Facile Polymerization of dNTPs Bearing Unnatural Base Analogues by DNA Polymerase α and Klenow Fragment (DNA Polymerase I)†

Molly Chiaramonte,‡ Chad L. Moore,‡ Kristi Kincaid, and Robert D. Kuchta*

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309

Received May 9, 2003; Revised Manuscript Received July 8, 2003

ABSTRACT: The high fidelity of DNA replication is largely dependent upon accurate incorporation of dNTPs by DNA polymerases. To study the mechanism underlying nucleotide selection, we synthesized four nucleotide analogues bearing the unnatural bases benzimidazole, 5-nitrobenzimidazole, 6-nitrobenzimidazole, and 5-nitroindole and analyzed their incorporation by three DNA polymerases. We have found that human DNA polymerase α (pol α) and the Klenow fragment of *Escherichia coli* DNA polymerase I (KF) incorporate all four nucleotide analogues opposite all four canonical bases up to 4000-fold more efficiently than an incorrect natural dNTP (i.e., rates that approach those of a correct, natural dNTP), even though the shape of any base pair formed between the analogue and the template likely does not resemble a normal base pair. While pol α preferentially incorporated the analogues opposite template pyrimidines, KF surprisingly preferred to polymerize them opposite template purines. Although neither pol α nor KF readily polymerized a natural dNTP opposite either 5- or 6-nitrobenzimidazole in the template strand, the enzymes did incorporate the analogues to generate novel base pairs. Both pol α and KF polymerized the analogues up to 140-fold more efficiently than dATP both across from abasic sites and as 3′-overhangs on blunt-ended templates. Although Maloney murine leukemia virus reverse transcriptase did not measurably incorporate the analogues, this enzyme bound the analogues with K_I's only slightly higher than the K_m for polymerization of the normal dNTP. The implications of these results with respect to how polymerases discriminate between correct and incorrect dNTPs are discussed.

Accurate replication of genomic DNA is one of the most critical aspects of survival for almost all organisms. Although DNA replication is a highly regulated task involving a huge cast of proteins, the error rates associated with this process, as low as 10^{-10}, are largely attributable to the DNA polymerases (*1*). Most DNA polymerases have very high base substitution fidelities during replication, displaying error rates as low as 10^{-4} to 10^{-6}, even without an accompanying 3′- to 5′-exonuclease (*2*−*4*). The job of the polymerase is to select the correct nucleotide, on the basis of its complementarity with the opposing base in the template strand, and incorporate it at the 3′-end of the growing daughter strand. This task is greatly complicated by two facts: the enzyme must discriminate between the four natural nucleotides which are structurally and chemically quite similar, and each polymerization event can involve any one of four template bases, causing the "correct" substrate to constantly vary.

While the methods by which DNA polymerases overcome these challenges are still a topic of vigorous debate, much of the general mechanism underlying polymerization has been determined (*5*−*7*). Initially, the polymerase·primer-template complex is in an "open" conformation with the helices closest to the 3′-end of the primer, the "fingers", rotated away from the DNA and making the active site readily accessible to the solvent. This appears to allow the dNTPs to be rapidly sampled until the correct partner for the template base becomes productively bound in the active site while at the same time providing an environment that could amplify the free energies associated with proper base pairing. With the correct base pair in hand, the polymerase undergoes a transformation to its "closed" form in which the fingers have rotated inward to clamp down on the DNA and bound nucleotide in the fully formed active site. In this closed form, the 3′-hydroxyl of the primer and the α-phosphate of the bound dNTP are positioned such that the enzyme can now efficiently catalyze phosphodiester bond formation. Finally, release of pyrophosphate and translocation of the polymerase serve to both place the next unpaired template base into the active site and regenerate the open state of the enzyme needed to begin a new round of incorporation.

One of the great mysteries of polymerase fidelity is the enzyme's transition from its open form, which can interact with all four natural nucleotides, to the closed catalytically active form in which, with rare exception, only the correct nucleotide is inserted. The seminal discovery of Watson and Crick that DNA forms a duplex by A base pairing with T and G base pairing with C led to the hypothesis that the hydrogen bonding in these two base pairs was the driving force behind nucleotide selection (*8*, *9*). However, differences in free energy between the matched and unmatched base pairs in solution cannot account for fidelity greater than 10^{-2} errors

† This work was supported by National Institutes of Health Grant GM54194 to R.D.K. K.K. is a Howard Hughes Medical Institute Predoctoral Fellow.

* To whom correspondence should be addressed. Telephone: (303) 492-7027. Fax: (303) 492-5894. E-mail: kuchta@spot.colorado.edu.

‡ Both of these authors contributed equally to these studies.

10.1021/bi034763l CCC: $25.00
Published on Web 08/12/2003

per insertion (*10*−*12*). This realization has resulted in intense interest in nucleotide selection and the proposal of many new hypotheses to explain fidelity.

Currently, one of the most widely accepted models for nucleotide selection by DNA polymerases is the active site geometry hypothesis, a variation of the induced-fit mechanism (*13*). The model proposes that the shape of the base pair formed between the incoming dNTP and the template base being replicated, as opposed to H-bond formation, is the critical determinant for fidelity. Indeed, dNTP analogues bearing isosteres of adenine (4-methylbenzimidazole) and thymine (2,4-difluorotoluene) incapable of normal H-bonding were shown to be relatively efficiently and selectively utilized by DNA polymerases, clearly demonstrating that formation of Watson−Crick H-bonds is not critical to either nucleotide polymerization or fidelity (*14*−*18*). This hypothesis proposes that the enzyme's active site is only partially formed when it is in the open conformation and sampling dNTPs; when the correct base-pairing partner binds across from the template base in the cleft, the switch to the closed conformation is either triggered or simply no longer prohibited. Since the closed form is required to complete the active site and only the proper base pair should induce the conformational change, this model provides an attractive solution to the problem of polymerase fidelity. This theory suggests a passive role for the DNA polymerase in which its selections are largely based on geometry, with additional contributions from various physicochemical properties of the bases (hydrophobicity, electrostatics, solvation, etc.) (*6*, *18*−*22*).

To further illuminate the mechanism of DNA polymerase nucleotide specificity, we characterized the polymerization of benzimidazole, 5-nitrobenzimidazole, 6-nitrobenzimidazole, and 5-nitroindole 2′-deoxyribofuranoside triphosphates by DNA pol α, KF,[1] and M-MuLV RTase. Both pol α and KF polymerized all four analogues opposite all four natural template bases at rates that approached those for a normal, correct base pair in many cases. Both polymerases also readily incorporated analogues across from abasic sites and as 3′-overhangs on blunt-ended templates. When either 5- or 6-nitrobenzimidazole was incorporated into the template strand, both pol α and KF formed novel analogue−analogue base pairs. These results have important implications with respect to inhibitor design and also suggest a novel mechanism for how polymerases discriminate between correct and incorrect dNTPs.

EXPERIMENTAL PROCEDURES

Materials. All reagents were of the highest quality commercially available. Unlabeled dNTPs were from Sigma and radiolabeled dNTPs from New England Nuclear. Unless they contained an unusual base or an abasic site, synthetic DNA oligonucleotides were purchased from Oligos, Etc., and their concentrations determined spectrally (*23*). The DNA oligonucleotide containing an abasic site was synthesized as previously described (*24*). T4 polynucleotide kinase, Klenow fragment (exo$^-$), and M-MuLV RTase were from New England BioLabs. Human DNA pol α (four-subunit complex) was expressed and purified as previously described (*25*).

Synthesis of 5- and 6-Nitrobenzimidazole 2′-Deoxyribofuranoside 5′-Triphosphates. 5-Nitrobenzimidazole was glycosylated using the procedure of Kazimierczuk et al. (*26*) using NaH and 1-chloro-2-deoxy-3,5-bis(*O*-*p*-toluoyl)-α-D-*erythro*-pentofuranose. The nucleoside isomers (protected 5- and 6-nitrobenzimidazole 2′-deoxyriboses) were separated by silica chromatography using 1:1 toluene−EtOAc. The protected nucleosides were identical in every respect to published data (*27*). The separate isomers were deprotected with NH_3-saturated MeOH and phosphorylated via the method of Ludwig (*28*). 5-Nitrobenzimidazole 2′-deoxyribonuceloside triphosphate: ^{31}P NMR (D_2O) −9.42 (d), −10.67 (d), −22.34 (dd); HRMS (FAB$^+$−glycerol) 607.9222 (calcd 607.9201, $MHNa_4^+$), 629.9046 (calcd 629.9021, MNa_5^+). 6-Nitrobenzimidazole 2′-deoxyribonucleoside triphosphate: ^{31}P NMR (D_2O) −10.32 (d), −11.04 (d), −21.87(dd); HRMS (FAB$^+$−glycerol) 563.9578 (calcd 563.9562, $MH_3Na_2^+$) 585.9388 (calcd 585.9382, $MH_2Na_3^+$) 607.9204 (calcd 607.9201, $MHNa_4^+$).

Synthesis of Benzimidazole 2′-Deoxyribofuranoside 5′-Triphosphate. The benzimidazole deoxyriboside was synthesized similarly to a previous report (*29*), except instead of generating the $HgBr^+$ salt of benzimidazole, we generated the Na^+ salt using NaH as previously described (*26*). Phosphorylation of the nucleoside was effected via the method of Ludwig (*28*). Benzimidazole 2′-deoxyriboside: ^{1}H NMR ([^{2}H]DMSO) 8.47 (1H, s, C2H), 7.70 (d, 1H, ArH), 7.66 (d, 1H, ArH), 7.24 (m, 2H, ArH), 6.36 (t, 1H, 1′H), 5.36 (d, 1H, OH), 4.98 (t, 1H, OH), 4.39 (m, 1H, 3′H), 3.86 (m, 1H, 4′H), 3.54 (m, 2H, 5′H), 2.60 (m, 1H, 2′H), 2.29 (m, 1H, 2′H). Benzimidazole 2′-deoxyriboside 5′-triphosphate: ^{31}P NMR (D_2O) −8.87 (d), −10.29 (d), −22.09 (dd).

Synthesis of 5-Nitroindole 2′-Deoxyribofuranoside 5′-Triphosphate. d5NITP was synthesized essentially as described previously (*30*). 5-Nitroindole 2′-deoxyriboside: ^{1}H NMR ([^{2}H]DMSO) 8.57 (1H, d, C2H), 8.02 (1H, dd, ArH), 7.87 (1H, d, ArH), 7.82 (1H, d, ArH), 6.80 (1H, d, ArH), 6.46 (1H, t, 1′H), 5.36 (1H, br s, OH), 4.98 (1H, br s, OH), 4.37 (1H, br s, 3′H), 3.85 (1H, m, 4′H), 3.53 (2H, m, 5′H), 2.49 (m, 2′H), 2.28 (1H, m, 2′H). 5-Nitroindole 2′-deoxyriboside 5′-triphosphate: ^{31}P NMR (D_2O) −9.61 (d), −10.35 (d), −22.25 (dd).

Synthesis of DNA Containing 5-Nitrobenzimidazole and 6-Nitrobenzimidazole. Synthesis of the phosphoramidites was performed essentially as described by Berger et al. (*31*). The nucleoside was coevaporated three times with pyridine, followed by vacuum drying overnight. The nucleoside (1 equiv), dimethoxytrityl chloride (DMTrCl, 1.5 equiv), and 4-(dimethylamino)pyridine (0.64 equiv) were dissolved in dry pyridine and heated to 65 °C. Once silica TLC (5:1 EtOAc−MeOH) showed complete consumption of starting material, the reaction was quenched by the addition of H_2O, and the mixture was partitioned between H_2O and EtOAc. The organic layer was washed with brine, dried over Na_2SO_4, and evaporated to dryness. The residue was then purified by silica chromatography (49:1 EtOAc−MeOH). Pro-

[1] Abbreviations: CldA, 9-β-D-2′-deoxyribofuranosyl-(2-chloroadenine); dBTP, 1-β-D-2′-deoxyribofuranosylbenzimidazole 5′-triphosphate; d5NBTP, 1-β-D-2′-deoxyribofuranosyl-(5-nitrobenzimidazole) 5′-triphosphate; d6NBTP, 1-β-D-2′-deoxyribofuranosyl-(6-nitrobenzimidazole) 5′-triphosphate; d5NITP, 1-β-D-2′-deoxyribofuranosyl-(5-nitroindole) 5′-triphosphate; FaraA, 9-β-D-arabinofuranosyl-(2-fluoroadenine); KF, Klenow fragment of *Escherichia coli* DNA polymerase I; M-MuLV RTase, Maloney murine leukemia virus reverse transcriptase; pol α, DNA polymerase α; Tris-HCl, tris(hydroxymethyl)-aminomethane hydrochloride.

tected 5-nitrobenzimidazole 2′-deoxyriboside: ^{1}H NMR ([^{2}H]DMSO) 8.72 (1H, s, NBI), 8.58 (1H, d, NBI), 8.06 (1H, dd, NBI), 7.95 (1H, d, NBI), 7.13−7.27 (9H, m, DMTr), 6.76 (4H, dd, DMTr), 6.50 (1H, t, 1′H), 5.47 (1H, d, 3′OH), 4.47 (1H, t, 3′H), 4.02 (1H, m, 4′H), 3.69 (6H, d, OMe), 3.14 (2H, m, 5′H), 2.78 (1H, m, 2′H), 2.45 (1H, m, 2′H). Protected 6-nitrobenzimidazole 2′-deoxyriboside: ^{1}H NMR ($CDCl_3$) 8.44 (1H, d, NBI), 8.28 (1H, s, NBI), 8.19 (1H, dd, NBI), 7.82 (1H, d, NBI), 7.32 (2H, d, DMTr), 7.16−7.24 (7H, m, DMTr), 6.73 (4H, dd, DMTr), 6.36 (1H, t, 1′H), 4.63 (1H, br m, 3′H), 4.22 (1H, m, 4′H), 3.73 (6H, d, OMe), 3.48 (1H, br s, 3′OH), 3.35−3.38 (2H, m, 5′H), 2.58−2.63 (2H, m, 2′H). The identity of protons was determined via COSY experiments.

To generate the phosphoramidites, the DMTr-protected nucleoside was coevaporated with toluene three times, followed by vacuum drying overnight. The nucleoside (1.0 equiv) was then dissolved in freshly distilled dichloromethane containing Hunig's base (2.0 equiv). Diisopropyl cyanoethyl chlorophosphoramidite (Aldrich) was then added, and the solution was allowed to stir at room temperature for approximately 3 h. The reaction was monitored by silica TLC (45:45:10 EtOAc−dichloromethane−triethylamine) and often had to be stopped prior to complete conversion of starting material to avoid formation of the H-phosphonate at longer times. The reaction was quenched with several drops of dry MeOH and evaporated to dryness. The residue was immediately subjected to flash chromatography (45:45:10 EtOAc−dichloromethane−triethylamine) to provide a pure phosphoramidite, which was dissolved in dry AcN (0.1 M final) and used immediately on an ABI394 DNA synthesizer.

To ensure that the synthesized templates actually contained either the 5-nitrobenzimidazole or 6-nitrobenzimidazole deoxyribonucleotides, purified synthetic oligonucleotides were first converted to nucleosides using with calf intestinal phosphodiesterase I and calf intestinal phosphatase (Sigma). Digested mixtures were directly subjected to HPLC analysis (Zorbax SB-Aq column, 30 mL gradient of 1−100% MeOH in H_2O). The UV chromatogram was compared to an authentic sample comprised of the four natural nucleosides and either 5- or 6-nitrobenzimidazole 2′-deoxyriboside in the same base ratio as the synthetic oligonucleotide. The chromatogram of the digested templates matched the UV chromatogram of the authentic samples in terms of relative chromatographic retention and UV spectrum of the nucleosides. Additionally, the UV spectrum of the 5- and 6-nitrobenzimidazole 2′-deoxyriboside derived from the synthetic oligonucleotides precisely matched that of authentic 5- and 6-nitrobenzimidazole 2′-deoxyriboside [For both the 5- and 6-nitrobenzimidazole nucleosides generated by digestion λ_{max} = 232 and 301 nm, identical (±2 nm) to the spectra obtained for authentic nucleoside.]

5′-End Labeling of Primers and Annealing of Primer-Template Pairs. DNA primers were 5′-^{32}P-labeled using polynucleotide kinase and [γ-^{32}P]ATP, gel purified, and annealed to the appropriate template as described previously (*32*, *33*). Stocks were stored at −20 °C.

Polymerization Assays with Pol α and KF. All kinetic data were determined under steady-state conditions. Assays contained enzyme, 1 μM 5′-^{32}P-labeled primer-template, 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM DTT, 0.05 mg/mL bovine serum albumin, and various concentrations of dNTPs and/or analogues. Polymerization reactions were initiated by the addition of enzyme, incubated at 37 °C for 5 min, and quenched by addition of an equal volume of gel loading buffer (90% formamide). Products were separated by denaturing gel electrophoresis (20% acrylamide, 8 M urea) and analyzed by phosphorimagery (Molecular Dynamics). Control experiments using both normal dNTPs and the analogues showed that the polymerization rates were constant over the time course of the assays. For pol α, all rates were normalized to the rates obtained in assays containing 2 nM pol α. For KF, all rates were normalized to the rates obtained in assays containing 0.0033 unit/μL KF.

DNA_A TCCATATCACAT(3′)
AGGTATAGTGTAATTCTTATCATCT

DNA_C TCCATATCACAT(3′)
AGGTATAGTGTACATCTTATCATCT

DNA_G TCCATATCACAT(3′)
AGGTATAGTGTAGATCTTATCATCT

DNA_T TCCATATCACAT(3′)
AGGTATAGTGTATATCTTATCATCT

DNA_{13-5N} TCCATATCACAT(3′)
AGGTATAGTGTAXATCTTATCATCT (X = 5-Nitrobenzimidazole)

DNA_{13-6N} TCCATATCACAT(3′)
AGGTATAGTGTAXATCTTATCATCT (X = 6-Nitrobenzimidazole)

DNA_{BLT} TCCATATCACAT(3′)
AGCAGGTATAGTGTA

DNA_{AB} TTTGAGAGAGAGA(3′)
AAACTCTCTCTCTXTCTGGGGGTAAAATGGGGCAAAATGGGGC

(X = Abasic site)

FIGURE 1: Sequences of the DNA primer-template pairs used. Primers are oriented 5′ to 3′ and templates 3′ to 5′ so that the complementary regions are easily visualized. The template base across from which the incoming dNTP will be incorporated is underlined.

Polymerization Assays with M-MuLV RTase. Polymerization assays were performed as above, except that the reaction buffer consisted of enzyme (final concentration 8 units/μL), 50 mM Tris (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, and 10 mM DTT.

RESULTS

Synthetic primer-templates of defined sequence were used to analyze the incorporation of unnatural nucleoside triphosphates (Figure 1). The primer-templates only differed in the identity of the 13th, and in one case the 13th and 14th, template bases. The primer covered the first 12 bases of the template; hence, the identity of the 13th template base specified the correct dNTP to be polymerized. The DNAs are named according to the 13th template base (e.g., adenine is the 13th base in DNA_A). Since we only wanted to observe the polymerization of one dNTP onto the primer, we needed to change the identity of both the 13th and 14th base in DNA_A to avoid having consecutive adenines in the template. Importantly, keeping the sequence of the flanking regions around the polymerization site constant minimized any potential effects of sequence context on the polymerization opposite each of the template bases.

Incorporation of the Four Natural dNTPs by Pol α, KF, and M-MuLV RTase. To provide a basis for understanding

Table 1: Kinetic Parameters for Incorporation of Correct Natural dNTPs on DNA_N by Pol α, KF, and M-MuLV RTase[a]

		pol α[b]			KF[b]			M-MuLV RTase[c]		
dNTP	DNA_N	V_{max} (SD)	K_m (SD) (μM)	V_{max}/K_m	V_{max} (SD)	K_m(SD) (μM)	V_{max}/K_m	V_{max} (SD)	K_m(SD) (μM)	V_{max}/K_m
dATP	DNA_T	21.6 (4.2)	2.3 (0.1)	9.4	26 (13)	0.2 (0.1)	124	110 (17)	148 (6)	0.7
dCTP	DNA_G	15.6 (2.5)	1.9 (0.7)	8.0	123 (63)	0.9 (0.5)	138	ND	ND	ND
dGTP	DNA_C	27.3 (5.5)	1.2 (0.2)	22.8	74 (28)	0.5 (0.5)	137	ND	ND	ND
dTTP	DNA_A	9.9 (0.2)	1.9 (0.9)	5.3	22 (13)	0.3 (0.3)	65	18 (3)	80 (3)	0.2

[a] Polymerization assays were performed as described in Experimental Procedures. SD = standard deviation; ND = not determined. Rates are given as percent extension per minute. [b] Kinetic constants were determined by measuring the percent of primer extended at increasing concentrations of the dNTP. Each dNTP was assayed a minimum of two times. Standard deviations were determined by averaging the values from each individual assay; V_{max}/K_m values are calculated by dividing the average V_{max} and K_m values. [c] Kinetic constants were determined as above for pol α and KF, except each dNTP was assayed once. Standard deviations were determined from the error in the linear fit.

Table 2: Kinetic Parameters for Misincorporation of Natural dNTPs on DNA_N by Pol α and KF[a]

		pol α[b]		KF	
dNTP	DNA_N	V_{max}/K_m (SD)	discrimination	V_{max}/K_m (SD)	discrimination
dCTP	DNA_T	<0.0003	$>3 \times 10^4$	0.008 (0.0002)	1.6×10^4
dGTP	DNA_T	0.001 (0.0002)	9×10^3	0.06 (0.005)	2.1×10^3
dTTP	DNA_T	≪0.0001	$>9 \times 10^4$	<0.0003	$>4 \times 10^5$
dATP	DNA_G	0.0003 (0.00002)	2.6×10^4	0.06 (0.003)	2.3×10^3
dGTP	DNA_G	0.0003 (0.00005)	2.6×10^4	0.035 (0.007)	3.9×10^3
dTTP	DNA_G	≪0.0001	$>8 \times 10^4$	0.03 (0.009)	4.6×10^3
dATP	DNA_C	0.02 (0.007)	1.2×10^3	0.03 (0.002)	4.6×10^3
dCTP	DNA_C	≪0.0001	$>2 \times 10^5$	≪0.0001	$>1.4 \times 10^6$
dTTP	DNA_C	0.0006 (0.0001)	3.8×10^4	≪0.0001	$>1.4 \times 10^6$
dATP	DNA_A	0.0008 (0.0001)	6.6×10^3	0.07 (0.04)	9.3×10^2
dCTP	DNA_A	<0.0003	$>1.8 \times 10^4$	0.04 (0.004)	1.6×10^3
dGTP	DNA_A	≪0.0001	$>5.3 \times 10^4$	0.006 (0.0003)	1.1×10^4

[a] Polymerization assays were performed as described in Experimental Procedures. SD = standard deviation. Discrimination reflects how much more efficiently the correct dNTP on each template is polymerized (Table 1) as compared to the incorrect dNTP. Kinetic constants were determined as described in Table 1, and standard deviations were determined by analyzing the error in the linear fit. The units for V_{max}/K_m are % elongated min^{-1} μM^{-1}. [b] Many mismatches were incorporated at levels that were either not detectable or too low for accurate determination of the kinetic parameters at up to 2 mM dNTP of interest. In these cases, $V_{max}/K_m < 0.0003$ was reported for dNTP incorporation that could be detected but not accurately measured, while $V_{max}/K_m \ll 0.0001$ was reported if incorporation of the dNTP was not detectable.

how efficiently a polymerase discriminates against the analogues, we first needed to quantify how efficiently the polymerases incorporated the natural dNTPs. Polymerization of correct dNTPs by pol α, KF, and M-MuLV RTase was measured on DNA_A, DNA_C, DNA_G, and DNA_T. The second-order rate constant for polymerization of each dNTP varied substantially for each polymerase, up to 5-fold (Table 1). Polymerization of the three noncognate dNTPs by each polymerase was also measured, and consistent with previous studies, these reactions were extremely slow (Table 2). The rates of the different misincorporation reactions varied substantially according to both the nature of the mismatch and the identity of the polymerase, again consistent with previous studies (*34*).

Incorporation of Nucleotide Analogues across from the Four Natural Bases by Pol α and KF. Using the set of four DNA_N primer-templates described above, we characterized the incorporation of four nucleotide analogues by pol α and KF (Table 3). The four analogues examined [2′-deoxyribofuranoside triphosphates of benzimidazole (dBTP), 5-nitrobenzimidazole (d5NBTP), 6-nitrobenzimidazole (d6NBTP), and 5-nitroindole (d5NITP) (Figure 2)] have a general structure similar to a purine; however, they lack all of the normal H-bonding groups found in the six-membered ring. Also, d5NITP lacks the ring nitrogen equivalent to N7 of a purine; hence, comparing its polymerization to that of d5NBTP, which retains this nitrogen, provides a measure of the importance of this nitrogen for selectivity. Both enzymes polymerized each of the analogues across from all four of the natural bases (Figure 3 and data not shown), even though structurally benzimidazole and indole more closely resemble a purine than a pyrimidine. This was not entirely unexpected as benzimidazole and indole derivatives have been reported to behave as somewhat universal bases in DNA melting temperature studies, and Smith et al. showed that d5NITP could be polymerized by KF (*30*, *35*, *36*). Most surprisingly, however, both polymerases incorporated these analogues with high efficiency and very different specificities (Table 3).

When compared with the insertion of the correct dNTP by pol α, the analogues were incorporated only 5−400-fold less efficiently, regardless of the identity of the opposing template base, and in some cases the incorporation efficiency approached that for a correct dNTP (e.g., d6NBTP opposite cytosine or thymine and d5NITP opposite cytosine). In contrast, pol α generally discriminated against incorrect, natural dNTPs by factors of 10^3 to $>10^5$. Thus, the discrimination against analogue incorporation by pol α was minimal when compared with the natural dNTPs. The enzyme did not greatly differentiate among the four analogues, with dBTP used slightly less efficiently than the others. Intriguingly, even though the analogues more closely resemble a purine than a pyrimidine regarding their size and shape, pol α exhibited only a slight preference to incorporate them as purines.

KF displayed stronger discrimination against the analogues, incorporating them 30−1300-fold less efficiently than the correct dNTP, and did not greatly differentiate between the four compounds. However, this level of discrimination

Table 3: Kinetic Parameters for Incorporation of dNTP Analogues on DNA_N by Pol α and KF[a]

		pol α				KF			
dNTP	DNA_N	V_{max} (SD)	K_m (SD) (μM)	V_{max}/K_m	discrimination	V_{max} (SD)	K_m (SD) (μM)	V_{max}/K_m	discrimination
d5NBTP	DNA_T	2.4 (0.8)	4.6 (4.1)	0.5	19	10 (5.4)	35.5 (30.4)	0.3	4.1×10^2
d5NBTP	DNA_G	1.8 (1.2)	9.5 (4.8)	0.2	40	53 (19)	30.4 (17.7)	1.7	81
d5NBTP	DNA_C	7.3 (4.1)	7.6 (3.4)	1.0	23	6.4 (3.9)	32.8 (27.0)	0.2	6.9×10^2
d5NBTP	DNA_A	18 (6.0)	26 (12)	0.7	13	12.6 (3.3)	282 (145.0)	0.04	3.1×10^3
d6NBTP	DNA_A	2.2 (0.7)	36 (12)	0.06	88	6.5 (4.2)	16.0 (8.5)	0.4	1.6×10^2
d6NBTP	DNA_G	12.9 (0.3)	38.7 (9.6)	0.3	27	31 (8.6)	85.3 (18.0)	0.4	3.5×10^2
d6NBTP	DNA_C	33.5 (0.7)	7.8 (0.6)	4.3	5.3	6.6 (0.2)	157 (23)	0.04	3.4×10^3
d6NBTP	DNA_T	10.5 (1.6)	35.0 (2.8)	0.3	31	6.8 (3.2)	91 (5.3)	0.07	9.3×10^2
d5NITP	DNA_T	6.4 (1.3)	12.3 (0.3)	0.5	19	21 (4.7)	57 (28)	0.4	3.1×10^2
d5NITP	DNA_G	4.8 (0.6)	17.3 (4.0)	0.3	27	15.2 (0.4)	32 (12)	0.5	2.8×10^2
d5NITP	DNA_C	11 (2.9)	4.0 (2.3)	2.7	8.4	17.4 (0.2)	43 (17)	0.4	3.4×10^2
d5NITP	DNA_A	5.2 (1.0)	17.3 (2.1)	0.3	18	117 (30)	56 (42)	2.1	31
dBTP	DNA_T	4.6 (0.5)	64 (25)	0.07	130	8.4 (3.0)	98 (43.0)	0.09	1.4×10^3
dBTP	DNA_G	2.3 (0.3)	100 (21)	0.02	400	121 (1.6)	70 (22)	1.7	81
dBTP	DNA_C	8.8 (2.0)	32.3 (7.6)	0.3	76	5.8 (0.1)	158 (5.9)	0.04	3.4×10^3
dBTP	DNA_A	4.8 (0.0)	110 (30)	0.04	130	23.1 (2.7)	107 (32)	0.2	3.3×10^2

[a] Polymerization assays were performed as described in Experimental Procedures. SD = standard deviation. The units for V_{max}/K_m are % elongated min^{-1} μM^{-1}. Discrimination reflects how much more efficiently the correct dNTP on each template is polymerized as compared to the analogue (Table 1). Kinetic constants were determined as described in Table 1 for KF and pol α.

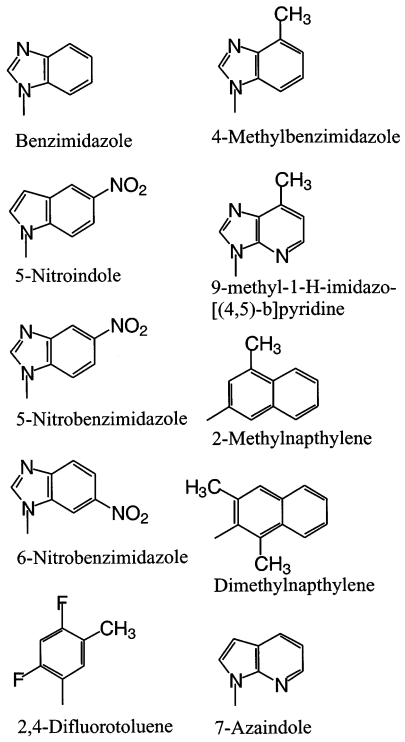

FIGURE 2: Structures of the base analogues discussed.

remains much less than the extent of discrimination against incorrect, natural dNTPs. In contrast to the results with pol α, KF displayed a slight preference to incorporate these analogues opposite purines in the template strand. Even though small, this favored insertion across from purines was quite astonishing in light of the structures of the analogues.

Inhibition of M-MuLV RTase. Unlike the rapid incorporation of the four analogues by pol α and KF, M-MuLV RTase very strongly discriminated against their polymerization. In fact, no detectable analogue polymerization was observed, even at elevated analogue concentrations (1 mM) and with extended incorporation times, indicating that this RTase discriminated against their polymerization by >10000-fold. However, the inability of RTase to polymerize these analogues did not result from an inability of the enzyme to bind

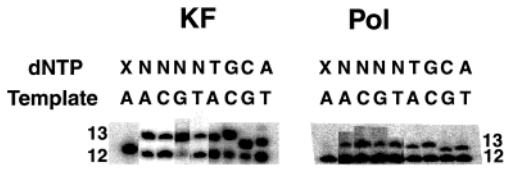

FIGURE 3: Incorporation of d6NBTP and the correct natural dNTPs across from the four natural bases by pol α and KF. The tested dNTP and the opposing base in the template strand are provided above the gel; X = no dNTP control, and N = d6NBTP. The assays contained either 50 μM d6NBTP or 10 μM correct natural dNTP and were performed as described in Experimental Procedures. The 12mer primer and 13mer extension products are labeled.

Table 4: K_I Values for the Inhibition of M-MuLV RTase Incorporation of dNTPs on DNA_N by Natural dNTPs and Analogues[a]

	K_I (SD) (μM)			
inhibitor	DNA_A	DNA_C	DNA_G	DNA_T
dATP	ND	2500 (500)	2200 (400)	ND
dCTP	2000 (100)	2800 (300)	ND	2500 (400)
dGTP	860 (120)	ND	1100 (300)	1300 (200)
d5NBTP	280 (30)	ND	ND	550 (70)
d6NBTP	100 (20)	ND	ND	110 (50)
d5NITP	190 (20)	490 (70)	240 (40)	340 (20)
dBTP	120 (10)	510 (100)	200 (50)	240 (40)

[a] Polymerization assays were performed as described in Experimental Procedures. SD = standard deviation; ND = not determined. For those normal dNTPs noted as "not determined", this was because they would be polymerized as a correct dNTP on the template (Figure 1). K_I values were determined from the IC_{50} values measured in assays containing either 25 μM dNTP being polymerized (inhibition by dATP, dCTP, and dGTP) or 100 μM dNTP being polymerized (inhibition by d5NBTP, d6NBTP, dBTP, and d5NITP).

them. Using the set of DNA_N primer-templates described above, we measured the ability of the analogues to inhibit the incorporation of natural dNTPs. In each case, inhibition was competitive with respect to the dNTP and relatively potent, as the K_m values for correct dNTP polymerization (ca. 100 μM) were similar to the K_I values measured for the analogues (100−500 μM, Table 4). Interestingly, the enzyme did not display any preferences with regard to either the template base or the inhibiting analogue.

Table 5: Kinetic Parameters for Incorporation of Natural dNTPs and Analogues on DNA_{13-5N} and DNA_{13-6N} by Pol α and KF[a]

dNTP	template	pol α[b] V_{max}/K_m (SD)	KF[b] V_{max}/K_m (SD)
dATP	DNA_{13-5N}	0.0014 (0.001)	≪0.0001
dCTP	DNA_{13-5N}	≪0.0001	≪0.0001
dGTP	DNA_{13-5N}	≪0.0001	≪0.0001
dTTP	DNA_{13-5N}	≪0.0001	≪0.0001
d5NBTP	DNA_{13-5N}	0.042 (0.01)	0.22 (0.02)
d6NBTP	DNA_{13-5N}	0.21 (0.04)	0.16 (0.03)
d5NITP	DNA_{13-5N}	0.14 (0.02)	0.46 (0.09)
dBTP	DNA_{13-5N}	0.006 (0.002)	≪0.0001
dATP	DNA_{13-6N}	0.003 (0.002)	≪0.0001
dCTP	DNA_{13-6N}	≪0.0001	≪0.0001
dGTP	DNA_{13-6N}	≪0.0001	≪0.0001
dTTP	DNA_{13-6N}	0.0011 (0.0003)	≪0.0001
d5NBTP	DNA_{13-6N}	0.06 (0.03)	0.19 (0.02)
d6NBTP	DNA_{13-6N}	0.98 (0.3)	≪0.0001
d5NITP	DNA_{13-6N}	0.45 (0.2)	1.3 (0.4)
dBTP	DNA_{13-6N}	0.019 (0.004)	≪0.0001

[a] Polymerization assays were performed as described in Experimental Procedures. SD = standard deviation. The units for V_{max}/K_m are percent elongated min^{-1} μM^{-1}. Kinetic constants were determined as described in Table 1 for pol α and KF. [b] Some dNTPs were incorporated at levels that were either not detectable or too low for accurate determination of the kinetic parameters at up to 2 mM dNTP of interest. In these cases, V_{max}/K_m was reported as ≪0.0001, as the smallest V_{max}/K_m that was measurable in our assays was 0.0003.

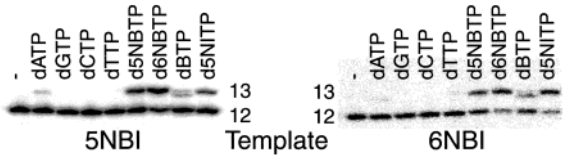

FIGURE 4: Polymerization of the natural dNTPs and analogues opposite 5-nitrobenzimidazole and 6-nitrobenzimidazole in the template. Assays contained pol α, either DNA_{13-5N} or DNA_{13-6N}, and 500 μM noted dNTP and were performed as described under Experimental Procedures. The 12mer primer and 13mer extension products are labeled.

Incorporation of Nucleotides across from 5- and 6-Nitrobenzimidazole by Pol α and KF. We incorporated two of the analogues, 5- and 6-nitrobenzimidazole, into the template strand and examined the ability of each enzyme to polymerize either a natural dNTP or an analogue opposite them. The primer-templates were identical to those described above, except that the 13th template base was either 5- or 6-nitrobenzimidazole [DNA_{13-5N} and DNA_{13-6N} , respectively (Figure 1)]. All of the polymerases incorporated the natural dNTPs poorly at best on these primer-templates, the most efficient polymerization of a natural dNTP being that of dATP on DNA_{13-6N} by pol α (ca. 3000-fold worse than forming a natural base pair; Table 5, Figure 4). Thus, pol α and KF clearly recognize the incipient base pair formed by any of the natural dNTPs and the nitrobenzimidazoles as incorrect in this orientation. In contrast, pol α and KF incorporated several of the analogues with efficiencies only 5−3800-fold worse than forming a Watson−Crick base pair. These novel analogue−analogue base pairs are quite surprising considering both their large size (similar to a purine−purine base pair) and that forming a planar base pair between some of the analogues could place the two electron-rich nitro groups in close proximity.

Incorporation of Nucleotides across from Abasic Sites and as 3′-Overhangs by Pol α and KF. Pol α and KF prefer to incorporate dATP in noninstructional situations such as across from abasic lesions or as 3′-overhangs on blunt-ended duplex DNA *(37)*. To explore polymerization of the analogues across from an abasic lesion, we used a primer-template containing a single abasic site (DNA_{AB}, Figure 1). To measure polymerization of nucleotides as 3′-overhangs, we annealed a 5′-^{32}P-labeled 12mer DNA primer to a 15mer DNA template such that the 3′-end of the primer formed a blunt end with the 5′-end of the template (DNA_{BLT}, Figure 1). The 3′-end of the template was designed to extend past the 5′-end of the primer in order to minimize the possibility that the polymerase would add nucleotides onto this 3′-hydroxyl. Consistent with previous studies, pol α and KF both incorporated dATP 6−1000-fold more efficiently on these templates than dCTP, dGTP, or dTTP (Figure 5, Table 6). KF polymerized d5NBTP, d6NBTP, and d5NITP on the abasic template, and d5NITP and d5NBTP on the blunt-ended template, with efficiencies similar to that for dATP on these templates. On the other hand, pol α polymerized the analogues slightly more efficiently opposite an abasic site (2−20-fold), and much more efficiently onto the blunt-ended template (10−150-fold), than it incorporated dATP in these situations.

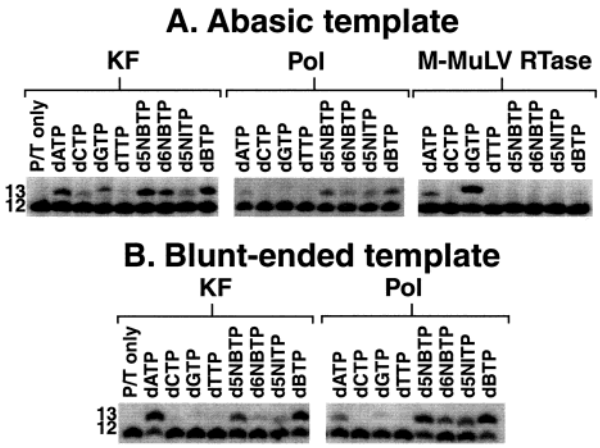

FIGURE 5: Incorporation of natural dNTPs and analogues across from an abasic site and as 3′-overhangs on a blunt-ended template by pol α, KF, and M-MuLV RTase. Polymerization assays were performed as described in Experimental Procedures and included either no dNTPs (P/T only) or 1 mM indicated natural or analogue dNTP. The sequences of the abasic primer-template (DNA_{AB}) and the blunt-ended primer-template (DNA_{BLT}) are shown in Figure 1. The 12mer primer and 13mer extension products are labeled on the gels.

M-MuLV RTase did not detectably incorporate any of the analogues on either the abasic or blunt-ended template, even at analogue concentrations as high as 1 mM. Thus, under both noninstructional and instructional conditions, this RTase does not measurably polymerize these analogues. Interestingly, when the natural dNTPs were tested with these noninstructional templates, this enzyme most readily polymerized dGTP opposite the abasic site instead of the more commonly used dATP and did not measurably incorporate any dNTP as a 3′-overhang onto the blunt-ended template (Figure 5 and data not shown).

Extension past Incorporated Analogues by Pol α and KF. Even though pol α and KF readily incorporated the various unnatural nucleotide analogues, elongation past these incorporations was extremely slow. As both polymerases inserted the analogues much more efficiently than they did an

Table 6: Kinetic Parameters for Incorporation of Natural dNTPs and Analogues on Abasic and Blunt-Ended Templates by Pol α and KF[a]

	V_{max}/K_m (SD)			
	abasic template		blunt-ended template	
dNTP	pol α	KF	pol α	KF
dATP	0.005 (0.0008)	0.42 (0.09)	0.007 (0.0005)	0.3 (0.02)
dCTP	≪0.0001	0.007 (0.003)	≪0.0001	<0.0003
dGTP	<0.0003	0.066 (0.004)	0.001 (0.0001)	<0.0003
dTTP	<0.0003	0.003 (0.001)	≪0.0001	<0.0003
d5NBTP	0.038 (0.004)	0.36 (0.04)	0.096 (0.004)	0.078 (0.02)
d6NBTP	0.079 (0.003)	0.28 (0.04)	1.02 (0.11)	<0.0003
d5NITP	0.098 (0.02)	0.81 (0.03)	0.46 (0.05)	0.48 (0.02)
dBTP	0.011 (0.0007)	<0.0003	0.076 (0.01)	<0.0003

[a] Polymerization assays were performed as described in Experimental Procedures. SD = standard deviation. The units for V_{max}/K_m are percent elongated min^{-1} μM^{-1}. Kinetic constants were determined as described in Table 1, and standard deviations were determined by analyzing the error in the linear fit. Because several of the compounds were incorporated at levels too low for accurate determination of kinetic data, $V_{max}/K_m < 0.0003$ was reported for dNTP incorporation that could be detected but not accurately measured, while $V_{max}/K_m \ll 0.0001$ was reported for no detectable incorporation.

incorrect dNTP, elongation past an incorporated analogue could be measured by simply including both the analogue triphosphate and the next correct dNTP in the reactions. Even when the concentration of the next correct dNTP was increased 10-fold (to 100 μM) and the concentration of enzyme was increased 4-fold over the standard reaction conditions, no detectable extension past the incorporated analogue was observed (Figure 6 and data not shown.)

Effects of Mn^{+2} on Incorporation and Elongation of the Nucleotide Analogues. Mn^{2+} has been shown to be highly mutagenic for both pol α and KF (*38*, *39*). Figure 6 shows that although the inclusion of 1 mM Mn^{2+} in the assays significantly increased the amount of d5NBTP incorporated by pol α and KF, it did not result in detectable polymerization past the incorporated analogue. Similar results were obtained with d5NITP, d6NBTP, and dBTP (data not shown). M-MuLV RTase remained unable to measurably polymerize the analogues, even when 1 mM Mn^{2+} was included in the assays (data not shown).

DISCUSSION

We have synthesized a series of novel dNTPs containing four different base analogues and tested them as substrates for DNA pol α, KF, and M-MuLV RTase. Whereas RTase did not readily polymerize any of these compounds, both pol α and KF incorporated them opposite all four of the canonical bases as well as at noninstructional/nonnatural template positions. The rates of incorporation of these nonnatural nucleotides opposite the natural bases were up to at least 4000-fold faster than the polymerases misincorporated natural dNTPs and, in some cases, only slightly less than those for the polymerization of the correct natural dNTP. In contrast to the rapid polymerization of the analogues, neither pol α nor KF detectably elongated them after incorporation, indicating that the "rules" for polymerization and elongation are likely very different.

Pol α and, to a lesser extent, KF exhibited a remarkable inability to discriminate against d5NBTP, d6NBTP, d5NITP, and dBTP, polymerizing them opposite all four normal bases.

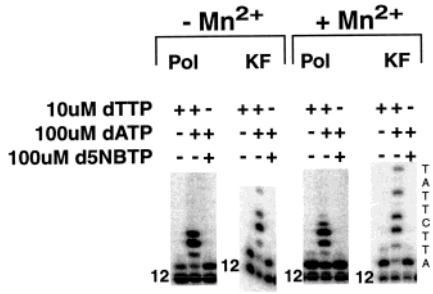

FIGURE 6: Effects of Mn^{2+} on polymerization of d5NTP on DNA_A. Polymerization assays were performed as described in Experimental Procedures and included either the first complementary dNTP (absent misincorporation, only +1 primer extension possible), the first and second complementary dNTPs (+2 extension possible), or d5NBTP and the second complementary dNTP (to test for extension past analogue incorporation). As noted, pol α and KF were tested in assays with and without 1 mM $MnCl_2$.

In contrast, these enzymes strongly discriminated against polymerization of incorrect natural dNTPs, indicating that the rapid polymerization of the analogues did not result from an unusually high misincorporation frequency on these primer-templates. The formation of H-bonds with the template bases probably does not account for the rapid polymerization of these nucleotides for several reasons. The H-bonds between the oxygens of the nitro group and an H-bond donor of the template base would likely be rather weak (*40*), and the drawn structures would vary substantially from the standard Watson−Crick geometry. Kool and coworkers have clearly shown that some DNA polymerases, including KF, do not require formation of H-bonds between the base of the incoming dNTP and the template base in order to catalyze phosphodiester bond formation (*15*, *18*). The nitroindole base is much more hydrophobic than a normal base (*41*, *42*), and it seems likely that the nitrobenzimidazoles are also more hydrophobic than a normal base, a feature that might allow them to bind more efficiently in the polymerase active site and stack onto the primer 3′-terminus. Indeed, and as described in greater detail below, it has been proposed that the reason adenine is so readily inserted in noninstructional situations is the greater hydrophobicity of adenine relative to the other normal bases. Interestingly, however, the greater hydrophobicity of these analogues did not enhance the ability of pol α and KF to polymerize them opposite a hydrophobic base analogue in the template. The rates of polymerization of the analogue triphosphates opposite 5-nitrobenzimidazole and 6-nitrobenzimidazole were very similar to the rates opposite the four natural bases (Tables 3 and 5).

Most surprising with respect to their rapid polymerization is the tremendous lack of shape complementarity of putative base pairs formed between the analogues and the canonical bases. Although the analogues are similar in overall shape to a normal purine, they contain extra mass at the Watson−Crick H-bonding face and lack an exocyclic group at C-6. One could imagine a base pair between either 5-nitroindole or 5-nitrobenzimidazole and either thymine or cytosine that is fairly close in size to a normal base pair, potentially accounting for their facile polymerization by KF and pol α. In contrast, any base pair between 6-nitrobenzimidazole and either thymine or cytosine likely requires at least some distortion away from normal Watson−Crick geometry. Placing the 6-nitrobenzimidazole and a pyrimidine in Watson−Crick

geometry would result in the electronegative O atoms of the nitro group and O-2 of the pyrimidine being in close proximity. Last, the generation of a base pair between any of the analogues and either adenine or guanine that resembles a normal base pair with respect to its geometry seems highly unlikely. Despite this, KF actually prefers to polymerize these analogues opposite a purine, and pol α polymerizes these analogues opposite purines much more readily than it polymerizes any normal, noncognate dNTPs.

KF polymerizes nucleotide analogues containing 4-methylbenzimidazole and difluorotoluene, isosteres of adenine and thymine, and the analogues described herein, with remarkably similar efficiency (compare the data in Table 3 and ref *15*). While detailed kinetic data for polymerization of 4-methylbenzimidazole and difluorotoluene nucleotides by pol α do not exist, this enzyme appears to actually prefer the more misshapen analogues (Table 3 and ref *17*). Together, these results suggest that, for KF and pol α, formation of a correctly shaped base pair is not critical for dNTP polymerization.

Potentially, the unique physicochemical properties of the nitrobenzimidazole and nitroindole bases vis à vis the canonical bases could account for the remarkable inability of the polymerases to discriminate against them. As noted earlier, these "bases" are made relatively electron deficient by the nitro group and are likely much more hydrophobic than the natural bases. If this feature (or features) does account for their facile polymerization, however, it would further indicate that shape is not a critical factor in determining whether these polymerases will incorporate a dNTP.

Additionally, previous studies by Schultz and co-workers suggest that the shape of the base pair between the incoming dNTP and the template base does not play a dominant role in determining whether a nucleotide is polymerized (*43*, *44*). They synthesized a number of nucleotides containing hydrophobic, aromatic base analogues that did not closely resemble the natural bases and tested them as substrates for KF. Many of them were polymerized opposite the natural bases as well as hydrophobic base analogues. Remarkably, even though any base pair formed between the incoming dNTP analogue and the template base likely has a geometry very different from that of a natural base pair, the rates of polymerization approached those observed for natural base pairs. For example, KF polymerized dNTPs containing the bases 2-methylnaphthylene and 7-azaindole opposite either a template adenine or a template dimethylnaphthylene only 4-fold and 10-fold slower, respectively, than it polymerized dTTP opposite a template adenine (Figure 2) (*44*).

The inability of the polymerases to identify nitrobenzimidazole and nitroindole as incorrect bases may actually result from their *lack* of similarity to a normal base. In discriminating between a correct dNTP and an incorrect dNTP, a polymerase could use two general mechanisms: specifically select *for* the correct dNTP (i.e., positive selectivity) or specifically select *against* the incorrect dNTP (i.e., negative selectivity). Mechanisms considered to date have generally focused on the polymerases obtaining fidelity by specifically selecting for the correct dNTP, based on either formation of Watson−Crick H-bonds or the correct active site geometry upon binding the correct dNTP. As noted above, however, it is difficult to rationalize the inability of a polymerase to discriminate against nitrobenzimidazole or nitroindole on the basis of a positive selection mechanism using either of these factors.

A negative selection mechanism would readily account for the inability of a DNA polymerase to discriminate against an unnatural analogue. The enzyme would bind the incoming dNTP, perhaps allowing the base of the dNTP and the template base to interact and adopt the lowest energy conformation. If the base pair formed is correct, the enzyme rapidly catalyzes phosphodiester bond formation. However, if the base pair formed is incorrect, some feature of this incorrect base pair prevents rapid phosphodiester bond formation. For example, if the enzyme attempted to polymerize dATP opposite a template deoxycytidylate, the 6-NH_2 group might be in the wrong location (vis à vis a correct A·T base pair), thereby blocking a conformational change necessary for rapid phosphodiester bond formation. In this model, therefore, the enzyme has developed mechanisms whereby the incorrect positioning of the ring substituents on the incoming dNTP (or, potentially, the template base) inhibits some process essential for catalysis. Two points about this model should be emphasized. First, the use of negative selectivity does not preclude the enzyme from simultaneously using positive selectivity to choose the correct dNTP [e.g., the apparent H-bond between N-3 of an incoming purine dNTP and Arg283 in pol *β* (*45*, *46*)]. Second, the enzyme does not block polymerization of an incorrect dNTP by specifically binding the incorrect dNTP. If it did, this would stabilize the bound, incorrect dNTP and would not result in discrimination against the incorrect dNTP.

Under negative selection, the inability of KF and pol α to effectively discriminate against the tested analogues results from their dissimilarity to a normal base. Compared to a normal purine, these compounds lack an exocyclic group at C6 and the free pair of electrons at N3, possess grossly altered exocyclic groups at N1 and C2, and, probably, display a very different *π*-electron distribution. If any of these features are critical for the polymerase to discriminate against an incorrect dNTP, a negative selection model predicts that the polymerase should not effectively discriminate against the analogue triphosphate.

Negative selection would also account for the ability of KF and pol α to effectively discriminate against polymerization of the natural dNTPs opposite a template 5- or 6-nitrobenzimidazole, even though these enzymes readily polymerize d5NBTP and d6NBTP opposite all four natural bases. When a natural dNTP binds opposite the template nitrobenzimidazole to form an E−DNA−dNTP ternary complex, some feature of the dNTP is located in a position that precludes polymerization. In contrast, the analogues lack this feature; hence, the polymerase can incorporate them opposite the nitrobenzimidazoles.

Importantly, precedence exists for DNA polymerases employing negative selectivity, albeit in the case of sugar fidelity. DNA polymerases very efficiently differentiate between ribose and 2′-deoxyribose, a situation where the concentration of the wrong substrate is 50−100-fold greater than the correct substrate. Tight binding of a nucleotide bearing a ribose sugar would require locating the 2′-hydroxyl in space occupied by an amino acid side chain of the protein (*47*−*51*). Reducing the size of this amino acid via mutagenesis dramatically reduces the ability of the mutant enzyme to discriminate against NTPs. Further evidence that at least

some DNA polymerases do not actively select for the two hydrogens at the 2′-C comes from studies of nucleotide analogues. A variety of analogues containing modified 2′-positions have been synthesized (e.g., 2′-deoxy-2′,2′-difluororibofuranosylcytosine, 2′,3′-dideoxydidehydrothymine, arabinofuranosylcytosine), and in many cases DNA polymerases readily polymerize these compounds (*51*−*53*). For example, both pol α and pol ϵ polymerize 2′-deoxy-2′,2′-difluororibofuranosylcytosine 5′-triphosphate only slightly less efficiently than dCTP, even though the electronic character of the 2′-C is vastly different in the two compounds (*53*). Presumably, the similar rates of polymerization result from the similar sizes of H and F.

Whereas base pair shape plays a relatively small role for simple polymerization of a dNTP, shape may be critical for polymerases to discriminate between a correct and incorrect dNTP. For example, KF discriminates against polymerization of 4-methybenzimidazole across from the natural bases by approximately 10^4-10^5-fold, levels similar to those at which KF discriminates against misincorporation of dATP (*15*). In contrast, KF discriminates against polymerization of benzimidazole across from the natural bases by only approximately 10^2-10^3-fold. Together, these data suggest that the exocyclic mass at C-4 of benzimidazole (equivalent to C-6 of a purine) is critical for the fidelity of KF.

Among the three polymerases we tested, only M-MuLV RTase demonstrated an ability to strongly discriminate against the analogues. This ability did not result from an inherently higher fidelity of M-MuLV RTase as compared to pol α and KF since the misincorporation frequencies of all three enzymes were similar. This extremely strong discrimination against polymerization also cannot be attributed to weak binding of the analogues by the enzyme; d5NBTP, d6NBTP, d5NITP, and dBTP all inhibited M-MuLV RTase with K_I's only slightly greater than the K_m for a correct dNTP. Interestingly, M-MuLV RTase also discriminated very strongly against polymerization of dNTP analogues containing either 4-methylbenzimidazole or 9-methyl-1*H*-imidazo-[4,5-*b*]pyridine as the base (*17*). While the actual reasons for this strong discrimination remain obscure, two potential explanations are that this enzyme requires hydrogen-bonding groups at C6 and/or N1 of the purine ring and that very hydrophobic bases are incapable of binding in the enzyme's active site in a productive conformation.

Unlike most polymerases that use dATP at noninstructional lesions, M-MuLV RTase instead uses dGTP across from an abasic site. It has been suggested that most polymerases use dATP because of the greater hydrophobicity of adenine relative to the other bases (*21*). Indeed, hydrophobicity could account for the rapid polymerization of the analogues by pol α and KF, both of which prefer dATP, under noninstructional conditions. The marked preference of M-MuLV RTase for dGTP indicates that hydrophobicity is not the primary driving force for this enzyme under noninstructional situations. Rather, these data suggest that specific interactions between the polymerase and dGTP direct its polymerization under noninstructional situations. M-MuLV RTase is not alone in preferring dGTP, since at least two other DNA polymerases, ι and η, also preferentially polymerize dGTP opposite abasic lesions (*54*, *55*).

A previous study using M-MuLV RTase reported that an abasic site completely blocked polymerization (*56*). Interestingly, this study used a deoxyribolactone to mimic an abasic site, whereas we used a tetrahydrofuran mimic. These different results indicate that the structure of the sugar at an abasic site can greatly influence M-MuLV RTase activity.

The remarkably rapid polymerization of nucleotides whose bases do not closely resemble any of the canonical bases has important implications for the design of nucleotide-based chemotherapeutics. Generally, these compounds contain a modified sugar and either a canonical base or a slightly modified base (e.g., 2-fluoroadenine in FaraA, 2-chloroadenine in CldA). Ribavirin probably contains the most unusual base, 1,2,4-triazole-3-carboxamide, although even in this case the 1,2,4-triazole-3-carboxamide is thought to resemble a normal base since the overall length of the base is similar to a purine and the carboxamide may form Watson−Crick H-bonds analogous to those formed by adenine or guanine (*57*). However, our results showing the rapid polymerization of dNTPs containing unusual bases suggest that it should be possible to generate a large number of novel nucleotide analogues with interesting pharmacological properties by utilizing bases that do not closely resemble the canonical bases.

REFERENCES

1. Drake, J. W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7160−7164.
2. Bebenek, K., and Kunkel, T. (1993) in *Reverse Transcriptase* (Skalka, A. M., and Goff, S., Eds.) pp 85−102, Cold Spring Harbor Laboratory Press, Plainview, NY.
3. Kunkel, T. A., and Bebenek, K. (1988) *Biochim. Biophys. Acta 951*, 1−15.
4. Roberts, J. D., and Kunkel, T. A. (1996) in *DNA Replication in Eukaryotic Cells*: *Concepts*, *Enzymes and Systems* (DePamphilis, M., Ed.) pp 217−247, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
5. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) *Biochemistry 36*, 11205−11215.
6. Johnson, K. A. (1993) *Annu. Rev. Biochem. 62*, 685−713.
7. Echols, H., and Goodman, M. F. (1991) *Annu. Rev. Biochem. 60*, 477−512.
8. Watson, J. D., and Crick, F. H. C. (1953) *Nature 171*, 737−738.
9. Watson, J. D., and Crick, F. H. C. (1953) *Nature 171*, 964−967.
10. Raszka, M., and Kaplan, N. O. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 2025−2029.
11. Mildvan, A. S. (1974) *Annu. Rev. Biochem. 43*, 357−399.
12. Loeb, L. A., and Kunkel, T. A. (1982) *Annu. Rev. Biochem. 51*, 429−457.
13. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry 30*, 526−537.
14. Modrich, P. (1997) *J. Biol. Chem. 272*, 24727−24730.
15. Morales, J. C., and Kool, E. T. (1998) *Nat. Struct. Biol. 5*, 950−954.
16. Morales, J. C., and Kool, E. T. (2000) *Biochemistry 39*, 2626−2632.
17. Morales, J. C., and Kool, E. T. (2000) *J. Am. Chem. Soc. 122*, 1001−1007.
18. Moran, S., Ren, R. X., and Kool, E. T. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 10506−10511.
19. Minnick, D. T., Grindley, N. D., Kunkel, T. A., and Joyce, C. M. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 1194−1199.
20. Kool, E. T. (1998) *Biopolymers 48*, 3−17.
21. Kool, E. T. (2002) *Annu. Rev. Biochem. 71*, 191−219.
22. Patel, P. H., Kawate, H., Adman, E., Ashbach, M., and Loeb, L. A. (2001) *J. Biol. Chem. 276*, 5044−5051.
23. Kirk, B. W., and Kuchta, R. D. (1999) *Biochemistry 38*, 10126−10134.
24. Randall, S. K., Eritja, R., Kaplan, B. E., Petruska, J., and Goodman, M. F. (1987) *J. Biol. Chem. 262*, 6864−6870.
25. Zerbe, L. K., Goodman, M. F., Efrati, E., and Kuchta, R. D. (1999) *Biochemistry 38*, 12908−12914.
26. Kazimierczuk, Z., Cottam, H. B., Revankar, G. R., and Robins, R. K. (1984) *J. Am. Chem. Soc. 106*, 6379−6382.
27. Seela, F., and Bourgeois, W. (1989) *Synthesis 12*, 912−918.

28. Ludwig, J. (1981) *Acta Biochim. Biophys. Acad. Sci. Hung. 16*, 131−135.
29. Papageorgiou, C., and Tamm, C. (1987) *Helv. Chim. Acta 70*, 138−141.
30. Smith, C. L., Simmonds, A. C., Felix, I. R., Hamilton, A. L., Kuman, S., Nampalli, S., Loakes, D., Hill, F., and Brown, D. M. (1998) *Nucleotides Nucleosides 17*, 541−554.
31. Berger, M., Luzzi, S. D., Henry, A. A., and Romesberg, F. E. (2002) *J. Am. Chem. Soc. 124*, 1222−1226.
32. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
33. Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., and Benkovic, S. J. (1987) *Biochemistry 26*, 8410−8417.
34. Kunkel, T. A., and Bebenek, K. (2000) *Annu. Rev. Biochem. 69*, 497−529.
35. Loakes, D., and Brown, D. M. (1994) *Nucleic Acids Res. 22*, 4039−4043.
36. Seela, F., Bourgeois, W., Rosemayer, H., and Wenzel, T. (1996) *Helv. Chim. Acta 79*, 488−498.
37. Strauss, B. S. (1991) *BioEssays 13*, 79−84.
38. Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H., and Kraut, J. (1996) *Biochemistry 35*, 12762−12777.
39. Beckman, R. A., Mildvan, A. S., and Loeb, L. A. (1985) *Biochemistry 24*, 5810−5817.
40. Loakes, D. (2001) *Nucleic Acids Res. 29*, 2437−2447.
41. Guckian, K. M., Schweitzer, B. A., Ren, R. X., Sheils, C. J., Tahmassei, D. C., and Kool, E. T. (2000) *J. Am. Chem. Soc. 122*, 2213−2222.
42. Kool, E. T. (2001) *Annu. Rev. Biophys. Biomol. Struct. 20*, 1−22.
43. Ogawa, A. K., Wu, Y., McMinn, D. L., Liu, J., Schultz, P. G., and Romesberg, F. E. (2000) *J. Am. Chem. Soc. 122*, 3274−3287.
44. Berger, M., Wu, Y., Ogawa, A. K., McMinn, D. L., Schultz, P. G., and Romesberg, F. E. (2000) *Biochemistry 28*, 2911−2914.
45. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science 264*, 1891−1903.
46. Werneburg, B. G., Ahn, F., Zhong, X., Hondal, R. J., Kraynov, V. S., and Tsai, M.-D. (1996) *Biochemistry 35*, 7041−7050.
47. Astatke, M., Ng, K., Grindley, N. D., and Joyce, C. M. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3402−3407.
48. Bonnin, A., Lazaro, J. M., Blanco, L., and Salas, M. (1999) *J. Mol. Biol. 290*, 241−251.
49. Gardner, A. F., and Jack, W. E. (1999) *Nucleic Acids Res. 27*, 2545−2553.
50. Cases-Gonzalez, C. E., Gutierrez-Rivas, M., and Menendez-Arias, L. (2000) *J. Biol. Chem. 275*, 19759−19767.
51. Yang, G., Franklin, M., Li, J., Lin, T. C., and Konigsberg, W. (2002) *Biochemistry 41*, 10256−10261.
52. Kuchta, R. D., Ilsley, D., Kravig, K. D., Schubert, S., and Harris, B. (1992) *Biochemistry 31*, 4720−4728.
53. Huang, P., Chubb, S., Hertel, L. W., Grindey, G. B., and Plunkett, W. (1991) *Cancer Res. 51*, 6110−6117.
54. Haracska, L., Washington, M. T., Prakash, S., and Prakash, L. (2001) *J. Biol. Chem. 276*, 6861−6866.
55. Zhang, Y., Yuan, F., Wu, X., Taylor, J. S., and Wang, Z. (2001) *Nucleic Acids Res. 29*, 928−935.
56. Berthet, N., Roupioz, Y., Constant, J.-F., Kotera, M., and Lhomme, J. (2001) *Nucleic Acids Res. 29*, 2725−2732.
57. Maag, D., Castro, C., Hong, Z., and Cameron, C. E. (2001) *J. Biol. Chem. 276*, 46094−46098.

BI034763L