Structural analysis of membrane-bound retrovirus capsid proteins

Eric Barklis^{1,2}, Jason McDermott¹,
Stephan Wilkens³, Eric Schabtach⁴,
M.F.Schmid⁵, Stephen Fuller⁶,
Sonya Karanjia¹, Zachary Love¹,
Russell Jones⁷, Yuanjui Rui⁸, Xiumin Zhao⁸
and David Thompson⁸

¹Vollum Institute and Department of Microbiology and ⁷Department of Pathology, Oregon Health Sciences University, Portland, OR 97201,

³Institute of Molecular Biology and ⁴Institute of Neuroscience, University of Oregon, Eugene, OR 97403, ⁵Verna and Marrs McLean Department of Biochemistry and W.M.Keck Center for Computational Biology, Baylor College of Medicine, Houston, TX 77030, USA,

⁶Structural Biology Programme, European Molecular Biology Laboratory, D-69012 Heidelberg, Germany and ⁸Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA

E.Barklis, J.McDermott, S.Wilkens and D.Thompson contributed equally to this work

²Corresponding author

We have developed a system for analysis of histidine-tagged (His-tagged) retrovirus core (Gag) proteins, assembled *in vitro* on lipid monolayers consisting of egg phosphatidylcholine (PC) plus the novel lipid DHGN. DHGN was shown to chelate nickel by atomic absorption spectrometry, and DHGN-containing monolayers specifically bound gold conjugates of His-tagged proteins. Using PC+DHGN monolayers, we examined membrane-bound arrays of an N-terminal His-tagged Moloney murine leukemia virus (M-MuLV) capsid (CA) protein, His-MoCA, and *in vivo* studies suggest that *in vitro*-derived His-MoCA arrays reflect some of the Gag protein interactions which occur in assembling virus particles. The His-MoCA proteins formed extensive two-dimensional (2D) protein crystals, with reflections out to 9.5 Å resolution. The image-analyzed 2D projection of His-MoCA arrays revealed a distinct cage-like network. The asymmetry of the individual building blocks of the network led to the formation of two types of hexamer rings, surrounding protein-free cage holes. These results predict that Gag hexamers constitute a retrovirus core substructure, and that cage hole sizes define an exclusion limit for entry of retrovirus envelope proteins, or other plasma membrane proteins, into virus particles. We believe that the 2D crystallization method will permit the detailed analysis of retroviral Gag proteins and other His-tagged proteins.

Keywords: electron microscopy/image analysis/lipid monolayer/retrovirus

Introduction

A number of cellular and viral components are present in all infectious mammalian retroviruses. Such viruses are

surrounded by host-derived lipid membrane envelopes, which carry receptor-binding/fusion proteins, the surface (SU) and transmembrane (TM) proteins, that are encoded by the retroviral *env* genes (Weiss *et al.*, 1984; Leis *et al.*, 1988). Within each mammalian retrovirus particle are two copies of the viral RNA genome with associated, cellularly derived tRNAs, which serve as primers during reverse transcription. Virus interiors also contain 1000–5000 copies of the viral Gag (group-specific antigen) proteins, and 10–100 copies of the viral *pol* gene products including the viral protease (PR), reverse transcriptase (RT), RNaseH and integrase (IN) (Weiss *et al.*, 1984). Components of C-type retroviruses, such as Moloney murine leukemia virus (M-MuLV), and lentiviruses, such as the human immunodeficiency virus (HIV), are delivered for assembly at the plasma membranes of infected cells (Weiss *et al.*, 1984), although assembly can occur at other intracellular membrane locations (Hansen *et al.*, 1990; Faেকে *et al.*, 1993; Wang *et al.*, 1993).

The one viral component that has been shown to be necessary and sufficient for C-type and lentivirus particle assembly is the Gag protein, which is synthesized as a precursor polyprotein (Pr^{Gag}), and normally cleaved into mature processed Gag proteins by the viral PR during or after budding. The proteolytic processing of Pr^{Gag} is required for infectivity, and results in a major change in the virus core, in which electron-dense material adjacent to the periphery of immature particles reorganizes into central spherical, cylindrical or cone-shaped structures (Weiss *et al.*, 1984; Stewart *et al.*, 1990). Three mature Gag proteins, matrix (MA), nucleocapsid (NC) and capsid (CA), are present in both C-type viruses and lentiviruses. During biosynthesis, MA is myristylated at its amino-terminus, and this fatty acid modification is necessary for membrane association of Pr^{Gag} proteins (Rein *et al.*, 1986). Genetic experiments have demonstrated that the HIV-1 matrix protein also interacts with the HIV Env protein complex (Yu *et al.*, 1992; Faেকে *et al.*, 1993; Wang *et al.*, 1993). However, the central 80–90% of HIV MA can be deleted without blocking assembly of virus particles (Faেকে *et al.*, 1993) that are conditionally infectious (Wang *et al.*, 1993). At or near the C-termini of lentivirus or C-type retrovirus, Pr^{Gag} proteins are Cys–His, zinc finger-containing NC domains. The NC domains have RNA-binding capabilities, but it is not clear whether NC completely accounts for the specificity of RNA encapsidation into virus particles (Zhang and Barklis, 1995). While MA interacts with viral membranes and Env proteins, and NC interacts with the viral RNA, the Gag CA domains appear to establish interprotein contacts that are essential to the oligomerization of Pr^{Gag} proteins, and assembly of Gag–Pol and Gag fusion proteins into virions (Hansen *et al.*, 1990; Jones *et al.*, 1990; Chazal *et al.*, 1994; Mammano *et al.*, 1994; Wang *et al.*, 1994; Craven

et al., 1995; Hansen and Barklis, 1995; Srinivasakumar *et al.*, 1995; McDermott *et al.*, 1996). Despite the apparent conservation of retrovirus capsid protein function, only a 20–30 residue section in the C-terminal half of CA, called the major homology region (MHR), is well conserved at the primary sequence level (Mammano *et al.*, 1994; Craven *et al.*, 1995; Hansen and Barklis, 1995; Clish *et al.*, 1996; McDermott *et al.*, 1996). In addition to the three conserved mature Gag proteins (MA, CA and NC), M-MuLV encodes a p12 protein between the MA and CA domains, and HIV encodes a p6 domain at the C-terminus of its Pr^{gag}: while these domains are essential to the infectivities of their respective viruses, they are not absolutely required for virus particle assembly (Crawford and Goff, 1984; Spearman *et al.*, 1994; Hansen and Barklis, 1995).

Three-dimensional (3D) structures of HIV MA (Matthews *et al.*, 1994; Hill *et al.*, 1996) and NC (Morrelet *et al.*, 1992) are available, as are partial structures of HIV CA (Clish *et al.*, 1996; Gitti *et al.*, 1996; Momany *et al.*, 1996). However, despite observations of preferred interprotein contacts in 3D crystals (Hill *et al.*, 1996), it is unclear how the individual Gag proteins fit into the immature or mature virus particle structures. Ordinarily, retrovirus structure analysis might proceed by X-ray diffraction analysis of 3D crystals of virus, or by electron microscopy (EM) and image analysis of homogeneous preparations of virus particles. Unfortunately, these approaches have proven difficult, since the virions do not crystallize, preparations do not appear homogeneous, and it is unclear whether mature or immature virus cores have helical or icosahedral symmetry, which would be a great help in image reconstruction (Choi *et al.*, 1991; Fuller *et al.*, 1995, 1996). In the absence of homogeneous, *in vivo*-derived virus preparations, several groups have undertaken the analysis of virus-like particles formed *in vitro* from Pr^{gag} proteins (Nermut *et al.*, 1994; Klikova *et al.*, 1995) or their derivatives (Ehrlich *et al.*, 1992; Campbell and Vogt, 1995; J.McDermott and E.Barklis, unpublished observations). Such studies have shown that Gag-derived monomer proteins can assemble to form higher order structures, but, for the most part, resolution has been such that the nature of Gag interprotein contacts remains unclear. Nevertheless one analysis of HIV Pr^{gag} proteins assembled at the plasma membranes of baculovirus vector-infected cells provided sufficient resolution to suggest that these proteins form a fullerene cage-like network at the membrane (Nermut *et al.*, 1994).

Because of the difficulties with traditional X-ray and EM approaches, and the current limited resolution available via analysis of standard *in vitro* assembly products, we have adapted a lipid monolayer approach (Uzgiris and Kornberg, 1983; Darst *et al.*, 1991a,b; Celia *et al.*, 1994; Olofsson *et al.*, 1994; Avila-Sakar and Chiu, 1996) to determine how Gag proteins may associate with each other at a membrane. Our approach (Zhao *et al.*, 1994), outlined in Figure 1, has been to employ a nickel-chelating lipid to facilitate the two dimensional (2D) crystallization of N-terminal histidine-tagged (His-tagged) Gag proteins at a lipid monolayer. This approach recently has been used, with modifications, to examine a variety of interfacial lipid–protein interactions (Schmitt *et al.*, 1994; Zhao *et al.*, 1994; Dietrich *et al.*, 1995, 1996; Kubalek *et al.*, 1995; Ng *et al.*, 1995; Frey *et al.*, 1996), and can be used to obtain

protein 3D structures, using image analysis resources developed over the past two decades (Unwin and Henderson, 1975; Fuller *et al.*, 1979; Baldwin *et al.*, 1988; Frank *et al.*, 1988; Henderson *et al.*, 1990; Schmid *et al.*, 1993; Kuhlbrandt *et al.*, 1994). The monolayer technique described in Figure 1 appears ideally suited for the analysis of retroviral Gag proteins, since their natural function is to oligomerize at the face of a membrane. Furthermore, at least for HIV, only the membrane-anchoring activity of the amino-terminal Gag matrix domain is required for virus particle assembly (Facke *et al.*, 1993; Wang *et al.*, 1993). In our current study, we have substituted a His-tag for the membrane-anchoring myristate moiety of the Gag protein, and have examined crystalline arrays of a His-tagged M-MuLV capsid protein (His-MoCA) formed on a monolayer of egg phosphatidyl choline (PC) and a novel nickel-chelating lipid, 1,2-di-*O*-hexadecyl-*sn*-glycero-3-[1'-(2''-*R*-hydroxy-3'-*N*-(5-amino-1-carboxypentyl)-iminodiacetic acid) propyl ether] (DHGN). The His-MoCA protein forms 2D crystals with reflections at 9.5 Å resolution, and can be classified as either orthorhombic ($a = 79.6$ Å, $b = 137.5$ Å, $\gamma = 90^\circ$) or hexagonal ($a = b = 79.7$ Å, $\gamma = 60^\circ$). Image reconstruction shows a cage-like array of proteins, similar to that seen for HIV-1 Pr^{gag} proteins (Nermut *et al.*, 1994). In 2D projections, the putative His-MoCA monomers form two different types of hexagonal units surrounding two different protein-free cage holes: a circular hole, with a diameter of 19.2 Å; and a triangular cage hole (length 28.0 Å, width 23.2 Å). Our results permit specific predictions concerning retrovirus particle structure and capsid protein interactions with MA, NC and Env proteins.

Results

Characterization of the nickel-chelating lipid, DHGN

To foster the development of a lipid monolayer approach for the analysis of retrovirus Gag protein interactions, we designed a nickel-chelating lipid to serve as a ligand to His-tagged Gag proteins. Our objective was to combine the nickel-chelating nitrilotriacetic acid (NTA; Hochuli *et al.*, 1987) group with an activated diacyl glycerol derivative (Thompson *et al.*, 1994). For this purpose, a convergent synthesis scheme was developed (Figure 2; Materials and methods), where compounds 3 and 5 (Figure 2) served as immediate precursors to DHGN (Figure 2, compound 6; Zhao *et al.*, 1994). The crude DHGN product was purified by two rounds of silica gel chromatography, producing DHGN, whose structure was confirmed by IR and NMR spectroscopy (Materials and methods).

While the presence of the NTA headgroup on DHGN suggested that the lipid should chelate nickel, it was necessary to test this assumption directly. To do so without excessive consumption of DHGN, we resorted to atomic absorption spectrometry using an instrument equipped with a graphite furnace, capable of detecting micromolar nickel concentrations of 20 µl samples (Materials and methods). For analysis of nickel binding, a simple extraction assay was employed. As shown in Figure 3A, limiting amounts of nickel sulfate in water were mixed with lipids in chloroform and, after phase separation, nickel levels were measured in each phase. When this protocol was

Fig. 1. Model for forming histidine-tagged protein arrays on a monolayer of nickel-chelating lipids at the air-buffer interface. Shown is a model, based on previous lipid monolayer studies (Uzgiris and Kornberg, 1983; Darst *et al.*, 1991a,b; Celia *et al.*, 1994; Olofsson *et al.*, 1994), in which His-tagged proteins bind to nickel-chelating lipids at the air-buffer interface. For our purposes, the proteins are N-terminally His-tagged retrovirus Gag proteins and the nickel-chelating lipid is DHGN, a dialkylglycerol derivative carrying an NTA (Hochuli *et al.*, 1987) headgroup. Analysis of 2D protein arrays, following previously established methods (Unwin and Henderson, 1975; Fuller *et al.*, 1979; Schmid *et al.*, 1993), involves adhering arrays to EM grids; photography of stained or unstained protein crystals at 20 000–100 000 $\times$ magnification; evaluation of 2D arrays by optical diffraction of EM negatives; digitization of negatives; and image processing to obtain structural details. In many circumstances, resolution is in the 10–30 Å range, sufficient for observation of protein–protein interactions, and in some cases atomic resolution has been achieved (Baldwin *et al.*, 1988; Henderson *et al.*, 1990; Kuhlbrandt *et al.*, 1994; Avila-Sakar and Chiu, 1996).

Fig. 2. Synthesis of the nickel-chelating lipid, DHGN. DHGN [compound 6; 1,2-di-*O*-hexadecyl-*sn*-glycero-3-(1'-(2''-*R*-hydroxy-3'-(5-amino-1-carboxypentyl)-iminodiacetic acid) propyl ether)] was prepared from starting materials *N* ϵ -benzyloxycarbonyl-L-lysine (compound 1) and 1,2-di-*O*-hexadecyl-*sn*-glycerol (compound 4) in a multistep synthesis. Compound 4 was converted to 1,2-di-*O*-hexadecyl-*sn*-glyceryl-3-*R*-glycidyl ether (compound 5) by alkylation with *R*-glycidyl-3-nitrobenzenesulfonate. Compound 3 [*N*-(5-amino-1-carboxypentyl)-iminodiacetic acid] was prepared from compound 1 as described previously (Hochuli *et al.*, 1987) by acetylation with bromoacetic acid to give *N*-(5-benzyloxycarbonylamino-1-carboxypentyl)-iminodiacetic acid (compound 2), which was hydrogenated in the presence of Pd-C catalyst to yield the compound 3 intermediate. DHGN was generated by epoxide ring opening of compound 5 via nucleophilic attack with compound 3. The crude DHGN product was purified by silica gel chromatography and the structure confirmed by IR and NMR spectroscopy.

used with control lipids OG (octyl glucoside) and DHP (dihexadecyl phosphate), Ni²⁺ remained in the aqueous phase (Figure 3B). In contrast, DHGN chelated nickel ions, resulting in their extraction to the organic phase. As expected, EDTA, which blocks His tag protein binding to NTA-resin, effectively competed with DHGN for Ni²⁺, inhibiting DHGN-mediated extraction of nickel.

The above experiments indicated that DHGN chelates Ni²⁺, but did not permit an assessment of whether DHGN, charged with nickel, was capable of binding His-tagged proteins. To test this, we modified the approach illustrated in Figure 1 by using His-tagged or untagged proteins conjugated to gold particles, rather than using free proteins in the buffer subphase. By this method, we reasoned it would be possible to quantitate protein binding to mono-

layers via gold particle counts. Thus, using standard procedures (Geoghegan and Ackerman, 1977; Horisberger and Clerc, 1985; see Materials and methods), gold conjugates were prepared to His-tagged activating transcription factor (His-ATF-gold) and His-tagged cAMP-responsive element modulator (His-CREM-gold). As a control, we conjugated bovine serum albumin (BSA-gold) and, to examine potential non-specific binding of a basic protein, cytochrome *c* (cytoC-gold; pI = 10.4) also was conjugated. Total gold particle conjugate concentrations were determined as described in Materials and methods, while binding to lipid monolayers was assessed by EM of glucose-embedded monolayers after binding incubations and washes (see Materials and methods). An example of our results is shown in Figure 4, where His-CREM-gold

Fig. 3. Atomic absorption analysis of nickel chelation to lipids. (A) Nickel was measured with a Perkin Elmer 603 atomic absorption spectrometer equipped with an HGA-400 programmer and a graphite furnace. For analysis of nickel binding to lipids, 250 μ M lipid solutions in 0.1 ml of chloroform were overlaid with an equal volume of 20 μ M nickel sulfate in distilled water. Samples were rocked for 1–18 h at 25°C, after which phases were separated by low speed centrifugation. Triplicate 20 μ l samples from aqueous (top) and organic (bottom) phases were measured and compared with nickel sulfate standards in distilled water to determine the nickel concentrations in each phase. (B) Nickel concentrations in aqueous (white) and organic (black) phases are as shown. In one experiment, EDTA was added to the aqueous phase to a final concentration of 50 mM. Neither EDTA nor any of the stock lipid solutions showed any nickel content above background prior to the phase fractionation experiments.

was incubated beneath a monolayer of PC only (Figure 4A) or PC plus DHGN (Figure 4B). As illustrated, on PC-only monolayers, the His-CREM–gold conjugate washed to the edges of the lacy carbon support. In contrast, when lipid monolayers consisted of PC plus DHGN, gold particles were observed over the surface of the monolayers, suggesting a specific binding of the His-tagged moiety to the DHGN headgroup. This interpretation was substantiated when gold particle counts were performed (Table I). All conjugates showed total gold counts of 100–300 particles/nl, and PC monolayer binding of 3–32 particles/1000 nm². However, binding of gold-labeled, His-tagged proteins to DHGN-containing monolayers was enhanced at least 10-fold, while binding of control conjugates was unimproved. These results support the notion that nickel-bound DHGN serves as a ligand for His-tagged proteins.

The success of the above experiments afforded us an avenue to evaluate the effects of incubation conditions on DHGN binding to His tags. Thus, lipid, salt, pH, ion and reducing agents were examined for their effects in gold conjugate binding studies (Table II). As expected, PC-only monolayers (Table II; 0% DHGN) and monolayers of PC plus DHGN that had not been charged with Ni²⁺ (20% DHGN no nickel) showed no appreciable His-tagged gold binding. We also observed that reduction of DHGN to PC ratios down to 1% DHGN reduced binding only slightly, whereas increasing the percentage to 50% reduced binding levels to 16% of the standard conditions. We interpret the results with 1% DHGN to indicate that this amount of DHGN (6×10^8 excess over gold particles) is still functionally saturating in our incubations, while we hypothesize that high levels of DHGN may disrupt PC

monolayers in some as yet undetermined fashion. While high levels of DHGN reduced binding, variation of subphase salt and pH conditions showed little effect on binding (Table II). However, certain divalent cations seemed to interfere with the DHGN–His tag interaction. In particular, CaCl₂ and FeCl₂ reduced binding levels 4- to 5-fold, and 5 mM NiCl₂ in the subphase eliminated binding in a more dramatic fashion than even EDTA. Finally, as might be expected from their effects on metal chelate chromatography (Hochuli *et al.*, 1987), dithiothreitol (DTT) and β -mercaptoethanol (β -Me) also inhibited His-tagged gold conjugate binding to DHGN-containing monolayers, which suggests that reducing agents must be used sparingly in studies with this lipid.

Analysis of M-MuLV capsid protein arrays

Provided with DHGN, which can bind His-tagged proteins, it was possible to examine whether His-tagged M-MuLV Gag proteins could form 2D crystals on DHGN-containing monolayers. Initially, we have chosen to test an N-terminally His-tagged M-MuLV capsid protein (His-MoCA) for several reasons: capsid–capsid interactions have been shown to be major determinants in retrovirus particle assembly (Hansen *et al.*, 1990, 1993; Jones *et al.*, 1990; Wang *et al.*, 1993, 1994; Mammano *et al.*, 1994; Craven *et al.*, 1995; Hansen and Barklis, 1995; McDermott *et al.*, 1996); tethering of His-MoCA to our model monolayer is analogous to myristate-mediated anchoring of Gag proteins *in vivo*; His-MoCA is not subject to interdomain proteolytic processing in bacteria; and His-MoCA may serve as a foundation for comparison with larger Gag-derived fragments in the future. Our hypothesis is that His-MoCA interactions reflect interactions between membrane-tethered capsid domains in assembling or immature M-MuLV particles, and *in vivo* experiments concerning the relevance of this interpretation are described below. As produced from the pET15B-MoCA plasmid in *Escherichia coli*, His-MoCA is a 35 kDa protein, composed of the 30 kDa M-MuLV capsid domain, a His tag-containing 25 residue N-terminal leader sequence, and a vector-derived, 21 residue C-terminal tail (see Materials and methods). The protein was isolated under non-denaturing conditions by two rounds of NTA–nickel chromatography (Hochuli *et al.*, 1987), in which the first round of chromatography typically gave a >90% pure protein preparation, and the second chromatographic step yielded >95% pure His-MoCA, based on Coomassie blue staining of SDS–PAGE gels. Because His-MoCA has no known enzymatic activity, it was not possible to employ an activity assay to determine whether the protein assumed a native conformation. However, to assess the conformational status of His-MoCA indirectly, we subjected His-MoCA and capsid proteins from virus particles to partial trypsin digestion and compared the products after SDS–PAGE and immunoblotting. Our results showed that one His-MoCA proteolytic product was the same size as *in vivo*-derived M-MuLV CA, and that both His-MoCA and M-MuLV CA gave partial proteolytic products of 27.9, 27.4, 26.5, 26.1, 25.6, 22.9, 20.7, 19.5 and 18.7 kDa (data not shown). These results suggested that the conformation of His-MoCA was similar to that of *bona fide* capsid protein from virus particles.

Using the purified His-MoCA protein, we employed the lipid monolayer approach depicted in Figure 1 for the

Fig. 4. Protein–gold conjugate binding to lipid monolayers. Histidine-tagged CREM proteins (His-CREM), conjugated to 20 nm gold particles as described in Materials and methods, were incubated under standard conditions beneath lipid monolayers composed of PC only (A) or PC plus nickel-charged DHGN (B). After incubations, monolayers were lifted onto lacy carbon grids, rinsed for 30 s in 1% glucose, dried, and viewed and photographed on the Philips 301 TEM. Carbon supports are clearly visible as the broad dark lines; monolayers are visible with occasional holes caused by beam damage, while gold particles appear as black dots. Note that with either monolayer, gold particles adhere or are washed to the carbon support regions; however, gold-conjugated His-CREM particles are present on the monolayer only when DHGN is present (B). The black bar in (A) indicates 10 nm.

Table I. Binding of untagged and His-tagged protein–gold conjugates to control and DHGN-containing lipid monolayers

Protein	Total particles per 10 nl	Particles bound per 1000 nm ²	
		PC only	PC+DHGN
BSA	198	32	22
CytoC	163	6	1
His-CREM	228	19	369
His-ATF	117	12	117

Cytochrome C (cytoC), bovine serum albumin (BSA) and histidine-tagged proteins His-CREM, and His-ATF were conjugated to 20 nm gold particles, purified and used to assess binding to lipid monolayers. Total gold particle numbers were quantitated by drying 2 μ l of 1:10 dilutions of gold conjugate stocks onto carbon–formvar grids, and counting all gold particles in 200 EM fields at 250 000 $\times$ magnification (1.06 μ m²) on a Philips 301 TEM to calculate gold particles per 10 nl of undiluted stocks. To assess binding to monolayers, particles were incubated beneath lipid monolayers containing phosphatidylcholine (PC) only, or PC plus nickel-charged DHGN (PC+DHGN), as described in Materials and methods. After incubations, monolayers were lifted onto lacy carbon grids, rinsed for 30 s in 1% glucose, dried, viewed and photographed. Binding was quantitated by counting particles on hole-free, carbon support-free monolayer areas in 20 fields at 91 000 $\times$ magnification to obtain particles bound per 1000 nm².

formation of 2D His-MoCA arrays (see Materials and methods). A typical example of a uranyl acetate-stained array is shown in Figure 5. As illustrated, His-MoCA forms extensive 2D arrays over the entire surface of PC+DHGN monolayers with few, if any, protein-free zones. Using standard conditions (Materials and methods), His-MoCA arrays formed reproducibly, and tolerated minor variations in temperature (23–30°C), pH (7.6–8.3) and ionic strength (200–300 mM NaCl) of the subphase solution. Arrays optimally formed at His-MoCA subphase concentrations of >1 mg/ml, and never formed on PC-only monolayers or on DHGN-containing monolayers in which the DHGN had not been pre-bound to nickel. We also observed that the quality of arrays varied, depending on the age and storage conditions of the diluted working stock of PC plus nickel-charged DHGN, possibly as a consequence of lipid oxidation, although lipids were stored under nitrogen gas at –80° C.

Because ~10–20% of the EM negatives from uranyl acetate-stained His-MoCA samples yielded discernable optical diffraction patterns (data not shown), we next examined unstained His-MoCA arrays and found that, while glucose and tannin embedding did not produce satisfactory results, cryo-preservation in vitreous water

Table II. Variables affecting His-tagged gold particle binding to lipid monolayers

Variations	Binding as % of standard conditions
Lipid/salt	
0% DHGN	0
1% DHGN	68
20% DHGN	100
50% DHGN	16
20% DHGN (no nickel)	7
0 mM NaCl	58
250 mM NaCl ^a	100
500 mM NaCl	119
pH/buffer	
pH 9.0, phosphate	90
pH 8.5, phosphate	142
pH 7.8, phosphate ^a	100
pH 7.4, phosphate	112
pH 7.0, phosphate	76
pH 6.5, phosphate	96
pH 6.0, phosphate	135
pH 5.5, phosphate	133
pH 5.0, phosphate	130
pH 7.8, phosphate ^a	100
pH 7.8, Tris-HCl	118
pH 7.8, HEPES	70
Ion/miscellaneous	
No additions ^a	100
0.1 mM ZnCl ₂	112
5 mM MgCl ₂	81
1 mM CaCl ₂	25
1 mM FeCl ₂	22
5 mM NiCl ₂	0
10 mM EDTA	7
5 mM DTT	0
1 mM DTT	9
5 mM β -Me	0
1 mM β -Me	9

Gold particle conjugates to His-ATF (His-ATF-gold) were incubated under lipid monolayers using a variety of different conditions to ascertain how different variables affect His tag binding to DHGN-containing monolayers. Incubations were performed as described in Materials and methods, and were compared with results from standard condition incubations (incubations marked with ^a) performed in parallel. Lipid variations involved using different percentages of nickel-charged DHGN in incubations, or employed DHGN which had not been pre-bound with Ni²⁺. Salt variations were performed using different concentrations of NaCl in the subphase; pH and buffer variations used 25 mM buffer substitutions in place of the standard 25 mM phosphate, pH 7.8 buffer; and ion, EDTA and reducing agent incubations involved addition of these agents to the subphase solutions of standard incubations. Results are listed as percentages of gold particle counts per unit area obtained in the parallel standard incubations.

(see Materials and methods) retained the His-MoCA array order. Consequently, four negatives (#717, #719, #255 and #260) of unstained His-MoCA arrays obtained from two different microscopes were chosen for further analysis. As an initial step, diffraction patterns were computed from scanned regions on each negative, an example of which is provided in Figure 6. As shown, the diffraction pattern appears either roughly hexagonal, or orthorhombic with systematic absences at $h + k$ odd positions: with film #717, reflections are apparent out to ~ 9.5 Å resolution (Figure 6, circled reflections). All four films gave qualitatively similar diffraction patterns, yielding unit cells of $a = 79.2 \pm 0.4$ Å, $b = 137.5 \pm 1.7$ Å, $\gamma = 89.72 \pm$

Fig. 5. M-MuLV His-MoCA two-dimensional arrays. His-MoCA proteins were incubated beneath a PC+DHGN monolayer, lifted onto a lacy EM grid, washed, uranyl acetate stained, and photographed on the Philips CM12. The lacy carbon support appears as thick gray lines, and the 2D protein array occurs over the entire area of the micrograph, with zones of differential darkness in staining, and occasional dark protein aggregates. The black label (60.0K D.1.5.3.) indicates the magnification of the negative (60 000 $\times$), and its length corresponds to 417 nm.

0.89° , or $a = 79.4 \pm 0.6$ Å, $b = 80.0 \pm 1.3$ Å, $\gamma = 59.89 \pm 0.73^\circ$, depending on the choice of indexing (Table III). After unbending and contrast transfer function (CTF) correction (see Materials and methods), film #717 was used as a reference for merging of the other three films. Films #255 and #260, which showed lower quality diffraction patterns, merged reasonably well to 23 Å resolution, using a primitive (p1) unit cell, while film #719, gave a phase residual of 21.0° at 14 Å in the p1 merge. Consistent with the appearance of calculated diffraction patterns (Figure 5), phase residuals from merges in p2, C222, p3 and p6 also were $<25^\circ$ (Table III), while the best merging for other space groups consistently gave phase residuals of $>50^\circ$. These results suggest that His-MoCA crystals obey centered orthorhombic (C222) or hexagonal (p6) spacing, although tilted images or higher resolution data may lead to an assignment in a lower symmetry space group.

To visualize how His-MoCA proteins are ordered on PC+DHGN monolayers, the unbent, CTF-corrected, Fourier filtered amplitudes and phases files for the four films were back-transformed, with no symmetry constraints (p1). The results of these steps are shown as 2D projections in Figure 7, where regions of higher electron density are depicted with darker coloring. All of the projections show a cage-like network of proteins (dark),

surrounding protein-free cage holes (light areas). In all cases, the projections show at least two types of cage holes: one hole type appears circular with a diameter of

Fig. 6. Calculated diffraction pattern for an unstained His-MoCA 2D crystal. An unstained His-MoCA crystalline area from a cryo sample frozen in vitreous ice (negative #717) was scanned at 2.67 Å/pixel, transformed, and displayed as a diffraction pattern, which appears hexagonal or orthorhombic, with systematic absences. The innermost six reflections are barely visible in this contrast range, while the six brightest reflections correspond to the 1,1; -1,2; -2,1; -1,-1; 1,-2; and 2,-1 reflections if $\gamma = 60^\circ$, or the 2,0; 1,3; -1,3; -2,0; -1,-3; and 1,-3 reflections if $\gamma = 90^\circ$. The two faint circled reflections are at ~ 9.5 Å resolution.

~ 19.2 Å, while the other is larger and roughly triangular (length 28.0 Å, width 23.2 Å). While films #255 and #260 (Figure 7A and B) do not indicate a packing arrangement of His-MoCA proteins around cage holes, the other two films show six electron-dense masses, possibly His-MoCA monomers (see Discussion) surrounding each cage hole. These results suggest that Gag protein hexamers serve as intermediates for M-MuLV assembly, and are schematized in Figure 7G. As shown, monomers are depicted as asymmetric hexagons with faces numbered 1–6 and inter-protein contacts at the 1–1 and 5–3 interfaces. Such an arrangement naturally leads to the occurrence of two types of cage holes, and the symmetry axes observed in the projections. The implications of these results with regard to virus assembly will be considered below (see Discussion).

Protein expression in mammalian cells

We initially have focused on arrays of the His-MoCA protein because it may serve as a basis for comparison with larger Gag-derived proteins in the future. As noted above, we envisage that His-MoCA interactions should provide information relating to capsid domain interactions in assembling or immature M-MuLV particles. However, since His-MoCA lacks matrix, p12 and nucleocapsid domains, it was important to establish how such deletions might affect Pr^{gag} proteins *in vivo*. To do so, we examined the expression of wild-type (wt) and deleted M-MuLV Gag proteins in transiently transfected mammalian cells. The plasmids used for transfections were as follows: pXMGPE, which encodes wild-type (wt) M-MuLV gag and pol genes; pXM2453T, which expresses an unprocessed, protease-minus (PR-) Gag protein; pXMG-

Table III. Unit cell parameters for His-MoCA 2D arrays

Film	717	719	255	260
Orthorhombic ^a unit cell dimensions	$a = 79.6$ Å $b = 137.3$ Å $a/b = 0.58$ $\gamma = 89.26^\circ$	$a = 78.9$ Å $b = 138.5$ Å $a/b = 0.57$ $\gamma = 90.48^\circ$	$a = 79.1$ Å $b = 138.9$ Å $a/b = 0.57$ $\gamma = 90.45^\circ$	$a = 80.8$ Å $b = 135.1$ Å $a/b = 0.60$ $\gamma = 90.00^\circ$
Hexagonal ^a unit cell dimensions	$a = 80.0$ Å $b = 79.1$ Å $a/b = 1.01$ $\gamma = 60.22^\circ$	$a = 80.0$ Å $b = 79.1$ Å $a/b = 1.01$ $\gamma = 60.22^\circ$	$a = 78.6$ Å $b = 80.3$ Å $a/b = 0.98$ $\gamma = 58.8^\circ$	$a = 79.0$ Å $b = 81.8$ Å $a/b = 0.97$ $\gamma = 60.31^\circ$
Phase residual for p1 merge		23 Å = 10.2° 17 Å = 15.3° 14 Å = 21.0°	23 Å = 18.9°	23 Å = 19.4°
Phase residual for p2 merge		23 Å = 10.8° 17 Å = 17.4° 14 Å = 21.6°	23 Å = 19.8°	23 Å = 19.9°
Phase residual for C222 merge		27 Å = 13.2°	27 Å = 15.2°	27 Å = 14.9°
Phase residual for p3 merge		23 Å = 16.8° 17 Å = 18.0° 14 Å = 19.9°	23 Å = 22.5°	23 Å = 14.4°
Phase residual for p6 merge		23 Å = 15.5° 17 Å = 16.3° 14 Å = 18.4°	23 Å = 23.6°	23 Å = 17.8°

Electron micrographs of unstained His-MoCA arrays were taken on the University of Oregon CM12 at 60 000 $\times$ magnification and scanned on the Baylor flat bed scanning densitometer (films 717 and 719), or on the EMBL CM200 and CCD-scanned (films 255 and 260). Crystalline areas from scanned images were boxed, transformed, displayed as diffraction patterns on SPECTRA (Schmid *et al.*, 1993) and indexed by hand as roughly orthogonal (orthogonal^a; $\gamma \sim 90^\circ$) or roughly hexagonal (hexagonal^a; $\gamma \sim 60^\circ$): unit cell dimensions derive from SPECTRA information headers of the raw, hand-indexed diffraction patterns. For merging, unbent transforms were used to calculate amplitudes and phases for reflections, and phases were CTF-corrected as described in Materials and methods. Film 717 was chosen as a reference film, phase origin-centered, and merging was performed using ORIGIN/TILT for space groups p1, p2, C222, p3 and p6. Phase residuals at given resolutions are as shown, and residuals for other space groups or C222, p3 and p6 using an uncentered reference were $>50^\circ$ at 23 or 27 Å resolution.

Fig. 7. Two-dimensional projections of His-MoCA arrays. Films #255 (A), #260 (B), #719 (C) and #717 (D–F) were scanned, transformed, unbent, CTF-corrected, filtered and back-transformed with no symmetry constraints to yield 2D projections (see Materials and methods). As shown, the darker regions correspond to areas of greater electron density, and white areas are indicative of protein-free ‘holes’ in the cage-like network. Corner brackets outline putative orthorhombic unit cells of $79.2 \text{ \AA} \times 137.5 \text{ \AA}$, and $20 \text{ \AA}$ marker bars are indicated. (E) is a larger scale version of (D), (F) is a contour map of (E), and (G) shows a schematic interpretation of the projections shown in (A–F). In this panel, capsid monomers are drawn as hexagons where the intrinsic asymmetry of each putative protein monomer is indicated by six different faces, 1–6. Bold small hexagons and triangles represent 6-fold and 3-fold crystallographic axes, respectively. Additional 2-fold axes are located at the contact (1–1) faces of the symmetric dimer (not shown). The dashed circles and triangles indicate the two types of pores, formed by the cage-like network of monomers. The two different distances between capsid monomers indicate the two different crystal contacts. Note that assembly of the network starting from either the symmetric or asymmetric dimer leads to the same symmetry in the final network.

Fig. 8. Assembly and release of Gag proteins from transfected mammalian cells. Cos7 tissue culture cells were transfected with the M-MuLV Gag protein expression constructs pXMGPE (lanes A and C), pXMΔMA+p12 (lanes B and D), pXM2453T (lanes E and G) and pXMpetMoCA (lanes F and H). Three days post-transfection, cell lysate (lanes A, B, E and F) and virus pellet (lanes C, D, G and H) samples were collected, fractionated by SDS-PAGE, electroblotted onto nitrocellulose and immunodetected with polyclonal goat anti-M-MuLV-CA antisera (lanes A–D) or monoclonal anti-M-MuLV-CA antibody (lanes E–H). Cell samples applied to gels corresponded to one-sixth of the total samples, while viral samples corresponded to one-half of the viral pellet. Mature capsid proteins produced from the pXMGPE and pXMΔMA+p12 constructs are indicated (CA), as are wt Pr^{gag} proteins produced by pXMGPE and pXM2453T (PrGag^{wt}), matrix plus p12-deleted Pr^{gag} proteins (PrGag^{ΔMA+p12}) and nucleocapsid-deleted Gag proteins (PrGag^{MoCA}). Vector-associated bands of 55 and 43 kDa in lanes A and C are also indicated. Estimates of the ratios of total Gag protein levels in total virus versus total cell samples, as a measurement of virus release, are as follows: pXMGPE, 4.0; pXMΔMA+p12, 3.3; pXM2453T, 5.0; pXMpetMoCA, 0.2.

PEΔMA+p12, encoding a myristylation signal-positive matrix plus p12-deleted version of pXMGPE; and pXMpetMoCA, a PR[−] vector with MA, p12 and CA domains, but which carries the same NC deletion and C-terminal Gag sequence as does the His-MoCA protein. Thus, along with wt and PR[−] controls, pXMGPE-ΔMA+p12 and pXMpetMoCA allowed us to assess the potential effects of N- and C-terminal deletions on Gag proteins expressed in cells.

For analysis, plasmids were transfected into Cos7 cells (Hansen and Barklis, 1995) and, 3 days post-transfection, virus particle and cell lysate samples were collected for Gag protein quantitation by anti-Gag immunoblotting of SDS-PAGE gels. Figure 8 shows the results of such analysis. Not surprisingly, cell lysate samples showed the expected 65 kDa Pr^{gag} proteins for the control wt (lane A) and PR[−] (lane E) constructs. With the polyclonal anti-CA antibody used with wt pXMGPE, we also observed a virus-encoded band at 55 kDa (possibly a partially processed MA-p12-CA moiety), and lower molecular weight cross-reactive bands (lane A). Consistent with our size expectations, cell samples for pXMGPEΔMA+p12 and pXMpetMoCA had Gag-reactive bands at ~40 and 57 kDa respectively (lanes B and F). With the wt virus pellet sample, three reactive bands (Pr^{gag}, CA and a 43 kDa band) were observed (lane C), and the total Gag protein level associated with virus particles was higher than that in transfected cells (compare lanes C and A). Similar

results were observed with pXM2453T (lane G versus E) and pXMΔMA+p12 (lane D versus B). The fact that the ΔMA+p12 protein appeared to assemble efficiently into virus particles was not surprising, since Δp12 M-MuLV Gag proteins previously have been shown to assemble virus particles (Crawford and Goff, 1984; Hansen and Barklis, 1995), and matrix-deleted HIV-1 assembles conditionally infectious virions (Wang *et al.*, 1993).

In contrast to the other three Gag variants, the NC-deleted protein expressed from the pXMpetMoCA construct was released from transfected cells at low efficiency, i.e. ~10- to 20-fold lower than the wt Gag protein (compare Figure 8, lanes H and F). Perhaps the pXMpetMoCA protein was not routed appropriately to the plasma membranes of transfected cells. Alternatively, budding and release of the plasma membrane-localized protein might have been affected. To help distinguish between these possibilities, we performed immunofluorescent localization of Gag proteins in Cos7 cells transfected with either pXM2453T or pXMpetMoCA, expressing wt and NC-deleted Pr^{gag} proteins, respectively. Interestingly, the localization patterns of the two proteins showed significant differences (see Figure 9). The wt Pr^{gag} proteins demonstrated a punctate staining pattern over the entire area of staining cells (Figure 9A–D and E–H). In contrast, the petMoCA proteins showed a bright, relatively homogeneous staining pattern, with occasional patches of extremely bright fluorescence at cell edges (Figure 9I–L

Fig. 9. Immunofluorescent localization of Gag proteins in mammalian cells. Cos7 cells on coverslips were transfected with either pXM2453T (A–H) or pXMpetMoCA (I–P), grown for 3 days and processed for immunofluorescent localization of Pr^{gag} proteins using an anti-p12 monoclonal antibody (Chesebro *et al.*, 1983; Hansen *et al.*, 1993) as the primary antibody. Cells were photographed on a Leica confocal scanning microscope, equipped with a Leitz microscope, and an argon/krypton laser using a Leitz 40× oil immersion objective and rhodamine filters. Each set of images (A–D, E–H, I–L and M–P) was taken in four 1.25 μm slices from the apical (A, E, I and M) to basolateral (D, H, L and P) surfaces. Fluorescence voltage settings were 950 for (A–H) and 760 for (I–P). The size bar in (P) indicates 50 μm.

and M–P). We interpret these observations to indicate that the petMoCA proteins accumulate at the plasma membranes of transfected cells, where their ability to form budding virus particles is reduced, relative to wt. These results are consistent with notion that retroviral nucleocapsid domains serve a function in the late stages of virus particle assembly (Spearman *et al.*, 1994; Wills *et al.*, 1994; Hansen and Barklis, 1995). They also suggest that while capsid protein interactions observed in our *in vitro*-generated His-MoCA arrays may be similar to those of Gag proteins at the inner face of the plasma membrane, they are not identical to those in budded immature M-MuLV particles. These observations are discussed below.

Discussion

We (Zhao *et al.*, 1994) and others (Schmitt *et al.*, 1994; Dietrich *et al.*, 1995, 1996; Kubalek *et al.*, 1995; Ng *et al.*, 1995; Frey *et al.*, 1996) have adapted a lipid monolayer system (Uzgiris and Kornberg, 1983) for the crystallization of His-tagged proteins, using nickel-binding lipids. The advantage of such a system is that potentially any His-tagged protein could be crystallized in two dimensions for structural analysis. Our lipid, DHGN, differs from nickel-chelating lipids developed by others in that it combines the NTA (Hochuli *et al.*, 1987) nickel-binding headgroup and a dialkyl glycerol tail (Thompson *et al.*, 1994). We have demonstrated that DHGN can bind nickel

(Figure 3), and that nickel-charged, DHGN-containing lipid monolayers bind His tag-conjugated gold particles (Figure 4, Table I). Gold particle studies appeared to be a convenient method by which to assess potential DHGN–His tag ligand interactions, and such experiments suggested that binding was tolerant of salt and buffer variations, but that divalent cations and reducing agents might impair binding (Table II). However, as 2D crystallization is not merely a matter of membrane affinity, the effects of cations and reducing agents ultimately must be tested in the context of each His-tagged protein.

Does DHGN facilitate the 2D crystallization of His-tagged proteins? In this regard, it is noteworthy that we have not observed His-MoCA arrays on monolayers consisting of only PC, or PC plus DHGN that had not been pre-incubated with nickel. We also have examined other His-tagged proteins and, with surprisingly little effort, apparent 2D arrays have been obtained with a number of them (HIV-1 CA, HIV-1 CA+NC, M-MuLV CA+NC, ATF), although the arrays are not yet as large as those of His-MoCA. Nevertheless, DHGN-containing monolayers are not universally applicable for His tag protein crystallization, as several His-tagged proteins (notably M-MuLV MA and CREM) have shown no evidence of array formation. An additional difficulty is that diluted, working stocks of lipid lose their ability to foster crystallization over time, possibly as a consequence of oxidation. Finally, one must emphasize that, while it is natural to assume that the His tag–nickel interaction leads to the monolayer binding that precedes array formation (Kubalek *et al.*, 1994; Schmitt *et al.*, 1994; Zhao *et al.*, 1994; Ng *et al.*, 1995; Dietrich *et al.*, 1995, 1996; Frey *et al.*, 1996), less obvious associations between proteins and lipids may dictate the process (Celia *et al.*, 1994; Olofsson *et al.*, 1994; Kubalek *et al.*, 1995). Despite these reservations, monolayers with an affinity for His tags appear to have great potential for analysis of histidine-tagged proteins.

Obviously, *in vitro* monolayer studies cannot substitute perfectly for studies on virus structures since cellular components and modified viral proteins are missing; there may be geometric restrictions on proteins at a monolayer versus a membrane which permits virus budding; and oligomerization may be driven by the lipid–protein complex, and could be dependent on the specific type of carrier lipid. However, several factors have suggested that lentivirus and C-type retroviral Gag proteins might be ideal candidates for lipid monolayer studies: they localize to the inner face of the plasma membrane, they oligomerize during assembly at the membrane, N-terminal deletion mutants are capable of virus particle assembly (Crawford and Goff, 1984; Faecke *et al.*, 1993; Wang *et al.*, 1993; Hansen and Barklis, 1995; Figure 8), and it has been possible to substitute the membrane-binding function of MA with that of a His tag. In our current study, we have focused on the M-MuLV capsid domain, which mediates critical Gag–Gag contacts in Pr^{gag} proteins (Hansen *et al.*, 1990; Jones *et al.*, 1990; Hansen and Barklis, 1995). The His-MoCA protein lacks the M-MuLV NC domain, and consequently cannot mimic NC and RNA associations which participate during virus particle assembly. Indeed, the reduced assembly efficiency of the pXMpetMoCA protein (Figure 8), and its localization pattern (Figure 9) are reminiscent of M-MuLV temperature-sensitive assembly

mutants (Weiss *et al.*, 1984): it is possible that these characteristics may contribute to the facility with which His-MoCA forms crystals.

Realizing the above caveats, we have undertaken the analysis of His-MoCA arrays as a necessary step in the analysis of membrane-bound Gag proteins. As shown in Figure 5, His-MoCA forms large arrays on DHGN-containing monolayers, and medium-resolution diffraction analysis suggests that MoCA crystals can be classified in C222 or hexagonal space groups (Figure 6, Table III). The results of this model system suggest that retroviral Gag proteins assemble cage-like networks at membrane faces (Figure 7). Such a cage is remarkable in its general similarity to the clathrin caging of endocytic vesicles, and to the ‘fullerene’ networks formed by HIV-1 Pr^{gag} proteins assembled on plasma membranes (Nermut *et al.*, 1994). Interestingly, the cage networks formed by HIV-1 Pr^{gag} were modeled as consisting of hexamer rings, with each ring sharing one subunit with a neighboring ring (Nermut *et al.*, 1994). Based on this model, Pr^{gag} subunits were modeled as rods of length 85 Å, cross-sectional diameter 34 Å, with cage hole–hole distances of 68 Å and an estimated 1890 Pr^{gag} molecules per virion (Nermut *et al.*, 1994).

Our model with the His-MoCA protein differs from that of Nermut *et al.* (1994) in that we observed two types of hexamer rings, sharing two subunits each, and creating two different types of cage holes (Figure 7). Cages had both circular holes with projection diameters of ~19.2 Å and a hole–hole spacing of 49 Å, as well as roughly triangular holes with lengths of 28.0 Å and widths of 23.2 Å (Figure 7). Assuming that hexamer subunits are His-MoCA monomers, they may be structured as rods of 26 Å diameter and 62–68 Å in length, given a specific volume of 0.74 cm³/g protein. All of the features observed in our projections are consistent with the scheme presented in Figure 7G, in which monomers are modeled with six faces with homodimer (1–1) and heterodimer (5–3) interprotein contacts. In this case, circular protein-free holes are bordered by six identical (4) faces, while triangular holes are edged by the monomer 2 and 6 faces. This scheme is compatible with array assembly by two pathways: (i) building of hexamer rings from monomers via (5–3) contacts, followed by network assembly through 1–1 contacts; or (ii) homodimer formation at the 1–1 faces, followed by network assembly mediated by 5–3 contacts. We prefer the latter (ii) pathway, because purified HIV and M-MuLV capsid proteins form stable dimers in solution, and retrovirus Gag cross-linking studies have revealed dimers, tetramers and higher-order oligomers, but a relative lack of the trimers (Ehrlich *et al.*, 1992; Hansen and Barklis, 1995) that would be expected from pathway (i).

Based on the projections in Figure 7, when immature M-MuLV particles are modeled as spheres with inner diameters of 110 nm, we calculate there would be 3960 protein monomers, 660 small cage holes and 1320 large holes per virion. This estimate assumes a matrix cross-sectional area similar to that of capsid, and is greater than what would be needed to pack an icosahedron of the same approximate size. While some studies suggest that immature retrovirions display icosahedral symmetry (Nermut *et al.*, 1994; Klikova *et al.*, 1995), we have no

incontrovertible evidence concerning such an interpretation. Our system may be biased against the formation of an icosahedral-compatible lattice since MoCA lacks the NC domain, or because we have constrained His-MoCA protein interactions within a planar monolayer. Alternatively, if immature M-MuLV particles have a high triangulation number, pentavalent units may be too rare to observe here. However, recent analyses of cryo-electron micrographs of immature HIV and HIV Gag particles have shown that particles are not icosahedral, but do display an ordered arrangement of Gag proteins with unit cell parameters similar to those reported here (S. Fuller, T. Wilk, H.-G. Krausslich and V. Vogt, in preparation). This suggests that the local arrangement of Gag proteins in the immature virus is dominated by interactions which are present in the crystals. As regards mature retrovirus Gag protein interactions, we realize that NC and RNA have been excluded in our studies, but conjecture that rod- or cone-shaped structures observed in some mature retrovirus types (Campbell and Vogt, 1995) may consist of hexamer rings of CA and NC wrapped around viral RNA.

Perhaps the most intriguing speculation resulting from our His-MoCA cage structure is that observed cage holes seem like a plausible site for the positioning of retrovirus envelope (Env) protein transmembrane regions and cytoplasmic tails. This hypothesis was forwarded by Nermut *et al.* (1994), who have modeled HIV-1 such that a fraction of all possible cage holes are Env protein-occupied, and is reminiscent of rotavirus VP4 insertion into capsid holes formed by the VP6 and VP7 proteins (Shaw *et al.*, 1993). Since His-MoCA lacks the M-MuLV Pr^{gag} MA and p12 domains, extrapolation of Env-Gag interactions in assembling immature virions may not be pertinent. However, since capsid domain interactions are a driving force in particle assembly (Hansen *et al.*, 1990, 1993; Jones *et al.*, 1990; Wang *et al.*, 1994; Hansen and Barklis, 1995), and matrix- and p12-deleted virions assemble with reasonable efficiency (Faecke *et al.*, 1993; Wang *et al.*, 1993; Hansen and Barklis, 1995; Figure 8), it does not seem unreasonable to expect that similar cage holes occur in assembling Pr^{gag} arrays. If this is the case, we would expect there to be a maximum of one glycoprotein knob per 12 Pr^{gag} monomers, if circular holes are filled, or two knobs per 12 monomers, if triangular holes are used. Our cage model permits several predictions: immature particle cores may be accessible to small lipophilic compounds; hexamers ought to form particle subunits and/or assembly intermediates; cage hole size should serve as an exclusion limit for entry of Env or other membrane proteins into virus particles; and Env proteins should pre-exist on membranes at which Gag proteins assemble cage networks. Some support for these predictions exists, in that studies have shown that NC domains in immature M-MuLV and HIV-1 particles can be cross-linked to form dimers, tetramers and pentamers or hexamers with bis-maleimide hexane (Hansen and Barklis, 1995; McDermott *et al.*, 1996). Additionally, differences in cage hole sizes may explain why M-MuLV Env proteins can be incorporated non-specifically into HIV particles (Wang *et al.*, 1993), but the reverse is not true. We anticipate that further study of Gag protein monolayer arrays combined with biochemical analysis of *in vivo*-derived particles will

provide a detailed view of retrovirus particle structures, and the mechanism of particle assembly.

Materials and methods

Lipid synthesis

A convergent synthetic sequence (see Figure 2) was used to prepare DHGN (compound 6) from 1,2-di-*O*-hexadecyl-*sn*-glyceryl-3-*R*-glycidyl ether (compound 5) and *N*-(5-amino-1-carboxypentyl)-iminodiacetic acid (compound 3). Compound 3 was prepared as described previously (Hochuli *et al.*, 1987). Briefly, *N*^ε-benzyloxycarbonyl-L-lysine (compound 1; Fluka) was acetylated with bromoacetic acid (Fluka) in 2 M sodium hydroxide to give *N*-(5-benzyloxycarbonylamino-1-carboxypentyl)-iminodiacetic acid (compound 2). Compound 2 was dissolved in 1 M sodium hydroxide and hydrogenated in the presence of Pd/C to yield the intermediate compound 3. For preparation of compound 5, 0.44 g (0.61 mmol) of 1,2-di-*O*-hexadecyl-*sn*-glycero-3-(3'-nitrobenzenesulfonate) (Thompson *et al.*, 1994) was hydrolyzed with aqueous tetrabutylammonium hydroxide in THF (1.5 ml, 2.3 mmol) to give 1,2-di-*O*-hexadecyl-*sn*-glycerol (compound 4), which was alkylated for 20 h at 22°C with (*R*)-glycidyl-3-nitrobenzenesulfonate under basic conditions (50 mg NaH in 5 ml THF) to give compound 5 in 98% yield. Condensation of compounds 5 and 3 in an aqueous CTAB (cetyltrimethyl ammonium bromide) solution yielded an oily crude product. Silica gel chromatography (first with 1:1 chloroform:methanol, followed by a second column using a chloroform:0–15% methanol gradient elution) produced DHGN as a white powder in 26% overall yield. Structures of compounds 5 and 6 were confirmed by IR and NMR spectroscopy (chemical shifts reported in p.p.m. using CDCl₃ as solvent) with results as follows: IR (compound 5): 2955, 2918, 2850, 1467, 1377, 1259 and 1118/cm; ¹H NMR (compound 5): 3.96 (m, 1H), 3.9–3.68 (m, 1H), 3.66–3.52 (m, 8H), 3.50–3.34 (m, 4H), 1.55 (s, 4H), 1.26 (s, 52H), 0.87 (t, 6H); ¹³C NMR (compound 5): 78.3, 72.7, 72.3, 72.1, 71.1, 70.8, 70.6, 32.1, 30.2, 30.0, 29.9, 29.8, 29.7, 29.6, 29.5, 26.3, 22.8, 14.2; IR (compound 6): 3409, 2924, 2854, 1735, 1467, 1378, 1215, 1113/cm; ¹H NMR (compound 6): 3.86 (m, 2H), 3.82–3.30 (m, 14H), 2.37 (t, 2H), 1.60 (m, 4H), 1.28 (m, 62H), 0.88 (t, 6H).

Atomic absorption spectrometry

Assay of the nickel-binding capabilities of the lipids DHGN, DHP (Aldrich) and OG (Boehringer) was achieved by atomic absorption spectrometry using a Perkin Elmer 603 spectrophotometer equipped with an HGA-400 programmer and a graphite furnace. Nickel sulfate standards were made as 0–100 μM solutions in distilled water, and 20 μl samples were flash-atomized at 2500°C, and assayed using a hollow cathode lamp set at a current of 10 mA, a wavelength of 232.0 nm and a slit width of 80 μm. Under these conditions, chloroform, lipid solutions and 50 mM EDTA in distilled water showed no absorbance over distilled water absorbance levels, and absorbance levels of nickel in water or chloroform were comparable. For analysis of nickel binding to lipids, 0.1 ml of lipid solutions (250 mM stock) in chloroform were overlaid with an equal volume of 20 μM nickel sulfate in distilled water. Samples were rocked for 1–18 h (mixing times within these limits did not appear to affect results) at room temperature, after which phases were separated by low-speed centrifugation. Absorption levels of triplicate 20 μl samples from aqueous (top) and organic (bottom) phases were measured and compared with calibration standards to obtain nickel concentrations in each phase. In competitive binding experiments to determine the nickel affinity of EDTA versus DHGN, EDTA first was added to the aqueous (Ni²⁺) phase to a final concentration of 50 mM prior to mixing with the DHGN–chloroform phase.

Protein purification and analysis

The M-MuLV CA protein was expressed as a His-tagged protein (His-MoCA) in *E. coli* strain BL21(DE3)/pLysS (Novagen) from the bacterial expression plasmid pET15B-MoCA. This plasmid contains a M-MuLV capsid-coding region cassette inserted into the *Bam*HI site of pET15B (Novagen). The cassette was constructed by PCR generation of a *Bam*HI site at the amino-terminus of the CA coding region and addition of a *Bam*HI linker at the carboxy-terminal *Msc*I site of this region. The respective N- and C-terminal juncture sequences of the cassette are GGAT/CCC, where the bold C is M-MuLV viral nucleotide 1266 (Shinnick *et al.*, 1981), and TTG/GCGGATCC, where the bold G is viral nucleotide 2055. When expressed in bacteria, the His-MoCA protein has an amino-terminal leader sequence of MGSSH HHHHH SGLV

APRGS HMLGD and a carboxy-terminal tail of ADPAA NKARK EAELA AATAEQ. The His-MoCA protein was purified from bacterial lysates by two cycles of standard non-denaturing nickel chelate chromatography (Hochuli *et al.*, 1987), and fraction purity was assessed by a combination of Coomassie staining and immunoblotting of electrophoretically separated proteins (Hansen *et al.*, 1990, 1993; Jones *et al.*, 1990; Zhang and Barklis, 1995). Once pure fractions were identified, they were desalted by two 2 h dialysis steps versus distilled water at 4°C, lyophilized, resuspended in distilled water and stored frozen at -80°C. Alternatively, highly concentrated fractions were desalted on Sephadex G25 spin columns equilibrated with 5 mM Tris pH 7.8, and stored at 4°C in 5 mM Tris, 16% glycerol, 0.02% sodium azide. Each desalting method yielded final protein concentrations of 0.2–2.0 mg/ml, and both appeared equally successful in 2D crystallization experiments.

Because the M-MuLV capsid protein has no enzymatic activity by which to monitor its native state, assessment of the conformational state of His-MoCA was made by comparison of its partial proteolysis profile with that of the capsid protein from M-MuLV virus particles. To do so, 1 µg samples of His-MoCA or equivalent samples (based on capsid immunoblotting) of purified M-MuLV particles (Hansen *et al.*, 1993) were incubated for 1 h at 30°C in TSE+T (10 mM Tris, pH 7.4; 100 mM NaCl, 1 mM EDTA, 0.25% Triton X-100) plus 0, 1, 10 or 100 µg/ml TPCK-treated trypsin (Sigma T-8642). Following incubations, phenylmethylsulfonyl fluoride (PMSF, 5 mM final concentration) was added and the samples were mixed with one volume of 2× sample buffer (Jones *et al.*, 1990), prior to heating at 100°C for 3–5 min. Capsid-derived protein fragments in samples were detected after 12.5% SDS-PAGE by immunoblotting (Hansen and Barklis, 1995) using a combination of rat monoclonal anti-M-MuLV-CA antibody HyR187 (Chesebro *et al.*, 1983) and polyclonal goat anti-M-MuLV-CA (National Cancer Institute) as the primary antisera.

Expression and detection of M-MuLV Gag proteins in mammalian cells

Plasmids for expression of M-MuLV Gag proteins in mammalian cells were constructed using standard methods (Maniatis *et al.*, 1982) from pXMGPE, a vector for transient expression of M-MuLV Gag, Pol and Env proteins in Cos7 cells, and pXM2453T, a related vector which expresses the full-length M-MuLV Gag protein (Pr65^{gag}), but not Pol or Env proteins (Hansen and Barklis, 1995). The previously described (Hansen and Barklis, 1995) vector pXM2051ΔNC is a derivative of pXM2453T, which has a complete deletion of the NC-coding region of Pr65^{gag}, and consequently expresses a truncated MA-p12-CA Gag protein. In addition to pXMGPE, pXM2453T and pXM2051ΔNC, the expression plasmids pXMpetMoCA and pXMGPEΔMA+p12 also were used. The plasmid pXMpetMoCA derives from pXMGPE, has intact MA-, p12- and CA-coding regions, is deleted for NC (and Pol and Env), and carries the CA C-terminal tail that is present on the plasmid pet15B-MoCA: it was generated by cloning the *XhoI*-*Clal* pet15B-MoCA fragment (encoding the CA C-terminal coding region) in place of the M-MuLV-derived *XhoI*-*Clal* fragment in pXMGPE. The vector pXMGPEΔMA+p12 is identical to pXMGPE, except that it possesses a deletion of all of the p12-coding region and most of MA, although it retains the amino-terminus of MA as well as the Gag myristylation signal. The deletion juncture sequence is 5' TTA ACG GAT CCC 3', where the 5' T and 3' C correspond to M-MuLV viral nucleotides 645 and 1268, respectively; the predicted pXMGPEΔMA+p12 Gag codons prior to the capsid coding region are MGQTV TTPLTD.

For analysis of pXM vector expression in Cos7 (African green monkey) cells, plasmids were transiently transfected into cells, and cell lysate and virus pellet samples were collected, fractionated by SDS-PAGE and immunoblotted to detect CA and Pr^{gag} proteins as described previously (Hansen and Barklis, 1995). Immunoblotting was performed using either rat monoclonal anti-M-MuLV-CA antibody HyR187 or mouse monoclonal anti-M-MuLV-p12 antibody Hy548 (Chesebro *et al.*, 1983) as the primary antisera. Immunolocalization studies were performed on transiently transfected cells following standard procedures (Hansen *et al.*, 1990, 1993; Wang *et al.*, 1994), using anti-M-MuLV-p12 antibody Hy548 as the first antibody and rhodamine-conjugated goat anti-mouse IgG (TAGO) at a 1:200 dilution as the second antibody. Stained samples were viewed on a Leitz Dialux 22 microscope or by confocal microscopy, using a confocal laser scanning microscope (Leica Lasertechnik GmbH, Heidelberg, Germany) equipped with a microscope (Fluorvert-FU; E. Lietz, Inc., Rockleigh, NJ), an argon/krypton laser, a Leitz 40× oil immersion objective, and rhodamine excitation and long-pass barrier filters. Images initially were stored on a Motorola 68040 computer using System OS9 2.4 software (Microware Systems Corp.,

Des Moines, IA), and were prepared for printing using Adobe Photoshop 2.0.1 on an Apple Performa 636CD.

Electron microscopy

Four electron microscopes were used in these studies. For routine screening and gold particle studies, the OHSU Pathology Department Philips 301 was used at an accelerating voltage of 60 kV and ambient temperature. Low-dose work was at ambient temperature for stained samples, or at -170°C using Gatan freeze stages equipped on the Portland VA Hospital JEOL JEM1200EX operated at 100 kV; the University of Oregon CM12 operated at 100 kV; or the EMBL-Heidelberg Philips CM200-FEG operated at 200 kV. For low dose photography using Kodak S0163 film, the search mode was performed at 1000–5000× magnification, focusing was at 100 000–150 000× and micrographs were taken at 30 000–60 000×, at 400–800 nm defocus for stained samples or 600–1000 nm defocus for unstained cryo samples (dose = 10–25 e Å⁻²).

Standard lipid monolayer incubations for crystallization of the bacterially expressed M-MuLV capsid protein, MoCA, followed previously established protocols (Uzgisir and Kornberg, 1983; Darst *et al.*, 1991a,b; Celia *et al.*, 1994) with a variety of modifications. For our standard lipid mix, a working (1×) stock of 200 µg/ml phosphatidylcholine (Avanti Polar Lipids) plus 50 µg/ml nickel-charged DHGN was used in 1:1 hexane:chloroform from 10× stocks. In some control experiments, the final proportion of DHGN in the 1× lipid mix varied from 0 to 50%, with the total lipid weight remaining constant. The 10× nickel-charged DHGN stock was prepared by mixing 2 µl of 100 mM nickel chloride (1 mM final) with 200 µl of 0.5 mg/ml DHGN (0.6 mM final) in 1:1 hexane:chloroform at room temperature for 1 h. Mixing was accomplished during this period by rocking on a rocker platform, and vortexing the mix three times for 30 s at 20 min intervals. Note that although all lipids were stored at -80°C under nitrogen gas, we found that working 1× stocks went bad for unknown reasons after 3–12 weeks, as observed by poor protein crystallization.

To set up monolayer incubations, 5 µl of 0.2–2.0 mg/ml MoCA protein in distilled water was mixed with an equal volume of 2× subphase buffer to yield a final mix 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), although considerable variations of subphase buffer conditions were employed in control experiments by alteration of 2× buffers (see Table II). Following the preparation of protein plus subphase buffer mixes, 10 µl drops were applied to 5 mm wells of ethanol-cleaned depression well slides (Carlson Scientific Inc., #101005), after which freshly vortexed 1× lipid mix (pre-loaded with Ni²⁺) was pipetted onto the surface of each drop. Slides with drops were placed on parafilm which was placed on top of pre-wetted filter paper (4 ml of distilled water) in plastic 150 mm Petri plates. Lids were placed on the plates and the plates were sealed with strips of parafilm prior to overnight incubations, which usually were at 30°C, although room temperature incubations also were successful.

After incubations, arrays were transferred to lacy carbon EM grids (Ted Pella 300 mesh copper 'Netmesh,' cat. no. 01883; or lacy carbon prepared on Ted Pella 300 mesh copper grids; or on Graticules, Ltd. 300 mesh, copper, 3.0 mm, 03DOO658HF35 grids) by placing grids on drop surfaces for 10–15 s. For samples to be stained, grids with adherent arrays were placed for 30–45 s on 50–100 µl water drops, wicked from the side, stained for 30 s on drops of 1.33% uranyl acetate (freshly diluted and filtered), wicked again and air dried. For unstained samples, arrays were lifted onto grids as above, transferred for 30–45 s onto water drops, carefully lifted with forceps, transferred to a cryo apparatus, wicked with filter paper under a stream of humidified, room temperature air, plunged into liquid ethane and stored in liquid nitrogen until viewing. Although alternate strategies were attempted for preparation of unstained samples (glucose or tannin embedding), they did not appear to give satisfactory results with our samples.

Gold particle studies

Standard protocols (Geoghegan and Ackerman, 1977; Horisberger and Clerc, 1985) were modified for preparation of gold particle conjugates to cytochrome *c* (Sigma), BSA (Sigma), purified His-ATF and purified His-CREM. Using 10 nm gold colloid (ICN), protein concentration and pH optima for conjugations were determined in small scale reactions (50 µl of colloid, 10 µl of protein in distilled water) by observation of color changes after addition of 10 µl of 10% NaCl. For large scale conjugations, 5 ml of colloid (pH optimized) was vortexed for 10 s with the optimized amount of protein in 0.2 ml of distilled water, and incubated for 10 min at room temperature. At 10 min, 5.1 µl of 10%

polyethylene glycol (PEG; 15 000–20 000) was added and vortexed. Large aggregates were removed by low speed centrifugation (20 min at 4°C; 2000 g), after which free protein was removed from the conjugates by centrifugation. To do so, the supernatant was centrifuged at 4°C for 45 min at 75 000 g (Beckman SW50.1; 25 000 r.p.m.), after which the gold conjugate pellet was resuspended gently in 5 ml of 10 mM Tris pH 7.4, 20 mM NaCl, 0.2 mg/ml PEG, and repelleted. The final gold conjugate pellets were resuspended in 100 µl of 10 mM Tris pH 7.4, 20 mM NaCl, 20% glycerol, 0.2 mg/ml PEG, 0.02% sodium azide, and frozen at –80°C in 20 µl aliquots. For use, aliquots were thawed and stored at 4°C for up to 4 weeks as a working stock of reagent.

Gold conjugates, prepared as described above, were used to test the protein binding capabilities of experimental and control lipid monolayers. To quantitate total gold particle numbers, 2 µl samples of 1:10 diluted gold solutions were dried onto carbon-formvar grids (SPI #3430C; copper, 300 mesh) and total gold particle amounts were counted in 200 fields at 250 000× (1.06 µm²) to calculate gold particles per 10 nl. For assay of gold conjugate binding to lipid monolayers, 1–5 µl of gold conjugates were mixed with the appropriate 2× subphase buffer (see above), overlaid with 1× lipid mix, incubated overnight at 30°C under the conditions described above, lifted onto lacy EM grids, washed once for 30 s on a 50 µl drop of distilled water, incubated for 30 s on a 50 µl drop of 1% glucose, wicked and air dried. To assess binding to monolayers, gold particles on hole-free, carbon support-free monolayer areas in at least 20 fields at 91 000× (4500 nm²) were counted to calculate the average number of particles bound per 1000 nm².

Image analysis

Initially, the quality of 2D protein arrays was evaluated by optical diffractions of EM negatives using the inverted laser diffractometer (Salmon and DeRosier, 1981) at the University of Oregon, the triply folded diffractometer at the EMBL-Heidelberg or the single-folded model at OHSU (Erickson *et al.*, 1978). Four negatives with acceptable diffraction patterns (negatives 255, 260, 717 and 719) then were digitized using either the flat-bed scanner at Baylor (negatives 717 and 719), which generated mrc format images, or with an Optronics cold CCD mounted onto a dissecting microscope, which generated tiff format images that were converted to pgm files with the XV image conversion program, and subsequently to mrc image files, using the PGM_FIMG program from Baylor. The resolutions of the scanned images were 2.67 Å/pixel for negative 717; 4.13 Å/pixel for 719; and 5.25 for negatives 255 and 260.

Following the generation of digitized mrc-format images, square regions for image analysis were obtained using BOXMRC, and Fourier transformed using FFTTRANS (Henderson *et al.*, 1990). Transforms were represented as diffraction patterns using SPECTRA (Schmid *et al.*, 1993), indexed by hand as either orthorhombic, or with $\gamma = 60^\circ$, after which the lattice was refined and unbent using LATREF, MMBOX and UNBEND (Baldwin *et al.*, 1988; Henderson *et al.*, 1990; Schmid *et al.*, 1993). The resulting unbent amplitudes and phases (APH) files were edited to remove all low signal-to-noise reflections of $IQ > 5$, and then corrected for phase inversions deriving from the CTF, using Thon rings as input for CTF_DETERMINE, and flipping phases by hand (Unwin and Henderson, 1975; Baldwin *et al.*, 1988; Henderson *et al.*, 1990). Merging of images and determination of phase residuals were performed using unbent, CTF-corrected APH files and the program ORIGINILTB. For all merges, the 717 film was used as the reference film, and for C222, p3 and p6 merging, a recursive search was conducted to obtain the 717 phase origin which gave the lowest phase residuals on merging. Generation and display of processed images from unbent, CTF-corrected APH files used the programs CREATE_TNF, FFTTRANS, ICE_SKEW and BOXMRC, in that order, while the contour map of the processed 717 image was created on an Apple Performa 636CD using Adobe Photoshop 2.0.1.

Programs for image analysis derive from the MRC suite of image processing programs (Unwin and Henderson, 1975; Baldwin *et al.*, 1988; Henderson *et al.*, 1990), modified for UNIX use and obtained from the EMBL-Heidelberg Structural Biology Programme and the Keck Center for Computational Biology at the Baylor College of Medicine. Processing was performed on an SGI Indigo2 running under UNIX, although we obtained comparable results when images were processed using SPIDER (Frank *et al.*, 1988) on VMS at the University of Oregon.

Acknowledgements

We have received an enormous amount of essential advice and assistance from a number of individuals throughout the undertaking of this work.

We are thankful to Jack Fellman for suggestions during the conception of this work, and to Richard Brennan for advice concerning data analysis. Paula Stenberg, Charles Meschul, Brent Gowan and Ed Gogol advised and assisted in a variety of EM techniques. Darrick Carter, Jackson Shea, Ken Fish and Steve Hardt helped with programing and computation, while Jenny Stegeman-Olson, Lori Farrell, Mike Yamauchi, Mark Hansen and Brent Berwin contributed to the cloning, expression, purification and analysis of His-tagged proteins. This research would not have been possible without crucial starting help to E.B. from the American Foundation for Aids Research (AmFAR grant 02301-17-RG). Our work was supported by grants 2R01 CA47088-07A3 from the National Cancer Institute and 5R01 GM52914-02 from the National Institute of General Medicine to E.B., Grant 9319099-MCB from the National Science Foundation to D.T., and grants NIH RR02250 and NSF BIR9500098 to M.F.S.

References

- Avila-Sakar, A.J. and Chiu, W. (1996) Visualization of β sheets and side chain clusters in 2-dimensional periodic arrays of streptavidin on phospholipid monolayers by electron crystallography. *Biophys. J.*, **70**, 57–68.
- Baldwin, J., Henderson, R., Beckman, E. and Zemlin, F. (1988) Images of purple membrane at 2.8 Å resolution obtained by cryo-electron microscopy. *J. Mol. Biol.*, **202**, 585–591.
- Campbell, S. and Vogt, V. (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.
- Celia, H., Hoermann, L., Schultz, P., Lebeau, L., Mallouh, V., Wigley, D., Wang, J., Mioskowski, C. and Oudet, P. (1994) Three-dimensional model of *Escherichia coli* gyrase B subunit crystallized in two-dimensions on novobiocin-linked phospholipid films. *J. Mol. Biol.*, **236**, 618–628.
- Chazal, N., Carriere, C., Gay, B. and Boulanger, P. (1994) Phenotypic characterization of insertion mutants of the human immunodeficiency virus type 1 Gag precursor expressed in recombinant baculovirus-infected cells. *J. Virol.*, **68**, 111–122.
- Chesebro, B., Britt, W., Evans, L., Wehrly, K., Nishio, J. and Cloyd, M. (1983) Characterization of monoclonal antibodies reactive with murine leukemia viruses: use in analysis of strains of Friend MCF and Friend ecotropic murine leukemia viruses. *Virology*, **127**, 134–148.
- Choi, H.-K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. and Wengler, G. (1991) Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature*, **354**, 37–43.
- Clish, C., Peyton, D. and Barklis, E. (1996) Spectroscopic study of an HIV-1 capsid protein major homology region peptide analog. *FEBS Lett.*, **378**, 43–47.
- Craven, R., Leure-duPree, A., Erdie, C., Wilson, C. and Wills, J. (1995) Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. *J. Virol.*, **69**, 4213–4227.
- Crawford, S. and Goff, S. (1984) Mutations in Gag proteins p12 and p15 of Moloney murine leukemia virus block early stages of infection. *J. Virol.*, **49**, 909–917.
- Darst, S., Ahlers, M., Meller, P., Kubalek, E., Blankenburg, R., Ribi, H., Ringsdorf, H. and Kornberg, R. (1991a) Two-dimensional crystals of streptavidin on biotinylated lipid layers and their interactions with biotinylated macromolecules. *Biophys. J.*, **59**, 387–396.
- Darst, S., Edwards, A., Kubalek, E. and Kornberg, R. (1991b) Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. *Cell*, **66**, 121–128.
- Dietrich, C., Schmitt, L. and Tampe, R. (1995) Molecular organization of histidine-tagged biomolecules at self-assembled lipid interfaces using a novel class of chelator lipids. *Proc. Natl Acad. Sci. USA*, **92**, 9014–9018.
- Dietrich, C., Boscheinen, O., Scharf, K., Schmitt, L. and Tampe, R. (1996) Functional immobilization of a DNA-binding protein at a membrane interface via histidine tag and synthetic chelator lipids. *Biochemistry*, **35**, 1100–1105.
- Ehrlich, L., Agresta, B. and Carter, C. (1992) Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro*. *J. Virol.*, **66**, 4874–4883.
- Erickson, H., Voter, W. and Leonard, K. (1978) Image reconstruction in electron microscopy: enhancement of periodic structure by optical filtering. *Methods Enzymol.*, **49**, 39–63.

- Faecke, M., Janetzko, A., Shoeman, R. and Krausslich, H. (1993) A large deletion in the matrix domain of the human immunodeficiency virus *gag* gene redirects virus particle assembly from the plasma membrane to the endoplasmic reticulum. *J. Virol.*, **67**, 4972–4980.
- Frank, J., Radermacher, M., Wagenknecht, T. and Verschoor, A. (1988) Studying ribosome structure by electron microscopy and computer-image processing. *Methods Enzymol.*, **164**, 3–35.
- Frey, W., Schief, W., Pack, D., Chen, C., Chilkot, A., Stayton, P., Vogel, V. and Arnold, F. (1996) Two-dimensional protein crystallization via metal-ion coordination by naturally occurring surface histidines. *Proc. Natl Acad. Sci. USA*, **93**, 4937–4941.
- Fuller, S.D., Capaldi, R. and Henderson, R. (1979) Structure of cytochrome *c* oxidase in deoxycholate-derived two-dimensional crystals. *J. Mol. Biol.*, **134**, 305–327.
- Fuller, S.D., Berriman, J., Butcher, S. and Gowen, B. (1995) Low pH induces swiveling of the glycoprotein heterodimers in the Semliki Forest virus spike complex. *Cell*, **81**, 715–725.
- Fuller, S.D., Butcher, S.J., Cheng, R.H. and Baker, T.S. (1996) Three-dimensional reconstruction of icosahedral particles: the uncommon line. *J. Struct. Biol.*, **116**, 48–55.
- Geoghegan, W. and Ackerman, A. (1977) Adsorption of horseradish peroxidase, ovomucoid and anti-immunoglobulin to colloidal gold for the indirect detection of concanavalin A, wheat germ agglutinin and goat anti-human immunoglobulin G on cell surfaces at the electron microscopic level: a new method, theory and application. *J. Histochem. Cytochem.*, **25**, 1187–1200.
- Gitti, R., Lee, B., Walker, J., Summers, M., Yoo, S. and Sundquist, W.I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science*, **273**, 231–235.
- Hansen, M. and Barklis, E. (1995) Structural interactions between retroviral Gag proteins examined by cysteine crosslinking. *J. Virol.*, **69**, 1150–1159.
- Hansen, M., Jelinek, L., Whiting, S. and Barklis, E. (1990) Transport and assembly of Gag proteins into Moloney murine leukemia virus. *J. Virol.*, **64**, 5306–5316.
- Hansen, M., Jelinek, L., Jones, R., Stegeman-Olsen, J. and Barklis, E. (1993) Assembly and composition of intracellular particles formed by Moloney murine leukemia virus. *J. Virol.*, **67**, 5163–5174.
- Henderson, R., Baldwin, J., Ceska, T., Zemlin, F., Beckmann, E. and Downing, K. (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.*, **213**, 899–929.
- Hill, C., Worthylake, D., Bancroft, D., Christensen, A. and Sundquist, W. (1996) Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *Proc. Natl Acad. Sci. USA*, **93**, 3099–3104.
- Hochuli, E., Dobeli, H. and Schacher, A. (1987) New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.*, **411**, 177–184.
- Horisberger, M. and Clerc, M.-F. (1985) Labelling of colloidal gold with protein A. *Histochemistry*, **82**, 219–223.
- Jones, T., Blaug, G., Hansen, M. and Barklis, E. (1990) Assembly of Gag- β -galactosidase proteins into retrovirus particles. *J. Virol.*, **64**, 2265–2279.
- 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**, 1830–1839.
- Kubalek, E., Le Grice, S. and Brown, P. (1995) Two-dimensional crystallization of histidine-tagged HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid. *J. Struct. Biol.*, **113**, 117–123.
- Kuhlbrandt, W., Wang, D.N. and Fujiyoshi, Y. (1994) Atomic model of plant light-harvesting complex by electron crystallography. *Nature*, **367**, 614–621.
- Leis, J. *et al.* (1988) Standardized and simplified nomenclature for proteins common to all retroviruses. *J. Virol.*, **62**, 1808–1809.
- Mammano, F., Ohagen, A., Hoglund, S. and Gottlinger, H. (1994) Role of the major homology region of the human immunodeficiency virus type 1 in virion morphogenesis. *J. Virol.*, **68**, 4927–4936.
- Maniatis, T., Fritsch, E. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Matthews, S. *et al.* (1994) Structural similarity between the p17 matrix protein of HIV-1 and interferon-gamma. *Nature*, **370**, 666–668.
- McDermott, J., Farrell, L., Ross, R. and Barklis, E. (1996) Structural analysis of human immunodeficiency virus type 1 Gag protein interactions using cysteine-specific reagents. *J. Virol.*, **70**, 5106–5114.
- Momany, C. *et al.* (1996) Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.*, **3**, 763–770.
- Morrelet, N., Julian, N., De Rocquigny, H., Maigret, B., Darlix, J.-L. and Roques, B. (1992) Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by ^1H NMR. *EMBO J.*, **11**, 3059–3065.
- Nermut, M., Hockley, D., Jowett, J., Jones, I., 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.
- Ng, K., Pack, D., Sasaki, D. and Arnold, F. (1995) Engineering protein-lipid interactions: targeting of histidine-tagged proteins to metal-chelating lipid monolayers. *Langmuir*, **11**, 4048–4055.
- Olofsson, A., Mallouh, V. and Brisson, A. (1994) Two-dimensional structure of membrane-bound annexin V at 8 Å resolution. *J. Struct. Biol.*, **113**, 199–205.
- Rein, A., McClure, M., Rice, N., Luftig, R. and Schultz, A. (1986) Myristylation site in Pr65Gag is essential for virus particle formation by Moloney murine leukemia virus. *Proc. Natl Acad. Sci. USA*, **83**, 7246–7250.
- Salmon, E. and DeRosier, D. (1981) A surveying optical diffractometer. *J. Microsc.*, **123**, 239–247.
- Schmitt, L., Dietrich, C. and Tampe, R. (1994) Synthesis and characterization of chelator-lipids for reversible immobilization of engineered proteins at self-assembled lipid interfaces. *J. Am. Chem. Soc.*, **116**, 8485–8491.
- Schmid, M., Dargahi, R. and Tam, M. (1993) SPECTRA: a system for processing electron images of crystals. *Ultramicroscopy*, **48**, 251–264.
- Shaw, A., Rothnagel, R., Chen, D., Ramig, R., Chiu, W. and Prasad, B. (1993) Three-dimensional visualization of the rotavirus hemagglutinin structure. *Cell*, **74**, 693–701.
- Shinnick, T., Lerner, R. and Sutcliffe, J. (1981) Nucleotide sequence of Moloney murine leukemia virus. *Nature*, **293**, 543–548.
- Spearman, P., Wang, J., Van der Heyden, N. and Ratner, L. (1994) Identification of human immunodeficiency virus type 1 Gag protein domains essential to membrane binding and particle assembly. *J. Virol.*, **68**, 3232–3242.
- Srinivasakumar, N., Hammariskjold, M. and 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.
- Stewart, L., Schatz, G. and Vogt, V. (1990) Properties of avian retrovirus particles defective in viral protease. *J. Virol.*, **64**, 5076–5092.
- Thompson, D., Svendsen, C., DiMeglio, C. and Anderson, V. (1994) Synthesis of chiral diether and tetraether phospholipids. Regiospecific ring opening of epoxides derived from asymmetric epoxidation. *J. Org. Chem.*, **59**, 2945–2955.
- Unwin, P.N.T. and Henderson, R. (1975) Molecular structure determination by electron microscopy of unstained crystalline specimens. *J. Mol. Biol.*, **94**, 425–440.
- Uzgiris, E. and Kornberg, R. (1983) Two-dimensional crystallization technique for imaging macromolecules with application to antigen-antibody-complement complexes. *Nature*, **301**, 125–129.
- Wang, C., Zhang, Y., McDermott, J. and Barklis, E. (1993) Conditional infectivity of a human immunodeficiency virus matrix domain deletion mutant. *J. Virol.*, **67**, 5550–5561.
- Wang, C., Stegeman-Olsen, J., Zhang, Y. and Barklis, E. (1994) Assembly of HIV Gag- β -galactosidase fusion proteins into virus particles. *Virology*, **200**, 524–534.
- Weiss, R., Teich, N., Varmus, H. and Coffin, J. (1984) *RNA Tumor Viruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Wills, J., Craven, R., Weldon, R., Nelle, T. and Erdie, C. (1994) An assembly domain of the Rous sarcoma virus Gag protein required late in budding. *J. Virol.*, **68**, 6605–6618.
- Yu, X., Yuan, X., Matsuda, Z., Lee, T. and Essex, M. (1992) The matrix protein of HIV-1 is required for incorporation of viral envelope protein into mature virions. *J. Virol.*, **66**, 5667–5670.
- Zhang, Y. and Barklis, E. (1995) Nucleocapsid protein effects on the specificity of retrovirus RNA encapsidation. *J. Virol.*, **69**, 5716–5722.
- Zhao, X., Rui, Y., Thompson, D. and Barklis, E. (1994) Interfacial crystallization of HIV capsid protein arrays. *American Chemical Society 46th Southeast Regional Meeting*, Abstract 73.

Received on September 12, 1996; revised on November 5, 1996