THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 271, No. 9, Issue of March 1, pp. 4872–4878, 1996

Printed in U.S.A.

Human Immunodeficiency Virus Reverse Transcriptase

FUNCTIONAL MUTANTS OBTAINED BY RANDOM MUTAGENESIS COUPLED WITH GENETIC SELECTION IN *ESCHERICHIA COLI**

(Received for publication, November 10, 1995, and in revised form, December 20, 1995)

Baek Kim, Tanya R. Hathaway, and Lawrence A. Loeb‡

From The Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology, Box 357705, University of Washington, Seattle, Washington 98195-7705

We describe catalytically active mutants of HIV RT (human immunodeficiency virus reverse transcriptase) generated by random sequence mutagenesis and selected in *Escherichia coli* for ability to complement the temperature-sensitive phenotype of a DNA polymerase I (Pol I^{ts}) mutant. We targeted amino acids Asp-67 through Arg-78 in HIV RT, which form part of the β3-β4 flexible loop and harbor many of the currently known mutations that confer resistance to nucleoside analogs. DNA sequencing of 109 selected mutants that complement the Pol I^{ts} phenotype revealed substitutions at all 12 residues targeted, indicating that none of the wild-type amino acids is essential. However, single mutations were not observed at Trp-71, Arg-72, and Arg-78, consistent with evolutionary conservation of these residues among viral RTs and lack of variation at these positions among isolates from patients. The mutations we recovered included most of those associated with drug resistance as well as previously unidentified mutations. Purification and assay of 14 mutant proteins revealed correlation between their DNA-dependent DNA polymerize activity *in vitro* and ability to complement the Pol I^{ts} phenotype. Activity of several mutants was resistant to 3′-azidothymidine triphosphate. We conclude that random sequence mutagenesis coupled with positive genetic selection in *E. coli* yields large numbers of functional HIV RT mutants. Among these are less active variants which are unlikely to be isolated from HIV-infected individuals and which will be informative of the roles of individual amino acids in the catalytic functions of the enzyme.

Infection by human immunodeficiency virus 1 (HIV-1)[1] is the pathogenic precursor to clinical development of acquired immunodeficiency syndrome (AIDS). HIV-1 contains a reverse transcriptase (HIV RT) that catalyzes synthesis of both a single- and double-stranded DNA copy of the viral genome, and is required for viral replication. As such, HIV RT provides a central target for chemotherapy of AIDS. Current chemotherapeutic agents that target HIV RT, such as AZT, ddC, and ddI, are incorporated into viral DNA but cannot be further elongated, thus terminating DNA synthesis (Mitsuya and Broder, 1986). Unfortunately, the exceptionally high mutation rate of the virus drives rapid emergence of resistant strains, rendering these nucleoside analogs ineffective (Boucher *et al.*, 1992; Larder and Kemp, 1989; Mitsuya and Broder, 1986). Single and multiple base substitutions within the HIV RT gene are responsible for emergence of these drug-resistant mutants that escape chemotherapy. In order to develop drugs that more effectively target HIV RT, we require a clearer understanding of the potential of different sites within the enzyme to produce drug-resistant mutations.

Many of the mutations that render HIV resistant to nucleoside analogs are located in the β3-β4 loop, β9-β10 turn, and β11a/b regions of HIV RT (Tantillo *et al.*, 1994). The β3-β4 region in the fingers domain is believed to function as a template grip that interacts with single-stranded (ss) DNA and RNA templates (Boyer *et al.*, 1994; Tantillo *et al.*, 1994). Mutations within the β3-β4 loop that confer drug resistance include K65R (resistant to ddC, ddI, and 3TC (Gu *et al.*, 1994)), D67N (resistant to AZT (Larder and Kemp, 1989), Larder and Kemp, 1989)), T69D (resistant to ddC (Fitzgibbon *et al.*, 1992)), K70R (resistant to AZT (Larder and Kemp, 1989)), and L74V (resistant to ddI (St. Clair *et al.*, 1991)). Recently, two additional mutations, V75I and F77L, have been isolated from patients treated with dideoxynucleosides (Shirasaka *et al.*, 1995). The diverse locations of mutations in HIV RT that render the virus resistant to the same drug has prompted the suggestion that many of these mutations act indirectly, at a distance. It is postulated that such mutations affect conformation near the catalytic site, modifying interactions between the fingers domain and the template DNA or RNA (Boyer *et al.*, 1994; Tantillo *et al.*, 1994) to enhance discrimination between natural substrates and inhibitors. Other drug-resistant mutations, not in β3-β4 loop region, such as M184V (Gu *et al.*, 1992), that are located close to the catalytic site are proposed to directly affect dNTP binding (Tantillo *et al.*, 1994).

We recently demonstrated that HIV RT can substitute for DNA polymerase I in *Escherichia coli* (Kim and Loeb, 1995a). In our genetic complementation system, expression of HIV RT allows mutant *E. coli* cells harboring a temperature-sensitive DNA polymerase I (Pol I^{ts}) to grow at non-permissive temperature. Pol I^{ts} cells expressing an inactive mutant (D186N) fail to grow at nonpermissive temperature, demonstrating the requirement for catalytically functional HIV RT. Based on the observation that complementation by HIV RT is specifically inhibited by nucleoside analogs such as AZT, we proposed that the system can be used to screen potential anti-HIV RT drugs in bacteria (Kim and Loeb, 1995b).

In the present study, we used the bacterial complementation

* This work was supported by National Institutes of Health Grant R35-CA-39903 (to L. A. L.) and by National Institutes of Health Molecular Training Grant 5T32-CA-09437-13 (to B. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 206-543-0557; Fax: 206-543-3967.

[1] The abbreviations used are: HIV RT, human immunodeficiency virus reverse transcriptase; AZT, 3′-azidothymidine; AZTTP, 3′-azidothymidine triphosphate; ddC, 2′,3′-dideoxycytidine; ddI, 2′,3′-dideoxyinosine; dNTP, deoxynucleoside triphosphate; IPTG, isopropyl-1-thio-β-D-galactopyranoside; IC_{50}, concentration yielding 50% inhibition; Pol I^{ts}, temperature-sensitive polymerase I.

system to select active HIV RT mutants from libraries that encode random sequence substitutions in the wild type gene. We replaced the 36 contiguous nucleotides near the β3-β4 loop that encode amino acids 67 through 78 with random sequences and selected from this library mutants that complement the Pol I^{ts} mutation. This positive genetic selection allowed us to recover most of the active HIV RT mutations thus far observed in naturally occurring variants and drug-resistant mutants. Our large library of active mutants includes previously unidentified mutations and mutations in the β3-β4 loop that confer resistance to AZTTP.

EXPERIMENTAL PROCEDURES

Strains—E. coli NM522 (Stratagene) was used for cloning and library construction. *E. coli* BL21(DE3) from Novagen was used for expression and purification of HIV RT. 2 × YT (Sambrook *et al*., 1989) was used for growth of these strains. The Pol I^{ts} strain used for genetic complementation is the *E. coli* B/r derivative SC18-12 (*rec*A718 *pol*A12 *uvr*155 *trp*E65 *lon*-11 *sul*A1) (Witkin and Roegner, 1992). Nutrient broth containing Difco nutrient broth (8 g/liter), NaCl (4 g/liter), and tetracycline (12.5 μg/ml) was used for culturing Pol I^{ts} cells. Nutrient agar plates containing Difco nutrient agar (23 g/liter), NaCl (5 g/liter), and tetracycline (12.5 μg/ml) were used for assessing genetic complementation.

Plasmids—pHIVRT is a derivative of pHSG576 that contains the wild-type HIV RT under control of *lac*P/O (Kim and Loeb, 1995a). pBK8 is a derivative of pBS-SK(+) (Stratagene) containing the HIV RT gene generated from digestion of pHIVRT with *Hin*dIII and *Eco*RI. A *Sac*I site at position 79 of the HIV RT gene in pBK8 was created by site-directed mutagenesis (Kunkel, 1985) to generate pBK13. To avoid mutations that might be created in other regions of the HIV RT gene during site-directed mutagenesis, the 310-bp segment of HIV RT in pBK13 flanked by *Bsr*GI and *Eco*RV restriction sites was replaced with the corresponding segment of pBK8 containing the *Sac*I site to generate pBK14. The 1.3-kilobase wild-type sequence between the *Bsr*GI and *Eco*RI sites of the HIV RT gene in pHIVRT was replaced with the corresponding 1.3-kilobase fragment generated by digestion of pBK14 with *Bsr*GI and *Eco*RI, generating pBK16. Thus, pBK16 is the same as pHIVRT, but contains a S*ac*I site at position 79 of the HIV RT gene. To avoid contamination with wild-type pHIVRT during genetic selection, a "stuffer" plasmid was constructed in which the 108-bp wild-type sequence of pBK16 between *Bsr*GI and *Sac*I was replaced with a 3,515-bp DNA fragment by digestion of phage λ DNA with *Bsr*GI and *Sac*I, followed by ligation, to generate pBK18 (*step 5*, Fig. 1). Chloramphenicol (30 μg/ml) was used for selection of the pHSG576 derivatives pHIVRT, pBK16, and pBK18 (Takeshita *et al*., 1987). Carbenicillin (50 μg/ml) was used for maintenance of the pBR322 derivatives pBK8, pBK13, and pBK14.

Construction of HIV RT Random Library—The HIV RT random library was constructed by annealing two single-stranded DNA oligomers. Oligomer 1 (Fig. 1, *steps 1* and *2*) was the 83-mer (5′-GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGC-3′) corresponding to the plus strand of the HIV RT sequence from amino acid residues 35 to 62. Oligomer 2 was the 79-mer (5′-CTTATTGAGCTCTCTGAAATCTACTAATTTTCTCCATTTAGTACTGTCTTTTTTCTTTATGGCAAATACTGGAGTATTG-3′) corresponding to the minus strand of the HIV RT sequence from amino acids 56 to 82. Oligomer 2 contained mutated sequences corresponding to amino acids 67 to 78 (underlined), with 12% random nucleotides and 88% wild-type nucleotide at each of the 36 randomized positions, and thus encoded an average of 4.3 nucleotide substitutions per primer. The two primers (1 μg each) were mixed in 50 μl of 200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl (5 × sequencing buffer (U. S. Biochemical Corp)) and annealed by incubating at 80 °C for 5 min, at 55 °C for 15 min, at 37 °C for 15 min, and at room temperature for 15 min (*step 1*, Fig. 1). The partially double-stranded oligonucleotide was extended by incubation with Sequenase Version 2.0 T7 DNA polymerase (U. S. Biochemical Corp., 0.6 unit/ml) and dNTPs (1 mM) for 2 h at 37 °C in a total volume of 54 μl (*step 2*, Fig. 1). The double-stranded oligonucleotides (2 μl) were then amplified in a polymerase chain reaction using Vent DNA polymerase (New England Biolabs) and 5′ and 3′ end primers in 100 μl of 20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100 at an annealing temperature of 28 °C for 30 cycles (*step 3*, Fig. 1). The 5′ primer was a 14-mer identical to the 5′ end of oligomer 1, and the 3′ primer was a 14-mer identical to the 5′ end of oligomer 2. The amplified DNA (5 μg) was digested with *Bsr*GI and *Sac*I (New England Biolabs) (*step 4*, Fig. 1), purified, and concentrated by centrifugation with a Microcone 30 filter (Amicon), and phenol extracted. The purified oligonucleotides were then used as the random nucleotide-containing inserts for construction of the HIV RT plasmid library. The stuffer plasmid pBK18 (10 μg) was also digested with *Bsr*GI and *Sac*I. The resulting 5.0-kilobase fragment was isolated from a 0.8% agarose gel by using an Ultrafree-MC 0.45-mm filter (Millipore) and 5 μg was ligated to the 108-bp insert by using T4 DNA ligase (Boehringer Mannheim) (*step 5*, Fig. 1). The ligation mixture was directly transformed into electrocompetent NM522 cells (Stratagene) (*step 6*, Fig. 1). A plasmid library was prepared from an overnight culture of transformed cells grown in 2 × YT containing 30 μg/ml chloramphenicol (*step 7*, Fig. 1).

Genetic Selection in E. coli—Electroporation of a 50-μl mixture of 8×10^9 Pol I^{ts} cells and 2 μg of plasmid library was carried out by using a Gene Pulsar apparatus (Bio-Rad; 25 microfarads, 2.1 V, 400 ohms). Transformed cells were plated onto nutrient agar selection plates containing chloramphenicol, tetracycline, and IPTG (1 mM) (about 200 cells per plate; *step 8*, Fig. 1). Transformation efficiency, measured after incubation at permissive temperature (30 °C) for 24 h, was typically 2×10^7 cells per 1 μg of DNA. For selection of active HIV RT mutants, plates were incubated at nonpermissive temperature (37 °C) for 18 h. Colonies obtained at 37 °C were regrown individually in 2 ml of nutrient broth containing chloramphenicol and tetracycline. To verify the phenotype, plasmids were purified from each of these cultures and retransformed into Pol I^{ts} cells by electroporation. About 200 re-transformed cells were plated onto each of two nutrient agar selection plates, followed by incubation of one plate at 30 °C and the other at 37 °C. After incubation for 18 h at 37 °C, the extent of complementation by each of the selected mutants was scored visually with respect to colony size (+ through +++, with +++ being equivalent to wild-type) and plating efficiency (number of colonies at 37 °C/number of colonies at 30 °C).

DNA Sequencing—Plasmid DNA, purified from Pol I^{ts} cells, was transformed into *E. coli* NM522. Approximately 0.4 μg of plasmid DNA was prepared from NM522 cells. The nucleotide sequence encoding amino acids 37 to 86 of HIV RT and including the randomized segment was established by polymerase chain reaction-based sequencing (fmol sequencing kit, Promega). The sequencing primer (5′-TTAAACAATGGCCATTGACAG) was a 22-mer encoding amino acid residues 21 to 28 of HIV RT.

Purification of Wild-type HIV RT—To construct a plasmid expressing wild-type HIV RT, a 1.6-kilobase DNA fragment of pBK14 was digested with *Nde*I and *Sal*I and cloned into pET28a (Novagene) to generate pBK34. pBK34 expresses wild-type HIV RT fused at its N terminus to 20 amino acids that includes a hexahistidine sequence. The HIV RT fusion protein was purified in a one-step metal chelation chromatographic procedure by using Ni^{2+} affinity resin and buffers (His Bind Resin and Buffer Kit, Novagen) according to a protocol modified from that of the supplier. Expression of HIV RT was induced by addition of 0.1 mM IPTG to 1 liter of log phase *E. coli* BL21 (DE3) grown in 2 × YT to an OD_{600} of 0.5. After incubation for 2 h, cells were harvested by centrifugation and the pellets were resuspended and frozen (−70 °C) with 30 ml of 1 × binding buffer and lysozyme (200 μg/ml, Sigma). Frozen cells were thawed and lysed on ice for approximately 2 h. The lysed cells were centrifuged (27,000 × *g*) and the supernatant solution was applied to a charged 10-ml His Bind column (1 × 10 cm). The resin was prepared by successive washes with deionized water (30 ml), 1 × charge buffer (30 ml), and 1 × binding buffer (45 ml). All chromatographic steps were carried out at 4 °C at a flow rate of 20 ml/h. Following application of the crude supernatant solution, the column was washed with 1 × binding buffer (30 ml) and 1 × wash buffer (10 ml). HIV RT was eluted with 1 × elute buffer (30 ml); 90% of the recovered HIV RT was released from the resin in the first 8 ml. Fractions containing purified HIV RT were analyzed by electrophoresis in a 12% SDS-polyacrylamide gel. Fractions containing HIV RT were combined (5 ml) and dialyzed against 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 200 mM NaCl, 10% glycerol for 16 h and then against the same buffer containing 1 mM dithiothreitol for an additional 16 h. The purity of HIV RT was greater than 95%, estimated by visual inspection of Coomassie Blue-stained gels. The concentration of purified HIV RT was determined by the dye-binding procedure of Bradford (Bradford, 1976) with the dye reagent and bovine serum albumin standard from Bio-Rad.

Purification of Mutant HIV RTs—The 108-bp fragments containing mutations in HIV RT were inserted into pBK34 between the *Bsr*GI and *Sac*I sites as follows. The DNA fragments between the *Bsr*GI and *Sac*I sites were polymerse chain reaction-amplified with the 5′ and 3′ end primers used in Fig. 1 (*step 3*). The amplified products were digested with *Bsr*GI and *Sac*I and inserted into pBK34 digested with the same

enzymes. Expression and purification of the mutant proteins were essentially as described for wild-type HIV RT. Mutant HIV RTs were purified from 200 ml of 2 × YT cultures on a 1-ml column; the volume of all buffers was one-fifth that specified for chromatography of wild-type HIV RT. Protein concentrations were estimated by visually comparing intensities of the 66- and 51-kDa bands with those of known amounts of wild-type HIV RT in Coomassie Blue-stained 12% SDS-polyacrylamide gels.

DNA Polymerase Activity—DNA-dependent DNA polymerase activity was assayed with a gapped DNA template. Purified HIV RT protein (0.1 μg) was incubated in a 50-μl reaction mixture containing activated calf thymus DNA (80 μg/ml), 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 8 mM $MgCl_2$, 5 mM dithiothreitol, 100 μg/ml bovine serum albumin, 5 μM each dATP, dCTP, dGTP, and dTTP, and 50 nM [*methyl*-^{3}H]dTTP (40 Ci/mmol, Moravek Biochemicals, Inc.). Incubation was at 37 °C for 20 min and incorporation was determined as described previously (Battula and Loeb, 1974). Incorporation of label from [*methyl*-^{3}H]dTTP was a linear function of amount of enzyme and incubation time.

RESULTS

Selection for Functional Mutants by Genetic Complementation—To construct an HIV RT plasmid library, we substituted a segment of the wild-type HIV RT gene with an oligomer containing 12% random nucleotides and 88% wild-type nucleotide at each of the 36 positions encoding amino acids 67–78 (Fig. 1, *steps 1–4*). The partially random inserts were then ligated into pBK18, generating approximately 5×10^5 plasmid-borne HIV RT mutants (Fig. 1, *step 5*). To identify active mutants, we employed a positive genetic selection system in which HIV RT complements the temperature-sensitive phenoytpe of an *E. coli* DNA polymerase I mutant (Kim and Loeb, 1995a). After transformation of the plasmid library into *E. coli* Pol I^{ts} cells (Fig. 1, *step 8*) a total of 1.2×10^4 cells were plated onto nutrient agar. Of these plated cells, 1400 formed colonies at 37 °C, a nonpermissive temperature for the bacterial host.

To confirm the temperature-independent phenotype of transformed cells that formed colonies at 37 °C, plasmids were prepared from 205 such colonies and retransformed individually into Pol I^{ts} cells. To assess the ability of each selected HIV RT mutant to substitute for *E. coli* DNA polymerase I, the plating efficiency and colony size were assayed following an 18-h incubation at 37 °C. Of the re-transformed plasmids, 176 (86%) showed greater than 75% plating efficiency (number of colonies at 37 °C relative to the number at 30 °C). Approximately 3% of transformants formed no colonies at 37 °C; this false positive value presumably reflects the background growth of Pol I^{ts} cells observed on high cell density plates (Kim and Loeb, 1995a). In separate experiments, Pol I^{ts} cells transformed with pBK18 (the stuffer plasmid without HIV RT) showed plating efficiency of 1.5% at 37 °C. Of the 176 retransformed cells that showed plating efficiency greater than 75%, 57 (32%) formed large colonies at 37 °C equivalent in size to that of Pol I^{ts} cells expressing wild-type HIV RT (+++), 81 (46%) formed colonies of medium size (++), and 38 (22%) produced small colonies.

Sequence Alterations in Mutant HIV RTs—The nucleotide and amino acid substitutions in 109 mutants selected at 37 °C, and in 18 non-selected mutants, were ascertained by DNA sequencing (Table I). Given an oligomer 36 nucleotides in length containing 12% random substitutions at each position, there should be an average of 4.3 nucleotide substitutions in the non-selected inserts; we detected an average of 5.0 substitutions per insert, suggesting that the oligomer did in fact contain approximately 12% random substitutions. Among selected mutants, the average number of nucleotide substitutions was 3.3 (Table I), suggesting that not all substitutions produced an active HIV RT that can complement the Pol ^{ts} phenotype. The average number of amino acid substitutions encoded by the non-selected and selected inserts was 4.1 and 2.6, respectively. Twelve of the 109 positive clones yielded inserts

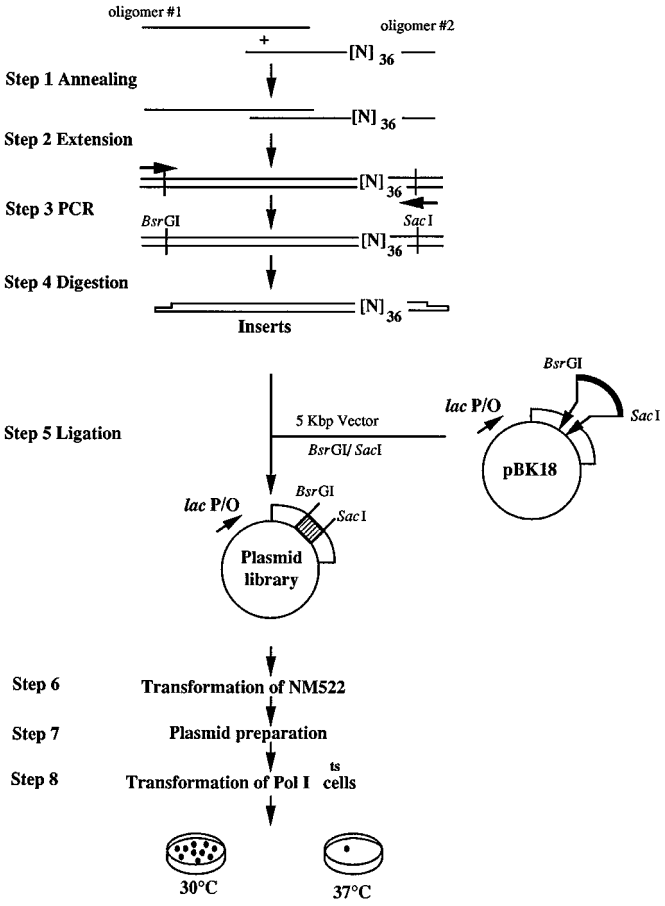

FIG. 1. **Scheme for random sequence mutagenesis of HIV RT coupled with genetic selection for functional mutants.** Construction of a plasmid-borne library of HIV RT variants containing random nucleotide substitutions at codons 67 through 78, and selection of active mutants by complementation of the temperature-sensitive phenotype of an *E. coli* DNA polymerase I mutant, is described under "Experimental Procedures." Synthesis of random nucleotide-containing oligonucleotides (*insets*) is illustrated in *steps 1–4*; $[N]_{36}$ in oligomer 2 denotes 36 contiguous residues containing 12% random nucleotides at each residue. Ligation of the inserts into the stuffer plasmid pBK18, to replace the wild-type HIV RT sequence at codons 67–78 and bring HIV RT under control of the *lac* promoter, is shown in *step 5*. Preparation of the plasmid library and selection for functional mutants in Pol I^{ts} cells is shown in *steps 6–8*.

that encoded no amino acid substitutions, including 6 inserts which had nucleotide substitutions (Table I). Among 97 active mutants, some contained as many as 6 amino acid substitutions within the randomized segment.

The levels of substitution among active mutants, together with the associated complementation efficiencies, are listed in Table I. Overall, there were fewer amino acid changes among mutants that exhibited high efficiency of complementation (large colony size equivalent to that of wild-type HIV RT) than among mutants with lower complementation efficiency. Thus, the average number of amino acid substitutions in the 37 mutants that formed large colonies at 37 °C (+++) was 1.0. Among the 49 mutants that formed medium sized colonies (++), the average number of amino acid substitutions was 3.0. Among the 23 mutants forming small colonies (+) the average number of substitutions was 4.1. All of the single mutants formed colonies that were the same size as that of wild type, whereas most of the active mutants having more than 4 amino acid changes produced smaller colonies at 37 °C.

Single Amino Acid Substitutions—Single amino acid substitutions were observed at 9 of the 12 positions randomized

TABLE I
Characterization of HIV RT mutant sequences

(A) Average number of substitutions in selected and non-selected mutants; (B) distribution of amino acid substitutions in selected mutants; (C) complementation efficiency and mean number of mutations.

A

	Nucleotide substitutions per sequence	Amino acid substitutions per sequence
Randomized oligomers (theoretical)	4.3[a]	
Non-selected mutants (n = 18)	5.0	4.1
Selected mutants (n = 109)	3.3	2.6

B

Number of substitutions	Number of mutants
0	12
1	18
2	23
3	24
4	21
5	9
6	2

C

Complementation efficiency[b]	Nucleotide substitutions per mutant	Amino acid substitutions per mutant
+++ (n = 37)	1.3	1.0
++ (n = 49)	3.9	3.0
+ (n = 23)	4.6	4.1

[a] 12% random substitutions at 36 positions = 0.12 × 36 = 4.3.
[b] +++, large colonies equivalent to wild type; ++, medium colonies; +, small colonies.

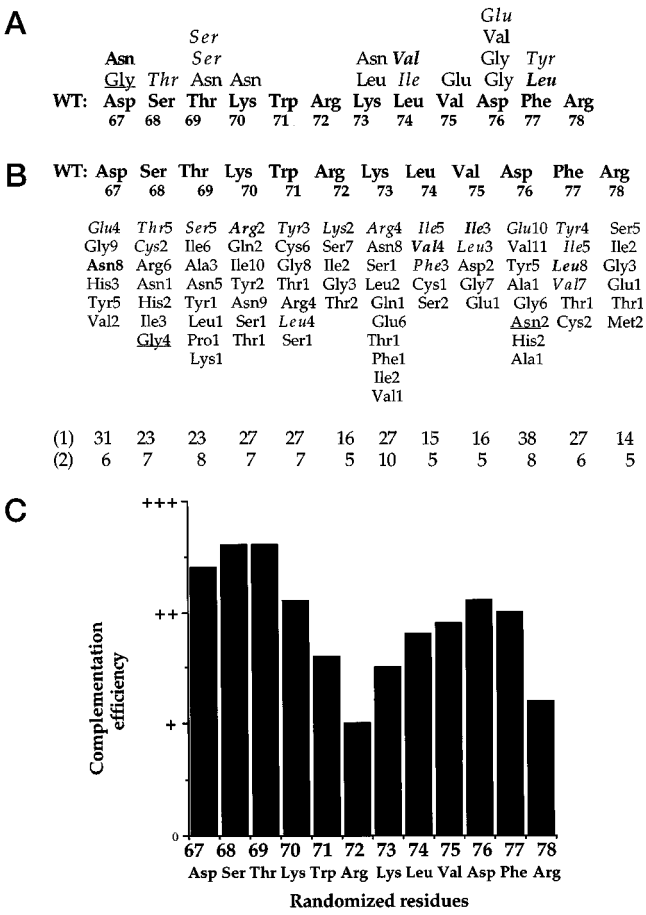

FIG. 2. **Amino acid substitutions in functional HIV RT mutants.** The amino acid sequences of 109 mutants that complemented the temperature-sensitive phenotype of Pol I[ts] cells were determined by DNA sequencing of the random nucleotide containing inserts. *Italicized* type indicates substitution of an amino acid similar to the wild-type. *Underlining* indicates a naturally occurring variation. **Bold** type indicates a substitution found, either singly or together with others, in drug resistant mutants. In *B*, the *number* to the *right* of individual mutant amino acids indicates the number of occurrences at the designated position. The number given at (*1*) below each position is the total number of substitutions at that position; the number at (*2*) is the number of different amino acids identified at that position. In *C*, complementation efficiency denotes the average colony size relative to wild type of all mutants containing substitutions at the designated residue; wild-type efficiency is denoted by +++ on the vertical scale.

(Fig. 2*A*). Mutations at 8 of these 12 positions have been observed in natural variants (positions 68 and 76) or in drug-resistant mutants (positions 67, 69, 70, 74, 75, and 77). As yet, mutations have not been found among natural variants or drug-resistant mutants at Trp-71, Arg-72, Lys-73, and Arg-78. Interestingly, Arg-71, Lys-72, and Trp-78 yielded no single mutants in this study.

Multiple Amino Acid Substitutions—Although individual randomized codons displayed varying yields of single amino acid substitutions (Fig. 2*A*), all 12 codons exhibited nearly equal probability of inclusion among multiple substitutions that complemented the Pol I[ts] phenotype (Fig. 2*B*). Most of the variant amino acids (italicized) have properties similar to those of the wild-type. Three mutations that were previously identified in drug-resistant mutants, D67N (AZT), K70R (AZT), and L74V (ddI), were isolated as either single or multiple mutations. In addition, we isolated two mutations, V75I and F76L, that were present as one of multiple substitutions in HIV RT from patients treated with multiple nucleoside inhibitors (Shirasaka *et al.*, 1995). In Fig. 2*C*, we show the average complementation efficiency (colony size) of all mutants that have substitutions at each of the randomized positions. Notably, mutations at positions that failed to produce single amino acid substitutions (71, 72, and 78) were associated with reduced complementation efficiency at 37 °C, suggesting that these residues are important for DNA-dependent DNA polymerase activity.

Purification of HIV RT Mutants—Wild-type and 13 mutant HIV RTs that complemented the Pol I[ts] phenotype were purified to near homogeneity. The mutants, all of which exhibited >75% plating efficiency and contained from one to six amino acid substitutions, are listed in Table II, together with complementation efficiency (colony size). The HIV RTs, expressed in *E. coli* as NH_2-terminal fusions with a "histidine tag," were purified by metal chelation chromatography. Approximately 95% of the purified wild-type preparation consisted of a 66-kDa protein (p66 subunit) and 5% of a 51-kDa protein (p51 subunit), estimated by visual inspection of a Coomassie Blue-stained 12% SDS-polyacrylamide gel (data not shown). Ten of the 13 mutant HIV RT preparations were similar in yield and purity to the wild-type (0.25–1 mg of total protein of 90–95% purity). Three of the mutant preparations were less pure (about 0.1 mg of total protein of 50–75% purity); reduced recovery of these three proteins could be due to instability, possibly correlated with an increased number of amino acid substitutions (5 to 6 per mutant). Presumably, the restricted time of induction with IPTG (2 h) limited proteolytic processing of p66 to p51.

DNA Polymerase Activity of Mutant HIV RTs Mutants—DNA-dependent DNA polymerase activity of purified HIV RTs, assayed on a gapped DNA template, is listed in Table II. These data demonstrate a positive correlation between complementation efficiency, estimated as colony size, and *in vitro* DNA polymerase activity. Thus the six single and two double mutants displayed a wild-type level of complementation efficiency together with enzymatic activity between 40 and 100% of wild-

TABLE II
DNA polymerase activity of purified HIV RT mutants

Mutant	Complementation efficiency[a]	Activity[b]
		%
WT	+++	100
D67G	+++	90
D67N	+++	85
T69N	+++	90
L74I	+++	100
F77Y	+++	100
F77V	+++	75
L74V/V75L	+++	40
D67N/K73T	++ ~ +++	45
D67N/W71G/K73N	+ ~ ++	15
D67N/R72S/K73E	++	10
R72T/L74V/D76Y/R78T	+	5–10
D67H/S68R/K70R/W71S/L74F	+	5
D67A/W71L/K73H/L74I/D76A/F77I	+	15
D67G/S68T/T69S/R72K/V75I	+ ~ ++	25

[a] +++, large colonies equivalent to wild type; ++, medium colonies; +, small colonies.
[b] DNA-dependent DNA polymerase activity was assayed as incorporation of label from [*methyl*]-^{3}H]TTP into activated DNA by 0.1 μg of purified protein at 37 °C for 20 min. Activity was normalized to that of wild-type HIV RT.

type. The 6 remaining mutants with larger numbers of substitutions displayed decreased complementation efficiency together with 5–25% of wild-type activity (Table II).

Inhibition of Mutant HIV RTs by AZTTP—AZT sensitivity, measured as inhibition of DNA-dependent DNA polymerase activity by AZT triphosphate, was determined for eight highly active mutants (D67N, D67G, T69N, L74I, F77Y, F77V, LV74/V75L, and D67N/K73T). Among the single mutants, a 18- and 16-fold decrease in IC_{50} for AZT triphosphate was observed for F77Y and F77V, respectively, relative to wild type (Fig. 3*A*). L74I showed a 13-fold decrease in IC_{50} (Fig. 3*B*). D67G showed an approximately 4-fold decrease, while no reduction in sensitivity was observed with two other substitutions at the same position (D67N and T69D). Previous studies with purified HIV RT from AZT-resistant mutants (*i.e.* D67N, K70R, T215Y, and K219Q) failed to demonstrate *in vitro* drug resistance even though viruses containing the same substitutions together with others showed a large increase in the IC_{50} for AZT in cultured cells (Larder and Kemp, 1989). Consistent with these observations, D67N/K73T showed decreased sensitivity to AZTTP even though D67N alone did not.

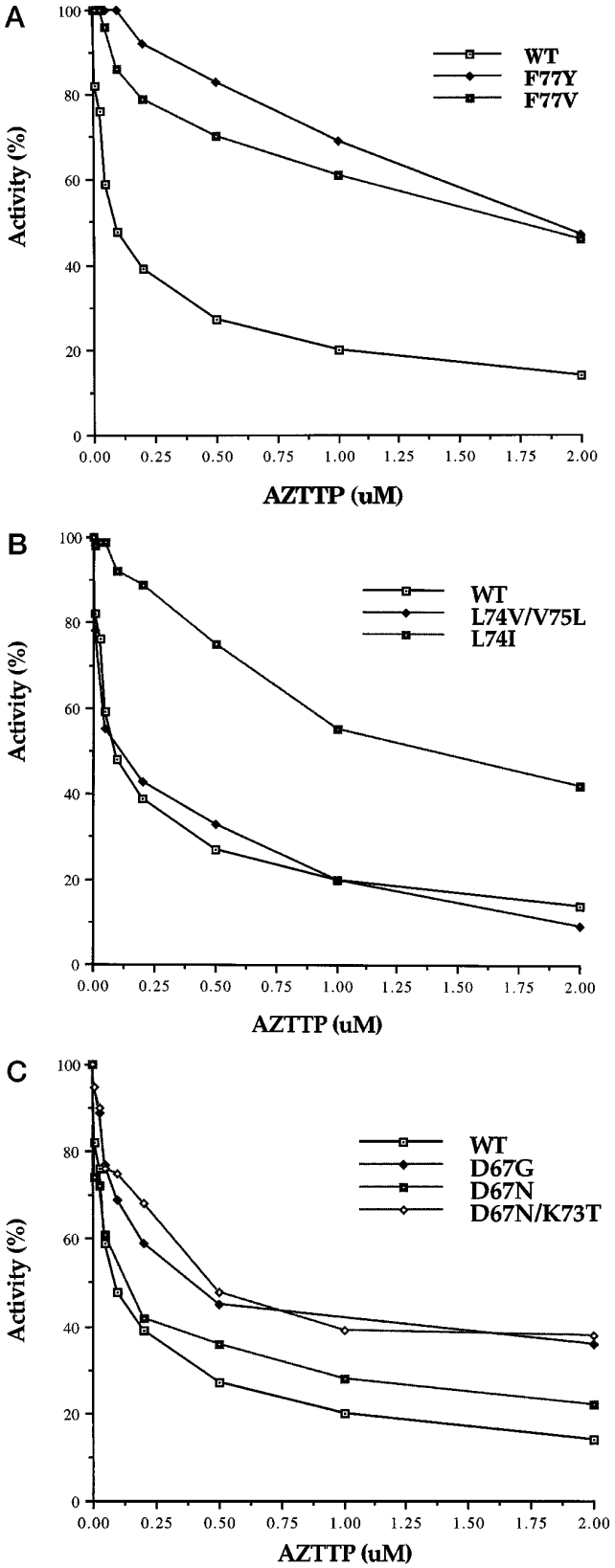

FIG. 3. **Inhibition of DNA polymerase activity of HIV RT mutants by AZTTP.** The DNA-dependent DNA polymerase activity of HIV RT mutants was assayed on a gapped DNA template in the presence of AZTTP as described under "Experimental Procedures."

DISCUSSION

In this work, we tested whether complementation of a replication defective *E. coli* DNA polymerase I mutant by HIV RT (Kim and Loeb, 1995a) can be used to select active HIV RT mutants from a population of variants containing random nucleotide substitutions (Horwitz and Loeb, 1986). In our functional complementation system, expression of HIV RT enables an *E. coli* mutant harboring temperature-sensitive DNA polymerase I to grow at a nonpermissive temperature (Kim and Loeb, 1995a). In these initial experiments with random sequence substitutions in HIV RT, we targeted a region that has little secondary structure and would likely permit a large number of substitutions that yield functional mutants. The crystal structure of HIV RT indicates that the 12 amino acid target forms a flexible loop between the β3 and β4 domains (Tantillo *et al.*, 1994). In addition, many viral isolates that are resistant to nucleoside analogs contain mutations within this region. The inherent flexibility of the target and the occurrence of variants in the natural host is consistent with our finding that all 12 residues are tolerant of substitutions, either singly or in combination with others.

We observed in the experiments described here that approximately 12% of Pol I^{ts} cells transformed with a library encoding random substitutions at amino acids 67–78 formed colonies at 37 °C. The fraction of a library that supports colony formation at nonpermissive temperature, *i.e.* the fraction of encoded sequences that can complement the Pol I^{ts} phenotype, depends on several factors. These include the upper and lower limits of intracellular HIV RT activity compatible with growth of the bacterial host (the window for selection), the proportion of random nucleotides relative to wild-type nucleotides within the substituted segment, and the function of the targeted amino acids in supporting DNA-dependent DNA polymerase activity. With respect to the first factor, it appears that the fraction of transformants capable of complementation can be adjusted by controlling the level of HIV RT expression. Thus, we observed that the fraction of transformants that form colonies at 37 °C decreased from 12 to 7% when IPTG was deleted from the selection plates (data not shown). By reducing the level of expression, we presumably limited recovery to mutants with relatively high specific activity. In the experiments described here where use of 1 mM IPTG gave 12% positive transformants, one-third (*i.e.* 4%) yielded colonies of wild-type size, suggestive of an intracellular HIV RT activity comparable to wild-type. That higher levels of HIV RT expression can be lethal to the host (Kim and Loeb, 1995a) indicates that there is an upper limit of activity compatible with growth. A second factor affecting the fraction of positive transformants is the degree of randomness in the targeted region, with greater randomness and more amino acid substitutions mitigating against levels of activity that promote growth and colony formation. In the present experiments, the target sequence was replaced with 12% random nucleotides at each position, and as a result the wild-type sequence should have occurred at a frequency of 1%, corresponding to 120 of the 12,000 cells plated. Assuming that the predicted 120 wild-type sequences were included in the 1400 positive transformants we recovered, we should have found 9 wild-type isolates (120/1400) among the 109 that we sequenced. That we found 6 wild-type sequences among the 109, approximately the expected occurrence, substantiates the designated composition of the library. Third, it is likely that the importance of the targeted amino acids for DNA polymerase activity affects the fraction of transformants selected, with substitution of residues having crucial functions mitigating against ability to complement. The high frequency of positive transformants that we recovered by targeting the β3-β4 flexible loop is consistent with the nonessential nature of the individual amino acids.

The amino acid replacement we observed among mutants that complement the Pol I^{ts} phenotype (Fig. 2) indicates that Trp-71, Arg-72, and Arg-78 are relatively intolerant of substitution compared with other residues in our 12-amino acid target. This pattern is consistent with substitutions recovered in HIV isolates and with sequence conservation among various viral RTs. Within the target region, substitution at two positions, S68G and D76N, occurs in natural HIV RT variants, and substitution at 6 positions (Asp-67, Thr-69, Lys-70, Leu-74, Val-75, and Phe-77) occurs in drug-resistant viral isolates. Of the 4 residues (Trp-71, Arg-72, Lys-73, and Arg-78) not included among viral mutations, Trp-71, Arg-72, and Arg-78 yielded no substitutions in the single mutants we sequenced (Fig. 2*A*); moreover, mutants with multiple substitutions including one or more of these three positions showed low complementation efficiency (Fig. 2*C*). These data indicate that Trp-71, Arg-72, and Arg-78 are relatively rarely mutable and suggest that they may be important in DNA-dependent DNA polymerase activity. In fact, a recent study showed that two mutations at Arg-72, including R72K, result in greatly decreased activity due to reduction in translocation (Sarafinos *et al*., 1995). Notably, sequence comparison among viral RTs reveals that Trp-71, Arg-72, and Arg-78, as well as Leu-74 and Asp-76 are highly conserved (Barber *et al*., 1990). Taken together, available data indicate that mutations at Arg-72 may be incompatible with viral replication in the natural host.

Among the mutants we analyzed, the number of amino acid substitutions is inversely correlated with colony size and with DNA polymerase activity on a gapped DNA template (Table II). These observations suggest that ability to complement in the *E. coli* system reflects the DNA-dependent DNA polymerase activity of HIV RT, with mutants having 10% of wild-type DNA-dependent DNA polymerase activity being selectable. It has been reported that an HIV RT mutant with 37% of wild-type activity, L74M, supports production of viable virions in cultured human cells, whereas another mutant, L74A, exhibiting 30% of wild-type activity, does not (Lacey and Larder, 1994). Thus, selection for mutants that support bacterial growth is apparently less stringent than selection for virion production in cell culture systems. In accord, the HIV RT mutants we selected that have significantly reduced DNA polymerase activity (Table II) are unlikely to be identified in viral populations because they presumably cannot support viral replication.

Six amino acids within the segment we targeted (Arg-67, Thr-69, Lys-70, Leu-74, Val-75, and Phe-77) have been found to be mutated in variants resistant to nucleoside analogs (Tantillo *et al*., 1994). We analyzed some of the mutations we isolated at these positions for AZTTP resistance. As illustrated in Fig. 3*A*, F77Y and F77V conferred an 18- and 16-fold decrease in inhibition by AZTTP, respectively. Although these substitutions have not been reported in viral isolates from patients, it was recently reported that individuals receiving therapy with a combination of nucleoside analogs developed a F77L mutation subsequent to a Q151M mutation. Thus, F77L is encompassed in the evolution of a set of five mutations, beginning with Q151M, followed by F77L and F116Y and later by A62V and V75I, that confers multiple drug resistance (Shirasaka *et al*., 1995). The L74I mutation we isolated exhibited an 8-fold reduction in inhibition by AZTTP; Fig. 3*B*). A similar mutation, L74V, has been isolated from patients after ddI treatment (St. Clair *et al*., 1991) and exhibited reduced sensitivity to ddGTP (Lacey *et al*., 1992). Thus Leu-74 is likely to be involved directly or indirectly in substrate selection. Interestingly, L74V displayed wild-type sensitivity to AZTTP in combination with V75L. The mutation D67N, in combination with additional mutations (K70R, T215Y, and K219Q) exhibits a more than 100-fold increased resistance to AZT in a cell culture system (Larder and Kemp, 1989). However, a purified HIV RT mutant containing all four mutations failed to show resistance to AZTTP *in vitro* (Skalka and Goff, 1993). In our study, D67N did not affect sensitivity to AZTTP, consistent with the observation that virions containing D67N as a single mutation did not show increased resistance to AZT *in vivo* (Larder *et al*., 1991; Larder and Kemp, 1989). However, we did find that both D67G and D67N/K73T exhibited increased resistance to AZTTP (Fig. 3*C*).

Many mutations in HIV RT have been observed in virions that are resistant to nucleoside analogs in patients and in cultured cells. However, purified HIV RT bearing the same mutations frequently does not exhibit resistance to the corresponding nucleoside triphosphates when assayed *in vitro*. Conversely, we isolated both single and multiple substitutions that render HIV RT resistant to AZTTP, but have not yet been identified in drug-resistant viral isolates. Considering the exceptionally high mutation rate of the virus (Preston *et al*., 1988; Roberts *et al*., 1989), the high error rate of the reverse tran-

scriptase *in vitro* (Preston *et al.*, 1988) and recent evidence for rapid viral replication during the course of HIV infection (Delwart *et al.*, 1993), it seems likely that all single base substitutions and many multiple substitutions would have arisen and been "tested" for resistance during AZT therapy in patients. If viruses containing the mutations we identified are in fact resistant to AZT, why have they not been selected? Some may not have a high enough DNA-dependent DNA polymerase activity to compete successfully (*e.g.* F77V, Table II). However, some (*e.g.* F77Y and L74I which may have wild-type levels of this activity) might interfere with additional activities or properties of HIV RT which are essential for production of virions, such as strand transfer/displacement, processivity, or replication fidelity, and might be inadequate for viral replication in the natural host.

Several conclusions can be drawn from our initial experiments on random sequence substitution of HIV RT and selection by genetic complementation. First, our positive genetic selection system can identify active HIV RT mutants. The mutants we characterized exhibit from 5 to 100% of the DNA-dependent DNA polymerase activity of the wild-type enzyme and include a majority of substitutions located within the target region that have been observed as natural variants and drug-resistant mutations. The ability to select active mutations and to characterize their phenotypes in an alternate bacterial host provides a powerful means for assessing the consequences of specific amino acid substitutions, and combinations of substitutions, on DNA polymerase activity. Second, random mutagenesis of HIV RT enables us to assess the mutability of each amino acid residue in the enzyme. Rarely mutable or immutable residues can be identified by this approach, and such residues can serve as potential targets for more efficacious anti-HIV drug therapy that precludes or delays emergence of resistant variants. Since random mutagenesis can be applied to large targets in HIV RT, and can also survey all possible combinations of amino acid substitutions within the target, the approach offers different parameters for identifying essential amino acids than does inactivation by substitution with alanine (Richardson and Richardson, 1990). Although substitutions within the flexible region we targeted in this study did not reveal immutable sites, it remains to be determined if other more structured domains contain such immutable (essential) residues. Finally, random mutagenesis creates HIV RT mutants which are unlikely to be recovered from the natural host because they cannot support viral replication. Phenotypic and biochemical analysis of such mutants, many of which will have new or altered biochemical properties, can provide insight into the involvement of individual amino acid residues in catalysis and into structure-function relationships within the enzyme.

Acknowledgments—We are grateful to Ann Blank and Margaret Black for generous counsel throughout the course of this study.

REFERENCES

Battula, N., and Loeb, L. A. (1974) *J. Biol. Chem.* **249,** 4086–4093
Boucher, C. B., O'Sullivan, E., Mulder, J. W., Ramantarsing, C., Kellam, P., Darby, G., Lange, J. M. A., Goudsmit, J., and Larder, B. A. (1992) *J. Infect. Dis.* **165,** 105–110
Boyer, P. L., Tantillo, C., Jacobo-Molina, A., Nanni, R. G., Ding, J., Arnold, E., and Hughes, S. H. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4882–4886
Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
Delwart, E. D., Shpaer, E. G., Louwagie, J., McCutchan, F. E., Grez, M., Rubsamen-Waigmann, H., and Mullins, J. I. (1993) *Science* **262,** 1257–1261
Fitzgibbon, J. E., Howell, R. M., Haberzettl, C. A., Sperber, S. J., Gocke, D. J., and Dubin, D. T. (1992) *Antimicrob. Agents Chemother.* **36,** 153–157
Gu, Z., Gao, Q., Li, X., Parniak, M. A., and Wainberg, M. A. (1992) *J. Virol.* **66,** 7128–7135
Gu, Z., Fletcher, R. S., Arts, E. J., Wainberg, M. A., and Parniak, M. A. (1994) *J. Biol. Chem.* **269,** 28118–28122
Horwitz, M. S., and Loeb, L. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 7405–7409
Kim, B., and Loeb, L. A. (1995a) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 684–688
Kim, B., and Loeb, L. A. (1995b) *J. Virol.* **69,** 6563–6566
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 488
Lacey, S. F., and Larder, B. A. (1994) *J. Virol.* **68,** 3421–3424
Lacey, S. F., Reardon, J. E., Furfine, E. S., Kunkel, T. A., Bebenek, K., Eckert, K. A., Kemp, S. D., and Larder, B. A. (1992) *J. Biol. Chem.* **267,** 15789–15794
Larder, B. A., and Kemp, S. D. (1989) *Science* **246,** 1155–1158
Larder, B. A., Kellam, P., and Kemp, S. D. (1991) *AIDS* **5,** 137–144
Mitsuya, H., and Broder, S. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 1911–1915
Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988) *Science* **242,** 1168–1171
Richardson, J. S., and Richardson, D. C. (1990) *Principles and Patterns of Protein Conformation,* Plenum Press, New York
Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A., and Kunkel, T. A. (1989) *Mol. Cell. Biol.* **9,** 469–476
Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Reference Manual,* Cold Spring Harbor Press, Cold Springs Harbor, NY
Sarafianos, S. G., Pandey, V. N., Kaushik, N., and Modak, M. J. (1995) *J. Biol. Chem.* **270,** 19729–19735
Shirasaka, T., Kavlick, M. F., Ueno, T., Gao, W.-Y., Kojima, E., Alcaide, M. L., Chokekijchai, S., Roy, B. M., Arnold, E., Yarchoan, R., and Mitsuya, H. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 2398–2402
Skalka, A. M., and Goff, S. P. (1993) *Reverse Transcriptase,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
St. Clair, M. H., Martin, J. L., Tudor-Williams, G., Bach, M. C., Vavro, C. L., King, D. M., Kellam, P., Kemp, S. D., and Larder, B. A. (1991) *Science* **253,** 1557–1559
Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A. J., and Arnold, E. (1994) *J. Mol. Biol.* **243,** 369–387
Witkin, E. M., and Roegner, M. V. (1992) *J. Bacteriol.* **174,** 4166–4168