Journal of Cellular Biochemistry 40:15–30 (1989)
Cellular Proteases and Control Mechanisms 101–116

Identification of HTLV-I Gag Protease and Its Sequential Processing of the Gag Gene Product

Masakazu Hatanaka and Seok Hyun Nam

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

The full-length provirus of human T-cell leukemia virus type I (HTLV-I) was isolated from MT-2, a lymphoid cell line producing HTLV-I. In transfected cells, structural proteins of HTLV-I, the gag and env products, were formed and processed in the same manner as observed in MT-2 cells. The nucleotide sequence was determined for a region between the *gag* and *pol* genes of the proviral DNA clone containing an open-reading frame. The deduced amino acid sequences show that this open-reading frame encodes a putative HTLV-I protease.

The protease gene (*pro*) of HTLV-I was investigated using a vaccinia virus expression vector. Processing of 53k *gag* precursor polyprotein into mature p19, p24, and p15 gag structural proteins was detectable with a recombinant plasmid harboring the entire *gag*- and protease-coding sequence. We demonstrated that the protease processed the *gag* precursor polyprotein in a *trans*-action. A change in the sequence Asp(64)-Thr-Gly, the catalytic core sequence among aspartyl proteases, to Gly-Thr-Gly was shown to abolish correct processing, suggesting that HTLV-I protease may belong to the aspartyl protease group. The 76k *gag-pro* precursor polyprotein was identified, implying that a *cis*-acting function of HTLV-I protease may be necessary to trigger the initial cleavage event for its own release from a precursor protein, followed by the release of p53 *gag* precursor protein. The p53 *gag* precursor protein is then processed by the *trans*-action of the released protease to form p19, p24, and p15.

Key words: retrovirus, adult T cell leukemia, aspartyl protease vaccinia virus, rivosomal frame shift, polyprotein transfection

It is known that virus-coded protease activity is essential in the formation of infectious retroviruses [1,2]. The complete nucleotide sequence reported in the literature, however, does not contain the open-reading frame for the HTLV-I protease, and the infectivity of the reported provirus of HTLV-I has not yet been established [3]. MT-2 cells are known to produce persistently HTLV-I viral particles and to have the ability to transform human lymphoid cells by co-cultivation. We determined that MT-2 cells contain various HTLV-I proviruses, including both defective and nondefective types [4]. A nondefective infectious proviral DNA of HTLV-I was isolated from MT-

Received April 26, 1988; accepted October 4, 1988.

2 cells, and the cells transfected by the cloned provirus were shown to produce the major proteins of HTLV-I and form syncytium [5]. We have found that the provirus isolated from MT-2 cells encodes a putative protease in an independent open-reading frame between the *gag* and *pol* genes. The molecular structure of the putative protease of HTLV-I was compared to those of other retroviruses obtained by sequence analysis [6]. The deduced amino acid sequences show significant homology with the known protease gene of other retroviruses and habor highly conserved amino acid sequences that are well conserved in other retroviral protease domains that had never been identified in the nucleotide sequence of HTLV-I in the original and other isolates [3,7]. Another full-length DNA clone of HTLV-I was independently isolated from extrachromosomal closed circular copies in chronically infected promyelocytic leukemia HL60 cells [8]. Both of the cloned provirus DNA harbor the same length nucleotide open-reading frame for putative HTLV-I protease except that a few base transitions occur in the reading frame. This strongly suggested that the open-reading frame encodes a biologically active HTLV-I protease. However, there is no direct evidence to support this, because sequence analysis of the authentic HTLV-I protease has been hampered by the scarcity of the proteolytic factor associated with virion. Furthermore, there are many reports on expression defect of viral gene products, caused by a few base transitions within the coding sequence of cloned virus DNA. Therefore, it is crucial to determine whether the open-reading frame is the protease gene of HTLV-I. We used the vaccinia virus expression vector system to characterize the biological function of the protease gene because of its high frequency of DNA transfection and subsequent efficient expression of the foreign genes in the transient expression assay. In this study, we provide evidence that proteolytic cleavage of *gag* precursor polyprotein does not depend on the action of cell-specific protease, but that it is accompanied by HTLV-I protease encoded in the virus gene. We also determined the biological properties of the protease, the possible mechanism for biosynthesis of HTLV-I protease, and the sequential processing of the precursor polyproteins in *cis*- and *trans*-acting manners.

MATERIALS AND METHODS

Materials

The restriction endonucleases, T4 ligase, Klenow fragment of DNA polymerase I, and M13 sequence kit were purchased from Takara Shuzo, Co., Ltd. Alpha-^{32}P-dCTP (3,000 Ci/mol) was obtained from Amersham, International, Amersham, England, and New England Nuclear, Bostom, MA. Other enzymes and chemicals were described previously [5].

Nucleotide Sequence Determination

The replication-competent HTLV-I provirus DNA was cloned from MT-2 cells and termed HTLV1C [5]. The appropriate fragments from HTLV1C DNA were sequenced by the procedure of Sanger et al. [9] with slight modification [10].

Homology Matrix Comparison and Sequence Alignment

Homology searches between amino acid sequences were made using the graphical matrix method, as described previously [11]. With the aid of these matrices, the sequences were aligned by manual inspection.

General DNA Method

Plasmid DNA was prepared by the alkali-sodium dodecyl sulfate (SDS) method of Birnboin and Doly [12] with slight modification. Restriction endonucleases and DNA-modified enzymes were used as specified by the manufacturers. All molecular biological manipulations were carried out by standard methods [13]. All DNA transformations were performed with competent bacteria of *E. coli* JM109. However, single-stranded DNAs for site-specific mutagenesis were produced by transformants of *E. coli* TG1 harboring subcloned M13RF DNAs or by infecting transformants of *E. coli* MV1184 harboring subcloned pUC118/119 DNAs with helper phage M13K07 as described by Vieira and Messing [14]. The sequencing of single-stranded M13 virion DNA or double-stranded plasmids was done by the chain-termination method described by Sanger et al. [9].

Construction of Plasmids

The detailed strategies for construction of various plasmids will be described elsewhere.

In Vitro Mutagenesis

Oligonucleotide-directed in vitro mutagenesis was performed using a mutagenesis kit supplied by Amersham, as described by Taylor et al. [15] except the reaction volume was reduced to one-half the original volume.

Recombinant Vaccinia Virus Preparation

The detailed procedure for the generation and purification of recombinant vaccinia virus was carried out as described by Shida et al. [16].

DNA Transfection

CV-I cells for transfection of plasmid DNA were maintained in Eagle's medium (MEM) 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 [17]. Briefly, 2–3 $\times$ 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 g of calcium phosphate precipitated recombinant plasmid DNA was added. After 30-min incubation at room temperature, fresh medium was added. Following 4 h incubation at 37°C, the medium was replaced by fresh medium and incubation was continued overnight.

Viral Protein Analysis

Twenty hours 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 95–100°C. Solubilized proteins were separated by SDS-polyacrylamide gel electrophoresis (12% polyacrylamide gel) and transferred onto nitrocellulose filters as described [18]. The blots were blocked in fetal calf serum at 4°C overnight, probed with the anti-*gag* monoclonal antibodies appropriately diluted with 3% bovine serum albumin solution containing 20 mM Tris, pH 8.0, 150 mM NaCl. After primary antibody reaction at room temperature for 3 h, secondary antibody reaction using rabbit anti-mouse IgG antibody was carried out under the same

conditions. Finally, protein A-Gold reaction and subsequent enhancement was carried out as specified by the manufacturer and the *gag*-related proteins expressed in CV-I cells transfected with the recombinant plasmids were revealed.

RESULTS AND DISCUSSION

Transfection of a Provirus DNA of HTLV-I

An HTLV-I provirus has been isolated from the leukemic cells of an ATL patient, and its sequence has been determined [3]. However, there has been no report of a complete provirus capable of producing its own viral products or virus particles. Therefore, we isolated the full-length provirus HTLV1C from the HTLV-I-producing cell line MT-2 and investigated its ability to form gene products in nonlymphoid cells. The results showed no significant difference in the production of *tat, gag,* and *env* gene products between nonlymphoid cell lines and MT-2 cells. In addition, the *trans*-acting activity in nonlymphoid cells was significantly high [5]. In HeLa cells, we observed syncytium formation. This is known to be caused by the *env* gene product of the virus particle. Processing of *gag* protein occurred in HOS cells [5], suggesting the expression of a protease gene similar to other retroviruses.

Identification of HTLV-I Protease Gene

A recombinant phage of the nondefective provirus was digested with *SstI* and *SmaI*. An 851-bp fragment produced by double-digestion of the restriction enzymes contained the *gag-pol* junction region. This fragment was redigested with appropriate restriction enzymes and subcloned into the polylinker site of M13 phage. This region was found to contain an open-reading frame of 702 bp that is in a different phase from the ones of *gag* and *pol* genes, spanning nucleotide 2052–2753 according to the sequence numbering by Seiki et al. [3]. This open-reading frame can encode a sequence of 234 amino acids that show a significant homology with the proteases of HTLV-II and BLV (Fig. 1). Moreover, there are two highly conserved amino acid regions that are identical with those of other retroviral protease domains [11]. The fact that the deduced amino acid sequence bears a significant homology with other retroviral proteases suggests that this open-reading frame possibly encodes a HTLV-I protease. Among these highly conserved amino acid regions, the second region contains a closely similar amino acid sequence around the active-site amino acids of the acid protease family, suggesting that the HTLV-I protease possesses a similar activity to those of acid proteases [11,19]. One additional short sequence, Leu-Val-Asp-Thr, is found at four amino acid sites before the second conserved region. This short sequence was conserved only in HTLV-I, HTLV-II, and BLV. Thus, it may be considered to be characteristic to the HTLV family proteases [19].

This protease-coding frame has not been identified in the nucleotide sequence of HTLV-I in the original and other isolates [3,7]. Recently, Inoue et al. [20] reported one nucleotide insertion at position 2298 after reexamining the sequence of HTLV-I ATK. In spite of the correction, this region is divided by two open-reading frames with an amber stop codon at position 2440. They suggested that HTLV-I protease may be synthesized by a translational readthrough of the amber codon in the middle of the frame or that the clone is biologically inactive [20]. In contrast, our HTLV1C was molecularly cloned from the HTLV-I-producing cell line MT-2 [4,5], and the HTLV1C DNA transfection experiment revealed the expression of the processed *gag*

```
HTLV-I    HSTPK KLHRG  GGLTS PPTLQ  QVF-- LNQDP  ASILP VIPLD
HTLV-II     GK KLLKG  GDLIS PHP--  ----- -DQD-  ISILP LIPLR
BLV       NPQIK KLI-E  GGLSA PQTVT  PITDP LSEAE  LECLL SIPLA

HTLV-I    PARRP VIKAQ  VDTQT SHPKT  IEA-- LLDTG  ADMTV LPIAL
HTLV-II   QQQQP ILGVR  ISVMG QTPQP  TQA-- LLDTG  ADLTV IPQTL
BLV       RSR-P SVAVY  LSGPW LQPSQ  NQALM LVDTG  AENTV LPQNW

HTLV-I    FSSNT PLKNT  SVLGA GGQTQ  DHFKL TSLPV  LI--- -RLPF
HTLV-II   VPGPV KLHDT  LILGA SGQTN  TQFKL LQTPL  HI--- -FLPF
BLV       LVRDY PRIPA  AVLGA GGVSR  NRYNW LQGPL  TLALK PEGPF

HTLV-I    RTTPI VLTSC  LVDTK NNWAI  IGRDA LQQCQ  GVLYL PEAKG
HTLV-II   RRSPV ILSSC  LLDTH NKWTI  IGRDA LQQCQ  GLLYL PDDPS
BLV       ITIP- ---KI  LVDTF DKWQI  LGRDV LSRLQ  ASISI PEEVR

HTLV-I    PPVIL PIQAP  AV-LG LEHLP  RPPEI SQFPL  NQNAS RPCNT
HTLV-II   PHQLL PIATP  NT-IG LEHLP  PPPQV DQFPL  NLSAS RP
BLV       PPMVG VLDAP  PSHIG LEHLP  VPPEV PQFPL  N

HTLV-I    WSGRP WRQAI  SNPTP GQEIT  QYSQL KRPME  PGDSS TTCGP

HTLV-I    LTL
```

Fig. 1. Alignments of amino acid sequences of the proteases and putative protease of HTLV-I. Amino acid residues are expressed as one-letter codes.

and *env* proteins and the syncytium formation on the transfected cells [5]. We compared our nucleotide sequence data with those of the corresponding region of HTLV-I ATK [3,20]. In the putative protease-coding frame, 11 nucleotides differ from the sequence of HTLV-I ATK, resulting in the change of six amino acids. Among the differences, the A to G transition at position 2440 is most noteworthy: the biologically active HTLV1C contains a tryptophan code, while the HTLV-I ATK clone has an amber codon. Apparently, another clone, HTLV-Ib, and the closely related simian retrovirus (STLV) have the same TGG codon in this region [7,20]. Taken together, the results suggest that the amber codon at position 2440 of HTLV-I ATK must have occurred by a point mutation (G to A) of the biologically active prototype HTLV-I, and, thus, the possibility that HTLV-I protease may be synthesized by suppressing the amber codon in the middle of the frame is excluded.

Within the open-reading frame of the putative HTLV-I protease gene, the first methionine residue is located at the 69th position from the amino-terminus; thus, the protease may be translated as a 76-kD gag-fused protein because we found the 76-kD protein in the cell extract or in the in vitro translation system when using total virion mRNA of HTLV-I [4].

Processing of Gag Precursor Polyprotein

The 2-kb RsaI-ApaI fragment harboring the entire *gag* and *pro* ORF was purified and introduced downstream of the p7.5k gene promoter, which expressed at both early and late times of the vaccinia virus infection cycle [21].

The constructed plasmid, p7.5gagN, as shown in Figure 2, was introduced into the 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 were

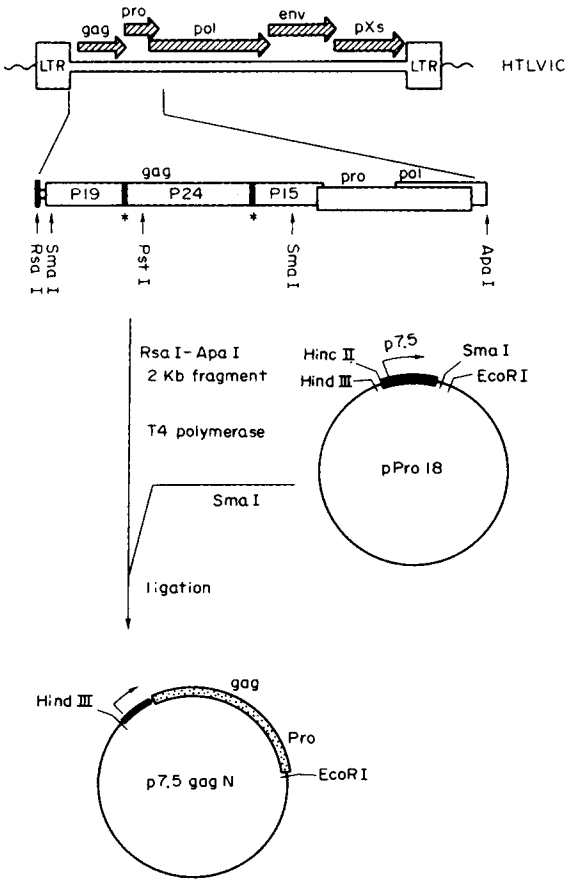

Fig. 2. Schematic representation of the HTLV-I genome and construction of p7.5gagN. The strategy used to insert the entire *gag*-and protease-coding sequence into the pPro18 expression vector is outlined. Symbols: striped arrows, *gag, pol, env,* and *pXs* genes of HTLV-I; solid bar, 7.5k gene promoter of vaccinia virus; stippled bar, *gag*- and protease-coding sequence; circle, pUC18 DNA sequence; *, the positions of the HTLV-I protease-mediated processing sites. Thin arrows show the direction of transcription.

examined by Western blotting analysis using the anti-*gag* monoclonal antibodies (Fig. 3).

Proteins 19k, 24k, and 15k were recognized by a mixture of anti-p19 monoclonal antibody (GIN7), anti-p24 monoclonal antibody (Nor-I), and anti-p15 monoclonal antibody (FR45), respectively. A small amount of 53k protein was recognized, which is the uncleaved *gag* precursor polyprotein. Protein 33k, which is considered to be the intermediate precursor protein consisting of p24 and p15 fused protein, was also detected.

All the *gag* proteins transiently expressed from the plasmid comigrated in SDS-polyacrylamide gel electrophoresis with those of corresponding HTLV-I *gag* proteins synthesized in HTLV-I-producing MT-2 cells or virion-associated *gag* proteins of HTLV-I. These results showed that HTLV-I *gag* gene products directed by cloned

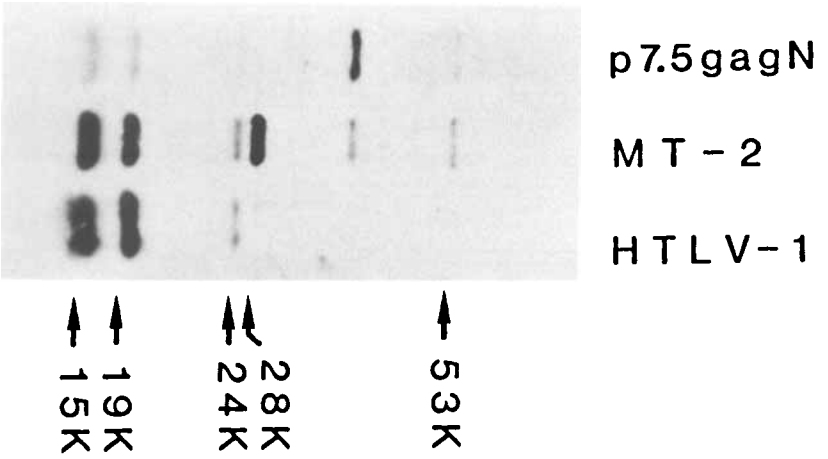

Fig. 3. Expression and processing of *gag* precursor polyprotein in CV-1 cells using vaccinia virus-derived expression vector. Plasmid p7.5gagN was constructed as shown in Figure 2. Proteins expressed from the plasmid were analyzed on a 12% SDS-polyacrylamide gel and electrophoretically transferred onto nitrocellulose filters. Transient expressions of *gag*-related proteins were analyzed by Western blotting analysis using anti-p19 monoclonal antibody (GIN7), anti-p24 monoclonal antibody (NOR-I), and anti-p15 monoclonal antibody (FR45), or their mixtures. Arrows indicate 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 p7.5gagN.

gag and *pro* ORF have the same antigenicity and molecular weights as those of processed mature *gag* proteins detected in chronically infected cells or virus particles.

Proteolytic Activity by the Putative Gene Product of proORF

Processing of *gag* precursor polyprotein is accomplished by HTLV-I protease encoded in proORF (Fig. 4). A deletion spanning from 2495 to 2611 did not abolish the correct processing of *gag* precursor polyprotein. However, processing was inhibited when the deletion was extended from 2228 to 2611 at the 5′ region of proORF, resulting in an accumulation of gag precursor polyprotein (Pr53gag). Translational termination of the protease by the insertion of 8 nt BglII linker at the 2495 site had no influence on correct processing of the 53k *gag* precursor polyprotein. These results indicate that a region spanning the nucleotide position 2228 to 2485 of proORF is important to process Pr53 *gag* into the mature *gag* products. In contrast, the downstream region from the 2495 site of proORF does not relate to the proteolytic function of protease. It is consistent with our previous data that the region spanning from nucleotide position 2228 to 2495 harbors the amino acid stretches that have been commonly conserved among proteases of the HTLV family retroviruses.

To examine further the involvement of strictly conserved amino acid stretches in the processing and maturation of *gag* protein, we constructed a plasmid that has the entire protease-coding sequence except for a region harboring one of the conserved amino acid stretches that was deleted from the coding sequence. The major immunoreactive band detected was 53k protein, which was not processed at all. The results

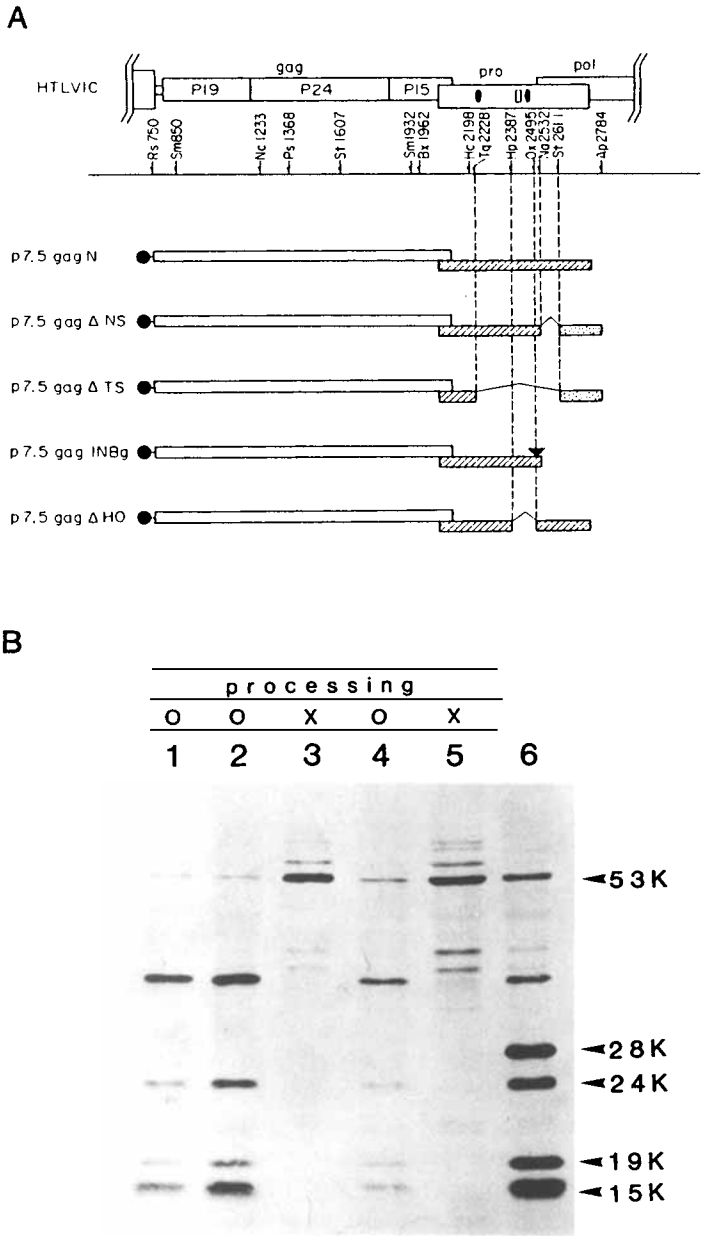

Fig. 4. Plasmids encoding the truncated protease molecules of HTLV-I and its expression. **A:** Map of part of the HTLV-I genome is given at the top. The nucleotide number indicated corresponds to that reported by Seiki et al. [3]. Several deletion mutants of the protease-coding sequence were constructed as described in Materials and Methods. Symbols: closed circles, 7.5k gene promoter of vaccinia virus; open bar, *gag*-coding sequence of HTLV-I; striped bar, protease-coding sequence of HTLV-I; ▼, insertion of 8nt phosphorylated BgIII synthetic linker at OxaNI 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, RsaI; Sm, SmaI; Nc, NcoI; Ps, PstI; St, StuI; Bx, BstXI; Hc, HincII; Tq, TaqI; Hp, HpaII; Ox, OxaNI; Na, NarI; Ap, ApaI. **B:** Western blotting analysis of *gag* proteins expressed from deletion mutants of protease-coding sequence. Evidence for processing is indicated by signals, represented by o (processing positive) or **X** (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. The extracts were prepared from cells transfected with p7.5gagN (**lane 1**), p7.5gagΔNS (**lane 2**), p7.5gagΔTS (**lane 3**), p7.5 gagINBg (**lane 4**), p7.5gagΔHO (**lane 5**), and MH2 cells (**lane 6**).

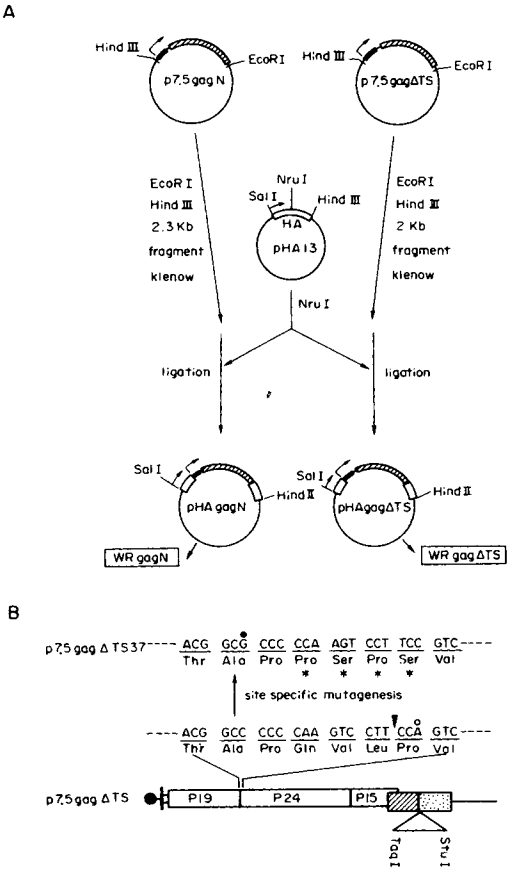

Fig. 5. Schematic representation of plasmid constructions for the establishment of enzyme-substrate reaction system in vivo. **A:** Construction of plasmid vectors for transfer of the *gag*- and protease-coding sequence of HTLV-I into vaccinia virus. Open bar indicates the hamagglutinin gene of vaccinia virus. Striped bar is the *gag*- and protease-coding sequence of HTLV-I. Closed bar indicates the 7.5k gene promoter derived from vaccinia virus. Arrows show the direction of transcription. **B:** Scheme for construction of plasmid p7.5gag△TS37. Symbols: close circle, 7.5k gene promoter of vaccinia virus; striped bar, truncated protease-coding sequence; stippled bar, 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.

indicated that HTLV-I protease is encoded in the proORF region and that conserved amino acid stretches play an important role in the expression of proteolytic activity.

Processing Defect of proORF-deletion Mutant Rescued by Supplying the Active Protease

Although it is certain that the processing event of the gag precursor polyproteins is co-related to the expression of the molecules encoded in the proORF, it was not demonstrated that HTLV-I protease cleaves the 53k *gag* precursor polyprotein directly. This could, however, be determined if the gag precursor polyprotein directed by the proORF-deletion mutant was able to be processed by a *trans*-action of a wild-

type protease. To elucidate this point, we attempted to establish an enzyme-substrate reaction system in CV-I cells. A DNA transfection experiment in vaccinia virus expression vector system is suitable for this kind of study because co-introduction of two independent recombinant DNA can be achieved with maximal efficiency by using the recombinant vaccinia virus as the helper virus. To make a distinction between the processed *gag* products of substrate and those derived from the protease-donor, site-specific mutagenesis was carried out to disrupt the essential amino acid sequence for proteolytic cleavage located between the p19 and p24 junction, that is, encoding Pro-Ser-Pro-Ser instead of Gln-Val-Leu-Pro (Fig. 5).

Western blotting analysis showed that a new 43k protein was identified together with the 15k protein. The identity of the 43k protein was examined by Western blotting analysis using anti-*gag* monoclonal antibodies, respectively (Fig. 6). As expected, the 43k protein was recognized by the anti-p19 monoclonal antibody (GIN 7), and anti-p24, by the monoclonal antibody (NorI). Anti-p15 monoclonal antibody (FR45) failed to recognize this protein. Thus, we concluded that proORF is the biologically functional HTLV-I protease gene (*pro*) and that HTLV-I protease is able to process the *gag* precursor polyprotein in a *trans* manner.

Biochemical Nature of HTLV-I Protease

Alignment of the amino acid sequence between 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 [6,22]. We have shown evidence that these conserved amino acid sequences play an important role in the expression of the proteolytic activity in this study.

The fact that conserved sequences are also found in the enzyme group of aspartyl proteases implies that aspartic acid residue located at L-L-D-T-G might be the catalytic center for HTLV-I protease action as it is for cellular aspartyl proteases such as pepsin. To identify the role of the aspartic acid positioned at 64 in the catalytic activity of HTLV-I protease, site-specific mutagenesis was carried out to exchange the amino acid from aspartic acid to glycine. We obtained the mutant plasmid p7.5gagN AspI (Fig. 7).

The major immunoreactive protein band was 53k *gag* precursor polyprotein, and no other specifically processed protein band was detectable. Therefore, it is suggested that HTLV-I protease may belong to the enzyme group of aspartyl protease. Among retroviral protease of the HTLV family, one additional short sequence is conserved between two well-defined conserved amino acid sequences [23]. This short sequence, represented by L-V-D-T, has partial resemblance to sequences surrounding the catalytic residue, aspartic acid 64. To examine the biological significance of the additional short conserved sequence in catalytic activity of protease, site-specific mutagenesis was carried out to exchange the amino acid position 125 from aspartic acid to glycine. The result indicated that one amino acid exchange from aspartic acid 125 to glycine also could lead to the inactivation of HTLV-I protease.

Mode of HTLV-I Protease Biosynthesis

We have suggested two possible translational regulation mechanisms for the synthesis of HTLV-I protease [6]: one is the generating of different mRNAs and the other is the escape of the *gag* termination codon by frameshifting. Correct processing of the *gag* precursor polyprotein in the vaccinia virus expression vector system,

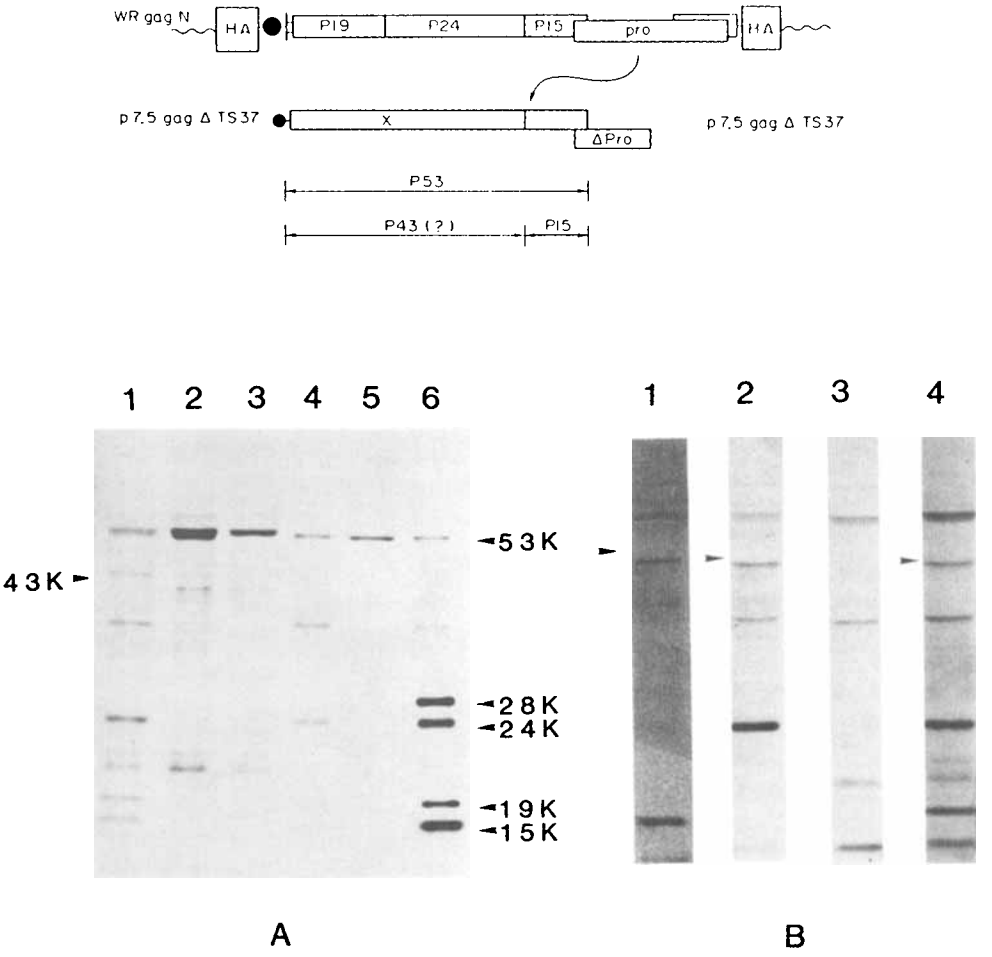

Fig. 6. Identification and charactarization of the 43k protein generated by enzyme-substrate reaction in vivo. **A:** Identification of the 43k processed protein from the mutant *gag* precursor polyprotein by *trans*-action of HTLV-I protease. The substrate donor, p7.5gagΔTS37, was expressed in CV-I cells using wild-type vaccinia virus as the helper virus (**lane 3**), or expressed using recombinant vaccinia virus harboring the protease-coding sequence (WRgagN) as the helper virus (**lane 1**). Protease-coding sequence defective recombinant vaccinia virus (WRgagΔTS) was also utilized as the helper virus for the control experiment (**lane 2**). Cell extracts were separated by 12% SDS-polyacrylamide gel, and subsequent Western blotting analysis was carried out using mixed anti-*gag* monoclonal antibodies of GIN7, NORI, and FR45. Producing of the active form of HTLV-I protease molecule from recombinant vaccinia virus was judged by the correct processing events of 53k *gag* precursor polyproteins (**lanes 4,5**). 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 43k protein. Proteins expressed from p7.5gagΔTS37 by the aid of WRgagN as the helper virus were separated by 12% SDS-polyacrylamide electrophoresis. Subsequent Western blotting analysis was carried out using individual anti-*gag* monoclonal antibodies GIN7, NORI, and FR45, respectively (**lanes 1,2,3**) or their mixture (**lane 4**). Arrows indicate 43k proteins that reacted with GIN7 or NORI. The possible processing event of the mutant *gag* precursor polyprotein is schematically shown at the top. X means the disruption of the conserved amino acid sequence for specific proteolysis. Wavy line indicates the vaccinia virus genome.

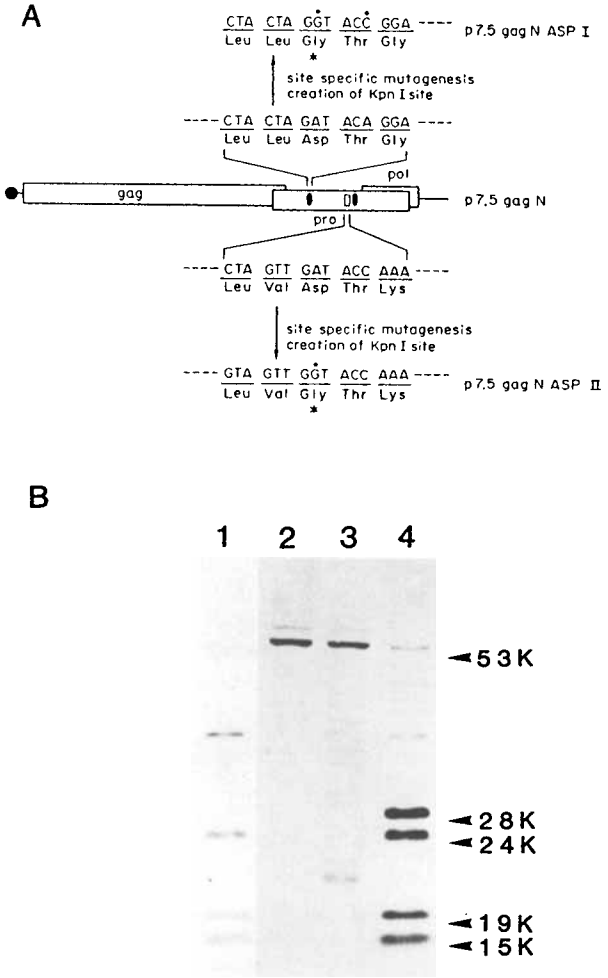

Fig. 7. Inhibition of proteolytic activity of HTLV-I protease resulted from the exchange of the putative catalytic residues by in vitro mutagenesis. **A:** Oligonucleotide-directed site-specific mutagenesis for generating p7.5gagNAspI and p7.5gagNAspII. Symbols: large closed circle, 7.5k gene promoter 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 resulted from the transitions of nucleotides; thin line, 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 a 12% SDS-polyacrylamide gel, and then Western blotting analysis was carried out 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**).

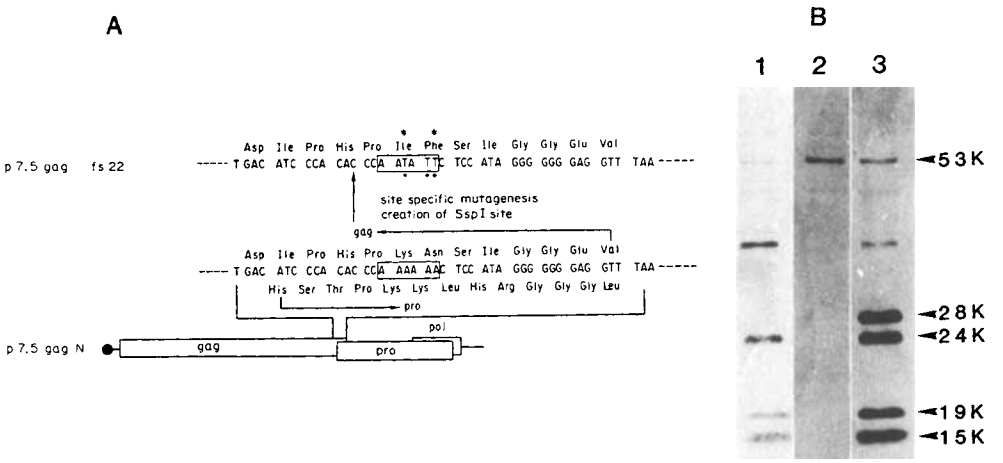

Fig. 8. Abolishment of the correct processing of *gag* precursor polyprotein by blocking the frameshifting event for the synthesis of HTLV-I protease. **A:** Oligonucleotide-directed site-specific mutagenesis for generating p7.5gagfs19. Symbols: □, six consecutive adenine residues presumably involved in the frameshifting event for synthesis of HTLV-I protease: ●, nucleotide transitions resulted from oligonucleotide-directed site-specific mutagenesis; *, exchanged amino acid codon caused by transition of nucleotides; large closed circle, 7.5k gene promoter of vaccinia virus. **B:** The extracts, prepared from cells transfected with p7.5gagN (**lane 2**) and gel and subsequent Western blotting analysis, was carried out using mixed anti-*gag* monoclonal antibodies. *gag*-related proteins synthesized in MT-2 cells were also analyzed under the same conditions (**lane 3**).

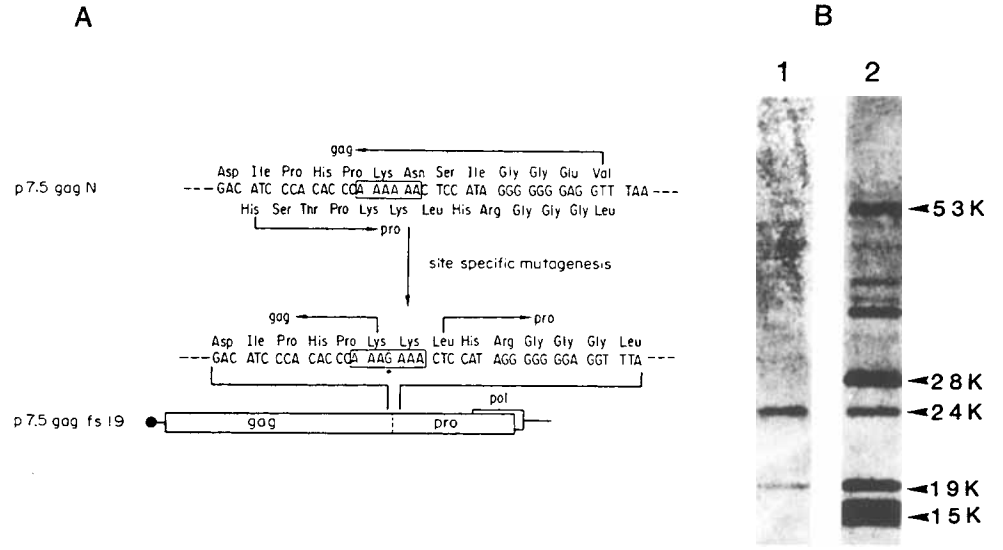

Fig. 9. **A:** Correctly processed *gag* proteins were generated from the *gag*-protease fusion protein. Symbols: □, six consecutive adenine residues of possible frameshifting site for synthesis of *gag-pro* fusion protein; ●, one nucleotide insertion introduced by oligonucleotide-directed site-specific mutagenesis. Vertical broken line indicates the junction between the *gag*- and protease-coding sequence in the fusion protein. **B:** Proteins expressed from p7.5gagfs19 were separated by 12% SDS-polyacrylamide gel and then Western blotting. Analysis carried out using mixed anti-*gag* shows the migration pattern of the authentic mature *gag* proteins of HTLV-I (**lane 1**). MT-2 cell lysate was used as the marker (**lane 2**).

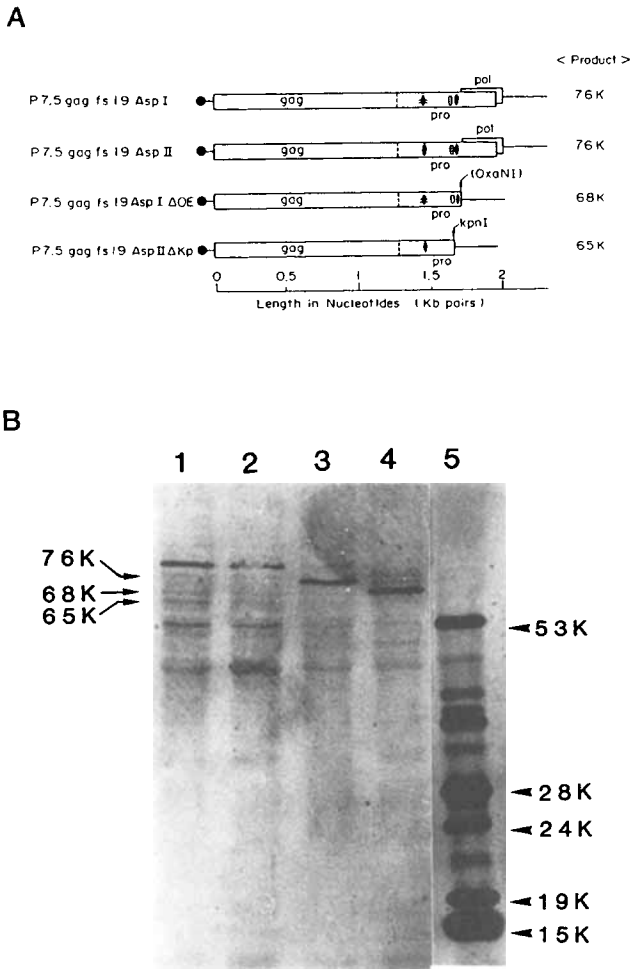

Fig. 10. Identification of 76k *gag*-protease fusion protein. **A:** Schematic representation of the mutant plasmids that express *gag*-protease fusion proteins. Symbols: closed circles, 7.5k gene promoter of vaccinia virus; ● , two conserved amino acid stretches between the retroviral proteases and the cellular enzyme group of aspartyl protease; □ , short amino acid residues conserved only HTLV family retroviral proteases; ⊣● or ⊣⊢ in vitro mutagenized amino acid residues that are described in Figure 7A; vertical broken line, the junction between the *gag*- and protease-coding sequence in fusion proteins; thin lines, pUC DNA sequence. **B:** Analysis of *gag*-protease fusion proteins expressed by the mutants by Western blotting. Proteins extracted from the transfected cell with p7.5gagfs19 AspI (**lane 1**) and p7.5gagfsAspII (**lane 2**), p7.5gagfs19AspIΔDE (**lane 3**), and p7.5gagfs19 AspIIΔKp (**lane 4**) were separated by 12% SDS-polyacrylamide gel and transferred to nitrocellulose filters and probed with mixed anti-*gag* monoclonal antibodies of GIN7, NORI, and FR45. MT-2 cell lysate was utilized as the marker of authentic mature *gag* proteins (**lane 5**).

together with the absence of the methionine codon at the 5′ terminus of proORF, strongly suggested the frameshift mechanism, because the lack of splices from higher eukaryotic mRNA has been well known in vaccinia virus [24]. To obtain direct evidence that HTLV-I protease is translated by avoiding the *gag* termination codon through frameshifting, we introduced the mutation at six consecutive adenine residues at the overlap between the *gag* and *pro* genes (Fig. 8). Three A to T transitions were generated without disruption of the phase of the *gag* gene. Processing of the *gag* precursor polyprotein was no longer observed. The results demonstrated that HTLV-I protease is translated from the identical mRNA for *gag* protein synthesis via ribosomal frameshift at six consecutive adenine residues located in the 3′ terminal region of the *gag* gene.

Therefore, it was assumed that a frameshift gave rise to a *gag-pro* fusion protein with protease activity. Such a fusion protein is expected to be 76k in size. Except for the in vitro translation system [4], our failure previously to detect this precursor polyprotein may have resulted from the rapid cleavage of the frameshift product when it included an active protease region. To identify *gag-pro* precursor polyprotein, one base insertion was introduced at the consecutive adenine residues located in the 3′ terminal region of the *gag* gene, resulting in *gag* and *pro* ORF in the same frame (Fig. 9). Protein 76k immunoreactive with anti-*gag* monoclonal antibodies was detected when inactive protease was formed (Fig. 10). This protein band could not be detected when active protease was expressed in CV-I cells. Instead, correct processing of precursor polyprotein occurred to form p19 and p24 (Fig. 9).

Sequential Processing of the *Gag* Gene Product

The results strongly suggest that 76k protein is the *gag-pro* precursor polyprotein and that HTLV-I protease is capable of triggering the first cleavage event for its own release from fusion protein (Stage I). The released protease catalyzes to form p19 and p33 from p53 *gag* precursor protein that accumulates in cells mainly from the ordinal translation of the *gag* ORF and somewhat from the product of p76 processing (Stage II). p33 is then processed to form p24 and p15 (Stage III). When p53 substrate has no adequate splitting site in Stage II, p15 can be also formed from the *gag* precursor, suggesting that Stage III is not mandatory for further processing (Stage IV). Stage II–IV are performed by the protease in *trans*-action (Table 1).

ACKNOWLEDGMENTS

This work was supported by The Kato Memorial Trust for Nambyo Research and grants-in-aid from the Ministry of Education, Science, and Culture of Japan.

TABLE I. Stages of Sequential Processing by HTLV-I Protease

Stage	Substrate	Product	Mode of action
I	p76	p53 + protease	*cis*
II	p53 (19 · 24 · 15)	p19 + p33 (24 · 15)	*trans*
III	p33	p24 + p15	*trans*
IV[a]	p53	p43 + p15 (19 · 24)	*trans*

[a]In artificial condition.

REFERENCES

1. von der Helm K: Proc Natl Acad Sci USA 74:911–915, 1977.
2. Katoh I, Yoshinaka Y, Rein A, Shibuya M, Okada T, Oroszlan S: Virology 145:280–292, 1985.
3. Seiki M, Hattori S, Hirayama Y, Yoshida M: Proc Natl Acad Sci USA 80:3618–3622, 1983.
4. Kobayashi N, Konishi H, Sabe H, Shigesada K, Noma T, Honjo T, Hatanaka M: EMBO J 3:1339–1343, 1984.
5. Mori K, Sabe H, Siomi H, Iino T, Tanaka A, Takeuchi K, Hirayoshi K, Hatanaka M: J Gen Virol 68:499–506, 1987.
6. Nam SH, Hatanaka M: Biochem Biophy Res Commun 139:129–135, 1986.
7. Ratner L, Josephs SF, Starcich B, Hahn B, Shaw GM, Gallo RC, Wong-Staal F: J Virol 54:781–790, 1985.
8. Hiramatsu K, Nishida J, Naito A, Yoshikura H: J Gen Virol 68:213–218, 1987.
9. Sanger F, Nicklen S, Coulson AR: Proc Natl Acad Sci USA 74:5463–5467, 1977.
10. Iino T, Tacheuchi K, Nam SH, Siomi H, Sabe H, Kobayashi N, Hatanaka M: J Gen Virol 67:1373–1379, 1986.
11. Toh H, Kikuno R, Hayashida H, Miyata T, Kugimiya W, Inoue S, Yuki S, Saigo K: EMBO J 4:1267–1272, 1985.
12. Birnboin HC, Doly J: Nucleic Acids Res 7:1513–1523, 1979.
13. Maniatis T, Fritsch EF, Sambrook J: "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1982.
14. Vieira J, Messing J: Meth Enzymol 153:3–11, 1987.
15. Taylor JW, Ott J, Eckstein F: Nucleic Acids Res 13:8765–8785, 1987.
16. Shida H, Tochikura T, Sato T, Konno T, Hirayoshi K, Seki M, Ito Y, Hatanaka M, Hinuma Y, Sugimoto M, Takahashi-Nishimaki F, Maruyama T, Miki K, Suzuki K, Morita M, Sashiyama H, Hayami M: EMBO J 6:3379–3384, 1987.
17. Graham FL, Van Der Eb AJ: Virology 52:456–467, 1973.
18. Towbin H, Staehelin T, Gordon J: Proc Natl Acad Sci USA 76:4350–4354, 1979.
19. Yasunaga T, Sagata N, Ikawa Y: FEBS Lett 199:145–150, 1986.
20. Inoue J, Watanabe T, Sato M, Oda A, Toyoshima K, Yoshida M, Seiki M: Virology 150:185–187, 1986.
21. Cochran MA, Puckett C, Moss B: J Virol 54:30–37, 1985.
22. Katoh I, Yasunaga T, Ikawa Y, Yoshinaka Y: Nature 329:654–656, 1987.
23. Hattori S, Kiyokawa T, Inakawa K, Shimizu F, Hashimura E, Seiki M, Yoshida M: Virology 136:338–347, 1984.
24. Venkatesan S, Baroudy BM, Moss B: Cell 125:805–813, 1981.