Cassette Mutagenesis of the Reverse Transcriptase of Human Immunodeficiency Virus Type 1

PAUL L. BOYER, ANDREA L. FERRIS, AND STEPHEN H. HUGHES*

*ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center,
P.O. Box B, Frederick, Maryland 21702-1201*

Received 19 June 1991/Accepted 22 October 1991

We constructed a series of *Bsp*MI cassettes that simplify the introduction of specific point mutations in the polymerase domain of human immunodeficiency virus type 1 reverse transcriptase. A series of point mutants were constructed by using these cassette vectors. The RNA-dependent DNA polymerase and RNase H activities of 20 point mutations in the conserved portion of the polymerase domain were assayed. All the mutations analyzed are conservative substitutions of evolutionarily conserved amino acids. The mutations were divided into four classes. The first class has little effect on either polymerase or RNase H activity. The second class affects RNase H but not polymerase activity, while the third class has a normal RNase H activity with diminished polymerase activity. The fourth class affects both activities.

Reverse transcriptase (RT) is one of the key enzymes in the life cycle of retroviruses and converts the single-stranded viral RNA genome into double-stranded DNA (42, 43). Since the human immunodeficiency virus (HIV) is the etiological agent for AIDS, HIV RT has been intensively studied. Drugs such as zidovudine are effective against AIDS because they specifically inhibit HIV RT. RTs are bifunctional, containing a DNA polymerase activity that can copy either RNA or DNA templates and an RNase H activity. The RT in HIV type 1 (HIV-1) virions has two protein subunits. The larger polypeptide has an apparent molecular mass of 66 kDa and contains both the RNA-dependent DNA polymerase and the RNase H activities. The second subunit is 51 kDa and is a proteolytic cleavage product of the 66-kDa protein or of a larger precursor (6, 8, 11, 19, 21). The function of the 51-kDa form remains obscure. The 51-kDa protein should contain most (or all) of the RNA-dependent DNA polymerase domain, but relatively low levels of RNA-dependent DNA polymerase activity have been observed when this protein is expressed in bacteria (11, 13, 29, 40). When the 51-kDa form of HIV-1 RT is fused to TrpE, it has been reported to have relatively high levels of polymerase activity; however, it is possible that the TrpE sequences stabilize the structure of the fusion protein (30). There is a report that a 51-kDa form of HIV-1 RT has a high level of activity in the *in situ* gel assay when it is expressed by a recombinant baculovirus (14). However, using the same type of *in situ* gel assay, two laboratories have reported that the 51-kDa form from virions has no detectable polymerase activity (8, 37). The 66-kDa form of HIV-1 RT can exist as both a monomer and a dimer. The dimeric form of the 66-kDa protein appears to have a higher level of RNA-dependent DNA polymerase activity than does the monomeric form (29, 31). The 66- and 51-kDa proteins can also interact to form a heterodimer complex with a high level of RNA-dependent DNA polymerase activity (23, 27, 31).

We have previously described plasmids that cause the expression of the 66- and 51-kDa forms of the HIV-1 RT enzyme in *Escherichia coli* (11, 15). These plasmids have been designated HIV-1 RT and CT-133, respectively. The

HIV-1 RT plasmid generates a 66-kDa protein that is soluble in *E. coli* extracts and is indistinguishable from viral RT in RT assays (11). The 51-kDa RT protein produced from the plasmid CT-133 has a carboxy-terminal end similar, but not identical, to the carboxy-terminal end of the 51-kDa form of RT arising from proteolytic cleavage of the 66-kDa protein in virions and has a low level of DNA polymerase activity (11, 15). The HIV-1 RT plasmid that induces the synthesis of the 66-kDa form of RT was modified to produce a series of *Bsp*MI cassettes within the DNA polymerase domain. The cassettes described here are logically similar to *Bsp*MI cassettes constructed for other genes (27, 38, 41). The cassettes take advantage of the properties of the restriction endonuclease *Bsp*MI, which cleaves DNA a discrete distance outside its recognition sequence and leaves a 4-bp protruding end (Fig. 1B). The HIV-1 RT cassettes contain two *Bsp*MI recognition sequences oriented in opposite directions. After *Bsp*MI digestion, a synthetic DNA fragment encoding the desired mutation can be ligated into the cassette. The cassettes described here allow any of the amino acids from amino acid (aa) 24 (Trp) to aa 223 (Lys) within the DNA polymerase domain of HIV-1 RT to be changed or removed. This region was chosen for study on the basis of evolutionary conservation (1, 26) and on the analysis of deletion, insertion, and point mutants created in our laboratory and by others (8, 17, 18, 23, 30). These studies indicated that the region of HIV-1 RT between three consecutive lysines (aa 64, 65, and 66) and two adjacent aspartic acids (aa 185 and 186) has several areas critical for the proper function of RT.

MATERIALS AND METHODS

Construction of *Bsp*MI cassettes. The *Bsp*MI restriction endonuclease sites were introduced into the HIV-1 RT coding region by the polymerase chain reaction (PCR) amplification (Fig. 1A), using the expression plasmid HIV-1 RT (11, 15) as the PCR substrate. The first PCR amplification used one primer that was complementary to sequences in the expression plasmid pUC12N upstream of the recognition site for the restriction endonuclease *Nco*I. The second primer contained a 6-bp G+C-rich "clamp" at the 5' end that was followed by the recognition site for the restriction endonu-

* Corresponding author.

cleave *EcoRI* and a recognition site for the restriction endonuclease *BspMI* oriented so that *BspMI* would cleave toward the *NcoI* site. The remainder of the primer was complementary to the sequence of HIV-1 RT. The product was digested with the restriction endonucleases *NcoI* and *EcoRI* and gel purified.

The second PCR amplification used a primer complementary to the sequence of pUC12N 3' of the *HindIII* recognition sequence (Fig. 1A). The second primer contained a 6-bp C+G-rich clamp at the 5' end that was followed by an *EcoRI* recognition site and a *BspMI* recognition site. This *BspMI* recognition site was oriented so that *BspMI* would cleave toward the *HindIII* site. The remainder of the primer contained sequences complementary to sequences in the HIV-1 RT coding region. The PCR product was digested with the restriction endonucleases *EcoRI* and *HindIII* and gel purified. The two PCR products were coligated into *NcoI*-*HindIII*-digested pUC12N to create a *BspMI* cassette. The cassette differs from the parental HIV-1 RT plasmid in that 100 bp of HIV-1 RT sequence were deleted. The deleted sequences were replaced with an *EcoRI* recognition site flanked by two *BspMI* recognition sites oriented in opposite directions (Fig. 1B).

Construction of HIV-1 RT mutations. The synthetic DNA fragments to be inserted into the *BspMI* cassettes were generated by hybridizing two complementary oligomers. The two oligomers to be hybridized were mixed and kinased by standard techniques (25). The oligomers were treated with polynucleotide kinase for 1 h at 37°C, and then sodium chloride was added to a final concentration of 50 mM and the oligomer mixture was heated to >90°C. The oligomers were allowed to hybridize by slow cooling to room temperature to give the synthetic DNA fragment.

BspMI cassette plasmids were digested with the restriction endonuclease *BspMI* and gel purified (Fig. 1C). As described in the Results, the 100-bp deletion in the cassettes was replaced with two synthetic DNA fragments. The two synthetic DNA fragments were ligated to the linearized cassette, and the resulting plasmid was used to transform *E. coli* DH5 α . Individual clones were examined with restriction endonucleases to verify that their overall structure was correct. Clones with the correct structure that expressed the RT protein were analyzed by dideoxynucleotide sequencing to verify the presence of the desired mutation.

Polyacrylamide gel electrophoresis. Bacteria containing the normal HIV-1 RT plasmid, mutant HIV-1 RT plasmids, or pUC12N were grown in NZY broth (25) supplemented with 100 μ g of ampicillin per ml for 12 to 16 h with shaking at 37°C. The plasmid pUC12N is a derivative of the expression plasmid pUC12N and contains an M13 origin of replication inserted into the *NdeI* site of pUC12N. The bacteria in 1 ml of each culture were collected by centrifugation, washed once (20 mM Tris [pH 7.5], 1 mM EDTA, 100 mM NaCl), and suspended in 0.2 ml of modified Laemmli sample

buffer (50 mM Tris [pH 7.5], 2.0% sodium dodecyl sulfate [SDS], 5.0% β -mercaptoethanol, 5.0% sucrose, 0.005% bromophenol blue). Samples were boiled for 4 min, and 30 μ l of extract was fractionated on a 9.0% SDS-polyacrylamide gel. The protein bands were visualized by staining with Coomassie brilliant blue.

RT assays. Bacteria containing the HIV-1 RT plasmid, mutant HIV-1 RT plasmids, or the pUC12N plasmid were grown and collected as described above. The RNA-dependent DNA polymerase assay has been described previously (11). Briefly, the bacterial pellets were washed and then resuspended in 0.25 ml of disruption buffer (0.2 ml NaCl, 20% glycerol, 1.0% Triton X-100, 1 mM EDTA, 2 mM dithiothreitol, 25 mM Tris, pH 8.0) and incubated at 4°C for 30 min. Insoluble debris was removed by sedimentation. The supernatants were stored at -20°C. The bacterial extract (20 μ l) was incubated in a 100- μ l reaction mixture [25 mM Tris (pH 8.0), 75 mM KCl, 8 mM MgCl₂, 2 mM dithiothreitol, 0.0005 units of poly(rC) · oligo(dG), 0.01 mM dGTP, 0.01 mM [α -³²P]dGTP (specific activity, 800 Ci/mmol)] for 30 min at 37°C. The assays were stopped by adding 0.2 ml of 0.2 M sodium PPi and 0.1 ml of 10 mg of denatured and sheared salmon sperm DNA per ml as a carrier. The labeled polymer was precipitated by the addition of 3.0 ml of ice-cold 10% trichloroacetic acid, collected on Whatman glass GF/C filters by suction filtration, and counted. We have validated this assay in reconstruction experiments in which purified HIV-1 RT (4) has been mixed with extracts of bacteria that do not express HIV-1 RT and extracts of bacteria expressing HIV-1 RT mutants that lack polymerase activity. Although the extraction buffer is not optimal for polymerase activity, the presence of the bacterial extract has no effect on the level of polymerase activity measured with a poly(rC) · oligo(dG) template-primer.

The 66-kDa/51-kDa heterodimers were assayed in a similar manner, except that 10 μ l of 66-kDa extract was mixed with 10 μ l of 51-kDa extract. Control extracts were generated by mixing 10 μ l of the 66-kDa extract with 10 μ l of extract from bacteria containing pUC12N. The extracts were incubated at 4°C for 30 min to allow heterodimer formation (4a), and then the extracts were assayed as described above.

In situ DNA-dependent DNA polymerase assay. The in situ DNA-dependent DNA polymerase assay was modified from the protocol of Spanos et al. (32, 35). Salmon sperm DNA was prepared as the polymerase substrate by digestion with DNase I. Bacteria were disrupted in SDS sample buffer containing 5 mM dithiothreitol. Following one cycle of freezing and thawing, the samples were separated on 10% SDS-polyacrylamide gels containing 50 μ g of DNase I-activated DNA substrate per ml and 2 mM EDTA. Gels were soaked in several changes of 50 mM Tris (pH 8)–1 mM EDTA at room temperature for 24 to 36 h to remove the SDS and allow the proteins to renature. Gels were preincubated in 25 mM Tris (pH 8)–8 mM MgCl₂–40 mM KCl for 12 h at

FIG. 1. (A) Construction of the *BspMI* cassettes. As described in Materials and Methods, PCR amplifications were used to delete 100 bp of the HIV-1 RT coding region and to replace the deleted segment with an *EcoRI* and two *BspMI* recognition sequences. The *BspMI* recognition sequences are oriented in opposite directions. The open boxes represent the vector pUC12N. The solid box represents the HIV-1 RT coding region. Abbreviations: B, *BspMI*; R, *EcoRI*; H, *HindIII*; N, *NcoI*. (B) Sequence across the *BspMI* recognition sites. The *BspMI* recognition sequences flank an *EcoRI* recognition site and are oriented in opposite directions. After digestion with *BspMI*, two noncomplementary 5' protruding ends are generated. The *EcoRI* recognition site (GAATTC) and the *BspMI* recognition sites (ACCTGC or GCAGGT) are shown in italics. (C) Construction of an HIV-1 RT mutant. As described in Materials and Methods, a *BspMI* cassette is digested with *BspMI*. The linearized cassette is then ligated to two double-stranded synthetic DNA fragments, one containing the desired alteration. The final clone resembles the parental HIV-1 RT clone with the exception of the mutation. Open boxes represent the pUC12N vector sequences. The solid box is the HIV-1 RT coding region. The x indicates a mutation.

FIG. 2. The six *Bsp*MI cassettes described in the Results are located in the RNA-dependent DNA polymerase domain of HIV-1 RT. The cassettes are consecutive and partially overlapping. Each cassette deletes 100 bp of the HIV-1 RT coding region (indicated by the gaps). Open boxes are pUC12N sequences, and the solid box is the HIV-1 RT coding region. Abbreviations: H, *Hind*III; N, *Nco*I.

room temperature; then 50 μ M dATP, dCTP, and dGTP and 1 μ Ci of [α - 32 P]dTTP per ml were added, and the gels were shaken overnight at 37°C. Reactions were stopped by repeated washing in 5% trichloroacetic acid–1% sodium PP_i at 4°C for at least 24 h. Gels were stained, dried, and exposed for autoradiography. Analysis of insertion and deletion mutants showed that for HIV-1 RT there is good agreement between this assay and the assays done in bacterial extracts (24).

RNase H assays. The in situ polyacrylamide gel assay for RNase H activity has been described previously (10). Briefly, the substrate for RNase H was an [α - 32 P]UTP-labeled RNA-DNA hybrid synthesized by using M13mp19 plus-strand DNA and RNA polymerase. Bacteria producing the HIV-1 RT or mutant HIV-1 RTs were lysed in sample buffer (1.0% SDS, 10% glycerol, 50 mM Tris [pH 7.0], 5 mM dithiothreitol) and fractionated on 10% SDS-polyacrylamide gels containing the [32 P]UTP-labeled RNA-DNA substrate. Following electrophoresis, the gels were incubated for at least 2 days in renaturation buffer (50 mM Tris [pH 8.0], 50 mM NaCl, 10 mM MgCl₂, 5.0 mM dithiothreitol). During this incubation, the RNase H renatured and hydrolyzed the 32 P-labeled RNA in the RNA-DNA hybrids. The gels were stained with Coomassie brilliant blue and exposed for autoradiography. RNase H activity caused clear bands to appear on an otherwise black background.

RESULTS AND DISCUSSION

Mutational analysis of HIV-1 RT. Six *Bsp*MI cassettes were constructed that simplify mutagenesis of the RNA-dependent DNA polymerase domain of the HIV-1 RT enzyme (Fig. 1A and 2). The cassettes are consecutive and partially overlap each other. Each *Bsp*MI cassette deletes 100 bp of HIV-1 RT coding sequence (Fig. 1A and 2). The six

cassettes allow the codons for any amino acid or combination of amino acids from aa 24 (Trp) to aa 223 (Lys) to be changed or removed. The section deleted within the *Bsp*MI cassettes was considered too large to be bridged by one synthetic DNA fragment. Two synthetic double-stranded DNA fragments of equal length (50 bp) were used to span the deletion (Fig. 1C). While the mutations described here have only one amino acid changed, the use of two synthetic DNA fragments will facilitate the creation of various combinations of multiple mutations.

Twenty amino acids that are conserved in a wide range of retroviral RTs (1, 26) were chosen for mutagenesis (Fig. 3). In most cases, the selected amino acid was replaced with a closely related amino acid to minimize the effect on the overall structure of RT. Table 1 summarizes the amino acid substitutions created within the RNA-dependent DNA polymerase of HIV-1 RT. As shown in Fig. 3, the 20 mutants are not randomly scattered but fall into four distinct groups or clusters. These clusters were designated group I through group IV and are similar to the clusters described by Larder et al. (17, 18).

All mutations were verified by dideoxynucleotide sequencing. The RTs produced by the mutated plasmids were analyzed by fractionation on SDS-polyacrylamide gels. As shown in Fig. 4, the amount of HIV-1 RT produced in the various *E. coli* strains is similar, although the mobility of some of the mutant proteins varies from that of wild-type HIV-1 RT. There is one strain (2866/2867, Ser-156→Ala) that appears to contain slightly lower levels of HIV-1 RT. Since this mutant shows a nearly wild-type level of polymerase activity (Table 1), the difference in level of the protein appears to be relatively modest. The RNA-dependent DNA polymerase and the RNase H activity levels were determined for each of the mutant forms of RT (Table 1). Loss of polymerase activity could be caused by the replacement of

FIG. 3. Protein sequence comparison between HIV-1 RT and other retroviral RTs. Only a part of the polymerase domain is shown in this figure. The amino acid numbers for HIV-1 RT are shown above the sequences. The four groups discussed in the Results are similar to the groups discussed by Larder et al. (18, 19) and are outlined by brackets (the figure is modified from Barber et al. [1] and is used with permission of the publisher, Mary Ann Liebert, Inc., New York). MMTV, mouse mammary tumor virus; VISNA, visna virus; EIAV, equine infectious anemia virus; RSV, Rous sarcoma virus; MoMLV, Moloney murine leukemia virus; HTLV-I, human T-cell leukemia virus type I; CAEV, caprine arthritis encephalitis virus; BLV, bovine leukemia virus; MPMV, Mason-Pfizer monkey virus.

an amino acid essential for catalytic activity, by distortion of the overall structure of RT, or by some combination of these mechanisms. Assaying both the polymerase and the RNase H functions of RT allows the mutants to be characterized more fully since any mutation in the polymerase domain that grossly distorts the overall structure of RT could also affect the RNase H domain. The mutants were divided into four classes based on their levels of enzymatic activity. The first class of mutants had normal levels of polymerase and RNase H activity; the second class had a significantly decreased RNase H activity but retained relatively high levels of polymerase activity; the third class had normal (or nearly normal) RNase H activity with a decreased polymerase activity; and the fourth class had both activities affected.

The first class of mutants is composed of amino acid substitutions which are functionally silent, either because the mutated amino acid was not required for RT activity or because the replacement amino acid can effectively substitute for the amino acid originally present in the protein. Further mutational analyses will be necessary to distinguish between these two possibilities. The second class of mutants includes those amino acid substitutions that retain a relatively high level of polymerase activity but have a diminished RNase H activity. Although these mutations are in the polymerase domain, it is possible that they distort the folding of RT in a manner that specifically affects RNase H activity and not the polymerase activity. In contrast to the well-studied RT of murine leukemia virus, which has polymerase and RNase H domains that are almost completely separate genetically (16, 22, 39), the polymerase and RNase H domains of both HIV-1 and HIV-2 RT appear to interact. Analysis of deletion and insertion mutations has shown that disruption of either the RNase H or the polymerase domain

of HIV RT usually disrupts both activities (10–12, 30). There are a small number of insertion and deletion mutations that affect only one of the two domains, and there are point mutations, which can be expected to have a less dramatic effect on the structure of the protein, that selectively affect only one domain (28, 33) (see also Table 1). Taken together, these data lead to the proposal that there are elements in both domains that are required for the proper folding of each domain. How the domains interact with each other is unclear.

Davies et al. (5) have solved the structure of an HIV-1 RT RNase H domain synthesized in *E. coli*. The HIV-1 RT RNase H domain is quite similar to the structure of *E. coli* RNase H. Moreover, the structure showed that the cleavage which converts the 66-kDa subunit of HIV-1 RT to the 51-kDa subunit of HIV-1 RT takes place not between the RNase H and polymerase domains, but within the RNase H domain. These results cast doubt on reports that a 15-kDa carboxy-terminal fragment released by cleavage with the viral protease retains significant RNase H activity (7, 8). These doubts are reinforced by a recent report that a 15-kDa fragment released when purified 66-kDa HIV-1 RT was cleaved with purified viral protease has, at most, 1/1,000th the RNase H activity of intact RT (36).

Even the inclusion of the additional amino acids necessary to make a complete HIV-1 RNase H domain does not appear to be sufficient to produce RNase H activity (2, 13). However, if the RNase H domain is combined with the 51-kDa HIV-1 RT subunit, the RNase H activity is reconstituted (5, 13). This implies that portions of the polymerase domain interact with the RNase H domain and are required for RNase H activity. Some or all of the mutations in this second

TABLE 1. Summary of the RNA-dependent DNA polymerase, DNA-dependent DNA polymerase (DDDP), and RNase H assays^a

Clone	Amino acid change	Polymerase activity (%)	In situ DDDP activity	In situ RNase H activity
HIV-1 RT	None	100	+++	+++
2852/2853	Pro-55→Gly	45	+++	—
2854/2855	Asn-57→Gln	<5	+	++
2428/2429	Phe-61→Trp	80	+++	+++
2430/2431	Ile-63→Ser	45	+++	+++
2432/2433	Lys-64→Arg	50	+++	++
2434/2435	Lys-65→Arg	<5	—	+++
2436/2437	Lys-66→Arg	90	+++	+++
2440/2441	Arg-72→Lys	45	+++	+++
2802/2803	Arg-78→Lys	50	ND ^b	++
2808/2809	Asp-110→Glu	<5	ND	+++
2810/2811	Asp-113→Glu	75	+++	++
2812/2813	Tyr-115→Phe	100	+++	++
2814/2815	Phe-116→Tyr	100	ND	+++
2860/2861	Pro-150→Gly	<5	—	—
2862/2863	Gln-151→Asn	100	ND	+++
2866/2867	Ser-156→Ala	80	+++	—
2818/2819	Tyr-183→Phe	30	ND	+++
2820/2821	Met-184→Tyr	20	ND	+++
2822/2823	Asp-185→Glu	<5	ND	+++
2824/2825	Asp-186→Glu	<5	ND	+++

^a The location and the amino acid substitution is listed for each mutant. The RNA-dependent DNA activity of the mutants was assayed in extracts as described in Materials and Methods. The parental HIV-1 RT polymerase activity was set at 100%, and the results for the mutated HIV-1 RT clones were normalized to this. The in situ DNA-dependent DNA polymerase assay is described in Materials and Methods. Activity is expressed as approximately equivalent to wild type (+++), moderate (++), weak (+), or undetectable (—). The RNase H activity for the mutants was estimated from an in situ gel assay (Materials and Methods). The results are expressed as +++ for strong (equal to wild-type) activity, ++ for moderate activity, + for weak activity, and — for no activity detected.

^b ND, not determined.

class may have altered amino acids required for interactions between the two domains of HIV-1 RT.

The third class consists of mutants with profoundly diminished polymerase activity and unaffected RNase H activity. This class consists of mutants in which an amino acid essential for polymerase activity has been replaced. The replacement amino acid appears not to distort the overall structure of RT since the RNase H activity is retained. The fourth class of mutants has lost most or all of the polymerase and the RNase H activities. The most likely explanation for the loss of both activities is that the folding of RT has been

substantially altered or distorted by the amino acid substitution.

As discussed above, the mutations were made in four distinct regions. The first region (aa 54 to 83) contains several amino acid residues that are conserved in most retroviral RTs (Fig. 3). Previous mutational studies (9) indicated that the motif IKKK centered at aa 65 in the HIV-1 RT enzyme may be of particular interest. As shown in Fig. 3, retroviral RTs seem to conserve a motif in which a hydrophobic residue (either Ile or Val) precedes at least two basic residues (usually Lys). We examined this motif by replacing each amino acid in the HIV-1 RT IKKK sequence. The requirement for a hydrophobic amino acid before the basic amino acids was tested by replacing the amino acid isoleucine with the uncharged hydrophilic amino acid Ser. This substitution decreased the polymerase activity by half but did not alter the RNase H activity. This suggests that hydrophobicity at this position is moderately important for HIV-1 RT since the introduction of an uncharged polar group reduces but does not eliminate polymerase activity. Full RNase H activity is retained, suggesting that the overall

FIG. 4. Expression of HIV-1 RT clones in *E. coli*. Bacteria containing the parental HIV-1 RT clone, a mutant HIV-1 RT clone, or the vector pUC112N were grown for 12 to 16 h at 37°C with agitation. The bacteria were collected by sedimentation in a microcentrifuge, and then lysed in 1× Spear's buffer. The protein extracts were fractionated on a 9.0% SDS-polyacrylamide gel and visualized by staining with Coomassie brilliant blue. The RT proteins expressed by nine of the mutants described in the text are shown. Bacteria containing the parental HIV-1 RT clone and bacteria containing the vector pUC112N (which does not express an RT enzyme) were included as controls. The arrow indicates the predicted size (66 kDa) of the parental HIV-1 RT protein. kd, kilodaltons.

structure of RT has not been significantly altered by this substitution.

Each lysine residue within the IKKK motif was replaced with the related basic amino acid Arg. Replacement of the Lys at aa 64 with an Arg decreased the polymerase activity by half and reduced the RNase H activity. Most retroviral RTs have lysine at this position, but both Rous sarcoma virus and bovine leukemia virus RTs have Arg (Fig. 3). The loss of activity may reflect an altered folding of RT since this mutant also has decreased RNase H activity. Lys-65 is highly conserved (Fig. 3). When this Lys is replaced with Arg, the result is an HIV-1 RT without measurable polymerase activity, but with full RNase H activity. The replacement of Lys-66 with Arg appears to have little effect on either the polymerase or RNase H activity. Only about one-half the retroviral RT enzymes listed in Fig. 3 have a Lys at this position, suggesting that this residue has not been strongly conserved. Both the evolutionary comparisons and the mutagenesis data suggest that this Lys residue is not necessary for full activity of the polymerase. The results from the polymerase and RNase H assays indicate that the IKKK motif is involved in the function of the HIV-1 RT polymerase but is probably not critical for the overall structure of RT.

Besides the IKKK motif, the region contains other important amino acids. Five amino acid substitutions (at aa 55, aa 57, aa 61, aa 72, and aa 78) yielded RTs with diminished polymerase activity. Three of these substitutions also reduced or eliminated the RNase H activity of the RT enzyme. These data suggest that the region is important for the proper folding of RT. One mutant (2852/2853, Pro-55→Gly) is of particular interest because it retained approximately one-half the original level of polymerase activity but completely lost RNase H function. It shares this characteristic with a second mutant, 2866/2867 (Ser-156→Ala). We considered the possibility that these mutants are unable to refold after the denaturation step in the RNase H assay. To investigate this possibility, we measured the DNA polymerase activity of 12 of the 20 mutants using an *in situ* gel assay. This assay, like the *in situ* RNase H assay, requires that the HIV-1 RT refolds properly after denaturation and gel electrophoresis. All the mutants we tested, including the two that showed high levels of DNA polymerase but lacked RNase H activity (Pro-55→Gly and Ser-156→Ala), showed DNA polymerase activity in the *in situ* gel assay that was similar to the level measured in extracts (Table 1). The agreement between the two polymerase assays rules out a number of simple artifacts that might arise from insolubility of the protein or from its inability to refold properly after denaturation, which strongly suggests that the differences we saw in the RNase H and DNA polymerase activities of these mutants are real.

To test whether these mutant RTs were less stable than wild-type RT, the mutants Pro-55→Gly, Ser-156→Ala, and HIV-1 RT were assayed for polymerase activity at 30, 37, and 43°C. The mutant and normal RTs had the same relative levels of polymerase activity at all three temperatures, indicating that the mutant RTs are not temperature sensitive (data not shown). The mutations may be in portions of the polymerase domain that are in contact with the RNase H domain and are required for proper RNase H folding and activity. On the basis of these results and the earlier studies (9, 30), the region between aa 54 and aa 83 appears to contain a number of amino acids essential for polymerase activity and for the overall structure of HIV-1 RT.

The second region (aa 108 to 123) contains several amino acids that may be directly involved in the polymerase function of HIV-1 RT. The two aspartic acid residues (aa 110

and 113) are conserved in a variety of retroviral RTs (Fig. 3). Substitution of Asp-110 with Glu completely abolishes polymerase activity but does not affect the RNase H function. In contrast, replacement of Asp-113 with Glu decreases the polymerase activity by 20% and slightly decreases RNase H activity. The slight decrease in RNase H activity with the aa 113 substitution may be due to a slightly altered folding of the RT. One other mutant in this region in which Tyr-115 was replaced with Phe also exhibits slightly diminished RNase H activity.

Three of the mutations in this group (Asp-110→Glu, Asp-113→Glu, and Tyr-115→Phe) are identical to mutations previously described (17, 18, 24). Although the previous reports did not include measurements of the RNase H activity for these mutants, the results we obtained for the polymerase activity are quite similar to the published values. This indicates that differences in the expression system and assay conditions, etc., do not materially affect the measurements of the polymerase activity and that we should be able to compare the rest of our data directly with what has been published.

The third region (aa 145 to 160) contains at least two amino acids that appear to be critical for the proper folding of RT. Replacement of Pro-150 with a Gly profoundly affects both the polymerase and the RNase H activities. This mutation may have distorted the overall structure of RT. Replacing Ser-156 with Ala gave a mutant with nearly normal polymerase activity but without RNase H activity. As stated previously, this effect is apparently not due to the inability of this mutant HIV-1 RT to renature in the gel since the mutant showed full DNA polymerase activity in an *in situ* assay (Table 1). We suggest that this position (156) is at or near an important contact between the DNA polymerase and RNase H domains. We also believe that this region has amino acids (Pro-150) that play a significant role in the proper folding of the entire molecule.

Larder et al. (17) have mutated Gln-151 to His, whereas we changed position 151 to Asn. They found that the substitution of His reduced polymerase activity to about one-third of the wild-type level (RNase H activity was not measured), while we found that the substitution of Asn has no detectable effect on polymerase or RNase H activity. We believe that these results can be explained by the fact that Glu and Asn are more similar than Gln and His.

The fourth region (aa 180 to 190) contains the motif YMDD (aa 183 to 186), which is conserved in RNA-dependent DNA polymerases (Fig. 3). All four of these amino acids appear to be critical for polymerase function, since the replacement of any of the four results in large decreases in activity. We made the conservative substitution of Glu for Asp at positions 185 and 186 and found that both of these mutations drastically reduce polymerase activity. These results are similar to results reported by Lowe et al. (24), who have also substituted Glu for Asp at positions 185 and 186. The substitution of Tyr-183 with Ser has been reported to nearly abolish polymerase function (17, 24). We found that a more conservative substitution (Phe) at this position leaves the RT with 30% of the polymerase activity. HIV-1 RT in which Met at aa 184 has been replaced with Leu has 5% the wild-type polymerase activity (17, 24). Substitution with Tyr yielded a mutant HIV-1 RT with 20% of wild-type level of polymerase (Table 1). RNase H activity was not measured in the previously published studies (17, 18, 24), but none of the mutations we tested affected RNase H activity. The retention of full RNase H activity suggests that this region does not play an important role in the overall struc-

ture of the protein. However, the extreme sensitivity of the polymerase to conservative substitutions in this region suggests that the YMDD motif is part of the active site.

Heterodimer activity. The mutant forms of the 66-kDa RT that lack polymerase activity were used to test for RNA-dependent DNA polymerase activity of the 51-kDa protein in a heterodimer complex. It is possible that the presence of the 66-kDa component in the heterodimer complex could induce polymerase activity in the 51-kDa component or that the presence of a normal 51-kDa component could rescue a defect in the 66-kDa component. A protein that is similar to the 51-kDa RT protein found in virions is produced by the plasmid CT-133 (11, 15). The HIV-1 RT made by cells that contain CT-133 has low levels of polymerase activity.

If the 51- and 66-kDa forms of HIV-1 RT are expressed separately and then mixed together, they can associate in vitro (3, 21, 31). We have shown that the CT-133 form of the 51-kDa subunit and the 66-kDa subunit will specifically associate when bacterial extracts containing these two forms are mixed together and that the resulting heterodimer can be purified to homogeneity (4a). Extracts containing the wild-type 66-kDa RT or one of the mutant HIV-1 RTs with low enzymatic activity (Asn-57→Gln, Lys-64→Arg, Asp-185→Glu, or Asp-186→Glu) were produced as described in Materials and Methods. To provide a baseline polymerase activity level, wild-type or mutant RT extracts were mixed with an extract from bacteria containing the parental plasmid pUC12N, which does not produce RT. The RT extracts were also mixed with the 51-kDa extract from bacteria containing the plasmid CT-133 and assayed for polymerase activity. Addition of the 51-kDa extract to the mutant HIV-1 RT extracts did not increase the level of polymerase activity relative to mixtures that contain extracts from the pUC12N control. In contrast, polymerase activity increased about 1.5-fold when extracts that contained the wild-type 66-kDa RT were combined with the 51-kDa RT (data not shown).

The four mutant HIV-1 RTs that were assayed with the 51-kDa protein may have decreased RNA-dependent DNA polymerase activity because an amino acid necessary for the polymerase activity of HIV-1 RT was replaced with an unsuitable amino acid. Alternatively, an amino acid substitution could distort the folding of the RT protein and inactivate the enzyme by causing structural changes. This distortion could affect the interaction of the mutant RT protein with the 51-kDa protein and prevent the appropriate heterodimer formation. Although two of the mutants used in this experiment, Asn-57→Gln and Lys-64→Arg, had diminished RNase H activity, two of the four mutants, Asp-185→Glu and Asp-186→Glu, had normal levels of RNase H activity, indicating that the RT proteins are not grossly distorted and should form heterodimers with the 51-kDa RT. Similar experiments involving heterodimer formation have also been done by LeGrice et al. (21). Two of the mutant forms of p66 they used, Asp-185→Asn and Asp-186→Asn, were quite similar to two of our mutants. They used metal chelation affinity chromatography to purify HIV-1 RT heterodimers that included a mutant p66 subunit and a wild-type p51 subunit. They also found that a heterodimeric HIV-1 RT with a wild-type p51 and a mutant p66 had extremely low levels of polymerase activity. Although the synthetic 51-kDa RT interacts with the 66-kDa form of RT, neither we nor LeGrice et al. (21) were able to show that the 51-kDa form makes a significant contribution to the polymerase activity of the complex. The function of the 51-kDa protein remains obscure.

ACKNOWLEDGMENTS

We are grateful to Marilyn Powers for the synthesis of the oligonucleotides and to Hilda Marusiodis for typing the manuscript.

This research was sponsored by the National Cancer Institute under contract N01-CO-74101 with ABL and by the National Institute of General Medical Sciences.

REFERENCES

1. Barber, A., A. Hizi, J. V. Maizel, Jr., and S. H. Hughes. 1990. HIV-1 reverse transcriptase: structure predictions for the polymerase domain. *AIDS Res. Hum. Retroviruses* 6:1061-1072.
2. Becerra, S. P., G. M. Clore, A. M. Gronenborn, A. R. Karlstrom, S. J. Stahl, S. H. Wilson, and P. T. Wingfield. 1990. Purification and characterization of the RNase H domain of HIV-1 reverse transcriptase expressed in recombinant *Escherichia coli*. *FEBS Lett.* 270:76-80.
3. Becerra, S. P., A. Kumar, M. S. Lewis, S. G. Widen, J. Abbotts, E. K. Karawya, S. H. Hughes, J. Shiloach, and S. H. Wilson. Protein-protein interactions of HIV-1 reverse transcriptase: implication of central and C-terminal regions in subunit binding. *Biochemistry*, in press.
4. Clark, P. K., A. L. Ferris, D. A. Miller, A. Hizi, K.-W. Kim, S. M. Deringer-Boyer, M. L. Mellini, A. D. Clark, Jr., G. F. Arnold, W. B. Leberer III, E. Arnold, G. M. Muschik, and S. H. Hughes. 1990. HIV-1 reverse transcriptase purified from a recombinant strain of *Escherichia coli*. *AIDS Res. Hum. Retroviruses* 6:753-764.
- 4a. Clark, P., and S. Hughes. Unpublished observations.
5. Davies, J. F., Z. Hostomska, Z. Hostomsky, S. R. Jordan, and D. A. Matthews. 1991. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* 252:88-95.
6. DiMarzo Veronese, F., T. D. Copeland, A. L. DeVico, R. Rahman, S. Oroszlan, R. C. Gallo, and M. G. Sarngadharan. 1986. Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. *Science* 231:1289-1291.
7. Hafkemeyer, P., E. Ferrari, J. Brecher, and U. Hubscher. 1991. The p15 carboxyl-terminal proteolysis product of the human immunodeficiency virus type 1 reverse transcriptase p66 has DNA polymerase activity. *Proc. Natl. Acad. Sci. USA* 88:5262-5266.
8. Hansen, J., T. Schulze, W. Mellert, and K. Moelling. 1988. Identification and characterization of the HIV-specific RNase H by monoclonal antibody. *EMBO J.* 7:239-243.
9. Hizi, A., A. Barber, and S. H. Hughes. 1989. Effects of small insertions on the RNA-dependent DNA polymerase of HIV-1 reverse transcriptase. *Virology* 170:326-329.
10. Hizi, A., S. H. Hughes, and M. Shaharabany. 1990. Mutational analysis of the ribonuclease H activity of human immunodeficiency virus 1 reverse transcriptase. *Virology* 175:575-580.
11. Hizi, A., C. McGill, and S. H. Hughes. 1988. Expression of soluble, enzymatically active, human immunodeficiency virus reverse transcriptase in *Escherichia coli* and analysis of mutants. *Proc. Natl. Acad. Sci. USA* 85:1218-1222.
12. Hizi, A., R. Tal, and S. H. Hughes. 1991. Mutational analysis of the DNA polymerase and ribonuclease H activities of human immunodeficiency virus type 2 reverse transcriptase expressed in *Escherichia coli*. *Virology* 180:339-346.
13. Hostomsky, Z., Z. Hostomska, G. O. Hudson, E. W. Moomaw, and B. Nides. 1991. Reconstitution *in vitro* of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase. *Proc. Natl. Acad. Sci. USA* 88:1148-1152.
14. Hu, Y.-W., and C. Y. Kang. 1991. Enzyme activities of four different forms of human immunodeficiency virus 1 *pol* gene products. *Proc. Natl. Acad. Sci. USA* 88:4596-4600.
15. Hughes, S. H., A. Ferris, and A. Hizi. 1990. Analysis of the reverse transcriptase of human immunodeficiency virus expressed in *Escherichia coli*, p. 297-307. In W. G. Laver and G. M. Air (ed.), *Use of X-ray crystallography in the design of antiviral agents*. Academic Press, Inc., San Diego, Calif.
16. Kotewicz, M. L., C. M. Samson, J. M. D'Alessio, and G. F. Gerard. 1988. Isolation of cloned Moloney murine leukemia

- virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* **16**:265–277.
17. Larder, B., S. Kemp, and D. J. M. Purifoy. 1989. Infectious potential of human immunodeficiency virus type 1 reverse transcriptase mutants with altered inhibitor sensitivity. *Proc. Natl. Acad. Sci. USA* **86**:4803–4807.
 18. Larder, B., D. J. M. Purifoy, K. Powell, and G. Darby. 1987. Site-specific mutagenesis of AIDS virus reverse transcriptase. *Nature (London)* **327**:716–717.
 19. Larder, B., D. J. M. Purifoy, K. Powell, and G. Darby. 1987. AIDS virus reverse transcriptase defined by high level expression in *E. coli*. *EMBO J.* **6**:3133–3137.
 20. LeGrice, S., R. Zehnle, and J. Mous. 1988. A single 66-kilodalton polypeptide processed from the human immunodeficiency virus type 2 *pol* polyprotein in *Escherichia coli* displays reverse transcriptase activity. *J. Virol.* **62**:2525–2529.
 21. LeGrice, S. F. J., T. Naas, B. Wohlgensinger, and O. Schatz. 1991. Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase. *EMBO J.*, in press.
 22. Levin, J. G., R. J. Crouch, K. Post, S. C. Hu, D. McKelvin, M. Zweig, D. L. Court, and B. I. Gerwin. 1988. Functional organization of the murine leukemia virus reverse transcriptase: characterization of a bacterially expressed AKR DNA polymerase deficient in RNase H activity. *J. Virol.* **62**:4376–4380.
 23. Lowe, D., A. Atiken, C. Bradley, G. Darby, B. Larder, K. Powell, D. Purifoy, M. Tisdale, and D. Stammers. 1988. HIV-1 reverse transcriptase: crystallization and analysis of domain structure by limited proteolysis. *Biochemistry* **27**:8884–8889.
 24. Lowe, D. M., V. Parmor, S. D. Kemp, and B. A. Larder. 1991. Mutational analysis of two conserved sequence motifs in HIV-1 reverse transcriptase. *FEBS Lett.* **282**:231–234.
 25. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 26. McClure, M. A., M. S. Johnson, D.-F. Feng, and R. F. Doolittle. 1988. Sequence comparisons of retroviral proteins: relative rates of change and general phylogeny. *Proc. Natl. Acad. Sci. USA* **85**:2469–2474.
 27. Mitra, G., G. Pauly, R. Kumar, G. Pei, S. H. Hughes, R. Moschel, and M. Barbacid. 1989. Molecular analysis of O⁶-substituted guanine-induced mutagenesis of *ras* oncogenes. *Proc. Natl. Acad. Sci. USA* **86**:8650–8654.
 28. Mizraki, V., M. T. Usdin, A. Harington, and L. R. Dudding. 1990. Site-directed mutagenesis of the conserved Asp443 and Asp498 carboxy-terminal residues of HIV-1 reverse transcriptase. *Nucleic Acids Res.* **18**:5359–5363.
 29. Muller, B., T. Restle, S. Weiss, M. Gautel, G. Sczakiel, and R. Goody. 1989. Co-expression of the subunits of the heterodimer of HIV-1 reverse transcriptase in *E. coli*. *J. Biol. Chem.* **264**:13975–13978.
 30. Prasad, V., and S. P. Goff. 1989. Linker insertion mutagenesis of the human immunodeficiency virus reverse transcriptase expressed in bacteria: definition of the minimal polymerase domain. *Proc. Natl. Acad. Sci. USA* **86**:3104–3108.
 31. Restle, T., B. Muller, and R. S. Goody. 1990. Dimerization of human immunodeficiency virus type 1 reverse transcriptase. *J. Biol. Chem.* **265**:8986–8988.
 32. Rucheton, M., M. N. Lelay, and P. Jeanteur. 1979. Evidence from direct visualization after denaturing gel electrophoresis that RNase H is associated with MSV-MuLV reverse transcriptase. *Virology* **97**:221–223.
 33. Schatz, O., F. V. Cromme, F. Gruninger-Leitch, and S. F. J. LeGrice. 1989. Point mutations in the conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function. *FEBS Lett.* **257**:311–314.
 34. Shaharabany, M., and A. Hizi. 1991. The DNA dependent and RNA dependent polymerase activities of the reverse transcriptases of human immunodeficiency viruses types 1 and 2. *AIDS Res. Hum. Retroviruses* **7**:883–888.
 35. Spanos, A., S. G. Sedgwick, G. T. Yarranton, U. Hubscher, and G. R. Banks. 1981. Detection of the catalytic activities of DNA polymerases and their associated exonucleases following SDS-polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:1825–1839.
 36. Stammers, D. K., M. Tisdale, S. Court, V. Parmar, C. Bradley, and C. K. Ross. 1991. Rapid purification and characterization of HIV-1 reverse transcriptase and RNase H engineered to incorporate a C-terminal tripeptide α -tubulin epitope. *FEBS Lett.* **283**:298–302.
 37. Starnes, M. C., W. Gao, R. Y. C. Ting, and Y.-C. Cheng. 1988. Enzyme activity gel analysis of human immunodeficiency virus reverse transcriptase. *J. Biol. Chem.* **263**:5132–5134.
 38. Stone, J. C., W. C. Vass, B. M. Willumsen, and D. R. Lowy. 1988. p21-*ras* effector domain mutants constructed by "cassette" mutagenesis. *Mol. Cell. Biol.* **8**:3565–3569.
 39. Tanese, N., and S. P. Goff. 1988. Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analyses and separate expression of the DNA polymerase and RNase H activities. *Proc. Natl. Acad. Sci. USA* **85**:1777–1781.
 40. Tisdale, M., P. Ertl, B. Larder, D. Darby, and K. Powell. 1988. Characterization of human immunodeficiency virus type 1 reverse transcriptase by using monoclonal antibodies: role of the C terminus in antibody reactivity. *J. Virol.* **62**:3662–3667.
 41. Tomic, M., I. Sunjevaric, E. S. Savtchenko, and M. Blumenberg. 1990. A rapid and simple method for introducing specific mutations into any position of DNA leaving all other positions unaltered. *Nucleic Acids Res.* **18**:1656.
 42. Varmus, H. 1988. *Retroviruses*. Science **240**:1427–1428.
 43. Weiss, R., N. Teich, H. Varmus, and J. Coffin (ed.). 1985. *RNA tumor viruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.