Processing of Avian Retroviral *gag* Polyprotein Precursors Is Blocked by a Mutation at the NC-PR Cleavage Site

HAIM BURSTEIN,¹ DIANE BIZUB,¹ MOSHE KOTLER,² GISELA SCHATZ,³ VOLKER M. VOGT,³ AND ANNA MARIE SKALKA^{1*}

The Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111¹; Department of Molecular Genetics, The Hebrew University, Hadassah Medical School, Jerusalem, Israel²; and Section of Biochemistry, Molecular and Cell Biology, Biotechnology Building, Cornell University, Ithaca, New York 14853³

Received 17 October 1991/Accepted 2 December 1991

The avian sarcoma and leukosis viruses (ASLV) encode a protease (PR) at the C terminus of gag which in vivo catalyzes the processing of both gag and gag-pol precursors. The studies reported here were undertaken to determine whether PR is able to cleave these polyproteins while it is still part of the gag precursor or whether the release of its N terminus to form free PR is necessary for full proteolytic activity. To address this question, we created a mutation that disrupts the PR cleavage site between the NC and PR coding regions of the gag gene. This mutation was introduced into a eukaryotic vector that expresses only the gag precursor and into an otherwise infectious clone of ASLV that carries the neo gene as a selectable marker. These constructs were expressed in monkey COS cells or in quail QT35 cells, respectively. Processing was impaired in both systems. Mutant particles were formed, but they contained no mature processed gag proteins. We observed only the uncleaved gag precursor polypeptide Pr76 in one case or Pr76 and a cleaved product of about 60 kDa in the other. Processing of the mutant gag precursor could be complemented in *trans* by PR from a wild-type construct, suggesting that the mutation did not induce gross structural alterations in its precursor. Our results suggest that the PR first must be released from its precursor before it can attack other sites in the gag and gag-pol polyproteins and that cleavage at the NC-PR boundary is a prerequisite for the initiation of the PR-directed processing.

The avian sarcoma and leukosis viruses (ASLV) encode a protease (PR) which is initially synthesized as part of both gag and gag-pol polyprotein precursors. Once synthesized, these polyproteins are transported to the plasma membrane and assemble into immature viral cores. Proteolytic processing of these precursors normally does not occur until the viral particles are enclosed in a plasma membrane envelope and bud from the host cell. The polyproteins are then cleaved by PR to produce the mature core components and virion-associated enzymes (for recent reviews, see references 5, 12, and 16). Recent studies have shown that the PR encoded at the C terminus of gag catalyzes the processing of both gag and gag-pol precursors during particle formation (1, 4, 11, 14); the PR embedded in the gag-pol precursor has little or no proteolytic activity (4, 14). It was not known, however, whether a wild-type PR can initiate the processing reaction while it is still part of the gag precursor or whether it must first be released. This report describes experiments designed to address this question.

A mutation at the C terminus of NC blocks processing of an NC-PR precursor fragment expressed in *Escherichia coli*. We have shown previously that an NC-PR fragment corresponding to the C-terminal portion of the Rous sarcoma virus (RSV) gag gene is partially processed when expressed in *E. coli*, giving rise to PR and NC products that appear to be identical to the virion proteins (7). Site-directed mutagenesis (10) of this bacterial expression vector (pNC-PR) was used to alter the sequence that comprises the PR cleavage site between the NC and PR coding regions (Fig. 1). The effect of these mutations was monitored conveniently by immunoblot analysis of bacterial lysates. Figure 2 shows results from

analysis of one of the mutations (cs22), in which the two

Proteolytic processing of the RSV gag precursor is blocked by the NC-PR cleavage site mutation. To determine whether the cs22 mutant gag protein also is refractory to cleavage, we first utilized a eukaryotic experimental system devised by Wills et al. (17). In this system, a myristoylated form of the RSV gag gene product ($Pr76^{myrl}$), whose amino-terminal 10 amino acids are derived from the product of the src oncogene ($p60^{v-src}$), is expressed by a simian virus 40-based

amino acids Ala and Val, normally located at positions P3 and P2 of the NC-PR recognition site, were removed and the Ser residue at position P1 was replaced by Ile (see schematic diagram in Fig. 1). None of these changes affects the PR coding region directly. E. coli cells transformed with either the wild type (pNC-PR) or the mutant plasmid construct (pNCcs22-PR) were grown to mid-log phase, and expression of the NC-PR fragments was induced by growth at 42°C as described previously (7). Cells were harvested 3 to 4 h postinduction, and the proteins in lysates were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). Immunoblot analysis showed that both constructs produced the expected NC-PR precursor of ca. 23 kDa, which reacted with both anti-PR (Fig. 2A, lanes 3 and 5, respectively) and anti-NC antisera (Fig. 2B, lanes 4 and 6, respectively). However, whereas the wild-type NC-PR precursor was partially processed to mature PR (Fig. 2A, lane 3) and NC proteins (Fig. 2B, lane 4), no processed products were observed with the cleavage-site (cs22) mutant (Fig. 2A and B, lanes 5 and 6, respectively). As has been shown previously (7), processing of the NC-PR polyprotein in these bacteria is directed by the PR; a derivative of the wild-type construct containing a mutation in the active site of PR [pNC-PR(D37S)] failed to show processed products (Fig. 2A and B, lanes 4 and 5, respectively).

^{*} Corresponding author.

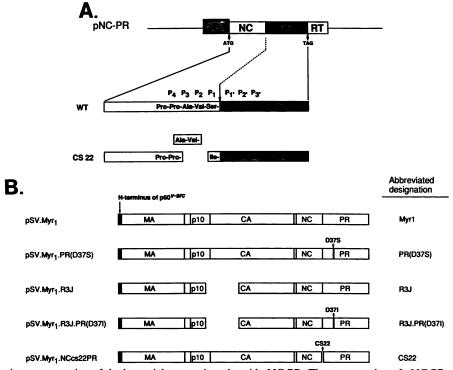


FIG. 1. (A) Schematic representation of the bacterial expression plasmid pNC-PR. The construction of pNC-PR and detailed analysis of the expression and proteolytic processing of the NC-PR polypeptide in *E. coli* have been described previously (7). Shown are the amino acid residues located at the NC-PR cleavage site region of the wild-type NC-PR fragment (WT). The amino acids Ala and Val, which were deleted from positions P3 and P2, respectively, and the Ile residue, which replaced the original Ser residue at position P1, are shown at the bottom (*cs22*). (B) Schematic diagrams of the RSV *gag* derivatives. All constructs were derived from pSV.Myr₁ (17). The hatched box at the amino terminus of each polypeptide represents the first 10 amino acids of $p60^{v-src}$. The five mature *gag* proteins are indicated.

vector (pSV.Myr₁) in COS-1 cells. The Pr76^{myrl} is produced in high levels in these cells; viruslike particles are assembled and released into the medium, and then their core protein precursors are processed in a manner indistinguishable from that of the authentic polyprotein (Pr76^{gag}) in virions produced from RSV-infected avian cells (17). The NC-PR cleavage-site mutation was introduced into the gag coding region by substituting a DNA fragment containing the mutation from pNCcs22-PR for the corresponding DNA fragment in $pSV.Myr_1$. The resulting construct, pSV.Myr₁.NCcs22-PR, as well as its parent, pSV.Myr₁, was introduced into COS-1 cells by transfection, and 48 h later the cells were exposed to [³⁵S]methionine for 2.5 h. Anti-RSV antiserum was used to immunoprecipitate gag-specific proteins from cell lysates (Fig. 3A) or from detergentdisrupted viruslike particles that had been harvested from the medium (Fig. 3B). As shown in Fig. 3B, both pSV.Myr₁ and pSV.Myr₁.NCcs22-PR produced gag proteins that appeared in the medium as components of viruslike particles (lanes 1 and 5, respectively) (17). However, whereas the wild-type gag was almost completely processed to mature capsid proteins (Fig. 3A and B, lanes 1), processing of the cs22 mutant gag precursor was severely impaired (Fig. 3A and B, lanes 5). An active site mutation (D37S) in the PR coding region also abolished processing (Fig. 3A and B, lanes 2; see also references 2 and 8), verifying that cleavage of the gag precursor in these particles is directed by the gag-encoded PR. These results suggest that the mutation which disrupts the NC-PR cleavage site also blocks processing at other sites in the gag precursor.

A second experimental system also was used to test the effects of altering the NC-PR cleavage site on polyprotein processing. The cs22 mutation was built into the BH-pol version of the RCAS neo vector (6), which encodes an infectious ASLV provirus that carries the selectable neo gene in place of src. This DNA was introduced into cells of the quail line QT35 by the transfection method of Chen and Okayama (3). Cells stably expressing the drug resistance marker were selected by growth in the presence of G418 for at least 1 month (13). Mass-selected cells and clones derived from them exhibited the same characteristics. Virus shed into the medium by these cells was collected by centrifugation through a 15% cushion of sucrose in 10 mM Tris-HCl (pH 7.5)-100 mM NaCl-1 mM EDTA containing 1 mM of the protease inhibitor phenylmethylsulfonyl fluoride. The virus pellet was dissolved immediately in SDS-containing sample buffer and analyzed by immunoblotting after SDS-PAGE. Rabbit antisera to purified virion proteins were used to detect gag products on the immunoblots. These antisera recognize the major gag proteins CA and NC. Two examples of analysis of the gag proteins incorporated into the cs22 virions are shown in Fig. 4. In the experiment shown in panel A, virus shed in a 7-h growth period was analyzed. The gag protein in the cs22 virions (lane 2), as well as in virions carrying a mutation (D37I; 7, 13) in the active site of the protease (lane 1), migrated predominantly as uncleaved Pr76. No mature CA or NC was observed. In contrast, virions carrying no mutation (lane 3 and 4) contained proteins that migrated indistinguishably from purified CA and NC.

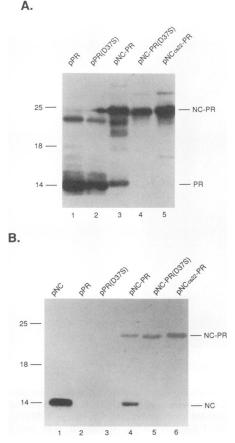


FIG. 2. Immunoblot analysis of NC-PR proteins and their processed products expressed in *E. coli*. Cell lysates were fractionated on an SDS-15% polyacrylamide gel, and virus-related proteins were detected with rabbit antiserum directed against ASLV PR (A) or NC (B) and ¹²⁵I-labelled protein G as described previously (7). Lysates from cells expressing NC (pNC), PR (pPR), and PR (D37S) [pPR(D37S)] were included as markers. The nature of the minor upper band in the pPR and pPR(D37S) lanes is unknown but may represent PR dimers. Molecular size markers in kilodaltons are indicated on the left. The positions of NC-PR, PR, and NC are indicated on the right.

In the independent experiment shown in panel B, a portion of the virus collected from cells producing D37I virions (lane 8) or cs22 virions (lane 10) was resuspended in protease buffer (1 M NaCl, 70 mM Tris acetate [pH 6.2], 0.7 mM EDTA, 5.3% glycerol, 1 mM dithiothreitol) in the presence of detergent (0.1% Triton X-100) and incubated for 5 h at 37°C. This incubation was intended to test the ability of immature cores of cs22 virions to undergo maturation spontaneously. The results (lanes 9 and 11) showed no evidence of residual PR activity in either preparation. In addition to Pr76, a gag-related polypeptide of molecular weight ca. 60 kDa was also detected in cs22 virions. This presumed proteolytic fragment was found in variable amounts in repeated experiments with cs22 but was never detected in PR(D37I)-containing virions. In no cases were the mature CA or NC proteins detected in either of these virions. Oertle and Spahr (11) observed a very similar phenotype for a mutant gag protein in which the N-terminal leucine residue of the PR domain was changed to arginine, thus presumably blocking cleavage at this site. When expressed in chicken

cells, this mutant failed to yield any mature gag proteins, giving rise to particles containing Pr76 plus a polypeptide of ca. 63 kDa that comigrated with a gag protein terminating exactly at the NC-PR boundary. In our hands, the 60-kDa species from cs22 also comigrated with the product of a termination mutant at this boundary (G. Schatz and V. M. Vogt, unpublished). We have not yet established whether (i) the 60-kDa polypeptide fragment is a product of PR cleavage, as is suggested by the absence of this band in D37I virions (and absence of a similar band in Oertle and Spahr's D37R PR substitution mutant, p15-1; 11), (ii) it is due to the action of a cellular protease, or (iii) it is attributable to some other process. In infections with wild-type virus in chicken cells, in addition to Pr76, two prominent intracellular gag polypeptides of molecular sizes ca. 66 kDa (Pr66) and 60 kDa (Pr60) commonly are found, both of which lack the tryptic peptides of PR (15). We have not yet determined whether either of these could be identical to the ca. 60-kDa species observed in the cs22 mutant. The 60-kDa product could not be detected in cs22 viruslike particles that were released from [³⁵S]methionine pulse-labelled COS-1 cells, even after a chase of as long as 8 h (data not shown). Since the cs22virions released from the QT35 cell lines also contain the gag-pol polyproteins, it is conceivable that the 60-kDa polypeptide resulted from proteolytic activity of the PR embedded in Pr180^{gag-pol}. However, experiments to date suggest that this embedded PR domain is devoid of activity (4, 14). Alternatively, it might be the result of a cellular protease that is more abundant in avian cells than in monkey cells.

The cs22-containing gag precursors can be processed in trans. Although it seemed unlikely, we wished to rule out the possibility that the cs22 mutation affected the overall folding of the Pr76 polypeptide in such a way that the cleavage sites at the boundaries of the mature gag proteins were unavailable to the PR. To test this, we used the eukaryotic expression system to perform a trans-complementation assay. COS-1 cells were cotransfected with pSV.Myr₁.NCcs22-PR and a derivative of pSV.Myr₁ (pSV.Myr₁.R3J; gift of J. W. Wills) that contains an internal deletion in the gag coding region. The gag polyprotein encoded by pSV.Myr₁.R3J lacks amino acid residues 192 to 364 which include 47 amino acids from the C terminus of p10 and 125 amino acids from the N terminus of CA. This large deletion does not disrupt particle formation or PR activity (15a). However, since the PR cleavage site between p10 and CA has been removed, processing of this gag polyprotein does not give rise to the normal CA product but to a fused protein of ca. 34 kDa (Fig. 3A, lane 3). The cs22 and R3J precursor polyproteins are of different lengths and are clearly distinguishable in the cell lysate (Fig. 3A, compare lanes 3 and 5). When the cs22 mutant and R3J gag proteins were coexpressed, the two uncleaved precursor molecules could be detected in the cell lysate (Fig. 3A, lane 6), confirming that both gag precursors were produced. In addition, a CA protein band was evident in both the cell lysate (Fig. 3A, lane 6) and in the medium (Fig. 3B, lane 6) which could have originated only from the cs22-containing gag precursor. From this result, we conclude that processing must have occurred in trans.

To confirm that the cs22 precursor was processed by the R3J-encoded PR, we performed a complementation assay with a derivative of R3J that contains a mutation (D37I) in the active site of its PR. Expression of this construct [pSV.Myr₁.R3J.PR(D37I)] alone in COS-1 cells produced *gag* precursors that formed particles, but the polyproteins were not cleaved (Fig. 3A and B, lanes 4). Coexpression of

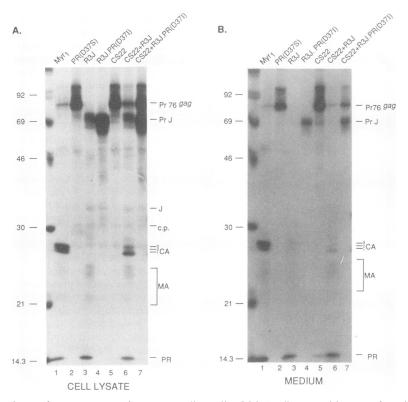


FIG. 3. Expression and release of *gag* precursors from mammalian cells. COS-1 cells were either transfected with single plasmid DNAs or cotransfected with two plasmid DNAs as indicated. Immunoprecipitated proteins were separated on SDS-12% polyacrylamide gels, and radioactive bands were detected by fluorography. (A) Cell lysate. (B) Medium-derived particles. Lanes: Myr₁, pSV.Myr₁; PR(D37S), pSV.Myr₁.PR(D37S); R3J, pSV.Myr₁.R3J; R3J.PR(D37I), pSV.Myr₁.R3J.PR(D37I); CS22, pSV.Myr₁.NCcs22PR. The positions of the molecular size markers in kilodaltons are indicated on the left. The positions of the precursor polyproteins Pr76^{seag} and PrR3J (PrJ) and their cleavage products are indicated on the right. The CA appears as a mixture of as many as three species, denoted CA1, CA2, and CA3 (1). The heterogeneous MA proteins are marked by a large bracket. J denotes a unique ca. 34-kDa processed protein characteristic of R3J; c.p. indicates cell proteins which appear as a background only in the cell lysate fraction.

pSV.Myr₁.R3J.PR(D37I) with pSV.Myr₁.NCcs22-PR gave rise to particles that contained uncleaved gag precursors of lengths characteristic of both sources (Fig. 3A and B, compare lanes 4, 5, and 7). Since the R3J.PR(D37I) precursor was not cleaved by the cs22-containing gag-encoded PR in this cotransfection, we conclude that the PR in the cs22-containing gag precursor is also inactive in trans. In agreement with this result, coexpression of the cs22 mutant and a PR active site mutant in the quail cell system yielded virus particles that also completely lacked mature gag proteins (data not shown). The results of these cotransfection assays show that the defect in the proteolytic processing of the cs22 mutant gag precursor is not due to gross alteration in its overall conformation. However, we cannot exclude the possibility that local conformational changes at the NC-PR border prevent proper folding of the PR domain or somehow interfere with the dimerization of subunits, which is required for PR activity. These mechanisms seem unlikely because other mutations at this site (preliminary data) appear to give the same phenotype. Further biochemical analysis of the properties of cs22 and additional cleavage site mutations may help to resolve this question.

The simplest interpretation of the results presented here is that the PR domain is unable to act as an efficient protease, unless it is first released from the polyprotein precursor in which it is embedded. How is release of the PR domain initiated? According to one model, PR itself is responsible

for this first proteolytic event. The observation that active site mutations in PR appear to completely block processing supports this model. One possibility is that the first proteolytic event can only occur intramolecularly, in cis. This would be consistent with our observation that cs22 cannot act in trans. Another possibility is that the PR domain in the gag precursor has a very low level of proteolytic activity that acts preferentially at the NC-PR cleavage site. It is known that peptide substrates containing this amino acid sequence are more readily cleaved by PR in vitro than any of the other peptides mimicking gag or pol cleavage sites. Also, topological constraints in the immature virions could prevent the PR domain from gaining access to sites other than that at the NC-PR boundary. With either mechanism, once enough PR is released, a cascade should ensue, leading to the production of more PR and eventually to the efficient cleavage of the gag and gag-pol polyprotein precursors. An alternative and more complicated model postulates that a cellular protease is responsible for initially releasing the PR domain in order to activate it, thus leading to the same cascade. The model does not require that this hypothetical protease cleave exactly at the NC-PR junction, but only that it cleave near enough to the junction to generate a PR-containing polypeptide that can fold and dimerize into an enzymatically active protease. This model is based on a number of ad hoc assumptions about a hypothetical cellular enzyme that is

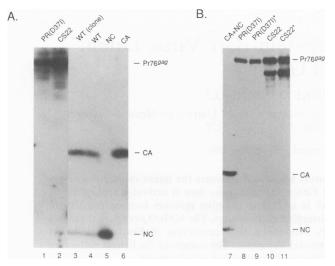


FIG. 4. Expression and release of gag precursors from avian cells. QT35 quail cells were transfected with wild-type, proteasemutant, or cleavage-site-mutant viral DNAs carrying the neomycin gene as a selectable marker. Virus particles were collected from G418 mass-selected cells or from selected and cloned cells (panel A, lane 3) as described in the text. (A) Virus pellets were dissolved in SDS-containing sample buffer. (B) Virus pellets were dissolved in protease digestion buffer in the presence of detergent (see text) and then were diluted immediately with SDS sample buffer (lanes 8 and 10), or they were incubated for 5 h at 37°C and then diluted (lanes 9 and 11; marked by an asterisk). The viral proteins were analyzed by SDS-PAGE and immunoblotting from a 15% polyacrylamide gel, with a mixture of rabbit antisera directed against the viral proteins CA and NC. Purified NC (lane 5) and CA (lane 6) or a mixture of purified NC and CA (lane 7) was used as a molecular weight marker. The positions of the precursor protein Pr76^{gag} and its cleavage products are indicated on the right.

packaged into virions and whose activity is triggered by release of the virions from the cell.

The applicability of our findings for cs22 to other retroviral systems remains to be ascertained. However, it may be relevant that in human immunodeficiency virus, fusion of PR to other viral or nonviral sequences can lead to drastic inhibition of activity in vitro and in *E. coli* (9). Enzymatic activity can be recovered after treatments that lead to polypeptide unfolding and refolding, suggesting that there may be alternative stable conformations of the PR domain that preclude enzymatic activity. Further experiments will be required to determine why the PR domain in the cs22mutant of ASLV is inactive and how the proteolytic cascade that underlies maturation in this and all other known retroviruses is initiated.

We thank John W. Wills for his generous gift of plasmids and antisera and for advice and suggestions during the course of these studies. Pat Roat provided excellent technical assistance.

This work was supported by National Institutes of Health grants CA20087 to V.M.V. and CA47486, CA-06927, and RR-05539 to

A.M.S. and by a grant from the Pew Charitable Trust and also by an appropriation from the Commonwealth of Pennsylvania.

REFERENCES

- 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.
- 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.
- Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. 7:2745-2752.
- Craven, R. C., R. P. Bennett, and J. W. Wills. 1991. Role of the avian retroviral protease in the activation of reverse transcriptase during virion assembly. J. Virol. 65:6205–6217.
- Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich. 1984. Protein biosynthesis and assembly, p. 513–648. In R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Hughes, S. H., J. J. Greenhouse, C. J. Petropoulos, and P. Sutrave. 1987. Adaptor plasmids simplify the insertion of foreign DNA into helper-independent retroviral vectors. J. Virol. 61: 3004-3012.
- Kotler, M., R. A. Katz, and A. M. Skalka. 1988. Activity of avian retroviral protease expressed in *Escherichia coli*. J. Virol. 62:2696–2700.
- Leis, J., D. Bizub, I. T. Weber, C. Cameron, R. Katz, A. Wlodawer, and A. M. Skalka. 1989. Structure function-analysis of retroviral aspartic protease, p. 235–243. *In* H. Kräusslich, S. Oroszlan, and E. Wimmer (ed.), Current communications in molecular biology: viral proteinases as targets for chemotherapy. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Louis, J. M., R. A. McDonald, N. T. Nashed, E. M. Wondrak, D. M. Jerina, S. Oroszlan, and P. T. Mora. 1991. Autoprocessing of HIV-1 protease using purified wild type and mutated fusion proteins expressed at high levels in *E. coli*. Eur. J. Biochem. 199:361-369.
- Morinaga, Y., T. Franceschini, S. Inouye, and M. Inouye. 1984. Improvement of oligonucleotide-directed site specific mutagenesis using double-stranded plasmid DNA. Bio/Technology 2:636-639.
- 11. 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.
- Skalka, A. M. 1989. Retroviral proteases: first glimpses at the anatomy of a processing machine. Cell 56:911-913.
- Stewart, L., G. Schatz, and V. M. Vogt. 1990. Properties of avian retroviral particles defective in viral protease. J. Virol. 64:5076-5092.
- 14. 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.
- 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.
- 15a.Wills, J. W. Personal communication.
- 16. Wills, J. W., and R. C. Craven. 1991. Form, function, and use of retroviral *Gag* proteins. AIDS 5:639–654.
- Wills, J. W., R. C. Craven, and A. Achacoso. 1989. Creation and expression of myristylated forms of Rous sarcoma virus gag protein in mammalian cells. J. Virol. 63:4331-4343.