Mutations within the Proteolytic Cleavage Site of the Rous Sarcoma Virus Glycoprotein That Block Processing to gp85 and gp37

LAUTARO G. PEREZ AND ERIC HUNTER*

Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294

Received 8 October 1986/Accepted 5 February 1987

We have investigated the specificity of the proteolytic cleavage of the Rous sarcoma virus glycoprotein precursor by introducing two mutations into the putative cleavage region (Arg-Arg-Lys-Arg). We show that neither a deletion of the cleavage sequence nor a glutamic acid for lysine substitution altered intracellular transport or surface expression of the env gene products. However, both the four-amino-acid deletion and the glutamic acid substitution block processing of the env precursor. Susceptibility of the glutamic acid-substituted env precursor to proteases indicated that tertiary protein structure was unaffected. While inhibitor experiments suggested that more than one endopeptidase might be capable of mediating the proteolytic cleavage, the results presented here point to the presence in the Golgi apparatus of a novel endopeptidase, required for retroviral glycoprotein cleavage, that has a high specificity for lysine-containing peptides.

The primary product of the Rous sarcoma virus (RSV) envelope gene (env) is a precursor protein of 95 kilodaltons (Pr95env). During biosynthesis, this precursor is proteolytically cleaved to yield the two glycoproteins gp85 and gp37, which remain associated by disulfide bonds as a complex on the viral membrane (3, 5, 6, 12, 17, 21, 26, 27). In Pr95env the cleavage site occurs after a sequence of four basic amino acids, arginine-arginine-lysine-arginine, which resembles the cleavage site of many hormone precursors and other retrovirus env precursors (9, 11, 14, 16, 28, 32, 36). In previous studies with radioactively labeled sugars and a simian virus 40 (SV40) expression vector, we found that the cleavage of Pr95env occurred in the absence of the remainder of the viral proteins in a late Golgi compartment in CV-1 cells (39), in agreement with previous studies of RSV-infected avian cells (21).

The processing of precursor polypeptides to produce biologically active proteins occurs even in unicellular eucaryotes, and at least two eucaryotic endoproteasomes, with specificity for a pair of basic amino acids (lysine-arginine or arginine-lysine) similar to those present in the cleavage site of the RSV env precursor, have been identified in Golgi compartments (8, 15). We have, therefore, investigated whether a protease with similar biological specificity is involved in the processing of the Pr95env. Using synthetic oligonucleotides and a subclone of the RSV env gene in the bacteriophage M13, we have constructed two mutants. In one the codon for lysine in the cleavage site was changed to a glutamic acid, and in the second the 12 nucleotides encoding the four basic amino acids of this region were deleted. The resulting mutant env genes have been inserted into an SV40 late-replacement vector for analysis of expression and transport. We have found that neither of these mutations altered biosynthesis, intracellular transport, or surface expression of the RSV glycoproteins in CV-1 cells. The deletion of the cleavage site results in a precursor protein that is not cleavable either intracellularly or by addition of extracellular proteases. The replacement of lysine for glutamic acid results in greater than 90% reduction in intracellular proteolytic processing of the env precursor, even though the uncleaved precursor is highly susceptible to cleavage by extracellular proteases. Our results show that the lysine present in the cleavage site of the env precursor is essential for the intracellular processing by the Golgi endopeptidase. This enzyme, unlike those described previously, appears to be specific for lysine-containing basic dipeptides and unable to cleave those containing only arginine.

MATERIALS AND METHODS

DNAs, viruses, and cells. The subcloning of the RSV env gene in the bacteriophage M13, the propagation of the clones, and the SV40 expression system for the glycoproteins has been described elsewhere (39; G. L. Davis and E. Hunter, J. Cell Biol., in press). African green monkey kidney cells (CV-1) were grown in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum. All the DNA plasmids were propagated in the DH-1 strain of Escherichia coli. Oligonucleotide site-directed mutagenesis and expression of the mutant genes in CV-1 cells from the late promoter of SV40. The procedure used for the mutagenesis of the cleavage region was essentially that described by Zoller and Smith (40), with the following modifications. Purification of the full-length double-stranded DNA from the single-stranded template was not made; instead, dilutions of the priming reaction were directly used to transform the CaCl2-treated JM 101 strain of E. coli. The plaques obtained were transferred to nitrocellulose filters, and the bound phage DNAs were hybridized to the mutagenic primer labeled with 32P. The phage of those plaques that remained as strong signals after the filters were washed at 60°C were sequenced by the chain-terminator method of Sanger et al. (33, 34). The EcoRI-XbaI restriction fragment containing the 3' end of the RSV env gene was recovered from the replicative form of mutant phage and substituted for the wild-type fragment in the SV40 expression vector. CV-1 cells were transfected with the mutant RSV env genes and a helper SV40 DNA as described previously (39). The viral stocks obtained were used to infect CV-1 cells and monitor the expression of the RSV env mutant glycoproteins.

Immunofluorescence staining of the infected cells, radio-labeling, immunoprecipitation, and sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Rabbit antibodies against
gp85 and gp37 were used for staining of fixed cells and for immunoprecipitation of the $^3$H-labeled proteins in the infected cells, as described previously (14, 39). For immunofluorescence staining of unfixed cells, a chicken antisera of RSV was used.

**Treatment of infected cells with protease inhibitors.** The lysosomal protease inhibitors chloroquine and leupeptin, and soybean trypsin inhibitor (SBTI) were added to the culture medium of infected cells in some experiments. Chloroquine and leupeptin were present at 100 and 20 $\mu$M final concentrations, respectively (1, 10, 22, 29). SBTI was added at a 100-$\mu$g/ml final concentration. The inhibitors were present in the cell medium 2 h before the $^3$H-leucine pulse-labeling and maintained during the chase period (Davis and Hunter, in press).

**External protease treatment of infected cells.** At 48 h postinfection, CV-1 cells in 35-mm plates were pulsed for 15 min with $^3$H-leucine and then chased for 3 h in complete medium to allow the accumulation of labeled proteins on the cell surface. After this period, cell monolayers were washed three times with serum-free medium, before the addition of 5, 10, or 20 $\mu$g of trypsin (Worthington Diagnostics, Freehold, N.J.) per ml, in serum-free medium at pH 9. The cells were incubated in the presence of trypsin at 4°C for 1 h. SBTI (200 $\mu$g/ml) was added to stop the digestions, and then cells were lysed and immunoprecipitated as described above.

**RESULTS**

**Site-directed mutagenesis of the cleavage region in Pr95env.** During biosynthesis of the RSV env gene products, the precursor polyepitide Pr95env is proteolytically cleaved to gp85 and gp37. Although the carboxy-terminal sequence of gp85 has not been determined experimentally, aminoterminal sequencing of gp37 located the cleavage site after a sequence of four basic amino acids: arginine-arginine-lysine-arginine (14; Fig. 1A). To modify the cleavage region of Pr95env two synthetic oligonucleotides were designed. A 20-mer was used to change the first A (nucleotide 6264) of the codon for lysine (AAA) to a G, creating a codon for glutamic acid (GAA). According to protein structure predictions, this change does not alter either the protein conformation (a turn), or the hydrophilicity of this region. This mutant was designated G1 (Fig. 1A). A 23-mer was used to delete the 12 nucleotides that encode the four basic amino acids adjacent to the cleavage site. This mutant was designated Dr1 (Fig. 1A). Figure 1B shows the DNA sequence of the mutated and wild-type regions. The env gene EcoRI-XbaI restriction fragments containing the modified cleavage region were substituted for the wild-type fragment in the SV40 env expression vector, as described elsewhere (Davis and Hunter, in press).

**Intracellular and surface expression of the modified env glycoprotein complexes in CV-1 cells.** The recombinant SV40 vector DNAs carrying the modified RSV env genes were cotransfected into CV-1 cells with helper SV40 DNA (dl1055) to obtain viral stocks. Expression was initially measured by indirect immunofluorescence staining of CV-1 monolayers infected with these viral stocks. Figure 2 (left) shows the pattern of intracellular fluorescence displayed by cells infected with SV40 vectors containing the wild-type and modified env genes. A similar fluorescence pattern was observed in both mutant and wild-type infected cells; bright reticulated cytoplasmic staining together with a clear ring around nucleus defined the rough endoplasmic reticulum, and the bright area adjacent to the nucleus defined the Golgi complex (39). These results indicate that the modifications to the cleavage site do not alter the expression and intracellular transport of the RSV env complex to the Golgi apparatus. Surface immunofluorescence staining of unfixed infected cells with chicken anti-RSV Prague C antisera showed immunofluorescence of equal intensity to that of the wild type (Fig. 2, right), proving that the modified viral glycoproteins were expressed on the cell surface in a similar manner to the wild-type viral complex. Thus, the changes created in the cleavage site of Pr95env did not affect intracellular transport or surface expression of the RSV env gene products.

**Processing of the modified Pr95env in CV-1 cells.** To compare the rate of processing of the Pr95env proteins carrying changes in the cleavage region with that of the wild-type precursor, CV-1 cells were pulse-labeled with $^3$H-leucine for 15 min and chased for periods of 0 to 8 h. The labeled
glycoproteins were immunoprecipitated with an anti-env rabbit antiserum and electrophoresed in a sodium dodecyl sulfate-polyacrylamide (10%) gel. Figure 3 shows the fluorogram of such an experiment. It can be seen that in wild-type-infected cells only the precursor protein Pr95\textsuperscript{env} was labeled after a 15-min pulse. This was processed to the final products (gp85 and gp37) during the chase periods (Fig. 3, left). An identically sized Pr95\textsuperscript{env} precursor was labeled during the 15-min pulse of mutant G1-infected cells, but in contrast to the wild type this precursor was processed to a higher-molecular-weight product during the chase (Fig. 3, center). A minor fraction of the G1 mutant protein was processed to gp85 and gp37, but most of it remained uncleaved, as the slower migrating form. The higher-molecular-weight form of the env gene product (approximately 120,000) would be expected if the unprocessed precursor were terminally glycosylated to yield a molecule equivalent in size to the sum of the terminally glycosylated gp85 and gp37 molecules. Indeed, unlike Pr95\textsuperscript{env}, the 120-kDa product can be labeled with [\textsuperscript{3}H]fucose, a terminal sugar that is normally only added to the env gene products after cleavage (39; data not shown). This result, therefore, constitutes the first evidence that lysine plays a critical role in the recognition of the cleavage site by the cellular endopeptidase that processes Pr95\textsuperscript{env}.

An analysis of the processing of the Dr1 mutant Pr95\textsuperscript{env}, in which the cleavage site is deleted, revealed that this precursor was similarly processed to the terminally glycosylated, slowly migrating form observed with mutant G1. However, in contrast to the latter, cleavage was completely abolished, and no gp85 or gp37 was observed (Fig. 3, right). This result indicates that the proteolytic cleavage of Pr95\textsuperscript{env} occurs at a highly specific sequence that is defined by the four deleted amino acids.

**FIG. 3.** Comparison of the rate of synthesis and processing of the wild-type and mutant env gene products expressed from the late promoter of SV40 in CV-1 cells. The fluorogram of a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (10% polyacrylamide) used to separate the immunoprecipitated labeled proteins is shown. Left, Synthesis and processing of env wild type; center, synthesis and processing of mutant G1; right, synthesis and processing of mutant Dr1. P, 15-min pulse with [\textsuperscript{3}H]leucine; P/C, 15-min pulse with [\textsuperscript{3}H]leucine followed by a chase of 2, 4, 6, or 8 h, as shown.

**FIG. 2.** Indirect immunofluorescence of CV-1 cells infected with an SV40 vector containing wild-type, G1, and Dr1 mutant env genes. Left, Internal immunofluorescence observed when ethanol-acetic acid-fixed cells were reacted first with a rabbit antibody to gp85 and then with a fluorescein-labeled goat anti-rabbit immunoglobulin G antibody. Right, Cell-surface immunofluorescence observed when unfixed cells were reacted sequentially with chicken antibody to RSV Prague C and rabbit antibody to chicken immunoglobulin G. In both sets of immunofluorescence staining experiments, the mutant env gene-expressing cells were indistinguishable from those of the wild-type env gene.
FIG. 4. Extracellular processing of mutant G1 Pr95<sup>env</sup> after treatment with trypsin. Infected cells were pulse-labeled for 15 min with [3H]leucine and then chased in complete medium for 3 h, as described in Materials and Methods, to allow surface expression of the labeled protein. Lanes: 1 to 4, wild-type env product; 5 to 8, G1 mutant Pr95<sup>env</sup>; 9 and 10, Dr1 mutant. The cells were then overlaid with 2 ml of serum-free medium containing 0, 5, 10, or 20 μg of trypsin per ml (lanes 1 and 5; 2 and 6; 3, 7, and 9; or 4, 8, and 10, respectively. Wild-type gp85 and gp37 were completely resistant to trypsin digestion even at the highest concentration (lane 4), whereas the G1 Pr95<sup>env</sup> was cleaved at the lowest (5 μg/ml) concentration of trypsin (lane 6).

mutant G1 (Fig. 4, lane 5), on the other hand, is cleaved to gp85- and gp37-sized products even at the lowest concentration of trypsin used (Fig. 4, lanes 6 to 8). These results support the idea that the substitution of glutamic acid for lysine in the cleavage region, rather than a lack of accessibility of the enzyme to the site, is responsible for the G1 phenotype. The Dr1 mutant polypeptide, in contrast, is completely resistant to trypsin digestion at the highest concentration (20 μg/ml), supporting the concept that the cleavage site is the sole trypsin-sensitive region in the mature glycoprotein.

Residual cleavage of the mutant G1 polypeptide occurs intracellularly. Because the mutant G1 Pr95<sup>env</sup> is accessible to external proteases (as demonstrated above), we were interested in determining the nature of the residual cleavage observed with this mutant protein. Three possibilities existed: the cleavage was mediated by extracellular serum protease, the normal protease possessed a broad specificity and was cutting at the remaining basic amino acid residues, or an alternative microsomal enzyme was involved. The following experiments were therefore carried out. Infected cells were pulse-labeled with [3H]leucine and chased in the presence of chloroquine, leupeptin, and SBTI, as described in Materials and Methods. Even when all three of these protease inhibitors were present during the pulse-chase period, the wild-type Pr95<sup>env</sup> was processed normally to gp85 and gp37 (Fig. 5A), although an increase in the mobility of both gp85 and gp37 was observed. The smaller apparent molecular weight of these two polypeptides probably reflects an inhibition of some terminal glycosylation events as it is seen with the ionophore monensin (2). In contrast to the wild type, residual cleavage of the G1 polypeptide was not observed in the presence of all three protease inhibitors (Fig. 5A). No gp85/gp37 bands were visible, and the terminal glycosylated uncleaved precursor showed a similar decrease in molecular weight to the wild-type cleaved products, so that it now comigrated with the Pr95<sup>env</sup> band (Fig. 5A). In the absence of inhibitors or in the presence of leupeptin and SBTI alone, the residual cleavage was observed (Fig. 5), suggesting that its inhibition was due to chloroquine. This drug does not affect cell-surface expression of the env gene products (Davis and Hunter, in press), but because it is impermeable to biological membranes and because it is known to inhibit the protease cathepsin B (22, 38), these results suggest that the residual processing of the G1 Pr95<sup>env</sup> occurs intracellularly rather than on the cell surface. Furthermore, the fact that the wild-type Pr95<sup>env</sup> is normally processed in the presence of all the protease inhibitors suggests that a proteolytic activity distinct from that involved in wild-type cleavage is responsible for the partial processing of G1 Pr95<sup>env</sup>.

DISCUSSION

In this study, we have attempted to define the cleavage recognition sequence for the cellular protease involved in the processing of the RSV Pr95<sup>env</sup> to the viral glycoproteins gp85 and gp37. The putative cleavage site in the RSV env precursor that is determined by the sequence of four basic amino acids (arginine-arginine-lysine-arginine) was modified by oligonucleotide-directed mutagenesis. Introducing a single base change, we changed the lysine residue present at the cleavage region to a glutamic acid. An oligonucleotide that looped out the 12 nucleotides encoding the four basic amino acid residues was used to delete the region. A computer

FIG. 5. Inhibition of the residual cleavage observed with mutant G1. CV-1 cells infected with SV40 vectors carrying the wild-type and mutant env genes were pretreated with either chloroquine, leupeptin, and SBTI (A) or leupeptin and SBTI alone (B) for 2 h before pulse-labeling with [3H]leucine and during the 3-h chase. P, 15-min pulse-labeling with [3H]leucine; C, chase in the absence of inhibitors; CI, chase in the presence of inhibitors.
RSV  Thr-Gly-Ile-Arg-Arg-Lys-Arg-Ser-Val-Ser
MvLV1  Arg-Ser-Asn-Arg-His-Lys-Arg-Glu-Pro-Val
MMTV  Asn-Leu-Ile-Arg-Ala-Lys-Arg-Phe-Val-Ala
MPMV  Phe-Leu-Gly-Lys-Ala-Lys-Arg-Ala-Ile-Gln
HIV   Val-Val-Gln-Arg-Glu-Lys-Ala-Val-Gly
FeLV  Lys-Ala-Val-Arg-Phe-Arg-Glu-Pro-Ile
HTLV-I Leu-Gly-Ser-Arg-Arg-Arg-Arg-Leu-Pro
BLV   Pro-Pro-Thr-Arg-Val-Arg-Arg-Arg
FPV   Ser-Lys-Lys-Glu-Glu-Gly-Leu-Phe
Influenza2  Ile-Glu-Ser-Arg--------------------------Gly-Leu-Phe

FIG. 6. Cleavage recognition sequences for the env gene products of the major retrovirus groups and for the HA of avian and human influenza viruses. The basic tetrapeptide that constitutes the proposed cleavage recognition sequence is shown for each virus in boldface type; flanking structural protein sequences are shown in lightface type. The virus names are abbreviated as follows: RSV, Rous sarcoma virus; MvLV1 Moloney murine leukemia virus; MMTV, mouse mammary tumor virus; MPMV, Mason-Pfizer monkey virus (35); HIV, human immunodeficiency virus; FeLV, feline leukemia virus; HTLV-I, human T lymphotropic virus type I; BLV, bovine leukemia virus; FPV, fowl plague (avian orthomyxovirus, Rostock strain [30]); Influenza2, human orthomyxovirus (H2 subtype, A/Jap/307/57 [7]). The sequences shown are excerpted from Weiss et al. (37) except where indicated otherwise.

analysis of the predicted protein structure of the Pr95env indicated that the amino acid substitution did not alter the protein structure of the cleavage region of the precursor or its hydrophobicity (13, 31). On the other hand, the deletion of the four basic residues from the env precursor changed the predicted protein structure of that region from a hydrophobic turn to a more hydrophobic B sheet while conserving the rest of the protein in its native structure. The above predictions were confirmed by the expression studies that indicated that biosynthesis, intracellular transport, and surface expression of these modified precursors were not impaired. However, processing of the modified env precursors expressed under the control of the late promoter of SV40 in CV-1 cells was significantly reduced. This finding proves that the posttranslational processing of Pr95env is not an essential step in its intracellular transport.

The results of the experiments presented here highlight the importance of the amino acid lysine in signalling recognition of the cleavage site by the cellular protease. They also provide evidence for an endopeptidase, present in Golgi, with high specificity for lysine-containing peptides. Thus, the endopeptidase that processes Pr95env in CV-1 cells appears to be different from those Golgi-associated enzymes previously reported to process hormone precursor polypeptides containing arginine-lysine or lysine-arginine dipeptides (8, 15) or an enkephalin-generating trypsin-like enzyme associated with bovine chromaffin granules (23, 24), since the latter proteases are able to process peptides containing an arginine-arginine pair as well as lysine-arginine dipeptide. The requirement of the protease for a lysine residue is consistent with the fact that most retroviral glycoprotein precursors possess a lysine-arginine pair at the junction of the outer membrane protein and the membrane-spanning protein (Fig. 6). Only four retroviruses do not follow this rule, feline leukemia virus, human T-cell lymphotropic virus types I and II, and bovine leukemia virus, which possess only arginine residues in the cleavage region. Whether the nature of the target cell or the pathway of maturation of these viral glycoproteins differs from the remainder of the retroviral family remains to be determined; viruses of the orthomyxovirus family that lack the lysine-arginine dipeptide demonstrate glycoprotein cleavage in only a limited number of cell types (18-20). In the interpretation of our results, we cannot rule out the possibility that the substitution of a negative charge (glutamic acid) for a positive charge (lysine) results in an inhibitory effect on the activity of the enzyme rather than in a lack of recognition of the lysine-deficient region. However, the presence of a glutamic acid residue within the tetrapeptide cleavage region of human immunodeficiency viruses (Fig. 6) makes this unlikely.

The fact that chloroquine, an inhibitor of cathepsin B protease activity in lysosomes, is able to block the residual cleavage of the mutant G1 Pr95env but does not block the processing of the wild-type env precursor suggests that more than one endopeptidase may be involved in the processing of Pr95env, the chloroquine-resistant endopeptidase activity being the most important. Leupeptin, a protease inhibitor also known to block cathepsin B activity of lysosomes (22), was unable to block the residual processing of G1 Pr95env. However, this probably reflects a more restricted access of leupeptin to cellular compartments than that of chloroquine, which is known to enter the Golgi compartments (22) in which processing of Pr95env occurs (39).

The hemagglutinin (HA) glycoprotein of avian influenza viruses is also proteolytically processed to HA1 and HA2 in a region of basic amino acids, an event that is required for full infectivity of virions but not for virus assembly or receptor-binding activity (4, 19). The effect of the mutations described here on the assembly of infectious RSV remains to be determined. Linial et al. (25) have reported on a nonconditional replication-defective mutant of RSV (SE521), with a mutation that affects the processing of the viral glycoprotein precursor. In SE521-infected cells, Pr95env remained uncleaved on the cell surface and was not incorporated into the viral particles released. However, the location of the mutation in the env gene of the mutant SE521 was not clearly defined and could have involved more than one region. Thus, the requirement for proteolytic processing of the envelope precursor for incorporation of the glycoprotein complex into viral particles also remains to be answered. The RSV processing-deficient mutant env genes described here will allow us to address these problems by substituting them for the wild-type env gene in an infectious provirus genome.

ACKNOWLEDGMENTS

We thank K. Shaw for excellent technical assistance and J. Engler for the synthesis of the oligonucleotides. For critical comments on this manuscript we thank G. Air, S. Chatterjee, R. W. Compans, and B. Sudath. We are grateful to M. Pollard for assistance in preparing this manuscript for publication.

This work was supported by Public Health Service grant CA-29884 from the National Institutes of Health. L.P. is a predoctoral trainee supported by an ODEPLAN-Chile fellowship. E.H. was the recipient of Research Career Development Award CA-00685 from the National Cancer Institute.

LITERATURE CITED


