Ubiquitin in retrovirus assembly: Actor or bystander?
Volker M. Vogt*
Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

The assembly of retroviruses is deceptively simple. Only the product of the gag gene is required for the formation of a virus-like particle (1). After its synthesis in the cytoplasm Gag is targeted to the plasma membrane, where Gag-Gag interactions, Gag-RNA interactions, Gag-membrane interactions, and perhaps Gag-host protein interactions lead to a bulging out of the nascent virus particle, first as a horseshoe-shaped and then as a lollipop-shaped structure that can be visualized by electron microscopy. The process of membrane envelopment of the virus is called budding. The last step in budding is a membrane fusion event that serves to pinch off the virus, releasing it into the medium. Late in assembly, perhaps concomitant with pinching off or just preceding it, a viral protease cleaves Gag at several sites, leading to a morphological change in the virus. This process is called maturation and is essential for the virus to become infectious.

In this issue of PNAS, three articles by independent groups provide an important new clue to the mechanism by which retroviruses bud from the plasma membrane (2–4). The articles implicate the cellular protein ubiquitin in a late step in budding. Ubiquitin is a small highly conserved 76-aa protein well known for its role in protein degradation by the proteasome (5). In proteins targeted for degradation, ubiquitin becomes coupled via its C-terminal glycine residue to the ε-aminogroup of lysine residues in the protein (or rarely to the N-terminal α-amino group), forming an isopeptide bond. Other ubiquitin proteins are similarly coupled to lysine residues in this first ubiquitin, and thus the target protein ends up carrying a ramified polyubiquitin chain that acts as a signal recognized by the proteasome (6). By mechanisms that are incompletely understood, the targeted protein is degraded while the ubiquitin molecules are recycled. A less well-known role for ubiquitin is in endocytosis of plasma membrane proteins (7, 8). For example, this process has been studied with yeast plasma membrane receptors (9) and with human growth hormone receptor (10). In yeast a single ubiquitin moiety becomes coupled to the cytoplasmic tail of the receptor in response to stimulation of the receptor, and this signals internalization by endocytosis. The enzymology of ubiquitin conjugation is complex. The first step is activation of ubiquitin by an ATP-requiring enzyme called E1, leading to conjugation of ubiquitin to E1 via a high-energy thioester bond. The ubiquitin is transferred by E1 to a second enzyme, E2. A third enzyme, E3, which may form a complex with E2, then transfers the ubiquitin to amino groups of the target protein. Although cells have only one E1, they may have dozens of different E2 and E3 enzymes, which confer specificity to the process of ubiquitination. Similarly, there are many different ubiquitin hydrolases that can cleave the isopeptide bonds between two ubiquitins or between ubiquitin and a target protein (11).

Despite lack of overall sequence similarity, the Gag proteins of retroviruses in different genera are functionally and structurally similar. All contain three common domains that become mature proteins upon proteolytic processing. The MA (matrix) domain contains a plasma membrane targeting sequence, and in the mature virus particle MA lines the inner face of the membrane. The CA (capsid) domain participates in protein–protein interactions in assembly, and the mature CA forms the core shell in the infectious virus. The NC (nucleocapsid) domain recognizes the packaging signal in the genomic RNA, and the mature NC coats and condenses the RNA genome. In addition to the three common mature proteins, Gag proteins have other polypeptide segments whose functions have not been fully elucidated and which vary among virus genera. Deletion analyses of HIV-1 and the avian Rous sarcoma virus (RSV) (Fig. 1), as well as a few other model systems, have shown that much of Gag is dispensable for budding, when budding is defined minimally as the efficient release of Gag in an enveloped and particulate form into the medium. The segments of Gag that are essential are called assembly domains (not to be confused with the units that are separated by cleavage, like MA or CA). These essential segments of Gag include a sequence called the late domain because it is required for efficient pinching off of the virus bud (12–15). It is the late domain that has a functional relationship with ubiquitin.

The core element of the late domains of RSV (16) (Fig. 1), murine leukemia virus (17), and Mason Pfizer monkey virus (18) is the sequence PPPY, which is found about 150–200 amino acid residues from the Gag N terminus in these and other retroviruses. The core element of the HIV-1 late domain, PTAP (15), is located in a different part of the protein, the C-terminal portion called p6 (Fig. 1). In a distantly related retrovirus of the same genus, equine infectious anemia virus (EIAV), the late domain also is near the C terminus of Gag, but has the sequence YPDL (19). Remarkably, these three core sequences together with immediately adjoining sequences can supply late domain function if inserted into an RSV Gag or an HIV-1 Gag missing its own late domain, and can do so at either the HIV-1 or the RSV location (13). The late domain function appears to be mediated by cellular proteins. PPPY (consensus PPxY) is a sequence recognized by a family of protein interaction modules called WW domains (20). YPDL (consensus YxxL) is recognized by the AP-2 complex that is involved in clathrin-mediated endocytosis (21). Indeed AP-2 colocalizes with EIAV budding structures by immunofluorescence techniques (22). The consequence of an HIV-1 late domain deletion is that the cells become covered with virus particles that remain tethered to the membrane by narrow stalks (14). This dramatic phenotype depends on the presence of an active protease (15), implying some functional connection between maturation and the final steps in budding.

Enter ubiquitin and ubiquitin ligase. The hint that ubiquitin might have something to do with retrovirus assembly came over a decade ago from the observation that RSV particles contain some 100 molecules of this protein in an unconjugated form, which represents an enrichment over the concentration of free ubiquitin in the cytoplasm of the cell (23). Similar amounts of free ubiquitin later were reported in HIV-1 and murine leukemia...
virus (MuLV) (24). In addition, a small percentage of the mature HIV-1 and MuLV Gag proteins that harbor the late domains was found to be ubiquitinated. The present papers by Patnaik et al. (2) and Schubert et al. (3) now establish a putative causal link between budding and ubiquitin, by showing that depletion of the intracellular pool of free ubiquitin inhibits budding. Both papers used the same experimental logic, treatment of cells with inhibitors that block proteasome function. This prevents recycling of the ubiquitin that is covalently coupled to proteins destined for degradation, locking up ubiquitin in a conjugated form. The rates of budding by RSV Gag (2) and HIV-1 (3) were found to be reduced by about 3-fold, with the reduction depending on the presence of a late domain. Although this level of inhibition is not large, its consistency in two different retrovirus systems is impressive. Some drug-treated RSV-infected cells had large aggregates of almost crystalline arrays of virus particles attached to the cell surface and attached to each other by thin stalks. The observation of viruses tethered to other viruses is reminiscent of the recently described phenotype of an MuLV late domain defect (25) and suggests that viruses bud sequentially from exactly the same sites on the plasma membrane. In the HIV-1 experiments the proteasome inhibitors not only reduced budding, but reduced proteolytic maturation of Gag to a similar extent. In contrast to HIV-1, RSV proteolytic maturation was not affected by the inhibitors. These apparently contradictory observations may relate to the fact that the kinetics of RSV budding are not affected by the viral protease, whereas there is at least one report that the kinetics of HIV-1 budding are slower in absence of protease function (26).

By what mechanism does proteasome inhibition adversely affect retrovirus budding? As discussed by Schubert et al. (27), one possibility derives from the buildup of defective Gag translation products that otherwise would be rapidly degraded by the proteasome. Even a few defective Gag proteins could be imagined to interfere with budding. Although this possibility has not been rigorously excluded, it does not explain the observation for RSV that genetically fusing a copy of ubiquitin onto the C terminus of Gag partially restored budding in the inhibitor-treated cells (2), a result analogous to that described previously for ubiquitin-dependent endocytosis of a yeast membrane receptor to which ubiquitin had been fused (9). Thus apparently it is the covalent attachment of ubiquitin to Gag that is important for budding, and not some other aspect of ubiquitin metabolism. However, it is troublesome that RSV Gag has not been found to be ubiquitinated, either previously (23) or in the present study (2). Moreover, although the level of ubiquitination of HIV-1 Gag decreased in inhibitor-treated cells, mutating the two lysine residues in p6 that are known to be acceptors of ubiquitin failed to compromise budding (28). For two reasons these results do not necessarily negate a critical role for ubiquitin-Gag conjugation. First, it is known that ubiquitin conjugation to proteins is not necessarily specific to particular lysine residues. If an E3 enzyme bound to the late domain of Gag, and then it searched for lysine acceptors nearby in three-dimensional space, the low level of any particular Gag ubiquitin conjugate might well have been missed in the published studies. Second, conjugation could well be transient, being reversed by one of the many ubiquitin hydrolases (11). Nevertheless, the case for Gag ubiquitination would remain weak were it not for the third paper in this issue (4).

One of the smallest assembly-competent HIV-1 Gag constructs yet created is only one-third the size of wild-type Gag (29). In studying the budding of this protein, Strack et al. (4) found that up to about one-half of all of the Gag molecules in the released virus-like particles were conjugated to one or more ubiquitins. Ubiquitination was decreased by proteasome inhibitors and absolutely depended on the presence of a late domain, with diverse core late domain sequences being able to supply this function. Among the most efficacious sequences was one from Ebola virus that contains partially overlapping PPxY and PTAP sequences. The authors noted from database analysis that several other viruses that become membrane enveloped in their life cycle, such as rhabdoviruses, also have closely juxtaposed PPxY and PTAP or PSAP sequences. Indeed the PPxY sequence in rhabdoviruses has been suggested previously to function in budding (30, 31).

What model best accounts for the observations in these three papers? The most obvious is that a ubiquitin ligase is pulled to the site of budding by interaction with the late domain, and there it conjugates ubiquitin to Gag and/or to other cellular proteins (Fig. 2). A relationship between ubiquitination and the late domain sequence PPxY has precedents outside of retrovirology. The protein Nedd4 is a multidomain polypeptide that includes a calcium-dependent lipid binding domain, an E3 ubiquitin ligase domain, and

---

**Fig. 1.** Structure of HIV-1 and RSV Gag proteins. Vertical lines denote sites of cleavage by the viral protease that liberates the mature proteins indicated. The vertical black bar represents the late domain, indicated by the core sequence PTAP or PPPY for these proteins.

---

**Fig. 2.** Speculative representation of late stage in retroviral budding. The Gag molecules have formed a spherical particle ready to pinch off from the plasma membrane. Some ubiquitin molecules (red star), both free and conjugated to Gag, are present in the nascent virus. A ubiquitin ligase (blue rectangle) is recruited to the site of budding by the Gag late domain, and it conjugates ubiquitin to some of the Gag molecules and perhaps to cellular proteins (green oval). Other cellular proteins (yellow triangle) may bind to these ubiquitin conjugates, setting up the machinery to pinch off the virus bud. (Adapted from ref. 20.)
several WW domains. Nedd4 is implicated in down-regulation of the epithelial sodium channel, and it acts by binding to PPxY motifs in subunits of the channel complex and then ubiquitinating them, leading to endocytosis and degradation (32). How ubiquitination of Gag or nearby cellular proteins might function is unknown. Presumably further proteins in turn would bind to the conjugated ubiquitin, as inferred for the ubiquitin involved in endocytosis of yeast receptors, ultimately recruiting cellular membrane remodeling machinery. This machinery probably shares elements with that used for retrovirus budding, as suggested by the observation that the late domain of EIAV interacts with AP-2 (22).

To establish the function of ubiquitin in retrovirus budding, several outstanding questions need to be answered. (a) What is the relevant target of ubiquitination? Does Gag ubiquitination play a role in budding, or is this modification simply acting as a reporter for the nearby presence of a ubiquitin ligase, whose real function is to ubiquitinate a cellular protein at the budding site? The data available to date do not yet make a persuasive case for the importance of modifying Gag by ubiquitin addition, but no other targets have been found or suggested. In the human growth hormone system, when all of the lysines in the cytoplasmic tail were mutated, the receptor was still internalized in a ubiquitin-dependent manner, implying that ubiquitination of the protein itself cannot be involved (33). (b) What is the origin of the free ubiquitin in virus particles? Are these molecules derived from hydrolysis of ubiquitin conjugated to Gag or conjugated to a cellular protein in the vicinity of the budding site? Alternatively, are the free ubiquitin molecules in the virus simply a footprint left close to the site of budding by a locally high concentration of proteins with ubiquitin binding domains? (c) Why are the HIV-1 mini-Gag proteins more heavily ubiquitinated than wild-type HIV-1 Gag? Is this a reflection of the presence of a ubiquitin hydrolase in normal particles but not mini-Gag particles, or is the size of the Gag protein or the structure of the particle the critical element? (d) What is the relationship between late domain function and proteolytic maturation of Gag, and why does proteolysis depend on late domain function in HIV-1 but not RSV? (e) Finally, what is the nature of the membrane fusion machinery implied by the late budding defect? Does the cortical cytoskeleton, which has been implicated indirectly in budding of HIV-1 (34), play a role in the failure of virus particles to pinch off completely?

Two other groups also have implicated ubiquitin in the function of late domains in retrovirus assembly, and their results provide further clues to its mechanism of action. Kikonyogol et al. (J. Leis, personal communication) found the E3 ubiquitin ligase Nedd4 to interact with the RSV late domain, and VerPlank et al. (C. Carter, personal communication) found the E2-like protein Tsg101 to interact with the HIV-1 late domain. Both groups also have evidence that these proteins can bind to their respective Gag partners in transfected cells. Although Tsg101 is clearly a ubiquitin ligase by overall sequence similarity, it lacks the active site Cys residue needed for ligation. The significance of the implied absence of E2 enzymatic function remains to be established, and suggests that the role of ubiquitination may be complex.

Because diverse viruses appear to have late domain sequences, we can anticipate that the principles learned from further studies of retroviral late domains and ubiquitination will have wide relevance. It may turn out, as has happened frequently in the history of virology, that deciphering how a virus has coopted cellular machinery will lead to insights into fundamental cellular processes that are otherwise difficult to approach.