JOURNAL OF VIROLOGY, Oct. 1988, p. 3718–3728
0022-538X/88/103718-11$02.00/0

Vol. 62, No. 10

Processing of *gag* Precursor Polyprotein of Human T-Cell Leukemia Virus Type I by Virus-Encoded Protease

SEOK HYUN NAM, MINORU KIDOKORO, HISATOSHI SHIDA, AND MASAKAZU HATANAKA*

Institute for Virus Research, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Received 14 March 1988/Accepted 13 June 1988

The biological activity encoded in the putative protease gene (*pro*) of human T-cell leukemia virus type I was investigated by using a vaccinia virus expression vector. The 53-kilodalton *gag* precursor polyprotein was processed into the mature p19, p24, and p15 *gag* proteins when the *gag* and protease-coding sequence was expressed under the control of a vaccinia virus promoter, suggesting that the protease may be synthesized through the mechanism of ribosomal frame shifting. The processing defect of a protease mutant could be complemented by cointroduction of a wild-type construct into the cell, demonstrating that the *pro* gene encodes the biologically active protease molecules which are capable of processing the *gag* precursor polyprotein in vivo in *trans*. A study involving the use of a variety of mutants constructed in vitro revealed that the protease consists of a nonessential carboxy-terminal region and a part essential for its activity, including the putative catalytic residue, aspartic acid. Furthermore, a cluster of adenine residues positioned at the overlapping region between the *gag* and *pro* genes was shown to be involved in the ribosomal frameshifting event for the synthesis of protease. To mimic the formation of the 76-kilodalton *gag-pro* precursor polyprotein formed by ribosomal slipping, the coding frames of the *gag* and *pro* gene were adjusted. The processing of the *gag-pro* precursor polyprotein depended on an intact protease gene, implying that a *cis*-acting function of human T-cell leukemia virus type I protease may be necessary to trigger the initial cleavage event that leads to the release of protease from the precursor protein.

Human T-cell leukemia virus type I (HTLV-I) is the first human type C exogenous retrovirus associated with adult T-cell leukemia (9, 31, 32, 50). HTLV-I has been molecularly cloned from the peripheral blood leukocytes of an adult T-cell leukemia patient, and subsequent characterization of the viral genome has revealed that the HTLV-I genome is composed of *gag*, *pol*, *env*, and *pX* genes, the order of which is similar to that in known animal retroviruses. The internal structural proteins of HTLV-I encoded in the *gag* gene are translated as a *gag* precursor polyprotein of 53 kilodaltons, which is subsequently processed into p19, p24 and p15, the mature *gag* structural proteins (8).

Proteolytic cleavage of the *gag* precursor polyprotein of murine and avian leukemogenic retroviruses has been shown to be carried out by a highly substrate-specific virion-associated protease, which plays an important role in infectivity of virus particles (17, 48). On the basis of comparison of the amino acid sequence of retrovirus-associated proteases with the DNA sequence of cloned viral genomes, it was concluded that the protease is encoded in the viral genome itself (36, 51–53).

We have molecularly cloned a novel full-length provirus genome of HTLV-I from the virus-producing cell line MT-2 (24). This clone, designated as HTLV1C, harbors an open reading frame (ORF) between the *gag* and the *pol* genes which corresponds to the protease-coding region of other retroviruses (26). This ORF had not been identified in the nucleotide sequence of the other isolates (33, 37). We have reported that amino acid sequences strictly conserved among other retroviral proteases are also found in the possible protease-coding sequence of HTLV1C. Recently, another full-length DNA clone of HTLV-I, also containing the ORF of the putative protease region, was independently isolated from extrachromosomal closed-circular copies in chronically infected promyelocytic leukemia HL60 cells (10). Both of the cloned viral DNAs harbor essentially the same ORFs for putative HTLV-I protease, except for a few base transitions. This strongly suggests that HTLV-I encodes it own protease. However, scarcity of the protease associated with virions, together with unavailability of efficient virus-producing cells, has hampered our ability to examine the properties of the HTLV-I protease.

Vaccinia virus propagates in the cytoplasm exclusively, and in infected cells a variety of genes are expressed in a regulated manner common to eucaryotic viruses and cells, except for the absence of splicing involved in mRNA synthesis. Recently, the usefulness of vaccinia virus expression vectors has been established (39). Such vectors express foreign genes at high levels. We decided to adopt the vaccinia virus system to investigate the biological function and mode of expression of the HTLV-I protease gene.

In this study we demonstrate that proteolytic cleavage of *gag* precursor polyprotein depends on HTLV-I protease, which is encoded in the putative protease gene region inferred from nucleotide sequencing. We also characterize the biosynthesis of the HTLV-I protease by using in vitro mutagenesis techniques.

MATERIALS AND METHODS

Materials. All restriction endonucleases, as well as DNA-modifying enzymes, the M13 sequencing kit, and the phosphorylated 8-nucleotide *Bgl*II DNA linker were purchased from Takara Shuzo Co., Kyoto, Japan. Restriction endonuclease *Bst*XI and *Oxa*NI were purchased from New England BioLabs, Inc., Beverly, Mass. Mung bean nuclease was obtained from Pharmacia, Uppsala, Sweden. Anti-p19 monoclonal antibody (GIN7), anti-p24 monoclonal antibody (NORI), anti-p15 monoclonal antibody (FR45), and partially purified HTLV-I virus proteins were kindly provided by Fujirevio Ltd., Tokyo, Japan. [α-^{32}P]dCTP was supplied by

* Corresponding author.

Du Pont, NEN Research Products, Boston, Mass. Protein A gold solution and its gold enhancement kit were purchased from Bio-Rad Laboratories, Richmond, Calif.

General DNA method. Plasmid DNA was prepared by a slight modification of the alkali-sodium dodecyl sulfate (SDS) method of Birnboim and Doly (2). Restriction endonucleases and DNA-modifying enzymes were used as specified by the manufacturers. All molecular-biological manipulations were carried out by standard methods (23). All DNA transformation were performed with competent *Escherichia coli* JM109 cells. However, single-stranded DNAs for site-specific mutagenesis were produced by using transformants of *E. coli* TG1 harboring subcloned M13 replicative-form (RF) DNAs or infecting transformants of *E. coli* MV1184 harboring subcloned pUC118 or pUC119 DNAs with helper bacteriophage M13KO7 as described by Vieira and Messing (J. Vieira and J. Messing, Methods Enzymol., in press). The sequencing of single-stranded M13 virion DNA or double-stranded plasmids was done by the chain termination method described by Sanger et al. (35).

Construction of plasmids. To construct p7.5gagN, we isolated the 2.7-kilobase (kb) *Sst*I-*Kpn*I fragment of HTLV1C that contains the 3′ half of the long terminal repeat, *gag*, and *pro* and the 5′ portion of *pol*. This fragment was digested with *Apa*I, and a 2.3-kb fragment was purified and then partially digested with *Rsa*I to remove the 3′ half of the long terminal repeat. Following treatment with T4 DNA polymerase, a 2-kb fragment was ligated with *Sma*I-digested pPro18 (38), and then plasmid DNA from an ampicillin-resistant transformant was tested to selected the desired plasmid by appropriate enzyme digestion.

To construct p7.5gagΔNS, p7.5gagN was digested to completion with *Stu*I, and the 1-kb *Stu*I fragment, which contains the entire p15 coding sequence and a portion of p24 and the protease-coding sequence, was purified. The 4-kb *Stu*I fragment harboring the pUC vector sequence was also isolated for further experiments. The 1-kb fragment was digested completely with *Nar*I and treated with mung bean nuclease to digest the 5′ protruding end of the *Nar*I site. The 0.9-kb fragment was ligated with the 4-kb *Stu*I fragment of p7.5gagN, and plasmid p7.5gagΔNS was obtained.

The identical *Stu*I fragments of p7.5gagM were used to construct p7.5gagΔTS. The 1-kb *Stu*I fragment was digested to completion with *Taq*I and then treated with the Klenow fragment of *E. coli* DNA polymerase I to fill the end. The enzyme-modified 0.6-kb fragment was recovered and ligated with 4-kb *Stu*I fragment as described above. The resulting plasmid was designated as p7.5gagΔTS.

To construct the p7.5gagINBg, we digested p7.5gagN completely with *Oxa*NI. After both ends have been filled with Klenow fragments, linearized p7.5gagN was ligated with the 5′-phosphorylated synthetic octanucleotide *Bgl*II linker, 5′-CAGATCTG-3′. Following complete digestion with *Bgl*II, the linearized plasmid was recircularized by ligation to obtain the plasmid p7.5gagINBg. Insertion of the linker DNA resulted in termination of the protease 10 amino acids downstream of the *Oxa*NI site.

To construct p7.5gagΔHO, p7.5gagN was partially digested with *Hinc*II and subsequently digested with *Oxa*NI to completion. The 0.3- and 4.7-kb fragments, harboring a portion of the protease-coding sequence and the remainder of the plasmid including the pUC sequence, respectively, were recovered from the gels. The 0.3-kb *Hinc*II-*Oxa*NI fragment was further digested with *Hpa*II and then treated with Klenow fragment to fill the end. The enzyme-modified 0.2-kb fragment was ligated with the 4.7-kb *Hinc*II-*Oxa*NI fragment of p7.5gagN, which was previously treated with Klenow fragment, to produce p7.5gagΔHO.

Strategies for construction of in vitro-mutagenized plasmids were as follows. Plasmids p7.5gagfs19AspI and p7.5gagfs19AspII were constructed by using p7.5gagfs19, p7.5gagNAspI, and p.7.5gagNAspII. The 0.6-kb *Hinc*II-*Eco*RI fragments harboring the in vitro-mutagenized protease-coding sequence were purified from p7.5gagNAspI or p7.5gagNAspII. These fragments were used to replace the corresponding region of p7.5gagfs19 to obtain p7.5gagfs19AspI and p7.5gagfs19AspII. To construct p7.5gagfs19ΔOE, p7.5gagfs19AspI was completely digested with *Oxa*NI and *Eco*RI and then treated with the Klenow fragment of *E. coli* DNA polymerase I to fill both ends. The 4.7-kb *Oxa*NI-*Eco*RI fragment was circularized by self-ligation to produce p7.5gagfs19ΔOE. This truncated *gag-pro* fusion protein may have 26 carboxy-terminal amino acids contributed by pPro18.

A novel restriction enzyme cleavage site, *Kpn*I, created at amino acid 125 of the protease was used to truncate the 3′ region of the protease-coding sequence in p7.5gagfs19AspII. After complete digestion with *Kpn*I, the 4.6-kb *Kpn*I fragment of p7.5gagfs19AspII was recircularized to generate p7.5gagfs19ΔKp. This truncated *gag-pro* fusion protein may have 32 carboxy-terminal amino acids contributed by pPro18.

In vitro mutagenesis. Oligonucleotide-directed in vitro mutagenesis was performed by using a mutagenesis kit supplied by Amersham International, Amersham, England, as described by Taylor et al. (43), except that the reaction volume was reduced to half of the original volume. To disrupt the essential amino acid sequence for proteolytic cleavage located between p19 and p24 junction, we excised a 0.6-kb *Bam*HI-*Pst*I fragment, harboring the entire p19 coding sequence and the 5′ portion of the p24 coding sequence, from p7.5gagN and then ligated it with M13mp11 RF DNA which had been previously digested with *Bam*HI and *Pst*I. Site-specific mutagenesis was performed by using the standard protocol specified by the manufacturer, with the 37-mer oligonucleotide 5′-TGAGCCTACGGCGCCCC CAAGTCCTTCCGTCATGCAT-3′. One base deletion at nucleotide 1179, together with one base insertion at nucleotide 1194, gave rise to the localized frame shifting at the p19 - p24 junction region, with simultaneous amino acid sequence changes from Pro-Gln-Val-Leu-Pro to Pro-Pro-Ser-Pro-Ser. Subcloned M13mp11 RF DNA that contained the mutagenized sequence was digested with *Bam*HI and *Nco*I, and then the 0.5-kb fragment was purified. This fragment was ligated with the 4.1-kb fragment of p7.5gagΔTS, which had been digested completely with *Bam*HI and *Nco*I, to produce p7.5gagΔTS37.

To construct p7.5gagfs19, the 0.9-kb *Sma*I-*Eco*RI fragment harboring the entire protease-coding sequence and the 3′ half of the p15 coding sequence was excised from p7.5gagN and then inserted at the corresponding restriction enzyme sites of M13mp10 RF DNA. Mutagenesis was carried out with the 19-mer oligonucleotide 5′-ACACCCAAA GAAACTCCAT-3′. Direct sequencing of the mutagenized region showed the creation of a new DNA sequence, AAA GAAAC, instead of AAAAAAC at six consecutive adenine residues positioned at the 3′ terminal region of the *gag* gene. The 0.2-kb *Bst*XI-*Hinc*II fragment was excised from subcloned M13mp10 RF DNA harboring the mutagenized target DNA by complete restriction enzyme digestions and then ligated with the 4.8-kb fragment of p7.5gagN, which had

been partially digested with *Hin*cII and subsequently digested with *Bst*XI to completion.

The 0.9-kb *Sma*I-*Eco*RI fragment described above was also inserted into the corresponding enzyme site of M13mp11 RF DNA, and in vitro mutagenesis was carried out with using 5'-CCTATGGAGAATATTGGGTGTG-3' as a mutagenic oligonucleotide. As a result of mutagenesis, the new DNA sequence represented by AATATTC was created instead of AAAAAAC, with simultaneous creation of a novel restriction enzyme cleavage site of *Ssp*I of six consecutive adenine residues positioned at the 3' terminal region of the *gag* gene. To construct p7.5gagfs22, we introduced the mutagenized DNA fragment into p7.5gagN by using a procedure identical to that used for construction of p7.5gagfs19.

To change the amino acid at position 64 of the protease from aspartic acid to glycine, we excised the 0.6-kb *Hin*cII-*Eco*RI fragment harboring the protease-coding sequence from p7.5gagN and ligated it with pUC118 linearized by previous *Hin*cII and *Eco*RI digestion. Mutagenesis was performed by using 5'-AAGCTCTACTAGGTACCGGAG CAGACATGACAGT-3' as a mutagenic oligonucleotide. The mutagenized clone was judged by the creation of a restriction enzyme site *Kpn*I at the novel glycine codon. A similar procedure was adopted to exchange the amino acid at position 125 from aspartic acid to glycine. The 0.6-kb *Hin*cII-*Eco*RI fragment described above was inserted at the corresponding restriction enzyme sites of pUC119. Mutagenesis was performed by using 5'-GTTTTTGGTACCAACTAGG-3' as a mutagenic oligonucleotide. The mutagenized clone was judged by the simultaneous creation of the novel restriction enzyme cleavage site of *Kpn*I at the novel glycine codon. To construct p7.5gagNAspI or p7.5gagNAspII, which contain the newly created glycine residues at amino acid positions 64 and 125 of the protease, respectively, we purified the 0.3-kb *Hin*cII-*Oxa*NI fragments from subcloned pUC118 or pUC119 DNA by complete restriction enzyme digestions and then ligated them to the 4.7-kb fragment of p7.5gagN, which had been partially digested with *Hin*cII and subsequently digested to completion with *Oxa*NI.

Recombinant vaccinia virus preparation. Vaccinia-HTLV-I recombinant plasmids for recombinant vaccinia virus were constructed as follows. Plasmid p7.5gagN was digested with *Hin*dIII and *Eco*RI, and the 2.3-kb fragment harboring the entire *gag* and protease-coding sequence including the promoter of the 7.5K gene was purified. The large fragment of *E. coli* DNA polymerase I (Klenow) was used to fill both ends of this fragment. To flank the fragment with the hemagglutinin gene of vaccinia virus, we inserted the fragment into pHA13 which had been linearized with *Nru*I and subsequent treatment of Klenow fragment to fill the ends (38). Candidate clones were tested for the orientation of the insert, and the desired plasmid was designated as pHAgagN.

The 2-kb *Hin*dIII-*Eco*RI fragment harboring the entire *gag* gene, including the truncated protease-coding sequence and the 7.5K gene promoter, was excised from p7.5gagΔTS. After treatment with the Klenow fragment of *E. coli* DNA polymerase I, this fragment was inserted in to the *Nru*I site of hemagglutinin gene of vaccinia virus as described above. The resultant plasmid was designated as pHAgagΔTS. The detailed procedure for generation and purification of recombinant vaccinia virus was described by Shida et al. (39).

DNA transfection. CV-I cells for transfection of plasmid DNA were maintained in Eagle minimal essential medium supplemented with 10% fetal calf serum at 37°C in humidified air with 5% CO_2. Transfection of plasmid DNAs for transient expression was performed as described (6). Briefly, 2×10^5 to 3×10^5 cells per 35-mm dish were preinfected with wild-type vaccinia virus at 30 PFU per cell and adsorbed for 1 h at 37°C. Then, 10 μg of calcium phosphate-precipitated recombinant plasmid DNA was added. After a 30-min incubation at room temperature, fresh medium was added. After a 4-h incubation at 37°C, the medium was replaced by fresh medium; incubation was continued overnight.

Viral protein analysis. At 20 h posttransfection with recombinant plasmid DNA, CV-I cells were washed twice with phosphate-buffered saline and then solubilized in SDS-sample buffer by heating for 3 min at 95 to 100°C. Solubilized proteins were separated by SDS-polyacrylamide gel electrophoresis (12% polyacrylamide) and transferred onto a nitrocellulose filter as described (45). The blots were blocked in fetal calf serum at 4°C overnight and then probed with the anti-*gag* monoclonal antibodies appropriately diluted with 3% bovine serum albumin solution containing 20 mM Tris (pH 8.0) and 150 mM NaCl. After the primary antibody reaction had continued at room temperature for 3 h, the secondary antibody reaction, with rabbit anti-mouse immunoglobulin G antibody, was performed under the same conditions. Finally, the protein A-gold reaction and subsequent enhancement were performed as specified by the manufacturer; the *gag*-related proteins expressed in CV-I cells transfected with the recombinant plasmids were revealed.

RESULTS

Processing of *gag* precursor polyprotein in a recombinant vaccinia virus expression vector system. The 2-kb *Rsa*I-*Apa*I fragment harboring the entire *gag* and *pro* ORF was purified and introduced downstream of the promoter of the vaccinia virus 7.5K gene, which is expressed at both early and late times in the vaccinia virus infection cycle (3). The construction of plasmid p7.5gagN is outlined in Fig. 1 and is described in Materials and Methods.

The plasmid was introduced into vaccinia virus-infected CV-I cells by using the calcium phosphate coprecipitation technique, and then transient expression of *gag* gene products in CV-I cells was examined by Western immunoblotting analysis with anti-*gag* monoclonal antibodies.

The 19-, 24-, and 15-kilodalton proteins (19K, 24K, and 15K proteins, respectively) were recognized by an anti-p19 monoclonal antibody (GIN7), an anti-p24 monoclonal antibody (NORI), and an anti-p15 monoclonal antibody (FR45), respectively (Fig. 2). Small amounts of a 53K protein were recognized by FR45, which is the uncleaved *gag* precursor polyprotein. A 36K protein, which is considered to be an intermediate precursor protein, was also detected. All the *gag* proteins transiently expressed from plasmid p7.5gagN comigrated in SDS-polyacrylamide gel electrophoresis with the corresponding HTLV-I *gag* proteins in extracts of HTLV-I-producing MT-2 cells and in isolated viruses. The 28K protein observed in MT-2 cells is not the processed product of *gag* precursor polyproteins, but is the *gag-pX* fusion protein consisting of p19, part of p24, and a peptide from the *pX* region, and is encoded by the defective proviruses (12). These results show that the cloned *gag* and *pro* ORF region is enough to synthesize all *gag* proteins correctly processed from the precursor polyprotein. As a matter of convenience, subsequent Western blot experiments were carried out by using the mixed anti-*gag* monoclonal antibodies GIN7, NORI, and FR45.

Proteolytic activity residues within the putative gene product of the *pro*ORF. To demonstrate that the processing of *gag*

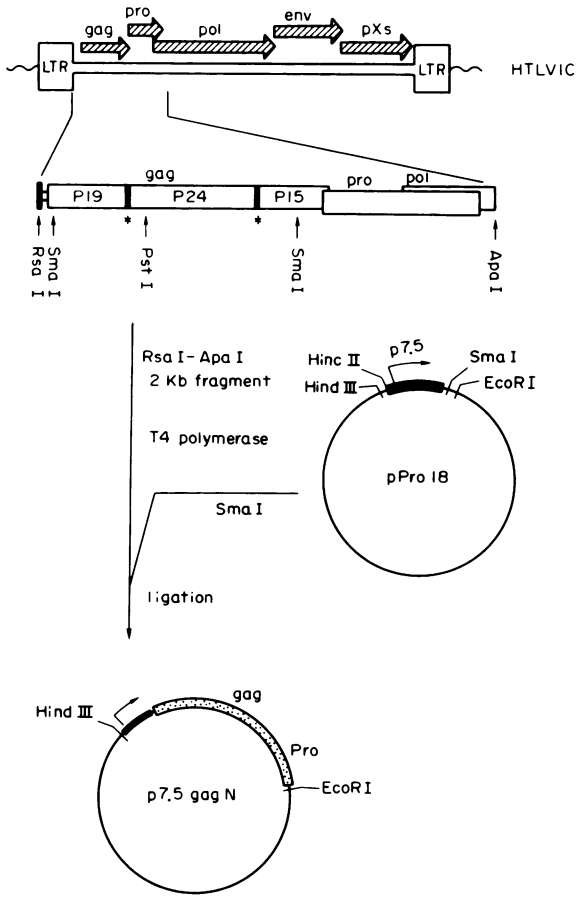

FIG. 1. Schematic representation of the HTLV-I genome and construction of p7.5gagN. The strategy used to insert the entire *gag* and protease-coding sequence into the pPro18 expression vector is outlined. Symbols: ▨, *gag*, *pol*, *env*, and *pXs* genes of HTLV-I; ▬, 7.5K gene promoter of vaccinia virus; ▩, *gag* and protease-coding sequence; circle, pUC18 DNA sequence; *, positions of the HTLV-I protease-mediated processing sites; →, direction of transcription.

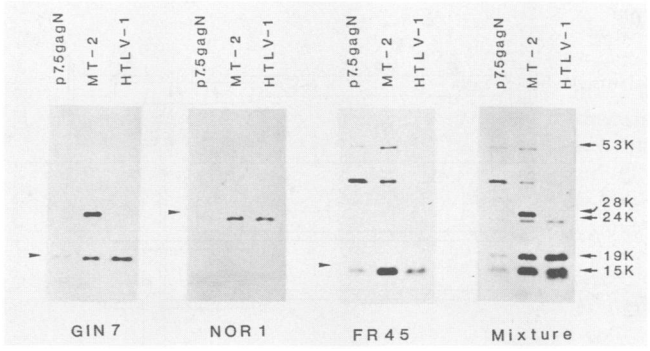

FIG. 2. Expression and processing of *gag* precursor polyprotein in CV-I cells by using vaccinia virus-derived expression vector. Plasmid p7.5gagN was constructed as shown in Fig. 1. Proteins expressed from the plasmid were analyzed on an SDS–12% polyacrylamide gel and electrophoretically transferred onto nitrocellulose filters. Transient expression of *gag*-related protein were analyzed by Western blotting analysis with anti-p19 monoclonal antibody (GIN7), anti-p24 monoclonal antibody (NORI), and anti-p15 monoclonal antibody (FR45), or their mixtures. Arrows indicated the mature *gag* proteins expressed from p7.5gagN. Virion-associated *gag* proteins (HTLV-I) and those synthesized in the chronically infected cells (MT-2) were also analyzed to confirm the authenticity of the mature *gag* proteins expressed from the p7.5gagN.

precursor polyprotein observed in the vaccinia virus expression system is accomplished by HTLV-I protease encoded in nucleotide sequences within the *pro* ORF, we deleted several regions of protease-coding sequence (Fig. 3A).

A deletion from the *Nar*I site to the *Stu*I site (p7.5gagΔNS) did not abolish the correct processing of *gag* precursor polyprotein (Fig. 3B, lane 2). However, processing was blocked when the deletion was extended to the *Taq*I site (p7.5gagΔTS), located at the 5′ region of the *pro* ORF (lane 3), and the *gag* precursor polyprotein (Pr53*gag*) accumulated. The minor immunoreactive bands of lower molecular mass were considered to be degraded *gag* products. Translational termination of the protease as a result of the insertion of an 8-nucleotide *Bgl*II linker at the *Oxa*NI site (p7.5gagINBg) had no effect on correct processing of the 53K *gag* precursor polyprotein (lane 4). These results imply that a region spanning nucleotide positions 2228 to 2495 of the *pro* ORF is important to the processing of Pr53*gag* into the mature *gag* products. By contrast, the region downstream from the *Oxa*NI site of *pro* ORF is not essential for proteolytic function. This is consistent with the observation that the region from nucleotides 2228 to 2495 encodes the amino acid stretch that is conserved among the HTLV family of retroviruses.

To further examine the involvement of conserved amino acid stretches in the cleavage of *gag* precursor, we constructed plasmid p7.5gagΔHO, which contains the entire protease-coding sequence, except that a region harboring one of the conserved amino acid stretches was deleted (Fig. 3A). The major immunoreactive band encoded by p7.5gagΔHO was the 53K protein, although nonspecifically degraded proteins were also observed (Fig. 3B, lane 5). These results further support the assignment of the *pro* ORF region as the coding sequence for HTLV-I protease.

The processing defect of the *pro* ORF deletion mutant is complemented by active protease. To obtain evidence that the intact protease molecule directly cleaves the *gag* precursor polyprotein, we needed to determine whether the *gag* precursor polyprotein encoded by the *pro* ORF deletion mutant was able to be processed in *trans* by wild-type protease in vivo. The vaccinia virus expression vector system is suitable for this kind of study, because two independent recombinant DNAs can be cointroduced by using the recombinant vaccinia virus as a helper virus. The detailed strategy for construction of plasmids is shown schematically in Fig. 4. As the efficient donor of the active form of protease, a recombinant vaccinia virus expressing the protease was constructed by using hemagglutinin gene-mediated homologous recombination as described in Materials and Methods. The resultant recombinant vaccinia virus, WRgagN, harboring the *gag* and *pro* ORF, produced processed *gag* products identical to those synthesized in MT-2 cells (Fig. 5A, lane 4). For the control experiments, recombinant vaccinia virus harboring the *gag* ORF alone was constructed and designated WRgagΔTS. No correct processing occurred in cells infected with WRgagΔTS (lane 5). The p7.5gagΔTS plasmid, encoding only the 53K *gag* precursor, was mutagenized in vitro to disrupt the essential amino acid sequence for proteolytic cleavage located between p19 and p24. The resultant plasmid, p7.5gagΔTS37, encoded a novel p19-p24 junction se-

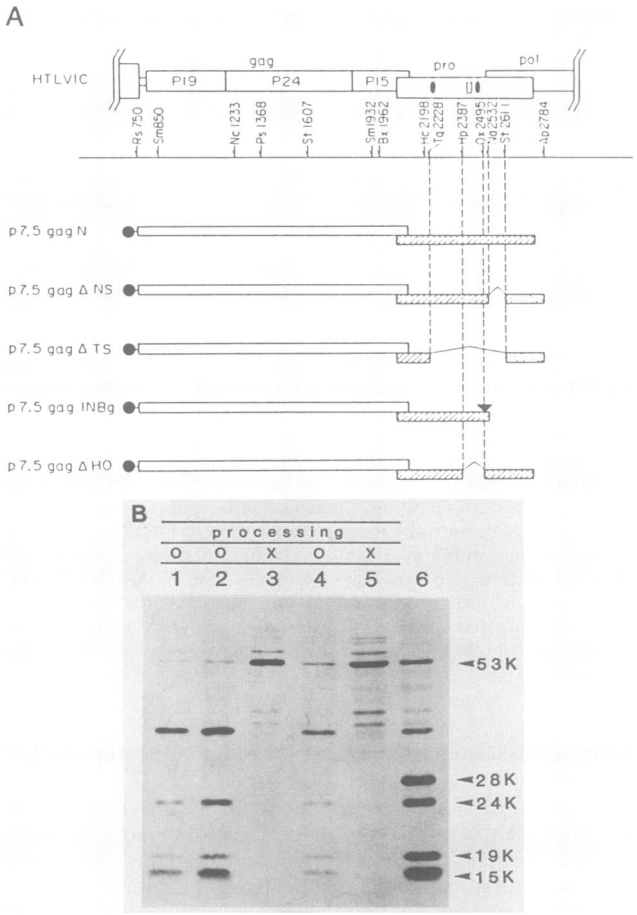

FIG. 3. Plasmids encoding the truncated protease molecules of HTLV-I and its expression. (A) Map of part of the HTLV-I genome is given at the top of the figure. The nucleotide number indicated in this figure corresponds to that reported by Seiki et al. (37). Several deletion mutants of protease-coding sequence were constructed as described in Materials and Methods. Symbols: ●, 7.5K gene promoter of vaccinia virus; ▭, *gag* coding sequence of HTLV-I; ▨, protease-coding sequence; ▩, reverse transcriptase coding sequence of HTLV-I; ▼, insertion of 8-nucleotide phosphorylated *Bgl*II synthetic linker at *Oxa*NI site; ⬮, two conserved amino acid stretches between the retroviral proteases and the cellular aspartyl proteases; □, short amino acid sequences conserved only among HTLV-family retroviral proteases. Abbreviations: Rs, *Rsa*I; Sm, *Sma*I; Nc, *Nco*I; Ps, *Pst*I; St, *Stu*I; Bx, *Bst*XI; Hc, *Hinc*II; Tq, *Taq*I; Hp, *Hpa*II; Ox, *Oxa*NI; Na, *Nar*I; Ap, *Apa*I. (B) Western blotting analysis of *gag* proteins expressed from deletion mutants of protease-coding sequence. Evidence for processing is indicated by signals, represented by ○ (processing positive) or × (processing negative). The mixed anti-*gag* monoclonal antibodies of GIN7, NORI, and FR45 were used to detect the HTLV-I *gag*-related proteins from cell lysates.

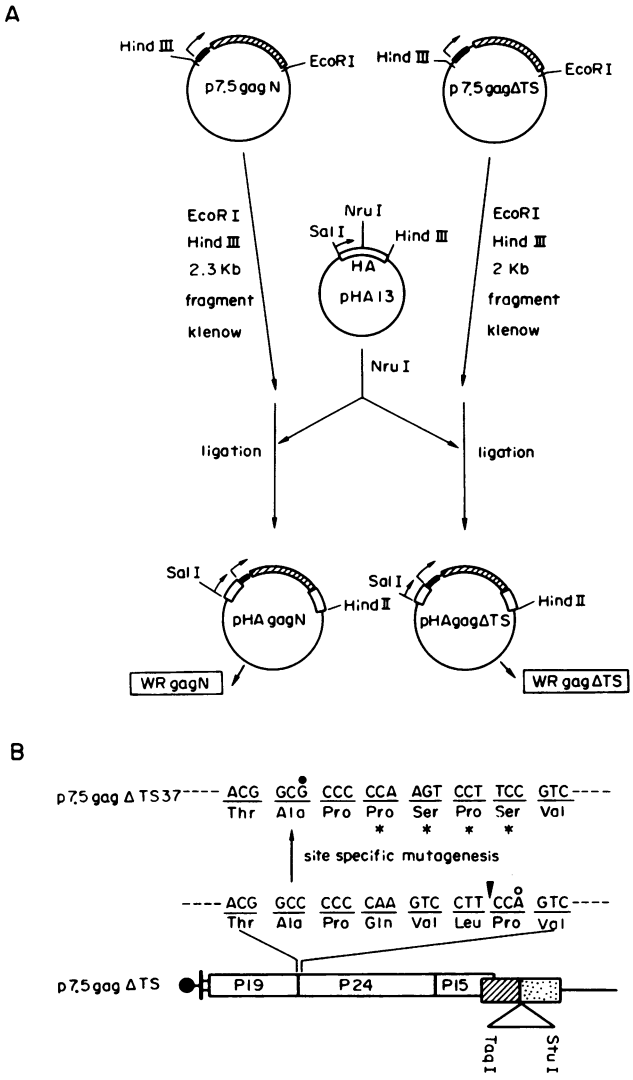

FIG. 4. Schematic representation of plasmid constructions for establishment of enzyme-substrate reaction system in vivo. (A) Construction of plasmid vectors for transfer of the *gag* and protease-coding sequence of HTLV-I into vaccinia virus. Symbols: ▭, hemagglutinin gene of vaccinia virus; ▨, *gag* and protease-coding sequence of HTLV-I; ▬, 7.5K gene promoter derived from vaccinia virus; the direction of transcription. (B) Scheme for construction of plasmid p7.5gagΔTS37. Symbols: ●, 7.5K gene promoter of vaccinia virus; ▨, truncated protease-coding sequence; ▩, 5′ portion of reverse transcriptase gene; ● or ○, insertion or deletion of nucleotide caused by oligonucleotide-directed site specific mutagenesis; *, exchanged amino acid residues caused by site-specific mutagenesis.

quence Pro-Ser-Pro-Ser, instead of the wild-type Gln-Val-Leu-Pro (Fig. 4B). This allowed the processed *gag* products directed by p7.5gagΔTS37 and those derived from WRgagN to be distinguished.

Western blotting analysis showed that p7.5gagΔTS37 directed the synthesis of a 53K *gag* precursor polyprotein with no evidence of the natural *gag* proteins (Fig. 5A, lane 3). When p7.5gagΔTS37 was introduced into CV-I cells infected with WRgagN, a new 43K protein was identified, although p19, p24, and p15 derived from WRgagN were evident (lane 1). This protein was no longer detected when WRgagΔTS was used as the helper virus (lane 2). We presume that the 43K protein represents a processed mutant *gag* precursor polyprotein that contains p19 plus p24 but not p15. The identity of the 43K protein was further examined by using individual anti-*gag* monoclonal antibodies. As expected, the 43K protein was recognized by anti-p19 monoclonal antibody (GIN7) and anti-p24 monoclonal antibody (NORI), (Fig. 5B, lanes 1 and 2). Anti-p15 monoclonal antibody (FR45) failed to recognize this protein (lane 3). Thus, we conclude that HTLV-I protease encoded in the *pro* ORF is able to process the *gag* precursor polyprotein in *trans*.

Importance of aspartic acid residues in HTLV-I protease.

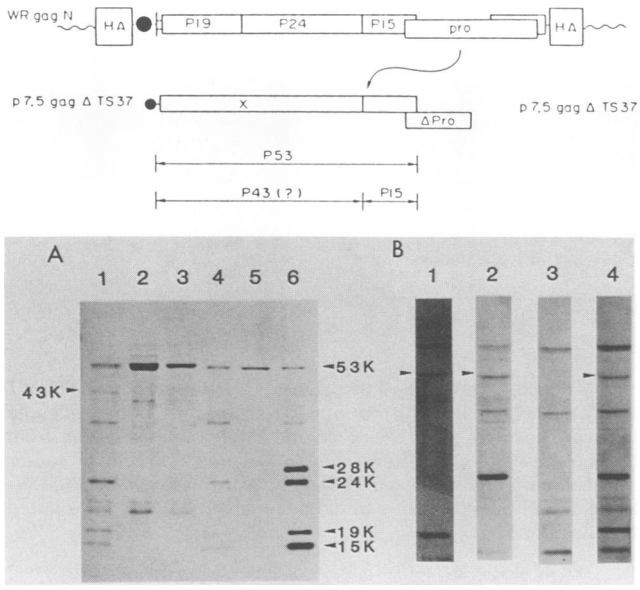

FIG. 5. Identification and characterization of 43K protein generated by the enzyme-substrate reaction in vivo. (A) Identification of the 43K processed protein from the mutant *gag* precursor polyprotein by an in *trans* action of HTLV-I protease. The substrate donor, p7.5gagΔTS37, was expressed in CV-I cells by using wild-type vaccinia virus as the helper virus (lane 3) or was expressed by using recombinant vaccinia virus harboring the protease-coding sequence (WRgagN) as the helper virus (lane 1). A protease-coding sequence-defective recombinant vaccinia virus (WRgagΔTS) was also used as the helper virus for the control experiment (lane 2). The cell extracts were separated on an SDS–12% polyacrylamide gel, and subsequent Western blotting analysis was carried out by using mixed anti-*gag* monoclonal antibodies of GIN7, NORI, and FR45. Production of the active form of the HTLV-I protease molecule from recombinant vaccinia virus was judged by the correct processing events of 53K *gag* precursor polyproteins (lanes 4 and 5). MT-2 cell lysate was used to show the correct mobilities of authentic mature *gag* proteins of HTLV-I in SDS-polyacrylamide gel electrophoresis (lane 6). (B) Characterization of the antigenicity of the 43K protein. Proteins expressed from p7.5gagΔTS37 with WRgagN as the helper virus were separated by SDS-polyacrylamide gel electrophoresis (12% polyacrylamide). Subsequent Western blotting analysis was carried out by using the individual anti-*gag* monoclonal antibodies, GIN7, NORI, and FR45, respectively (lanes 1, 2, and 3), or a mixture (lane 4). Arrows indicate 43K proteins reacted with GIN7 or NORI. The possible processing event of the mutant *gag* precursor polyprotein is schematically shown at the top of the figure. × indicates the disruption of the conserved amino acid sequence for specific proteolysis. ∿, Vaccinia virus genome.

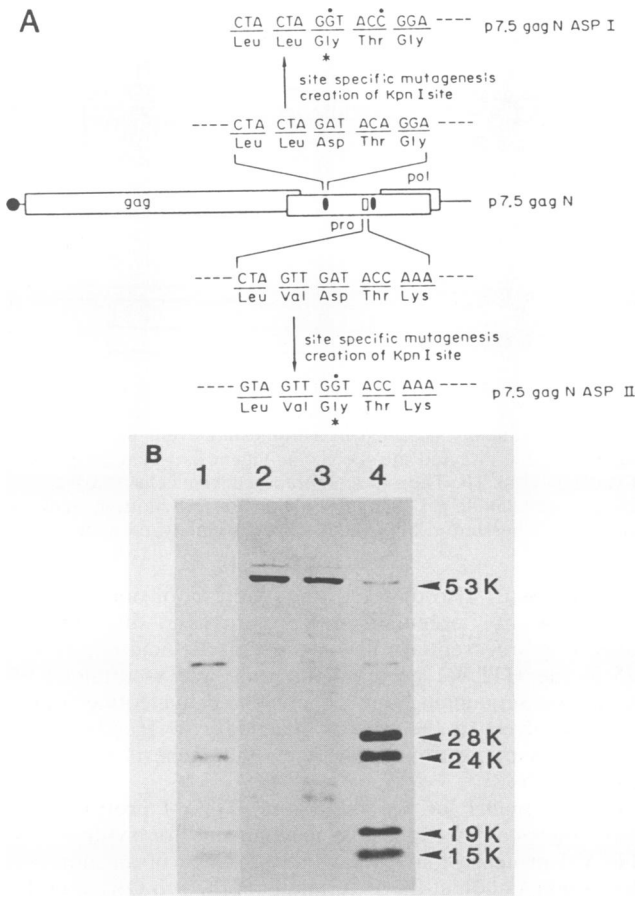

FIG. 6. Inhibition of proteolytic activity of HTLV-I protease resulting from the exchange of the putative catalytic residues by in vitro mutagenesis. (A) Oligonucleotide-directed site-specific mutagenesis for generating p7.5gagNAspI and p7.5gagNAspII. Symbols: ●, 7.5K gene promoter of vaccinia virus; ⬮, two conserved amino acid stretches between the retroviral proteases and cellular enzyme group of aspartyl proteases; □, short amino acid stretch conserved only among HTLV-family retroviral proteases; ●, nucleotide transitions caused by oligonucleotide-directed site-specific mutagenesis; *, exchanged amino acid residues resulting from the transitions of nucleotides; —, pUC DNA sequence. (B) Expression of wild-type and mutant plasmids. The extracts, prepared from cells transfected with p7.5gagN (lane 1), p7.5gagNAspI (lane 2), and p7.5gagNAspII (lane 3), were analyzed on an SDS–12% polyacrylamide gel, and then Western blotting analysis was carried out by using mixed anti-*gag* monoclonal antibodies of GIN7, NORI, and FR45. MT-2 cell lysate was used as the marker of the authentic mature *gag* products (lane 4).

Alignment of the amino acid sequences of HTLV-I and other retroviral proteases revealed that amino acid sequences are strictly conserved in two stretches, L-L-D-T-G and I-I-G-R-D (16, 26). We have shown with the evidence above that the region containing these conserved amino acid sequences plays an important role in the expression of proteolytic activity (Fig. 3).

The fact that L-L-D-T-G is found in the group of aspartyl proteases, which includes pepsin and others, suggests that the aspartic acid residue located in this sequence might be the catalytic center for HTLV-I protease. To identify the role of the aspartic acid positioned at residue 64 in the catalytic activity of HTLV-I protease, we performed site-specific mutagenesis to change the amino acid from aspartic acid to glycine. The resultant mutant plasmid, p7.5gagNAspI (Fig. 6A), directed the synthesis of the 53K *gag* precursor polyprotein but no other specifically processed proteins (Fig. 6B, lane 2), providing further evidence that HTLV-I protease belongs to the group of aspartyl proteases.

In the retroviral proteases of the HTLV family, an additional short sequence is conserved between the two well-defined conserved amino acid sequences (89). This short sequence, L-V-D-T (HTLV-I protease sequence positions 123 through 126), partially resembles that surrounding the catalytic residue, aspartic acid 64. To examine the biological significance of the additional short conserved sequence in the catalytic activity of the protease, we performed site-specific mutagenesis out to change the aspartic acid to glycine at amino acid 125 (Fig. 6A). The mutant plasmid

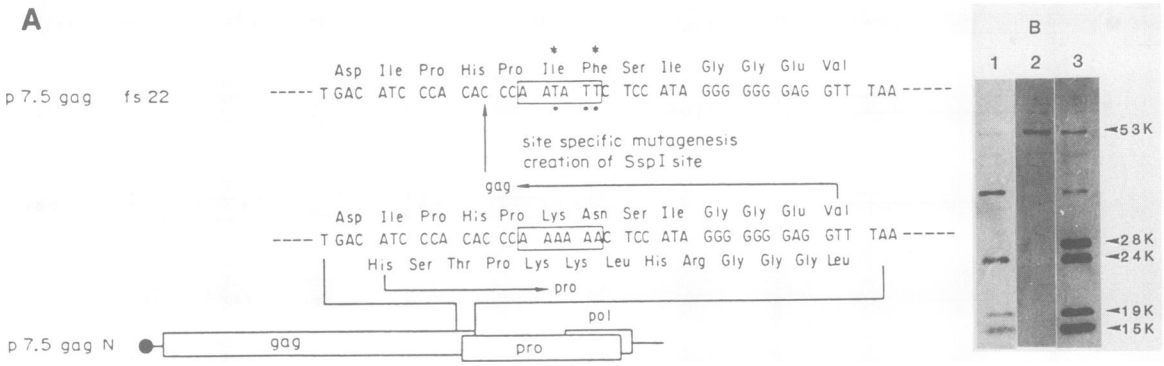

FIG. 7. Abolition of the correct processing of *gag* precursor polyprotein by blocking the frameshift event for the synthesis of HTLV-I protease. (A) Oligonucleotide-directed site-specific mutagenesis for generating p7.5gagfs19. The box encloses six consecutive adenine residues presumably involved in a frameshift event for synthesis of HTLV-I protease. Symbols: •, nucleotide transitions resulting from oligonucleotide-directed site specific mutagenesis; *, exchanged amino acid codon caused by transition of nucleotides; ●, 7.5K gene promoter of vaccinia virus. (B) The extracts, prepared from cells transfected with p7.5gagN (lane 2) and p7.5gagfs22 (lane 1), were analyzed on an SDS–12% polyacrylamide gel, and subsequent Western blotting analysis was carried out with mixed anti-*gag* monoclonal antibodies. *gag*-related proteins synthesized in MT-2 cells were also analyzed under the same conditions (lane 3).

p7.5gagNAspII also directed the synthesis of only unprocessed 53K *gag* precursor polyprotein (Fig. 6B, lane 3), suggesting an essential role for the aspartic acid at position 125 in the HTLV-I protease function. We ascertained by nucleotide sequencing that no unwanted nucleotide transitions occurred in the mutant plasmids, p7.5gagAspI and p7.5gagNAspII, during the in vitro mutagenesis procedure (data not shown).

Possible model for biosynthesis of HTLV-I protease. We have suggested two possible mechanisms for synthesis of HTLV-I protease that explain the absence of an initiation methionine codon at the 5′ terminus of the *pro* ORF and the difference in its reading frame from that of the *gag* gene (26): one is that the mature mRNA encoding the protease is generated by splicing, and the other is that the *gag* termination codon is not read owing to an upstream frameshift. Correct processing of the *gag* precursor polyprotein in the vaccinia virus expression vector system strongly suggested the frameshift mechanism, because vaccinia virus propagates exclusively in the cytoplasm of the host cells and does not possess its own machinery for splicing (46). Recently, ribosomal frame shifting in mouse mammary tumor virus has been demonstrated (11, 15). To obtain more direct evidence for frame shifting, we introduced mutations of three of the six consecutive adenines (nucleotide positions 2064 through 2069) located in the overlapping region of the *gag* and *pro* ORF. The mutant plasmid obtained was designated p7.5gagfs22 (Fig. 7). These consecutive adenine residues were expected to be involved in putative frame shifting by analogy with other retroviruses such as bovine leukemia virus. In mutant p7.5gagfs22, three A-to-T transitions were generated without disruption of the phase of the *gag* gene (Fig. 7A). The result of Western blotting analysis is shown in Fig. 7B. Processing of *gag* precursor polyprotein was no longer observed (lane 2), indicating that these residues are essential for generation of active protease. One may speculate that the mature protease is generated as follows: a frame shift gives rise to a *gag-pro* fusion protein with protease activity, which is expected to be 76 kilodaltons in size. This *gag-pro* fusion protein is rapidly cleaved by its own or a cellular protease to generate more free protease molecules that can act in *trans* on the 53K *gag* precursor polyprotein. To substantiate this scheme, it is important to demonstrate the presence of the 76K precursor polyprotein and then to identify which protease is involved in its cleavage. In HTLV-I-producing cells, the 76K protein is hardly detected, probably because of inefficient frame shifting and/or rapid processing. To mimic the naturally occurring ribosomal frameshifting event, we introduced a 1-base insertion to align the coding frames of the *gag* and *pro* genes. The insertion was at the consecutive adenine residues located in the 3′ terminal region of *gag* gene of the protease-negative or protease-positive plasmids (p7.5gagNAspI, p7.5gagNAspII, and p7.5gagN) (Fig. 8 and 9). The resultant plasmids were named p7.5gagfs19AspI p7.5gagfs19AspII, and p7.5gagfs19. Although 76K proteins immunoreactive with anti-*gag* monoclonal antibodies were produced in the cells transfected by p7.5fs19AspI and p7.5gagfs19AspII (Fig. 9B, lanes 1 and 2), p7.5gagfs19 directed and synthesis of the processed products of p19 and p24 (Fig. 8B, lane 1). The p15 (or related peptide) was not detected by the anti-p15 monoclonal antibody used in this study, probably because of the conformational change induced by the mutation.

These results can be interpreted to mean that the HTLV-I-encoded protease can trigger the first cleavage event for its own release form the 76K *gag-pro* precursor form. However, we cannot exclude the possibility that a cell protease initially activates the viral protease by cleaving it from an inactive precursor. To verify that the 76K protein band is not artifactually cross-reacting with anti-*gag* monoclonal antibodies, we constructed plasmids with serial truncations from the 3′ terminus of the *pro* gene (Fig. 9A). The size of the programmed products, 68K and 65K (directed by p7.5gagfs19AspIΔOE and p7.5gagfs19AspIIΔKp, respectively; Fig. 9A and B, lanes 3 and 4), were as expected for the truncated *gag* polypeptides. These results confirmed that 76K protein observed is the *gag*-pro precursor polyprotein.

DISCUSSION

Retroviral proteases responsible for processing *gag* precursor polyproteins play an important role in the production of infectious virus particles. In a previous communication, we reported that the HTLV-I genome cloned from a virus-producing cell line MT-2 harbored an ORF for the putative HTLV-I protease (26). In this study we examined the biological function and properties of HTLV-I *pro* ORF by using the vaccinia virus expression vector. This system is

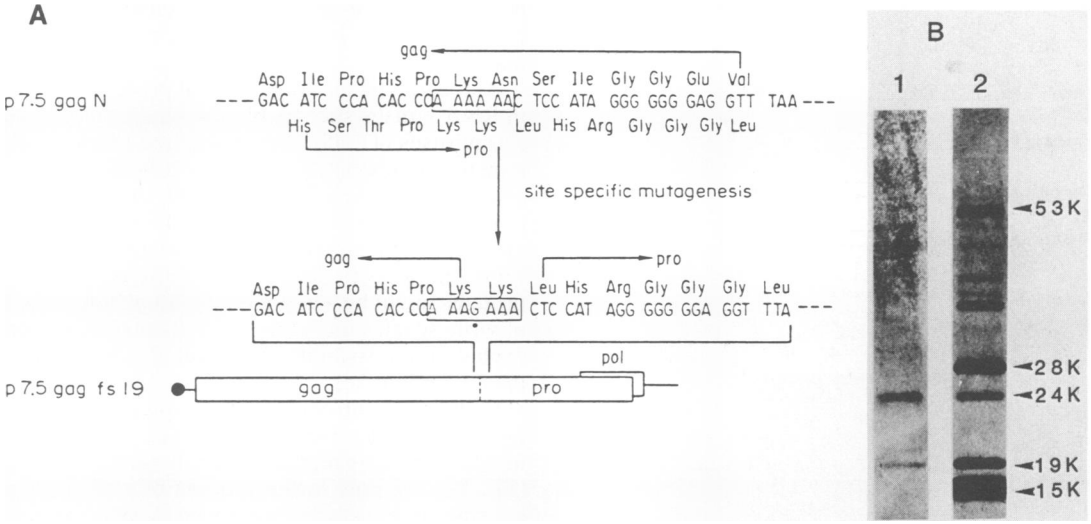

FIG. 8. (A) Correctly processed *gag* proteins were generated from the *gag*-protease fusion protein. The box encloses six consecutive adenine residues of possible frameshift site for synthesis of *gag-pro* fusion protein. Symbols: •, one nucleotide insertion introduced by oligonucleotide-directed site-specific mutagenesis; ●, junction between the *gag* and protease-coding sequence in the fusion protein. (B) Proteins expressed from p7.5gagfs19 was separated on an SDS–12% polyacrylamide gel, and then Western blotting analysis was carried out with mixed anti-*gag* monoclonal antibodies (lane 1). MT-2 cell lysate was used as the marker to show the migration pattern of the authentic mature *gag* proteins of HTLV-I (lane 2).

characterized by a high frequency of DNA transfection and by the capacity to efficiently express foreign gene products in eucaryotic cells.

We used a combination of two approaches to demonstrate that the *pro* ORF is the biologically functional protease gene of HTLV-I. The first method involved deletion of appropriate regions in the putative protease-coding sequence. The results revealed that deletion of a region containing the amino acid sequences strictly conserved among all retroviral proteases resulted in loss of proteolytic activity (Fig. 3). The second method is the establishment of an enzyme-substrate reaction system in vivo. We showed that the processing defect of the *gag* precursor polyprotein, directed by the *pro* ORF deletion mutant, could be complemented by cointroduction of a recombinant DNA constructed to have wild-type *pro* ORF. This result provides direct evidence that the *pro* ORF is the biologically functional HTLV-I protease gene and that the active form of the HTLV-I protease molecule processes the *gag* precursor polyprotein in *trans* (Fig. 5).

A significant structural difference between HTLV-I protease and other retroviral proteases is the length of the *pro* ORF. The 3′ region of the putative HTLV-I protease, extending from the second conserved amino acid stretch to the carboxy-terminal residue, is more than 46 amino acids longer than the analogous region of other members of the HTLV family (26). The data presented in this paper demonstrate that this region is dispensable for proteolytic activity of the HTLV-I protease, despite the high level of amino acid homology found in this region among the members of the HTLV family (Fig. 3). The biological function of this region remains unclear.

A survey of the amino acid sequence of the HTLV-I protease revealed the presence of the two strictly conserved regions common to the retroviral proteases. The unique core sequence, D-T-G, constituting one of the conserved sequences, was identical to that of aspartyl protease, which has been predicted as the catalytic center for protease activity (30, 44). We introduced a mutation in this aspartic acid, at position 64 in HTLV-I protease, and found that the mutation completely abolished cleavage. Independent evidence suggesting that the retroviral protease is an aspartyl protease has recently been obtained enzymatically with aspartyl protease inhibitors (16). Furthermore, we have found that one Asp-to-Gly change at amino acid position 125 in HTLV-I protease also abolished the processing of the *gag* precursor. The aspartic acid at amino acid 125 is located within a short amino sequence L-V-D-T, which is not found among the proteases of animal type B, C, or D retrovirus or lentivirus or among cellular aspartyl proteases. Interestingly, this sequence, L-V-D-T, or its homolog is conserved only in the protease of the HTLV, family including HTLV-I, HTLV-II, and bovine leukemia virus (34, 49). Molecular aspects of HTLV-I protease biosynthesis have not been elucidated so far. The coding frame of the *pro* gene is different from that of the *gag* gene, and no methionine codon was found at the 5′ terminal of the *pro* ORF. Therefore, either mRNA processing or translational frame shifting would be required to produce functional protease. Correct processing of the *gag* precursor polyprotein in the vaccinia virus expression vector system strongly suggests that a frameshift mechanism would be involved in protease synthesis, because this system does not support the splicing event of mRNA. In fact, characteristic nucleotide sequences were found in the overlapping region between the *gag* and *pro* genes, which contains a guanine-plus-cytosine-rich hairpin structure and six consecutive adenine residues. Both characteristic sequences have been thought to be important for ribosomal frame shifting on mRNA in procaryotes and eucaryotes (1, 5, 19). We directed site-specific mutations into the six consecutive adenine residues to create a new sequence, AATATT. The processing defect of this mutant may be caused by the blocking of ribosomal frame shifting responsible for synthesis of the protease, suggesting that the *gag-pro* fusion protein would be translated as a precursor of a protease via ribosomal frame shifting. Our data are compatible with those for mouse mammary tumor virus (11, 15).

The *gag-pro* precursor polyprotein is expected to be 76 kilodaltons in size. However, a 76K protein has not been

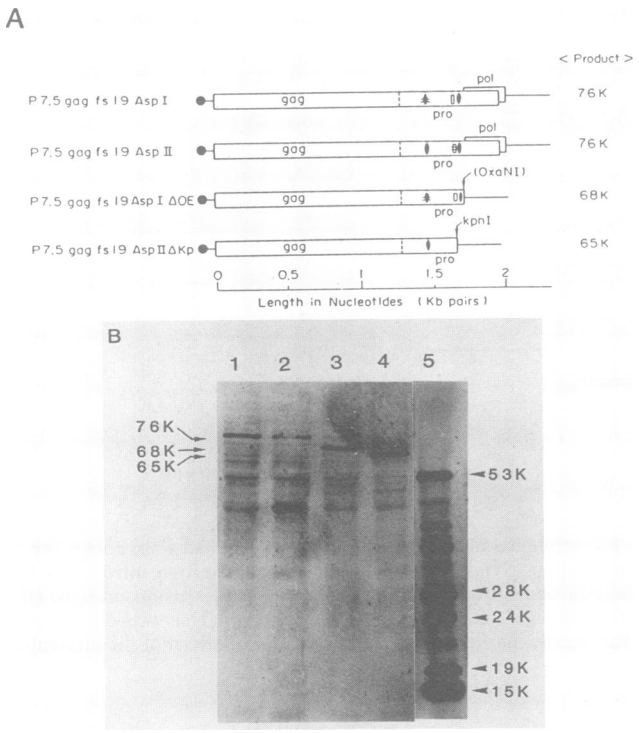

FIG. 9. Identification of 76K *gag*-protease fusion protein. (A) Schematic representation of the mutant plasmids which express *gag*-protease fusion proteins. Symbols: ●, 7.5K gene promoter of vaccinia virus; ⬮, two conserved amino acid stretches between the retroviral proteases and the cellular enzyme group of aspartyl protease; □, short amino acid residues conserved only among HTLV-family retroviral proteases; ◆ or ⊟, in vitro-mutagenized amino acid residues described in Fig. 6A; ¦, junction between the *gag* and protease-coding sequences in fusion proteins; —, pUC DNA sequence. (B) Analysis of *gag*-protease fusion proteins expressed from the mutants by Western blotting. Proteins extracted from transfected cell with p7.5gagfs19AspI (lane 1), p7.5gagfs19AspII (lane 2), p7.5gagfs19AspIΔOE (lane 3), and p7.5gagfs19AspIIΔKp (lane 4) were separated on an SDS–12% polyacrylamide gel, transferred to a nitrocellulose filter, and probed with mixed anti-*gag* monoclonal antibodies of GIN7, NORI, and FR45. MT-2 cell lysate was used as the marker of authentic mature *gag* proteins (lane 5).

identified in the chronically infected cell line. In vitro translation system programmed with virion-associated RNA allowed the detection of trace amounts of a 76K protein by immunoprecipitation with serum from an adult T-cell leukemia patient (18). Considering that the virion-associated RNA directs only the synthesis of *gag* gene-related products in the in vitro translation system, we presume that the 76K protein observed in in vitro-translated products is the *gag-pro* fusion protein. Two reasons for failure to detect this precursor are conceivable: rapid self-cleavage of the frameshift product or the inefficiency of ribosomal frame shifting for synthesis of *gag-pro* precursor polyproteins, or both.

To examine the action of protease on the 76K *gag-pro* precursor polyprotein, we introduced an insertion mutation to align both frames of the *gag* and *pro* genes, producing a detectable amount of this precursor. Although this *gag-pro* fusion protein is not exactly the same as the precursor protein produced by ribosomal frame shift, it should be very similar (Fig. 9). Correct processing of the 76K protein could be observed only when it retained an active protease domain (Fig. 8). This result strongly suggests that HTLV-I protease is capable of triggering the first cleavage event for its own release from the *gag-pro* precursor, probably in a *cis*-acting manner.

Autocatalytic cleavage of viral protease has been reported in RNA viruses such as picornaviruses (7, 28, 40) but not in retroviruses (47, 48). According to a recent study, however, human immunodeficiency virus protease expressed in *Saccharomyces cerevisiae* and *E. coli* exhibited autoprocessing of the *gag* precursor model (4, 21). Our results with HTLV-I protease also support the autocatalytic cleavage model. Thus, it appears likely that self-cleavage of constituent proteins might be a common phenomenon in the broad range of virus species including retroviruses. However, we cannot exclude the possibility that a cellular protease activates the inactive precursor.

Deduced amino acid sequence alignment of *gag-pro* fusion protein among retroviruses implies that protease might be involved in virus replication or integration into host genome (17). We are now in the process of establishing an in vitro enzyme assay system involving purified protease and the purified *gag* precursor product of HTLV-I. The system should provide us with the precise enzymatic properties of HTLV-I protease. Such knowledge will facilitate the development of effective antiviral therapeutic agents.

ACKNOWLEDGMENTS

We thank Y. Ito and M. Maki for making helpful suggestions and K. Igarashi, Takeda Chemical Industries, Ltd., for synthesizing oligonucleotides.

This work was supported by grants-in-aid from the Ministry of Education, Science, and Culture of Japan.

ADDENDUM

Jacks et al. (14) have recently shown by amino acid sequencing and in vitro mutagenesis that the consecutive uridine residues located upstream of the putative guanine-plus-cytosine-rich hairpin structure are involved in ribosomal frame shifting for the synthesis of *gag-pol* fusion protein in human immunodeficiency virus type I. Their result is consistent with our data for locating the candidate site of ribosomal frame shift.

In addition, several groups have recently reported independent results which support an enzymatic role for the conserved Asp in the D-T-G sequence in retroviral proteases, homologous to the catalytic center in cellular aspartyl proteases. Changes of this Asp to Ala in the human immunodeficiency virus protease (25) and to Ile in the avian sarcoma-leukosis virus protease (20) have been shown to abolish cleavage of the human immunodeficiency virus and avian sarcoma-leukosis virus protease-containing viral precursor proteins, respectively. Thus, any one of three changes at the site (Asp to Gly, Ala, or Ile) appears to destroy protease activity in three different viral proteases.

LITERATURE CITED

1. **Atkins, F., R. F. Gesteland, B. R. Reid, and C. W. Anderson.** 1979. Normal tRNAs promote ribosomal frameshifting. Cell **18:** 1119–1131.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
3. **Cochran, M. A., C. Puckett, and B. Moss.** 1985. In vitro mutagenesis of the promoter region for a vaccinia virus gene: evidence for tandem early and late regulatory signals. J. Virol. **54:**30–37.
4. **Debouck, C., J. G. Gorniak, J. E. Strickler, T. D. Meek, B. W.**

Metcalf, and M. Rosenberg. 1987. Human immunodeficiency virus protease expressed in *Escherichia coli* exhibits autoprocessing and specific maturation of the *gag* precursor. Proc. Natl. Acad. Sci. USA **84**:8903–8906.

5. **Fox, T. D., and B. Weiss-Brummer.** 1980. Leaky $^{+}1$ and $^{-}1$ frameshift mutations at the same site in a yeast mitochondrial gene. Nature (London) **288**:60–63.
6. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus DNA. Virology **52**: 456–467.
7. **Hanecak, R., B. L. Semler, H. Ariga, C. W. Anderson, and E. Wimmer.** 1984. Expression of a cloned gene segment of poliovirus in *E. coli*: evidence for autocatalytic production of the viral proteinase. Cell **37**:1063–1073.
8. **Hattori, S., T. Kiyokawa, K. Inakawa, F. Shimizu, E. Hashimura, M. Seiki, and M. Yoshida.** 1984. Identification of *gag* and *env* gene products of human T-cell leukemia virus (HTLV). Virology **136**:338–347.
9. **Hinuma, Y., K. Nagata, M. Hanaoka, M. Nakai, T. Matsumoto, K. Kinoshita, S. Shirakawa, and I. Miyoshi.** 1981. Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. Proc. Natl. Acad. Sci. USA **78**: 6476–6480.
10. **Hiramatsu, K., J. Nishida, A. Naito, and H. Yoshikura.** 1987. Molecular cloning of the closed circular provirus of human T cell leukaemia virus type I: a new open reading frame in the *gag-pol* region. J. Gen. Virol. **68**:213–218.
11. **Hizi, A., L. E. Henderson, T. D. Copeland, R. C. Sowder, C. V. Hixson, and S. Oroszlan.** 1987. Characterization of mouse mammary tumor virus *gag-pro* gene products and the ribosomal frameshift site by protein sequencing. Proc. Natl. Acad. Sci. USA **84**:7041–7045.
12. **Iino, T., K. Takeuchi, S. H. Nam, H. Siomi, H. Sabe, N. Kobayashi, and M. Hatanaka.** 1986. Structural analysis of p28 adult T-cell leukaemia-associated antigen. J. Gen. Virol. **67**: 1373–1379.
13. **Inoue, J., T. Watanabe, M. Sato, A. Oda, K. Toyoshima, M. Yoshida, and M. Seiki.** 1986. Nucleotide sequence of the protease-coding region in an infectious DNA of simian retrovirus (STLV) of HTLV-I family. Virology **150**:187–195.
14. **Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. I. Barr, and H. E. Varmus.** 1988. Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. Nature (London) **331**:280–283.
15. **Jacks, T., K. Townsley, H. E. Varmus, and J. Majors.** 1987. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus *gag*-related polyproteins. Proc. Natl. Acad. Sci. USA **84**:4298–4302.
16. **Katoh, I., T. Yasunaga, Y. Ikawa, and Y. Yoshinaka.** 1987. Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor. Nature (London) **329**:654–656.
17. **Katoh, I., Y. Yoshinaka, A. Rein, M. Shibuya, T. Okada, and S. Oroszlan.** 1985. Murine leukemia virus maturation: protease region required for conversion from "immature" to "mature" core form and for virus infectivity. Virology **145**:280–292.
18. **Kobayashi, N., N. Yamamoto, Y. Koyanagi, J. Schneider, G. Hunsmann, and M. Hatanaka.** 1984. Translation of HTLV (human T-cell leukemia virus) RNA in a nuclease-treated rabbit reticulocyte system. EMBO J. **3**:321–325.
19. **Kohno, T., and J. R. Roth.** 1978. A *Salmonella* frameshift suppressor that acts at runs of A residues in the messenger RNA. J. Mol. Biol. **126**:37–52.
20. **Kotler, M. R., A. Katz, and A. M. Skalka.** 1988. Activity of avian retroviral protease expressed in *Escherichia coli*. J. Virol. **62**:2696–2700.
21. **Kramer, R. A., M. D. Schaber, A. M. Skalka, K. Ganguly, F. Wong-Staal, and E. P. Reddy.** 1986. HTLVIII *gag* protein is processed in yeast cells by the virus *pol* protease. Science **231**: 1580–1584.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Mori, K., H. Sabe, H. Siomi, T. Iino, A. Tanaka, K. Takeuchi, K. Hirayoshi, and M. Hatanaka.** 1987. Expression of a provirus of human T cell leukaemia virus type I by DNA transfection. J. Gen. Virol. **68**:499–506.
25. **Mous, J., E. P. Heimer, and S. T. J. LeGrice.** 1988. Processing protease and reverse transcriptase from human immunodeficiency virus type I polyprotein in *Escherichia coli*. J. Virol. **62**: 1433–1436.
26. **Nam, S. H., and M. Hatanaka.** 1986. Identification of a protease gene of human T-cell leukemia virus type I (HTLV-I) and its structural comparison. Biochem. Biophys. Res. Commun. **139**: 129–135.
27. **Oroszlan, S., and T. D. Copeland.** 1985. Primary structure and processing of *gag* and *env* gene products of human T-cell leukemia virus HTLV-I_{CR} HTLV-I_{ATK}. Curr. Top. Microbiol. Immunol. **115**:221–223.
28. **Palmenberg, A. C., and R. R. Rueckert.** 1982. Evidence for intramolecular self-cleavage of picornaviral replicase precursors. J. Virol. **41**:244–249.
29. **Pearl, L. H., and W. R. Taylor.** 1987. Sequence specificity of retroviral proteases. Nature (London) **328**:482.
30. **Pearl, L. H., and W. R. Taylor.** 1987. A structural model for the retroviral proteases. Nature (London) **329**:351–354.
31. **Poiesz, B. J., F. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna, and R. C. Gallo.** 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. Proc. Natl. Acad. Sci. USA **77**:7415–7419.
32. **Poiesz, B. J., F. W. Ruscetti, M. S. Reitz, V. S. Kalyanaraman, and R. C. Gallo.** 1981. Isolation of a new type C retrovirus (HTLV) in primary uncultured cells of a patient with Sezary T-cell leukemia. Nature (London) **294**:268–271.
33. **Ratner, L., S. F. Josephs, B. Starcich, B. Hahn, G. M. Shaw, R. C. Gallo, and F. Wong-Staal.** 1985. Nucleotide sequence analysis of a variant human T-cell leukemia virus (HTLV-Ib) provirus with a deletion in pX-I. J. Virol. **54**:781–790.
34. **Sagata, N., T. Yasunaga, J. Tsuzuku-Kawamura, K. Ohishi, Y. Ogawa, and Y. Ikawa.** 1985. Complete nucleotide sequence of the genome of bovine leukemia virus: its evolutionary relationship to other retroviruses. Proc. Natl. Acad. Sci. USA **82**:677–681.
35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
36. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
37. **Seiki, M., S. Hattori, Y. Hirayama, and M. Yoshida.** 1983. Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. Proc. Natl. Acad. Sci. USA **80**:3618–3622.
38. **Shida, H.** 1986. Nucleotide sequence of the vaccinia virus hemagglutinin gene. Virology **150**:451–462.
39. **Shida, H., T. Tochikura, T. Sato, T. Konno, K. Hirayoshi, M. Seki, Y. Ito, M. Hatanaka, Y. Hinuma, M. Sugimoto, F. Takahashi-Nishimaki, T. Maruyama, K. Miki, K. Suzuki, M. Morita, H. Sashiyama, and M. Hayami.** 1987. Effect of the recombinant vaccinia virus that express HTLV-I envelope gene on HTLV-I infection. EMBO J. **6**:3379–3384.
40. **Strebel, K., and E. Beck.** 1986. A second protease of foot-and-mouth disease virus. J. Virol. **58**:893–899.
41. **Tanaka, Y., Y. Koyanagi, T. Chosa, N. Yamamoto, and Y. Hinuma.** 1983. Monoclonal antibodies reactive with both p28 and p19 of adult T-cell leukemia virus-specific polypeptides. Gann **74**:327–330.
42. **Tanaka, Y., B. Lee, T. Inoi, H. Tozawa, N. Yamamoto, and Y. Hinuma.** 1986. Antigens related to three core proteins of HTLV-I (p24, p19 and p15) and their intracellular localizations, as defined by monoclonal antibodies. Int. J. Cancer **37**:35–42.
43. **Taylor, J. W., J. Ott, and F. Eckstein.** 1985. The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. Nucleic Acids Res. **13**:

8765–8785.

44. **Toh, H., R. Kikuno, H. Hayashida, T. Miyata, W. Kugimiya, S. Inoue, S. Yuki, and K. Saigo.** 1985. Close structural resemblance between putative polymerase of a *Drosophila* transposable genetic element 17.6 and *pol* gene product of Moloney murine leukaemia virus. EMBO J. **4**:1267–1272.
45. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
46. **Venkatesan, S., B. M. Baroudy, and B. Moss.** 1981. Distinctive nucleotide sequences adjacent to multiple initiation and termination site of early vaccinia virus gene. Cell **125**:805–813.
47. **Vogt, V. M., R. Eisenman, and H. Diggelmann.** 1975. Generation of avian myeloblastosis virus structural proteins by proteolytic cleavage of a precursor polypeptide. J. Mol. Biol. **96**:471–493.
48. **vonder Helm, K.** 1977. Cleavage of Rous sarcoma viral polypeptide precursor into internal structural proteins *in vitro* involves viral protein p15. Proc. Natl. Acad. Sci. USA **74**:911–915.
49. **Yasunaga, T., N. Sagata, and Y. Ikawa.** 1986. Protease gene structure and *env* gene variability of the AIDS virus. FEBS Lett. **199**:145–150.
50. **Yoshida, M., I. Miyoshi, and Y. Hinuma.** 1982. Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. Proc. Natl. Acad. Sci. USA **79**:2031–2035.
51. **Yoshinaka, Y., I. Katoh, T. D. Copeland, and S. Oroszlan.** 1985. Translational readthrough of an amber termination codon during synthesis of feline leukemia virus protease. J. Virol. **55**:870–873.
52. **Yoshinaka, Y., I. Katoh, T. D. Copeland, and S. Oroszlan.** 1985. Murine leukemia virus protease is encoded by the *gag-pol* gene and is synthesized through suppression of an amber termination codon. Proc. Natl. Acad. Sci. USA **82**:1618–1622.
53. **Yoshinaka, Y., I. Katoh, T. D. Copeland, G. W. Smythers, and S. Oroszlan.** 1986. Bovine leukemia virus protease: purification, chemical analysis and in vitro processing of *gag* precursor polyproteins. J. Virol. **57**:826–832.