BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

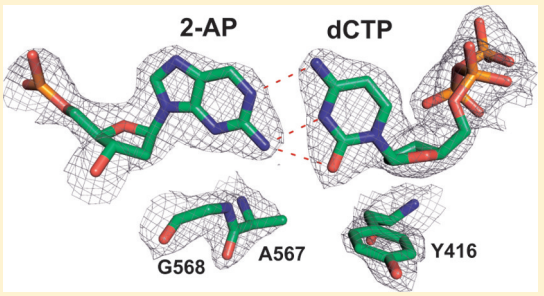
Structure of the 2-Aminopurine-Cytosine Base Pair Formed in the Polymerase Active Site of the RB69 Y567A-DNA Polymerase

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

[‡]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8024, United States

ABSTRACT: The adenine base analogue 2-aminopurine (2AP) is a potent base substitution mutagen in prokaryotes because of its enhanceed 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 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 atoms 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 the 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 facilitate 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.



DNA polymerases incorporate nucleotides with a high degree of 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 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 surveillance steps described above. 2-Aminopurine (2AP), 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 AT base pair by using O2 instead of 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 enhance the 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

Received: September 19, 2011
Revised: October 22, 2011
Published: October 24, 2011

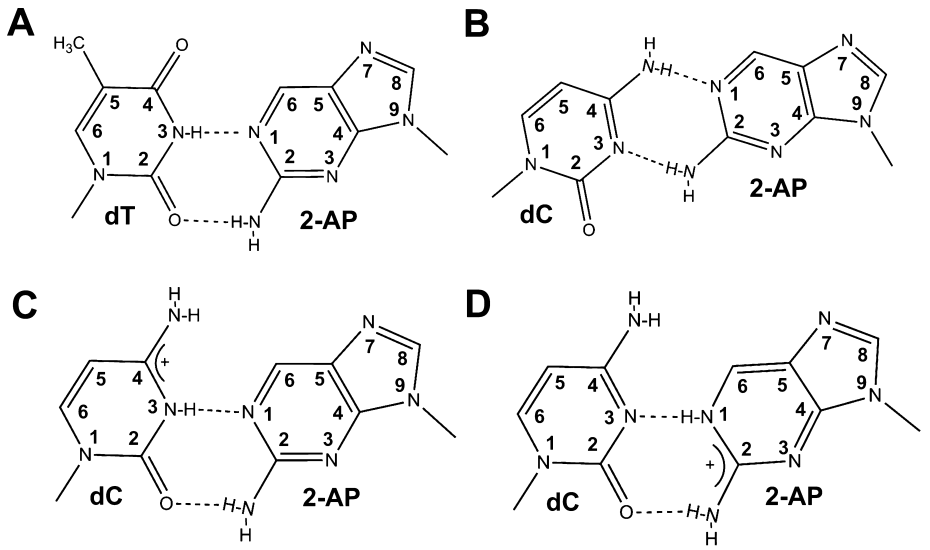


Figure 1. Proposed structures of 2AP-T (A) and 2AP-C (B–D) base pairs.

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 enhances 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 an increased level of misincorporation, the Y567A-DNA polymerase also has an enhanced 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 enhance the 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 at position 1 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.

At least three or four 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 at position 1 (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

DNA	Sequence
	2AP (P) in the +1 templating position
	+1
1	3′ B-CCCTTCGTGCAGTAGCCATTAATPGATCGATGGTTT GGGAAGCACGTCATCGGTAATTA OH/dd
2	3′ B-CCCTTCGTGCAGTAGCCGCCGCTPGATCGATGGTTT GGGAAGCACGTCATCGGCGGCGA OH/dd
3	3′ GTTAATTAATTAATPCTT AATTAATTAATTA OH/dd
4	3′ GGCGCGCCGCCGCTPCTT CGCGCGGCGGCGA OH/dd
	2AP (P) at the terminal position of the primer
5	3′ B-CCCTTCGTGCAGTAGCCATTATAGATCGATGGTTT GGGAAGCACGTCATCGGTAATP
6	3′ CGTGCAGTAATTGCCACGGATCGATGGTTT GCACGTCATTAACGGTP
7	3′ CGTGCAGTAGCCATGCCGGATCGATGGTTT GCACGTCATCGGTA$_{ATP}$

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

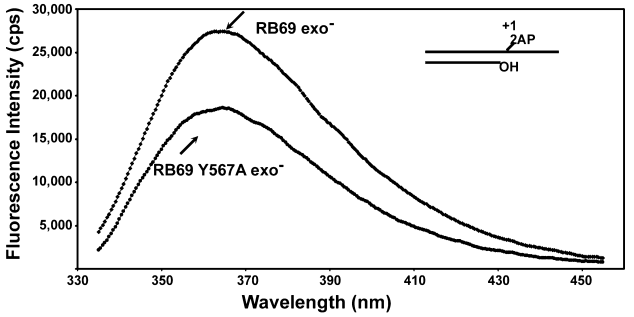
2AP is placed in the terminal position of the primer strand, which indicates that 2AP at the primer end in exonuclease complexes is also unstacked,[26,30−32] as observed in the crystalline state.[33−35]

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 at position 1 of the template strand if the primer-terminal region is G+C-rich as opposed to A+T-rich.[26] As G+C richness favors the 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.

Mg^{2+}-dependent dTTP binding opposite template 2AP produces a rapid quench of 2AP fluorescence that is observed in reactions with chain-terminated DNA substrates (2′,3′-

dx.doi.org/10.1021/bi2014618|*Biochemistry* 2011, 50, 10136–10149

A



B

DNA polymerase	Fluorescence lifetimes (ns) τ_1	τ_2	τ_3
RB69 exo⁻	0.23 (0.46)	4.6 (0.16)	10.4 (0.38)
RB69 Y567A exo⁻	0.71 (0.38)	5.3 (0.42)	9.2 (0.20)

Figure 3. Fluorescence emission spectra and fluorescence lifetimes for 2AP in binary complexes formed with the RB69 wild-type and Y567A-DNA polymerases. (A) Fluorescence emission spectra determined for binary complexes formed with DNA 1 (described in Figure 2). 2AP is located at position 1 of the template strand as illustrated by the cartoon. Complexes were excited at 315 nm as described in Experimental Procedures. (B) Fluorescence lifetimes determined for binary complexes formed with the RB69 wild-type and Y567A-DNA polymerases and DNA 1 as described in Experimental Procedures and by Hariharan et al.[26]

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 μM. 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 of 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 of 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 of <150 μM, but dFTP is bound because it is an effective competitive inhibitor of dTTP-Mg^{2+}-induced quenching with a K_i of ~68 μM.[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 apparent K_d ($K_{d,app}$) for incorporation of dTMP opposite template 2AP to form complex V is >10-fold higher (367 μM) 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 analogue, 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 that 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 (PP_i) 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 an 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 weakened 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 previously.[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 purified via polyacrylamide gel electrophoresis. 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 from 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 2–2.5-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 dideoxynucleotide and the reaction mixtures contained $MgCl_2$ (10 mM), which is essential for nucleotide binding.[25] Equilibrium dissociation constants for binding of a nucleotide opposite template 2AP were determined as described

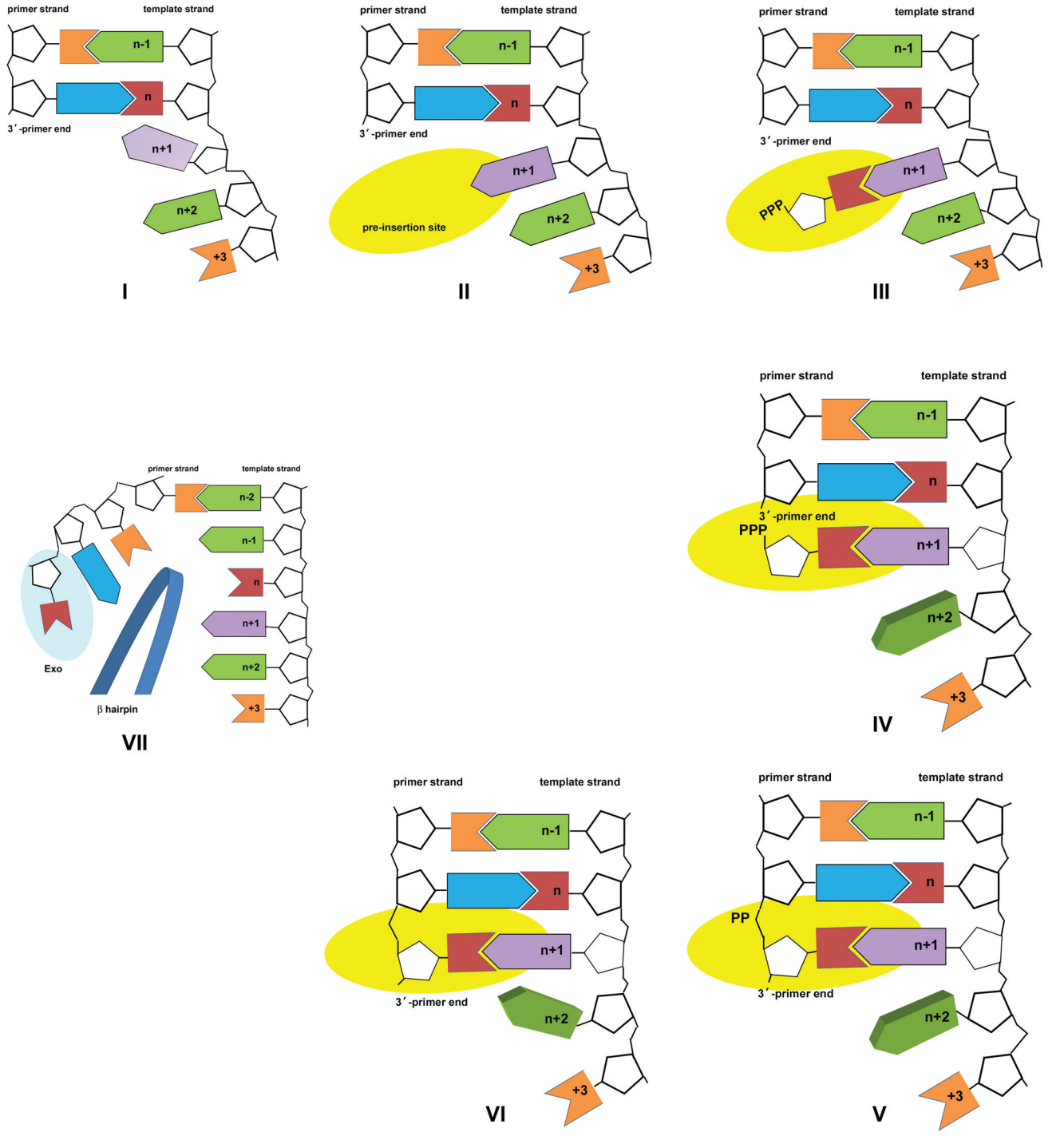


Figure 4. Proposed nucleotide incorporation pathway catalyzed by the phage T4 and RB69 DNA polymerases.

previously.[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 buffer described above 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 analysis 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′-^{32}P-labeled DNA substrate. Each reaction mixture contained 66 mM Tris-HCl (pH 7.5), and each reaction was performed at 22 °C by mixing equal volumes of the 200 nM DNA substrate with 2 μM DNA polymerase, 20 mM $MgCl_2$, and varying concentrations of dNTPs. The final concentrations were 100 nM DNA substrate, 1 μM DNA polymerase, and 10 mM $MgCl_2$. 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.

dx.doi.org/10.1021/bi2014618|*Biochemistry* 2011, 50, 10136–10149

The band intensities were determined using a Molecular Dynamics Phosphorimager and analyzed with Imagequant. Data from the single-turnover experiments were fit to a single-exponential equation as described previously.[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 via 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 $CaCl_2$, and 10% (w/v) polyethylene glycol 350 monomethyl ether (PEG350 MME). Cube-shaped crystals were stabilized and cryoprotected by being transferred to the stabilization solution containing 100 mM sodium cacodylate buffer (pH 6.5), 20% (w/v) PEG350 MME, and 100 mM $CaCl_2$. Just prior to being frozen in liquid nitrogen, the crystals were transformed into the same cyroprotectant 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 $P2_12_12_1$ with slightly different cell dimensions.

All six structures were determined 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 Protein Data Bank (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 PDB.

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 the sake of simplicity, the RB69 and T4 exo$^-$ DNA polymerases are termed wild type and the RB69 exo$^-$ Y567A-DNA polymerase is termed the mutant or Y567A-DNA polymerase. Complexes were formed with 200 nM DNA that was labeled at position 1 of the template strand with 2AP [DNA 1 (Figure 2)] and 500 nM enzyme. A significantly lower fluorescence intensity was observed for the Y567A-DNA polymerase that did not increase upon addition of more enzyme. The high fluorescence intensity of binary pol complexes is due to formation of a large population of fluorescent species with an ~10 ns lifetime that is produced by unstacking 2AP at position 1 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 (τ = 5.3 ns; 42%). The reduced level of 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 at position 1 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 an increased level of 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 a high fluorescence intensity because the terminal 2AP is unstacked in the exonuclease active site;[33,34] an ~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 [DNA 2 (Figure 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 the level of formation of exo complexes, a 9.7 ns fluorescence lifetime indicative of exo complexes is still detected, but the amplitude is reduced from 100%, which is observed for the wild-type DNA polymerase without the amino acid substitutions, to 17% 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 the fluorescence intensity (Figure 5A). A likely explanation for the apparent reduced level of binary exo complexes is that the Y567A-DNA polymerase has enhanced ability to form less fluorescent binary pol complexes, which has been observed previously for mutant T4 DNA polymerases that have a weakened 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, a 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 three 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 polymerase can form exo complexes as efficiently as the exo$^-$ wild-type DNA polymerase only if formation of pol complexes is prevented by three terminal mismatches. These observations are consistent

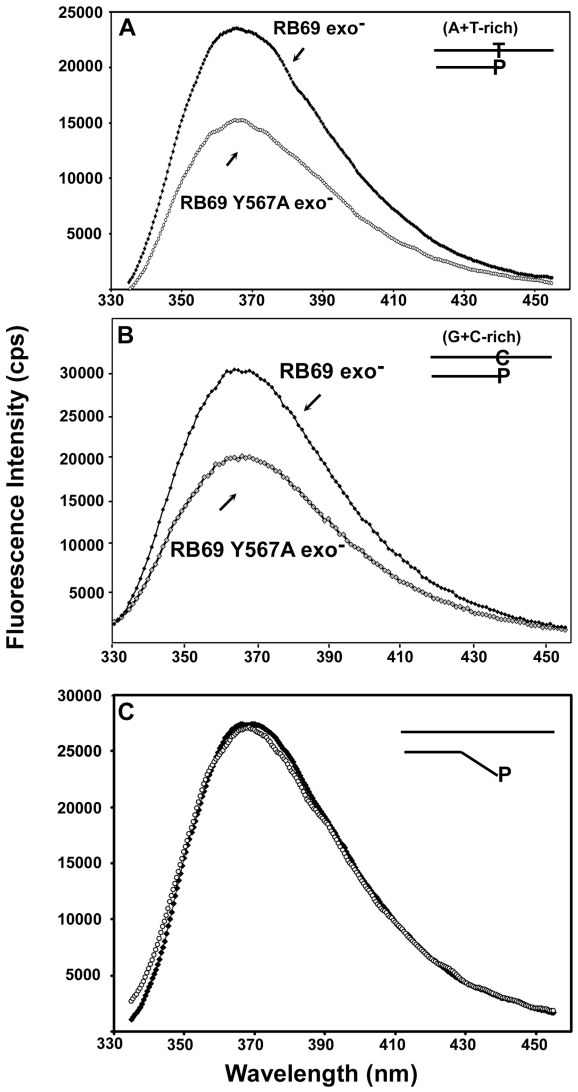


Figure 5. Fluorescence emission spectra for exonuclease complexes formed with the RB69 exo⁻ wild type and exo⁻ Y567A-DNA polymerases and DNA substrates labeled with 2AP (P) at the 3′-end of the primer strand. (A) DNA polymerase complexes formed with DNA 5, which has a terminal 2AP-T base pair (Figure 2). (B) DNA polymerase complexes formed with DNA 6, which has a terminal 2AP-C base pair (Figure 2). (C) DNA polymerase complexes formed with DNA 7, which has three preformed terminal mismatches. Similar emission spectra were observed for the RB69 exo⁻ DNA (◆) and exo⁻ Y567A-DNA (○) polymerases.

with the proposal that the Y567A substitution affects the partitioning of DNA between the polymerase and exonuclease active sites in favor of an increased level of 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), binary pol complexes are also expected to be formed with DNA 1 in which the primer end is matched and 2AP is in templating position 1 (Figure 3). Thus, the lower level of fluorescence intensity observed for complexes formed with the Y567A-DNA polymerase compared to that of the wild-type DNA polymerase and DNA 1 does not appear to be due to formation of binary exo complexes, but instead to an increased level of formation of less fluorescent binary pol complexes (complex II). The increase in the amplitude of the intermediate fluorescence lifetime species (τ = 5.3 ns) for the mutant DNA polymerase is consistent with this proposal (Figure 3B).

The RB69 Y567A-DNA Polymerase Is Less Able To Discriminate against Binding dCTP and Incorporating dCMP opposite Template 2AP. Mg^{2+}-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 $MgCl_2$. Addition of incremental amounts of dTTP to the mixture quenches the fluorescence intensity via the 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 μM for DNA substrates that are relatively A+T-rich [DNA 1 (Figure 2)] or G+C-rich [DNA 2 (Figure 2)] in the primer-terminal region (Table 1). Wild-type T4 and RB69 DNA polymerases strongly discriminate against binding of dCTP as unphysiologically high concentrations are needed to observe quenching of 2AP fluorescence; the K_d is ~3 mM (Table 1). Thus, the wild-type T4 and RB69 DNA polymerases discriminate against binding dCTP by a factor of at least ~100 compared to dTTP (30 μM compared to 3 mM).

In contrast, the Y567A-DNA polymerase is significantly less able to discriminate against binding dCTP opposite template 2AP (Table 1). With the A+T-rich DNA substrate, the K_d for dCTP binding at ~25 μM was <3-fold higher than the K_d for dTTP binding at ~9 μM, which indicates that dCTP is bound opposite template 2AP almost as easily as dTTP. Fluorescence quench curves are shown in Figure 6. An increased level of discrimination, however, was observed with the G+C-rich DNA; the K_d for dCTP binding at ~317 μM was >20-fold higher than the K_d for dTTP binding at ~15 μM (Table 1). Even though an increased level of discrimination was observed with the G+C-rich DNA against binding dCTP-Mg^{2+}, the K_d of ~317 μM is still well below the value of >3 mM 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 ~35 μM is lower than the apparent K_d ($K_{d,app}$) for the overall reaction for dTMP incorporation (367 μM) (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 ~9 μM compared to the $K_{d,app}$ of ~16 μM. The $K_{d,app}$ for incorporation of dCMP by the Y567A-DNA polymerase was higher than the K_d values for dCTP binding with both DNA substrates, but highest for the G+C-rich DNA; the $K_{d,app}$ for incorporation of dCMP at 2 mM approached the values of ~3 mM 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).

dx.doi.org/10.1021/bi2014618|*Biochemistry* 2011, 50, 10136–10149

Table 1. Kinetic Parameters for Binding and Incorporation of Nucleotides opposite Template 2AP with "Long" DNA Substrates

	A+T-rich (DNA 1)				G+C-rich (DNA 2)			
DNA pol	K_{dbind}(dTTP) (μM)	k_{pol}(dTTP)	K_{dbind}(dCTP)	k_{pol}(dCTP)	K_{dbind}(dTTP) (μM)	k_{pol}(dTTP)	K_{dbind}(dCTP)	k_{pol}(dCTP)
T4 exo$^-$	35.0 ± 1.2[a]	rapid[a]	>3 mM[a,b]	na[c]	28.7 ± 0.6	rapid	~3 mM[b]	na[c]
RB69 exo$^-$	31.1 ± 0.9	–	~3 mM[b]	na[c]	30.0 ± 0.4	–	~3 mM[b]	na[c]
RB69 Y567A exo$^-$	9.2 ± 0.8	–	24.8 ± 2.1 μM	nd[d]	15.1 ± 1.4	–	317 ± 18 μM	nd[d]

	A+T-rich (DNA 1)				G+C-rich (DNA 2)			
DNA pol	$K_{dincorp,app}$(dTMP) (μM)	k_{pol}(dTMP) (s^{-1})	$K_{dincorp,app}$(dCMP)	k_{pol}(dCMP)[e] (s^{-1})	$K_{dincorp,app}$(dTMP) (μM)	k_{pol}(dTMP)[e] (s^{-1})	$K_{dincorp,app}$(dCMP)	k_{pol}(dCMP) (s^{-1})
T4 exo$^-$	367 ± 36[a]	165 ± 3[a]	>3 mM[a,b]	2.3 ± 0.1[a]	nd	nd	nd	nd
RB69 Y567A exo$^-$	16 ± 3	240 ± 11	170 ± 50 μM	6.4 ± 0.6	180 ± 75	190 ± 11	2 mM[b]	~6

	A+T-rich (DNA 1)		G+C-rich (DNA 2)	
DNA pol	$k_{pol}/K_{dincorp}$(dTMP) (μM^{-1} s^{-1})	$k_{pol}/K_{dincorp}$(dCMP) (μM^{-1} s^{-1})	$k_{pol}/K_{dincorp}$(dTMP) (μM^{-1} s^{-1})	$k_{pol}/K_{dincorp}$(dCMP) (μM^{-1} s^{-1})
T4 exo$^-$	0.45[f]	<0.0008	nd	nd
RB69 Y567A exo$^-$	15	0.04	1	~0.003

	dTMP/dCMP	
DNA pol	A+T-rich (DNA 1)	G+C-rich (DNA 2)
T4 exo$^-$	>560	nd
RB69 Y567A exo$^-$	375	333

[a]Data from ref 13. Rapid nucleotide binding occurs within the dead time of the stopped-flow instrument. [b]The high apparent K_d value likely includes nonspecific interactions, and thus, the true K_d may be higher. [c]Not applicable. The rate of nucleotide binding cannot be determined. [d]Not determined. [e]Overall nucleotide incorporation rate. [f]The $k_{pol}/K_{dincorp}$(dTMP) for incorporation of dTMP opposite template A is 10.3 μM^{-1} s^{-1}.[13]

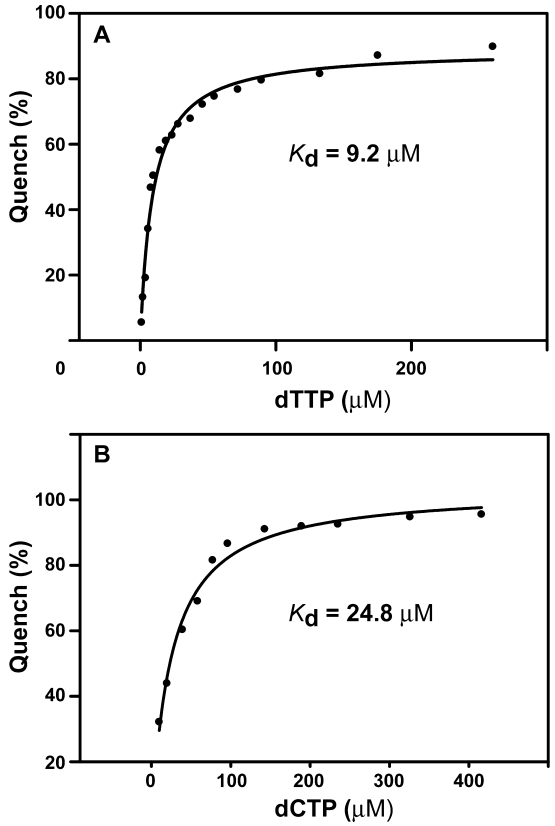


Figure 6. Nucleotide binding assays. (A) Formation of ternary DNA pol–DNA–dTTP complexes with the Y567A-DNA polymerase as a function of dTTP concentration. Incremental increases in dTTP concentration produce incremental increases in fluorescence quenching. (B) Formation of ternary DNA pol–DNA–dCTP complexes with the Y567A-DNA polymerase as a function of dCTP concentration.

The experiments described 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 [DNAs 3 and 4 (Figure 2)] 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) than for the G+C-rich DNA (~373 μM). $K_{d,app}$ 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 values of >2 mM 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 the Ability To Discriminate against Binding rUTP opposite Template 2AP. Because the Y567A-DNA polymerase is less able 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 (Table1), 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 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

Table 2. Kinetic Parameters for Binding and Incorporation of Nucleotides opposite Template 2AP with "Short" DNA Substrates

	A+T-rich (DNA 3)				G+C-rich (DNA 4)			
DNA pol	K_{dbind}(dTTP)	k_{pol}(dTTP)	K_{dbind}(dCTP)	k_{pol}(dCTP)	K_{dbind}(dTTP)	k_{pol}(dTTP)	K_{dbind}(dCTP)	k_{pol}(dCTP)
RB69 Y567A exo$^-$	nd[a]	nd	53 ± 11 μM	nd	nd	nd	373 ± 88 μM	nd

	A+T-rich (DNA 3)				G+C-rich (DNA 4)			
DNA pol	$K_{dincorp,app}$(dTMP) (μM)	k_{pol}(dTMP) (s^{-1})	$K_{dincorp,app}$(dCMP)	k_{pol}(dCMP)[b] (s^{-1})	$K_{dincorp,app}$(dTMP) (μM)	k_{pol}(dTMP) (s^{-1})	$K_{dincorp,app}$(dCMP)	k_{pol}(dCMP)[b] (s^{-1})
RB69 exo$^-$	114 ± 26	379 ± 28	>2 mM[c]	slow[c]	200 ± 45	201 ± 16	>2 mM[c]	slow[c]
RB69 Y567A exo$^-$	67.3 ± 8	289 ± 9	298 ± 65 μM	9.3 ± 0.7	263 ± 78	233 ± 27	361 ± 52 μM	2.5 ± 0.1

	A+T-rich (DNA 3)		G+C-rich (DNA 4)	
DNA pol	$k_{pol}/K_{dincorp}$(dTMP) (μM^{-1} s^{-1})	$k_{pol}/K_{dincorp}$(dCMP) (μM^{-1} s^{-1})	$k_{pol}/K_{dincorp}$(dTMP) (μM^{-1} s^{-1})	$k_{pol}/K_{dincorp}$(dCMP) (μM^{-1} s^{-1})
RB69 exo$^-$	3.3	<0.001−0.003	1.0	<0.001−0.003
RB69 Y567A exo$^-$	4.3	0.03	0.9	0.007

	dTMP/dCMP	
DNA pol	A+T-rich (DNA 3)	G+C-rich (DNA 4)
RB69 exo$^-$	>1000	>1000
RB69 Y567A exo$^-$	143	129

[a]Not determined. [b]Overall nucleotide incorporation rate. [c]The kinetic parameters for the misincorporation of dCMP opposite template 2AP cannot be determined accurately because of nonspecific interactions; however, the apparent K_d is >2 mM, and incorporation is slow, on the order of 2−6 s^{-1}.

Table 3. K_d Values for Binding dUTP and rUTP opposite Template 2AP[a]

DNA pol	K_d(dUTP) (μM)	K_d(rUTP)
T4 exo$^-$	74.1 ± 2	>3 mM
RB69 Y567A exo$^-$	21 ± 1	522 ± 55 μM

[a]Experimental conditions are the same as those for Table 1.

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 to 3.19 Å for the A+T-rich DNA and from 2.24 to 2.32 Å for the G+C-rich DNA (Figure 7). The R_{free} values 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 N2 of 2AP and 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 8A−D). The bases are nearly coplanar.

Several H-bonded structures have been proposed for the 2AP-C base pair (Figure 1B−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 hydrogen bonding between N1 of 2AP and N3 of cytosine (Figure 8E,F), and thus, 2AP-C base pairs involving protonated bases (Figure 1C,D) are unlikely. Instead, apparent nonideal H-bonds are observed. One possible H-bond is between a proton of the C4 amino group of cytosine and N1 of

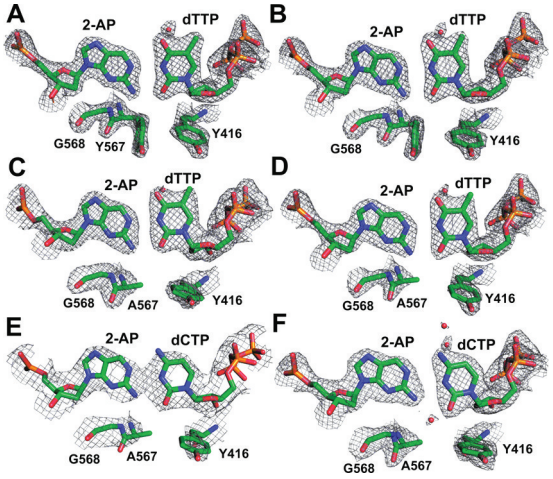


Figure 7. Close-up views of nascent base pairs formed in the polymerase active sites of the wild-type RB69 and RB69 Y567A-DNA polymerases: (A) 2.09 Å resolution structure of the wild-type DNA polymerase−A+T-rich DNA 3−dTTP ternary complex contoured at 2.0σ, (B) 2.24 Å resolution structure of the wild-type DNA polymerase−G+C-rich DNA 4−dTTP ternary complex contoured at 2.0σ, (C) 2.44 Å resolution structure of the Y567A-DNA polymerase−A+T-rich DNA 3−dTTP ternary complex contoured at 1.5σ, (D) 2.25 Å resolution structure of the Y567A-DNA polymerase−G+C-rich DNA 4−dTTP ternary complex contoured at 2.0σ, (E) 3.19 Å resolution structure of the Y567A-DNA polymerase−A+T-rich DNA 3−dCTP ternary complex contoured at 1.0σ, and (F) 2.32 Å resolution structure of the Y567A-DNA polymerase−G+C-rich DNA 4−dCTP ternary complex contoured at 1.5σ.

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 O2 of cytosine.

A rigid H-bond network involving the γ-OH groups of Y567, Y416, and T622, four ordered water molecules, and the

dx.doi.org/10.1021/bi2014618|*Biochemistry* 2011, 50, 10136−10149

Table 4. Data Collection and Refinement Statistics for Ternary Complexes

	dTTP vs 2AP				dCTP vs 2AP	
	AT-rich (wt)	GC-rich (wt)	AT-rich (Y567A)	GC-rich (Y567A)	AT-rich (Y567A)	GC-rich (Y567A)
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
unit cell (Å)						
a	74.95	74.68	74.84	74.40	76.01	74.76
b	119.88	120.47	120.99	120.07	121.60	120.36
c	130.43	130.99	127.37	130.89	125.20	130.82
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)	50.0−3.19 (3.30−3.19)	50.0−2.32 (2.36−2.32)
no. of unique reflections	64604	54679	41995	52104	18719	48550
redundancy	3.9 (3.6)	4.1 (4.0)	3.9 (3.8)	4.0 (4.0)	3.7 (3.7)	4.0 (4.0)
completeness (%)	98.5 (97.0)	99.8 (99.9)	99.9 (100.0)	96.1 (95.9)	97.1 (97.7)	99.8 (100.0)
R_{merge} (%)	11.6 (95.2)	11.3 (78.0)	9.8 (89.9)	8.9 (81.9)	7.9 (76.9)	9.8 (91.4)
I/σ	11.0 (0.8)	12.4 (2.0)	11.8 (1.1)	13.4 (1.6)	16.5 (1.2)	13.1 (1.1)
			Final Model			
no. of amino acid residues	903	903	903	903	903	903
no. of water molecules	350	285	162	254	0	263
no. of Ca^{2+} ions	5	5	5	5	5	5
no. of template nucleotides	18	18	18	18	18	18
no. of primer nucleotides	13	13	13	13	13	13
no. of dNTPs	1	1	1	1	1	1
R (%)	20.2 (30.5)	20.2 (36.2)	20.7 (28.9)	20.9 (41.9)	23.2 (36.2)	20.3 (24.7)
R_{free} (%)	24.3 (33.9)	23.8 (42.6)	25.5 (34.1)	25.4 (47.4)	30.0 (39.5)	26.0 (30.1)
root-mean-square deviation						
bond lengths (Å)	0.008	0.008	0.008	0.008	0.006	0.008
bond angles (deg)	1.122	1.116	1.138	1.132	1.035	1.124
PDB entry	3SQ2	3SQ4	3SUN	3SUO	3SUQ	3SUP

dx.doi.org/10.1021/bi2014618 |*Biochemistry* 2011, 50, 10136−10149

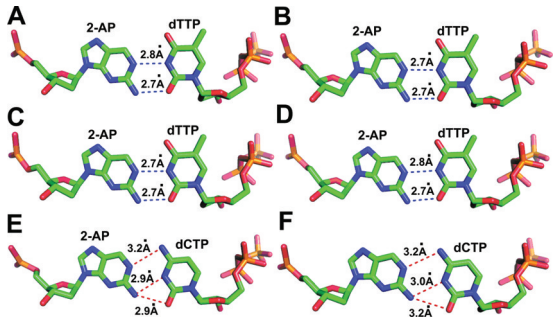


Figure 8. Predicted H-bonds between 2AP-T and 2AP-C base pairs formed in the polymerase active sites of the wild-type RB69 and RB69 Y567A-DNA polymerases. Interbase distances are consistent with two H-bonds formed for the 2AP-T base pair as proposed (Figure 1A). One H-bond is between N1 of 2AP and N3 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 A+T-rich DNA 3 (A) and G+C-rich DNA 4 (B) and for complexes formed with the Y567A-DNA polymerase and A+T-rich DNA 3 (C) and G+C-rich DNA 4 (D). A wobble-type 2AP-C base pair was observed for complexes formed with the Y567A-DNA polymerase, dCTP, and the A+T-rich DNA 3 (E) and the G+C-rich DNA 4 (F). The apparent nonideal 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 N3 as well as O2 of dCTP are shown as red dashed lines.

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-bond network was disrupted by the

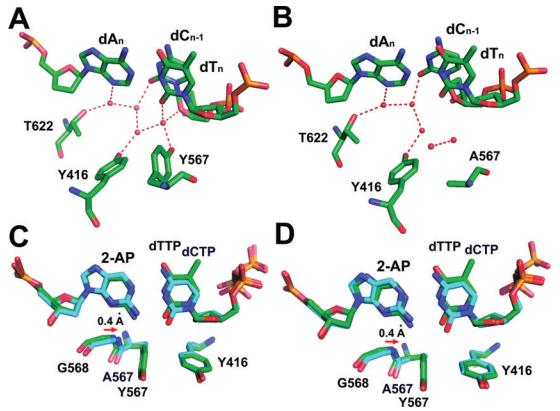


Figure 9. Flexible nucleotide binding pocket produced by the Y567A substitution. The H-bond network in the nucleotide binding pocket formed with the RB69 wild-type DNA polymerase, G+C-rich DNA 4, and dTTP (A) and the disrupted H-bond network formed with the Y567A-DNA polymerase, G+C-rich DNA 4, and dCTP (B). A similar disrupted H-bond network was observed with dTTP. Superimposition of the nucleotide binding pockets of the wild-type and Y567A-DNA polymerases with bound dTTP or dCTP with the G+C-rich DNA 4 (C) or A+T-rich DNA 3 (D), respectively. Note that residue G568 is shifted 0.4 Å toward Y416 and that the distance between N3 of 2AP and G568 increases in the complex formed with the Y567A-DNA polymerase.

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,D). In addition, the distance between N3 of 2AP and the α-CH group of G568 was increased to 3.9 Å via replacement of Y567 with alanine compared to 3.5 Å observed for the wild-type DNA polymerase. The disruption of the H-bond network and the more flexible nucleotide binding pocket caused by the Y567A substitution are predicted to increase the level of accommodation of the 2AP-C base pair.

DISCUSSION

2AP has a long history of providing insights into the fidelity of DNA replication from Freese's observations that 2AP is a potent base substitution mutagen[8] to Drake's observations that antimutator phage T4 DNA polymerases reduce the level of 2AP mutagenesis while mutator DNA polymerases increase the level of mutagenesis[45] 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 the extent of proofreading. Here we provide new information about the ability of the RB69 DNA polymerase to form 2AP-dTTP and 2AP-dCTP base pairs within the polymerase active site, the role of A+T and G+C richness of the primer-terminal region in formation of the mutagenic 2AP-dCTP base pair, and 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 at position 1 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 an increased level of formation of less fluorescent binary pol complexes [complex II (Figure 4)] or an increased level of formation of binary exo complexes [complex VII (Figure 4)], which are also only weakly fluorescent when formed with DNA substrates labeled with 2AP at position 1 of the template strand. An increased level of formation of exo complexes was ruled out because the Y567A-DNA polymerase is less able to form exo complexes compared to the wild-type DNA polymerase (Figure 5A,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 showed that the Y567A-DNA polymerase has an enhanced 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 three-way equilibrium between complex I/VI (pretranslocation) and complex II (post-translocation, nucleotide preinsertion), which prepares for nucleotide incorporation; between complex I/VI and 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 three-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 ϕ29 DNA polymerase.[48] The Y567A

dx.doi.org/10.1021/bi2014618|*Biochemistry* 2011, 50, 10136−10149

substitution disrupted the H-bond network, and the space left by substitution of alanine with tyrosine was replaced by two water molecules (Figure 9A,B). Because translocation to form complex II requires transient release of minor groove interactions, the disrupted H-bond 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 in the weakened 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 Cα hydrogen of G568 and the N3 hydrogen acceptor of purines.[47] Because minor groove interactions are proposed in general to be important for determining the accuracy of nucleotide binding,[49] the increased distance between G568 and N3 of 2AP, from 3.5 Å in the wild-type DNA polymerase to 3.9 Å in the mutant (Figure 9C,D), is expected to contribute to the observed reduction in base selectivity. Note, however, that weakened minor groove interactions produced by the Y567A substitution did not substantially reduce the level of 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 of 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 the level of partitioning of the primer terminus from the exonuclease to the polymerase active site and reduces the level of 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 three-way competition among elongation (complex II), proofreading (complex VII), and dissociation, but theY567A-DNA polymerase appears to favor formation of the post-translocated/preinsertion 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 the level of 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 preinsertion site appears to more readily accommodate dCTP to form complex III if the primer-terminal region is A+T-rich 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, refs 7, 23, 50, and 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 polymerase primer−template substrates. More local strand separation is predicted for A+T-rich 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 an increased level of breathing in the primer-terminal region may take place in the context of the polymerase active site. For an increased level of breathing to facilitate dCTP binding opposite template 2AP to form complex III, either the reverse equilibrium with complex II is weakened or the forward equilibrium with complex IV is strengthened (Figure 4).

Higher fidelity is also reported for *E. coli* DNA pol II with G+C-rich 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 level of 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 ~170 μM with the long, A+T-rich DNA to ~2 mM with the long, G+C-rich DNA (Table 1), but similar values of ~300−360 μM 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 three-way kinetic competition model if enzyme dissociation is more likely with the shorter DNA substrates. The level of binding of T4 DNA polymerase 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 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 the effects of DNA sequence on the accuracy of nucleotide incorporation can be detected with the long DNA substrates, but increased levels of dissociation may mask these effects with the shorter DNAs.

For all reactions, the k_{pol} rates were low, <9 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 approximately 0.03−0.04 μM^{-1} s^{-1} for the longer and shorter A+T-rich DNA substrates and approximately 10-fold lower for the longer and shorter G+C-rich DNA substrates (0.003−0.007 μM^{-1} s^{-1}). Note that an increased level of 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, an ~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. Nonideal H-bonds are observed between template 2AP and the incoming dCTP (Figure 8E,F), and dCTP is tilted ~5° toward the primer-terminal base (ddA) (data not shown). Low rates of 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 binding of dCTP 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 (refs 15 and 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 O2 and N3 of cytosine. Free energy calculations have not been conducted 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, incorporation of dTMP opposite template 2AP was favored 20−25-fold over incorporation of dCMP.[55] We observed a higher apparent level of discrimination with the RB69 Y567A-DNA polymerase from >100-fold with the shorter DNAs (Table 2) to >300-fold with the longer DNAs (Table 1), which may reflect differences in the DNA polymerases and/or in the enzyme assays; however, a much higher level of 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 previous experiments predicted.

The major significance of our studies, however, is not in identifying the mutagenic 2AP-dCTP base pair but in demonstrating the importance of dynamic interactions between a DNA polymerase and 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 binding of nucleotide 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 an increased level of formation of post-translocated complexes is correlated with an increased level of extension of mismatched primer ends.

AUTHOR INFORMATION

Corresponding Author

*L.J.R.-K.: telephone, (780) 492-5383; fax, (780) 494-9234; e-mail, linda.reha-krantz@ualberta.ca. W.K.: e-mail, william.konigsberg@yale.edu.

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.K. interpreted data and contributed to writing the paper.

Funding

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.K.).

ACKNOWLEDGMENTS

We thank Likui Zhang and Neil Johnson for helpful discussions.

ABBREVIATIONS

2AP, 2-aminopurine; dd, 2′,3′-dideoxynucleotide; dFTP, difluorotoluene deoxynucleoside triphosphate; exo, exonuclease; $K_{d,app}$, apparent K_d; pol, polymerase.

REFERENCES

(1) Tsai, Y.-C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry 45*, 9675−9687.

(2) Moran, S., Ren, R. X.-F., and Kool, E. T. (1997) A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence specificity. *Proc. Natl. Acad. Sci. U.S.A. 94*, 10506−10511.

(3) Chiaramonte, M., Moore, C. L., Kincaid, K., and Kuchta, R. D. (2003) Facile polymerization of dNTPs bearing unnatural base analogues by DNA polymerase α and Klenow fragment (DNA polymerase I). *Biochemistry 42*, 10472−10481.

(4) Hariharan, C., Bloom, L. B., Helquist, S. A., Kool, E. T., and Reha-Krantz, L. J. (2006) dynamics of nucleotide incorporation; snapshots revealed by 2-aminopurine fluorescence studies. *Biochemistry 45*, 2836−2844.

(5) Joyce, C. M., and Benkovic, S. J. (2004) DNA polymerase fidelity: Kinetics, structure and checkpoints. *Biochemistry 43*, 14317−14324.

(6) Rothwell, P. J., and Waksman, G. (2007) A pre-equilibrium before nucleotide binding limits fingers subdomain closure by Klentaq1. *J. Biol. Chem. 282*, 28884−28892.

(7) Datta, K., Johnson, N. P., and von Hippel, P. H. (2010) DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A. 107*, 17980−17985.

(8) Freese, E. (1959) The specific mutagenic effect of base analogues on phage T4. *J. Mol. Biol. 1*, 87−105.

(9) Hopkins, R. L., and Goodman, M. F. (1979) 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. 135*, 1801−1822.

(10) Bessman, M. J., Muzyczka, N., Goodman, M. F., and Schnaar, R. L. (1974) 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. 88*, 409−421.

(11) Bessman, M. J., and Reha-Krantz, L. J. (1977) Studies on the biochemical basis of spontaneous mutation. V. Effect of temperature on mutation frequency. *J. Mol. Biol. 116*, 115−123.

(12) Bloom, L. B., Otto, M. R., Eritja, R., Reha-Krantz, L. J., Goodman, M. F., and Beechem, J. M. (1994) Pre-steady-state kinetic analysis of sequence-dependent nucleotide excision by the 3′-exonuclease activity of bacteriophage T4 DNA polymerase. *Biochemistry 33*, 7576−7586.
(13) Fidalgo da Silva, E., Mandal, S. S., and Reha-Krantz, L. J. (2002) Using 2-aminopurine fluorescence to measure incorporation of incorrect nucleotides by wild type and mutant bacteriophage T4 DNA polymerases. *J. Biol. Chem. 277*, 40640−40649.
(14) Law, S. M., Eritja, R., Goodman, M. F., and Breslauer, K. J. (1996) Spectroscopic and calorimetric characterization of DNA duplexes containing 2-aminopurine. *Biochemistry 35*, 12329−12337.
(15) Sowers, L. C., Boulard, Y., and Fazakerley, G. V. (2000) Multiple structures for the 2-aminopurine-cytosine mispair. *Biochemistry 39*, 7613−7620.
(16) Sherer, E. C., and Cramer, S. J. (2001) Quantum chemical characterization of the cytosine: 2-Aminopurine base pair. *J. Comput. Chem. 22*, 1167−1179.
(17) Hogg, M., Cooper, W., Reha-Krantz, L., and Wallace, S. S. (2006) Kinetics of error generation in homologous B-family DNA polymerases. *Nucleic Acids Res. 34*, 2528−2535.
(18) Bebenek, A., Dressman, H. K., Carver, G. T., Ng, S., Petrov, V., Yang, G., Konigsberg, W. H., Karam, J. D., and Drake, J. W. (2001) Interacting fidelity defects in the replicative DNA polymerase of bacteriophage RB69. *J. Biol. Chem. 276*, 10387−10397.
(19) Graziewicz, M. A., Bienstock, R. J., and Copeland, W. C. (2007) 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. 16*, 2729−2739.
(20) Beckman, J., Wang, M., Blaha, G., Wang, J., and Konigsberg, W. H. (2010) Substitution of Ala for Tyr567 in RB69 DNA polymerase allows dAMP to be inserted opposite 7,8-dihydro-8-oxoguanine. *Biochemistry 49*, 8554−8563.
(21) Xia, S., Wang, M., Lee, H. R., Sinha, A., Blaha, G., Christian, T., Wang, J., and Konigsberg, W. (2011) 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. 406*, 558−570.
(22) Reha-Krantz, L. J. (2010) DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochim. Biophys. Acta 1804*, 1049−1063.
(23) Ward, D. C., Reich, E., and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem. 244*, 1228−1237.
(24) Frey, M. A., Sowers, L. C., Millar, D. P., and Benkovic, S. J. (1995) 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 34*, 9185−9192.
(25) Mandal, S. S., Fidalgo da Silva, E., and Reha-Krantz, L. J. (2002) Using 2-aminopurine fluorescence to detect base unstacking in the template strand during nucleotide incorporation by the bacteriophage T4 DNA polymerase. *Biochemistry 41*, 4399−4406.
(26) Hariharan, C., and Reha-Krantz, L. J. (2005) Using 2-aminopurine fluorescence to detect bacteriophage T4 DNA polymerase-DNA complexes that are important for primer extension and proofreading reactions. *Biochemistry 44*, 15674−15684.
(27) Rachofsky, E. L., Seibert, E., Stivers, J. T., Osman, R., and Ross, J. B. A. (2001) Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine. *Biochemistry 40*, 957−967.
(28) Jean, J. M., and Hall, K. B. (2001) 2-Aminopurine fluorescence quenching and lifetimes: Role of the base stacking. *Proc. Natl. Acad. Sci. U.S.A. 98*, 37−41.
(29) Neely, R. K., Daujotyte, D., Grazulis, S., Magennis, S. W., Dryden, D. T. F., Klimasauskas, S., and Jones, A. C. (2005) Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes. *Nucleic Acids Res. 33*, 6953−6969.
(30) Beechem, J. M., Otto, M. R., Bloom, L. B., Eritja, R., Reha-Krantz, L. J., and Goodman, M. F. (1998) Exonuclease−polymerase active site partitioning of primer−template DNA strands and equilibrium Mg^{2+} binding properties of bacteriophage T4 DNA polymerase. *Biochemistry 37*, 9095−9103.
(31) Tleugabulova, D., and Reha-Krantz, L. J. (2007) Probing DNA polymerase-DNA interactions: Examining the template strand in exonuclease complexes using 2-aminopurine fluorescence and acrylamide quenching. *Biochemistry 46*, 6559−6569.
(32) Subuddhi, U., Hogg, M., and Reha-Krantz, L. J. (2008) 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 47*, 6130−6137.
(33) Shamoo, Y., and Steitz, T. A. (1999) Building a replisome from interacting pieces: Sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell 99*, 155−166.
(34) Hogg, M., Aller, P., Konigsberg, W., Wallace, S. S., and Doublié, S. (2004) Crystallographic snapshots of a replication DNA polymerase encountering an abasic site. *EMBO J. 23*, 1483−1493.
(35) Hogg, M., Aller, P., Konigsberg, W., Wallace, S. S., and Doublié, S. (2007) 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. 282*, 1432−1444.
(36) Franklin, M. C., Wang, J., and Steitz, T. A. (2001) Structure of the replication complex of a pol α family DNA polymerase. *Cell 105*, 657−667.
(37) Zhang, H., Beckman, J., Wang, J., and Konigsberg, W. (2009) RB69 DNA polymerase mutants with expanded nascent base-pair-binding pockets are highly efficient but have reduced base selectivity. *Biochemistry 48*, 6940−6950.
(38) Marquez, L. A., and Reha-Krantz, L. J. (1996) 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. 271*, 28903−28911.
(39) Reha-Krantz, L. J. (2009) The use of 2-aminopurine fluorescence to study DNA polymerase function. In DNA Replication Methods and Protocols. *Methods Mol. Biol. 521*, 381−396.
(40) Navaza, J. (2001) Implementation of molecular replacement in AMoRe. *Acta Crystallogr. D53*, 240−255.
(41) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2216−2132.
(42) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240−255.
(43) DeLano, W. L. (2002) *The PyMol Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
(44) Yang, G., Franklin, M., Li, J., Lin, T. C., and Konigsberg, W. (2002) A conserved Tyr residue is required for sugar selectivity in a pol α DNA polymerase. *Biochemistry 41*, 10256−10261.
(45) Drake, J. W., Allen, E. F., Forsberg, S. A., Preparata, R.-M., and Greening, E. O. (1969) Spontaneous mutation. *Nature 221*, 1128−1132.
(46) Greive, S. J., and von Hippel, P. H. (2005) Thinking quantitatively about transcriptional regulation. *Nat. Rev. Mol. Cell Biol. 6*, 221−232.
(47) Wang, M., Xia, S., Blaha, G., Steitz, T. A., Konigsberg, W. H., and Wang, J. (2011) Insights into base selectivity from the 1.8 Å resolution structure of an RB69 DNA polymerase ternary complex. *Biochemistry 50*, 581−590.
(48) Berman, A. J., Kamtekar, S., Goodman, J. L., Lazaro, J. M., deVega, M., Blanco, L., Salas, M., and Steitz, T. A. (2007) Structures of ϕ29 DNA polymerase complexes with substrate: The mechanisms of translocation in B-family polymerases. *EMBO J. 26*, 3496−3505.
(49) Seeman, N. D., Rosenberg, J. M., and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A. 73*, 804−808.

(50) Hochstrasser, R. A., Carver, T. E., Sowers, L. C., and Millar, D. P. (1994) Melting of a DNA helix terminus within the active site of a DNA polymerase. *Biochemistry 33*, 11971−11979.

(51) Jose, D., Datta, K., Johnson, N. P., and von Hippel, P. H. (2009) Spectroscopic studies of position-specific DNA "breathing" fluctuations at replication forks and primer-template junctions. *Proc. Natl. Acad. Sci. U.S.A. 106*, 4231−4236.

(52) Wang, Z., Lazaeov, E., O'Donnell, M., and Goodman, M. F. (2002) 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. 277*, 4446−4454.

(53) Wang, F., and Yang, W. (2009) Structural insight into translesion synthesis by DNA pol II. *Cell 139*, 1279−1289.

(54) Delagoutte, E., and von Hippel, P. H. (2003) Function and assembly of the bacteriophage T4 DNA replication complex: Interactions of the T4 DNA polymerase with various model DNA constructs. *J. Biol. Chem. 278*, 25435−25447.

(55) Watanabe, S. M., and Goodman, M. F. (1982) 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. 79*, 6429−6433.