Probing T7 DNA Polymerase with Variously Sized Base Pairs

Tae Woo Kim, Luis G. Brieza, Tom Ellenberger, and Eric T. Kool

From the Department of Chemistry, Stanford University, Stanford, California 94305 and the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Hypotheses on the origins of high fidelity in replicative DNA polymerases have recently focused on the importance of geometric or steric effects in this selectivity. Here we reported a systematic study of the effects of base pair size in T7 DNA polymerase (pol), the replicative enzyme for bacteriophage T7. We varied base pair size in very small (0.25 Å) increments by use of a series of nonpolar thymidine shape mimics having gradually increasing size. Steady-state kinetics were evaluated for the Sata exonuclease-deficient mutant in a 1:1 complex with thioredoxin. For T7 pol, we studied insertion of natural nucleotides opposite variably sized T analogs in the template and, conversely, for variably sized dTTP analogs opposite natural template bases. The enzyme displayed extremely high selectivity for a specific base pair size, with drops in efficiency of as much as 280-fold for increases of 0.4 Å beyond an optimum size approximating the size of a natural pair. The enzyme also strongly rejected pairs that were smaller than the optimum by as little as 0.3 Å. The size preferences with T7 DNA pol were generally smaller, and the steric rejection was greater than DNA pol I Klenow fragment, correlating with the higher fidelity of the former. The hypothetical effects of varied active site size and rigidity are discussed. The data lend direct support to the concept that active site tightness is a chief determinant of high fidelity of replicative polymerases and that a less rigid (looser) and larger active site can lead to lower fidelity.

The DNA polymerase encoded by gene 5 of the T7 phage has long been studied as a model for DNA replication in general (1–3). The T7 polymerase works together with primase-helicase and single-stranded binding protein to replicate the circular 40-kb genome of the phage. To replicate a genome of this size, the polymerase recruits an Escherichia coli protein, thioredoxin, which forms a tight 1:1 complex with the polymerase (pol) (4), and strongly increases processivity (5). The T7 pol-thioredoxin complex (abbreviated as T7 pol here) is a high efficiency and high fidelity polymerase of the A family, sharing homology and structural features with other enzymes of this family such as DNA pol I and Taq polymerase (6). The polymerase synthesizes DNA processively at a rate of ~500 nucleotides/s, and typically proceeds several thousand nucleotides before dissociating (5).

Another important feature of T7 DNA polymerase is its high fidelity, which is necessary for conservation of the phage genome as it is copied. Crystal structures of the enzyme bound to DNA and dNTP substrates have shown a high packing density of the protein around the DNA being synthesized (7), resembling other A family enzyme structures (8, 9). In general, it has been hypothesized that a “tight” fit by high fidelity polymerases around the incipient base pair might help enforce a correct Watson–Crick geometry and aid in rejecting geometrically incorrect mismatches (10–14). However, testing this geometric and steric concept with standard DNA bases and their close analogs can be complicated by the fact that they have widely varying sizes, shapes, and functional groups.

We recently developed a set of nucleoside analogs that are designed to test steric effects in a systematic way, with incremental size differences over a 1.0 Å range (15). These molecules are all thymidine analogs (Fig. 1), and their bases retain the general shape of natural thymine, but all are nonpolar and lack hydrogen bonding ability (16). The sizes of the analogs are varied in sub-angstrom increments by use of the atomic series H, F, Cl, Br, and I. Of these five, the F-substituted analog (difluorotoluene (17, 18)) is a nearly exact shape and size mimic for natural thymine, whereas the H analog is smaller, and the others are progressively larger. An initial study testing these analogs with the Klenow fragment of E. coli DNA pol I (Kf exo−) showed that this enzyme was highly sensitive to size changes across this series (19). Most interestingly, the most efficiently processed analog was the dichloro compound (L), which forms base pairs with adenine that are larger than a natural pair by ~0.6 Å. It was suggested that tightness, as defined by active site rigidity, plays a strong role in determining the fidelity of that enzyme. However, it is unknown whether other polymerases use this same mechanism. Moreover, it remains to be seen whether enzymes with higher or lower fidelity than Kf would exhibit correspondingly different active site steric responses in association with these differences in fidelity.

Finally, it remains unclear what the best definition of polymerase “tightness” should be: should it refer to active site rigidity (i.e. ability to move (or to resist moving) in response to steric properties of a substrate), or should it refer to a static size (i.e. larger or smaller) preference for its optimum substrate?

The present study is aimed at addressing these issues. Here we investigate the effects of varied DNA substrate size on the synthesis of DNA by the high fidelity T7 DNA polymerase. We find that, like Kf exo−, the T7 enzyme is highly sensitive in kinetic efficiency to size changes across this analog series. However, distinct from Kf, the T7 enzyme shows an...
even stronger dependence on substrate size and also shows a preference for smaller substrates. We discuss the functional mechanism by which this enzyme achieves high fidelity, and we also introduce an experimental strategy to measure individual contributions of the two chief parameters of polymerase tightness, namely active site rigidity and active site size.

MATERIALS AND METHODS

Thymidine Analogs—The dNTP derivatives of the thymidine analogs were prepared as described (19). Phosphoramidite derivatives of dH, dL, dA, and dI were prepared as described previously (18). The phosphoramidite derivative of analog dF was purchased from Glen Research.

Oligodeoxynucleotides—DNA oligonucleotides were synthesized on an Applied Biosystems 394 synthesizer using standard β-cyanoethylphosphoramidite chemistry. Oligomers were synthesized in DMT-off mode, deprotected in concentrated NH₄OH (55 °C, 8 h), purified by preparative 20% denaturing PAGE, isolated by the crush and soak method, and quantitated by absorbance at 260 nm. Molar extinction coefficients were calculated by the nearest neighbor method (20). Values for oligonucleotides containing non-natural residues were obtained in the following way. The molar extinction coefficients at 260 nm for dH, dF, dL, and dI were taken as 250, 1000, 250, 500, and 3900 M⁻¹ cm⁻¹, respectively. The individual extinction coefficients for all the bases in a given oligomer were summed and compared with the sum from the corresponding sequence in which the non-natural residues were replaced by T. Because in most cases the content of non-natural residues in the sequences is low, this estimation method is unlikely to generate large errors in concentration.

Steady-state Kinetics Studies—Primer (final concentration, ~100 nM) was labeled with 5'-[γ-³²P]ATP (catalog number PB01218 from Amersham Biosciences) and T4 polynucleotide kinase (catalog number 18004-010 from Invitrogen) and purified using ethanol precipitation (21). Labeled primer (final concentration, ~20 nM), template, and unlabeled primer were mixed in 5X reaction buffer and gave a final total concentration of primer-template of ~200 nM. The primer-template duplexes were annealed by heating to 95 °C and slow cooling to 4 °C over 1 h.

Steady-state kinetics for single-nucleotide insertions were carried out as described (22). Briefly, insertion reactions were carried out in 10-μl volumes in the reaction buffer: 50 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 50 mM NaCl, and 5 mM dithiothreitol for exonuclease-free T7 5A7A DNA polymerase (3). A 1× stock solution of duplex was mixed with 1× T7 5A7A DNA polymerase for 2 min at 37 °C, and the reaction was initiated by adding a 1× solution of the appropriate dNTP. Concentrations of dNTP ranged from 1 to 50 nM for efficient pairings with a template base and 50–500 nM for inefficient pairings. Enzyme concentration and reaction time were adjusted in different dNTP reactions to give 1–20% incorporation in time periods ≤50 min.

The following enzyme concentrations (nM) and times (min) were used (N → M denotes dNTP inserted across from base M in the template strand): H → A, 40, 30; F → A, 2, 3, 1 → A, 2, 0.5; B → A, 2, 1, 6; I → A, 2, 30; T → A, 2, 0.67; T → G, 40, 1; H → T, 160, 15; F → T, 40, 30; L → T, 40, 4; B → T, 40, 4; I → A, 40, 2; A → H, 2, 5; A → F, 2, 0.83; A → L, 2, 2; A → B, 2, 8; A → I, 40, 2; T → A, 2, 0.67; T → H, 40, 1; T → F, 2, 10; T → L, 40, 3; T → B, 40, 3; T → I, 40, 50; G → T, 40, 3.

Reactions were quenched with 15 μl of loading buffer I (80% formamide, 1× TBE (89 mM Tris borate (pH 8.3), 2 mM EDTA), 0.05% xylene cyanol, and bromphenol blue). Extents of reaction were determined by running quenched reaction samples on a 20% denaturing polyacrylamide gel to separate unreacted primer from insertion products. Radioactivity was quantified using a PhosphorImager (Amersham Biosciences) and the ImageQuant program. Relative velocity v was measured as the ratio of the extended product (Iext) to remaining primer (Iprim) as follows 

\[ v = \frac{I_{ext}}{I_{prim} \times t}, \]

where t represents the reaction time, and normalized for the lowest enzyme concentration used. The apparent \( K_m \) and \( V_{max} \) values were obtained from Hanes-Woolf plots.

RESULTS

To test the effects of small size differences in the incipient base pair on the functioning of T7 DNA polymerase, we employed the series of thymidine analogs shown in Fig. 1. Base pairing studies of these compounds in DNA (in the absence of enzymes) have shown that they all behave similarly, and their pairing properties are consistent with their nonpolar character (16). All five compounds are somewhat destabilizing (compared with a native base pair) when placed opposite adenine in the center of an oligonucleotide duplex, and they show little pairing selectivity with any of the four possible natural partners, with a slight preference for adenine among the four. This preference has been attributed to the stronger stacking ability of adenine among the natural bases. The general lack of selectivity and the similar behavior of these compounds in DNA alone serves as a base line for enzymatic studies; any observed polymerase selectivity in choosing among natural partners, and any differences among the five analogs, would necessarily arise from the influence of the enzyme rather than from the energetics of base pairing alone.

Our initial experiments were carried out qualitatively and were aimed at surveying the effects of the different analogs with all possible natural partners. We employed the 5A7A mutant of T7 DNA polymerase, which has extremely low exonuclease activity (3), in 1:1 complex with thioredoxin. We tested the analogs as both incoming nucleotide (dNTP) derivatives (Fig. 2A) and as template bases (Fig. 2B), and we evaluated single nucleotide additions by gel electrophoresis. All five analogs were active as dNTP or template substrates for the enzyme but clearly required different conditions (see Fig. 2 legend), indicating a wide variation in reaction efficiency. Also clear from the relative intensities of the product bands was that there were clear preferences among natural partners for pairing with the analogs in the polymerase active site. Adenine was by far the preferred replication partner, and this was true whether the thymidine analogs were dNTP derivatives or were present in the template. The second most prevalent products resulted from mispairing of the analogs with thymine. This is different from the natural base thymine, which is commonly mispaired with guanine during enzymatic synthesis of DNA. Overall, the data show considerable symmetry, in that the selectivities for analogs as incoming dNTP derivatives closely resemble those with insertion of natural dNTPs opposite the analogs in the template strand.

To quantify the effects of small size variations in the incipient base pair on efficiency of this enzyme, we carried out steady-state kinetics measurements for single nucleotide insertion reactions with these analogs. The data are assembled in Tables 1 and 2. We measured the kinetics for insertion of the thymidine analogs opposite adenine in the template DNA and for insertion of dATP opposite the variably sized thymine analogs in the template. The corresponding natural base pairings (dTTP opposite A and dATP opposite T) were also measured for comparison.

The results showed that catalytic efficiencies varied markedly across this series. As an incoming nucleotide, the most efficient analog was the dichloro derivative dL.TP. Its efficiency (as \( V_{max}/K_m \)) was only 4-fold below that of the natural congener dTTP. The dL.TP analog was 17-fold more efficient than previously reported for dFTP, which is the closest
size mimic of the natural nucleotide. Activity fell off steeply for larger and smaller analogs, with the largest (dITP) being 320-fold less efficient than the optimum, and the smallest (dHTP) being 6700-fold less efficient than the optimum. For the converse experiments with analogs in the template strand, the most efficient template base was the F analog, which was 14-fold less efficient than T. The smallest analog (H) was 48-fold less efficient than the optimum, whereas the largest (I) was less efficient by 280-fold. These size versus efficiency data are presented as plots in Fig. 3, A and B. The data were further analyzed by plotting $V_{\text{max}}$ and $K_m$ values as a function of size (see supplemental Fig. S1). Results showed that varying the size either of the incoming nucleotide or of the template base had a similar effect on both parameters. $V_{\text{max}}$ values varied strongly and were much greater for the optimally sized analogs (F or L) and dropped off precipitously for smaller and larger analogs. The $K_m$ values stayed relatively constant but showed a clear optimum (minimal $K_m$) for F or L, consistent with more favorable interactions in the active site for those cases.

We also evaluated size effects on selectivity and fidelity in the active site by measuring the kinetics of the most efficient misinsertions, in which the analogs are paired with thymine. We used the most efficient mismatch to measure minimum fidelity (rather than averaging all mismatches) because the most frequent mismatch defines the majority of
Probing T7 DNA pol with Variably Sized Base Pairs

TABLE 1
Steady-state kinetics for single nucleotide insertions with T7 DNA polymerase (exo−)
The conditions used are as follows: 200 nM 23-mer/28-mer primer-template duplex, 2, 40, or 160 nM T7(exo−), 10 mM Tris-HCl (pH 7.5), 1 mM MgCl₂, 10 mM NaCl, and 1 mM dithiothreitol, incubated for 30 s to 30 min at 37 °C in a reaction volume of 10 μl. Standard deviations are given in parentheses.

| Nucleoside triphosphate | Template base (X) | Kₘ | Vₘₐₓ | Efficiency (Vₘₐₓ/Kₘ) | Relative efficiency |
|-------------------------|------------------|----|------|----------------------|--------------------|
| H                       | A                | 47 (11) | 0.022 (0.002) | 4.8 × 10² | 3.7 × 10⁻⁵ |
| F                       | A                | 63 (22) | 11 (3) | 1.9 × 10⁵ | 1.5 × 10⁻² |
| L                       | A                | 12 (4)  | 37 (2) | 3.2 × 10⁶ | 2.5 × 10⁻¹ |
| B                       | A                | 36 (8)  | 14 (2) | 4.1 × 10⁷ | 3.2 × 10⁻³ |
| I                       | A                | 35 (4)  | 0.35 (0.02) | 1.0 × 10⁴ | 7.7 × 10⁻⁴ |
| T                       | A                | 2.2 (0.9) | 25 (1) | 1.3 × 10⁷ | 1 |
| H                       | T                | 400 (470) | 0.036 (0.030) | 9.0 × 10¹ | 6.9 × 10⁻⁶ |
| F                       | T                | 160 (20) | 0.016 (0.002) | 1.0 × 10² | 7.7 × 10⁻⁶ |
| L                       | T                | 33 (5)  | 0.28 (0.03) | 8.6 × 10³ | 6.6 × 10⁻⁴ |
| B                       | T                | 20 (9)  | 0.35 (0.10) | 1.9 × 10⁴ | 1.5 × 10⁻³ |
| I                       | T                | 36 (10) | 0.47 (0.07) | 1.4 × 10⁴ | 1.1 × 10⁻³ |
| T                       | G                | 63 (6)  | 0.95 (0.5) | 1.6 × 10⁴ | 1.2 × 10⁻³ |

Data are normalized for the lowest enzyme concentration used.

TABLE 2
Steady-state kinetics for single nucleotide insertions with T7 DNA polymerase (exo−)
The conditions used are as follows: 200 nM 23-mer/28-mer primer-template duplex, 2 or 40 nM T7(exo−), 10 mM Tris-HCl (pH 7.5), 1 mM MgCl₂, 10 mM NaCl, and 1 mM dithiothreitol, incubated for 40 s to 50 min at 37 °C in a reaction volume of 10 μl. Standard deviations are given in parentheses.

| Nucleoside triphosphate | Template base (X) | Kₘ | Vₘₐₓ | Efficiency (Vₘₐₓ/Kₘ) | Relative efficiency |
|-------------------------|------------------|----|------|----------------------|--------------------|
| A                       | H                | 33 (10) | 36 (1.1) | 1.1 × 10⁵ | 1.5 × 10⁻³ |
| A                       | F                | 4.8 (1.6) | 24 (2) | 5.3 × 10⁶ | 7.2 × 10⁻² |
| A                       | L                | 6.9 (1.6) | 11 (1) | 1.7 × 10⁸ | 2.3 × 10⁻² |
| A                       | B                | 36 (3)  | 3.3 (0.4) | 9.3 × 10⁴ | 1.3 × 10⁻³ |
| A                       | I                | 36 (10) | 0.66 (0.06) | 1.9 × 10⁴ | 2.6 × 10⁻⁴ |
| A                       | T                | 0.59 (0.37) | 26 (1) | 7.4 × 10⁷ | 1 |
| T                       | H                | 180 (70) | 4.1 (1.0) | 2.3 × 10⁴ | 3.1 × 10⁻⁴ |
| T                       | F                | 360 (190) | 1.2 (0.5) | 3.5 × 10⁵ | 4.7 × 10⁻⁵ |
| T                       | L                | 65 (13) | 0.26 (0.03) | 4.0 × 10⁵ | 5.4 × 10⁻⁵ |
| T                       | B                | 74 (14) | 0.28 (0.02) | 3.8 × 10⁵ | 5.1 × 10⁻⁵ |
| T                       | I                | 730 (60) | 0.032 (0.004) | 4.4 × 10⁴ | 5.9 × 10⁻⁷ |
| G                       | T                | 230 (40) | 0.26 (0.03) | 1.1 × 10⁴ | 1.5 × 10⁻⁵ |

Data are normalized for the lowest enzyme concentration used.

mutations in DNA replication. The numerical data are given in Tables 1 and 2. Results with varying dNTP sizes showed that mismatching efficiency increased with size (see plot in supplemental Fig. S2), but there was a broad maximum efficiency with the largest analogs. The selectivity for an adenine-containing template over a thymine template is plotted in Fig. 3C; the maximum is found at the difluoro compound (dFTP).

For fidelity in the converse situation, with thymidine analogs in the template, the current experiments show that the T7 enzyme has a strong preference for specifically sized partners for a given template base. When the incoming nucleotide was varied in size, the optimum (dLTP) was considerably more efficient than analogs only slightly larger or smaller. For example, dFTP (with bond lengths only 0.28 Å shorter) was 16-fold less efficient; dBTP (with bond lengths 0.16 Å longer) was 8-fold less efficient. Further expansion or reduction in size led to even more dramatic differences; decreasing the size from dFTP to dHTP (0.28 Å difference) led to a 400-fold efficiency drop, and increasing it from dBTP to dITP (0.20 Å difference) led to a 41-fold drop. The magnitude of this steric sensitivity is significantly greater than that observed previously for the Kf polymerase (19). For example, a decrease in size from dFTP to dITP caused a 58-fold decrease in the efficiency of Kf, and increasing size from dBTP to dITP led to a 11-fold drop.

A general comparison of the response to varying thymidine analog sizes between T7 pol and Kf polymerases reveals two significant differences in these enzymes (see supplemental Fig. S3 for overlaid plots). First is the somewhat steeper drop in efficiency for substrates that are slightly too large or too small, as mentioned above. We propose that this is best explained by greater rigidity of the T7 polymerase active site in the closed, catalytically active conformation; this is discussed below. The second difference is that the T7 enzyme has a smaller optimum size for an analog in the templating position (supplemental Figs. S3, B and D). This high fidelity enzyme prefers difluorotoluene as a substrate, yielding both higher efficiency and fidelity with this size than the Kf.

DISCUSSION

The current data shed light on the functional basis for the high fidelity of T7 DNA polymerase. Our data confirm the high selectivity for insertion of natural nucleotides; the insertion of dGTP opposite T is less favorable than dATP opposite T by a large factor of 67,000. By comparison, the Klenow fragment of E. coli DNA pol I (Kf) has a fidelity of 1,800 against G → T mismatch synthesis (18). In the T7 enzyme, this high fidelity is accompanied by a high rate of processive synthesis (∼500 nucleotides/s, compared with ∼10 nucleotides/s for Kf) and high processivity (>1,000 versus ∼10 nucleotides for Kf), due in large part to the thioredoxin component (5).
enzyme, which shows a kinetic preference for the larger dichloro analog (19). The difference in bond lengths for these two substrates is 0.38 Å. Thus, the effective active site size for the T7 enzyme appears to be 0.3–0.4 Å smaller than for Kf.

High fidelity DNA polymerases undergo an induced fit conformation change prior to forming the new phosphodiester bond (24, 25). It has been proposed that, on closing of the fingers domain, high fidelity polymerases tightly surround the incipient base pair to closely enforce base pair selectivity, whereas low fidelity repair and lesion bypass polymerases may have more “open” active sites (13, 26, 27) that can accept varied geometries of mismatched or damaged pairs with less energetic cost. However, such general descriptions do not make clear how fidelity is enforced in the steric sense. We hypothesize that there are two chief parameters that might affect the kinetics of base pairing and dNTP insertion, the preferred size of an active site and its rigidity. We illustrate these concepts with two hypothetical plots in which these parameters are varied (Fig. 4).

First we address the effects of different active site sizes on the insertion kinetics for variably sized analogs (Fig. 4A). Here we define “active site” as both the residues that come in contact with the incipient base pair, as well as those parts of the protein that are structurally affected by motions in those residues. One might envision a highly rigid active site that could not easily flex outward or compress inward around different-sized substrates once the polymerase closed around an incipient base pair. A close structural correspondence between the active site and canonical Watson-Crick pairs would then strongly disfavor mispairing of the incoming dNTP in the closed conformation of the polymerase, particularly if DNA substrates projected beyond the boundaries of the binding pocket in any direction. Note that there must also be a mechanism for rejecting base pairs that are too small, such as T-T. Such small pairs might be rejected by the high thermodynamic cost of the void (loss of favorable contacts) left by the too-small substrate. Alternatively, if the nucleobases carried water of solvation (28), such pairing might be sterically rejected as too large (13).

Now we imagine what would be the result if different polymerases, in their closed or most active form, had active sites fixed to different sizes (Fig. 4B). First, using the present series of size-varied analogs, we might see distinctive peaks at different optimal substrate sizes corresponding to more or less room in the active site of each polymerase. This is in fact what we have observed in a comparison of the very high fidelity T7 enzyme with the somewhat lower fidelity Kf enzyme (supplemental Fig. S3 B and D). We also predict that an enzyme with an active site size that is larger than natural Watson-Crick base pairs would show less tendency to reject sterically larger mismatches. Thus polymerases with similar rigidities, but distinct optimal size preferences, could in principle show very different fidelities of templated DNA synthesis with natural bases.

A second factor contributing to the variable fidelities of different
Probing T7 DNA pol with Variably Sized Base Pairs

polymerases might also be envisioned (Fig. 4B), namely varied rigidity. In this mechanism, different polymerases might exhibit the same active site size preferences, as shown by the similar optima in Fig. 4B. For these polymerases, the active site structure surrounding the incipient pair might look the same in a crystallographic snapshot of the closed conformation, with similarly close contacts to the incipient base pair in both high fidelity and low fidelity enzymes. However, if active site rigidity plays a role in substrate selection, a high fidelity enzyme might be more rigid and produce large energetic penalties for flexing outward to relieve a clash or inward to fill a void while maintaining other interactions with the substrate that contribute to catalysis in the closed conformation of the polymerase. Conversely, a low fidelity enzyme might be relatively “loose” or flexible in the region of the base binding pocket, moving outward or inward with much lower energetic cost, thus accepting the varied shapes and sizes of mismatches more readily. Note that differences in rigidity between two or more enzymes (Fig. 4B) would be clearly distinguishable from differences in substrate size preference (Fig. 4A) in plots of size versus kinetic efficiency.

Both mechanisms, active site size and rigidity, may contribute substantially to the fidelity of templated DNA synthesis by T7 pol and other enzymes. The steady-state kinetics data show both a larger size preference of T7 pol relative to Kf (supplemental Figs. S3, B and D) in the template strand, and larger energetic penalties for increasing size and decreasing size for the larger and smaller substrates. Thus size-varied analogs provide functional evidence that the T7 polymerase/thioredoxin holoenzyme has both a somewhat smaller and more rigid active site than the Kf enzyme.

Most interestingly, both the T7 enzyme and Kf show surprisingly high efficiency for base pairs that are larger than canonical T-A pairs. For example, the L-A pair, which has nearly wild-type efficiency with Kf (19) and reasonably high efficiency with T7, is larger than a T-A pair by ~0.5–0.6 Å, due to both the longer bond lengths and the lack of a hydrogen bonding contraction. We have hypothesized (19) that for Kf, this larger active site size may be a mechanism by which Kf (as pol I) contributes to the evolution of E. coli, by allowing a rare but finite mutation rate so that the bacteria can adapt to changing environments. However, the cost of this mechanistic and structural adaptation is that this larger active site size would yield lower efficiency with natural DNA than would be possible if the active site were smaller. DNA polymerase I and its derivative Kf primarily function in short gap synthesis during DNA repair, so the need for rapid synthesis and high processivity may not be pressing. However, as a replicative polymerase, the T7 pol must function in a highly accurate and efficient manner to copy the phage genome (29). We suggest that the smaller active site size preference of T7 pol is a mechanism by which this can be achieved, and this also results in the lower frequency of mismatched base pairings.

The kinetics experiments with the varied size probes provide data that are complementary to that from structural studies. The current probe series offers functional information at a resolution (~0.2 Å) that exceeds that available for published x-ray crystal structures, which have typically been determined at 2 Å resolution or lower. In addition, the kinetic data directly relate to the rate-limiting transition state of the polymerase reaction, whereas structural studies yield information about the ground state and do not directly give measures of energetics.

However, published crystal structures of DNA polymerases give important information about the protein-DNA contacts that are likely to be involved in maintaining steric control over matched versus mismatched (or damaged) pairs. Most interestingly, enzyme-DNA co-crystal structures of A family polymerases and error-prone Y family polymerases show essentially the same DNA C-1’–C-1’ interstrand distances (7–9, 26), and similarly close protein-DNA contacts on opposite sides of the two DNA strands. One of the chief differences between the A and Y family polymerases is the appearance of more space around the incipient base pair in a Y family enzyme (23, 26). We and others have hypothesized that more space might provide for adjustments to the differences in size at lower energetic cost, by allowing some freedom of motion without strong steric penalties. By contrast, the greater density in the closed form of the T7 enzyme (7) may be consistent with the notion of greater rigidity by providing closer packing of chains. Future kinetics studies with size-varied probes, such as those used here, would be useful in determining whether the low fidelity of Y family polymerases is associated with a larger active site functional size preference, or greater flexibility, or both.

REFERENCES

1. Dunn, J. J., and Studier, F. W. (1983) J. Mol. Biol. 166, 477–535
2. Engler, M. J., and Richardson, C. C. (1983) J. Biol. Chem. 258, 11197–11205
3. Patel, S. S., Wong, I., and Johnson, K. A. (1991) Biochemistry 30, 511–525
4. Mark, D. F., and Richardson, C. C. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 780–784
5. Tabor, S., Huber, and H. E., and Richardson, C. C. (1987) J. Biol. Chem. 262, 16212–16223
6. Jung, G. H., Leavitt, M. C., Hsieh, J. C., and Ito, J. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 8287–8291
7. Dublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Nature 391, 251–258
8. Li, Y., Korolev, S., and Waksman, G. (1998) EMBO J. 17, 7514–7525
9. Johnson, S. I., and Beese, L. S. (2004) Cell 116, 803–816
10. Echols, H., and Goodman, M. F. (1991) *Annu. Rev. Biochem.* **60**, 477–511
11. Moran, S., Ren, R. X., and Kool, E. T. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10506–10511
12. Kunkel, T. A., and Behnenk, K. (2000) *Annu. Rev. Biochem.* **69**, 497–529
13. Kool, E. T. (2002) *Annu. Rev. Biochem.* **71**, 191–219
14. Beard, W. A., and Wilson, S. H. (1998) *Chem. Biol.* **5**, R7–R13
15. Kim, T. W., and Kool, E. T. (2004) *Org. Lett.* **6**, 3949–3952
16. Kim, T. W., and Kool, E. T. (2005) *J. Org. Chem.* **70**, 2048–2053
17. Schweitzer, B. A., and Kool, E. T. (1994) *J. Org. Chem.* **59**, 7238–7242
18. Moran, S., Ren, R. X.-F., Rumney, S., and Kool, E. T. (1997) *J. Am. Chem. Soc.* **119**, 2056–2057
19. Kim, T. W., Delaney, J. C., Essigmann, J. M., and Kool, E. T. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15803–15808
20. Cavaluzzi, M. J., and Rorer, P. N. (2004) *Nucleic Acids Res.* **32**, e13
21. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning*, Vol. 2, 3rd Ed., pp. 10.17 and 10.20, Cold Spring Harbor Laboratory Press, Woodbury, NY
22. Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) *Crit. Rev. Biochem. Mol. Biol.* **28**, 83–126
23. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekova, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) *J. Biol. Chem.* **280**, 29750–29764
24. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry* **30**, 526–537
25. Double, S., Sawaya, M. R., and Ellenberger, T. (1999) *Struct. Fold. Des.* **7**, R31–R35
26. Yang, W. (2005) *FEBS Lett.* **579**, 868–872
27. Harris, D., Kaushik, N., Pandey, P. K., Yadav, P. N., and Pandey, V. N. (1998) *J. Biol. Chem.* **273**, 33624–33634
28. Gao, J. (1994) *Biophys. Chem.* **51**, 253–261
29. Richardson, C. C. (1983) *Cell* **33**, 315–317