

Differential Proteolytic Processing Leads to Multiple Forms of the CA Protein in Avian Sarcoma and Leukemia Viruses

R. BLAKE PEPINSKY,^{1*} IOANNIS A. PAPAYANNOPOULOS,¹ E. PINGCHANG CHOW,¹
NEEL K. KRISHNA,² REBECCA C. CRAVEN,² AND VOLKER M. VOGT³

Department of Protein Chemistry, Biogen, Inc., Cambridge, Massachusetts 02142¹; Department of Microbiology and Immunology, Hershey Medical Center, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033²; and Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853³

Received 6 April 1995/Accepted 18 July 1995

The CA (capsid) protein of avian sarcoma and leukemia viruses occurs in multiple species. Only one form has been previously characterized biochemically. We have now determined that the mature CA protein of avian sarcoma and leukemia viruses exists as three species with different C termini, ending in amino acid residues A-476, A-478, and M-479 of the Gag precursor, respectively. These structures were deduced from a combination of cyanogen bromide peptide mapping, sequence analysis of tryptic peptides, and electrospray mass spectrometry. The three forms of CA were detected in the same ratios in Rous sarcoma virus and avian myeloblastosis virus and therefore are likely to represent a common feature of members of this genus of avian retroviruses. The only previously reported CA species, CA^{M-479}, accounts for only about 36% of the total CA protein, while CA^{A-476} and CA^{A-478} account for 55 and 9%, respectively. From the analysis of peptides cleaved *in vitro* by PR, the viral protease, we infer that the cleavage site between A-476 and A-477 not only is recognized by PR but is the preferred site. We were unable to determine if A-478/A-479 is a cleavage site for PR or alternatively if CA^{A-478} results from further processing of CA^{M-479} by a carboxypeptidase. To study the biological significance of residues A-477 to M-479, we constructed genetically altered viruses in which deletions removed either residues 477 to 479 or 477 to 488. The resulting virus particles appeared to assemble with normal efficiencies, but the latter mutant showed slowed proteolytic processing. Neither of the mutants was infectious.

The internal structural proteins of retroviruses are formed by proteolytic processing of the Gag precursor polypeptide in the last stages or after budding of the virus particle from the infected cell. In all retroviruses the proteins cleaved from Gag include MA (matrix protein), CA (capsid protein), and NC (nucleocapsid protein). In many viruses additional proteins also are generated from the segments of Gag between MA and CA or distal to NC. Cleavage of Gag leads to a morphological change in the virus particle, called maturation, and is essential for the virus to become infectious. It is believed that in the mature particle CA forms the outer shell of the core, which encapsulates the viral RNA genome associated with NC as well as the replication enzymes reverse transcriptase and integrase. In the newly infected cell, the core is delivered to the cytoplasm, where reverse transcription occurs. In viruses of the murine leukemia virus genus, CA appears to remain associated with the newly synthesized viral DNA (5), while in the lentivirus genus this has not been observed (6, 13).

In the avian sarcoma and leukemia viruses (ASLV), typified by Rous sarcoma virus (RSV), the major Gag proteins are MA, p10, CA, NC, and PR, in order from the N terminus to the C terminus. Proteolytic processing also generates several small peptides, from between MA and p10 and between CA and NC, which are referred to as p2 and SP (for spacer), respectively. Two peptides are released from the p2 region (p2a and p2b) and are found in virions (26). The p2b region of Gag has recently been shown to have a late role in particle budding (37). The presence of SP was inferred previously from a comparison of the nucleotide sequence (30) with N- and C-termi-

nal analyses for CA and NC (3). However, to date the SP peptide has not been identified in virions. The published C-terminal analysis of CA, based on carboxypeptidase digestions, previously was interpreted to imply a single cleavage site at residue M-479 (3).

A more detailed view of CA processing has been developed recently from pulse-chase labeling and immunoprecipitation of CA proteins coupled with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (2, 10). Processing of CA proceeds sequentially with the first C-terminal cleavage believed to occur at M-488/A-489, yielding a species composed of CA plus SP, designated CA1. With time, this protein disappears and is replaced by a CA doublet (CA2 and CA3) slightly higher in the gel. It was not known why two proteins are formed, nor had the exact C terminus of either been determined. Curiously, despite its larger predicted size, CA1 migrates faster in SDS-PAGE than the doublet that is inferred to result from proteolysis of CA1. The biological function of the SP region is unknown. Genetically altered virus containing a deletion of the 27 nucleotides corresponding to SP yielded particles that appeared normal by morphology, by RNA content, and by proteolytic processing but that were noninfectious (10).

In studies to characterize ASLV Gag products, in part completed more than 10 years ago (25), we had observed that mature viruses contain multiple forms of the CA protein that differ by proteolytic processing at the C terminus. We have now definitively identified three forms of CA in mature virions. In diverse ASLV, approximately two-thirds of the CA molecules are different from the previously accepted species CA^{M-479}. These results account for the doublet of polypeptide bands observed recently by SDS-PAGE. We constructed two mutant viruses carrying deletions in this region. The mutants were found to undergo proteolytic maturation but to be noninfectious.

* Corresponding author. Mailing address: Department of Protein Chemistry, Biogen Inc., 14 Cambridge Center, Cambridge, MA 02142. Phone: (617) 252-9200, ext. 3310. Fax: (617) 679-2616.

tious. Together with previously published analyses of ASLV CA (10), these results suggest that correct C-terminal processing of CA represents a critical step in the formation of infectious virus.

MATERIALS AND METHODS

Virus. Primary chicken embryo fibroblasts from 12-day embryos were infected either with the Prague C strain of RSV (RSV PrC) or with avian myeloblastosis virus (AMV) and grown as monolayers at 39°C in Dulbecco's modified Eagle medium containing 50 U of penicillin per ml, 50 µg of streptomycin per ml, and 10% fetal bovine serum. Metabolically labeled viruses were purified from the supernatants of infected cells as described previously (27). Unlabeled AMV from leukemic chicken plasma was purchased from Life Sciences, Inc. Unlabeled RSV PrC, in the form of a concentrated suspension from supernatants of virus-infected tissue culture cells, was obtained from the Biological Carcinogenesis Branch of the National Cancer Institute. RSV PrC-infected C/O cells expressing the endogenous provirus ev-1 were a gift of R. Eisenman. Virus collected from these cells represented roughly a 1:1 mixture of RSV and ev-1-derived components.

HPLC analysis. Aliquots of PrC or AMV containing approximately 100 µg of total protein were dissolved in 150 µl of 6 M guanidine hydrochloride–12.5 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 8.0)–5 mM dithiothreitol and incubated at 37°C for 1 h. The lysate was subjected to reverse-phase high-pressure liquid chromatography (HPLC) on a C₄ column (catalog no. 214TP104; column dimensions, 0.46 cm [inside diameter] by 25 cm; Vydac) at ambient temperature. Bound components were eluted with a 30-min 0 to 70% gradient of acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 1.0 ml/min. The column effluent was monitored at 280 nm. Fractions were collected at 0.5-min intervals. Aliquots of protein-containing fractions (50 µl) were dried in a Speed Vac concentrator, dissolved in electrophoresis sample buffer, and analyzed by SDS-PAGE.

ES-MS. Electrospray-mass spectrometry (ES-MS) was carried out on a Fisons VG Quattro II triple quadrupole MS equipped with an ES ion source. A volume of 20 µl of HPLC-purified CA in 50% water–50% acetonitrile (with 0.1% trifluoroacetic acid) was directly infused into the ion source at a rate of 5 µl/min. Scans were obtained throughout the sample infusion. All ES-MS data were acquired and stored in profile mode and were processed by using the VG MassLynx data system.

C-terminal analysis of AMV CA protein. HPLC-purified AMV CA protein was dried in a Speed Vac concentrator (Savant) and resuspended in 200 µl of 100 mM ammonium bicarbonate–0.1 mM CaCl₂. The sample was digested with trypsin for 16 h at 37°C. Trypsin was added in three installments at 0, 4, and 12 h, with each installment containing 2% of the total protein by weight. Twenty percent of the digest was directly analyzed by reverse-phase HPLC on a C₁₈ column (2.1 by 250 mm; Vydac). The column was developed with a 90-min gradient (0 to 60% acetonitrile) in 0.1% trifluoroacetic acid at 0.3 ml/min. The effluent was monitored at 214 and 280 nm. Fractions were collected at 0.5-min intervals. The remainder of the sample was dried in a Speed Vac concentrator and subjected to C-terminal fragment isolation on an anhydrotypsin column by the protocol supplied by Pierce. Both effluent and eluted fractions were analyzed by HPLC. Selected peaks were subjected to N-terminal sequence analysis in an Applied Biosystems model 470 sequenator. Phenylthiohydantoin-amino acids were analyzed on line in a model 120A PTH amino acid analyzer. Samples for sequencing were loaded on Polybrene-treated discs.

CNBr mapping and SDS-PAGE. Details of the CNBr mapping method have been previously described (27). Gel slices (2 by 1.5 by 1.4 mm) containing 0.5 µg of the CA protein or strips of gel containing the regions of interest were incubated at 23°C for 1 h with 10 mg of CNBr per ml in 0.1 N HCl–0.1% 2-mercaptoethanol. CNBr-treated gel samples were washed twice for 5 min each with water, once for 5 min with 0.25 M Tris-HCl (pH 6.8), and once for 10 min at 37°C with electrophoresis sample buffer and loaded onto SDS-urea slab gels. The cleavage products were subjected to SDS-PAGE in polyacrylamide gels containing 8 M urea.

Viral proteins were subjected to SDS-PAGE by using the Laemmli gel system and detected by staining with Coomassie brilliant blue, silver staining (39), or fluorography as indicated. Prior to electrophoresis, samples were heated at 65°C for 10 min in electrophoresis sample buffer (50 mM Tris HCl [pH 6.8], 2% SDS, 12.5% glycerol, 0.1% bromophenol blue). To locate polypeptides to be excised from a gel for CNBr mapping, bands were visualized without fixation either by brief emersion of the gels in ice-cold 0.5 M KCl–0.1% 2-mercaptoethanol or by rapid surface staining of parallel tracks with Coomassie blue (9). To facilitate the identification of modified variants, the gel slices were cut orthogonal to band. The partial resolution of the CA species that is achieved in the electrophoresis step is critical to the analysis. Gel slices were stored at –20°C.

In vitro digestions with AMV PR. AMV PR was purified from virus by a one-step organic extraction method that was previously described (21). The protease was stored at –20°C in the extraction cocktail, which we found to be an effective method for retaining activity and preventing aggregation. Prior to use, the methanol was removed and the protein concentrated approximately fivefold in a Speed Vac concentrator (Savant). A series of synthetic peptides that con-

tained the potential cleavage sites for CA (LTDQGIA, LTDQGI A A A M, and LTDQGI A A A M S S A I) were obtained from Research Genetics (Huntsville, Ala.). Each peptide was preparatively purified on a C₄ column and was determined to be greater than 95% pure on the basis of analytical HPLC analyses. The peptides were dried in a Speed Vac concentrator, stored at –20°C, and then suspended in dimethyl sulfoxide immediately prior to use. The peptides were diluted into PR digestion buffer (50 mM sodium acetate [pH 5.5], 900 mM NaCl) so that the final concentration of dimethyl sulfoxide was 20% by volume. The protease was added, and samples were incubated at 37°C for 48 h. The digestion mixture (30 µl), containing 5 µg of peptide and 10 µg of PR, was diluted with 150 µl of 0.1% trifluoroacetic acid. The cleavage products were fractionated by reverse-phase HPLC on a C₄ column under the same conditions described above for fractionating the intact viral proteins. The column effluent was monitored at 214 nm, and fractions were collected at 0.5-min intervals. Selected samples were sequenced by tandem MS where specific mass ions were selected in the first MS and then fragmented by collision-induced dissociation in the second MS to determine the structure (4).

Construction and expression of spacer deletion mutants. Two deletion mutations were constructed by in vitro mutagenesis of the RSV PrC *gag* gene cloned into M13 (38). The deletions ΔSP3 and ΔSP12 eliminate the peptide segments AAM and AAMSSAIQPLIM, respectively. The sequences of the mutagenic oligonucleotides are 5'-GGATCAAGGCATAGCCTCGTCTGTATCCAGC-3' for ΔSP3 and 5'-GGATCAAGGCATAGCCGAGTAGTCAATAGAGA G-3' for ΔSP12. Progeny phage from the mutagenesis reactions were screened by sequencing of the single-stranded DNA. For each mutation, the *gag* genes from two or more independent isolates were subcloned into the proviral vector RCAN.HISV (11) by using the *SacI*-*HpaI* sites, replacing the wild-type gene. ΔSP, a deletion of the nine residues SSAIQPLIM described previously (10), in this paper is called ΔSP9 for clarity. The viral genomes were transfected into QT6 quail cells, and the ability of the mutant Gag proteins to direct particle assembly was evaluated by labeling with [³⁵S]methionine followed by immunoprecipitation with anti-RSV serum, as described previously (10).

RESULTS

Detection of a second CA component by cyanogen bromide mapping. We have determined that the mature ASLV virion contains multiple forms of CA. This heterogeneity was detected first by a two-dimensional analysis in which the virion proteins were subjected to SDS-PAGE, treated in situ with cyanogen bromide to partially cleave the polypeptide at methionine residues (27), and then reanalyzed by SDS-PAGE. For [³⁵S]cysteine-labeled proteins from RSV PrC, this analysis revealed first that the four major Gag proteins, MA, CA, NC, and PR, have unique patterns of cleavage products, as expected (Fig. 1A, lanes c, b, e, and f). The digestion profile of CA (lane b) shows four major products: intact CA at 24 kDa and fragments at 16, 10, and 6 kDa. Intact CA appears as a discrete spot, while each of the cleavage products is split into two components, with the separation of the components becoming progressively larger with decreasing fragment size. An arrowhead marks the position of the 6-kDa fragments, which appear to be separated by about 500 Da. The pattern of fragments thus suggests that mature CA exists both as a 27.0-kDa species and a 26.5-kDa species. Although these species are not distinct in the undigested CA, partial resolution in the cleavage products is indicated by the skewed appearance of the spots. The results of the subsequent studies, described below, verify that these species correspond to the doublet of bands recently observed by radioimmune precipitation analysis of Gag-expressing cells (2, 10). The MA protein also was observed to be split into two components (lane c). In this case the shift in mobility was shown previously to be caused by partial phosphorylation (27).

To examine the generality of the two CA components, we also analyzed CNBr cleavage patterns for CA from two other ASLV. One was a recombinant RSV that carries the *gag* gene of the endogenous virus ev-1 (1) (Fig. 1B, lane g). The CA encoded by the endogenous viruses ev-1 and ev-2 is known to migrate more slowly than RSV CA (31). The ev-1 and PrC profiles are very similar, both exhibiting evidence of the two CA components. However, the ev-1 profile is simpler because

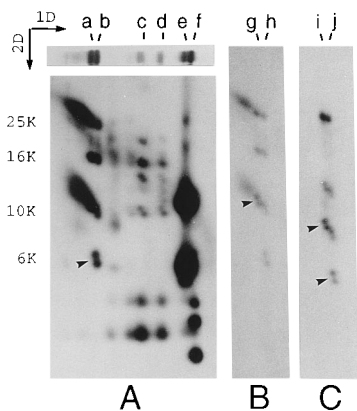


FIG. 1. Detection of second form of CA protein in ASLV by two-dimensional SDS-PAGE. Samples containing a mixture of [35 S]cysteine-labeled proteins from RSV PrC and from ev-1 were dissolved in electrophoresis sample buffer and subjected to SDS-PAGE. Gel strips containing the entire viral profile were treated with CNBr, and the cleavage products were electrophoresed in an orthogonal direction into an SDS-urea polyacrylamide gel. Labeled fragments were detected by fluorography. Only the relevant regions from the gels are shown. (A) 5- to 35-kDa region of the gel. The two-dimensional (2D) cleavage data are displayed in the lower panel. Values on the left indicate the apparent masses of fragments based on their electrophoretic mobility. The small image in the upper panel is a typical one-dimensional (1D) SDS-polyacrylamide gel profile shown to facilitate identification of the Gag components. (B) Lighter exposure of CA region from the two-dimensional gel shown in panel A. (C) CA region from a two-dimensional gel that had been developed by using the Laemmli system in place of the urea gel. Lanes: a, g, and i, ev-1 CA; b, h, and j, PrC CA; c, PrC MA; d, ev-1 MA; e, NC; f, PR. Arrowheads indicate the positions of fragments discussed in the text.

its CA contains one less Met residue; the residue at position 396 is not a Met as in RSV. The presence of the second component is readily apparent from the 10-kDa cleavage fragment (denoted with an arrowhead). Similar results were obtained with CA from AMV (Fig. 2, lane j). A striking feature of cleavage profiles was that the ratio of the CA components always was similar. The ratio also appeared to be independent of the age of the virus preparation (data not shown).

To better understand the nature of the difference between the two CA components, and to localize the difference on the CA polypeptide backbone, we performed a series of CNBr mapping studies on CA metabolically labeled with amino acids. Cysteine is ideally suited for this type of analysis because it occurs only once in CA, in the very C-terminal CNBr fragment. As a consequence, the smallest labeled fragment on SDS-PAGE will be the C-terminal fragment, and all larger labeled fragments will be partial cleavage products that end in the C-terminal fragment. This nesting of labeled fragments allows the N- to C-terminal order of CNBr cleavage products to be deduced readily. Thus, we infer for [35 S]cysteine-labeled RSV CA (Fig. 2, lane g) that the 6-kDa fragment is located at the C terminus, the 10-kDa fragment contains the 6-kDa fragment plus an additional 4 kDa upstream (toward the N-terminal side), the 16-kDa fragment contains the 10-kDa fragment plus an additional 6 kDa upstream, and intact CA contains the remainder of the sequence that is upstream of the 16-kDa fragment. These assignments are consistent with predicted masses of fragments based on the amino acid sequence. All of the cysteine-labeled CNBr fragments show the two components, implying that the two components differ at or near the C terminus of CA. Similar conclusions were drawn from mapping data of [35 S]cysteine-labeled AMV CA (Fig. 2, lane j). Like the ev-1 CA profile, the cleavage profile for AMV CA contains only three prominent CNBr fragments, since in AMV residue

396 also is not a methionine. All three AMV fragments showed the two components of CA.

Since the Gag proteins are generated by proteolysis, the possibility of an additional processing site in CA seemed the most likely explanation for the two components of CA. To investigate this possibility, we devised a simple test for monitoring the integrity of the previously accepted C terminus of CA, Met-479, using CNBr mapping of [3 H]methionine-labeled protein. CNBr cleaves to the C-terminal side of methionine residues. Thus for most proteins, the C-terminal CNBr fragment should be unlabeled. However, since the last residue in CA itself is a methionine, the C-terminal CNBr fragment should retain radioactivity. If alternative proteolytic processing is responsible for the two CA components, and this leads to the loss of Met-479 because of cleavage farther upstream, then the larger of the two CA components should be labeled while the smaller one should not. Indeed this is what we observed (Fig. 2, lane h). When [3 H]methionine-labeled CA was cleaved with CNBr and analyzed, the upper of the spots corresponding to the C-terminal fragment was radioactive and the lower one was devoid of radioactivity. The differences in labeling intensities of the C-terminal spot and the spot corresponding to the next smallest cleavage product are consistent with the methionine content of the fragments, representing one and five methionine residues, respectively.

Primary structure of CA components. Advances in MS by using ES ionization have revolutionized the ability to evaluate protein structure. The masses of proteins of up to 100,000 Da can be determined to an accuracy of better than $\pm 0.005\%$, corresponding to an error of less than 1 Da for a protein of 250 amino acid residues (32). In this technique, the protein molecule becomes multiply charged, usually because of protonation

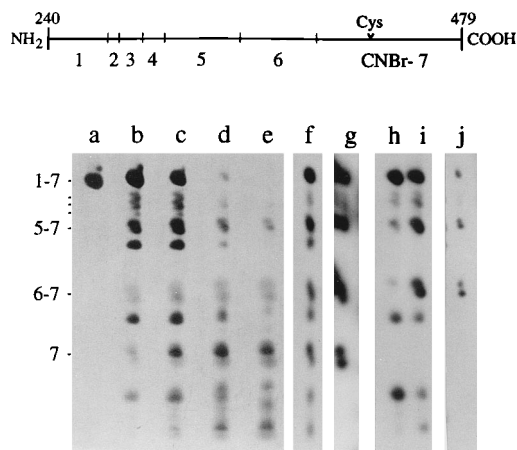


FIG. 2. Localization of site of modification by CNBr mapping. Samples containing unlabeled RSV PrC CA protein (lanes a to e) or radioactive CA from PrC (lanes f to g) and from AMV (lanes h to j) were treated with CNBr and analyzed by SDS-urea PAGE. Cleavage products were visualized by silver staining (lanes a to e) or fluorography (lanes f to j). Lanes: a, no CNBr; b, 3 mg of CNBr per ml; c, 10 mg of CNBr per ml; d, 30 mg of CNBr per ml; e, 100 mg of CNBr per ml; f and i, [14 C]glycine-labeled CA; g and j, [35 S]cysteine-labeled CA; h, [3 H]methionine-labeled CA. The numerical designations on the left mark the positions of the C-terminal (CNBr fragment 7)-containing cleavage products. The positions of fragments 4 to 7, 5 to 7, and 6 and 7 are indicated with hash marks. The schematic at the top of the figure summarizes important structural features of the RSV CA protein. Six internal methionines at positions 276, 283, 296, 309, 352, and 396 dissect the protein at positions indicated (|). Numbering is based on the position of the residues in the gag precursor. The resulting CNBr fragments are denoted 1 to 7, where 1 is the N-terminal fragment and 7 is the C-terminal fragment. Cys-431 (Cys), the unique target of the cysteine labeling, is also indicated.

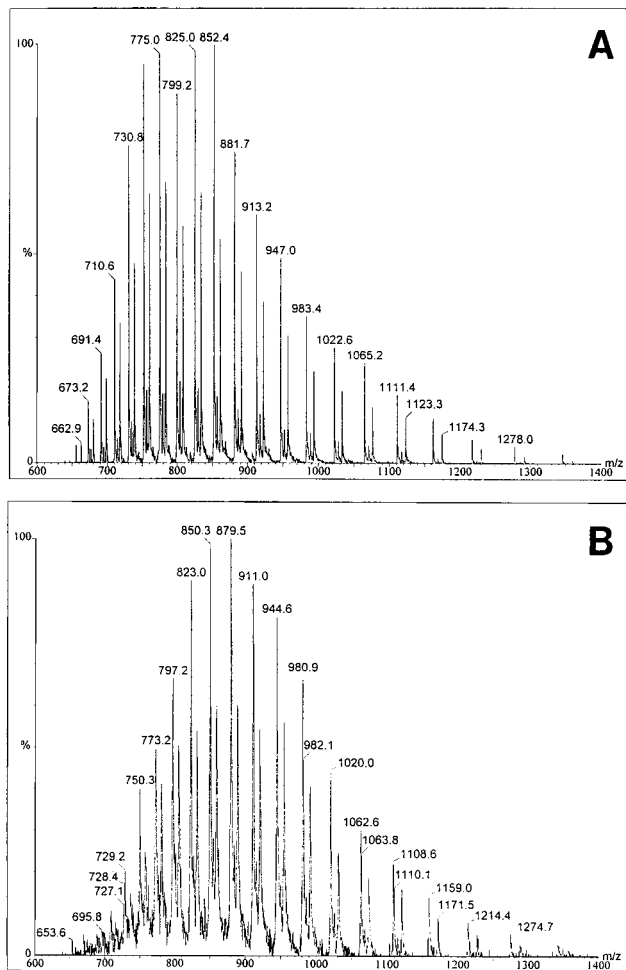


FIG. 3. ES-MS data for CA proteins. The CA proteins from RSV and AMV were analyzed by MS on a VG Quattro II quadrupole instrument. The m/z data that were collected are shown. (A) RSV PrC CA; (B) AMV CA.

at basic residues. The mass-to-charge ratio for each charge state is measured by the MS, and from these data the molecular mass of the protein can be computed. Previously, we described a simple one-step purification for ASLV CA, using reverse-phase HPLC on a C_4 column (26). The final product was greater than 95% pure by SDS-PAGE. When the HPLC-purified CA from RSV PrC was subjected to ES-MS, two major components were observed (Fig. 3A and 4A). The 25,815-Da value agrees exactly with the mass predicted for CA^{M-479} by the nucleotide sequence, 25,815.3 Da. The 25,542.0-Da value differs in mass by 273 Da, which is exactly the mass of Ala-Ala-Met. We therefore infer that the second major CA component is generated by proteolytic processing at a previously unrecognized cleavage site three amino acid residues upstream of M-479. A minor component with a mass of 25,684 Da also was observed, differing from CA^{M-479} by 131 Da. This difference corresponds to the mass of a methionine residue, and thus it is likely that the minor CA component simply is missing the C-terminal methionine residue at position 479. We therefore conclude that the three components are all generated by proteolysis. Since the mass differences for the three peaks are small relative to the mass of the intact CA protein, the ratio of species is likely to be a good measure of the ratio of the products in the virion. On the basis of peak

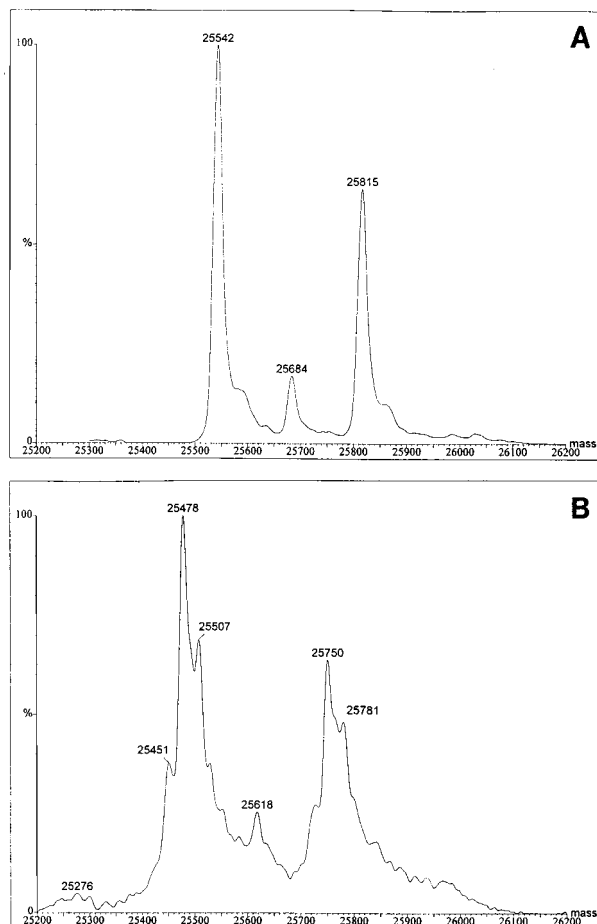


FIG. 4. Molecular mass spectra for CA proteins. The ES-MS m/z data that had been collected and are shown in Fig. 3 were transformed into molecular mass spectra by using the VG data system software. (A) RSV PrC CA; (B) AMV CA.

heights, CA is composed of 55% CA^{A-476}, 9% CA^{A-478}, and 36% CA^{M-479}.

The AMV CA protein also was subjected to ES-MS (Fig. 3B and 4B). Unlike RSV PrC, AMV is a mixture of closely related viruses, including the helper viruses myeloblastosis-associated viruses A and B in addition to the actual replication-defective transforming virus. Thus, a mixture of CA species might be anticipated. In fact, we observed at least three different CA proteins in the spectrum. For each, a corresponding peak that differed in mass by 273 Da was identified. Similarly, a minor component, which differed in mass by 132 Da, was observed. We thus infer that the viruses that make up the AMV stock contain the same forms of CA as those found for the cloned RSV PrC, namely CA species ending in GIAAAM-479, GIAAA-478, and GIA-476. The ratios of the three components were also similar to those for RSV, with the GIA-476 component being predominant.

As an independent confirmation of the identities of the different CA components, we used a combination of tryptic peptide mapping and N-terminal sequencing of peptides. To rapidly screen for peptides of interest, the tryptic digest was loaded onto an anhydrotrypsin column, and then all the peptides that failed to bind were characterized. Anhydrotrypsin binds peptides with Lys or Arg at their C termini, and therefore only tryptic fragments from the C terminus of CA should fail to be retained on the column. The anhydrotrypsin flowthrough

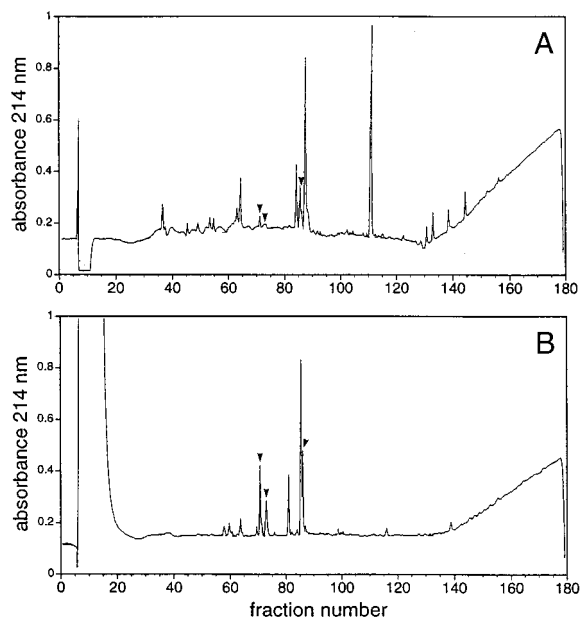


FIG. 5. Purification of C-terminal tryptic peptides from AMV CA. HPLC-purified AMV CA (25 μ g) was digested with trypsin and either analyzed by reverse-phase HPLC directly (A) or after enriching for the C-terminal peptides on an anhydrotrypsin column (B) as described in the text. The column effluent was monitored at both 214 and 280 nm (only the 214-nm data are shown), and 150- μ l fractions were collected. The peak fractions from the anhydrotrypsin-treated digest were submitted to N-terminal sequencing. (A) 50 pmol of digest, (B) 200 pmol of digest.

fraction showed five major peaks (Fig. 5B), all of which were subjected to sequencing. The sequence obtained from fraction 86 (TAPLTDQGIAAAM) is that predicted for CA^{M-479}, as had been inferred many years ago from carboxypeptidase digestions (3). The sequences obtained from fractions 70 (TAPLTDQGIA) and 72 (TAPLTDQGIAAA) are those predicted for the species CA^{A-476} and CA^{A-478}, respectively. The two other peaks contained internal peptides that had been generated by a chymotryptic activity, a common problem with trypsin digestions.

Characterization of PR cleavage site by in vitro proteolysis.

We performed in vitro cleavage studies in order to ascertain that the A-476/A-477 cleavage site in fact is a target of the avian viral protease. The synthetic peptide LTDQGIAAAMSSAI was synthesized and used as a substrate. This peptide contains all three predicted cleavage sites at the end of CA. The products generated were analyzed by reverse-phase HPLC, with the peptides LTDQGIA and LTDQGIAAAM being used as markers, as these are the expected N-terminal fragments resulting from cleavages at A-476/A-477 and at M-479/S-480, respectively. The substrate peptide was found to be hydrolyzed predominantly at A-476 (Fig. 6B), since the main digestion product comigrated with the peptide standard LTDQGIA. On the basis of the relative peak heights, at least 80% of the cleavage occurred at this site. Sequencing by tandem MS confirmed that peak a had the predicted sequence LTDQGIA. Peak b was also characterized by sequencing and shown to contain the complementary half of the peptide, AAMSSAI. While the minor peak in Fig. 6B corresponded in position to peptide LTDQGIAAAM, the fact that a peak at this position was already present in the protease preparation indicates that not all of this material can be the product of cleavage at M-479/S-480. In the experiment shown, this minor peak did

increase slightly after digestion of LTDQGIAAAMSSAI with PR (compare Fig. 6B and C), but in other digests no such increase was observed (data not shown). In sum, these results indicate a very strong preference of PR for cleavage at A-476/A-477.

The near absence of cleavage of the peptide at M-479/S-480 is surprising. To test if the peptide cleavage results correlate with in vitro processing of a native protein, we also digested a fragment of Gag that had been purified after expression in *Escherichia coli*. This protein, CA-NC, is exactly the segment of Gag corresponding to the CA, spacer, and NC domains. The CA species generated after incubation for various times with PR were characterized by ES-MS (unpublished results). The first CA product generated was CA^{M-488}, as also found in vivo. Then after longer incubation times CA^{A-476} and CA^{M-479} were formed. The ratio of the CA^{A-476} and CA^{M-479} products was 3:1 at low protease concentrations and 4:1 with higher amounts of protease. We infer from these data that cleavage of the peptide substrates approximates cleavage of proteins. A possible explanation for the differences in the ratios of CA^{A-476} to CA^{M-479} after maturation of virions compared with digestion in vitro is that CA^{M-479} is further processed by PR in vitro. To test this possibility we treated the peptide LTDQGIAAAM with PR and then looked for formation of LTDQGIA by the reverse-phase HPLC assay. There was no evidence for processing, even after a 96-h incubation with PR (data not shown).

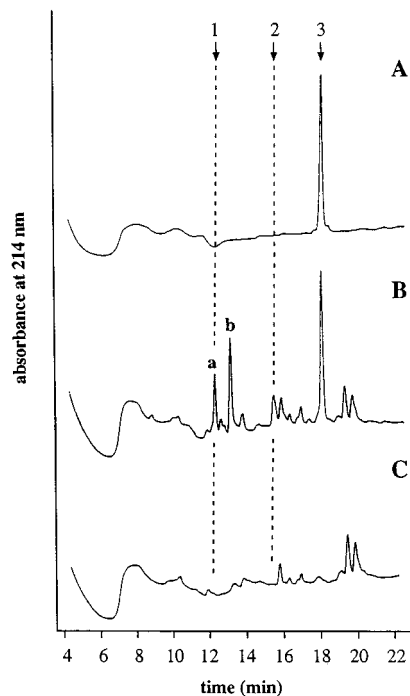


FIG. 6. AMV PR cleavage data for peptide LTDQGIAAAMSSAI. A synthetic peptide containing potential C-terminal cleavage sites of CA was incubated at 37°C for 48 h with AMV PR, without added protease, or the protease was incubated without added peptide and the extent of cleavage was assessed by reverse-phase HPLC on a C₄ column. The column effluent was monitored at 214 nm, and fractions were collected at 0.5-min intervals. Relevant column fractions were subjected to sequencing by ES-MS. Partial chromatograms from the HPLC runs are shown. (A) peptide alone; (B) peptide + PR; (C) PR alone. The arrows designated 1, 2, and 3 mark the positions of a series of synthetic peptides that were used to analyze the cleavage data and refer to peptides LTDQGIA, LTDQGIAAAM, and LTDQGIAAAMSSAI, respectively. LTDQGIAAAMSSAI peptide cleavage products that were identified by sequencing are denoted a and b in panel B.

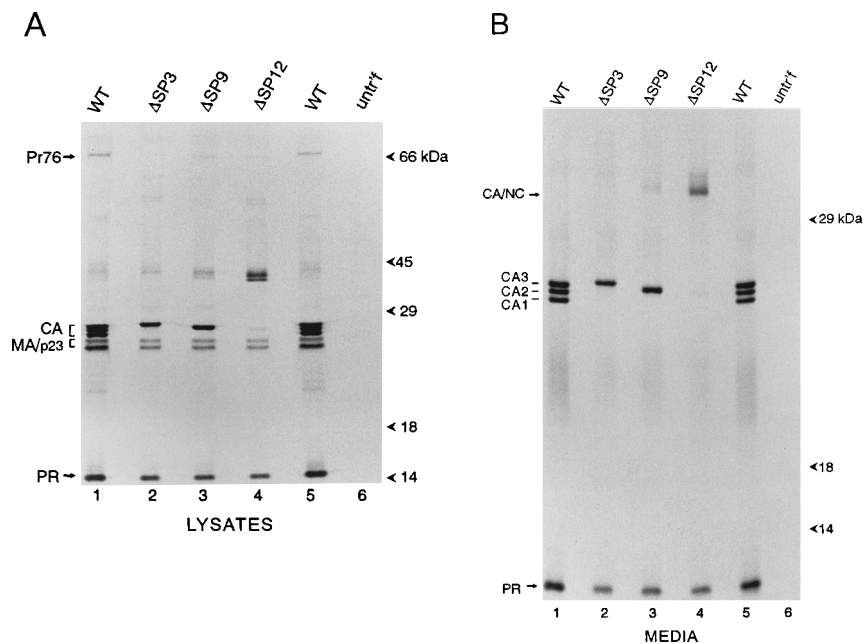


FIG. 7. Effect of spacer peptide deletions on virion release. Mutations were created in the RSV proviral vector pBH.RCAN.HiSV to precisely delete the peptides that are removed from the C terminus of CA during maturation. If the new junction acts as a PR recognition site, then the Δ SP12 mutation should result in formation of CA^{A-476} species. The Δ SP9 mutant, previously described (10), is predicted to form the CA^{M-479} species. The Δ SP3 mutation deletes the three amino acids in between the A-476/A-477 and M-479/S-480 cleavage sites; this protein species does not occur naturally. Plasmid DNA was introduced into quail cells of the QT6 line by calcium phosphate-mediated transfection. Twelve hours later, proteins were labeled with [³⁵S]methionine for 3 h and then analyzed by immunoprecipitation with an anti-RSV serum followed by SDS-PAGE. (A) Gag proteins from lysates of transfected cells were analyzed on a 10-cm-long 12% polyacrylamide gel (37) that was run at 4°C to keep the CA bands tight; (B) particles from the labeling medium were analyzed as described for panel A except that the gel was 25-cm long and contained 13% polyacrylamide. Only the portion of the gel below the 45-kDa marker is shown. The positions of molecular weight standards are indicated on the right. WT, wild type; untrf, untransfected cells.

Properties of virions with spacer peptide deletions. Previous studies had shown that RSV maturation in both avian and mammalian cells is a sequential process. The first CA species visible by SDS-PAGE (CA1), which we had inferred to correspond to CA-SP and thus to arise from cleavage at M-488/A-489, is further modified, presumably by cleavage, to yield a tight doublet of bands called CA2 and CA3 (2, 10) (Fig. 7B, lanes 1 and 5). The fact that the larger CA1 migrates faster than CA2 and CA3 originally made this identification somewhat tenuous. The observation that a mutant called Δ SP (here renamed Δ SP9 for clarity), with a deletion of the nine-residue spacer segment S-480 to M-488 exhibited only a single CA form with mobility similar to that of CA3, provided support for the conclusion that CA3 corresponds to CA^{M-479}. This reasoning would predict that the new capsid protein described above, CA^{A-476}, should be CA2. To test this prediction, we constructed the mutant Δ SP12, in which the entire 12-amino-acid spacer segment from A-477 through M-488 is deleted. If the newly created junction QGIA/AVVN is recognized and cleaved by PR, then the only form of CA produced should be CA2. Indeed, expression in quail cells of the Δ SP12 Gag protein resulted in the formation of a small amount of a protein with the mobility of CA2, which was visible both in cell lysates (Fig. 7A, lane 4) and in virions shed into the medium (Fig. 7B, lane 4). However, most of the CA domain remained incompletely processed in this mutant, as evidenced by the small amount of mature CA and by the presence of a product of ca. 35 kDa. We presume this product to be CA-NC, on the basis of its size and its reactivity with both anti-CA and anti-NC antisera (data not shown). Thus, the newly created site must be cleaved inefficiently by PR.

We also constructed a mutant with the three residues A-477

to M-479 deleted. Expression of the Δ SP3 Gag protein resulted in formation of a single CA species, presumably because of cleavage at the unaltered site M-488/A-489 (Fig. 7, lanes 2). The migration of this mutant CA species was slower than that of CA2 and CA3, consistent with its size. We therefore presume that the anomalously rapid migration of CA1 in SDS-PAGE may be traced to the segment of polypeptide encompassing the deletion. However, further work would be needed to establish definitively if the mutant CA band formed in fact has the C terminus predicted. Since the relative amounts of Gag protein shed into the medium in Δ SP3, Δ SP12, and Δ SP9 all are similar to that in the wild type, none of these mutants appears to grossly affect virus assembly in this system.

There is considerable uncertainty about the structure of the capsid shell believed to be composed of CA protein and to surround the NC-RNA ribonucleoprotein in retroviruses. This shell appears to be quite labile and has not been isolated in a clean form and in good yield from ASLV or indeed from most retroviruses. However, when ASLV virions are treated under defined conditions with nonionic detergents, a portion of CA (primarily the mature CA2 and CA3 forms) is observed to remain in a sedimentable form (10, 33). The capsid in the spacer mutant Δ SP9 was previously reported to be more easily disruptable by detergent than wild-type virus, as evidenced by this sedimentation assay (10). We have extended these observations to the new mutants Δ SP12 and Δ SP3. These mutant virions as well as wild-type and Δ SP9 virions were treated with 1% Triton X-100 and centrifuged, and the pellet fraction was displayed by SDS-PAGE. Both of the new mutants were identical in behavior to Δ SP9, with less than 2% of the total CA being found in the pellet, under conditions where the parallel wild-type virus showed approximately 30% of the total amount

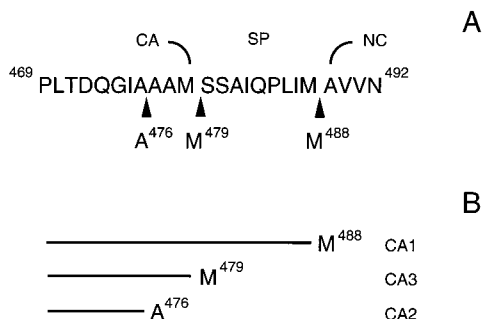


FIG. 8. Schematic summary of cleavage data. (A) Amino acid sequence of C terminus of CA and spacer region that connects CA and NC. Arrowheads denote the major protease cleavage sites. (B) Predicted structures for CA1, CA2, and CA3.

of CA2 plus CA3 in the pellet (data not shown). Thus, diverse mutants that disrupt the C terminus of CA and/or the spacer region make the capsid shell more sensitive to detergent.

To address the biological significance of the complicated C-terminal processing of ASLV CA, we assayed the Δ SP3 and Δ SP12 mutants for infectivity. Earlier results had demonstrated the lack of infectivity of Δ SP9 (10). The mutations were introduced into an otherwise infectious viral DNA carrying a dominant selectable marker for hygromycin resistance (11). The three mutant genomes and the wild-type genome as a control were transfected into QT6 quail cells, and 3 days later the cells were selected for hygromycin resistance. The wild-type virus was observed to spread efficiently from the transfected cells to neighboring ones, and almost all of the cells became hygromycin resistant. By contrast, no hygromycin-resistant colonies in excess of the background (mock-transfected cells) were seen in any of the cultures that received Δ SP3, Δ SP9, or Δ SP12. Although this evaluation is not quantitative, previous experience with this rapid screening method has shown that replication of mutants with 100-fold lower infectivity can still be detected in this way. Hence, we conclude that for all the spacer mutants, infectivity is either absent or greatly attenuated. Although the interpretation is complicated because of the incomplete cleavage of the CA domain, these results suggest that the C terminus of CA and the spacer region has an important function for virus structure or function.

DISCUSSION

We have demonstrated that in mature ASLV, the CA protein exists as three species, CA^{A-476}, CA^{A-478}, and CA^{M-479} (Fig. 8). The proteins ending in M-479 and A-476 appear to correspond to the CA3 and CA2 species, respectively, previously detected in RSV (10). Only the longest form, CA^{M-479}, had been characterized biochemically prior to this work (3). The ratio of the three species appears to be the same in diverse ASLV and to be independent of the virus purification scheme, suggesting that the ratio is biologically important. At least the major forms of CA that end at A-476 and M-479 are generated by PR-mediated cleavage, and the sequences at these sites conform to the loose PR recognition consensus sequences (7). Proteolytic processing at the C-terminal end of CA begins with cleavage at M-488/A-489, leading to the intermediate species CA-SP (CA1), which is not found in mature virions (Fig. 8). We envision that subsequent cleavage takes place either at A-476/A-477 or at M-479/S-480, leading to CA^{A-476} (CA2) or CA^{M-479} (CA3), respectively. Presumably further trimming of the CA^{M-479} molecule to produce CA^{A-476} does not occur or

occurs slowly. This latter interpretation is based on two observations. First, under the conditions tested, PR does not cleave the synthetic peptide LTDQGI A AAM in vitro. Second, in the Δ SP9 mutant the CA3 species does not show evidence of further processing to yield CA2. In pulse/chase experiments with wild-type virus, the ratio of CA2 to CA3 appears to remain constant in periods up to 3 h (2), and this ratio reflects the abundance of the different proteins in mature virus (this study). The observation that the minor CA^{A-478} species was not detected in any of the in vitro cleavage studies suggests that it is not the consequence of PR-mediated cleavage. We speculate that it is generated from CA^{M-479} by a carboxypeptidase activity. This idea would be consistent with the observation that in virions the C-terminal Met residue of p10 is missing (27a). The significance of these events is unknown.

Several factors help explain why the major species CA^{A-476} had not been characterized previously. First, the original C-terminal analysis performed with carboxypeptidase, which identified Ala-Met-C as the C terminus of CA (3), was not quantitative enough to have allowed the multiple cleavage sites to be discerned. The predominance of free methionine in the released amino acids could have resulted from the greater accessibility of the C terminus in CA^{M-479}, while the free alanine interpreted to arise from the second residue from the C terminus probably resulted in part from the two CA species ending in Ala. Second, the 1% mass difference between CA^{A-476} and CA^{M-479} is too small to be easily detected by SDS-PAGE under most standard conditions. This doublet of bands is readily apparent only when small amounts of protein are run on gels, when a light exposure of radioactive CA is analyzed, or when gels are run under special conditions, such as in the cold. The detection of the CA^{A-476} component was facilitated by evaluation of CNBr fragments, since the mass difference is proportionally larger for smaller polypeptides. However, only by use of ES-MS were we able to exactly define the several forms of CA. High-resolution ES-MS no doubt will be of great utility in the molecular characterization of other proteins in the future and has already provided information on the molecular structures of the other Gag proteins in ASLV (25a).

Although the full significance of multiple processing sites between CA and NC is not known, the available evidence suggests that the regulated cleavage of peptides from the end of CA plays a role in the proper organization of the virion core during the final stages of maturation. Deletions of peptides in the Δ SP3, Δ SP9, and Δ SP12 mutants had no detrimental effect on budding by the Gag protein, but the particles released were noninfectious. The greater-than-normal sensitivity of the CA protein to detergent extraction suggests that intermolecular interactions that anchor the mature CA proteins in the wild-type core are lacking in these mutants. We speculate that the spacer peptides and their regulated cleavage are important in directing the dynamic processes that allow such interactions to be formed.

Certain features of the CA maturation of RSV also are found in other retroviruses. In human immunodeficiency virus type 1 (HIV-1), the first cleavage at the p2/NC junction yields CA-p25, which is further processed by removal of a 14-residue peptide to form CA-p24 (16, 23, 35). These two species would thus appear to be analogous to ASLV CA^{M-488} (CA1) and CA^{A-476} (CA2), respectively. A third cleavage site in HIV-1, comparable to the M-479/S-480 site of ASLV, also has been reported previously (17). Indeed, the CA/NC junction regions of ASLV and HIV-1 show a limited degree of sequence similarity (GIA/AAM/SSAIQPLIM/AV for ASLV and ARL/AEAM/SQVTNPATIM/IG for HIV-1, where the skills indi-

cate cleavage sites). The significance of these similarities remains to be explored.

The effects of spacer peptide mutations on HIV-1 are more striking than those in ASLV. Deletion of the 14-residue spacer from HIV-1 Gag has been reported by one laboratory to destroy infectivity without preventing particle release (28), whereas another group has found dramatic effects on both budding capacity and particle morphology (22). Furthermore, a point mutation that prevented cleavage of p2 from CA/p25 also disturbed the particle morphology (14). Thus, although such a role has not been observed in ASLV, the HIV-1 data suggest an influence of the spacer peptide on budding or on the organization of the Gag protein in the budding particle. Although the spacer peptides of other viruses have not been studied extensively, cleavage at multiple sites between CA and NC has been demonstrated in simian immunodeficiency virus (15), feline immunodeficiency virus (12), bovine immunodeficiency virus (34), and equine infectious anemia virus (19), providing further evidence of the importance of this arrangement for the replication of diverse retroviruses. The only retrovirus in which CA has been clearly shown to have a unique C terminus is murine leukemia virus (18, 24), although in this virus heterogeneities in CA have been noted by isoelectric focusing (20).

Precisely what function the individual CA species have in the virion is a matter of speculation. The three-dimensional structure of the CA protein might shed light on this question. The fact that diffracting crystals have been grown both for HIV-1 CA complexed with an Fab antibody fragment (29) and for RSV CA (21a) encourages the expectation that these structures will be identified soon. To understand CA function it will be necessary to unravel the structure of the capsid shell, which is generally believed to be composed of about 2,000 CA molecules. Numerous models have been drawn for retroviral capsids, with capsids of C-type retroviruses such as ASLV usually being depicted as icosahedra and those of lentiviruses such as HIV-1 being depicted as cones. While a cone shape is often visible in thin-section electron micrographs, icosahedral shapes have been visualized only rarely and are based, to a large extent, on conjecture. Regardless of the exact structure of the capsid, however, the CA molecules must pack together in a regular fashion to form it. Thus, the principles of quasiequivalence as formulated for icosahedral viruses (8) may also apply even if the structure is not a simple icosahedron. We hypothesize that in ASLV and perhaps other retroviruses the extreme C-terminal segment of CA acts to modulate contacts with neighboring protein molecules. For example, by analogy with icosahedron-shaped viruses, one form of CA could make the most-stable contacts on a 6-fold axis while another could make the contacts on a 5-fold axis. Testing of this model will require determination of a three-dimensional structure for the viral capsid itself or perhaps of a smaller structure that could be formed *in vitro*.

ACKNOWLEDGMENTS

We thank Betsy O'Brine-Greco for assistance in the tryptic peptide mapping and Joyce Ho for assistance in sequencing the tryptic peptides.

V.M.V. is supported by Public Health Service grant CA-20081. R.C.C. and N.K.K. are supported by Public Health Service grant CA-47482 to John W. Wills, in whose laboratory a portion of this work was conducted, and by funds from the Pennsylvania State School of Medicine.

REFERENCES

- Astrin, S. M. 1978. Endogenous viral genes of white leghorn chickens: common site of residence and sites associated with specific phenotypes of viral gene expression. *Proc. Natl. Acad. Sci. USA* **75**:5941-5945.
- Bennett, R. P., S. Rhee, R. C. Craven, E. Hunter, and J. W. Wills. 1991. Amino acids encoded downstream of *gag* are not required by Rous sarcoma virus protease during Gag-mediated assembly. *J. Virol.* **65**:272-280.
- Bhown, A. S., J. C. Bennett, and E. Hunter. 1980. Alignment of the peptides derived from acid-catalyzed cleavage of an aspartylprolyl bond in the major internal structural polypeptide of avian retroviruses. *J. Biol. Chem.* **255**:6962-6965.
- Biemann, K., and I. A. Papayannopoulos. 1994. Amino acid sequencing of proteins. *Acc. Chem. Res.* **27**:370-378.
- Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus. 1989. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* **3**:469-478.
- Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G. Tarpley, and M. Stevenson. 1993. Association of integrase, matrix and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. USA* **90**:6125-6129.
- 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 the Rous sarcoma virus and human immunodeficiency virus type 1 proteases. *J. Biol. Chem.* **268**:11711-11720.
- Caspar, D. L. D., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**:1-24.
- Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli. 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* **252**:1102-1106.
- Craven, R. C., A. E. Leure-duPree, C. R. Erdie, C. B. Wilson, and J. W. Wills. 1993. Necessity of the spacer peptide between CA and NC in the Rous sarcoma virus Gag protein. *J. Virol.* **67**:6246-6252.
- Dong, J., J. W. Dubay, L. G. Perez, and E. Hunter. 1992. Mutations within the proteolytic cleavage site of the Rous sarcoma virus glycoprotein define a requirement for dibasic residues for intracellular cleavage. *J. Virol.* **66**:865-874.
- Elder, J. H., M. Schnolzer, C. S. Hasselkus-Light, M. Henson, D. A. Lerner, T. R. Phillips, P. C. Wagaman, and S. B. Kent. 1993. Identification of proteolytic processing sites within the Gag and Pol polyproteins of feline immunodeficiency virus. *J. Virol.* **67**:1869-1876.
- Farnet, C. M., and W. A. Haseltine. 1991. Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J. Virol.* **65**:1910-1915.
- Göttlinger, H. G., J. G. Sodroski, and W. A. Haseltine. 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **86**:5781-5785.
- Henderson, L. E., R. E. Benveniste, R. Sowder, T. D. Copeland, A. M. Schultz, and S. Oroszlan. 1988. Molecular characterization of *gag* proteins from simian immunodeficiency virus (SIV_{MNe}). *J. Virol.* **62**:2587-2595.
- Henderson, L. E., M. A. Bowers, R. C. Sowder, S. A. Serabyn, D. G. Johnson, J. W. Bess, L. O. Arthur, D. K. Bryant, and C. Fenselau. 1992. Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. *J. Virol.* **66**:1856-1865.
- Henderson, L. E., T. D. Copeland, R. C. Sowder, A. M. Schultz, and S. Oroszlan. 1988. Analysis of proteins and peptides from sucrose banded HTLV-III, p. 135-147. *In* D. Bolognesi (ed.), *Human retroviruses, cancer and AIDS: approaches to prevention and therapy*. Alan R. Liss, Inc., New York.
- Henderson, L. E., R. Sowder, T. D. Copeland, G. Smythers, and S. Oroszlan. 1984. Quantitative separation of murine leukemia virus proteins by reversed-phase high-pressure liquid chromatography reveals newly described *gag* and *env* cleavage products. *J. Virol.* **52**:492-500.
- Henderson, L. E., R. C. Sowder, G. W. Smythers, and S. Oroszlan. 1987. Chemical and immunological characterizations of equine infectious anemia virus *gag*-encoded proteins. *J. Virol.* **61**:1116-1124.
- Ikuta, K., and R. B. Luftig. 1987. Differences in the pI heterogeneity of virion and intracellular Moloney murine leukemia virus p30s. *J. Gen. Virol.* **68**:487-498.
- 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.
- 21a. Kovari, L., and M. Rossmann. Personal communication.
- Kräusslich, H.-G., M. Fäcke, A.-M. Heuser, J. Konvalinka, and H. Zentgraf. 1995. The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *J. Virol.* **69**:3407-3419.
- Mervis, R. J., N. Ahmad, E. P. Lillehoj, M. G. Raum, F. H. R. Salazar, H. W. Chan, and S. Venkatesan. 1988. The *gag* gene products of human immunodeficiency virus type 1: alignment within the *gag* open reading frame, iden-

- tification of posttranslational modifications, and evidence for alternative *gag* precursors. *J. Virol.* **62**:3993–4002.
24. **Oroszlan, S., L. E. Henderson, J. R. Stephenson, T. D. Copeland, C. W. Long, J. N. Ihle, and R. V. Gilden.** 1978. Amino- and carboxyl-terminal amino acid sequences of proteins coded by *gag* gene of murine leukemia virus. *Proc. Natl. Acad. Sci. USA* **75**:1404–1408.
 25. **Pepinsky, R. B.** 1983. PhD thesis. Cornell University, Ithaca, N.Y.
 - 25a. **Pepinsky, R. B., et al.** Unpublished data.
 26. **Pepinsky, R. B., R. J. Mattaliano, and V. M. Vogt.** 1986. Structure and processing of the p2 region of avian sarcoma and leukemia virus *gag* precursor polyproteins. *J. Virol.* **58**:50–58.
 27. **Pepinsky, R. B., and V. M. Vogt.** 1984. Fine-structure analyses of lipid-protein and protein-protein interactions of *gag* protein p19 of the avian sarcoma and leukemia viruses by cyanogen bromide mapping. *J. Virol.* **52**:145–153.
 - 27a. **Pepinsky, R. B., and V. M. Vogt.** Unpublished data.
 28. **Pettit, S. C., M. D. Moody, R. S. Wehbie, A. H. Kaplan, P. V. Nantermet, C. A. Klein, and R. Swanstrom.** 1994. The p2 domain of human immunodeficiency virus type 1 *gag* regulates sequential proteolytic processing and is required to produce fully infectious virions. *J. Virol.* **68**:8017–8027.
 29. **Prongay, A. J., T. J. Smith, M. G. Rossmann, L. S. Ehrlich, C. A. Carter, and J. McClure.** 1990. Preparation and crystallization of a human immunodeficiency virus p24-Fab complex. *Proc. Natl. Acad. Sci. USA* **87**:9980–9984.
 30. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. *Cell* **32**:853–869.
 31. **Shaikh, R., M. Linial, J. Coffin, and R. Eisenman.** 1978. Recombinant avian oncoviruses. I. Alterations in the precursor to the internal structural proteins. *Virology* **87**:326–338.
 32. **Smith, R. D., J. A. Loo, R. R. Ogorzalek-Loo, M. Busman, and H. R. Udseth.** 1991. Principles and practice of electrospray ionization-mass spectrometry for large polypeptides and proteins. *Mass Spectrom. Rev.* **10**:359–451.
 33. **Stromberg, K., N. E. Hurley, N. L. Davis, R. R. Rueckert, and E. Fleissner.** 1974. Structural studies of avian myeloblastosis virus: comparison of polypeptides in virion and core component by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Virol.* **13**:513–528.
 34. **Tobin, G. J., R. C. Sowder, D. Fabris, M. Y. Hu, J. K. Battles, C. Fenselau, L. E. Henderson, and M. A. Gonda.** 1994. Amino acid sequence analysis of the proteolytic cleavage products of the bovine immunodeficiency virus Gag precursor polypeptide. *J. Virol.* **68**:7620–7627.
 35. **Tritch, R. J., Y. E. Cheng, F. H. Yin, and S. Erickson-Viitanen.** 1991. Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 Gag polyprotein. *J. Virol.* **65**:922–930.
 36. **Weldon, R. A., Jr., C. R. Erdie, M. G. Oliver, and J. W. Wills.** 1990. Incorporation of chimeric Gag protein into retroviral particles. *J. Virol.* **64**:4169–4179.
 37. **Wills, J. W., C. E. Cameron, C. B. Wilson, Y. Xiang, R. P. Bennett, and J. Leis.** 1994. An assembly domain of the Rous sarcoma virus Gag protein required late in budding. *J. Virol.* **68**:6605–6618.
 38. **Wills, J. W., R. C. Craven, and J. A. Achacoso.** 1989. Creation and expression of myristylated forms of Rous sarcoma virus Gag protein in mammalian cells. *J. Virol.* **63**:4331–4343.
 39. **Wray, W., T. Boulikas, V. P. Wray, and R. Hancock.** 1981. Silver staining of proteins in polyacrylamide gels. *Anal. Biochem.* **118**:187–203.