Analysis of Cleavage Site Mutations between the NC and PR Gag Domains of Rous Sarcoma Virus

GISELA SCHATZ,¹ IVA PICHOVA,² AND VOLKER M. VOGT^{1*}

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853,¹ and Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague, Czech Republic²

Received 5 July 1996/Accepted 19 September 1996

In retroviruses, the viral protease (PR) is released as a mature protein by cleavage of Gag, Gag-Pro, or Gag-Pro-Pol precursor polypeptides. In avian sarcoma and leukemia viruses (ASLV), PR forms the C-terminal domain of Gag. Based on the properties of a mutation (*cs22*) in the cleavage site between the upstream NC domain and the PR domain, the proteolytic liberation of PR previously was inferred to be essential for processing of Gag and Pol proteins. To study this process in more detail, we have analyzed the effects that several mutations at the NC-PR cleavage site have on proteolytic processing in virus-like particles expressed in COS and quail cells. Mutant Gag proteins carrying the same mutations also were synthesized *in vitro* and tested for processing with purified PR. In both types of studies, N-terminal sequencing of the liberated PR domain was carried out to exactly identify the site of cleavage. Finally, synthetic peptides corresponding to the mutant proteins were assessed for the ability to act as substrates for PR. The results were all consistent and led to the following conclusions. (i) *In vivo*, if normal processing between NC and PR is prevented by mutations, limited cleavage occurs at a previously unrecognized alternative site three amino acids downstream, i.e., in PR. This N-terminally truncated PR is inactive as an enzyme, as inferred from the global processing defect in *cs22* and a similar mutant. (ii) In Gag proteins translated *in vitro*, purified PR cleaves this alternative site as rapidly as it does the wild-type site. (iii) Contrary to previously accepted rules describing retroviral cleavage sites, an isoleucine residue placed at the P1 position of the NC-PR cleavage site does not hinder normal processing. (iv) A proline residue placed at the P2 position in this cleavage site blocks normal processing.

In retroviruses, the internal structural proteins (products of the *gag* gene) and enzymes involved in replication (products of the *pro* and *pol* genes) are synthesized as polyprotein precursors. Late in the assembly pathway, as the virus particle is about to finish budding from the plasma membrane or just thereafter, a process called maturation occurs, in which the polyproteins are cleaved into their several constituent mature proteins. This proteolytic processing coincides with distinctive morphological changes and is essential for infectivity of the particles. The agent of processing is the viral protease (PR), which itself is synthesized as a polyprotein. In most retroviruses, such as human immunodeficiency virus type 1 (HIV-1), PR is embedded in the Gag-Pol (also called Gag-Pro-Pol) protein. In the avian sarcoma and leukemia viruses (ASLV; prototype Rous sarcoma virus), PR comprises the C-terminal domain of Gag. Although ASLV Gag-Pol also contains the PR domain, the PR derived from Gag is necessary and sufficient to carry out all the processing steps needed to create an infectious virion (30).

By structure and by catalytic mechanism, retroviral PRs belong to the class of aspartic proteases (reviewed in reference 6). All of these enzymes function exclusively as dimers, the active site being created at the dimer interface. In cellular aspartic proteases, a single polypeptide with two homologous domains folds into two lobes of similar structure, while in retroviral PRs, separate, identical subunits come together to form the dimer. The three-dimensional structure of retroviral PRs is known to high resolution for HIV-1 PR (reviewed in reference 34) and for ASLV PR (8). The key structural feature

that holds the dimer together is a four-stranded antiparallel β -sheet composed of the first several residues from the N terminus and the last several residues from the C terminus. Interference with these interactions, for example, in HIV-1 by a high concentration of competing homologous peptides (1a, 35), prevents dimerization. The N-terminal amino acids of PR thus are of crucial importance for activity.

Proteolytic processing in retroviruses somehow is delayed until late in virus assembly. The mechanism underlying this regulation has yet to be unraveled. However, several clues have emerged. First, control of dimerization is likely to play an important role. This inference is based in part on the observation that expression of polyproteins with a linked PR, composed of two sequential PR domains connected by a short flexible peptide segment, leads to premature processing and thus abrogation of assembly (2, 15). Second, apparently the PR domain must be proteolytically liberated in order to carry out its function. Mutations at the cleavage site that creates the N terminus of PR may fully or partially block Gag and Gag-Pol processing. This effect has been observed both in the HIV-1 (36) and ASLV (3, 17, 32) systems. The segment of polypeptide immediately upstream of the PR domain may inhibit enzymatic activity, just as the upstream sequences in the zymogens of cellular aspartic proteases restrain their activity. More direct evidence for such an effect has been found for both HIV-1 (37) and ASLV (28).

Three mutations at the junction of the ASLV NC and PR domains of Gag that lead to a processing defect have been described. In one, an Arg residue was substituted for the Leu residue at the P1' position in the cleavage site, i.e., the N-terminal residue of PR (17). When this construct was expressed in chicken cells, no mature Gag proteins were generated. However, a significant proportion of the Gag polyprotein Pr76 was observed to be processed to a polypeptide with an

* Corresponding author. Mailing address: Section of Biochemistry, Molecular and Cell Biology, Biotechnology Bldg., Cornell University, Ithaca, NY 14853. Phone: (607) 255-2443. Fax: (607) 255-2428. E-mail: vmv1@cornell.edu.

approximate size of 63 kDa. In the second mutation, called *cs22*, the normal P2 Val and P3 Ala residues upstream of the cleavage site were deleted and the P1 Ser residue was replaced by an Ile residue. This resulted in a processing defect in *Escherichia coli*, with NC-PR remaining fused but CA and MA being liberated from the Gag precursor (1a, 12). The third mutation, *cs4*, in which only the P1 Ser is replaced by an Ile residue, had a phenotype similar to that of *cs22* in *E. coli* (1a). When the *cs22* construct was expressed in monkey COS cells, no Gag processing was noted (3). However, expression in quail cells led to limited processing, yielding a ca. 63-kDa protein (here called P63) presumably identical to that seen in the P1' Arg mutant. Again no mature CA or MA proteins were observed.

Because of its apparent importance for activation of proteolysis, we set out to characterize processing at the NC-PR site of ASLV Gag in more detail, with particular reference to the provenance of the 63-kDa Gag protein. The results demonstrate that this protein is generated not only in avian cells but also in COS cells. Amino acid sequence analyses of PR-related products in mutants with defects in the NC-PR cleavage site imply that P63 is formed by processing at a previously unrecognized cleavage site three residues downstream from that used in wild-type Gag. The PR domain that is liberated by this cleavage apparently is inactive. Our results also show that contrary to accepted dogma on the specificity of retroviral PRs, an Ile residue in the P1 position of the NC-PR cleavage site does not block correct processing, either in proteins or in synthetic peptides. In the *cs22* mutation, it is apparently the Pro residue at P2, not the Ile residue at P1, that is responsible for the cleavage block.

MATERIALS AND METHODS

DNA constructions and in vitro translation. All plasmids were constructed by using common subcloning techniques and were propagated in *E. coli* DH5 α . Constructions of the cleavage site mutations *cs4* (12) and *cs22* and the wild-type control (3) have been described previously. The *np1* mutation was constructed by PCR mutagenesis and built into the final vector pSVMyr1A (27) for expression in COS-1 cells and into the BH-pol version of the full-length infectious clone RCASneo, as for *cs22* (3), for expression in the quail line QT35. An intermediate vector based on pBluescript (Stratagene), which places the *gag* genes with the various cleavage site mutations under control of the T7 promoter, was used as the template for in vitro transcription and translation of [³⁵S]methionine-labeled Gag protein, using a TNT lysate kit (Promega).

Cell culture and transfection. Monkey COS-1 cells were maintained in low-glucose Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (Sigma Chemical Company), 1% vitamins, 20 mM L-glutamine, 100 U of penicillin per ml, and 100 μ g of streptomycin sulfate per ml (all from Life Technologies). Quail QT35 cells were grown in the same medium with the addition of 1% dimethyl sulfoxide (DMSO). For transfection, cells were seeded at a density of 5×10^6 per 10-cm-diameter plate in transfection medium (growth medium with fetal bovine serum replaced by 10% Nu-Serum [Collaborative Research]), grown for ca. 18 h, and washed with phosphate-buffered saline (PBS), and then a cocktail of 5 to 10 μ g of DNA and 200 μ g of DEAE-dextran (Sigma) per ml in PBS was added. The plates were incubated at 37°C for 20 to 60 min and shocked with 10% DMSO in transfection medium for 1 to 2 min, and the DMSO was replaced with transfection medium containing 100 μ M chloroquine. After a further incubation of 4 h, the chloroquine medium was removed and fresh transfection medium was added. The cells were then grown for 2 days. For analysis of proteins by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE), labeling with [³⁵S]methionine was carried out with 2 ml of methionine-free DMEM (100 μ Ci/ml) per 10-cm-diameter plate for 30 min. The labeling medium was replaced with fresh DMEM, incubation continued for various times, labeled particles were collected, and the cells were lysed. For N-terminal sequencing, 600 μ Ci of [³⁵S]methionine per ml was added per 10-cm-diameter plate.

Immunoprecipitation and analysis of proteins. Lysis of cells was carried out in ice-cold 20 mM Tris-HCl (pH 8)–50 mM NaCl–1 mM EDTA–0.5% Triton X-100–0.5% sodium deoxycholate–1 mM phenylmethylsulfonyl fluoride. After removal of cell debris and nuclei by centrifugation for 20 s in a microcentrifuge, 2 μ l of rabbit antiserum to PR was added, and the mixture was incubated for 1 h at 4°C. The PR had been purified from avian myeloblastosis virus (AMV) and was contaminated with a small amount of CA, and hence this antiserum also reacts with CA. Then 35 μ l of a 50% (vol/vol) slurry of protein A-Sepharose beads (Sigma) in radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH

FIG. 1. Diagram of mutants and peptides. The schematic structure of the ASLV Gag protein is shown at the top, with its acetylated N terminus and the major domains MA, p10, CA, NC, and PR marked. The vertical lines indicate the known cleavage sites in Gag. The amino acid sequences at the junctions between the NC and PR domains in wild-type (wt) and mutant proteins are shown below. Dashes indicate the sites at which cleavage occurs in the protein and in the peptides (Table 1).

8], 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate) was added, and incubation continued for 4 h to overnight at 4°C. The adsorbed proteins were collected and washed three times in radioimmunoprecipitation assay buffer, dissolved in SDS-containing sample buffer, and analyzed by standard SDS-PAGE. Proteins were visualized by fluorography of the dried gel after soaking the gel in 0.7 M sodium salicylate. For amino acid sequencing, polypeptides in the gel were transferred electrophoretically to Immobilon P membranes (Millipore). The band corresponding to radioactive PR was located by autoradiography, excised with a razor blade, and submitted for N-terminal sequencing. The released fraction after each cycle of the Edman degradation was counted in scintillation fluid.

Cleavage assays in vitro. The standard reaction conditions for processing of proteins by PR were 100 mM morpholine ethanesulfonic acid (pH 6.8), 825 mM NaCl, 0.25 mM EDTA, 1 mM dithiothreitol, 6% glycerol, and 1 mM phenylmethylsulfonyl fluoride, plus or minus 0.7 μ M PR in 50 μ l (corresponding to 1 μ g of protein), for 1 h at 39°C. PR was purified from AMV (primarily a mixture of myeloblastosis-associated viruses A and B; Life Sciences) by extraction with chloroform-methanol (9), and the stock solution was stored at –20°C at a concentration of 2 mg/ml. The peptides used in the assays were synthesized by solid-phase methods. The standard reaction conditions for peptide cleavage were 200 mM NaPO₄ (pH 6.0), 2 M NaCl, and 400 μ M peptide, plus or minus 0.7 μ M PR in 50 μ l, for 2 h at 37°C. In experiments to approximately gauge the K_m , the concentration of peptide was varied between 20 and 500 μ M. The course of the cleavage reaction was followed by high-pressure liquid chromatography. Peptide products generated by PR were characterized by amino acid composition analysis and in some cases by amino acid sequencing.

RESULTS

Processing in transfected cells. To further illuminate processing at the junction of the ASLV NC and PR domains, we constructed Gag expression vectors with two new mutations and also reexamined the original *cs22* mutant (3). The sequences of the mutants differed from that of the wild type in several ways (Fig. 1). Mutant *np1* (EGPILAMT) closely resembles *cs22* (EPPILAMT), both having an Ile residue at the P1 position of the cleavage site instead of the wild-type Ser residue (PPAVSLAMT) and having the two upstream residues Ala and Ser deleted. In addition *np1* has a Pro-to-Gly change. The mutant *cs4*, which was analyzed previously in an *E. coli* expression system and found to be leaky in what was termed autocatalytic cleavage (12) but had not been studied in eucaryotic cells, has only the Ile change at P1 (PPAVILAMT). Peptides corresponding to these and related sequences were synthesized as well (see below).

The Gag proteins containing these mutations, or the active-site PR mutation D37I (29) as a control, were analyzed in a transient transfection assay in COS cells. The transfected cells

FIG. 2. Cleavage of mutant Gag proteins in vivo. Pairs of plates of COS cells transiently transfected with mutant or wild-type (wt) DNAs were labeled with [35 S]methionine for 45 min. The cells were either lysed immediately after labeling with detergent (pulse) or grown in nonradioactive medium for 16 h and then lysed (chase). In the latter case, virus particles in the medium were collected by centrifugation. Postnuclear lysates were subjected to immune precipitation with anti-PR serum, and the radioactive polypeptides were separated by SDS-PAGE, transferred to an Immobilon membrane, and analyzed by autoradiography. Odd-numbered lanes 1 to 11, lysates of pulse-labeled cells; even-numbered lanes 2 to 12, lysates of chased cells; lanes 13 to 17, virions from medium. The appearance of the CA polypeptide is due to the presence of anti-CA antibodies in the rabbit anti-PR serum.

were pulse-labeled with [35 S]methionine, or pulse-labeled and chased overnight, and then cell lysates as well as released virions were collected. The proteins were subjected to immune precipitation with a rabbit antiserum that recognizes PR and CA, and analysis of both fractions was carried out by SDS-PAGE and fluorography (Fig. 2). As expected, Gag carrying the D37I mutation showed no processing in cell lysates (lanes 1 and 2) or in collected virions (lane 13). Mutants *np1* and *cs22* also showed an absence of the mature protein CA, which is the most readily visible and diagnostic product of normal proteolytic processing (lanes 3 to 6, 14, and 15). By comparison, the wild-type Gag yielded the expected CA band (actually a triplet of bands corresponding to three closely related CA species differing at their C termini [19]) (lanes 9, 10, and 17). In contrast to the previously reported results with transfected COS cells (3), we found that both mutants also showed evidence of a 63-kDa Gag protein in virions (lanes 14 and 15), which was absent in the PR-defective mutant (lane 13). On the other hand, the pattern of Gag products in mutant *cs4* (lanes 7, 8, and 16) appeared identical to that in the wild type. This result originally seemed surprising, since β -branched amino acids in the P1 position are never present in natural cleavage sites in any retroviruses (20) and in several cases have been shown to block cleavage in model peptides (7, 13, 14, 24) as well as in proteins (11). To confirm this result and to test its generality, we also built the *cs4* mutation into a full-length clone of Rous sarcoma virus in which the *neo* gene replaced *src*. Quail clones selected for expression of *neo* by G418 drug resistance also produced *cs4* virus particles that were indistinguishable from wild-type particles on SDS-PAGE (data not shown). The same result was obtained with independent DNA clones that were subsequently resequenced to verify the presence of the mutation.

The P63 in *cs22* previously was suggested to result from cleavage at or near the NC-PR junction (3, 32). Thus, the

complementary product of this cleavage ought to be PR or a slightly modified PR if processing occurs at a nearby sequence. A polypeptide was observed to migrate at the position of PR by SDS-PAGE in both the *cs22* and *np1* mutants (Fig. 2). PR is relatively poor in its ability to be recognized by antisera, and we reasoned that miscleavage at its N terminus might further reduce its antigenicity or lead to its degradation in cells or virions. These factors, in addition to the incomplete cleavage of Gag, might account for the lesser amount of this product (lanes 3 to 6, 14, and 15) compared with that of wild-type PR (lanes 9, 10, and 17).

We used N-terminal amino acid sequencing to precisely define the location at which processing occurs in the NC-PR cleavage site mutants. Since the amounts and purity of the PR species recovered after transfection would make normal sequencing very difficult, sequencing was performed instead on products radioactively labeled with [35 S]methionine. In the wild-type PR, Met residues occur at positions 3 and 5, providing a distinctive signature for PR. Transfected COS cells were grown for 24 h with [35 S]Met, and PR-related products were isolated by immune precipitation of cell lysates and of virions with the same antiserum that recognizes PR and CA. After SDS-PAGE of the dissolved immune precipitates, radioactive bands migrating at the position of PR were submitted to Edman degradation. The released residue at each cycle was counted by scintillation counting. Wild-type PR as a control showed the expected peaks at cycle 3 and cycle 5 (Fig. 3A). PR from the *cs4* mutant showed a pattern similar to that of the wild type, implying directly that most of the cleavage events still take place at the same site despite the presence of Ile at P1 (Fig. 3B). By contrast, PR from the *cs22* mutant (Fig. 3C) showed a single peak of released radioactivity at cycle 2. We interpret this peak of activity to represent the Met residue at position 5, which would imply that the cleavage that generates the P63 protein occurs three amino acids downstream of the wild-type site, at the sequence AVSLAM-TMEH.

Processing of in vitro-translated proteins. As an independent approach to study the properties of the NC-PR cleavage site mutations, [35 S]Met-labeled Gag proteins were expressed by in vitro translation in a reticulocyte lysate. Previous results had shown that wild-type ASLV Gag is stable in this system, i.e., does not undergo autocatalytic cleavage (23a). This stability in part is a consequence of the low concentration of protein in such translation systems. It also is possible that activation of proteolysis requires virion formation, which presumably does not occur in the reticulocyte lysate. After adjustment to acid pH and high salt, the conditions preferred by retroviral PRs, the labeled translation extracts were incubated with purified PR. The extent of proteolysis as a function of time was visualized by SDS-PAGE and autoradiography after immune precipitation of the reaction mixtures with anti-PR antiserum. The immune precipitation is a key step because PR migrates near the position of globin, which is present in large quantities in the translation reaction and thus interferes with resolution of radioactive polypeptides in this portion of the gel.

All of the mutant Gag proteins were similar to the wild-type protein in ability to act as substrates for exogenous PR (Fig. 4). The long incubation time of 120 min led to fully or nearly fully processed proteins (lanes 4, 7, 10, and 14). As controls, parallel incubation of two unprogrammed translation mixtures (lanes 4 and 11) did not significantly change the profile of bands compared with the time zero controls (lanes 1, 5, 8, and 12), confirming previous results with wild-type Gag that processing was due to the added PR and not to autocatalytic cleavage of the Gag proteins themselves. In each of the mutants, a PR species was prominent even at 15 min, suggesting that in vitro,

FIG. 3. Amino acid sequence of PR polypeptides. [^{35}S]methionine-labeled PR was obtained from COS cells transfected with wild-type (w.t.) or mutant constructs (panels A to C) or from in vitro translation reactions that were incubated with PR (panels D to F). In each case, the crude lysates (A to C) or in vitro translation mixtures (D to F) were immune precipitated with anti-PR serum, and the radioactive polypeptides were separated by SDS-PAGE and then transferred to an Immobilon P membrane. The bands corresponding to PR were subjected to Edman degradation, and the amount of radioactivity released after each cycle was counted. (A and D) PR from wild-type Gag; (B and E) PR from *cs4* mutant Gag; (C and F) PR from *cs22* mutant Gag. The N-terminal sequence of the wild-type PR is shown below panels A and D.

the polypeptide segment joining the NC and PR domains is similarly accessible to PR in the wild type and mutant proteins. In ASLV, the first or one of the first cleavages in maturation in vivo is known to occur near the NC-PR junction (33).

To pinpoint the processing site, we used the amino acid sequencing procedure described above. PR from the wild type (Fig. 3D) and from *cs4* (Fig. 3E) yielded methionine residues at positions 3 and 5. PR from *cs22* (Fig. 3F) yielded a methionine residue at position 2. Thus, the patterns of released radioactivity were the same as those found for the PR species derived from cells labeled in vivo.

We draw the following conclusions from these results. First, the limited processing in the *cs22* and *np1* mutants is not restricted to a particular cell type, although *E. coli* appears to be an exception (1). Second, in eucaryotic cells the site where this processing occurs is three amino acid residues downstream of the normal site, and cleavage there occurs not only in virus maturation but also in vitro. Third, an Ile residue at the P1 position by itself is insufficient to block correct processing, both in vitro and during particle maturation in living cells.

Processing of synthetic peptides. Polypeptide folding may affect the accessibility of a potential cleavage site to PR. To gain more direct information on the role of the primary amino acid sequence in cleavage at the NC-PR junction, we prepared several synthetic octapeptides and one longer peptide (Table 1). The peptides were designed to mimic the normal and mutant cleavage sites as well as to allow us to explore which residues are responsible for blocking cleavage in *cs22* and *np1*. Pep-wt contains the wild-type sequence and served as the control. Pep-*cs4* differs from the wild type only in the substitution of Ile for Ser at the P1 position. Pep-*cs22* and Pep-*np1* are octapeptides containing the sequence of *cs22* and *np1*, respectively. Pep-*cs22L* is a dodecapeptide with the indicated mutation but also extending downstream to include the site where hydrolysis was inferred in the proteins described above. Pep-a, Pep-b, and Pep-c differ from Pep-*cs22* in one or two positions. The peptides were incubated at pH 6.0 with 2 M NaCl for 2 h at 37°C with purified PR at a concentration of 0.7 μM , and the resulting products were analyzed and quantified by high-pressure liquid chromatography (Table 1). The identities of the

FIG. 4. Cleavage of mutant Gag proteins in vitro. Mutant or wild-type (wt) Gag proteins were translated in vitro in a rabbit reticulocyte lysate system in presence of [³⁵S]methionine. Samples of 5 μ l of the translation mixtures were diluted with 20 μ l of cleavage buffer and incubated (Incub) for 0 min, 15 min, or 120 min at 37°C in the presence (+) or absence (–) of 1 μ g of PR. The reaction mixtures were subjected to immune precipitation with anti-PR serum and then analyzed by SDS-PAGE and autoradiography.

peaks in the chromatographic profile were established by amino acid composition and N-terminal amino acid sequence analysis. For those peptides that were good substrates for PR, the maximum rates of hydrolysis (k_{cat} values) are approximately proportional to percent cleavage at the 2-h time point shown in Table 1, since the concentration of the peptides was higher than the K_m (data not shown).

The results from these analyses support and extend the conclusions drawn from maturation of Gag mutants in cells and the cleavage of Gag protein in vitro. Pep-wt was found to be processed into the two expected fragments PAVS and LAMT, and the rate of this reaction was normalized to 100% (Table 1). Pep-cs4 was processed at the same site and with only about a twofold-lower rate, yielding products PAVI and LAMT, directly confirming the ability of avian PR to hydrolyze the peptide bond following an Ile residue. This finding contradicts a previous finding that a cs4 decapeptide could not be cleaved by PR (1), a discrepancy that might be accounted for by different cleavage conditions. Pep-cs22 and Pep-np1 were not processed by PR. No cleavage was observed even upon prolonged incubations and with a 10-fold-higher concentration of PR. By contrast, Pep-cs22L was an excellent substrate, with a relative cleavage rate of 45. However, in this case the products were peptides EPPILAM and TMEHK, supporting the identification of the downstream processing site in the Gag proteins described above. Pep-a, which differs from Pep-cs22 only at the P4 position, was also refractory to processing, ruling out that the Glu residue at this position is responsible for the cleavage defect. By contrast, Pep-b, which has the cs22 sequence except for substitution of the wild-type Val residue in the P2 position, was processed with a relative cleavage rate of 36. Similar results were obtained for Pep-c, a similar peptide with two changes compared with cs22. We interpret these results as direct evidence that the Pro residue at position P2 is the cause of the phenotype of the cs22 and np1 mutations. The importance of the P2 residue in determining the ability of peptides to be cleaved in vitro by ASLV PR has been inferred previously, although the effect of a Pro substitution had not been tested (7, 11).

DISCUSSION

The experiments presented here clarify and extend previously published results on the properties of NC-PR cleavage site mutants like cs22. If proper processing at this site cannot occur, cleavage is redirected to the neighboring site three amino acids downstream in the PR domain, generating P63, a Gag fragment of about 63 kDa, plus an N-terminally truncated, inactive PR. We interpret the absence of mature MA and CA Gag proteins (3) and of mature RT and IN Pol proteins (32) in cs22 particles to be due to the inability of the truncated PR to dimerize. Although it has not been demonstrated directly that the missing three amino acid residues prevent dimerization, there is abundant indirect evidence favoring this explanation. In HIV-1, high concentrations of synthetic peptides that mimic the sequences tied up in the four-stranded β -sheet prevent dimerization (1a, 35), and autoproteolysis after residue 5 leads to an inactive PR (26). In ASLV, even small changes at the N terminus of PR appear to be deleterious to activity, as suggested by the report that a minor form of PR found after expression in *E. coli* is missing the N-terminal Leu residue and is inactive (22). The importance of the exact N terminus of ASLV PR also is suggested by two other observations. First, PR engineered to begin with an initiating N-terminal Met residue placed immediately proximal to the normal N-terminal Leu residue does not have full activity (7, 28). Second, PR-Pol fusion proteins with a similar initiating Met and expressed in insect cells do not manifest any proteolytic activity (31).

It is unclear what factors influence the extent of processing at the alternative NC-PR site in cs22 and in the other similar mutants. In a previous study on the phenotype of the cs22 mutation in eucaryotic cells (3), partial cleavage of the Gag protein was noted in avian cells but not in COS cells. However, in retrospect, minor amounts of radioactive bands corresponding both to P63 as well as to the truncated PR are apparent in Fig. 3 of that publication. The source of the quantitative differences between those amounts and the amounts of these proteins observed here, which correspond to at least one-quarter of all Gag molecules in virions, is unknown. According to the most commonly presented model for the initial processing events in retrovirus maturation (as drawn, for example, in Fig. 7 in reference 32), two PR domains on a pair of Gag molecules first dimerize to form an active enzymatic unit. This unit then acts *in trans* to sever the NC-PR connection in an adjoining Gag molecule in the budding virion. In the case of cs22 and the other cleavage site mutants discussed here, in this model each cleavage event *in trans* would remove one Gag molecule from

TABLE 1. Cleavage of peptides by PR

Peptide	Sequence ^a	Relative rate of cleavage ^b
Pep-wt	PAVS-LAMT	100
Pep-cs22	EPPILAMT	<1
Pep-np1	EGPILAMT	<1
Pep-cs22L	EPPILAM-TMEHK	45
Pep-cs4	PAVI-LAMT	54
Pep-a	APPILAMT	<1
Pep-b	EPVI-LAMT	36
Pep-c	APVI-LAMT	78

^a Hyphens indicate sites of cleavage in those peptides that are substrates for PR.

^b Digestions were carried out for 2 h at 37°C in 0.2 M NaPO₄ (pH 6)–2 M NaCl–400 μ M peptide with 0.7 μ M AMV PR in 50 μ l. Cleavage products were separated on a C₁₈ reverse-phase column.

the population that is competent to dimerize in its C-terminal PR domain. Thus, cleavage eventually would stop. How far it would proceed depends on the mobility of Gag molecules in the maturing virion and on other geometrical factors that are unknown.

An alternative model for the initial cleavages in virus maturation postulates that the dimerized PR domains in a pair of Gag molecules form an active enzymatic unit that severs the NC-PR connection in *cis*, i.e., on one of the Gag molecules themselves. In the crystal structure of retroviral PRs, the active site is on one side of the enzyme while the N and C termini are on the other, suggesting that an enzyme could not reach its own termini. However, recent kinetic evidence as well as molecular modeling suggest that cleavage in *cis* is possible and occurs in some situations in vitro with purified HIV-1 PR derivatives containing upstream sequences (5, 16). Furthermore, in cotransfection experiments, the ASLV *cs22* Gag protein was not observed to cleave a different and marked Gag molecule expressed in the same cells and packaged into the same virions (3). The presence of low amounts of P63 and the truncated PR, as noted above, thus could be taken to suggest cleavage in *cis*.

The biological significance of processing at the LAM-TME site in the cleavage site mutants is uncertain. However, it is important to note that this processing is not limited to mutant Gag proteins. In wild-type virus, approximately 7% of the population of PR molecules found in virions is composed of truncated, presumably inactive molecules generated by cleavage at this site (18). Conceivably this cleavage represents a first step in proteolytic processing, perhaps functioning to loosen the structure of the shell of Gag molecules or the PR domains in Gag, to then facilitate correct processing at the NC-PR junction. Mutational analysis may provide a way to address if this processing in fact is important in maturation.

It had been widely accepted for years that retroviral PRs cannot cleave, or can cleave only extremely poorly, sequences in which a β -branched amino acid occurs at the P1 position. This conclusion was derived originally from the complete absence of Ile or Val residues in known cleavage sites in viral (20) and nonviral (23) proteins. Assays with model peptides and mutant proteins seemed to confirm this conclusion for HIV-1 PR (10, 24) and HIV-2 PR (21) as well as for ASLV PR (13, 14). However, in determining the ability of a substrate sequence to be cleaved by PR, the amino acid residues at each of the seven or eight positions in a cleavage site do not act entirely independently of each other (25). Our finding that Ile is well tolerated by ASLV PR in the context of the sequence PAVILAM does not necessarily imply that Ile is tolerated in general in the P1 position. One factor that helps explain the ready cleavage after the large Ile residue at the mutant *cs4* site is that ASLV PR, unlike HIV-1 PR, has a larger binding pocket for the P1 amino acid of the substrate, with a Gly residue at the equivalent position in the PR polypeptide in which an Ile residue occurs in HIV-1 PR (4). However, even HIV-1 PR has been reported to cleave peptides with β -branched amino acids in the P1 site, in the context of the same avian NC-PR cleavage site we have studied here, with k_{cat} values for the peptides PPAVILAMTMR and PPAVLAMTMR listed as about 10-fold lower than for the wild-type sequence (25). But interpretation of these data may be complicated by the presence of the second cleavage site after the methionine residue that we report here, since it is unclear from the data presented to what extent the authors defined the sequence of the cleavage products. Bovine leukemia virus PR recently also has been found to process peptides containing Ile in P1 position (7a). In any case, the observation that avian PR correctly and efficiently pro-

cesses the mutant *cs4* cleavage site between the NC and PR domains of ASLV Gag should provide a cautionary note for predictions on the ability of particular sequences to be substrates for retroviral proteases.

ACKNOWLEDGMENTS

We thank A. M. Skalka for the gift of some of the plasmids used in this study, J. Konvalinka for helpful discussions, and J. Velek for synthesis of peptides.

This work was supported by grant CA-20081 from the USPHS and by grant A4055503 from the Academy of Sciences of the Czech Republic.

REFERENCES

- Arad, G., R. Bar-Meir, N. Almog, M. Chorev, and M. Kotler. 1995. Avian sarcoma leukemia virus protease linked to the adjacent Gag polyprotein is enzymatically active. *Virology* **214**:439–444.
- Babé, L. M., J. Rosé, and C. S. Craik. 1992. Synthetic "interface" peptides alter dimeric assembly of the HIV-1 and 2 proteases. *Protein Sci.* **1**:1244–1253.
- Burstein, H., D. Bizub, and A. M. Skalka. 1991. Assembly and processing of avian retroviral Gag polyproteins containing linked protease dimers. *J. Virol.* **65**:6165–6172.
- Burstein, H., D. Bizub, M. Kotler, G. Schatz, V. M. Vogt, and A. M. Skalka. 1992. Processing of avian retroviral Gag polyprotein precursors is blocked by a mutation in the NC-PR cleavage site. *J. Virol.* **66**:1781–1785.
- Cameron, C. E., B. Grinde, P. Jacques, J. Jentoft, J. Leis, A. Wlodawer, and I. T. Weber. 1993. Comparison of the substrate binding pockets of Rous sarcoma virus and human immunodeficiency virus type 1 proteinases. *J. Biol. Chem.* **268**:11711–11720.
- Co, E., G. Koelsch, Y. Lin, E. Ido, J. A. Hartsuck, and J. Tang. 1994. Proteolytic processing mechanisms of a miniprecursor of the aspartic protease of human immunodeficiency virus type 1. *Biochemistry* **33**:1248–1254.
- Davies, D. R. 1990. The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Chem.* **19**:189–215.
- Grinde, B., C. E. Cameron, J. Leis, I. T. Weber, A. Wlodawer, H. Burstein, and A. M. Skalka. 1992. Analysis of substrate interactions of the Rous sarcoma virus wild-type and mutant proteases and human immunodeficiency virus 1 protease using a set of systematically altered peptide substrates. *J. Biol. Chem.* **267**:9491–9498.
- Hrusková, O. Personal communication.
- Jaskólski, M., M. Miller, J. K. M. Rao, J. Leis, and A. Wlodawer. 1990. Structure of the aspartic protease from Rous sarcoma retrovirus refined at 1 Å resolution. *Biochemistry* **29**:5889–5898.
- Johnson, S. P., M. Veigl, T. Vanaman, and J. Leis. 1983. Cyanogen bromide digestion of the avian myeloblastosis virus pp19 protein: isolation of an amino-terminal peptide that binds to viral RNA. *J. Virol.* **45**:876–881.
- Jupp, R. A., L. H. Philip, J. S. Mills, S. F. J. LeGrice, and J. Kay. 1991. Mutating P2 and P1 residues at cleavage junctions in the HIV-1 pol polyprotein. *FEBS Lett.* **283**:180–184.
- Konvalinka, J., I. Blaha, R. Skrabana, J. Sedlacek, I. Pichova, F. Kapralko, V. Kostka, and P. Strop. 1991. Subsite specificity of the proteinase from myeloblastosis associated virus. *FEBS Lett.* **282**:73–76.
- Kotler, M., R. A. Katz, and A. M. Skalka. 1988. Activity of avian retroviral protease expressed in *Escherichia coli*. *J. Virol.* **62**:2696–2700.
- Kotler, M., R. A. Katz, W. Danho, J. Leis, and A. M. Skalka. 1988. Synthetic peptides as substrates and inhibitors of a retroviral PR. *Proc. Natl. Acad. Sci. USA* **85**:4185–4189.
- Kotler, M., W. Danho, R. A. Katz, J. Leis, and A. M. Skalka. 1989. Avian retroviral protease and cellular aspartic proteases are distinguished by activities on peptide substrates. *J. Biol. Chem.* **264**:3428–3435.
- Kräusslich, H.-G. 1991. Human immunodeficiency virus proteinase dimer as a component of the viral polyprotein prevents particle assembly and viral infectivity. *Proc. Natl. Acad. Sci. USA* **88**:3213–3217.
- Louis, J. M., N. T. Nashed, K. D. Parris, A. R. Kimmel, and D. M. Jerina. 1994. Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. *Proc. Natl. Acad. Sci. USA* **91**:7970–7974.
- Oertle, S., and P.-F. Spahr. 1990. Role of the Gag polyprotein precursor in packaging and maturation of Rous sarcoma virus genomic RNA. *J. Virol.* **64**:5757–5763.
- Pepinsky, R. B., I. A. Papayannopoulos, S. Campbell, and V. M. Vogt. 1996. Analysis of Rous sarcoma virus Gag proteins by mass spectrometry indicates trimming by host exopeptidase. *J. Virol.* **70**:3313–3318.
- Pepinsky, R. B., I. A. Papayannopoulos, E. P. Chow, N. K. Krishna, R. C. Craven, and V. M. Vogt. 1995. Differential proteolytic processing leads to multiple forms of the CA protein in avian sarcoma and leukemia viruses. *J. Virol.* **69**:6430–6438.
- Pettit, S. C., J. Simsic, D. D. Loeb, L. Everitt, C. A. Hutchison III, and R. Swanstrom. 1991. Analysis of retroviral protease cleavage sites reveals two

- types of cleavage sites and the structural requirements of the P1 amino acid. *J. Biol. Chem.* **266**:14539–14547.
21. Phylip, L. H., A. D. Richards, J. Kay, J. Konvalinka, P. Strop, I. Blaha, J. Velek, V. Kostka, A. R. Ritchie, A. V. Broadhurst, W. G. Farmerie, P. E. Scarborough, and B. M. Dunn. 1990. Hydrolysis of synthetic chromogenic substrates by HIV-1 and HIV-2 proteinases. *Biochem. Biophys. Res. Commun.* **171**:439–444.
 22. Pichova, I., P. Strop, J. Sedlacek, F. Kapralk, V. Benes, M. Travnicek, L. Pavlickova, M. Soujcek, V. Kostka, and S. Foundling. 1992. Isolation and biochemical characterization and crystallization of the p15Gag proteinase of myeloblastosis associated virus expressed in *E. coli*. *Int. J. Biochem.* **24**:235–242.
 23. Poorman, R. A., A. G. Tomasselli, R. L. Heinrikson, and F. J. Kézdy. 1991. A cumulative specificity model for proteases from human immunodeficiency virus types 1 and 2, inferred from statistical analysis of an extended substrate data base. *J. Biol. Chem.* **266**:14554–14561.
 - 23a. Puterman, D., G. Schatz, and V. Vogt. Unpublished observations.
 24. Richards, A. D., L. H. Phylip, W. G. Farmeries, P. E. Scarborough, A. Alvarez, B. N. Dunn, P.-H. Hirel, J. Konvalinka, P. Strop, L. Pavlickova, V. Kostka, and J. Kay. 1990. Sensitive soluble chromogenic substrates for HIV-1 proteinase. *J. Biol. Chem.* **265**:7733–7736.
 25. Ridky, T. W., C. E. Cameron, J. Leis, T. Copeland, A. Wlodawer, I. T. Weber, and R. W. Harrison. 1996. Human immunodeficiency virus, type 1 protease substrate specificity is limited by interactions between substrate amino acids bound in adjacent enzyme subsites. *J. Biol. Chem.* **271**:4709–4717.
 26. Rosé, J. R., R. Salto, and C. S. Craik. 1993. Regulation of autoproteolysis of the HIV-1 and HIV-2 proteases with engineered amino acid substitutions. *J. Biol. Chem.* **268**:11939–11945.
 27. Sakalian, M., J. W. Wills, and V. M. Vogt. 1994. Efficiency and selectivity of RNA packaging by Rous sarcoma virus gag deletion mutants. *J. Virol.* **68**:5969–5981.
 28. Séllos-Moura, M., and V. M. Vogt. 1996. Proteolytic activity of the NC-PR fragment of avian sarcoma and leukemia virus Gag protein expressed in *E. coli*. *Virology* **221**:335–345.
 29. Stewart, L., G. Schatz, and V. M. Vogt. 1990. Properties of avian retrovirus particles defective in viral protease. *J. Virol.* **64**:5076–5092.
 30. Stewart, L., and V. M. Vogt. 1991. *trans*-acting viral protease is necessary and sufficient for activation of avian leukosis virus reverse transcriptase. *J. Virol.* **65**:6218–6231.
 31. Stewart, L., and V. M. Vogt. 1993. Reverse transcriptase and protease activities of avian leukosis virus Gag-Pol fusion proteins expressed in insect cells. *J. Virol.* **67**:7582–7596.
 32. Stewart, L., and V. M. Vogt. 1994. Proteolytic cleavage at the Gag-Pol junction in avian leukosis viruses: differences in vitro and in vivo. *Virology* **204**:45–59.
 33. 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.
 34. Wlodawer, A., and J. W. Erickson. 1993. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* **62**:543–585.
 35. Zhang, Z. Y., R. A. Poorman, L. L. Maggiora, R. L. Heinrikson, and F. J. Kézdy. 1991. Dissociative inhibition of dimeric enzymes: kinetic characterization of the inhibition of HIV-1 protease by its carboxyl terminal tetrapeptide. *J. Biol. Chem.* **266**:15591–15594.
 36. Zybarth, G., H.-G. Kräusslich, K. Partin, and C. Carter. 1994. Proteolytic activity of novel human immunodeficiency virus type 1 proteinase proteins from a precursor with a blocking mutation at the N terminus of the PR domain. *J. Virol.* **68**:240–250.
 37. Zybarth, G., and C. Carter. 1995. Domains upstream of the protease (PR) in human immunodeficiency virus type 1 Gag-Pol influence PR autoprocessing. *J. Virol.* **69**:3878–3884.