doi:10.1006/jmbi.2000.4030 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2000) 302, 121–133

JMB

Three-dimensional Organization of Retroviral Capsid Proteins on a Lipid Monolayer

Jason McDermott, Keith Mayo and Eric Barklis*

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

We have used a method for the two-dimensional crystallization of retroviral structural proteins to obtain a three-dimensional structure of negatively stained, membrane-bound, histidine-tagged Moloney murine leukemia virus (M-MuLV) capsid protein (his-MoCA) arrays. Tilted and untilted micrographs from crystals formed by purified his-MoCA proteins incubated beneath lipid monolayers containing nickel-chelating lipids were used in 3D reconstructions. The 2D crystals had unit cell dimensions of $a = 72.6$ Å, $b = 72.5$ Å and $\gamma = 119.5°$, but appeared to have no intrinsic symmetry ($p1$) in 3D, in contrast to the trigonal or hexagonal appearance of their 2D projections. Membrane-bound his-MoCA proteins showed a strand-like organization, apparently with dimer building blocks. Membrane-proximal regions, or putative N-terminal domains (NTDs), dimerized with different partners than the membrane-distal putative C-terminal domains (CTDs). Evidence also suggests that CTDs can adopt alternate orientations relative to their NTDs, forming interstrand connections. Our results are consistent with helical-spiral models for retrovirus particle assembly, but are not easily reconcilable with icosahedral models.

*Corresponding author

Keywords: retrovirus; capsid; Gag; membrane; electron microscopy

Introduction

Retrovirus assembly is directed by the polyprotein precursor protein, PrGag. For mammalian retroviruses, PrGag consists of matrix (MA), capsid (CA), and nucleocapsid (NC) domains, plus other less well conserved domains, such as $p12$ in the case of Moloney murine leukemia virus (M-MuLV), and $p6$ in human immunodeficiency virus type 1 (HIV-1). There are two forms of retrovirus particles: immature, in which the PrGag proteins are localized adjacent to the membrane envelope; and mature, where CA and NC proteins have been released by proteolytic cleavage into a centralized ribonucleoprotein structure (Swanstrom & Wills, 1997). Although deletion or mutation of certain portions of CA have been shown to be compatible with the cellular release of virus-like particles (Wang & Barklis, 1993; Weldon & Wills, 1993; Reicin *et al.*, 1995; Borsetti *et al.*, 1998; Wang *et al.*, 1998), interprotein contacts mediated by CA appear crucial to the assembly of infectious viruses (Hsu *et al.*, 1985; Goff & Lobel, 1987; Strambio-de-Castillia & Hunter, 1992; Wang & Barklis, 1993; Mammano *et al.*, 1994; Craven *et al.*, 1995; Srinivasakumar *et al.*, 1995; Alin & Goff, 1996; McDermott *et al.*, 1996; Zhang *et al.*, 1996). High-resolution structures for CA have been determined by X-ray or NMR methods for Rous sarcoma virus (RSV; Campos-Olivas *et al.*, 2000), equine infectious anemia virus (EIAV; Jin *et al.*, 1999), human T-cell leukemia virus (HTLV; Khorasanizadeh *et al.*, 1999), and HIV-1 (Gamble *et al.*, 1996, 1997; Gitti *et al.*, 1996; Momany *et al.*, 1996; Berthet-Colominas *et al.*, 1999). Despite limited sequence homology, each of these structures displays a two-domain organization with the independently folding domains joined by a flexible linker region. The N-terminal domains (NTDs) are composed of seven α-helices in a bundle-like arrangement, and the C-terminal domains (CTDs) adopt a globular organization formed by four α-helices. In each of these structures the orientation of the CTD relative to the NTD appears to be flexible (Berthet-Colominas

Abbreviations used: M-MuLV, Moloney murine leukemia virus; NTD, N-terminal domains; CTD, C-terminal domains; MA, matrix; CA, capsid; NC, nucleocapsid; RSV, Rous sarcoma virus; EIAV, equine infectious anemia virus; HTLV, human T-cell leukemia virus; MHR, major homology region; EM, electron microscopy.

E-mail address of the corresponding author: barklis@ohsu.edu

0022-2836/00/010121–13 $35.00/0

et al., 1999; Jin *et al*., 1999; Khorasanizadeh *et al*., 1999; Campos-Olivas *et al*., 2000). Despite the global structural conservation of capsid proteins, CA contains only one stretch of residues that are highly conserved among retroviruses, the major homology region (MHR; Strambio-de-Castillia & Hunter, 1992; Mammano *et al*., 1994; Craven *et al*., 1995; Swanstrom & Wills, 1997). The MHR has been shown to form an α-helix which overlaps the linker region and the CTD, and stabilizes the structure of the CTD by forming hydrogen bonds with other CTD α-helices.

High-resolution structures derived from X-ray and NMR studies provide a great deal of information about the structure of these proteins but less about how the proteins associate to form retrovirus particles. However, electron microscopy (EM) techniques have been used effectively to investigate retrovirus structure. Studies performed on whole immature and mature virus particles of HIV-1 (Nermut *et al*., 1994, 1998; Fuller *et al*., 1997) and M-MuLV (Yeager *et al*., 1998) have shown that PrGag proteins form cage-like structures with hole-to-hole spacings of 66-70 Å for HIV-1, and Fourier spacings of 45 Å for M-MuLV. Studies also have been performed on the *in vitro* assembly products of CA proteins (Ehrlich *et al*., 1992; Gross *et al*., 1998; Schwedler *et al*., 1998) as well as CA-NC fusion proteins (Campbell & Vogt, 1995, 1997; Campbell & Rein, 1999; Ganser *et al*., 1999; Zuber *et al*., 2000) in the absence of membranes. Such studies have shown that the capsid domain is capable of directing the assembly of rod-like or spherical structures *in vitro*, depending on the presence of both N and C-terminal extensions to the CA domain as well as buffer conditions. Since lentiviruses, such as HIV-1, as well as type-C retroviruses, such as M-MuLV, assemble at the plasma membranes of host cells, it is of interest to examine the structures formed by Gag proteins assembled on membranes. To address this problem, we have developed a technique to crystallize histidine-tagged (his-tagged) retroviral PrGag proteins on lipid monolayers containing nickel-chelating lipids to mimic immature virus particle assembly. Previously we obtained two-dimensional projection structures of membrane-bound HIV-1 (Barklis *et al*., 1998) and M-MuLV (Barklis *et al*., 1997) CA proteins, which formed cage-like structures similar in appearance and dimensions as that of PrGag proteins assembled *in vivo* into immature virus particles (Nermut *et al*., 1994, 1998; Fuller *et al*., 1997; Yeager *et al*., 1998). We have extended these studies by obtaining a three-dimensional structure of 2D his-tagged M-MuLV capsid protein (his-MoCA) crystals formed on a lipid monolayer. Interestingly, the hexagonal cage appearance of 2D projections is not a consequence of a hexagonal 3D arrangement of proteins. Rather, his-MoCA proteins organized on membranes as strands of proteins, apparently from dimer building blocks. In our structure, membrane-proximal regions, presumably NTDs, formed dimers with different neighbors than membrane-distal, putative CTD dimers. Indeed, CTDs apparently adopted three alternate orientations relative to their respective NTDs. These results suggest that retroviral CA MHR and spacer regions may be well conserved because they have to satisfy several alternate conformation constraints. Our observations also are consistent with helical-spiral models for retrovirus particle assembly (Campbell & Vogt, 1995), but are not immediately compatible with icosahedral models of retrovirus structure.

Results

Projection structure of membrane-bound M-MuLV capsid proteins

To avoid limitations inherent in the analysis of retroviral particles (see the Introduction), we adapted the lipid monolayer method of 2D protein crystallization (Uzgiris & Kornberg, 1983) to analyze interactions between membrane-bound M-MuLV Gag proteins. The method employs lipids with nickel-binding head groups (DHGN or DOGS; Barklis *et al*., 1997; Wilson-Kubalek *et al*., 1998), which together with phosphatidylcholine (PC), are used to form monolayer membranes on which his-tagged, membrane-associated proteins can crystallize. To investigate the interactions of Gag proteins, we employed an N-terminally his-tagged M-MuLV capsid protein (his-MoCA; Barklis *et al*., 1997). Association of the protein his-tag with the nickel-chelating lipid monolayer mimics the *in vivo* association of PrGag proteins and lipid bilayers (Swanstrom & Wills, 1997), and allows lateral and rotational movement during formation of crystalline arrays. These arrays were then lifted onto EM grids, stained, and examined by EM for the presence of crystals (Barklis *et al*., 1997, 1998). Crystalline areas on EM negatives then were scanned and digitized for image processing.

An example of a membrane-bound array of his-MoCA proteins is provided in Figure 1(a), which shows a 140.9 nm × 140.9 nm crystalline patch. The crystalline order is apparent in the corresponding diffraction pattern (power spectrum; Figure 1(b)). As observed previously, his-MoCA 2D crystals exhibit a hexagonal or trigonal appearance, and the barely visible innermost reflections give a unit cell size of 72.6 Å × 72.5 Å, with a gamma angle of 119.5° (Table 1). This unit cell size is slightly smaller than that previously observed with ice-embedded his-MoCA crystals (79.6 Å × 79.6 Å; Barklis *et al*., 1997), which may be due to shrinkage on drying in negative stain (Stoops *et al*., 1992). To obtain a 2D projection reconstruction of membrane-bound his-MoCA proteins, 61 untilted diffraction patterns were boxed, indexed, unbent, CTF-corrected and merged, assuming no symmetry constraints (*p*1). Amplitude and phase values for each reflection were vector averaged, and averaged values were used in back-transformations to yield the (Figure 1(c)) signal-enhanced 225 Å × 225 Å projection reconstruction,

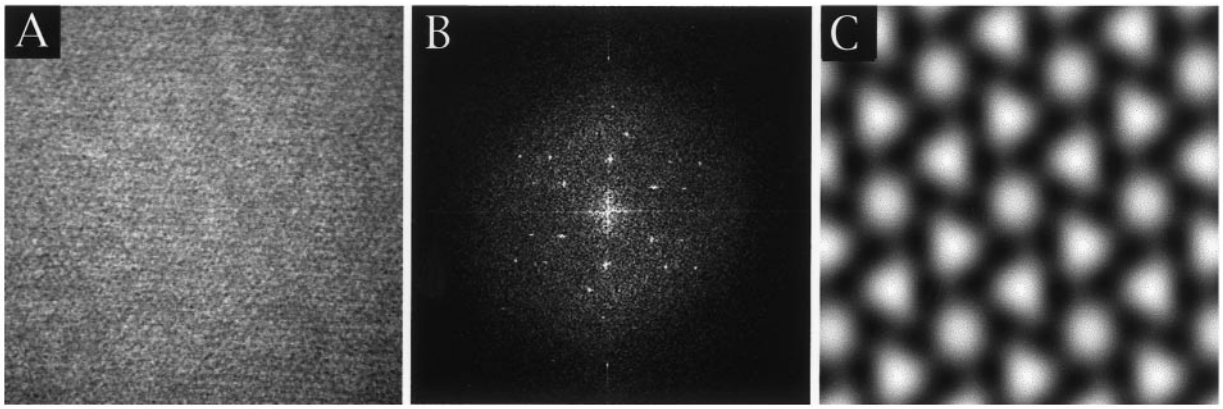

Figure 1. Projection structure of membrane-bound his-MoCA proteins. (a) Micrograph of a his-MoCA array formed on a PC + DHGN lipid monolayer, lifted onto an electron microscope grid and negatively stained with uranyl acetate as described in Materials and Methods. The image was taken at 52,000× magnification and represents a 140.9 nm × 140.9 nm patch. Protein areas appear light and protein-free areas appear dark. (b) The calculated diffraction pattern from the image in (a) is displayed as a power spectrum. Reflections were visible to 23 Å and were indexed with a γ^* angle of approximately 60°, giving a unit cell size of $a = 72.5$ Å, $b = 72.6$ Å. The brightest six reflections are at 0.0276 $Å^{-1}$. (c) A two-dimensional reconstruction of membrane-bound his-MoCA proteins is shown as a 224 Å × 224 Å panel in which electron dense (protein) regions are dark. Data merging was performed using 61 untilted images. Merging was performed to a resolution of 20 Å with no symmetry assumed (*p*1), using high signal to noise reflections ($IQ \leqslant 4$) for the phase origin search. The overall phase residual of the merge was 13.9°, and the amplitude and phase values for each reflection were vector-averaged and then back-transformed to give the projection reconstruction.

in which protein areas appear dark, and protein-free areas are white. As illustrated, the proteins formed a regular cage-like arrangement, consisting of circular and triangular protein-free holes, each surrounded by six electron-dense protein blobs. This arrangement is consistent with previous observations (Barklis *et al*., 1997). Additionally the 41.6 Å spacing between nearest neighbor holes corresponds well with the major 44-45 Å Fourier spacing observed in immature M-MuLV particles (Yeager *et al*., 1998), after correction for the unit cell size reduction observed in negatively stained *versus* cryo-embedded samples (see above).

Inspection of Figure 1(b) and (c) suggests that membrane-bound his-MoCA proteins formed crystals which may be compatible with hexagonal or trigonal space group assignments. Indeed, statistical analysis of diffraction patterns from 59 untilted images showed agreement (as indicated by low phase residuals) with trigonal (*p*3) and hexagonal (*p*6) space groups, in addition to the primitive (*p*1) space group. All other 2D crystal space groups gave significantly higher internal symmetry phase residuals.

Table 1. Untilted unit cell parameters

Parameter	Measurement
A. *Dimensions*	
a	72.6 (± 1.7) Å
b	72.5 (± 1.7) Å
γ	119.5 (± 1.5)°
B. *Space group fit*	
Space group	Phase residual (deg.)
*p*1	21.6 ± 3.0
*p*3	16.7 ± 8.6
*p*6	18.9 ± 8.0

Diffraction patterns from untilted crystalline his-MoCA images were indexed, boxed, unbent, and CTF-corrected as described in Materials and Methods. Average real space unit cell dimensions were obtained from 62 patterns. For space group fitting, phases of reflections to 15 Å resolution from 59 diffraction patterns were compared for internal consistency *versus* all two-dimensional crystal space groups using the program ALLSPACE. The phase residual for the *p*1 space group is based on the signal-to-noise ratio of the observed amplitudes, since no internal phase comparisons are possible with no symmetry constraints. Phase residuals of 0° indicate perfect matching, while residuals of 90° are random. Note that all other space groups gave average phase residuals of >25°, and that *p*1, *p*3, and *p*6 values derive from an average of 28, 22, and 58 comparisons, respectively.

Space group assignment for membrane-bound his-MoCA crystals

As a prerequisite to 3D reconstruction from individual tilted and untilted images, it is necessary to align the diffraction patterns from crystals at different tilt angles to generate lattice lines which can be binned into a reciprocal space, three-dimensional lattice (Amos *et al*., 1992). We chose to do this with phase-centered references in all three indicated space groups, *p*1, *p*3, and *p*6, since merging with higher symmetry (*p*6) gives a greater sampling of the data, while merging with lower symmetry (*p*1) requires fewer assumptions about the data. How-

ever, *p*1 merging presented a special problem, since with crystals that have nearly identical *a* and *b* lengths, such as ours, it can be difficult to determine which set of possible *p*1 axes in one diffraction pattern correspond to the same *p*1 set in another. Arbitrarily choosing axes by eye for each image is equivalent to applying 6-fold symmetry to the merge. To circumvent this problem, all six axes sets for each image were statistically compared in merges, and sets which gave the lowest merge phase residuals were chosen as the *a* and *b* axes for their respective diffraction patterns (see Materials and Methods). A comparison of averaged phase residuals from such an optimization (Figure 2(a)) showed a high level of match for one set of axes for each image, with decreasing values for the other five potential axes sets. This result is compatible with *p*1 space group symmetry, but less clearly so with *p*3 or *p*6 symmetry, which would be expected to yield either three or six approximately equally good normalized phase match values. Other analysis also argued against *p*3 or *p*6 space group symmetry assignments for his-MoCA crystals. Comparison of phase residuals *versus* tilt angles in his-MoCA *p*1, *p*3, and *p*6 merges (Figure 2(b)) showed a clear distinction. In particular, non-zero tilt angles yielded rising phase residuals for *p*3 and *p*6 merges, but relatively constant *p*1 residuals (Figure 2(b)). Thus, the *p*3 and *p*6 space group assignments do not appear accordant with the his-MoCA crystals. The organization of the proteins viewed in a projection perpendicular to the membrane (Figure 1(c)) has a hexagonal appearance, but this apparent symmetry is not retained in three dimensions.

Three-dimensional organization of membrane-bound M-MuLV capsid proteins

Based on the results shown in Figure 2, amplitude and phase data from a total of 218 tilted and untilted his-MoCA crystalline images were merged as described in Materials and Methods to obtain amplitudes and phases along reciprocal space lattice lines. Phase residuals for the merge suggested that the data were good to about 20 Å resolution, and the phase residuals increased at higher resolution (Table 2). To make the three-dimensional reciprocal lattice, lattice line Z^* values were vector-averaged in 1/600 Å width bins. Because missing cone information (Fuller *et al.*, 1979; Barth *et al.*, 1988; Glaeser *et al.*, 1989) affected the completeness

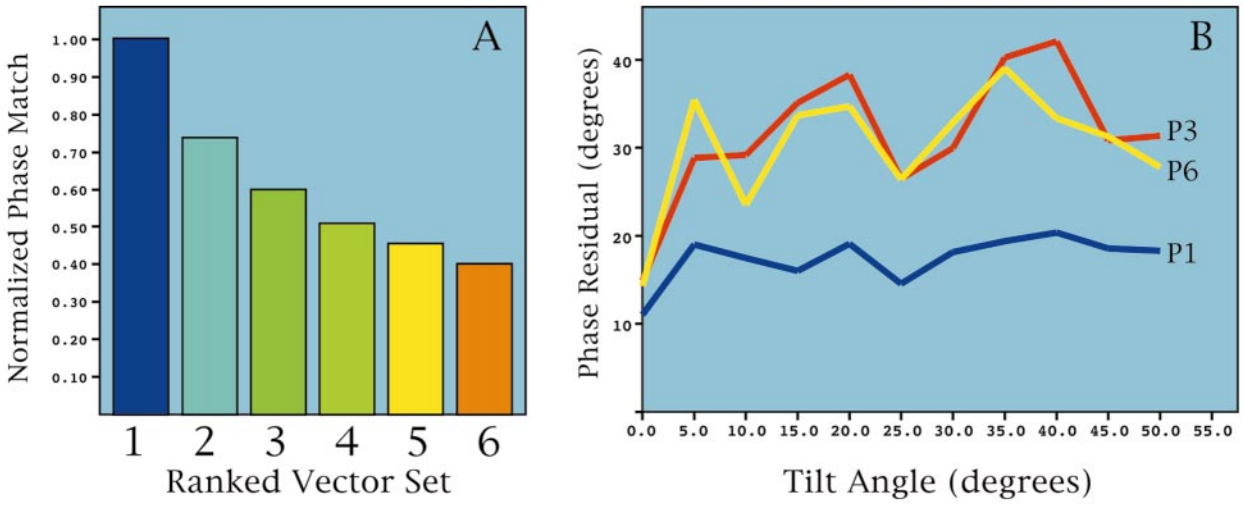

Figure 2. Space group assignment of his-MoCA crystals. (a) Analysis of the consequences of indexing diffraction patterns using the six possible $\gamma^* = 60^\circ$ axes sets. Each diffraction pattern was indexed in the six possible indexing schemes, and all indexing schemes for each image were tested in *p*1 merges to obtain phase residuals for each indexing scheme for each image. The six phase residuals obtained from indexing schemes for each image were ranked as follows: the indexing scheme yielding the best (lowest) phase residual was ranked as vector set 1, and given a normalized phase match score of 1.0. The other five indexing schemes received normalized phase match scores corresponding to the reciprocal of the ratio of their phase residual *versus* that of vector set 1, and were ranked accordingly. This operation resulted in the generation of six phase match ranked vector sets for each of the 232 images. For display purposes, all top, second, third, fourth, fifth, and sixth place vector sets were averaged. Ideally a strictly primitive (*p*1) diffraction pattern will show only one high phase match value; a trigonal (*p*3) pattern will show three high values, and all six vector sets should give equally high values for hexagonal (*p*6) diffraction patterns. (b) The full index-optimized MoCA dataset, consisting of 232 tilted and untilted images, was merged assuming *p*1, *p*3, and *p*6 space group symmetry as described in Materials and Methods. For each of the three merges, phase residual values from individual images were grouped according to their tilt angles and averaged. Results for *p*1 (blue), *p*3 (red), and *p*6 (yellow) merges were graphed to depict the relationship between tilt angle (abscissa) and phase residual (ordinate). Note that phase residuals of 0° indicate perfect agreement, while residuals of 90° indicate random correlation.

Table 2. Analysis of merged three-dimensional data

A. *Phase residuals versus resolution*

Resolution range (Å)	Overall phase residual (deg.)
>27.4	20.9
27.4-23.7	23.4
23.7-19.4	17.5
19.4-15.0	42.1

B. *Vector addition analysis*

Multiplicity	Frequency	*Q* value
1	117	1.00
2	84	0.922
3	41	0.918
4	26	0.912
5	20	0.931
6	20	0.937
7	11	0.927
8	6	0.950
9	10	0.971
10	11	0.921
11-20	18	0.942
21-100	15	0.911

Merging of three-dimensional data was performed on index-optimized amplitude and phase (APH) files from 218 images using ORIGTILTC, assuming no symmetry constraints (*p*1). For the purposes of phase residual *versus* resolution comparisons, merging was performed to 15 Å, using $IQ \leqslant 4$ reflections. Results were obtained from ORIGTILTC output: phase residuals of 0° indicate perfect matching, while residuals of 90° are random. For 3D reconstructions and vector addition analysis, merging was performed to 20 Å, using $IQ \leqslant 4$ reflections, and reflections of $IQ \leqslant 7$ were binned into l values of $Z^* = 1/600$ Å. Of 525 possible reflections, values were available for 365 (70% completeness). Multiplicity is defined as the number of amplitude plus phase vector measurements contributing to averaged amplitude plus phase values in the same *hkl* bin. Frequency represents the number of times a given multiplicity was encountered in the 365 reflection data set. The *Q* value is defined as the vector sum for an *hkl* bin divided by the scalar sum of measurements contributing to that bin. Thus the maximum *Q* value is 1.000, while a random *Q* value equals the reciprocal of the square-root of the number of vectors added (multiplicity). Note that while a *Q* value of 1.000 for multiplicity = 1 reflections is insignificant, random *Q* values for multiplicities of 2 to 10 drop from 0.707 to 0.316.

of the low *h*,*k*/high *l* reflections, reconstructions were filtered in the Z^* direction to 35 Å. The resulting reconstruction included 365 of 525 possible lattice points (70% completeness). To evaluate the consistency of the data, *Q* values for each reflection (Robinson *et al.*, 1988) were calculated (Table 2). The *Q* value corresponds to the vector-averaged amplitude divided by the scalar-averaged amplitude, and can vary between zero and 1. Although high *Q* values are meaningless for binned *h*,*k*,*l* values represented by single phase and amplitude measurements (a random *Q* value is inversely proportional to the square-root of the number of vectors added), they provided a gauge of data agreement when multiple measurements were compared. As shown in Table 2, observed *Q* values were high, even when over 100 vector measurements were compared, suggesting that reflections derived from different images agreed closely with each other.

Initial analysis of back-transformed 3D volumes showed that protein density centered in a 100 Å slab within the 72.5 Å × 72.5 Å *p*1 unit cell. For comparison with the 2D projection in Figure 1(c), examination of negative stain accumulation areas corresponding to protein-free holes in a view perpendicular to the membrane (Figure 3(a)) was helpful. Three different types of protein-free zones occurred in an apparent hexagonal or trigonal arrangement (Figure 3(a)), reminiscent of Figure 1(c). Looking at the protein areas from the same view gives a different impression (Figure 3(b)). Instead of hexamers surrounding the protein-free holes, the holes appear to be formed by staggered lines of C-shaped protein units. When viewed nearly parallel with the membrane (Figure 4(a)), the membrane-proximal domains of the C-shaped units from Figure 3(b), which we take to be his-MoCA NTDs, appear to form dimers (Figure 4(b)). The connections between the putative

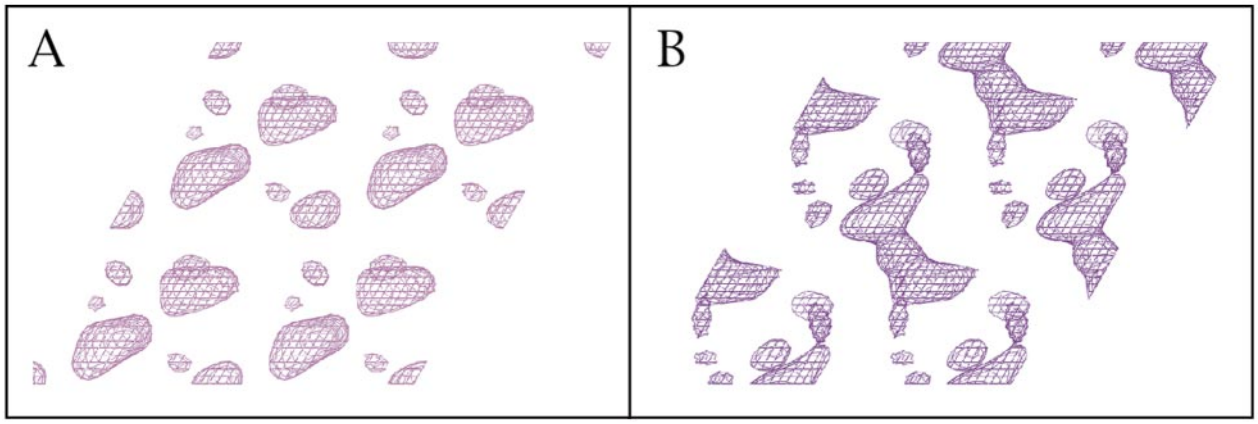

Figure 3. Membrane-bound his-MoCA protein reconstruction viewed perpendicular to the membrane. The his-MoCA dataset was merged into a *p*1 unit cell, binned, filtered, and back-transformed as described in Materials and Methods. Protein density was located in the central 100 Å thick *z*-axis slab, and volumes were viewed using XFIT. Shown in the Figure are approximately four unit cells ($a = b = 144$ Å) viewed along the *z*-axis, with the membrane face closest to the viewer, at a thickness of 100 Å. (a) The volume contour level was set at $\sigma = -1.0$ to show protein-free zones. (b) The contour level was set at $\sigma = 2.5$ to show protein regions.

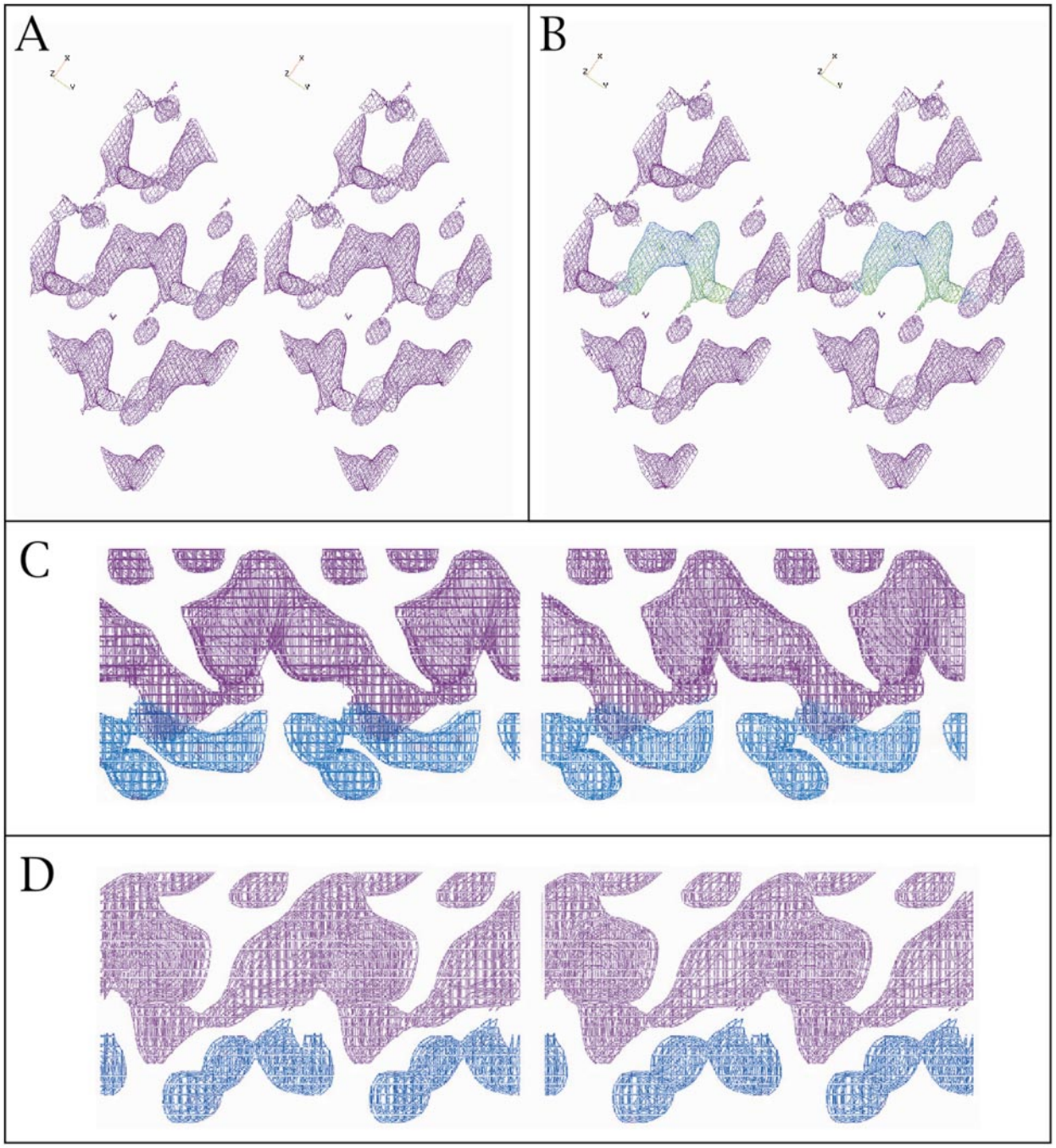

Figure 4. his-MoCA protein-proteins interactions. Shown is the 3D his-MoCA structure from the same 3D density map and using the same *a*,*b*,*c* dimensions as in Figure 3, viewed in stereo with $\sigma = 2.0$, at a slight angle to the *z*-axis, (a) by itself, and (b) with the members of the putative N-terminal dimer (NTD) outlined. (c) Structure is viewed in stereo along the *x*-axis, at a $\sigma = 2.0$, with the membrane face towards the top of the page; (d) is viewed similarly but rotated 60° clockwise around the *z*-axis. (c) and (d) Additional membrane-distal cross-strand densities, rendered in blue.

NTD dimers occur in membrane-distal regions, which appear to represent CTDs (Figure 4(a), (b)). The overall look of the structure is that of chains of CA proteins, alternately linked by putative NTD and CTD dimers (Figure 4(a)-(d)).

At the density levels used in Figure 4, membrane-bound molecules appeared to form linear chains, but there also are an unaccounted pair of peanut-shaped densities in the membrane-distal zone (Figure 4(c), (d)). Viewed from an angle slightly off of perpendicular to the membrane, where the linear strings of proteins are oriented lower-left to upper-right, interconnections between strands mediated by these membrane-

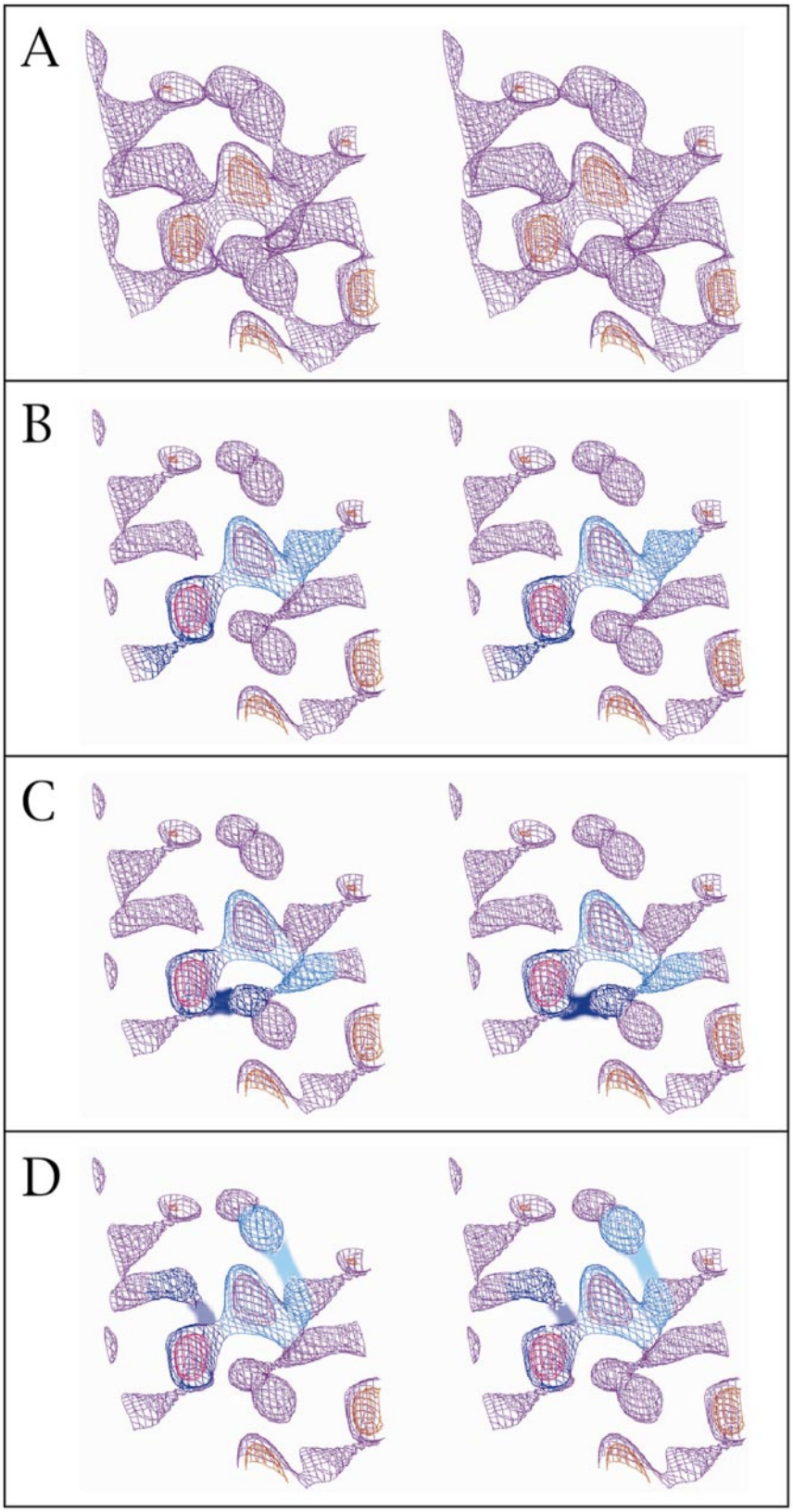

Figure 5. Alternate conformations for the his-MoCA C-terminal domain (CTD). The structure of the membrane-distal portion of the map described in Figure 3 is shown so as to facilitate putative CTD viewing. In all stereo panels, *a* and *b* dimensions were 100 Å, and thicknesses were 50 Å. (a) Alternate CTD interactions shown with the red contour level at $\sigma = 3.2$ and the purple contour level at $\sigma = 1.5$. (b)-(d) Alternate CTD interactions shown with red contour levels at $\sigma = 3.2$, and purple contour levels at $\sigma = 2.0$. Putative N-terminal domain (NTD) dimers are shown with alternate CTD locations outlined in dark and light blue to depict (b) intrastrand and (c) and (d) cross-strand connections.

distal regions can be seen, using two different density levels (Figure 5). Assuming that membrane-proximal densities are NTDs and membrane distal densities represent CTDs, there appear to be three alternate CTD dimer linkages between NTDs. One set of putative CTD dimers

appears to link NTDs within long strands (Figure 5(b)); one set appears to connect neighbor strands in one direction (Figure 5(c)); and one set appears to connect neighbor strands in the other direction (Figure 5(d)). Assuming our interpretation is correct, at least six types of CA monomers (Figure 5(b)-(d)) appear to occupy positions in the crystalline lattice. Because the crystal represents an average, all these alternatives are superimposed in the final reconstruction (Figure 5(a); Smith *et al.*, 1986).

Because relatively little information is available concerning the M-MuLV capsid protein, few structural comparisons can be performed with previous studies. However, distances between protein-free areas in the 3D reconstruction (41.3 Å; Figure 3(a)) correspond to 2D projection hole-to-hole distances (Figure 1(c)) and are compatible with major Fourier spacings determined using whole M-MuLV particles (Yeager *et al.*, 1998). The *Z* distance from the membrane-proximal to membrane-distal densities in our structures is 75-100 Å, slightly shorter than the length assigned to CA in cryo preparations of immature M-MuLV particles (Yeager *et al.*, 1998), and could be due to the negative stain used in our preparation, or the packing of our two-dimensional crystals relative to packing in virus particles. No high-resolution M-MuLV capsid structures are available, but hand-fitting the CA structure from EIAV (Jin *et al.*, 1999) shows a reasonably good match to one lobe from a putative his-MoCA dimer (Figure 6). The EIAV NTD structure overlaps with membrane-his-MoCA regions, while the CTD from the EIAV structure seems to have a similar angle as the his-MoCA CTD region(s), relative to the NTD. Another distinction between the structures of EIAV and his-MoCA is that the EIAV CTD is more pronounced than its putative M-MuLV counterparts. This may be due to inaccurate assignment of M-MuLV CTDs; differences between the structures of M-MuLV and EIAV CAs; or a relative reduction of M-MuLV CTD to NTD density because the CTD density is spread over three alternate conformations.

Discussion

As seen previously (Barklis *et al.*, 1997; Zuber & Barklis, 2000) the projection structure (Figure 1) of the his-MoCA crystals appears trigonal or hexagonal, with hole-to-hole spacings similar to the Fourier spacings observed in immature virions (Yeager *et al.*, 1998). This result may not be surprising, since our model system is meant to duplicate CA

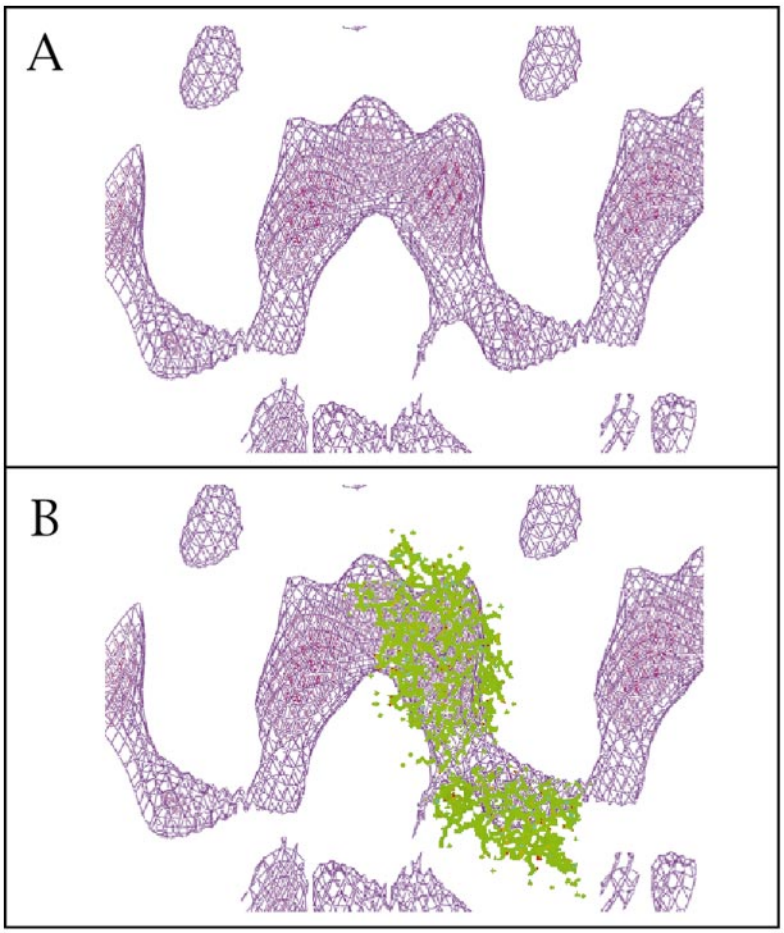

Figure 6. Fitting of the EIAV capsid structure to his-MoCA structures. The electron density map described in Figure 3 was used as a template to fit the high-resolution structure of the EIAV capsid protein monomer as described in Materials and Methods. The map is shown rotated 40° off of parallel with the membrane, clipped to a 37.5 Å slab at $\sigma = 2.0$. (a) A centered putative NTD dimer and its ''strand-type'' CTD interactions. (b) The structure of the EIAV capsid protein was manually overlaid to the his-MoCA electron density, to bring the NTD of the EIAV CA into register with one lobe of the putative his-MoCA NTD dimer.

protein interactions in immature virions. However, analysis of merges from tilted images (Figure 2) indicated that although the projection structure has a hexagonal appearance, the three-dimensional structure is not trigonally or hexagonally symmetric. In contrast, merging of 3D datasets using a primitive (*p*1) space group yielded low-phase residuals at all tilt angles, and *Q* values approaching 1 (Figure 2, Table 2). As shown in Figures 3 and 4, the membrane-bound CA structure appears to consist of strands of proteins in which putative NTDs and CTDs dimerize with different partners. From the appearance of Figures 3-5, membrane-proximal domains do not appear to possess 2-fold symmetry, as observed for head-to-head HIV-1 CA NTD dimers (Gamble *et al*., 1996; Momany *et al*., 1996). However, our current resolution does not permit us to distinguish whether membrane-proximal domains associate as head-to-tail or modified head-to-head dimers. Also of note are membrane-distal densities in close proximity to the putative CTDs, which make intra-strand connections (Figures 4 and 5). These densities connect with membrane-proximal densities at moderate density cut-offs ($\sigma = 1.5$), and do not appear to be artifacts, as they occurred in all our 3D renderings. The extra densities also appear to join neighbor membrane-proximal domains in alternate fashions (Figure 5). Based on X-ray and NMR structures of other retroviral capsid proteins, we believe it is unlikely that three membrane-distal densities are associated with each putative membrane-proximal monomer. Rather, we believe the densities represent alternative CTD conformations and that all conformations are superimposed in our final average reconstructions. This structural interpretation suggests that the CTDs can adopt at least three different conformations relative to their NTDs (Figure 5). The fact that putative CTD regions are somewhat less dense relative to the NTDs than might be expected from known crystal structures of CAs could be explained by the division of CTD electron density among multiple conformations (Smith *et al*., 1986). Moreover, the flexibility between putative M-MuLV NTDs and CTDs that we observe has been suggested for several retrovirus capsid proteins (Berthet-Colominas *et al*., 1999; Jin *et al*., 1999; Khorasanizadeh *et al*., 1999; Campos-Olivas *et al*., 2000). In this regard it is of interest that the most conserved regions of retroviral CAs, the spacer region and N-terminal portion of the MHR, are expected to occupy regions which adopt alternate conformations in our model. The requirement of this region to adopt multiple conformations may contribute to its sequence stability.

The appearance of the his-MoCA proteins in strands is reminiscent of that seen in retroviral CA and CA-NC helical rod assemblies (Ehrlich *et al*., 1992; Campbell & Vogt, 1995, 1997; Gross *et al*., 1998; Campbell & Rein, 1999; Ganser *et al*., 1999; Zuber *et al*., 2000), and is compatible with the spiral model of retroviral capsid structure assembly (Campbell & Vogt, 1995). Regardless of the relative orientations of NTDs and CTDs, if the two domains dimerize with different partners as we observe (Figures 4-6), at least one additional interaction would be necessary to connect strands to assemble a reproducible two-dimensional network, but our current data do not indicate what this interaction might be. It also is unclear whether there is a consistent repeated arrangement of CTDs relative to NTDs in our crystals. Our data are compatible with two interpretations: that alternate connections between CTDs are randomly distributed throughout crystals; or that there is a regular sequence of strand interconnections, but that reflections from such a large unit cell were too faint to see in our crystals. Each of the two possibilities poses problems for understanding the mechanism of Gag protein assembly into sheets, spheres or rods. Furthermore, it is uncertain what might be the effects of MA or NC domains on the structure of membrane-bound CA arrangements. These uncertainties may be resolved by examination of single *versus* multi-domain Gag proteins in higher resolution cryo-EM studies.

Materials and Methods

Protein crystallization and microscopy

Proteins were expressed and purified as described (Barklis *et al*., 1997). Briefly, the his-tagged M-MuLV CA protein (his-MoCA) was expressed in *Escherichia coli* strain BL21(DE3)/pLysS (Novagen) from a bacterial expression construct containing the capsid coding region inserted in the *Bam*HI site of pet15B (Novagen). The protein, consisting of a 25-residue leader containing a histidine tag, CA, and a vector-derived 21-residue C-terminal tail (Barklis *et al*., 1997), was purified by two rounds of nickel chelate chromatography. Purified proteins (>95% pure) were desalted by dialysis, lyophilized, resuspended in distilled water to 0.2-20 mg/ml and stored at −80 °C. Alternatively, purified fractions from nickel-chelate chromatography were desalted on Sephadex G25 spin columns, and stored at −80 °C or 4 °C in a buffer of 5 mM Tris, 16% (v/v) glycerol, and 0.02% (w/v) sodium azide. Phosphatidylcholine (PC; Avanti Polar Lipids), and the nickel chelating lipids 1,2-di-*O*-hexadecyl-sn-glycero-3-[1′-(2″-R-hydroxy-3′-*N*-(5-amino-1-carboxypentyl)-iminodiacetic acid) propyl ether] (DHGN; Barklis *et al*., 1997), and DOGS (Kubalek *et al*., 1994; Wilson-Kubalek *et al*., 1998) were stored as 10× stock solutions in chloroform under nitrogen gas.

Monolayer crystallization incubations were performed as described (Barklis *et al*., 1997, 1998). his-MoCA protein at a concentration of 0.2-2.0 mg/ml in 10 µl of a subphase buffer (25 mM sodium phosphate (pH 7.8 or 8.3); 250 mM NaCl; 5 mM sodium acetate (pH 7.6); 20% (v/v) glycerol) was overlaid with 1 µl of a 1:1 (v/v) mix of hexane/chloroform containing 200 µg/ml PC plus 50 µg/ml nickel-charged DHGN or DOGS. Crystallization incubations on glass slide depression wells or Teflon wells were kept overnight at 30 °C in sealed, distilled water-humidified plates, after which they were subjected to optional 0.5-3 hours 4 °C post-incubations. Crystalline arrays were lifted onto lacey EM grids, and grids were washed for 30-60 seconds in distilled water, stained for 45-60 seconds with 1.33% (w/v) uranyl acetate, wicked,

and air-dried. Monolayer arrays were observed and photographed at the Oregon Health Sciences University (OHSU) Philips CM120 transmission electron microscope (TEM). Photographs at 52,000× were taken using the CM120 low-exposure protocol, with the search mode at 2500-5000× , and the focus mode at 135,000× . For this TEM, at a particular zoom-mode magnification, the characteristic tilt axis to image *x*-axis angle was calculated using a pair of images from a single area, one untilted and the other at 30-60°. Distinctive points on each negative were assigned *x*,*y*-coordinates, and parameters from lattice vectors defined by three such points, from each image, were used as input (with tilted *versus* untilted lattice vectors switched) to the program EMTILT (Shaw & Hills, 1981) to back-calculate tilt to *x*-axis angles. At 52,000×, the OHSU CM120 TEM tilt to *x*-angle was determined to be 116.5(±7.1)°.

Digitization, transformation, and initial indexing

Negatives were scanned using an Optronics DEI-470 CCD mounted on a Fisher Stereomaster dissecting microscope equipped with a 0.5× objective lens. Raw scanned images were converted to the MRC image format by converting the raw TIFF images to PGM format, and then to MRC format using the MRC PGM_FIMG program (Crowther *et al*., 1996) and inverting about the *y*-axis of the image to correct for inversion introduced by the PGM_FIMG program, as described (Barklis *et al*., 1997). Care was taken to specify a standard methodology for the scanning to ensure that all images were in the same orientation (negatives face down, label imprinted from the microscope to the left). Image inversions throughout specimen handling and image processing, the net result being one image inversion, were as follows: by lifting the specimen onto the grid, by the specimen holder of the electron microscope, by the TEM lenses during magnification, by the MRC conversion from a TIFF image format to an MRC image format, and by the subsequent correction.

Images were Fourier-transformed using the ICE_FFTRANS function, part of the ICE suite of MRC programs (Crowther *et al*., 1996; Hardt *et al*., 1996). Transforms were represented as power spectra using the program SPECTRA (Schmid *et al*., 1993), and were hand-indexed in a hexagonal fashion (γ^* approximately 60°). Lattices were refined and unbent with SPECTRA versions of LATREF, MMBOX, and UNBEND (Schmid *et al*., 1993) programs. The resulting amplitude and phase (APH) files were corrected for phase inversions resulting from the contrast transfer function (CTF) using the CTF_APPLY program from the ICE suite. To determine suitability of the lattices for each space group, APH files from untilted, CTF-corrected images were used as input to the program ALLSPACE (Valpuesta *et al*., 1994; Crowther *et al*., 1996).

Merging, optimization, and lattice line binning

For merging of 3D data sets, untilted image f005c, which gave ALLSPACE phase residuals of 21.7°, 6.1°, and 11.3° for the *p*1, *p*3, and *p*6 space groups, respectively, was chosen as an initial reference. Reference files were created by applying phase shifts derived from the ALLSPACE output to APH file reflections to center phase origins, prior to merging on the asymmetric units. By convention (Hahn, 1983), the *p*1 reference was converted to positive *h* index form, while *p*3 and *p*6 reference files were of $h+$, $k+$ form, in which symmetry-related reflections were combined by vector averaging. APH files from tilted and untilted images were merged in three dimensions using the program ORIGTILTC for *p*1 and *p*3 and ORIGTILTB for *p*6 (Baldwin *et al*., 1988; Henderson *et al*., 1990). Merging was performed to a resolution of 20 Å, with a *Z*-window of 0.002 $Å^{-1}$, a *Z* thickness of 150 Å, $a = b = 72.5$ Å, and $\gamma = 119.6°$, using reflections with a signal to noise factor (*IQ*) of 4 or better for phase origin searches. The previously derived tilt axis to *x*-axis angle was added to the *x*-axis to a^* angle from each image, as calculated from the power spectra during indexing, to give tilt axis to a^* angles for each file in the merge. Tilt angles were as indicated on the CM120 goniometer, and were cross-checked with the program EMTILT (Shaw & Hills, 1981) for high tilt angles.

For *p*1 merges, an optimization step similar to that used by Stoylova *et al.* (1999) was performed to facilitate indexing. Since $\gamma^* = 60°$ and the *a* and *b*-axes were approximately equal in length, each diffraction pattern could be indexed in six ways. Based on an initial indexing (*h*,*k*) the transformations to the h', k' sets were as follows: set 1, identity, (h, k); set 2, ($h + k$, $-h$); set 3, (k, $-[h + k]$); set 4, ($-h$, $-k$); set 5, ($-[h + k]$, h); set 6, ($-k$, $h + k$). For optimization, each diffraction pattern was rotated to generate the six index APH file sets, and each set was tested in merges with all previously optimized APH files: for each image, the APH set yielding the lowest overall phase residual was incorporated into merges. A reduction from an overall phase residual of 20.75° to an overall phase residual of 15.8° was observed after index optimization.

Z^* values from output of the ORIGTILT programs were binned using a *Z* thickness of 600 Å ($Z^* = 0.001666$), four to six times the expected sample thickness (Amos *et al*., 1982), to get *l* index values. Reflections from each bin were vector averaged to give single *h*,*k*,*l* indexed reflections in HKL files. The quality of the averaged reflections was monitored by *Q* values (Robinson *et al*., 1988) that were calculated by dividing the vector addition-derived amplitudes by the simple scalar addition product of amplitudes from the corresponding reflections: a *Q* value of 1.0 signifies a maximum match value, while a random *Q* value corresponds to the reciprocal of the square-root of the number of vectors added (Robinson *et al*., 1988).

Back transformation and 3D image representation

Three-dimensional back-transformations of *p*1, *p*3, and *p*6 HKL files were performed using the XFFT program from the XTALVIEW suite (McRee, 1992), after HKL file conversion into XTALVIEW PHS files. Volume viewing of XTALVIEW MAP files was performed using the XFIT program. For the final *p*1 merged set, images from the index-optimized *p*1 merge with phase residuals greater than 30° were removed, and the resulting 218 image set was re-optimized giving a dataset with an overall phase residual of 14.79°. The merged, binned HKL file was filtered to 22.5 Å in the *x*-*y* plane (based on phase residuals) and then to 35 Å in the *z* dimension (based on completeness of low-resolution lattice lines). After filtering, HKL files were converted to XTALVIEW PHS and then MAP files for 3D representation by XFIT. Rotational matrices for the volume seen in the Figures are as follows: Figure 3(a) and (b), *x*[− 1, 0, 0], *y*[0, 1, 0], *z*[0, 0, −1]; Figure 4(a) and (b), *x*[0.5719, 0.8095, −0.1328], *y*[0.8203, −0.5664, 0.0797], *z*[− 0.0107, −0.1545, −0.9879];

Figure 4(c), *x*[0, 1, 0], *y*[0, 0, 1], *z*[1, 0, 0]; Figure 4(d), *x*[0.866, 0.5, 0], *y*[0, 0, −1], *z*[− 0.5, 0.866, 0]; Figure 5(a)-(d), *x*[− 0.0751, 0.9770, −0.1997], *y*[0.8272, 0.1729, 0.5346], *z*[0.5568, −0.1250, −0.8212]; Figure 6(a) and (b), *x*[0.5031, 0.8624, 0.0558], *y*[0.7491, −0.4030, −0.5257], *z*[− 0.4309, 0.3063, −0.8488]. Figure 6(a) and (b) were front-clipped at 29.2 Å and back-clipped at −8.3 Å from the origin. The EIAV monomer fitting was performed manually by overlaying a monomer from the EIAV capsid structure, PDB accession number 1EIA (Jin *et al.*, 1999), such that the N-terminal domain of the EIAV capsid was overlaid by eye as well as possible to a lobe of the putative N-terminal dimer density in all three dimensions.

Acknowledgments

We are grateful to Mike Schmid and Stephen Fuller for continual image processing help and advice; and to Elizabeth Kubalek-Wilson for the nickel-chelating lipid, DOGS. Sonya Karanjia contributed invaluable assistance in protein purification and crystallization; Robin Barklis performed tilt axis determinations; and Doug Huseby and Guy Zuber provided helpful advice. This work was supported by NIH grants GM52914 and GM60170 to E.B., and by NIH pre-doctoral training support (A107472) to J.M.

References

Alin, K. & Goff, S. P. (1996). Amino acid substitutions in the CA protein of Moloney murine leukemia virus that block early events in infection. *Virology,* **222**, 339-351.

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

Baldwin, J., Henderson, R., Beckman, E. & Zemlin, F. (1988). Images of purple membrane at 2.8 Å resolution obtained by cryo-electron microscopy. *J. Mol. Biol.* **202**, 585-591.

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

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

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

Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F. & 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.

Borsetti, A., Ohagen, A. & Gottlinger, H. G. (1998). The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. *J. Virol.* **72**, 9313-9317.

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

Campbell, S. & 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. & 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. & 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.

Craven, R. C., Leure-duPree, A. E., Weldon, R. A. J. & Wills, J. W. (1995). Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. *J. Virol.* **69**, 4213-4227.

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

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

Fuller, S. D., Capaldi, R. A. & Henderson, R. (1979). Structure of cytochrome *c* oxidase in deoxycholate-derived two-dimensional crystals. *J. Mol. Biol.* **134**, 305-327.

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

Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I. & 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., Schwedler v., U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I. & 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. & 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. & Sundquist, W. I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science,* **273**, 231-235.

Glaeser, R. M., Tong, L. & Kim, S. H. (1989). Three-dimensional reconstructions from incomplete data: interpretability of density maps at ''atomic'' resolution. *Ultramicroscopy,* **27**, 307-318.

Goff, S. P. & Lobel, L. I. (1987). Mutants of murine leukemia viruses and retroviral replication. *Biochim. Biophys. Acta,* **907**, 93-123.

Gross, I., Hohenberg, H., Huckhagel, C. & 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.

Hahn, T. (1983). Editor of *International Tables for Crystallography*: *Space Group Symmetry*, D. Reidel Pub. Co., Dordrecht, Holland.

Hardt, S., Wang, B. & Schmid, M. F. (1996). A brief description of I.C.E.: The integrated crystallographic environment. *J. Struct. Biol.* **116**, 68-70.

Henderson, R., Baldwin, J., Ceska, T., Zemlin, F., Beckmann, R. & Downing, K. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899-929.

Hsu, H.-W., Schwartzberg, P. & Goff, S. P. (1985). Point mutations in the *p*30 domain of the *gag* gene of Moloney murine leukemia virus. *Virology,* **142**, 211-214.

Jin, Z., Jin, L., Peterson, D. L. & 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. & 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.

Kubalek, E. W., Grice, S. F. L. & Brown, P. O. (1994). Two-dimensional crystallization of histidine-tagged, HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid. *J. Struct. Biol.* **113**, 117-123.

Mammano, F., Ohagen, A., Hoglund, S. & Gottlinger, H. G. (1994). Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J. Virol.* **68**, 4927-4936.

McDermott, J., Farrell, L., Ross, R. & Barklis, E. (1996). Structural analysis of human immunodeficiency virus type 1 Gag protein interaction using cysteine-specific reagents. *J. Virol.* **70**, 5106-5114.

McRee, D. E. (1992). A visual protein crystallographic software system for X11/XView. *J. Mol. Graph.* **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. & Rossman, M. G. (1996). Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.* **3**, 763-770.

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

Nermut, M. V., Hockley, D. J., Bron, P., Thomas, D., Zhang, W.-H. & 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.

Reicin, A. S., Paik, S., Berkowitz, R. D., Luban, J., Lowy, I. & Goff, S. P. (1995). Linker insertion mutations in the human immunodeficiency virus type 1 *gag* gene: effects on virion particle assembly, release, and infectivity. *J. Virol.* **69**, 642-650.

Robinson, J. P., Schmid, M. F., Morgan, D. G. & Chiu, W. (1988). Three-dimensional structural analysis of tetanus toxin by electron crystallography. *J. Mol. Biol.* **200**, 367-375.

Schmid, M., Dargahi, R. & Tam, M. (1993). SPECTRA: a system for processing electron images of crystals. *Ultramicroscopy,* **48**, 251-264.

Schwedler, U. K. V., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H., Davis, D. R. & Sundquist, W. I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555-1568.

Shaw, P. J. & Hills, G. J. (1981). Tilted specimen in the electron microscope: a simple specimen holder and the calculation of tilt angles for crystalline specimens. *Micron,* **12**, 279-282.

Smith, J. L., Hendrickson, W. A., Honzatko, R. B. & Sheriff, S. (1986). Structural heterogeneity in protein crystals. *Biochemistry,* **25**, 5018-5027.

Srinivasakumar, N., Hammarskjold, M.-L. & Rekosh, D. (1995). Characterization of deletion mutations in the capsid region of human immunodeficiency virus type 1 that affect particle formation and Gag-Pol precursor incorporation. *J. Virol.* **69**, 6106-6114.

Stoops, J. K., Baker, T. S., Schroeter, J. P., Kolodziej, S. J., Niu, X. D. & Reed, L. J. (1992). Three-dimensional structure of the truncated core of the *Saccharomyces cerevisiae* pyruvate dehydrogenase complex determined from negative stain and cryoelectron microscopy images. *J. Biol. Chem.* **267**, 24769-24775.

Strambio-de-Castillia, C. & Hunter, E. (1992). Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. *J. Virol.* **66**, 7021-7032.

Stoylova, S., Ford, R. & Holzenburg, A. (1999). Cryo-electron crystallography of small and mosaic 2-D crystals: an assessment of a procedure for high-resolution data retrieval. *Ultramicroscopy,* **77**, 113-128.

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

Uzgiris, E. & Kornberg, R. (1983). Two-dimensional crystallization technique for imaging macromolecules with application to antigen-antibody-complement complexes. *Nature,* **301**, 125-129.

Valpuesta, J. M., Carrascosa, J. L. & Henderson, R. (1994). Analysis of electron microscopy images and electron diffraction patterns of thin crystals of ø29 connectors in ice. *J. Mol. Biol.* **240**, 281-287.

Wang, C.-T. & Barklis, E. (1993). Assembly, processing, and infectivity of human immunodeficiency virus type 1 Gag mutants. *J. Virol.* **67**, 4264-4273.

Wang, C.-T., Lai, H.-Y. & Li, J.-J. (1998). Analysis of human immunodeficiency virus type 1 gag coding sequences capable of virus-like particle assembly and release. *J. Virol.* **72**, 7950-7959.

Weldon, R. A. J. & Wills, J. W. (1993). Characterization of a small (25 kilodalton) derivative of the Rous sarcoma virus Gag protein competent for particle release. *J. Virol.* **67**, 5550-5561.

Wilson-Kubalek, E. M., Brown, R. E., Celia, H. & Milligan, R. A. (1998). Lipid nanotubes as substrates for helical crystallization of macromolecules. *Proc. Natl Acad. Sci. USA,* **95**, 8040-8045.

Yeager, M., Wilson-Kubalek, E. M., Weiner, S. G., Brown, P. O. & 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**, 7299-7304.

Zhang, W., Hockley, D. J., Nermut, M. V., Morikawa, Y. & Jones, I. M. (1996). Gag-Gag interactions in the C-terminal domain of human immunodeficiency virus type 1 p24 capsid antigen are essential for Gag particle assembly. *J. Gen. Virol.* **77**, 743-751.

Zuber, G. & 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. & Barklis, E. (2000). Assembly of retrovirus capsid-nucleocapsid proteins in the presence of membranes or RNA. *J. Virol.* **74**, 7431-7441.

Edited by M. Summers

(Received 17 April 2000; received in revised form 18 June 2000; accepted 18 June 2000)