form the RecA/ssDNA nucleoprotein filament. LexA binds to the
RecA/ssDNA nucleoprotein filament and autolytically cleaves, inactivating itself as a repressor and leading to the expression of the SOS genes. There are at least 57 genes in the SOS regulon, which includes those coding for the Y family polymerases DinB and UmuD’c (Simmons 2008). DinB inserts a base across from the lesion, and replicates several nucleotides at potentially a lower fidelity and higher mutagenic cost to the cell. The benefit of rescuing the stalled replication fork apparently outweighs the cost of mutagenesis. This mutagenic cost is about a 100-fold increase in the rate of base substitution mutations at the site of DNA damage, but can also include untargeted mutagenesis of undamaged DNA downstream from the lesion (Tang et al. 2000).

In addition to induction by DNA damage, the SOS response can also be induced by β-lactam antibiotics. Transcription of the dinB gene is induced by the inhibition of cell wall biosynthesis at different levels, independent of the LexA repressor/RecA regulatory system (Perez-Capilla et al. 2005). This suggests that transcription of the dinB gene can be considered a general stress response mechanism (Perez-Capilla et al. 2005). This is also seen in dinB-dependent increased mutability in starving, non-dividing cells in a phenomenon known as adaptive mutagenesis. In general E. coli lacking SOS induced polymerases are less viable when grown in the presence of wild-type E. coli (Foster 2000; McKenzie et al. 2001; Kobayashi et al. 2002; Yeiser et al. 2002). Furthermore, the increased mutagenic cost of DNA replication by DinB may be a contributing factor to antibiotic resistance.

The dinB gene product was first demonstrated to be a DNA-dependent DNA polymerase in 1999 (Wagner et al. 1999). DinB was found to lack intrinsic 3’-5’ exonuclease proofreading activity and to exhibit distributive synthesis. DinB is able to
accommodate misaligned or bulged primer-template structures (Wagner et al. 1999). It was also determined that DinB shared a high degree of sequence homology with *E. coli* UmuC and *Saccharomyces cerevisiae* Rev1 (Ohmori et al. 1995). These enzymes were demonstrated to possess translesion synthesis activity and were placed in a family, then known as the UmuC-like superfamily, which became the Y family of DNA polymerases (Wagner et al. 1999).

**IV. Y family Polymerase Structure and Function**

There are two Y family polymerases in *E. coli* and four in mammalian cells; homologs of UmuC and DinB are found in other eubacteria as well as archaea (Yang 2005). Structurally the Y family polymerases have five conserved sequence motifs but their overall length can vary considerably (Yang 2005). Homologs from higher order organisms tend to be twice as large as their bacterial and archaeal counterparts, and Y family polymerases as a whole have not been demonstrated to have intrinsic 3′ → 5′ exonuclease activity.

The domains of both replicative and Y family DNA polymerases are metaphorically named for the parts of a right hand, including the thumb, palm, and finger domains. The Y family polymerases are characterized by rather small finger and thumb domains which result in an open and solvent-accessible active site in the palm domain (Yang 2005). The Y family specific ‘little finger’ domain seems to play an important role in substrate selection and DNA binding (Boudsocq et al. 2004; Beuning et al. 2006).

The Y family active site is considerably different from that of the replicative polymerases. The active site of replicative polymerases contains an ‘O-helix’ that only
allows a correct base pair to be formed. The Y family polymerases lack the O-helix, and the active site of these polymerases is generally considered to be more flexible and open in order to accommodate lesions on the DNA strand that the replicative polymerases cannot bypass. A crystal structure of *E. coli* DinB has yet to be solved, however several homology models have been constructed based on the crystal structures of homologs from *Sulfolobus solfataricus* (Dpo4) and *Sulfolobus acidocaldaricus* (Dbh).

A structure of Dpo4 in complex with DNA containing a thymine – thymine (T-T) cyclobutane pyrimidine dimer showed that both thymines could be accommodated in the active site (Ling et al. 2003). This study showed that there was a difference in the efficiency and fidelity of adding a base across from the first T of the dimer compared to addition across from the second. Dpo4 replicates past the 3′ T of the T-T dimer with modest efficiency, whereas it replicates past the second base faster, and it is more likely to add a deoxyadenosine triphosphate across from it than the first thymine in the cyclobutane dimer (Ling et al. 2003).

V. DinB Specificity

DNA is subject to a variety of insults from environmental sources, endogenous metabolic byproducts and spontaneous changes (Figure 1.2). DNA adducts and deaminated bases are just a few of the common examples of damage. An abasic (AP) site, in which the base is missing can also be detrimental to DNA synthesis. DinB has been shown to bypass AP sites *in vitro* generating -1 frameshift mutations on normal template DNA (Kobayashi et al. 2002).

The potent mutagenic lesion benzo[a]pyrene, a large cyclic hydrocarbon, provides one of the first instances of a link between an environmental toxin and cancer that was
discovered (Butlin 1892). Benzo[a]pyrene is found in cigarette smoke, overcooked or charred animal food products, and is a byproduct of incomplete combustion. The presence of DinB significantly increases bypass of the benzo[a]pyrene adduct (Seo et al. 2006). Furthermore DinB discriminates against some isomers of BPDE, suggesting that even subtle changes in the conformation of adducts can alter how they are processed by Y family polymerases.

DinB has been shown to bypass other adducts as well. For example, DinB bypasses $N^2$-(1-carboxylethyl)-2′-deoxyguanine ($N^2$-CEdG) (Yuan et al. 2008). This adduct is a common byproduct of glycolysis, occurring in one in $10^7$ bases in melanoma cells. Both human pol κ (the DinB ortholog in eukaryotes) and DinB insert dCTP opposite $N^2$-CEdG accurately and efficiently (Yuan et al. 2008). The properties of lesion bypass by DinB do not always correlate with those of pol κ. The $C^8$-dG adducts of N-2-acetylaminofluorene (2-AAF) and aminofluorine (2-AF), which are prototypical aromatic amides and carcinogens (Suzuki et al. 2001), were bypassed efficiently and accurately by human Pol κ, but these adducts effectively blocked insertion and extension by DinB (Suzuki et al. 2001). DinB has also been shown to bypass certain DNA-protein cross linking adducts that form in DNA following exposure to acrolein, but can also occur during routine nucleic acid metabolism (Minko et al. 2008). Acrolein is a potent toxin, and has tumor initiating properties, but it is also an endogenous byproduct of fatty acid metabolism (Minko et al. 2008). The $N^2$-dG adduct of acrolein is $\gamma$-hydroxypropano-deoxyguanosine ($\gamma$-HOPdG) and DinB inserts dCTP across from this lesion (Minko et al. 2008). However there were some cases of misincorporation events of dATP and dTTP across from the lesion DNA (Minko et al. 2008).
VI. Cellular Interactions of DinB

DinB is conserved in all domains of life, and is the most abundant polymerase in *E. coli* during times of DNA damage induced stress (Godoy et al. 2007). It has been shown that an important role of DinB is to bypass efficiently via translesion synthesis $N^2$-deoxyguanine adducts (Jarosz et al. 2006). However DinB also increases the frequency of deleterious -1 frameshift mutations during mutagenesis (Brotcorne-Lannoye et al. 1986; McKenzie et al. 2001). These two properties seem to be counterproductive to one another. DinB is physically regulated through protein-protein interactions with UmuD, UmuD' and RecA, which were all previously known to regulate UmuC, another Y family polymerase (Godoy et al. 2007). It appears that the deleterious -1 frameshift mutator activity of DinB is a result of the number of molecules of DinB present in the cell in relation to the amount of UmuD present (Godoy et al. 2007). This regulation may explain the dual nature of the polymerase activity of DinB, which accurately bypasses bulky $N^2$-dG adducts, but is also responsible for highly mutagenic -1 frameshift mutations. DinB also has been shown to interact with the processivity clamp and the specific residues that are involved in this interaction have been identified as $^{346}$GlnLeuValLeuGlyLeu$^{351}$ (Lenne-Samuel et al. 2000). DinB containing a mutation of this sequence showed a decrease in β-clamp stimulation *in vitro*, while *in vivo* it could not perform TLS on a plasmid containing a benzo[a]pyrene adduct. This indicates the clamp has a major role *in vivo* for processing of the substrate by DinB (Lenne-Samuel et al. 2000; Wagner et al. 2000).
VII. Significance

DNA damage can lead to gene mutations and carcinogenesis. The mutagenic function of the Y family polymerases also contributes to resistance to β-lactam antibiotics, which can induce the SOS response without causing DNA damage. Early work suggests that mutagenesis is an active process involving the Y family DNA polymerases that bypass chromosomal lesions. DNA damage would normally stall DNA replication by interfering with the replicative polymerases. However the Y family polymerases copy lesion-containing DNA and replication continues (Jarosz et al. 2007). This raises the question of whether Y family polymerases are “unfaithful copiers or specialized polymerases” (Jarosz et al. 2007), which is still a subject of much debate. The Y family polymerases are conserved throughout nature, which demonstrates the evolutionary importance of chromosomal replication integrity, and the mechanisms to process DNA damage. From an evolutionary standpoint it is beneficial to have the ability to replicate DNA during times of environmental and chemotoxic stresses, even at a potentially mutagenic cost.

VIII. Summary

This work seeks to understand the ability of DinB, a specialized Y family polymerase, to both insert and extend DNA across from different lesions. Through DNA replication assays we demonstrate differences in the kinetics of DinB containing single mutations to insert dCTP across from N²-furfuryl-dG (Chapter 2). There is precedence for single amino acid changes to affect DinB’s activity and here we examine additional amino acids predicted to be important for lesion bypass (Wagner et al. 1999; Jarosz et al. 2006).
DinB is shown to be inefficient in completing TLS with a DNA modification in the form of a cytosine analog; 1,3-diaza-2-oxophenothiazine (tC) (Stengel et al. 2007; Sandin et al. 2009). This modification has been shown to be efficiently bypassed, and its triphosphate version also incorporated into DNA, by the Klenow fragment, which is an A family polymerase. The authors speculate the reason for this ability is the flexibility of the KF active site (Sandin et al. 2009), but the Y family polymerases have an open and flexible active site, which accounts for their innate ability to accommodate bulky adducts in their active site. However we show that even though KF can both bypass and incorporate this modified base effectively, DinB is unable to bypass tC (Chapter 3). The findings presented here add to the evidence that the Y family polymerases have cognate substrates and contribute to our understanding of their specificity.

IX. References


Chapter 2: Determination of DinB Specificity

I. Introduction

DNA is constantly subjected to damage, including: thymine-thymine cyclobutane pyrimidine dimer induced by exposure to UV light, the benzo[a]pyrene adduct at $N^2$ deoxyguanosine from an environmental toxin such as cigarette smoke or combustion engine exhaust, or a furfuryl adduct also in the $N^2$ position of dG from an internal source such as the metabolism of RNA. Replicative polymerases usually cannot copy over a lesion in the DNA template strand. The damage stalls the DNA replication machinery leading to the induction of the SOS response.

Figure 2.1 Homology model of DinB (Jarosz et al. 2006). Residues in the DinB active site predicted to be important for functions based on calculations by THEMATICIS (A) (Ondrechen et al. 2001), and evolutionarily conserved alignment are shown with color coded highlighted residues (B). Template DNA is shown in blue, primer is shown in red and incoming nucleotide is shown colored by atom.

*E. coli* DinB (polymerase IV) is a Y family polymerase that is conserved throughout all domains of life and is capable of copying damaged DNA templates (Wagner et al. 1999; Ohmori et al. 2001). In particular, DinB has a 15-fold preference to
bypass \(N^2\)-furfuryl-dG in comparison to undamaged DNA (Jarosz et al. 2006). It has been shown that a single amino acid change eliminates the translesion synthesis activity of \(E.\ coli\) DinB (Jarosz et al. 2006). DinB Phe13Val is unable to bypass \(N^2\)-furfuryl dG, but is proficient for DNA synthesis on undamaged DNA (Jarosz et al. 2006). Mammalian pol \(\kappa\) has shown the same preference to insert dC across from the lesion indicating this function is conserved evolutionarily (Jarosz et al. 2006).

Based on the structure of Dbh (\(Sulfolobus acidocaldarius\)), a DinB ortholog, it was hypothesized that a specific ‘pocket’ in the enzyme is involved in the accommodation of the lesion in the active site (Jarosz et al. 2006). Phenylalanine 13 lies in this pocket and is known as the steric gate residue. Ribonucleotide discrimination is a function of the steric gate found in DNA polymerases. The steric gate prevents the incorporation of ribonucleotides by blocking nucleotides with a 2’OH from being incorporated. The steric gate residue of Y family DNA polymerases is most frequently tyrosine or a phenylalanine (DeLucia et al. 2006; Jarosz et al. 2006; Shurtleff et al. 2009). Mutation of the steric gate residue in DinB had a marked effect on the ability of DinB to discriminate against ribonucleotides, increasing the frequency of misincorporation from \(<10^{-5}\) to \(10^{-3}\) (Jarosz et al. 2006). Another important residue near the steric gate is tyrosine at position 79 which is conserved among all orthologous DinBs. Tyr 79 is hypothesized to regulate the function of the steric gate. Mutations at this position prevent DinB from extending more than a few nucleotides beyond a lesion and result in extreme cellular sensitivity to nitrofurazone (Jarosz 2009). Previous mutation frequency studies on other DinB variants have been done (Wagner et al. 1999) and single point mutations have an effect on the frequency of mutations as compared to wild-type DinB. Cells expressing
wild-type DinB had a mutation frequency reported of $68,466.5 \times 10^7$ (Wagner et al. 1999). Mutations Asp103Asn, Asp8His, Arg49Phe, Asp8Ala, Arg49Ala, Asp103Ala, and Glu104Ala led to much lower mutation frequencies of 21.2, 37.1, 19.6, 18.4, 18.4, 51.3 and 80.9 (all $\times 10^7$), respectively (Wagner et al. 1999). As demonstrated in the past, single mutations can have a large effect on the activity of DinB.

The main goal of this research is to determine which of the amino acid residues of DinB are important for its ability to copy damaged DNA. We are using the THEMATICS program developed by Ondrechen et al., as well as multiple sequence alignments in order to determine which residues are the most likely to be important for efficiency and specificity (Figure 2.1) (Ondrechen et al. 2001). Both of these methods have been used to identify amino acid residues to change in order to test their roles in activity and specificity. Alanine scanning mutagenesis was performed as an initial test of the role of each residue. Additionally, conservative amino acid substitutions were made at some positions in DinB (Figure 2.2). The ability of the resulting DinB variants to incorporate nucleotides across from $N^2$-furfuryl-deoxyguanosine has been determined.
II. Materials and Methods

a. Expression and Purification of Wild Type DinB and Variants

Site directed mutants of DinB were constructed using a QuikChange kit (Stratagene) and the DinB expression plasmid, pDFJ1 (Beuning et al. 2006). The mutants were subsequently confirmed by DNA sequence analysis (Mass. General Hospital Core Facility, Cambridge, MA). *E. coli* BL21 competent cells were then transformed with the respective plasmids for expression. The bacterial cells were cultured in LB broth supplemented with ampicillin (100 mg/mL) overnight. The starter culture was used to inoculate 1 L of LB broth supplemented with ampicillin (100 mg/mL), which was incubated while shaking at 37 °C until its optical density at 600 nm reaches between 0.80 and 1.00. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final
concentration of 1 mM to induce the cells to express the protein of interest. The cells were then transferred to a 30 °C incubator with shaking for 3 hours. Expression of the protein was monitored by SDS-PAGE. The cells were harvested by centrifugation at 6150 \( \times \) g for 10 min, the pellet was collected and stored at -80 °C at least overnight. The cells were then thawed on ice, at 4 °C overnight in Sa buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM \( \beta \)-mercaptoethanol) with EDTA free protease inhibitor cocktail tablet (Roche) and 10 μg/mL PMSF. The cells were lysed with lysozyme and treated with DNase (Beuning et al. 2006). The cells were then carried through one freeze-thaw cycle and the cell debris was removed by centrifugation at 12,300 \( \times \) g at 4 °C for 1 h. The supernatant contains the desired protein that is then purified by using fast protein liquid chromatography (FPLC). It is possible that these mutants are contaminated with small amounts of wild-type DinB.

The protein was purified first on MonoS ion exchange column (GE Healthcare) in buffers Sa (50 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM \( \beta \)-mercaptoethanol) using a gradient to 100% Sb (50 mM HEPES, pH 7.5, 1 M NaCl, 2 mM \( \beta \)-mercaptoethanol) over 5 column volumes. DinB elutes as the only peak in the gradient and the fractions are routinely assayed via SDS-PAGE. Those fractions which contain the protein were combined, and mixed in a 1:1 ratio with buffer PSa (50 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM \( \beta \)-mercaptoethanol, 1 M (NH\(_4\))\(_2\)SO\(_4\)) and were then loaded onto a phenyl sepharose column (GE Healthcare) for further separation. The separation occurs in a gradient to 100% buffer PSb (50 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM \( \beta \)-mercaptoethanol) over 5 column volumes. The protein is then dialyzed two times against a buffer of 50 mM HEPES, pH 7.5, 100 mM NaCl, 10% glycerol, 2 mM \( \beta \)-
mercaptoethanol, for at least two hours each. The protein concentration was determined by using the Bradford assay. The volume and concentration obtained in DinB purifications are generally in the micromolar range so concentrating the fractions is not necessary if the protein is to be used in primer extension assays. DinB is stored in single-use aliquots at -80 °C.

b. DNA Synthesis and Purification

Unmodified DNA and DNA containing adducted bases were synthesized at Northeastern University by the Beuning group and the Rozners group. First, $N^2$-F-dG is site-specifically incorporated into a growing chain of DNA via solid phase DNA synthesis (Figure 2.3). Then the base was modified with furfurylamine, with fluorine as the leaving group, according to published procedures (Figure 2.4) (DeCorte et al. 1996). A series of deprotection steps completes the synthesis (data not shown).
Figure 2.3 Solid phase DNA synthesis. The base 2-F-dG with a fluorine that will act as a leaving group is incorporated in order to add the lesion (DeCorte et al. 1996).
Figure 2.4 Addition of the furfuryl lesion to deoxyguanosine triphosphate (Sproat 1995).

The synthesized DNA is gel purified by polyacrylamide gel electrophoresis, visualized by UV shadowing, and excised from the gel. The DNA is extracted from the gel by crushing the gel slice in a Poly-Prep Chromatography column (BioRad) with 4 mL “crush and soak” buffer (0.5 M ammonium acetate; 0.1 M EDTA pH 8). It is incubated overnight with shaking at 4 °C. The liquid is drained by gravity through the column, and 8 mL of 100% ethanol plus 400 μL 3 M sodium acetate is added and DNA is precipitated at -20 °C overnight. It is then spun down at 8200 x g for 40 min at 4 °C at which point a pellet is apparent. The ethanol is decanted and 1 mL 80% ethanol is added to wash the pellet. The pellet is dried by centrifugation under vacuum. It is then reconstituted in
double distilled Milli-Q water that has been sterile filtered and autoclaved. The concentration is determined via UV absorbance at 260 nm.

c. Primer Extension Assays

We performed primer extension assays on a 61-mer DNA template

5’GGTTACTCAGATCAGGCCTGCGAAGACCTXGGCGTCCGGCTGCAGGTACTATCATATGC3’

Where X is either dG or N2-furfuryl-dG. The primers include:

Running Start (30 mer) 5’GCATATGATAGTACAGCTGCAGCCGGACGC3’
Standing Start (31 mer) 5’GCATATGATAGTACAGCTGCAGCCGGACGC3’

Running start primer allows incorporation of one nucleotide before the lesion and was used in the cases where primer extension activity was being assessed. Standing start primer, in which the first nucleotide incorporated is opposite the lesion, was used in the instances where the kinetics of incorporation of a single nucleotide was measured.

Running start reactions are used with a mix of all four dNTPs to determine the ability of DinB to bypass the lesion of interest and extend to the end of the template. Standing start primer reactions are used to determine the ability of DinB to add a specific nucleotide (in this case dC opposite dG or N2-furfuryl-dG) across from the lesion. The DNA primer and DNA template are combined to a final ratio of 1:1 (10 nM) and were annealed in an annealing buffer (20 mM HEPES pH 7.5; 5 mM Mg(OAc)2), heated for 2 min at 95 °C, incubated at 50 °C for 50 min, and then cooled to 37 °C.

The reactions are carried out in a reaction buffer containing final concentrations of 10 nM 32P-labeled primer/template in 30 mM HEPES pH 7.5, 20 mM NaCl, 7.5 mM MgSO4, 2 mM β-mercaptoethanol, 1% BSA and 4% glycerol. A zero point is taken out
prior to addition of dNTP and reactions are initiated by the addition of dNTP (Beuning et al. 2006). Reaction volumes are 30 μL, though at times can be doubled. Time points are typically taken up to 30-60 min and are quenched with 85% formamide, 50 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue (Beuning et al. 2006). Products from the quenched reactions are analyzed on denaturing (8 M urea) 10-12% polyacrylamide gels, which are subsequently imaged on a Molecular Dynamics storage phosphor imaging screen. Data are analyzed using ImageQuant software (GE Healthcare) (Beuning et al. 2006). This allows for the analysis of kinetic data and determination of $V_{\text{max}}$ and $K_M$ by assessing the percentage of primer extended at various time points taken in the experiment. It also allows for the examination of the fidelity of DinB variants.

d. Nitrofurazone Sensitivity Assays

DinB confers cellular resistance to nitrofurazone (Jarosz et al. 2007). In vivo assays for nitrofurazone (NFZ) sensitivity are performed on LB-agar plates supplemented with 100 μg/mL ampicillin and NFZ at varying concentrations. For in vivo experiments, the DinB-encoding plasmid pYG768 is used; pWSK30 is used as the vector control (Kim et al. 1997). Large LB-agar plates are poured with final concentrations of NFZ between 0 and 3.0 μg/mL. AB1157 ΔdinB competent cells are transformed with the plasmid, serially diluted in 0.85% saline solution and plated. Plates were incubated overnight at 37 °C and the colony forming units were counted to assess survival.

III. Results

We are making conservative substitutions as well as single alanine substitutions at DinB residues shown in Figure 2.1 and assaying them for activity and specificity. All residues shown are highly conserved or predicted by THEMATICS to be important for
activity or specificity. Leu100Ile is a conserved residue in virtually all DinB orthologs. Finally, we are constructing a single-cysteine DinB variant by mutating each solvent-accessible cysteine to serine, and assaying the resulting proteins for lesion bypass activity.

a. Activity of Wild Type DinB

Activity of DinB and the site-directed variants was determined via primer extension assays as described above. We first determined the appropriate DinB concentration to use in assays of its activity in bypassing the $N^2$-furfuryl-dG lesion. DinB demonstrates the ability to bypass the lesion at a range of dNTP (Figure 2.5). $V_o$ was found to be in the linear range when the concentration of DinB was between 0.5 nM to 2 nM, as the higher concentrations were virtually saturated upon addition of dNTPs, which was kept constant at 0.5 mM for this analysis.

![Figure 2.5](image.png)

**Figure 2.5** Activity of WT DinB on lesion containing DNA at various dNTP concentrations
We assayed incorporation of dCTP opposite $N^2$-furfuryl-dG by wild-type DinB (Figures 2.6 and 2.7). By varying the concentration of dCTP, we were able to determine the kinetic constants $k_{\text{cat}}$ and $K_M$ (Table 2.1).

**Figure 2.6** Initial rate of dCTP incorporation across from $N^2$-furfuryl-dG by wild-type DinB as a function of dCTP concentration. The range of dCTP concentration was varied from 0.5 μM to 1.0 mM.

**Figure 2.7** Double reciprocal plot of dCTP incorporation by WT DinB across from $N^2$-furfuryl dG.
b. Activity and Kinetics of Ser55Thr and Cys66Ser

The ability of DinB Ser55Thr and DinB Cys66Ser to bypass the \( N^2 \)-furfuryl dG lesion was determined by using primer extension assays. The Ser55Thr mutant was able to bypass the lesion, but not as efficiently as wild-type DinB, which fully extended the primer by 10 minutes. The Cys66Ser mutant demonstrated virtually no activity based on its lack of extension of the primer, and thus was not assayed by single nucleotide incorporation.

![Figure 2.8](image)

**Figure 2.8** Primer extension activity of Ser55Thr and Cys66Ser with wild-type DinB as a control. [DinB] 25 nM, [dNTP] 1 mM. Time points (min) vary in these experiments with WT: 0,0.5,1,2,10; Ser55Thr: 0,0.5,1,2,5,10; and Cys66Ser 0,0.5,1,2,10.

The kinetics of incorporation of dCTP across from \( N^2 \)-furfuryl dG by the mutant Ser55Thr were assessed as described above. The initial rates were plotted against the concentration of dCTP, and only the linear portion of the rate data was used (Figure 2.9). The parameters \( k_{cat} \) and \( K_M \) are reported in Table 2.1.
Figure 2.9 Initial rates of incorporation of dCTP by Ser55Thr across from $N^2$-furfuryl-dG as a function of dCTP concentration.

Figure 2.10 Double reciprocal plot of dCTP incorporation by Ser55Thr across from $N^2$-furfuryl dG.
c. Leu102Ile Kinetics

Leu102Ile does not fall into the residues listed as either a THEMATICS positive residue or an alignment based residue of interest. However this residue is conserved among all DinBs and it is in close proximity to the amino acid residues implicated by these methods, as well as to the incoming nucleotide (Figure 2.19). It is interesting in that the activity of DinB is seemingly enhanced by the simple change of the position of the methyl group in isoleucine compared to leucine.

**Figure 2.11** Initial rates of incorporation of dCTP by DinB Leu102Ile across from \(N^2\)-furfuryl-dG plotted as a function of dCTP concentration.
Figure 2.12 Double reciprocal plot of dCTP incorporation by DinB Leu102Ile across from \(N^2\)-furfuryl dG.

**d. Ser42Ala Kinetics**

DinB Ser42Ala was assayed to determine the kinetics of dCTP incorporation. Exchanging a serine with an alanine leads to a loss of a hydroxyl group that may be important in hydrogen bonding with the DNA or the dNTP substrates, or another amino acid residue on the protein that may become closer in proximity to position 42 during a conformational change as DinB binds DNA. The initial rate of insertion is plotted against the concentration of dCTP (Figure 2.13), and the kinetic constants \(K_m\) and \(V_{max}\) were derived from the double reciprocal plot (Figure 2.14).
Figure 2.13 Initial rates of insertion of dCTP across from $N^2$-furfuryl dG by DinB Ser42Ala as a function of dCTP concentration.

Figure 2.143 Double reciprocal plot of initial rate of insertion versus substrate concentration of insertion of dCTP across from $N^2$-furfuryl dG by DinB Ser42Ala.
e. Val40Ala Kinetics

Val40Ala is a residue that was identified through the alignment based method of identifying possible important DinB residues and was mutated to alanine. The kinetics of incorporation of dCTP across from $N^2$-furfuryl dG by DinB Val40Ala were assessed as described above. The initial rates were plotted as a function of dCTP concentration (Figure 2.15), and $K_m$ and $V_{max}$ values were derived from a double reciprocal plot (Figure 2.16).

![Graph showing the initial rates of insertion of dCTP across from $N^2$-furfuryl dG by DinB Val40Ala as a function of dCTP concentration.]

**Figure 2.15** Initial rates of insertion of dCTP across from $N^2$-furfuryl dG by DinB Val40Ala as a function of dCTP concentration.
Figure 2.16 Double reciprocal plot of initial rate of insertion dCTP across from $N^2$-furfuryl dG by DinBVal40Ala versus substrate concentration.

f. Met57Ala

DinB Met57Ala has been assayed in depth in both *in vitro* and *in vivo*. Cells that are grown with this variant of DinB show decreased survival to treatment with nitrofurazone (Figure 2.17). The *in vivo* work was done by Khadijah Balfour-Jeffrey, an undergraduate student in the summer of 2009. Since the cells harboring DinB Met57Ala were sensitive to nitrofurazone, the methionine at position 57 was of considerable interest.
Figure 2.17 Survival assay of *E. coli* expressing DinB Met57Ala. pWSK30 is the empty vector, pGY768 is DinB⁺. Cells that express DinB Met57Ala do not survive as well as cells with the empty vector.

In *vitro* primer extension assays were then done to determine if the enzyme was active, and two different purifications of DinB Met57Ala demonstrated that Met57 is an important residue for the functionality of DinB. DinB Met57Ala seems to be completely devoid of polymerase activity. DinB Met57Ala cannot replicate past undamaged DNA or damaged DNA but wild-type DinB is proficient with both substrates (Figure 2.18). Since DinB Met57Ala showed no extension activity on either damaged or undamaged templates, no further kinetic studies of single nucleotide incorporation were done. It seems as if this single amino acid change has a very large effect on the ability of the enzyme to perform its function of bypassing lesions both in the cell and *in vitro*. 
Figure 2.18 Met57Ala and Wild-type DinB activity assay via primer extension with a mixture of all four dNTPs. The concentration of DinB in each reaction is 25 nM. (A) Assay of wild type DinB and Met57Ala; the wild type extends the primer past the lesion to the end of template. (B) Assay of wild-type DinB and Met57Ala on the undamaged template. Wild-type DinB was assayed with lower dNTP concentration than Met57Ala because wild-type DinB is expected to be more active.
IV. Discussion

DinB can efficiently and accurately utilize $N^2$-furfuryl dG as a template while incorporating dCTP across from the lesion as it polymerizes DNA (Jarosz et al. 2006). However, as shown by the data presented above, DinB variants can also incorporate dCTP across from $N^2$-furfuryl dG, but with a wide range of efficiencies (Table 1). These kinetic parameters were calculated using a double reciprocal plot of initial rate ($V_o$) of catalysis versus the concentration of the substrate, dCTP. $K_m$ is the Michaelis-Menten constant, and is the concentration of substrate needed to reach one half of the maximum velocity ($V_{max}$) of the reaction. $k_{cat}$ is known as the turnover number of an enzyme and is a measure of the number of substrate molecules converted to product per unit time. It is calculated by dividing the $V_{max}$ by the total molar enzyme concentration, which in these reactions is 1 nM.
<table>
<thead>
<tr>
<th></th>
<th>$k_{\text{cat}}$ (min$^{-1}$)</th>
<th>$K_M$ (μM)</th>
<th>$k_{\text{cat}}/K_M$ (μM$^{-1}$min$^{-1}$)</th>
<th>Relative $k_{\text{cat}}/K_M$ (μM$^{-1}$min$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-Type</td>
<td>1.33 x 10$^8$</td>
<td>166</td>
<td>8.00 x 10$^5$</td>
<td>1.00</td>
</tr>
<tr>
<td>Val40Ala</td>
<td>2.32 x 10$^8$</td>
<td>98.7</td>
<td>2.35 x 10$^6$</td>
<td>2.94</td>
</tr>
<tr>
<td>Leu102Ile</td>
<td>2.10 x 10$^9$</td>
<td>0.412</td>
<td>5.09 x 10$^9$</td>
<td>6.36 x 10$^3$</td>
</tr>
<tr>
<td>Ser55Thr</td>
<td>3.99 x 10$^8$</td>
<td>2465</td>
<td>1.62 x 10$^5$</td>
<td>2.02 x 10$^{-1}$</td>
</tr>
<tr>
<td>Ser42Ala</td>
<td>1.57 x 10$^8$</td>
<td>1072</td>
<td>1.47 x 10$^5$</td>
<td>1.38 x 10$^{-1}$</td>
</tr>
</tbody>
</table>

**Table 2.1** Kinetic parameter results calculated for wild type DinB and variants. In all experiments the total enzyme concentration is kept constant at 1 nM and the substrate concentration is varied.

**Figure 2.19** DinB active site with the positions of the residues of interest highlighted. Val40 in yellow; Leu102 in gray; Ser55 in orange; Ser42 in purple; and Met57 in Blue; incoming nucleotide triphosphate shown in blue/teal/red/brown, colored by atom.
Modifying a single amino acid residue is known to effect the translesion synthesis activity of DinB (Jarosz et al. 2006), and we show here that additional variants of DinB affect its catalytic efficiency (Table 1). DinB variant Leu102Ile has a larger impact on the catalytic efficiency of the enzyme, increasing the efficiency by over 6000-fold, while DinB Ser55Thr and Ser42Ala have small decreases in efficiency. According to the homology model of DinB (Jarosz et al. 2006) (Figure 2.1), valine 40 and leucine 102 are close to the incoming nucleotide, in a pocket of residues implicated by sequence alignment to be near the incoming base (Figure 2.19). It is likely that the effect demonstrated is related to the proximity of these amino acids to the incoming base. The crystal structures of Dpo4 (a Y family polymerase from archaea) suggest the importance of residues in this region, which lie in the finger domain (Ling et al. 2001; Johnson et al. 2005; Wong et al. 2008). However Ser55Thr has a five-fold effect and it is predicted based on the model to be near the incoming nucleotide. Met57 is also in this region, and DinB Met57Ala was assayed both in vitro and in vivo with strikingly similar results of inactivity. This residue warrants further investigation and will be discussed in Chapter 4.

V. References


Chapter 3: Fluorescent Cytosine Analog Bypass

I. Introduction

Fluorescent nucleotide analogs are used to study the structure and dynamics of nucleic acids, the interactions of proteins with nucleic acids, and the mechanisms of DNA replication, damage tolerance and repair (Sandin et al. 2009). Such studies have traditionally been carried out with tethered dyes, but fluorescent base analogs have major advantages in that they can pi stack with surrounding bases and can hydrogen bond in a canonical nature with their complementary base, and so not perturb the structural integrity of the DNA (Figure 3.1) (Sandin et al. 2009). They can potentially give more useful information for DNA incorporation if a fluorescence acceptor is placed on a DNA polymerase for example, and their position is readily observable. This is in contrast to the classical tethered dyes, whose position cannot be definitively ascertained, and thus may be less useful in fluorescence resonance energy transfer (FRET) experiments where accurate distance determination may be the goal (Sandin et al. 2009).
Figure 3.1 Hydrogen bonding between (A) canonical guanosine and cytosine and (B) guanosine and the fluorescent cytosine analog (Stengel et al. 2007).

Virtually all existing base analogs have fluorescence quantum yields that are dependent on their environment (Sandin et al. 2009). However, the quantum yield of tC (and its family of analogs) is the least affected by the surrounding bases and is about 25-50 times greater than 2-aminopurine, a commonly used base analog (Sandin et al. 2009). Base flipping, in which the base conformation is extrahelical, can alter the fluorescence intensity and complicate interpretation of fluorescence experiments. The rate of base flipping in the tC analogs is not fast enough to impact the spectral output, as opposed to 2-aminopurine, for which the base flipping phenomenon must be corrected (Sandin et al. 2009). Originally DNA polymerization was measured with the use of small tethered fluorophores, which have fast rotational dynamics in solution but that slow down when they are bound to a large moiety such as a DNA strand or a protein, leading to an increase in fluorescence anisotropy (Sandin et al. 2009).
The Klenow fragment is the polymerase domain of A family DNA polymerase I, and has shown the ability to both bypass and incorporate the cytosine analog 1,3-diaza-2-oxophenothiazine (tC) (Stengel et al. 2007). Kinetic constants $K_m$ and $V_{max}$ were determined for the incorporation of these cytosine analogs into DNA by Klenow fragment (KF) and the $K_d$ was also determined using fluorescence anisotropy. Published data give a $K_D$ for the KF-DNA binary complex of 5 nM (Dahlberg et al. 1991), and that for the binding of the KF-DNA binary complex to the four different dNTPs of ~4-20 μM depending on the base and the environmental context (Eger et al. 1991). Fluorescence anisotropy, which measures the rotational freedom of a molecule during the lifetime of the excited state, was used to determine a $K_d$ value of 64 nM (+/- 15) for the binding of dtCTP by Klenow fragment. KF incorporates the unnatural fluorescent analog 100 times more efficiently than that of the natural base (Sandin et al. 2009). Competition assays were also completed, and these assays demonstrated that the Klenow fragment was indeed binding the analogs of cytosine in the same binding pocket as the natural dCTP (Sandin et al. 2009).

The authors indicated that this was the first time that the insertion of a size-expanded analog was incorporated at a higher efficiency than that of the natural substrate during a polymerization event. This efficiency is demonstrated by a lower $K_M$ value with a primarily unaffected $V_{max}$ value for the dtCTP analog versus dCTP (Sandin et al. 2009). Secondly, they took advantage of this very bright base analog as a biochemical tool for the study of DNA polymerases and used it to probe the dynamics of KF. They suggest that the Klenow fragment has a more flexible active site than previously imagined, enough so that it can handle the additional steric bulk of the extra benzene ring of the
tricyclic analog (Sandin et al. 2009). Even though tCTP was useful for studying the dynamics of KF, we show here that DNA containing this modification is not a substrate for DinB.

This is contradictory to the current thinking of the differences in the active sites of the families of polymerases. The current model suggests that the replicative polymerases have rigid active sites, containing the O-helix, a domain which closes the active site upon correct base incorporation, among other features that ensure fidelity (Beese et al. 1993). The Y family polymerases lack the O-helix and have rather small finger and thumb domains which allow for the active site in the palm domain to be open and solvent accessible (Yang 2005).

II. Materials and Methods

The assays of primer extension with the dNTP mix as well as single nucleotide incorporation were carried out as described in Chapter 2 with exceptions noted here. Most reaction volumes were doubled and reactions were carried out to longer times (60 min) due to the apparent inability of DinB to polymerize DNA that contained the tC analog. All of the reactions were performed in low light. The reactions were quenched as described in Chapter 2. We performed primer extension assays on a 61-mer DNA template with the tC analog incorporated at position 32.

\[5'\text{GGTTACTCAGATCAGGCCTGCGAAGACCTtCGCGTCCGGCTGCAGCTGTAC}\]
\[\text{TATCATATGC3'}\]

The primers used in these experiments were as follows:

- Running Start (30 mer)
  \[5'\text{GCATATGATAGTACAGCTGCAGCCGGACGC}\]
- Standing Start (31 mer)
  \[5'\text{GCATATGATAGTACAGCTGCAGCCGGACGCC}\]
Primer 22 (22 mer)
$5'$GCATATGATAGTACAGCTGCAG$3'
Primer 27 (27 mer)
$5'$GCATATGATAGTACAGCTGCAGCCGGA$3'
MismatchExtension;A,C,G,T (32 mer)
$5'$GCATATGATAGTACAGCTGCAGCCGGACGCCN$3'

III. Results

a. Activity of DinB on Template Containing tC Modification

We first assayed DinB activity on $^{32}$P- labeled running start primer annealed to the template containing the tC modification. Wild-type DinB appears to add nucleotides up to and possibly across from the tC nucleotide analog (Figure 3.2). Even though single nucleotide resolution was not achieved here, stepwise addition of two nucleotides is apparent, but the primer was not extended beyond the nucleotide opposite the modification.

Figure 3.2 Primer extension activity of wild-type DinB on DNA containing tC. Primer extension experiment on template tC annealed to running start, which is one nucleotide away from the lesion. (A) Time points are 0, 1, 10, 30 min; [dNTP] = 0.5 mM. (B) Time points are 0, 0.5, 1, 5, 10, 45 min; [dNTP] = 1 mM.

Next, primer extension activity was assayed with the template annealed to standing start primer. In this experiment, the first nucleotide incorporated is opposite the modified base. Three different concentrations of DinB were used: at 1 nM DinB there is
very little activity, while at both 10 nM DinB and 25 nM DinB there is incorporation opposite tC. Even at higher dNTP concentrations (1 mM), DinB could only extend the primer up to the position opposite the modification. Because the experiment was carried out with a mixture of all four dNTPs (at a concentration of 1 mM), it is unknown which nucleotide is being added (data not shown).

We next determined which nucleotide was incorporated opposite tC by DinB. We assayed single nucleotide incorporation efficiency in a series of reactions, each containing only one of the standard nucleotides. DinB incorporated dGTP (Figure 3.3), but not any of the other nucleotides opposite tC (data not shown).

![Figure 3.3 Incorporation of dGTP opposite tC by DinB.](image)

Since it seemed that DinB could insert dGTP across from but not extend past the tC analog, we probed whether DinB could extend from a nucleotide in a base pair with the modification. For these experiments, we used a set of mismatch extension primers, in which the 3’ nucleotide of the primer is involved in a base pair with the tC modification. A primer extension assay was carried out using each mismatch extension primer with wild-type DinB and with the KF. KF was able to extend past all of the mismatches, while
DinB was not able to extend from any nucleotide base paired with the tC modification (Figure 3.4).

![Figure 3.4 Primer extension assay by DinB wild type on MismatchExtension primer. The primer has one of the four dNTPs opposite the lesion. All primers were assayed with the Klenow fragment as a control. KF extended in all cases, however DinB did not extend under any conditions (dTTP not shown, in this case our observations were consistent with those shown).](image)

Since it was clear that DinB could not extend from a nucleotide opposite the tC analog, we wanted to determine if DinB could extend past the modification when given a larger running start. We carried out a primer extension assay using a shorter primer. Primer22 was annealed to tC-containing template. We carried the reactions out to 60 minutes. DinB adds nucleotides up to and across from the tC nucleotide. Standing start

63
primer was used as a marker to indicate the position just before the modification. One nucleotide beyond the standing start primer would indicate addition across from tC, and any bands beyond that would indicate bypass (Figure 3.5). We compared the activity of DinB on analogous unmodified DNA (Primer22:TemplateC) (Figure 3.5). This reaction demonstrated the ability of DinB to polymerize DNA to the full length of the template. Thus, the presence of tC prevents DinB from reaching the end of the template (Figure 3.5). There is also stalling near position 26 in the template for unknown reasons, although the presence of two cytosines adjacent to each other at position 26 and 27 may be a slight impediment to the enzyme.

Figure 3.5 DinB extension of primer22 annealed to tC-DNA template. A: unmodified DNA; Extension of TemplateC DNA with the unmodified cytosine instead of tC. DinB is able to extend the length of the primer to the template. B: tC DNA. DinB extends up to and adds across from the modification, but does not extend past it to any great extent. Wild-type DinB [25 nM] extension from Primer22; dNTP [0.5 mM]; Time points (min) in both experiments are: 0, 0.5, 1, 2,5, 10, 15,20,30,45,60.
b. Kinetics

Since we demonstrated the ability of DinB to insert dGTP across from tC but not to extend DNA beyond that, we wanted to determine the efficiency of incorporating dGTP across from the modification. This analysis was completed for a range of concentrations of dGTP from 0.5 μM to 1 mM. The initial velocity \( (V_o) \) is plotted against the concentration of substrate (Figure 3.6). A Lineweaver-Burk double-reciprocal plot was used to determine \( V_{\text{max}} \) and \( K_M \) (Figure 3.7, Table 3.1). We will now compare these values to parameters determined for incorporation of dGTP opposite unmodified C.

![Figure 3.6 Plot of initial velocities of insertion of dGTP opposite tC as a function of dGTP concentration.](image-url)
Figure 3.7 Double reciprocal plot of concentration of substrate versus initial velocity of the enzymatic reaction of insertion of dGTP across from tC.

<table>
<thead>
<tr>
<th></th>
<th>$k_{\text{cat}}$ (min$^{-1}$)</th>
<th>$K_M$ (μM)</th>
<th>$k_{\text{cat}}/ K_M$ (μM$^{-1}$min$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dGTP Insertion</td>
<td>$4.45 \times 10^6$</td>
<td>461</td>
<td>$9.66 \times 10^3$</td>
</tr>
</tbody>
</table>

Table 3.1 Kinetic values for the insertion of dGTP across from DNA template containing tC modified cytosine.

IV. Discussion

We show that DinB can insert dGTP accurately opposite tC, but is unable to extend from the resulting primer terminus. This is in sharp contrast to KF, which efficiently utilizes templates containing tC (Figure 3.5) (Stengel et al. 2007) The kinetic parameters of dtCTP incorporation opposite a dG by the Klenow fragment have been established with a $K_m$ of 0.021 μM, $V_{\text{max}}$ of 5.4%/min and a $V_{\text{max}}/K_m$ of 257.1%/min*μM (Sandin et al. 2009).
In order to remain consistent with the published data the kinetic parameters are reported here as a percentage of primer extended. DinB inserts dGTP across from the tC modified base in a DNA template strand with $V_{\text{max}}$ of 0.011%/min, $K_M$ of 5.71 $\mu$M and a $V_{\text{max}}/K_M$ of 0.0019%/min*μM. We also report the values in terms of molar concentration of primer extended by multiplying the ratio of extended primer to unextended primer and multiplying it by the concentration (10 nM) of primer in the reaction vessel (Table 3.1).

The $K_M$ for the incorporation of dGTP by DinB across from a tC modification is much higher, meaning a high concentration of substrate is needed to reach half of the maximum velocity of the reaction. Thus, DinB is far less active than KF on DNA substrates containing the tC modification. We plan to obtain the triphosphate form of tC so that we can repeat these experiments with DinB adding dtCTP across from a dG, to determine whether DinB is able to incorporate a modified nucleotidie.

Y family polymerases have the ability to accommodate bulky lesions, even as large as benzo[a]pyrene, in their active sites. Yet here it seems that DinB cannot synthesize DNA beyond a relatively simple tricyclic analog of cytosine, an analog which can efficiently and faithfully be incorporated and bypassed by KF (Stengel et al. 2007; Sandin et al. 2009). This seems to imply that there is a high level of substrate specificity for these polymerases. This may also suggest that DinB discriminates against modifications or adducts in the major versus minor groove of DNA.

V. References


Chapter 4: Future Consideration

I. Introduction

The data presented in this thesis contributes to our knowledge of substrate recognition by *E. coli* DinB (pol IV). By constructing single point mutations in DinB, we hope to elucidate which amino acids have the greatest impact on the functionality of DinB. We have employed two methods in order to identify candidate amino acids. The first is a computational assessment of active site residues utilizing THEMATICS (Ondrechen et al. 2001), which uses theoretical titration curve perturbations to identify important active site residues. The second method uses multiple sequence alignments to find conserved residues throughout the DinB subfamily. Thus far we have found one variant (Met57Ala) that causes cells expressing it to be as sensitive to nitrofurazone as cells lacking DinB. Also, this same variant is inactive in primer extension assays *in vitro*. A single amino acid change, other than changes in the catalytic residues, effecting this enzyme is not without precedent (Jarosz et al. 2006), and shown here is the enhanced efficiency of certain variants (Val40Ala, Leu102Ile) as well as variants that show little to no enhancement (Ser55Thr, Ser42Ala). There is relatively little known about the structure of DinB, because no crystal structure has been solved. A crystal structure would be an enormous breakthrough as possible conformational dynamics of the enzyme along with DNA and the incoming nucleotide could at last be demonstrated.
II. DinB Variants

These experiments will be repeated. The next step to finalizing this work is to continue to purify all of the identified variants of interest. Many of them are already in process. Additionally, the Ondrechen group is developing new computational methods to improve active site residue determination and to determine if amino acids somewhat distant from the active site (second or third shell) can affect the efficiency and specificity of the enzyme (Tong et al. 2008; Tong et al. 2009). Variants identified through this analysis will be assayed for activity and kinetic parameters will be determined using single nucleotide incorporation assays. This will initially be done for undamaged DNA, as well as DNA containing an $N^2$-dG lesion.

In combination with in vitro primer extension assays, in vivo assays on the sensitivity of cells containing these DinB variants will be done as well. It is a good way to assess if the in vitro findings mimic what is occurring at the cellular level. It is prudent to test the variants both ways so that an accurate description of the events of TLS can be determined.

III. Fluorescent Cytosine Analog

In order to accurately compare the findings of the ability of the Klenow fragment to incorporate dtCTP (Stengel et al. 2007; Sandin et al. 2009) with the ability of DinB to incorporate the same modified nucleotide, we first have to acquire the triphosphate form of 1,3-diaza-2-oxophenothiazine (dtCTP). We also plan to test the ability of the available DinB variants to bypass tC. Since the variants may have enhanced bypass abilities, it is worth determining if the variants can bypass the tC modification. Furthermore, we have obtained a new set of primers that are similar to the MismatchExtension (N) primers
discussed in Chapter 3. These primers have a dG annealed across from the tC modified base, and then from +1 base to +5 bases downstream from the lesion. It would be interesting to see if this modified base has any effect on the ability of DinB to polymerize DNA if it is beginning downstream from the tC modified base. These experiments will determine how far from the active site DinB must be before it can overcome the disruption induced by the tC modification.

IV. References


