THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 5, pp. 4011–4020, February 4, 2011

Human DNA Polymerase β Mutations Allowing Efficient Abasic Site Bypass*[S]

Received for publication, August 19, 2010, and in revised form, November 5, 2010 Published, JBC Papers in Press, November 24, 2010, DOI 10.1074/jbc.M110.176826

Sonja Gieseking, Konrad Bergen, Francesca Di Pasquale, Kay Diederichs, Wolfram Welte, and Andreas Marx[1]

From the Departments of Chemistry and Biology, Konstanz Research School Chemical Biology, University of Konstanz, 78464 Konstanz, Germany

The DNA of every cell in the human body gets damaged more than 50,000 times a day. The most frequent damages are abasic sites. This kind of damage blocks proceeding DNA synthesis by several DNA polymerases that are involved in DNA replication and repair. The mechanistic basis for the incapability of these DNA polymerases to bypass abasic sites is not clarified. To gain insights into the mechanistic basis, we intended to identify amino acid residues that govern for the pausing of DNA polymerase β when incorporating a nucleotide opposite to abasic sites. Human DNA polymerase β was chosen because it is a well characterized DNA polymerase and serves as model enzyme for studies of DNA polymerase mechanisms. Moreover, it acts as the main gap-filling enzyme in base excision repair, and human tumor studies suggest a link between DNA polymerase β and cancer. In this study we employed high throughput screening of a library of more than 11,000 human DNA polymerase β variants. We identified two mutants that have increased ability to incorporate a nucleotide opposite to an abasic site. We found that the substitutions E232K and T233I promote incorporation opposite the lesion. In addition to this feature, the variants have an increased activity and a lower fidelity when processing nondamaged DNA. The mutations described in this work are located in well characterized regions but have not been reported before. A crystallographic structure of one of the mutants was obtained, providing structural insights.

The DNA of every cell in the human body gets damaged more than 50,000 times a day (1). The relation between DNA damage and repair has a significant effect on various cancers, neurological aberrations, and the process of premature aging (2). DNA polymerases are key enzymes that function in maintaining the integrity of the encoded genetic information in DNA replication, DNA repair, DNA recombination, and the bypassing of damages in DNA. Therefore they are central to the aforementioned interplay (3).

The most frequent DNA damage observed under physiological conditions are abasic sites resulting from spontaneous hydrolysis of the bond that connects the sugar and the nucleobase in DNA (4). Guanine and adenine nucleobase residues are cleaved most efficiently, resulting in the abasic sugar moiety (AP; see Fig. 1*A*) with the loss of the genetic information stored in the nucleobase (5). Because these lesions are devoid of genetic information, they are potentially mutagenic. The bulk of this damage is removed by base excision repair pathway, which uses the sister strand to guide incorporation of the right nucleotide in place of the lesion (6, 7). DNA polymerase β acts as the main gap-filling enzyme in base excision repair (8). The enzyme governs for selecting the right nucleotide complementary to the undamaged templating nucleotide (9).

DNA polymerases have been grouped in seven different families named A, B, C, D, X, and Y and reverse transcriptases. The grouping depends on sequence homology and structural similarity (10). DNA polymerase β belongs to the X family. The single polypeptide is the smallest eukaryotic DNA polymerase (39 kDa), containing 335 amino acid residues (11). It is folded into discrete domains and subdomains (12). The enzyme consists of an N-terminal lyase domain (8 kDa), which is connected to the DNA polymerase domain (31 kDa) by a short protease-sensitive segment (13).

The important role of DNA polymerase β in the maintenance of genome integrity is emphasized by the finding that ~30% of all tumors studied to date harbor mutations in the DNA polymerase β gene (14). The tumor-associated mutations are found in the lyase as well as in the DNA polymerase domain. The variants show altered enzyme properties such as fidelity and lyase function (14–20).

On the other hand DNA polymerase β serves as a very useful model for studying basic mechanisms of DNA polymerase function such as activity and fidelity. The enzyme has been studied extensively in a functional and structural context in the past (13). Several crystal structures are available, including those of the enzyme in ternary complex with various substrates (9, 21–24).

Lesions like abasic sites that remain undetected by the repair systems or are formed during the S phase remain a great challenge to the replication machinery. Abasic sites block proceeding DNA synthesis by several DNA polymerases including those that are involved in DNA replication and base excision repair such as DNA polymerase β (25). Stalling of DNA replication might be essential to allow repair processes to occur that ensure faithful DNA repair without introducing mutations. However, approximately a decade ago, novel specialized DNA polymerases were identified to be capable of

* This work was supported by the Deutsche Forschungsgemeinschaft within SPP1170.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1.

The atomic coordinates and structure factors (code 3OGU) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed: Dept. of Chemistry, Universität Konstanz, Universitätsstrasse 10, D-78464 Konstanz, Germany. Fax: (+49) 7531–88-5140; E-mail: andreas.marx@uni-konstanz.de.

bypassing DNA lesion (26). These enzymes are found in a variety of organisms and, for example, are involved in the suppression of skin cancer in humans. The translesion synthesis pathways are believed to occur after chromosomal replication thus, outside the S phase (27, 28). The mechanistic basis of the proficiency of some DNA polymerases to bypass certain lesions and the failure of other DNA polymerase are still under investigation.

To identify amino acid residues that govern for the pausing of DNA polymerase β when incorporating a nucleotide opposite to an abasic site, we generated and screened a library of more than 11,000 human DNA polymerase β variants randomly mutated using error-prone PCR. By using a high throughput screening system, we selected mutants that have an increased propensity to bypass abasic sites in comparison with the wild-type enzyme. In addition to this feature, the variants have an increased activity and a somewhat lower fidelity when processing nondamaged DNA. The mutations described in this work are located in well characterized regions but have not been reported before. A crystallographic structure of one of the mutants was obtained providing structural insights.

EXPERIMENTAL PROCEDURES

Library Construction—The human DNA polymerase β gene (NM_002690) was purchased codon optimized for expression in *Escherichia coli* from Geneart AG (Regensburg, Germany). It is noteworthy that codon optimization resulted in changes within the gene sequence without affecting the amino acid sequence (see supplemental Fig. S1). The gene was cloned into the plasmid pGDR11 using the restriction enzymes BamHI and SalI. The insert was sequenced to confirm that it does not contain mutations. Random mutations were inserted by error-prone PCR. PCRs were performed using 5 units of *Taq* DNA polymerase (Fermentas) in the reaction buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 7 mM $MgCl_2$, 0.05 mM $MnCl_2$, 1 mM dCTP/dTTP, 0.2 mM dATP/dGTP), 350 pg/μl plasmid pGDR11(hDNApolβ), and 200 nM of each primer (forward,5′-CAC CAT CAC CAT CAC CAT ACG GAT CCG ATG-3′; reverse, 5′-GCT AAT TAA GCT TGG CTG CAG GTC GAC TTA-3′) in a 50-μl total volume. 20 cycles of PCR (95 °C for 1 min, 63 °C for 1 min, and 72 °C for 2.5 min) were conducted. The PCR product was isolated by electrophoresis in a 0.8% agarose gel and purified (MinElute™ gel extraction kit; Qiagen). After digestion with BamHI and SalI, the restriction fragment was ligated with the digested and dephosphorylated (Antarctic phosphatase; NEB) plasmid pGDR11 by using T4 DNA ligase (Fermentas) for 16 h at 16 °C. The resulting plasmids were transformed into *E. coli* XL10 Gold. The clones were picked from LB agar plates and separately grown overnight in 384-well plates containing 150 μl of LB medium with 100 μg/ml carbenicillin. hDNApolβ mutants were parallel expressed in 1-ml cultures in 96-well plates by inducing during exponential growing phase (A_{600} = 0,6) with 1 mM isopropyl β-D-thiogalactopyranoside for 4 h. The cells were pelleted by centrifugation (40 min, 4000 rpm, 4 °C) and lysated in 300 μl of lysis buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.1 mg/ml lysozyme). After incubating on ice for 45 min, 900 μl of storage buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM DTT, 0.1 mM EDTA) was added followed by another 45 min on ice and centrifugation (40 min, 4000 rpm, 4 °C). The lysate was directly used for screening.

Screening—Reaction mixtures for library screening contained 2 mM $MnCl_2$, 200 μM of each dNTP, 100 μg/ml BSA, 150 nM primer (P20thioAG, 5′-CGT TGG TCC TGA AGG AGG *A*T-3′, where the asterisks indicate the positions of phosphorothiolate) and 150 nM template (T90A, 5′-CCG TCA GCT GTG CCG TCG CGC AGC ACG CGC CGC CGT GGA CAG AGG ACT GCA GAA AAT CAA CCT ATC CTC CTT CAG GAC CAA CGT ACA GAG-3′; or T90F, 5′-CCG TCA GCT GTG CCG TCG CGC AGC ACG CGC CGC CGT GGA CAG AGG ACT GCA GAA AAT CAF CCT ATC CTC CTT CAG GAC CAA CGT ACA GAG-3′; F indicates an abasic site analog, respectively) in buffer (50 mM Tris-HCl, pH 7.8, 70 mM KCl, 1 mM DTT, 100 μg/ml BSA, 5% (v/v) glycerol) in 20 μl of total volume. The reaction mixtures were dispensed in 384 well plates using an automated liquid handling device (Hamilton Microlab Star) followed by the addition of 5 μl of lysate solution. As a positive control, the DNA polymerase β wild-type (WT) was used; as a negative control, the catalytic inactive DNA polymerase β D256A (29) in analogously treated cultures of *E. coli* XL10 Gold cells was used. After incubation at 37 °C for 30 min, the reaction was stopped by adding 45 μl/well SYBR® Green I-containing solution (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 4× SYBR® Green I). Fluorescence intensities were quantified using a plate reader (Polarstar Optima, BMG Labtechnologies GmbH) with excitation at 485 nm and emission at 520 nm.

Purification of DNA Polymerase β—The WT and mutants of the DNA polymerase β were expressed and purified as described (30). The purity of the proteins was >95% as controlled by SDS-PAGE. The concentrations were determined by the Bradford assay.

Primer Extension Studies—Reactions contained primer (P20, 5′-CGT TGG TCC TGA AGG AGG AT-3′; or P23, 5′-CGT TGG TCC TGA AGG AGG ATA GG-3′), template (T33G, 5′-AAA TGC GCC TAT CCT CCT TCA GGA CCA ACG TAC-3′; or T33F, 5′-AAA TGC FCC TAT CCT CCT TCA GGA CCA ACG TAC-3′; F indicates an abasic site analog) and purified enzyme in storage buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 50% (v/v) glycerol), 200 μg/ml BSA, 200 μM dNTPs (N) or single nucleotide dATP (A), dCTP (C), dGTP (G), or dTTP (T), and 1× reaction buffer (50 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM KCl, 10 mM $MgCl_2$, 2 mM DTT, 1% (v/v) glycerol) in a final volume of 20 μl. The concentrations of the respective enzyme and the primer template complexes are given in the figure legends. The reaction mixtures were incubated for the times indicated in the corresponding figures legends at 37 °C and subsequently quenched using PAGE loading solution (80% (v/v) formamide, 20 mM EDTA, 0.025% (w/v) bromphenol blue, 0.025% (w/v) xylene cyanol). The samples were denatured at 95 °C for 5 min analyzed by 12% PAGE containing 8 M urea. Visualization was performed using phosphorimaging.

Kinetics—The rate of single turnover, single-nucleotide incorporation was determined using the different primer/33-

mer template DNA duplexes using rapid quench technology by using a KinTek RQF-3 rapid quench flow apparatus (KinTek Corp.). Reaction mixtures contained 200 nM primer (P23 or P24); 200 nM template (T33G or T33F), 2 μM purified enzyme in storage buffer, 400 μg/ml BSA, 1× reaction buffer, and different concentrations of dNTP as indicated in the figure legends. The reaction was initiated by mixing equal volumes (15 μl each) of the enzyme reaction mixture with a solution containing dNTP in 1× reaction buffer. For reaction times ranging from 0.05 to 10 s, the rapid quench instrument was used. The reactions were quenched using 0.3 M EDTA prior to mixing with a PAGE loading solution (80% (v/v) formamide, 0.025% (w/v) bromphenol blue, 0.025% (w/v) xylene cyanol). For reaction times longer than 10 s, a manual quench was performed. Samples of the quenched reactions were denatured at 95 °C for 5 min and analyzed by 12% PAGE containing 8 M urea. Quantification was performed using phosphorimaging. The product bands were quantified using the Bio-Rad Quantity One software. For pre-steady state analysis, experimental data were fit by nonlinear regression using the program GraphPad Prism4. The data were fit to a one phase exponential equation [primer + 1] = $A^*(1 - \exp(-k_{obs}{}^*t))$. The observed catalytic rates (k_{obs}) were plotted against the used dNTP concentration. These data were fitted to a hyperbolic equation to determine the K_d of the incoming nucleotide. The incorporation efficiency is given by k_{pol}/K_d. The kinetic data result from multiple independently conducted experiments.

DNA Binding Assay—The DNA binding affinity was measured using a gel mobility shift assay (31, 32). P23 and T33G or T33F were annealed as described above. Up to 12 protein concentrations ranging from 0.125 nM up to 5 μM, expected to bracket the dissociation constant (K_d), were incubated with 0.1 nM primer template complex in buffer containing 10 mM Tris, pH 7.5, 6 mM $MgCl_2$, 100 mM NaCl, 10% glycerol, and 0.1% Nonidet P-40 (30). After incubation for 10 min at 20 °C, the samples were loaded onto a nondenaturing 8% polyacrylamide gel (in 7.5 mM Tris borate, 2 mM $MgCl_2$, 0.1 mM EDTA) (33), which run at 300 V for 1 h. After loading, the voltage was reduced to 150 V, and the gel was run for 1.5 h. Quantification was performed using phosphorimaging. The product bands were quantified using the Bio-Rad Quantity One software. The dissociation constant K_d was derived by fitting the experimental data using nonlinear regression with the program GraphPad Prism4 using the equation: $Y = [(m_1{}^*X)]/(K_d + X) + c$, where m_1 is a scaling factor, and c is the apparent minimum Y value as described (30).

DNA Polymerase Activity Determination—For activity determination, primer extension reactions were carried out with varying concentrations of DNA polymerase β WT and mutants. The reaction contained 150 nM primer (P23), 150 nM template (T33GCG), 200 μM dNTP, 200 μg/ml BSA, and 1× reaction buffer (50 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM KCl, 10 mM $MgCl_2$, 2 mM DTT, 1% (v/v) glycerol). The reactions were performed at 37 °C for 10 min, subsequently quenched using PAGE loading solution. The products were separated on a 12% denaturing PAGE. Phosphorimaging was carried out, and the obtained intensities for each band were transformed into dNTP conversion. Activity was measured by quantifying the intensity of each band produced by the respective DNA polymerase with a PhosphorImager (Bio-Rad). From this quantification, the amount of incorporated nucleotides was calculated. The total amount of incorporated nucleotides for each reaction equals the sum of incorporated nucleotides of each band. The presented results are from measurements that were independently repeated at least three times.

Crystallization—The enzyme mutant was expressed in *E. coli* BL21 AI carrying the CDS in a pET21b vector. The recombinant enzyme was purified as described earlier (34) with slight modification. Protein concentration was determined by absorbance at 280 nm (ϵ = 23,380 M^{-1} cm^{-1}) and Bradford assay. Enzyme solutions were concentrated up to 20 mg/ml (VivaSpin) and stored at −80 °C. The crystals of the binary complex were obtained as described by Pelletier *et al.* (24) against a reservoir solution containing 100 mM MES[2] (pH 6.3), 10% PEG 3350, and 100 mM NaCl in hanging drop vapor diffusion plates (Nextal). The data were collected at PXI (PSI, Villigen, Switzerland) at the Pilatus 6 M detector. Data integration and reduction were performed using XDS (35, 36). Crystals had similar lattice constants as Protein Data Bank code 9ICW (24), and structure determination was therefore initiated by rigid body refinement of 9ICW, using the PHENIX suite (37). As a model for refinement, Protein Data Bank code 9ICW (24) for the binary complex was chosen. Subsequent refinement and model building was performed using the PHENIX suite and Coot (38). Molecular graphics were prepared in PyMol (39).

RESULTS

Screening—To evolve human DNA polymerase β variants with increased abilities to bypass an abasic site, we first randomized its gene by error-prone PCR. A library of 11,520 entities was generated by picking colonies. The proteins were expressed and lysated in multiwell plates. For screening of activity, we first employed a screen based on primer extension using an undamaged template in 384-well plates (40, 41). The enzymes were screened directly after lysis without further purification steps. As controls we used colonies expressing the wild-type enzyme and the catalytic inactive DNA polymerase β mutant D256A (29). DNA polymerase activity was monitored through termination by the addition of EDTA and quantification of synthesized double-stranded DNA through staining with SYBR® Green I. Active mutants (39% of all screened entities) were pooled and subsequently tested for the ability to bypass an abasic site using a similar screen as described above with the difference that a template containing the abasic site analog F (Fig. 1*A*) was employed. We identified two DNA polymerase β mutants that exhibited promising properties. We refer to these enzymes as 1L19 and 5P20, referring to their location in the library. Sequencing revealed that each enzyme has four mutations (for 1L19, S44R, F99L, E123K, and T233I; and for 5P20,

[2] The abbreviation used is: MES, 4-morpholineethanesulfonic acid.

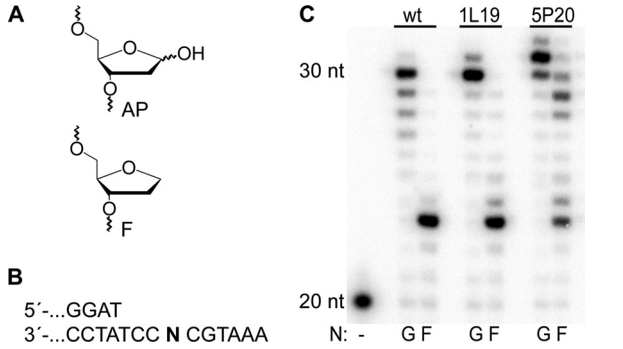


FIGURE 1. **Studies comparing WT human DNA polymerase β and mutants 1L19 and 5P20.** *A*, structures of an abasic site (*AP*) and the abasic site analog F. *B*, partial primer template sequence used in primer extension experiment. *C*, primer extension assay under running start conditions. Side by side comparisons of DNA synthesis using a canonical template (*N* = dGMP, G) and abasic site containing template (*N* = F) were conducted. All of the experiments were conducted employing the same condition using 500 nM of the respective enzyme and 150 nM of primer template complex.

S2G, F99L, E232K, and V269M). These mutants were subsequently purified to homogeneity and were subjected to further characterization.

Primer Extension and Abasic Site Bypass by WT and Mutants—To compare the ability to bypass F, the purified WT and mutated enzymes were subsequently used in different primer extension studies under identical reaction conditions. In a running start experiment employing all four dNTPs, the DNA polymerases extended the radioactively labeled primer strand by three nucleotides before approaching the abasic site in the template strand (Fig. 1*B*). The WT enzyme only sluggishly incorporates a nucleotide opposite the abasic site as shown by a faint band at the respective position after PAGE analysis (Fig. 1*C*). The synthesis of longer products resulting from bypass of the abasic site could not be detected. Nevertheless, when an undamaged template was used, full-length extension was observed. In the undamaged template, a G was chosen instead of the abasic site because it has been shown that G is the predominant source of abasic sites under physiological conditions (5).

Next, the mutants were investigated. Both mutant enzymes show pausing opposite the abasic site as well; however, they were able to synthesize longer products resulting from bypass of the abasic site contrasting the finding for the WT enzyme.

Next we investigated which nucleotide is preferentially incorporated opposite the abasic site. Thus, the primer template complex was designed in a way that the 3′-end of the primer strand terminates opposite the template (Fig. 2*A*). Two different sets of experiments were conducted. In one series, the enzyme concentration was equal to the concentration of the primer template complex (Fig. 2*B*). In the second series, the enzyme concentration was increased to be 10-fold higher than the primer template complex to ensure that all of the primer template complexes are bound to the enzyme (Fig. 2*C*). In all single-nucleotide incorporation experiments, it became clear that the mutants incorporate a nucleotide opposite an abasic site more preferentially than the wild-type enzyme, regardless of the ration of the enzyme to the primer template complex (Fig. 2, *B–D*).

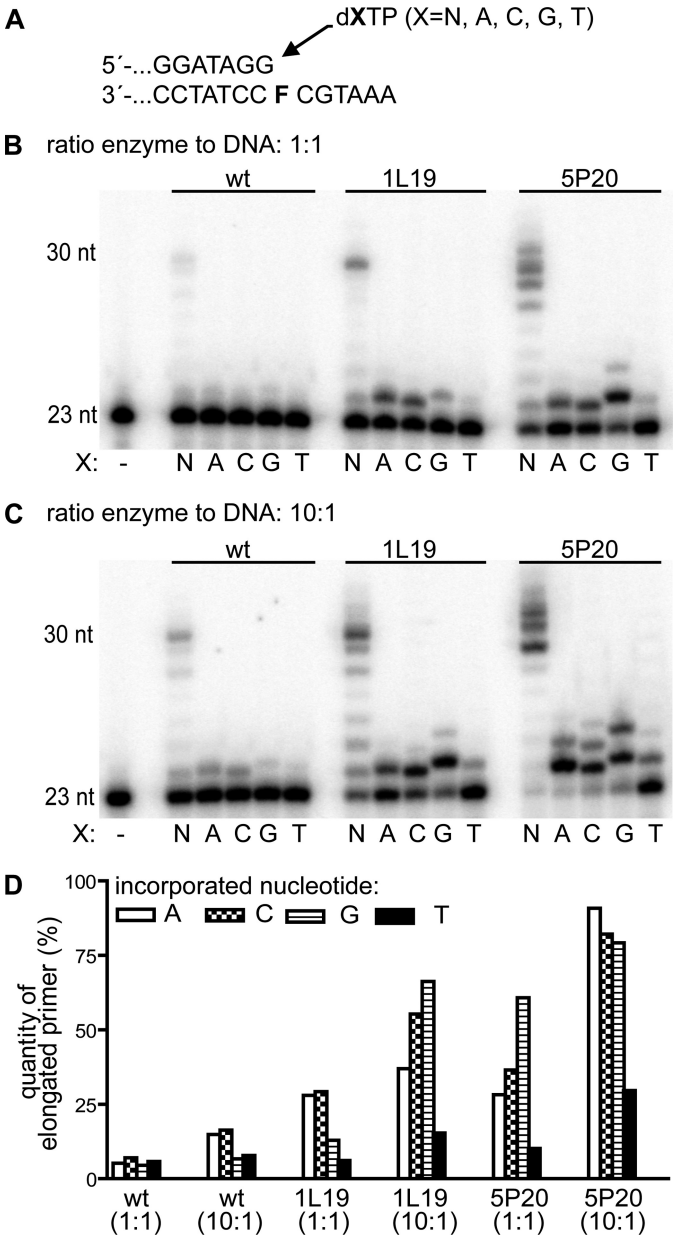


FIGURE 2. **Primer extension studies comparing WT human DNA polymerase β and mutants 1L19 and 5P20.** *A*, partial primer template sequence used in primer extension experiment. *B*, PAGE analysis of single nucleotide incorporation opposite the abasic site. The reactions have been executed either with all four dNTP (*N*) or one of the four single nucleotides (*A*, dATP; *C*, dCTP; *G*, dGTP; and *T*, dTTP) at 200 μM and equal concentrations of enzyme and primer template complex (200 nM). Primer template sequences are indicated in the figure. *C*, same as in *B* employing an enzyme concentration of 2 μM and 200 nM primer template complex. *D*, quantified primer extension depicted in *B* and *C* by the respective enzyme using the indicated dNTP.

Kinetics of WT and Mutant Enzymes—For quantification of the aforementioned results, we measured nucleotide incorporation using a quench flow device. Using this method, the incorporations opposite an undamaged G and the stable abasic site analog F was investigated. Measuring the incorporation rate k_{pol} and the affinity of the binary polymerase-primer template complex for the incoming nucleotide (K_d) led to the incorporation efficiency (k_{pol}/K_d) (Fig. 3 and Table 1).

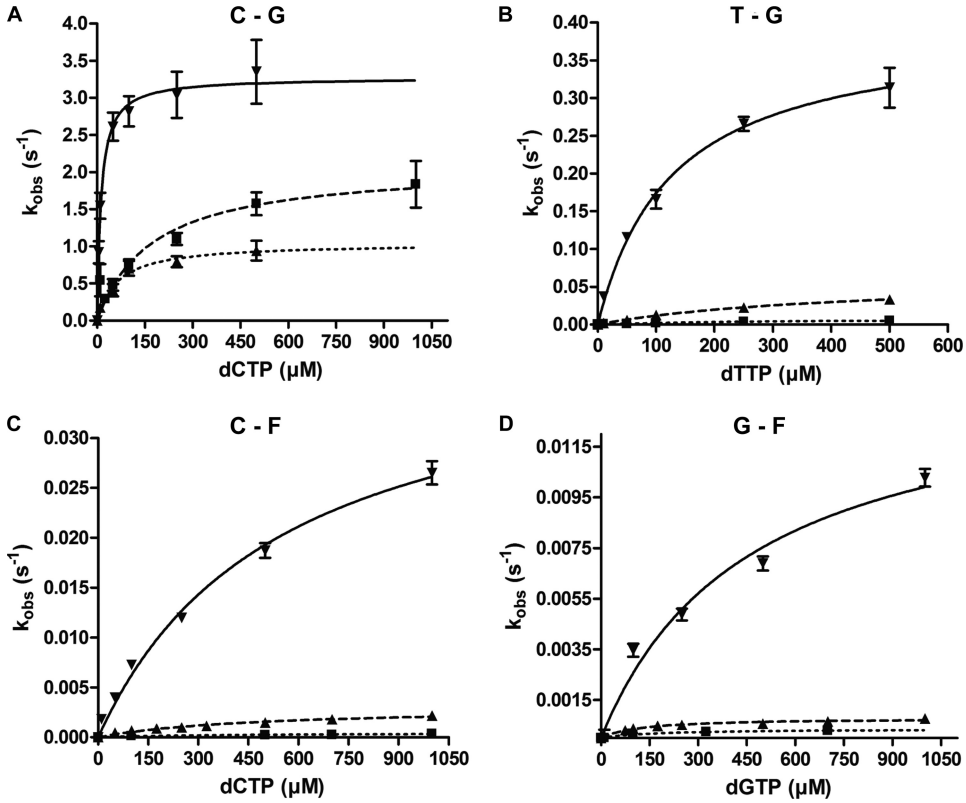


FIGURE 3. **Kinetics of nucleotide incorporation by WT (■) and the 1L19 (▲) and 5P20 (▼) mutant enzymes employing different primer template complexes, respectively.** The graphs show the dependence of the pre-steady state rates on the nucleotide concentration. The k_{obs} values were plotted *versus* the nucleotide concentration. *A*, incorporation of dCMP opposite template G. *B*, mismatch formation by incorporation of dTMP opposite template G. *C*, incorporation of dCMP opposite abasic site analog F in the template. *D*, incorporation of dCMP opposite F in the template.

TABLE 1
Kinetic data of nucleotide incorporation

	Template:dNTP	k_{pol}	K_d	k_{pol}/K_d'
		s^{-1}	μM	$_M{}^{-1}\,s^{-1}$
WT	G:dCTP	2.09 ± 0.30	173 ± 72	12,055
	G:dTTP	0.01[a]	334 ± 66	27
1L19	G:dCTP	1.04 ± 0.07	62.5 ± 13.6	16,677
	G:dTTP	0.06[a]	417 ± 48	146
5P20	G:dCTP	3.27 ± 0.07	12.2 ± 1.6	267,621
	G:dTTP	0.39 ± 0.01	126 ± 11	3,122
WT	F:dCTP	$0.42 \times 10^{-3} \pm 0.09 \times 10^{-3}$	302 ± 192	1
	F:dGTP	$0.36 \times 10^{-3} \pm 0.02 \times 10^{-3}$	161 ± 29	2
1L19	F:dCTP	$3.04 \times 10^{-3} \pm 0.03 \times 10^{-3}$	468 ± 102	6
	F:dGTP	$0.79 \times 10^{-3} \pm 0.04 \times 10^{-3}$	124 ± 23	6
5P20	F:dCTP	$40.4 \times 10^{-3} \pm 3.29 \times 10^{-3}$	546 ± 91	74
	F:dGTP	$14.5 \times 10^{-3} \pm 2.31 \times 10^{-3}$	461 ± 163	31

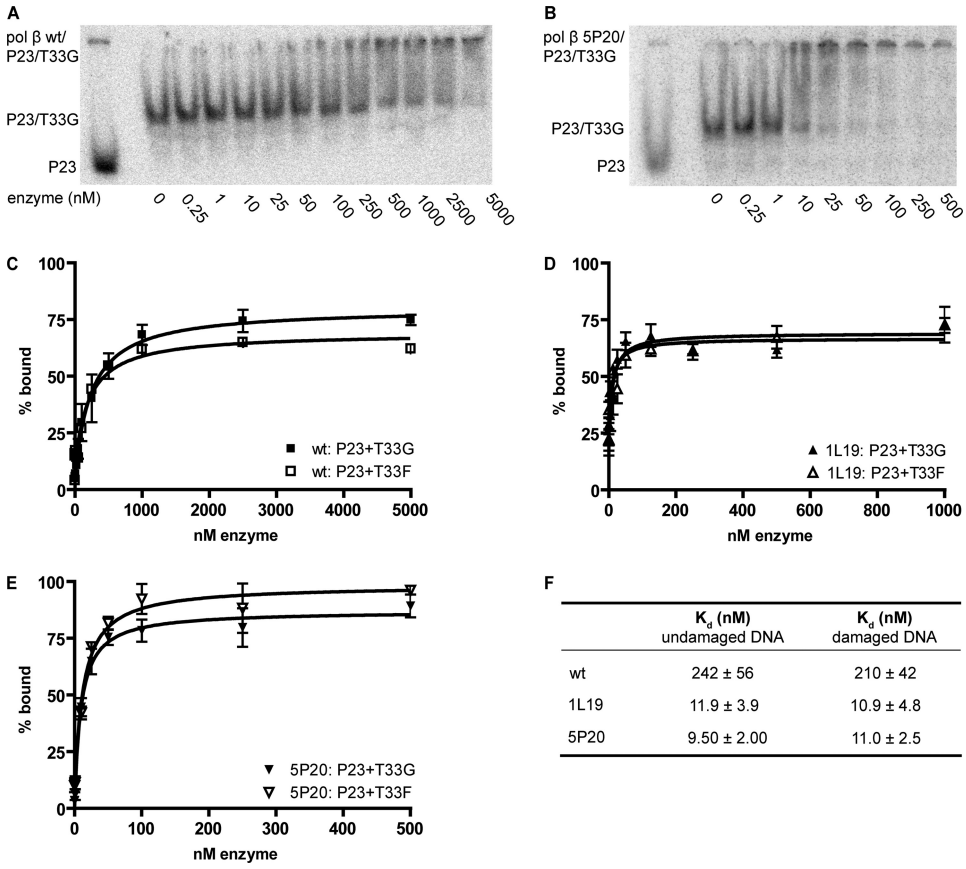
[a] The derivations in multiple experiments were below significance.

Kinetic experiments were performed by examining single nucleotide extension of a 23-mer primer annealed to a 33-mer template using an excess of enzyme over DNA substrate. In the match case, when dCMP was incorporated opposite to G, the incorporation rates were 2.09 ± 0.30 s^{-1} for WT, 1.04 ± 0.07 s^{-1} for 1L19, and 3.27 ± 0.07 s^{-1} for 5P20, respectively (Fig. 3*A* and Table 1). These results are in the same range as recently published data (42). The K_d values of the enzymes differ more significantly among the enzymes than the k_{pol} values. The affinity of the WT polymerase-primer template complex toward the incoming dCTP (K_d) was ~14-fold weaker in comparison with 5P20. This yielded a 22-fold higher incorporation efficiency of the mutant 5P20 in the match case. In contrast the incorporation efficiency of the WT and mutant 1L19 was in the same order.

Because dTMP was misincorporated opposite dG preferentially (data not shown), we investigated this combination in more detail. In this case, the incorporation rates were 0.01 s^{-1} for WT, 0.06 s^{-1} for 1L19, and 0.39 ± 0.01 s^{-1} for 5P20, respectively (Fig. 3*B* and Table 1). The K_d values showed only marginal differences (WT, 334 ± 66 μM; 1L19, 417 ± 48 μM; and 5P20, 126 ± 11 μM). The selectivity of a DNA polymerase can be defined by the ratio for the efficiencies for incorporation of a canonical nucleotide *versus* the misincorporation of a noncanonical nucleotide. Considering this, it appears that 1L19 has a 4-fold and 5P20 a 5-fold reduced selectivity compared with the WT enzyme.

The single nucleotide incorporation opposite F was analyzed next using dCTP and dGTP. In these reactions, the incorporation rates of the enzymes were significantly reduced compared with incorporation opposite the canonical template (Fig. 3, *C* and *D*, and Table 1). Nevertheless, distinctive differences of the incorporation rates were detectable, whereas the K_d values are in the same range. Taking together, 5P20 incorporated dCMP opposite F more than 50-fold and dGMP ~15-fold more efficiently than the WT enzyme. 1L19 exhibits a 3–6-fold increase in efficiency compared with WT depending on the nucleotide (Table 1).

Affinity of WT and Mutants for DNA—To determine whether the mutant enzymes had an altered affinity for the primer template complexes, we performed a gel mobility shift assay (Fig. 4) using approaches described earlier (30–33). The K_d values for undamaged and damaged DNA of WT enzyme



	K_d (nM) undamaged DNA	K_d (nM) damaged DNA
wt	242 ± 56	210 ± 42
1L19	11.9 ± 3.9	10.9 ± 4.8
5P20	9.50 ± 2.00	11.0 ± 2.5

FIGURE 4. **DNA binding affinity of the WT, 5P20, and 1L19 for DNA with undamaged (P23 + T33G) and damaged (P23 + T33F) primer template complex.** *A* and *B*, nondenaturating PAGE analysis of binding to undamaged primer template complex by the WT enzyme (*A*) and mutant 5P20 (*B*), respectively. The mixtures contained 0.1 nM primer template complex and varying enzyme concentrations (0.125 nM to 5 μM) and were incubated for 10 min at 20 °C. *C–E*, the graphs show the fractions of primer template complexes bound to the respective enzyme depending on the enzyme concentrations (nM) for WT (*C*), mutant 1L19 (*D*), and mutant 5P20 (*E*), respectively. *F*, determined K_d (nM) values for binding to the undamaged and damaged primer template complexes by the respective enzyme. *pol*, polymerase.

were similar (242 ± 56 and 210 ± 42 nM, respectively). The K_d values of the mutant enzymes were similar as well for undamaged and damaged template (1L19, 11.9 ± 3.9 and 10.9 ± 4.8 nM; 5P20, 9.5 ± 2.0 and 11.0 ± 2.5 nM). Thus, the mutants appear to interact with the primer template complex with the same affinity that is ~20-fold higher that the WT enzyme.

Specific Activity of WT and Mutants—Next we investigated the specific activity of the enzymes using an assay that detects nucleotide incorporation and extension. Thus, in primer extension the amount of dNTPs that is incorporated within 10 min depending on the enzyme concentration was investigated. We found that the WT enzyme exhibits a specific activity of 37.3 ± 2.3 dNTP/fmol enzyme, whereas the activity of the mutant 1L19 was 3-fold increased (87.2 ± 5.6) (Fig. 5*A*), and the mutant 5P20 (360 ± 9) was 10-fold increased (Fig. 5*B*).

Contribution of Mutations to Abasic Site Bypass Proficiency—To determine whether all mutations of 1L19 and 5P20 are required to render the enzymes proficient in bypassing an active site, every single mutation as well as combinations of double and triple mutations were generated by site-directed mutagenesis. All of the mutants were subsequently purified to homogeneity and were utilized in primer extension reactions employing identical reaction conditions. An unmodified (T33A) and a damaged (T33F) template containing an abasic site were used in combination with a 20-nucleotide primer (P20) (Fig. 6). We found that all of the variants of the 1L19 mutant that contain the mutation T233I showed the highest proficiency for bypassing an abasic site (Fig. 6*A*). Interestingly, when comparing all mutants derived from the 5P20 scaffold, it turned out that a single mutation (E232K) was sufficient for more bypass of the abasic site. However, the combination of this mutation with V269M results into a more efficient translesion enzyme. Interestingly, the V269M mutation alone is not sufficient to significantly increase lesion bypass ability compared with WT DNA polymerase β. Thus, the double mutant E232K/V269M appears to be as efficient as the parental 5P20 enzyme containing four mutations (Fig. 6*B*). However, every respective single mutant has significantly reduced abasic site bypass proficiency.

Crystallographic Structure of 5P20—To gain structural insights into the most active mutant 5P20, we conducted crystallization trials. Crystals of the 5P20 in complex with a 6-bp DNA duplex fragment were obtained as described (24) and diffracted to 1.85 Å (Table 2). Refinement led to good density for the majority of the enzyme, except for the first nine residues. Mutated sites could be identified in the electron density (see omit map in Fig. 7), except for S2G, which is located in the unresolved N-terminal region. F99L is located near the second helix-hairpin-helix (residues 101–106) motif without

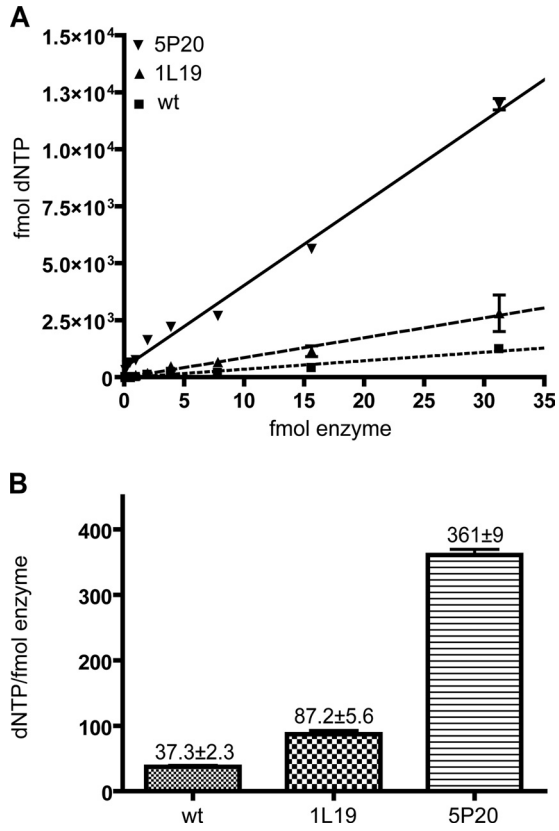


FIGURE 5. **Measuring the specific enzyme activity of WT human DNA polymerase β (■) and mutants 1L19 (▲) and 5P20 (▼).** *A*, amount of incorporated dNTPs depending on the amount of enzyme. *B*, depiction of the specific activity of WT human DNA polymerase β and mutants 1L19 and 5P20.

disturbing the metal binding ability, because the region has an average *B* of 24.9 Å^2, and the Na^+ ion shows a *B* value of 27.6 Å^2 and occupancy of 1. E232K is located in the template-binding loop region near residues 230–234. Mutation V269M is found near the hydrophobic hinge. In the complex, Lys-232 showed a conformer that enables a water-mediated contact to the phosphate group of template base 3 (dTMP), thereby forming an additional stabilizing contact to the substrate.

DISCUSSION

DNA polymerase β acts as the key gap-filling enzyme in base excision repair (8). Alterations in the proper function of the enzyme such as fidelity and processing of DNA lesions are crucial for genome integrity (14–20). Additionally, DNA polymerase β is a well studied DNA polymerase that can be regarded as a model enzyme for studies of DNA polymerase mechanisms in general (13). To identify amino acid residues that are critical for preventing the enzyme in incorporating a nucleotide opposite an abasic site, we employed a screening system based on arrayed primer extension reactions using SYBR® Green I fluorescence for double-stranded DNA product detection. From a more than 11,000-member library of randomly mutagenized DNA polymerase β entities, we could identify two variants, referred to as 5P20 and 1L19, which exhibit distinct alterations of enzyme properties.

In radiometric primer extension studies, we found that both mutants exhibit an increased propensity for incorporating a nucleotide opposite an abasic site. Single nucleotide incorporation and kinetic measurements (Figs. 2 and 3 and Table 1) disclose that dCMP and dGMP are most preferentially incorporated opposite the abasic site by all enzymes albeit

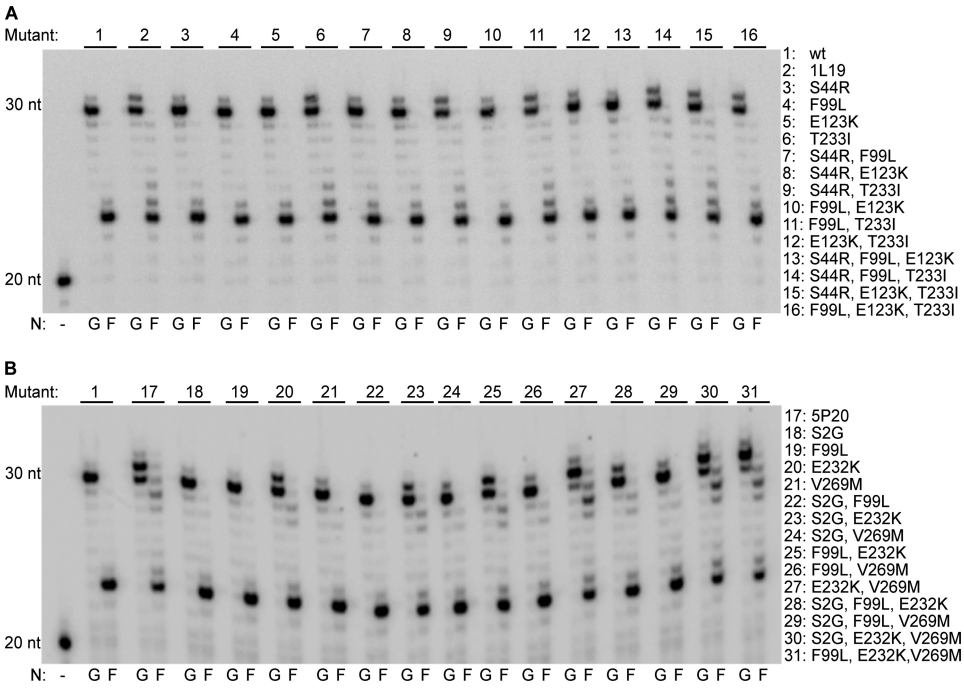


FIGURE 6. **Analyses of all combinations of mutations deriving from 1L19 and 5P20.** Running start primer extension studies using the 15 possible mutations of 1L19 (*A*) and 5P20 (*B*) in comparison with the WT enzyme as indicated on the *right side* of the figure. Primer template sequences are identical as in Fig. 1*A*. Side by side comparisons of DNA synthesis along undamaged template (indicated with *G* along the bottom of the gel image) and abasic site-containing template indicated with *F* were conducted. Respective mutations as well as combinations of mutations are indicated at the *right side*. All of the experiments were conducted employing the same condition and enzyme concentration.

TABLE 2
Crystallographic data and refinement statistics of 5P20 6bp

Protein Data Bank code	3OGU
Data collection	
a	175.5 Å
b	56.79 Å
c	47.61 Å
Resolution[a]	47,67 (1.96)-1.85 Å
R_{meas} [a,b]	6.5% (107.1%)
Completeness[a]	98.7% (92.4%)
$<I/\sigma>$[a]	16.13 (1.43)
Wilson B	33.04 Å^2
No. of observed reflections[a]	257039 (28089)
No. of unique reflections[a]	41386 (6158)
Wavelength	1.000 Å
Refinement	
Root mean square deviations	
Bond lengths	0.003 Å
Bond angles	0.7°
R_{work} [a,c]	19.7% (31.8%)
R_{free} [a,d]	23.3% (34.5%)
Average *B* factor	
Protein	40.2 Å^2
DNA	36.5 Å^2
Solvent	48.9 Å^2
Ramachandran	
Favored	97.8%
Allowed	1.9%
Outlier	0.3%

[a] The values in parentheses correspond to the highest resolution shell.
[b] For derivation and usage of R_{meas}, see Ref. 49.
[c] $R_{work} = 100 \times \Sigma||F_{obs}| - |F_{calc}||/\Sigma F_{obs}$.
[d] R_{free} for a subset of about 5% withheld from refinement.

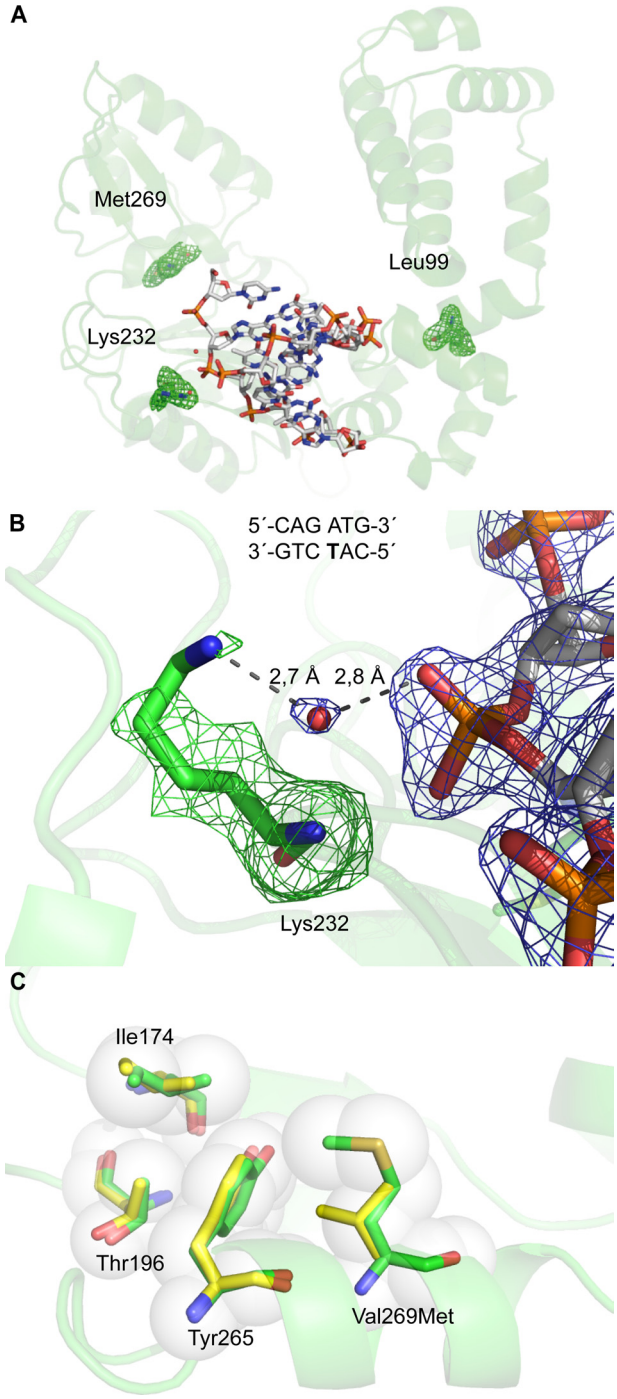


FIGURE 7. **Crystal structure of 5P20 mutant bound to DNA duplex.** *A*, overview of the mutated sites. *B*, detailed depiction of mutated residue Lys-232. The interaction of Lys-232 with the phosphate moiety of thymidine 3 (*bold*) in the templating strand via an H_2O molecule is highlighted. Lys-232 is shown with the omit map ($\sigma = 2.5$), and H_2O and DNA are shown in model density ($2mF_o - DF_c$). Omit maps were created using simulated annealing. *C*, outer lining of the hinge region containing mutation V269M. Shown is a superposition with the wild-type enzyme (Protein Data Bank code 9ICW, *yellow*) complexed to the same dsDNA. Molecular graphics were made using PyMol.

with varied efficiencies. The incorporation efficiency of dCMP by 5P20 was 53- and 14-fold increased for dGMP. 1L19 showed a 5-fold increase in incorporating efficiency of dCMP and 3-fold for dGMP compared with the WT DNA polymerase β. Although abasic sites are considered to be noninstructive, *in vitro* and *in vivo* studies of the abasic site or the stabilized tetrahydrofuran analog F have shown that adenine, and to a lesser extent guanine, is most frequently incorporated opposite the lesion by DNA polymerases from family A and B and has been termed "A rule" (43–46). However, it has been shown that DNA polymerases β use different, sequence-dependent mechanisms that might compete with the A rule when bypassing abasic sites (47). The preferred incorporation of dCMP and dGMP of all enzymes studied here indicates that the enzymes loop out the abasic site and use a mechanism in which nucleotides downstream to the abasic site are used as templates as reported before (47, 48). This incorporation signature suggests that the enzyme uses the dG moiety two positions downstream of the abasic site as template preferentially in this sequence context. Both DNA polymerase β variants have four mutations. The substitution F99L exists in common to both variants. The 20-fold decrease in K_d exhibited by the mutants represents an increase in DNA binding affinity. The tighter binding to primer template complex might increase the ability of the mutant enzymes to promote bypass of the lesion.

Testing all possible mutation combinations of the respective mutant showed that the substitution E232K is sufficient for the developing of the new properties in 5P20. V269M potentiates this effect, whereas S2G and F99L show no significant influence (Fig. 6*B*). Glu-232 is located in the palm domain, and Val-269 belongs to the so called "hydrophobic hinge" of the palm-thumb interface (17, 18). Val-269 is not part of the active site of DNA polymerase β and has no direct contact with either the primer template duplex or incoming dNTP (17, 23). Both relevant mutations in 5P20 are located near the amino acids, which get in contact to nucleobases or the phosphate backbone by hydrogen bonds. The exchange of

Glu at position 232 with Lys, an anionic amino acid against a cationic one, leads to a significant alteration in the side chain. The positive charge could interact with the anionic sugar phosphate backbone of the template, thereby increasing the affinity for the DNA substrate and stabilizing conformations that are required for nucleotide incorporation opposite an aberrant template site such as an abasic lesion. Indeed our structural studies indicate a water-mediated hydrogen bonding network of Lys-232 with a phosphate moiety in the template strand. Thus, an increase of the electrostatic interaction with the DNA could be one of the reasons for the increased bypass activity of 5P20. The relevance of positively charged amino acid on lesion bypass has been recently discussed. Patel *et al.* (50) isolated a variant of *Taq* DNA polymerase, I614K, which exhibited higher lesion bypass ability than the WT. Furthermore, Gloeckner *et al.* (51) found a mutant of *Klen-Taq* DNA polymerase containing an exchange of methionine to lysine, which showed a higher activity in bypassing DNA damages as well. Both mutants carry mutations of an apolar amino acid to a lysine near the template strand or the active site. Interestingly, the presence of a large positively charged DNA-binding surface has been discussed recently to promote DNA lesion bypass by human DNA polymerase η, an enzyme that is able to perform translesion synthesis (52).

The hydrophobic hinge, where the second relevant mutation in 5P20 (V269M) is located, is lined on the inside with hydrophobic residues Leu-194, Ile-260, Phe-272, and Tyr-296 and on the outside with residues Ile-174, Thr-196, Tyr-265, and Val-269. All of these amino acids build up stabilizing hydrophobic interaction (24). The roles of some amino acids in this region were determined by Sweasy and co-workers (53). By studying position 260, the influence of the hydrophobic nature of the amino acid on the activity and fidelity of the enzyme could be outlined. It was further postulated that position 265 is involved in the conformational change when the enzyme switches to the closed conformation. Furthermore, the aforementioned relevance of hydrophobicity on the polymerase fidelity has been elucidated (54–56). Together with Tyr-271 and Asp-276, Phe-272 forms the dNTP-binding pocket of DNA polymerase β (49). This residue seems to be critical for maintaining fidelity during the binding of the dNTP (57). Inside DNA polymerase family X, to which DNA polymerase β belongs, the residues building the nucleotide-binding pocket are highly conserved. Equivalent to position 269, exclusively nonpolar amino acids are present at the respective sites, mostly leucine. However, DNA polymerase β harbors a valine residue at the respective position. The mutation V269M in 5P20 induces a significant change of the constitution of the side chain, whereas the polarity remains the same. Although the isopropyl side chain in valine imposes considerable steric strain near the amino acid backbone, the methionine side chain is more flexible in this regard. It might be that the gain in flexibility by the V269M mutation influences the switch from open to closed conformation in which the region of the "hydrophobic hinge" is involved, thereby favoring nucleotide incorporation opposite aberrant template sides.

In mutant 1L19, the substitution T233I appears to be sufficient for the increased nucleotide incorporation opposite abasic sites. Structural data indicate that Thr-233 builds up hydrogen bonds to the DNA backbone via the hydroxyl functionality in the amino acid side chain (58). By substituting polar threonine by nonpolar and aliphatic isoleucine, the hydrogen bond can no longer be formed. How this loss of hydrogen bond forming capability translates into more efficient incorporation efficiency opposite an abasic site is unclear.

In summary, from a library of more than 11,000 entities of random human DNA polymerase β mutants, we have identified two variants that have increased ability to incorporate a nucleotide opposite an abasic site and have reduced fidelity. This study provides insights into the impact of amino acids residues that are involved in proper function of DNA polymerases. We identified residues that are remote from the active site and have significant impact on the activity, fidelity, and incorporation propensity opposite abasic sites of the enyzme. These findings support the model of adaptive mutations (59). This model postulates that the inherent plasticity of the DNA polymerase allows for mutations to change the substrate scope of the enzyme without compromising the general activity of the enzyme. Thereby, beneficial factors could be provided that allow for adapting to changes in the environment. Additionally, the knowledge of effects of amino acid impact on abasic site bypass and fidelity might provide insights into the role of tumor-associated mutants found in DNA polymerase β.

Acknowledgment—We would like to thank the DFG for funding within SPP 1170 and the beamline staff of the Swiss Light Source (SLS) for support.

REFERENCES

1. Saul, R. L., and Ames, B. N. (1986) *Basic Life Sci.* **38,** 529–535
2. Kulkarni, A., and Wilson, D. M., 3rd (2008) *Am. J. Hum. Genet.* **82,** 539–566
3. Hubscher, U., Maga, G., and Spadari, S. (2002) *Annu. Rev. Biochem.* **71,** 133–163
4. Lindahl, T. (1993) *Nature* **362,** 709–715
5. Loeb, L. A., and Preston, B. D. (1986) *Annu. Rev. Genet.* **20,** 201–230
6. Lindahl, T., and Wood, R. D. (1999) *Science* **286,** 1897–1905
7. Nilsen, H., and Krokan, H. E. (2001) *Carcinogenesis* **22,** 987–998
8. Seeberg, E., Eide, L., and Bjørås, M. T. (1995) *Trends Biochem. Sci.* **20,** 391–397
9. Beard, W. A., Shock, D. D., Batra, V. K., Pedersen, L. C., and Wilson, S. H. (2009) *J. Biol. Chem.* **284,** 31680–31689
10. Filée, J., Forterre, P., Sen-Lin, T., and Laurent, J. (2002) *J. Mol. Evol.* **54,** 763–773
11. Idriss, H. T., Al-Assar, O., and Wilson, S. H. (2002) *Int. J. Biochem. Cell Biol.* **34,** 321–324
12. Kumar, A., Widen, S. G., Williams, K. R., Kedar, P., Karpel, R. L., and Wilson, S. H. (1990) *J. Biol. Chem.* **265,** 2124–2131
13. Beard, W. A., and Wilson, S. H. (2006) *Chem. Rev.* **106,** 361–382
14. Dalal, S., Chikova, A., Jaeger, J., and Sweasy, J. B. (2008) *Nucleic Acids Res.* **36,** 411–422
15. Starcevic, D., Dalal, S., and Sweasy, J. B. (2004) *Cell Cycle* **3,** 998–1001
16. Lang, T., Maitra, M., Starcevic, D., Li, S. X., and Sweasy, J. B. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101,** 6074–6079
17. Dalal, S., Hile, S., Eckert, K. A., Sun, K. W., Starcevic, D., and Sweasy, J. B. (2005) *Biochemistry* **44,** 15664–15673

18. Sweasy, J. B., Lang, T., Starcevic, D., Sun, K. W., Lai, C. C., Dimaio, D., and Dalal, S. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 14350–14355
19. Sweasy, J. B., Lauper, J. M., and Eckert, K. A. (2006) *Radiat. Res.* **166,** 693–714
20. Lang, T., Dalal, S., Chikova, A., DiMaio, D., and Sweasy, J. B. (2007) *Mol. Cell Biol.* **27,** 5587–5596
21. Batra, V. K., Beard, W. A., Shock, D. D., Pedersen, L. C., and Wilson, S. H. (2008) *Mol. Cell* **30,** 315–324
22. Krahn, J. M., Beard, W. A., Miller, H., Grollman, A. P., and Wilson, S. H. (2003) *Structure* **11,** 121–127
23. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) *Biochemistry* **36,** 11205–11215
24. Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H., and Kraut, J. (1996) *Biochemistry* **35,** 12742–12761
25. Sobol, R. W., Horton, J. K., Kühn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K., and Wilson, S. H. (1996) *Nature* **379,** 183–186
26. Prakash, S., Johnson, R. E., and Prakash, L. (2005) *Annu. Rev. Biochem.* **74,** 317–353
27. Karras, G. I., and Jentsch, S. (2010) *Cell* **141,** 255–267
28. Daigaku, Y., Davies, A. A., and Ulrich, H. D. (2010) *Nature* **465,** 951–955
29. Menge, K. L., Hostomsky, Z., Nodes, B. R., Hudson, G. O., Rahmati, S., Moomaw, E. W., Almassy, R. J., and Hostomska, Z. (1995) *Biochemistry* **34,** 15934–15942
30. Kosa, J. L., and Sweasy, J. B. (1999) *J. Biol. Chem.* **274,** 3851–3858
31. Minnick, D. T., Astatke, M., Joyce, C. M., and Kunkel, T. A. (1996) *J. Biol. Chem.* **271,** 24954–24961
32. Casas-Finet, J. R., Kumar, A., Karpel, R. L., and Wilson, S. H. (1992) *Biochemistry* **31,** 10272–10280
33. Astatke, M., Grindley, N. D., and Joyce, C. M. (1995) *J. Biol. Chem.* **270,** 1945–1954
34. Abbotts, J., SenGupta, D. N., Zmudzka, B., Widen, S. G., Notario, V., and Wilson, S. H. (1988) *Biochemistry* **27,** 901–909
35. Kabsch, W. (2010) *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
36. Kabsch, W. (2010) *Acta Crystallogr. D Biol. Crystallogr.* **66,** 133–144
37. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954
38. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
39. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, Palo Alto, CA
40. Summerer, D., Rudinger, N. Z., Detmer, I., and Marx, A. (2005) *Angew. Chem. Int. Ed. Engl.* **44,** 4712–4715
41. Strerath, M., Gloeckner, C., Liu, D., Schnur, A., and Marx, A. (2007) *ChemBioChem* **8,** 395–401
42. Ahn, J., Werneburg, B. +., and Tsai, M. D. (1997) *Biochemistry* **36,** 1100–1107
43. Sagher, D., and Strauss, B. (1983) *Biochemistry* **22,** 4518–4526
44. Schaaper, R. M., Kunkel, T. A., and Loeb, L. A. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80,** 487–491
45. Strauss, B. S. (2002) *DNA Repair* **1,** 125–135
46. Obeid, S., Blatter, N., Kranaster, R., Schnur, A., Diederichs, K., Welte, W., and Marx, A. (2010) *EMBO J.* **29,** 1738–1747
47. Efrati, E., Tocco, G., Eritja, R., Wilson, S. H., and Goodman, M. F. (1997) *J. Biol. Chem.* **272,** 2559–2569
48. Efrati, E., Tocco, G., Eritja, R., Wilson, S. H., and Goodman, M. F. (1999) *J. Biol. Chem.* **274,** 15920–15926
49. Diederichs, K., and Karplus, P. A. (1997) *Nat. Struct. Biol.* **4,** 269–275
50. Patel, P. H., Kawate, H., Adman, E., Ashbach, M., and Loeb, L. A. (2001) *J. Biol. Chem.* **276,** 5044–5051
51. Gloeckner, C., Sauter, K. B., and Marx, A. (2007) *Angew. Chem. Int. Ed. Engl.* **46,** 3115–3117
52. Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-Maiques, S., Gregory, M., Lee, J. Y., Masutani, C., Lehmann, A. R., Hanaoka, F., and Yang, W. (2010) *Nature* **465,** 1044–1048
53. Starcevic, D., Dalal, S., and Sweasy, J. B. (2005) *Biochemistry* **44,** 3775–3884
54. Clairmont, C. A., Narayanan, L., Sun, K. W., Glazer, P. M., and Sweasy, J. B. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96,** 9580–9585
55. Shah, A. M., Li, S. X., Anderson, K. S., and Sweasy, J. B. (2001) *J. Biol. Chem.* **276,** 10824–10831
56. Opresko, P. L., Sweasy, J. B., and Eckert, K. A. (1998) *Biochemistry* **37,** 2111–2119
57. Li, S. X., Vaccaro, J. A., and Sweasy, J. B. (1999) *Biochemistry* **38,** 4800–4808
58. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science* **264,** 1891–1903
59. Patel, P. H., and Loeb, L. A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 5095–5100