DNA polymerase β nucleotide-stabilized template misalignment fidelity depends on local sequence context

Michael J. Howard, Nisha A. Cavanaugh, Vinod K. Batra, David D. Shock, William A. Beard, and Samuel H. Wilson

From the Genome Integrity and Structural Biology Laboratory, NIEHS, National Institutes of Health, Research Triangle Park, North Carolina 27709

Edited by Patrick Sung

DNA polymerase β has two DNA-binding domains that interact with the opposite sides of short DNA gaps. These domains contribute two activities that modify the 5′ and 3′ margins of gapped DNA during base excision repair. DNA gaps greater than 1 nucleotide (nt) pose an architectural and logistical problem for the two domains to interact with their respective DNA termini. Here, crystallographic and kinetic analyses of 2-nt gap-filling DNA synthesis revealed that the fidelity of DNA synthesis depends on local sequence context. This was due to template dynamics that altered which of the two template nucleotides in the gap served as the coding nucleotide. We observed that, when a purine nucleotide was in the first coding position, DNA synthesis fidelity was similar to that observed with a 1-nt gap. However, when the initial templating nucleotide was a pyrimidine, fidelity was decreased. If the first templating nucleotide was a cytidine, there was a significantly higher probability that the downstream template nucleotide coded for the incoming nucleotide. This dNTP-stabilized misalignment reduced base substitution and frameshift deletion fidelities. A crystal structure of a binary DNA product complex revealed that the cytidine in the downstream template nucleotide to occupy the coding position. These results indicate that DNA polymerase β can induce a strain in the DNA that modulates the position of the coding nucleotide and thereby impacts the identity of the incoming nucleotide. Our findings demonstrate that “correct” DNA synthesis can result in errors when template dynamics induce coding ambiguity.

To maintain faithful DNA synthesis, DNA polymerases have evolved to select a dNTP from a pool of structurally similar molecules that preserve Watson–Crick base-pairing with the coding template base (i.e. the template base immediately downstream of the growing primer terminus). This is achieved through enzyme–DNA interactions that position the polymerase active site in the proper register with the nascent base pair occluding the active site. Accordingly, subsequent to nucleotide insertion, the enzyme must translocate downstream by 1 nt prior to the next insertion to maintain the proper register. Because the insertion reaction is reversible and enzyme translocation does not occur, PPi, can drive the reaction in reverse (i.e. pyrophosphorolysis), removing the primer terminus obstructing the active site.

DNA repair intermediates include short gapped DNA substrates. These short gaps are generally utilized by DNA polymerases from the X family. These enzymes bind avidly to the 5′ margin of the gap, utilizing an 8-kDa DNA-binding domain (Fig. S1). This domain includes a helix–hairpin–helix motif that interacts with the DNA backbone via a monovalent cation in a sequence-independent manner. DNA polymerase (pol) β, an X family member, is a critical enzyme of the base excision repair pathway and well-suited for filling short DNA gaps (1–5 nt) (1). It has also been implicated recently to participate in an error-prone alternative nonhomologous end joining pathway (2). Pol β is the simplest mammalian pol, comprised of an 8-kDa lyase domain and a 31-kDa polymerase domain (Fig. S1A). Each domain contributes an enzymatic activity at the 5′ and 3′ DNA gap termini (deoxyribose phosphate lyase and DNA synthesis, respectively). The amino-terminal lyase domain targets the enzyme to the 5′ margin of a DNA gap (3). Additionally, the helix–hairpin–helix motifs in the lyase and polymerase domains stabilize the noncontiguous DNA strand downstream and upstream of the gap, respectively (Fig. S1B).

The polymerase domain is composed of functionally distinct subdomains: DNA binding (D), catalytic (C), and nascent base pair binding (N) (4). Several global conformational changes occur when pol β binds substrates. The most notable change occurs when the N subdomain of the binary enzyme–DNA complex closes around the nascent base pair upon binding dNTP. This subdomain repositioning is accompanied by subtle protein side-chain and DNA adjustments (5).

Pol β has moderate fidelity that is both DNA sequence– and DNA gap size–dependent (6). For a single-nucleotide gap, pol β misinserts 1 nt for 104-107 insertions (7). Previous studies
have shown that X family pols (e.g. pol λ and pol μ) exhibit low frameshift fidelity (8–10). Pol β can bypass an abasic site by using the downstream nucleotide (11) and can reposition a templating pyrimidine, but not purine, from a coding template position (12). These attributes suggest that template ambiguity exists at the polymerase active site and that the incoming nucleotide may preferentially utilize a downstream template nucleotide rather than the templating nucleotide nearest the primer terminus. This has been referred to as dNTP-stabilized misalignment (13). Pol μ can utilize this strategy during nonhomologous end joining (14). If the downstream nucleotide codes for correct nucleotide insertion, then the mutagenic result would be either deletion of a single nucleotide or an apparent base substitution error when realignment of the primer terminus occurs, permitting insertion of an additional nucleotide (i.e. dislocation) (15). In this study, we characterize the sequence dependence of the fidelity of 2-nt gap filling and structurally capture a frameshift intermediate.

**Results**

**Fidelity of 2-nt gap-filling DNA synthesis**

The catalytic efficiency of single-nucleotide insertion catalyzed by pol β with 1- and 2-nt gapped DNA substrates in different sequence contexts was measured (Fig. 1 and Tables S1–S4). In general, the catalytic efficiency of inserting the cor-

---

**Figure 1. Discrimination plots for 2-nt gap-filling DNA synthesis by pol β.** Steady-state catalytic efficiencies \( k_{cat}/K_m,\text{dNTP} \) are plotted on a log scale so that the distance between the efficiencies of correct and incorrect insertions is a measure of discrimination (29). The first column in each plot illustrates the catalytic efficiencies and discrimination for 1-nt gap filling; the templating base in the gap is indicated. The other columns show the catalytic efficiencies and discrimination for 2-nt gap filling; the templating sequence in the gap is indicated. A cartoon of the 2-nt gapped DNA substrate is shown above each plot, highlighting the identity of the first templating base and showing that the downstream template nucleotide \( (N) \) was varied. A, the identity of the first templating nucleotide is deoxyadenosine \( (A) \). B, the identity of the first templating nucleotide is deoxycytidine \( (C) \). C, the identity of the first templating nucleotide is deoxyguanosine \( (G) \). D, the identity of the first templating nucleotide is thymidine \( (T) \). The data are tabulated in Tables S1–S4, respectively. The identities of the incoming nucleotides are dATP (red), dCTP (green), dGTP (black), and dTTP (blue).
rect nucleotide with the 2-nt gapped DNA substrate compared with 1-nt gapped DNA substrate is only modestly altered when a purine occupies the first templating position (AN and GN, where N represents an alternate downstream templating nucleotide). In contrast, the catalytic efficiency for correct insertion decreases 4- to 9-fold for CN gaps and 11- to 40-fold for TN sequences as compared with the 1-nt gapped DNA substrates.

To determine whether the fidelity of nucleotide insertion catalyzed by pol β is altered with the 2-nt gapped DNA substrates, we measured the catalytic efficiency for nucleotide misinsertion. This analysis revealed a striking trend for 2-nt gapped substrates containing a CN sequence where the catalytic efficiency for misinsertion increased significantly (100-fold) and depended on the identity of the downstream templating base. The misinsertion efficiency increased for the incoming nucleotide that was complementary to the downstream templating base but not for those that were not. For instance, the catalytic efficiency for misinsertion increased significantly (100-fold) and depended on the identity of the downstream templating base. However, this appeared to be primarily due to the loss in efficiency of correct dAMP insertion. In contrast, when the initial templating nucleotide was a purine, the catalytic efficiencies for correct insertion in a 2-nt gap were not significantly altered compared with the 1-nt gap. These data indicate that pol β fidelity with gapped substrates greater than 1 nt is sequence-dependent. Furthermore, the observation that the incoming dNTP can utilize the coding potential of the downstream templating base suggests that a dNTP-stabilized template misalignment intermediate is formed that alters fidelity.

**Influence of lyase domain interactions on fidelity**

In the crystal structure of pol β bound to 1-nt gap DNA, the lyase domain interacts with downstream duplex DNA (16). To ascertain whether these interactions are important for the fidelity of DNA synthesis, we altered these interactions and measured the fidelity within the CG sequence. If the lyase domain anchors pol β to the 5’-phosphate in the gap and influences the coding template register, then reducing or eliminating these interactions should increase fidelity (i.e. reduce dNTP-stabilized misalignment). To alter 5’-phosphate binding, we examined the fidelity of a triple-lysine mutant, where the lysines that coordinate the 5’-phosphate (Fig. 51C) were substituted with alanine (K35A/K68A/K72A (3KA)) (17). The catalytic efficiency for dGTP insertion with the 2-nt CG gapped substrate was reduced ~6-fold for the mutant compared with the WT enzyme, whereas dCTP misinsertion decreased ~3-fold (Figs. 1B and 2B and Tables S1 and S5). These results suggest that interactions with the lyase domain and the 5’-phosphate in the gap are not fully responsible for dNTP-stabilized misalignment synthesis.

Another approach to modulate downstream interactions with duplex DNA was to remove the downstream oligonucleotide to eliminate the gapped DNA structure. This substrate is referred to as CG-open, where the first two templating nucleotides are C and G, respectively (Fig. 2A). Removal of the downstream oligonucleotide modestly increases the fidelity of dCTP misinsertion from 250 (CG 2-nt gap) to 480 (CG-open). This is due to an 18-fold decrease in efficiency for the incorrect insertion (dCTP) coupled with a 9-fold decrease in correct (dGTP) insertion relative to a 2-nt gapped DNA substrate (Figs. 1B and 2B and Tables S2 and S5). These results indicate that 5’-phosphate and downstream duplex interactions have very little effect on dNTP-stabilized misalignment.

**RB69exo⁻ fidelity of the CG-open substrate**

To determine whether properties inherent to the CG sequence give rise to the observed dNTP-stabilized misalignment synthesis, we examined the fidelity of the B family replicative pol. This polymerase serves as a high-fidelity model replicative polymerase and does not make specific interactions with gapped DNA. We measured the catalytic efficiency for dNTP insertion using an exonuclease-deficient mutant with the CG-open sub-
Mutagenically correct nucleotide insertion

Figure 3. Single-turnover analysis of mutagenic 2-nt gap-filling DNA synthesis. Single-turnover primer extension was assayed and quantified as described under “Experimental procedures.” A, illustration of the 5′-FAM–labeled 2-nt gapped DNA substrate. The sequence in the gap is 3′-CG-5′, and dCTP was added to initiate DNA synthesis. B, gel image illustrating the time course for 2-nt gap-filling (shown above). The substrate primer band (S) is extended by two dCMP insertions (+2 band). An intermediate band (+1) accumulates at short time intervals (<10 s). C, plot of the time course for the product bands (+1 and +2 as well as their sum, Total P). The progress curves were fit to either a single-exponential (+2 time course: \( A = -0.93, k = 0.07 s^{-1}, C = 0.89 \); Total P time course: \( A = -0.88, k = 0.30 s^{-1}, C = 0.91 \)) or double-exponential equation (+1 time course: \( A_1 = -0.78, k_1 = 0.49 s^{-1}, A_2 = 0.75, k_2 = 0.09 s^{-1}, C = 0.05 \)). D, illustration of the 5′-FAM–labeled 2-nt gapped DNA substrate. The sequence in the gap is 3′-GC-5′, and dGTP was added to initiate DNA synthesis. E, gel image illustrating the time course for 2-nt gap-filling (shown above). The substrate primer band (S) is extended by at least two dGMP insertions (P). An intermediate band can be observed that accumulates at later intervals. F, plot of the time course for all product bands (Total P). The progress curve was fitted to a single-exponential equation (\( A = -0.90, k = 0.0007 s^{-1}, C = 0.90 \)).

To further explore the mechanistic basis of the enhanced catalytic efficiency with the CN substrates, dCTP insertion with a CG 2-nt gapped DNA substrate was measured under single-turnover conditions (i.e. pol/DNA = 10; Fig. 3). These reactions were performed with saturating concentrations of enzyme and dCTP so that substrate binding was not rate-limiting. The observed rate constant (\( k_{\text{obs}} \)) for dCTP insertion with a 2-nt gapped DNA was significantly higher than the corresponding rate constant for misinsertion into a 1-nt gapped DNA with a template C (0.3 s\(^{-1}\) and 0.009 s\(^{-1}\), respectively; Fig. S2A, left plot). In contrast, \( k_{\text{obs}} \) for primer extension of a GC 2-nt gapped DNA substrate with dGTP was comparable with that observed with a 1-nt gapped DNA substrate (template G; Fig. S2A, right plot). These data are consistent with the sequence-dependent effects observed in the steady-state reactions and indicate that the incoming nucleotide can base-pair with the downstream templating nucleotide in a CG 2-nt gap nucleotide as opposed to direct misincorporation of dCTP opposite C.

Crystalllographic structures of 2-nt gapped DNA intermediates

To provide molecular insight into the interactions of pol β with 2-nt gapped DNA substrates, crystals of binary enzyme–DNA complexes were grown with DNA substrates that included different sequence contexts in a 2-nt gap. To create the 2-nt gap, the primer strand was 1 nt shorter (9-mer) than that used to create a 1-nt gap (10-mer). The nucleotide complementary to the second template base in the gap was soaked into the crystal to form a ternary substrate complex and initiate DNA synthesis. Not surprisingly, the structures revealed that the incoming nucleotide formed a Watson–Crick base pair with the downstream templating nucleotide. The structures were nearly identical to that observed with a 1-nt gapped ternary substrate complex, except that the primer strand was 1 nt shorter. Thus, there was no indication that the 2-nt gapped DNA substrate was altered to hasten nucleotide insertion (i.e. the shortened primer strand was too far from the incoming nucleotide). For example, soaking dCTP into a crystal of a DNA binary complex with two G residues in the gap resulted in a closed ternary substrate complex (Table S6 and Fig. S3, A and B). The
A

Figure 4. Structure of a stable binary pol β-DNA complex with an extra-
helical templating cytosine. A, pol β ribbon representation of the binary pol
β-DNA complex with an extrahelical templating cytosine (purple) superim-
posed with an open 1-nt gapped DNA substrate (gray; PDB code 3ISB)
(18). The root mean square deviation over 320 Cα is 0.54 Å. B, DNA structure of
the resulting product after insertion of dCMP into a 2-nt gap (5′-CG-3′).
The inserted cytosine (yellow) forms a stable base pair with the downstream tem-
plate G (orange) but is distant from α-helix N (purple ribbon), which has repo-
sitioned itself to an open state following dCMP insertion. The substrate
primer terminus base pair is colored green, and the extrahelical cytosine is
also yellow. The inset illustrates the sequence alignment, color code, and gen-
eral oligonucleotide arrangement.

B

incoming dCTP is base-paired with the downstream G in the gap,
but because the primer strand was 1 nt shorter, O3′ of the primer
terminus was too far (~7 Å) from Po of the incoming nucleotide,
precluding DNA synthesis. Because pol β can efficiently fill 2-nt
gaps, the structure represents a trapped precatalytic intermediate
because of crystallographic constraints.

As an alternative approach, a new 2-nt gapped DNA sub-
strate was created with a longer template (17-mer rather
than the 16-mer used to create the 1-nt gapped DNA sub-
strate). This approach conserves the upstream and down-
stream duplex nature of the DNA substrate used for the 1-nt
gap substrate. In addition, the structure was cocrystallized
with all necessary components. The sequence in the gap
was CG, and dCTP was included. The resulting structure was a
“nick” product complex after a single dCMP insertion (Fig.
4). The first templating C was in an extrahelical position to
permit the downstream G to occupy the coding template
position. The global pol conformation was open compared
with the open conformation of the binary DNA complex
with a 1-nt DNA gap (18) (Fig. 4A; root mean square deviation
= 0.54 Å). The active-site Mg2+ ions and PPi had disso-
ciated, and the nascent GC base pair was severely buckled,
weakening the hydrogen bonds between these complemen-
tary bases (Fig. 4B). The poor density for the extrahelical C
indicates that this nucleoside is highly dynamic and lacks
stabilizing enzyme interactions (Fig. S4).

dNTP-stabilized misalignment leading to a base substitution
error

For the 2-nt gapped DNA substrate with the CT sequence
context, dATP rapidly filled the gap (i.e. two dAMP insertions)
(Fig. 5). In this case, the n+1 product accumulated only to a
minor extent, suggesting that the downstream T served as the
coding template base less often than that observed in the CG
sequence context (Fig. 3C). The modestly elevated rate of inser-
tion, compared with direct misinsertion into a 1-nt gap (Fig.
S2B), is consistent with only a proportion of the product being
generated through dATP-stabilized misalignment. Subsequent
to dATP-stabilized misalignment, the primer terminus must
rapidly realign to create a mismatched primer terminus (primer

Figure 5. Single-turnover analysis of mutagenic 2-nt gap-filling DNA
synthesis. Single-turnover primer extension was assayed and quantified as
described under “Experimental procedures.” A, illustration of the 5′-FAM-
labeled 2-nt gapped DNA substrate. The sequence in the gap is 3′-CT-5′, and
dATP was added to initiate DNA synthesis. B, gel image illustrating the time
course for 2-nt gap-filling (shown above). The substrate primer band (5′) is
extended by two dAMP insertions (+2 band). A weak intermediate band (+1)
is also observed. C, plot of the time course for the product bands (+1 and +2
as well as their sum, Total P). The progress curves were fit to either a single-
exponential (+2 time course: A = −0.84, k = 0.23 s−1, C = 0.82; Total P time
course: A = −0.82, k = 0.28 s−1, C = 0.82) or double-exponential equation
(+1 time course: A1 = −0.07, k1 = 2.1 s−1, A2 = 0.06, k2 = 0.10 s−1, C = 0.02).
The inset expands the time course for the n+1 band.
Mutagenically correct nucleotide insertion

terminus A opposite C) that is extended with a second dAMP insertion opposite T. This creates nicked DNA with a matched 3′-terminus. If this DNA product is ligated, it will generate a base substitution error. In contrast to the product structure determined for the CG gapped DNA substrate, crystals were not formed with the CT gapped DNA, probably because of the dynamic nature of events with this sequence context and the constraints necessary to form crystals.

Discussion

DNA polymerases must select the correct incoming nucleotide that is complementary to the templating base to maintain high fidelity. An essential aspect of this selection is that the proper template base must direct nucleotide insertion. In reiterated DNA sequences (e.g. mono-, di-, and trinucleotide repeats), polymerases can be error-prone because of DNA slippage events that realign nucleotides in the primer stem. Because this occurs in runs of a reiterated sequence, correct base pairing is maintained, albeit with fewer base pairs, and generally leads to frameshift insertion or deletion errors. In contrast to upstream slippage events that require DNA melting and annealing, downstream DNA template dynamics can alter the proper register of the single-stranded template strand that can confound proper identification of the coding template nucleotide. It has been noted previously that pols can utilize the downstream template nucleotide to direct nucleotide insertion, and this has been referred to as dNTP-stabilized misalignment (13, 15). This event has been clearly demonstrated for pol β when it encounters an abasic site centrally positioned in a 5-nt DNA gap (11). Rather than insert dATP opposite an abasic site, as observed for most pols, pol β utilizes the downstream templating base to direct DNA synthesis. If DNA synthesis continues, then this results in a −1 deletion. If on the other hand, the template strand realigns, then further DNA synthesis would generate a base substitution error (Fig. 6).

In this study, we characterized the fidelity of 2-nt gap-filling DNA synthesis by pol β to ascertain which nucleotides in the template strand may be prone to repositioning and thereby impact fidelity. Previous studies with pol β indicated that an adenine mispair at the primer terminus can displace a templating cytosine (19), but not adenine (20), in 1-nt gapped DNA. In the former case, the templating cytosine is moved to an extra-helical position. Similarly, spectroscopic analysis of pol β binding to 1-nt DNA gaps indicated that addition of divalent metals can form an active enzyme when pyrimidines serve as the templating nucleotides, but not purines (12). To accomplish this, the templating pyrimidine must be displaced to permit the primer terminus to move into the nascent base pair binding pocket. This behavior correlates with a robust reverse reaction, pyrophosphorolysis, observed with a template pyrimidine in the gap. Because the reverse reaction requires the primer terminus to occupy the nascent base pair binding pocket, the templating pyrimidine in the gap must be removed from the coding template position.

The results presented here indicate that the fidelity of 2-nt gap filling is highly dependent on the sequence in the gap. The distance between the respective insertion efficiencies plotted in Fig. 1 is a representation of the fidelities for the apparent misinsertions and provides insight into the origin of the altered discrimination. In general, when the first templating nucleotide in the gap is a purine (A or G), then the catalytic efficiency for correct nucleotide insertion is hardly affected when going from a 1- to a 2-nt gap (Fig. 1, A and C). Likewise, the catalytic efficiencies for misinsertion are not significantly affected by gap size; accordingly, fidelity is not significantly altered.

In contrast, when the first templating base in a 2-nt gap is a pyrimidine (C or T), then there is a significant decrease in fidelity. This is the result of two changes: a general decrease in the efficiency of correct insertion when a 2-nt gap is utilized and an increase in catalytic efficiency of the incoming incorrect nucleotide when it is complementary to the downstream templating base (Fig. 1, B and D), suggesting that the downstream templating nucleotide provided the coding base (i.e. dNTP-stabilized misalignment; Fig. 6, path ii) (13). This behavior was most pronounced in the CG gap. The fidelity of insertion of dCTP opposite C in a 1-nt gap is 290,000 ($(k_{cat}/K_m)_{dCTP}$)/$(k_{cat}/K_m)_{dCTP}$, but in the 2-nt gap it is 250 (1160-fold lower). This is primarily due to a 290-fold increase in the insertion efficiency of dCTP in the CG gap (Table S2). These results are consistent with a dCTP-stabilized misalignment where the downstream G in the CG gap is utilized as the coding template base. Additionally, the stability of this nascent base pair is consistent with the rapid accumulation of primer extended by 1 nt (Fig. 3, B and C), indicating that a step after initial insertion is slow. The crystallographic structure of the binary product complex after insertion of a single dCMP (Fig. 4) is also consistent with a stable misaligned intermediate resulting in a slow primer realignment step. The structure reveals that the first templating C is extrahelical and dynamic (Fig. S4). If this intermediate is ligated, then the result would be a −1 deletion mutation (Fig. 6, path ii, ligation). In contrast to the extrahelical displacement of C observed here, pol μ preferentially utilizes the downstream templating nucleotide in a 2-nt gap even when the first templating nucleotide is a purine (14). This structure is similar to that of pol β with 2 G residues in the gap (Fig. S3B). In the pol μ structure, however, the primer terminus is well-positioned for catalysis through compression of the templating nucleotides near the active site. This strain in the template strand is relieved after nucleotide insertion.

Over the past 20 years, high-resolution crystal structures of pol β with gapped DNA substrates have been determined with 16-mer templates. This presumably optimizes crystal packing because attempts to crystallize pol β with longer substrates have been unsuccessful. The 2-nt CG-gapped 17-mer substrate utilized here produced a low-resolution structure by flipping the cytosine in the original gap outside of the DNA helix, generating a semistable 16-mer template mimic. The low resolution of the complex probably reflects the stress introduced into the active site with this premutagenic DNA intermediate. In contrast, a stable 16-mer intermediate mimic could not be formed with the CT substrate. This interpretation is consistent with the
kinetics of gap-filling DNA synthesis with these two substrates and failure to solve a structure with the CT 2-nt gap substrate. In contrast to the dCTP-stabilized misalignment in the CG context described above, the dATP-stabilized misalignment in the CT context does not result in a stable product complex (Fig. 5). Because the rate of dAMP insertion into a 1-nt gap with a templating C is similar to that observed for misinsertion into the CT 2-nt gap (Fig. S2B), the poor accumulation of the n nucleotide product is consistent with slow misinsertion followed by rapid extension. A closer examination of the results illustrated in Fig. 5 suggests that both misinsertion and dATP-stabilized misalignment are occurring during CT-gap filling. The time course for the n+1 band indicates that a population of enzymes can rapidly insert dAMP ($k_{obs} = 2 \text{ s}^{-1}$) into the 2-nt CT gap (Fig. 5C). In this instance, however, the amplitude (i.e. the population of enzymes utilizing the downstream T residue, ~10%) is smaller than that observed in Fig. 3 for the CG gap (n+1 band amplitude ~80%). Additionally, the decay of this intermediate suggests that rapid primer realignment and correct insertion occur to fill the 2-nt gap.

The high rate of dAMP misinsertion opposite C or into the CT 2-nt gap is also noteworthy (Fig. S2B). A crystallographic ternary substrate complex structure with a templating C and incoming dATP analog indicates that the template C residue is removed from the active site and positioned upstream upon nucleotide binding, effectively creating an abasic site (21). This is consistent with the relatively robust misinsertion rate relative to other misinsertions and similar to the rate observed for dAMP insertion opposite an authentic abasic site (18).

Taken together, the results suggest that the ability of the first templating base to stack with upstream DNA and/or form van der Waals contact with pol β influences template dynamics, which, in turn, will dictate which path will prevail (Fig. 6). In the
Mutagenically correct nucleotide insertion

case of a dNTP-stabilized misalignment (Fig. 6, path ii), the stability of the nascent base pair will influence the competition between subsequent steps (ligation versus realignment). Although the influence of such intermediates on DNA ligase efficiency has not been examined in detail, these intermediates can be ligated after pol β-dependent synthesis on short gaps (9).

Although DNA synthesis fidelity is compromised during 2-nt gap filling relative to that observed in a 1-nt gap when the first templating nucleotide is a pyrimidine, fidelity is not significantly altered when a purine occupies the first templating position. As suggested above, this is probably due to stabilizing interactions: stacking with the primer terminus base pair and hydrogen bonding with the incoming nucleotide. In this situation, the downstream templating nucleotide must be accommodated either in an extrahelical position or in an alternative structure. Ternary substrate complex structures of template strand is bent 90° at the coding templating nucleotide, stabilized either in an extrahelical position or in an alternative conformation. As suggested above, this is probably due to stabilizing interactions: stacking with the primer terminus base pair and hydrogen bonding with the incoming nucleotide. In this situation, the downstream templating nucleotide must be accommodated either in an extrahelical position or in an alternative binding pocket. The first possibility seems unlikely because the template strand is bent 90° at the coding templating nucleotide, occluding this region. Ternary substrate complex structures of pol λ, a close relative of pol β, with a 2-nt gap indicate that the downstream noncoding template nucleotide is sequestered in a pocket situated between the polymerase and lyase domains (22). This was termed template scrunching (Fig. 6, path i). However, for longer DNA gaps, other strategies must be utilized to handle a greater number of single-stranded template nucleotides. In that situation, perhaps the downstream duplex is partially melted to permit the template strand to follow its normal trajectory. The single-stranded complementary strand could extend beyond the 5′-phosphate–binding site because there are positively charged residues in this region that could provide a good binding pocket for the negatively charged backbone of the displaced downstream strand (Fig. S5). This general scenario is consistent with the minimal effect of altering interactions with the downstream duplex on catalytic efficiency and fidelity (Fig. 2).

X family pols interact with duplex DNA on opposite sides of short DNA gaps. For pol β, the amino-terminal lyase domain avidly binds to the downstream 5′-phosphate through three lysine residues. Removing these lysines by alanine substitution (i.e., the 3KA mutant) or removing the downstream oligonucleotide did not significantly alter the infidelity of dCTP misinsertion in the CG context (Fig. 2 and Table S5). The replicative β family RB69exo− pol exhibited high fidelity in this sequence context. Although the pol β lyase domain also has single-stranded DNA–binding activity (23), these results suggest that the incoming nucleotide base plays an important role in directing template choice, in conjunction with the nature of the first templating base.

Historically, pol fidelity or discrimination has been characterized in terms of selecting the correct dNTP from a pool of structurally similar molecules. Here we examine the ability to select the correct templating nucleotide. In this scenario, if the wrong templating nucleotide is used to select the incoming nucleotide, then correct nucleotide selection may lead to a base substitution or deletion error. Template base selection errors appear to be limited to short gap-filling rather than replicative pols.

Experimental procedures

DNA substrates

DNA oligonucleotides were from Integrated DNA Technologies (Coralville, IA) and were either HPLC-purified (6-carboxyfluorescein (FAM)–labeled) or gel-purified. The sequences of the gapped DNA substrates were as follows: 14-mer primer, 5′-CTG CAG CTG ATG CG-3′; 18-mer downstream oligonucleotide, 5′-GTA CGG ATC CCC GGG TAC-3′; 34-mer template, 3′-GAC GTC GAC TAC GCN NCA TGC CTA GGG GCC CAT G-5′; N represents A, C, G, or T. Oligonucleotides were resuspended in 10 mM Tris-HCl (pH 7.4) and 1 mM EDTA, and the concentrations were determined by UV absorbance at 260 nm using the extinction coefficients provided by Integrated DNA Technologies. 2-nt gapped DNA substrates were prepared by annealing primer with 10% excess downstream oligonucleotide and template (i.e., 1:1:1:1.1 molar ratio, respectively) and stored in 20 mM Tris-HCl (pH 7.4) and 50 mM KCl. The annealing reactions were performed in a thermal cycler by heating at 95 °C for 5 min followed by cooling to 10 °C (1 °C/min).

Enzymes

WT and mutant pol β were prepared as described previously (24). The concentration was determined by absorbance at 280 nm using an extinction coefficient of 23,380 M⁻¹ cm⁻¹. RB69exo− was kindly provided by William Konigsberg (Yale).

Steady-state kinetic assays

For correct nucleotide insertion, steady-state kinetic parameters for gap-filling reactions were determined by initial velocity measurements as described previously (4). Unless noted otherwise, enzyme activities were determined using a standard reaction mixture containing 50 mM Tris-HCl (pH 7.4), 100 mM KCl, 5 mM MgCl₂, 1 mM DTT, 0.1 mg/ml BSA, 200 nM gapped DNA, and 10% glycerol at 37 °C. Enzyme concentrations and reaction time intervals were chosen so that substrate depletion or product inhibition did not influence initial velocity measurements. Reactions were initiated by addition of enzyme and quenched with 2× quench buffer (200 mM EDTA, 80% formamide, ~0.1% bromphenol blue, and 0.1% xylene cyanol). Reaction mixtures were then heated at 95 °C for 5 min, placed on ice, and loaded onto prerun 22% denaturing polyacrylamide gels. The gels were scanned using a Typhoon scanner and quantified using ImageQuant TL (GE Healthcare Life Sciences). Steady-state kinetics parameters and standard errors represent the results of least squares curve fitting of duplicated measurements to either the Michaelis–Menten equation (k_cat and K_m) or a modified form to extract catalytic efficiencies (k_cat/K_m). The catalytic efficiencies for most misinsertion reactions were determined by using sub-saturating concentrations (25–200 μM) of nucleotides. When [dNTP] ≪ K_m, then k_obs ≈ (k_cat/K_m)[dNTP].

Single-turnover assay

To follow complete 2-nt gap filling, single-turnover assays (enzyme:DNA) were performed with 1 μM pol β, 50 nM DNA, and 1 mM dNTP in standard reaction buffer employing a KinTek
Model RQF-3 chemical quench-flow apparatus (KinTek Corp., State College, PA). Samples were quenched, resolved, and quantified as described above. Progress curves were fit to either a single-exponential ($k_{\text{obs}} = A \times \exp(-kt) + C$) or double-exponential ($k_{\text{obs}} = A_1 \times \exp(-k_1t) + A_2 \times \exp(-k_2t) + C$) equation, as indicated, where $A$ is the amplitude, $k$ the rate constant, and $C$ the end point.

Crystallography

Ternary complex crystals with 2-nt gapped DNA were obtained by soaking the DNA binary complex crystals with dCTP as described earlier (25, 26). The sequences of the template strand (16-mer), primer strand (9-mer), and phosphorylated downstream primer strand (5-mer) were 5′-CCG AGC GCC CAT CAG C-3′, 5′-GCT GAT GCG C-3′, and 5′-pGTC GG-3′, respectively. The nucleotides in the gap are underlined. The oligonucleotides were annealed at a ratio of 1:1:1 by heating at 90 °C for 10 min and cooling to 4 °C (1 °C/min) using a PCR thermocycler, resulting in a 1 mM mixture of 2-nt gapped DNA. This annealed mixture was crystallized by sitting drop vapor diffusion at 18 °C by mixing 2 M PEG3350, 350 mM sodium acetate, and 50 mM imidazole (pH 7.5).

Binary product DNA complex crystals were obtained by cocystalization of annealed oligonucleotides, creating 2-nt gapped DNA, pol β, and dCTP. The sequences of the template strand (17-mer), primer strand (10-mer), and phosphorylated downstream primer strand (5-mer) were 5′-CCG AGC GCC GCA TCA GC-3′, 5′-GCT GAT GCG C-3′, and 5′-pGTC GG-3′, respectively. The nucleotides in the gap are underlined.

Data were collected at 100 K on a charge coupled device detector system mounted on a MiraMax®-007HF (Rigaku Corp.) rotating anode generator. Data were integrated and reduced with HKL2000 software (27). All crystals belong to the space group $P2_1$. The structures were solved by molecular replacement using PDB codes 4KLM (binary) and 2FMS (ternary) as reference models. The structures were refined using PHENIX (28), and manual model building was done using Coot. The crystallographic statistics are reported in Table S6.


Acknowledgments—We thank Lars Pedersen and Andrea Kamiński of the X-ray crystallography core laboratory at NIEHS, National Institutes of Health for discussions. Molecular graphics images were produced using the Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by National Institutes of Health Grant P41 GM-103311).

References

20. Batra, V. K., Beard, W. A., Pedersen, L. C., and Wilson, S. H. (2016) Structures of DNA polymerase mispaired DNA termini transitioning to pre-
Mutagenically correct nucleotide insertion


