Effects of non-catalytic, distal amino acid residues on activity of E. coli DinB (DNA Polymerase IV)

ARTICLE *in* ENVIRONMENTAL AND MOLECULAR MUTAGENESIS · DECEMBER 2012

Impact Factor: 2.63 · DOI: 10.1002/em.21730 · Source: PubMed

6 AUTHORS, INCLUDING:

Ramya Parasuram
Dana-Farber Cancer Institute
4 PUBLICATIONS **27** CITATIONS
SEE PROFILE

Eriks Rozners
Binghamton University
63 PUBLICATIONS **849** CITATIONS
SEE PROFILE

Mary Jo Ondrechen
Northeastern University
88 PUBLICATIONS **2,128** CITATIONS
SEE PROFILE

Environmental and Molecular Mutagenesis 53:766–776 (2012)

Research Article

Effects of Non-catalytic, Distal Amino Acid Residues on Activity of *E. coli* DinB (DNA Polymerase IV)

Jason M. Walsh,[1] Ramya Parasuram,[1] Pradyumna R. Rajput,[2] Eriks Rozners,[3] Mary Jo Ondrechen,[1,4] and Penny J. Beuning[1,4]*

[1]*Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts*
[2]*Program in Bioinformatics, Northeastern University, Boston, Massachusetts*
[3]*Department of Chemistry, Binghamton University, State University of New York, Binghamton, New York*
[4]*Center for Interdisciplinary Research on Complex Systems, Northeastern University, Boston, Massachusetts*

DinB is one of two Y family polymerases in *E. coli* and is involved in copying damaged DNA. DinB is specialized to bypass deoxyguanosine adducts that occur at the N^2 position, with its cognate lesion being the furfuryl adduct. Active site residues have been identified that make contact with the substrate and carry out deoxynucleotide triphosphate (dNTP) addition to the growing DNA strand. In DNA polymerases, these include negatively charged aspartate and glutamate residues (D8, D103, and E104 in *E. coli* DNA polymerase IV DinB). These residues position the essential magnesium ions correctly to facilitate nucleophilic attack by the primer hydroxyl group on the α-phosphate group of the incoming dNTP. To study the contribution of DinB residues to lesion bypass, the computational methods THEMATICS and POOL were employed. These methods correctly predict the known active site residues, as well as other residues known to be important for activity. In addition, these methods predict other residues involved in substrate binding as well as more remote residues. DinB variants with mutations at the predicted positions were constructed and assayed for bypass of the N^2-furfuryl-dG lesion. We find a wide range of effects of predicted residues, including some mutations that abolish damage bypass. Moreover, most of the DinB variants constructed are unable to carry out the extension step of lesion bypass. The use of computational prediction methods represents another tool that will lead to a more complete understanding of translesion DNA synthesis. Environ. Mol. Mutagen. 53:766–776, 2012.

Key words: DNA damage; homology model; computational predictions of activity; deoxyguanosine adduct; translesion synthesis

INTRODUCTION

DNA polymerases are generally highly efficient and accurate enzymes that use polymeric DNA as templates to synthesize new DNA. The DNA polymerase active site consists of negatively charged aspartate and glutamate residues that bridge the metal ions (2 Mg^{2+}) and position them into the correct orientation to activate and facilitate nucleophilic attack of the primer 3′-hydroxyl group onto

Additional Supporting Information may be found in the online version of this article.

Abbreviations: KSI, ketosteroid isomerase, NH, nitrile hydratase, dNTP, deoxy nucleotide triphosphate, PGI, phosphoglucose isomerase, POOL, partial order optimum likelihood, THEMATICS, theoretical microscopic anomalous titration curve shapes, TLS, translesion synthesis, XPV, xeroderma pigmentosum-variant.

Grant sponsor: National Science Foundation; Grant Number MCB-0843603; Grant Number MCB-1158176 (to M.J.O. and P.J.B.); Grant Number MCB-0845033 (Career Award to P.J.B.); Grant sponsor: NU Office of the Provost (to P.J.B.); Grant sponsor: The American Cancer Society; Grant Number RSG-12-161-01-DMC (to P.J.B.); Grant sponsor: Camille and Henry Dreyfus Foundation; Grant Sponsor: Research Corporation for Science Advancement (to P.J.B.).

Pradyumna R. Rajput is currently at IB Technology Solutions Inc., 201 Spear Street, San Francisco, CA 94105

*Correspondence to: Penny J. Beuning, Department of Chemistry and Chemical Biology, Northeastern University, 360 Huntington Ave, 102 Hurtig Hall, Boston, MA 02115. E-mail: beuning@neu.edu

Received 25 May 2012; provisionally accepted 8 August 2012; and in final form 6 August 2012

DOI 10.1002/em.21730

Published online 4 October 2102 in Wiley Online Library (wileyonlinelibrary.com).

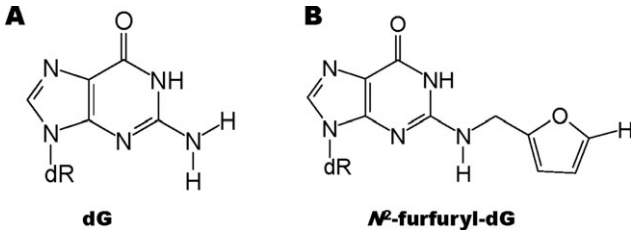

Fig. 1. Structural comparison of (**A**) natural deoxyguanosine and (**B**) N^2-furfuryl-deoxyguanosine (N^2-f-dG). DinB demonstrates a ~16 fold preference to incorporate dCTP opposite N^2-f-dG [Jarosz et al., 2006] and thus N^2-f-dG is considered the cognate lesion of DinB.

the α-phosphate group of the incoming deoxynucleotide triphosphate (dNTP) [Berdis, 2009]. The metal ions also help stabilize the negative charge that accumulates on both this ternary complex transition state and on the pyrophosphate product of DNA synthesis [Berdis, 2009].

DNA is constantly subjected to insults from environmental and endogenous sources [Friedberg et al., 2006]. The resulting damaged DNA generally cannot be copied efficiently or accurately by replicative DNA polymerases [Friedberg et al., 2006]. Thus, in addition to multiple DNA repair processes, organisms in all domains of life harbor specialized DNA polymerases that can copy damaged DNA in a process known as translesion synthesis (TLS) [Goodman, 2002; Friedberg, 2005; Prakash et al., 2005; Friedberg et al., 2006; Walsh et al., 2011b]. Y family DNA polymerases are capable of carrying out TLS, which is an important mechanism of DNA damage tolerance, and also exhibit a relatively high error frequency when copying undamaged DNA. Because of their roles in DNA damage tolerance and in mutagenesis, Y family DNA polymerases are implicated in cancer [Cordonnier et al., 1999; Bresson et al., 2002; Pages et al., 2002; Loeb et al., 2008; Lange et al., 2011], in resistance to DNA-damaging cancer chemotherapy agents [Bassett et al., 2004; Albertella et al., 2005; Albertella et al., 2005a; Chen et al., 2006; Ho et al., 2011], and are proposed to play a role in antibiotic resistance [Napolitano et al., 2000; Miller et al., 2004; Cirz et al., 2005; Cirz et al., 2006, 2007]. Notably, defects in Y family DNA pol η can cause the cancer-prone disease xeroderma pigmentosum-variant (XPV) [Kraemer et al., 1985; Johnson et al., 1999; Masutani et al., 1999; Itoh et al., 2000; Tanioka et al., 2007; Inui et al., 2008; Biertumpfel et al., 2010].

DinB (DNA polymerase IV) is one of two Y family DNA polymerases in *Escherichia coli* (*E. coli*) [Friedberg et al., 2002]. DinB incorporates nucleotides across from deoxyguanine adducts at the N^2 position, namely carboxyethyl [Yuan et al., 2008], benzo[a]pyrene [Shen et al., 2002], and the furfuryl adduct [Jarosz et al., 2006], with the latter being bypassed preferentially to that of natural dG in the template (Fig. 1) [Jarosz et al., 2006]. The N^2-furfuryl-dG lesion is thought to arise from nitrofurazone [Jarosz et al., 2006], which was used as an antibiotic in livestock until it was found to leave residue in edible tissues and was suspected of causing tumors in laboratory animals [Smith et al., 1998]. However, by analogy to the plant hormone kinetin (N^6-furfuryl-dA), the furfuryl moiety could also arise due to oxidation of ribose [Amasino et al., 2005; Barciszewski et al., 2007]. Kinetin has also been found in DNA extracted from human cells and in human urine [Barciszewski et al., 1996; Barciszewski et al., 2000]. Incorporation of dCTP opposite N^2-furfuryl-dG is efficient and accurate for both DinB and its mammalian ortholog Pol κ [Jarosz et al., 2006].

In addition to understanding the range of DNA damage that can be bypassed by DinB, it is of interest to determine how specific amino acids contribute to the specialized ability to bypass DNA damage. There are several important residues that have previously been identified in DinB [Walsh et al., 2011b], including the ''steric gate'' residue (F13), mutation of which eliminates the enhanced ability of DinB to bypass its cognate lesion N^2-furfuryl-dG [Jarosz et al., 2006]. The steric gate residues found in many DNA polymerases [Astatke et al., 1998; Bonnin et al., 1999; Ogawa et al., 2001; Yang et al., 2002; DeLucia et al., 2003; Shurtleff et al., 2009; Sherrer et al., 2010], including DinB [Jarosz et al., 2006], prevent ribonucleotide incorporation [Patel et al., 2000; Brown et al., 2011]. Residue Y79, which is near the steric gate residue and highly conserved in the DinB family, is thought to modulate the function of the steric gate residue [Jarosz et al., 2009]. Changing this residue has a profound effect on the ability of DinB to carry out the extension step of TLS, in which the primer is extended sufficiently beyond the lesion so that replicative DNA polymerases can resume replication [Jarosz et al., 2009]. This extension step of TLS may be as significant in overall bypass as the addition of a nucleotide opposite the lesion [Fujii et al., 2004]. The importance of positions other than the aromatic residues F13 and Y79 has been probed; specifically, mutations at position D8 (D8A, D8H), R49 (R49A, R49F), D103 (D103A, D103N), and E104 (E104A) confer a much lower mutation frequency (ranging from ~850 to ~3600 fold) than wild-type DinB [Wagner et al., 1999]. This is expected as D8, D103, and E104 are the highly conserved residues that coordinate the critical divalent cations [Wagner et al., 1999; Ling et al., 2001]. In addition, R49 lies in a loop region that is predicted to be in close proximity to the incoming nucleotide [Walsh et al., 2011b].

To systematically identify other residues that could contribute to polymerase activity, we applied computational methods to predict the functionally important residues in DinB. **The**oretical **M**icroscopic **A**nomalous **Ti**tration **C**urve **S**hapes (THEMATICS) uses perturbations in theoretical titration curves of ionizable residues to predict catalytic residues [Ondrechen et al., 2001; Ko et al.,

2005; Wei et al., 2007]. **P**artial **O**rder **O**ptimum **L**ikelihood (POOL [Tong et al., 2009; Somarowthu et al., 2011b]) is a monotonicity-constrained maximum likelihood machine learning methodology that incorporates the electrostatic features of THEMATICS as well as other input data, such as surface geometric properties and phylogenetic information, to rank all of the amino acid residues of DinB according to their probability of functional importance. Since there is no known experimentally determined crystal structure available for DinB, these methods were performed with a DinB model, which is based on homology to Dpo4 [Fiala et al., 2007]. POOL with THEMATICS and INTREPID [Sankararaman et al., 2009] input features has been verified as a successful predictor of catalytic and binding residues [Tong et al., 2009; Somarowthu et al., 2011b] using the CSA-100 [Bartlett et al., 2002; Porter et al., 2004; Sankararaman et al., 2009], a benchmark dataset of 100 non-homologous enzymes with experimental annotations of functionally important residues. Intriguingly, in addition to residues that immediately surround the substrate, other residues that are more distant from the substrate are often predicted to be important for activity. Such distal residues, although they do not come into direct contact with the substrate, may influence specificity or efficiency, possibly by their contact with those amino acid residues that do contact the substrate. These predictions of participation in catalysis by distal residues has been verified experimentally for nitrile hydratase (NH) [Brodkin et al., 2011], phosphoglucose isomerase (PGI) and ketosteroid isomerase (KSI) [Somarowthu et al., 2011a]. We typically define residues in contact with the substrate as first shell, those in contact with the first shell as second shell and those in contact with the second shell as third shell. Certain second-shell residues have been shown to have a profound impact on PGI function, while distal residues contribute little to the activity of KSI [Somarowthu et al., 2011a], as predicted. Therefore, changing residues that are distally located from the active site can impact different enzymes differently. While the residues predicted by THEMATICS/POOL tend to be highly conserved residues, the use of these electrostatics-based methods for the prediction of active residues has distinctive advantages over simple selection of conserved residues. The chief advantage is selectivity [Wei et al., 2007]. Residues may be conserved for a variety of reasons and thus not all conserved residues are directly involved in catalysis. The specificity scores achieved by THEMATICS [Wei et al., 2007] and POOL [Tong et al., 2009; Somarowthu et al., 2011b] are substantially higher than those obtained from simple sequence conservation-based methods. Indeed, as was demonstrated in earlier studies of NH [Brodkin et al., 2011] and of KSI and PGI [Somarowthu et al., 2011a], most of the residues predicted by THEMATICS prove to be important for catalysis, whereas negative control residues, i.e., conserved nearby residues that are not predicted by THEMATICS, are shown not to play a significant role in catalytic efficiency.

To our knowledge, systematic characterization of the contributions of remote residues to Y-family DNA polymerase activity has not yet been carried out, although remote residues in DNA pol η are associated with the cancer-prone disorder xeroderma pigmentosum-variant (XPV) [Broughton et al., 2002; Tanioka et al., 2007; Biertumpfel et al., 2010]. Moreover, a systematic and thorough analysis of the roles of residues remote from the active site in DNA polymerase β has identified residues that are important for nucleotide selection and template alignment [Yamtich and Sweasy, 2010]. As one example, a triad of residues Arg333 Glu316 Arg182 in the fingers domain of pol β may disrupt packing of side chains in the hinge region and therefore alter the conformational change required for efficient DNA replication [Murphy et al., 2011]. Distal mutations in two separate domains in *Thermus aquaticus* DNA polymerase I can impact its ability to replicate DNA; S543N in the thumb region can hinder template-dependent pausing during replication [Ignatov et al., 1999] and a number of mutations along the O helix (F667L, A661E, I665T) change the mutation spectrum [Suzuki et al., 2000]. Because of the large size of the DNA substrate and extensive contacts between DNA polymerases and DNA, it can be difficult to assign rigid definitions of shells. For the purposes of this work, we defined the first shell residues as the residues that are in contact with the incoming nucleotide substrate and the second-shell residues as those that are in contact with the first shell residues.

THEMATICS [Ondrechen et al., 2001; Ko et al., 2005; Wei et al., 2007] and POOL [Tong et al., 2009; Somarowthu et al., 2011b] were used to rank all amino acids of DinB in order of their probability of functional importance. Typically, the top 8–10% of the rank-ordered residues are used for functional residue prediction. Of the top 8%, we chose to focus on the top 10 residues. Among the top 10 POOL-predicted residues are seven that are in direct contact with the substrate and three second- or third-shell residues (Table I). The identified residues are H6, D103, E104, D8, K150, K157, Y106, R49, D10, and K146 in POOL rank order (Table I). As mutations at D8, R49, D103, and E104 have been previously characterized [Wagner et al., 1999], we constructed DinB variants that harbor the single mutations H6L, K150A, K157A, K157I, Y106A, Y106F, D10E, D10N, and K146A, and assessed the impact on catalytic efficiency of DinB with its cognate lesion N^2-furfuryl-dG. We find a range of contributions of these remote residues to activity, from modest effects to nearly complete loss of activity. Moreover, we find that these DinB variants have greatly diminished ability to carry out the extension step of TLS.

TABLE I. Top 8% of POOL Predicted Residues

Rank	No.	Residue	Shell	Distance from dNTP (Å)	Closest first shell residue	Domain	Consurf score
1	**6**	**HIS**	**second**	**7.6**	**E104**	**Palm**	**−1.054**
2	***103***	***ASP***	***first***	***2.4***		**Palm**	**−1.053**
3	***104***	***GLU***	***first***	***3.7***		**Palm**	**−1.052**
4	***8***	***ASP***	***first***	***3.2***		**Palm**	**−1.053**
5	**150**	**LYS**	**first**	**4.7**		**Palm**	**−1.051**
6	**157**	**LYS**	**first**	**2.5**		**Palm**	**−1.051**
7	**106**	**TYR**	**second**	**12.0**	**K150**	**Palm**	**−0.407**
8	**49**	**ARG**	**first**	**1.7**		**Finger**	**−1.053**
9	**10**	**ASP**	**first**	**4.7**		**Palm**	**−1.053**
10	**146**	**LYS**	**third**	**13.6**	**K150**	**Palm**	**−1.006**
11	9	MET	first	3.3		Palm	
12	98	GLU	second	12.9	E104	Palm	
13	158	PRO	second	5.8	K157	Palm	
14	79	TYR	second	5.3	F12	Palm	
15	54	ARG	first	2.4		Finger	
16	139	SER	second	6.8	D8	Palm	
17	66	CYS	third	14.8	R54	Finger	
18	57	MET	second	5.8	R54	Finger	
19	46	TYR	second	7.1	R49	Finger	
20	153	SER	second	7.3	K157	Palm	
21	330	ARG	remote	14.3		Little finger	
22	11	CYS	first	2.6		Palm	
23	101	SER	first	4.3		Palm	
24	248	THR	remote	11.5		Little finger	
25	12	PHE	first	2.0		Palm	
26	141	GLY	second	11.1	K150	Palm	
27	67	PRO	third	19.2	R54	Finger	
28	108	ASP	third	18.7	Y106	Palm	

Residues in bold, italic type are involved in metal coordination. Bold residues are the top ten of the POOL-predicted residues.

MATERIALS AND METHODS

Computational Methods

The DinB structure has not yet been determined, so a homology model was built with DNA and incoming nucleotide bound using the structure of the Dpo4 polymerase from *Sulfolobus solfataricus* with PDB ID 2imw [Fiala et al., 2007] as the template using the homology feature in the YASARA/WHATIF suite of programs [Krieger et al., 2002]. The model was refined with energy minimization using YAMBER3. The overall Quality Z-score of the final model was calculated to be −1.033 by YASARA. There are no outlier residues and 96.8% of residues are in the favored region of the Ramachandran plot [Lovell et al., 2003]. Because the DNA in the 2imw structure is a blunt-end DNA molecule, the DNA and the incoming nucleotide were then modeled from 1jx4 [Ling et al., 2001]. The YASARA molecular modeling software was used to add missing atoms to the ligands bound in the final DinB model. The coordinates for the model are included as Supporting Information.

Ligands are removed prior to the analysis using THEMATICS. The electrostatic properties from THEMATICS [Ondrechen et al., 2001] and the phylogenetic tree-based scores from INTREPID [Sankararaman et al., 2009] (http://phylogenomics.berkeley.edu/intrepid/) were used as input features to POOL. POOL then generated an output file of all residues in the model, rank-ordered according to their probability of functional importance.

Shell definitions were established as follows: the distance between the incoming nucleotide and the predicted residues was calculated using YASARA. Any residue that is within a distance of 5Å of the incoming nucleotide was considered to belong to the first shell, i.e., in direct contact with the incoming nucleotide. Any other residue that is within 5Å of any of the first shell residues is considered to be in the second shell, and so on.

Experimental Methods

DinB was purified as described previously by Beuning et al. [Beuning et al., 2006] and stored in single-use aliquots at −80°C. The DNA template containing a single N^2-furfuryl-dG adduct was prepared as described [DeCorte et al., 1996; Jarosz et al., 2006]. DNA sequences used in primer extension assays are as follows: standing start primer, 31-mer 5′-GCATATGATAGTACAGCTGCAGCCGGACGCC-3′; MatchC primer, 32-mer, 5′-GCATATGATAGTACAGCTGCAGCCGGACGCCC-3′; and template containing N^2-furfuryl-dG (N), 61-mer, 5′-GGTTACTCA GATCAGGCCTGCGAAGACCTNGGCGTCCGGCTGCTGTACTATCAT ATGC-3′. DNA primer (standing start or MatchC) and the template containing N^2-furfuryl-dG were combined to a final ratio of 1:1 (100 nM) and annealed in annealing buffer [20 mM Hepes (pH 7.5) and 5 mM $Mg(OAc)_2$] by heating for 2 min at 95°C, incubating at 50°C for 60 min, and then cooling to 37°C. The reactions were carried out with 100 nM ^{32}P-end-labeled primer/template in a reaction buffer containing final concentrations of 30 mM Hepes (pH 7.5), 20 mM NaCl, 7.5 mM $MgSO_4$, 2 mM β-mercaptoethanol, 1% bovine serum albumin, and 4% glycerol [Beuning et al., 2006]. Experiments were carried out with dNTP concentrations ranging from 1 μM-5000 μM, unless otherwise noted, and DinB concentration varied based on the extent of activity of each variant. An aliquot for the zero point was removed prior to addition of dNTP, and reactions were initiated by the addition of dNTP [Beuning et al., 2006]. The final reaction volumes were 30 μL. Time points were typically taken from 0.5 min to 30 min and reactions were quenched with 85% formamide, 50 mM ethylenediaminetetraacetic acid, 0.025% xylene cyanol, and 0.025% bromophenol blue [Beuning et al., 2006]. To determine kinetic parameters, conditions were chosen such that less than 25% of the substrate was converted to product. Quenched reaction products were analyzed on denaturing (8 M urea) 16% polyacrylamide gels, which were subsequently imaged on a Molecular Dynamics storage phosphor imaging screen with a Storm 860 imager. ImageQuant software (GE Healthcare) was used to analyze data. Kinetic parameters V_{max} and K_m were determined by assessing the percentage of primer extended at various time points in the experiment and were derived using GraphPad Prism® nonlinear regression analysis software [Segel, 1993]. DNA binding experiments were conducted by titrating 25 μM annealed DNA into a 60 μL solution of 1 μM DinB as previously described [Walsh et al., 2011a].

RESULTS

Table I shows the top 8% of the POOL predicted residues from the homology model of DinB. Of these, 12 residues belong to the first shell, 11 to the second shell, three to the third shell, and two are in more distant shells. Among the POOL predicted residues, there are 20 in the palm domain, six in the finger domain, and two in the little finger domain.

When used as a standalone method without POOL and other input features, THEMATICS predicts two clusters of residues using a statistical cutoff of 0.99 [Wei et al., 2007]. The first cluster has eight residues H6, D8, E104, Y106, K150, K157, D103, and R49 of which H6 and Y106 are second-shell residues. These residues represent the eight top-ranked residues also predicted by POOL. THEMATICS also predicts a second cluster of two resi-

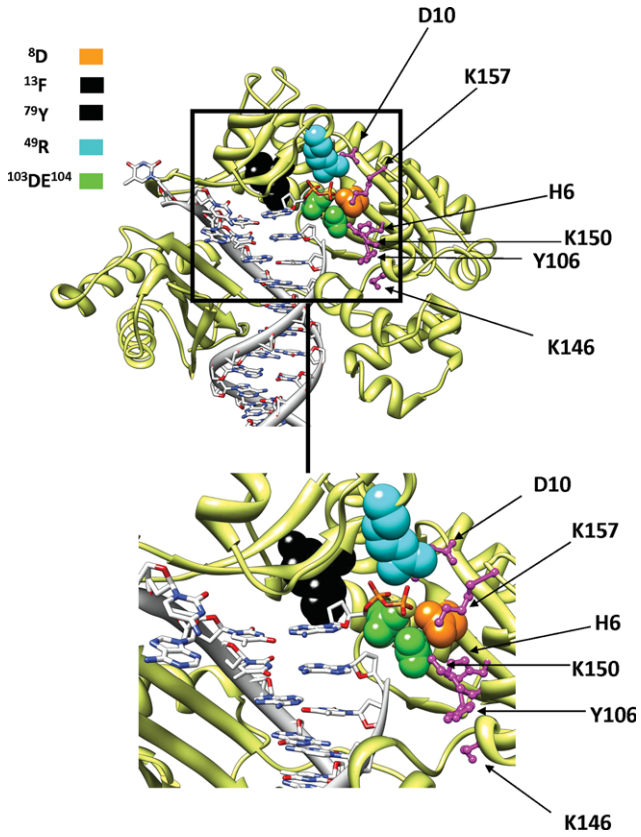

Fig. 2. Homology model of *E. coli* polymerase IV DinB. Highlighted residues include the identified distal residues (shown in magenta sticks) and their position in relation to known important residues (shown in space filling), including the main active site residues [Wagner et al., 1999; Jarosz et al., 2006; Jarosz et al., 2009] D8 (orange), D103 and E104 (green); the steric gate and its modulator, F13 and Y79, respectively (black); and the important residue R49 (light blue) [Wagner et al., 1999]. Images were rendered using the UCSF Chimera package [Pettersen et al., 2004].

dues, R35 and D252 that is somewhat distant from the active site. These residues are not ranked high in the POOL predictions, but may play a role in interaction of the finger and little finger domains with DNA upon DNA binding and catalysis. In this paper, we have investigated experimentally the top 10 residues predicted by POOL, of which three are residues outside the first shell.

Some of the highly POOL-ranked residues are the catalytic residues D8, D103, and E104, as well as R49, which have previously been shown to be important for activity, and other residues that are near the incoming nucleotide [Wagner et al., 1999]. Therefore, POOL correctly predicts the importance of active site residues. These methods also identified a number of distal residues, those not directly involved in binding the substrate, that are likely to contribute to activity. In this work, we focused on the 10 most highly ranked POOL-predicted residues (Table I). The predicted residues studied here form a crescent-shaped cluster around, and including, the metal-coordinating residues D8, D103, and E104 (Fig. 2). All of these residues are ionizable and may have either a steric or an electronic impact on the ability of DinB to catalyze dCTP addition across from N^2-furfuryl-dG, its cognate lesion. For every variant constructed, primer extension with a mixture of nucleotides as well as fidelity of incorporation was determined. Each variant that was active incorporated only dCTP opposite template N^2-furfuryl-dG (Fig. 3). For those variants that were inactive, we also determined their ability to bind DNA. We find that mutations at these amino acid positions have a variety of impacts on DinB function, as described in detail below.

H6L

Histidine-6 is the highest POOL-ranked residue. Histidine-6 is a second-shell residue that lies in the middle of the crescent near Y106 and K157 and is in close proximity to E104. Leucine was chosen as a reasonably close steric mimic of histidine. Compared with wild-type DinB, the H6L variant showed a small but statistically significant decrease (8.8-fold) in catalytic efficiency of dCTP incorporation opposite N^2-furfuryl-dG, with a 5-fold decrease in k_{cat} and a 1.7-fold increase in K_m. The H6L variant also showed weak extension of the MatchC primer/N^2-furfuryl-dG-containing template DNA, but did not extend to the end of the template, in contrast to wild-type DinB (Fig. 4).

D10E, D10N

Aspartic acid-10 is a first-shell residue that is 9th in the POOL rankings and is located near the catalytic metal ions, their coordinating residues D8, D103, and E104, and the important aromatic residues F13 [Jarosz et al., 2006] and Y79 [Jarosz et al., 2009] that are known to have a role in DinB specificity. Specifically, D10 is near the steric gate residue F13, which is important for TLS [Jarosz et al., 2006], and Y79, which modulates the activity of F13 and plays a role in TLS extension [Jarosz et al., 2009]. We constructed the mutation D10E to preserve the negative charge and D10N to preserve the length of the original amino acid. We did not observe any measurable activity with the D10E variant, although it could still bind DNA (Table II), but the D10N variant was weakly active at insertion of dCTP opposite N^2-furfuryl-dG. D10N showed a ~25 fold decrease in activity of insertion and was not able to fully extend to the end of the template, even when N^2-furfuryl-dG was correctly paired with the MatchC primer (Fig. 4). The position of this residue is likely the main reason that such large effects were observed.

Y106A, Y106F

Tyrosine-106 is a second-shell residue that is near H6, K146, and K150, as well as the first-shell catalytic residue E104. Tyrosine has the ability to be involved in hydrogen

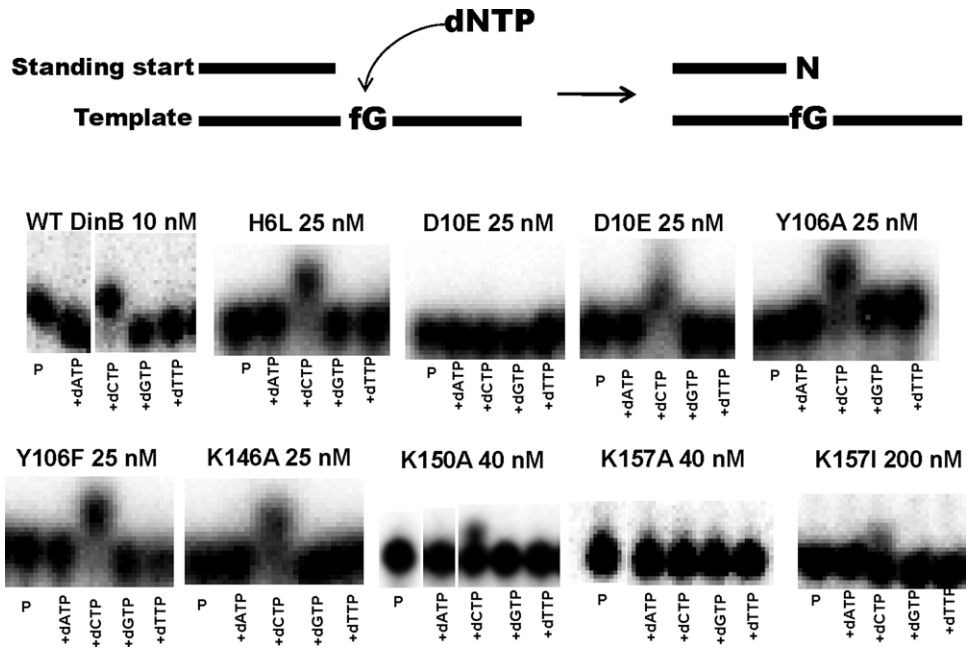

Fig. 3. Schematic of single nucleotide incorporation of each of the four deoxynucleotide triphosphates (dATP, dCTP, dGTP, dTTP) with a P indicating primer only, reactions in which no deoxynucleotide triphosphates were added. No misincorporation of dATP, dGTP, dTTP opposite N^2-f-dG was observed, while dCTP addition opposite N^2-f-dG was observed for the variants that were active. The concentration of polymerase is noted above each set of reactions.

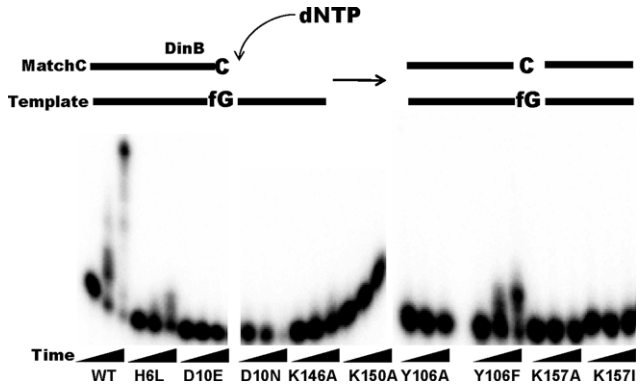

Fig. 4. Schematic of primer extension assays with primer/template DNA containing N^2-f-dG correctly base paired with C. Wild-type DinB extended to the end of the template; only H6L and Y106F showed appreciable extension, but they did not complete extension to the end of the template. The other variants could not extend to the end of the template. DinB was present at 25 nM.

bonding as well as pi stacking, so we first modified this position to alanine to eliminate both of those properties. The effects of this Y106A modification on overall catalytic efficiency were large, resulting in a decrease in k_{cat}/K_m of 29 fold. Interestingly, this overall decrease in catalytic efficiency resulted from only a 1.6-fold increase in K_m, but an 18-fold decrease in k_{cat}.

Modifying this position to phenylalanine (Y106F), which retains the aromatic pi system, gives catalytic efficiency closer to that of wild-type DinB, with a small but statistically significant 5.8-fold decrease. This mutation resulted in a ~3-fold increase in the value of K_m and a 2-fold decrease in k_{cat}. This suggests that the pi system at this position contributes to catalytic efficiency of DinB and that the hydroxyl group is also important for activity. DinB Y106F was the most proficient of the variants at extension from the damaged site (Fig. 4).

K146A

Among the residues tested, lysine-146 is the furthest from either the DNA substrate or the known catalytic residues that ranked in the top 10 POOL-predicted residues. Because it is behind second-shell residue Y106 with respect to the substrate, it is defined as a third-shell residue. Even though it is considered a third-shell residue and therefore only two other residues are between it and the substrate, K146 lies in a solvent exposed area on α helix F. DinB K146A gave a 1.6-fold reduction in K_m and an 11.4-fold decrease in k_{cat}. However, we could not detect extension of the primer to the end of the template when DinB K146A was presented with a mixture of all four dNTPs, which may suggest an inability to extend from the new primer terminus. To test this possibility, K146A was also assayed for activity with a DNA substrate in which the N^2-furfuryl adduct was correctly paired with a C on the primer (MatchC). DinB K146A was not able to extend the primer when the terminus involved a base pair containing the template N^2-furfuryl-dG (Fig. 4). Such extension could only be detected at very high concentrations (1.5 μM) of DinB K146A.

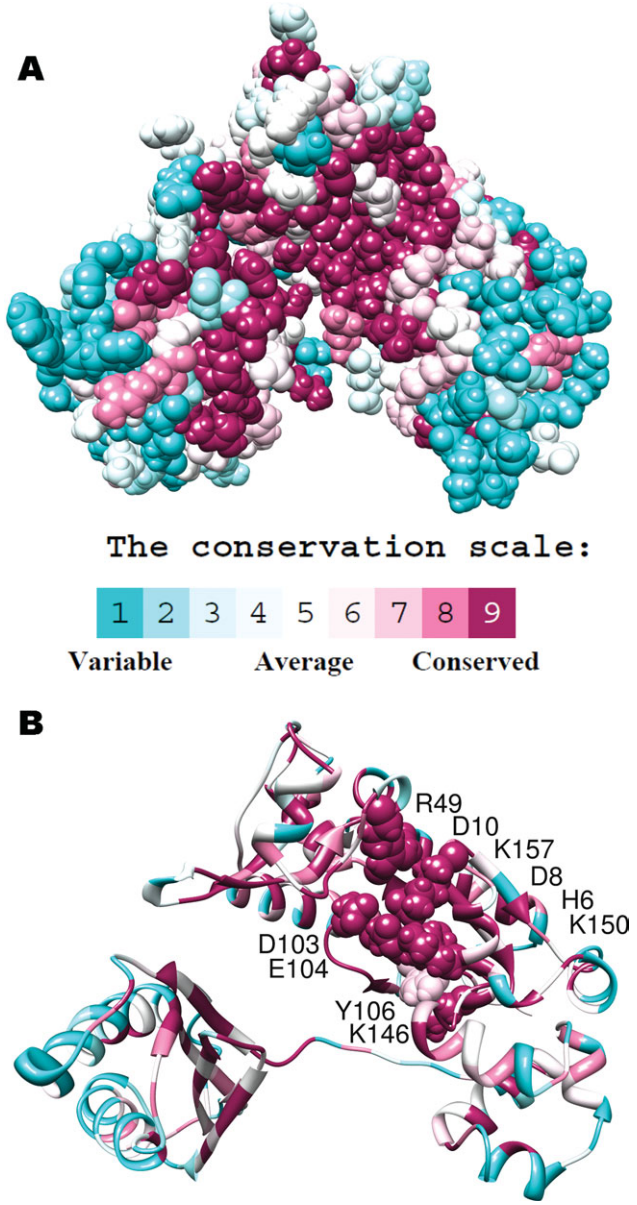

Fig. 5. (**A**) ConSurf model [Glaser et al., 2003] showing which amino acid residues are conserved in the DinB family. The degrees of conservation are color coded as blue being the least conserved to deep magenta being the most conserved. (**B**) The top 10 POOL-predicted residues in DinB that are highly conserved, except for Y106, are highlighted. The substrates have been removed for clarity. Images were rendered using the UCSF Chimera package [Pettersen et al., 2004].

K150A, K157A, K157I

Based on our homology model, lysine-150, and lysine-157 are both very near the active site residues D8, D103, and E104, as well as the DNA and incoming nucleotide substrates. According to POOL, these two residues are ranked 5th and 6th, respectively. Changing either of these residues to alanine had significant effects on activity. DinB K150A was active, but with a 15-fold decrease in catalytic efficiency relative to that of wild-type DinB. Based on the homology model we hypothesize that this residue with its positively charged side chain may stabilize the backbone phosphate of the nucleotide at the end of the primer.

Lysine-157 is closer than K150 to the catalytic residues and may be involved in binding to the phosphate of the incoming nucleotide. The variant K157A was not catalytically active; however, DinB K157A was found to bind DNA with similar affinity to that of wild-type DinB (Table II), suggesting that the overall conformation of DinB is not greatly perturbed by this mutation. Introducing an isoleucine at this position restored weak activity for incorporation of dCTP opposite N^2-furfuryl-dG, suggesting that the chain length is a factor in the importance of this residue. Among the variants that exhibited detectable activity, K157I gave the largest change observed, with a ~59 fold reduction in catalytic efficiency. None of these K157 mutations were able to exhibit any further extension on DNA substrates from a MatchC primer (Fig. 4).

DISCUSSION

We used the THEMATICS and POOL programs, which predict functionally important residues, to uncover and probe residues that contribute to activity in DinB. In this work, we focused on the residues ranked in the top 10 of the predicted residues. These functional site predictors have been used to correctly predict important residues outside of the first shell in several different enzymes [Brodkin et al., 2011; Somarowthu et al., 2011a]. In each of the previous cases, the substrates were small molecules. In this work, we tested two aspects of the prediction: (1) the performance with an enzyme such as DinB that makes extensive contacts with its large DNA substrate and (2) the performance of the prediction based on a homology model. The top 10 POOL-predicted residues are all in or near the catalytic site and include known metal-coordinating residues (D8, D103, and E104) as well as R49, which is known to be critical for function [Wagner et al., 1999]. The predicted residues include other residues near the substrate, which are likely to interact directly with the incoming nucleotide, as well as second- and third-shell residues that are found to contribute to activity. Notably in this case, the largest effects observed upon mutation of the predicted residues were the catalytically inactive first-shell variants K157A and D10E; whereas the largest effect of a variant that could accomplish catalysis was the first-shell variant K157I, followed by a second-shell variant Y106A.

The K146 residue is a third-shell residue that is located near the well-conserved ''FLAKIA'' motif in Y family DNA polymerases that includes the highly-conserved residue K150 (Fig. 5). DinB K146 appears to be involved

TABLE II. Kinetic Parameters for dCTP Incorporation opposite N^2-f-dG by DinB and Binding Affinity of DinB to DNA

DinB construct	K_m (μM)	k_{cat} (min^{-1})	V_{max} (μM min^{-1})	k_{cat}/K_m (min^{-1} μM^{-1})	Fold decrease	$K_{D,\ DNA}$ (μM)	Fold increase
Wild Type	49 ± 22	24	0.24 ± 0.070	0.49	1	0.33 ± 0.07	1
His6Leu	81 ± 21	4.5	0.045 ± 0.030	0.056	8.8		
Lys150Ala	27 ± 5.9	0.92	0.023 ± 0.0040	0.034	15		
Lys157Ala	ND		ND			0.49 ± 0.10	1.5
Lys157Ile	130 ± 34	1.1	0.027 ± 0.010	0.0083	59		
Tyr106Ala	78 ± 35	1.3	0.026 ± 0.0030	0.017	29		
Tyr106Phe	150 ± 44	12	0.12 ± 0.062	0.085	5.8		
Asp10Glu	ND		ND			0.51 ± 0.13	1.5
Asp10Asn	46 ± 4.7	0.90	0.022 ± .0050	0.020	25		
Lys146Ala	30 ± 7.2	2.1	0.053 ± 0.009	0.07	7.1		

ND, not detected.

in interactions with E98 and N235, the latter of which is part of the peptide linker that tethers the little finger domain to the polymerase domain in Y family DNA polymerases. Therefore, K146 could play a role in positioning the little finger domain. K146, as well as other variants assayed here, also appear to play an important role in the extension step of TLS, which can be as important in bypass as insertion opposite the lesion [Jarosz et al., 2009]), because primer extension to the full length of the template demonstrated by the K146 variant was very poor. The specific effect of a third-shell residue on extension is unexpected.

Distal residues can have a variety of effects on the structure [Chalissery et al., 2007; Chaptal et al., 2007], substrate binding affinity [Heitman et al., 2006; Carpten et al., 2007; Chalissery et al., 2007; Mirza et al., 2007], stability [Hong et al., 2007], and function [El Omari et al., 2006; Dupre et al., 2007; Klyuyeva et al., 2008] of a wide variety of different enzymes [Lee et al., 2011]. The semantics of defining the ''shell'' to which a specific amino acid residue belongs can be challenging, in particular in enzymes with large substrates, such as DNA polymerases. In DNA polymerases, a larger proportion of the enzyme can make contacts with the DNA substrate than an enzyme with a small molecule substrate. Aminoacyl-tRNA synthetases (aaRSs) are examples of enzymes that have very large nucleic acid substrates and whose activity is affected by mutations of distant residues [Lue et al., 2007]. The aaRSs must bind to and activate their cognate amino acids and tRNAs, and then transfer the activated aminoacyl adenylate to the tRNA [Arnez et al., 1997; Lee et al., 2011]. There is a need for communication between distinct domains in aaRSs for efficient and accurate aminoacylation [Alexander et al., 2001]. Interestingly, the mutation L570F in *E. coli* leucyl-tRNA synthetase has a significant impact on the function of this enzyme, resulting in a ~30 fold decrease in activity from that of the wild type [Lue et al., 2007]. Crystal structure studies of *Thermus thermophilus* LeuRS show that the equivalent residue does not influence the interaction between the protein and the amino acid substrate [Cusack et al., 2000]. Rather, it is hypothesized that this residue comes into close proximity to highly conserved M576 which is located in a functionally important loop area in the protein that is responsible for the modulation of the active site conformation that is essential for substrate recognition [Weimer et al., 2009]. Thus, the remote residue L570 contributes substantially to activity via its ability to mediate a conformational change.

A more closely related protein to DinB is mammalian DNA pol η, a Y family DNA polymerase encoded by the XPV gene [Johnson et al., 1999; Masutani et al., 1999]. Defective DNA pol η can cause the autosomal recessive disease XPV, which is characterized by UV-light-induced deterioration of a variety of tissues, as well as malignant skin carcinomas and melanomas [Kraemer et al., 1985]. Some of the pol η variants that are associated with XPV involve mutations that are distal to the polymerase active site triad including D13, D115, and E116 [Biertumpfel et al., 2010]. Many missense mutations in pol η are found in XPV patients: R111H, A117P, T122P, G263V, R361S [Tanioka et al., 2007]; W174C, K220E [Inui et al., 2008]; K535E, K589T [Itoh et al., 2000]. Some of these missense mutations along with others (A264P, F290S, and G295R) identified by Biertümpfel, et al. were mapped onto the crystal structure of pol η [Biertumpfel et al., 2010], and are located relative to the catalytic triad in similar positions to the highly ranked DinB residues examined here. In particular, DNA pol η K220 is aligned with DinB K146.

Through the use of predictive computations, we have demonstrated that residues that are removed from the active site in a DNA damage bypass polymerase have an impact on the catalytic efficiency of that enzyme. The ability of the TLS polymerase DinB to add dCTP opposite its cognate N^2-furfuryl-dG lesion is altered when point mutations are introduced distal to the active site residues. In particular, we observed reductions in catalytic efficiency from ~6 to 60 fold as well as instances of complete loss of activity. Using only computational methods to guide us, we identified residues that are important for TLS activity. These residues are postulated to impact

the active site residues through steric, pi systems, or electrostatic interactions that are either lost or reduced in the mutations made. Identifying these distal residues and characterizing their contributions to activity may lead to a better understanding of the enzymes as well as to controlled manipulation of enzyme activity and specificity [Somarowthu et al., 2011a].

AUTHOR CONTRIBUTIONS

PJB and MJO designed the study. RP modeled the DNA, carried out calculations for the residue predictions, and prepared figures. PRR constructed the homology model. ER synthesized DNA constructs. JMW collected and analyzed the experimental data and prepared the manuscript draft, tables and figures. JMW, RP, PJB, and MJO edited the manuscript. All authors approved the final manuscript.

ACKNOWLEDGMENTS

The authors acknowledge the current and former members of the Beuning DNA Damage and Repair laboratory, especially Dr. Jana Sefcikova, Dr. Srinivas Somarowthu, Lisa Hawver, and Nicholas DeLauteur. They also acknowledge the Ondrechen Computational Biology research group (Northeastern University), as well as the Strauss and Godoy-Carter research groups (Northeastern University) for use of their phosphorimagers. P.J.B. is a Cottrell Scholar of the Research Corporation for Science Advancement.

REFERENCES

Albertella MR, Green CM, Lehmann AR, O'Connor MJ. 2005. A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. Cancer Res 65:9799–9806.

Albertella MR, Lau A, O'Connor MJ. 2005a. The overexpression of specialized DNA polymerases in cancer. DNA Repair (Amst) 4:583–593.

Alexander RW, Schimmel P. 2001. Domain-domain communication in aminoacyl-tRNA synthetases. Prog Nucleic Acid Res Mol Biol 69:317–349.

Amasino R. 2005. 1955: Kinetin arrives: the 50th anniversary of a new plant hormone. Plant Physiol 138:1177–1184.

Arnez JG, Moras D. 1997. Structural and functional considerations of the aminoacylation reaction. Trends Biochem Sci 22:211–216.

Astatke M, Ng K, Grindley ND, Joyce CM. 1998. A single side chain prevents *Escherichia coli* DNA polymerase I (Klenow fragment) from incorporating ribonucleotides. Proc Natl Acad Sci U S A 95:3402–3407.

Barciszewski J, Massino F, Clark BF. 2007. Kinetin–a multiactive molecule. Int J Biol Macromol 40:182–192.

Barciszewski J, Mielcarek M, Stobiecki M, Siboska G, Clark BF. 2000. Identification of 6-furfuryladenine (kinetin) in human urine. Biochem Biophys Res Commun 279:69–73.

Barciszewski J, Siboska GE, Pedersen BO, Clark BF, Rattan SI. 1996. Evidence for the presence of kinetin in DNA and cell extracts. FEBS Lett 393:197–200.

Bartlett GJ, Porter CT, Borkakoti N, Thornton JM. 2002. Analysis of catalytic residues in enzyme active sites. J Mol Biol 324:105–121.

Bassett E, King NM, Bryant MF, Hector S, Pendyala L, Chaney SG, Cordeiro-Stone M. 2004. The role of DNA polymerase eta in translesion synthesis past platinum-DNA adducts in human fibroblasts. Cancer Res 64:6469–6475.

Berdis AJ. 2009. Mechanisms of DNA polymerases. Chem Rev 109:2862–2879.

Beuning PJ, Simon SM, Godoy VG, Jarosz DF, Walker GC. 2006. Characterization of *Escherichia coli* translesion synthesis polymerases and their accessory factors. Methods Enzymol 408:318–340.

Biertumpfel C, Zhao Y, Kondo Y, Ramon-Maiques S, Gregory M, Lee JY, Masutani C, Lehmann AR, Hanaoka F, Yang W. 2010. Structure and mechanism of human DNA polymerase eta. Nature 465:1044–1048.

Bonnin A, Lazaro JM, Blanco L, Salas M. 1999. A single tyrosine prevents insertion of ribonucleotides in the eukaryotic-type phi29 DNA polymerase. J Mol Biol 290:241–251.

Bresson A, Fuchs RP. 2002. Lesion bypass in yeast cells: Pol eta participates in a multi-DNA polymerase process. EMBO J 21:3881–3887.

Brodkin HR, Novak WR, Milne AC, D'Aquino JA, Karabacak NM, Goldberg IG, Agar JN, Payne MS, Petsko GA, Ondrechen MJ, Ringe D. 2011. Evidence of the participation of remote residues in the catalytic activity of co-type nitrile hydratase from *Pseudomonas putida*. Biochemistry 50:4923–4935.

Broughton BC, Cordonnier A, Kleijer WJ, Jaspers NG, Fawcett H, Raams A, Garritsen VH, Stary A, Avril MF, Boudsocq F, Masutani C, Hanaoka F, Fuchs RP, Sarasin A, Lehmann AR. 2002. Molecular analysis of mutations in DNA polymerase eta in xeroderma pigmentosum-variant patients. Proc Natl Acad Sci U S A 99:815–820.

Brown JA, Suo Z. 2011. Unlocking the sugar "steric gate" of DNA polymerases. Biochemistry 50:1135–1142.

Carpten JD, Faber AL, Horn C, Donoho GP, Briggs SL, Robbins CM, Hostetter G, Boguslawski S, Moses TY, Savage S, Uhlik M, Lin A, Du J, Qian YW, Zeckner DJ, Tucker-Kellogg G, Touchman J, Patel K, Mousses S, Bittner M, Schevitz R, Lai MH, Blanchard KL, Thomas JE. 2007. A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. Nature 448:439–444.

Chalissery J, Banerjee S, Bandey I, Sen R. 2007. Transcription termination defective mutants of Rho: Role of different functions of Rho in releasing RNA from the elongation complex. J Mol Biol 371:855–872.

Chaptal V, Vincent F, Gueguen-Chaignon V, Monedero V, Poncet S, Deutscher J, Nessler S, Morera S. 2007. Structural analysis of the bacterial HPr kinase/phosphorylase V267F mutant gives insights into the allosteric regulation mechanism of this bifunctional enzyme. J Biol Chem 282:34952–34957.

Chen YW, Cleaver JE, Hanaoka F, Chang CF, Chou KM. 2006. A novel role of DNA polymerase eta in modulating cellular sensitivity to chemotherapeutic agents. Mol Cancer Res 4:257–265.

Cirz RT, Chin JK, Andes DR, de Crecy-Lagard V, Craig WA, Romesberg FE. 2005. Inhibition of mutation and combating the evolution of antibiotic resistance. PLoS Biol 3:e176.

Cirz RT, Jones MB, Gingles NA, Minogue TD, Jarrahi B, Peterson SN, Romesberg FE. 2007. Complete and SOS-mediated response of *Staphylococcus aureus* to the antibiotic ciprofloxacin. J Bacteriol 189:531–539.

Cirz RT, O'Neill BM, Hammond JA, Head SR, Romesberg FE. 2006. Defining the *Pseudomonas aeruginosa* SOS response and its role

in the global response to the antibiotic ciprofloxacin. J Bacteriol 188:7101–7110.

Cordonnier AM, Fuchs RP 1999. Replication of damaged DNA: Molecular defect in xeroderma pigmentosum variant cells. Mutat Res 435:111–119.

Cusack S, Yaremchuk A, Tukalo M. 2000. The 2A crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. EMBO J 19:2351–2361.

DeCorte BL, Tsarouhtsis D, Kuchimanchi S, Cooper MD, Horton P, Harris CM, Harris TM. 1996. Improved strategies for postoligomerization synthesis of oligodeoxynucleotides bearing structurally defined adducts at the N^2 position of deoxyguanosine. Chem Res Toxicol 9:630–637.

DeLucia AM, Grindley ND, Joyce CM. 2003. An error-prone family Y DNA polymerase (DinB homolog from *Sulfolobus solfataricus*) uses a 'steric gate' residue for discrimination against ribonucleotides. Nucleic Acids Res 31:4129–4137.

Dupre ML, Broyles JM, Mihic SJ. 2007. Effects of a mutation in the TM2–TM3 linker region of the glycine receptor alpha1 subunit on gating and allosteric modulation. Brain Res 1152:1–9.

El Omari K, Liekens S, Bird LE, Balzarini J, Stammers DK. 2006. Mutations distal to the substrate site can affect varicella zoster virus thymidine kinase activity: Implications for drug design. Mol Pharmacol 69:1891–1896.

Fiala KA, Brown JA, Ling H, Kshetry AK, Zhang J, Taylor JS, Yang W, Suo Z. 2007. Mechanism of template-independent nucleotide incorporation catalyzed by a template-dependent DNA polymerase. J Mol Biol 365:590–602.

Friedberg EC, Walker GC, Siede WW, Wood RD, Schultz RA, Ellenberger T. 2006. DNA repair and mutagenesis, 2nd edition. Washington, D.C.: ASM Press. 1118 p.

Friedberg EC. 2005. Suffering in silence: The tolerance of DNA damage. Nat Rev Mol Cell Biol 6:943–953.

Friedberg EC, Wagner R, Radman M. 2002. Specialized DNA polymerases, cellular survival, and the genesis of mutations. Science 296:1627–1630.

Fujii S, Fuchs RP. 2004. Defining the position of the switches between replicative and bypass DNA polymerases. EMBO J 23:4342–4352.

Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N. 2003. ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 19:163–164.

Goodman MF. 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. Annu Rev Biochem 71:17–50.

Heitman LH, Mulder-Krieger T, Spanjersberg RF, von Frijtag Drabbe Kunzel JK, Dalpiaz A, Ijzerman AP. 2006. Allosteric modulation, thermodynamics and binding to wild-type and mutant (T277A) adenosine A1 receptors of LUF5831, a novel nonadenosine-like agonist. Br J Pharmacol 147:533–541.

Ho TV, Guainazzi A, Derkunt SB, Enoiu M, Scharer OD. 2011. Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases. Nucleic Acids Res 39:7455–7464.

Hong BS, Senisterra G, Rabeh WM, Vedadi M, Leonardi R, Zhang YM, Rock CO, Jackowski S, Park HW. 2007. Crystal structures of human pantothenate kinases. Insights into allosteric regulation and mutations linked to a neurodegeneration disorder. J Biol Chem 282:27984–27993.

Ignatov KB, Bashirova AA, Miroshnikov AI, Kramarov VM. 1999. Mutation S543N in the thumb subdomain of the Taq DNA polymerase large fragment suppresses pausing associated with the template structure. FEBS Lett 448:145–148.

Inui H, Oh KS, Nadem C, Ueda T, Khan SG, Metin A, Gozukara E, Emmert S, Slor H, Busch DB, Baker CC, DiGiovanna JJ, Tamura D, Seitz CS, Gratchev A, Wu WH, Chung KY, Chung HJ, Azizi E, Woodgate R, Schneider TD, Kraemer KH. 2008. Xeroderma pigmentosum-variant patients from America, Europe, and Asia. J Invest Dermatol 128:2055–2068.

Itoh T, Linn S, Kamide R, Tokushige H, Katori N, Hosaka Y, Yamaizumi M. 2000. Xeroderma pigmentosum variant heterozygotes show reduced levels of recovery of replicative DNA synthesis in the presence of caffeine after ultraviolet irradiation. J Invest Dermatol 115:981–985.

Jarosz DF, Cohen SE, Delaney JC, Essigmann JM, Walker GC. 2009. A DinB variant reveals diverse physiological consequences of incomplete TLS extension by a Y-family DNA polymerase. Proc Natl Acad Sci U S A 106:21137–21142.

Jarosz DF, Godoy VG, Delaney JC, Essigmann JM, Walker GC. 2006. A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. Nature 439:225–228.

Johnson RE, Kondratick CM, Prakash S, Prakash L. 1999. hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 285:263–265.

Klyuyeva A, Tuganova A, Popov KM. 2008. Allosteric coupling in pyruvate dehydrogenase kinase 2. Biochemistry 47:8358–8366.

Ko J, Murga LF, Andre P, Yang H, Ondrechen MJ, Williams RJ, Agunwamba A, Budil DE. 2005. Statistical criteria for the identification of protein active sites using theoretical microscopic titration curves. Proteins 59:183–195.

Kraemer KH, Slor H. 1985. Xeroderma pigmentosum. Clin Dermatol 3:33–69.

Krieger E, Koraimann G, Vriend G. 2002. Increasing the precision of comparative models with YASARA NOVA—A self-parameterizing force field. Proteins 47:393–402.

Lange SS, Takata K, Wood RD. 2011. DNA polymerases and cancer. Nat Rev Cancer 11:96–110.

Lee J, Goodey NM. 2011. Catalytic contributions from remote regions of enzyme structure. Chem Rev 111:7595–7624.

Ling H, Boudsocq F, Woodgate R, Yang W. 2001. Crystal structure of a Y-family DNA polymerase in action: A mechanism for error-prone and lesion-bypass replication. Cell 107:91–102.

Loeb LA, Monnat RJ Jr. 2008. DNA polymerases and human disease. Nat Rev Genet 9:594–604.

Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. 2003. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 50:437–450.

Lue SW, Kelley SO. 2007. A single residue in leucyl-tRNA synthetase affecting amino acid specificity and tRNA aminoacylation. Biochemistry 46:4466–4472.

Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, Yuasa M, Araki M, Iwai S, Takio K, Hanaoka F. 1999. The XPV (*xeroderma pigmentosum* variant) gene encodes human DNA polymerase eta. Nature 399:700–704.

Miller C, Thomsen LE, Gaggero C, Mosseri R, Ingmer H, Cohen SN. 2004. SOS response induction by beta-lactams and bacterial defense against antibiotic lethality. Science 305:1629–1631.

Mirza M, Robinson P, Kremneva E, Copeland O, Nikolaeva O, Watkins H, Levitsky D, Redwood C, El-Mezgueldi M, Marston S. 2007. The effect of mutations in alpha-tropomyosin (E40K and E54K) that cause familial dilated cardiomyopathy on the regulatory mechanism of cardiac muscle thin filaments. J Biol Chem 282:13487–13497.

Murphy DL, Jaeger J, Sweasy JB. 2011. A triad interaction in the fingers subdomain of DNA polymerase beta controls polymerase activity. J Am Chem Soc 133:6279–6287.

Napolitano R, Janel-Bintz R, Wagner J, Fuchs RP. 2000. All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. EMBO J 19:6259–6265.

Ogawa M, Tosaka A, Ito Y, Yoshida S, Suzuki M. 2001. Enhanced ribonucleotide incorporation by an O-helix mutant of *Thermus aquaticus* DNA polymerase I. Mutat Res 485:197–207.

Ondrechen MJ, Clifton JG, Ringe D. 2001. THEMATICS: A simple computational predictor of enzyme function from structure. Proc Natl Acad Sci U S A 98:12473–12478.

Pages V, Fuchs RP. 2002. How DNA lesions are turned into mutations within cells? Oncogene 21:8957–8966.

Patel PH, Loeb LA. 2000. Multiple amino acid substitutions allow DNA polymerases to synthesize RNA. J Biol Chem 275:40266–40272.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612.

Porter CT, Bartlett GJ, Thornton JM. 2004. The Catalytic Site Atlas: A resource of catalytic sites and residues identified in enzymes using structural data. Nucleic Acids Res 32:D129–133.

Prakash S, Johnson RE, Prakash L. 2005. Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function. Annu Rev Biochem 74:317–353.

Sankararaman S, Kolaczkowski B, Sjölander K. 2009. INTREPID: A web server for prediction of functionally important residues by evolutionary analysis. Nucleic Acids Research 37:W390–W395.

Segel I. 1993. Enzyme Kinetics: Behavior and analysis of rapid equilibrium and steady-state enzyme systems. New York, NY:Wiley-Interscience.

Shen X, Sayer JM, Kroth H, Ponten I, O'Donnell M, Woodgate R, Jerina DM, Goodman MF. 2002. Efficiency and accuracy of SOS-induced DNA polymerases replicating benzo[a]pyrene-7,8-diol 9,10-epoxide A and G adducts. J Biol Chem 277:5265–5274.

Sherrer SM, Beyer DC, Xia CX, Fowler JD, Suo Z. 2010. Kinetic basis of sugar selection by a Y-family DNA polymerase from *Sulfolobus solfataricus* P2. Biochemistry 49:10179–10186.

Shurtleff BW, Ollivierre JN, Tehrani M, Walker GC, Beuning PJ 2009. Steric gate variants of UmuC confer UV hypersensitivity on *Escherichia coli*. J Bacteriol 191:4815–4823.

Somarowthu S, Brodkin HR, D'Aquino JA, Ringe D, Ondrechen MJ, Beuning PJ. 2011a. A tale of two isomerases: Compact versus extended active sites in ketosteroid isomerase and phosphoglucose isomerase. Biochemistry 50:9283–9295.

Somarowthu S, Yang H, Hildebrand DG, Ondrechen MJ. 2011b. High-performance prediction of functional residues in proteins with machine learning and computed input features. Biopolymers 95:390–400.

Suzuki M, Yoshida S, Adman ET, Blank A, Loeb LA. 2000. Thermus aquaticus DNA polymerase I mutants with altered fidelity. Interacting mutations in the O-helix. J Biol Chem 275:32728–32735.

Tanioka M, Masaki T, Ono R, Nagano T, Otoshi-Honda E, Matsumura Y, Takigawa M, Inui H, Miyachi Y, Moriwaki S, Nishigori C. 2007. Molecular analysis of DNA polymerase eta gene in Japanese patients diagnosed as xeroderma pigmentosum variant type. J Invest Dermatol 127:1745–1751.

Tong W, Wei Y, Murga LF, Ondrechen MJ, Williams RJ. 2009. Partial order optimum likelihood (POOL): Maximum likelihood prediction of protein active site residues using 3D Structure and sequence properties. PLoS Comput Biol 5:e1000266.

Wagner J, Gruz P, Kim SR, Yamada M, Matsui K, Fuchs RP, Nohmi T. 1999. The dinB gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. Mol Cell 4:281–286.

Walsh JM, Bouamaied I, Brown T, Wilhelmsson LM, Beuning PJ. 2011a. Discrimination against the cytosine analog tC by *Escherichia coli* DNA polymerase IV DinB. J Mol Biol 409:89–100.

Walsh JM, Hawver LA, Beuning PJ. 2011b. *Escherichia coli* Y family DNA polymerases. Front Biosci 17:3164–3182.

Wei Y, Ko J, Murga L, Ondrechen MJ. 2007. Selective prediction of interaction sites in protein structures with THEMATICS. BMC Bioinformatics 8:119.

Weimer KM, Shane BL, Brunetto M, Bhattacharyya S, Hati S. 2009. Evolutionary basis for the coupled-domain motions in *Thermus thermophilus* leucyl-tRNA synthetase. J Biol Chem 284:10088–10099.

Yamtich J, Sweasy JB. 2010. DNA polymerase family X: function, structure, and cellular roles. Biochim Biophys Acta 1804:1136-1150.

Yang G, Franklin M, Li J, Lin TC, Konigsberg W. 2002. A conserved Tyr residue is required for sugar selectivity in a Pol alpha DNA polymerase. Biochemistry 41:10256–10261.

Yuan B, Cao H, Jiang Y, Hong H, Wang Y. 2008. Efficient and accurate bypass of N^2-(1-carboxyethyl)-2′-deoxyguanosine by DinB DNA polymerase *in vitro* and *in vivo*. Proc Natl Acad Sci U S A 105:8679–8684.