Pre-Steady-State Kinetic Studies of the Fidelity of *Sulfolobus solfataricus* P2 DNA Polymerase IV†

Kevin A. Fiala‡ and Zucai Suo*,‡,§

Department of Biochemistry, Ohio State Biochemistry Program, Ohio State Biophysics Program, Molecular, Cellular and Developmental Biology Program, and Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210

Received September 26, 2003; Revised Manuscript Received December 23, 2003

ABSTRACT: *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4) is a thermostable archaeal enzyme and a member of the error-prone and lesion-bypass Y-family. In this paper, for the first time, the fidelity of a Y-family polymerase, Dpo4, was determined using pre-steady-state kinetic analysis of the incorporation of a single nucleotide into an undamaged DNA substrate 21/41-mer at 37 °C. We assessed single-turnover (with Dpo4 in molar excess over DNA) saturation kinetics for all 16 possible nucleotide incorporations. The fidelity of Dpo4 was estimated to be in the range of $10^{-3}-10^{-4}$. Interestingly, the ground-state binding affinity of correct nucleotides (70−230 μM) is 10−50-fold weaker than those of replicative DNA polymerases. Such a low affinity is consistent with the lack of interactions between Dpo4 and the bound nucleotides as revealed in the crystal structure of Dpo4, DNA, and a matched nucleotide. The affinity of incorrect nucleotides for Dpo4 is ∼2−10-fold weaker than that of correct nucleotides. Intriguingly, the mismatched dCTP has an affinity similar to that of the matched nucleotides when it is incorporated against a pyrimidine template base flanked by a 5′-template guanine. The incoming dCTP likely skips the first available template base and base pairs with the 5′-template guanine, as observed in the crystal structure of Dpo4, DNA, and a mismatched nucleotide. The mismatch incorporation rates, regardless of the 5′-template base, were ∼2−3 orders of magnitude slower than the incorporation rates for matched nucleotides, which is the predominant contribution to the fidelity of Dpo4.

Cellular DNA is continually under assault from a wide array of DNA-damaging agents such as reactive oxygen species, UV light, ionizing radiation, and radiomimetic chemicals (*1*). Various forms of DNA lesions stall the replication machinery by impeding the progression of replicative polymerases. Over the past few years, a novel superfamily of DNA polymerases, now known as the Y-family, has been identified, and its members collectively demonstrate an ability to bypass DNA lesions, in either an error-free or error-prone manner (*2*−*4*). The members of the Y-family are distributed among the three kingdoms of life and include *Escherichia coli* DNA polymerases IV (also known as DinB) and V (also known as UmuC), yeast DNA polymerase η (yPolη),[1] human DNA polymerases η (hPolη), κ (hPolκ), and ι (hPolι), and *Solfolobus solfataricus* DNA polymerases Dbh and Dpo4 (*5*). Dpo4, the focus of this paper, is shown to preferentially traverse abasic sites, followed by cisplatin−DNA adducts, *cis-syn* thymine−thymine (TT) dimers, 6−4 TT dimers, and acetylaminofluorene−DNA adducts (*6*).

Dpo4 is a thermostable DNA polymerase which can catalyze synthesis at 37 °C, exhibiting a misincorporation fidelity in the range of $10^{-3}-10^{-4}$ with undamaged DNA substrates as revealed by steady-state kinetic analysis (*6*, *7*) and by the implementation of a forward mutation assay (*7*). The latter assay scores a variety of substitution and frameshift errors generated during the copying of a lacZ template in a single-stranded gap of M13mp2 DNA. Surprisingly, the forward mutation assays also reveal that Dpo4 generates C·C mismatches at an unusually high rate (4×10^{-2}) and preferentially at cytosines flanked by a 5′-template guanine (*7*). Interestingly, Dbh, a Dpo4 homologue from *S. solfataricus* strain P1, makes C·C mismatches most frequently (*8*). Except for the unusually high rate of C·C mismatches, the error rates of the other 11 mismatches for Dpo4 are remarkably similar to those for hPolκ (*7*, *9*). Dpo4 also demonstrates very low frameshift fidelity and frequently generates deletions of even noniterated nucleotides, especially cytosine flanked by a 5′-template guanine (*7*). The frameshift error rate of Dpo4 (*7*) is similar to those of hPolκ (*9*) and hPolη (*10*), while being much higher than those of human DNA polymerases α (*11*), β (*12*), and γ (*13*). Interestingly, *E. coli* DNA polymerase IV (*14*, *15*) and *S. solfataricus* Dbh (*16*) are also found to predominantly generate a −1 frame-

† This work was supported in part by American Chemical Society Petroleum Research Fund Grant PRF38364-G4, by American Cancer Society Grant IRG-98-278-03, and by the TriLink Biotechnologies Research Funding Program (Z.S.). K.A.F. was supported by the National Institutes of Health Chemistry and Biology Interface Program at The Ohio State University (Grant T32 GM08512-08).

* To whom correspondence should be addressed. Telephone: (614) 688-3706. Fax: (614) 292-6773. E-mail: suo.3@osu.edu.

‡ Department of Biochemistry and Ohio State Biochemistry Program.

§ Ohio State Biophysics Program, Molecular, Cellular & Developmental Biology Program, and Comprehensive Cancer Center.

[1] Abbreviations: Ap, 2-aminopurine; BSA, bovine serum albumin; CPD, *cis-syn* cyclobutane pyrimidine dimer; Dbh, DinB homologue; dNTP, deoxynucleoside 5′-triphosphate; Dpo4, *S. solfataricus* P2 DNA polymerase IV; DTT, dithiothreitol; hPolγ, human mitochondrial DNA polymerase γ; hPolη, human DNA polymerase η; hPolι, human DNA polymerase ι; hPolκ, human DNA polymerase κ; $MgAc_2$, magnesium acetate; Pol, polymerase; Py, pyrimidine; rPolβ, rabbit DNA polymerase β; TBE, Tris/borate/EDTA electrophoresis buffer; yPolη, yeast DNA polymerase η.

10.1021/bi0357457 CCC: $27.50 © 2004 American Chemical Society
Published on Web 01/29/2004

shift at a pyrimidine flanked by a 5′-template guanine. On the basis of these findings, a "dNTP-stabilized" misalignment mechanism is hypothesized to be responsible for making −1 frameshift mutations for these Y-family polymerases (*7*). This mechanism suggests the incoming nucleotide, dCTP, forms a correct base pair with a downstream template base G on a "looped out" template strand instead of forming an incorrect base pair with the first available template base (*14*). A mutagenic intermediate, important for verifying this mechanism, is observed in the crystal structure of Dpo4 in a ternary complex with DNA and a mismatched incoming nucleotide (*17*).

In addition to that of Dpo4, the fidelities of several Y-family polymerases, including yPolη (*16*, *18*), hPolη (*10*, *19*), hPolκ (*9*, *20*), hPolι (*21*−*23*), *E. coli* polymerase IV (*14*), and Dbh (*16*), have been estimated by steady-state kinetic analyses. Although the steady-state kinetic methods can be used to estimate the polymerase fidelity and substrate specificity, the kinetic and thermodynamic basis for the fidelity of the Y-family polymerases has not been established because of the inherent limitations of steady-state kinetics (*24*). Like other polymerases, the mechanistic basis of enzyme fidelity can only be established by employing pre-steady-state kinetic methods (*24*). So far, pre-steady-state kinetic studies on the novel Y-family have been used to determine the fidelity of yPolη based on one correct and one incorrect nucleotide incorporation (*25*), but the other 14 possible incorporations were not analyzed. In this paper, we chose Dpo4 as a model enzyme from the Y-family and employed pre-steady-state kinetic techniques to study, for the first time, the replication fidelity and its kinetic and thermodynamic basis, based on all 16 possible nucleotide incorporations into undamaged DNA substrates. The dNTP-stabilized misalignment mechanism will also be examined. In the following paper (*26*), we establish the kinetic mechanism of DNA polymerization catalyzed by Dpo4.

EXPERIMENTAL PROCEDURES

Materials. These chemicals were purchased from the following companies: [γ-^{32}P]ATP from Perkin-Elmer Life Sciences (Boston, MA), dNTPs from Gibco-BRL (Rockville, MD), calf intestine alkaline phosphatase from Fermentas (Hanover, MD), T4 polynucleotide kinase from USB (Cleveland, OH), and Biospin columns from Bio-Rad Laboratories (Hercules, CA).

Cloning and Purification of Dpo4. The dpo4 gene from *S. solfataricus* P2 was cloned into the *Nde*I and *Xho*I sites of pET22b. Dpo4 fused to a C-terminal His_6 tag was expressed in *E. coli* strain BL21(DE3). An overnight culture of *E. coli* expression strain BL21(DE3) carrying this plasmid was used to inoculate Luria-Bertani medium containing 100 μg/mL ampicillin, and the cells were grown at 22 °C. After the OD_{600} reached 0.5, cultures were induced with 0.2 mM IPTG and incubated at 22 °C until the OD_{600} reached 1.7. Cells were harvested (4000 rpm for 15 min) and resuspended in buffer A [10 mM $KHPO_4$ (pH 7.0), 50 mM NaCl, 10 mM $MgAc_2$, 10% glycerol, and 0.1% 2-mercaptoethanol]. Cells were then lysed by passing them through a French press at 16 000 psi twice, and the resulting lysate was cleared by ultracentrifugation (35 000 rpm for 40 min). Cleared lysate was incubated at 78 °C for 12 min to precipitate thermolabile *E. coli* proteins which were then removed by ultracentrifugation. The supernatant containing predominantly Dpo4 was pooled and incubated overnight at 4 °C with nickel−NTA superflow resin (Qiagen). The supernatant was removed by centrifugation in a swing-bucket centrifuge (2500 rpm for 10 min), and the Dpo4-bound nickel resin was packed into a column. Bound proteins were eluted through a linear gradient from 25 to 350 mM imidazole in buffer B [10 mM $KHPO_4$ (pH 7.0), 0.35 M NaCl, 5−500 mM imidazole, 2.5 mM $MgAc_2$, 10% glycerol, and 0.1% 2-mercaptoethanol]. Dpo4-containing fractions were pooled and applied to a MonoS column (Pharmacia Biotech) and eluted via a three-step gradient (from 50 to 240 mM NaCl, 240 mM NaCl, and from 240 to 1000 mM NaCl) in buffer C [50 mM Tris-HCl (pH 7.0), 10−1000 mM NaCl, 2 mM EDTA, 10% glycerol, and 0.1% 2-mercaptoethanol]. Fractions containing Dpo4 were pooled, dialyzed against buffer D [50 mM Tris-acetate (pH 7.5), 50 mM NaAc, 1 mM DTT, 0.5 mM EDTA, and 10% glycerol] twice, and concentrated using a Centriprep YM-30 apparatus (Millipore). The concentrated Dpo4 was finally dialyzed against the storage buffer [50 mM Tris-acetate (pH 7.5), 50 mM NaAc, 1 mM DTT, 0.5 mM EDTA, and 50% glycerol]. Dpo4 was purified to >95% purity on the basis of staining SDS−PAGE gels with Coomassie Blue R-250 (Figure 1) and resulted in a yield of approximately 50 mg from 9 L of initial *E. coli* culture. The concentration of the purified Dpo4 was measured spectrophotometrically at 280 nm using the calculated extinction coefficient of 24 058 M^{-1} cm^{-1}.

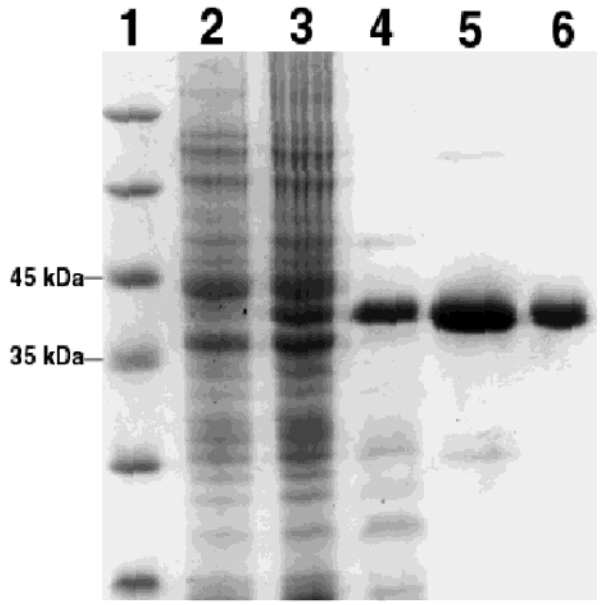

FIGURE 1: Purification of *S. solfataricus* Dpo4. SDS−PAGE analysis and subsequent Coomassie Blue staining of proteins at each step of the purification are shown: lane 1, protein marker; lane 2, crude extracts of noninduced cells; lane 3, crude extracts of IPTG-induced cells; lane 4, soluble fraction after cleared lysate was incubated at 78 °C for 12 min; lane 5, eluate from the Ni affinity column; and lane 6, eluate from the MonoS cation-exchange column.

Synthetic Oligonucleotides. The DNA substrates listed in Table 1 were purchased from either Integrated DNA Technologies or TriLink Biotechnologies and purified by denaturing polyacrylamide gel electrophoresis (18% acrylamide and 8 M urea), and the concentration was determined by the UV absorbance at 260 nm with the following

Table 1: DNA Substrates

D-1	5′-CGCAGCCGTCCAACCAACTCA-3′ 3′-GCGTCGGCAGGTTGGTTGAGTAGCAGCTAGGTTACGGCAGG-5′
D-6	5′-CGCAGCCGTCCAACCAACTCA-3′ 3′-GCGTCGGCAGGTTGGTTGAGTGGCAGCTAGGTTACGGCAGG-5′
D-7	5′-CGCAGCCGTCCAACCAACTCA-3′ 3′-GCGTCGGCAGGTTGGTTGAGTTGCAGCTAGGTTACGGCAGG-5′
D-8	5′-CGCAGCCGTCCAACCAACTCA-3′ 3′-GCGTCGGCAGGTTGGTTGAGTCGCAGCTAGGTTACGGCAGG-5′
D-12	5′-CGCAGCCGTCCAACCAACTCA-3′ 3′-GCGTCGGCAGGTTGGTTGAGTCACAGCTAGGTTACGGCAGG-5′

extinction coefficients: $\epsilon = 194\,100\ M^{-1}\ cm^{-1}$ for 21-mer, $\epsilon = 396\,700\ M^{-1}\ cm^{-1}$ for D-1 41-mer, $\epsilon = 394\,200\ M^{-1}\ cm^{-1}$ for D-6 41-mer, $\epsilon = 392\,200\ M^{-1}\ cm^{-1}$ for D-7 41-mer, $\epsilon = 389\,500\ M^{-1}\ cm^{-1}$ for D-8 41-mer, and $\epsilon = 392\,400\ M^{-1}\ cm^{-1}$ for D-12 41-mer.

Labeling and Annealing of the DNA Substrates. The primer strand 21-mer was 5′-labeled with ^{32}P by incubation with T4 polynucleotide kinase and [γ-^{32}P]ATP for 1 h at 37 °C. After the kinase had been inactivated via heating of the reaction mixture for 5 min at 95 °C, unreacted [γ-^{32}P]ATP was subsequently removed by centrifugation via a Biospin-6 column (Bio-Rad). The 5′-^{32}P-labeled 21-mer was then annealed with the corresponding nonradiolabeled DNA 41-mer at a molar ratio of 1.0:1.1 to form the DNA complex of 21/41-mer. Mixtures to be annealed were denatured at 95 °C for 8 min, and then cooled slowly to room temperature for several hours.

Buffers. All experiments that included Dpo4, if not specified, were performed in buffer R [50 mM HEPES (pH 7.5 at 37 °C), 5 mM $MgCl_2$, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/mL BSA]. All reactions were carried out at 37 °C.

Rapid-Quench Experiments. Experiments were carried out in a rapid chemical quench flow apparatus (KinTek). The apparatus contained a computer-controlled stepping motor and was modified for small reaction volumes (15 μL). Invariably, the experiments were carried out by allowing the enzyme and DNA to be preincubated in buffer R. An aliquot of this solution (15 μL) was loaded into a sample loop and rapidly mixed with an equal volume of a solution containing nucleotide in buffer R from a second sample loop. The reactions were quenched with 90 μL of 0.37 M EDTA (final concentration) after time intervals ranging from 5 ms to several minutes. All concentrations reported in this paper refer to concentrations during the reaction following rapid mixing.

Measurement of Substrate Specificity. A preincubated solution of Dpo4 and DNA at fixed concentrations was mixed with varying concentrations of Mg^{2+}·dNTP (5−2400 μM) in buffer R at 37 °C to start the reaction. The reaction at each concentration of Mg^{2+}·dNTP was terminated with 0.37 M EDTA at varying times ranging from milliseconds to minutes. The reaction products were analyzed by sequencing gel analysis. The time course of product formation was fit to a single-exponential equation for each concentration of Mg^{2+}·dNTP (see Data Analysis) to give the observed rate of nucleotide incorporation. The observed rates extracted from a series of time courses of product formation were plotted against the concentrations of Mg^{2+}·dNTP, and the data were fit to a hyperbola (see Data Analysis) to give the equilibrium dissociation constant of dNTP (K_d) and the maximum rate for incorporation of dNTP (k_p). The substrate specificity (k_p/K_d) was then calculated.

Product Analysis. Reaction products were analyzed by sequencing gel electrophoresis (17% acrylamide, 8 M urea, and 1× TBE running buffer) and quantitated with a Phosphorimager 445 SI (Molecular Dynamics).

Data Analysis. Data were fit by nonlinear regression using KaleidaGraph (Synergy Software). The single-turnover experimental data were fit to eq 1 (single-exponential equation)

$$[\text{product}] = A[1 - \exp(-k_{obs}t)] \quad (1)$$

where A represents the reaction amplitude or initial concentration of the binary complex of the enzyme and DNA and k_{obs} the observed single-turnover rate. Data from the measurement of K_d of dNTP were fit to eq 2 (hyperbola)

$$k_{obs} = k_p[\text{dNTP}]/([\text{dNTP}] + K_d) \quad (2)$$

where k_p is the maximum rate of dNTP incorporation. The substrate specificity and polymerase fidelity were calculated as k_p/K_d and $(k_p/K_d)_{incorrect}/(k_p/K_d)_{correct}$, respectively.

RESULTS

Protein Purification. Dpo4 (352 amino acid residues and 40.2 kDa) is one of the few Y-family polymerases which can be copiously overexpressed and purified from *E. coli* in high yield (*6*). We modified the purification protocol published previously (*6*) by adding a C-terminal hexahistidine tag to wild-type Dpo4 for the convenience of protein purification. The C-terminal hexahistidine-tagged Dpo4 was overexpressed in *E. coli* and purified through a series of three steps: heat denaturation at 78 °C for 12 min, Ni affinity chromatography, and MonoS cation-exchange chromatography (Figure 1). The heat denaturation step precipitated most of the *E. coli* proteins, but the thermostable Dpo4 remained soluble. The purified Dpo4 was ~88% active on the basis of the results of an active site titration [see the following paper (*26*)]. To evaluate the effect of the C-terminal hexahistidine tag on the activity of Dpo4, we also purified wild-type Dpo4 lacking the C-terminal hexahistidine tag following the published purification protocol (*6*) (data not shown). We subsequently performed kinetic experiments with the purified wild-type Dpo4, and it was found to incorporate a single nucleotide at a rate of $4.7 \pm 0.3\ s^{-1}$ which is similar to the rate of $3.8 \pm 0.2\ s^{-1}$ observed with the C-terminal hexahistidine-tagged Dpo4 under similar reaction conditions. These results demonstrated the C-terminal hexahistidine tag had little effect on the activity of Dpo4. This is not surprising since the crystal structure reveals that the C-terminus of Dpo4 is exposed on the protein surface and is distal from the active site and the enzyme-bound DNA (*17*). It should be noted that the C-terminal hexahistidine-tagged Dpo4 was used in all experiments reported below and in the following paper (*26*).

Optimization of Reaction Conditions. Unless otherwise stated, all reactions in this paper were performed at 37 °C, rather than at 80 °C, the optimal growth temperature of S. *solfataricus.* This was done primarily for four reasons. First, the rapid chemical-quench apparatus used in our studies cannot be operated at 80 °C without causing irreversible damage to the apparatus itself. Second, with regard to DNA

melting temperatures, the primer would have to be significantly longer than 21 nucleotides to allow the annealed DNA substrate to remain stable at 80 °C. The products generated by incorporation of single nucleotides into a longer primer would become increasingly difficult to separate from the unreacted substrate by sequencing gel electrophoresis. The difficulty incurred in product analysis would likely hinder data acquisition and thus the overall veracity of the results. Third, kinetic data obtained at 37 °C allow us to make direct comparisons with other polymerases assayed at the same temperature. Fourth, we attempted to measure nucleotide incorporation rates at more physiologically relevant temperatures for Dpo4 by performing manual-quench experiments for incorporation of a mismatched nucleotide, due to the aforementioned limitations in instrument operation at elevated temperatures and much slower incorporation of a mismatched nucleotide versus a matched one. Experiments involving the misincorporation of dATP into D-1 were initially carried out at 60 °C (with the appropriate adjustment of the pH of the reaction buffer to 7.5 at 60 °C); however, even at the shortest time intervals (7 s) we could accomplish manually, we found all substrate (21-mer) was converted to product (22-mer). We then decided to carry out the same experiments at a slightly reduced temperature, 50 °C (with the appropriate adjustment of the buffer pH to 7.5 at 50 °C). These experiments yielded rates of incorporation of incorrect nucleotides that were on average 24-fold faster than the rates observed at 37 °C (see below). Our observations were not unprecedented. Dbh, a homologous polymerase from *S. solfataricus* strain P1, has been shown to have a 40-fold higher nucleotide incorporation rate at 65 °C than at 22 °C (*16*). Therefore, if the rate increase for correct nucleotide incorporation is similar to what we observed for incorrect nucleotides, the incorporation of a correct nucleotide catalyzed by Dpo4 at 80 °C should be too fast to be measured by a rapid chemical-quench apparatus since the apparatus is limited by its mixing dead time of 1−3 ms. However, it is not unprecedented to study the kinetics of a polymerase at a nonphysiological temperature. For example, the kinetic studies of T7 phage DNA polymerase were performed at 20 °C, rather than at 37 °C, and the overall results were not affected except the reaction rates were slower (*27*−*29*).

To optimize the reaction conditions, all components were kept constant while the $MgCl_2$ concentration, NaCl concentration, and pH were individually varied. Using a rapid chemical-quench apparatus, a preincubated solution of Dpo4 and 5′-[^{32}P]D-1 (Table 1) was mixed with dTTP at 37 °C to initiate the reaction. Since Dpo4 was in 3-fold molar excess over the DNA substrate, ~90% of D-1 should be complexed to Dpo4 on the basis of the measured affinity of 10.6 nM in the following paper (*26*). The reactions were quenched with EDTA at various time intervals. The DNA product 22-mer and unreacted substrate 21-mer were separated by sequencing gel electrophoresis, and quantitated using a Phosphorimager. The data were fit to eq 1 (see Experimental Procedures). The single-turnover rate varied dramatically with the concentration of Mg^{2+}, with the concentration of NaCl, and with the buffer pH. The optimal conditions for Dpo4 polymerization were 5 mM $MgCl_2$ (Figure 2A), 0−75 mM NaCl (Figure 2B), and pH 7.5 (Figure 2C). These data demonstrated a significant decrease in the single-turnover rates when assayed outside optimum conditions. Thus, the opti-

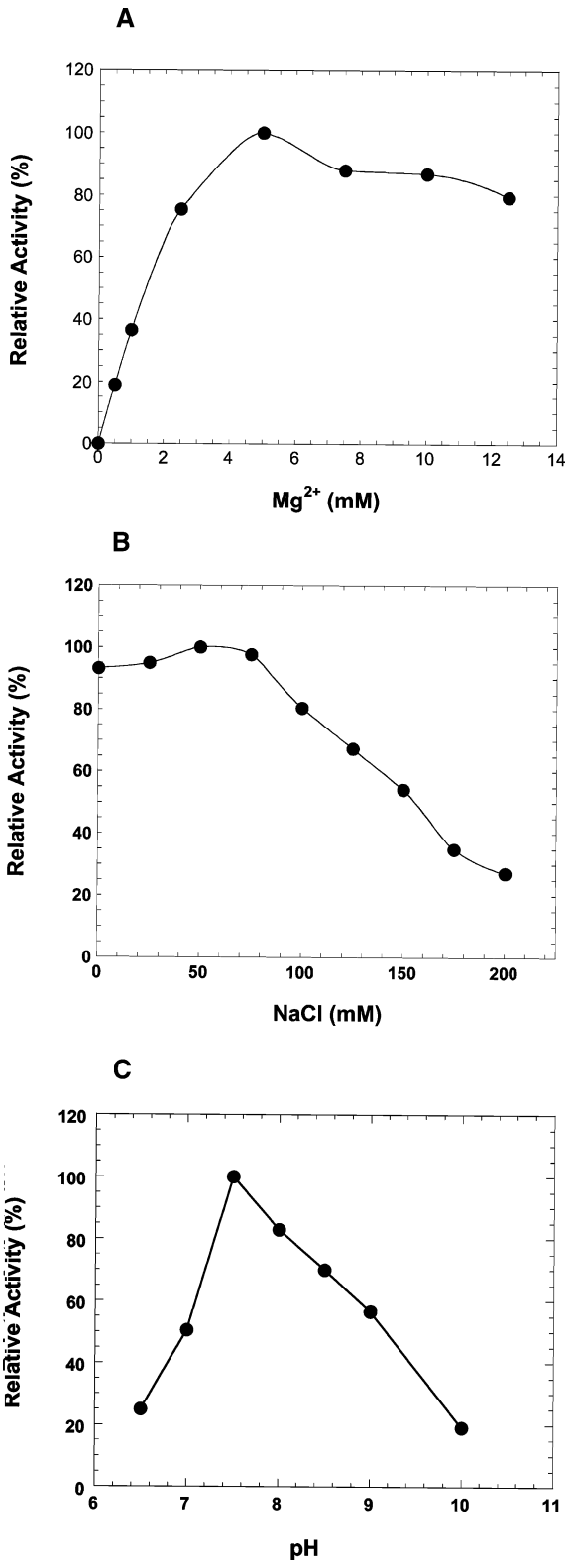

FIGURE 2: Effects of Mg^{2+} concentration, NaCl concentration, and buffer pH on the enzymatic activity of Dpo4. A preincubated solution of D-1 5′-labeled with ^{32}P (30 nM) and a 3-fold excess of Dpo4 (120 nM) was rapidly mixed with the correct nucleotide (100 μM dTTP) for various time intervals under single-turnover conditions. Concentrations of all components were held constant while the (A) Mg^{2+} concentration, (B) NaCl concentration, or (C) buffer pH was varied. Activity in panel C was assayed in 25 mM MES-NaOH buffer between pH 6.0 and 7.0, in 25 mM Tris-HCl buffer for pH 8.0 and 8.5, and in 25 mM glycine-NaOH buffer for pH 9.0 and 10.0.

Scheme 1

$$E{\bullet}DNA_n + dNTP \underset{}{\overset{K_d}{\rightleftharpoons}} E{\bullet}DNA_n{\bullet}dNTP \overset{k_p}{\rightleftharpoons} E'{\bullet}DNA_{n+1}{\bullet}PPi$$

mized reaction buffer contains 50 mM HEPES (pH 7.5), 5 mM $MgCl_2$, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/mL BSA (buffer R).

Substrate Specificity of the Correct Nucleotide. An incoming nucleotide specifically binds at the polymerase active site only after the binding of DNA to the enzyme, to form the E·DNA binary complex (*30*). The ground-state binding affinity of dTTP (K_d) for the binary complex (Scheme 1) was measured through the dTTP concentration dependence of the single-turnover rate (k_{obs}) under conditions where the concentration of the enzyme was 3-fold greater than the DNA concentration to ensure almost all of the DNA molecules were bound by Dpo4 (see above). A preincubated solution of Dpo4 (120 nM) and radiolabeled D-1 (30 nM) was mixed with increasing concentrations of dTTP in buffer R. EDTA was then added to the reaction mixtures to quench the polymerization at various time intervals. The DNA products and remaining substrate 21-mer were separated by sequencing gel electrophoresis, and quantitated using a Phosphorimager. The product concentration was plotted against the reaction time, and the data were fit to eq 1 (see Experimental Procedures) to yield a single-turnover rate at each concentration of dTTP (Figure 3A). The single-turnover rates were then plotted against dTTP concentrations (Figure 3B). The data were subsequently fit to eq 2 (see Experimental Procedures) to yield a k_p of 9.4 ± 0.3 s^{-1} for the maximum dTTP incorporation rate and a K_d of 230 ± 17 μM for the binding of dTTP. Since the value of K_d was unusually high for a correct nucleotide binding to a polymerase·DNA complex, we repeated the measurement and confirmed that this was not an experimental artifact. We also measured and obtained similar K_d and k_p values when DNA was in molar excess over Dpo4 (data not shown). This suggested the single-turnover method used in Figure 3 did not affect the measured K_d values. To further discredit a possible experimental anomaly, the affinity of the correct nucleotide for human polymerase lambda was assayed independently using the same single-turnover method and gave a low K_d of 2.4 μM (K. Fiala and Z. Suo, unpublished results). These observations are consistent with the finding that free DNA polymerase either does not bind a nucleotide prior to the binding of DNA or binds a nucleotide nonspecifically and weakly (*30*). Additionally, in the case in which Dpo4 was binding nucleotides nonspecifically, the relatively small amount of excess enzyme (90 nM) should not affect the experimental results since the dTTP concentrations were up to 4 orders of magnitude greater than the enzyme excess as seen in Figure 3. Alternatively, the K_d value could be affected by the concentration of free Mg^{2+} in the reaction buffer since dTTP also binds Mg^{2+}. We repeated the experiments in Figure 3 with the concentration of free Mg^{2+} kept constant at 5 mM by adding extra $MgCl_2$ to compensate for the binding of Mg^{2+} by dTTP. The results from these experiments showed similar values of K_d and k_p for incorporation of dTTP into D-1 (data not shown). Thus, the affinity of correct nucleotide dTTP for Dpo4, regardless of Mg^{2+} compensation, was more than 10- and 50-fold weaker than the affinities of replicative polymerases (*27*−*29*, *31*−*33*).

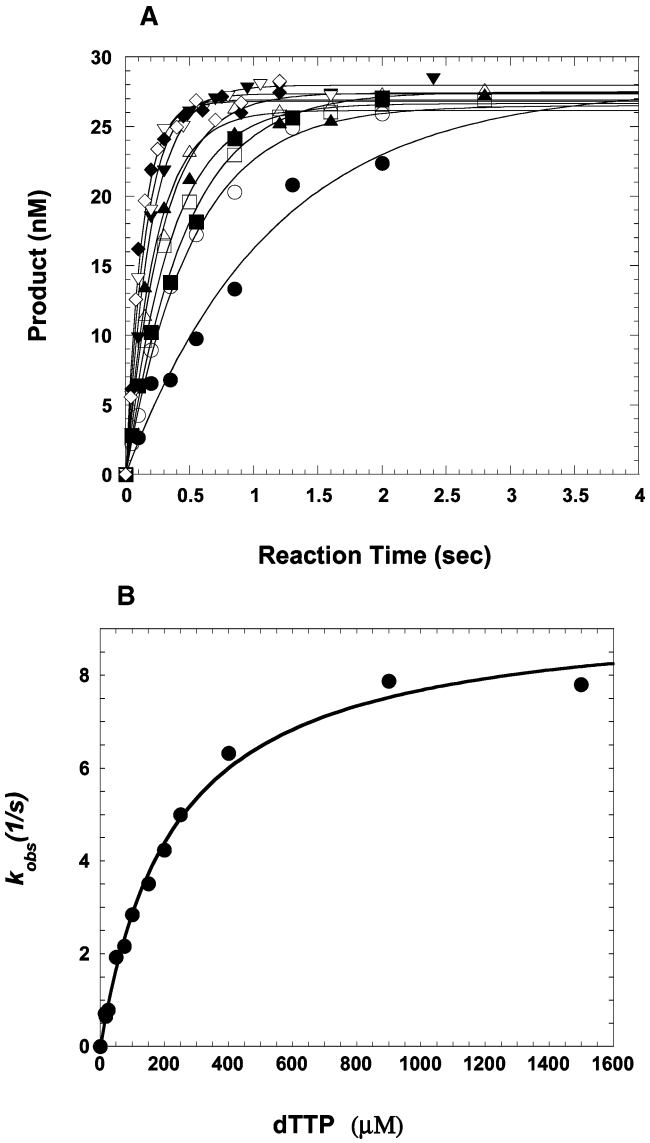

FIGURE 3: Concentration dependence on the pre-steady-state rate of correct nucleotide incorporation. (A) A preincubated solution of Dpo4 (120 nM) and D-1 5′-labeled with ^{32}P (30 nM) was mixed with increasing concentrations of Mg^{2+}·dTTP [17 (●), 50 (○), 75 (■), 100 (□), 150 (△), 200 (▲), 250 (▼), 400 (▽), 900 (◆), and 1500 μM (◇)] for various time intervals. The solid lines are the best fits to the single-exponential equation (eq 1). (B) The single-exponential rates obtained from the above data fitting were plotted as a function of dTTP concentration. The rate data were then fit to the hyperbolic equation (eq 2), yielding a k_p of 9.4 ± 0.3 s^{-1} and a K_d of 230 ± 17 μM.

Additionally, the value of the substrate specificity (k_p/K_d) was calculated to be 0.0409 μM^{-1} s^{-1} (Table 2) for the incorporation of dTTP into D-1.

Similar single-turnover measurements were performed with the incorporations of dCTP into D-6, dATP into D-7, and dGTP into D-8 (data not shown), and the measured kinetic parameters k_p, K_d, and k_p/K_d are listed in Table 2. The ground-state binding affinity of dCTP (70 μM) is 2−3-fold higher than the affinity of the other three correct nucleotides. This could be due to the downstream template guanine (Table 1), leading to two consecutive dCTP incorporations with different polymerization kinetics.

Substrate Specificity of the Incorrect Nucleotide. Pre-steady-state kinetic analysis of a single incorrect dCTP

Table 2: Pre-Steady-State Kinetic Parameters of Dpo4

dNTP	K_d (μM)	k_p (s^{-1})	k_p/K_d (μM^{-1} s^{-1})	fidelity[a]
template A (D-1)				
dTTP	230 ± 17	9.4 ± 0.3	4.09×10^{-2}	
dATP	578 ± 188	0.006 ± 0.001	9.86×10^{-6}	2.4×10^{-4}
dCTP	1036 ± 95	0.013 ± 0.001	1.25×10^{-5}	3.1×10^{-4}
dGTP	1150 ± 312	0.007 ± 0.001	6.00×10^{-6}	1.5×10^{-4}
template G (D-6)				
dCTP	70 ± 8	7.6 ± 0.2	1.09×10^{-1}	
dATP	334 ± 128	0.009 ± 0.001	2.69×10^{-5}	2.5×10^{-4}
dTTP	1283 ± 150	0.077 ± 0.004	6.00×10^{-5}	5.5×10^{-4}
dGTP	131 ± 42	0.008 ± 0.001	6.11×10^{-5}	5.6×10^{-4}
template T (D-7)				
dATP	206 ± 46	16.1 ± 0.9	7.82×10^{-2}	
dTTP	1941 ± 589	0.034 ± 0.006	1.75×10^{-5}	2.2×10^{-4}
dCTP	309 ± 84	0.026 ± 0.002	8.41×10^{-5}	1.1×10^{-3}
dGTP	935 ± 129	0.066 ± 0.004	7.06×10^{-5}	9.0×10^{-4}
template C (D-8)				
dGTP	171 ± 15	9.4 ± 0.2	5.50×10^{-2}	
dATP	617 ± 113	0.016 ± 0.001	2.59×10^{-5}	4.7×10^{-4}
dCTP	192 ± 48	0.034 ± 0.002	1.77×10^{-4}	3.2×10^{-3}
dTTP	913 ± 181	0.011 ± 0.001	1.20×10^{-5}	2.2×10^{-4}
template C (D-12)				
dCTP	1228 ± 266	0.005 ± 0.001	4.23×10^{-6}	

[a] Calculated as $(k_p/K_d)_{incorrect}/(k_p/K_d)_{correct}$.

incorporation into D-1 (Table 1) was carried out in a manner analogous to those experiments described above for the specificity of correct nucleotide incorporation. Dpo4 (120 nM), preincubated with radiolabeled D-1 (30 nM), was reacted with varying concentrations of dCTP in buffer R. The reactions were quenched with 0.37 M EDTA and analyzed by sequencing gel electrophoresis, and the products were quantitated using a Phosphorimager. The observed single-turnover rate at each concentration of dCTP was obtained through the fit of the time course of product formation to eq 1 (Figure 4A). The observed reaction rates were then plotted against the concentrations of dCTP, and the data were fit into eq 2 (see Experimental Procedures) to yield k_p, K_d, and substrate specificity (k_p/K_d) values of 0.013 ± 0.001 s^{-1}, 1036 ± 95 μM, and 1.25 × 10^{-5} μM^{-1} s^{-1}, respectively (Figure 4B).

With DNA substrates D-1 and D-6 to D-8, the kinetic parameters for the incorporations of all possible incorrect single nucleotides were determined under single-turnover conditions (Table 2). Overall, the incorrect nucleotide incorporations have 2−10-fold lower ground-state binding affinities (K_d) and 2−3 orders of magnitude lower incorporation rates in comparison to the four correct nucleotide incorporations. Surprisingly, the ground-state binding affinity of the mismatched C·T (D-7) and C·C (D-8) pairs is slightly lower than that of correct A·T and G·C base pairs, although the maximum incorporation rates still differ by several hundred fold. Interestingly, the downstream template base for these two incorrect incorporations is a guanine, which could form a matched base pair if the incoming nucleotide dCTP skips the first available template base (Table 1). To provide further evidence for such a "dNTP-stabilized" misalignment mechanism (*14*), we changed the downstream template base G (D-8) to an A (D-12) and measured the kinetic parameters for single dCTP incorporation into D-12 under the same single-turnover conditions (data not shown). The sequence of the D-12 template (Table 1) prevents the "dNTP"-stabilized dCTP incorporation. The measured K_d (1228 ± 265 μM) is ~6-fold larger than the K_d of

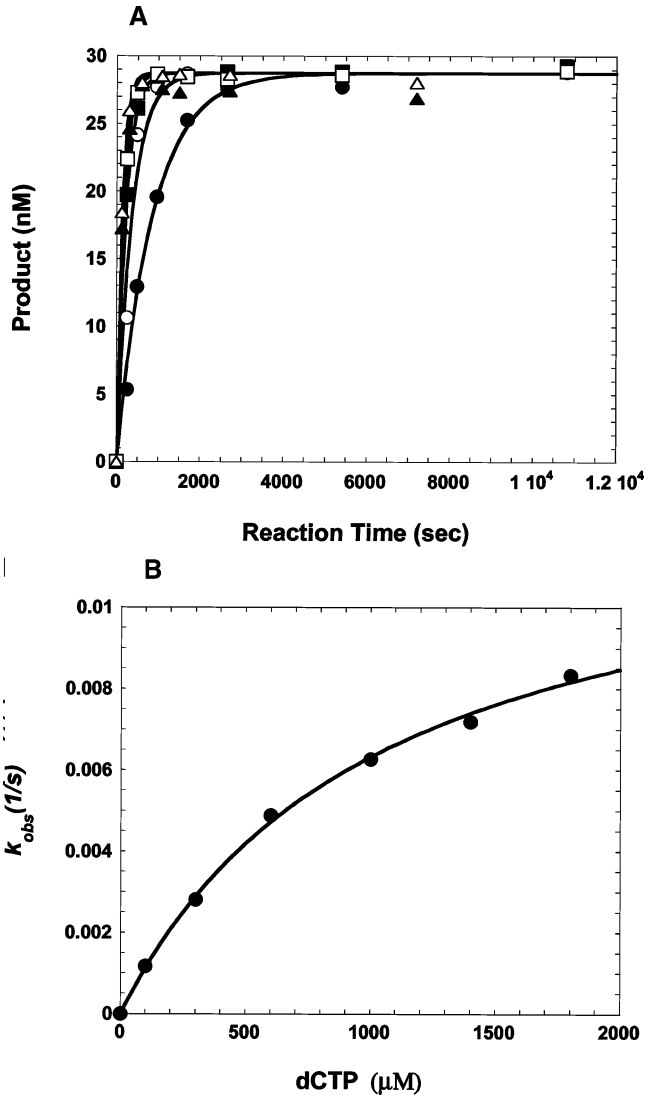

FIGURE 4: Concentration dependence on the pre-steady-state rate of incorrect nucleotide incorporation. (A) A preincubated solution of Dpo4 (120 nM) and D-1 5′-labeled with ^{32}P (30 nM) was mixed with increasing concentrations of Mg^{2+}·dCTP [100 (●), 300 (○), 600 (■), 1000 (□), 1400 (▲), and 1800 μM (△)] for various time intervals. The solid lines are the best fits to the single-exponential equation. (B) The single-exponential rates obtained from the above data fitting were plotted as a function of dCTP concentration. The rate data were then fit to the hyperbolic equation, yielding a k_p of 0.013 ± 0.001 s^{-1} and a K_d of 1036 ± 95 μM.

incorporation of dCTP into D-8 (192 μM), supporting this mutagenic mechanism. Interestingly, the above observations did not occur to the mismatched C·A base pair in D-1, although the next template base is also a guanine (Table 1). This observation suggests the dNTP-stabilized misalignment mechanism is sequence-dependent.

The values of substrate specificity and fidelity for all 12 incorrect incorporations were calculated and listed in Table 2. The fidelity of Dpo4 with undamaged DNA is in the range of 10^{-3}−10^{-4}, similar to what was estimated previously by steady-state kinetic analysis (*6*).

DISCUSSION

Although *S. solfataricus* Dpo4 is an archaeal polymerase, its structural and functional features are most likely conserved throughout the Y-family, and thus, it serves as a model for

Table 3: Comparison of the Fidelity of Dpo4 with the Replicative, Repair, and Lesion-Bypass DNA Polymerases

polymerase	fidelity $[(k_p/K_d)_{incorrect}/(k_p/K_d)_{correct}]$	K_d difference $[(K_d)_{incorrect}/(K_d)_{correct}]$	k_p difference $[(k_p)_{correct}/(k_p)_{incorrect}]$
Dpo4[a]	1.5×10^{-4} to 3.2×10^{-3}	1.1−18.1	2.4×10^2 to 1.7×10^3
yPolη[b]	1.24×10^{-3}	5.4	1.5×10^2
rPolβ[c]	6.3×10^{-5} to 2.0×10^{-4}	4.8−43.4	83.3 to 2.9×10^3
E. coli Pol I[d]	1.7×10^{-4} to 5.3×10^{-4}	1.7−4.2	5.0×10^3 to 2.4×10^4
hPolγ[e]	1.8×10^{-6} to 2.9×10^{-4}	72−454	38.5 to 3.4×10^3
T7 DNA Pol[f]	2.6×10^{-7} to 6.7×10^{-6}	200−400	2.0×10^3 to 4.0×10^3

[a] At 37 °C (this work). [b] At 37 °C (*25*); the numbers are derived from the incorporations of correct and incorrect nucleotides, not from the 16 possible incorporations. [c] At 37 °C (*33*). [d] At 22 °C, not including the fidelity contribution from the 3′−5′ and 5′−3′ exonucleases (*35*). [e] At 37 °C, not including the fidelity contribution from the 3′−5′ exonuclease activity (*34*). [f] At 20 °C, not including the fidelity contribution from the 3′−5′ exonuclease activity (*28*).

assessing both the lesion bypass properties and reduced replication fidelity inherent to the family (*17*). Using the approach of single-nucleotide incorporation under single-turnover conditions, we measured the substrate specificity of incoming nucleotides and calculated the fidelity of Dpo4 on undamaged DNA substrates. The fidelity was determined to be in the range of 10^{-3}−10^{-4} (Table 2), similar to what was estimated previously by steady-state kinetic analysis (*6*, *7*) and a forward mutation assay (*7*). The fidelity of a polymerase, defined in eq 3, is inversely proportional to the K_d difference and to the k_p difference between the incorporations of a correct versus an incorrect nucleotide.

$$\text{fidelity} = (k_p/K_d)_{incorrect}/(k_p/K_d)_{correct} = [(K_d)_{incorrect}/(K_d)_{correct}]^{-1}[(k_p)_{correct}/(k_p)_{incorrect}]^{-1} = (K_d \text{ difference})^{-1}(k_p \text{ difference})^{-1} \quad (3)$$

Both the ground-state binding affinity K_d and the maximum incorporation rate k_p influence the overall fidelity. The kinetic and thermodynamic basis of the fidelity of Dpo4 is discussed below.

Contribution of the Ground-State Binding Affinity of an Incoming Nucleotide to the Fidelity of Dpo4. The ground-state binding affinity of mismatched nucleotides (K_d) for Dpo4 is, on average, 9.6-fold (1.1−18.1-fold) lower than those of matched nucleotides (Tables 2 and 3), corresponding to a $\Delta\Delta G$ ($= -RT \ln K$) equal to 1.6 kcal/mol (*27*). The average K_d difference (Table 3) and the corresponding $\Delta\Delta G$ value are lower than the average values reported for replicative polymerases such as T7 DNA polymerase (300-fold and 3.9 kcal/mol, respectively, at 20 °C) (*27*) and human DNA polymerase γ (hPolγ) (263-fold and 4.0 kcal/mol, respectively) (*34*), but similar to the values for another Y-family member, yPolη (5.4-fold and 1.0 kcal/mol, respectively) (*25*), and repair enzymes, including rabbit DNA polymerase β (rPolβ) (24.1-fold and 2.3 kcal/mol, respectively) (*33*) and *E. coli* DNA polymerase I (Klenow) (3.0-fold and 0.75 kcal/mol, respectively) (*35*). Previous analysis of DNA melting experiments found that the free energy differences between primer terminal correct and incorrect base pairs are less than 1.0 kcal/mol at 37 °C. This suggests the selection of a matched nucleotide by the two Y-family polymerases, Dpo4 and yPolη, and the two repair DNA polymerases, rPolβ and Klenow, in the ground-state is determined primarily by its intrinsic ability to base pair with the template base, with little contribution from the respective loose polymerase active site as illustrated in the crystal structures (*17*, *36*). In contrast, the replicative DNA polymerases significantly stabilize the binding of correct nucleotide through hydrogen bonding and van der Waals interactions in relatively tight enzyme active sites. In these polymerases, the difference in ground-state binding affinity of correct and incorrect nucleotides contributes ~100-fold selection to the fidelity (Table 3). In addition, Dpo4 (70−230 μM) binds correct nucleotides with an affinity at least 10-fold lower than those of replicative polymerases. For example, the K_d values for correct nucleotide incorporation for hPolγ (*34*) and T7 DNA polymerase (*27*) are 0.8 and 18 μM, respectively. The low ground-state binding affinity observed for matched nucleotides with Dpo4 further suggests a minimal interaction between Dpo4 and the nascent base pair.

Scheme 2

$$E{\bullet}DNA_n + dNTP \overset{1}{\rightleftharpoons} E{\bullet}DNA_n{\bullet}dNTP \overset{2}{\rightleftharpoons} E'{\bullet}DNA_n{\bullet}dNTP \overset{3}{\rightleftharpoons} E'{\bullet}DNA_{n+1}{\bullet}PPi$$

"open" "closed"

The difference in the ground-state binding affinity and the selection of nucleotides at the initial binding step between the low-fidelity and high-fidelity polymerases can be explained by the structural differences at the respective polymerase active sites based on the X-ray crystal structures of Dpo4 (*17*) and other DNA polymerases (*37*−*41*). The ternary structure of Dpo4, DNA, and a matched nucleotide in the "closed" state (E′·DNA·dNTP, Scheme 2) reveals the enzyme active site is relatively open and unusually accessible to solvent in comparison to the active sites of replicative polymerases (*17*). The weak interaction between Dpo4 and the nascent base pair is expected to be further exacerbated in the "open" state (E·DNA·dNTP, Scheme 2) since the finger and little finger domains of Dpo4 have not undergone the conformational change to form the closed state which necessarily increases the processivity and confers relaxed interactions between Dpo4 and the replicating base pair. The energetically favorable entropic contribution from the exclusion of water molecules around the nascent base pair at the Dpo4 active site is diminished due to its greater solvent accessibility (*42*). Additionally, the Dpo4 amino acid residues (G41, A42, A44, A57, and G58) in direct contact with the replicating base pair are all small (*17*), unlike the residues found in high-fidelity polymerases, which often involve the aromatic ring of Tyr or Phe, or the guanidinium group of Arg, in formation of planar stacking interactions with the incoming base (*37*−*41*). Except for one water-mediated hydrogen bond, the nascent base pair is unrestrained in either

minor or major groove and has no specific polar interactions with Dpo4 (*17*).

Interestingly, yPolη binds the correct nucleotide ($K_d = 2.4$ μM) (*25*) with much higher affinity than Dpo4. yPolη, a Y-family member like Dpo4, is expected to possess a loose active site. However, recently, two structures of Dpo4 in a ternary complex with a DNA substrate containing a *cis-syn* cyclobutane pyrimidine dimer (CPD) and an incoming nucleotide were crystallized (*43*). From these structures and the available structure of yPolη alone (*36*), a theoretical model of a Polη−CPD complex was generated. Structure-based sequence alignment suggests that several amino acid residues in yPolη make its active site less solvent accessible than that of Dpo4 (*43*). For example, the larger side chains of Gln55 and Ile60 residues in yPolη purportedly interact with the replicating base pair in the minor groove, while the corresponding residues in Dpo4, Val32 and Ala44, are less capable of such an interaction due to their relatively small side chains. Additionally, Arg73 of yPolη, corresponding to Ala57 of Dpo4, has been purported to help orient the incoming nucleotide through charge interactions conferred via the side chain, while the aliphatic portion of this residue shields the major groove of the replicating base pair and the triphosphate moiety of the incoming nucleotide (*43*). Overall, these and other subtle structural differences between the Dpo4−CPD complex and the theoretical model of the yPolη−CPD complex may produce significantly different active site environments whose characteristics most likely affect nucleotide binding affinities.

Finally, although the reaction temperature in our measurements (37 °C) was much lower than the physiological temperature (80 °C) of *S. solfataricus*, the temperature difference should not affect the equilibrium dissociation constant K_d ($=k_{off}/k_{on}$) since both association (k_{on}) and dissociation (k_{off}) will be equally faster at higher temperatures. Interestingly, the ground-state binding affinity of nucleotides matched to *S. solfataricus* Dbh (0.2−1.0 mM) (*16*) and *Thermus aquaticus* DNA polymerase (35−57 μM) (*44*) is also low compared to those of other replicative DNA polymerases. More studies are required to investigate the low affinity of matched nucleotides for these thermostable DNA polymerases.

Contribution of the Rate of Nucleotide Incorporation to the Fidelity of Dpo4. With Dpo4, the rates of incorporation of incorrect nucleotides [$(k_p)_{incorrect}$] are approximately 2−3 orders of magnitude slower than the rates of incorporation of correct nucleotides [$(k_p)_{correct}$] (Table 2). The contribution to the overall replication fidelity from the k_p difference [$(k_p)_{correct}/(k_p)_{incorrect}$] is ∼100-fold greater than the contribution from the K_d difference [$(K_d)_{incorrect}/(K_d)_{correct}$] (Table 3). The data obtained from similar studies with yPolη reveal a similar predominant contribution to fidelity from the k_p difference (Table 3) (*25*). Likewise, the only other DNA polymerases to be extensively studied using pre-steady-state kinetics (Table 3) show the same 10−100 fold greater contribution to fidelity from the k_p difference. These observations suggest that the steps after the initial binding of nucleotide to form the E·DNA·dNTP complex (Scheme 2) provide a more stringent selection for the incorporation of correct versus incorrect nucleotides at the polymerase active site than the initial binding. One of the following two steps has been postulated to limit the incorporation of a nucleotide: an enzyme conformational change (step 2) or the chemistry step (step 3) (Scheme 2). The former deals with the previously proposed "induced-fit" model for replicative DNA polymerases (*27*, *31*). In this model, the binding of a correct nucleotide will induce a rate-limiting protein conformational change to form a closed conformation followed by a fast chemical incorporation step. The conformational change from the open to the closed state serves as a kinetic barrier to improving selectivity. This step is largely a function of Watson−Crick base pairing geometry required by the polymerase active site (*45*). The latter model is based on the results gleaned from studies of rPolβ (*46*), which show that the protein conformational change is actually faster than the slow chemistry step. In the following paper (*26*), our kinetic data for Dpo4 suggest that step 2 in Scheme 2 is rate-limiting for the incorporation of a correct nucleotide while step 3 limits the incorporation of an incorrect nucleotide. More discussion about the contribution of the protein conformational change to the overall fidelity of Dpo4 can be found in the following paper (*26*).

Incidentally, the k_p for incorporation of a correct nucleotide by Dpo4 (9.4−16.1 s^{-1}) (Table 2) is much higher than the rate reported for Dbh (0.06−2.3 min^{-1}), a Dpo4 homologue from a different strain of *S. solfataricus*, measured by Potapova *et al.* (*16*) using single-nucleotide incorporation under "single-turnover" conditions at 22 °C. This discrepancy cannot be logically reconciled by the different reaction temperatures. Most likely, the slow incorporation rates observed with Dbh are due to the fact that this group did not preincubate the enzyme with DNA prior to mixing the enzyme with dNTP to initiate reaction time courses (*16*). The order in which components are mixed to start nucleotide incorporation reactions significantly affects the pre-steady-state kinetic results (*30*). Moreover, the Dbh reaction buffer (pH 8.5 with 10 mM $MgCl_2$ and no additional salt) was different from buffer R (pH 7.5 with 5 mM $MgCl_2$ and 50 mM NaCl) used in our studies.

Differential Factor between Low- and High-Fidelity Polymerases. Surprisingly, all four low-fidelity polymerases, including the two Y-family members Dpo4 and yPolη and the two DNA repair polymerases Klenow and rPolβ, when compared to the two high-fidelity replicative polymerases T7 DNA polymerase and hPolγ, have a 10−100-fold lower K_d selection factor, while the values of the k_p difference are similar (Table 3). Although the overall contribution to the fidelity from the k_p difference is more substantial, this suggests that the discrimination at the initial nucleotide binding step determines whether a DNA polymerase is either a high- or low-fidelity polymerase (Table 3). The selection provided by nucleotide ground-state binding affinity is dictated by the tightness conferred by the respective polymerase active site. Similar proclivities in selection during the subsequent nucleotide incorporation steps after the initial binding step are justifiable since all polymerases use a catalytic mechanism based on two metal ions, and since all polymerases have structurally similar palm domains containing the three conserved active site carboxylates (one Asp and two Glu residues) (*47*). However, more DNA polymerases need to be kinetically examined to determine whether these conclusions can be globally applied.

Scheme 3

"dNTP"-Stabilized Misalignment Model (Scheme 3). The ground-state binding affinity of the mismatched dCTP for D-7 (C·T) and D-8 (C·C) is very similar to that of a correct nucleotide, although the k_p differs by 619- and 276-fold (Table 2), respectively. Interestingly, the downstream template base, in both cases, is a guanine. These data suggest the mismatched dCTP somehow base pairs with the downstream guanine to form the canonical C·G base pair by "skipping" the first available template base from the primer 3′-terminus. This type of base pairing necessarily increases the ground-state binding affinity of mismatched dCTP, leading to a subsequent decrease in the polymerase fidelity ($1.1-3.2 \times 10^{-3}$) at these specific template sequences. Interestingly, such a mechanism of nucleotide incorporation is not plausible for replicative DNA polymerases, due to their inherent inability to accommodate two template bases at the active site simultaneously. To provide further evidence for this mechanism, we designed a substrate (D-12) that mimicked the D-8 substrate except that the downstream guanine in D-8 was replaced with an adenine. The affinity of mismatched dCTP (1228 ± 266 μM) for the complex of D-12 and Dpo4 indeed drops to the level of those of other mismatched base pairs (Table 2). This result not only provided additional evidence for the dNTP-stabilized misalignment mechanism shown in Scheme 3 but also showed this type of dCTP incorporation dominated direct misincorporations (C·T and C·C). Once dCTP was incorporated, the next correct nucleotide incorporated into D-7 or D-8 could be dGTP, leading to a −1 frameshift, or dCTP, resulting in substitution after realignment of the primer terminus (*7*). These pathways are currently being studied by our group to determine which is more favored. Interestingly, a forward mutation assay has found that Dpo4 frequently generates deletions and substitutions of even noniterated nucleotides, especially cytosines flanked by a 5′-template guanine (*7*). These results are consistent with our pre-steady-state kinetic results. The dNTP-stabilized misalignment mechanism is also consistent with the ternary structure of Dpo4, DNA, and a mismatched ddGTP, showing Dpo4 simultaneously accommodates two template bases at its active site and the incoming ddGTP skips the first available template base and base pairs with the template base cytosine (*17*). In addition, when *E. coli* DNA polymerase IV catalyzes incorporation of a single nucleotide into a template 5′-CAp sequence where the next available template base is 2-aminopurine (Ap), the incorporation of dGTP increases the fluorescence intensity of 2-aminopurine due to G·C base pairing which forces 2-aminopurine to become extrahelical or "looped-out", while the incorporation of dTTP quenches the 2-aminopurine fluorescence signal due to the canonical T·Ap base pairing (*14*). The results from these 2-aminopurine fluorescence intensity change experiments support the unusual base pairing in the dNTP-stabilized misalignment mechanism.

Intriguingly, the rates of incorrect incorporation (k_p) of dCTP into D-7 and D-8 (Table 2) were not affected by the dNTP-stabilized misalignment. These kinetic results are consistent with the mismatched ternary structure of Dpo4 (*17*). In this ternary structure, the base at the 3′-end of the primer is tilted up, while the incoming guanine moiety is tilted down, thereby orienting the primer 3′-hydroxyl group and the α-phosphate of the incoming nucleotide less than 4 Å apart. Additionally, this crystal structure reveals the sugar−phosphate bond of the incoming ddGTP is mobile. The mobility observed for the incoming nucleotide coupled with the observed unconventional base pairing arrangement may justify the slow incorporation rate observed for dCTP incorporation via the dNTP-stabilized misalignment mechanism.

Additionally, the dNTP-stabilized misalignment mechanism appears to be sequence-dependent. The binding affinity of mismatched dCTP for a template adenine flanked by a 5′-template guanine (D-1) was not increased. We speculate that we did not observe an increase in affinity in this specific case was due to the difficulty of the bulky base adenine, rather than a less bulky pyrimidine base as in D-7 and D-8, in looping out at the enzyme active site, therefore allowing the incoming dCTP to base pair with the downstream template guanine. Interestingly, 50% of the one- or two-base deletion errors made by hPolκ (*9*) and *E. coli* DNA polymerase IV (*15*) are deletions of a template pyrimidine (Py) in a 5′-GPy sequence. Dbh, a homologue of Dpo4, has also shown a remarkable preference for skipping a pyrimidine rather than a purine, when the skipped nucleotide is positioned immediately 3′ to a template G residue, generating a single-base deletion (*16*). Therefore, the dNTP-stabilized misalignment mechanism could apply to all Y-family polymerases when they replicate through 5′-GPy "hotspots". It will be interesting to see if these Y-family polymerases also use the dNTP-stabilized misalignment mechanism to incorporate dGTP into the 5′-CPy sequences. 2-Aminopurine fluorescence intensity studies with *E. coli* polymerase IV (*14*) and the mismatched Dpo4 ternary crystal structure (*17*) have shown this is possible. We are currently studying the incorporation of dGTP into 5′-CPy-containing sequences.

The high frameshift rates at 5′-GPy, and possibly other sequence contexts, could be detrimental to genome stability especially for higher organisms, although, alternatively, substitutions resulting from realignment of the DNA substrate after incorporation via the dNTP-stabilized misalignment mechanism could arguably promote genetic diversity. Two factors could decrease such a high frameshift frequency *in vivo*. The first, previously mentioned, is the realignment of the primer 3′-terminus after nucleotide incorporation following the dNTP-stabilized misalignment mechanism. This will eliminate frameshift mutations and increase substitution rates. The second factor is the inaccessibility of these frameshift hotspots to Y-family polymerases because of their low processivity coupled with the tight association of DNA replicative polymerases with DNA.

ACKNOWLEDGMENT

We thank Dr. Roger Woodgate at the National Institutes of Health (Bethesda, MD) for providing us with the initial clone of Dpo4.

REFERENCES

1. Jackson, S. P. (1998) *Biochem. Soc. Trans. 27*, 1−13.
2. Lehmann, A. R. (2002) *Mutat. Res. 509*, 23−34.

3. Goodman, M. F. (2002) *Annu. Rev. Biochem. 71*, 17−50.
4. Boudsocq, F., Ling, H., Yang, W., and Woodgate, R. (2002) *DNA Repair 1*, 343−358.
5. She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W., and Van der Oost, J. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 7835−7840.
6. Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001) *Nucleic Acids Res. 29*, 4607−4616.
7. Kokoska, R. J., Bebenek, K., Boudsocq, F., Woodgate, R., and Kunkel, T. A. (2002) *J. Biol. Chem. 277*, 19633−19638.
8. Silvian, L. F., Toth, E. A., Pham, P., Goodman, M. F., and Ellenberger, T. (2001) *Nat. Struct. Biol. 8*, 984−989.
9. Ohashi, E., Bebenek, K., Matsuda, T., Feaver, W. J., Gerlach, V. L., Friedberg, E. C., Ohmori, H., and Kunkel, T. A. (2000) *J Biol Chem 275*, 39678−39684.
10. Matsuda, T., Bebenek, K., Masutani, C., Rogozin, I. B., Hanaoka, F., and Kunkel, T. A. (2001) *J. Mol. Biol. 312*, 335−346.
11. Thomas, D. C., Roberts, J. D., Sabatino, R. D., Myers, T. W., Tan, C. K., Downey, K. M., So, A. G., Bambara, R. A., and Kunkel, T. A. (1991) *Biochemistry 30*, 11751−11759.
12. Osheroff, W. P., Beard, W. A., Wilson, S. H., and Kunkel, T. A. (1999) *J. Biol. Chem. 274*, 20749−20752.
13. Longley, M. J., Nguyen, D., Kunkel, T. A., and Copeland, W. C. (2001) *J. Biol. Chem. 276*, 38555−38562.
14. Kobayashi, S., Valentine, M. R., Pham, P., O'Donnell, M., and Goodman, M. F. (2002) *J. Biol. Chem. 277*, 34198−34207.
15. Wagner, J., and Nohmi, T. (2000) *J. Bacteriol. 182*, 4587−4595.
16. Potapova, O., Grindley, N. D., and Joyce, C. M. (2002) *J. Biol. Chem. 277*, 28157−28166.
17. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) *Cell 107*, 91−102.
18. Washington, M. T., Johnson, R. E., Prakash, S., and Prakash, L. (1999) *J. Biol. Chem. 274*, 36835−36838.
19. Johnson, R. E., Washington, M. T., Prakash, S., and Prakash, L. (2000) *J. Biol. Chem. 275*, 7447−7450.
20. Zhang, Y., Yuan, F., Xin, H., Wu, X., Rajpal, D. K., Yang, D., and Wang, Z. (2000) *Nucleic Acids Res. 28*, 4147−4156.
21. Zhang, Y., Yuan, F., Wu, X., and Wang, Z. (2000) *Mol. Cell. Biol. 20*, 7099−7108.
22. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, L. (2000) *Nature 406*, 1015−1019.
23. Tissier, A., McDonald, J. P., Frank, E. G., and Woodgate, R. (2000) *Genes Dev. 14*, 1642−1650.
24. Johnson, K. A. (1992) *Enzymes 20*, 1−61.
25. Washington, M. T., Prakash, L., and Prakash, S. (2001) *Cell 107*, 917−927.
26. Fiala, K. A., and Suo, Z. (2004) *Biochemistry 43*, 2116−2125.
27. Patel, S. S., Wong, I., and Johnson, K. A. (1991) *Biochemistry 30*, 511−525.
28. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry 30*, 526−537.
29. Donlin, M. J., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry 30*, 538−546.
30. Bryant, F. R., Johnson, K. A., and Benkovic, S. J. (1983) *Biochemistry 22*, 3537−3546.
31. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) *J. Biol. Chem. 267*, 25988−25997.
32. Johnson, A. A., and Johnson, K. A. (2001) *J. Biol. Chem. 276*, 38097−38107.
33. Ahn, J., Werneburg, B. G., and Tsai, M. D. (1997) *Biochemistry 36*, 1100−1107.
34. Johnson, A. A., and Johnson, K. A. (2001) *J. Biol. Chem. 276*, 38090−38096.
35. Kuchta, R. D., Benkovic, P., and Benkovic, S. J. (1988) *Biochemistry 27*, 6716−6725.
36. Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001) *Mol. Cell 8*, 417−426.
37. Doublie, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature 391*, 251−258.
38. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) *Science 282*, 1669−1675.
39. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) *Biochemistry 36*, 11205−11215.
40. Eom, S. H., Wang, J., and Steitz, T. A. (1996) *Nature 382*, 278−281.
41. Li, Y., Korolev, S., and Waksman, G. (1998) *EMBO J. 17*, 7514−7525.
42. Petruska, J., Sowers, L. C., and Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1559−1562.
43. Ling, H., Boudsocq, F., Plosky, B. S., Woodgate, R., and Yang, W. (2003) *Nature 424*, 1083−1087.
44. Brandis, J. W., Edwards, S. G., and Johnson, K. A. (1996) *Biochemistry 35*, 2189−2200.
45. Yang, W. (2003) *Curr. Opin. Struct. Biol. 13*, 23−30.
46. Showalter, A. K., and Tsai, M. D. (2002) *Biochemistry 41*, 10571−10576.
47. Steitz, T. A. (1999) *J. Biol. Chem. 274*, 17395−17398.

BI0357457