Volume 17 Number 4 1989 Nucleic Acids Research

Fidelity of human immunodeficiency virus type I reverse transcriptase in copying natural DNA

J.Weber and F.Grosse

Department of Chemistry, Max-Planck-Institute for Experimental Medicine, Hermann-Rein-Strasse 3, D-3400 Göttingen, FRG

Received December 22, 1988; Accepted January 20, 1989

ABSTRACT

Reverse transcriptase from the human immunodeficiency virus type I (HIV-1) was expressed in *E. coli* and purified to near homogeneity. The enzyme was shown to contain reverse transcriptase, DNA polymerase and ribonuclease H activities. The DNA polymerase activity converted singly-primed ΦX174 (+) DNA into the double-stranded form. Two third of the replication product is ligatable to covalently closed circular DNA (RFIV-form DNA) indicating that DNA synthesis by HIV reverse transcriptase can proceed until the enzyme matches the 5'-end of a pre-existing primer molecule. The *in vitro* accuracy of HIV reverse transcriptase was measured with the ΦX174*am*16 reversion assay to be 1/7,400. Reversion rates for the individual mispairs were determined from pool bias studies to be 1/8,000 for the dGMP:$T_{template}$ mismatch, 1/35,000 for the dGMP:$A_{template}$ mismatch, 1/45,000 for the dAMP:$G_{template}$ mismatch, 1/73,000 for the dCMP:$T_{template}$ mispair, 1/140,000 for the dCMP:$A_{template}$ mispair, and 1/180,000 for the dGMP:$G_{template}$ mismatch. The dTMP:$T_{template}$ mispair was below the detection limit of the assay indicating a reversion rate of less than 1/300,000 for this particular mispair.

INTRODUCTION

The human immunodeficiency virus HIV is the etiologic agent of the acquired immune deficiency syndrome (AIDS) (1,2). As a member of the group of retroviruses the HIV genome codes for a particular retroviral enzyme, the HIV reverse transcriptase (EC 2.7.7.49). Reverse transcriptase (RT) is a unique enzyme because it catalyzes both, RNA-directed DNA synthesis and DNA-directed DNA synthesis. Due to its RNA-directed DNA synthesis capability it is the only enzyme that can convert viral RNA into DNA, which in turn is a prerequisite for viral gene expression and further amplification of the viral genome. Since the reverse transcriptase is crucial for viral growth and such an activity apparently does not exist in non-infected human cells, it has become the main target for antiviral therapies.

A critical feature of reverse transcriptases from retroviruses is their rather

low accuracy of synthesizing DNA. The error frequencies may be as high as 1/700 per nucleotide polymerized (3). This might give a rational for the frequently observed high variability of the retroviral genome, which is particularly evident with HIV (4). Such a low fidelity might also be harmful for the processes of cellular DNA replication and DNA repair, if the DNA polymerase activity of HIV RT also participates in cellular DNA metabolizing processes. Therefore, studying the accuracy of HIV reverse transcriptase might lead to important clues not only for the genetic variability of the HIV genome but also to the accuracy of DNA replication in infected cells.

In the present paper we describe the expression cloning of the HIV *pol*-gene and the purification of bacterially expressed HIV RT. We show that HIV RT, in addition to its RNA-directed DNA synthesizing capacity, is also active with singly-primed single-stranded DNA as a template. The efficient replication of natural bacteriophage DNA allowed the determination of the base substitution fidelities. With the ΦX174 amber 16 reversion assay, the frequency of all the seven evaluable base-pairing mismatches were determined. The results might help to an understanding of the high mutation frequency of HIV which is observed *in vivo*.

MATERIALS AND METHODS

Materials

The *E. coli* strain JM109 and plasmid pSS20°c were described in (5). Bacterial indicator strains *E. coli* C (wild type, *sup°*), *E. coli* CQ_2 *(supE)*, *E. coli* C600 *(supF,* for the preparation of spheroplasts) and phage ΦX174*am*16 were as described in (6). ΦX174*am*16 single- and double-stranded DNA were kindly provided by Dr. G. Krauss, Universität Bayreuth, FRG. The 14-base oligonucleotides 5´-CAGTAACTTTTCCC-3´ and 5´-CCCAGCCTGAATCT-3´ were synthesized on an Applied Biosystems model 380A oligonucleotide synthesizer, deprotected, 5´–phosphorylated (7) and hybridized in a 10:1 ratio (primer/template) to ΦX174*am*16 single-stranded DNA. Activated DNA, with a 20% acid solubility, was prepared as described (8). Polynucleotides and oligonucleotide primers were obtained from Pharmacia, radioactive nucleoside triphosphates were from Amersham. T4 DNA ligase and *E. coli* RNA polymerase were kind gifts of Dr. F. Eckstein, this department, and Dr. H. Sternbach, this department, respectively. Column materials, DE52 and P11, were purchased from Whatman. Heparin-Sepharose was prepared as described (9). All other chemicals were from commercial sources.

Cloning and Expression of HIV RT in *E. coli.*

The RT expression vector was constructed by inserting the HIV *pol*-gene from position 2093 (BglII position) to 5819 (SalI position) of the Ratner sequence (10) into the BglII/SalI-sites of the multilinker of pEMBL19 (11,12), excising the gene from pEMBL19 by BglII/XbaI cleavage, and subsequently inserting the BglII/XbaI fragment into the corresponding sites of the expression vector pSS20*c (5), yielding the plasmid pJS3.7. To express the *pol*- gene, *E. coli* JM109, transfected with pJS3.7, was grown in NZCYM-medium (10 g NZ-amine, 1 g Casamino acids, 5 g yeast extract, 2 g $MgSO_4$ X 7 H_2O, 5 g NaCl, 1 ml of 1% (w/v) thiamine hydrochloride per litre bidistilled water) overnight at 25°C to an optical density (at 600 nm) of approximately 2. The expression of the fusion protein was induced by adding IPTG to a final concentration of 8 mM.

Purification of HIV RT from *E. coli.*

5 hrs after IPTG-induction cells were collected by centrifugation and resuspended to a density of about 20 OD_{600} into 50 mM Tris/HCl (pH 7.5), 1 mM EDTA, 7 mM 2-mercaptoethanol, 0.4 M NaCl. The suspension was treated with 100 µg/ml lysozyme for 30 min at 25°C, chilled to 0°C, and sonicated on ice with 10-20 bursts of 5 sec duration each with 55 sec chilling intervals. The sonicated suspension was centrifuged for 30 min at 16,000 x g, decanted from the pellet and dialyzed 2 times against a 20-fold volume each of buffer A (50 mM Tris/HCl (pH 7.0), 1 mM EDTA, 3 mM 2-mercaptoethanol, 0.1% (v/v) Nonidet P-40, and 10% (v/v) glycerol). The dialyzed crude extract was passed over DEAE-cellulose (DE52, equilibrated with buffer A, 1 ml wet volume/ml crude extract) to remove nucleic acids and then loaded onto a second DEAE-cellulose (equilibrated with buffer A, 1 ml wet volume/ml crude extract). After washing with two column volumes of buffer A, RT was eluted with 2 column volumes of buffer B (buffer A +100 mM KCl). Fractions with RT activity were collected and directly loaded onto a phosphocellulose column (P11, equilibrated with buffer B, 0.1 column volume/ml crude extract). The column was washed with 2 volumes of buffer B and then eluted with 10 column volumes of a linear gradient from 100 to 500 mM KCl in buffer A. Activity-containing fractions were diluted to a conductivity equivalent to that of buffer B. The diluted fraction was loaded on a heparin-Sepharose (equilibrated with buffer B, 0.1 column volume/ml crude extract). The affinity column was washed with 2 column volumes buffer B and eluted with 10 column volumes of a linear gradient containing 100 and 500 mM KCl, respectively. Activity-containing fractions were combined and dialyzed against buffer C (50 mM

potassium phosphate, pH 7.8, 2 mM 2-mercaptoethanol, 0.1% Nonidet P-40 and 50% (v/v) glycerol). The dialyzed fraction was stored at -20°C. About 10% of the activity was lost after storage for six month.

RNA-dependent DNA Polymerase Assay (RT Assay).

The assay mixture contained in a 200 µl reaction volume 50 mM Tris/HCl (pH 8.0), 80 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, 25 µg/ml poly(rA), 10 µg/ml oligo$(dT)_{12-18}$, 100 µM (^{32}P-α)dTTP (50 cpm/pmol), 0.05% (v/v) Nonidet P-40 and 2 µl enzyme. After 2 to 10 min at 37°C the reaction was stopped by spotting an aliquot of the assay mixture onto a glass fiber filter (BA55, Schleicher and Schüll, Dassel, FRG) and immersing the filter into ice-cold 10% (w/v) trichloroacetic acid. After 5 min the filter discs were transferred to a suction device and washed 5 times with 1 ml each of 1 M HCl, followed by 2 times with 1 ml each of 96% ethanol. Filters were dried with a fan. The incorporated radioactivity was determined by toluene-based scintillation counting. 1 unit of reverse transcriptase activity is defined as the amount of enzyme that catalyzes the incorporation of 1 nmol dTMP into poly(rA):oligo(dT) in 10 min at 37°C.

DNA-dependent DNA Polymerase Assay.

Assays using poly(dA-dT) or activated DNA as template-primers were performed in 50 µl assay mixture containing 50 mM Tris/HCl (pH 8.0), 50 mM NaCl, 10 mM $MgCl_2$, 0.05% Nonidet P-40, 100 µM (nucleotide) DNA, 100 µM (each) dNTPs and 0.2 µg RT. For studying singly-primed single-stranded DNA replication a 14-mer oligonucleotide, complementary to region AB5282 to AB5296 of the ΦX174 genome (13), was annealed in a 10-fold molar excess to the ΦX174 DNA (+) strand DNA. The assay mixture (100 µl) contained 50 mM Tris/HCl (pH 8.0), 50 mM NaCl, 10 mM $MgCl_2$, 0.05% Nonidet P-40, 1 mM rATP, 15 units T4 DNA ligase, 60 µM (nucleotide) singly-primed ΦX174 (+) strand DNA, 100 µM (each) dNTPs and 1.2 µg RT. When pool biases were applied, the concentration of one of the four dNTPs was raised to 1 mM and to 3 mM. After 22 h at 37°C, 10 µl of the assay mixture was withdrawn, heated to 70°C, and then analyzed on a 1% agarose gel (containing 0.2 µg/ml ethidium bromide). Within this time > 90% of the DNA was replicated; about 2/3 of the material formed covalently closed circular DNA (RFIV), 1/3 formed open circular double-stranded DNA (RFII) (Figure 2). Only material with > 90% RFII/RFIV DNA was used for transfection.

Fidelity Assay.

Fidelity measurements were performed as described (6,14,15) with the modifications given in (16). Following RFII/IV formation of ΦX174*am*16 DNA, HIV

RT and T4 DNA ligase were inactivated by heating to 70°C for 10 min. Spheroplasts of *E. coli* C600 *(supF)* were transfected with 0.05 to 0.2 fmol and with 50-100 fmol of RFII/IV DNA to measure the expression of the progeny and revertant phages, respectively. After incubation at 30°C for 10 min, the spheroplasts were mixed with soft agar containing *E. coli* CQ_2 *(supE)* for the expression of all progenies, and with soft agar containing *E. coli* C *(sup°)* for the expression of revertants at the amber codon. The mixtures were plated on L-agar. *E. coli* CQ_2 was incubated at 37°C, *E. coli* C at 30°C. To analyse the individual mutational pathways, plaques were taken from the revertant plates and re-grown on *E. coli* C *(sup°)* at different temperatures. Incubation at 39°C allows to differentiate between wild-type and related phages (large plaques), ts_{38} (small plaques) and $ts_{34/35}$ (no plaques). Phages that are forming large plaques at 39°C were re-plated and grown at 43°C. Wild-type and pseudo-wild-type phages yield large plaques, ts_{43} progenies give rise to small plaques, whereas ts_{42} do not grow at this temperature. Phages that do not grow at 39°C were re-plated and grown at 34.5°C. ts_{35} progenies form middle-size plaques and ts_{34} progenies form pin-point-sized plaques at this temperature. Reversion frequencies were calculated from the ratio of (number of revertants to a distinct phenotype) ÷ ((total number of progeny phages) x (minus strand expression rate)).

The minus strand expression was determined by hybridizing a 5´-phosphorylated 14-mer oligonucleotide, complementary to region AB5271 to AB5285 of ΦX174*am* 16 DNA, elongating it to RFII/IV in the presence of HIV RT and T4 DNA ligase, transfecting the DNA to *E. coli* C600, and plating the resulting progenies onto *E. coli* C and *E. coli* CQ_2, respectively. This particular 14-mer introduces a $G_{template}$:T mispair at position AB5276 of the amber 16 codon, leading to pseudo-wild-type progenies. The ratio between pseudo-wild-type progenies and total progenies was determined to be 0.27.

Preparation of Radioactively Labeled (^{3}H)Poly(rA):poly(dT).

The reaction mixture (0.1 ml) contained 20 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 0.5 mM (^{3}H)ATP (1,000 cpm/pmol), 0.1 mg/ml poly(dT) and 20 ng RNA polymerase from *E. coli*. The reaction was completed after 5 min incubation at 37°C. After phenolization and precipitation with ethanol (7), the DNA:RNA hybrid was dissolved in 0.1 ml sterile 10 mM Tris/HCl, pH 7.8, 1 mM EDTA.

Ribonuclease H Assay.

Reaction mixtures (50 µl) contained 20 mM Tris/HCl, pH 7.3, 75 mM KCl, 5 mM $MnCl_2$ (or 8 mM $MgCl_2$), 4 mM dithiothreitol, and 20 µM (nucleotide) poly(dT):(^{3}H)poly(rA). Incubation was for 10 min at 37 °C. Reactions were

terminated by spotting aliquots onto GF/C filter discs (Whatman) and immersing the discs into ice-cold 10% trichloroacetic acid. The filters were washed and dried as described above. Acid precipitable radioactivity was determined by scintillation counting. One unit of RNase H activity is defined as the amount of enzyme that catalyzes the hydrolysis of 1 nmol (^{3}H)poly(rA), hybridized to poly(dT), in 10 min at 37°C.

RESULTS

Purification of Bacterially Expressed HIV RT from *E. coli* Crude Extracts.

The HIV *pol*-gene from position 2093 to 5819 of the Ratner sequence (10) was inserted into the multicloning site of the expression vector pSS20*c (5). Upon IPTG induction, a fusion protein consisting of β-galactosidase, a collagenase linker and HIV protease/RT/endonuclease is expressed in *E. coli*. From the fusion protein the co-expressed HIV protease liberates HIV RT. After opening the cells the bulk of nucleic acids and some unrelated proteins were removed by passing the crude extract through DEAE cellulose. The flow-through material was loaded onto a second DEAE cellulose column. RT containing proteins were eluted from this column batch-wise with 100 mM KCl. The eluate was directly applied to a phosphocellulose column. Elution was achieved with a gradient from 100 to 500 mM KCl. RT-containing fractions were collected and dialyzed to a conductivity equivalent to 100 mM KCl. This material was loaded onto a heparin-Sepharose column. A gradient from 100 to 500 mM KCl eluted nearly homogeneous RT at a position of about 250 mM. From 2 l *E. coli* culture, about 3-4 mg of pure enzyme was obtained. A summary of the purification procedure is shown on Table I and in Figure 1.

Enzymatic Activities of the Expression Cloned RT.

The purified HIV RT was most active with poly(rA):oligo(dT) as template-primer, displaying a specific activity of 4,010 nmol TMP incorporated per mg enzyme in 10 min at 37°C. Activity was 4-fold less with poly(rC):oligo(dG). Long stretches of single-stranded DNA templates were replicated with about 1/5 the rate observed for poly(rA)-directed DNA synthesis. Short-gapped DNA (activated DNA) as well as double-stranded DNA (poly(dA-dT):poly(dA-dT)), on the other hand, were poor templates (Table II). DNA synthesis was dependent on the presence of a divalent cation. 5 to 10 mM Mg^{2+} was found to be optimal; Mn^{2+} could replace Mg^{2+} at an optimal concentration of 0.5 mM $MnCl_2$. In addition to its DNA synthesizing capability, purified RT possesses significant amounts of ribonuclease H activity (17). The specific activity for ribonuclease H was determined with (^{3}H)poly(rA):poly(dT) as 1850 nmol

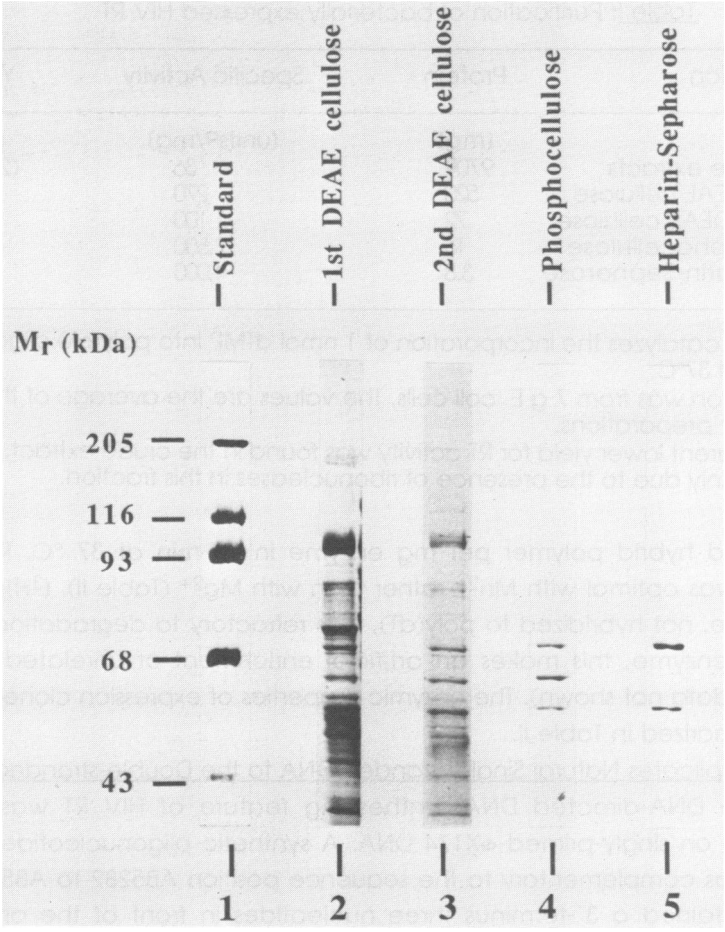

Figure 1: SDS gel analysis of the polypeptide structure of the expression cloned HIV RT during purification from *E. coli* crude extracts. Lane 1: Molecular weight markers were rabbit muscle myosin (205,000), E. coli β-galactosidase (116,000), rabbit muscle phosphorylase b (93,000), bovine serum albumin (68,000), and chicken ovalbumin (43,000). Lane 2: Flow-through material of the first DEAE-cellulose containing 10 μg protein. Lane 3: 5 μg HIV RT-containing fractions from the second DEAE cellulose. Lane 4: 2 μg HIV RT-containing fractions from the phosphocellulose eluate. Lane 5: 1 μg of heparin - Sepharose-purified HIV RT. The proteins were separated on an 8.5% acrylamide/0.17% methylenebisacrylamide gel in the presence of 0.1% sodium dodecylsulfate as a denaturant (40). After electrophoresis, the gel was stained with Coomassie.

Table I: Purification of bacterially expressed HIV RT

Fraction	Protein	Specific Activity	Yield
	(mg)	(units[a]/mg)	(%)
Crude extracts	970[b]	36	(25)[c]
1st DEAE cellulose	520	270	100
2nd DEAE cellulose	72	1,100	56
Phosphocellulose	16	2,500	29
Heparin-Sepharose	3.8	4,000	11

[a]1 unit RT catalyzes the incorporation of 1 nmol dTMP into poly(rA):oligo(dT) in 10 min at 37°C.
[b]Purification was from 7 g E. coli cells. The values are the average of three different preparations.
[c]An apparent lower yield for RT activity was found in the crude extract. This was mainly due to the presence of ribonucleases in this fraction.

degraded hybrid polymer per mg enzyme in 10 min at 37 °C. RNase H activity was optimal with Mn^{2+} rather than with Mg^{2+} (Table II). (^{3}H)poly(rA) alone, i. e. not hybridized to poly(dT), was refractory to degradation by the purified enzyme, this makes an artificial enrichment of unrelated RNases unlikely (data not shown). The enzymic properties of expression cloned HIV RT are summarized in Table II.

HIV RT Replicates Natural Single-stranded DNA to the Double-stranded Form.

The DNA-directed DNA synthesizing feature of HIV RT was further analysed on singly-primed ΦX174 DNA. A synthetic oligonucleotide primer, which was complementary to the sequence position AB5282 to AB5296 (13) and contained a 3´-terminus three nucleotides in front of the amber 16 codon, was hybridized to ΦX174*am*16 DNA and subsequently elongated by HIV RT in the presence of T4 DNA ligase and ATP. The primer was elongated by HIV RT with a linear kinetics for about 20 h (data not shown). An analysis of the replication products on agarose gels revealed that during this time replication was driven to near completeness (Figure 2). The observed conversion of approximately 70% of the open circular DNA (RFII) into covalently closed DNA (RFIV) indicates that HIV RT is able to synthesize until it matches the 5´ end of a pre-existing primer.

Fidelity of HIV RT on In Vitro Replication of ΦX174*am* 16 DNA.

ΦX174*am*16 DNA, replicated *in vitro* by HIV RT, was used for transfecting suppressor tRNA-containing *E. coli* spheroplasts. This produces viable and infectious progeny phages, which in turn can be used for the infection of

Table II: Enzymatic activities of purified HIV RT

		Specific Activities[a]
Reverse Transcriptase:		
	Poly(rA):oligo(dT)	4,010 U/mg
	Poly(rC):oligo(dG)	1,048 U/mg
DNA-directed DNA polymerase:		
	Poly(dA-dT):poly(dA-dT)	144 U/mg
	Activated DNA	177 U/mg
	Singly-primed ΦX174 DNA	718 U/mg
Ribonuclease H:		
	5 mM $MnCl_2$	1,850 U/mg
	8 mM $MgCl_2$	210 U/mg

[a]1 U is defined as the amount of enzyme that catalyzes the incorporation/degradation of 1 nmol nucleotides in 10 min at 37°C.

different *E. coli* strains. Progeny infection of suppressor tRNA-containing *E. coli* cells yields the total number of viable ΦX174*am*16 phages that have been obtained after *in vitro* replication of the single-stranded DNA, whereas infection of wild-type *E. coli* cells scores only the number of revertants at the amber codon (6,14,15). The ratio between plaques formed on *E. coli* C (wild-type) and *E. coli* CQ_2 *(supF),* divided by the minus strand expression factor (0.27, cf. Materials and Methods) to account for repair processes inside the spheroplast, gives the overall fidelity of RT at the amber codon. With three different enzyme preparations, fidelities between 1/5,000 to 1/7,000 were determined. This is more than 20-fold above the frequency of spontaneous mutations of uncopied ΦX174*am*16 DNA, indicating that errors are indeed introduced during *in vitro* DNA synthesis by HIV RT. In a further control experiment, the overall fidelity of human lymphoblast DNA polymerase α was measured to be better than 1/1,000,000 (Bialek et al., submitted for publication).

Seven mutational pathways contribute to the overall fidelity at the amber 16 codon: dGMP:$T_{template}$, dCMP:$T_{template}$, and dTMP:$T_{template}$ mismatches at the first nucleotide, dGMP:$A_{template}$ and dCMP:$A_{template}$ at the second nucleotide, and dGMP:$G_{template}$ and dAMP:$G_{template}$ at third nucleotide all lead to mutated progenies that grow on wild-type *E. coli* strains. Moreover, each of the individual revertant types displays a different growth characteristic at different temperatures. Thus, by measuring the

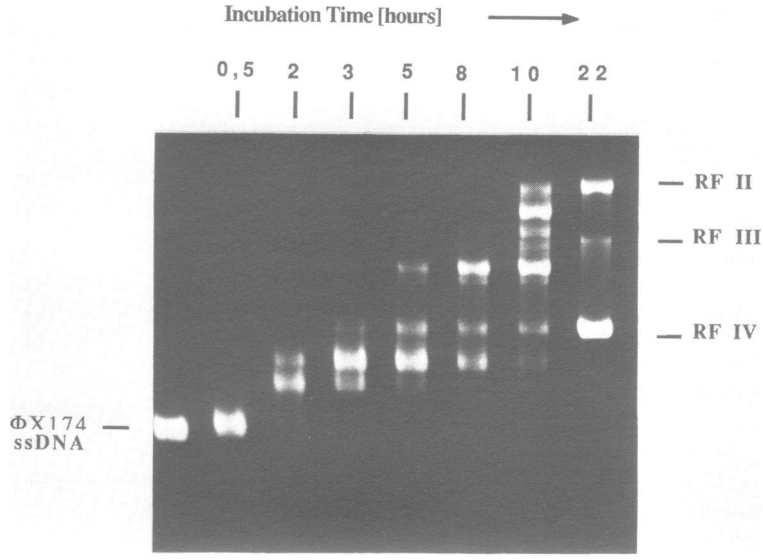

Figure 2: Replication of singly-primed ΦX174*am*16 DNA by purified HIV-1 reverse transcriptase. 2 µg oligonucleotide-primed ΦX174 (+) strand DNA was replicated at 37 °C in 100 µl assay mixture containing 20 mM Tris/HCl (pH 8.0), 50 mM NaCl, 10 mM $MgCl_2$, 0.05% Nonidet P-40, 1 mM rATP, 15 units T4 DNA ligase and 100 µM (each) dNTPs. After the times indicated 10 µl aliquots were withdrawn, heated to 70°C, and then analyzed on a 1% agarose gel (containing 0.2 µg/ml ethidium bromide). Bands were visualized with an UV-lamp.

temperature to which each of the revertants is able to grow one can evaluate the error rate for each of the seven mispairs (6,14,15). Furthermore, reversions can be forced into distinct pathways by biasing the concentrations of the four dNTPs during *in vitro* replication. For example, increasing dGTP over the other three dNTPs will favour the occurrence of dGMP:$T_{template}$, dGMP:$A_{template}$, and dGMP:$G_{template}$ revertants, which are able to grow at ≥ 44°C (*ψ wt*), ≤ 38°C (*ts_{38}*), and ≤ 35°C (*ts_{35}*), respectively. Thus, by performing replication assays under biased dNTP concentrations and scoring only the phages of interest it is possible to measure rate laws for the occurrence of each of the seven revertants as a function of the nucleotide bias employed. This procedure has been performed for the seven evaluable mismatches. With the exception of the dTMP:$T_{template}$ mispair, all the the biases studied gave a linear response in the occurrence of the relevant phenotypes (Table III). The error frequencies for the occurrence of individual mismatches under

Table III: Effect of pool biases on the fidelity of HIV RT.

Pool Bias applied	Ratio	Mispair induced	Phenotype evaluated	Reversion frequency[a]	Estimated error rate[b]
				(x 10^{-5})	
G↑ A↓	1:1	G:$T_{template}$	*ψwt*	2.17[c]	
	10:1			16.0	
	30:1			98.0	1/8,000
C↑ A↓	1:1	C:$T_{template}$	*wt*	2.17[c]	
	10:1			4.49	
	30:1			12.6	1/73,000
T↑ A↓	1:1	T:$T_{template}$	*ts43*	0.73	
	10:1			5.16	
	30:1			1.97	< 1/300,000[d]
G↑ T↓	1:1	G:$A_{template}$	*ts42*	0.318	
	10:1			2.76	
	30:1			21.7	1/35,000
C↑ T↓	1:1	C:$A_{template}$	*ts34*	0.188	
	10:1			1.18	
	30:1			5.6	1/140,000
A↑ C↓	1:1	A:$G_{template}$	*ts38*	0.14	
	10:1			4.8	
	30:1			17.2	1/45,000
G↑ C↓	1:1	G:$G_{template}$	*ts35*	0.043	
	10:1			2.6	
	30:1			4.65	1/180,000

[a]The reversion frequency is given by (number of the evaluated mutant) ÷ (number of total plaques) ÷ 0.27.

[b]The error rates were calculated from the slopes of the straight lines of plots of the reversion frequency over the nucleotide pool bias applied.

[c]Without applying pool biases, only the sum of *ψwt* and true wild-type is evaluable.

[d]Since no measurable pool bias effect was observable, only the number of revertants at the 30-fold bias was considered for the estimation of the error rate.

unbiased conditions were derived from the slopes of the straight lines obtained from a graph of the measured reversion frequency over the pool bias applied upon in vitro replication of ΦX174*am*16 DNA (6). These error frequencies are summarized in Table III.

DISCUSSION

A prerequisite for the measurement of the *in vitro* accuracy of HIV RT is the accessibility of sufficient amounts of purified enzyme. Therefore, we have expressed and purified to near homogeneity an enzymatically active HIV-1 RT in *E. coli.* The expression cloned RT was obtained in a nearly homogeneous form and shown to be similar, in its structure and its enzymatic properties, to earlier preparations from either virus particles (18,19) or other expression cloned preparations (20-25). We demonstrate here for the first time that HIV RT is also active with single-stranded natural DNA as a template. A nearly complete replication of singly-primed bacteriophage DNA was achieved *in vitro.* This might suggest that HIV RT can also participate in processes of cellular DNA metabolism, such as DNA replication, DNA repair and DNA recombination.

The DNA synthesizing capability of HIV RT allowed the measurement of fidelities with natural DNA as a template. The nucleotide substitution rate at the amber 16 codon of bacteriophage ΦX174 was determined to be between 1/5,000 to 1/7,000. This is in good agreement with the nucleotide substitution rate determined *in vivo* for DNA replication at a different amber locus by avian myeloblastosis virus RT (1/5,000) (26) and *in vitro* (1/17,000) by using a different reversion system (Φ*X174am3*) (27). During the preparation of this manuscript an accuracy of 1/4,000 was reported for HIV RT by using the Φ*X174am3* reversion assay (28). Furthermore, an error rate of 1/18,000 was found with an M13-based amber codon reversion assay (29). A fidelity of around 1/10,000 is characteristic for polymerases that lack a proofreading function (30). Since HIV RT is devoid of a proofreading exonuclease (29), the measured accuracy was expected. It is noteworthy to point out that the observed fidelity at the amber 16 codon is not exceptionally low.

The Φ*Xam16* assay utilized for this study allows not only the measurement of an overall accuracy at a certain codon but also the characterization of seven individual mutational pathways. Such an analysis revealed that the error rates strongly depend on the base pairing under consideration. dGMP:$T_{template}$ mispairs were most prominent with a frequency of 1/8,000. This was also expected since theoretical considerations and model building studies predict a prevalence of dGMP:$T_{template}$ mispairs (31). Furthermore, DNA polymerases without a 3′→5′exonuclease form dGMP:$T_{template}$ mispairs with the highest frequency (6,32-34). Also not quite unexpected was the finding that dAMP:$G_{template}$ and dGMP:$A_{template}$ mispairs, which cause transversion mutations, were second most prominent with relative frequencies of around

1/40,000. Somewhat surprising was the relatively high frequency of 1/73,000 for the pyrimidine:pyrimidine mispair dCMP:$T_{template}$, which is in contrast to predictions based on models for the occurrence of mispairs (31). However, the high frequency of dCMP:$T_{template}$ mispairs is not without precedence because an even 24-fold higher error frequency (1/3,000) has been observed for purified chicken DNA polymerase β (34). dCMP:$A_{template}$ mispairs were formed by HIV RT with a frequency of 1/140,000. AMV RT formed the same type of mispair at the amber 3 locus of ΦX174 DNA with a frequency of 1/17,000 (27). With poly(rA) as a template, frequencies between 1/17,000 to 1/38,000 were observed for the dCMP:$A_{template}$ mispair with different HIV RT isolates (35). This is 4- to 8-fold higher than our measurement at the amber 16 locus. However, higher error rates are frequently observed with synthetic homopolymer templates for many DNA polymerases (30). Therefore, the higher error rate of HIV RT with a poly(rA) template does not necessarily mean that HIV RT copies RNA with a significantly lower accuracy than DNA templates.

The mutation rate of RNA genomes has been estimated to be more than a million times greater than that of DNA genomes (36). The HIV genome seems to display an even higher variability, whereby the amino-terminal part of the *gag*- protein and the entire envelope gene contain hypervariable regions (4, 37, 38). A recent computer-assisted comparison of the genomes from nine different retroviruses revealed that recombinational events contribute to the hypervariability of the *env*-gene (39). On the other hand, the observation of a prevalence of point mutations over frame shift mutations (4, 38) suggests an inaccurate DNA polymerase being responsible for most of the mutations. In fact, in a direct comparison of fidelities of three different reverse transcriptases, HIV RT was demonstrated to be threefold less accurate than RTs from other retroviruses (35). After finishing this work an accuracy of 1/1,700 was reported for HIV RT by using a forward mutation assay. This was tenfold lower than the simultaneously studied accuracies of RTs from avian myeloblastosis virus and from Moloney murine leukemia virus (29). By contrast, in the same study an accuracy of 1/18,000 was measured for HIV RT and the other two retroviral RTs by using an M13-based reversion assay. The low accuracy with the forward mutation assay was shown to be mainly due to a preponderance of G:T mispairs at certain defined locations, leading to misinsertion rates of as high as 1/70. Furthermore, frameshift mutations occured with frequencies of up to 1/250 (29). The observed discrepancies in error rates measured with several reversion assays (this study,

(28, 29)) and the forward mutation assay point to a remarkable nucleotide sequence dependence of the replication fidelity of HIV RT. The molecular basis for this is not yet understood.

ACKNOWLEDGEMENTS

We are grateful to F. Benseler for preparing the oligonucleotides. We thank F. Eckstein, M. Kruhoeffer and D. Gauss for many helpful suggestions and critically reading the manuscript.

REFERENCES

1. Barré-Sinoussi, F. Cherman, J. C., Rey, R., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C. Vezinet-Brun, F., Rouzioux, C, Rozenbaum, W. and Montagnier, L. (1983) Science 220, 868-871.
2. Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D. L., Sidhu, G., Stahl, R. E., Zolla-Pazner, S., Leibowitch, J. and Popovic, M. (1983) Science 220, 865-867.
3. Battula, N. and Loeb, L. A. (1974) J. Biol. Chem. 249, 4086-4093.
4. Starcich, B. R., Hahn, B. H., Shaw, G. M., McNeely, P. D., Modrow, S., Wolf, H., Parks, E. S., Parks, W. P., Josephs, S. F., Gallo, R. C. and Wong-Staal, F. (1986) Cell 45, 637-648.
5. Schottissek, S. and Grosse, F. (1988) Gene 62, 55-64.
6. Grosse, F., Krauss, G., Knill-Jones, J. W. and Fersht, A. R. (1983) EMBO J. 2, 1515-1519.
7. Maniatis, T. Fritsch, E. F. and Sambrook, J. (1982) in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
8. Grosse, F. and Krauss, G. (1980) Nucleic Acids Res. 8, 5703-5714.
9. Nasheuer, H.-P. and Grosse, F. (1987) Biochemistry 26, 8458-8466.
10. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitwhoen, E. A., Baumeister, K., Ivanoff, L. Petteway, S. R., Pearson, , M. L., Lautenberger, J. A., Papas, T. S. Ghrayeb, J., Chang, N. T., Gallo, R. C. and Wong-Staal, F. (1985) Nature 313, 277-284.
11. Dente, L., Sollazzo, M., Baldari, C., Cesarini, G. and Cortese, R. (1985) in: DNA cloning, Vol. 1, (Glover, D. M., ed.), pp101-107, IRL Press Oxford.
12. Weber, J. (1987) Diplom-thesis, Universität Göttingen.
13. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchinson, C. A., III, Slocombe, P. M., Smith, M., Drouin, J., Friedman, T. and Smith, A. J. H. (1978) in Denhardt, D. T., Dressler, D. and Ray, D. S. (eds.), The Single Strand Phages, Cold Spring Harbor Press, NY, pp. 659-669.
14. Fersht, A. R. and Knill-Jones, J. W. (1981) Proc. Natl. Acad. Sci. USA, 78, 4251-4255.
15. Fersht, A. R. and Knill-Jones, J. W. (1983) J. Mol. Biol., 165, 633-654.
16. Hussy, P., Maass, G., Tümmler, B. Grosse, F. and Schomburg, U. (1986) Antimicrob. Agents Chemother., 29, 1073-1078.
17. Hansen, J., Schulze, T., Mellert, W. and Moelling, K. (1988) EMBO J. 7, 239-243.
18. Di Marzo Veronese, F., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C. and Sarngadharan, M. G. (1986) Science 231, 1289-1291.
19. Wondrak, E. M., Löwer, J. and Kurth, R. (1986) J. gen Virol. 67, 2791-2797.

20. Tanese, N., Sodroski, J., Haseltine, W. and Goff, S. P. (1986) J. Virol. 59, 743-745.
21. Farmerie, W. G., Loeb, D. D., Casavant, N. C., Hutchinson III, C. A., Edgell, M. H. and Swanstrom, R. (1987) Science 236, 305-308.
22. Hansen, J., Schulze, T. and Moelling, K. (1987) J. Biol. Chem. 262, 12393-12396.
23. Hizi, A., McGill, C. and Hughes, S. H. (1988) Proc. Natl. Acad. Sci. USA, 85, 1218-1222.
24. Larder, B., Purifoy, D., Powell, K. and Darby, G. (1987) EMBO J. 6, 3133-3137.
25. Le Grice, S. F. J., Beuck, V. and Mous, J. (1987) Gene 55, 95-103.
26. Dougherty, J. P. and Temin, H. M. (1988) J. Virol. 62, 2817-2822.
27. Kunkel, T. A., Eckstein, F., Mildvan, A. S., Koplitz, R. M. and Loeb, L. A. (1981) Proc. Natl. Acad. Sci. USA, 78, 6734-6738.
28. Preston, B. D., Poiesz, B. J. and Loeb, L. A. (1988) Science 242, 1168-1171.
29. Roberts, J. D., Bebenek, K. and Kunkel, T. A. (1988) Science 242, 1171-1173.
30. Loeb, L. A. and Kunkel, T. A. (1982) Annu. Rev. Biochem. 51, 429-457.
31. Topal, M. D. and Fresco, J. R. (1976) Nature, 263, 285-289.
32. Kunkel, T. A. and Loeb, L. A. (1981) Science, 213, 765-767.
33. Rienitz, A., Grosse, F. , Blöcker, H., Frank, R. and Krauss, G. (1985) Nucleic Acids Res. 13, 5685-5695.
34. Kunkel, T. A. and Alexander, P. S. (1986) J. Biol. Chem. 261, 160-166.
35. Takeuchi, Y., Nagumo, T. and Hoshino, H. (1988) J. Virol. 62, 3900-3902.
36. Gojobori, T. and Yokoyoma, S. (1985) Proc. Natl. Acad. Sci. 82, 4198-4201.
37. Li, W.-H., Tanimura, M. and Sharp, P. M. (1988) Mol. Biol. Evol. 5, 313-330.
38. Hahn, B. H., Shaw, G. M., Taylor, M. E., Redfield, R. R., Markham, P. D., Salahuddin, S. Z., Wong-Staal, F., Gallo, R. C., Parks, E. S. and Parks, W. P. (1986) Science 232, 1548-1553.
39. McClure, M. A., Johnson, M. S., Feng, D.-F. and Doolittle, R. F. (1988) Proc. Natl. Acad. Sci. 85, 2469-2473.
40. Laemmli, U. K. (1970) Nature, 227, 680-685.