Articles

Varying DNA Base-Pair Size in Subangstrom Increments: Evidence for a Loose, Not Large, Active Site in Low-Fidelity Dpo4 Polymerase

Shin Mizukami, Tae Woo Kim, Sandra A. Helquist, and Eric T. Kool*

Department of Chemistry, Stanford University, Stanford, California 94305

Received September 28, 2005; Revised Manuscript Received December 7, 2005

ABSTRACT: We describe the first systematic test of steric effects in the active site of a Y-family DNA polymerase, Dpo4. It has been hypothesized that low-fidelity repair polymerases in this family more readily accept damaged or mismatched base pairs because of a sterically more open active site, which might place lower geometric constraints on the incipient pair. We have tested the origin of low fidelity by use of five nonpolar thymidine analogues that vary in size by a total of 1.0 Å over the series. The efficiency and fidelity of base-pair synthesis was measured by steady-state kinetics for single-nucleotide insertions. Analogues were examined both as incoming deoxynucleoside triphosphate (dNTP) derivatives and as template bases. The results showed that Dpo4 preferred to pair the thymidine shape mimics with adenine and, surprisingly, the preferred size was at the center of the range, the same optimum size as recently found for the high-fidelity Klenow fragment (Kf) of *Escherichia coli* DNA Pol I. However, the size preference with Dpo4 was quite small, varying by a factor of only 30−35 from most to least efficient thymidine analogue. This is in marked contrast to Kf, which showed a rigid size preference, varying by 1100-fold from best to worst. The fidelity for the non-hydrogen-bonding analogues in pairing with A over T, C, or G was much lower in Dpo4 than in the previous high-fidelity enzyme. The data establish that, unlike Kf, Dpo4 has very low steric selectivity and that steric effects alone cannot explain the fidelity (albeit low) that Dpo4 has for a correct base pair; the findings suggest that hydrogen bonds may be important in determining the fidelity of this enzyme. The results suggest that the low steric selectivity of this enzyme is the result of a conformationally flexible or loose active site that adapts with small energetic cost to different base-pair sizes (as measured by the glycosidic C1′−C1′ distance), rather than a spatially large active site.

The Y-family DNA polymerases are distinct among polymerase enzymes in their low fidelity of base-pair assembly and bypass (*1*−*3*). Enzymes in this class are noted for their moderate activity overall (*4*) and their relatively high efficiency of mismatched base-pair synthesis compared to correctly matched pairs (*1*−*3*). In addition to processing mismatches, a number of these enzymes have also been demonstrated to process and extend damaged DNA bases more efficiently than higher fidelity polymerases (*5*−*7*). It is generally hypothesized that the Y-family enzymes play a role in lessening the acute toxicity of damage and mispairing that arise in the cell (*1*−*3*). High-fidelity replicative polymerases tend to stall with high frequency at these lesion sites in DNA, and if such stalling persisted, it would lead to cell death. Thus, the low-fidelity Y-family enzymes are believed to aid in continuing DNA synthesis past these problem points, allowing other DNA damage repair machinery to correct these errors once replication is completed. Understanding the mechanistic and structural features that make Y-family polymerases different from other polymerases is therefore important for understanding how DNA damage and mutations arise during the lifespan of an organism and how such problems can be addressed without serious health consequences.

A chief issue in understanding these low-fidelity enzymes is how they adapt to process undamaged DNA and varied mismatches and damaged pairs in the same active site. Recent mechanistic and structural studies of the very different high-fidelity polymerases have led to the hypothesis that they may achieve their high selectivity of nucleotide choice by maintaining a tight steric control of the base-pair shape (*8*−*11*). Conversely, it has been suggested, primarily from structural studies, that the low-fidelity enzymes may allow mismatched and damaged pair shapes more readily by presenting a sterically more "open" active site, which might present a lower cost to accepting these nonstandard shapes (*1*−*3*, *12*−*20*).

Structures of polymerases complexed with template−primer DNA and incoming nucleotides are especially instructive in suggesting mechanistic differences. To date, two Y-family enzymes have been cocrystallized with DNA,

* To whom correspondence should be addressed. E-mail: kool@stanford.edu.

10.1021/bi051961z CCC: $33.50
Published on Web 02/07/2006

namely, Pol ι and Dpo4 polymerase. In the present paper, we address the latter, a DinB homologue from the crenarchaeon *Sulfolobus solfataricus* (*16*−*20*). This polymerase has been cocrystallized with template−primer and incoming nucleotides at 1.7−2.5 Å resolution, and both mismatched pairs and correctly matched pairs have been captured in the active site. Perhaps the most distinctive feature of these structures is the greater degree of "openness" around the active site in comparison with published structures of high-fidelity polymerases. The Dpo4 structures showed considerably less protein structure around the major groove of the incipient base pair and less compact structure in the minor groove as well.

However, those structural studies have raised a number of important questions that remain unanswered. First and most important is how exactly does this enzyme adapt to incorrectly shaped (mostly larger) substrates (*21*)? Does an "open" active site imply that it presents a greater amount of working space, thus making it sterically larger than other active sites? Or is it different in its flexibility, adapting to different substrates by shifting side chains and backbones subtly? Structural studies alone cannot answer these issues adequately for two reasons: first, although different substrates might be captured in X-ray structures, the static structures do not give the thermodynamic cost of structural changes, which is of course what controls kinetic efficiency. Second, the resolution of such structures is low compared to the known sensitivity of steric effects. For example, the Lennard−Jones approximation gives a cost of several kcal/mol for steric compression of less than 1 Å (*22*). This is an energy that, at a transition state, is sufficient to explain the observed base-pair selectivity, in a distance considerably less than the resolution of existing crystal structures.

For these reasons, it is important to address the functional differences among these enzymes mechanistically and at higher structural resolution. To test steric effects systematically, we have developed a series of nucleotide analogues having gradually increasing size (Figure 1) (*23*). These are thymidine analogues in which the natural polar heterocycle is replaced by a nonpolar structure meant to mimic the shape but gradually increase size. Size in this series is increased by starting with the smallest hydrogen substituents (replacing the two 2,4-exocyclic oxygens of thymine) and increasing the bulk gradually by replacing them with the halogen series F, Cl, Br, and I. Thus, size is increased in several increments over a total 1.0 Å range, giving resolution on a 25 pm average bond length change. When paired opposite adenine, these thymidine analogues are expected to give close to natural base-pair geometry (*24*), but the added size is expected to widen the base pair (as measured by the glycosidic distance from C1′−C1′ across a pair) in subtle increments.

A recent study with this steric series in a high-fidelity replicative enzyme, the Klenow fragment (Kf) of DNA Pol I, showed that the enzyme was remarkably strongly affected by the base-pair size, with a three-orders-of-magnitude range of DNA base-pair synthesis activity with steric changes of only 0.5 Å (*25*). Interestingly, the enzyme was not only efficient at rejecting pairs that are too large, but it also rejected pairs that are too small at least as selectively. The results suggested that at least one high-fidelity enzyme tightly controls complementary base-pair size by presenting a sterically rigid active site.

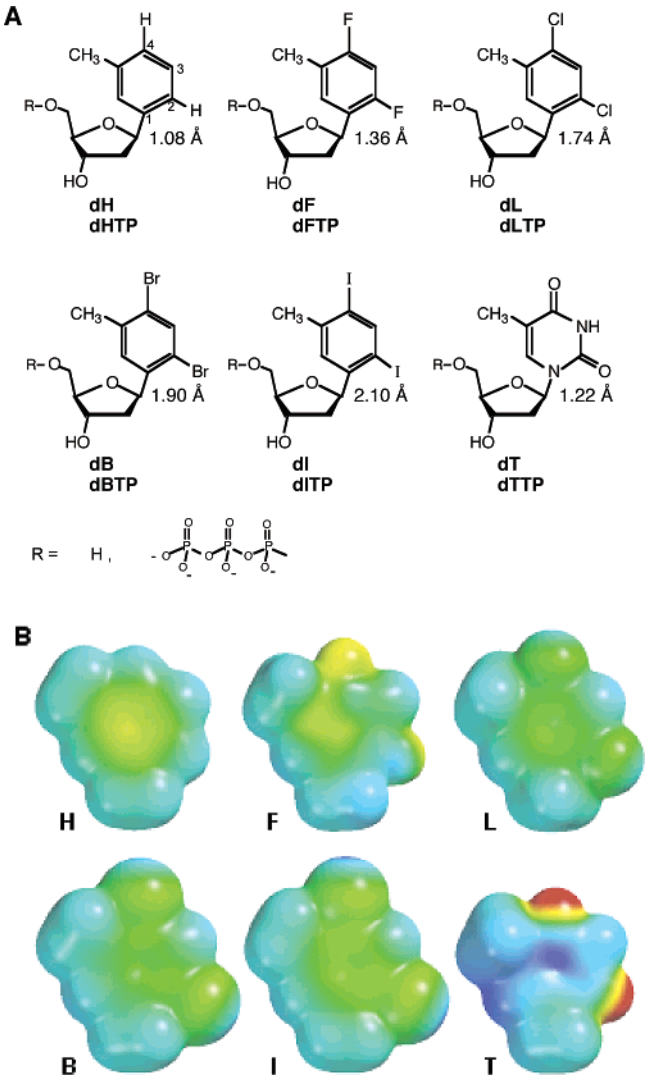

FIGURE 1: Structures of thymidine analogues having gradually increasing size, with thymidine shown for a comparison. (A) Chemical structures. Varied bond lengths at the 2 and 4 positions are shown. Numbering of positions is as indicated on the first structure. (B) Space-filling models of the bases with a methyl group at the 1 position in the place of deoxyribose.

Here, we employ this shape-mimicking thymidine analogue series to address the mechanistic similarities and differences between the low-fidelity enzyme, Dpo4, and a high-fidelity one. We directly compare kinetic behavior as a function of the size for the new Y-family enzyme and the previous A-family one, and we find evidence that the error-prone enzyme accepts varying-size substrates not by presenting a sterically larger and open active site but rather by flexibly adapting its pocket to its substrate.

MATERIALS AND METHODS

Modified Oligonucleosides and Triphosphate Derivatives. Modified oligodeoxynucleotides including H, F, L, B, and I were prepared as described (*23*). The 5′-triphosphate delivatives (dHTP, dFTP, dLTP, dBTP, and dITP) were prepared following the published procedure (*25*). The sequence of the primer was 5′-TAATACGACTCACTATAGGGAGA-3′ (23-mer), and the sequence of the template was 5′-ACTGX-TCTCCCTATAGTGAGTCGTATTA-3′ (28-mer), where X was A, G, T, C, H, F, L, B, or I.

Purification of Dpo4. Escherichia coli expressing Dpo4 was a gift from Dr. R. Woodgate (National Institutes of Health). The protein was purified according to the published method (*26*). The concentration was quantitated by the Bradford method.

Single-Nucleotide Insertion Reactions. Primer 5′ termini were labeled using [γ-^{32}P]ATP (Amersham Bioscience) and T4 polynucleotide kinase (Invitrogen). The labeled primer (~20 nM), template (100 μM), and unlabeled primer (50 μM) were mixed in 2× reaction buffer and gave a final total concentration of primer−template of 20 μM. The reaction buffer (1×) contained 40 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 100 μM nucleoside triphosphate, 10 mM dithiothreitol, 250 μg/mL bovine serum albumin (BSA), and 2.5% glycerol. The primer−template duplexes were annealed by heating to 90 °C and cooling slowly to 4 °C over 1 h. A 2× concentrated stock solution of duplex (2.5 μL) was mixed with Dpo4 (2.5 μL), and the reaction was initiated by adding a solution of the appropriate dNTP (5 μL). The enzyme concentration and reaction time were adjusted in different dNTP reactions to give less than 20% incorporation. Reactions were quenched with 15 μL of 95% formamide/10 mM EDTA containing 0.05% xylene cyanol and 0.05% bromophenol blue. Extents of the reaction were determined by running quenched reaction samples on a 20% polyacrylamide/7 M urea gel.

Steady-State Kinetic Analysis. Radioactivity was quantified using a Phosphorimager (Molecular Dynamics) and the ImageQuant Program. Reaction velocity v (M min^{-1}) was defined as $v = [S]I_{ext}/[(I_{prim} + I_{ext})t]$, where [S] is the concentration of triphosphate and I_{ext} and I_{prim} are the intensities of the extended product and the remaining primer, respectively. The k_{cat} and K_m values were obtained from Hanes−Woolf plots.

RESULTS

To probe the effects of varied base-pair size in the active site of Dpo4 polymerase, we employed a recently reported series of thymidine analogues that have gradually increasing size (Figure 1) (*17*). We prepared oligonucleotide templates in which the compounds replace natural thymidine immediately downstream of a primer-binding site, and we also prepared deoxynucleoside triphosphate (dNTP) derivatives, to allow them to be inserted opposite natural DNA bases in template−primer complexes. For a comparison, we also studied natural base pairs and mismatches. Polymerase-catalyzed insertion was monitored by use of 5′-radiolabeled primers; the addition of a nucleotide at the 3′ end was observed by polyacrylamide gel electrophoresis.

Our initial experiments compared the qualitative differences with the increasing-size series in the template strand immediately downstream from the primer. We observed that dATP could be inserted opposite all five members of the series, from the smallest analogue H (albeit weakly) to the largest analogue I (Figure 2A). This is considerably different than previously observed for DNA Pol I (Kf), which showed strong differences across the series (*25*). On the basis of relative intensities of the product band, the most efficient templates for Dpo4 fell in the middle of the series, with the smallest and largest showing a somewhat lower intensity. The overall efficiencies were significantly lower than the

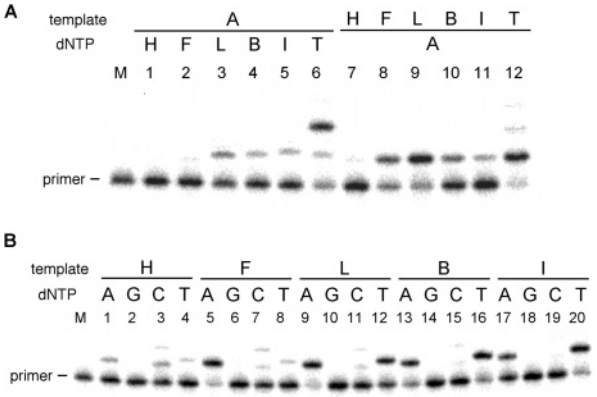

FIGURE 2: Autoradiograms showing single-nucleotide insertions by Dpo4 DNA polymerase with varied-size dNTP analogues. (A) PAGE gel showing the relative efficiencies for enzymatic pairing with adenine. (B) PAGE gel showing the selectivity of the analogues in the template strand with each of the four natural dNTPs. Conditions: 5 μM primer/template duplex, 300 μM dNTP, 100 nM Dpo4, 60 min, and 37 °C.

efficiency for insertion of dATP opposite the natural base thymine, because the same product band intensity for this natural pair was produced at a lower dATP concentration.

We also carried out the converse experiments in the same way, examining insertion of varying-size dTTP analogues opposite adenine in the template. The autoradiogram of the PAGE gel showed again (Figure 2B) that all five analogues could be paired opposite adenine, similar to the previous experiments. The relative band intensities showed that the three largest analogues were more efficiently inserted opposite A than were the smallest two; this is in marked contrast to previous results with Kf, which showed strong substrate ability for dLTP but much less for dBTP and dITP. As observed previously, the efficiencies for analogue insertion were significantly lower than for the insertion of natural dTTP opposite adenine.

To evaluate quantitative differences across this series, we then carried out steady-state kinetics studies for single-nucleotide insertions, by varying the dNTP concentration and reaction time and quantitatively evaluating product amounts by digital phosphorimagery. The data for varying template base sizes are given in Table 1, and data for varied dNTP size are given in Table 2. We carried out measurements of natural pairs and mismatches for a comparison to the analogue data. Examination of the data for the natural pairs and mismatches confirms that Dpo4, a Y-family repair enzyme, has significantly lower fidelity than Kf, an A-family replicative enzyme, in agreement with previous kinetics studies (*27*). When A is the template base being addressed, the order of dNTP insertion efficiency was T > C ≥ A ≥ G and the fidelity was demonstrated in a 1660-fold preference for insertion of dTTP over dCTP. When T was in the template strand, the order was dATP > dGTP > dCTP ~ dTTP, with a smaller 750-fold preference for dATP over dGTP. This is a lower fidelity than for Kf, which has a 1780-fold preference for dATP over dGTP using the same primer−template (see the Supporting Information) (*28*).

The data were further broken down into k_{cat} and K_m components with incoming nucleotide analogues (see the plot in the Supporting Information). The trend of size effects on k_{cat} shows a peak at the L analogue and a similar ~10-fold

Table 1: Steady-State Kinetic Efficiencies for Single-Nucleotide Insertions by Dpo4 Polymerase[a]

dNTP	template base	k_{cat} (min^{-1})	K_M (μM)	effciency, k_{cat}/K_M (M^{-1} min^{-1})	relative efficiency[b]
dATP	H	0.26 ± 0.03	270 ± 20	9.7×10^2	3.3×10^{-4}
	F	3.3 ± 1.1	210 ± 70	1.6×10^4	5.4×10^{-3}
	L	7.3 ± 2.9	210 ± 80	3.4×10^4	1.2×10^{-2}
	B	1.9 ± 0.3	340 ± 40	5.5×10^3	1.9×10^{-3}
	I	0.86 ± 0.25	190 ± 50	4.5×10^3	1.5×10^{-3}
	T	44 ± 1.4	15 ± 4	2.9×10^6	1
dCTP	H	0.45 ± 0.4	290 ± 80	1.6×10^3	5.4×10^{-4}
	F	0.36 ± 0.06	220 ± 30	1.6×10^3	5.6×10^{-4}
	L	0.23 ± 0.08	430 ± 160	5.4×10^2	1.8×10^{-4}
	B	0.13 ± 0.02	280 ± 70	4.5×10^2	1.5×10^{-4}
	I	0.12 ± 0.03	320 ± 110	3.8×10^2	1.3×10^{-4}
	T	0.99 ± 0.24	470 ± 130	2.1×10^3	7.2×10^{-4}
dGTP	H	0.021 ± 0.001	460 ± 80	4.5×10^1	3.2×10^{-6}
	F	0.026 ± 0.004	390 ± 40	6.7×10^1	2.3×10^{-5}
	L	0.025 ± 0.010	390 ± 200	6.5×10^1	2.2×10^{-5}
	B	0.022 ± 0.008	470 ± 210	4.6×10^1	1.6×10^{-5}
	I	0.039 ± 0.016	810 ± 350	4.8×10^1	1.7×10^{-5}
	T	2.5 ± 0.4	640 ± 90	3.9×10^3	1.3×10^{-3}
dTTP	H	0.23 ± 0.06	760 ± 250	3.0×10^2	1.0×10^{-4}
	F	0.33 ± 0.04	1000 ± 300	3.2×10^2	1.1×10^{-4}
	L	1.8 ± 0.5	830 ± 180	2.2×10^3	7.4×10^{-4}
	B	3.1 ± 1.5	540 ± 80	5.7×10^3	1.9×10^{-3}
	I	2.4 ± 0.6	240 ± 130	1.0×10^4	3.5×10^{-3}
	T	1.0 ± 0.1	660 ± 60	1.5×10^3	5.3×10^{-4}

[a] The template contained variable-size thymidine analogues. Conditions: 5 μM template−primer DNA; 10, 100, 300, or 500 nM enzyme; 40 mM Tris-HCl buffer (pH 8.0); 5 mM $MgCl_2$; 10 mM dithiothreitol; 250 μg/mL BSA; 2.5% glycerol; and 37 °C. [b] Relative to insertion of dATP opposite T.

Table 2: Steady-State Kinetic Efficiencies for Single-Nucleotide Insertions by Dpo4 Polymerase[a]

dNTP	template base	k_{cat} (min^{-1})	K_M (μM)	efficiency, k_{cat}/K_M (M^{-1} min^{-1})	relative efficiency[b]
dHTP	A	0.11 ± 0.05	620 ± 250	1.7×10^2	8.6×10^{-5}
dFTP		0.16 ± 0.04	270 ± 110	6.0×10^2	3.0×10^{-4}
dLTP		0.83 ± 0.10	160 ± 30	5.2×10^3	2.6×10^{-3}
dBTP		0.25 ± 0.02	110 ± 20	2.3×10^3	1.1×10^{-3}
dITP		0.15 ± 0.06	57 ± 11	2.7×10^3	1.3×10^{-3}
dTTP		43 ± 15	21 ± 3	2.0×10^6	1
dHTP	C	0.0036 ± 0.0002	260 ± 50	1.4×10^1	6.8×10^{-6}
dFTP		0.020 ± 0.003	490 ± 100	4.1×10^1	2.0×10^{-5}
dLTP		0.064 ± 0.003	160 ± 20	4.0×10^2	2.0×10^{-4}
dBTP		0.025 ± 0.001	89 ± 17	2.8×10^2	1.4×10^{-4}
dITP		0.030 ± 0.002	130 ± 20	2.3×10^2	1.1×10^{-4}
dTTP		0.22 ± 0.03	370 ± 70	6.0×10^2	3.0×10^{-4}
dHTP	G	0.0033 ± 0.0003	320 ± 70	1.0×10^1	5.1×10^{-6}
dETP		0.015 ± 0.002	420 ± 70	3.5×10^1	1.7×10^{-5}
dLTP		0.036 ± 0.004	120 ± 40	2.9×10^2	1.4×10^{-4}
dBTP		0.017 ± 0.002	35 ± 31	4.9×10^2	2.4×10^{-4}
dITP		0.056 ± 0.005	110 ± 30	5.2×10^2	2.6×10^{-4}
dTTP		0.34 ± 0.05	470 ± 100	7.3×10^2	3.6×10^{-4}
dHTP	T	0.013 ± 0.004	600 ± 160	2.1×10^1	1.1×10^{-5}
dFTP		0.095 ± 0.045	760 ± 160	1.3×10^2	6.2×10^{-5}
dLTP		0.29 ± 0.06	210 ± 80	1.4×10^3	6.8×10^{-4}
dBTP		0.22 ± 0.04	140 ± 40	1.5×10^3	7.6×10^{-4}
dITP		0.33 ± 0.14	140 ± 50	2.4×10^3	1.2×10^{-3}
dTTP		1.0 ± 0.1	660 ± 60	1.5×10^3	7.7×10^{-4}

[a] Incoming nucleotide analogues of variable size were used. Conditions: 5 μM template−primer DNA; 10, 100, 300, or 500 nM enzyme; 40 mM Tris-HCl buffer (pH 8.0); 5 mM $MgCl_2$; 10 mM dithiothreitol; 250 μg/mL BSA; 2.5% glycerol; and 37 °C. [b] Relative to insertion of dTTP opposite A (M^{-1} min^{-1}).

drop in activity at smaller and larger analogues. The trend with K_m values is different, showing a gradual drop in K_m as the analogues grow larger and more hydrophobic from dHTP to dITP.

The analogue kinetic data give useful information about the effects of small changes in size. The data show surprising differences from previous data for the high-fidelity Kf enzyme (Figure 3) (*25*). Interestingly, in both cases, the most efficient pairing involved the dichloro analogue, which lies at the middle of the series in size. This is virtually the same result as was seen for the high-fidelity Kf enzyme. However, in contrast to this, the range of efficiencies as a function of size for Dpo4 was very small compared to that seen for Kf. With A as the template base, the series dNTP analogues varied by only a factor of 30 across the series, whereas Kf showed a much larger range of 2200-fold from the maximum (dLTP) to the minimum (dHTP). The converse experiments, with varied-size analogues in the template, again showed a very small range of efficiencies as a function of size. In this case, the range was only a factor of 35 from the best template

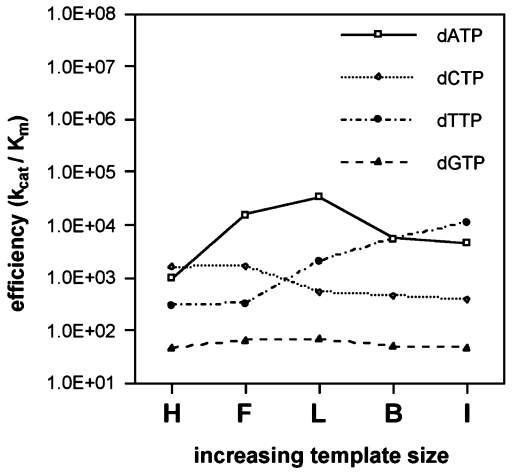

FIGURE 3: Plot of base-pair and mismatch efficiencies for Dpo4 polymerase with varied-size thymidine analogues in the template. Efficiencies are for single-nucleotide insertions in the steady state.

(L) to the worst template (H), which is again small compared to the previously studied Kf enzyme (which had a range of 180-fold).

In general, the data show that the most efficient base pairings involving T analogues opposite A were significantly lower in efficiency than the corresponding natural base pair, by a factor of ca. 2−2.5 orders of magnitude. The nucleotide dATP was inserted opposite the dichloro analogue 390-fold less efficiently than opposite T, while the converse pairing (dLTP opposite A) was 85-fold less efficient than the natural pairing (dTTP opposite A). This is again in marked contrast to the high-fidelity Kf enzyme, for which the dichlorotoluene−adenine pairings gave efficiencies that were nearly the same as the natural T−A pairing (*25*).

We also measured kinetics data for mismatched analogue pairings, to give information about the effects of nucleobase size on fidelity. In this case, because the analogues act as thymine surrogates, mismatches were defined as pairings of the analogues with T, C, or G. The data show (Table 1) that in most cases the most efficient mismatch involved the pairing of T with the analogue, which is consistent with previous results for the Kf enzyme. However, the results are markedly different from that high-fidelity enzyme, in that the fidelity for processing of the thymidine analogues was found to be very low with the Dpo4 enzyme. For the varied size dNTP analogues, the difference for an A template versus a T template was only 7.9-fold at the highest fidelity case. This compares to a fidelity of 250-fold for the Kf enzyme using these same analogues. The converse data, with varied-size template bases, showed a similar result: the maximum fidelity was 16-fold, compared to 1100-fold for Kf.

DISCUSSION

Our data comprise a systematic test of base-pair size effects in a Y-family DNA polymerase. The results show that the low-fidelity Dpo4 enzyme is functionally rather insensitive to base-pair size in kinetic terms, over a 1.0 Å range of bond-length increments. This is in marked contrast to previous data with the Kf of DNA Pol I, which demonstrated exquisite sensitivity to base-pair size (*25*). Kf polymerase showed differences of as much as 2200-fold in efficiency across the series, whereas Dpo4 polymerase displayed differences of only 35-fold. Although Dpo4 is indeed lower in fidelity than Kf, Dpo4 still retains almost 3 orders of magnitude of fidelity against mismatched pairs and its fidelity is less than that of Kf only by a factor of less than 5. Despite this, it discriminates by steric size much less than Kf. It is possible that, as has been suggested for the Y-family enzymes Pol η (*29*) and Pol κ (*30*), hydrogen bonding may be more critical in Dpo4 for positioning of the incoming nucleotide such that the triphosphate group is aligned correctly. This hypothesis would be consistent with our observation that even for the optimally sized base pair (L−A), the efficiency with Dpo4 is considerably lower than for a natural pair. This is quite different from what has been observed for Kf and related DNA polymerases, for which the L−A pair is virtually as efficient as a natural pair, suggesting little or no requirement for hydrogen bonding in those A-family enzymes.

We have hypothesized that polymerases might be most sensitive to "base-pair size" along the axis between glycosidic C1′ carbons (*11*); models suggest that the main effect in DNA of varying base size with the current analogues is to vary this glycosidic distance. What is the mechanism by which Dpo4 can function with relatively small dependence on this base-pair dimension? We consider two possible explanations: (i) either the enzyme presents a sterically open, larger active site with considerable space to accept varying base-pair sizes with only small steric penalties or (ii) the enzyme surrounds the incipient base pair closely but is flexible, opening the active site wider for larger pairs and collapsing inward around smaller pairs, with only small energetic cost for these flexible movements.

Closer examination of the data sheds light on these two possibilities. Figure 4 shows an overlay of data for Dpo4 with previous data for the Kf enzyme, plotting efficiency versus size for the various T analogue−A base pairs. Importantly, the data show virtually the same size trends for both enzymes, with even similar shapes to the plots. This is true both for the varied template and varied dNTPs. Both enzymes show a maximum in efficiency at the same size, the dichloro analogue, and both show drops in efficiency for smaller and larger analogue pairs. Both enzymes show a greater penalty for sizes that are too small than for those that are too large. These results establish that the functional active-site size (i.e., incipient base-pair size) for Dpo4 is identical to that of Kf, within 0.25 Å (the resolution of our probe series). If the functional active-site space were larger along the interglycosidic axis, our data would be expected to show a larger preferred size than for the high-fidelity enzyme; yet they did not. For this reason, the data are much more consistent with the notion that the active site of Dpo4 is functionally more flexible, rather than larger, than that of Kf.

We conclude that, in functional terms, the Dpo4 enzyme has a moderately sized but flexible ("loose") active site. Examination of the published structures of ternary complexes of Dpo4 with template−primer and dNTP (*17*, *18*) can suggest adaptations that may be made for pairs that are larger or smaller than the natural one. In the structures, the enzyme makes only van der Waals contacts with the sugars or bases of the incipient pair, except for hydrogen bonds to phosphate groups, which are not base-specific. Although the structure has been described as sterically "open", the glycosidic distance of the pair is limited by direct protein contacts with

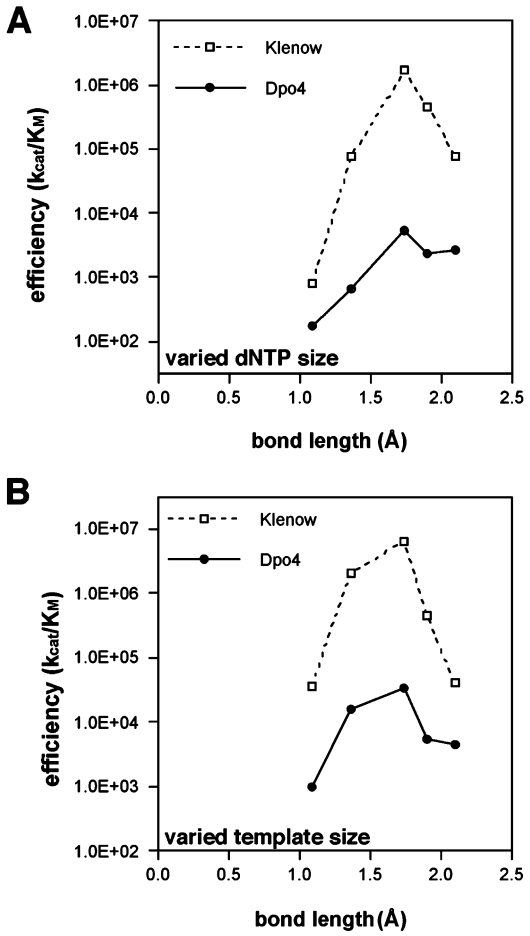

FIGURE 4: Comparison of base-pair size effects in Dpo4 polymerase (a low-fidelity enzyme) and Kf (a high-fidelity enzyme). Shown are efficiencies for pairing nonpolar thymine analogues with adenine. (A) Trend of varied dNTP size, with adenine in the template. (B) Trend with varied template base size, using dATP as the incoming nucleotide.

the opposite strands near the incipient pair. There are apparent contacts with the sugar of the dNTP residue and contacts with the sugar of the template nucleotide on the opposite strand. This may place limits the glycosidic C1′−C1′distance in the Dpo4 structure, a distance similar to that in a Klentaq ternary structure (*31*), an enzyme closely related to Kf. For Dpo4 to adapt to pairs larger or smaller than natural ones, these protein groups (notably, Arg247, Thr250 on the template side, and Asp105 and Tyr12 on the dNTP side) must move inward or expand outward. The outward expansion is likely driven by steric repulsion, while the inward collapse might be driven by the avoidance of a void (unfilled space), which is energetically costly in water (*32*). At present, these structural details must be considered speculative until more structural and functional data become available. We suggest that similar movements also likely occur with a high-fidelity enzyme such as Kf, but our data suggest that they are much more energetically prohibitive for Kf, thus leading to a rejection of the pair rather than the costly conformational adjustment. Thus, in simple terms, we would characterize the Dpo4 enzyme as functionally loose or flexible, while Kf is tight or rigid. It therefore seems likely that the structural difference in Dpo4, which clearly shows less congestion and more openness in the protein in the vicinity of the DNA, enables the active-site pocket residues to move inward or outward with relative freedom, collapsing or expanding the pocket as needed, its role in processing damaged or mismatched pairs.

Virtually all of the fidelity of the high-fidelity Kf enzyme in selection of incoming nucleotides can be explained by its high sensitivity to alterations in the base-pair size (*25*). In contrast, the present results show that Dpo4 exhibits quite a low fidelity in the choice of a partner for the thymine shape mimics, and interestingly, the fidelity among non-hydrogen-bonding analogue pairs is considerably lower than for hydrogen-bonding pairs. For example, a natural T−T mismatch is 1300-fold lower in efficiency than a T−A pair with Dpo4, whereas the analogous L−T mismatched pair is only 3.5-fold less efficient than a "correct" L−A pair. The efficiency data show that correct hydrogen-bonded pairing is required for high activity in this enzyme, and we suggest that hydrogen bonding is also important for fidelity in this case, by enhancing the activity of the correct pair while having little effect on mismatches. Our data show that mismatched pairs have similar activities whether they can undergo hydrogen bonding or not. Thus, although polymerases display structural similarities from class to class, it is quite clear that at least one A-family enzyme and one Y-family enzyme have functionally diverged markedly in their mechanisms of base-pair synthesis and discrimination.

Fiala and Suo have carried out more detailed studies of Dpo4 polymerase kinetics using natural nucleotides with mismatched and correctly matched partners (*33*). Their pre-steady-state data suggest that the rate-limiting step for correctly matched partners is a conformational change prior to the chemical-bond-forming step, while mismatch pair synthesis rates are limited by the bond-forming chemistry. Future kinetics studies involving the present size-varied analogue series would be useful in determining which step is rate-limiting.

Although these experiments shed some light on the mechanism by which Dpo4 may adapt to mismatched and damaged pairs, some important questions still remain. For example, would other Y-family polymerases or other families of DNA polymerases respond similarly to varied base-pair sizes, or would they differ because of specialized biological roles? In addition, because the biological function of this and other Y-family enzymes may be not only the synthesis of pairs involving damage or mismatching but also extension beyond such damaged pairs, what is the steric sensitivity of these enzymes in the extension step? More work will be needed to address these issues.

ACKNOWLEDGMENT

We are grateful to Dr. Roger Woodgate (NICHD) for providing the plasmid encoding the polymerase. We thank the U.S. National Institutes of Health (GM072705) for financial support. S. M. acknowledges a JSPS Postdoctoral Fellowship for Research Abroad, and T. W. K. acknowledges a KOSEF Fellowship.

SUPPORTING INFORMATION AVAILABLE

Figure S1, histograms of efficiency with natural pairs and mismatches; Figure S2, plots of increasing dNTP size on efficiency; Figure S3, plots of k_{cat} and K_m size dependence. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Goodman, M. F. (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes, *Annu. Rev. Biochem. 71*, 17−50.
2. Boudsocq, F., Ling, H., Yang, W., and Woodgate, R. (2002) Structure-based interpretation of missense mutations in Y-family DNA polymerases and their implications for polymerase function and lesion bypass, *DNA Repair 1*, 343−358.
3. Tippin, B., Pham, P., and Goodman, M. F. (2004) Error-prone replication for better or worse, *Trends Microbiol. 12*, 288−295.
4. Beard, W. A., Shock, D. D., Vande Berg, B. J., and Wilson, S. H. (2002) Efficiency of correct nucleotide insertion governs DNA polymerase fidelity, *J. Biol. Chem. 277*, 47393−47398.
5. Masutani, C., Kusumoto, R., Iwai, S., and Hanaoka, F. (2000) Mechanism of accurate transition synthesis by human DNA polymerase *η*, *EMBO J. 19*, 3100−3109.
6. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, R. (2000) Eukaryotic polymerases *ι* and *ζ* act sequentially to bypass DNA lesions, *Nature 406*, 1015−1019.
7. Zhang, Y. B., Yuan, F. H., Wu, X. H., Wang, M., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., and Wang, Z. (2000) Error-free and error-prone lesion bypass by human DNA polymerase *κ in vitro*, *Nucleic Acids Res. 28*, 4138−4146.
8. 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 selectivity, *Proc. Natl. Acad. Sci. U.S.A. 94*, 10506−10511.
9. Goodman M. F. (1997) Hydrogen bonding revisited: Geometric selection as a principal determinant of DNA replication fidelity, *Proc. Natl. Acad. Sci. U.S.A. 94*, 10493−10495.
10. Kunkel, T. A., and Bebenek, K. (2000) DNA replication fidelity, *Annu. Rev. Biochem. 69*, 497−529.
11. Kool, E. T. (2002) Active site tightness and substrate fit in DNA replication, *Annu. Rev. Biochem. 71*, 191−219.
12. Silvian, L. F., Toth, E. A., Pham, P., Goodman, M. F., and Ellenberger, T. (2001) Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*, *Nat. Struct. Biol. 8*, 984−989.
13. Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001) Structure of the catalytic core of *S. cerevisiae* DNA polymerase *η*: Implications for translesion DNA synthesis, *Mol. Cell. 8*, 417−426.
14. Zhou, B. L., Pata, J. D., and Steitz, T. A. (2001) Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain, *Mol. Cell. 8*, 427−437.
15. Nair, D. T., Johnson, R. E., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) Replication by human DNA polymerase *ι* occurs by Hoogsteen base-pairing, *Nature 430*, 377−380.
16. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4: Analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N2-ethenoguanine, *J. Biol. Chem. 280*, 29750−29764.
17. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: A mechanism for error-prone and lesion-bypass replication, *Cell 107*, 91−102.
18. Trincao, J., Johnson, R. E., Wolfle, W. T., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) Dpo4 is hindered in extending a G·T mismatch by a reverse wobble, *Nat. Struct. Mol. Biol. 11*, 457−462.
19. Ling, H., Boudsocq, F., Plosky, B. S., Woodgate, R., and Yang, W. (2003) Replication of a *cis-syn* thymine dimer at atomic resolution, *Nature 424*, 1083−1087.
20. Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M., and Yang, W. (2004) Crystal structure of a benzo[*a*]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase, *Proc. Natl. Acad. Sci. U.S.A. 101*, 2265−2269.
21. Choi, J. Y., and Guengerich, F. P. (2005) Adduct size limits efficient and error-free bypass across bulky N2-guanine DNA lesions by human DNA polymerase *η*, *J. Mol. Biol. 352*, 72−90.
22. Maitland, G. C., Rigby, M., Smith, E. B., Wakeham, W. A. (1981) *Intermolecular Forces*, Clarendon Press, Oxford, U.K.
23. Kim, T. W., and Kool, E. T. (2005) A series of nonpolar thymidine analogues of increasing size: DNA base pairing and stacking properties, *J. Org. Chem. 70*, 2048−2053.
24. Guckian, K. M., Krugh, T. R., and Kool, E. T. (1998) Solution structure of a DNA duplex containing a replicable difluorotoluene-adenine pair, *Nat. Struct. Biol. 5*, 954−959.
25. Kim, T. W., Delaney, J. C., Essigman, J. M., and Kool, E. T. (2005) Probing the active site tightness of DNA polymerase in subangstrom increments, *Proc. Natl. Acad. Sci. U.S.A. 102*, 15803−15808.
26. Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., Yang, W., and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication, *J. Biol. Chem. 279*, 32932−32940.
27. Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001) *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): An archaeral DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol*η*, *Nucleic Acids Res. 29*, 4607−4616.
28. Moran, S., Ren, R. X. F., Rumney, S., and Kool, E. T. (1997) Difluorotoluene, a nonpolar isostere for thymine, codes specifically and efficiently for adenine in DNA replication, *J. Am. Chem. Soc. 119*, 2056−2057.
29. Washington, M. T., Helquist, S. A., Kool, E. T., Prakash, L., and Prakash, S. (2003) Requirement of Watson−Crick hydrogen bonding for DNA synthesis by yeast DNA polymerase *η*, *Mol. Cell. Biol. 23*, 5107−5112.
30. Wolfle, W. T., Washington, M. T., Kool, E. T., Spratt, T. E., Helquist, S. A., Prakash, L., and Prakash, S. (2005) Evidence for a Watson−Crick hydrogen bonding requirement in DNA synthesis by human DNA polymerase *κ*, *Mol. Cell. Biol. 25*, 7137−7143.
31. Korolev, S., Nayal, M., Barnes, W. M., Di Cera, E., and Waksman, G. (1995) Crystal structure of the large fragment of *Thermos aquaticus* DNA polymerase I at 2.5 Å resolution-structural basis for thermostability, *Proc. Natl. Acad. Sci. U.S.A. 92*, 9264−9268.
32. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein-structure to cavity-creating mutations and its relation to the hydrophobic effect, *Science 255*, 178−183.
33. Fiala, K. A., and Suo, Z. (2004) Mechanism of DNA polymerization catalyzed by *Sulfolobus solfataricus* P2 DNA polymerase IV, *Biochemistry 43*, 2116−2125.

BI051961Z