VIROLOGY 175, 456–464 (1990)

Immunologic and Proteolytic Analysis of HIV-1 Reverse Transcriptase Structure[1]

ANDREA L. FERRIS,* AMNON HIZI,† STEPHEN D. SHOWALTER,‡ SERGIO PICHUANTES,§ LILIA BABE,§ CHARLES S. CRAIK,§ AND STEPHEN H. HUGHES*,[2]

*BRI-Basic Research Program, ‡Program Resources, Inc., NCI-Frederick Cancer Research Facility, P.O. Box B, Frederick, Maryland 21701; †Sackler School of Medicine, Tel Aviv University, Ramat Aviv, Israel; and §University of California, San Francisco, San Francisco, California 94143

Received July 28, 1989; accepted December 1, 1989

HIV-1 virions contain two reverse transcriptase polypeptides that have apparent molecular weights of 66 and 51 kDa. The 51-kDa form lacks the carboxy-terminal sequences found in the 66-kDa form, and is believed to be a proteolytic digestion product. We have treated purified 66-kDa reverse transcriptase with viral and nonviral proteases. The digestion products were characterized by their ability to react with monoclonal antibodies known to recognize particular segments of the HIV-1 reverse transcriptase. The approximate location of the segments recognized by the monoclonal antibodies was determined by testing the ability of the antibodies to recognize a series of amino- and carboxy-terminal-deleted forms of HIV-1 reverse transcriptase. The segments recognized are not uniformly distributed along the primary amino acid sequence of HIV-1 reverse transcriptase. We suggest that these segments are probably on the surface of the properly folded form of reverse transcriptase. Of the tested proteases, only the viral protease was able to cleave the 66-kDa form to the 51-kDa form without producing additional cleavage products, suggesting that the viral protease cleaves the 66-kDa protein to the 51-kDa form in virions.

INTRODUCTION

HIV-1 virions contain two reverse transcriptase (RT) polypeptides that have apparent molecular weights of 66 and 51 kDa. (As explained under Results, we find the 51-kDa form has an apparent molecular weight of 54 kDa. We use the term 51 kDa to be consistent with the literature.) These two polypeptides have identical amino termini, but the larger form has additional carboxy-terminal sequences (Lightfoote *et al.*, 1986; Di Marzo Veronese *et al.*, 1986). The 66-kDa form has been shown to have both RNA-dependent DNA polymerase and RNase H enzymatic activities (Hansen *et al.*, 1987, 1988; Larder *et al.*, 1987; Hizi *et al.*, 1988; LeGrice *et al.*, 1988). Genetic analyses and sequence comparisons have demonstrated that the RNA-dependent DNA polymerase activity is in the amino terminus, and that the RNase H activity is in the carboxy terminus of the Moloney murine leukemia virus (MuLV) RT, and that the overall organization of HIV-1 RT is the same as the MuLV RT (Johnson *et al.*, 1986; Kotewicz *et al.*, 1988; Tanese and Goff, 1988). Although the 51-kDa form of HIV-1 RT contains almost precisely the amino acids that should form the RNA-dependent DNA polymerase domain, the function(s) of the 51-kDa form of HIV-1 RT (if any) has remained obscure. The HIV-1 RT obtained from virions has been separated into 66- and 51-kDa components only under denaturing conditions. The denatured 66-kDa protein was fractionated by electrophoresis, then refolded *in situ*, and regained RNA-dependent DNA polymerase activity. The 51-kDa form present on the same gel lacked this activity (Hansen *et al.*, 1988).

Several laboratories have reported making fully active 66-kDa forms of HIV-1 RT in bacteria (Hansen *et al.*, 1987; Larder *et al.*, 1987; Hizi *et al.*, 1988; LeGrice *et al.*, 1987, 1988; Tanese *et al.*, 1986; Farmerie *et al.*, 1987; Mous *et al.*, 1988). The bacterial expression systems were also used to express truncated proteins similar to the 51-kDa form found in virions. The 51-kDa recombinant proteins have levels of RNA-dependent DNA polymerase activity that are at most only a few percent of the activity possessed by the 66-kDa form (Hizi *et al.*, 1988; Tisdale *et al.*, 1988). It has been suggested that the 66- and 51-kDa proteins are the two subunits of a dimer, and that this heterodimer is more active in RNA-dependent DNA polymerase assays than is the pure 66-kDa form (Lowe *et al.*, 1988).

The precise nature of the processing events that produce the 51-kDa form of RT is still unclear. In virions, the *pol* gene products, including the RT proteins, are derived by proteolytic processing from a *gag–pol* polyprotein by a virally encoded protease that is also part of the polyprotein (Weiss *et al.*, 1982, 1985). It is usually assumed that the 51-kDa form of RT found in virions is

[1] The U.S. Government's right to retain a nonexclusive royalty-free license in and to the copyright covering this paper, for governmental purposes, is acknowledged.

[2] To whom reprint requests should be addressed.

0042-6822/90 $3.00

created by the action of the viral protease on the 66-kDa form of RT or on some intermediate in the processing of the *gag–pol* polyprotein. However, there is no direct evidence to show that this conjecture is correct.

We have cleaved a highly purified preparation of the 66-kDa form of HIV-1 RT isolated from a recombinant strain of *Escherichia coli* with two highly purified preparations of HIV-1 protease that derive from recombinant strains of *E. coli* and yeast (Babe *et al.*, 1989). Both of the recombinant HIV-1 proteases cleave the 66-kDa form of HIV-1 RT to yield a product that is either extremely similar or identical to the 51-kDa form isolated from virions. Some commercially available nonviral proteases make related cleavages in the 66-kDa RT protein, although none of the proteases we have tested shows as much specificity as does the viral protease.

The positions of the proteolytic cleavage products in the HIV-1 RT molecule were identified using a series of monoclonal antibodies (mAbs) that recognize specific segments of HIV-1 RT. The segments recognized by the mAbs are not uniformly distributed along the primary sequence of HIV-RT. We have proposed a model for the structure of HIV-1 RT, and find, with one possible exception, that the segments recognized by the mAbs were predicted to be on the surface of the properly folded HIV-1 RT.

MATERIALS AND METHODS

Construction of the plasmids expressing deleted forms of HIV-1 RT

The plasmid we use to express HIV-1 RT has been described (Hizi *et al.*, 1988). In brief, the ends of the region encoding HIV-1 RT were modified using synthetic DNA segments to introduce an initiation and a termination codon at the sites in the HIV-1 genome that encode the amino acids where HIV-1 RT is normally cleaved from the polyprotein precursor. This modified segment was inserted into the expression plasmid pUC12N. The HIV-1 RT made in *E. coli* has two additional amino-terminal amino acids, methionine and valine, when compared with the HIV-1 RT isolated from virions. All the deletions are named for the number of amino acids removed by the deletion. AT designates amino-terminal deletions and CT designates carboxy-terminal deletions.

Several carboxy-terminal deletion mutants (CT 8, CT 16, CT 23, and CT 133) and one amino-terminal deletion mutant (AT 23) have been described (Hizi *et al.*, 1988). Additional carboxy-terminal deletion mutants were constructed by cleaving the expression plasmid with *Asp*718 and a second restriction endonuclease that cleaves in the HIV-1 RT coding region. The ends of the DNA were filled in using the large (Klenow) fragment of *E. coli* DNA polymerase I and the ends ligated together. The second restriction endonuclease used to construct CT 58 was *Nsi*I, for CT 250 *Pfl* MI was used. In all cases, the small DNA segment encoding the carboxy terminus of HIV-1 RT between the sites for *Asp*718 and *Sal*I is out of frame and does not contribute HIV-1 RT-related amino acids. For some of these deletion mutants, the DNA sequence around the deletion was confirmed by directly sequencing the deleted plasmids.

Amino-terminal deletions were made by digesting the expression plasmid with *Nco*I and with a second restriction endonuclease that recognizes an internal site. To maintain the correct reading frame, an appropriate double-stranded synthetic DNA segment was used to join the *Nco*I recognition site to the internal site. For the mutant AT 13, the second restriction endonuclease was *Sma*I, for AT 126 it was *Acc*I, for AT 139 it was *Eco*RV, and for AT 221 it was *Bst*XI. In some cases, the synthetic oligonucleotides used to join the second site to *Nco*I encoded a short segment of HIV-1 RT sequence adjacent to the internal restriction site. The plasmids were introduced into competent DH-5 cells (Bethesda Research Labs). Clones containing deleted plasmids were identified by digesting plasmid DNAs with appropriate restriction endonucleases.

The MuLV/HIV RT chimera (MH6-2) was constructed by ligating an *Nco*I to *Eco*RV segment from the MuLV RT expression plasmid (Hizi and Hughes, 1988) to the HIV-1 RT expression plasmid completely digested with *Nco*I and partially digested with *Kpn*I. The *Eco*RV site was joined to the *Kpn*I site with synthetic oligonucleotides. The resulting chimeric protein contains 500 amino acids from MuLV RT at the amino terminus and 137 amino acids from HIV-1 RT at the carboxy terminus.

Virus

Purified virus (strains IIIB, RF, and PH) was obtained from Dr. Larry Arthur, Program Resources Inc. (Frederick, MD).

Production of mAbs

Eight-week-old BALB/c mice were inoculated intraperitoneally with purified RT (66-kDa form) in 0.5 ml of emulsion with complete Freund's adjuvant. Two weeks later, the animals were boosted with a 0.5 ml subcutaneous inoculation of the purified RT in incomplete Freund's adjuvant. The following week, blood was obtained from the animals by orbital bleed and the serum tested by ELISA to determine anti-RT activity. Two weeks following the second injection, the mice were given a final intravenous boost of 0.1 ml and the

spleens of mice showing the strongest reactivity were removed 4 days later.

The cell fusion protocol has been described (Showalter *et al.*, 1981). Briefly, spleens were minced, passed through sterile gauze into a conical 50-ml tube, and washed twice. Cells were counted and mixed at a 5:1 ratio with actively growing, washed NS-1 cells (Kohler *et al.*, 1976). The cell mixture was pelleted and gently resuspended over a 1-min period in 1 ml of 50% polyethylene glycol 4000 (Merck) per 1.6×10^8 lymphocytes. Two milliliters of RPMI-1640 containing 15% fetal bovine serum was added, followed by 10 ml of complete medium. Cells were pelleted, gently resuspended in 22 ml of complete medium, and dispensed into 96-well Costar plates (~100 μl/well). Cells were grown in RPMI containing hypoxanthine–aminopterin–thymidine (HAT; GIBCO) for 30 days, and then maintained on medium without selection.

Beginning about Day 10, media from actively growing hybridomas were tested by ELISA for anti-RT activity. Cells from positive wells were expanded to 24-well plates (Costar) and cloned by limiting dilution in 96-well plates on a feeder layer of compatible thymocytes. Cloned cells were passaged in adult BALB/c mice primed with 0.5 ml pristane (2,6,10,14-tetramethyl-petadecane; Aldrich Chemical Co.) by intraperitoneal injection of 3×10^6 cells. The ascites fluids were harvested 7–10 days after injection and clarified, and the immunoglobulin was purified for further use.

Proteolysis

Four commercially available proteases (Sigma Chemical Co., St. Louis, MO) were incubated with purified HIV-1 RT in a buffer of 25 m*M* Tris–HCl, 50 m*M* NaCl, and 8 m*M* $MgCl_2$, pH 8. To digest 60 μg of RT the following amounts of the proteases were used: 0.135 μg V8 (0.074 units of activity), 0.045 μg chymotrypsin (2.7×10^{-3} units), 0.022 μg papain (2.86×10^{-3} units), and 0.075 μg trypsin (0.75 units). Reactions were incubated at 23° for 3 or 24 hr.

For most experiments, 5 μg of purified RT was combined with 40 ng of purified viral protease in a buffer of 50 m*M* sodium phosphate, pH 6.5, 10 m*M* NaCl, 5 m*M* DTT, and 0.05% Triton X-100. The reaction mix was incubated at 37° for 0, 1, 2, 3, 4, 5, 6, and 22 hr. Purified RT was also incubated for 22 hr without added protease. For the digestion shown in Fig. 5D, 5 μg of RT was incubated at 37° for 16 hr, with and without 300 ng of purified viral protease. HIV-1 protease was purified from both *E. coli* and yeast expression strains. A description of the strains and of the purification will be published elsewhere (Babe *et al.*, 1989).

Western transfer and immunostaining

For immunodetection of RT in bacterial strains, the protein from *E. coli* grown in 40–60 μl of culture media was dissolved in SDS sample buffer and separated on polyacrylamide gels. Proteins were transferred to nitrocellulose paper using an ABN PolyBlot (American Bionetics, Hayward, CA). Blocking and washing were performed using 0.3% gelatin in 137 m*M* NaCl, 2.68 m*M* KCl, 8.1 m*M* Na_2HPO_4, and 1.5 m*M* KH_2PO_4. Transfers were probed with a 1:50 to 1:400 dilution of mAb followed by 20 μg of alkaline phosphatase-conjugated goat anti-mouse IgG (Kirkegaard and Perry, Gaithersburg, MD). Alkaline phosphatase activity was detected using the chromogenic substrates BCIP and NBT (BRL, Gaithersburg, MD).

RESULTS

Analysis of segments of HIV-1 RT recognized by the mAbs

Monoclonal antibodies were produced following the injection of mice with purified 66 kDa HIV-1 RT made by a recombinant strain of *E. coli* (Hizi *et al.*, 1988). The purification of the RT will be reported elsewhere (manuscript in preparation). Antibodies that react with HIV-1 RT were initially detected with an ELISA assay using purified HIV-1 RT. For the epitope mapping experiments, a secondary screening was performed to identify those mAbs that react in an immunoblot. mAbs that recognize recombinant HIV-1 RT also react in immunoblots with RT obtained from three different strains of HIV-1 (Fig. 1). The segments recognized by the mAbs were mapped using a series of amino-terminal and carboxy-terminal deletion mutants. The plasmids used to express the deleted versions of the HIV-1 RT all derive from a plasmid that expresses the 66-kDa form of the HIV-1 RT (Hizi *et al.*, 1988). The bacterial strains carrying plasmids with the deletion mutations produce HIV-1 RT-related proteins of the appropriate size (Fig. 2). In some cases, the specificity of the mapping was confirmed using strains of *E. coli* that express RT from MuLV and strains that express hybrid proteins composed of segments from MuLV and HIV-1 RT (Fig. 3C and data not shown). The ability of the mAbs to recognize RT proteins produced by the various strains of *E. coli* was determined by immunoblot (Figs. 3A–3C).

We have found that the mAbs recognize several different segments of the HIV-1 RT. These recognition sites are not uniformly distributed along the primary amino acid sequence of the protein (Fig. 3D). Although we have obtained mAbs that recognize sequences near the amino and carboxy termini, the majority of our monoclonals and the majority of those reported by Tis-

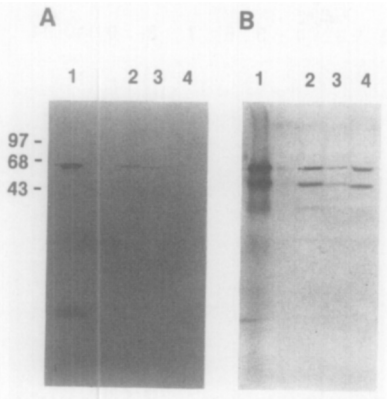
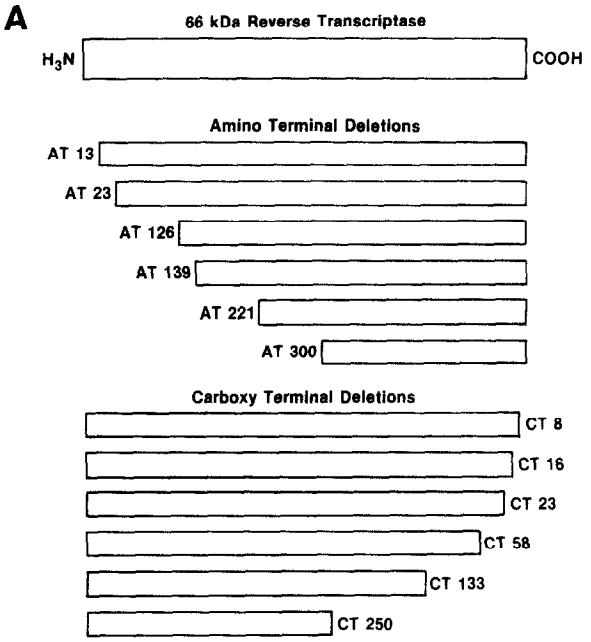


FIG. 1. Representative immunoblots showing the reaction of two mAbs made against recombinant RT with recombinant RT and with HIV particle lysates from three different strains. (A) mAb 50 binds an epitope near the carboxy terminus of the RT (Fig. 3C). The segment of RT containing this epitope is missing from the 51-kDa form. Therefore, this antibody reacts with the 66-kDa but not with the 51-kDa protein. Lanes: 1, partially purified HIV-1 RT from *Escherichia coli;* 2, disrupted virions of the III B strain; 3, disrupted virions of the RF strain; and 4, disrupted virions of the PH strain. (B) mAb 21 binds a segment near the center of the RT (Fig. 3D). This segment is present in both the 66- and 51-kDa forms. Lanes: 1, partially purified HIV-1 RT from *E. coli;* 2, disrupted III B virions; 3, disrupted RF virions; and 4, disrupted PH virions.

dale *et al.* (1988) react with a relatively small, central region of the RT. The mAbs we produced were elicited and screened using nondenatured HIV-1 RT. We suggest that the majority of our mAbs react with segments on the surface of the correctly folded 66-kDa form of HIV-1 RT (see Discussion).

The immunoblots also provide information about the RT-related proteins that accumulate in the various *E. coli* strains. Most of the expression strains produce, in addition to the expected protein, smaller and, in some cases, larger forms that are specifically recognized by some of the mAbs. Control experiments have shown that these are not *E. coli* proteins (Figs. 3A–3C). In theory the smaller than expected forms could derive from premature termination, from internal initiation of translation, or from proteolytic degradation. Analyses with a mAb that recognizes the amino terminus of the full-length HIV-1 RT protein (Fig. 3A) demonstrated that none of the smaller forms contain the sequences found at the amino terminus of the normal HIV-1 RT protein. This suggests that, at least for the full-length clone and the carboxy-terminal deletion mutants, the smaller forms are probably not the result of premature termination. The DNA sequence encoding HIV-1 RT contains several internal ATGs, but these are not associated with a consensus sequence for the initiation of protein synthesis in *E. coli* (Shine and Dalgarno, 1974). If the shorter forms derive from proteolytic cleavage, the cleavage is quite distinct from the proteolytic event that gives rise to the truncated form seen in virions and from the proteolytic events that produce a similar protein during purification of the HIV-1 RT. The smaller form of



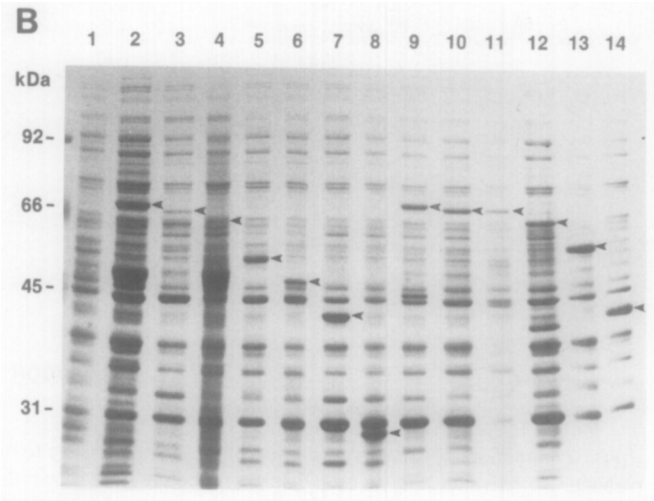


FIG. 2. The deleted forms of HIV-1 RT are shown relative to the full-length RT. (A) Mutants are named by the number of amino acids missing relative to the wild-type RT and the by end of the protein [amino terminus (AT) vs carboxy terminus (CT)] from which they were lost. Construction of the plasmids producing these deleted forms is described under Materials and Methods. The relative size of the deletions is shown. (B) Coomassie brilliant blue-stained SDS–polyacrylamide gel shows the proteins (arrows) produced by DH5 cells carrying the pUC12N plasmid with or without an RT insert. The various strains of *E. coli* were grown overnight in NZY broth (25) containing 100 μg/ml ampicillin. Cells from 40–60 μl of culture fluid were recovered by centrifugation, solubilized in SDS sample buffer, and loaded on each lane. The plasmids carried by the strains were as follows. Lanes: 1, pUC12N, no insert; 2, pUC12N with full-length RT insert; 3, AT 13; 4, AT 23; 5, AT 126; 6, AT 139; 7, AT 221; 8, AT 300; 9, CT 8; 10, CT 16; 11, CT 23; 12, CT 58; 13, CT 133; and 14, CT 250.

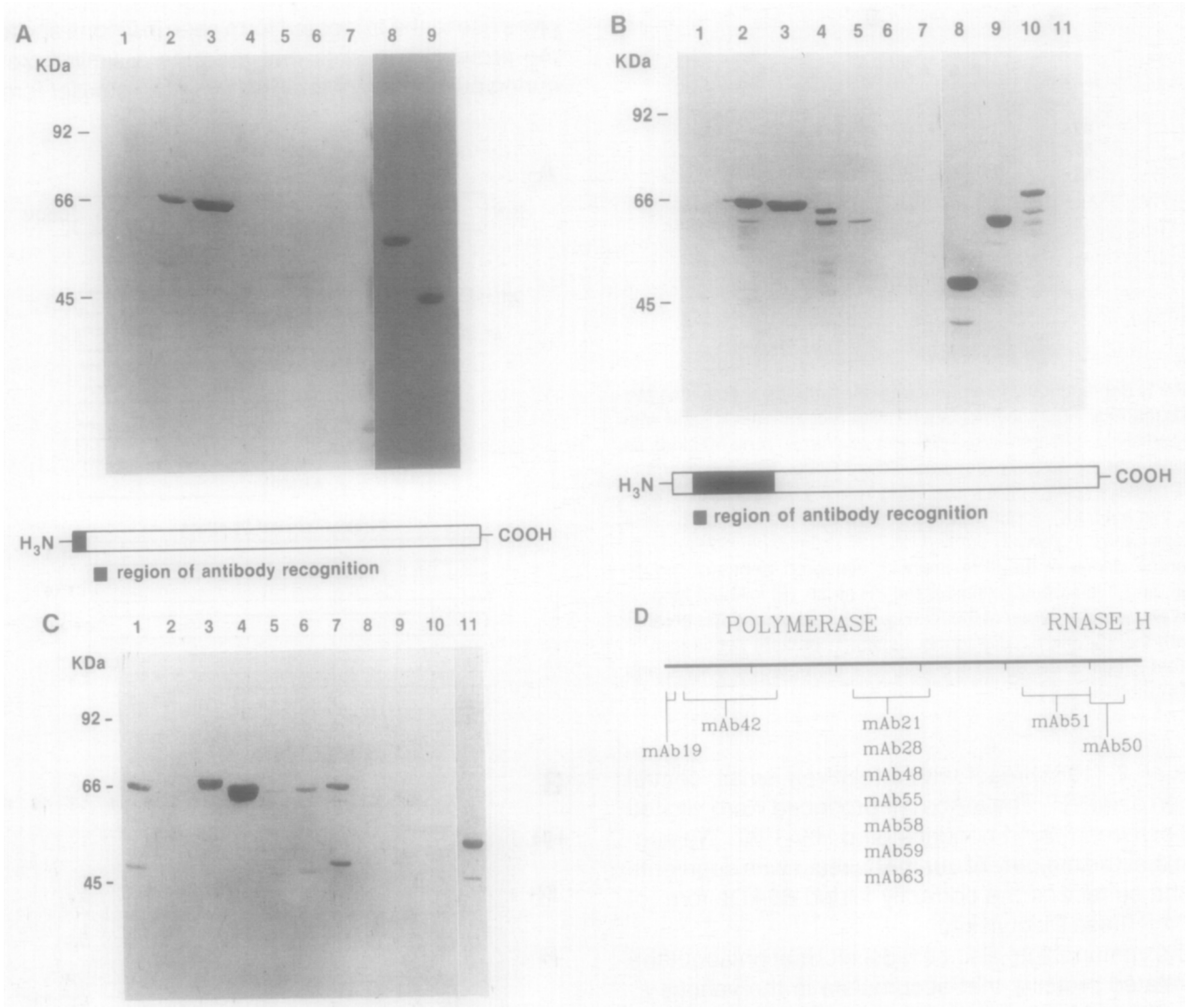


FIG. 3. Immunoblots showing the differential binding of mAbs to native and mutant forms of HIV-1 RT. *E. coli* DH5 cells carrying the pUC12N plasmid with or without RT or an RT deletion mutant were grown in NZY broth containing 100 μg/ml of ampicillin, collected by centrifugation, and solubilized in SDS sample buffer. Proteins from cells in 40–60 μl of culture fluid were loaded onto each lane of a polyacrylamide gel, separated, and transferred to nitrocellulose paper. Transfers were probed with mAbs and binding was visualized by alkaline phosphatase staining. (A) mAb 19. Lanes: 1, pUC12N, no insert; 2, pUC12N with full-length RT insert (pUC12N/RT); 3, purified RT; 4, AT 13; 5, AT 23; 6, AT 139; 7, AT 221; 8, CT 133; and 9, CT 250. (B) mAb 42. Lanes: 1, pUC12N; 2, pUC12N/RT; 3, purified RT; 4, AT 13; 5, AT 23; 6, AT 139; 7, AT 221; 8, CT 250; 9, CT 133; 10, CT 58; and 11, AT 126. (C) mAb 50. Lanes: 1, pUC12N/RT; 2, pUC12N with MuLV RT insert (19); 3, MuLV/HIV-1 RT chimera MH6-2; 4, purified RT; 5, CT 23; 6, CT 16; 7, CT 8; 8, CT 58; 9, CT 250; 10, CT 133; and 11, AT 126. (D) A map showing binding regions of 11 mAbs. The upper bar in the figure schematically depicts the 560 amino acids that constitute the 66-kDa form of RT. The marked intervals represent blocks of 100 amino acids. The brackets shown beneath the bar represent the segments bound by the various mAbs. Four antibodies (mAb 19, 42, 50, and 51) recognize distinct segments of the HIV-1 RT. The majority of the antibodies (mAb 21, 28, 48, 55, 58, 59, and 63) bind to the central portion of the RT within the RNA-dependent DNA polymerase domain.

RT in HIV-1 virions, usually called 51 kDa, has the same amino terminus as the 66-kDa protein (Lightfoote *et al.*, 1986). Although we refer to this protein as the 51-kDa form, we find that it migrates to the position expected for a protein of 54 kDa on our SDS–polyacrylamide gels. By contrast, the most prominent of the small forms that accumulate in the *E. coli* strain expressing the full-length 66-kDa HIV-1 RT migrates as a protein

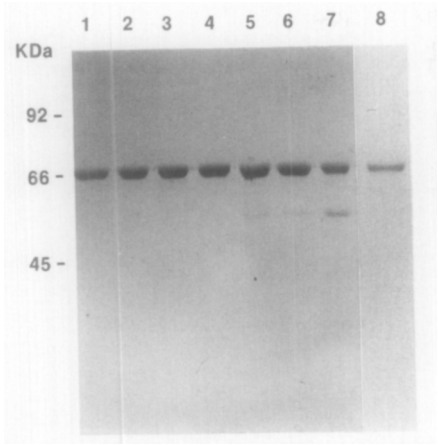


FIG. 4. Coomassie brilliant blue-stained polyacrylamide gel showing the results of incubating purified HIV-1 RT with purified HIV-1 protease at 37° over a time course of 0–22 hr. Each lane contains 2.7 μg of protein. Lanes: 1, starting material, RT, and protease with no incubation; 2, 1 hr; 3, 2 hr; 4, 3 hr; 5, 4 hr; 6, 5 hr; 7, 22 hr; and 8, control, RT incubated 22 hr at 37° without added protease.

of approximately 47 kDa, and has the same (or nearly the same) carboxy terminus as the 66-kDa protein (Fig. 3C). This 47-kDa form is missing the amino-terminal segment present in the 66-kDa form (Fig. 3A). At least some of the strains carrying deleted plasmids produce corresponding proteins that also lack the expected amino terminus.

The larger forms of HIV-1 RT could arise from readthrough. There may be readthrough from plasmid sequences upstream of the *lac*Z ATG that is the expected start for translation. Alternatively, the termination codon inserted at the end of the RT coding regions could be suppressed.

Proteolytic digestion of HIV-1 RT

Purified 66-kDa HIV-1 RT was digested with both viral and nonviral proteases. To investigate the cleavage of the RT by the viral protease, purified 66-kDa HIV-1 RT was mixed with cloned, purified HIV-1 protease (Babe *et al.*, 1989) and incubated overnight at 37°. Samples were taken at various times during the incubation. Aliquots of these samples were fractionated on SDS–polyacrylamide gels. The 66-kDa protein and a cleavage product with an apparent molecular weight of 54 kDa (usually called 51 kDa) were visualized by staining with Coomassie brilliant blue (Fig. 4). Additional aliquots were fractionated on gels, transferred to nitrocellulose paper, and probed with several of the mAbs (Fig. 5). The 51-kDa protein reacted with the amino-terminal-specific mAb 19 and failed to react with the carboxy-terminal-specific mAb 50. Since the first 13 amino acids of HIV-1 RT must be present for mAb19 to bind the RT, this demonstrates that the 51-kDa protein produced *in vitro* has the same amino terminus as the 66-kDa protein but lacks the carboxy terminus. The 51-kDa protein from HIV-1 virions has the same apparent molecular weight and reacts identically when probed with these mAbs (see Figs. 1 and 5D). These results suggest that cleavage occurs either at or very near the same site in the *in vitro* digestion and in virions. The 66-kDa form can be cleaved to a 51-kDa form by a protease present in *E. coli* (Lowe *et al.*, 1988). Although the HIV-1 protease made in *E. coli* was purified by HPLC, we considered the possibility that a contaminating bacterial protease could have been responsible for cleaving the HIV-1 RT *in vitro*. To demonstrate that the HIV-1 protease was responsible for the cleavage seen *in vitro*, the 66-kDa form of HIV-1 RT was cleaved with a highly purified HIV-1 protease synthesized by a recombinant strain of yeast (Pichuantes *et al.*, 1989). The results were similar to those obtained with the HIV-1 protease purified from *E. coli* (data not shown). The *in vitro* digestion of HIV-1 RT with recombinant protease demonstrates that the only site in the 66-kDa form cleaved by the viral protease is the site giving rise to the 51-kDa form. Intermediate species between 66 kDa and 51 kDa were not observed (Fig. 5). We were unable to detect a small carboxy-terminal fragment either in disrupted virions or in the *in vitro* digests with either of the mAbs that react with this portion of HIV-1 RT (Fig. 5 and data not shown).

The specificity of the cleavage event may reside in the viral protease, in the way the 66-kDa HIV-1 RT is folded or both. Lowe *et al.* (1988) have reported that a 66-kDa HIV-1 RT purified from *E. coli* tends to break down to a mixture of 66- and 51-kDa forms. We have noticed similar breakdown in preparations that have not been rigorously purified (see Fig. 1). The cleavage site that yields this 51-kDa form must be near the site cleaved by the viral protease. Immunoblots probed with mAb 19 and 50 reveal that this 51-kDa form also has an intact amino terminus and is missing the carboxy terminus (data not shown). Although the protease responsible for this cleavage is presumably of *E. coli* origin, it has not been identified. These results suggest that the site (or region) cleaved by the viral protease is unusually sensitive to proteolysis. We have digested the 66-kDa protein with several other proteases. Although none of the enzymes we tested shows the specificity of the viral protease, both papain and trypsin cleave near the site recognized by the viral protease and at additional sites (Fig. 6). In contrast to the data reported by Lowe *et al.* (1988), we did not find chromotrypsin particularly effective in mimicking the cleavage made in virions. This unexpected difference may result

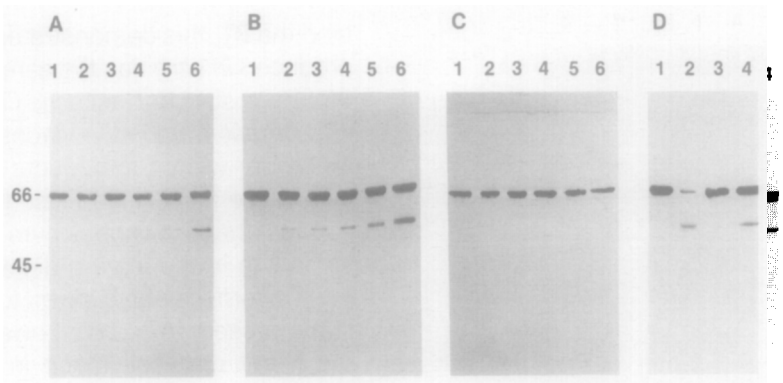


FIG. 5. Immunoblots showing the recognition of the HIV-1 RT proteolysis products shown in Fig. 4 by different mAbs (A–C). D shows a comparison of the electrophoretic mobility of the products of *in vitro* proteolysis of HIV-1 RT by recombinant HIV-1 protease with RT proteins in HIV-1 virions. All lanes except D2 were loaded with 2.7 μg of protein. The proteins were separated and transferred to nitrocellulose paper. Transfers were probed with mAbs and binding visualized by alkaline phosphatase staining. (A) mAB 19. Lanes: 1, control, RT incubated 22 hr at 37° without protease; 2, 2 hr; 3, 3 hr; 4, 4 hr; 5, 5 hr; and 6, 22 hr. (B) mAb21. Lanes: 1, control, 22 hr without protease; 2, 3 hr; 3, 4 hr; 4, 5 hr; 5, 6 hr; and 6, 22 hr. (C) mAb 50. Lanes: 1, control, 22 hr without protease; 2, 3 hr; 3, 4 hr; 4, 5 hr; 5, 6 hr; and 6, 22 hr. (D) mAb21. Lanes: 1, starting material, RT before incubation; 2, HIV-IIIB virions disrupted in SDS sample buffer; 3, control, RT incubated 16 hr at 37° without protease; 4, RT incubated 16 hr at 37° with protease.

from the different incubation temperatures used by each group (see Discussion).

DISCUSSION

Retroviral reverse transcriptases contain two enzymatic activities, a DNA polymerase that can copy either RNA or DNA templates and an RNase H. These activities are associated with distinct segments of the reverse transcriptase. The DNA polymerase domain comprises approximately the amino-terminal three-quarters of HIV-1 RT, and the RNase H domain comprises the remaining one-quarter at the carboxy terminus (Johnson *et al.*, 1986). In the MuLV RT, the two domains have been expressed independently of each other and each is fully active. For HIV-1 RT it has not been possible to express a protein with fully active RNA-dependent DNA polymerase without the RNase H domain (Hizi *et al.*, 1988; Tisdale *et al.*, 1988). Moreover, polymerase activity is sensitive to small deletions or insertions in the RNase H domain (Hizi *et al.*, 1989).

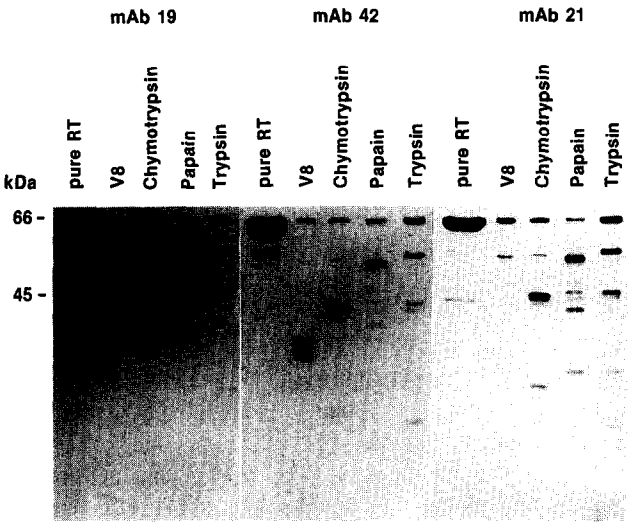


FIG. 6. Immunologic detection of proteolytic fragments generated by nonviral proteases. Four proteases, V8 (2.25 μg/ml), chymotrypsin (0.75 μg/ml), papain (0.375 μg/ml), and trypsin (1.25 μg/ml), were incubated with purified RT (0.38 mg/ml) for 24, 3, 24, and 3 hr, respectively, at 23°. The digestion products were separated on polyacrylamide gels, transferred to nitrocellulose paper, and probed with three different mAbs, 19, 42, and 21. Positive reactions were visualized by alkaline phosphatase staining. Lanes containing only pure RT have 14 μg of protein. All other lanes contain 5 μg of protein.

The HIV-1 RT isolated from virions is a mixture of two forms. The larger 66-kDa form contains both the polymerase and the RNase H segments, while the 51-kDa form has only the polymerase domain. Other well-studied retroviral reverse transcriptases do not appear to undergo a similar cleavage event (Weiss *et al.*, 1982, 1985). The sequences in the 66-kDa protein that are thought to be cleaved to yield the 51-kDa protein are not cleaved by the HIV-1 protease if presented as a synthetic peptide (Billich *et al.*, 1988). Nonviral proteases have been reported to cleave the 66-kDa form of HIV-1 RT to yield a protein similar to the 51-kDa protein (Lowe *et al.*, 1988). Of the proteases we have tested, the viral protease displays the most selectivity in cleaving the 66-kDa form of HIV-1. Unrelated nonviral proteases can cleave the 66-kDa form at sites near the cleavage site for the viral protease but are less selective and make additional cleavages. This provides direct support for the hypothesis that the cleavage in viri-

ons that produces the 51-kDa protein is made by the viral protease. These data also suggest that the specificity of the cleavage event depends both on the protease and on the structure of the 66-kDa protein. There is a region of the 66-kDa protein that is susceptible to the viral protease and to unrelated proteases. We have not seen as much specificity in the cleavage of the 66-kDa HIV-1 RT when we incubated the protein with nonviral proteases at 23° as Lowe *et al.* (1988) observed when they performed digestions at 0°. They reported additional cleavages when the incubation temperature was raised from 0 to 21° and the incubation time was increased. However, the cleavage of RT in virions occurs at 37°. This suggests that the relative specificity of the viral protease plays a significant role in correctly defining the cleavage of the 66-kDa protein.

We have been unable to detect a 15-kDa (or smaller) carboxy-terminal fragment after the *in vitro* cleavage of 66-kDa HIV-1 RT with the viral protease either by Coomassie brilliant blue staining or by immunoblotting. In addition, a 15-kDa fragment was not detected in disrupted virions by immunoblotting. Hansen *et al.* (1988) have reported detecting a 15-kDa RNase H activity in immunoprecipitates from extracts of HIV-1 virions. We are in the process of investigating this apparent discrepancy.

We have proposed a model for the structure of the HIV-1 RT based on sequence comparisons, mutational analysis, and biochemical data (Barber *et al.*, manuscript in preparation). The proteolytic data supports the model. In addition, the antibody binding regions reported by Tisdale *et al.* (1988) and by us are (with one possible exception) within segments of the HIV-1 RT we have predicted would lie on the surface of the molecule.

We have suggested that the amino terminus and the central region between amino acids 190 and 340 are on the surface of the properly folded RNA-dependent DNA polymerase domain. RNase H activity appears to be part of a separate domain. Because the RNase H domain is small and is connected to the polymerase domain by a protease-sensitive segment, much of the carboxy terminus of HIV-1 RT should be on the surface. All of the mAbs, with the possible exception of mAb 42, recognize segments we predict are on the surface. The binding site for mAb 42 has not been mapped precisely and may be near the amino terminus on the surface of the properly folded form of HIV-1 RT. We are now attempting to define the binding site of this antibody more precisely.

ACKNOWLEDGMENTS

We are grateful to H. Marusiodis for preparing the manuscript, and to Laura Fornwald for excellent technical assistance. Research was sponsored by the National Cancer Institute, DHHS, under Contract No. NO1-CO-74101 with BRI and the National Institute of General Medical Sciences. The contents of this publication do not necessarily reflect the views or policies of the Department of Health & Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

REFERENCES

BABE, L., PICHUANTES, S., BARR, P., and CRAIK, C. S. (1989). HIV-1 protease: Bacterial expression, purification and characterization. *J. Cell. Biochem.*, in press.

BILLICH, S., KNOOP, M.-T., HANSEN, J., STROP, P., SEDLACK, J., MERTZ, R., and MOELLING, K. (1988). Synthetic peptides as substrates and inhibitors of human immune deficiency virus-I protease. *J. Biol. Chem.* **263,** 17,905–17,908.

DI MARZO VERONESE, F., COPELAND, T. D., DEVICO, A. L., RAHMAN, R., OROSZLAN, S., GALLO, R. C., and SARNGADHARAN, M. G. (1986). Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. *Science* **231,** 1289–1291.

FARMERIE, W. G., LOEB, D. D., CASAVANT, N. C., HUTCHINSON, C. A., III, EDGELL, M. H., and SWANSTROM, R. (1987). Expression and processing of the AIDS virus reverse transcriptase in *Escherichia coli*. *Science* **236,** 305–308.

HANSEN, J., SCHULZE, T., and MOELLING, K. (1987). RNase H activity associated with bacterially expressed reverse transcriptase of human T-cell lymphotropic virus III/lymphadenopathy-associated virus. *J. Biol. Chem.* **262,** 12,393–12,396.

HANSEN, J., SCHULZE, T., MELLERT, W., and MOELLING, K. (1988). Identification and characterization of HIV-specific RNase H by monoclonal antibody. *EMBO J.* **7,** 239–243.

HIZI, A., and HUGHES, S. H. (1988). Expression in *Escherichia coli* of a Moloney murine leukemia virus reverse transcriptase whose structure closely resembles the viral enzyme. *Gene* **66,** 319–323.

HIZI, A., MCGILL, C., and HUGHES, S. H. (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.

HIZI, A., BARBER, A., and HUGHES, S. H. (1989). Effects of small insertions on the RNA-dependent DNa polymerase activity of HIV-1 reverse transcriptase. *Virology* **170,** 326–329.

JOHNSON, M. S., MCCLURE, M. A., FENG, D. F., GRAY, J., and DOOLITTLE, R. (1986). Computer analysis of retroviral *pol* genes: Assignment of enzymatic functions of specific sequences and homologies with nonviral enzymes. *Proc. Natl. Acad. Sci. USA* **83,** 7648–7652.

KOHLER, G., HOWE, S. C., and MILSTEIN, C. (1976). Fusion between immunoglobulin-secreting and nonsecreting myeloma cell lines. *Eur. J. Immunol.* **6,** 292–295.

KOTEWICZ, M. L., SAMPSON, C. M., D'ALESSIO, J. M., and GERARD, G. F. (1988). Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* **16,** 265–277.

LARDER, B., PURIFOY, D., POWELL, K., and DARBY, G. (1987). AIDS virus reverse transcriptase defined by high level expression in *Escherichia coli*. *EMBO J.* **6,** 3133–3137.

LEGRICE, S., BENCK, V., and MOUS, J. (1987). Expression of biologically active T-cell lymphotropic virus type III reverse transcriptase in *Bacillus subtilis*. *Gene* **55,** 95–103.

LEGRICE, S. F. J., ZEHNLE, R., and MOUS, J. (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.

LIGHTFOOTE, M., COLIGAN, J., FOLKS, T., FAUCI, A. S., MARTIN, M. A., and VENKATESAN, S. (1986). Structural characterization of reverse

transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus. *J. Virol.* **60,** 771–775.

LOWE, D. M., ATIKEN, A., BRADLEY, C., DARBY, G. K., LARDER, B. A., POWELL, K. L., PURIFOY, D. J. M., TISDALE, M., and STAMMERS, D. K. (1988). HIV-1 reverse transcriptase: Crystallization and analysis of domain structure by limited proteolysis. *Biochemistry* **27,** 8884–8889.

MANIATIS, T., FRITSCH, E. F., and SAMBROOK, J. (1982). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

MOUS, J., HEIMER, E. P., and LEGRICE, S. F. (1988). Processing protease and reverse transcriptase from human immunodeficiency virus type I polyprotein in *Escherichia coli. J. Virol.* **62,** 1433–1436.

PICHUANTES, S., BABE, L., BARR, P., and CRAIK, C. (1989). Recombinant HIV-1 protease secreted by *Saccharomyces cerevisiae* correctly processes myristylated gag polyproteins. *Proteins: Struct. Funct. Genet.,* in press.

SHINE, J., and DALGARNO, L. (1974). The 3′ terminal sequence of *E. coli* 16S ribosomal RNA: Complementarity to nonsense triplets in ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71,** 1342–1346.

SHOWALTER, S. D., ZWEIG, M., and HAMPAR, B. (1981). Monoclonal antibodies to Herpes simplex virus type I proteins, including the immediate-early protein, ICP 4. *Infect. Immun.* **34,** 684–692.

TANESE, N., SODROSKI, J., HASELTINE, W., and GOFF, S. P. (1986). Expression of reverse transcriptase activity of human T-lymphotropic virus type III (HTLV-III/LAV) in *Escherichia coli. J. Virol.* **59,** 743–745.

TANESE, N., and GOFF, S. P. (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.

TISDALE, M., ERTL, P., LARDER, B. A., PURIFOY, D. J. M., DARBY, G., and POWELL, K. L. (1988). Characterization of human immunodeficiency virus type I reverse transcriptase by using monoclonal antibodies: Role of the C terminus in antibody reactivity and enzyme function. *J. Virol.* **62,** 3662–3667.

WEISS, R., TEICH, N., VARMUS, H. E., and COFFIN, J. (1982, 1985). "RNA Tumor Viruses." Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.