Probing minor groove recognition contacts by DNA polymerases and reverse transcriptases using 3-deaza-2′-deoxyadenosine

Cynthia L. Hendrickson, Kevin G. Devine and Steven A. Benner*

Department of Chemistry and Department of Anatomy and Cell Biology, University of Florida, Gainesville, FL 32611-7200, USA

ABSTRACT

Standard nucleobases all present electron density as an unshared pair of electrons to the minor groove of the double helix. Many heterocycles supporting artificial genetic systems lack this electron pair. To determine how different DNA polymerases use the pair as a substrate specificity determinant, three Family A polymerases, three Family B polymerases and three reverse transcriptases were examined for their ability to handle 3-deaza-2′-deoxyadenosine (c³dA), an analog of 2′-deoxyadenosine lacking the minor groove electron pair. Different polymerases differed widely in their interaction with c³dA. Most notably, Family A and Family B polymerases differed in their use of this interaction to exploit their exonuclease activities. Significant differences were also found within polymerase families. This plasticity in polymerase behavior is encouraging to those wishing to develop a synthetic biology based on artificial genetic systems. The differences also suggest either that Family A and Family B polymerases do not share a common ancestor, that minor groove contact was not used by that ancestor functionally or that this contact was not sufficiently critical to fitness to have been conserved as the polymerase families diverged. Each interpretation is significant for understanding the planetary biology of polymerases.

INTRODUCTION

An interesting area of contemporary nucleic acid research seeks to develop artificial genetic systems composed of non-standard nucleotides (1–3). Molecular recognition is well established in several of these (4) and some artificial genetic systems are presently exploited in FDA-approved tests for monitoring the load of HIV and hepatitis virus in individual patients in the clinic. More broadly, artificial genetic systems offer the next step towards the development of a synthetic biology, where artificial chemical systems capable of Darwinian evolution generate, in laboratory selection environments, advanced biological behaviors, including inheritance and evolution (5).

At present and undoubtedly for some time into the future, synthetic biologists will manipulate artificial genetic systems using enzymes that have evolved in nature to handle standard nucleic acids. This is especially true for polymerases. For this reason, information is needed concerning features of standard nucleic acids that are recognition elements for natural polymerases. This information, in turn, will influence the design of artificial genetic systems.

One possible recognition element is the unshared pair of electrons that each of the four standard nucleobases in natural DNA (adenine, guanine, cytosine and thymine) presents to the minor groove of the double helix (6). More formally called 'electron density', this pair is carried by N^3 in the standard purines and the keto group at position 2 in the standard pyrimidines.

This pair of electrons is a hydrogen bond acceptor and can therefore interact with a hydrogen bond donating group presented by a polymerase to the minor groove. Because it is present in all standard nucleobases, this pair can be the basis of a 'common site' interaction between the polymerase and whatever nucleobase is present in the active site at any point in the polymerase catalytic cycle. Indeed, the electron pair appears to be the only such contact that all standard nucleobases can make in the same way. Therefore, the interaction between the unshared electron pair and the polymerase is expected to be used by polymerases generally to enforce the geometry of the base pair without discriminating between different substrates. This might be a key to polymerase fidelity.

Crystallography has found evidence for such hydrogen bonding interactions in the minor groove for various polymerases (Fig. 1), including Taq (7,8) and Bst (9) from the A evolutionary family of polymerases and RB69 from the B family of polymerases (10). The residues from Taq and Bst that form hydrogen bonds with the minor groove are conserved within most known Family A DNA polymerases

*To whom correspondence should be addressed. Tel: +1 352 392 7773; Fax: +1 352 392 7918; Email: benner@chem.ufl.edu
Present addresses:
Kevin G. Devine, Department of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland
Cynthia L. Hendrickson, Nuclear Medicine Department, Warren G. Magnuson Clinical Center, National Institutes of Health, Bethesda, MD 20892, USA
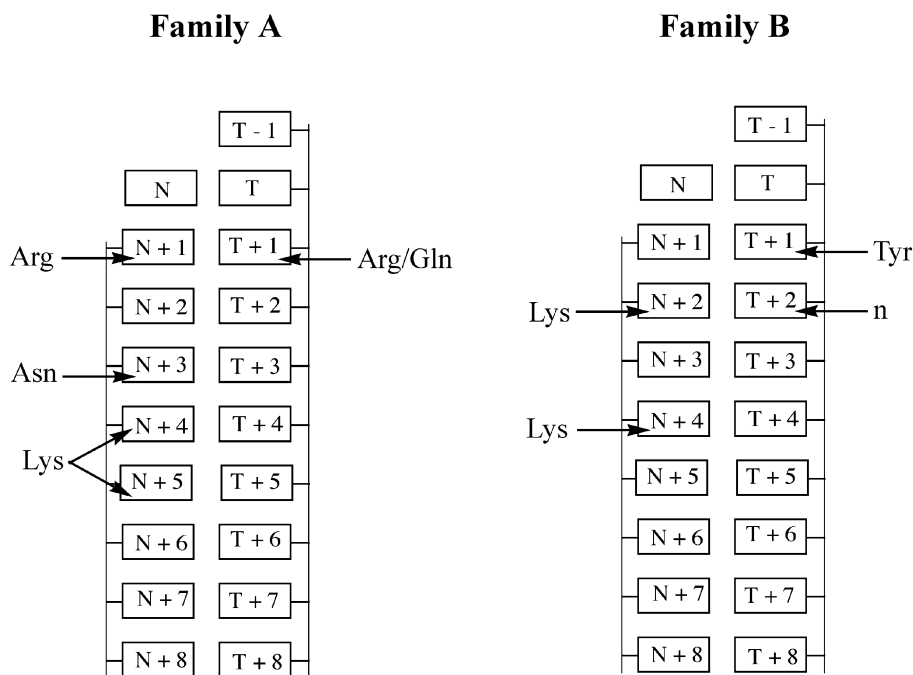
Family A   **Family B**

Figure 1. Interactions between Family A and Family B DNA polymerase residues and the minor groove unshared pair of electrons suggested by x-ray crystallography. Family A DNA polymerases form hydrogen bonds between an Arg (573 in Taq), an Asn (583 in Taq), and a Lys (540 in Taq) and the minor groove of the elongating strand, and hydrogen bonds between an Arg and a Gln with the template strand (7,8). Family B polymerases form hydrogen bonds between a Lys (706 in RB69) and position N+2 of the elongating strand, via a water molecule between another Lys or Arg (Lys[734] in RB69) and the elongating strand at N+4, and between a Tyr (567 in RB69) and the template strand. A main-chain amide (n; at position 394 in RB69) also interacts with the minor groove of the template strand (10).

(11). These consist of (i) an Arg (position 573 in Taq) that forms a hydrogen bond with a nucleotide immediately after incorporation and its template complement (N+1, T+1) (Fig. 1), (ii) a Gln (position 754 in Taq) that can also form a hydrogen bond with the template at position T+1, (iii) an Asn (position 583 in Taq) that forms a hydrogen bond with the elongating DNA strand three sites from the site of triphosphate addition (N+3) and (iv) a Lys (position 540 in Taq) that can form hydrogen bonds with the nucleotides four and five positions away from the site of triphosphate addition (N+4, N+5).

Experiments that mutate Family A polymerases also suggest a minor groove interaction. Mutation of the Arg at position 668 of Klenow fragment, which forms a hydrogen bond with N+1 and T+1, reduces catalytic efficiency (12,13) and the polymerase's ability to recognize mismatches (14).

Minor groove contacts are also suggested for Family B polymerases by crystallography (Fig. 1). For example, a Lys at position 706 in RB69 forms a hydrogen bond with a nucleotide on the elongating strand two positions from the site of triphosphate addition (N+2) (Fig. 1). A second Lys (position 734 in RB69) forms a bond in the minor groove to a nucleotide four positions away from the addition site via a water molecule (N+4). A Tyr (position 567 in RB69) forms a bond to the minor groove of the template strand via a water molecule one position away from the site of addition (T+1). The Lys that interacts with the unshared electron pair carried by nucleobase N+1 and the Tyr that interacts with nucleobase T+1 are conserved among Family B polymerases. The Lys that interacts in the minor groove with nucleobase N+4 is an Arg

in most archaebacterial and mammalian Family B DNA polymerases (11,15).

Reverse transcriptases (RTs) appear to be analogous in their fold to Family A and Family B DNA polymerases (16–18). The possible minor groove contacts appear to be different, however. The crystal structures for HIV-RT (19,20) and Moloney murine leukemia virus RT (21) are particularly relevant. The structures of HIV-RT suggest that Tyr183 forms a hydrogen bond with the minor groove of the elongating strand at a nucleotide two positions from the site of addition. This residue is conserved among the RTs, including MMLV-RT, but the crystal structure of MMLV-RT does not show any interaction between the aligned Tyr and the DNA. Interestingly, the crystal structure of MMLV-RT indicates that a different residue, Arg116, interacts with the minor groove electrons of the template strand at position T+2 and T+3. This residue is also conserved among RTs, including HIV-RT (Arg78); the HIV-RT residue appears to interact with the phosphate backbone of the template strand at position T+1, however (21).

Further evidence for minor groove hydrogen bonding comes from the DNA analogs lacking either the purine N^3 or the pyrimidine 2-keto group as polymerase substrates. Cosstick *et al.*, for example, reported that 3-deaza-2′-deoxyadenosine triphosphate did not support PCR with Taq polymerase (22). Likewise, a thymidine analog lacking the 2-keto group did not support PCR and was poorly incorporated by Taq and Klenow fragment in primer extension experiments (23). More recently, Spratt reported a decrease in the rate of incorporation of dNTPs by the Klenow fragment of DNA

Table 1. Oligonucleotide sequences

Oligonucleotide	Sequence
SS primer	5′-GCGTAATACGACTCACTATAG-3′
T template	3′-CGCATTATGCTGAGTGATATCTGCGCAGAG-5′
C template	3′-CGCATTATGCTGAGTGATATCCGCGCAGAG-5′
G template	3′-CGCATTATGCTGAGTGATATCGGCGCAGAG-5′
A template	3′-CGCATTATGCTGAGTGATATCAGCGCAGAG-5′
RS primer	5′-GCGTAATACGACTCACTATG-3′
RS T template	3′-CGCATTATGCTGAGTGATACCTGCGCTGTG-5′

polymerase I from *Escherichia coli* when 3-deaza-2′-deoxyguanosine was present at the primer terminus or in the template opposite either the terminal primer base or the incoming dNTP (13).

More significantly modified nucleotide analogs have been used to probe polymerase–substrate interactions (13,24–26). Occasionally, the heterocycle is altered to destroy inter-strand hydrogen bonding potential, as well as the potential of the polymerase to interact with the minor groove electron pair. While the larger number of modifications in some of these analogs makes interpretation difficult, these studies have suggested that the interaction between a polymerase and an unshared pair of electrons in the minor groove is as essential to DNA replication as Watson–Crick hydrogen bonding between the nucleotides (26,27).

If polymerases interact with a nucleobase hydrogen bond acceptor group as a way to identify or exclude substrates, the interaction is problematic for those seeking to expand the genetic alphabet artificially and develop a synthetic biology (5). Many (but not all) heterocycles that enable expanded genetic information systems lack this hydrogen bond accepting pair.

To assess whether DNA polymerases use the hydrogen bond accepting pair as a substrate specificity determinant and the extent to which this use is conserved, three Family A DNA polymerases, three Family B DNA polymerases and three RTs were examined for their ability to incorporate 3-deaza-2′-deoxyadenosine triphosphate (c^3dATP) into a growing DNA strand by template-directed polymerization. c^3dATP is a good steric analog of 2′-deoxyadenosine triphosphate and presents the same hydrogen bonding pattern to a complementary strand as dATP. c^3dATP presents a C-H bond to the minor groove, however, allowing it to serve as a relatively specific probe for this interaction.

MATERIALS AND METHODS

Deoxyribonucleotides and oligonucleotides

2′-Deoxy-3-deazaadenosine (c^3dA) was prepared by the method of Cosstick *et al.* (22) and converted to the triphosphate by the method of Ludwig and Eckstein (28). Oligonucleotides (Table 1) were purchased from Integrated DNA Technologies (Coralville, IA). Radiolabeled primer was prepared by incubating primer (1 nmol), [γ-^{32}P]ATP (50 µCi), kinase buffer (5 µl), T4 polynucleotide kinase (50 U) and H$_2$O to a final volume of 50 µl for 30 min at 37°C. The kinase was inactivated by incubation at 80°C for 10 min. The radiolabel primer was purified using a Qiagen Qiaquick Nucleotide

Removal Kit and the primer was eluted from the column in H$_2$O (100 µl).

Polymerases

Taq, Klenow fragment, Klenow fragment (exo⁻), Bst, T7, GB-D (Deep Vent), GB-D (exo⁻) [Deep Vent (exo⁻)], Tli (Vent) and Tli (exo⁻) [Vent (exo⁻)] polymerases were purchased from New England Biolabs. Tfl and Tth polymerases were purchased from Promega. Pfu and Pfu (exo⁻) polymerases were purchased from Stratagene.

Primer extension assays with c^3dATP

SS primer (250 pmol ^{32}P-labeled, 2 nmol unlabeled) (Table 1) and T template (3 nmol) (c^3dATP assay) were annealed by heating to 95°C for 5 min, then cooling to room temperature over 1 h. Single nucleotide incorporation reactions were initiated by adding polymerase (1 µl) to reaction mixtures (9 µl) containing reaction buffer (provided with the polymerases by the suppliers; see Supplementary Material for a list of their contents), annealed primer and template (1.5 µM primer, 2 µM template) and either dATP or c^3dATP (100 µM). The reactions were incubated at 72 [with GB-D (exo⁻), Pfu (exo⁻), Taq, Tli (exo⁻)], 65 (with Bst) or 37°C [with Klenow (exo⁻)] for 2 min, followed by the addition of 10 µl of quenching buffer (95% formamide, 40 mM EDTA, 0.05% xylene cyanol and bromophenol blue). To observe primer extension beyond the incorporation of c^3dATP, the primer extension reactions were repeated, except following the 2 min incubation with the single dNTP, 1 µl of dNTP mix (1 nM dCTP, dGTP and TTP) was added, followed by another 2 min incubation prior to the addition of quenching buffer (10 µl).

Polymerase dissociation assays with c^3dATP

RS T template (600 pmol) was annealed to ^{32}P-labeled RS primer (1 nmol) in H$_2$O (final volume 100 µl) by heating to 95°C for 5 min, then cooling to room temperature over 1 h. A DNA trap was prepared in the same way with RS T template (30 nmol) and unlabeled RS primer (30 nmol) in H$_2$O (final volume 60 µl).

To determine the effectiveness of the trap, test reactions were run [1 µl primer/template, 1 µl reaction buffer (provided with the polymerases by the suppliers), 3 µl DNA trap, 2 µl of 1 mM dNTPs, 2 µl H$_2$O] and initiated by the addition of 1 µl polymerase and incubated at either 72 [Taq, Tli (exo⁻), GB-D (exo⁻) and Pfu (exo⁻)], 65 (Bst) or 37°C [Klenow fragment (exo⁻)] for 20 s before the addition of quenching buffer (10 µl).

Reactions without trap were prepared [1 µl primer/template, 1 µl reaction buffer (provided with the polymerases by the suppliers), 2 µl 1 mM dNTPs (dATP, dCTP, dGTP, TTP or c^3dATP, dCTP, dGTP, TTP), 5 µl H$_2$O], initiated by the addition of 1 µl of polymerase, incubated at either 72, 65 or 37°C for 20 s and stopped by the addition of quenching buffer (10 µl).

For the reactions with trap, polymerase (1 µl), primer/template (1 µl) and reaction buffer (1 µl, provided with the polymerases by the suppliers) were incubated together (1 min, 25°C) then added to the reaction mixture [2 µl of 1 mM dNTPs (dATP, dCTP, dGTP, TTP or c^3dATP, dCTP, dGTP, TTP), 3 µl DNA trap, 2 µl H$_2$O], followed by incubation for 20 s and addition of quenching buffer (10 µl).

Mismatch incorporation assays

SS primer (250 pmol ^{32}P-labeled, 2 nmol unlabeled) and either T, C, G or A template (3 nmol) were annealed by heating to 95°C for 5 min, then cooling to room temperature over 1 h.

Single nucleotide incorporation reactions were initiated by adding polymerase (1 μl) to reaction mixtures (9 μl) containing 1 μl of reaction buffer (provided with the polymerases by the suppliers; see Supplementary material), annealed primer and template (1.5 μM primer, 2 μM template) and either dATP, c[3]dATP or TTP (100 μM). The reactions were incubated at 72 [GB-D (exo⁻), Pfu (exo⁻), Taq, Tli (exo⁻)], 65 (Bst) or 37°C [Klenow (exo⁻)] for 2 min, followed by the addition of 10 μl of quenching buffer (95% formamide, 40 mM EDTA, 0.05% xylene cyanol and bromophenol blue).

Primer extension assays with exonuclease-active DNA polymerases

SS primer (250 pmol ^{32}P-labeled, 2 nmol unlabeled) and T template (3 nmol) (c[3]dATP assay) were annealed by heating to 95°C for 5 min, then cooling to room temperature over 1 h. Reactions were initiated by adding polymerase (1 μl) to reaction mixtures (9 μl) containing 1 μl of reaction buffer (provided with the polymerases by the suppliers), annealed primer and template (1.5 μM primer, 2 μM template) and dNTPs (100 μM). The mixtures were incubated at 37 (Klenow fragment, T7) or 72°C (Tli, GB-D, Pfu) for 2 min followed by the addition of quenching buffer (95% formamide, 40 mM EDTA, 0.05% xylene cyanol and bromophenol blue).

PAGE analysis

All reactions were analyzed by separating an aliquot of each quenched reaction mixture (2 μl) on denaturing polyacrylamide gels (7 M urea, 20% 40:1 acrylamide:bisacrylamide) and analyzed with a MolecularImager® system.

RESULTS

Screen of polymerases for the incorporation of c[3]dATP

In these experiments, the template/primer combinations were designed to require that the polymerase first add a 2′-deoxyadenosine to the primer. Further elongation required the incorporation of a 2′-deoxycytidine, followed by the sequence d(GCGTCTC). Full-length product required the addition of 9 nt. This design permitted a single template/primer combination to be used to test incorporation, elongation and some features of fidelity.

To test the ability of Family A DNA polymerases to incorporate c[3]dATP, Taq, Bst and Klenow fragment (exo⁻) polymerases were challenged to extend the template/primer combination using either dATP or c[3]dATP in a single base extension experiment. Incubation was done for 2 min at 72, 65 and 37°C, respectively.

To test the ability of Family A polymerases to elongate a primer terminated in c[3]dATP, the incubation was continued for a further 2 min following the addition of dCTP, dGTP and TTP. The reaction products were analyzed by PAGE. The results are shown in Figure 2.

As expected in the positive control, Taq efficiently incorporated dATP opposite a T at the first position of the template, indicated by the band at position N+1. A faint band at position
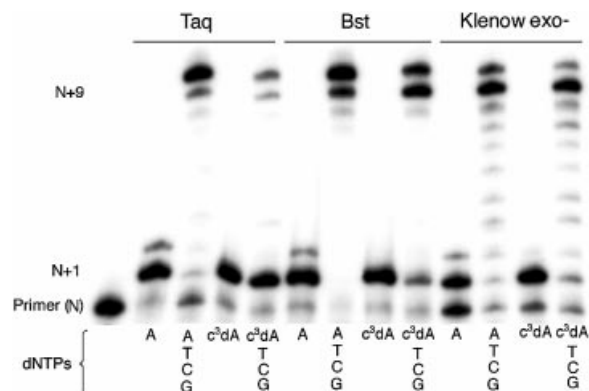

Figure 2. Incorporation of c[3]dATP by three Family A polymerases [Taq, Bst and Klenow (exo⁻)]. 20% polyacrylamide gel. Incubations used the SS primer (5′-GCGTAATACGACTCACTATAG-3′), the T Template (3′-CGCATTATGCTGAGTGATATCTGCGCAGAG-5′) and either dATP or c[3]dATP (indicated below lane) alone for 2 min or with dATP or c[3]dATP for 2 min, followed by addition of dCTP, dGTP and TTP and an additional 2 min incubation. Unextended primer is at position N; addition of dATP or c[3]dATP yields a product band at position N+1. Full-length product is at position N+9. The band at position N+10 is due to the addition of an extra, untemplated nucleotide by the polymerase (see ref. 29).

N+2 indicated that some dATP was also incorporated as a mismatch opposite dG, the second nucleotide in the template, or in an untemplated extension. With the addition of the other standard nucleotides, full-length product was generated (N+9). Taq also added an additional untemplated nucleotide, giving a band at position N+10 (29).

Taq polymerase incorporated as much c[3]dATP as dATP opposite T in the initial 2 min incubation, as indicated by a band at position N+1. This suggested that Taq does not require an unshared electron pair at the 3-position of the incoming triphosphate. The N+1 product largely remains, however, even after addition of all of the other triphosphates in the second 2 min incubation, indicating that Taq does not efficiently extend a primer having a c[3]dA at the 3′-end. This suggests that Taq makes a contact with a minor groove electron pair in the nucleobase standing at the end of the primer. This is consistent with the interpretation, made from the crystal structure, that a hydrogen bonding contact is made to this electron pair by Arg573 (8).

The loss of this interaction does not, however, entirely prevent further extension. Faint bands are seen at positions N+9 and N+10 when Taq is provided a full complement of triphosphates. This suggests a capacity of Taq to elongate a primer lacking this electron pair, albeit inefficiently and with unknown fidelity. No bands that might indicate an abortion of elongation are seen between N+1 and N+9, however, suggesting that once the c[3]dA is not at the 3′-end of the primer, elongation is efficient.

The Bst and Klenow (exo⁻) polymerases, also from Family A, also added c[3]dATP opposite T in the first step of the primer elongation reaction. The two polymerases differed from Taq, and from each other, in their ability to extend a primer carrying a c[3]dA at its 3′-end. With Bst, the absence of an intense band at position N+1 and the presence of strong bands corresponding to N+9 and N+10 full-length products suggested more efficient extension of the primer after the addition of c[3]dATP. Klenow fragment (exo⁻) of DNA polymerase I

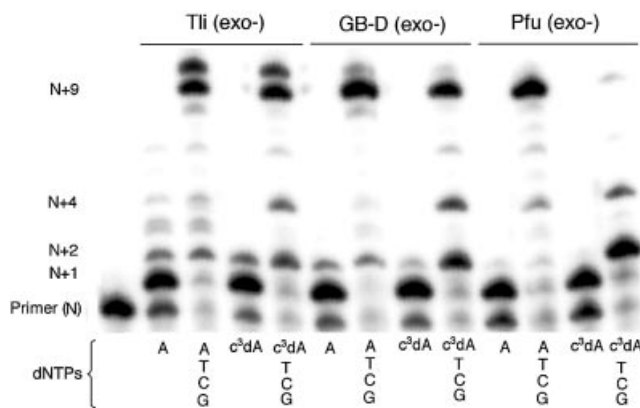**Figure 3.** Incorporation of c^{3}dATP by three Family B polymerases [Tli (exo$^-$), GB-D (exo$^-$) and Pfu (exo$^-$)]. 20% polyacrylamide gel. Incubations were with SS primer (5′-GCGTAATACGACTCACTATAG-3′), T Template (3′-CGCATTATGCTGAGTGATATCTGCGCAGAG-5′) and either dATP or c^{3}dATP (indicated below lane) alone for 2 min or with dATP or c^{3}dATP for 2 min, followed by addition of dCTP, dGTP and TTP and incubation for an additional 2 min. Unextended primer is at position N; addition of dATP or c^{3}dATP yields a product band at position N+1. Full-length product is at position N+9. The band at position N+10 is due to the addition of an extra, untemplated nucleotide by the polymerase.

Figure 4. Incorporation of c^{3}dATP by three reverse transcriptases (HIV-RT, AMV-RT and MMLV-RT). 20% polyacrylamide gel. Incubations were with SS primer (5′-GCGTAATACGACTCACTATAG-3′), T Template (3′-CGCATTATGCTGAGTGATATCTGCGCAGAG-5′) and either dATP or c^{3}dATP (indicated below lane) alone for 2 min or with dATP or c^{3}dATP for 2 min, followed by addition of dCTP, dGTP and TTP and incubation for an additional 2 min. Unextended primer is at position N; addition of dATP or c^{3}dATP yields a product band at position N+1. Full-length product is at position N+9. The band at position N+10 is due to the addition of an extra, untemplated nucleotide by the reverse transcriptase.

gave even less evidence for pausing at N+1 after c^{3}dA was incorporated, indicating even less of a need for the polymerase to identify its substrate by making an interaction with the unshared electron pair in the minor groove. These differences are remarkable considering that Arg573 in Taq, assigned a role in forming a hydrogen bond to the N^3 electron pair, is conserved in the homologous positions in the other Family A polymerases. Even though the amino acid is conserved, it appears that its impact on the ability of the polymerase to accept a primer terminated with c^{3}dA is not.

The differences seen between the three Family A polymerases could be due to the different assay temperatures used. Each polymerase was examined at its optimal temperature, 72°C for Taq, 65°C for Bst and 37°C for Klenow fragment (exo$^-$). The ability to accept c^{3}dATP thus correlated inversely with the reaction temperature. If the minor groove interactions were more important for polymerization at higher temperatures, then an explanation could be offered that explains the different ability of Family A polymerases to accept c^{3}dATP.

Alternatively, the different behaviors of the three polymerases may arise from differential dissociation of the polymerase from the primer–template complex. All three polymerases may dissociate from the primer–template complex with the same frequency when the interaction with the minor groove is lost, but the probability of finding an annealed primer and template to bind to and elongate may decrease with increasing temperature.

Family B polymerases Tli (exo$^-$), GB-D (exo$^-$) and Pfu (exo$^-$) were screened for their ability to accept c^{3}dATP (Fig. 3). In the positive controls, Tli (exo$^-$) polymerase incorporated dA opposite the first template T (N+1). Some misincorporation of dA opposite dG, dC and dG was also observed in the 2 min incubation, generating additional product bands at positions N+2, N+3 and N+4. Completing the inventory of triphosphates in the second 2 min incubation yielded full-
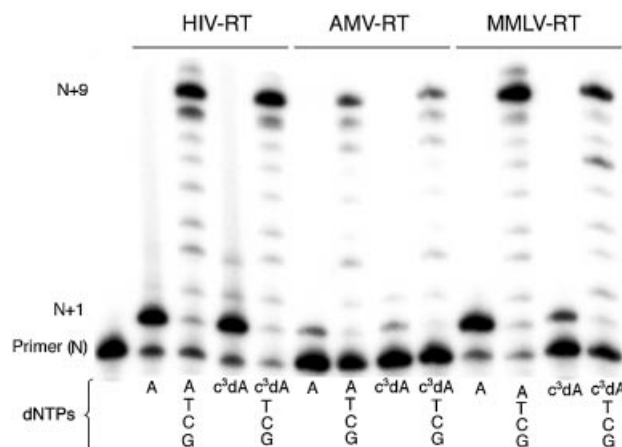
length product (N+9) and product with one additional untemplated nucleotide (N+10).

Incubation with c^{3}dATP generated a product band indicating incorporation of c^{3}dA opposite the first template T (N+1) (Fig. 3). A faint band at N+2 suggested misincorporation via mismatching between c^{3}dA and dG. GB-D (exo$^-$) and Pfu (exo$^-$), two other Family B polymerases, produced analogous results for single base extension.

The three Family B polymerases did not perform analogously in the second 2 min incubation, where they were challenged to extend the primer ending with c^{3}dA. With Tli (exo$^-$), addition of dCTP, dGTP and TTP resulted in full-length product, but with bands at N+2 and N+4 that are greater in intensity than observed with dATP. This suggested that Tli (exo$^-$) pauses in primer extension when the primer lacks an unshared electron pair at positions N+2 and N+4. This is consistent with the interpretation, made from the crystal structure, that the unshared electron pair from the N+2 nucleobase accepts a hydrogen bond from a polymerase Lys (position 706 in RB69), while the unshared electron pair carried by the N+4 nucleobase accepts a hydrogen bond from another polymerase Lys (position 734 in RB69) via a water molecule.

The extent of pausing by GB-D (exo$^-$) and Pfu (exo$^-$) was different. With GB-D (exo$^-$), the N+2 and N+4 bands indicating early termination were slightly greater in intensity than those observed with Tli (exo$^-$). With Pfu (exo$^-$), however, elongation was nearly entirely terminated at position N+2 when c^{3}dA was present at position N+1. This suggested that an interaction with the N3 unshared electron pair was more critical as a specificity determinant for GB-D (exo$^-$) than for Tli (exo$^-$) and a nearly absolute specificity determinant for Pfu (exo$^-$).

Three RTs were also challenged to incorporate c^{3}dA opposite a T in the template. PAGE (Fig. 4) indicated no

detectable difference in incorporation of c^{3}dA by HIV-RT or AMV-RT in the first 2 min incubation. This is illustrated by the presence of bands having comparable intensities at the N+1 position regardless of whether dATP or c^{3}dATP was present in the incubation. MMLV-RT incorporated c^{3}dA with lower efficiency under these conditions. After incorporation of the analog, all three RTs extended the primer to full length (N+9) in the presence of the other dNTPs without any significant accumulation of intermediates.

Polymerase dissociation assays with c^{3}dATP

The differences in the ability of various polymerases to produce full-length product might reflect differences at the incorporation step. Alternatively, the differences may arise from different levels of processivity in the polymerases. Processivity is the property of a polymerase that arises when the rate for further elongation is greater than the rate of dissociation of the polymerase from the primer–template complex.

To determine whether the loss of hydrogen bonding in the minor groove results in a loss of processivity, a modified primer extension assay was used. In this assay, primer extension experiments were performed in the presence of a 300-fold excess of unlabeled primer (a trap). In the absence of the trap, primer can be extended to yield full-length product without processivity. If a polymerase dissociates from the radiolabeled primer–template complex, it can recapture the partially elongated complex and resume elongation. In the presence of a 300-fold excess of unlabeled primer, however, the disassociated polymerase will only rarely again find the partially elongated complex and resume elongation. Therefore, the labeled bands observed in PAGE analysis correspond to products generated by the polymerase processively.

In this assay, the polymerase was incubated with template and primer complex in the absence of dNTPs for 20 s. The incubation mixture was then diluted with a solution containing the trap, dCTP, dGTP, TTP and dATP or c^{3}dATP. Incubation was continued for 20 s before quenching and analysis.

Family A polymerase dissociation assay with c^{3}dATP

With Taq without a trap (Supplementary Material, Fig. S1, lane 3), most of the radiolabeled primer is converted to N+9 and N+10 products, as before. In the presence of trap and dATP (Supplementary Material, Fig. S1, lane 4), the radiolabeled primer is again largely extended to full-length product (N+9), indicating processive elongation. The additional untemplated nucleotide is not added to generate an N+10 product. This suggests that the addition of an untemplated nucleotide is not done processively.

Quantitation of the bands in lane 4 (Supplementary Material, Fig. S15) indicated that ~55% of the primer (N–1) was not extended at all, suggesting that it was not bound by polymerase upon addition of the dNTPs. Of the remaining 45% with dC incorporated (to give the band labeled N), only 20% is extended further by the incorporation of dATP (N+1). This indicates that ~25% of the polymerase complex dissociated before a second nucleotide was added. A small amount (~5–10%) of labeled primer was lost with each subsequent nucleotide addition.

With Taq using c^{3}dATP in the presence of trap (Supplementary Material, Fig. S1, lane 6), no change in processivity was observed until after incorporation of c^{3}dATP (N+1). At that point, 20% of the primer was lost to the trap before it is extended past c^{3}dA with dCTP, compared to 5% with dATP. At positions N+2 to N+6, loss of the primer containing c^{3}dATP to the trap was again the same as loss of the primer containing dATP.

Results with Bst were similar (Supplementary Material, Fig. S2). With dATP and trap, after addition of the first nucleotide, dCTP, and with every subsequent addition, only small amounts (5–10%) of the radiolabeled primer were lost to the trap. With c^{3}dATP, 30% of the primer terminated with c^{3}dA was lost to the trap to give radiolabeled product N+1. Another 30% of the radiolabeled primer was lost to the trap after extension to give N+2 product. Only a small loss of extending product was observed with each subsequent addition.

With the Klenow fragment (exo⁻), trap and dATP, only 15% of the radiolabeled primer was not elongated at all (Supplementary Material, Fig. S3). Again, a loss of 5–10% of the labeled primer to the trap is seen at each additional step. With c^{3}dATP, an extra 10% of the radiolabeled primer is lost to the trap before incorporation of c^{3}dATP, and an additional 40% is lost from the polymerase to the trap after c^{3}dA is added to the 3′-end of the primer. Additional loss of primer is within the normal range after each nucleotide addition.

Prior to incorporation of the first nucleotide in the presence of trap, far less Taq polymerase was bound to the primer and template than Bst or Klenow (exo⁻). After the first nucleotide is added to the primer, however, the relative rates of dissociation of the three polymerases at each step were comparable, both with dATP (Fig. 5A) and with c^{3}dATP (Fig. 5B).

Family B polymerase dissociation assay with c^{3}dATP

The dissociation assay was also used to test for processivity within Family B polymerases. For Tli (exo⁻) in the presence of trap, radiolabeled bands were generated at positions N to N+6 (Supplementary Material, Fig. S4). Only 58% of the radiolabeled primer was initially bound by the polymerase, based on the incorporation of dCTP (to give a band of length N), 18% dissociated after incorporation of dATP (to give a band of length N+1). Approximately 5–10% of the labeled primer was lost to the trap with each subsequent addition.

Extension by Tli (exo⁻) with c^{3}dATP in the absence of trap resulted in extension primarily to position N+7, with only a trace of (N+9) product being formed, with additional bands observed both prior to incorporation of c^{3}dATP at position N and after incorporation of the analog indicated by bands at positions N+1, N+2 and N+4. With trap and c^{3}dATP, bands were also generated at positions N, N+1, N+2 and N+4, but with no extension beyond position N+4. Compared to the extension with dATP and trap, the only increase in dissociation with c^{3}dATP occurs after positions N+2 and N+4. After position N+2, 5% is lost with dATP, 20% with c^{3}dATP; after position N+4, 10% is lost with dATP and no further extension is observed with c^{3}dATP. GB-D (exo⁻) extension with dATP and c^{3}dATP also only differs after position N+2 (Supplementary Material, Fig. S5). Like Tli (exo⁻), the loss after N+2 is 3% with dATP and 20% with c^{3}dATP. With Pfu
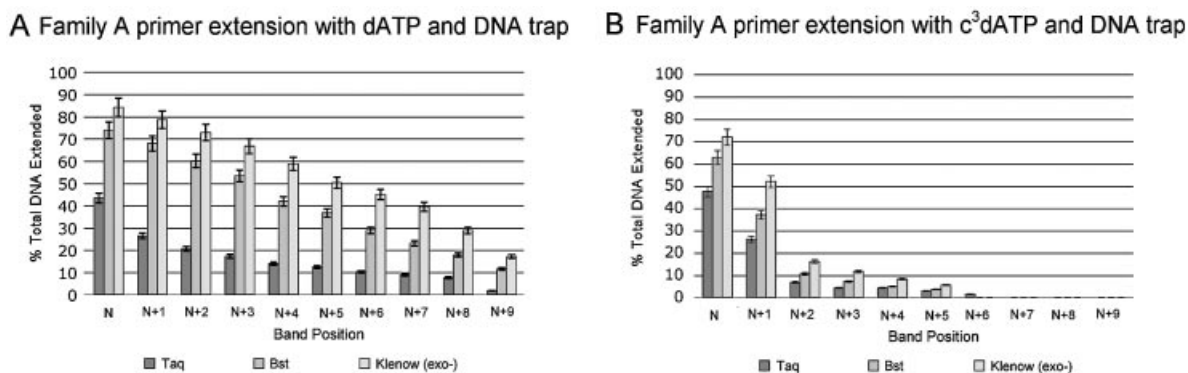
Figure 5. (**A**) Loss of primer upon incubation with dATP, dCTP, dGTP and TTP, trap and Family A polymerases. The extension of the primer by the first nucleotide, dCTP opposite dG, is observed least with Taq, more so with Bst and most with Klenow fragment (exo⁻). A small loss of extended DNA is observed after the addition of each subsequent nucleotide. Due to the small amount of primer initially extended by 1 nt by Taq, very little full-length product is observed. (**B**) Loss of DNA primer extension after addition of each nucleotide with c3dATP, dCTP, dGTP and TTP. Again, extension of the primer by the first nucleotide, dGTP opposite dC, is observed least with Taq, more so with Bst and most with Klenow fragment (exo⁻). The relative loss of extended DNA with each subsequent nucleotide addition does not differ between polymerases.
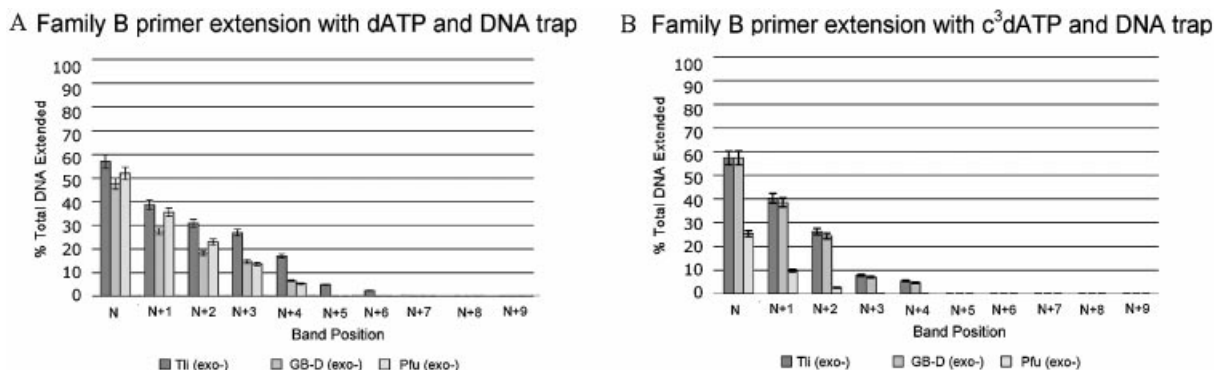

Figure 6. (**A**) Loss of DNA primer extension after addition of each nucleotide with dATP, dCTP, dGTP and TTP. (**B**) Loss of DNA primer extension after addition of each nucleotide with dc³ATP, dCTP, dGTP and TTP.

(exo⁻), only a small amount of primer is extended by dCTP in the presence of c³dATP. No other significant differences in the rate of Pfu dissociation with dATP and c³dATP are seen (Supplementary Material, Fig. S6).

A comparison between the three Family B polymerase primer extensions with dATP and trap indicates that GB-D (exo⁻) and Pfu (exo⁻) have a modestly higher rate of dissociation after each nucleotide incorporation than Tli (exo⁻) (Fig. 6A). This difference is also observed with c³dATP in the presence of trap, most significantly with Pfu (exo⁻). This difference in dissociation is exaggerated at positions N+2 and N+4 (Fig. 6B).

Mismatch incorporation of c³dATP

Interaction between the polymerase and the minor groove unshared pair of electrons might be hypothesized to be important in proofreading. To determine if a mismatch involving c³dA is recognized less efficiently than a mismatch involving dA, mismatches for both were forced by incubating SS primer and T, A, C and G templates for 4 min with polymerase and either dATP or c³dATP.

Contrary to this hypothesis, dA and c³dA were misincorporated opposite dA, dC and dG with equal frequency by Family A polymerases (Supplementary Material, Fig. S7). With Family B DNA polymerases, c³dATP was incorporated opposite a template dC by Tli (exo⁻), GB-D (exo⁻) and Pfu (exo⁻) almost as efficiently as opposite T, shown by the almost full extension of the primer (N) by at least 1 nt (Supplementary Material, Fig. S8). This data suggests that interactions with the minor groove unshared pair of electrons do play a role in fidelity with these Family B polymerases, but not with Family A polymerases. The lack of difference in incorporation of dATP and c³dATP in purine–purine or pyrimidine–pyrimidine mismatches suggests that the minor groove electron pairs do not have a role in exclusion of these mismatches. It is likely that these mismatches could be excluded based on their size.

The incorporation of c³dATP opposite template dC, dG and dA by RTs was also examined (Supplementary Material, Fig. S9). Based on the loss of intensity of the band at position N, the only significant difference in mismatch incorporation observed was an increase in incorporation of c³dATP opposite a template dA by HIV-RT. An increase in mismatch

incorporation of c^3dATP opposite dA is not observed to any significant extent with AMV-RT or MMLV-RT.

Polymerase exonuclease activity with c^3dATP

To further understand the role of the minor groove interactions in polymerase fidelity, the incorporation of c^3dATP by Family A and Family B polymerases with exonuclease activity was explored.

Of the Family A polymerases examined here, only the Klenow fragment has native exonuclease activity. Both the exonuclease-deficient and the exonuclease-active variants were incubated with c^3dATP, TTP, dCTP and dGTP. The exonuclease-active Klenow was also incubated with dATP and the other standard nucleotides as a control. The results are shown in Figure 7.

As noted above, Klenow (exo⁻) incorporates c^3dA opposite T in a template as well as it incorporates dATP. In contrast, no incorporation of c^3dA is seen at all with Klenow (exo⁺). This suggests that when the Klenow polymerase fails to find the minor groove unshared pair of electrons at position N+1, it directs the product to the exonuclease pathway.

To determine whether the behavior observed with Klenow fragment is present in other Family A polymerases with exonuclease activity, the DNA polymerase from T7 bacteriophage was also examined (Supplementary Material, Fig. S10). This polymerase is a distant Family A relative of Klenow fragment (11). Incubation of T7 with the natural nucleotides resulted in full-length product. When dATP was replaced with c^3dATP, however, the primer was degraded. This suggests that the failure of T7 to find the minor groove
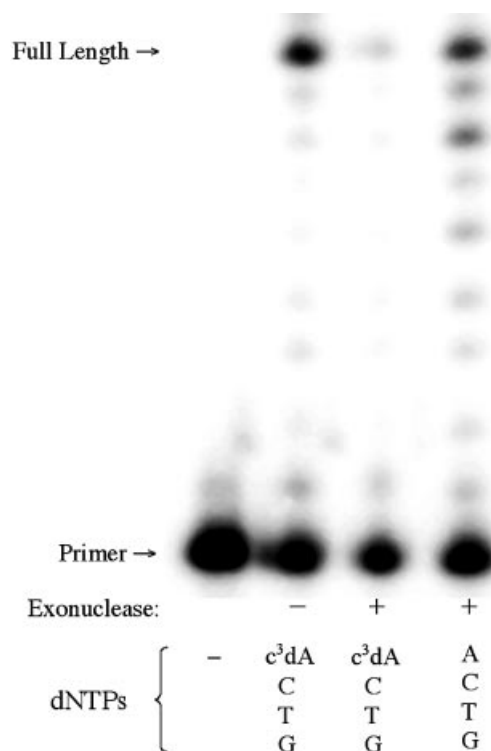

Figure 7. Incorporation of c^3dA by the Klenow fragment of DNA polymerase 1 with (+) and without (−) 3′→5′ exonuclease activity. Triphosphates present are indicated at the bottom of each lane. SS primer (5′-GCGTAATACGACTCACTATAG-3′) and T template (3′-CGCATTATGCTGAGTGATATCTGCGCAGAG-5′) were used.
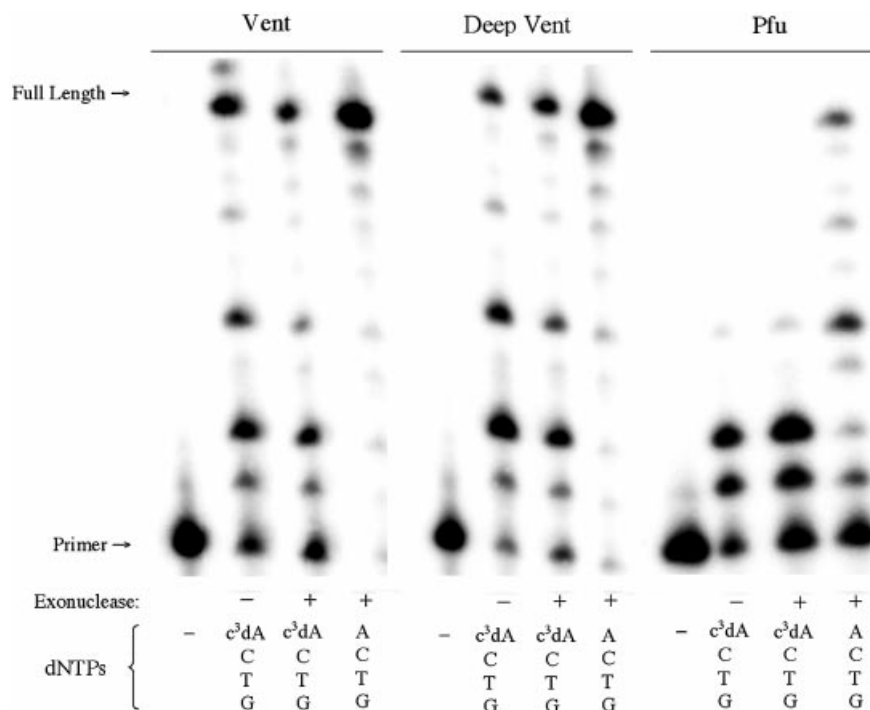

Figure 8. Incorporation of c^3dATP by Tli, GB-D and Pfu with (+) and without (−) 3′→5′ exonuclease activity. SS primer (5′-GCGTAATACGACTCACTATAG-3′) and T template (3′-CGCATTATGCTGAGTGATATCTGCGCAGAG-5′) were used. Nucleotides present are indicated at the bottom of each lane.

unshared pair of electrons at the primer terminus is also sufficient to direct the product into the exonuclease pathway.

All three Family B polymerases examined here have native $3' \rightarrow 5'$ exonuclease activity. Primer extension experiments were performed in parallel with both the exo+ and exo− variants to determine if the absence of the minor groove interaction with c^3dATP directs the product to the exonuclease. Figure 8 shows no discernible difference in incorporation of c^3dA by the exonuclease-deficient and exonuclease-active versions of all three polymerases. This indicates that loss of the minor groove contact is not sufficient to direct the primer to the exonuclease site in Family B polymerases.

DISCUSSION

The most striking results from these experiments are the differences in the patterns with which Family A polymerases, Family B polymerases and RTs accept c^3dA and the extent to which these differences are manifested by different members of each family. These observations have both technological and scientific implications. From a technological perspective, they show a remarkable plasticity of polymerases with respect to this particular specificity determinant. This suggests that native polymerases, perhaps with a small number of amino acid replacements, can accept a large number of non-standard nucleotide analogs. This encourages the development of artificial genetic information systems.

It seemed unlikely that the increased basicity of c^3dA relative to dA [the pK_a of protonated c^3dA (7.0) is 3.5 units higher than for protonated A] (30) might be a source of its exclusion by some polymerases. Although the pH of the buffers varied to meet the specifications of the suppliers, they were typically 8.8, implying that only a few percent of the c^3dA is expected to be protonated. Only with HIV-RT was the pH of the buffer (7.2) near the pK_a of c^3dA. HIV-RT accepted c^3dA rather well. Likewise, the known decrease in the stability of duplexes containing c^3dA (31) should be considered in assessing the outcome. With only a single c^3dA in the stack, however, it seems unlikely that this decrease is significant.

From a scientific perspective, however, the variation is particularly interesting. The Family A and Family B polymerases seem to differ dramatically in their use of the potential for minor groove interactions to address a common problem, fidelity, that is important to their function. In Family A, minor groove scanning appears to be a mechanism to decide whether to use an exonuclease activity. In Family B polymerases, minor groove scanning appears to be incidental to the use of exonuclease activity to determine the fate of a product.

The same is true for mismatch avoidance. The only significant increase in misincorporation occurs with c^3dATP opposite a template dC in Family B polymerases. Apparently, the absence of the purine N^3 in c^3dATP diminishes the ability of the Family B polymerases to detect an A–C mismatch.

Perhaps naively, one might expect that the need for high fidelity in a polymerase would create a demand for many contacts between the polymerase and its substrates. Contacts to minor groove unshared electron pairs would seem to be ideal to meet this demand. We might therefore expect polymerases to make this contact. The fact that they do not is noteworthy.

The differences between polymerases is especially significant from an evolutionary perspective. If Family A and Family B polymerases share a common ancestor, if the ancestor exploited contacts with unshared pairs of electrons in the minor groove to achieve fidelity and if fidelity has been conserved as an important phenotype in the time since Family A and Family B polymerases diverged, then one might expect these contacts, and their phenotypic consequences, to have been conserved as well. The fact that they are not suggests either that Family A and Family B polymerases do not share a common ancestor, that fidelity through minor groove contact was not a feature of that ancestor or that this feature was not sufficiently critical to fitness to have been constrained from divergence as the polymerase families diverged. Each of these interpretations is significant to those interested in the planetary biology of polymerases (32).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

REFERENCES

1. Switzer,C.Y., Moroney,S.E. and Benner,S.A. (1989) Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.*, **111**, 8322–8323.
2. Kool,E.T. (2000) Roles of Watson-Crick and minor groove hydrogen bonds in DNA replication. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 93–102.
3. McMinn,D.L., Ogawa,A.K., Wu,Y., Liu,J., Schultz,P.G. and Romesberg,F.E. (1999) Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly stable self-pairing hydrophobic base. *J. Am. Chem. Soc.*, **121**, 11585–11586.
4. Geyer,C.R., Battersby,T.R. and Benner,S.A. (2003) Rules for designing artificial genetic systems. The nucleobases. *Structure*, **11**, 1485–1498.
5. Benner,S.A. (2003) Synthetic biology. *Nature*, **421**, 118.
6. Seeman,N.C. Rosenberg,J.M. and Rich,A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA*, **73**, 804–808.
7. Eom,S.H., Wang,J. and Steitz,T.A. (1996) Structure of Taq polymerase with DNA at the polymerase active site. *Nature*, **382**, 278–281.
8. Li,Y., Korolev,S. and Waksman,G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.*, **17**, 7514–7525.
9. Kiefer,J.R., Mao,C., Braman,J.C. and Beese,L.S. (1998) Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature*, **391**, 304–307.
10. Franklin,M.C., Wang,J.M. and Steitz,T.A. (2001) Structure of the replicating complex of a Pol alpha family DNA polymerase. *Cell*, **105**, 657–667.
11. Braithwaite,D.K. and Ito,J. (1993) Compilation, alignment and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res.*, **21**, 787–802.
12. Singh,K. and Modak,M.J. (2003) Presence of 18-Å long hydrogen bond track in the active site of *Escherichia coli* DNA polymerase I (Klenow fragment). *J. Biol. Chem.*, **278**, 11289–11302.
13. Spratt,T.E. (2001) Identification of hydrogen bonds between *Escherichia coli* DNA polymerase I (Klenow Fragment) and the minor groove of DNA by amino acid substitution of the polymerase and atomic substitution of the DNA. *Biochemistry*, **40**, 2647–2652.
14. Thompson,E.H.Z., Bailey,M.F., van der Schans,E.J.C., Joyce,C.M. and Millar,D.P. (2002) Determinants of DNA mismatch recognition within the polymerase domain of the Klenow fragment. *Biochemistry*, **41**, 713–722.
15. Wang,J., Sattar,A.K.M.A., Wang,C.C., Karam,J.D., Konigsberg,W.H. and Steitz,T.A. (1997) Crystal structure of a pol α family replication DNA polymerase from bacteriophage RB69. *Cell*, **89**, 1087–1099.

16. Joyce,C.M. and Steitz,T.A. (1994) Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.*, **63**, 777–822.

17. Patel,P.H. and Loeb,L.L. (2001) Getting a grip on how DNA polymerases function. *Nature Struct. Biol.*, **8**, 656–659.

18. Steitz,T.A., Smerdon,S., Jäger,J., Wang,J., Kohlstaedt,L.A., Friedman,J.M., Beese,L.S. and Rice,P.A. (1993) Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of *Escherichia coli* DNA polymerase I. *Cold Spring Harb. Symp. Quant. Biol.*, **58**, 495–504.

19. Huang,H., Chopra,R., Verdine,G.L. and Harrison,S.C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, **282**, 1669–1675.

20. Jacobo-Molina,A., Ding,J., Nanni,R.G., Clark,A.D.,Jr, Lu,X., Tantillo,C., Williams,R.L., Kamer,G., Ferris,A.L., Clark,P., Hizi,A.; Hughes,S.H. and Arnold,E. (1990) Crystal structure of human immunodeficiency virus Type-1 reverse transcriptase complexed with double-stranded DNA at 3.0 Angstrom resolution shows bent DNA. *Proc. Natl Acad. Sci. USA*, **90**, 6320–6324.

21. Najmudin,S., Coté,M.L., Sun,D., Yohannan,S., Montano,S.P., Gu,J. and Feorgiadis,M.M. (2000) Crystal structures of an N-terminal fragment from moloney murine leukemia virus reverse transcriptase complexed with nucleic acid: functional implications for template-primer binding to the fingers domain. *J. Mol. Biol.*, **296**, 613–632.

22. Cosstick,R., Li,X. Tuli,D.K. Williams,D.M., Connolly,B.A. and Newman,P.C. (1990) Molecular recognition in the minor groove of the DNA helix. Studies on the synthesis of oligonucleotides and polynucleotides containing 3-deaza-2′-deoxyadenosine. Interaction of the oligonucleotides with the restriction endonuclease EcoRV. *Nucleic Acids Res.*, **18**, 4771–4778.

23. Guo,M.-J., Hildbrand,S., Leumann,C.J., McLaughlin,L.W. and Waring,M.J. (1998) Inhibition of DNA polymerase reactions by pyrimidine nucleotide analogues lacking the 2-keto group. *Nucleic Acids Res.*, **26**, 1863–1869.

24. Hoops,G.C., Zhang,P., Johnson,W.T., Paul,N., Bergstrom,D.E. and Davisson,V.J. (1997) Template directed incorporation of nucleotide mixtures using azole-nucleobase analogs. *Nucleic Acids Res.*, **25**, 4866–4871.

25. Wu,U., Ogawa,A.K., Berger,M., McMinn,D.L., Schultz,P.G. and Romesberg,F.E. (2000) Efforts toward expansion of the genetic alphabet: optimization of interbase hydrophobic interactions. *J. Am. Chem. Soc.*, **122**, 7621–7632.

26. Morales,J.C. and Kool,E.T. (1999) Minor groove interactions between polymerase and DNA: more essential to replication than Watson-Crick hydrogen bonds? *J. Am. Chem. Soc.*, **121**, 2323–2324.

27. Morales,J.C. and Kool,E.T. (2000) Functional hydrogen-bonding map of the minor groove binding tracks of six DNA polymerases. *Biochemistry*, **39**, 12979–12988.

28. Ludwig,J. and Eckstein,F. (1989) Rapid and efficient synthesis of nucleoside 5′-o-(1-thiotriphosphates), 5′-triphosphates and 2′,3′-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.*, **54**, 631–635.

29. Hu,G.X. (1993) DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3′-end of a DNA fragment. *DNA Cell Biol.*, **12**, 763–770.

30. Minikawa,N., Kojima,N. and Matsuda,A. (1999) Nucleosides and nucleotides. 184. Synthesis and conformational investigation of antifixed 3-deaza-3-halopurine ribonucleotides. *J. Org. Chem.*, **64**, 7158–7172.

31. Lever,C., Li,X., Cosstick,R., Ebel,S. and Brown,T. (1993) Thermodynamic stability and drug-binding properties of oligodeoxyribonucleotide duplexes containing 3-deazaadenine–thymine base-pairs. *Nucleic Acids Res.*, **21**, 1743–1746

32. Benner,S.A., Caraco,M.D., Thomson,J.M. and Gaucher,E.A. (2002) Planetary biology. Paleontological, geological and molecular histories of life. *Science*, **293**, 864–868.