Virology **298,** 30–38 (2002)
doi:10.1006/viro.2002.1452

Hexagonal Organization of Moloney Murine Leukemia Virus Capsid Proteins

Keith Mayo, Jason McDermott,[1] and Eric Barklis[2]

Vollum Institute and Department of Microbiology, Oregon Health Sciences University, Portland, Oregon 97201-3098

Received August 30, 2001; returned to author for revision December 11, 2001; accepted March 5, 2002

To help elucidate the mechanisms by which retrovirus structural proteins associate to form virus particles, we have examined membrane-bound assemblies of Moloney murine leukemia virus (M-MuLV) capsid (CA) proteins. Electron microscopy and image reconstruction techniques showed that CA dimers appear to function as organizational subunits of the cage-like, membrane-bound protein arrays. However, new three-dimensional (3D) data also were consistent with hexagonal (*p*6) assembly models. The *p*6 3D reconstructions of membrane-bound M-MuLV CA proteins gave unit cells of $a = b = 80.3$ Å, $c = 110$ Å, $\gamma = 120°$, in which six dimer units formed a cage lattice. Neighbor cage hole-to-hole distances were 45 Å, while distances between hexagonal cage holes corresponded to unit cell lengths (80.3 Å). The hexagonal model predicts two types of cage holes (trimer and hexamer holes), uses symmetric head-to-head dimer building blocks, and permits the introduction of lattice curvature by conversion of hexamer to pentamer units. The M-MuLV CA lattice is similar to those formed in helical tubes by HIV CA in that hexamer units surround cage holes of 25–30 Å, but differs in that M-MuLV hexamer units appear to be CA dimers, whereas HIV CA units appear to be monomers. These results suggest that while general assembly principles apply to different retroviruses, clear assembly distinctions exist between these virus types.

INTRODUCTION

With the recent availability of high-resolution structures for retrovirus structural (Gag) proteins (Berthet-Colominas *et al.,* 1999; Campos-Olivas *et al.,* 2000; Conte *et al.,* 1997; Gamble *et al.,* 1996, 1997; Gitti *et al.,* 1996; Jin *et al.,* 1999; Khorasanizadeh *et al.,* 1999; McDonnell *et al.,* 1998; Momany *et al.,* 1996), a central question concerning retrovirus assembly is how the three conserved Gag protein domains—matrix (MA), capsid (CA), and nucleocapsid (NC)—organize to form virus particle structures (Swanstrom and Wills, 1997). Examination of virus particle radial density profiles (Fuller *et al.,* 1997; Yeager *et al.,* 1998) has validated and extended previous electron microscopy (EM) studies (Bolognesi *et al.,* 1978; Nermut *et al.,* 1994), indicating that matrix, capsid N-terminal (NTD), and C-terminal domains (CTD) and nucleocapsid domains are arranged radially from immature retrovirus particle envelopes to their central cores. Although structural characterization of retroviruses has proven difficult due to virus heterogeneity, EM analysis at low resolution has demonstrated that subunits assemble into cage-like lattices to form virus particles (Nermut *et al.,* 1994, 1998).

The nature of subunit associations in Gag protein cages has been investigated in several *in vitro* assembly systems for primate viruses such as HIV, for murine retroviruses such as Moloney murine leukemia virus (M-MuLV), and for avian retroviruses such as Rous Sarcoma virus (RSV) (Aggresta and Carter, 1997; Barklis *et al.,* 1997, 1998; Campbell and Rein, 1999; Campbell and Vogt, 1995, 1997; Ehrlich *et al.,* 1992; Ganser *et al.,* 1999; Grattinger *et al.,* 1999; Gross *et al.,* 1998, 2000; Kingston *et al.,* 2000; Klikova *et al.,* 1995; Li *et al.,* 2000; McDermott *et al.,* 2000; Nermut *et al.,* 1994, 1998; Scarlata *et al.,* 1998; von Schwedler *et al.,* 1998; Zuber and Barklis, 2000; Zuber *et al.,* 2000). Such studies have been instrumental in defining the influences of membranes, nucleic acids, and Gag protein domains on virus-like particle assembly and shape, and regular structures assembled *in vitro* have been used in EM reconstructions to identify subunit interactions (Barklis *et al.,* 1997, 1998; Li *et al.,* 2000; McDermott *et al.,* 2000; Zuber and Barklis, 2000; Zuber *et al.,* 2000). In particular, HIV capsid proteins in helical tubes formed under high-salt conditions were shown to organize into hexamers surrounding protein-free 25-Å diameter holes, spaced at 107-Å distances (Li *et al.,* 2000). With M-MuLV, we have shown that histidine-tagged (his-tagged) capsid (his-MoCA) and capsid plus nucleocapsid (his-CANC) proteins can assemble into two-dimensional (2D) crystalline arrays on membranes containing nickel-chelating lipids (Barklis *et al.,* 1997; McDermott *et al.,* 2000; Zuber and Barklis, 2000; Zuber *et al.,* 2000). These studies suggested that his-MoCA and his-MoCANC proteins assemble on membranes in hexameric cages (Barklis *et al.,* 1997; Zuber and Barklis, 2000; Zuber *et al.,* 2000), and three-dimensional (3D) reconstructions performed with no symmetry constraints

[1] Current address: Department of Microbiology, University of Washington, Seattle, WA 98166.

[2] To whom correspondence and reprint requests should be addressed. E-mail: barklis@ohsu.edu.

0042-6822/02 $35.00

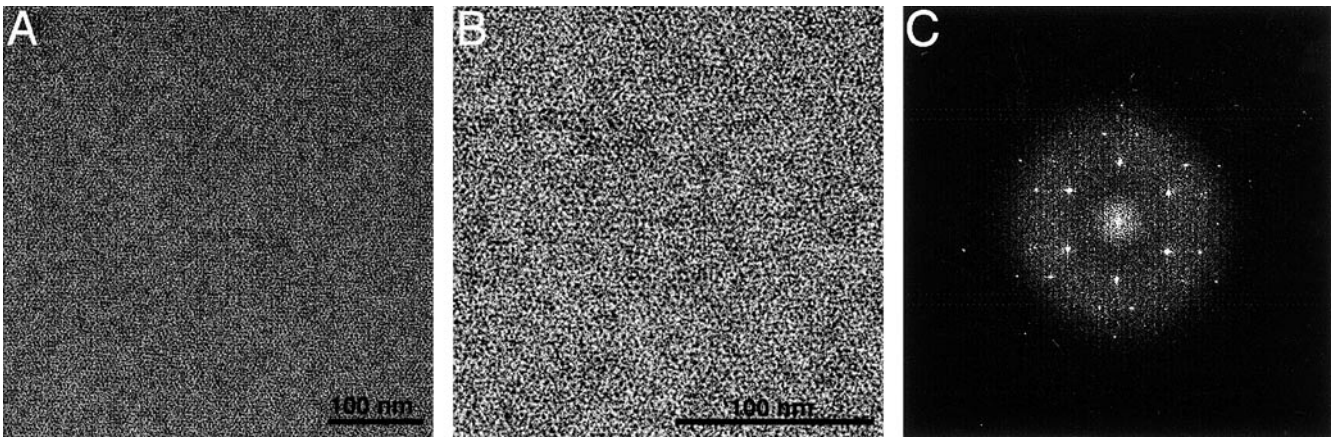

FIG. 1. Assembly of M-MuLV capsid proteins on lipid membranes. Purified his-tagged M-MuLV capsid (his-MoCA) proteins were incubated beneath lipid monolayers of egg PC plus the nickel-chelating lipid DHGN as described under Materials and Methods. After incubations, membranes plus associated proteins were lifted onto grids, negatively stained, and imaged by EM. (A) A close-to-focus (700-nm defocus) image (A121121) of membrane-bound his-MoCA proteins, with proteins in white and protein-free regions dark. (B) A magnified, contrasted version of the image to show the protein pattern, although the crystal lattice is not discerned readily under these imaging conditions. To reveal the crystal lattice, the calculated diffraction pattern of the original image is displayed as a power spectrum in C, showing a hexagonal organization of the protein assembly. In this panel, the innermost (1,0) reflections are not apparent, while the inner bright six reflections correspond to the hexagonally indexed 1,1 reflections and are at a resolution of 1/40.17 Å. For this image, the radius at which the power spectrum fades, corresponding to the first inversion of the CTF, is at about 1/16 Å.

($p1$) showed that CA dimers appear to be assembly building blocks, but did not show evidence of hexamer formation (McDermott *et al.,* 2000).

Because the apparent 2D hexagonal symmetry of membrane-bound his-MoCA proteins was not reflected by 3D data previously collected from highly tilted images (McDermott *et al.,* 2000), we decided to collect and analyze an improved data set from multiple tilt series. Our new data set proved to be compatible with hexagonal ($p6$) symmetry and shows some interesting features. Notably, membrane-bound head-to-head capsid protein dimers appear to organize as spokes around hexameric cage holes, forming a network of hexameric and trimeric cage holes. This protein arrangement suggests how trimeric (Fass *et al.,* 1996) M-MuLV envelope proteins might be coordinated at cage hole sites, either in a fixed fashion or with two degrees of freedom. The arrangement also allows for the formation of closed spherical virus particles through the conversion of hexameric centers to pentameric centers.

RESULTS AND DISCUSSION

Analysis of membrane-bound M-MuLV capsid protein projections

To examine the manner by which retrovirus Gag proteins organize on membranes, we have developed a method where his-tagged Gag proteins can be assembled *in vitro* onto monolayers containing nickel-chelating lipids (Barklis *et al.,* 1997, 1998; McDermott *et al.,* 2000). It has been shown previously that deletion of the retrovirus Gag protein amino-terminal matrix domain is compatible with particle assembly, as long as a membrane anchoring function has been provided (Wang *et al.,* 1993; Barklis *et al.,* 1997; Reil *et al.,* 1998). Thus, an advantage of our *in vitro* method is that membrane binding of N-terminally his-tagged Gag proteins mimics the normal *in vivo* assembly process. Moreover, 2D electron density projections obtained from *in vitro* incubations of his-tagged proteins have proven comparable to results obtained at lower resolution from immature virus particles (Barklis *et al.,* 1997, 1998; Fuller *et al.,* 1997; Yeager *et al.,* 1998). Thus, to analyze the 3D membrane organization of amino-terminally his-tagged M-MuLV capsid protein (his-MoCA), proteins in buffer were incubated beneath lipid monolayers, after which membranes and associated proteins were lifted onto grids and visualized by EM.

Although direct images of membrane-bound his-MoCA proteins appeared relatively featureless (Fig. 1A), higher contrast, magnified images showed a regular arrangement of protein (light) and protein-free (dark) areas (Fig. 1B). Evidence for the regular organization of membrane-bound his-MoCA proteins was apparent when diffraction patterns were obtained from EM images. As shown in Fig. 1C, patterns demonstrated a distinctly hexagonal appearance, and indexing of multiple diffraction patterns yielded an average unit cell of $a = 80.35 \pm 3.77$ Å, $b = 80.25 \pm 3.30$ Å; $\gamma = 120.14° + 1.21°$. Given the hexagonal appearance of the diffraction patterns, it was not surprising that the phases of the reflections were consistent with the presence of twofold ($p2$), threefold ($p3$), and sixfold ($p6$) axes of symmetry. This was demonstrated by the observation that acceptably low ($<30°$; Amos *et al.,* 1982; Crowther *et al.,* 1996)

TABLE 1

Space Group Analysis of 2D His-MoCA Crystals

Space group	Phase residual value (°)				
	Internal consistency	Overall merge	0–15° tilts	20–35° tilts	40–55° tilts
*p*1	19.8 ± 5.0	19.4 ± 2.2	20.5 ± 2.4	16.5 ± 2.1	18.4 ± 1.8
p1opt	19.8 ± 5.0	15.5 ± 1.9	17.6 ± 1.9	13.8 ± 2.1	13.5 ± 1.5
*p*2	27.7 ± 14.1	19.5 ± 2.3	22.0 ± 2.7	16.2 ± 2.0	19.8 ± 1.6
*p*3	16.7 ± 8.3	19.4 ± 2.2	20.7 ± 2.3	15.8 ± 2.3	21.4 ± 2.0
*p*6	18.5 ± 10.1	19.1 ± 2.2	20.7 ± 2.4	21.8 ± 2.3	28.6 ± 2.0

Note. Phase data from 56 untilted and tilted diffraction patterns of his-MoCA crystals were used to evaluate space group fits by calculation of phase residual values, which are inversely related to goodness of fit (0° indicates a perfect fit; 90° is a random fit). Internal consistencies were determined from five untilted crystals to 15 Å using the program ALLSPACE (Crowther *et al.,* 1996), while overall merge values and values for low (0–15°), medium (20–35°), and high (40–55°) tilt images were obtained using the programs ORIGTILTB (p2, p3, p6) or ORIGTILTC (p1, p1opt) (Crowther *et al.,* 1996; McDermott *et al.,* 2000). Note that all other space groups gave internal consistency phase residual values of >30°, and that p1opt refers to the indexing of reflections into a primitive, but rotationally optimized (McDermott *et al.,* 2000) space group set.

phase residuals were obtained not only when it was assumed that proteins assembled into unit cells with no (primitive; *p*1) symmetry, but also when symmetry-related reflections were compared for internal consistency with the *p*2, *p*3, and *p*6 space groups (Table 1).

From diffraction patterns such as that shown in Fig. 1C, it was relatively straightforward to calculate 2D electron density maps for membrane-bound his-MoCA proteins. To do so, amplitude and phase data from indexed diffraction patterns were back-transformed (see Materials and Methods) to give signal-to-noise enhanced density maps of proteins viewed perpendicular to the membrane. As shown (Fig. 2A), proteins (dark) formed a cage network surrounding two types of protein-free (white) cage holes. These appeared either circular and hexagonally coordinated or roughly triangular and trigonally symmetric (Fig. 2A). In accordance with unit cell data from diffraction patterns, the spacing between similarly oriented, identical features is 80.3 Å.

While 2D maps of membrane-bound proteins were obtained readily, 3D reconstruction required several additional steps. In particular, images from crystals viewed at different tilts were collected so as to obtain a 3D database of projections. After that, diffraction patterns from untilted and tilted images had to be aligned, merged, averaged, and then used for construction of 3D reciprocal space lattices, which could be back-transformed to generate 3D density maps (see Materials and Methods). When these steps were performed without any symmetry constraints (*p*1), or assuming no symmetry, but statistically optimizing for rotational alignment (p1opt), we found that overall data merging yielded low (15–20°) phase residuals and that phase residuals remained low with high-tilt data (Table 1). However, we also merged our data assuming that proteins assembled with twofold (*p*2), threefold (*p*3), and sixfold (*p*6) symmetry. As usual, this was accomplished by mapping amplitude and phase data from symmetry-related reflections into 3D lattices that were one-half (*p*2), one-third (*p*3), and one-sixth (*p*6) the size of the *p*1 map (Amos *et al.,* 1982; McDermott *et al.,* 2000). In performing *p*2, *p*3, and *p*6 merges, low (<30°) phase residuals also were obtained (Table 1), suggesting that his-MoCA proteins assemble onto membranes with two-, three-, or sixfold symmetry. Indeed, the slight statistical preference for primitive (*p*1 and p1opt) data merges was less than observed previously (McDermott *et al.,* 2000), possibly due to closer-to-focus imaging, better alignment of series images, and more accurate tilt axis determination (Materials and Methods).

Based on phase residual analyses (Table 1, and see above), the assemblies formed on membranes by his-MoCA proteins were compatible with *p*6 space group symmetry and also the lower symmetry *p*2 and *p*3 space groups. Moreover, features present in 2D projections from untilted images were obscured in projections obtained from *p*1 and p1opt 3D reconstructions, but were retained in higher symmetry reconstructions. In particular, the appearance of cage holes in Fig. 2A was maintained to a greater degree in projections of *p*3 and *p*6 3D reconstructions (Figs. 2E and 2F) than from reconstructions using lower symmetry space groups (Figs. 2B–2D). The loss of such details may be a consequence of the database size, the sequential nature of merging programs, and/or the missing cone of information which occurs in 2D crystallographic studies (Amos *et al.,* 1982; Barth *et al.,* 1988; Crowther *et al.,* 1996; McDermott *et al.,* 2000).

3D structure of hexagonally organized his-MoCA proteins

Because hexagonal 3D merges yielded low phase residuals (Table 1) and appeared to reproduce 2D projections faithfully (Fig. 1), it was of interest to examine this electron density map in more detail. Thus, a new

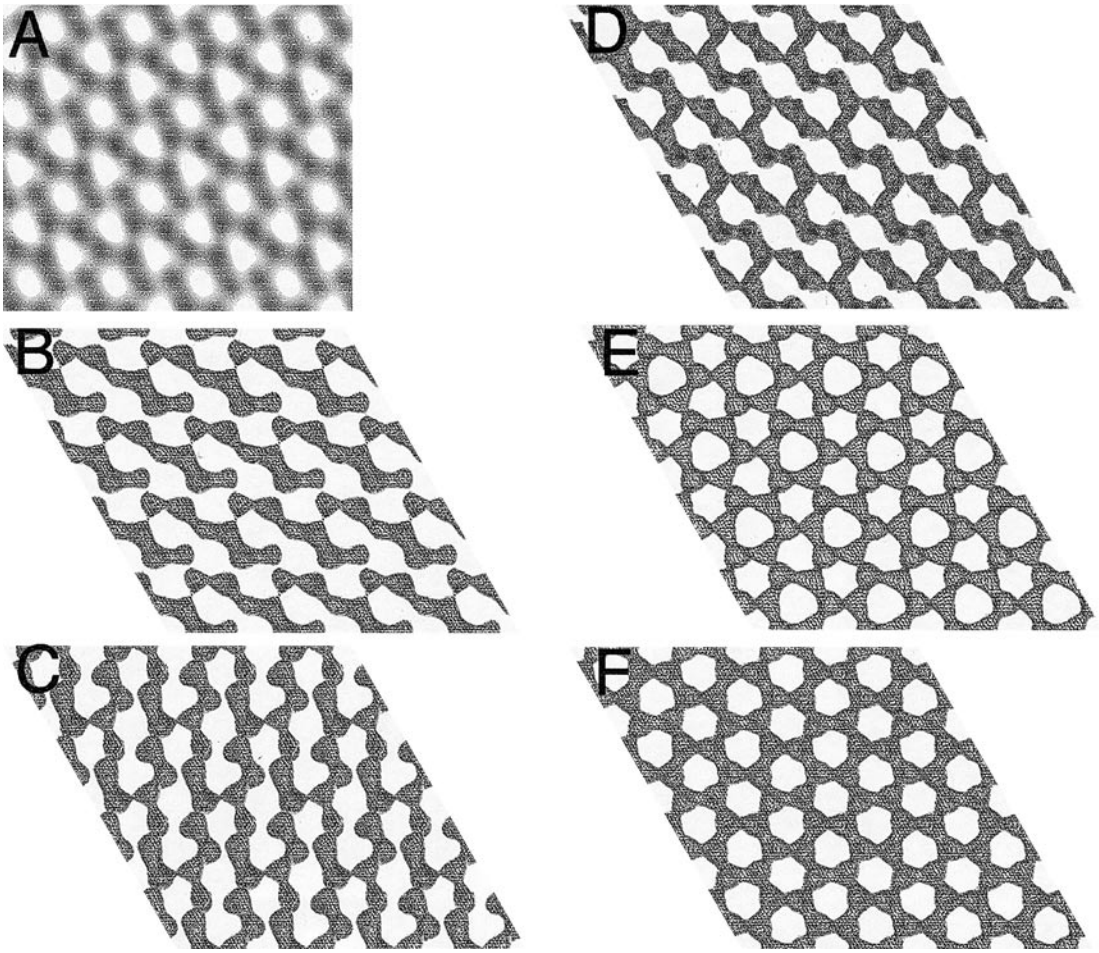

FIG. 2. Comparison of symmetry constraints on image reconstructions. Shown is a 320 Å × 270 Å back-transformation of zero tilt reference film 121120 (A), along with projections of 3D reconstructions from data merges applying different symmetry constraints (B–F). Projection reconstructions are viewed perpendicular to and above membranes and represent 4 × 4 unit cells of $a = b = 80.3$ Å, $c = 110$ Å, $\gamma = 120°$. Symmetry constraints were as follows: (B) no symmetry constraints, but images were rotationally aligned to obtain optimized phase residuals (p1opt); (C) no symmetry constraints (*p*1); (D) constrained by a twofold axis of symmetry perpendicular to the membrane plane (*p*2); (E) trigonal (*p*3) symmetry; (F) hexagonal (*p*6) symmetry. All 3D reconstructions used the same 56 image files and were filtered to 20 Å in the *h*, *k* plane and 30 Å perpendicular to the *h*, *k* plane. Reconstruction projections are centered on the *p*6 phase center, using an electron density contour level of 2σ. The *p*6 phase centers on the 2D back-transformation (A) are close to the lower left and upper right corners, and the contrast of the original back-transformation was inverted and adjusted to 52% of the original maximum brightness so that protein areas appear at a similar grayscale level as in B–F.

data set merge, applying *p*6 symmetry, was performed using IQ ≤ 6 (Crowther *et al.,* 1996; McDermott *et al.,* 2000) reflections. As shown in Table 2, the *p*6 merge gave good phase residuals to 20-Å resolution, dropping off slightly at 15 Å. To maintain a high degree of completeness, and because lower resolution normally is obtained out of the crystalline plane, the merge was filtered to 19 Å in the *h*, *k* plane and to 30 Å in 1. The resulting reciprocal space lattice was 88.4% complete. Additionally, there was a high reproducibility of phase measurements averaged in each *h*, *k*, *l* bin, as indicated by *Q* factors (ratios of vector to scalar sums; McDermott *et al.,* 2000) close to a value of 1.

The 3D his-MoCA *p*6 reconstruction is shown in Fig. 3. When viewed perpendicular and above the lipid membranes, corresponding to a view from the outside to the inside of a virus particle, hexagonal (Fig. 3A) projections show two different types of cage holes, similar to our 2D projections (Fig. 2A). In agreement with previous results (Barklis *et al.,* 1997; McDermott *et al.,* 2000; Zuber and Barklis, 2000; Zuber *et al.,* 2000), cage hole diameters were 27–37 Å, and nearest neighbor hole-to-hole spacings were 45 Å, while like-to-like hole spacings were 80 Å. Assuming that barbell-shaped units represent dimers, the appearance of the *p*6 units (Fig. 3A) is consistent with that of head-to-head dimers. The relationship between putative dimer units could be examined by lowering the electron density contouring levels from 3σ to 2.5σ and tilting the maps such the upper edges are closer to (Fig. 3B), equidistant from (Fig. 3C), or further (Fig. 3D) from the viewer than the bottom edges. With these contours and tilts, the contacts between putative dimers are more evident, and it appears that each of the barbell unit ends makes two additional connections with their neigh-

TABLE 2

Analysis of the His-MoCA Hexagonal 3D Reconstruction

Merge phase residuals (°)	
to 30 Å	18.9
to 25 Å	18.3
to 20 Å	19.1
to 15 Å	24.4
Completeness:	88.4%
Vector addition analysis	
1 vector	$Q = 1.000$
	$N = 16$
2–4 vectors	$Q = 0.907$
	$N = 58$
5–9 vectors	$Q = 0.931$
	$N = 59$
10–19 vectors	$Q = 0.895$
	$N = 23$
≥20 vectors	$Q = 0.928$
	$N = 11$

Note. Diffraction data from a total of 56 untilted and tilted his-MoCA crystalline images were corrected for contrast transfer function variations (Crowther *et al.,* 1996; McDermott *et al.,* 2000) and merged using ORIGTILTB and IQ ≤ 6 reflections (Crowther *et al.,* 1996; McDermott *et al.,* 2000), assuming hexagonal (*p*6) symmetry: merge phase residuals were determined to 30, 25, 20, and 15 Å. Amplitude plus phase vector data for lattice lines to 19 Å in the *h*, *k* plane were filtered to 30 Å in the *Z** direction to generate 3D *h*, *k*, *l* lattices for reconstructions. Completeness is defined as the percentage of the theoretical *h*, *k*, *l* lattice for which values were obtained. For vector addition analyses, the number of amplitude and phase vectors (multiplicity) averaged to contribute to the same *h*, *k*, *l* bin is given in the left column, while *N* indicates the number of times a given multiplicity was encountered in the *h*, *k*, *l* reflection sets. The *Q* value is defined as the vector sum for an *h*, *k*, *l* bin divided by the scalar sum of measurements contributing to that bin and provides an evaluation of phase agreement: the maximum value is 1, while a random value equals 1 divided by the square root of the number of vectors added (multiplicity). Note that maximum and random *Q* values are equal when the multiplicity is 1, but when two or more vectors are added, random *Q* values decrease to <0.71.

bors (around the hexagonally coordinated cage holes). When density maps were tilted 90°, so that the view was parallel to that of the membrane, front- and back-clipping permitted observation of individual dimer units (Fig. 3E). As anticipated, the heights of the units were equal (50 Å), as were the membrane projection plane lengths (50 Å) and widths (20 Å). These dimensions are consistent with those of CA NTD dimers from several different retroviruses (Berthet-Colominas *et al.,* 1999; Campos-Olivas *et al.,* 2000; Gamble *et al.,* 1996; Gitti *et al.,* 1996; Jin *et al.,* 1999; Khorasanizadeh *et al.,* 1999; Momany *et al.,* 1996), supporting the notion that his-MoCA 2D crystals employ NTD dimer building blocks.

The appearance of dimers in hexagonal reconstructions (Figs. 2 and 3) agrees with our previous reconstructions, using an optimized primitive space group lattice (McDermott *et al.,* 2000). However, the fitting of our new data set with the *p*6 space group has generated a new density map, which is informative with regard to the organization of retrovirus Gag proteins. The hexagonal model of membrane-bound his-MoCA proteins (Fig. 4) requires a total of three distinct monomer association domains and employs symmetric head-to-head dimers. Assuming that Gag matrix domains align above each CA

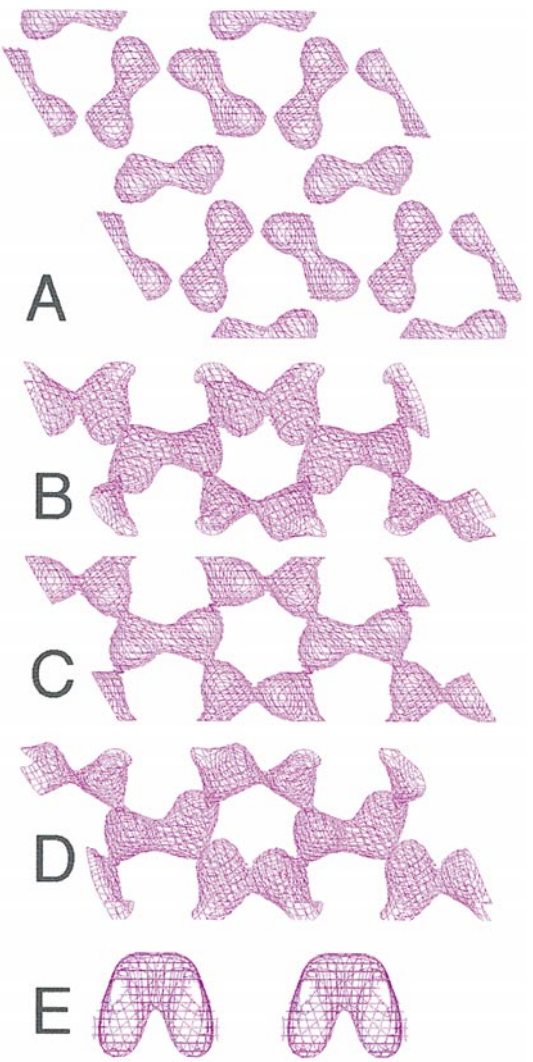

FIG. 3. Three-dimensional reconstruction of membrane-bound M-MuLV capsid proteins. Amplitude and phase data from 56 untilted and tilted images of his-MoCA proteins assembled on membranes were merged assuming hexagonal symmetry to generate 3D electron density maps at a resolution of 19 Å parallel to the membrane and 30 Å perpendicular to the membrane. Unit cells were $a = b = 80.3$ Å, $c = 110$ Å, $\gamma = 120°$; A shows 2 × 2 unit cells and is scaled at $\sigma = 3$, while other panels are size-scaled similarly, but they show 2 × 1 unit cells and are contoured at $\sigma = 2.5$. Unit cell orientations are as follows: A is viewed perpendicular to the membrane from the membrane side ($X = 1.000, 0.000, 0.000$; $Y = 0.000, 1.000, 0.000$; $Z = 0.000, 0.000, 1.000$); B is viewed tilted 20° up from A ($X = 1.000, 0.000, 0.000$, $Y = 0.000, 0.9397, 0.3420$; $Z = 0.000, -0.3420, 0.9397$); C is oriented similarly to A; D is viewed tilted 20° down from A ($X = 1.000, 0.000, 0.000$; $Y = 0.000, 0.9397, -0.3420$; $Z = 0.000, 0.3420, 0.9397$); and E is viewed in the plane of the membrane (membrane side up: $X = 1.000, 0.000, 0.000$; $Y = 0.000, 0.000, 1.000$; $Z = 0.000, -1.000, 0.000$) and is back-clipped to −9.7 Å and front-clipped to +9.7 Å to show two dimers per panel.

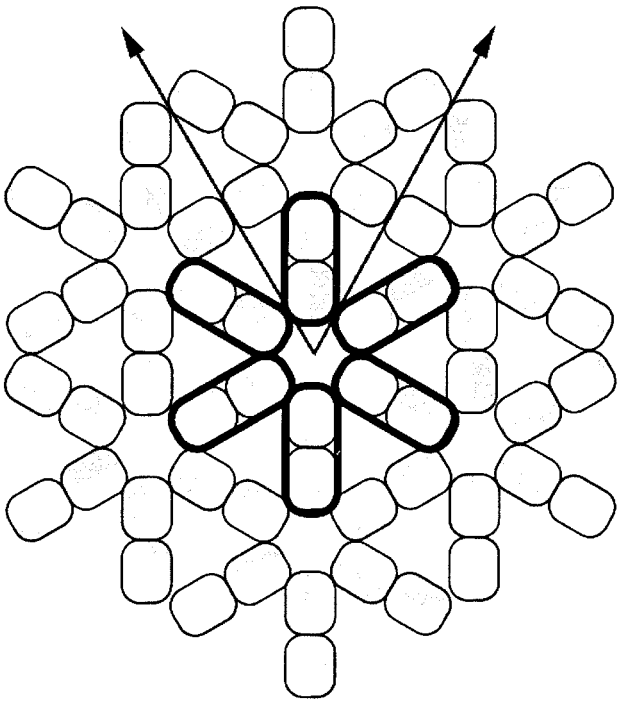

FIG. 4. Hexagonal model for capsid protein arrangements on membranes. A hexagonal model for M-MuLV capsid protein arrangements on membranes is diagrammed, depicting CA monomers as rounded, two-toned rectangles and outlining six dimer units. The hexagonal model utilizes head-to-head, presumably NTD-to-NTD dimer building blocks, but requires two additional association regions per monomer to tesselate into the hexagonal lattice. Assuming that Gag matrix domains align above capsid domains, envelope protein (SU/TM) trimers could be coordinated in triangular cage holes or with two degrees of freedom in hexagonal cage holes. Putative fivefold symmetry axes required for icosahedral symmetry could be generated from this model by removal of a 60° wedge (arrowed lines) and forming connections between units at the opposite edges of the removed wedge.

unit, they could coordinate trimeric (Fass *et al.,* 1996) envelope (Env; SU-TM) protein complexes in a fixed fashion via triangular cage holes or with two degrees of freedom via hexameric cage holes. Furthermore, removal of 60° wedges (Fig. 4, arrowed lines) from hexameric centers and connecting the remaining contacts makes it possible to induce curvature by formation of fivefold symmetry axes and is consistent with the assembly of strict icosahedra or other closed structures.

How does the membrane-bound his-MoCA structure relate to Gag protein assemblies in virions? Because we have substituted the his-tag-nickel chelating lipid interaction for the natural precursor Gag (PrGag) protein membrane anchor, we believe that the his-MoCA organization represents that observed for PrGag proteins in immature virus particles. In this regard, it is noteworthy that the 45-Å nearest neighbor hole-to-hole spacing corresponds well to the major Fourier spacing observed with immature M-MuLV particles (Yeager *et al.,* 1998). Furthermore, we previously observed that major hole-to-hole spacing projection structures of membrane-bound HIV CA proteins were similar to those reported for immature HIV particles (Nermut *et al.,* 1994, 1998), supporting the notion that the approach can mimic *in vivo* assemblies faithfully. Although the his-MoCA proteins do not retain the M-MuLV nucleocapsid domain, which is needed for efficient *in vivo* particle assembly (Barklis *et al.,* 1997; Bowzard *et al.,* 1998), we previously showed that his-tagged capsid plus nucleocapsid (his-MoCANC) proteins yielded projection structures similar to those of their his-MoCA counterparts (Zuber *et al.,* 2000). Moreover, membrane-bound his-MoCANC protein lattices were closely related to his-MoCANC plus RNA tube lattices assembled in the absence of membranes, suggesting that the two forms organize in roughly equivalent fashions (Zuber *et al.,* 2000). The hexagonal model (Fig. 4) compares interestingly with the recently reported model of HIV CA proteins assembled *in vitro* into helical tubes (Li *et al.,* 2000). In each case, hexagonally coordinated cage holes are surrounded by putative NTD monomers (Figs. 2 and 3; Li *et al.,* 2000). However, while M-MuLV hexamer holes appear connected via putative NTD dimer associations at distances of 80 Å, HIV hexamer holes appear coordinated by NTD monomers which connect through CTD interactions with hole-to-hole distances of 107 Å (Li *et al.,* 2000). This difference may arise from naturally different organizational patterns of the two proteins or may reflect the fact that the arrangements represent immature versus mature particle forms. Alternatively, assembly patterns may be affected by the presence of the N-terminal his-tag in our studies or differences in assembly techniques or incubation conditions. We believe that analysis of additional retrovirus Gag assemblies at higher resolution will help reveal the general principles of how these proteins organize into virus structures.

MATERIALS AND METHODS

Protein crystallization, microscopy, and image collection

The his-tagged M-MuLV capsid (his-MoCA) proteins were produced from the bacterial expression plasmid pET15B-MoCA, which encodes the M-MuLV CA domain with an amino-terminal his-tag (MGSSH HHHHH SSGLV APRGS HMLGD) and a plasmid-derived carboxy-terminal tail (ADPAA NKARK EAELA AATAEQ). Proteins were expressed in *Escherichia coli* strain BL21(DE3)/LysS, purified to >90% homogeneity by one or two rounds of nickel-chelate chromatography, and stored in aliquots at 0.2–2.0 mg/ml in water at −80°C prior to use (Barklis *et al.,* 1997; McDermott *et al.,* 2000). Crystallization incubations were for 16–18 h in humidified petri plates at room temperature or 30°C and consisted of 10 μl of 0.1–1.0 mg/ml protein in 1× subphase buffer (25 mM sodium phosphate, pH 7.8; 250 mM NaCl; 10 mM imidazole; 5 mM sodium acetate, pH 7.6; 20% glycerol) beneath 1 μl of

200 μg/ml phosphatidyl choline (PC; Avanti Polar Lipids) plus 50 μg/ml of the nickel-chelating lipids DHGN (Barklis *et al.,* 1997) or DOGS (McDermott *et al.,* 2000). After incubations plus optional 0.5- to 1.5-h 4°C postincubations, lipid monolayers and associated proteins were lifted onto lacy EM grids, washed for 30–45 s on 50- to 100-μl water drops, wicked from the side, stained for 30 s on drops of 1.33% uranyl acetate, wicked, and dried.

Membrane-bound protein arrays were photographed and digitized at 37,000–70,000× under low-dose conditions on a Philips CM120/Biotwin TEM, equipped with a Gatan 1024 × 1024 pixel multiscan 794 CCD camera. Images were converted to TIFF and then MRC (Crowther *et al.,* 1996) format, and crystalline areas were viewed and boxed using the MRC program BOXMRC (Crowther *et al.,* 1996).

Transformation, indexing, and projection analysis

Diffraction patterns were calculated from boxed crystalline images using the MRC FFTRANS program (Crowther *et al.,* 1996) and were visualized and indexed using the ICE and SPECTRA MRC program interfaces (Barklis *et al.,* 1997, 1998). The membrane-bound his-MoCA crystals were conveniently indexed in a hexagonal ($a^* = b^*$, $\gamma^* = 60°$) fashion, and amplitudes and phases for indexed reflections were obtained as APH files by sequential use of the LATREF, MMBOX, and UNBEND programs (Amos *et al.,* 1982; Barklis *et al.,* 1997; Crowther *et al.,* 1996). To correct for changes in the contrast transfer function (CTF; Amos *et al.,* 1982) in APH files, diffraction pattern Thon rings were used in conjunction with the program CTF_DETERMINE (Barklis *et al.,* 1997; McDermott *et al.,* 2000) to calculate image defocus parameters.

Generation of Fourier filtered 2D projections from CTF-corrected APH files was accomplished as described previously (Barklis *et al.,* 1997; McDermott *et al.,* 2000), by sequential application of the programs CREATE_TNF, FFTRANS, ICE_SKEW, and BOXMRC. Space group analysis of untilted APH files employed the program ALLSPACE (Crowther *et al.,* 1996), which evaluates the consistency of symmetry-related reflections for all possible 2D space groups. The output of ALLSPACE includes a phase residual value for each space group, which indicates goodness of fit (a 0° phase residual is a perfect fit; a 90° value is a random fit). Also included in the ALLSPACE output are optimized phase origin center locations for each space group, which were used to create untilted reference APH files for 3D data merging.

Merging of 2D data sets and 3D reconstruction

For merging of 3D data sets, the zero tilt reference image 121120 was centered on the best hexagonal ($p6$) phase origin (determined from ALLSPACE; see above), which also was the best phase origin for the $p2$ and $p3$ space groups. References for the primitive ($p1$ and p1opt; McDermott *et al.,* 2000) and $p2$, $p3$ and $p6$ merges were generated by vector averaging of related reflections for the respective asymmetric $p1$, $p2$, $p3$, and $p6$ units. Separate $p1$, $p2$, $p3$, and $p6$ merges of the same 56 untilted and tilted APH files were performed using a Z^* window of 0.002 $Å^{-1}$, IQ ≤ 6 reflections, and ORIGTILTB, which computes amplitude and phase values along h, k, Z^* lattice lines (Crowther *et al.,* 1996; McDermott *et al.,* 2000). As described previously (McDermott *et al.,* 2000), we also performed an optimized $p1$ (p1opt) merge, which is the same as the $p1$ data merge, except that the rotational alignment of each APH file is optimized prior to merging.

To obtain 3D maps, Z^* values from ORIGTILTB merges were binned using a Z thickness of 600 Å to obtain l index values, and reflections from each bin were vector averaged to give single amplitude and phase values for each h, k, l index in HKL files (McDermott *et al.,* 2000). The HKL files then were filtered to a resolution of 30 Å in l and either 19 Å (Table 2, Fig. 3) or 20 Å (Table 1, Fig. 2) in the h, k plane. Three-dimensional back-transformations and volume renderings were performed using the XTALVIEW suite as described previously (McDermott *et al.,* 2000): HKL files were converted to PHS files, back-transformed to MAP files with XFFT, and viewed with XFIT. Density contour levels of 2–3 standard deviations (σ) were used for viewing, and orientation matrices are provided in the figure legends.

Statistical evaluation of 3D reconstructions used several criteria. Phase residuals for overall merges were obtained from ORIGTILT output and provide a measure of the extent to which phase data from different images match (0°, perfect match; 90°, random match). Phase residuals also were averaged for images collected at different tilt angles to evaluate the quality of high-tilt data (Table 1) and in resolution rings, for resolution determinations (Table 2). Another measure used for analysis of phase agreement in 3D reconstructions was the Q value, which is defined as the vector sum for an h, k, l bin divided by the scalar sum of measurements contributing to that bin (McDermott *et al.,* 2000). The maximum possible Q value is 1, which is achieved when there is perfect phase agreement or when only one vector contributes to a bin. A random Q value equals one divided by the square root of the number of vectors added (multiplicity; m), so when $m > 1$, an evaluation of phase agreement can be made by comparing the Q value to $m^{-1/2}$. For the hexagonal ($p6$) 3D reconstruction, Q values and the completeness (percentage of the theoretical h, k, l lattice for which values were obtained) are given in Table 2.

ACKNOWLEDGMENTS

We are grateful to Ayna Alfadhli and Eric Steel for help in protein purification and to Doug Huseby, Matt Foster, and Liam Finlay for EM

and image reconstruction assistance. We thank Michael F. Schmid for his continued patient advice and appreciate the grant support (GM 52914 and GM 60170) of the National Institute of General Medical Sciences (NIGMS).

REFERENCES

Aggresta, B., and Carter, C. (1997). Cyclophilin A-induced alterations of human immunodeficiency virus type 1 CA protein in vitro. *J. Virol.* **71,** 6921–6927.

Amos, L. A., Henderson, R. H., and Unwin, P. N. T. (1982). Three-dimensional structure determination by electron microscopy of two-dimensional crystals. *Prog. Biophys. Mol. Biol.* **39,** 183–231.

Barklis, E., McDermott, J., Wilkens, S., Fuller, S., and Thompson, D. (1998). Organization of HIV-1 capsid proteins on a lipid monolayer. *J. Biol. Chem.* **273,** 7177–7180.

Barklis, E., McDermott, J., Wilkens, S., Schabtach, E., Schmid, M. F., Fuller, S., Karanjia, S., Love, Z., Jones, R., Rui, Y., Zhao, Z., and Thompson, D. (1997). Structural analysis of membrane-bound retrovirus capsid proteins. *EMBO J.* **16,** 1199–1213.

Barth, M., Bryan, R. K., Hegerl, R., and Baumeister, W. (1988). Estimation of missing cone data in three-dimensional electron microscopy. *Scanning Microsc. Suppl.* **2,** 277–284.

Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F., and Cusack, S. (1999). Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO J.* **18,** 1124–1136.

Bolognesi, D. P., Montelaro, R., Frank, H., and Schafer, W. (1978). Assembly of type C oncornaviruses: A model. *Science* **199,** 183–186.

Bowzard, J., Bennett, R., Krishna, N., Ernst, S., Rein, A., and Wills, J. (1998). Importance of basic residues in the nucleocapsid sequence for retrovirus Gag assembly and complementation rescue. *J. Virol.* **72,** 9034–9044.

Campbell, S., and Rein, A. (1999). *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J. Virol.* **73,** 2270–2279.

Campbell, S., and Vogt, V. M. (1995). Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol.* **69,** 6487–6497.

Campbell, S., and Vogt, V. M. (1997). *In vitro* assembly of virus-like particles with Rous sarcoma virus Gag deletion mutants: Identification of the p10 domain as a morphological determinant in the formation of spherical particles. *J. Virol.* **71,** 4425–4435.

Campos-Olivas, R., Newman, J. L., and Summers, M. F. (2000). Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. *J. Mol. Biol.* **296,** 633–649.

Conte, M., Klikova, M., Hunter, E., Ruml, T., and Matthews, S. (1997). The three-dimensional solution structure of the matrix protein from the type D retrovirus, the Mason–Pfizer monkey virus, and implications for the morphology of retroviral assembly. *EMBO J.* **16,** 5819–5826.

Crowther, R., Henderson, R., and Smith, J. (1996). MRC image processing programs. *J. Struct. Biol.* **116,** 9–16.

Ehrlich, L. S., Agresta, B. E., and Carter, C. A. (1992). Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro*. *J. Virol.* **66,** 4874–4883.

Fass, D., Harrison, S., and Kim, P. S. (1996). Retrovirus envelope domain at 1.7 angstrom resolution. *Nat. Struct. Biol.* **3,** 465–469.

Fuller, S. D., Wilk, T., Gowen, B. E., Krausslich, H. G., and Vogt, V. M. (1997). Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.* **7**(10), 729–738.

Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **87,** 1285–1294.

Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **278,** 849–853.

Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999). Assembly and analysis of conical models for the HIV-1 core. *Science* **283,** 80–83.

Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273,** 231–235.

Grattinger, M., Hohenberg, H., Thomas, D., Wilk, T., Muller, B., and Krausslich, H. (1999). *In vitro* assembly properties of wild-type and cyclophilin-binding defective human immunodeficiency virus capsid proteins in the presence and absence of cyclophilin A. *Virology* **257,** 247–260.

Gross, I., Hohenberg, H., Huckhagel, C., and Krausslich, H.-G. (1998). N-terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. *J. Virol.* **72,** 4798–4810.

Gross, I., Hohenberg, H., Wilk, T., Wiegers, K., Grattinger, M., Muller, B., Fuller, S., and Krausslich, H. G. (2000). A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* **19,** 103–113.

Jin, Z., Jin, L., Peterson, D. L., and Lawson, C. L. (1999). Model for lentivirus capsid core assembly based on crystal dimers of EIAV p26. *J. Mol. Biol.* **286,** 83–93.

Khorasanizadeh, S., Campos-Olivas, R., and Summers, M. F. (1999). Solution structure of the capsid protein from the human T-cell leukemia virus type-1. *J. Mol. Biol.* **291,** 491–505.

Kingston, R., Fitzon-Ostendorp, T., Eisenmesser, E., Schatz, G., Vogt, V., Post, C., and Rossmann, M. G. (2000). Structure and self-association of the Rous sarcoma virus capsid protein. *Structure* **8,** 617–628.

Klikova, M., Rhee, S., Hunter, E., and Ruml, T. (1995). Efficient in vivo and in vitro assembly of retroviral capsids from Gag precursor proteins expressed in bacteria. *J. Virol.* **69,** 1093–1098.

Li, S., Hill, C., Sundquist, W., and Finch, J. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407,** 409–413.

McDermott, J., Mayo, K., and Barklis, E. (2000). Three-dimensional organization of retroviral capsid proteins on a lipid monolayer. *J. Mol. Biol.* **302,** 121–133.

McDonnell, J., Fushman, D., Cahill, S., Zhou, W., Wolven, A., Wilson, C., Nelle, T., Resh, M., Wills, J., and Cowburn, D. (1998). Solution structure and dynamics of the bioactive retroviral M domain from Rous sarcoma virus. *J. Mol. Biol.* **279,** 921–928.

McRee, D. E. (1992). A visual protein crystallographic software system for X11/XView. *J. Mol. Graphics* **10,** 44–46.

Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Erlich, L. S., Summers, M. F., Carter, C., and Rossmann, M. G. (1996). Crystal structure of dimeric HIV-1 capsid protein. *Nat. Struct. Biol.* **3,** 763–770.

Nermut, M. V., Hockley, D. J., Bron, P., Thomas, D., Zhang, W.-H., and Jones, I. M. (1998). Further evidence for hexagonal organization of HIV gag protein in prebudding assemblies and immature virus-like particles. *J. Struct. Biol.* **123,** 143–149.

Nermut, M. V., Hockley, D. J., Jowett, J. B. M., Jones, I. M., Garreau, M., and Thomas, D. (1994). Fullerene-like organization of HIV gag-protein shell in virus-like particles produced by recombinant baculovirus. *Virology* **198,** 288–296.

Reil, H., Bukovsky, A., Gelderblom, H., and Gottlinger, H. (1998). Efficient HIV-1 replication can occur in the absence of the viral matrix protein. *EMBO J.* **17,** 2699–2708.

Scarlata, S., Ehrlich, L., and Carter, C. (1998). Membrane-induced alterations in HIV-1 Gag and matrix protein–protein interactions. *J. Mol. Biol.* **277,** 161–169.

Swanstrom, R., and Wills, J. W. (1997). Synthesis, assembly, and processing of viral proteins. *In* "Retroviruses" (J. M. Coffin, S. H. Hughes, and H. E. Varmus, Eds.), pp. 263–334. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

von Schwedler, U. K., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H.,

Davis, D. R., and Sundquist, W. I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17,** 1555–1568.

Wang, C., Zhang, Y., McDermott, J., and Barklis, E. (1993). Conditional infectivity of a human immunodeficiency virus matrix domain deletion mutant. *J. Virol.* **67,** 7067–7076.

Yeager, M., Wilson-Kubalek, E. M., Weiner, S. G., Brown, P. O., and Rein, A. (1998). Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: Implications for retroviral assembly mechanisms. *Proc. Natl. Acad. Sci. USA* **95**(13), 7299–7304.

Zuber, G., and Barklis, E. (2000). Atomic force microscopy and electron microscopy analysis of retrovirus Gag proteins assembled in vitro on lipid bilayers. *Biophys. J.* **78,** 373–384.

Zuber, G., McDermott, J., Karanjia, S., Zhao, W., Schmid, M. F., and Barklis, E. (2000). Assembly of retrovirus capsid–nucleocapsid proteins in the presence of membranes or RNA. *J. Virol.* **74,** 7431–7441.