Insights into DNA Polymerization Mechanisms from Structure and Function Analysis of HIV-1 Reverse Transcriptase†

Premal H. Patel,‡,§ Alfredo Jacobo-Molina,‡,§,|| Jianping Ding,§ Chris Tantillo,§ Arthur D. Clark, Jr.,§ Reetta Raag,§ Raymond G. Nanni,§ Stephen H. Hughes,⊥ and Edward Arnold*,§

Center for Advanced Biotechnology and Medicine and Department of Chemistry, Rutgers University, 679 Hoes Lane, Piscataway, New Jersey 08854-5638, and ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, P.O. Box B, Building 539, Frederick, Maryland 21702-1201

Received September 29, 1994; Revised Manuscript Received February 17, 1995®

ABSTRACT: When the single-stranded RNA genome of HIV-1 is copied into double-stranded DNA, the viral enzyme reverse transcriptase (RT) catalyzes the addition of approximately 20 000 nucleotides; however, the precise mechanism of nucleotide addition is unknown. In this study, we attempt to integrate the genetic data and biochemical mechanism of DNA polymerization with the structure of HIV-1 RT complexed with a dsDNA template-primer. The first step of polymerization involves the physical association of a polymerase with its nucleic acid substrate. A comparison of the structures of HIV-1 RT in the presence and absence of DNA indicates that the tip of the p66 thumb moves approximately 30 Å upon DNA binding. This conformational change permits numerous interactions between residues of α-helices H and I in the thumb subdomain and the DNA. Measurements of DNA binding affinity for nucleic acids with double-stranded DNAs that have an increasing number of bases in the template overhang and molecular modeling suggest that portions of β3 and β4 within the fingers subdomain bind single-stranded regions of the template. Measurements of nucleotide incorporation efficiency (k_{cat}/K_m) show that the binding and incorporation of the next complementary nucleotide are not dependent on the length of the template overhang. Molecular modeling of an incoming nucleotide triphosphate (dTTP), based in part on the position of mercury atoms in a RT/DNA/Hg-UTP/Fab structure, suggests that portions of secondary structural elements αC−β6, αE, β11b, and β9−β10 determine the topology of the dNTP-binding site. These results also suggest that nucleotide incorporation is accompanied by a protein conformational change that positions the dNTP for nucleophilic attack. Nucleophilic attack by the oxygen atom of the 3′-OH group of the primer strand could be metal-mediated, and Asp185 may be directly involved in stabilizing the transition state. The translocation step may be characterized by rotational as well as translational motions of HIV-1 RT relative to the DNA double helix. Some of the energy required for translocation could be provided by dNTP hydrolysis and could be coupled with conformational changes within the nucleic acid. A structural comparison of HIV-1 RT, Klenow fragment, and T7 RNA polymerase identified regions within T7 RNA polymerase which are not present in the other two polymerases that might help this polymerase to remain bound with nucleic acids and contribute to the ability of the T7 RNA polymerase to polymerize processively.

Human immunodeficiency virus type 1 (HIV-1)[1] reverse transcribes its RNA genome into double-stranded DNA by a complex mechanism involving RNA- and DNA-templated polymerase and RNase H activities [for review, see Coffin (1990)]. These activities are carried out by HIV-1 reverse transcriptase (RT), a multifunctional enzyme composed of two asymmetric subunits of 66 kDa and 51 kDa [p66 and p51; for review, see Jacobo-Molina and Arnold (1991)]. During reverse transcription of the viral genome, HIV-1 RT also catalyzes strand transfer (Luo & Taylor, 1990; Peliska & Benkovic, 1992) and strand displacement (Huber et al., 1989; Hottinger et al., 1994) reactions, and the enzyme exhibits poor fidelity (Preston et al., 1988; Roberts et al., 1988; Ji & Loeb, 1992; Yu & Goodman, 1992; Patel & Preston, 1994) and low processivity (Huber et al., 1989; Abbotts et al., 1993; Klarmann et al., 1993). High-resolution structures of RT complexed with relevant intermediates may help to understand the enzymatic mechanisms that underlie these activities.

Because of its pivotal role in the HIV-1 life cycle and as the target of the only clinically approved anti-retroviral agents, there have been extensive biochemical, structural, and site-directed mutagenesis analyses of HIV-1 RT. This wealth of data renders HIV-1 RT an ideal candidate for studies on the mechanism of polymerization. At least three laboratories (Kati et al., 1992; Reardon, 1992, 1993; Hsieh et al., 1993) have studied the pathway for polymerization using rapid-quench, pre-steady-state kinetic analysis. The

† The work in E.A.'s laboratory has been supported by NIH Grants AI26790 and AI36144, Janssen Research Foundation, a Johnson & Johnson Focused Giving Award, and a grant from the American Foundation for AIDS Research (to A.J.-M.). Research in S.H.H.'s laboratory is sponsored by the National Cancer Institute, DHHS, under Contract No. N01-CO-4600 with ABL and the National Institute of General Medical Sciences.

‡ Contributed equally to the work.

§ Rutgers University.

|| Current address: Centro de Biotecnología y Departamento de Química, Instituto Tecnológico de Monterrey, Sucursal de Correos "J", Monterrey NL 64849, Mexico.

⊥ NCI-Frederick Cancer Research and Development Center.

® Abstract published in *Advance ACS Abstracts,* April 1, 1995.

[1] Abbreviations: dNTP, 2′-deoxynucleotide 5′-triphosphate; HIV-1; human immunodeficiency virus type 1; PP_i, pyrophosphate; RT, reverse transcriptase; UTP, uridine 5′-triphosphate.

0006-2960/95/0434-5351$09.00/0

Table 1: Oligonucleotides

5'-GTCCCTGTTCGGGCGCCA CAGGGACAAGCCCGCGGTA	19/18
5'-GTCCCTGTTCGGGCGCCA CAGGGACAAGCCCGCGGTATC	21/18
5'-GTCCCTGTTCGGGCGCCA CAGGGACAAGCCCGCGGTATCGAT	24/18
5'-GTCCCTGTTCGGGCGCCA CAGGGACAAGCCCGCGGTATCGATCGATCG	30/18
5'-GTCCCTGTTCGGGCGCCA GGTCAGGGACAAGCCCGCGGTATCGAT	27/18

proposed pathway for a single nucleotide incorporation involves the sequential binding of RT with DNA, and then with dNTP, followed by a nucleophilic attack that leads to the formation of a phosphodiester bond, with subsequent release of the pyrophosphate moiety.

The folding of HIV-1 RT was first described from a 3.5 Å resolution structure (Kohlstaedt et al., 1992). This study showed that the overall HIV-1 RT structure is related to that of the large (Klenow) fragment of *Escherichia coli* DNA polymerase I in that both enzymes are morphologically related to a human right hand. The polymerase domain was found to contain four distinct subdomains (i.e., fingers, palm, thumb, and connection). While this study described the overall folding of the HIV-1 RT, the enzyme was not bound to a relevant substrate and therefore did not provide details regarding the mechanism of DNA polymerization.

We have reported a structure of HIV-1 RT complexed with a dsDNA template-primer at 3.0 Å resolution (Jacobo-Molina et al., 1993). This structure showed which elements of RT interact with the primer strand ($\beta12-\beta13$ connecting loop of HIV-1 RT) and the template strand ($\beta4$, αB, $\beta8-\alpha E$ connecting loop, and $\beta5a$) which we termed the "primer grip" and the "template grip", respectively. In the present study, we attempt to integrate into a structural framework biochemical and genetic data regarding DNA polymerase mechanisms prior and subsequent to the DNA- and dNTP-binding steps, and we describe the structural elements of HIV-1 RT that are likely to be involved in each step of DNA polymerization.

EXPERIMENTAL PROCEDURES

Materials. HIV-1 RT was purified to homogeneity as a p66/p51 heterodimer and concentrated to 40.3 mg/mL (340 μM) as previously described (Jacobo-Molina et al., 1991; Clark et al., 1995). The concentration of the purified protein was determined by a modified procedure of Lowry et al. (1951), and the obtained results are in good agreement with the concentration of active RT molecules determined by active site titration experiments. DNA oligonucleotides (Table 1) were purchased from the DNA synthesis facility at Rutgers University and were purified by ion-exchange chromatography (Jacobo-Molina et al., 1991; Clark et al., 1995). The concentrations of the purified oligonucleotides were determined by reading UV absorbance at 260 nm. T4 polynucleotide kinase, 2′-deoxyribonucleotide 5′-triphosphates (dNTPs), and poly(rA)•oligo(dT)$_{12-18}$ were purchased from Pharmacia LKB Biotechnology. [γ-^{32}P]ATP for labeling the 5′-ends of DNA was obtained from Amersham.

Measurement of RT−DNA Binding Affinity. The 5′-^{32}P-labeled 18-mer primer was prepared by standard protocols (Sambrook et al., 1989) and was hybridized in a 1:1 molar ratio with templates of varying lengths (yielding double-stranded DNAs with template overhangs up to 12 bases in length; Table 1) in 67 mM HEPES (pH 7.3) and 200 mM KCl as described (Preston et al., 1988). Most duplex DNA substrates (2 to 100 nM) were preincubated with HIV-1 RT (14 nM), Tris•HCl (pH 8; 20 mM), DTT (2 mM), and BSA (0.1 mg/mL) at room temperature for 5 min without the addition of divalent cations. The 19-mer/18-mer duplex (19/18 dsDNA) was preincubated in the presence and absence of divalent cations (Mn^{2+} at 5 mM). All reactions were initiated by the addition of the four dNTPs (500 μM each), $MgCl_2$ (5 mM), and a poly(rA)•oligo(dT)$_{12-18}$ (0.1 mg/mL) trap, and the reactions were stopped after a 5-min incubation at 37 °C with the addition of EDTA (pH 8) to 7.5 mM in the absence of Mn^{2+} preincubation and to 15 mM in the presence of Mn^{2+}. Incubations in which HIV-1 RT was initially preincubated with poly(rA)•oligo(dT)$_{12-18}$ (0.1 mg/mL) and subsequently initiated by the addition of dNTPs (500 μM each), $MgCl_2$ (5 mM), and DNA (2 and 100 nM) resulted in no extension of the end-labeled nucleic acid substrates. Thus, the nonradiolabeled trap bound to the free enzyme molecules and restricted polymerization to a single cycle (Patel & Preston, 1994). Products were separated by 20% polyacrylamide−urea gel electrophoresis and quantified by the use of a phosphorimager (Molecular Dynamics). Incubations with each concentration of DNA were done in duplicate, and the amount of products synthesized in a single cycle of polymerization (burst amplitudes) was averaged for the two sets of reactions and graphed with respect to the amount of duplex DNA initially reacted. The graph was fitted to the equation:

$$[E\cdot TP] = 0.5(K_{d_1} + E_t + TP) - 0.5\sqrt{(K_{d_1} + E_t + TP)^2 - 4E_tTP}$$

where [E•TP], K_{d_1}, E_t, and TP represent the productive RT−DNA complex concentration, equilibrium dissociation constant for RT−DNA binding, total enzyme concentration, and total DNA concentration, respectively (Carroll et al., 1991).

The K_{d_1} of the dsDNA with varying template overhang lengths ranged from 5 to 38 nM (Figure 2), indicating that the affinity of RT for a duplex with a 12-base overhang is greater than that with a single-base overhang. The K_{d_1} of 19/18 dsDNA in the presence of $MnCl_2$ showed that binding was 4−5 times tighter in the presence than in the absence of preincubation with a metal; this is consistent with the 10-fold increase in HIV-1 RT binding affinity to dsDNA in the presence of $MgCl_2$ reported by Modrich and co-workers (Hsieh et al., 1993). Incubations done in the presence of the divalent cation also yielded higher burst amplitudes; these increased product yields were most prevalent at low DNA concentrations. While incubations with 2 nM 19/18 dsDNA resulted in 20% extension of the substrate, incubations with the 2 nM 19/18 dsDNA•Mn^{2+} complex resulted in a 60% extension after a single cycle of polymerization.

Although RT binding affinity of the 27/18 dsDNA nucleic acid (containing a single 3′-OH group that can serve as a substrate for extension) was similar to that of 24/18 dsDNA (containing two such 3′-OH groups), the incubations with 27/18 dsDNA yielded a 20% higher concentration of "active" enzyme molecules, suggesting that ~20% of 24/18 dsDNA

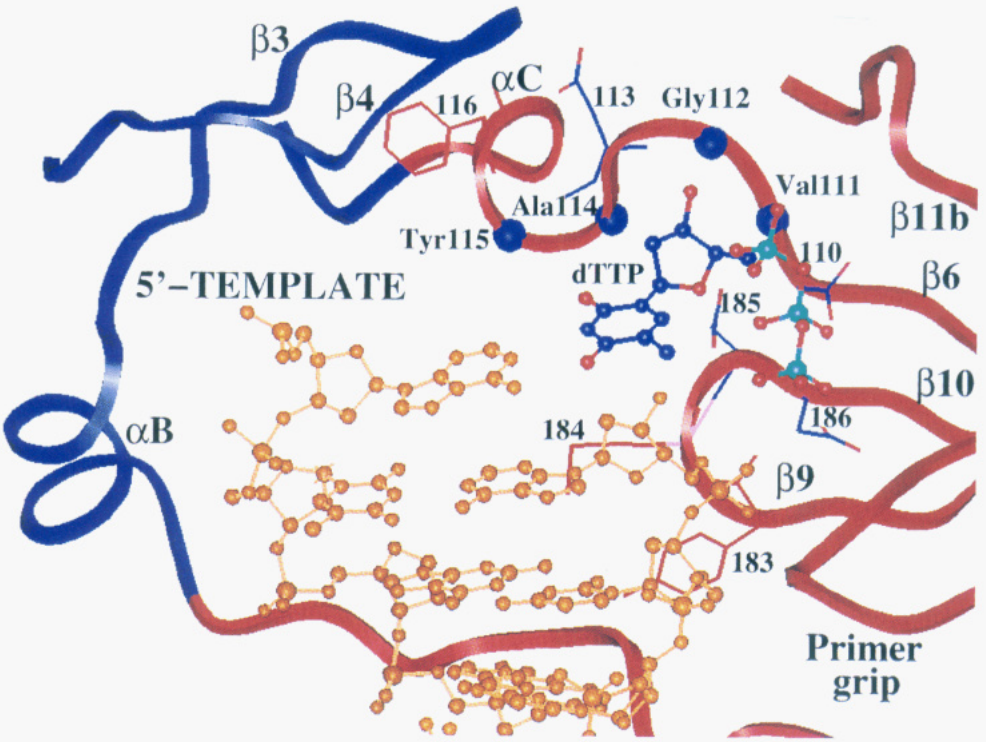


FIGURE 1: Ribbon diagram of the HIV-1 polymerase catalytic site. Secondary structural elements within the palm (colored in red) and fingers (colored in blue) subdomains that are in position to interact with the incoming dTTP and the DNA are highlighted. Side chains of several amino acids that likely have important roles during the dNTP-binding and incorporation steps are shown.

are oriented such that the blunt end is in the polymerase catalytic site (Patel & Preston, 1994). A comparison of [E·TP] from active site titration experiments with E_t demonstrated that the enzyme was ~90% active. This high specific activity may be due to the care taken in our purification steps to produce HIV-1 RT for crystallization purposes. Since significant amounts of products were formed after a single cycle of polymerization with all DNA oligonucleotides, the rate of polymerization is faster than the rate of enzyme dissociation even when there is only a one-base template overhang.

Measurement of dTTP Incorporation Efficiency. Time-course experiments were conducted with each nucleic acid construct to determine the conditions at which HIV-1 RT incorporates nucleotides at a maximum steady-state rate as described (Patel & Preston, 1994). The steady-state Michaelis−Menten parameters k_{cat} and K_m were then measured by incubating limiting amounts of HIV-1 RT (30 pM) with 2 nM indicated nucleic acids in the presence of varying concentrations of dTTP (5 nM to 5 μM) at 37 °C as described (Boosalis et al., 1987).

Molecular Modeling of Docked dNTP. Biochemical data suggest that polymerases undergo a conformational change subsequent to dNTP binding. The structural analysis described here is based on modeling the dNTP at two locations separated by ~4 Å which could correspond to bound dNTP before and after a conformational change. The criteria used to dock the dNTP into the RT-DNA complex at the first site include the following: (1) the incoming base should form Watson−Crick hydrogen bonds with the 5′ template base and (2) the β- and γ -phosphates chelate a Mg^{2+} that is shared with Asp110 and Asp186, consistent with known coordination geometry. It should be noted that the electron density for the RT−DNA complex indicates that the template overhang is positioned over the center of the first duplex base pair. Modeling of the dNTP bound at the second site was based on electron density maps identifying the mercury position of the Hg-UTP derivative which was soaked into the RT/dsDNA/Fab crystals as described (Arnold et al., 1992; Jacobo-Molina et al., 1993).

RESULTS AND DISCUSSION

HIV-1 RT Polymerase Active Site. The location of the polymerase active site in HIV-1 RT was first described at 7 Å resolution by determining the structure of HIV-1 RT bound to DNA and a mercurated UTP (Arnold et al., 1992). The heavy metal covalently attached to an incoming nucleotide (Hg-UTP) yielded a strong signal in the electron density at a position not found with mercury-containing reagents that did not contain a nucleotide. The positions of mercury in the RT/dsDNA/Hg-UTP/Fab structure have since been resolved at 3.5 Å resolution and found to be near the 3′ primer terminus; in addition, the RT/dsDNA/Fab structure has been partially refined at 2.8 Å resolution (Ding et al., unpublished results). At this resolution, the nucleotide-binding site can be characterized. The amino acid side chains in close proximity to a modeled incoming dNTP include those in secondary structural elements αC−β6, αE, β9−β10, and β11b (amino acid residues 110−117, 160−161, 183−186, and 219−221; Figure 1). These residues, which are well conserved within the *pol* genes of many retroviruses, are involved in forming the topology of the nucleotide-binding site.

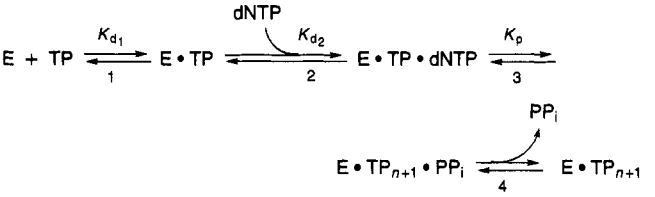
The nucleotide-binding site is located in the palm subdomain of the p66 subunit of HIV-1 RT. Among the most conserved elements of this region in the RTs of HIV and other retroviruses in the lentivirus subgroup is the YMDD motif (Tyr-Met-Asp-Asp amino acid residues 183−186) located in β9−β10 (Kohlstaedt et al., 1992; Jacobo-Molina

et al., 1993). This motif, along with Asp110 (located on β6), is likely to be involved in binding the Mg^{2+}−dNTP complex. The structure of HIV-1 RT complexed with DNA suggests that Asp185, Asp110, and Asp186 are in position to chelate a Mg^{2+}−dNTP complex directly (Figure 1). The recently published structure of DNA polymerase β with nucleic acid substrate bound shows Mg^{2+} at the polymerase active site chelated in an analogous fashion with Asp256, Asp192, and Asp190 (Pelletier et al., 1994). [However, the template-primer substrate is in the opposite orientation (see Appendix).] This aspartic acid triad is well conserved among nucleic acid polymerases (Delarue et al., 1990) and is present in reverse transcriptases from a broad class of retroviruses (Poch et al., 1989). Mutant RTs in which these residues are substituted, including Asp110Gln, Asp110Glu, Asp185His, Asp185Glu, Asp186Asn, and Asp186Glu, were inactive (Larder et al., 1989; Boyer et al., 1992). These studies suggest that the precise geometry of the aspartic acid side chains is critical for enzymatic activity since a single methylene extension of the carboxylic group (Asp to Glu) at any site in the aspartic triad produced inactive enzymes.

During the dNTP-binding step and prior to nucleotide incorporation, polymerases must select the complementary base. Although constraints imposed by Watson−Crick base pairing are presumably major factors in selecting the correct nucleotide, reverse transcriptases have fidelities varying by over 100-fold with identical DNA templates [reviewed in Preston and Garvey (1992)], suggesting that the structure of the polymerase active site must also be involved in the selection of the proper base. As described above, the residues of the secondary structural elements αC−β6, αE, β9−β10, and β11b are in positions where they could affect the topology of the dNTP-binding site and can be directly involved in fidelity. Specifically, Tyr115 in αC may interact directly with the base of the incoming nucleotide. This residue is likely to be analogous to Tyr865 of mammalian polymerase α (Dong et al., 1993) and Tyr766 of Klenow fragment (Carroll et al., 1991), which are thought to be involved in fidelity.

Along with binding complementary dNTPs, the polymerase active site must also position the growing primer and the template chains. The amino acids that interact with the 3′ nucleotides of the primer strand include residues within the primer grip (i.e., Met230 and Gly231) and the Tyr-Met-Asp residues of the YMDD motif. Met184Val mutants of HIV-1 RT have been shown to confer resistance to the nucleoside analogs 3′-thiacytidine (3TC) and 5-fluoro-3′-thiacytidine (FTC) [Schinazi, 1993; Tisdale et al., 1993; for a compilation of drug-resistance mutations, see Schinazi et al. (1994); for review in context of HIV-1 RT structure, see Tantillo et al. (1994)]. Amino acid residues near the polymerase active site that position the template strand include those in the template grip: residues Gln151 to Trp153 and Leu74 to Arg78 and Glu89. Mutations within the template grip have also been shown to decrease drug sensitivity [e.g., Leu74Val (St. Clair et al., 1991) and Glu89Gly (Prasad et al., 1991)]. Mutations that affect the residues that interact with the DNA could also affect fidelity indirectly by altering the position or the conformation of the template-primer at the active site. These changes in the topography of the active site could lead to an enhanced discrimination between nucleotides and their analogs and give rise to a drug-resistant phenotype (Boyer et al., 1994b).

Scheme 1[a]



[a] E, TP, dNTP, and PP$_i$ represent HIV-1 RT, nucleic acid template-primer, deoxyribonucleotide, and inorganic pyrophosphate, respectively.

Mutations of key amino acid residues involved in nucleic acid interactions may also affect other polymerase characteristics, including processivity. By affecting fidelity and processivity, mutations within the DNA binding cleft that decrease drug sensitivity may also compromise (or enhance) polymerase function and attenuate (or enhance) viral replication.

Pathway of DNA Polymerization. During replication of the retroviral RNA genome, HIV-1 RT must polymerize approximately 20 000 nucleotides (almost 50% of these events are RNA-templated) to generate double-stranded DNA [for review, see Coffin (1990)]. The steps involved during nucleotide addition and the rates at which each step occurs have been determined (Majumadar et al., 1988; Kati et al., 1992; Reardon, 1992, 1993; Hsieh et al., 1993). A general mechanism for a single nucleotide addition is shown in Scheme 1. During DNA synthesis (Scheme 1), (1) RT binds with its template-primer, (2) the appropriate dNTP binds to the RT−nucleic acid complex, (3) a nucleophilic attack results in phosphodiester bond formation, and (4) the pyrophosphate is released.

Biochemical experiments show that template-primer binding to HIV-1 RT and the subsequent dNTP-binding steps occur rapidly and efficiently. The presence of a burst of products in reactions confined to a single cycle of polymerization shows that DNA and dNTP binding are faster steps than enzyme dissociation, even from template-primers with only a single template−base overhang (Figure 2). The rate-limiting step is either phosphodiester bond formation or the putative RT conformational change preceding nucleotide incorporation (k_p; Kati et al., 1992, 1993; Hsieh et al., 1993). Overall, HIV-1 RT can complete the four steps in ~0.05 s (Kati et al., 1992; Reardon, 1992). If the reaction continued in a distributive fashion, then the rate of enzyme turnover, specifically the rate of RT dissociation, becomes rate limiting (see below).

The faithful insertion of nucleotides during polymerization involves dynamic interactions between HIV-1 RT and its nucleic acid and dNTP substrates. During these interactions, RT has been proposed to undergo at least three significant conformational changes. Tryptophan fluorescence quenching studies of HIV-1 RT in the presence and absence of DNA suggest that RT first changes conformation upon binding to nucleic acids (Divita et al., 1993). Polymerases are thought to undergo a second conformational change subsequent to dNTP binding immediately prior to the chemical catalysis step. This putative conformational change is thought to facilitate the interaction of the incoming nucleotide with the 3′-OH of the primer (Patel et al., 1991; Kati et al., 1992; Hsieh et al., 1993). The final conformational change involves translocation toward the new 3′-OH primer terminus.

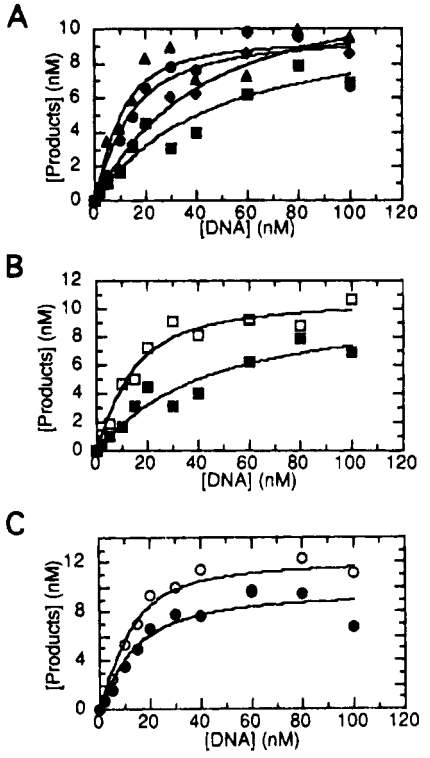


FIGURE 2: Equilibrium dissociation constant for the polymerase–dsDNA complex. HIV-1 RT was preincubated for 5 min at room temperature in the presence of 2, 5, 10, 15, 20, 30, 40, 60, 80, and 100 nM DNA. Reactions were initiated and products analyzed as described. (A) The K_{d_1} of 19/18 (■), 21/18 (◆), 24/18 (●), and 30/18 (▲) template-primer constructs were 38 ± 16 nM, 32 ± 7 nM, 8 ± 3 nM, and 5 ± 2 nM, respectively (±curve fit error). (B) The K_{d_1} values of 19/18 preincubated for 5 min in the absence (■) and presence (□) of 5 mM $MnCl_2$ at room temperature were 38 ± 16 nM and 12 ± 4 nM, respectively. (C) The K_{d_1} of 27/18 (○) construct containing a single 3′-OH group (to prevent RT binding to blunt ends) was 5 ± 2 nM . The active enzyme concentration in incubations containing 27/18 was 12 ± 1 nM and in incubations containing 24/18 (●) was 10 ± 1 nM.

Template-Primer Binding. The first step of polymerization involves the physical association of the polymerase with its nucleic acid substrate. HIV-1 RT binds double-stranded nucleic acids tightly with a K_{d_1} of 5–38 nM (Figure 2), and fluorescence quenching studies suggest that RT binds to DNA by a two-step mechanism (Divita et al., 1993). The structural data are consistent with these results. The structure of HIV-1 RT in the absence of DNA shows the p66 thumb subdomain folded over the p66 palm subdomain, coming in contact with the fingers subdomain (Figure 3; Raag et al., 1994). In this structure, the nucleic acid binding cleft is ~15 Å in diameter, which is large enough to accommodate a single-stranded but not a double-stranded DNA. This may have biological significance since it is possible that, prior to polymerization, HIV-1 RT may bind first to single-stranded RNA or DNA and then slide until it encounters the primer 3′-OH. This two-step DNA binding mechanism in which the rate of the first step is dependent on RT concentration and the rate of the second step is independent of RT concentration is consistent with changes in fluorescence pattern observed upon DNA binding (Divita et al., 1993).

After RT binds dsDNA, the tip of the p66 thumb travels approximately 30 Å away from the fingers subdomain (Figure 3). We have recently solved the structure of RT bound to the nonnucleoside inhibitor α-APA (Ding et al., 1995). This structure is similar to the structure of the HIV-1 RT/nevirapine complex (Smerdon et al., 1994). In both of these structures, the position of the thumb is nearly the same as in the RT/dsDNA/Fab structure. Binding a nonnucleoside inhibitor alters the position and presumably also affects the mobility of the thumb subdomain. This could inhibit, for example, enzymatic turnover (Gopalakrishnan & Benkovic, 1994) or could lower the processivity of the enzyme by influencing translocation (see below).

The RT–DNA interactions involve elements primarily from the thumb, fingers, and palm subdomains of the p66 subunit (Jacobo-Molina et al., 1993; Ding et al., 1994). Amino acid residues in helices H and I of the thumb subdomain have extensive interactions with the DNA. Other interactions occur between the β12–β13 hairpin (primer grip) and β4, αB, β8–αE, and β5a (template grip) and the base pairs closest to the polymerase active site. This portion of the DNA is in the A conformation (Jacobo-Molina et al., 1993). The interactions between the residues of the thumb, palm, and fingers subdomains with the first six base pairs may displace key water molecules. Since DNA hydration can affect the B to A transition, the loss of the water could influence the propensity of the DNA near the polymerase active site to assume an A form. The rest of the molecule may simply assume the B form since, in the absence of strong DNA–protein interactions, this is the preferred conformation. The template-primer also assumes an A-form/B-form hybrid geometry in a complex with pol β (Pelletier et al., 1994),

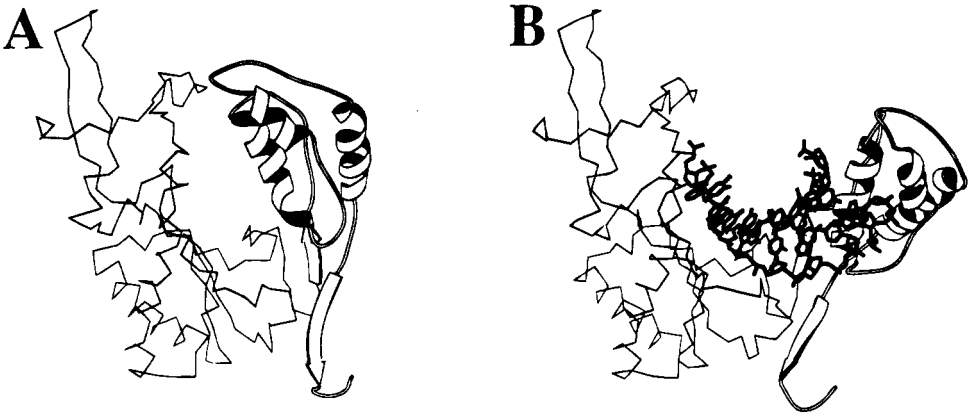
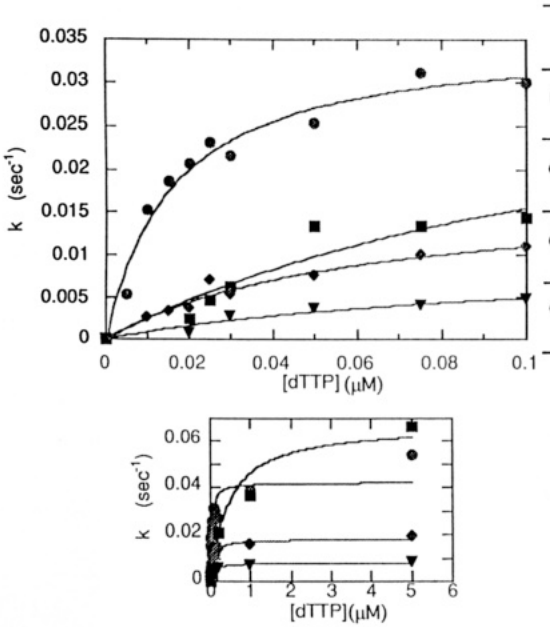


FIGURE 3: Structure of HIV-1 in the (A) absence and (B) presence of DNA. The fingers, palm, and thumb subdomains (highlighted in ribbons) of HIV-1 RT are shown. In the absence of DNA, the tip of the thumb subdomain interacts with the β3–β4 portion of the fingers subdomain. In the presence of dsDNA the thumb travels 25–30 Å away from the fingers subdomain (Raag et al., 1994).



T/P	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m ($sec^{-1}\mu M^{-1}$)
r19/d18	0.042 ±0.0033	0.026 ±0.0072	1.6
d19/d18	0.067 ±0.0059	0.46 ±0.12	0.15
d21/d18	0.018 ±0.00067	0.064 ±0.0072	0.28
d30/d18	0.0079 ±0.00044	0.072 ±0.013	0.11

FIGURE 4: Efficiency of dTTP incorporation. HIV-1 RT (30 pM) was incubated in the presence of 0.005, 0.010, 0.015, 0.020, 0.025, 0.030, 0.050, 0.075, 0.10, 0.20, 1.0, and 5.0 μM dTTP and 2.0 nM indicated nucleic acid. Reactions were conducted for either 200 s for r19/d18 (●), d19/d18 (■), and d21/d18 (◆) or 600 s for d30/d18 (▼), and products were quantified as described above. The k_{cat}/K_m values obtained from curve fits of the plots directly measure the efficiency of dTTP incorporation at these sites. The k_{cat} value represents the rate of RT disassociation after the 19-mer is synthesized. The rate of disassociation slows 6−8-fold as the template length is increased from a 1-base to a 12-base overhang. Thus, the increased binding affinity observed for dsDNA containing longer template overhangs (Figure 2) is largely due to slower disassociation rates (as opposed to faster association rates). The disparity in the K_m values for single-templated nucleotide incorporation onto the nucleic acid substrates does not necessarily reflect differences in dNTP binding affinity (K_{d_2}; $K_m = K_{d_2}k_{cat}/k_p$) and is more likely attributed to the differences in dissociation rates (k_{cat}).

and it is possible that HIV-1 RT does not constrain template-primers to A-form geometry in order to accommodate RNA templates. This unusual nucleic acid structure may have general relevance for polymerization by diverse polymerases.

The measured binding affinity of RT with DNA duplexes that have varying template extensions suggests that additional amino acid residues contact the template strand. The binding affinity of RT for 27/18 dsDNA and 24/18 dsDNA (both with six-base overhangs; Table 1) is 5−8 times greater than that for a 19/18 dsDNA (with a single-base overhang; K_{d_1} of 5−8 nM and 38 nM, respectively); binding affinity is not enhanced significantly when the size of the overhang is increased from 6 to 12 nucleotides (Figure 2A). In comparison, the K_{d_1} for blunt-end DNA is ~50 nM (Patel and Preston, manuscript in preparation). The enhanced binding affinity of duplexes containing overhangs of up to six nucleotides likely results from an increase in protein-DNA contacts. Repeating this experiment with the 19/18 dsDNA substrate in the presence of a divalent cation also yielded lower K_{d_1} values, indicating that divalent cations promote RT−DNA interactions.

A combination of biochemical data and molecular modeling suggests that side chains of residues near β4 (amino acids 65−74) within the p66 fingers subdomain interact with single-stranded templates extending approximately six nucleotides (Boyer et al., 1994a). This region is highly conserved among retroviral RTs and corresponds to the homologous region A (Larder et al., 1987). There are several residues with basic side chains in this region which may form stable interactions with the phosphates of an extended template overhang (unpublished results and Wöhrl et al., 1995). Furthermore, enzymatic footprinting of HIV-1 RT complexed with DNA combined with structural analysis suggests that the fingers subdomain of p66 protects about seven nucleotides of the single-stranded template (Wöhrl et al., 1995). We are currently studying the structure of RT bound to nucleic acid duplexes with longer template overhangs to determine the precise nature of the interactions between the template and the fingers subdomain.

Other regions that lie near the template-primer include the connection subdomain and the RNase H domain of p66 (Jacobo-Molina et al., 1993; Ding et al., 1994). Enzymatic footprinting in combination with molecular modeling of HIV-1 RT−DNA complexes indicates that RT protects about six base pairs (toward the 3′-end of the template strand) of dsDNA extended beyond the RNase H active site in the RT/dsDNA/Fab structure (Wöhrl et al., 1995). These studies also suggest that when duplexes extend beyond the RNase H active site, the p51 thumb subdomain contributes either directly or indirectly to template-primer binding.

dNTP Binding. After binding of HIV-1 RT to nucleic acids, the next steps of polymerization include binding the appropriate dNTP and the nucleotidyl-transfer reaction which lead to incorporation of the nucleotide. To characterize the efficiency of nucleotide incorporation, the steady-state k_{cat} and K_m values of dTTP incorporation were measured using template-primers with varying template lengths as described above. The plots of rates of nucleotide incorporation as a function of dTTP concentration exhibited typical Michaelis−Menten saturation kinetics (Figure 4). The second-order k_{cat}/K_m value obtained from fitting a hyperbolic curve directly measures the efficiency of dTTP incorporation, and the results show that whether the DNA duplex contained a 1-base, a 3-base, or a 12-base template overhang, the incorporation efficiency for a single residue was essentially the same (Figure 4). These results demonstrate that HIV-1 RT incorporates the last templated base to form a blunt-end dsDNA as efficiently as it adds a nucleotide when there is a

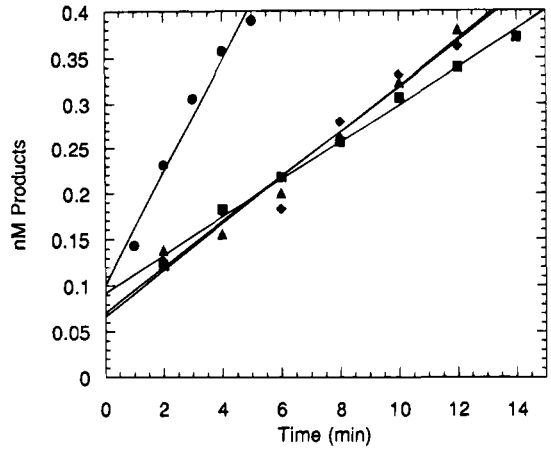


FIGURE 5: Rate of ddTMP incorporation in the presence of the next complementary nucleotide(s). HIV-1 RT (30 pM, assuming the enzyme is quantitatively active) was incubated with 27/18 dsDNA (2 nM; Table 1) in the presence of ddTTP (250 μM; ●); ddTTP (250 μM) and dATP (physiologically relevant 50 μM; ■); ddTTP (250 μM) and dATP and dGTP (50 μM each; ◆); and ddTTP (250 μM) and dATP, dGTP, and dCTP (50 μM each; ▲). Incubations containing ddTTP were stopped after every minute up to 10 min, and incubations containing ddTTP and the next complementary dNTP(s) were stopped after every 2 min up to 20 min. All incubations resulted in the extension of 27/18 dsDNA by one nucleotide. The amount of products synthesized (<20% of the DNA incubated to ensure steady-state conditions) was plotted as a function of time. The 27/18 dsDNA substrate was constructed such that it contained a six-base overhang and lacked a blunt end (which can efficiently be extended by HIV-1 RT) at the opposite end.

12-base single-stranded template extension. This suggests that the RT-DNA contacts that promote the binding of duplexes with longer template overhangs (Figure 2) do not significantly influence the incorporation of a single templated nucleotide, although the selection of nucleotide analogs can be affected (Boyer et al., 1994a). In contrast, nucleotide incorporation onto a 19-mer/18-mer with an RNA template was 5−10 times more efficient than onto a 19-mer/18-mer with a DNA template.

To determine what regions of RT are involved in dNTP binding and incorporation, an incoming dTTP was modeled into the putative nucleotide-binding site (Figure 1). In this model, the incoming nucleotide can base pair with the template, and portions of β6, the loop connecting β6 and αC, the β9−β10 hairpin, portions of αE, and portions of β11b form the dNTP-binding site. Specifically, Tyr115 of αC interacts directly with the base of the incoming nucleotide, and Asp185 interacts with the α-phosphate, while Asp110 and Asp186 interact with the β- and γ-phosphates via bivalent coordination of a Mg^{2+} ion. It is also possible that, during processive synthesis, in addition to binding the next complementary nucleotide, two (or more) complementary nucleotides are bound to the RT/DNA complex.

To test if binding of several successive dNTPs kinetically favors polymerization, the rate of dideoxythymidine nucleotide incorporation onto 27/18 dsDNA was measured in the presence and absence of the next complementary dNTP. Product formation (27/19 dsDNA with a 3′-H) was quantified and graphed as a function of incubation time (Figure 5). The four steady-state rates measure the rate of RT dissociation from the chain-terminated template-primer (RT-27/19 in complex with a dsDNA with 3′-H instead of a 3′-OH), from the RT-27/19 (3′-H) dsDNA−dATP complex (which cannot undergo chemical catalysis), from the putative RT-27/19 (3′-H) dsDNA−dATP−dGTP complex, or from the putative RT-27/19(3′-H) dsDNA−dATP−dGTP−dCTP complex, respectively. The k_{cat} value for ddTMP incorporation, obtained by dividing the slope by enzyme concentration (30 pM, assuming the enzyme is quantitatively active), equaled 0.034 s^{-1} in incubations containing ddTTP alone and 0.012 s^{-1} in incubations containing ddTTP and the next complementary nucleotide (dATP at 50 μM). The rate of RT dissociation from an RT-27/19 (3′-H) dsDNA−dATP complex is 3 times slower than the dissociation from an RT-27/19 (3′-H) dsDNA complex. These results are similar to published values (Muller et al., 1991; Kati et al., 1992) and indicate that binding of a single dNTP (at the aspartic triad) stabilizes the RT−DNA−dNTP complex. This enhanced stability of the complex could be due to the additional protein−nucleic acid contacts created after dNTP binds (see below). There was no measurable difference in the k_{cat} for ddTMP incorporation in the presence of the next correct dNTP with those in the presence of the next two and three complementary dNTPs (0.012, 0.014, 0.014 s^{-1}, respectively), indicating that the presence of multiple nucleotides does not further stabilize the RT−DNA−dNTP complex and that there is no apparent kinetic advantage for HIV-1 RT to bind multiple successive nucleotides (beyond a single dNTP) during polymerization.

Larder et al. (1987) identified two regions of homology which are conserved among retroviral RTs that correspond to the putative dNTP-binding site (region B, Leu109 to Phe116 in β6 and αC, and region E, Ile180 to Gly190 in β9−β10). Mutations at these sites disrupt the catalytic activity of HIV-1 RT. The mutations Asp110Gln (Larder et al., 1987) and Asp110Glu (Boyer et al., 1992), as mentioned before, demonstrated that Asp110 is an essential amino acid. In contrast, amino acid substitutions Asp113Gly (Larder et al., 1987) and Asp113Glu (Larder et al., 1989; Boyer et al., 1992) yielded mutant enzymes with more than 50% the activity of wild type; however, these mutants are resistant to either AZTTP, PFA, and/or ddTTP (Larder et al., 1987, 1989; Lowe et al., 1991). These data show that although Asp113 is not critical for the activity of the enzyme, this residue is involved in distinguishing normal nucleotides from analogs. Other mutations at the putative dNTP-binding site include Ala114Ser (Larder et al., 1987), Ala114Gly (Larder et al., 1989), Tyr115Asn (Larder et al., 1989), Tyr115His (Larder et al., 1989), Tyr115Phe (Larder et al., 1989; Boyer et al., 1992), Tyr115Val (Boyer et al., 1994a), Phe116Tyr (Boyer et al., 1994a), and Phe116Val (Boyer et al., 1994a). Both Ala114Ser and Ala114Gly reduced RT activity (80% and 22% that of wild type, respectively), and both mutant RTs were resistant to inhibition by AZTTP and PFA. Introducing nonhydrophobic amino acids reduced RT activity and were highly resistant to AZTTP and PFA; however, substituting other hydrophobic amino acids for Tyr115 and Phe116 yielded enzymes that behaved like wild type. These data suggest that hydrophobic interactions may be involved in the interaction between Tyr115, Phe116, and the incoming dNTP, as the modeling experiments suggest (Figures 1 and 6).

Conformational Change Preceding Chemical Catalysis. Just as HIV-1 RT undergoes a conformational change upon binding with DNA, it is thought that polymerases undergo a second structural change following dNTP binding. Such a structural change could promote the selection of a complementary nucleotide, enhancing the fidelity of the enzyme

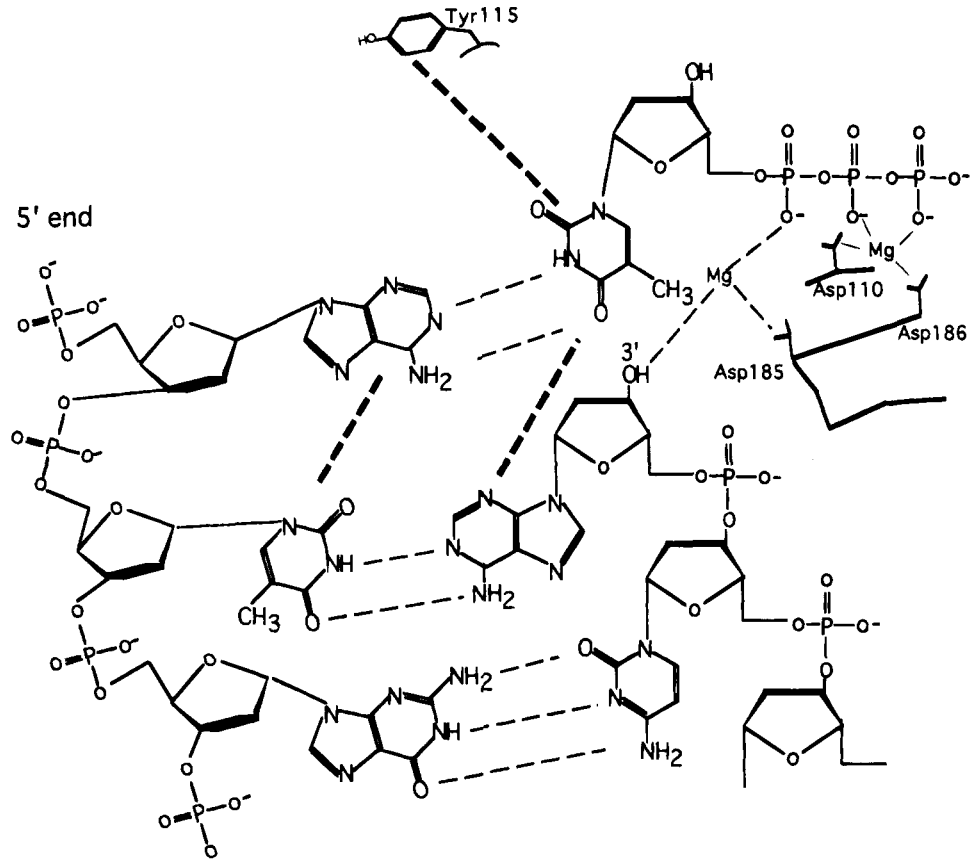


FIGURE 6: Schematic representation of the dNTP-binding and conformational change steps. The 5′ template base (adenosine) within the RT/dsDNA/Fab structure protrudes into the helix axis and stacks on top of both the adenosine and thymidine of the base pair below. The position of the template adenosine and the aspartic acid triad (Asp185, Asp110, and Asp186) is such that the incoming nucleotide (dTTP) can Watson–Crick base pair with the template and chelate (via coordination with divalent cations) to the aspartic acid triad. However, the newly formed base pair bulges 3–4 Å out of the DNA double helix. The next proposed step is for the base pair to slide 3–4 Å toward the helix axis. This sliding mechanism would not only stabilize stacking interactions between the last two base pairs and the incoming nucleotide with Tyr115 but the sliding action will also bring the α-phosphate of the dTTP in close proximity to the primer 3′-OH group.

(Mizrahi et al., 1985; Wong et al., 1991; Johnson, 1993). It has not been possible to dissect the conformational change from the chemistry step using kinetic techniques (Kuchta et al., 1987; Patel et al., 1991; Reardon, 1993). These analyses are difficult because it has not yet been possible to determine unambiguously the magnitude by which phosphorothiolate nucleotide analogs slow phosphodiester bond formation (Benkovic & Schray, 1971; Herschlag et al., 1991). There are two lines of evidence, however, that do suggest that polymerases can undergo a conformational change. Pulse–chase, pulse–quench studies with bacteriophage T7 DNA polymerase (Patel et al., 1991) and HIV-1 RT [Hsieh et al., 1993; but see Reardon 1993] demonstrate that there is a lag between the dNTP-binding and incorporation steps, suggesting that there is time for a conformational change between these steps. A second set of studies, discussed previously, showed that dissociation of HIV-1 RT from a duplex with a chain-terminated primer is significantly slower in the presence of the next correct nucleotide, suggesting that, after binding with a complementary nucleotide, HIV-1 RT changes to a slowly dissociating conformation (Figure 5; Muller et al., 1991; Kati et al., 1992).

Because a high-resolution structure of HIV-1 RT complexed with DNA and a dNTP does not yet exist, it would be premature to try to give a detailed account of how the conformation of RT might change subsequent to dNTP binding. The position of the Hg-UTP in the RT–DNA–Hg-UTP complex (Jacobo-Molina et al., 1993) and the computer-assisted modeling of dTTP into the RT–DNA complex (Figure 1) show that there could be two distinct sites (separated by ~4 Å), where the incoming dNTP binds, which could correspond to bound dNTP before and after a conformational change. In our model, during the initial dNTP-binding step, the base pair within the RT structure (Figure 1) is not stacked directly on top of the base pair below it but instead bulges 3–4 Å out of the double helix. This protrusion would cause the α-phosphate of the incoming dNTP to be positioned too far away to directly interact with the primer 3′-OH group. Before a phosphodiester bond could be formed, the new base pair would need to slide toward the DNA double helix and to stack favorably with the base pair below (Figure 6). This would increase the stacking interactions between the new base pair and both the adjacent base pair and Tyr115 and would as a consequence be energetically favorable. This conformational change would also bring the α-phosphate of the incoming nucleotide in close proximity to the primer 3′-OH, facilitating bond formation. If a noncomplementary nucleotide binds, then this sliding motion would be inhibited. This structural change could occur during the lag between the dNTP-binding and incorporation steps. If the 3′-OH of the primer were replaced with a 3′-H in the model, then, upon completion of the sliding motion, the incoming dNTP could still make extensive contact with the DNA (via Watson–Crick base

pairing) and with the enzyme (via coordination with the Mg^{2+} and stacking interactions with base pairs and Tyr115) without being incorporated into the primer. Such extensive contacts would stabilize the RT−DNA−dNTP complex and could explain why this complex is slow to dissociate.

The predicted sliding motion of the newly formed base pair in the sliding induced-fit model of conformational change would be driven by favorable stacking interactions of the bases. Other elements within HIV-1 RT that would need to alter position to accommodate the sliding nucleotide include the side chains of Tyr115, Asp185, Asp186, and Asp110. However, the conformational change in the protein need not be significant; for example, a simple rotation of these side chains would be sufficient to retain the interactions between the sliding nucleotide and these residues. In addition, there may be significant conformational changes within the duplex portion of the template-primer subsequent to dNTP binding, and these changes could contribute to translocation (see below). Overall, the function of the conformational change we have postulated would be to facilitate the interaction between the 3′-OH group of the primer and an α-phosphate of incoming nucleoside triphosphate; this motion could also restrict the conformations and structures of the incoming dNTP, promoting the efficiency of incorporation of the proper nucleotide.

Phosphodiester Bond Formation. Subsequent to the dNTP binding, the polymerase (if it is in the proper conformation) can catalyze formation of the phosphodiester bond. The structure of HIV-1 RT complexed with a template-primer permits a prediction of the mechanism of phosphodiester bond formation (Figure 7). In this model, Asp110 and Asp186 chelate one of the two Mg^{2+} ions proposed to be involved in the polymerization reaction. One Mg^{2+} ion is thought to be bound to the β- and γ-phosphates of the incoming nucleotide (dTTP) in a bidentate fashion. This Mg^{2+} ion could complex with Asp110 and Asp186. In a two-metal mechanism, Asp185 could chelate the second Mg^{2+}, forming an α-phosphate−Mg^{2+}−Asp185 complex. The interaction of the α-phosphate and the magnesium ion would enhance the electropositive character of the phosphorus atom, facilitating nucleophilic attack by the oxygen atom of the 3′-OH of the primer terminus. The nucleophilic attack by the oxygen atom would generate an intermediate with a pentavalent phosphorus which, following inversion of configuration, would yield, as products, an elongated primer and PP_i. This two-metal phosphoryl-transfer mechanism differs from a mechanism involving three metals proposed for polymerization by Klenow fragment (Steitz, 1993) and is consistent with the two-metal mechanism proposed for DNA pol β (Pelletier et al., 1994). Similar phosphoryl-transfer mechanisms have been postulated for other enzymes, including alkaline phosphatase (Kim & Wyckoff, 1991), and this may be a general mechanism for a class of phosphoryl-transfer reactions [summarized in Steitz and Steitz (1993)]. After catalysis, coordination of PP_i with Asp110 and Asp186 could be broken either by simple diffusion or by competitive binding at this site by the next dNTP.

Translocation. After a polymerase has incorporated a nucleotide, it can either dissociate from the DNA or translocate toward the new 3′-primer terminus. The molecular and biochemical bases of translocation have not yet been determined. In the structure of the HIV-1 RT−DNA complex, α-helices H and I of the thumb subdomain "grip"

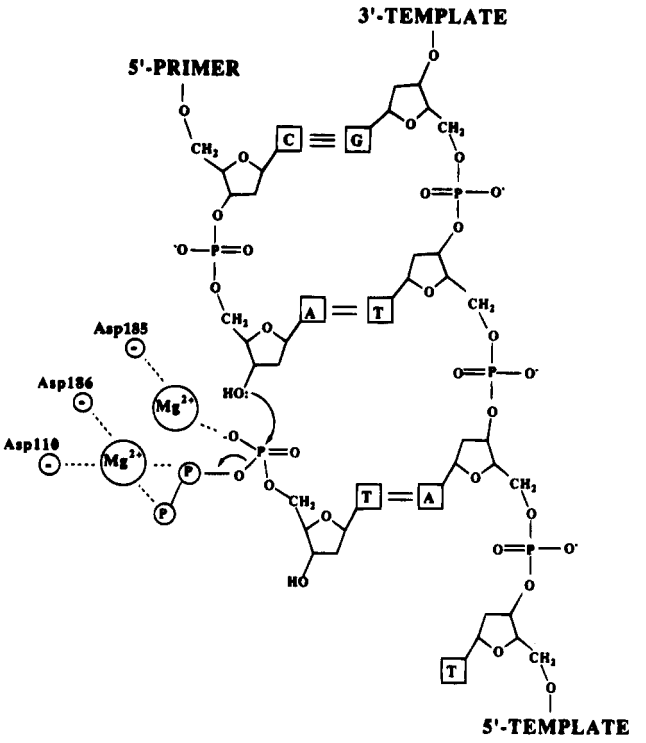


FIGURE 7: Schematic representation of the proposed mechanism of phosphodiester bond formation by HIV-1 RT. The chemistry step is proposed to be metal-mediated, involving two divalent cations separated by ~4 Å (Mg^{2+} in the model can also be substituted by Mn^{2+} or Zn^{2+}). One of the metals is shown coordinating Asp110 and Asp186 with the β- and γ-phosphates of the incoming dNTP, while the other metal chelates with Asp185 and the α-phosphate. The second metal is in position to directly participate in the nucleophilic attack step by contributing to the electropositive character of the α-phosphate. The resulting PP_i may transiently bind to Asp110 and Asp186 via Mg^{2+} coordination prior to release. Arrows indicate the net movement of the electrons.

the DNA (Figure 8). To retain numerous resulting contacts with the DNA, RT would need to interact continuously with the DNA. The majority of contacts with RT involve the sugar−phosphate backbone of the DNA and are therefore not specific, permitting the polymerase to move relatively freely. Thus, the resulting overall binding affinity will remain approximately constant. Since the double helix is a spiral, the motion of the protein relative to the DNA must also be a spiral.

After the translocation step, the precise interactions of the polymerase active site would need to be re-formed at the new primer terminus. We believe that the polymerase moves down the nucleic acid in steps of ~3.4 Å (along the axis of the DNA) and that this movement involves a ratchet-type mechanism. Consistent with a ratchet-type mechanism, nucleotide incorporation usually occurs within milliseconds; however, the dissociation of RT from the nucleic acid after incorporation can take several seconds. The energy for the movement of the protein could come from dNTP hydrolysis; in addition, the transition of the constrained A-form DNA to the relaxed B-form may also supply some of the energy and dictate the overall direction (5′ to 3′ with respect to the primer) of translocation for HIV-1 RT. Nucleotide hydrolysis coupled with A- to B-form DNA transition may supply the energy for translocation for other polymerases such as the rat DNA polymerase β, which also contains a template-primer that assumes an A-form/B-form hybrid geometry

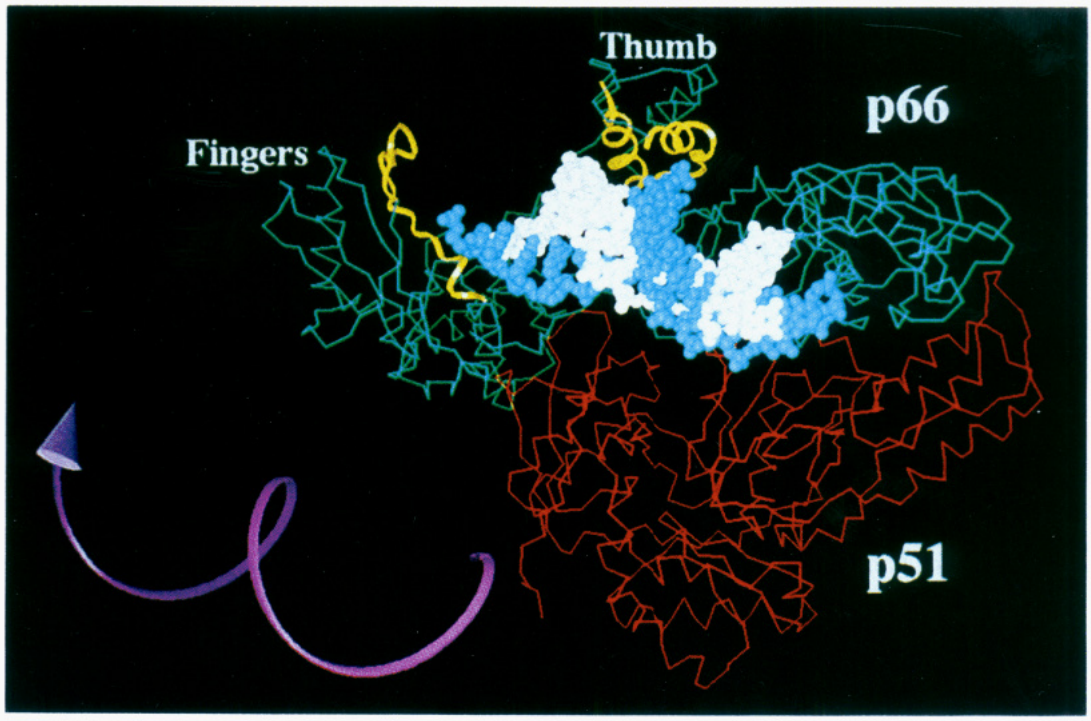


FIGURE 8: Structural representation showing the direction of polymerase translocation. The regions of RT that likely interact with the DNA, including the $\beta3-\beta4$ region of the fingers subdomain, are highlighted in golden ribbons. To retain these interactions, HIV-1 RT would need to translocate horizontally and rotationally around the double helix. The spiraling vector indicates the direction of protein translocation relative to the template-primer. The p66 subunit of HIV-1 RT is colored in green and the p51 subunit in red. The template strand is colored in blue while the primer strand is shown in white.

(Pelletier et al., 1994; Sawaya et al., 1994). It is not clear, however, at what point during polymerization translocation will occur.

Processive Polymerization. It is obvious from the previous discussions that there must be considerable coordination of the activity and structure of HIV-1 RT during the incorporation of a single nucleotide. For productive infection, HIV-1 RT must efficiently catalyze ~20 000 of these reactions. This means that the intricate coordination between the four polymerase subdomains (fingers, palm, thumb, and connection), the DNA, and the incoming dNTPs must be repeated efficiently. Biochemical data show that some polymerases are more processive and can catalyze the addition of many nucleotides before dissociating from the nucleic acid substrate. HIV-1 RT and Klenow fragment are examples of enzymes with moderate to low processivity (Eckert & Kunkel, 1993; Klarmann et al., 1993) while T7 RNA polymerase exhibits high processivity.

Why certain polymerases stall and/or dissociate readily (i.e., what steps in Scheme 1 are affected) at specific nucleic acid sites is unknown. We compared the structure of a processive polymerase, T7 RNA polymerase (Sousa et al., 1993), with that of the relatively distributive Klenow (Freemont et al., 1988) and HIV-1 RT (Jacobo-Molina et al., 1993). A comparison of the size of the thumb and the fingers subdomains of HIV-1 RT, Klenow, and T7 RNA polymerase shows that while HIV-1 RT and Klenow have fingers and thumb subdomains of approximately the same length, the fingers and thumb subdomains of the processive T7 RNA polymerase are nearly twice as long. The large fingers and thumb subdomains allow T7 RNA polymerase to comfortably accommodate a modeled DNA and yet retain a relatively narrow distance between the tip of the thumb and fingers subdomains (20 Å). The distance between the tips of the thumb and fingers subdomains in the HIV-1 RT-DNA complex is 30 Å. This means that while the opening between fingers and thumb subdomains in the HIV-1 RT–DNA complex is wide enough to allow for the dissociation of DNA, in T7 RNA polymerase this opening is smaller than the diameter of the DNA double helix (25 Å). It is conceivable that the large fingers and thumb subdomains allow T7 RNA polymerase to wrap around DNA and prevent its dissociation (Sousa et al., 1993).

In addition, T7 RNA polymerase contains two secondary structural elements that fold across the palm subdomain which could play a role in processivity (Figure 9). A modeled DNA (in an orientation consistent with structural and biochemical data) fits comfortably under the two secondary structural elements (Figure 9B). In addition, nine base pairs separate the polymerase active site from the larger of the two elements, and biochemical data show that there is a transition from the distributive to processive mode after polymerization of RNA $\geq$ 9-mer (Sousa et al., 1993). These data suggest that these secondary structural elements serve as processivity clamps and may be responsible for the efficient nucleic acid extension of T7 transcripts (longer than nine nucleotides).

It is possible that all processive polymerases possess structural elements which enable nucleic acids to remain bound. This may be true for polymerases involved directly in DNA replication and for some repair polymerases, for example, pol β, which reportedly contains a processivity clamp (Pelletier et al., 1994). For other repair polymerases such as *E. coli* DNA polymerase I, processive polymerization may not be necessary. If the polymerases of the organism are not inherently processive, and high processivity is required, other proteins which promote polymerase–nucleic acid binding may be recruited. This is the mechanism by

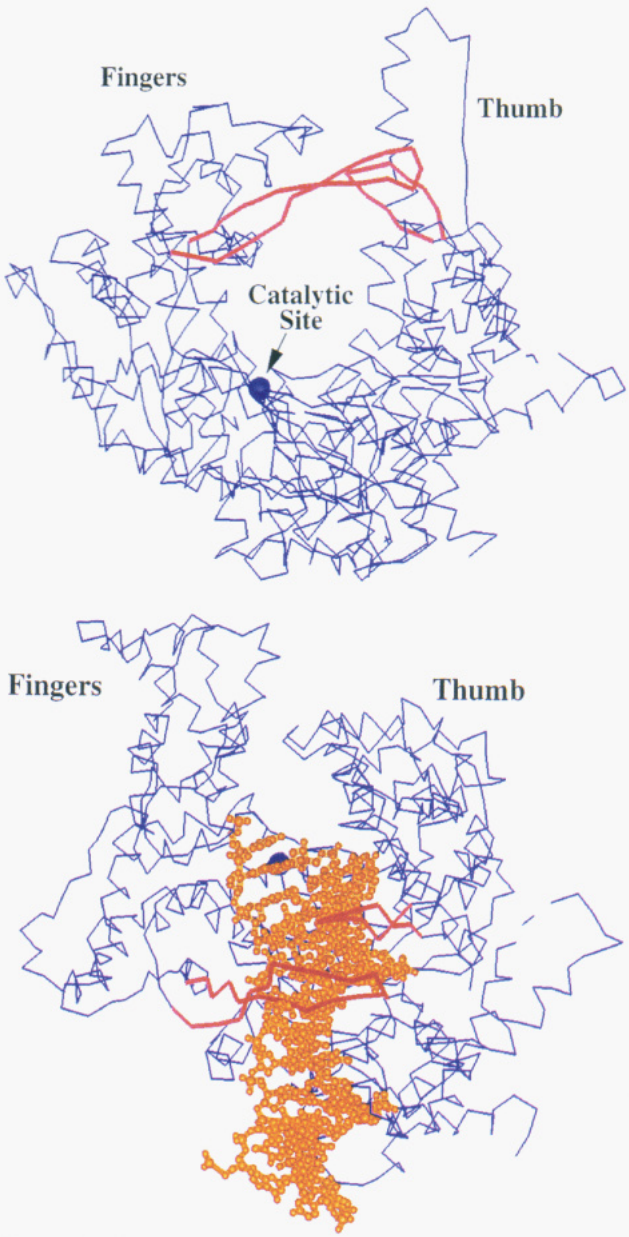


FIGURE 9: The α-carbon backbone of T7 polymerase in (A, top) the absence of and (B, bottom) the presence of a modeled dsDNA template-primer. This polymerase structure has a morphology similar to that of a human right hand. Using this analogy, the T7 polymerase (colored in blue) is oriented such that either (A) the back or (B) the top-back portion of the T7 RNA polymerase structure is in the foreground. The fingers and thumb subdomains in both orientations point toward the top of the page. This structure contains secondary structural elements (colored in red) that fold over the palm subdomain. (B) The modeled DNA (highlighted in gold) fits comfortably underneath these elements; thus, this region of T7 RNA polymerase may serve as a processivity clamp. The polymerase active site (located in the palm subdomain) is depicted by a blue sphere. The α-carbon backbone of T7 polymerase corresponds to coordinates from Sousa et al. (1993), and the DNA was positioned on the basis of the coordinates of the structure of HIV-1 RT complexed with DNA (Jacobo-Molina et al., 1993) after superpositioning the active sites of the two enzymes.

which the accessory factors thioredoxin and proliferating cellular nuclear antigen (PCNA) augment processivity of bacteriophage T7 DNA polymerase (Huber et al., 1987) and eucaryotic polymerase δ (Ng et al., 1993), respectively. Since HIV-1 RT does not possess a processivity clamp, and the virus seemingly needs processive polymerization for efficient replication, HIV-1 RT may also recruit other proteins to increase the efficiency of its RNA- and DNA-templated polymerization activities. These proteins could be of either cellular or viral origin and would likely function by the polymerase in such a way that the binding cleft opening is reduced, interfering with the dissociation of the RT−DNA complex.

ACKNOWLEDGMENT

We thank the members of the Arnold and Hughes laboratories for their contributions to the structural studies of HIV-1 reverse transcriptase. We also thank Gail Ferstandig Arnold, Paul Boyer, Wade Huber, Terri Larsen, Stuart Le Grice, and Brad Preston for helpful discussions and assistance; Birgit Roy for technical assistance; Jim Champoux and Larry Loeb for critical reading of the manuscript; and Rui Souza and B. C. Wang for providing Cα coordinates for T7 RNA polymerase. We also thank the staff at the Cornell High Energy Synchrotron Source for continuing support of these studies.

APPENDIX

Orientation of the DNA. The structure of rat DNA polymerase β (pol β) has been described in the presence (Pelletier et al., 1994) and absence (Sawaya et al., 1994) of template-primer. If one assumes that the polymerase active sites are structurally and topologically conserved, the orientation of the complexed DNA within pol β differs by 180° relative to the orientation of the DNA in a complex of HIV-1 RT with DNA (Jacobo-Molina et al., 1993) and complexes of Klenow fragment with DNA (Freemont et al., 1988; Beese et al., 1993). Since the conservation of the active site architecture suggests that these polymerases are evolutionarily related, this presents an apparent paradox. It was suggested that since pol β appears to be in a conformation relevant for polymerization, the DNA in the other polymerase structures is bound in a mode that would not be physiologically relevant (Pelletier et al., 1994; Sawaya et al., 1994). Since it is clear from biochemical experiments that the RNase H domain cleaves RNA ~18 nucleotides downstream from the site of polymerization (Gopalakrishnan et al., 1992; Kati et al., 1992), in agreement with the orientation of nucleic acids and the distance between the polymerase and RNase H active sites reported by Jacobo-Molina et al., Pelletier et al. have suggested, on the basis of the data obtained with RNase H, that the orientation of the enzyme is opposite when it uses RNA versus DNA templates.

However, there is a wealth of data to show that HIV-1 RT binds both RNA/DNA and DNA/DNA substrates in the same orientation and makes it quite clear that the orientation is the same as seen in the X-ray crystal structure. For example, DNase I and S1 nuclease footprints show that the regions of dsDNA protected by HIV-1 RT are in agreement with the duplex DNA being oriented as it is in the HIV-1 RT/dsDNA/Fab structure (Wöhrl et al., 1995). Hydroxyl radical footprints (Metzger et al., 1993) are also in agreement with the orientation reported by Jacobo-Molina et al. The two modes of binding suggested by Pelletier et al., if correct, would predict marked biochemical differences between RNA- and DNA-templated polymerizations. However, numerous biochemical measurements of the properties of the polymerase with RNA and DNA template show that nucleic acid binding affinity (Kati et al., 1992; Yu & Goodman,

1992), processivity (Klarmann et al., 1993), and fidelity (Preston et al., 1988; Roberts et al., 1988; Ji & Loeb, 1992; Yu & Goodman, 1992) are essentially unchanged whether RNA/DNA or DNA/DNA serve as substrates for HIV-1 RT. In addition, the vast majority of hundreds of site-directed mutations in HIV-1 RT had similar enzymatic activity whether DNA or RNA served as the template (Boyer et al., 1994a; Boyer et al., 1994b; unpublished data of the authors). Together, these data demonstrate that HIV-1 RT binds both RNA/DNA and DNA/DNA within the same cleft in the orientation as described (Jacobo-Molina et al., 1993).

What then is the biologically relevant orientation of the DNA in pol β? Unlike HIV-1 RT, pol β is not a replicative polymerase; rather it functions to bridge short gaps within DNA, and the only other known polymerase that is similar in sequence to pol β is terminal deoxynucleotidyl transferase (Delarue et al., 1990). Since pol β appears to be a unique polymerase, it is possible that it has evolved (through either convergent or divergent evolution) a unique mode of binding nucleic acids to fulfill its role in the repair of damaged DNA. It is, therefore, also possible that the orientation of the nucleic acid in the pol β/DNA structure, although unusual, is biologically relevant. However, pol β was not cocrystallized with a physiologically relevant gapped intermediate, and the DNA in the structure is not directly situated within the cleft; rather the DNA sits at an angle and runs into portions of the enzyme (Pelletier et al., 1994). The biological relevance of this conformation was questioned by the authors of the paper describing the pol β structure. Under these circumstances, it is worth considering the possibility that, under physiological conditions of DNA repair, the DNA is bound to pol β in an opposite orientation to that described in the structure of Pelletier et al.

REFERENCES

Abbotts, J., Bebenek, K., Kunkel, T. A., & Wilson, S. H. (1993) *J. Biol. Chem. 268,* 10312−10323.

Arnold, E., Jacobo-Molina, A., Nanni, R. G., Williams, R. L., Lu, X., Ding, J., Clark, A. D., Jr., Zhang, A., Ferris, A. L., Clark, P., Hizi, A., & Hughes, S. H. (1992) *Nature 357,* 85−89.

Beese, L. S., Derbyshire, V., & Steitz, T. A. (1993) *Science 260,* 352−355.

Benkovic, S. J., & Schray, K. J. (1971) in *The Enzymes: Group Transfer, Part A* (Boyer, P. D., Ed.) pp 201−238, Academic Press, New York.

Boosalis, M. S., Petruska, J., & Goodman, M. F. (1987) *J. Biol. Chem. 262,* 14689−14696.

Boyer, P. L., Ferris, A. L., & Hughes, S. H. (1992) *J. Virol. 66,* 1031−1039.

Boyer, P. L., Ferris, A. L., Clark, P., Whitmer, J., Frank, P., Tantillo, C., Arnold, E., & Hughes, S. H. (1994a) *J. Mol. Biol. 243,* 472−483.

Boyer, P. L., Tantillo, C., Jacobo-Molina, A., Nanni, R. G., Ding, J., Arnold, E., & Hughes, S. H. (1994b) *Proc. Natl. Acad. Sci. U.S.A. 91,* 4882−4886.

Carroll, S. S., Cowart, M., & Benkovic, S. J. (1991) *Biochemistry 30,* 804−813.

Clark, A. D., Jr., Jacobo-Molina, A., Clark, P., Hughes, S. H., & Arnold, E. (1995) Crystallization of HIV-1 reverse transcriptase in the presence and absence of nucleic acid substrates, inhibitors, and an antibody Fab fragment, in *Methods in Enzymology,* Academic Press, New York (in press).

Coffin, J. M. (1990) Retroviridae and Their Replication, in *Virology* (Fields, B. N., Knipe, D. M., Chanock, R. M., et al., Eds.) pp 1437−1500, Raven, New York.

Delarue, M., Poch, O., Tordo, N., Moras, D., & Argos, P. (1990) *Protein Eng. 6,* 461−467.

Ding, J., Jacobo-Molina, A., Tantillo, C., Lu, X., Nanni, R. G., & Arnold, E. (1994) *J. Mol. Recogn. 7,* 157−161.

Ding, J., Das, K., Tantillo, C., Zhang, W., Clark, A. D., Jr., Jessen, S., Lu, X., Hsiou, Y., Jacobo-Molina, A., Andries, K., Pauwels, R., Moereels, H., Koymans, L., Janssen, P. A. J., Smith, R. H., Jr., Koepke, M. J., Michejda, C. J., Hughes, S. H., & Arnold, E. (1995) *Structure 3,* 365−379.

Divita, G., Muller, B., Immendorfer, U., Gautel, M., Rittinger, K., Restle, T., & Goody, R. S. (1993) *Biochemistry 32,* 7966−7971.

Dong, Q., Copeland, W. C., & Wang, T. S.-F. (1993) *J. Biol. Chem. 268,* 24163−24174.

Eckert, K. A., & Kunkel, T. A. (1993) *J. Biol. Chem. 268,* 13462−13471.

Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., & Steitz, T. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85,* 8924−8928.

Gopalakrishnan, V., & Benkovic, S. (1994) *J Biol. Chem. 269,* 4110−4115.

Gopalakrishnan, V., Peliska, J. A., & Benkovic, S. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89,* 10763−10767.

Herschlag, D., Piccirilli, J. A., & Cech, T. R. (1991) *Biochemistry 30,* 4844−4854.

Hottinger, M., Podust, V. N., Thimmig, R. L., McHenry, C., & Hubscher, U. (1994) *J. Biol. Chem. 269,* 986−991.

Hsieh, J., Zinnen, S., & Modrich, P. (1993) *J. Biol. Chem. 268,* 24607−24613.

Huber, H., Tabor, S., & Richardson, C. C. (1987) *J. Biol. Chem. 262,* 16224−16232.

Huber, H. E., McCoy, J. M., Seehra, J. S., & Richardson, C. C. (1989) *J. Biol. Chem. 264,* 4669−4678.

Jacobo-Molina, A., & Arnold, E. (1991) *Biochemistry 30,* 6351−6361.

Jacobo-Molina, A., Clark, A. D., Jr., Williams, R. L., Nanni, R. G., Clark, P., Ferris, A. L., Hughes, S. H., & Arnold, E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88,* 10895−10899.

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., & Arnold, E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 6320−6324.

Ji, J., & Loeb, L. A. (1992) *Biochemistry 31,* 954−958.

Johnson, K. A. (1993) *Annu. Rev. Biochem. 62,* 685−713.

Kati, W. M., Johnson, K. A., Jerva, L. F., & Anderson, K. S. (1992) *J. Biol. Chem. 267,* 25988−25997.

Kim, E. E., & Wyckoff, H. W. (1991) *J. Mol. Biol. 218,* 449−464.

Klarmann, G. J., Schauber, C. A., & Preston, B. D. (1993) *J. Biol. Chem. 268,* 9793−9802.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1992) *Science 256,* 1783−1790.

Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987) *Biochemistry 26,* 8410−8417.

Larder, B. A., Purifoy, D. J. M., Powell, K. L., & Darby, G. (1987) *Nature 327,* 716−717.

Larder, B., A., Kemp, S. D., & Purifoy, D. J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86,* 4803−4807.

Lowe, D. M., Parmar, V., Kemp, S. D., & Larder, B. A. (1991) *FEBS Lett. 282,* 231−234.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193,* 265.

Luo, G., & Taylor, J. (1990) *J. Virol. 64,* 4321−4328.

Majumdar, C., Abbotts, J., Broder, S., & Wilson, S. H. (1988) *J. Biol. Chem. 263,* 15657−15665.

Metzger, W., Hermann, T., Schatz, O., Le Grice, S. F. J., & Heumann, H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 5909−5913.

Mizrahi, V., Henrie, R. N., Marlier, J. F., Johnson, K. A., & Benkovic, S. J. (1985) *Biochemistry 24,* 4010−4018.

Muller, B., Restle, T., Reinstein, J., & Goody, R. S. (1991) *Biochemistry 30,* 3709−3715.

Ng, L., McConnell, M., Tan, C.-K., Downey, K. M., & Fisher, P. A. (1993) *J. Biol. Chem. 268,* 13571−13576.

Patel, P. H., & Preston, B. D. (1994) *Proc. Natl. Acad. Sci. U.S.A 91,* 549−553.

Patel, S. S., Wong, I., & Johnson, K. A. (1991) *Biochemistry 30,* 511−525.
Peliska, J. A., & Benkovic, S. J. (1992) *Science 258,* 1112−1118.
Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., & Kraut, J. (1994) *Science 264,* 1891−1903.
Poch, O., Sauvaget, I., Delarue, M., & Tordo, N. (1989) *EMBO J. 8,* 3867−3874.
Prasad, V. R., Lowy, I., Santos, T. D. L., Chiang, L., & Goff, S. P. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88,* 11363−11367.
Preston, B. D., & Garvey, N. (1992) *Pharm. Technol. 16,* 34−51.
Preston, B. D., Poiesz, B. J., & Loeb, L. A. (1988) *Science 242,* 1168−1171.
Raag, R., Nanni, R. G., Clark, A. D., Jr., Ding, J., Jacobo-Molina A., Lu, X., Tantillo, C., Hughes, S. H., & Arnold, E. (1994) *American Crystallographic Association Annual Meeting Abstracts,* p 44 Abstract B03, Atlanta, GA.
Reardon, J. E. (1992) *Biochemistry 31,* 4473−4479.
Reardon, J. E. (1993) *J. Biol. Chem. 268,* 8743−8751.
Roberts, J. D., Bebenek, K., & Kunkel, T. A. (1988) *Science 241,* 1171−1173.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboraty Press, Plainview, N.Y.
Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S., & Kraut, J. (1994) *Science 264,* 1930−1935.
Schinazi, R. F. (1993) *Perspect. in Drug Discovery Des. 1,* 151−180.
Schinazi, R., Larder, B., & Mellors, J. (1994) *Int. Antiviral News 2,* 72−75.
Smerdon, S. J., Jager, J., Wang, J., Kohlstaedt, L. A., Chirino, A. J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91,* 3911−3915.
Sousa, R., Chung, Y. J., Rose, J. P., & Wang, B.-C. (1993) *Nature 364,* 593−599.
St. Clair, M. H., Martin, J. L., Tudor-Williams, G., Bach, M. C., Vavro, C. L., King, D. M., Kellam, P., Kemp, S. D., & Larder, B. A. (1991) *Science 253,* 1557−1559.
Steitz, T. A. (1993) *Curr. Opin. Struct. Biol. 3,* 31−38.
Steitz, T. A., & Steitz, J. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 6498−6502.
Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A. J., & Arnold, E. (1994) *J. Mol. Biol. 243,* 369−387.
Tisdale, M., Kemp, S. D., Parry, N. R., & Larder, B. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 5653−5656.
Wöhrl, B. M., Tantillo, C., Arnold, E., & Le Grice, S. F. J. (1995) *Biochemistry 34,* 5343−5350.
Wong, I., Patel, S. S., & Johnson, K. A. (1991) *Biochemistry 30,* 526−537.
Yu, H., & Goodman, M. F. (1992) *J. Biol. Chem. 267,* 10888−10896.

BI9422900