2418–2426 *Nucleic Acids Research, 2001, Vol. 29, No. 11*

Enzymatic properties of rat DNA polymerase β mutants obtained by randomized mutagenesis

Adonis Skandalis* and Lawrence A. Loeb[1]

The Department of Biological Sciences, Brock University, 500 Glenridge Avenue, St Catharines, Ontario L2S 3A1, Canada and [1]The Joseph Gottstein Memorial Cancer Research Laboratory, Departments of Pathology and Biochemistry, University of Washington School of Medicine, Box 357705, Seattle, WA 98195-7705, USA

Received December 1, 2001; Revised March 28, 2001; Accepted April 12, 2001

ABSTRACT

We have used random sequence mutagenesis to generate mutants of DNA polymerase β in an effort to identify amino acid residues important for function, catalytic efficiency and fidelity of replication. A library containing 100 000 mutants at residues 274–278 in the N-helix of the thumb subdomain of the polymerase was constructed and screened for polymerase activity by genetic complementation. The genetic screen identified 4000 active pol β mutants, 146 of which were sequenced. Each of the five positions mutagenized tolerated substitutions, but residues G274 and F278 were only found substituted in combination with mutations at other positions. The least conserved residue, D276, was replaced by a variety of amino acids and, therefore, does not appear to be essential for function. Steady-state kinetic analysis, however, demonstrated that D276 may be important for catalytic efficiency. Mutant D276E exhibited a 25-fold increase in catalytic efficiency over the wild-type enzyme but also a 25-fold increase in G:T misincorporation efficiency. We present a structural model that can account for the observations and we discuss the implications of this study for the question of enzyme optimization by natural selection.

INTRODUCTION

Many structural and sequence elements of DNA polymerases are highly conserved across kingdoms, from viruses to mammals (1–4). This high degree of conservation is generally interpreted as evidence that stringent evolutionary pressures have led to the optimization of polymerase enzymatic properties such as replication fidelity, catalytic efficiency and processivity. Is it possible then to engineer polymerases with significantly improved performance over the wild-type variant or are wild-type polymerases already as efficient as they can be? The answer is unclear because of our poor understanding of how exhaustively natural selection can explore new variants, particularly those resulting from multiple mutations (5–7). Moreover, there is uncertainty about the exact nature of the selection pressures on polymerases in the context of their cellular function to maintain homeostasis. For example, the catalytic efficiency of replicative polymerases in bacteria may be determined and restricted not by inherent limitations of the enzymes, but by the synchronization requirements necessitated by the passive coupling of DNA replication and cell division (8). Likewise, DNA-repair-associated polymerases may be under selection pressure to exhibit reduced fidelity in order to by-pass potentially lethal DNA lesions (9).

An experimental approach for addressing these issues is the use of random mutagenesis to generate large numbers of mutant proteins (10,11). It involves the replacement of wild-type nucleotide sequences with randomized sequences in plasmid-encoded genes. Random mutagenesis, coupled with positive genetic selection for functional polymerases, is a powerful simulation of natural evolution and one in which the selection pressures are well controlled. It can elucidate the effect on enzymatic function of both single and multiple amino acid substitutions, and it allows the identification of immutable residues as well as residues that can tolerate specific substitutions and, thus, it can provide important empirical confirmation to mechanistic models.

In order to investigate the enzymatic potential of DNA polymerases and assess the evolutionary pressures acting upon them, we have performed random mutagenesis on the rat DNA polymerase β (pol β). Pol β is a DNA polymerase whose primary role is thought to be gap-filling in base excision repair and the removal of 5′-deoxyribose phosphate (12,13). It is the smallest known eukaryotic polymerase (39 kDa) and does not have any proofreading exonuclease activity. The structure of pol β is well understood based on crystallographic studies with a variety of substrates (4,14–18). Pol β consists of a 31 kDa C-terminal domain with nucleotidyl transfer activity and an 8 kDa domain with 5′-deoxyribose phosphodiesterase, lyase, activity (4,13). Like other polymerases, the 31 kDa domain can be visualized as thumb, fingers and palm sub-domains (2,19,20). The small size of pol β and the lack of exonuclease activity, as well as the wealth of information about its structure, have made it an important model for the study of mechanisms of DNA replication and fidelity and have enabled the identification of amino acid residues potentially critical for catalysis. Site-directed mutagenesis of residues such as Y271, N279 and R283 results in mutants with reduced catalytic efficiency and mostly lower fidelity (21–23). The use of genetic screens for lower replication fidelity mutants has also

*To whom correspondence should be addressed. Tel: +1 905 688 5550; Fax: +1 905 688 1855; Email: askandal@spartan.ac.brocku.ca

enabled the identification of other amino acid substitutions that reduce the fidelity of DNA synthesis by pol β (24,25). Pol β mutants with significantly higher catalytic efficiency have not, to our knowledge, been reported.

In the present study, we have randomized the nucleotide sequence of residues 274–278 located in the N-helix (15) in order to systematically analyze the role of this region in polymerase function and to identify mutants with increased catalytic efficiency. N-helix residues, located in the thumb subdomain of pol β, are important for function and are in close contact with the incoming dNTP (15). We have taken advantage of a genetic screen based on the ability of pol β to complement the function of the bacterial polymerase I (pol I) to identify mutants that maintain activity and can support bacterial growth as well as the wild-type pol β does (26). Approximately 4000 pol β mutants (4%) of the library were functional as indicated by the complementation assay. This analysis enabled us to identify residues important for polymerase function and also identify a mutant, D276E, with increased catalytic efficiency but also decreased fidelity relative to the wild-type. Overall, this study supports the hypothesis that in the pol β sites studied, the wild-type residues that evolved by natural selection effect an optimal equilibrium among the polymerase's different enzymatic activities, such as fidelity and catalytic efficiency. Improvement in the performance in one property may lead to deterioration in another.

MATERIALS AND METHODS

Bacterial strains, growth media and plasmids

Escherichia coli B/r strain SC18-12 with genotype *recA718 polA12 uvrA155 trpE65 lon-11 sulA1* was used for the pol β complementation studies. *Escherichia coli* strain DH5α [*deoR endA1 gyrA96 phoA hsdR17* ($r_K^-m_K^+$), *recA1 relA1 supE44 thi-1* Δ(*lacZYA-argFU169*), ϕ*80dlacXM15*, *F*$^-$, λ$^-$] was used in cloning experiments and plasmid analysis. *Escherichia coli* strain BL21(DE3) [*F-ompT hsdSB*(r_B^- m_B^-) *gal dcm* (DE3)] was used for pol β over-expression. The *E.coli* strains NR9099, MC1061 and CSH50 F′ used in the M13 mutation assay have genotypes as described previously (27). Bacterial cultures were grown in Luria–Bertani (LB) medium unless otherwise indicated. Cells used for genetic complementation were grown in nutrient broth or nutrient agar prepared according to the manufacturer's instructions (Difco) and supplemented with chloramphenicol (30 μg/ml), tetracyclin (12.5 μg/ml) and isopropyl β-D-thiogalactoside (IPTG, 1 mM) when appropriate. The plasmid pβL is a pHSG576 derivative containing the entire rat pol β cDNA (28). Pol β sequences were subcloned between the *Nde*I and *Sal*I sites of the expression vector pET28a (Novagen). The vector provided a C-terminal affinity histidine tag used for protein purification. Initial analysis of the pol I and pol β activities of the bacterial strains used was performed as described previously (29). Briefly, samples of 1×10^6 cells were induced by IPTG, lysed by boiling and electrophoresed on 12% SDS–polyacrylamide gel containing gapped DNA. The polymerases in the gel were renatured at 37°C, [α-^{32}P]dCTP was added and incorporation was visualized by autoradiography.

Random mutagenesis and complementation assay

To avoid contamination due to incomplete cleavage by restriction endonucleases during the generation of the randomized library of pol β, we used a modified pβL vector which carries an inactive copy of pol β. Specifically, we generated a pβL derivative vector (pβLD) which carries a pol β gene that has been inactivated by the substitution of the region between the *Cla*I and *Sph*I site with a 500 bp *Cla*I/*Sph*I fragment from pBR322. The oligonucleotide POLB2R, 5′-CCA GGG CAT GCG CTC TCA TAT TCT TAT T(A/N A/N A/N G/N A/N T/N G/N T/N C/N A/N C/N T/N T/N C/N C/N)AG TGA AGT AGA GAA CAC CAC-3′, synthesized by Keystone Laboratories contained, in the sequence in parentheses, 73% of the wild-type nucleotide, indicated by the letter code, and 9% each of the other, indicated by N. The partially complementary oligonucleotide POLB2F, 5′-CAC AGG AGA ATC GAT ATC AGG TTG ATC CCC AAA GAT CAG TAC TAC TGT GGT GTT CTC TAC TT-3′ was synthesized by Operon Technologies. Equimolar amounts (20 pmol) of the two oligonucleotides were mixed together and extended by 25 cycles of PCR (95°C for 30 s, 55°C for 30 s and 72°C for 45 s) to a full duplex DNA of 109 bp using a Promega PCR kit according to the manufacturer's instructions. The product DNA was digested with *Cla*I and *Sph*I and it was subsequently gel purified and ligated using T4 DNA ligase into a *Cla*I/*Sph*I pβLD restricted vector to generate the pol β random library. The ligation reaction mixture was then ethanol precipitated and used to transform electrocompetent DH5α cells by an electroporator (Bio-Rad). To estimate the library size, an aliquot of the transformed cells was plated in LB agar plates with chloramphenicol at 37°C. The remaining cells were inoculated into 500 ml of LB liquid medium with chloramphenicol and were used for plasmid preparation.

DNA sequencing, protein purification and activity assays

DNA sequencing was carried out using the fmol cycle sequencing system (Promega) using [γ-^{32}P]-end-labeled primers. Pol β proteins were overexpressed and purified using the His-Tag system (Novagen). Mutant pol β sequences were amplified using oligonucleotide primers that generated the novel restriction sites *Nde*I and *Sal*I. The primers used were HISNDE, 5′-GGC TCG AGC ATA TGC TCG TGG AAC TCG CAA ACT TTG AGA AGA ACG TGA GCC AGG CGA TCC ACA AGT-3′, and HISSAL2, 5′-CTA GTC GAC TCA TTC ACT CCT GTC CTT GGG CTC CCG GTA-3′. High fidelity *Pfu* DNA polymerase was used for this procedure according to the manufacturer's instructions. The PCR products were gel purified, restricted by digestion with *Nde*I and *Sal*I, and were subcloned into the vector pET28a as described above. The nucleotide sequence of the entire pol β gene of individual clones was determined to verify that no mutations were introduced into the gene during this procedure. Pol β protein was purified by Nickel chelation chromatography according to the manufacturer's suggestions with the following modifications. Cells were lysed by freezing for 12 h at –70°C and slow thawing on ice in the presence of 10 mg/ml lysozyme. The stringency of the washes of the non-specifically adsorbed proteins was increased by using 100 mM imidazole solution. All the preparations consisted of >95% pol β as estimated by visual inspection of Coomassie Blue-stained 12% SDS–polyacrylamide

gels. Amounts of proteins were estimated by the Bradford assay (Bio-Rad) using a bovine serum albumin standard. The polymerase activity of the mutants was assayed using activated calf thymus DNA as described previously (30).

Primer extension assay with gap substrates

A 46mer oligonucleotide template (P607G) 5′-G CGC GGA AGC TTG GCT GCA GAA TAT TGC TAG CGG GAA TTC GGC GCG-3′ annealed to a 15mer (P637) 5′-CGC GCC GAA TTC CCG-3′ ^{32}P-labeled at the 5′ end, and to a downstream primer (GAP1P) 5′-P-CAA TAT TCT GCA GCC AAG CTT CCG CGC-3′, designed to generate a 4 nt gap opposite the GATC template sequence. Template, gap oligonucleotide (1.5 pmol of each) and labeled primer (0.8 pmol) were heated to 75°C for 5 min, 55°C for 10 min, 37°C for 2 min and then placed on ice. Annealing results in the following gapped template construct:

5′-CGCGCCGAATTCCCG P-CAATATTCTGCA-GCCAAGCTTCCGCGC-3′

3′-GCGCGGCTTAAGGGC *G* A T C GTTATAAGACGTC-GGTTCGAAGGCGCG-5′.

The gapped template was visualized using a 20% non-denaturing polyacrylamide gel. The assay mixture for DNA synthesis contained 50 mM Tris–HCl pH 8.5, 8 mM $MgCl_2$, 100 mM KCl, 3 mM DTT, 200 µg/ml bovine serum albumin, 0.07–0.7 nM pol β and 260 µM dNTPs in a final volume of 15 µl. In some experiments dCTP was excluded to assess the ability of the polymerases to misincorporate and misextend past G. The reactions were carried out at 37°C for 5 min and were terminated by the addition of 1.5 µl of 500 mM EDTA. Reaction products were resolved in a 14% denaturing polyacrylamide gel.

M13mp2 *lacZ* forward mutation fidelity assay

The mutation frequencies of mutant and wild-type polymerases were assayed as described previously (27). The non-coding strand of the *lacZ*α gene contained in 200 ng of gapped M13mp2 DNA was copied by pol β in a reaction mixture containing 50 mM Tris–HCl pH 8.5, 8 mM $MgCl_2$, 20 µM each dNTP, 100 mM KCl, 3 mM DTT and 200 µg/ml bovine serum albumin. After incubation at 37°C for 15 min, the DNA was transfected and the blue or white plaques were scored.

Single nucleotide incorporation and steady-state kinetics

This assay was performed as described for the primer extension assay, except a single nucleotide was added at various concentrations. The products of the reaction were resolved by electrophoresis in a denaturing 14% polyacrylamide gel. The extent of elongation was quantified by PhosphorImager analysis (Molecular Dynamics, #400S). The apparent k_m and V_{max} values were calculated from Lineweaver–Burk plots (31).

RESULTS

Generation of a pol β plasmid library

The nucleotides spanning positions 822–836 of the pol β gene in the plasmid pβL were replaced by a randomized sequence comprising, at each position, 73% wild-type nucleotides and 9% of each of the other three nucleotides. This construction was designed to yield 1% wild-type sequence at the amino acid level. The procedure generated ~1 × 10^5 plasmid-borne, pol β mutants (Table 1).

Table 1. Summary of pol β library parameters before and after selection by genetic complementation

Size of randomized region	15 nucleotides
Randomization at each nucleotide	73% wild-type residue, 27% other
Unselected library size	100 000
Unselected mutants sequenced	22
Average amino acid changes per mutant	2.4
Complementing mutants	4000
Complementing mutants sequenced	146
Average amino acid changes per complementing mutant	0.9

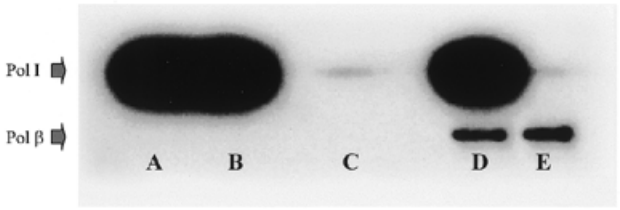

Figure 1. Polymerase gel activity assay. The band location corresponds to the electrophoretic mobility of the protein DNA pol I or β (as indicated) and the band area corresponds to the amount of polymerase activity. (A) DNA pol I activity in the bacterial strain DH5α in the absence of IPTG. (B) DNA pol I activity in the bacterial strain DH5α in the presence of IPTG. (C) Bacterial strain SC18-12 grown at the non-permissive temperature 37°C in the absence of IPTG. Only residual activity of DNA pol I is detectable. (D) Bacterial strain DH5α transfected with a vector containing the wild-type pol β, grown in the presence of IPTG. The activities of both pol I (top) and pol β (bottom) are detectable. (E) Bacterial strain SC18-12 grown at 37°C in the presence of IPTG. The activity of pol β is detectable.

Genetic complementation and sequence analysis of mutants

Active mutants were identified by a genetic complementation assay based on the *E.coli* strain SC18-12 which is unable to grow at 37°C unless pol β can substitute for the temperature-sensitive pol(A) gene (26). To confirm that pol β is expressed in these cells and that pol I is temperature sensitive we conducted a polymerase activity assay. DH5α, the bacterial strain used in the plasmid library manipulations in some experiments, exhibited a prominent band with polymerase activity that corresponded to the electrophoretic mobility of DNA pol I (Fig. 1A). Upon addition of IPTG no other band with polymerase activity was observed (Fig. 1B). After transfection with the wild-type pol β construct and induction with IPTG, two bands of polymerase activity were detected, one corresponding to pol I and the other to pol β (Fig. 1D). SC18-12 exhibited negligible pol I activity at 37°C (Fig. 1C) but pol β was expressed in these cells under IPTG control (Fig. 1E).

The plasmids harboring the random sequence were introduced into the bacteria which were then grown at 30 or 37°C.

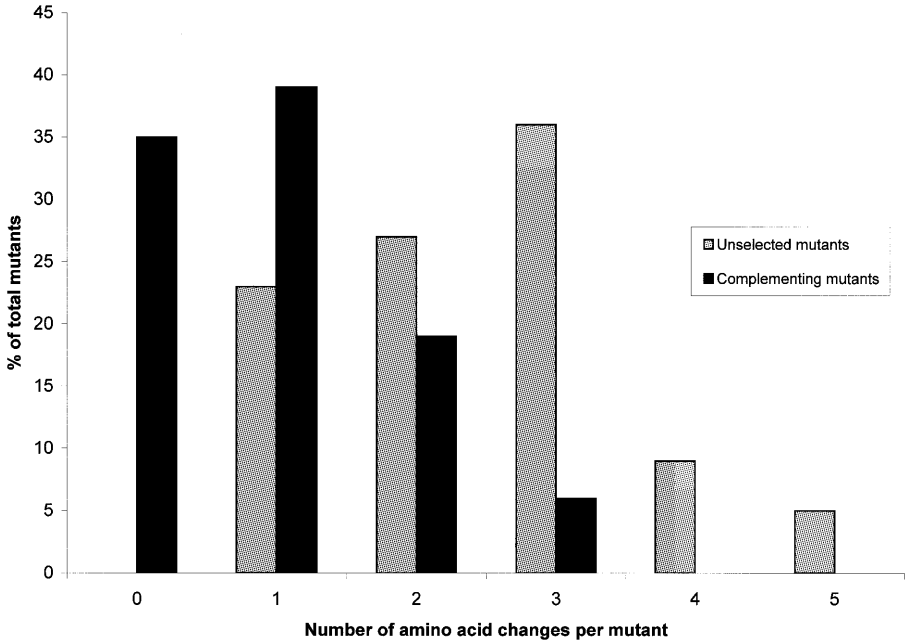

Figure 2. Distribution of the number of amino acid changes per pol β mutant in the library generated by random mutagenesis, before and after screening by genetic complementation.

Nucleotide sequencing of 22 unselected mutants, grown at 30°C, revealed that there were, on average, 2.4 amino acid substitutions per mutant in the library (Fig. 2). None of the unselected mutants sequenced contained the wild-type nucleotide sequence. Approximately 4% of the cells formed colonies at the non-permissive temperature of 37°C indicating that 96% of the mutant polymerases were unable to complement the pol I functions (Table 1). The degree of complementation can be estimated in this system by the cloning efficiency, per μg of DNA, and the colony size of transformed isolates. Only pol β mutants that complimented growth at least 75% as well as the wild-type pol β were designated as complementing. This limit was experimentally determined to be optimal in order to minimize false positive results. We have sequenced 146 complementing pol β mutant genes which had, on average, 0.9 amino acid substitutions. Approximately one-third (52) of the complementing mutants had no changes at the amino acid level, 58 had one substitution, 28 had two substitutions and eight had three substitutions (Fig. 2). At the nucleotide level, 41 isolates had wild-type sequences. The most conserved amino acid was G274 which was mutated only in mutants with triple substitutions (Fig. 3). The least conserved amino acid was D276 which was found mutated in 60 of the 146 mutants. One particular mutant, D276E, represented >10% of the complementing mutants analyzed at the amino acid level. Moreover, this mutant was the only one superior to the wild-type in the degree of complementation, as judged by both bacterial colony size and cloning efficiency. Substitutions at the other amino acids were generally conservative. S275 was found substituted only by C or N which are hydrophilic and uncharged. I277 was found mostly substituted by hydrophobic uncharged residues such as L and V. F278 was only found substituted by hydrophobic residues in double or triple mutants.

M13mp2 forward mutation assay

The mutation frequency of the wild-type polymerase was 1.6×10^{-2}. Polymerase D276A/I277L had a mutation frequency of 7×10^{-3} and D276E a mutation frequency of 2.5×10^{-2} (Table 2).

Primer extension

A primer/template oligonucleotide substrate containing a single-strand gap of 4 nt was used to determine the functional characteristics of the mutant polymerases. Polymerases were added to the primer/template at concentrations that yield equal polymerase activities as determined by activated calf thymus DNA activity assays. We monitored the ability of each mutant polymerase to replicate the template sequence in the presence of all four dNTPs. We also monitored the ability of polymerases to replicate the template in the absence of dCTP which is the first nucleotide to be incorporated (Fig. 4) to estimate replication fidelity. In fidelity experiments the concentration of polymerase was varied to produce equal extension when all four dNTPs were present. Fourteen mutant purified polymerases were tested: D276E, A, G, V, K and Q; L277E, L and V; S275C/D276A, D276A/L277L, D276A/L277N, D276H/I277R/F278Y and D276H/I277K/F278Y. All tested polymerases were able to perform DNA replication to completion. In the absence of dCTP, all but one of the polymerases tested misincorporated a non-complementary nucleotide opposite G to a lesser extent than the wild-type pol β. The highest fidelity was exhibited by pol β D276A/I277L. The only polymerase exhibiting lower fidelity of misincorporation than wild-type was pol β D276E.

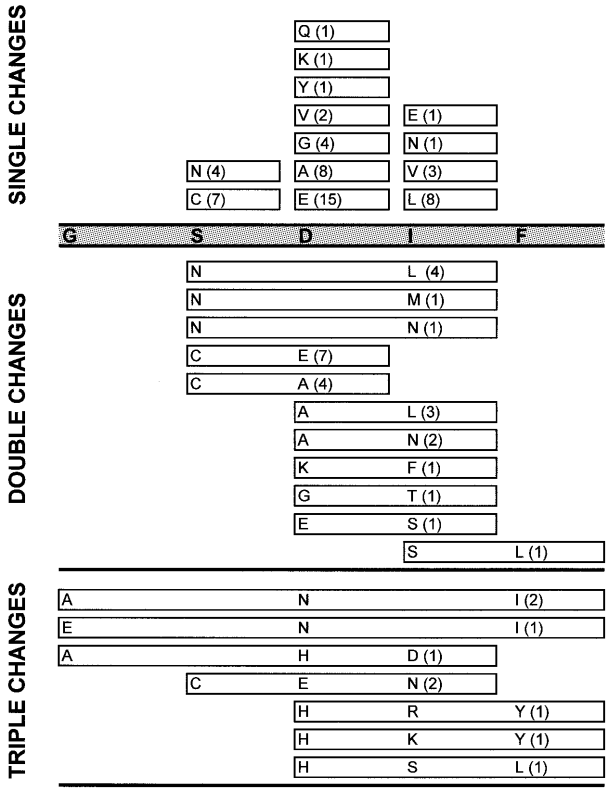

Figure 3. Amino acid substitutions characterized in complementing mutants. The wild-type sequence is shaded. Above the wild-type sequence are listed mutants with single amino acid substitutions and below are listed mutants with double and triple amino acid substitutions. The number in parentheses indicates the number of times that particular substitution was recovered among the 146 mutants sequenced.

Steady-state kinetics

Catalytic efficiency (k_{cat}/K_m). We performed steady-state kinetic analysis of single-nucleotide incorporation (dCTP, dGTP, dTTP or dATP) opposite a template G by the wild-type and the D276A/I277L and D276E mutant pol βs (Table 3). (i) dCTP:G. The wild-type enzyme exhibited K_m of 21 μM and a k_{cat} of 1.7 s^{-1} for single dCTP incorporation. The kinetic values for the wild-type enzyme are in good agreement with those published previously (22,25,32). The K_m of mutant D276A/I277L was 10-fold higher (217 μM) and the k_{cat} 4-fold lower (0.40 s^{-1}). As a result, the catalytic efficiency (k_{cat}/K_m) of D276A/I277L was 40-fold lower (8×10^3 $s^{-1}M^{-1}$) than the wild-type (2×10^4 $s^{-1}M^{-1}$). Mutant D276E also had a lower k_{cat} (0.53 s^{-1}) than wild-type but the 100-fold lower K_m (0.28 μM) boosted its catalytic efficiency to 25-fold higher than wild-type. (ii) dTTP:G. In the G:T mismatch, the wild-type enzyme exhibited catalytic efficiency of 2×10^{-1} $s^{-1}M^{-1}$. D276A/I277L exhibited 25-fold lower catalytic efficiency, whereas D276E exhibited 400-fold higher catalytic efficiency. (iii) dATP:G. In the A:G mismatch, the wild-type enzyme exhibited catalytic efficiency of 3×10^{-3} $s^{-1}M^{-1}$. D276A/I277L exhibited 3-fold lower catalytic efficiency, whereas D276E exhibited 1000-fold higher catalytic efficiency. (iv) dGTP:G. In the G:G mismatch, the wild-type enzyme exhibited catalytic efficiency of 6×10^{-4} $s^{-1}M^{-1}$. D276A/I277L exhibited essentially no difference in catalytic efficiency whereas D276E exhibited 1167-fold higher catalytic efficiency.

Table 2. Mutational frequencies (MF) of mutant and wild-type pol β in the M13mp2 gap system

	WT	D276E	D276A/I277L
Number of white colonies	20	32	5
Number of light blue colonies	36	54	9
Total number of colonies	1898	2015	1154
MF ($\times 10^{-2}$)[a]	1.6	2.5	0.7

WT, Wild-type.
[a]Corrected for no-polymerase controls

Misincorporation efficiency f_{ins}. Misincoporation efficiency is the ratio of the catalytic efficiency of incorrect nucleotide incorporation over the catalytic efficiency of the correct nucleotide incorporation (Table 4). f_{ins} for D276A/I277L was no different to the wild-type for G:T mismatch but it was 12.5- and 40-fold higher for the G:A and G:G mismatches, respectively. D276E was essentially no different to the wild-type for G:A or G:G mismatches but it was 25-fold higher for G:T.

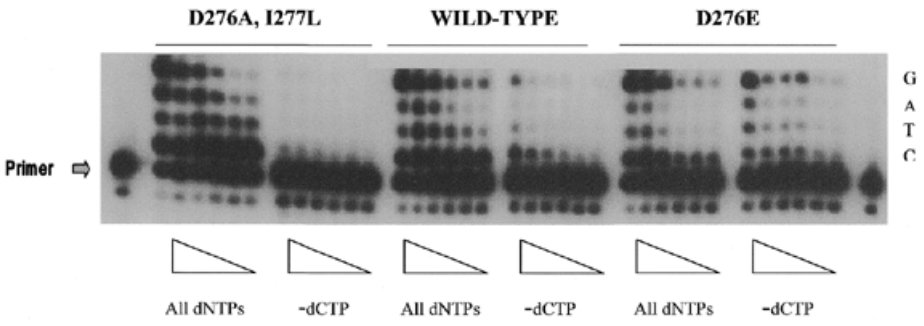

Figure 4. Primer extension assay on a 4 nt gap opposite a GATC template sequence in the presence and absence of dCTP. The sequence of the growing strand is indicated. The triangles are indicating decreasing concentration of dNTPs from left to right in all lanes.

Table 3. Steady-state kinetic parameters for single dNTP, correct or incorrect, incorporation into a gapped oligo substrate by wild-type, mutant D276A/I277L and mutant D276E pol β

Pol β	$K_{m(app)}$ (μM)[a]	k_{cat} (s^{-1})[b]	k_{cat}/K_m ($s^{-1}M^{-1}$)
dCTP:G			
WT	21.0 (±1.9)	1.7	8×10^4
D276A/I277L	217 (±18)	0.40	2×10^3
D276E	0.28 (±0.01)	0.53	2×10^6
dTTP:G			
WT	134 (±9)	2×10^{-5}	2×10^{-1}
D276A/I277L	230 (±19)	2×10^{-6}	8×10^{-3}
D276E	5 (±0.7)	4×10^{-4}	8×10^1
dATP:G			
WT	500 (±51)	2×10^{-6}	3×10^{-3}
D276A/I277L	343 (±27)	4×10^{-7}	1×10^{-3}
D276E	9 (±0.5)	3×10^{-5}	3.1
dGTP:G			
WT	273 (±15)	2×10^{-7}	6×10^{-4}
D276A/I277L	330 (±18)	2×10^{-7}	5×10^{-4}
D276E	11 (±0.4)	8×10^{-6}	7×10^{-1}

WT, wild-type.
[a]Mean values of three replicates. Numbers in parentheses indicate the standard deviation.
[b]Calculated using total protein concentration.

Table 4. Misincorporation efficiency (f_{ins}) of a single dNTP opposite G in a gapped substrate by wild-type, mutant D276A/I277L and mutant D276E pol β

	Mispair		
	G:T	G:A	G:G
f_{ins}[a]			
WT	2×10^{-6}	8×10^{-5}	1×10^{-5}
D276A/I277L	3×10^{-6}	1×10^{-3}	4×10^{-4}
D276E	5×10^{-5}	4×10^{-5}	9×10^{-6}
Fold increase in the f_{ins} over wild-type			
D276A/I277L	1.5	12.5	40
D276E	25	0.5	0.9

[a]f_{ins} is the ratio of $k_{cat}/K_{m(app)}$ for the incorrect versus the correct nucleotide. $k_{cat\,(rel)}/K_{m(app)}$ values are reported in Table 3.

DISCUSSION

Elucidating why enzymes have the form and properties observed today is almost an impossible task. A biochemist may define an 'optimized' enzyme as one with a turnover equal to the diffusion rate of the substrate, but enzyme evolution is shaped not only by catalytic efficiency but also by enzyme stability, substrate specificity and flexibility to respond to regulatory elements. Yet another factor that may affect the current form of enzymes is historical contingency. The sequence of the ancestral form may impose constraints and limitations on the present form since the rate of mutation allows natural selection to explore mostly single mutations. We have used random mutagenesis of the dNTP binding domain of the rat DNA pol β to explore whether the wild-type residues that evolved by natural selection offer optimal enzymatic activities, such as fidelity and catalytic efficiency.

Random mutagenesis is a technique that offers not only the capability to explore multiple mutations but also allows us to control the nature of selection pressures on the polymerase. In the cell, even a small enzyme like pol β has a variety of functions. The principal function of DNA pol β seems to be in 'short patch' base excision repair (33), but it may also play a role during 'long patch' base excision repair (34,35), meiosis (36), nucleotide excision repair (37), and possibly in DNA replication. The pol β lyase activity, which removes the 5′-deoxyribose phosphate during gap repair, seems to be at least as important as its replicative activity (13).

The natural substrate of pol β during short patch base excision repair is a single nucleotide gapped substrate with a 5′-PO_4 at the downstream margin of the gap. The 5′-PO_4 appears to interact with the 8 kDa domain of the polymerase which also interacts with downstream nucleotides in a non-sequence-specific manner (38). This interaction is necessary for the polymerase to fill the gap processively. It has been demonstrated that the catalytic efficiency (k_{cat}/K_m) is >2500 times greater than in templates without a gap (32,39). The 8 kDa interaction with downstream elements is also thought to affect fidelity by assisting in the bending of the template 90° at the active site (18). Pol β is 10–100-fold more likely to misinsert a nucleotide in a non-gapped substrate (32). Initial experiments during this study confirmed these observations and only data obtained using gapped substrates are reported here.

Another important mechanistic feature of pol β is the flexible thumb subdomain. The thumb subdomain is hinged along the axis of the M helix and it 'closes' when the ternary complex of pol β, template/primer and dNTP forms. This movement brings the residues of the N-helix into very close proximity with the incoming dNTP and particularly residue D276. It has been reported that this residue may come within 3.5 Å of the C4 of the base of the incoming dNTP but crystal data are ambiguous about its exact position (16,18). The close proximity of the N-helix to the incoming dNTP indicated to us that this site may have significant influence on polymerase catalytic efficiency and replication fidelity. To explore the role of this site on polymerase function in detail, we used random mutagenesis to substitute residues G274, S275, D276, I277 and F278. To identify mutants exhibiting increased catalytic efficiency, only mutants that supported bacterial growth in the complementation screen as well as the wild-type pol β were considered active.

Random library

We constructed a library of 1×10^5 mutants in a segment spanning nucleotides 822–834. The library was screened for the ability of mutants to complement the bacterial pol I *in vivo*. Using this genetic selection protocol, most (96%) of the mutants in this region were unable to complement pol I and therefore were considered functionally inactive. The large

number of inactive mutants observed is consistent with the hypothesis that this region is indeed important for pol β function (15,22). Sequence analysis of 146 complementing mutants revealed that 35% of the mutants were wild-type at the amino acid level. The high incidence of wild-type sequences reflects, in part, the design parameters of the library, since the degree of randomization was chosen so as to generate a library that contained 1% wild-type sequences at the amino acid level. Therefore, given that 4% of the library was able to complement, at least 25% of the mutants were expected to be wild-type. On average there were only 0.9 amino acid substitutions per mutant. This is considered further confirmation that the complementation assay was effective in discriminating between active and inactive pol βs, since the unselected library contained 2.4 amino acid substitutions per mutant.

It should be noted that an efficient genetic selection to identify mutants that maintain function is essential for studies of structure and function since a large number of mutants are generated, the majority non-functional. Functional complementation in bacterial systems offers the best, and often the only, practical method of enzymatic activity screening. However, the interactions and function of a eukaryotic enzyme in a bacterial system are not always well understood and additional assays are needed to evaluate enzymatic function. In this study, all enzymes that complemented pol I for bacterial growth were shown to be active in assays of polymerase activity. However, it has not been ascertained whether some mutants that would be active in eukaryotic cells failed to complement bacterial growth.

Amino acid substitutions

The most conserved residue observed was G274 (Fig. 3). This side chain is thought to form a *cis*-peptide bond which acts as a hinge between helices M and N (16). Considering the postulated movement of this region by as much as 12 Å during polymerization, it is perhaps not surprising that a hinge residue between helices M and N such as G274 is very conserved (18). Nonetheless, it can be replaced in triple substitution combinations which also contain either D276N or D276H mutations. It is not apparent how these mutations can compensate for substitution of the G274. Residue S275 was substituted only by the hydrophilic and non-charged amino acids C and N, demonstrating the requirement for these properties at the site. Based on its proximity to the incoming dNTP, it was unexpected that the least conserved residue was D276. The conservative substitution D276E which maintains a charged, hydrophilic amino acid at this position, comprised 10% of all the complementing mutants. D276 was also found substituted in most single and double mutants and in all triple mutants characterized (Fig. 3). Hence, despite its close proximity to the incoming dNTP, D276 is clearly not required for activity. I277 was also poorly conserved overall, but the predominant single substitution was a conservative replacement, I277L. All double and triple mutants had D276 or I277 substituted. F278 was the second most immutable residue next to G274.

These results demonstrate the ability of random mutagenesis, coupled with a strong genetic screening assay, to provide information about how essential specific amino acid residues are for enzyme activity and the ability of multiple mutations to compensate for substitutions at immutable residues. They also offer insights into the amino acid properties important for function.

Characterization of mutants

The analysis of amino acid residue mutability revealed that positions D276 and I277 had few constraints on permitted substitutions and were clearly not essential for function. Such analysis, however, cannot reveal the influence of these residues on catalytic efficiency and replication fidelity. For example, mutant D276E was superior to the wild-type pol β in supporting bacterial growth in the complementation assay. This observation led us to further explore the impact of mutations on D276 and I277 on the polymerase function in order to understand the structure–function relationship at these sites. Fourteen mutant enzymes were purified and used to extend a primer over a 4 nt template gap in the absence of dCTP, the first nucleotide templated for. Several mutants at the D276 position were less efficient than the wild-type in misinserting another nucleotide opposite G and extending the mismatch to completion (data not shown). The only mutant identified in this screen that showed higher efficiency of replication past the G than the wild-type was D276E. This was unexpected because of the conservative nature of this substitution; the main difference between D and E is the slightly larger size of E by ~1.5 Å. D276E was the single most numerous mutant in the library (15% of all the mutants with amino acid changes). The least efficient enzyme in replicating the gap in the absence of dCTP was D276A/I277L (Fig. 4). The relative efficiencies of mutants D276E and D276A/I277L in misincorporation and misextension were also reflected in the mutation frequencies observed in the forward M13mp2 assay with all four dNTPs present although the differences were very small (Table 2). Overall, the mutation frequencies observed were higher than those previously reported (22) but nonetheless the D27E exhibited relatively higher mutation frequency than the wild-type whereas the D276A/I277L exhibited relatively lower mutation frequency compared to the wild-type.

Steady-state kinetic analysis of single nucleotide incorporation (dCTP), opposite a template G by the wild-type pol β and the D276A/I277L and D276E mutants, yielded surprising results. D276E exhibited a 25-fold increase in catalytic efficiency over the wild-type. To the best of our knowledge this is the highest reported increase in catalytic efficiency for any polymerase mutant (7). The most significant difference between D276E and the wild-type polymerase was their $K_{m(app)}$ for dCTP; the K_m of D276E was nearly 100-fold lower than in the wild-type (Table 3). In contrast, D276A/I277L exhibited 40-fold less catalytic efficiency than the wild-type. The increase in catalytic efficiency exhibited by D276E raises the question of why this substitution has not been favored by natural selection. Analysis of the misincorporation efficiency of the three polymerases offered a possible explanation (Fig. 4). D276E showed a 25-fold increase in G:T misincorporation efficiency but no increase was evident for G:A or G:G mismatches. These results indicate that natural selection may favour fidelity over catalytic efficiency in maintaining D276. In contrast to D276E, mutant D276A/I277L showed increased misincorporation efficiency for both G:A and G:G, but not for G:T. This result is unexpected because D276A/I277L was less efficient in replicating a 4 nt template gap in the absence of dCTP (Fig. 4). It also displayed a lower mutation frequency

than both D276E and wild-type (Table 2). One possibility is that D276A/I277L is unable to extend mispairs once they form.

Fidelity mutants of pol β have been identified previously, and invariably they have exhibited lower catalytic efficiency than wild-type. Y265C, a mutant in the M-helix, was identified using a genetic screen which was based on the ability of mutants to generate Trp^+ revertants (25). This mutant also exhibited 2.5-fold lower K_m values than wild-type, but its k_{cat} was nearly 50-fold lower. Mutants at residues N279 and R283, generated by site-directed mutagenesis, had also altered fidelity (22,40). The substitutions R283A, R283L and R283K led to a dramatic decrease in fidelity and catalytic efficiency. The mutant R283A exhibited a 160-fold decrease in fidelity in a reversion assay and a 5000-fold decrease in catalytic efficiency. The substitution N279A led to a 19-fold decrease in catalytic efficiency and exhibited an increase in fidelity based on a reversion assay. All mutants exhibited increased K_m values. These data clearly illustrate the importance of helices M and N in fidelity, but also the fact that the low fidelity mutant identified here, D276E, is the only pol β mutant to date which exhibits an increase in its catalytic efficiency compared to wild-type.

Our results are consistent with a role for D276 within the induced fit mechanism suggested for pol β (18). According to this hypothesis, proper alignment of the substrates in the active site is only achieved when the correct dNTP is present. The correct fit is then the limiting step for the nucleotidyl transfer. The electron density map derived from the structural studies is of insufficient resolution to determine the orientation of D276 (16,18,41). However, the data from this study support the model that during the 'closing' of the thumb subdomain, D276 is oriented towards, and makes contact with, the residue R40 of the 8 kDa domain, thus helping to determine the size of the active site (Fig. 5). A residue such as E, which essentially replicates the molecular bonding propensities of D but is 1.5 Å longer, may increase the size of the active site enough to stabilize G:T mispairs but not the bulkier G:A and G:G mispairs. A consequence of this model is that random mutagenesis of the 8 kDa region near R40 may yield catalytic efficiency and fidelity mutants.

In conclusion, our identification of a mutant polymerase exhibiting a 25-fold increase in catalytic activity demonstrates that the pol β design has the potential for significantly higher catalytic efficiency. The importance of higher efficiency is illustrated by the ability of D276E to support faster bacterial growth than the wild-type form. Yet this substitution has not been chosen by natural selection in mammalian cells, probably because for pol β, a DNA repair enzyme, higher fidelity is more important than catalytic efficiency. The inability of this and other studies to identify DNA polymerase mutants that exhibit higher catalytic efficiency or higher replication fidelity without detrimental effect on other properties supports the hypothesis that the wild-type residues that evolved by natural selection effect an optimal equilibrium among the polymerase's different enzymatic activities. Improvement in the performance of one property may lead to deterioration in another.

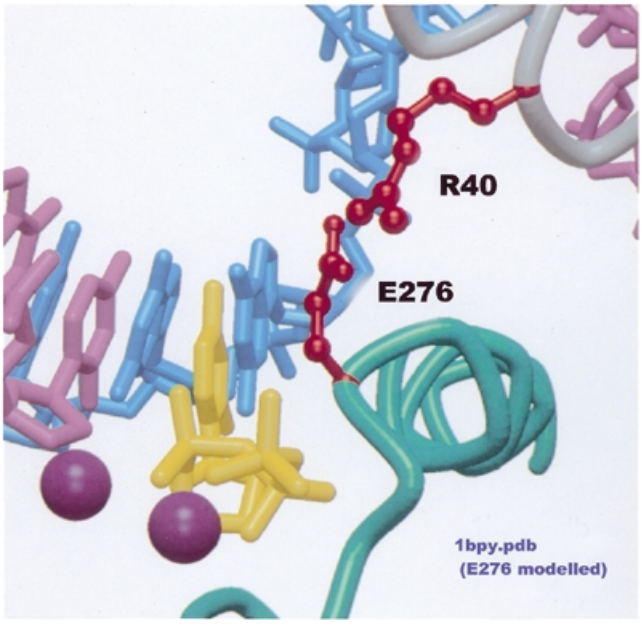

Figure 5. Graphic modeling of the interaction between E276 and R40 of the 8 kDa domain in the closed polymerase configuration. The molecule colored yellow is the incoming dNTP, the molecule colored blue is the DNA template, and the molecule colored pink is the primer DNA. The purple spheres represent the Mg^{2+} ions. This figure was made using MOLESCRIPT and Raster 3D.

ACKNOWLEDGEMENTS

We are grateful to Michael Fry, Premal Patel, Lance Encel, Ellie Adman and Miriam Richards for helpful discussions and editorial comments. This study was supported by an O.I.G. grant, R35-CA-39903 from the National Institutes of Health, and by an RO1 grant, AG-01751.

REFERENCES

1. Beese,L.S., Friedman,J.M. and Steitz,T.A. (1993) Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. *Biochemistry*, **32**, 14095–14101.
2. Joyce,C.M. and Steitz,T.A. (1994) Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.*, **63**, 777–822.
3. Steitz,T.A. (1999) DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.*, **274**, 17395–17398.
4. Beard,W.A. and Wilson,S.H. (2000) Structural design of a eukaryotic DNA repair polymerase: DNA polymerase β. *Mutat. Res.*, **460**, 231–244.
5. Ochman,H., Lawrence,J.G. and Groisman,E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.
6. Miller,J.H. (1998) Mutators in *Escherichia coli*. *Mutat. Res.*, **409**, 99–106.
7. Patel,P.H. and Loeb,L.A. (2000) DNA polymerase active site is highly mutable: evolutionary consequences. *Proc. Natl Acad. Sci. USA*, **97**, 5095–5100.
8. Grossman,N., Rosner,E. and Ron,E.Z. (1989) Termination of DNA replication is required for cell division in *Escherichia coli*. *J. Bacteriol.*, **171**, 74–79.
9. Johnson,R.E., Washington,M.T., Haracska,L., Prakash,S. and Prakash,L. (2000) Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature*, **406**, 1015–1019.
10. Suzuki,M., Baskin,D., Hood,L. and Loeb,L.A. (1996) Random mutagenesis of *Thermus aquaticus* DNA polymerase I: concordance of immutable sites *in vivo* with the crystal structure. *Proc. Natl Acad. Sci. USA*, **93**, 9670–9675.
11. Landis,D.M., Gerlach,J.L., Adman,E.T. and Loeb,L.A. (1999) Tolerance of 5-fluorodeoxyuridine resistant human thymidylate synthases to alterations in active site residues. *Nucleic Acids Res.*, **27**, 3702–3711.
12. Singhal,R.K., Prasad,R. and Wilson,S.H. (1995) DNA polymerase β conducts the gap-filling step in uracil-initiated base excision repair in a bovine testis nuclear extract. *J. Biol. Chem.*, **270**, 949–957.

13. Sobol,R.W., Prasad,R., Evenski,A., Baker,A., Yang,X.P., Horton,J.K. and Wilson,S.H. (2000) The lyase activity of the DNA repair protein β-polymerase protects from DNA-damage-induced cytotoxicity. *Nature*, **405**, 807–810.
14. Davies,J.F.,II, Almassy,R.J., Hostomska,Z., Ferre,R.A. and Hostomsky,Z. (1994) 2.3 A crystal structure of the catalytic domain of DNA polymerase β. *Cell*, **76**, 1123–1133.
15. Pelletier,H., Sawaya,M.R., Kumar,A., Wilson,S.H. and Kraut,J. (1994) Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science*, **264**, 1891–1903.
16. Pelletier,H., Sawaya,M.R., Wolfle,W., Wilson,S.H. and Kraut,J. (1996) Crystal structures of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry*, **35**, 12742–12761.
17. Sawaya,M.R., Pelletier,H., Kumar,A., Wilson,S.H. and Kraut,J. (1994) Crystal structure of rat DNA polymerase β: evidence for a common polymerase mechanism. *Science*, **264**, 1930–1935.
18. Sawaya,M.R., Prasad,R., Wilson,S.H., Kraut,J. and Pelletier,H. (1997) Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry*, **36**, 11205–11215.
19. Joyce,C.M. and Steitz,T.A. (1995) Polymerase structures and function: variations on a theme? *J. Bacteriol.*, **177**, 6321–6329.
20. Huang,H., Chopra,R., Verdine,G.L. and Harrison,S.C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, **282**, 1669–1675.
21. Osheroff,W.P., Beard,W.A., Yin,S., Wilson,S.H. and Kunkel,T.A. (2000) Minor groove interactions at the DNA polymerase β active site modulate single-base deletion error rates. *J. Biol. Chem.*, **275**, 28033–28038.
22. Beard,W.A., Osheroff,W.P., Prasad,R., Sawaya,M.R., Jaju,M., Wood,T.G., Kraut,J., Kunkel,T.A. and Wilson,S.H. (1996) Enzyme–DNA interactions required for efficient nucleotide incorporation and discrimination in human DNA polymerase β. *J. Biol. Chem.*, **271**, 12141–12144.
23. Kraynov,V.S., Werneburg,B.G., Zhong,X., Lee,H., Ahn,J. and Tsai,M.D. (1997) DNA polymerase β: analysis of the contributions of tyrosine-271 and asparagine-279 to substrate specificity and fidelity of DNA replication by pre-steady-state kinetics [published erratum appears in *Biochem. J.* (1997) **325**, 815]. *Biochem. J.*, **323**, 103–111.
24. Clairmont,C.A. and Sweasy,J.B. (1998) The Pol β-14 dominant negative rat DNA polymerase β mutator mutant commits errors during the gap-filling step of base excision repair in *Saccharomyces cerevisiae*. *J. Bacteriol.*, **180**, 2292–2297.
25. Washington,S.L., Yoon,M.S., Chagovetz,A.M., Li,S.X., Clairmont,C.A., Preston,B.D., Eckert,K.A. and Sweasy,J.B. (1997) A genetic system to identify DNA polymerase β mutator mutants. *Proc. Natl Acad. Sci. USA*, **94**, 1321–1326.
26. Sweasy,J.B. and Loeb,L.A. (1992) Mammalian DNA polymerase β can substitute for DNA polymerase I during DNA replication in *Escherichia coli*. *J. Biol. Chem.*, **267**, 1407–1410.
27. Bebenek,K. and Kunkel,T.A. (1995) Analyzing fidelity of DNA polymerases. *Methods Enzymol.*, **262**, 217–232.
28. Sweasy,J.B. and Loeb,L.A. (1993) Detection and characterization of mammalian DNA polymerase β mutants by functional complementation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **90**, 4626–4630.
29. Blank,A., Kim,B. and Loeb,L.A. (1994) DNA polymerase δ is required for base excision repair of DNA methylation damage in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **91**, 9047–9051.
30. Suzuki,M., Avicola,A.K., Hood,L. and Loeb,L.A. (1997) Low fidelity mutants in the O-helix of *Thermus aquaticus* DNA polymerase I. *J. Biol. Chem.*, **272**, 11228–11235.
31. Boosalis,M.S., Petruska,J. and Goodman,M.F. (1987) DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. *J. Biol. Chem.*, **262**, 14689–14696.
32. Chagovetz,A.M., Sweasy,J.B. and Preston,B.D. (1997) Increased activity and fidelity of DNA polymerase β on single-nucleotide gapped DNA. *J. Biol. Chem.*, **272**, 27501–27504.
33. Wilson,S.H. (1998) Mammalian base excision repair and DNA polymerase β. *Mutat. Res.*, **407**, 203–215.
34. Klungland,A. and Lindahl,T. (1997) Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.*, **16**, 3341–3348.
35. Dianov,G.L., Prasad,R., Wilson,S.H. and Bohr,V.A. (1999) Role of DNA polymerase β in the excision step of long patch mammalian base excision repair. *J. Biol. Chem.*, **274**, 13741–13743.
36. Plug,A.W., Clairmont,C.A., Sapi,E., Ashley,T. and Sweasy,J.B. (1997) Evidence for a role for DNA polymerase β in mammalian meiosis. *Proc. Natl Acad. Sci. USA*, **94**, 1327–1331.
37. Oda,N., Saxena,J.K., Jenkins,T.M., Prasad,R., Wilson,S.H. and Ackerman,E.J. (1996) DNA polymerases α and β are required for DNA repair in an efficient nuclear extract from *Xenopus* oocytes. *J. Biol. Chem.*, **271**, 13816–13820.
38. Prasad,R., Beard,W.A. and Wilson,S.H. (1994) Studies of gapped DNA substrate binding by mammalian DNA polymerase β. Dependence on 5′-phosphate group. *J. Biol. Chem.*, **269**, 18096–18101.
39. Osheroff,W.P., Jung,H.K., Beard,W.A., Wilson,S.H. and Kunkel,T.A. (1999) The fidelity of DNA polymerase β during distributive and processive DNA synthesis. *J. Biol. Chem.*, **274**, 3642–3650.
40. Ahn,J., Werneburg,B.G. and Tsai,M.D. (1997) DNA polymerase β: structure-fidelity relationship from pre-steady-state kinetic analyses of all possible correct and incorrect base pairs for wild type and R283A mutant. *Biochemistry*, **36**, 1100–1107.
41. Pelletier,H., Sawaya,M.R., Wolfle,W., Wilson,S.H. and Kraut,J. (1996) A structural basis for metal ion mutagenicity and nucleotide selectivity in human DNA polymerase β. *Biochemistry*, **35**, 12762–12777.