Published in final edited form as:

Biochemistry. 2011 November 22; 50(46): 10136–10149. doi:10.1021/bi2014618.

Structure of the 2-Aminopurine-Cytosine Base Pair Formed in the Polymerase Active Site of the RB69 Y567A-DNA Polymerase†

Linda J. Reha-Krantz^{*,§}, Chithra Hariharan[§], Usharani Subuddhi[§], Shuangluo Xia[¶], Chao Zhao[¶], Jeff Beckman[¶], Thomas Christian[¶], and William Konigsberg^{*,¶}

Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven Connecticut 06520-8024

Abstract

The adenine base analog 2-aminopurine (2AP) is a potent base substitution mutagen in prokaryotes because of increased ability to form a mutagenic base pair with an incoming dCTP. Despite more than 50 years of research, the structure of the 2AP-C base pair remains unclear. We report the structure of the 2AP-dCTP base pair formed within the polymerase active site of the RB69 Y567A-DNA polymerase. A modified wobble 2AP-C base pair was detected with one H-bond between the N1 of 2AP and a proton from the C4 amino group of cytosine and an apparent bifurcated H-bond between a proton on the 2-amino group of 2-aminopurine and the ring N3 and O2 of cytosine. Interestingly, a primer-terminal region rich in AT base pairs, compared to GC base pairs, facilitated dCTP binding opposite template 2AP. We propose that increased flexibility of the nucleotide binding pocket formed in the Y567A-DNA polymerase and increased ‘breathing’ at the primer-terminal junction of A+T-rich DNA facilitates dCTP binding opposite template 2AP. Thus, interactions between DNA polymerase residues with a dynamic primer-terminal junction play a role in determining base selectivity within the polymerase active site of RB69 DNA polymerase.

Keywords

DNA polymerase-DNA interactions; 2-aminopurine; nucleotide binding; 2-aminopurine-C mutagenic base pair; 2-aminopurine fluorescence; DNA replication fidelity; base selectivity

DNA polymerases incorporate nucleotides with high accuracy despite the challenge of having to correctly assess four substrates: dCTP binding opposite template G, dGTP binding opposite template C, dTTP binding opposite template A, and dATP binding opposite template T. DNA polymerases achieve high substrate specificity by evaluating the shape of the newly forming base pair, by utilizing H-bonds that form between the base in the template strand and the base of the incoming nucleotide, by interacting with phosphate groups, and by probing chemical features of the bases and sugars (1-7). Even if an incorrect nucleotide is bound in the polymerase active site, nucleotide incorporation will occur only if

[†]Supported by grants from the Canadian Institutes of Health Research (L.J.R-K.) and by U.S. Public Health Service Grant ROI-GM063276-09 (W.H.K.).

^{*}To whom correspondence should be addressed. L.J.R-K.: Telephone: (780) 492-5383. Fax: (780) 494-9234. linda.reha-krantz@ualberta.ca. W.H.K. william.konigsberg@yale.edu. .

[§]University of Alberta.

[¶]Yale University.

[‡]With regard to author contributions, L.J.R-K., C.H., and U.S. performed fluorescence experiments and interpreted data, L.J.R-K. wrote the paper, S.X. and C.Z. performed structural and kinetic studies, J.B. and T.C. performed kinetic experiments, and W.H.K. interpreted data and contributed to writing the paper.

the reactants are in the correct position for chemistry to occur. In the unlikely event that an incorrect nucleotide is incorporated, the limited ability of the DNA polymerase to extend a mismatched primer-terminus leads to exonucleolytic proofreading.

Base analogue mutagens partially escape the above surveillance steps. 2-aminopurine (2AP)¹, for example, templates the incorporation of dCMP more frequently than an adenine and, as a consequence, is a potent base substitution mutagen in bacteria and bacteriophage T4 infections (8, 9). 2AP base pairs with T with two H-bonds, but differs from the A – T base pair by using the O2 instead of the O4 of thymine, which allows minor groove interactions to influence base pair formation (Figure 1). Even though the 2AP-T base pair is not mutagenic, the bacteriophage T4 DNA polymerase discriminates in forming and extending the 2AP-T base pair (10-13), but much stronger discrimination is observed for the mutagenic 2AP-C base pair.

Several structures have been proposed for the 2AP-C base pair, including rare tautomers, protonated base pairs, and a neutral wobble structure, but the structure of the mutagenic base pair is still unclear (Figure 1; 1, 14-16). These studies employed UV absorbance, fluorescence, NMR and quantum chemical characterization of the 2AP-C base pair, but in the absence of the DNA polymerase and, thus, in the protein environment in which the 2AP-C base pair is formed *in vivo*. Because wild type DNA pols discriminate against forming the 2AP-C base pair, we have used a mutant DNA polymerase, the bacteriophage RB69 Y567A-DNA polymerase, to increase formation of the 2AP-dCTP base pair within the polymerase active site.

Bacteriophage RB69 DNA polymerase is a family B DNA polymerase. A residue analogous to Y567 is conserved in the polymerase active site of family B DNA polymerases including the closely related bacteriophage T4 DNA polymerase (17) and in the polymerase active sites of family A and X DNA polymerases. A strong mutator phenotype is observed for the RB69 Y567A-DNA polymerase (18) and a cysteine substitution for the analogous tyrosine Y955 in the human mitochondrial DNA polymerase, a family A DNA polymerase, is linked to parkinsonism and ophthalmoplegia (19). Structural and kinetic studies of the RB69 Y567A-DNA polymerase demonstrate that the Y567A substitution increases the flexibility of the nucleotide binding pocket, which in turn increases the ability of the mutant DNA polymerase to incorporate dAMP opposite 7,8-dihydro-8-oxoguanine (8-oxoG) (20) and dTMP opposite template G (21). Besides increased misincorporation, the Y567A-DNA polymerase also has an increased ability to extend mismatches (21), which is necessary for a mismatch to escape DNA polymerase proofreading (22). From these observations we reasoned that the Y567A-DNA polymerase would also increase formation of the mutagenic 2AP-dCTP base pair, which we tested here by using the intrinsic fluorescence of 2AP and by X-ray crystallography.

2AP as the free base, nucleoside or nucleotide is highly fluorescent, but 2AP fluorescence in DNA is quenched by base stacking interactions (23); however, protein-induced base unstacking can produce large increases in 2AP fluorescence intensity as observed for complexes formed with the T4 and RB69 DNA polymerases with DNA substrates in which 2AP is positioned in the +1 position of the template strand (24, 25). 2AP-containing DNA substrates are described in Figure 2. The 2AP emission spectrum for complexes formed with the exonuclease-deficient D222A/D327A-RB69 DNA polymerase (RB69 *exo*⁻) is shown in Figure 3A. With excitation at 315 nm, a broad peak of fluorescence emission is detected at ~365 nm that is 15-fold higher than the fluorescence intensity observed for unbound DNA.

¹Abbreviations: 2AP, 2-aminopurine; dd, 2',3'-dideoxynucleotide; dFTP, difluorotoluene deoxynucleoside triphosphate; *exo*, exonuclease; *K*_{d,app}, *K*_d apparent; pol, polymerase

At least 3 to 4 fluorescence lifetimes for 2AP are observed in binary T4 (26) and RB69 DNA pol complexes (Figure 3B) that range from 0.05 to 8 ns or more. Although each lifetime represents a mean of a distribution, lifetimes can be correlated with distinct 2AP conformations (26-30). Because one conformation is characterized by a fluorescence lifetime that is similar to the ~10 ns lifetime observed for the free 2AP nucleoside in solution (27), the longest lifetime species is attributed to the fully unstacked state. Thus, the T4 and RB69 DNA polymerases induce conformational changes in the template strand that unstack 2AP in the +1 position (Figure 3B), which is illustrated as complex I in Figure 4. Highly fluorescent (long lifetime) species are also observed for T4 and RB69 DNA exonuclease complexes formed with DNA substrates in which 2AP is placed in the terminal position of the primer strand, which indicates that 2AP at the primer-end in exonuclease complexes is also unstacked (26, 30-32) as observed in the crystalline state (33-35).

Several experiments suggest that the highly fluorescent complex I is in rapid equilibrium with the less fluorescent complex II (Figure 4). For example, less fluorescence intensity is observed for complexes formed with the T4 DNA polymerase and DNA labeled with 2AP in the +1 position of the template strand if the primer-terminal region is G+C- compared to A+T-rich (26). As G+C-richness favors formation of polymerase over exonuclease complexes (11, 12), the lower level of fluorescence intensity observed indicates the presence of less fluorescent polymerase complexes compared to the highly fluorescent complex I.

Mg²⁺-dependent dTTP binding opposite template 2AP produces a rapid quench in 2AP fluorescence that is observed in reactions with chain-terminated DNA substrates (2',3'-dideoxynucleotide, dd) and as the initial rapid phase in reactions when the primer-terminus is extended by dTMP incorporation (13). The K_d for dTTP binding opposite template 2AP is ~31 μ M. We propose that the quenched ternary DNA pol-DNA-dTTP complex (complex IV) that is trapped with chain-terminated DNA substrates resembles a closed complex in which the fingers domain and residues in the polymerase active site form a tight-fitting nucleotide binding pocket as observed in structural studies of the RB69 DNA polymerase ternary complex (36). The rapid quench in 2AP fluorescence observed for dTTP binding to complex II, however, is likely due to formation of the less fluorescent open ternary complex (complex III), which drives the equilibrium from complex I to complex III.

We obtained evidence for complex III indirectly by using dFTP (F = difluorotoluene), a dTTP analogue that does not form H-bonds with a templating A, but is nevertheless incorporated by DNA polymerases with high fidelity (2). Addition of dFTP to T4 DNA pol binary complexes does not quench 2AP fluorescence at concentrations <150 μ M, but dFTP is bound because it is an effective competitive inhibitor of dTTP-Mg²⁺-induced quenching with a K_i of ~68 μ M (4). Thus, both dTTP and dFTP bind to complex II to form the ternary complex III, but only dTTP binding can trap quenched ternary complexes (complex IV). In contrast, the K_d for dCTP binding opposite template 2AP is at least 3 mM and dCTP is not an inhibitor of dTTP binding (13). Thus, dNTPs have ready access to complex II, but only the correct incoming dNTP is stably bound to form the open ternary complex (complex III). Subsequent conformational changes form the closed ternary complex (complex IV).

The K_d apparent ($K_{d,app}$) for dTMP incorporation opposite template 2AP to form complex V is >10-fold higher (367 μ M) than the K_d for dTTP binding and is also ~10-fold higher than the $K_{d,app}$ observed for nucleotide incorporation reactions without the base analog, which demonstrates that the T4 DNA polymerase discriminates in the incorporation of dTMP opposite template 2AP compared to incorporation opposite template A (13). The k_{pol} for incorporation of dTMP opposite 2AP, however, is the same as observed for standard nucleotide incorporation reactions, which indicates that the 2AP-T base pair formed is correctly positioned for optimal reactivity. In contrast, the observed k_{pol} for incorporation of

dCMP is very low, similar to the rate for enzyme dissociation. Release of pyrophosphate (PPi) produces complex VI, which resembles complex I except that the primer-end has been extended by a single nucleotide. Complex VI is a pivotal complex that is the starting point for another cycle of nucleotide incorporation or, if the incorporated nucleotide is not correct, the starting point for initiation of the exonucleolytic proofreading pathway and formation of the binary exo complex (complex VII) (22).

We examined the ability of the RB69 Y567A-DNA polymerase to form base pairs with 2AP in the template position with incoming dTTP or dCTP at each step of the nucleotide incorporation pathway as described in Figure 4. *In vivo*, the mutagenic 2AP-C base pair is most likely to occur when 2AP is in the template rather than as the incoming nucleotide (9). As expected, the Y567A substitution decreased the ability of the RB69 DNA polymerase to discriminate against dCTP binding opposite template 2AP; however, we were surprised by the large magnitude of the defect. In some reactions, ternary pol complexes (complex IV) were formed with the Y567A-DNA polymerase almost as easily with dCTP as with dTTP. The ability of the Y567A-DNA polymerase to readily form the 2AP-dCTP base pair provided the means to use X-ray crystallography to examine the mismatch within the polymerase active site.

EXPERIMENTAL PROCEDURES

DNA Polymerases

Expression, purification and characterization of wild type and mutant RB69 DNA polymerases have been described (37).

DNA Substrates

The DNA substrates are described in Figure 2. The oligonucleotides used for the longer DNA substrates (DNAs 1 and 2) were purchased from IDT and the oligonucleotides used for the shorter DNA substrates (DNAs 3 and 4) were synthesized by the W. M. Keck Foundation Biotechnology Resource Laboratory, Yale University, and PAGE purified. DNAs used to form exo complexes (DNAs 5-7) were described previously (26, 38). The longer DNA substrates were synthesized with biotin (B) at the 3'-end of the template strand to prevent DNA polymerase binding at the blunt end and, thus, to direct the DNA polymerase to the primer-terminus junction.

Fluorescence Intensity Experiments

Samples were excited at 315 nm to minimize excitation of tryptophan residues and fluorescence emission was monitored from 330 to 450 nm (26, 39). A 2 nm band-pass was used for both the excitation and emission monochromators. Solutions of complexes were formed with 200 nM 2AP-labeled DNA and a 2 to 2.5-fold excess of DNA polymerase in buffer containing 25 mM HEPES (pH 7.6), 50 mM NaCl, and 1 mM DTT. For dTTP and dCTP titration experiments, the primer-end was terminated with a dideoxy nucleotide and the reactions contained MgCl₂ (10 mM), which is essential for nucleotide binding (25). Equilibrium dissociation constants for nucleotide binding opposite template 2AP were determined as described (4). All fluorescence experiments were performed at 20 °C.

Fluorescence Lifetime Determinations

Solutions of complexes (1 μM 2AP-labeled DNA and 2 μM DNA polymerase) were formed in the above buffer with the addition of 0.5 mM EDTA. Solutions were excited at 315 nm using the frequency-doubled output from a pulsed-dye laser (PTI). Fluorescence emission was monitored at 368 nm with a 5 nm band-pass; a 320 nm long-pass filter was inserted

between the cuvette and emission monochromator. Decay curves and analyses procedures were described previously (26, 31). Experiments were performed at 20 °C.

Rapid Chemical Quench Assays

Single turnover experiments were performed under conditions in which the enzyme concentration was 10 times that of the 5'-³²P-labeled DNA substrate. Each reaction contained 66 mM Tris-HCl (pH 7.5) and was performed at 22 °C by mixing equal volumes of the 200 nM DNA substrate with 2 μM DNA polymerase, 20 mM MgCl₂ and varying concentrations of dNTPs. The final concentrations were 100 nM DNA substrate, 1 μM DNA polymerase, and 10 mM MgCl₂. The reactions were quenched in 0.5 M EDTA using the KinTek rapid quench instrument (Model RQF-3). Products were separated by electrophoresis in 20% acrylamide gels containing 8 M urea. The band intensities were determined using a Molecular Dynamics Phosphorimager and analyzed with Imagequant software. Data from the single turnover experiments were fit to a single exponential equation as described (21).

Crystallization of Exonuclease-deficient D222A/D327A-RB69 DNA Polymerase (Wild Type) and Exonuclease-deficient Y567A-DNA Polymerase Ternary Complexes with 2AP-dTTP and 2AP-dCTP Base Pairs

DNA polymerase (120 μM) was mixed in an equimolar ratio with freshly annealed DNA substrate. Nucleotide (dTTP or dCTP) was added to give a final concentration of 3 mM. Crystals of the ternary complexes were grown under oil in a microbatch procedure by mixing equal volumes of the ternary complex solution with a solution containing 100 mM sodium cacodylate buffer (pH 6.5), 125 mM CaCl₂, and 10% (wt/vol) polyethylene glycol 350 monomethyl ether (PEG350 MME). Cubic-shaped crystals were stabilized and cryoprotected by transfer to the stabilization solution containing 100 mM sodium cacodylate buffer (pH 6.5), 20% (wt/vol) PEG350 MME and 100 mM CaCl₂. Just prior to freezing in liquid nitrogen, the crystals were transformed into the same cryoprotectant solution but containing 30% PEG350 MME.

X-ray Diffraction Data Collection, Structure Determination, and Refinement

X-ray diffraction data were collected at a wavelength of 0.959 Å and at 110 K at NECAT, beamline 24ID-E (Advanced Photon Source, Argonne National Laboratory, Argonne, IL). The data were processed using the HKL2000 program suite. All crystals belonged to orthorhombic space group *P*2₁2₁2₁ with slightly different cell dimensions.

All six structures were solved by the automated molecular replacement method AMoRe (40) as implemented in CCP4, starting with the wild type RB69 DNA polymerase structure of the ternary complex with PDB entry 3NCI. The primer-template duplex and the incoming dTTP or dCTP were built into residual electron density maps, which were phased with the partially refined RB69 DNA polymerase structure using Coot (41). The structures were refined using Refmac5 (42). All figures were made using PYMOL (43). Coordinates and structure factors for the wild type RB69 DNA polymerase ternary complexes with the 2AP-dTTP base pair with the A+T- and G+C-rich DNA substrates have been deposited in the Protein Data Bank.

RESULTS

The Y567A Substitution Affects Partitioning of DNA between the Polymerase and Exonuclease Active Sites and the Equilibrium between Complexes I and II

Fluorescent binary complexes were formed with the RB69 DNA polymerase that lacks exonuclease activity (exo⁻) due to alanine substitutions for two essential aspartate residues in the exonuclease active site (D222 and D327) and a similarly exonuclease-deficient RB69

DNA polymerase that also has the Y567A substitution (Figure 3A). For simplicity, the RB69 and T4 exo^- DNA polymerases are referred to as wild type and the RB69 exo^- Y567A-DNA polymerase is referred to as the mutant or Y567A-DNA polymerase. Complexes were formed with 200 nM DNA that was labeled in the +1 position of the template strand with 2AP (DNA 1, Figure 2) and 500 nM enzyme. Significantly less fluorescence intensity was observed for the Y567A-DNA polymerase that did not increase with addition of more enzyme. The high fluorescence intensity of binary pol complexes is due to formation of a large population of fluorescent species with a ~ 10 ns lifetime that is produced by unstacking 2AP in the +1 template position in the template strand (complex I, Figure 4). Long lifetime fluorescent species in the 10 ns range were observed for both DNA polymerases (Figure 3B), but the amplitude was reduced from 38% for the wild type RB69 DNA polymerase to 20% for the mutant DNA polymerase, which was accompanied by an increase in the amplitude of a species with an intermediate lifetime ($\tau = 5.3$ ns, 42%). The reduced formation of the longest lifetime fluorescent species observed for the mutant DNA polymerase accounts for the observed reduction in 2AP fluorescence intensity.

DNA polymerases form binary pol complexes, complexes I and II, and binary exo^- complexes (complex VII). Complexes II and VII formed with DNA labeled with 2AP in the +1 position of the template strand are less fluorescent than the highly fluorescent complex I (25, 26). Thus, the reduced fluorescence intensity observed for complexes formed with the Y567A-DNA polymerase (Figure 3A) could be due to increased formation of less fluorescent binary pol complexes (complex II) or exo^- complexes (complex VII).

To determine if the Y567A substitution affects formation of binary exo^- complexes, complexes were formed with DNA substrates labeled with 2AP at the primer terminus (DNAs 5-7, Figure 2). Binary exo^- complexes labeled with 2AP at the primer terminus have high fluorescence intensity because the terminal 2AP is unstacked in the exonuclease active site (33, 34); a ~ 10 ns lifetime is observed for this species (30-32). The wild type T4 DNA polymerase preferentially forms fluorescent binary exo^- complexes with DNA substrates in which 2AP at the primer terminus is paired with template T (Figure 2, DNA 2) and, thus, recognizes the terminal 2AP-T base pair as a mismatch (10-12, 25, 26, 38). Although the D222A and D327A substitutions in the exonuclease active site reduce formation of exo^- complexes, a 9.7 ns fluorescence lifetime indicative of exo^- complexes is still detected, but the amplitude is reduced from 100 percent that is observed the wild type DNA polymerase without the amino acid substitutions to 17 percent for the exo^- DNA polymerase (32). Fluorescent species characterized by shorter fluorescence lifetimes indicate populations of complexes in which 2AP is imperfectly stacked with one or more adjacent bases as well as complexes in which 2AP is fully base-stacked. The addition of the Y567A substitution further reduces fluorescence intensity (Figure 5A). A likely explanation for the apparent reduced level of binary exo^- complexes is that the Y567A-DNA polymerase has increased ability to form less fluorescent binary pol complexes, which has been observed previously for mutant T4 DNA polymerases that have reduced ability to initiate the proofreading pathway (26, 30-32, 38).

We next tested if the Y567A-DNA polymerase recognizes the terminal 2AP-C base pair as a mismatch. Complexes were formed with DNA 6 (Figure 2) and again lower fluorescence intensity was detected with the Y567A-DNA polymerase (Figure 5B), which suggests that the terminal 2AP-C base pair is not always recognized as a mismatch by the Y567A-DNA polymerase. When this experiment was repeated with DNA 7, which has 3 preformed mismatches at the primer-end (Figure 2), similar fluorescence emission spectra were observed for both the exo^- wild type and Y567A-DNA polymerases (Figure 5C). Thus, the exo^- Y567A-DNA pol can form exo^- complexes as efficiently as the exo^- wild type DNA pol only if formation of pol complexes is prevented by three terminal mismatches. These

observations are consistent with the proposal that the Y567A substitution affects partitioning of DNA between the polymerase and exonuclease active sites in favor of increased formation of binary pol complexes under conditions in which the wild type enzyme favors formation of binary exo complexes.

Because the Y567A-DNA polymerase appears to favor formation of binary pol complexes more than the wild type DNA polymerase with DNAs with 2AP-T and 2AP-C terminal mismatches (Figure 5A, B), formation of binary pol complexes are also expected to be formed with DNA 1 in which the primer-end is matched and 2AP is in the +1 templating position (Figure 3). Thus, the lower level of fluorescence intensity observed for complexes formed with the Y567A-DNA polymerase compared to the wild type DNA polymerase and DNA 1 does not appear to be due to formation of binary exo complexes, but instead to increased formation of less fluorescent binary pol complexes (complex II). The increase in the amplitude of the intermediate fluorescence lifetime species ($\tau = 5.3$ ns) for the mutant DNA polymerase is consistent with this proposal (Figure 3B).

The RB69 Y567A-DNA Polymerase has Reduced Ability to Discriminate against Binding dCTP and Incorporating dCMP Opposite Template 2AP

Mg²⁺-dependent dTTP binding can be measured in reactions with chain-terminated DNA substrates (DNAs 1 and 2, Figure 2) by the decrease in 2AP fluorescence intensity as the concentration of dTTP is increased (4, 13, 25, 26). A mixture of highly fluorescent binary pol complexes (complex I), less fluorescent binary pol complexes (complex II), exo complexes, etc. are formed in the presence of 10 mM MgCl₂. Addition of incremental amounts of dTTP to the mixture quenches fluorescence intensity by formation of complex IV in which 2AP is fully base-stacked with a fluorescence lifetime of 0.6 ns (26). The K_d for dTTP binding can be determined from fluorescence quench curves (4). K_d values for dTTP binding by the RB69 exo⁻ DNA polymerase are similar to the values determined for the T4 exo⁻ DNA polymerase and are in the range of 30 μ M for DNA substrates that are relatively A+T- (DNA1, Figure 2) or G+C-rich (DNA2, Figure 2) in the primer terminal region (Table 1). Wild type T4 and RB69 DNA polymerases strongly discriminate against binding dCTP as unphysiologically high concentrations are needed to observed quench of 2AP fluorescence; the K_d is $\sim$ 3 mM (Table 1). Thus, the wild type T4 and RB69 DNA polymerases discriminate against binding dCTP by at least a factor of $\sim$ 100 compared to dTTP (30 μ M compared to 3 mM).

In contrast, the Y567A-DNA polymerase has significantly less ability to discriminate against binding dCTP opposite template 2AP (Table 1). With the A+T-rich DNA substrate, the K_d for dCTP binding at $\sim$ 25 μ M was $<$ 3-fold higher than the K_d for dTTP binding at $\sim$ 9 μ M, which indicates that dCTP is bound opposite template 2AP almost as easily as dTTP. Fluorescence quench curves are shown in Figure 6. Increased discrimination, however, was observed with the G+C-rich DNA; the K_d for dCTP binding at $\sim$ 317 μ M was $>$ 20-fold higher than the K_d for dTTP binding at $\sim$ 15 μ M (Table 1). Even though increased discrimination was observed with the G+C-rich DNA against binding dCTP-Mg²⁺, the K_d of $\sim$ 317 μ M is still well below the $>$ 3 mM value observed for the wild type T4 and RB69 DNA polymerases.

The K_d for dTTP binding by the wild type T4 exo⁻ DNA polymerase at $\sim$ 35 μ M is less than the K_d apparent ($K_{d,app}$) for the overall reaction for dTMP incorporation (367 μ M) (Table 1). This trend was observed for reactions with dTTP and the Y567A-DNA polymerase for the G +C-rich, but not for the A+T-rich DNA as the K_d for dTTP binding was $\sim$ 9 μ M compared to the $K_{d,app}$ of $\sim$ 16 μ M. The $K_{d,app}$ for dCMP incorporation by the Y567A-DNA polymerase was higher than the K_d s for dCTP binding with both DNA substrates, but highest for the G +C-rich DNA; the $K_{d,app}$ for incorporation of dCMP incorporation at 2 mM approached the

~3 mM values observed for the wild type T4 and RB69 DNA polymerases. The newly forming 2AP-dCTP base pair in the polymerase active site of the Y567A-DNA polymerase with A+T- and G+C-rich DNAs, however, appears not to be in optimal position for the chemistry step because k_{pol} values were >30-fold lower for the incorporation of dCMP (~6 s⁻¹) than for incorporation of dTMP (~200 s⁻¹) (Table 1).

The experiments in Table 1 were performed with relatively long DNA substrates, DNAs 1 and 2 (Figure 2). The experiments were repeated with the shorter DNA substrates (Figure 2, DNAs 3 and 4) that were used for structural studies (Table 2). Again, less discrimination was observed for the Y567A-DNA polymerase in binding dCTP opposite template 2AP with the A+T-rich DNA (~53 μM) compared to the G+C-rich DNA (~373 μM). $K_{\text{d,app}}$ values for incorporation of dCMP opposite 2AP were similar for the A+T- and G+C-rich DNAs, 298 and 361 μM respectively and less than the >2 mM values observed for the wild type DNA polymerase. As observed with the longer DNAs, k_{pol} rates were low and in the range of the enzyme dissociation rate.

The RB69 Y567A-DNA Polymerase Retains Ability to Discriminate against Binding rUTP opposite Template 2AP

Because the Y567A-DNA polymerase has reduced ability to discriminate against binding dCTP opposite template 2AP (Table 1), we wanted to know if relaxed nucleotide specificity extended to the sugar. Sugar specificity was tested in nucleotide binding reactions with rUTP and dUTP with template 2AP in DNA 1 (Figure 2). While a ~2-fold increase was observed for binding dUTP opposite template 2AP by wild type or Y567A-DNA polymerases (Table 3) compared to binding dTTP (Table 1), strong discrimination was detected for rUTP binding by wild type and Y567A-DNA polymerases. Thus, while the Y567A substitution reduces base selectivity, sugar specificity is largely retained which is expected if the sugar-gate residue – Y416 (5, 44), remains in position to interact with the deoxyribose of the incoming nucleotide in the nucleotide binding pocket formed with the Y567A-DNA polymerase.

Structural Studies of the 2AP-T and 2AP-C Base Pairs in the Nucleotide Binding Pocket of Wild Type and the Y567A-DNA Polymerases

In an attempt to provide a structural framework that would explain how dCTP is bound almost as readily as dTTP opposite template 2AP by the Y567A-DNA polymerase, we determined the crystal structures of three pairs of ternary complexes. One pair was with the wild type RB69 DNA polymerase, the A+T- or G+C-rich DNA substrates (DNAs 3 and 4, Figure 2), and dTTP. The second pair was like the first except with the Y567A-DNA polymerase. The third pair was like the second except that dCTP was used instead of dTTP. These six structures were obtained with resolutions ranging from 2.09 Å to 3.19 Å for the A+T-rich DNA and from 2.24 Å to 2.32 Å for the G+C-rich DNA (Figure 7). The free R factors for complexes with the A+T-rich DNA ranged from 24.3% to 30% and from 23.8% to 26.0% for the G+C-rich DNA (Table 4).

Two H-bonds were expected for the 2AP-T base pair; one between the ring nitrogens of the bases, N1 of 2AP and N3 of thymine, and the second between the N2 of 2AP and the O2 of thymine (Figure 1A). Interbase distances of ~ 2.7 Å, which are consistent with the presence of two H-bonds in these positions, were observed for the 2AP-T base pairs in ternary complexes formed with the wild type and Y567A-DNA polymerases with both the A+T- and G+C-rich DNA substrates (Figure 8, A-D). The bases are nearly coplanar.

Several H-bonded structures have been proposed for the 2AP-C base pair (Figure 1, B-D; 14-16). The long interbase distances for the 2AP-C base pair observed in ternary complexes

with the Y567A-DNA polymerase rule out the possibility of H-bonding between N1 of 2AP and N3 of cytosine (Figure 8, E and F) and, thus, 2AP-C base pairs involving protonated bases (Figure 1, C and D) are unlikely. Instead, apparent non-ideal H-bonds are observed. One possible H-bond is between a proton of the C4 amino group of cytosine and the N1 of 2AP as proposed for the wobble or neutral 2AP-C base pair (Figure 1B). In addition, bifurcated H-bonds are possible between a proton of the C2 amino group of 2AP and N3 of cytosine, as proposed for the neutral 2AP-C base pair, and between the O2 of cytosine.

A rigid H-bonding network involving the γ -OH groups of Y567, Y416, T622, four ordered water molecules, and the penultimate nucleotide at the primer-template junction was observed for ternary complexes formed with the wild type RB69 DNA polymerase, the G +C-rich DNA substrate, and dTTP (Figure 9A). The H-bonding network was disrupted by the Y567A substitution (Figure 9B), which created more flexibility in the nucleotide binding pocket. Residue G568 was shifted by 0.4 Å laterally toward Y416 in the Y567A-DNA polymerase ternary complex formed with dCTP compared to the ternary complex formed with the wild type DNA polymerase and dTTP (Figure 9C and D). In addition, the distance between the N3 of 2AP and the α C-H of G568 was increased to 3.9 Å by replacing Y567 with alanine compared to 3.5 Å observed for the wild type DNA polymerase. The disruption of the H-bonding network and the more flexible nucleotide binding pocket caused by the Y567A substitution is predicted to increase accommodation of the 2AP-C base pair.

DISCUSSION

2AP has a long history in providing insights into the fidelity of DNA replication that began with Freese's observations that 2AP is a potent base substitution mutagen (8), to Drake's observations that antimutator phage T4 DNA polymerases reduce 2AP mutagenesis while mutator DNA polymerases increase mutagenesis (45), and to observations by Bessman and others (10-12, 26, 38) that 2AP mutagenesis depends on the A+T- and G+C-richness of the primer-terminal region because A+T-richness in the duplex region of the primer-template junction increases proofreading. Here we provide new information on the ability of the RB69 DNA polymerase to form 2AP-dTTP and 2AP-dCTP base pairs within the polymerase active site, on the role of A+T- and G+C-richness of the primer-terminal region in formation of the mutagenic 2AP-dCTP base pair, and on how the Y567A substitution affects replication fidelity at multiple steps in the nucleotide incorporation pathway.

The Y567A-DNA polymerase forms fewer highly fluorescent binary pol complexes with 2AP placed in the +1 position of the template strand (complex I) than observed for the wild type DNA polymerase (Figure 3). The lower fluorescence intensity could be due to either increased formation of less fluorescent binary pol complexes (complex II, Figure 4) or to increased formation of binary exo complexes (complex VII, Figure 4), which are also only weakly fluorescent when formed with DNA substrates labeled with 2AP in the +1 position of the template strand. Increased formation of exo complexes was ruled out because the Y567A-DNA pol has reduced ability to form exo complexes compared to the wild type DNA polymerase (Figure 5A and B) unless there are three preformed mismatches at the primer-terminus (Figure 5C). Thus, the Y567A-DNA polymerase favors partitioning the primer-terminus to the polymerase rather than to the exonuclease active site, which is consistent with previous findings that the Y567A-DNA polymerase has increased ability to extend mismatches (21).

Why does the Y567A-DNA polymerase form more complex II? We propose that DNA polymerase interactions with the primer-terminal region affect a 3-way equilibrium between complex I/VI (pre-translocation) with complex II (post-translocation, nucleotide pre-insertion), which prepares for nucleotide incorporation; between complex I/VI with complex

VII (exonuclease complex), which prepares for proofreading; and with the possibility of enzyme dissociation. This proposal is similar to the model proposed for transcription in which the RNA polymerase is in a 3-way kinetic competition for elongation, editing, and termination (dissociation) (46). For the RB69 DNA polymerase, five ordered water molecules are observed to interact with the minor groove of the three terminal base pairs of the RB69 DNA polymerase (47). The water molecules serve as extensions of conserved amino acid residues as observed for the phi29 DNA polymerase (48). The Y567A substitution disrupted the H-bonding network and the space left by substitution of alanine for tyrosine was replaced by two water molecules (Figure 9A and B). Since translocation to form complex II requires transient release of minor groove interactions, the disrupted hydrogen bonding network and increased flexibility of the nucleotide binding pocket of the Y567A-DNA polymerase appear to favor translocation to form complex II. Preferred formation of complex II will shift the equilibrium away from formation of exonuclease complexes (complex VII), which is observed by the decreased ability of the mutant DNA polymerase to form exo complexes (Figure 5).

The increased flexibility in the nucleotide binding pocket produced by the Y567A substitution also affects an unusual nonpolar-polar interaction between the C α hydrogen of G568 and the N3 hydrogen acceptor of purines (47). Since minor groove interactions are proposed in general to be important for determining the accuracy of nucleotide binding (49), the increased distance between G568 and the N3 of 2AP from 3.5 Å in the wild type DNA polymerase to 3.9 Å in the mutant (Figure 9C and D) is expected to contribute to the observed reduction in base selectivity. Note, however, that reduced minor groove interactions produced by the Y567A substitution did not substantially reduce discrimination against ribose compared to deoxyribose nucleotides (Table 3). Structural studies show that Y416, which interacts with the deoxyribose of the incoming dNTP (44), maintains the sugar contact within the nucleotide binding pocket of the mutant as observed for the wild type DNA polymerase (Figures 7 and 9C, D).

While the increased flexibility of the nucleotide binding pocket of the Y567A-DNA polymerase is predicted to accommodate the 2AP-dCTP base pair more easily than the pocket formed by the wild type DNA polymerase, structural studies do not provide clear evidence why dCTP is bound more easily opposite template 2AP if the primer-terminal region is rich in AT compared to GC base pairs. The 2AP-C interbase distances may be closer in complexes formed with the A+T-rich DNA (Figures 7E and 8E) than with the G+C-rich DNA (Figures 7F and 8F), but the resolution of the ternary complexes formed with the A+T-rich DNA is not sufficient to draw this conclusion.

We propose that in addition to the increased flexibility in the nucleotide binding pocket provided by the Y567A substitution that additional flexibility is provided by a primer-terminal region rich in AT base pairs that further reduces base selectivity. This proposal may seem counterintuitive because G+C-richness in the primer-terminal region increases partitioning of the primer terminus from the exonuclease to the polymerase active site and reduces proofreading while A+T-richness does the opposite (10-12, 26); however, these studies examined the effects of DNA sequence on the fate of the pivotal complex I/VI in determining the winner of the 3-way competition between elongation (complex II), proofreading (complex VII) and dissociation. But the Y567A-DNA polymerase appears to favor formation of the post-translocated/pre-insertion complex II, as discussed above.

Complex II is in equilibrium with complex I/VI and complex III (Figure 4). G+C-richness in the duplex region of the primer-template junction also appears to increase formation of complex II by the wild type T4 DNA polymerase (26), but the Y567A-DNA polymerase favors formation of complex II even when the duplex region of the primer terminus is A+T-

rich (Figure 3). Once complex II is formed, the pre-insertion site appears to more readily accommodate dCTP to form complex III if the primer-terminal region is A+T- rather than G+C-rich (Table 1). This observation suggests that 'breathing' in the duplex region at the primer-template junction facilitates dCTP binding. Our proposal is supported by reports (for example 7, 23, 50, 51) that local conformational changes have been observed at the ends of duplex DNA and at the junction of single- and double-stranded DNA of primer-templates. More local strand separation is predicted for A+T- compared to G+C-rich DNAs, but we do not detect significant differences in the emission spectra or fluorescence lifetimes for the A+T- and G+C-rich DNA 1 and 2 substrates (Figure 2) in the absence of DNA polymerase (26). This could mean that 2AP is not in position to report on breathing or that increased breathing in the primer-terminal region may take place in the context of the polymerase active site. In order for increased breathing to facilitate dCTP binding opposite template 2AP to form complex III, either the reverse equilibrium with complex II is reduced or the forward equilibrium with complex IV is increased (Figure 4).

Higher fidelity is also reported for *E. coli* DNA pol II with G+C- compared to A+T-rich DNA substrates (52). This DNA sequence effect was explained by a lower activation energy for elongation with the A+T-rich DNA, which reduces the competition in forming exonuclease complexes. By analogy to the RB69 Y567A-DNA polymerase, *E. coli* DNA pol II may also not have tight, water molecule-mediated minor groove interactions with the primer-terminal region. In support of this proposal, the template strand is observed to be looped out in some *E. coli* DNA pol II structures (53), which would require relatively loose protein interactions that allow strand separation within the primer-terminal region.

We also observed an effect of DNA substrate length on the accuracy of nucleotide incorporation. The $K_{d,app}$ for dCMP incorporation by the Y567A-DNA polymerase increased from $\sim 170 \mu\text{M}$ with the long, A+T-rich DNA to $\sim 2 \text{ mM}$ with the long, G+C-rich DNA (Table 1), but similar values of $\sim 300\text{--}360 \mu\text{M}$ were observed with the shorter A+T- and G+C-rich DNAs (Table 2). These observations can also be understood in the context of the 3-way kinetic competition model if enzyme dissociation is more likely with the shorter DNA substrates. T4 DNA polymerase binding to primer-template DNA increases with the length of the single-stranded template overhang up to about five nucleotides (54). The short DNAs (DNAs 3 and 4, Figure 2) have only a four-nucleotide overhang, which means that the activation energy for dissociation is expected to be lower than for the long DNA substrates. If this is true then effects of DNA sequence on the accuracy of nucleotide incorporation can be detected with the long DNA substrates, but increased dissociation may mask these effects with the shorter DNAs.

For all reactions the k_{pol} rates were low, $< 9 \text{ s}^{-1}$ (Tables 1 and 2). The efficiency of misincorporation of dCMP opposite template 2AP, as measured by the k_{pol}/K_d incorporation ratio, was about $0.03 - 0.04 \mu\text{M}^{-1} \text{ s}^{-1}$ for the longer and shorter A+T-rich DNA substrates and about 10-fold lower for the longer and shorter G+C-rich DNA substrates ($0.003 - 0.007 \mu\text{M}^{-1} \text{ s}^{-1}$). Note that increased discrimination observed for G+C-rich DNAs was observed in parallel for the incorporation of dTMP and dCMP opposite template 2AP; thus, the dTMP/dCMP incorporation ratio is the same for A+T- and G+C-rich DNAs, ~ 300 -fold preference for dTMP for the longer DNA substrates (Table 1) and ~ 100 -fold preference with the shorter DNAs (Table 2).

Even though dCTP forms a stable base pair with 2AP that can be captured in quenched ternary complexes in solution (Figure 6B) and in crystal structures of ternary complexes formed with the Y567A-DNA polymerase (Figure 7), the 2AP-dCTP base pair does not have an optimal geometric shape. Non-ideal H bonds are observed between template 2AP and the incoming dCTP (Figure 8E and F) and dCTP is tilted about 5° towards the primer-

terminal base (ddA) (data not shown). Low rates for dCMP incorporation suggest that structure is also not optimal for chemistry in the transition state, which indicates that the chemistry step can serve as a final fidelity gate. Thus, stable dCTP binding opposite template 2AP in the ground state does not ensure that reactants will be optimally aligned for phosphodiester bond formation in the transition state.

This raises the question about the identity of the chemically reactive 2AP-dCTP base pair and the possibility of a minor conformation with increased reactivity (15, 16 and references cited therein). Our structural studies indicate the possibility of formation of a new type of wobble base pair in which there is the potential of a bifurcated H-bond between the C2 amino proton of 2AP and the O2 and N3 of cytosine. Free energy calculations have not been done for the new 2AP-C base pair, but if the bifurcated H-bond increases stability then the equilibria between this and more active configurations would be decreased. If true, then subtle conformation changes in the transition state needed for optimal chemistry may be impeded.

In kinetic experiments with DNA polymerase α , dTMP incorporation opposite template 2AP was favored 20- to 25-fold over incorporation of dCMP (55). We observed higher apparent discrimination with the RB69 Y567A-DNA polymerase from >100-fold with the shorter DNAs (Table 2) and >300-fold with the longer DNAs (Table 1), which may reflect differences in the DNA polymerases and/or in the enzyme assays; however, much higher discrimination is expected for the wild type T4 and RB69 DNA polymerases, which suggests that the mutagenic 2AP-dCTP base pair may be formed less frequently than predicted from previous experiments.

The major significance of our studies, however, is not the identification of the mutagenic 2AP-dCTP base pair but in demonstrating the importance of dynamic interactions between a DNA polymerase with a dynamic primer-terminal junction at all steps in the nucleotide incorporation pathway (Figure 4). Local DNA sequence has long been known to affect the fidelity of DNA replication on the pivotal complex I/VI, which is at the crossroads of the nucleotide incorporation and proofreading pathways. Our studies extend previous findings by demonstrating a role for breathing at the primer-terminus in determining the accuracy of nucleotide binding to form ternary DNA polymerase/DNA/dNTP complexes and in nucleotide incorporation. Our studies also demonstrate that alterations in minor groove interactions with the primer-terminus produced by the Y567A substitution affect partitioning of DNA between pre- and post-translocated polymerase complexes and that increased formation of post-translocated complexes is correlated with increased extension of mismatched primer ends.

Acknowledgments

We thank Likui Zhang and Neil Johnson for helpful discussions.

REFERENCES

1. Tsai Y-C, Johnson KA. A new paradigm for DNA polymerase specificity. *Biochemistry*. 2006; 45:9675–9687. [PubMed: 16893169]
2. Moran S, Ren RX-F, Kool ET. A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence specificity. *Proc. Natl. Acad. Sci. U.S.A.* 1997; 94:10506–10511. [PubMed: 9380669]
3. Chiaramonte M, Moore CL, Kincaid K, Kuchta RD. Facile polymerization of dNTPs bearing unnatural base analogues by DNA polymerase α and Klenow fragment (DNA polymerase I). *Biochemistry*. 2003; 42:10472–10481. [PubMed: 12950174]

4. Hariharan C, Bloom LB, Helquist SA, Kool ET, Reha-Krantz LJ. dynamics of nucleotide incorporation; snapshots revealed by 2-aminopurine fluorescence studies. *Biochemistry*. 2006; 45:2836–2844. [PubMed: 16503638]
5. Joyce CM, Benkovic SJ. DNA polymerase fidelity: kinetics, structure and checkpoints. *Biochemistry*. 2004; 43:14317–14324. [PubMed: 15533035]
6. Rothwell PJ, Waksman G. A pre-equilibrium before nucleotide binding limits fingers subdomain closure by KlenTaq1. *J. Biol. Chem.* 2007; 282:28884–28892. [PubMed: 17640877]
7. Datta K, Johnson NP, von Hippel PH. DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 2010; 107:17980–17985. [PubMed: 20921373]
8. Freese E. The specific mutagenic effect of base analogues on phage T4. *J. Mol. Biol.* 1959; 1:87–105.
9. Hopkins RL, Goodman MF. Asymmetry in forming 2-aminopurine-hydroxymethylcytosine heteroduplexes; a model giving misincorporation frequencies and rounds of DNA replication from base-pair populations *in vivo*. *J. Mol. Biol.* 1979; 135:1801–1822.
10. Bessman MJ, Muzyczka N, Goodman MF, Schnaar RL. Studies on the biochemical basis of spontaneous mutation. II. The incorporation of a base and its analogue into DNA by wild-type and antimutator DNA polymerases. *J. Mol. Biol.* 1974; 88:409–421. [PubMed: 4616089]
11. Bessman MJ, Reha-Krantz LJ. Studies on the biochemical basis of spontaneous mutation. V. Effect of temperature on mutation frequency. *J. Mol. Biol.* 1977; 116:115–123. [PubMed: 592380]
12. Bloom LB, Otto MR, Eritja R, Reha-Krantz LJ, Goodman MF, Beechem JM. Pre-steady-state kinetic analysis of sequence-dependent nucleotide excision by the 3'-exonuclease activity of bacteriophage T4 DNA polymerase. *Biochemistry*. 1994; 33:7576–7586. [PubMed: 8011623]
13. da Silva, E. Fidalgo; Mandal, SS.; Reha-Krantz, LJ. Using 2-aminopurine fluorescence to measure incorporation of incorrect nucleotides by wild type and mutant bacteriophage T4 DNA polymerases. *J. Biol. Chem.* 2002; 277:40640–40649. [PubMed: 12189135]
14. Law SM, Eritja R, Goodman MF, Breslauer KJ. Spectroscopic and calorimetric characterization of DNA duplexes containing 2-aminopurine. *Biochemistry*. 1996; 35:12329–12337. [PubMed: 8823167]
15. Sowers LC, Boulard Y, Fazakerley GV. Multiple structures for the 2-aminopurine-cytosine mispair. *Biochemistry*. 2000; 39:7613–7620. [PubMed: 10858312]
16. Sherer EC, Cramer SJ. Quantum chemical characterization of the cytosine: 2-aminopurine base pair. *J. Computational Chem.* 2001; 22:1167–1179.
17. Hogg M, Cooper W, Reha-Krantz L, Wallace SS. Kinetics of error generation in homologous B-family DNA polymerases. *Nucleic Acids Res.* 2006; 34:2528–2535. [PubMed: 16687658]
18. Bebenek A, Dressman HK, Carver GT, Ng S, Petrov V, Yang G, Konigsberg WH, Karam JD, Drake JW. Interacting fidelity defects in the replicative DNA polymerase of bacteriophage RB69. *J. Biol. Chem.* 2001; 276:10387–10397. [PubMed: 11133987]
19. Graziewicz MA, Bienstock RJ, Copeland WC. The DNA polymerase γ Y955C disease variant associated with PEO and parkinsonism mediates the incorporation and translesion synthesis opposite 7,8-dihydro-8-oxo-2'-deoxyguanosine. *Hum. Mol. Genet.* 2007; 16:2729–2739. [PubMed: 17725985]
20. Beckman J, Wang M, Blaha G, Wang J, Konigsberg WH. Substitution of Ala for Tyr567 in RB69 DNA polymerase allows dAMP to be inserted opposite 7,8-dihydro-8-oxoguanine. *Biochemistry*. 2010; 43:8554–8563. [PubMed: 20795733]
21. Xia S, Wang M, Lee HR, Sinha A, Blaha G, Christian T, Wang J, Konigsberg W. Variation in mutation rates caused by RB69 pol fidelity mutants can be rationalized on the basis of their kinetic behavior and crystal structures. *J. Mol. Biol.* 2011; 406:558–570. [PubMed: 21216248]
22. Reha-Krantz LJ. DNA polymerase proofreading: multiple roles maintain genome stability. *Biochimica et Biophysica Acta*. 2010; 1804:1049–1063.
23. Ward DC, Reich E, Stryer L. Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* 1969; 244:1228–1237. [PubMed: 5767305]

24. Frey MA, Sowers LC, Millar DP, Benkovic SJ. The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of *Escherichia coli* polymerase I and bacteriophage T4 DNA polymerase. *Biochemistry*. 1995; 34:9185–9192. [PubMed: 7619819]
25. Mandal SS, da Silva Fidalgo, Reha-Krantz LJ. Using 2-aminopurine fluorescence to detect base unstacking in the template strand during nucleotide incorporation by the bacteriophage T4 DNA polymerase. *Biochemistry*. 2002; 41:4399–4406. [PubMed: 11914087]
26. Hariharan C, Reha-Krantz LJ. Using 2-aminopurine fluorescence to detect bacteriophage T4 DNA polymerase-DNA complexes that are important for primer extension and proofreading reactions. *Biochemistry*. 2005; 44:15674–15684. [PubMed: 16313170]
27. Rachofsky EL, Seibert E, Stivers JT, Osman R, Ross JBA. Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine. *Biochemistry*. 2001; 40:957–967. [PubMed: 11170417]
28. Jean JM, Hall KB. 2-Aminopurine fluorescence quenching and lifetimes: role of the base stacking. *Proc. Natl. Acad. Sci. U.S.A.* 2001; 98:37–41. [PubMed: 11120885]
29. Neely RK, Daujotyte D, Grazulis S, Magennis SW, Dryden DTF, Klimasauskas S, Jones AC. Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes. *Nucleic Acids Res.* 2005; 33:6953–6969. [PubMed: 16340006]
30. Beechem JM, Otto MR, Bloom LB, Eritja R, Reha-Krantz LJ, Goodman MF. Exonuclease-polymerase active site partitioning of primer-template DNA strands and equilibrium Mg^{2+} binding properties of bacteriophage T4 DNA polymerase. *Biochemistry*. 1998; 28:9095–9103.
31. Tleugabulova D, Reha-Krantz LJ. Probing DNA polymerase-DNA interactions: examining the template strand in exonuclease complexes using 2-aminopurine fluorescence and acrylamide quenching. *Biochemistry*. 2007; 46:6559–6569. [PubMed: 17497891]
32. Subuddhi U, Hogg M, Reha-Krantz LJ. Use of 2-aminopurine fluorescence to study the role of the β hairpin in the proofreading pathway catalyzed by the phage T4 and RB69 DNA polymerases. *Biochemistry*. 2008; 47:6130–6137. [PubMed: 18481871]
33. Shamoo Y, Steitz TA. Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell*. 1999; 99:155–166. [PubMed: 10535734]
34. Hogg M, Aller P, Konigsberg W, Wallace SS, Doublié S. Crystallographic snapshots of a replication DNA polymerase encountering an abasic site. *EMBO J.* 2004; 23:1483–1493. [PubMed: 15057283]
35. Hogg M, Aller P, Konigsberg W, Wallace SS, Doublié S. Structural and biochemical investigation of the role in proofreading of a β hairpin loop found in the exonuclease domain of a replicative DNA polymerase of the B family. *J. Biol. Chem.* 2007; 282:1432–1444. [PubMed: 17098747]
36. Franklin MC, Wang J, Steitz TA. Structure of the replication complex of a pol α family DNA polymerase. *Cell*. 2001; 105:657–667. [PubMed: 11389835]
37. Zhang H, Beckman J, Wang J, Konigsberg W. RB69 DNA polymerase mutants with expanded nascent base-pair-binding pockets are highly efficient but have reduced base selectivity. *Biochemistry*. 2009; 48:6940–6950. [PubMed: 19522539]
38. Marquez LA, Reha-Krantz LJ. Using 2-aminopurine fluorescence and mutational analysis to demonstrate an active role of bacteriophage T4 DNA polymerase in strand separation required for 3'→5'-exonuclease activity. *J. Biol. Chem.* 1996; 271:28903–28911. [PubMed: 8910538]
39. Reha-Krantz LJ. The use of 2-aminopurine fluorescence to study DNA polymerase function. *In* DNA Replication Methods and Protocols. *Methods in Molecular Biology*. 2009; 521:381–396. [PubMed: 19563118]
40. Navaza J. Implementation of molecular replacement in AMoRe. *Acta Crystallogr.* 2001; D53:240–255.
41. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr.* 2004; D60:2216–2132.
42. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* 1997; D53:240–255.

43. DeLano, WL. The PyMol Molecular Graphics System. Delano Scientific; San Carlos, CA: 2002. <http://www.pymol.org>
44. Yang G, Franklin M, Li J, Lin TC, Konigsberg W. A conserved Tyr residue is required for sugar selectivity in a pol alpha DNA polymerase. *Biochemistry*. 2002; 41:10256–10261. [PubMed: 12162740]
45. Drake JW, Allen EF, Forsberg SA, Preparata R-M, Greening EO. Spontaneous mutation. *Nature*. 1969; 221:1128–1132. [PubMed: 4975273]
46. Greive SJ, von Hippel PH. Thinking quantitatively about transcriptional regulation. *Nature Reviews Mol. Cell Biol.* 2005; 6:221–232.
47. Wang M, Xia S, Blaha G, Steitz TA, Konigsberg WH, Wang J. Insights into base selectivity from the 1.8 Å resolution structure of an RB69 DNA polymerase ternary complex. *Biochemistry*. 2011; 50:581–590. [PubMed: 21158418]
48. Berman AJ, Kamtekar S, Goodman JL, Lazaro JM, deVega M, Blanco L, Salas M, Steitz TA. Structures of ϕ 29 DNA polymerase complexes with substrate: the mechanisms of translocation in B-family polymerases. *EMBO J.* 2007; 26:3496–3505.
49. Seeman ND, Rosenberg JM, Rich A. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A.* 1976; 73:804–808. [PubMed: 1062791]
50. Hochstrasser RA, Carver TE, Sowers LC, Millar DP. Melting of a DNA helix terminus within the active site of a DNA polymerase. *Biochemistry*. 1994; 33:11971–11979. [PubMed: 7918416]
51. Jose, D.; Datta, K.; Johnson, NP.; von Hippel, PH. Spectroscopic studies of position-specific DNA “breathing” fluctuations at replication forks and primer-template junctions. 2009.
52. Wang Z, Lazaeov E, O'Donnell M, Goodman MF. Resolving a fidelity paradox. Why *Escherichia coli* DNA polymerase II makes more base substitution errors in AT- compared with GC-rich DNA. *J. Biol. Chem.* 2002; 277:4446–4454. [PubMed: 11733526]
53. Wang F, Yang W. Structural insight into translesion synthesis by DNA pol II. *Cell*. 2009; 139:1279–1289. [PubMed: 20064374]
54. Delagoutte E, von Hippel PH. Function and assembly of the bacteriophage T4 DNA replication complex: interactions of the T4 DNA polymerase with various model DNA constructs. *J. Biol. Chem.* 2003; 278:25435–25447. [PubMed: 12700227]
55. Watanabe SM, Goodman MF. Kinetic measurements of 2-aminopurine-cytosine and 2-aminopurine-thymine base pairs as a test of DNA polymerase fidelity mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 1982; 79:6429–6433. [PubMed: 6959128]

FIGURE 1.
Proposed structures of 2AP-T (A) and 2AP-C base pairs (B-D).

DNA	Sequence
2AP (P) in the +1 templating position	
1	3' B-CCCTTCGTGCAGTAGCCATTAAT ⁺¹ P _{GATCGATGGTTT} GGGAAGCACGTCATCGGTAATTA _{OH/dd}
2	3' B-CCCTTCGTGCAGTAGCCGCCGCTP _{GATCGATGGTTT} GGGAAGCACGTCATCGGCGGCGA _{OH/dd}
3	3' GTTAATTAATTAATP _{CTT} AATTAATTAATTA _{OH/dd}
4	3' GGCGCGCCGCCGCTP _{CTT} CGCGCGGCGGCGA _{OH/dd}
2AP (P) at the terminal position of the primer	
5	3' B-CCCTTCGTGCAGTAGCCATTATAGATCGATGGTTT GGGAAGCACGTCATCGGTAATP
6	3' CGTGCAGTAATTGCCACGGATCGATGGTTT GCACGTCATTAACGGTP
7	3' CGTGCAGTAGCCATGCCGGATCGATGGTTT GCACGTCATCGGTAATP

FIGURE 2.

DNA substrates with 2AP (P) in the template strand (DNAs 1-4) or in the primer strand (DNAs 5-7).

FIGURE 3.

Fluorescence emission spectra and fluorescence lifetimes for 2AP in binary complexes formed with the RB69 wild type and Y567A-DNA polymerases. Panel A: Fluorescence emission spectra were determined for binary complexes formed with DNA 1 (described in Figure 2). 2AP is located in the +1 position of the template strand as illustrated by the cartoon. Complexes were excited at 315 nm as described in Materials and Methods. Panel B: Fluorescence lifetimes were determined for binary complexes formed with the RB69 wild type and Y567A-DNA polymerases and DNA1 as described in Material and Methods and by Hariharan *et al.* (26).

FIGURE 4.
Proposed nucleotide incorporation pathway catalyzed by the phage T4 and RB69 DNA polymerases.

FIGURE 5.

Fluorescence emission spectra for exonuclease complexes formed with the RB69 *exo*⁻ wild type and *exo*⁻ Y567A-DNA polymerases and DNA substrates labeled with 2AP (P) at the 3'-end of the primer strand. Panel A: DNA polymerase complexes formed with DNA 5, which has a terminal 2AP-T base pair (Figure 2). Panel B: DNA polymerase complexes formed with DNA 6, which has a terminal 2AP-C base pair (Figure 2). Panel C: DNA polymerase complexes formed with DNA 7, which has three preformed terminal mismatches. Similar emission spectra were observed for the RB69 *exo*⁻ (black diamonds) and *exo*⁻ Y567A- (open circles) DNA polymerases.

FIGURE 6.

Nucleotide binding assays. Panel A: Formation of ternary DNA pol-DNA-dTTP complexes with the Y567A-DNA polymerase as a function of dTTP concentration. Incremental increases in dTTP concentration produces incremental increases in fluorescence quench. Panel B: Formation of ternary DNA pol-DNA-dCTP complexes with the Y567A-DNA polymerase as a function of dCTP concentration.

FIGURE 7.

Close-up views of nascent base pairs formed in the polymerase active sites of the RB69 wild type and Y567A-DNA polymerases. Panel A: 2.09-Å resolution of the wild type DNA polymerase/A+T-rich DNA3/dTTP ternary complex contoured at 2.0 σ . Panel B: 2.24-Å resolution of the wild type DNA polymerase/G+C-rich DNA4/dTTP ternary complex contoured at 2.0 σ . Panel C: 2.44-Å resolution of the Y567A-DNA polymerase/A+T-rich DNA3/dTTP ternary complex contoured at 1.5 σ . Panel D: 2.25-Å resolution of the Y567A-DNA polymerase/G+C-rich DNA4/dTTP ternary complex contoured at 2.0 σ . Panel E: 3.19-Å resolution of the Y567A-DNA polymerase/A+T-rich DNA3/dCTP ternary complex contoured at 1.0 σ . Panel F: 2.32-Å resolution of the Y567A-DNA polymerase/G+C-rich DNA4/dCTP ternary complex contoured at 1.5 σ .

FIGURE 8.

Predicted H-bonds between 2AP-T and 2AP-C base pairs formed in the polymerase active sites of the RB69 wild type and Y567A-DNA polymerases. Interbase distances are consistent with two H-bonds formed for the 2AP-T base pair as proposed (Figure 1A). One H-bond is between N1 of 2AP and N3 of dTTP and the second is between a proton of the C2 amino group of 2AP and O2 of thymine as shown by dashed blue lines. Similar structures were observed for complexes formed with the wild type DNA polymerase and the A+T-rich DNA3 (panel A) and the G+C-rich DNA4 (panel B) and for the Y567A-DNA polymerase and the A+T-rich DNA3 (panel C) and the G+C-rich DNA4 (panel D). A wobble-type 2AP-C base pair was observed for complexes formed with the Y567A-DNA polymerase, dCTP and the A+T-rich DNA3 (panel E) and the G+C-rich DNA4 (panel F). The apparent non-ideal H-bond between N1 of 2AP and a proton from the C4 amino group of dCTP and the apparent bifurcated H-bond between a proton of the C2 amino group of 2AP and the N3 as well as O2 of dCTP are shown as red dashed lines.

FIGURE 9.

A flexible nucleotide binding pocket is produced by the Y567A substitution. The hydrogen-bonding network in the nucleotide binding pocket formed with the RB69 wild type DNA polymerase, G+C-rich DNA4, and dTTP (panel A) and the disrupted hydrogen-bonding network formed with the Y567A-DNA polymerase, G+C-rich DNA4, and dCTP (panel B); similar disrupted hydrogen-bonding network was observed with dTTP. Superimposition of the nucleotide binding pockets of the wild type and Y567A-DNA polymerases with bound dTTP or dCTP, respectively with the G+C-rich DNA4 (panel C) or with the A+T-rich DNA 3 (panel D). Note that residue G568 is shifted 0.4 Å towards Y416 and that the distance between N3 of 2AP and G568 increases in the complex formed with the Y567A-DNA polymerase.

Table 1
Kinetic parameters for binding and incorporation of nucleotides opposite template 2AP with 'long' DNA substrates

DNA pol	A+T-rich (DNA 1)					G+C-rich (DNA 2)				
	$K_{d\text{bind}}(\text{dTTP})$	$k_{\text{pol}}(\text{dTTP})$	$K_{d\text{bind}}(\text{dCTP})$	$k_{\text{pol}}(\text{dCTP})$		$K_{d\text{bind}}(\text{dTTP})$	$k_{\text{pol}}(\text{dTTP})$	$K_{d\text{bind}}(\text{dCTP})$	$k_{\text{pol}}(\text{dCTP})$	
T4exo ⁻	$35.0 \pm 1.2 \mu\text{M}^a$	rapid ^d	$>3 \text{ mM}^{a,b}$	na ^c		$28.7 \pm 0.6 \mu\text{M}$	rapid	$\sim 3 \text{ mM}^b$	na ^c	
RB69exo ⁻	$31.1 \pm 0.9 \mu\text{M}$	$\sim$	$\sim 3 \text{ mM}^b$	na ^c		$30.0 \pm 0.4 \mu\text{M}$	$\sim$	$\sim 3 \text{ mM}^b$	na ^c	
RB69Y567Aexo ⁻	$9.2 \pm 0.8 \mu\text{M}$	$\sim$	$24.8 \pm 2.1 \mu\text{M}$	nd ^d		$15.1 \pm 1.4 \mu\text{M}$	$\sim$	$317 \pm 18 \mu\text{M}$	nd ^d	
	$K_{d\text{incorp, app}}(\text{dTMP})$	$k_{\text{pol}}(\text{dTMP})^e$	$K_{d\text{incorp, app}}(\text{dCMP})$	$k_{\text{pol}}(\text{dCMP})^e$		$K_{d\text{incorp, app}}(\text{dTMP})$	$k_{\text{pol}}(\text{dTMP})^e$	$K_{d\text{incorp, app}}(\text{dCMP})$	$k_{\text{pol}}(\text{dCMP})^e$	
T4exo ⁻	$367 \pm 36 \mu\text{M}^a$	$165 \pm 3 \text{ s}^{-1} a$	$>3 \text{ mM}^{a,b}$	$2.3 \pm 0.1 \text{ s}^{-1} a$		nd	nd	nd	nd	
RB69Y567Aexo ⁻	$16 \pm 3 \mu\text{M}$	$240 \pm 11 \text{ s}^{-1}$	$170 \pm 50 \mu\text{M}$	$6.4 \pm 0.6 \text{ s}^{-1}$		$180 \pm 75 \mu\text{M}$	$190 \pm 11 \text{ s}^{-1}$	2 mM^b	$\sim 6 \text{ s}^{-1}$	
	$k_{\text{pol}} / K_{d\text{incorp}}(\text{dTMP}) \mu\text{M}^{-1} \text{ s}^{-1}$	$k_{\text{pol}} / K_{d\text{incorp}}(\text{dCMP}) \mu\text{M}^{-1} \text{ s}^{-1}$				$k_{\text{pol}} / K_{d\text{incorp}}(\text{dTMP}) \mu\text{M}^{-1} \text{ s}^{-1}$	$k_{\text{pol}} / K_{d\text{incorp}}(\text{dCMP}) \mu\text{M}^{-1} \text{ s}^{-1}$			
T4exo ⁻	0.45 ^f	<0.0008				nd	nd		nd	
RB69Y567Aexo ⁻	15	0.04				1			~ 0.003	
			dTMP/dCMP				dTMP/dCMP			
T4exo ⁻		>560					nd			
RB69Y567Aexo ⁻		375					333			

^aData from Fidalgo *et al.* (13). Rapid nucleotide binding occurs within the dead time of the stopped flow instrument.

^bHigh K_d apparent value likely includes non-specific interactions and, thus, the true K_d may be higher.

^cNot applicable. The rate of nucleotide binding cannot be determined.

^dNot determined.

^eOverall nucleotide incorporation rate.

^f $k_{\text{pol}} / K_{d\text{incorp}}(\text{dTMP}) \mu\text{M}^{-1} \text{ s}^{-1}$ for incorporation of dTMP opposite template A is 10.3 (13).

Table 2
Kinetic parameters for binding and incorporation of nucleotides opposite template 2AP with ‘short’ DNA substrates

DNA pol	A+T-rich (DNA 3)					G+C-rich (DNA 4)				
	$K_{\text{dbind}}(\text{dTTP})$	$k_{\text{pol}}(\text{dTTP})$	$K_{\text{dbind}}(\text{dCTP})$	$k_{\text{pol}}(\text{dCTP})$		$K_{\text{dbind}}(\text{dTTP})$	$k_{\text{pol}}(\text{dTTP})$	$K_{\text{dbind}}(\text{dCTP})$	$k_{\text{pol}}(\text{dCTP})$	
RB69Y567Aexo ⁻	nd ^a	nd	53 ± 11 μM	nd		nd	nd	373 ± 88 μM	nd	
	$K_{\text{dincorp,app}}(\text{dTMP})$	$k_{\text{pol}}(\text{dTMP})^b$	$K_{\text{dincorp,app}}(\text{dCMP})$	$k_{\text{pol}}(\text{dCMP})^b$		$K_{\text{dincorp,app}}(\text{dTMP})$	$k_{\text{pol}}(\text{dTMP})^b$	$K_{\text{dincorp,app}}(\text{dCMP})$	$k_{\text{pol}}(\text{dCMP})^b$	
RB69 exo ⁻	114 ± 26 μM	379 ± 28 s ⁻¹	>2mM ^c	slow ^c		200 ± 45 μM	201 ± 16 s ⁻¹	>2mM ^c	slow ^c	
RB69Y567Aexo ⁻	67.3 ± 8 μM	289 ± 9 s ⁻¹	298 ± 65 μM	9.3 ± 0.7 s ⁻¹		263 ± 78 μM	233 ± 27 s ⁻¹	361 ± 52 μM	2.5 0.1 s ⁻¹	
	$k_{\text{pol}} / K_{\text{dincorp}}(\text{dTMP}) \mu\text{M}^{-1} \text{s}^{-1}$	$k_{\text{pol}} / K_{\text{dincorp}}(\text{dCMP}) \mu\text{M}^{-1} \text{s}^{-1}$				$k_{\text{pol}} / K_{\text{dincorp}}(\text{dTMP}) \mu\text{M}^{-1} \text{s}^{-1}$	$k_{\text{pol}} / K_{\text{dincorp}}(\text{dCMP}) \mu\text{M}^{-1} \text{s}^{-1}$			
RB69exo ⁻	3.3		<0.001-0.003			1.0		<0.001-0.003		
RB69Y567Aexo ⁻	4.3		0.03			0.9		0.007		
			dTMP/dCMP				dTMP/dCMP			
RB69exo ⁻			>1000					>1000		
RB69Y567Aexo ⁻			143					129		

^aNot determined.

^bOverall nucleotide incorporation rate.

^cThe kinetic parameters for the misincorporation of dCMP opposite template 2AP cannot be determined accurately because of non-specific interactions, but the K_{d} apparent is >2mM and incorporation is slow on the order of 2 – 6 s⁻¹.

Table 3

K_d s for binding dUTP and rUTP opposite template 2AP

DNA pol	K_d dUTP	K_d rUTP
T4 exo ⁻	$74.1 \pm 2 \mu\text{M}$	>3 mM
RB69 Y567A exo ⁻	$21 \pm 1 \mu\text{M}$	$522 \pm 55 \mu\text{M}$

Experimental conditions are the same as for Table 1.

Table 4

Data Collection and Refinement Statistics for Ternary Complexes

	ATrich (wt)	GCrich (wt)	ATrich (Y567A)	GCrich (Y567A)	dTTP vs 2AP	dTTP vs 2AP	dTTP vs 2AP	dTTP vs 2AP	dCTP vs 2AP	GCrich (Y567A)
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell (Å)										
a	74.95	74.68	74.84	74.40	74.40	74.40	74.40	74.40	74.40	74.40
b	119.88	120.47	120.99	120.07	120.07	120.07	120.07	120.07	120.07	120.07
c	130.43	130.99	127.37	130.89	130.89	130.89	130.89	130.89	130.89	130.89
Resolution (Å)	50.0-2.09 (2.13-2.09)	50.0-2.24 (2.28-2.24)	50.0-2.44 (2.48-2.44)	50.0-2.25 (2.29-2.25)	50.0-2.25 (2.29-2.25)	50.0-2.25 (2.29-2.25)	50.0-2.25 (2.29-2.25)	50.0-2.25 (2.29-2.25)	50.0-2.32 (2.36-2.32)	50.0-2.32 (2.36-2.32)
No of unique reflections	64,604	54,679	41,995	52,104	52,104	52,104	52,104	52,104	48,550	48,550
Redundancy	3.9 (3.6)	4.1 (4.0)	3.9 (3.8)	4.0 (4.0)	4.0 (4.0)	4.0 (4.0)	4.0 (4.0)	4.0 (4.0)	4.0 (4.0)	4.0 (4.0)
Completeness (%)	98.5 (97.0)	99.8 (99.9)	99.9 (100.0)	96.1 (95.9)	96.1 (95.9)	96.1 (95.9)	96.1 (95.9)	96.1 (95.9)	99.8 (100.0)	99.8 (100.0)
R _{merge} (%)	11.6 (95.2)	11.3 (78.0)	9.8 (89.9)	8.9 (81.9)	8.9 (81.9)	8.9 (81.9)	8.9 (81.9)	8.9 (81.9)	9.8 (91.4)	9.8 (91.4)
I / σ	11.0 (0.8)	12.4 (2.0)	11.8 (1.1)	13.4 (1.6)	13.4 (1.6)	13.4 (1.6)	13.4 (1.6)	13.4 (1.6)	13.1 (1.1)	13.1 (1.1)
Final model										
Amino acid residues	903	903	903	903	903	903	903	903	903	903
Water molecules	350	285	162	254	254	254	254	254	263	263
Ca ²⁺ ions	5	5	5	5	5	5	5	5	5	5
Template nucleotides	18	18	18	18	18	18	18	18	18	18
Primer nucleotides	13	13	13	13	13	13	13	13	13	13
dNTP	1	1	1	1	1	1	1	1	1	1
R (%)	20.2 (30.5)	20.2 (36.2)	20.7 (28.9)	20.9 (41.9)	20.9 (41.9)	20.9 (41.9)	20.9 (41.9)	20.9 (41.9)	20.3 (24.7)	20.3 (24.7)
R _{free} (%)	24.3 (33.9)	23.8 (42.6)	25.5 (34.1)	25.4 (47.4)	25.4 (47.4)	25.4 (47.4)	25.4 (47.4)	25.4 (47.4)	26.0 (30.1)	26.0 (30.1)
r.m.s.d										
Bond length (Å)	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
Bond angles (°)	1.122	1.116	1.138	1.132	1.132	1.132	1.132	1.132	1.124	1.124
PDB access code	3SQ2	3SQ4	3SUN	3SUO	3SUO	3SUO	3SUO	3SUO	3SUP	3SUP