Human DNA Primase Uses Watson−Crick Hydrogen Bonds To Distinguish between Correct and Incorrect Nucleoside Triphosphates†

Chad L. Moore,‡ Aleksandra Zivkovic,§ Joachim W. Engels,§ and Robert D. Kuchta*,‡

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309, and Institut für Organische Chemie, Johann Wolfgang Goethe Universität, Marie Curie Strasse 11, D-60439 Frankfurt am Main, Germany

Received May 6, 2004; Revised Manuscript Received July 9, 2004

ABSTRACT: Human DNA primase synthesizes short RNA primers that DNA polymerase α further elongates. Primase readily misincorporates the natural NTPs and will generate a wide variety of mismatches. In contrast, primase exhibited a remarkable resistance to polymerizing NTPs containing unnatural bases. This included bases whose shape was almost identical to the natural bases (4-aminobenzimidazole and 4,6-difluorobenzimidazole), bases shaped very differently than a natural base [e.g., 5- and 6-(trifluoromethyl)benzimidazole], bases much more hydrophobic than a natural base [e.g., 4- and 7-(trifluoromethyl)-benzimidazole], bases of similar hydrophobicity as a natural base but with the Watson−Crick hydrogen-bonding groups in unusual positions (7-*β*-D-guanine), and bases capable of forming only one Watson−Crick hydrogen bond with the template base (purine and 4-aminobenzimidazole). Primase only polymerized NTP analogues containing bases capable of forming hydrogen bonds between the equivalent of both N-1 and the exocyclic group at C-6 of a purine NTP (2-fluoroadenine, 2-chloroadenine, 3-deazaadenine, and hypoxanthine) and N-3 and the exocyclic group at C-4 of a pyrimidine. These data indicate that human primase requires the formation of Watson−Crick hydrogen bonds in order to polymerize a NTP, a situation very different than what is observed with some DNA polymerases. The implications of these results with respect to current theories of how polymerases discriminate between right and wrong (d)NTPs are discussed.

One of the key features of nucleotide polymerizing enzymes that remains unresolved is how they distinguish between correct and incorrect (d)NTPs. The similar shape, hydrophilicity, and charge of NTPs and dNTPs present a daunting challenge of determining a correct substrate in the face of numerous incorrect binding events. This rapid binding and release of incorrect (d)NTPs must occur through the tens, hundreds, or thousands of base pairs replicated by a processive polymerase (*1*). By understanding the mechanism of substrate selectivity in differing polymerases, clues to the origin of polymerase fidelity can be deciphered.

A traditional view of substrate selection based on hydrogen-bonding potential between the template and incoming nucleotide has largely been discounted as a primary model for fidelity (*2*−*4*). Calculation of free energy differences between correct and incorrect base pairs suggests that hydrogen bonding is not a sufficient explanation for high-fidelity substrate selection. Given that the change in free energy in aqueous solution is approximately 0.2−4 kcal/mol between correct and mismatched base pairs (suggesting an error rate of ∼1 in 100), the hydrogen-bonding theory does not account for the typically observed error rates of 10^{-4}−10^{-6} for high-fidelity polymerases (*5*). Additionally, (d)NTP polymerization by some polymerases does not require the formation of hydrogen bonds between the incoming (d)NTP and template base being replicated (*6*). This has led to the development of new models for polymerase fidelity, which do not heavily rely on hydrogen bonds.

One theory suggests a modified induced-fit theory ("shape selectivity") (*7*). Correct base pairs adopt a shape that the polymerase recognizes as correct, resulting in rapid (d)NTP polymerization. In contrast, incorrect base pairs adopt unusual shapes that do not allow for rapid polymerization. Alternatively, others have posited that base pair fidelity is due to primarily hydrophobic interactions (*8*). It has also been proposed that some DNA polymerases obtain fidelity primarily by discriminating against the wrong (d)NTP, as opposed to selecting for the correct (d)NTP (*9*, *10*). While each model has proponents and biochemical data, none has proven universal in explaining substrate choice of all the polymerases studied, nor have the data been entirely consistent between the two models.

Human DNA primase synthesizes short RNA primers that DNA polymerase α (pol α)[1] then elongates (*1*). The rate-limiting step in primer synthesis is the initiation reaction (i.e., dinucleotide synthesis), after which all further polymerization

† This work was supported by National Institutes of Health Grant GM54194 to R.D.K.

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

‡ University of Colorado.

§ Johann Wolfgang Goethe Universität.

[1] Abbreviations: 2-Cl-ATP, 2-chloro-ATP; 2-F-ATP, 2-fluoro-ATP; AMV reverse transcriptase, avian myeloblastosis virus reverse transcriptase; FC, flash chromatography; KF, Klenow fragment of *Escherichia coli* DNA polymerase I; MMLV-RT, murine Maloney leukemia virus reverse transcriptase; pol α, DNA polymerase α; pol *β*, DNA polymerase *β*; pol *η*, DNA polymerase *η*; TEAB, triethylammonium bicarbonate (pH 7.5); TEAAc, triethylammonium acetate (pH 6.0); TLC, thin-layer chromatography.

10.1021/bi0490791 CCC: $27.50
Published on Web 08/28/2004

Table 1: Synthetic DNAs of Defined Sequence Used in These Studies

dC_{40}	5′-CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
$d(TC)_{30}$	5′-TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC
$d(C_3T)_{15}$	5′-TCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCC

events are fast (*11*). Once primers reach a length of 8−10 nucleotides, primase transiently stops elongating the primer (*12*, *13*). If pol α and dNTPs are present, pol α rapidly elongates this primer. In the absence of pol α, however, primase will polymerize another 8−9 nucleotides onto the primer. Frequently, primase will synthesize abortive products that do not reach a length that pol α readily elongates (i.e., ≥7 nucleotides long). When this occurs, primase extremely rapidly reinitiates primer synthesis to "try again" to synthesize a primer that pol α can elongate. Thus, inhibitors that decrease the ability of primase to convert a dinucleotide into a primer ≥7 nucleotides long will result in primase synthesizing more moles of primers; albeit they will be only short primers (*14*).

Eukaryotic primase exhibits remarkably low fidelity, with error frequencies occasionally in excess of 1 in 30 (*15*, *16*). Additionally, primase can polymerize multiple, consecutive incorrect NTPs (*16*). Previously, we found that NTP analogues containing bases such as benzimidazole and dichlorobenzimidazole potently inhibited primase (*14*). Surprisingly, however, primase did not incorporate significant amounts of these NTP analogues.

To better understand these data, potentially identify bases that give more potent primase inhibition, and elucidate how primase determines whether to polymerize a NTP, we examined a series of NTP analogues. Primase bound NTP analogues containing bases that varied significantly in terms of shape, hydrophobicity, and potential for forming Watson−Crick hydrogen bonds. However, polymerization only occurred when the incoming NTP and the template base could form at least two Watson−Crick hydrogen bonds. The implications of these data regarding polymerase fidelity mechanisms are discussed.

EXPERIMENTAL PROCEDURES

Materials

Human DNA primase was overexpressed and purified from *Escherichia coli* as previously described (*17*). 2-Cl-ATP and purine riboside were purchased from Sigma. 2-Flouroadenosine was a generous gift from Dr. Jack A. Secrist (Southern Research Institute). 3-Deaza-ATP was a generous gift from Merck Research. (4,6-Difluoro-1-*N*-benzimidazol-1-yl)-*β*-D-ribofuranose was prepared as previously described (*18*). All chemicals used were of the highest commercial grade. Fluoro compounds were obtained from Chempure, HNO_3 was obtained from Merck (d = 1.52, 100%, p.a.), and $(CH_3COO)_2O$ was obtained from Fluka (d = 1.08, p.a.). Silica gel 60 (40−63 μm) for flash chromatography (FC) and silica gel 60 F_{254} TLC plates were from Merck. When necessary, anhydrous solvents were purchased from Fluka or Aldrich and used without further purification. Synthetic oligonucleotides of defined sequence were from Oligos Etc., Inc., or Biosearch Technologies Inc. (Table 1).

Methods

NMR Spectroscopy. NMR spectroscopy was performed on Varian INOVA 400 MHz (^{1}H and ^{19}F) or 500 MHz (^{1}H, GOESY) and Bruker-AM 250, Bruker-WH270 (^{1}H, ^{13}C), and Bruker-AMX400 (^{1}H, ^{13}C) spectrometers.

Synthesis and Characterization of Nucleotide Triphosphates. Nucleotide triphosphates used in these studies were synthesized as described previously (*14*). The synthesis and characterization of nucleosides and nucleotides (numbering corresponds to Chart 1) containing the aromatic bases 5,6-difluorobenzimidazole (**17**), 4,6-difluorobenzimidazole (**18**), 4-(trifluoromethyl)benzimidazole (**19**), 7-(trifluoromethyl)-benzimidazole (**20**), 5-(trifluoromethyl)benzimidazole (**21**), 6-(trifluoromethyl)benzimidazole (**22**), 4-nitrobenzimidazole (**23**), N^7-guanosine (**24**), N^9-purine (**25**), N^9-1-deazapurine (**26**), 4-aminobenzimidazole (**27**), 2-chloroadenosine (2-Cl-ATP), and 2-fluoroadenosine (2-F-ATP) are described in Supporting Information.

Primase Assay. Reactions (10 μL) typically contained 50 mM tris(hydroxylmethyl)aminomethane, pH 7.9 (HCl salt), 5 mM $MgCl_2$, 60 μM ssDNA template (total nucleotide), 0.05 mg mL^{-1} bovine serum albumin,1 mM dithiothreitol, and 100−800 μM [α-^{32}P]NTPs. Reactions were initiated by adding enzyme and incubating at 37 °C for 1 h. Assays were quenched by adding 2.5 volumes of gel-loading buffer (90% formamide), and the products were separated by denaturing polyacrylamide gel electrophoresis (20% polyacrylamide, 8 M urea) and analyzed by phosphorimagery (Molecular Dynamics). IC_{50} values were calculated by nonlinear regression analysis using the curve fit feature of Kaliedagraph. Percent inhibition data were fit to the equation:

$$\% \text{ activity} = V_{\text{uninhibited}}/[1 + ([I]/IC_{50})]$$

RESULTS

To better understand how primase interacts with the base of a NTP with respect to both binding and polymerization, we synthesized a series of NTP analogues containing modified bases and measured their effects on primase (Chart 1). The analogues chosen varied substantially in terms of shape, electron distribution, hydrogen-bonding capacity, and hydrophilicity.

We initially measured the ability of these compounds to inhibit either ATP or GTP polymerization on two templates, poly(dT) and $(dC)_{40}$, respectively. Table 2 shows that all of the compounds inhibited both ATP and GTP polymerization. Since the biological role of primase is to synthesize primers that pol α can elongate, we measured the ability of these analogues to inhibit the synthesis of primers at least 7 nucleotides long (i.e., those that pol α efficiently elongates). In each case, the analogues inhibited ATP polymerization more potently than they inhibited GTP polymerization. A complete kinetic analysis of inhibition by compounds **17**, **18**, and **22** showed that inhibition was competitive with respect to the NTP (data not shown). The similarity in structure of **17**, **18**, and **22** and the rest of the NTP analogues suggests that all of the analogues will inhibit primase competitively with respect to the NTP.

Surprisingly, all of the compounds inhibited primase activity with remarkably similar potency. Changing the shape

Chart 1: Structure of Nucleotide Analogues Used in These Studies

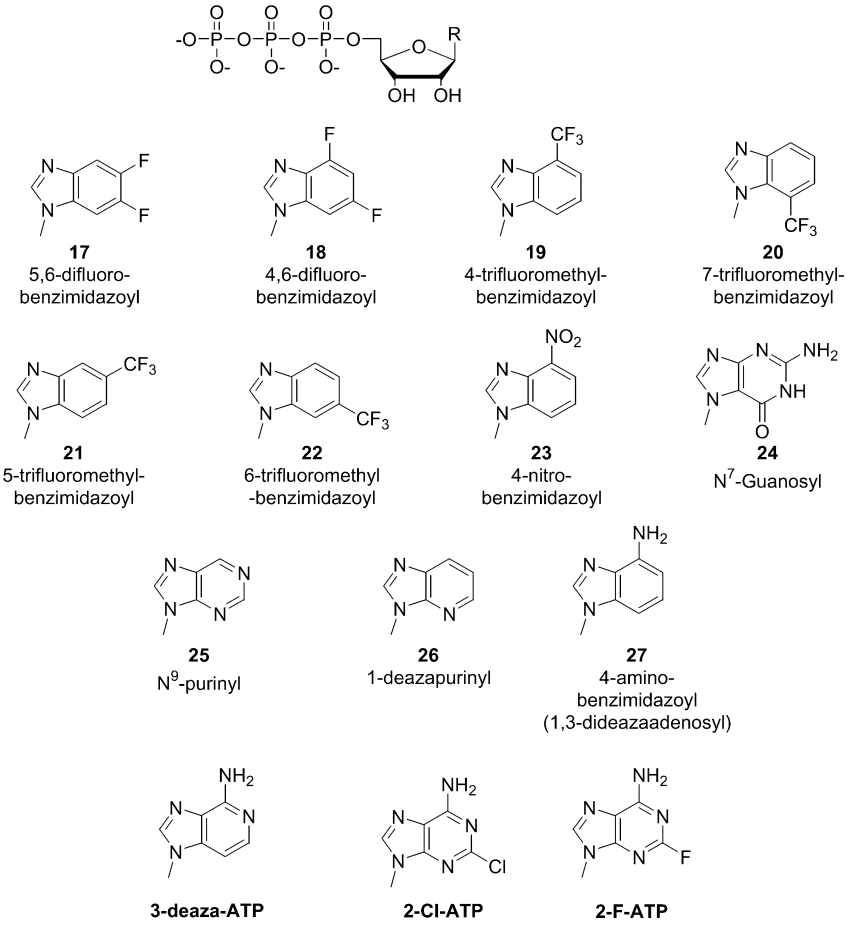

Table 2: IC_{50} Values (μM) of Nucleotide Analogues[a]

compound	IC_{50} vs ATP	IC_{50} vs GTP
17	19 ± 3.5	29 ± 3.7
18	170 ± 53	156 ± 27
19	87 ± 16	82 ± 22
20	21 ± 4.2	15 ± 4.7
21	97 ± 22	39 ± 18
22	22 ± 4.1	22 ± 3.8
23	34 ± 13	19 ± 1.6
24	NM[b]	17 ± 3.7
25	35 ± 7.2	28 ± 4.9
26	18 ± 6.7	32 ± 14
27	54 ± 14	28 ± 7.2
2-Cl-ATP	408 ± 117	196 ± 30
2-F-ATP	800 ± 300	148 ± 45
3-deaza-ATP	700 ± 400	178 ± 29

[a] Experiments were conducted with 60 μM (total nucleotide) ssDNA template and 200 μM [α-^{32}P]NTP. [b] NM = not measured.

of the base from one that closely resembled a normal base (e.g., 4,6-difluorobenzimidazole closely resembles guanine, and 4-aminobenzimidazole is an isostere of adenine) to one whose shape varied substantially from that of a normal base [e.g., 5- and 6-(trifluoromethyl)benzimidazole and guanine attached to ribose via N-7] did not greatly affect inhibition. We measured inhibition by NTPs containing bases capable of forming a single Watson−Crick hydrogen bond to either a template cytidine or thymine (4-aminobenzimidazole and purine, respectively). Curiously, the ability to form a single Watson−Crick hydrogen bond did not enhance the ability of these compounds to inhibit primase relative to bases incapable of forming any Watson−Crick hydrogen bonds. Additionally, changing the base from one that is much more hydrophobic than a normal base (e.g., the fluorinated and trifluoromethylated benzimidazoles) to one that should have similar hydrophilicity as a normal base (guanine attached to ribose via N-7) did not enhance inhibition.

While each of these compounds inhibited primase, primase did not detectably polymerize any of these compounds on either poly(dT) or $(dC)_{40}$. Figure 1 shows that adding increasing concentrations of **23** to primase assays containing [α-^{32}P]GTP and $(dC)_{40}$ inhibits primer synthesis but does not result in the appearance of products of altered electrophoretic mobility. Previously, we have observed that altering the base composition of a primer significantly alters the electrophoretic mobility of the primers, especially the dinucleotides and trinucleotides (ref *16* and see below). Thus, these data suggest that while **23** inhibits primase, it does not serve as a substrate. Similar results were obtained with all of the compounds in Chart 1 using either poly(dT) or $(dC)_{40}$ as the template (data not shown), suggesting that primase did not use any of these NTP analogues as substrates.

To further demonstrate that primase strongly discriminated against polymerization of these NTP analogues, we examined primase activity under conditions where assays lacked one of the "required" NTPs. Primase activity was measured on the template $d(TC)_{30}$ in the presence of only either [α-^{32}P]-ATP or [α-^{32}P]GTP. Indicative of primase's remarkable infidelity, primase synthesized the pppApA and pppGpG dinucleotides in the presence of only ATP or GTP, respectively (Figure 2). Adding **26** to assays resulted in decreased

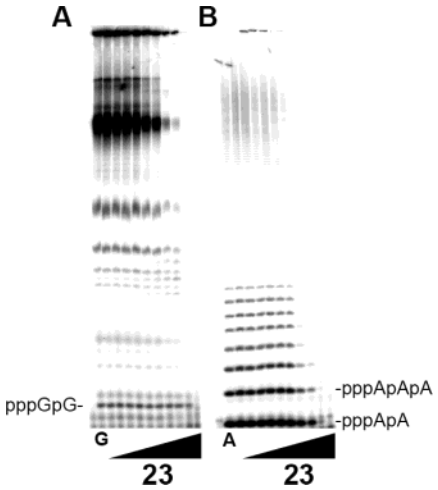

FIGURE 1: Effect of compound **23** on primase activity. Assays were performed with 60 μM template and 200 μM [α-^{32}P]NTP. The electrophoretic mobility of a dinucleotide is noted next to each panel. (A) GTP polymerization on dC_{40} was measured in the presence of 0.2, 0.4, 2.2, 4.4, 22, 44, 220, 440, 2200, and 4400 μM **23**. (B) ATP polymerization on poly(dT) was measured in the presence of **23**. The concentrations were identical to those used in panel A.

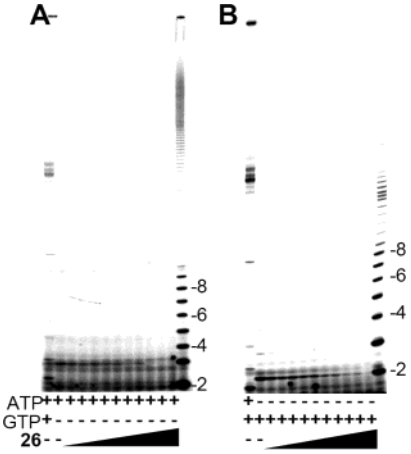

FIGURE 2: Primase incorporates **26** much less efficiently than it misincorporates a natural, incorrect NTP. Assays were performed with 60 μM $d(TC)_{30}$ and 200 μM [α-^{32}P]NTP. The electrophoretic mobility of primers of known length is shown to the right of each panel. (A) Test of **26** as a specific GTP analogue. Lane 1 shows full primer synthesis in the presence of 200 μM GTP. Lanes 2−11 contain no GTP and increasing **26** concentrations from 0 to 1040 μM. (B) Test of **26** as a specific ATP analogue. Lane 1 shows full primer synthesis in the presence of 200 μM ATP. Lanes 2−11 contain no ATP and increasing **26** concentrations from 0 to 1040 μM. In both panels lane 2 shows significant misincorporation of the noncognate purine across from a template C and T, respectively.

amounts of dinucleotide synthesis using either ATP or GTP, but no new products of altered electrophoretic mobility were observed. Similar results were obtained with all of the other analogues in Chart 1 (data not shown), again indicating that primase strongly discriminates against their polymerization.

We also examined polymerization in assays containing the template $d(C_3T)_{15}$ and only [α-^{32}P]GTP. We previously found that as the length of the newly synthesized primer increases, the ability of primase to misincorporate NTPs likewise increases, presumably due to the increased stability of the primer-template duplex (*16*). Since primers will be either 2 or 3 nucleotides long when primase encounters a template

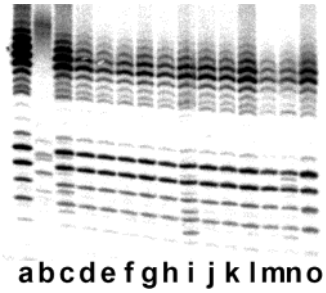

FIGURE 3: Primase fails to incorporate significant amounts of the analogues even under forcing conditions. Assays were performed with 60 μM $d(C_3T)_{15}$ template, 200 μM [α-^{32}P]GTP, and the specified amount of natural or analogue NTP. Assays contained the following: (a) [α-^{32}P]GTP only, (b) 200 μM ATP, (c) 200 μM CTP, (d) 100 μM **19** [4-(trifluoromethyl)benzimidazole], (e) 50 μM **20** [7-(trifluoromethyl)benzimidazole], (f) 200 μM **18** (4,6-difluorobenzimidazole), (g) 100 μM **21** (5-trifluorobenzimidazole), (h) 50 μM **22** [6-(trifluoromethyl)benzimidazole], (i) 100 μM **23** (4-nitrobenzimidazole), (j) 200 μM **27** (4-aminobenzimidazole), (k) 50 μM **17** (5,6-difluorobenzimidazole), (l) 100 μM **26** (1-deazapurinyl), (m) 100 μM **24** (N^7-guanosyl), (n) 100 μM **25** (N^9-purinyl), and (o) 200 μM 2-Cl-ATP.

base for which it lacks the cognate NTP in assays containing the template $d(C_3T)_{15}$ and [α-^{32}P]GTP, we reasoned that primase would be more likely to incorporate the analogues on this template. In the absence of any analogues, primase readily misincorporates GTP and continues primer synthesis past the template thymidylates (Figure 3). Including most of the analogues at levels sufficient to inhibit primase activity by 70−90%, no new products of altered electrophoretic mobility were formed. In two cases, **23** and **25**, trace amounts of new products of altered electrophoretic mobility were formed, indicative of primase polymerizing **23** and **25**. Importantly, the amount of these products was 120−330-fold less than the amount of GTP misincorporation in assays containing only GTP, even though the concentration of **23** and **25** was half that of the incorrect GTP. Thus, while these compounds are substrates, they are extremely poor substrates. The appearance of products of altered electrophoretic mobility in assays containing **23** and **25** is also consistent with previous work showing that altering the base composition of primers will alter their electrophoretic mobility (*16*, *19*).

The almost complete lack of analogue NTP polymerization was particularly surprising in light of the tremendous infidelity of primase (as described below, human primase appears to be as inaccurate as the previously characterized calf primase) (*16*). This lack of incorporation did not result from a lack of analogue NTP binding to primase. These compounds inhibited the incorporation of both correct and incorrect natural NTPs. One common feature of the unnatural bases is the inability to completely satisfy Watson−Crick hydrogen-bonding requirements of the opposite strand. For example, **25** and **27** can satisfy just one of the Watson−Crick hydrogen bonds found in an A·T base pair. In contrast, the natural bases can form Watson−Crick base pairs with both cognate and noncognate template bases by using either the major or minor tautomer of the base.

To further test the hypothesis that formation of Watson−Crick hydrogen bonds is important for primase to polymerize a NTP, we examined polymerization of 2-fluoro-ATP (2-F-ATP), 2-chloro-ATP (2-Cl-ATP), and 3-deaza-ATP. Since the electron-withdrawing properties of the halogens will

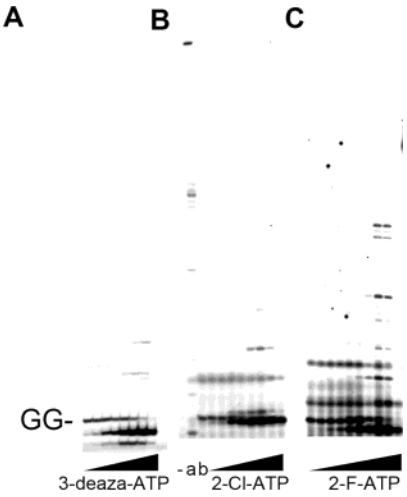

FIGURE 4: Primase polymerizes 3-deaza-ATP, 2-Cl-ATP, and 2-F-ATP. Assays were performed with 60 μM d(TC)$_{30}$ and 200 μM [α-^{32}P]GTP. Lanes are (−) no enzyme control, (a) addition of 200 μM ATP, and (b) GTP only. Panel A shows the increased incorporation of 3-deaza-ATP at concentrations from 0 to 5000 μM. The products were separated on a 40% acrylamide gel in order to resolve the pppGpG dinucleotide from the dinucleotide containing 3-deaza-ATP. Panel B shows the effect of increasing 2-Cl-ATP concentrations from 0 to 3200 μM. Panel C shows the effect of increasing 2-F-ATP concentrations from 0.1 to 1700 μM.

reduce the pK_a of N-1, fluorine at C-2 reduces the pK_a by ca. 3 pH units, while chlorine at C-2 reduces the pK_a of N-1 by ca. 4.4 pH units, the greatly reduced electron density at N-1 is expected to reduce the ability of N-1 to form a Watson−Crick hydrogen bond with a template thymidylate (*20*, *21*). Alternatively, removing N-3 and replacing it with CH at that position increases the pK_a of N-1 by 4.8 pH units (*22*). The increased electron density in this case should at least not affect hydrogen bonding with template thymidylate and potentially could stabilize the A·T base pair.

Both 2-F-ATP and 2-Cl-ATP inhibited primase activity on a poly(dT) template. In assays containing 200 μM ATP, the IC$_{50}$s were 800 and 408 μM, respectively. However, no bands of altered electrophoretic mobility were observed, suggesting that primase does not readily incorporate either compound. Conversely, assays containing 250 or 500 μM 3-deaza-ATP produced several bands of altered electrophoretic mobility, indicating that primase was capable of incorporating this analogue into primers of lengths 2 and 3 (data not shown). Incorporation was also examined on the template d(TC)$_{30}$ in assays containing only [α-^{32}P]GTP and increasing analogue concentration (Figure 4). The pppGpG dinucleotide decreases and a new band of altered electrophoretic mobility appears in each titration, indicating that primase will polymerize both 2-F-ATP, 2-Cl-ATP, and 3-deaza-ATP. Interestingly, whereas primase did not further elongate dinucleotides containing 2-chloroadenine or 3-deaza-ATP, it readily elongated dinucleotides containing 2-fluoroadenine.

We quantified the ability of primase to polymerize ATP, 3-deaza-ATP, 2-F-ATP, and 2-Cl-ATP on the template d(TC)$_{30}$. Primase synthesizes only the pppGpG dinucleotide when assays contain just [^{32}P]GTP. Thus, we can compare the ability of primase to polymerize ATP, 3-deaza-ATP, 2-F-ATP, and 2-Cl-ATP by determining how much of each of these compounds must be present in assays to cause primase to synthesize equal amounts of a dinucleotide containing one guanylate and one adenylate (or adenylate analogue).[2] In assays containing 200 μM GTP, including only 0.8 and 1.8 μM ATP and 3-deaza-ATP, respectively, resulted in primase synthesizing equimolar amounts of the two dinucleotides. In contrast, adding either 2-Cl-ATP or 2-F-ATP to the assays resulted in equimolar synthesis of the two dinucleotides only when the concentration of these NTPs reached 11 and 66 μM, respectively.

Human Primase Readily Misincorporates Natural NTPs. We previously demonstrated that DNA primase from calf thymus had very low fidelity and readily generated a wide spectrum of mismatches (*16*). To ensure that the human enzyme was similar, we briefly surveyed human primase for the ability to generate a range of mismatches. Figures 2 and 3 show that human primase also generates a wide spectrum of mismatches, including A·C (Figure 2, panel A), G·T (Figure 2, panel B), and C·T or C·C (Figure 3, lane c). Importantly, these mismatches were generated at relatively low ratios of the incorrect to correct NTP. Thus, both human and calf primase have low fidelity.

DISCUSSION

We examined the ability of human primase to bind and polymerize both NTP analogues and incorrect, natural NTPs. Primase bound a wide variety of NTP analogues with similar affinity. However, whereas primase very readily misincorporated the natural NTPs, it discriminated very strongly against polymerization of the NTP analogues. The only exceptions were those NTPs that could still form Watson−Crick hydrogen bonds with both the exocyclic group at C-4 and N(H)-3 of the pyrimidine.

Human primase readily misincorporates the natural NTPs (*12*). It generated G·T, A·C, C·T, and C·C mismatches, similar to the ability of calf thymus primase to generate a wide variety of mismatches. Importantly, human primase generated these mismatches at relatively low concentrations of the noncognate NTP, demonstrating that the human enzyme likely has high error frequencies. Consistent with this idea, we previously found that human primase can incorporate consecutive incorrect NTPs (*19*).

Human primase misincorporates NTPs during initiation of new primers much more readily than calf primase. In the presence of only either GTP or ATP, initiation of primer synthesis on the template d(TC)$_{30}$ requires misincorporation (see above). Even at elevated concentrations of just ATP or GTP, the calf enzyme did not generate detectable amounts of dinucleotide or longer products (*16*). In contrast, human primase synthesized large amounts of pppNpN dinucleotide in the presence of only 200 μM ATP or GTP. The efficiency of misincorporation during dinucleotide synthesis can be estimated from the relative amount of ATP needed to decrease GTP misincorporation by 50%. In the presence of both ATP and GTP, primase can synthesize either the pppGpG dinucleotide via misincorporation or a correct dinucleotide containing both ATP and GTP. Since primase generates 50% correct and 50% incorrect product at a GTP:

[2] Quantitation of ATP incorporation requires comparison of pppGpG to dinucleotide plus longer products since AG/GA dinucleotide can be extended whereas the GG dinucleotide is poorly extended to longer products.

ATP ratio of 250, primase misincorporates GTP only ca. 250-fold less well than it incorporates the correct NTP.[3]

Even though primase readily incorporates the natural NTPs opposite noncognate template bases, it is remarkably resistant to polymerizing NTP analogues containing unnatural bases. Of the base analogues that we have tested, primase only polymerized four of them at rates similar to or faster than it polymerized an incorrect, natural NTP: hypoxanthine, 3-deazaadenine, 2-chloroadenine, and 2-fluoroadenine. On the template $d(TC)_{30}$, primase would prefer to misincorporate GTP (or ATP) at least 10-fold more readily than it will misincorporate any of the other NTP analogues other than 3-deaza-ATP.[4] In assays containing [α-^{32}P]GTP and the template $d(C_3T)_{15}$, primase could polymerize trace amounts of **23** and **25**, as evidenced by the appearance of new bands of altered mobility. However, this only occurred at concentrations of **23** and **25** that significantly inhibited primase activity, again indicating that primase greatly prefers to misincorporate GTP opposite a template deoxythymidylate rather than incorporate these analogues.

The almost complete absence of analogue polymerization argues that shape complementarity between the incoming NTP and template base does not play a significant role in determining whether primase will polymerize the NTP. The base analogues that primase does not incorporate range in size from slightly smaller than a canonical purine (benzimidazole and purine) to slightly larger than a normal purine (trifluorobenzimidazole and nitrobenzimidazole). Furthermore, primase did not significantly polymerize analogues whose shape is virtually identical (4-aminobenzimidazole vs adenine) or only slightly altered (4,6-difluorobenzimidazole vs guanine or 2-fluoroadenine) from the shape of NTPs that primase polymerizes. Additional evidence that shape is not a critical factor comes from the facile misincorporation of the canonical NTPs, since these base pairs will necessarily lack shape complementarity.

The hydrophobicity of the base of the incoming NTP also does not appear to play a major role in determining whether primase polymerizes the incoming NTP. Potentially, the hydrophobic nature of the base could enhance polymerization either via stacking interactions with the neighboring base and/or direct interactions with the enzyme. However, primase did not polymerize the benzimidazole-derived base analogues, even though they are much more hydrophobic than either adenine or guanine. We also considered the possibility that these analogues were too hydrophobic by synthesizing and testing 7-β-D-GTP. Since the base is identical to that found on the normal 9-β-D-GTP, the hydrophobicity and stacking ability of 7-β-D-GTP and 9-β-D-GTP should be similar. The lack of incorporation of 7-β-D-GTP suggests that the lack of incorporation of the benzimidazole-derived bases did not result from their hydrophobicity.

Alternatively, formation of hydrogen bonds between the Watson−Crick hydrogen-bonding groups at N-1 and C-6 of the incoming purine NTP and the template base (N-3 and the exocyclic group at C-4 of a pyrimidine) may be essential for primase to polymerize the NTP. In the entire set of NTPs tested, primase only readily polymerized those that retained Watson−Crick hydrogen-bonding groups at the equivalent of both purine N-1 and C-6 (i.e., the four natural NTPs, ITP, 3-deaza-ATP, 2-F-ATP, and 2-Cl-ATP). Remarkably, primase did not even polymerize NTPs whose bases lacked only one of these hydrogen-bonding groups, purine and 4-aminobenzimidazole. Formation of a hydrogen bond between the C-2 oxygen of a pyrimidine and a group on the purine is not critical for polymerization since primase can generate A·T, G·C, and I·C base pairs.

Consistent with the importance of Watson−Crick hydrogen-bonding groups, primase polymerized 2-F-ATP and 2-Cl-ATP significantly less efficiently than it polymerized ATP. The electron-withdrawing properties of the fluorine and chlorine significantly decrease the pK_a of N-1 (*20*, *21*). Importantly, this decreased proton affinity at N-1 should inhibit the formation of hydrogen bonds at N-1. If the hydrogen-bonding capacity of N-1 is critical for primase to polymerize a purine NTP, then primase should polymerize 2-F-ATP and 2-Cl-ATP less efficiently than ATP, and indeed, the observed result. It should be noted that we cannot rule out an alternative possibility, namely, that the extra mass at C-2 inhibits polymerization of these compounds.

Further evidence in support of the critical role of Watson−Crick hydrogen bonds comes from the facile polymerization of 3-deaza-ATP. In contrast to the effects of F or Cl at the 2-position, replacing N-3 with a CH greatly increases the pK_a of N-1 (*22*). This change should enhance the ability of N-1 to form a hydrogen bond, and indeed, primase polymerizes 3-deaza-ATP almost as well as ATP.

If formation of hydrogen bonds between N-1 and the exocyclic group at C-6 and the complementary positions on a template pyrimidine (N-3 and the exocyclic group at C-4) is so important for polymerization, why does primase misincorporate NTPs at such a high rate? The natural bases typically can exist in both a major and minor tautomer (Figure 5), and using these two tautomers a single base can form Watson−Crick hydrogen bonds with all four natural bases. For example, the major tautomer of guanine can form hydrogen bonds with both cytosine and adenine (Figure 5), while the minor tautomer can form hydrogen bonds with both thymine and guanine. In order for tautomerization to account for the high rate of NTP misincorporation, primase would likely have to stabilize the minor tautomer. In aqueous solution, the minor tautomer of the natural bases is energetically much less stable than the major tautomer (*23*). However, previous studies have shown that the equilibrium between the major and minor tautomers can be altered by varying the solvent (*24*, *25*), thereby suggesting that an enzyme active site could likewise alter this equilibrium. Support for the hypothesis that an enzyme active site can alter the tautomeric state of a "base pair" comes from recent work of Johnson and Beese (*26*). They examined the structure of a primer-template containing a G·T mismatch bound to the thermophillic *Bacillus* DNA polymerase I. Depending upon the position of the mismatch within the primer-template, the structure of the G·T mismatch suggested that the

[3] The relative fold difference should be considered only an estimate. On the template $d(TC)_{30}$, formation of a pppGpG dinucleotide could result from either the GTP that will become the 5′-terminal nucleotide being mismatched or the GTP that will become the second nucleotide of the primer being mismatched. Because human primase does not initiate synthesis at defined sites, it is not possible to differentiate between these possibilities.

[4] This value is based on the assumption that we could detect a new band that had an intensity 10-fold less than the intensity of the pppGpG (or pppApA) dinucleotide. In practice, we can routinely detect bands with an intensity 20−30-fold less than the pppGpG and pppApA bands.

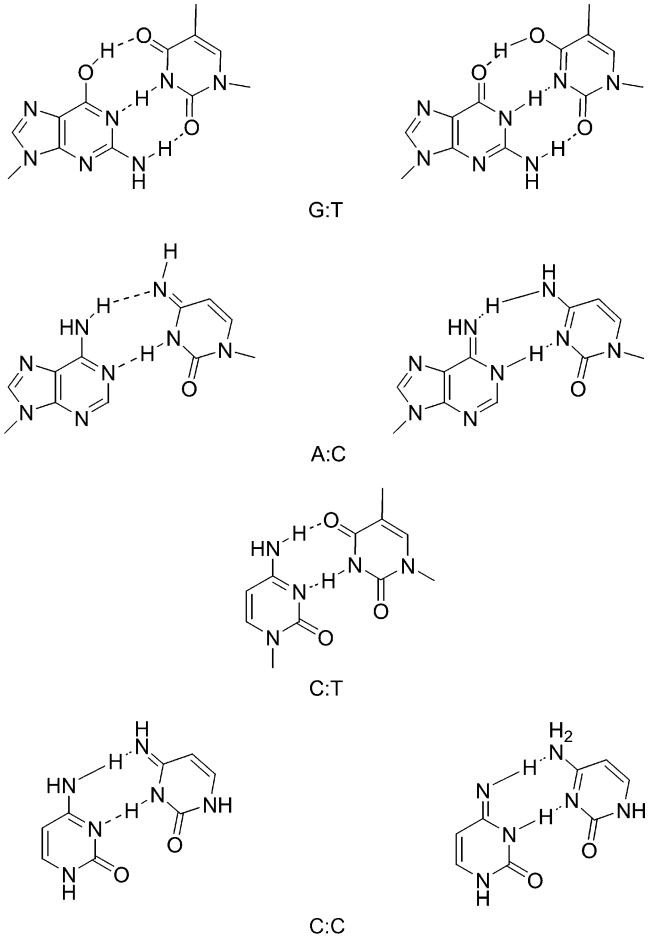

FIGURE 5: Possible minor tautomer hydrogen bond pairs generated in the active site of DNA primase. The minor tautomer is displayed as either template base or incoming base, since the enzyme could isomerize either base to stabilize a mismatch.

tautomerization state of the bases varied. Thus, a nucleotide selection model that emphasizes the formation of Watson−Crick hydrogen bonds could account for both the ability of primase to discriminate against NTPs whose bases lack Watson−Crick hydrogen-bonding groups and the high rate of NTP misincorporation.

If, as seems likely, formation of Watson−Crick hydrogen bonds between the incoming NTP and the template base is critical for primase to polymerize the NTP, how does the enzyme detect the formation of a hydrogen bond? One possibility would be for the enzyme to sense the movement of the hydrogen-bonding proton from very near the template or NTP base to a position intermediate between the two. However, we are not aware of any examples of this strategy, and any attempt by the enzyme to measure a hydrogen bond in this manner would at least partially disrupt the hydrogen bond. Alternatively, and more likely, the presence of unsatisfied Watson−Crick hydrogen bond donors and acceptors in the enzyme active site might prevent NTP polymerization. For example, if the active site of primase were relatively hydrophobic, an unsatisfied hydrogen-bonding group on a template or NTP base would be energetically unfavorable. This could preclude proper positioning and/or binding of the incoming NTP, thereby preventing incorporation. Only upon formation of the Watson−Crick hydrogen bonds would the incoming NTP assume the proper orientation to allow polymerization.

Formation of a hydrogen bond between primase and the minor groove of the 3′-terminal base of the primer strand may be critical for efficient addition of the next correct NTP. While the effects of either 2-Cl or 2-F on the electronic properties of N-3 have not been measured, we anticipate that these halogens will reduce the ability of N-3 to form hydrogen bonds. The higher efficiency with which primase elongates primers containing 2-fluoroadenine would result from the weaker electron-withdrawing effects of the F. While there is no direct structural evidence for such a hydrogen-bonding interaction in primase, insights can be obtained by examining the evolutionarily related enzyme pol β (*27*). Tyr271 of pol β appears to form a hydrogen bond with the minor groove base at the primer terminus (*28*). If primase also normally forms a similar hydrogen bond, it would not be unexpected that interfering with this hydrogen bond should interfere with primer elongation.

The role of Watson−Crick hydrogen bonds during dNTP polymerization by DNA polymerases appears to vary substantially depending upon the polymerase. Some polymerases, including pol α, T7 DNA polymerase, and the Klenow fragment of DNA pol I (*E. coli*), very readily incorporate dNTPs bearing nonpolar base analogues (*4*, *29*). Thus, these polymerases clearly do not require the formation of Watson−Crick hydrogen bonds in order to polymerize the dNTP. Other polymerases, including pol β, AMV reverse transcriptase, MMLV reverse transcriptase, and pol η, do not readily polymerize dNTPs containing hydrophobic bases incapable of forming Watson−Crick hydrogen bonds (*29*). While there are multiple reasons why a polymerase might not polymerize an analogue, one possibility is that formation of Watson−Crick hydrogen bonds is essential for nucleotide polymerization. Indeed, it was suggested that this might be the reason pol η did not incorporate a dNTP bearing the base difluorotoluene (*30*).

Primase and pol α, two enzymes that form a tightly bound complex, appear to use fundamentally different mechanisms in determining whether to polymerize a (d)NTP. Whereas primase effectively discriminates against nucleotides that cannot form Watson−Crick hydrogen bonds, pol α readily polymerizes a variety of nucleotides whose bases neither can form Watson−Crick hydrogen bonds nor closely resemble the shape of a canonical base (*9*). Indeed, pol α generally polymerizes these dNTP analogues 1−3 orders of magnitude more rapidly than it polymerizes an incorrect natural dNTP, and in some cases the rates of analogue polymerization approach those for a correctly base-paired natural dNTP. This included many of the bases that primase did not polymerize [benzimidazole, 5-(trifluoromethyl)benzimidazole, etc.]. Similarly, Klenow fragment will frequently polymerize dNTPs bearing nonpolar base analogues whose shape does not closely resemble a canonical base at rates that approach those of a normal dNTP. *Importantly, these data indicate that (d)NTP polymerases have evolved at least two, and probably more, mechanisms to discriminate between correct and incorrect nucleotides*.

Both primase and pol β are members of the class X superfamily of nucleotide polymerases. Interestingly, neither enzyme readily polymerizes hydrophobic (d)NTP analogues incapable of forming Watson−Crick hydrogen bonds (*29*). While the precise mechanism by which these two enzymes discriminate between correct and incorrect have not been

defined, these similarities suggest that they will likely use similar mechanisms. Curiously, however, these two enzymes exhibit very different levels of fidelity. Pol β discriminates quite effectively against incorrect dNTPs, while primase discriminates very poorly against noncognate NTPs. Indeed, the fidelity of primase is more similar to the lesion bypass/error prone Y class polymerases. Thus, it will be interesting to ask if differences in fidelity mechanisms reflect the evolutionary family to which the polymerase belongs (A, B, etc.) or if the fidelity mechanism used by a given polymerase reflects its level of fidelity (high or low).

SUPPORTING INFORMATION AVAILABLE

Detailed synthesis for the nucleosides 1′-deoxy-1′-[4-(trifluoromethyl)-1-*N*-benzimidazol-1-yl]-β-D-ribofuranose, 1′-deoxy-1′-[4-(trifluoromethyl)-3-*N*-benzimidazol-1-yl]-β-D-ribofuranose, 1′-deoxy-1′-[5-(trifluoromethyl)-1-*N*-benzimidazol-1-yl]-β-D-ribofuranose, 1′-deoxy-1′-[5-(trifluoromethyl)-3-*N*-benzimidazol-1-yl]-β-D-ribofuranose, 1′-deoxy-1′-(4-nitro-1-*N*-benzimidazol-1-yl)-β-D-ribofuranose, and 1′-deoxy-1′-(5,6-difluoro-1-*N*-benzimidazol-1-yl)-β-D-ribofuranose and mass spectral characterization of nucleotide triphosphates. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kornberg, A., and Baker, T. (1992) *DNA Replication*, 2nd ed., W. H. Freeman, San Francisco.
2. Kool, E. T. (1998) *Biopolymers 48*, 3−17.
3. Kunkel, T. A., and Bebenek, K. (2000) *Annu. Rev. Biochem. 69*, 497−529.
4. Kool, E. T. (2001) *Annu. Rev. Biophys. Biomol. Struct. 30*, 1−22.
5. Kunkel, T. A. (2004) *J. Biol. Chem. 279*, 16895−16898.
6. Goodman, M. F. (2002) *Annu. Rev. Biochem. 71*, 17−50.
7. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry 30*, 526.
8. Wu, Y., Ogawa, A. K., Berger, M., McMinn, D. L., Schultz, P. G., and Romesberg, F. E. (2000) *J. Am. Chem. Soc. 122*, 7621−7632.
9. Chiaramonte, M., Moore, C. L., Kincaid, K., and Kuchta, R. D. (2003) *Biochemistry 42*, 10472−10481.
10. Kincaid, K., Beckman, J., Zivkovic, A., Halcomb, R. L., Engels, J. W., and Kuchta, R. D. (2004) *Biochemistry* (submitted for publication).
11. Sheaff, R. J., and Kuchta, R. D. (1993) *Biochemistry 32*, 3027−3037.
12. Arezi, B., and Kuchta, R. D. (2000) *Trends Biochem. Sci. 25*, 572−576.
13. Sheaff, R. J., Kuchta, R. D., and Ilsley, D. (1994) *Biochemistry 33*, 2247−2254.
14. Moore, C. L., Chiaramonte, M., Higgins, T., and Kuchta, R. D. (2002) *Biochemistry 41*, 14066−14075.
15. Zhang, S. S., and Grosse, F. (1990) *J. Mol. Biol. 216*, 475−479.
16. Sheaff, R. J., and Kuchta, R. D. (1994) *J. Biol. Chem. 269*, 19225−19231.
17. Arezi, B., Kirk, B. W., Copeland, W. C., and Kuchta, R. D. (1999) *Biochemistry 38*, 12899−12907.
18. Parsch, J., and Engels, J. W. (2000) *Helv. Chim. Acta 83*, 1791−1808.
19. Zerbe, L. K., Goodman, M. F., Efrati, E., and Kuchta, R. D. (1999) *Biochemistry 38*, 12908−12914.
20. Broom, A. D., Amarnath, V., Vince, R., and Brownell, J. (1979) *Biochim. Biophys. Acta 563*, 508−517.
21. Major, D. T., Laxer, A., and Fischer, B. (2002) *J. Org. Chem. 67*, 790−802.
22. Ryder, S. P., Oyelere, A. K., Padilla, J. L., Klostermeier, D., Millar, D. P., and Strobel, S. A. (2001) *RNA 7*, 1454−1463.
23. Cullis, P. M., and Wolfenden, R. (1981) *Biochemistry 20*, 3024−3028.
24. Blas, J. R., Luque, F. J., and Orozco, M. (2004) *J. Am. Chem. Soc. 126*, 154−164.
25. Beak, P., and Fry, F. S. (1973) *J. Am. Chem. Soc. 95*, 1700−1702.
26. Johnson, S. J., and Beese, L. S. (2004) *Cell 116*, 803−816.
27. Kirk, B. W., and Kuchta, R. D. (1999) *Biochemistry 38*, 7727−7736.
28. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science 264*, 1891−1903.
29. Morales, J. C., and Kool, E. T. (2000) *J. Am. Chem. Soc. 122*, 1001−1007.
30. Washington, M. T., Helquist, S. A., Kool, E. T., Prakash, L., and Prakash, S. (2003) *Mol. Cell. Biol. 23*, 5107−5112.

BI0490791