Published online September 1, 2005

Nucleic Acids Research, 2005, Vol. 33, No. 15 ***4865–4873***
doi:10.1093/nar/gki779

Nucleotide modification at the γ-phosphate leads to the improved fidelity of HIV-1 reverse transcriptase

Brent A. Mulder[1,4], Steve Anaya[1], Peilin Yu[2], Keun Woo Lee[1], Anvy Nguyen[1,4], Jason Murphy[3], Richard Willson[1,3,4], James M. Briggs[1,4], Xiaolian Gao[1,2] and Susan H. Hardin[1,4,*]

[1]Department of Biology and Biochemistry, University of Houston, Houston TX 77204-5001, [2]Department of Chemistry, University of Houston, Houston TX 77204-5003, [3]Department of Chemical Engineering, University of Houston, Houston, TX 77204-4004, USA and [4]VisiGen Biotechnologies, Inc., 2575 West Bellfort, Suite 250, Houston, TX 77054, USA

Received May 3, 2005; Revised and Accepted August 2, 2005

ABSTRACT

The mechanism by which HIV-1 reverse transcriptase (HIV-RT) discriminates between the correct and incorrect nucleotide is not clearly understood. Chemically modified nucleotides containing 1-amino-naphthalene-5-sulfonate (ANS) attached to their γ-phosphate were synthesized and used to probe nucleotide selection by this error prone polymerase. Primer extension reactions provide direct evidence that the polymerase is able to incorporate the gamma-modified nucleotides. Forward mutation assays reveal a 6-fold reduction in the mutational frequency with the modified nucleotides, and specific base substitutions are dramatically reduced or eliminated. Molecular modeling illustrates potential interactions between critical residues within the polymerase active site and the modified nucleotides. Our data demonstrate that the fidelity of reverse transcriptase is improved using modified nucleotides, and we suggest that specific modifications to the γ-phosphate may be useful in designing new antiviral therapeutics or, more generally, as a tool for defining the structural role that the polymerase active site has on nucleotide selectivity.

INTRODUCTION

HIV-1, the causative agent of AIDS, has evolved a variety of ways to defeat human host defenses (1). One of these involves evading the immune system by inaccurately replicating its genome. HIV-1 reverse transcriptase (HIV-RT) is the error-prone polymerase responsible for replicating the single-stranded RNA genome of the virus and producing a double-stranded proviral DNA (1,2). HIV-RT's low fidelity is attributed to its lower discrimination during polymerization and the ease at which it extends a mispaired primer terminus (3,4). The resulting high rate of genetic variability, coupled with the rapid accumulation of virions, facilitates escape from the host immune response and promotes the rapid development of drug resistant mutants under single drug regimens. The need for more potent HIV inhibitors is urgent. A clearer understanding of the enzymatic properties governing fidelity is essential for the creation of more effective antiviral therapies.

Analysis of crystal structures of various DNA polymerases, including HIV-RT, reveals a common architecture resembling a partially open right hand (5,6). Each polymerase displays its own characteristic kinetic properties which can be used to assess fidelity. Only recently well-established concepts surrounding nucleotide selection have been challenged by alternate theories with data showing that selection of the correct nucleotide is governed by more than hydrogen-bonding (7,8). One proposed mechanism for polymerase fidelity is the inherent 'geometric selection' for the proper Watson–Crick base pair at the enzyme's active site (9–12). In this model a correct nucleotide is distinguished by its fit within the active site and by any free energy differences which may be amplified by partial water exclusion (10,13–15). Analysis of amino acid substitutions support a strictly confined active site and suggest that polymerase fidelity may be a function of the pocket's flexibility (16). Thus, polymerases with higher fidelities would have a more rigid (i.e. less flexible) pocket than those with lower fidelities.

*To whom correspondence should be addressed. Tel: +1 713 743 2686; Fax: +1 713 743 2636; Email: shardin@uh.edu

It is estimated that on average one error is introduced into the HIV-1 genome at each viral replication cycle (17–19). This high mutation rate ($\sim 3.4 \times 10^{-5}$) suggests that HIV-RT's binding pocket may be quite flexible. Others have reported that altering amino acids in or around the nucleotide binding pocket affects fidelity (20–23). We discovered that fidelity can be manipulated by engineering a phosphate group within the nucleotide substrate.

MATERIALS AND METHODS

Recombinant HIV-RT preparation

Individual *Escherichia coli* cell strains carrying pRT and p6RT51 were grown in a 3 litre:1 litre ratio, respectively, in PSD187 media [10 g/l Bacto Tryptone, 10 g/l yeast extract, 10 g/l glycerol, 5 g K_2HPO_4 and 5 ml/l of trace metals solution (4 g $MgSO_4 \cdot 7H_2O$, 0.2 g NaCl, 0.4 g $FeSO_4 \cdot 7H_2O$ and 0.2 g $MnSO_4 \cdot 4H_2O$ per 100 ml of solution adjusted to pH 3.0 with H_2SO_4)] to an OD of 1–2 at 37°C. Protein was induced by the addition of 210 μg/ml of IPTG and grown for an additional 2–3 h at 37°C. Cell cultures were centrifuged at 5 000 × g then lysed by adding 50 ml of lysis suspension [5 mM spermidine, 0.1% Triton X-100 and 25 mM $NaPO_4$ (pH 7.0)] to each, vortexing, and passing lysed solution through a French press at ~11 000 psi at 4°C. The solution was then centrifuged at 10 000 *g* for 30 min at 4°C. Supernatant was decanted, filtered through Whatman Filter paper no.1, and loaded onto a pre-equilibrated 8 ml Talon Superflow IMAC column (Clontech). Fifteen additional column volumes of running buffer 1 [25 mM NaPO4 and 300 mM NaCl (pH 7.0)] were loaded at a linear velocity of 100 cm/h and the column was washed with twenty column volumes of running buffer 1. A gradient of 15 column volumes from 0 to 500 mM imidazole was used to elute HIV-RT, which occurred at ~100 mM imidazole. Protein fractions were pooled and dialyzed against RT buffer [50 mM Tris, 25 mM NaCl, 1 mM EDTA and 10% glycerol (pH 7.0)]. Further purification consisted of isocratically running HIV-RT over a 1 ml DEAE Sepharose FF Hitrap column (Amersham Biosciences), followed by a 1 ml SP Sepharose FF Hitrap column (Amersham Biosciences). The protein was pooled, dialyzed against buffer containing 50 mM Tris, 25 mM NaCl, 1 mM EDTA and 50% glycerol (pH 7.0), and stored at −80°C.

Oligodeoxyribonucleotides

A DNA primer, (TOP) d(GGTACTAAGCGGCCGCATG), and a DNA template, (BotC6C) d(CCCCCCCCCATGCGGCCCTTAGTACC), were used in primer extension reactions. The DNA primer, (5′Gap) d(ACAGGATTTTCGCCTGC), was used in DNA sequencing reactions. The oligonucleotides $d(G_{15})$ and $d(C_{30})$ were used as primer-template for determining the specific activity of the enzyme. All oligodeoxyribonucleotides were synthesized by MWG Biotech with high-performance liquid chromatography (HPLC) purification.

Primer extension reactions

Reaction conditions were according to (24) with the following modifications. Reactions were prepared by combining 5 μl of Solution B [160 μM KCl, 100 mM Tris–HCl (pH 8.0), 20 mM DTT and 12 mM $MgCl_2$] with 5 μl of Solution A which contains 15:1 molar excess of primer-template to enzyme. The reactions contained 0.01 U of recombinant HIV-RT isolated from *E.coli* (pRT and p6HRT51 plasmids, generous gifts from S.F. LeGrice), where 1 U is equal to the incorporation of 1 nmol of [α-^{32}P]dGTP (PerkinElmer) into an oligo (dG_{15})/(dC_{30}) primer-template at 37°C over a 15 min period. Reactions were terminated by adding 20 μl of stop solution (99% deionized formamide, 0.1% xylene cyanol, 0.1% bromophenol blue and 20 mM EDTA, pH 8.0). Samples were heated at 96°C for 3 min, placed on ice, loaded onto a 20% polyacrylamide-urea gel, and electrophoresed for ~2 h at 50 W. Intensities of unextended primer band (I_o) versus extended primer band (I_1) were quantified with a Fuji MacBas 1000 phosphorimager and version 3.0 image software.

Thin layer chromatography

Thin layer chromatography was performed to examine ANS-dGTP incorporation by HIV-RT. Modifications to the primer extension reaction include: (i) the amount of duplex in the primer extension reaction was increased to 100 pmol to generate detectable amounts of ANS-pyrophosphate, a product of the incorporation event, for UV visualization, (ii) the template oligonucleotide, (BotC6C) d(CCCCCCCCA TGCGGCCG-CTTAGTACC), is complementary to the primer oligo(TOP) and has a homopolymeric overhang of seven Cs (sequence underlined), and (iii) ANS-dGTP was used at a final concentration of 100 μM. To monitor the migration of ANS-PPi, the ANS-dGTP was treated with phosphodiesterase 1 (PDE) (Worthington) for 10 min at room temperature in 1× PDE1 buffer. One microliter of the 10 μl primer extension reaction was spotted onto a PEI Cellulose F plate (Merck), and the spotted reactions were air dried and placed under foil to protect them from light. The plate was then incubated in a closed chamber with 50 ml of running buffer (70% ethanol and 6.6 mM NaCl) for 30 min, and scanned with a Bio-Rad Gel Documentation System under UV mode with Bio-Rad Quantity One software (Bio-Rad Laboratories, Inc., USA).

Synthesis of (5-sulfonate)-naphthalenyl-amino-γ-phosphoramidate-β,α-diphosphate-deoxynucleotides (γ-ANS-dATP)

The four γ-P-aminonaphalene-5-sulfonate dNTP compounds, A, C, G and T, were synthesized using a modified procedure for the synthesis of the ANS-ATP (25). The preparation of gamma ANS-dATP is shown in Figure 1. 1-Aminonaphthalene-5-sulphonic acid (ANS) and 1-(3-dimethylaminopropyl)-3-ethyl-carbodiimide hydrochloride (DEC) (Lancaster); dNTP (Sigma), and DEAE-Sephadex ion exchange resin (A-25-120) (Aldrich).

ANS (447 mg, 2 mmol and 40 eq.) was dissolved in 10 ml H_2O. The pH of the solution was adjusted to 5.8 using NaOH (1 N) and insoluble materials were removed by a syringe filter, yielding a saturated solution of sodium ANS (0.18–0.2 M). 5′-triphosphate-2′-deoxyadenosine disodium salt (0.05 mmol) and DEC (2 mmol) were mixed in 4 ml H_2O (12.5 mM dATP) at 22°C. The two solutions were mixed and the reaction continued for an additional 2.5 h. The pH of the reaction was kept at 5.7–5.8 by periodic addition of 0.1 N HCl. The reaction

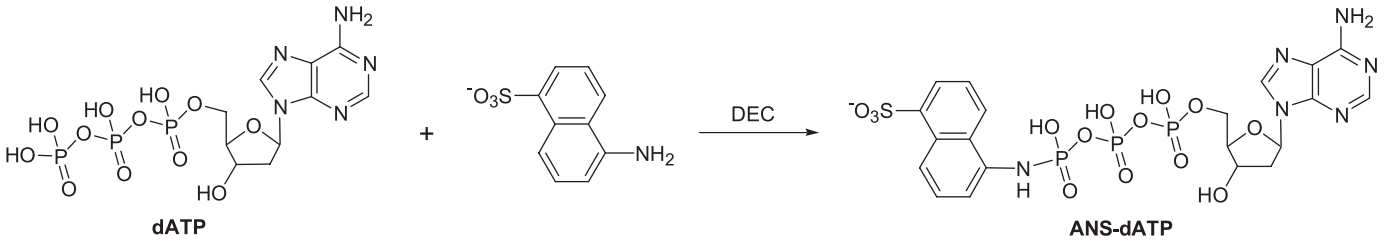

Figure 1. Synthesis of ANS-dATP. Chemical attachment of ANS to the gamma-phosphate position of 2′ deoxyadenosine 5′ triphosphate (dATP). dCTP, dGTP and dTTP were similarly modified. All structures were confirmed via NMR.

solution was then diluted to 50 ml and adjusted to pH 7.4 using 0.05 M triethylammonium bicarbonate buffer (TEAB, pH ~7.5). The DEAE-Sephadex anion exchange column used for purification was pre-equilibrated with 100 ml TEAB (pH 7.5, 1 M), 100 ml aqueous sodium bicarbonate (1 M), and 100 ml TEAB (pH 7.5 and 0.05 M). The reaction mixture was then applied to the ion exchange column at 5°C. The column was eluted with a linear gradient of TEAB (pH 7.5) from 0.05 to 0.9 M. The product fractions were collected at ~0.7 M buffer in a total volume of about 10 ml, and detected to emit brilliant blue fluorescence at 366 nm excitation. These fractions were pooled, freeze dried and co-evaporated twice with H_2O/ethanol (70/30). The dried residue was dissolved in water and lyophilized to give γ-ANS modified dATP as powder. UV (CH_3OH) λ_{max} 245 nm. [^{31}P] NMR(^{1}H decoupled, 600 MHz, D_2O, Me_3PO_4 external reference, 293 K, pH 6.1) δ (p.p.m.) −12.60, −14.10 and −25.79. [^{31}P] NMR (dATP, Na^+, D_2O 293 K) δ (p.p.m.) −11.53 (γ), −13.92 (α) and (β) −24.93. The concentrations of γ-ANS modified dNTPs were obtained from 260 nm absorption readings after multiplying by a fraction factor that was derived from the ratio of the absorption at 260 nm for free 1-aminonapthalene-5-sulfonate (ANS) versus dATP (for γ-ANS-dATP the factor is 0.72). The synthesis of γ-ANS-dCTP, γ-ANS-dGTP and γ-ANS-dTTP followed similar procedures. Each of the ANS-dNTPs was resuspended in MQ H_2O for use.

In vitro forward mutation assay

The forward mutation bacterial strains and phage (generous gifts from T. Kunkel) and methodology are described (26). Two independent 25 µl reactions for each nucleotide substrate, dNTPs or ANS-dNTPs, were performed in the presence of 0.01 U of recombinant HIV-RT at 37°C for 1 h. The nucleotide substrates used for gap-filling were at a 500 µM concentration, and all reactions were analyzed by agarose gel electrophoresis (0.8% gel, 25 mA and 18 h) to confirm that gap-filling synthesis proceeded to completion. The X-gal substrate used for blue/white screening was replaced with S-gal/LB agar blend, (Sigma), which produces black (wild-type) and clear (mutant) plaques to facilitate mutant screening. All candidate mutant plaques were re-screened in the presence of wild-type to confirm their phenotype. The background mutation frequency was determined by electroporating unfilled gapped-duplex DNA and screening for mutants as described. Single-stranded M13 DNA was purified from each mutant plaque using an M13 ssDNA kit (Qiagen) as per manufacturer's instructions.

Calculation of mutational frequencies

Mutational frequencies were calculated by dividing the total number of mutants detected for a substrate (dNTPs or ANS-dNTPs) by the total number of plaques screened for that substrate. This mutational frequency was then corrected by subtracting the background mutational frequency of 2.2×10^{-4}. The corrected mutational frequency was then multiplied by the percentage of mutations represented in the particular class. As an example, the mutational frequency for the dNTPs single-base substitution class was calculated by dividing the number of mutants for this class (178) by the total number of mutants detected (211) and then multiplying by the corrected mutational frequency for the dNTPs.

DNA sequencing

Automated DNA sequencing reactions were performed in a 10 µl volume using 50 ng of purified phage ssDNA, 2 µl of ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit Mix v2.0, and 1.6 pmol of primer (5′Gap). The cycling regimen was as follows: 96°C for 10 s, 50°C for 1 min and 60°C for 4 min for a total of 30 cycles. Sequencing reactions were precipitated, lyophilized and resuspended in 3 µl loading buffer (50 mg/ml Blue Dextran, 5 parts deionized formamide, 1 part 25 mM EDTA at pH 8.0). The samples were heated to 96°C for 2 min and one-half of the reaction was loaded onto a 36 cm 5.3% RapidGel-XL (US Biochemical). Data were collected in 2× mode on an ABI PRISM 377 DNA Sequencer and manually edited using Sequencher v3.0 or v3.1 (GeneCodes).

Molecular modeling

The DISCOVER and HOMOLOGY modules of the INSIGHTII program were used for the modeling studies. The CVFF forcefield and an implicit solvent model using a distance dependent dielectric constant ($\varepsilon = 4$ r) were introduced for the calculations (Insight II Version 2000, Accelrys Inc.). A 12 Å cutoff was applied to truncate intermolecular interactions. Minimization of the initial system was carried out for 100 steps using the steepest descent algorithm. Since the molecular system is very large, most of it was held rigid/fixed during the minimizations and molecular dynamics simulations. The ANS-dTTP, two Mg^{2+} and 235 residues, mostly in the finger and palm domains of the p66 domain that are neighbors of the ANS-dTTP, were allowed to move. After initial minimization, the system temperature was raised to 400 K for 5 ps and then cooled to 300 K for 5 ps. The final structures after the simulated annealing were minimized for 1000 steps using the

conjugate gradient scheme. The partial charges for each atom of the sulfate moiety in ANS were derived from the electrostatic potential fitted charges from a PM3 quantum mechanical geometry optimization. The net charges of the oxygen atoms in the sulfate were very similar to the oxygen atoms in the DNA phosphate, −0.80 and −0.85, respectively. The original X-ray structure contains three structural defects in the p66 domain, namely missing coordinates for residues 1–2, 218–230 and 430–440 which were repaired using the HOMOLOGY module of the INSIGHTII program. These regions were held rigid in the modeling studies since they exist on the opposite side of the dNTP and DNA binding sites so the repairs are not anticipated to have a significant impact on these results.

RESULTS

HIV-RT incorporates the ANS-dNTPs

Both an extended primer and a pyrophosphate group are products of a nucleotide incorporation event. Thus, to determine whether the gamma-modified nucleotides are incorporated by HIV-RT, both the accumulation of tagged-PPi and extended DNA strands were monitored during the course of the extension reaction. Specifically, thin layer chromatography (TLC) was used to monitor the production of ANS-pyrophosphate (ANS-PPi) during the incorporation reaction. TLC exploits the chemical differences of intact ANS-dGTP and ANS-PPi via their differential migration in solvent. The intact ANS-dGTP migrates close to the solvent front (Figure 2A, lane 1), whereas ANS-PPi is more slowly migrating, staying near the bottom of the plate (Figure 2A, lane 2). Extension reaction timepoints demonstrate that as the reaction progresses, an increase in signal for the ANS-PPi is observed (Figure 2A, lanes 3–7) along with an accumulation of extended products (Figure 2B, lanes 9–13). To eliminate the possibility that production of the ANS-PPi spot might be due to reaction conditions promoting ANS-dGTP decomposition instead of HIV-RT incorporation, a 60 min reaction lacking HIV-RT was spotted. As shown, the ANS-dGTP remains intact (Figure 2A, lane 8). HIV-RT's ability to incorporate the ANS-dGTP against a template containing a run of seven complementary bases was also assessed using a gel-based primer extension assay. Primer extension products were electrophoresed on a 20% denaturing polyacrylamide gel where each incorporated base can be defined by its difference in mobility. Extension reactions mimicking those performed in the TLC analysis displayed extended products over the same time period, further confirming that nucleotides modified at their gamma-phosphate with ANS are incorporated by HIV-RT (Figure 2B). Additionally, to address the possibility that even a minimal amount (<1%) of contaminating natural dNTP carried over from synthesis of the ANS-dGTP could lead to erroneous results in a primer extension assay, the ANS-dGTP was pre-treated with shrimp alkaline phosphatase to eliminate any natural dNTPs. Results from this treatment did not produce differences in the amounts of ANS-dGTP primer extension product nor did it have any detrimental effects on the integrity of the ANS-dGTP, data not shown.

HIV-RT mutational frequencies using natural or ANS-modified dNTPs

Analysis of primer extension reactions comparing natural and ANS-modified dNTPs in the context of matched and all possible mismatched configurations demonstrated that the modified nucleotides are incorporated by HIV-RT and that the presence of the ANS-tag appeared to reduce misincorporation, relative to natural nucleotides (data not shown). These reactions also demonstrated that this apparent fidelity improvement required attachment to the gamma-phosphate of the nucleotide. To more thoroughly investigate this result a forward mutation assay was conducted with HIV-RT to determine the enzyme's fidelity and preference for certain error types using either all natural or all ANS-modified nucleotides (4,28). Two independent gap-filling reactions were performed with each substrate and analyzed by agarose gel electrophoresis to ensure complete synthesis. In the assay containing all four natural dNTPs, HIV-RT produced 169 mutant plaques out of the 11 334 plaques screened, (mutational frequencies of 1.4 and 1.8% for screen no. 1 and no. 2, respectively), resulting in an overall mutational frequency

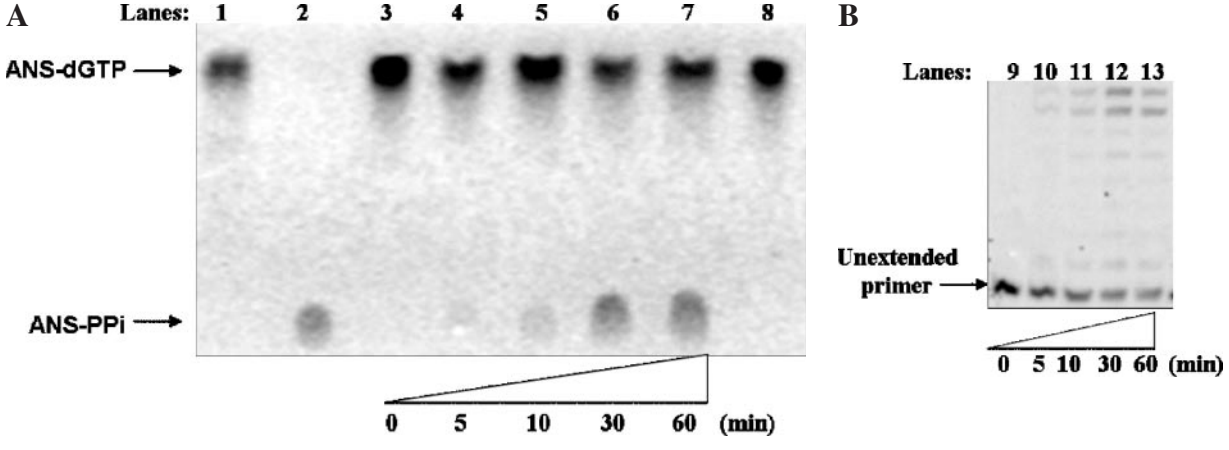

Figure 2. Analysis of HIV-RT's incorporation of ANS-dGTP. (**A**) TLC. The mobility of ANS-dGTP (lane 1) and ANS-PPi produced by PDE treatment of ANS-dGTP (lane 2) are used as controls to monitor migration distances. A timecourse for primer extension reactions at 0, 5, 10, 30 and 60 min, respectively (lanes 3–7). Control conditions to examine ANS-dGTP stability in reaction conditions lacking enzyme (lane 8). (**B**) Gel electrophoresis. Primer extension reactions at 0, 5, 10, 30 and 60 min, respectively (lanes 9–13). Unextended primer is indicated at the bottom of the gel image.

of 149×10^{-4}. Whereas the assay containing all four ANS-dNTPs produced only 26 mutant plaques out of the 10 053 plaques screened, (mutational frequencies of 0.2 and 0.5% for screen no. 1 and no. 2, respectively), resulting in an overall mutational frequency of 26×10^{-4}. The background mutant frequency (2.2×10^{-4}) was subtracted from each and a comparison of these frequencies revealed a 6-fold reduction in mutational frequency using the ANS-dNTPs.

The specific mutations generated by HIV-RT in the *lacZ* α-complementation region with either substrate (dNTPs or ANS-dNTPs) are illustrated (Figure 3). For both cases, the distribution of mutations is non-random and the appearance of hot spots (three or more independent mutations at one location) is evident. A lower incidence of hot spots is observed with the ANS-dNTPs, but this may reflect the overall reduction in mutation frequency observed using the ANS-modified nucleotides (Figure 3). Comparatively, the largest reduction in mutation frequency is at positions −35 and −36 where there were 88 errors introduced by reverse transcriptase using natural dNTPs, but only six errors when RT incorporated ANS-dNTPs. Typically hot spots are found within a run of two or more of the same template base, and result in a change to the base identity of the 5′ neighboring base of the run (29). This is generally supported by the data for both the dNTPs and ANS-dNTPs. The other hot spot observed with both natural and modified nucleotides occured at position +145, a region not contained within a homopolymeric sequence (Figure 3). Interestingly, the dG:ANS-dATP mispair, the largest hot spot observed for the ANS-dNTPs, lies at the center of a 12-base perfect (18-base imperfect) palindrome.

The mutational classes found for both error spectra consist of single-base substitutions, single-base frameshifts, and complex errors (i.e. multiple mutations) (Table 1). Single-base substitutions predominate and a 6-fold difference is observed for these errors occurring with the dNTPs versus the ANS-dNTPs. The dNTP substitutions that occur most frequently are the mispairs: dT:dGTP, dG:dATP and dG:dTTP (Table 1). The ANS-dNTPs mispairs of dT:ANS-dGTP and dG:ANS-dATP are observed most frequently, and each is reduced to 10- to 3-fold relative to natural nucleotides, respectively. However, no dG:ANS-dTTP mispairs are detected.

The preferential site of introduction of a frameshift error in both cases is within runs of a common template base. However, no frameshift hotspots are observed for the ANS-dNTPs, possibly because of the low number of mutants recovered. Complex errors account for 16 of the 169 screened dNTP mutants; the majority having at least one error within a hot spot. Only one of the 26 ANS-dNTP mutants contains non-adjacent, multiple mutations, and neither error is located within a region corresponding to a hot spot.

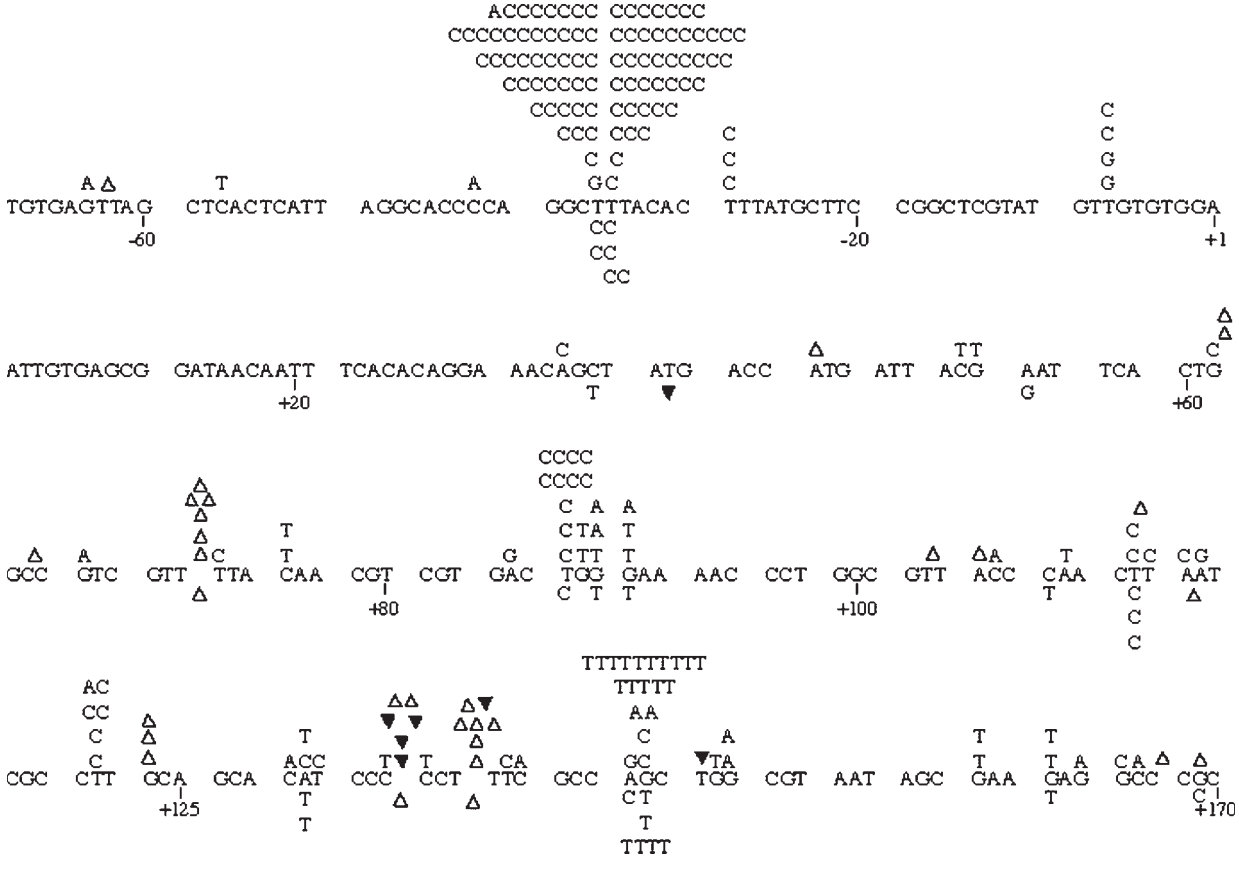

Figure 3. Results of forward mutation assays containing either natural or gamma-modified dNTPs. The identity of the mutation is shown above (natural dNTPs) or below (ANS-dNTPs) the sequence of the *lacZ* alpha-complementation region. A filled, inverted triangle represents an insertion. An empty triangle represents a deletion.

Table 1. *In vitro* forward mu tation assay error results for HIV-RT utilizing either natural or gamma-modified dNTP substrates

Mutational class	dNTPs Number of mutants	Mutational frequency ($\times 10^{-4}$)	ANS-dNTPs Number of mutants	Mutational frequency ($\times 10^{-4}$)	
Single-base substitutions	178	124	26	20	
Single-base frameshifts	33	23	5	4	
Multiple mutations	38		5		
Mutation: From→To					Fold reduced w/ANS
T_C	113	93	10	9	10
G_T	26	21	9	8	3
A_T	2	2	2	2	1
A_C	3	2	1	1	2
C_T	6	5	2	2	3
G_C	4	3	1	1	3
A_G	4	3	1	1	3
G_A	10	8	0	CD	>8
T_G	3	2	0	CD	>3
C_A	5	4	0	CD	>4
T_A	2	2	0	CD	>2
C_G	0	CD	0	CD	CD

The results are the sum of two independent reactions per substrate at a concentration of 500 μM each. If a clone contained more than one mutation, each mutation was counted as an individual event. 'Multiple mutants' indicates the number of phage containing more than one mutation. CD: cannot be determined.

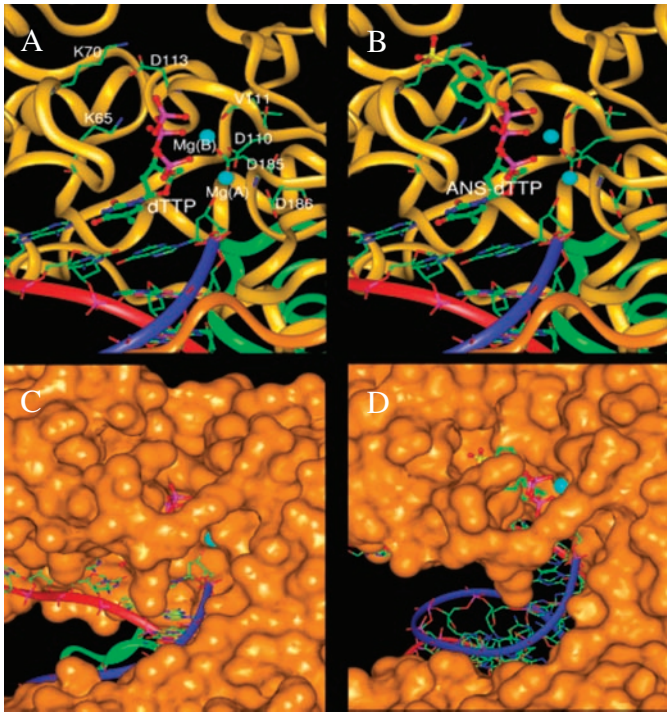

Figure 4. Modeling dTTP and ANS-dTTP in the HIV-RT complex. The backbone structures of p66 and p51 are rendered in brown and green ribbons, respectively. The lighter brown region represents residues from 1 to 235 (finger and palm domains) which were allowed to move during the simulations. The DNA template and primer backbone structures are rendered in red and blue ribbons, and two Mg^{2+} ions are rendered as cyan spheres. In each figure, the following colors were used for the different atoms: C-green, O-red, N-blue, P-pink, and S-yellow. (**A**) The X-ray structure of the dTTP was displayed with the neighboring residues of the two Mg^{2+} ions and the gamma-phosphate moiety: K65, K70, D110, V111, D113, D185 and D186. The dTTP structure is adjacent to the primer DNA in ball and stick. (**B**) For the initial structure of ANS-dTTP, the ANS group was manually attached to an oxygen atom of the gamma-phosphate group. The ten membered ring and sulfate group were placed without atomic overlaps against the existing atoms of the p66 domain of the HIV-RT structure. Forcefield parameters were developed for the ANS linked to the gamma-phosphate via an oxygen atom, instead of an NH group, to simplify the force field parameterization process. (**C**) Solvent accessible surface renderings for panel A is shown. (**D**) Solvent accessible surface renderings for panel B is shown. Viewing angles in panels C and D were slightly changed from those of A and B to more clearly show the locations of gamma-phosphate and ANS groups.

Molecular modeling of ANS-dTTP within the HIV-RT binding pocket

A plausible starting conformation of ANS-dNTP in the dNTP binding pocket and the function of the ANS group were investigated using molecular modeling studies. For the modeling study, the 3.2 Å X-ray structure of HIV-RT solved by Huang *et al*. was employed (PDB code: 1RTD) (30). This structure contains a DNA template, primer, and a dTTP containing two Mg^{2+} ions in addition to the p66 and p51 domains that comprise the heterodimer RT. The core region around the dTTP is rendered in Figure 4A. Based on this ternary complex, the ANS-dTTP structure was manually constructed using the BUILDER module of the INSIGHTII program. The ANS moiety was introduced to one of the three oxygen atoms of the gamma-phosphate group (Figure 4B). While the available room for ANS was small, it could be accommodated in a few conformations. The three oxygen atoms in the sulfate group were placed near residue K70 of the p66 domain. Before introducing the ANS group to the dTTP, the solvent accessible surface of the p66 domain was investigated to examine the available volume and its location. The surface diagram of the X-ray structure clearly shows that some amount of volume exists at the end of the gamma-phosphate of the dTTP (Figure 4C), eventually extending to the solvent. Interestingly, another surface map of the initial ANS-dTTP model (Figure 4D) shows that there is a small amount of room near the residue K70 which seems to be perfect for introduction of the sulfate group into the ANS structure.

In an attempt to identify possible conformations of the manually built ANS-dTTP ternary complex, a simulated annealing method was employed. During the simulations, most of the structure was held fixed except for the ANS-dTTP, two magnesium ions and 235 residues. After initial minimization, the system temperature was increased to 400 K and then cooled to room temperature (300 K). Subsequently, the structure was again minimized. The final structure represents a refined starting conformation of the ANS-dTTP complex, useful for initial analyses and as a starting point for further detailed modeling studies. For comparison purposes, the same approach was applied to the dTTP X-ray structure with the RT and DNA structures.

The refined dTTP structure after simulated annealing and minimization is shown in Figure 5A. The overall shape of the pyrimidine ring is slightly distorted from that in the X-ray structure (Figure 4A); the gamma-phosphate group moved away from the side chain N-atom of K70. However, the

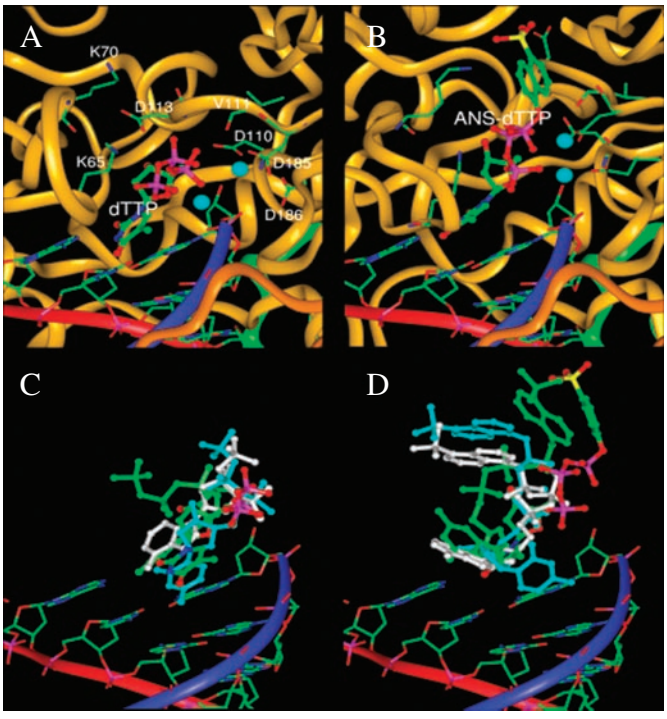

Figure 5. The modeled structures of dTTP and ANS-dTTP within the HIV-RT ternary complex. The backbone structures of p66 and p51 are rendered in brown and green ribbons, respectively, and the DNA template and primer backbone structures are rendered in red and blue ribbons. (**A**) The final conformation of the modeled structure containing dTTP displayed at the same viewing angle as in Figure 4A. (**B**) The final conformation of the modeled structure containing ANS-dTTP displayed at the same viewing angle as in Figure 4B. (**C**) Comparison of the dTTP structure with the modeled, non-complementary structures. (**D**) Comparison of the ANS-dTTP structure with the modeled, non-complementary structures. For clarity, panels C and D do not show the p66 and p51 domains or the two Mg^{2+} ions. In panels C and D the dNTP and ANS-dNTP structures for A, C, and G are colored in white, cyan, and green, respectively, while the rest of the atoms are colored as in Figure 4.

positions of the sugar ring and the alpha phosphate group are well conserved. Although our simulated annealing approach caused minor structural distortions, it did not introduce unphysical change to the ternary complex. Of course, this was in part by design since most of the protein (76.4%) was held rigid. For the ANS-dTTP structure (Figure 5B), the sugar ring and the pyrimidine ring of the ANS-dTTP are very similar to those of the X-ray structure (Figure 4B). Interestingly, the side chain of K70 maintains the interaction with the termini of the substrates. The backbone geometries of residues from 68 to 72 significantly changed their positions to adjust to the new interactions. The ANS group is well aligned along the valley of the p66 domain. This result supports the idea that the effect of adding ANS to the dTTP structure can be compared to an extension of the phosphate chain.

Further computational experiments were performed to model the interaction of non-complementary DNA bases with and without the ANS moiety. A, C and G were introduced in place of 'T' to investigate the effect of the ANS group in the fidelity enhancement of the DNA synthesis reaction. Based on the refined structures of the dTTP and ANS-dTTP (Figure 5A and 5B), the remaining dNTP and ANS-dNTP structures for A, C and G were manually built by manipulating the pre-existing pyrimidine ring using the BUILDER module of INSIGHTII. After initial energy minimizations, the identical simulated annealing procedures were applied to each structure to obtain refined conformations for each of the ternary complexes, and these refined conformations of mispair complexes with or without ANS were compared with those of the complementary base, T (Figures 5C and D).

The conformations of dNTPs without ANS (Figure 5C) are predicted to be more similar to each other, in contrast to those of the ANS-dNTPs (Figure 5D). In particular, the positions of the sugar ring and the alpha phosphate, both of which are critical positions in the DNA synthesis reaction, are more closely aligned with each other in the models containing unmodified dNTPs. However with the addition of the ANS group (Figure 5D), large differences in the binding conformation between the complementary and non-complementary bases are predicted.

DISCUSSION

Highly accurate enzymatic DNA synthesis requires that the DNA polymerase recognizes and transmits information about the identity of the bases within a template sequence, either before or immediately after nucleotide incorporation. Current models suggest that the geometry of the active site plays a role in fidelity by distinguishing the dimensions of a correct versus and an incorrect base pair (8,12,31,32). Nucleotide analogs containing base (33–36), sugar (37–42), or phosphate (43–46) modifications have been designed to probe the active site architecture. Incorporation analyses demonstrate that each polymerase has a defined tolerance for a particular nucleotide modification. One explanation for this variability is attributed to the flexibility within each polymerase's nucleotide binding pocket.

Our results show that in an M13-based forward mutation assay HIV-RT was able to incorporate and extend all four ANS-dNTPs. We demonstrated that HIV-RT incorporates nucleotides with ANS attachment at the γ-phosphate via a TLC assay. Differences in the mobility of the intact ANS-dGTP and ANS-PPi permit identification of nucleotide incorporation, which was detected by increasing intensities of tagged-PPi with increased reaction time (Figure 2). Primer extension analyses (data not shown) reveal that the ANS-dNTPs are incorporated at slower rates than their natural dNTP counterparts. Specifically, a natural dNTP typically produces fully-extended products at approximately 10-fold lower concentration than the ANS-dNTPs. For a misincorporated nucleotide to become a mistake, the DNA polymerase must be able to extend the terminal mispair (48). Fewer errors were made with the ANS-dNTP substrates than with the dNTPs (Figure 3) for each class of mutation (Table 1). Comparison of the specific error types supports the trend initially detected via primer extension reactions (data not shown): gamma-phosphate modifications more dramatically reduce misincorporation of purine nucleotides (Table 1). However, misincorporation of ANS-pyrimidines (whose frequency was already low) is also either reduced or eliminated. For example, the G:dTTP mispair is only detected with the natural substrate. A key feature of the wobble base pair dG:dT is its use of alternate hydrogen bonding patterns and its ability to adopt orientations in which either the template base or incoming

dNTP base laterally shifts (49). Given that the G:ANS-dTTP mispair is not observed, the accommodation of a shift may not be possible in the presence of the ANS-attachment, suggesting a mechanism by which the modification improves fidelity.

HIV-RT's low error frequency with the ANS-dNTPs resulted in a reduction in the number of hotspots (29). ANS-dNTP mutations occur at 19 positions, and 79% of these are located at the same spots as the dNTP mutations. Primer-template misalignments are suspected to be the underlying cause of mutations at sites where there are runs of a common template base (27,50,51), which is where a majority of the ANS-dNTP errors were found. However, the largest ANS-dNTP hotspot (+145) is not located within a run of common template bases, but rather at the center of a palindrome.

A molecular model of the HIV-RT binding pocket with an ANS-dTTP positioned at the primer-template was constructed. The incorporation of a nucleotide is a dynamic process which involves the polymerase changing its conformation from an 'open' to a 'closed' configuration. Critical residues within the fingers domain bend inward toward the palm domain and make contact with the incoming nucleotide, resulting in the formation of the dNTP binding pocket (30). Our model predicts that the active site amino acid residues 68–72 sterically clashes with the ANS-attachment at the γ-phosphate. Crystal structure analysis of a trapped catalytic complex reveals that the incoming triphosphate portion of the dNTP is positioned to interact with K65, R72 and D113 (30). The clash of the ANS-tag with these amino acid sidechains may play a role in nucleotide discrimination by minimizing alternative nucleotide orientations that are required for incorporation of a mismatched nucleotide. This hypothesis is supported by the host of multinucleoside resistance mutants which have been identified within this region of the fingers domain (52).

ANS attachment to the γ-phosphate of a dNTP enhances HIV-RT's discrimination between correct and incorrect nucleotides and demonstrates that base discrimination can be modulated by 'non-informational' portions of the nucleotide. A number of alternative base pairing arrangements are proposed for mispairing events, and each requires an altered base alignment. Molecular modeling suggests that ANS attachment to the γ-phosphate may sterically inhibit mispair formation such that a mispair may not easily adopt an orientation that is conducive to nucleotide incorporation. Lack of contact between the phosphates and any of the elements necessary for catalysis (e.g. critical amino acid side chains, divalent cations, or the 3′-OH of the terminal base) could impair the chemical step of the reaction.

Our data demonstrate that γ-modified dNTPs improve the accuracy at which nucleotides are inserted into a DNA strand and our molecular modeling studies predict significant differences in the positions of the sugar ring and the alpha phosphate group in non-complementary base pairings involving γ-modified dNTPs. Together these results suggest that the γ-phosphate modification may act as a signal amplifier of base identity. The conformation of the ANS group is sensitive to the internal environment (presence of a mispair), and this conformational information is transmitted along the attached phosphate groups to the sugar ring position. This effect may be the key reason for the fidelity increase in the DNA synthesis reaction.

These findings may have both therapeutic and biotechnology applications. Probing the active site with nucleotide attachments that differ in size, charge, or length of linker will provide valuable information about the polymerase's fidelity and potentially lead to the development of more potent antiviral therapies. Biotechnological applications for nucleotides modified at the gamma-phosphate may include increased fidelity of *in vitro* DNA replication, such as more accurate production of first strand cDNA, and as new substrates for next generation sequencing methods.

ACKNOWLEDGEMENTS

We thank Dr Shiao-Chun Tu for stimulating discussions and Dr Hongyi Wang for additional syntheses of modified nucleotides. Financial support was provided by the National Institutes of Health, the Welch Foundation, and the Defense Advanced Research Projects Agency. Funding to pay the Open Access publication charges for this article was provided by the University of Houston.

Conflict of interest statement. B. Mulder, R.C. Wilson, J. Briggs, S. Hardin and K. Lee all hold stock in VisiGen Biotechnologies, Inc., the developer and manufacturer of modified nucleotides for use in a varity of biotech applications. X. Gao is a minor stock holder of VisiGen. The other authors have not declared any conflicts.

REFERENCES

1. Varmus,H. (1988) Retroviruses. *Science*, **240**, 1427–1435.
2. Weiss,R., Teich,N., Varmus,H. and Coffin,J.M. (1985) *RNA Tumor Viruses, 2nd edn*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
3. Kati,W.M., Johnson,K.A., Jerva,L.F. and Anderson,K.S. (1992) Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.*, **267**, 25988–25997.
4. Yu,H. and Goodman,M.F. (1992) Comparison of HIV-1 and avian myeloblastosis virus reverse transcriptase fidelity on RNA and DNA templates. *J. Biol. Chem.*, **267**, 10888–10896.
5. Steitz,T.A. (1999) DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.*, **274**, 17395–17398.
6. Kohlstaedt,L.A., Wang,J., Friedman,J.M., Rice,P.A. and Steitz,T.A. (1992) Crystal structure at 3.5 a resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*, **256**, 1783–1790.
7. Matray,T.J. and Kool,E.T. (1999) A specific partner for abasic damage in DNA. *Nature*, **399**, 704–708.
8. Kool,E.T. (2002) Active site tightness and substrate fit in DNA replication. *Annu. Rev. Biochem.*, **71**, 191–219.
9. Sloane,D.L., Goodman,M.F. and Echols,H. (1988) The fidelity of base selection by the polymerase subunit of DNA polymerase III holoenzyme. *Nucleic Acids Res.*, **16**, 6465–6475.
10. Goodman,M.F. and Fygenson,K.D. (1998) DNA polymerase fidelity: from genetics toward a biochemical understanding. *Genetics*, **148**, 1475–1482.
11. Goodman,M.F. (1997) Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. *Proc. Natl Acad. Sci. USA*, **94**, 10493–10495.
12. Echols,H. and Goodman,M.F. (1991) Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.*, **60**, 477–511.
13. Law,S.M., Eritja,R., Goodman,M.F. and Breslauer,K.J. (1996) Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. *Biochemistry*, **35**, 12329–12337.
14. Petruska,J., Goodman,M.F., Boosalis,M.S., Sowers,L.C., Cheong,C. and Tinoco,I.,Jr (1988) Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc. Natl Acad. Sci. USA*, **85**, 6252–6256.

15. Kunkel,T.A. and Bebenek,K. (2000) DNA replication fidelity. *Annu. Rev. Biochem.*, **69**, 497–529.
16. Harris,D., Kaushik,N., Pandey,P.K., Yadav,P.N. and Pandey,V.N. (1998) Functional analysis of amino acid residues constituting the dNTP binding pocket of HIV-1 reverse transcriptase. *J. Biol. Chem.*, **273**, 33624–33634.
17. Coffin,J.M. (1995) HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy. *Science*, **267**, 483–489.
18. Preston,B.D., Poiesz,B.J. and Loeb,L.A. (1988) Fidelity of HIV-1 reverse transcriptase. *Science*, **242**, 1168–1171.
19. Ji,J.P. and Loeb,L.A. (1992) Fidelity of HIV-1 reverse transcriptase copying RNA *in vitro*. *Biochemistry*, **31**, 954–958.
20. Fisher,T.S. and Prasad,V.R. (2002) Substitutions of Phe61 located in the vicinity of template 5′-overhang influence polymerase fidelity and nucleoside analog sensitivity of HIV-1 reverse transcriptase. *J. Biol. Chem.*, **277**, 22345–22352.
21. Lewis,D.A., Bebenek,K., Beard,W.A., Wilson,S.H. and Kunkel,T.A. (1999) Uniquely altered DNA replication fidelity conferred by an amino acid change in the nucleotide binding pocket of human immunodeficiency virus type 1 reverse transcriptase. *J. Biol. Chem.*, **274**, 32924–32930.
22. Harris,D., Yadav,P.N. and Pandey,V.N. (1998) Loss of polymerase activity due to Tyr to Phe substitution in the YMDD motif of human immunodeficiency virus type-1 reverse transcriptase is compensated by Met to Val substitution within the same motif. *Biochemistry*, **37**, 9630–9640.
23. Loya,S., Bakhanashvili,M., Tal,R., Hughes,S.H., Boyer,P.L. and Hizi,A. (1994) Enzymatic properties of two mutants of reverse transcriptase of human immunodeficiency virus type 1 (tyrosine 181–>isoleucine and tyrosine 188–>leucine), resistant to nonnucleoside inhibitors. *AIDS Res. Hum. Retroviruses*, **10**, 939–946.
24. Wainberg,M.A., Drosopoulos,W.C., Salomon,H., Hsu,M., Borkow,G., Parniak,M., Gu,Z., Song,Q., Manne,J., Islam,S. *et al.* (1996) Enhanced fidelity of 3TC-selected mutant HIV-1 reverse transcriptase. *Science*, **271**, 1282–1285.
25. Yarbrough,L.R., Schlageck,J.G. and Baughman,M. (1979) Synthesis and properties of fluorescent nucleotide substrates for DNA-dependent RNA polymerases. *J. Biol. Chem.*, **254**, 12069–12073.
26. Bebenek,K. and Kunkel,T.A. (1995) Analyzing fidelity of DNA polymerases. *Methods Enzymol.*, **262**, 217–232.
27. Kunkel,T.A. and Soni,A. (1988) Mutagenesis by transient misalignment. *J. Biol. Chem.*, **263**, 14784–14789.
28. Dauber-Osguthorpe,P., Roberts,V.A., Osguthorpe,D.J., Wolff,J., Genest,M. and Hagler,A.T. (1988) Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins*, **4**, 31–47.
29. Bebenek,K., Abbotts,J., Roberts,J.D., Wilson,S.H. and Kunkel,T.A. (1989) Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. *J. Biol. Chem.*, **264**, 16948–16956.
30. 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.
31. Kool,E.T. (2001) Hydrogen bonding, base stacking, and steric effects in dna replication. *Annu. Rev. Biophys Biomol. Struct.*, **30**, 1–22.
32. Kool,E.T. (1998) Replication of non-hydrogen bonded bases by DNA polymerases: a mechanism for steric matching. *Biopolymers*, **48**, 3–17.
33. 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.
34. Horlacher,J., Hottiger,M., Podust,V.N., Hubscher,U. and Benner,S.A. (1995) Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Natl Acad. Sci. USA*, **92**, 6329–6333.
35. Bebenek,K., Boyer,J.C. and Kunkel,T.A. (1999) The base substitution fidelity of HIV-1 reverse transcriptase on DNA and RNA templates probed with 8-oxo-deoxyguanosine triphosphate. *Mutat. Res.*, **429**, 149–158.
36. Lutz,M.J., Held,H.A., Hottiger,M., Hubscher,U. and Benner,S.A. (1996) Differential discrimination of DNA polymerase for variants of the non-standard nucleobase pair between xanthosine and 2,4-diaminopyrimidine, two components of an expanded genetic alphabet. *Nucleic Acids Res.*, **24**, 1308–1313.
37. Marx,A., MacWilliams,M.P., Bickle,T.A., Schwitter,U. and Giese,B. (1997) 4′-Acylated thymidines: a new class of DNA chain terminators and photocleavable DNA building blocks. *J. Am. Chem. Soc.*, **119**, 1131–1132.
38. Summerer,D. and Marx,A. (2001) Dna polymerase selectivity: sugar interactions monitored with high-fidelity nucleotides this work was supported by a grant from the deutsche forschungsgemeinschaft. We thank Professor Dr Michael Famulok for his continuing support. *Angew Chem. Int. Ed. Engl.*, **40**, 3693–3695.
39. Mu,L., Sarafianos,S.G., Nicklaus,M.C., Russ,P., Siddiqui,M.A., Ford,H.,Jr, Mitsuya,H., Le,R., Kodama,E., Meier,C. *et al.* (2000) Interactions of conformationally biased north and south 2′-fluoro-2′, 3′-dideoxynucleoside 5′-triphosphates with the active site of HIV-1 reverse transcriptase. *Biochemistry*, **39**, 11205–11215.
40. Marx,A., Amacker,M., Stucki,M., Hubscher,U., Bickle,T.A. and Giese,B. (1998) 4′-Acylated thymidine 5′-triphosphates: a tool to increase selectivity towards HIV-1 reverse transcriptase. *Nucleic Acids Res.*, **26**, 4063–4067.
41. Vastmans,K., Froeyen,M., Kerremans,L., Pochet,S. and Herdewijn,P. (2001) Reverse transcriptase incorporation of 1,5-anhydrohexitol nucleotides. *Nucleic Acids Res.*, **29**, 3154–3163.
42. Marx,A., Spichty,M., Amacker,M., Schwitter,U., Hubscher,U., Bickle,T.A., Maga,G. and Giese,B. (1999) Probing interactions between HIV-1 reverse transcriptase and its DNA substrate with backbone-modified nucleotides. *Chem. Biol.*, **6**, 111–116.
43. Martynov,B.I., Shirokova,E.A., Jasko,M.V., Victorova,L.S. and Krayevsky,A.A. (1997) Effect of triphosphate modifications in 2′-deoxynucleoside 5′-triphosphates on their specificity towards various DNA polymerases. *FEBS Lett.*, **410**, 423–427.
44. Alexandrova,L.A., Skoblov,A.Y., Jasko,M.V., Victorova,L.S. and Krayevsky,A.A. (1998) 2′-Deoxynucleoside 5′-triphosphates modified at alpha-, beta- and gamma-phosphates as substrates for DNA polymerases. *Nucleic Acids Res.*, **26**, 778–786.
45. Victorova,L., Sosunov,V., Skoblov,A., Shipytsin,A. and Krayevsky,A. (1999) New substrates of DNA polymerases. *FEBS Lett.*, **453**, 6–10.
46. Arzumanov,A.A., Semizarov,D.G., Victorova,L.S., Dyatkina,N.B. and Krayevsky,A.A. (1996) Gamma-phosphate-substituted 2′-deoxynucleoside 5′-triphosphates as substrates for DNA polymerases. *J. Biol. Chem.*, **271**, 24389–24394.
47. Ricchetti,M. and Buc,H. (1990) Reverse transcriptases and genomic variability: the accuracy of DNA replication is enzyme specific and sequence dependent. *EMBO J.*, **9**, 1583–1593.
48. Mendelman,L.V., Petruska,J. and Goodman,M.F. (1990) Base mispair extension kinetics. Comparison of DNA polymerase alpha and reverse transcriptase. *J. Biol. Chem.*, **265**, 2338–2346.
49. Gmeiner,W.H. and Walberer,B.J. (2001) *Base Pairing in DNA: Unusual Patterns*. Encyclopedia of Life Sciences, Nature Publishing Group.
50. Streisinger,G., Okada,Y., Emrich,J., Newton,J., Tsugita,A., Terzaghi,E. and Inouye,M. (1966) Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harb. Symp. Quant. Biol.*, **31**, 77–84.
51. Kunkel,T.A. (1986) Frameshift mutagenesis by eucaryotic DNA polymerases *in vitro*. *J. Biol. Chem.*, **261**, 13581–13587.
52. Larder,B.A., Bloor,S., Kemp,S.D., Hertogs,K., Desmet,R.L., Miller,V., Sturmer,M., Staszewski,S., Ren,J., Stammers,D.K. *et al.* (1999) A family of insertion mutations between codons 67 and 70 of human immunodeficiency virus type 1 reverse transcriptase confer multinucleoside analog resistance. *Antimicrob. Agents Chemother.*, **43**, 1961–1967.