Human foamy virus (HFV) is the prototypic spumavirus that has distinct features of gene expression different from those of other known retroviruses. Foamy viruses (FVs) as complex retroviruses express genes from two different promoters (3, 13, 15), and unlike other retroviruses, HFV and feline FV (FeFV) express subgenomic pol-specific transcripts (1, 4, 16, 27). Although recombinant forms of the HFV Pol proteins have been shown previously to express active HFV reverse transcriptase (RT), RNase H (RH), and integrase (IN), to date, proteolytic processing of HFV Pol proteins has not been analyzed, and the precise interdomain cleavage sites of HFV Pol are unknown (2, 8, 14, 19, 21). Genetic analysis has shown that the HFV protease (PR) is absolutely required for infectivity and processing. In control experiments, HFV protease-deficient mutant proteins in which the active site Asp was replaced by an Ala residue were used to rule out unspecific processing by nonviral proteases. Specific proteolytic cleavage products were isolated, and the cleavage sites were analyzed by amino acid sequencing. Peptides spanning the resulting cleavage sites were chemically synthesized and assayed with HFV protease, and the cleaved peptides were subjected to mass spectrometry. The cleavage site sequences obtained were in complete agreement with the amino-terminal sequences from amino acid sequencing of authentic cleavage products of the HFV Pol proteins. Analysis by fast-protein liquid chromatography of a short version of the active HFV protease revealed that the enzyme predominantly formed dimeric molecules.

To map the precise residues that flank a given cleavage site, the following approaches were employed: (i) autocatalytic processing of His-tagged HFV Pol proteins shortened at various defined positions, (ii) in vitro PR assays with incubation of an appropriate HFV Pro-Pol protein that contained a suspected virus-specific cleavage site and subsequent isolation of sufficient amounts of the resulting cleavage products for amino acid microsequencing, and (iii) in vitro PR assays with synthetic peptides that were chosen from the regions flanking the putative HFV-specific cleavage sites. To rule out any unspecific proteolytic cleavages, PR-deficient mutant proteins in which the Asp of the active center of the viral PR was replaced by an Ala residue were generated. In these control experiments, the PR(D/A) mutant proteins were expressed in parallel, and the cleavage pattern was directly compared to that of the authentic HFV Pol proteins. Three His-tagged Pol polyproteins differing only in the COOH-terminal regions, namely viral inserts 2, 3, and 6 (Fig. 1), were cloned into the pET22b plasmid vector by PCR, expressed in Escherichia coli BL21(DE3) cells, purified by affinity chromatography on Ni$^{2+}$-chelate columns, and reacted with a polyclonal antiserum directed against HFV PR (20). The immunoblot showed that the full-length proteins were expressed and affinity purified, and the NH$_2$-terminal PR domain reacted with the PR antiserum (not shown). The apparent molecular masses of the PR-RT-RH-DIN and PR-ΔRT Pol proteins were in close agreement with the calculated values. As expected, the PR-ΔRH protein migrated slightly faster, consistent with previous reports (8). Remarkably, the PR subdomain was retained in the three Pol proteins, as shown by the positive immunoreaction with a polyclonal antisem directed against PR.

Upon close inspection, however, it was observed that in total bacterial lysates, various expressed HFV Pol proteins were invariably accompanied by autocatalytic processing (Fig. 2). To show autocatalytically proteolytic processing, the pET22b plasmids containing the HFV pol inserts were expressed in E. coli BL21(DE3) cells, and whole bacterial lysates analyzed by immunoblotting. The 72-kDa HFV PR-RT-ΔRH protein was partially cut to a smaller Pol protein of about 67 kDa (Fig. 2A, lane 1), which was a specific cleavage product, since it was not detected in the control reaction with the PR-inactive D/A mutant protein (lane 2). The same specific cleavage product as that in lane 1 was identified upon autocatalytic processing of
PR-RT-RH-DIN (arrowhead 1 + 3 in Fig. 2A, lane 3). During expression of the HFV PR-RT-RH-DIN protein, another cleavage event must have occurred, since a novel band with a size of about 85 kDa was observed in lane 3 (arrow 3). The additional cleavage products were not observed because of their small molecular sizes of about 3 to 4 kDa. Again, the cleavage products were not detectable in the PR(D/A) mutant protein (lane 4). In order to obtain protein bands suitable for amino acid sequencing, Pol proteins were expressed with longer COOH-terminal extensions that contain another PR(D/A) sequence (Fig. 1) reactive with antiserum against HFV PR (20). The viral inserts are schematically shown in Fig. 1 as proteins 4 and 5. Bacterial expression and autocatalytic processing of the PR-RT-DIN-TH-PRO(D/A) and of PR-RT-RH-DIN-TH-PRO(D/A) proteins is illustrated in Fig. 2B in parallel with the double PR-deficient mutants. Both recombinant proteins were autocatalytically processed to the cleavage product of about 67 kDa observed previously (arrowhead 1 + 3) that corresponds to the PR-RT domains shared by both proteins (lanes 1 and 3). In addition, a relatively large protein with a size of about 85 kDa was detectable in lane 3 (labeled with arrow 3), as expected for a cleavage between the RH and the IN domains of PR-RT-RH-DIN-TH-PRO(D/A) with one of the PR domains active.

It is noteworthy that two protein bands with sizes of about 27 kDa were additionally identified (double arrowheads 1 and 3). These sizes are consistent with the calculated values for the COOH-terminal extensions of inserts 4 and 5. Importantly, a cleavage product with a relatively high intensity of about 27 kDa was observed after autoprocessing of insert 5. The PR-inactive D/A mutant proteins showed unspecific bands that did not comigrate with the actual cleavage products in lanes 1 and 3. To determine the cleavage sites, the reaction products of 27 kDa were affinity purified and subjected to amino acid microsequencing (data not shown). The results revealed that the site where the amino terminus of the integrase was cleaved from the RH domain consisted of the sequence (NH₂)-CNTKKPNLDA. The amino-terminal part of the RT-RH cleavage site obtained by amino acid sequence analysis was (NH₂)-YTDGSAIKS (data not shown). Both sequences are unique and occur at the appropriate locations in the HFV Pol protein sequence deduced from nucleotide sequencing of the infectious HFV DNA (12).

To independently confirm and prove the authenticity of the

<table>
<thead>
<tr>
<th>Virus</th>
<th>Location of cleavage site in FV protein:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PR-RT</td>
</tr>
<tr>
<td>HFV</td>
<td>HWEN↓VGH</td>
</tr>
<tr>
<td>SFV-1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>HWEN↓VGH</td>
</tr>
<tr>
<td>SFV-3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HWEN↓VGH</td>
</tr>
<tr>
<td>BFV&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CWEN↓QGH</td>
</tr>
<tr>
<td>FeFV</td>
<td>SWEN↓QGH</td>
</tr>
<tr>
<td>Consensus</td>
<td>XWEN↓QGH</td>
</tr>
</tbody>
</table>

<sup>a</sup> Reported previously by Pfrepper et al. (20).
<sup>b</sup> SFV-1 and SFV-3, simian foamy virus types 1 and 3.
<sup>c</sup> BFV, bovine foamy virus.
cleavage sites, peptides that span the processing site were synthesized and assayed in vitro by either the HFV PR-D/His or TH-PRO in the presence of EDTA (20). Synthetic peptides that correspond to the RH-IN and the RT-RH sites were subjected to PR assays, and the cleavage products were analyzed by matrix-assisted laser desorption ionization (MALDI) mass spectrometry (20). This analysis revealed that proteolytic cleavage of the two peptides occurred at the sites marked by the vertical arrows TQGSYVVN\(^2\)CNTKK and PEGV\(^2\)YTDGSR, respectively, in agreement with the known HFV sequences. The resulting cleavage sites are compiled in Table 1. In parallel, apparent reaction products after incubation with HFV PR-deficient D/A mutant proteins were also subjected to analysis by MALDI mass spectrometry; mass peaks with sizes that corresponded to specific reaction products were not found.

So far, two processing sites of the HFV Pol protein have been identified, namely those between the RT and RH and the RH and IN domains. To determine which of the different proteolytically active forms of PR-Pol was the shortest version, a 179-amino-acid-long recombinant Pol protein, PR-\(\Delta RT\)-His (protein 6 in Fig. 1) that starts from the first Met residue of Pol and extends to residue 159, followed by a 14-amino-acid-long stretch of a vector-derived peptide sequence and a hexa-His tag was subjected to analysis. The PR-\(\Delta RT\)-His protein was bacterially expressed and purified in parallel with the corre-

FIG. 2. Bacterial expression and concomitant autocatalytic processing of authentic and PR-deficient mutant Pol polyproteins. (A) Immunoblots of the PR-RT-\(\Delta RH\) (lane 1) and the corresponding PR-deficient D/A mutant proteins (lane 2); PR-RT-RH-\(\Delta IN\) and the corresponding PR(D/A) mutant proteins (lanes 3 and 4) reacted with antiserum against HFV PR. Arrowhead 1+3 marks the proteolytic cleavage products with sizes of about 67 kDa in lanes 1 and 3; arrow 3 indicates the reaction product with a size of 85 kDa in lane 3; M, marker proteins with apparent molecular masses shown to the left. For the structures of the HFV inserts, see Fig. 1. (B) Immunoblot of PR-RT-\(\Delta RH\)-TH-PRO(D/A) (lane 1) and the corresponding PR-deficient double D/A mutant proteins (lane 2) reacted with antibody against HFV PR. Arrowhead 1+3 marks the PR-RT cleavage product p67, and arrow 1 marks a product with a size of about 27 kDa derived from the carboxy-terminal part of insert 4 (Fig. 1). Panel B also represents an immunoblot of processed PR-RT-RH-\(\Delta IN\)-TH-PRO(D/A) and the corresponding PR(D/A) double mutant proteins (lanes 3 and 4). Arrowhead 1+3 points to the p67 PR-RT cleavage product, arrow 3 points to the p85 reaction product, and arrowhead 3 marks a product with a size of about 27 kDa derived from the COOH-terminal regions of HFV insert 5. Double PR(D/A) mutant proteins served as controls (lanes 2 and 4).

FIG. 3. Autocatalytic processing of PR-\(\Delta RT\)-His proteins (viral insert 6). (A) Gel electrophoretic analysis and staining with Coomassie blue of PR-deficient D/A mutant PR-\(\Delta RT\)-His. Lane 1, PR-\(\Delta RT\)-His protein; lane 2, band of the PR-\(\Delta RT\)-His substrate with a size of about 21 kDa (arrow) and HFV PR as cleavage product of 17 kDa (arrowhead). The product cleaved at residue 166 (from within the vector backbone protein sequence) with a size of 19.6 kDa is marked with an open circle) (see panel B); the double arrowhead marks the two comigrating faint cleavage products of low molecular mass used for N-terminal sequencing. M, low-molecular-mass markers. (B) Schematic drawing of the proteolytically processed PR-\(\Delta RT\)-His protein cleaved at sites 143 and 166 resulting in four cleavage products; the two smaller peptides comigrated under the conditions used. Both the 3.5- and 1.5-kDa bands were microsequenced. N- and C-terminal residues are in the one-letter amino acid code.
acrylamide gel electrophoresis of both protein peaks eluted showed that not only the monomeric but also the dimeric PR-ART-His peak comigrated as a protein band of 21 kDa, as expected. This result is in agreement with the properties of other retroviral enzymes of the aspartic PR family that were unambiguously shown to exist as active homodimers (23, 25, 26). We identified three cleavage sites of the HFV Pol polyprotein by means of immunoblots with HFV PR-specific antisera, isolation of defined cleavage products from soluble, affinity-purified recombinant proteins that contained authentic HFV Pol sequences, and subsequent amino acid sequencing. Residues spanning the cleavage sites were chemically synthesized; the resulting peptides were subjected to proteolysis by the HFV PR (10) and analyzed by MALDI mass spectrometry. The PR-deficient D/A mutant proteins were consistently employed to control the specificity of the proteolytic cleavages observed. Our data are consistent with the molecular sizes of virtually all HFV Pol proteins reported to occur in HFV-infected cells (5, 8, 16, 18). The mature HFV Pol proteins encompass the integrase, the p85RT that also contains the RH domain, and the p67RT. Since an RH with the calculated value of 17 kDa was not detectable in HFV-infected cells, it is likely that the active HFV RT consists of a heterodimer of p85/p67 comparable to those of other retroviruses (7). Our results open the way to prove this hypothesis. Since only N-terminal and no C-terminal sequencing was carried out, we cannot rule out whether additional spacer peptides exist.

As to the HFV PR, our data show that distinct forms of PR were proteolytically active. One short form of HFV PR, PR143, was capable of cleaving the HFV Gag precursor into p70 and p3 and the two peptides that link the RT-RH and the RH-IN domains (data not shown). One of the open questions is which form of the proteolytically active HFV PR is responsible for the individual steps of Pol processing in vivo.

It is noteworthy that forms of the HFV Pol proteins in which the PR domain was removed were not consistently observed in infected cells. The cleavage between the RH domain and the IN domain was faster and more efficient than those of the two other sites that were cleaved with suboptimal efficiency. This was also observed during proteolysis of the corresponding junction peptides. It is worth mentioning that processing of the site spanning the PR and RT was even less efficient. The poor cleavage efficiency of this site might be one reason that only minute amounts of PR143 are available for processing of FV Gag proteins that in virus-infected cells has been reported to be invariably incomplete (14, 27). It seems possible that the PR-RT cleavage might not be required for virus infectivity. The inefficient cleavage between residues 143 and 144 may also be reflected by the fact that the PR143 did not cleave the corresponding peptides of various lengths that span this site (data not shown). Close examination of the predicted secondary structures of HFV PR by the EMBL phd program (22) showed that FV PR-specific residues from 121 through 144 form a stable alpha helix. A search program in the data banks revealed that part of this HFV PR sequence from residues 128 to 139, KTLFVKYDNLWQ, was highly homologous to an alpha-helical sequence, KKLTKYDNLE (identical residues are underlined), of the galactose-1-phosphate-uridyltransferase as determined by X-ray crystallography (11, 24). The three-dimensional structure of an FV PR will be required to solve this question and related issues. Table 1 shows that FV PRs seem to prefer Val at the P2 or P2’ position, the scissile bond being P1 and P1’ (23, 26). A comparison of the flap regions of well-studied retroviral PRs with those of the HFV PR illustrates large differences and relatively few common features (6, 17, 23, 26).
We thank Helmut Bannert for excellent technical assistance, Jennifer Reed for critically reading the manuscript, and Harald zur Hausen for support.

REFERENCES