Structural Analysis of Retroviral Capsid Proteins

by

Jason McDermott

A DISSERTATION

Presented to the Department of Molecular

Microbiology and Immunology and

to the Oregon Health Sciences University

School of Medicine

in partial fulfillment of

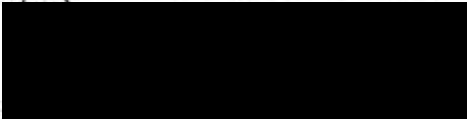*the requirements for the degree of*

Doctor of Philosophy

May, 2000

School of Medicine

Oregon Health Sciences University

〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜

CERTIFICATE OF APPROVAL

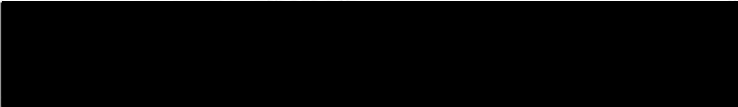〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜〜

This is to certify that the Ph.D. thesis of

Jason McDermott

has been approved

Committee Chair: Dr. Buddy Ullman

Member: Dr. Richard Brennan

Member: Dr. Jay Nelson

Member: Dr. Fred Heffron

5/18/00

Member: Dr. Eric Barklis

TABLE OF CONTENTS

INDEX OF TABLES AND FIGURES

Chapter 4

Chapter 5

ACKNOWLEDGEMENTS

I would first like to thank the people who have made this all possible. Without their support, nurturing, guidance, encouragement and love, this would not have been possible, my parents, Dennis and Brygida McDermott. I would also like to thank my grandfather, Gideon Kramer, for his undying curiosity and voracious quest for the 'big picture'. In addition, the acknowledgements section for this Thesis, which is a culmination of 21 years of schooling, would be lacking without mention of a number of very special teachers, most still slogging the trenches on the front lines of education, some still in Yelm, WA; my 8th grade science teacher, Rich Sauers, my high school English teachers, Lester Krupp and Holly Breidenbach, my high school politics teacher, Ed Berg, my high school math and physics and calculus teacher, and orienteering coach, Mike Tharp, my high school biology and chemistry teacher (I've never done more coursework than in A.P. Biology), Virginia Reed, my Reed College plant ecology professor, Bert Brehm, and many others along the way. Each of these educators has left an indelible mark on my current way of thinking and world (as well as scientific) view.

I would also like to thank all of the co-authors, contributors and lab members, both past and present who have made this work possible, Zac Love for the mental sparring, and coffee, for helping get it all done. Thanks also to my Thesis committee, Dr. Buddy Ullman, Dr. Richard Brennan, Dr. Fred (Z-serve) Heffron, and Dr. Jay Nelson. Special thanks to my mentor, Dr. Eric Barklis. In addition to invaluable lessons in pigmania, croquet, and the finer points of being a scumbag, he has helped shape my scientific education, sharpened my intellect, and been my friend through my serving nearly every position in his lab over the course of nine years. Thank you for this opportunity.

Lastly, I would like to dedicate this work to an incredible woman who is not only a co-author on several of these manuscripts, but is also, as I discovered working by her side, my life partner. To many future joint ventures and joint happiness. May we continue to lead each other down new paths and discover new sunsets together.

To my wife, Sonya Karanjia.

ABBREVIATIONS

2D ---------------------------------- two dimensional

3D ---------------------------------- three dimensional

AIDS -------------------------------- acquired immunodeficiency syndrome

BMH -------------------------------- bis-maleimido hexane

CA ----------------------------------- capsid

CTD --------------------------------- (capsid) C-terminal domain

CTF --------------------------------- contrast transfer function

DHGN ------------------------------ 1,2-di-O-hexadecyl-*sn*-glycero-3-(1'-(2''-R-hydroxy-3''-N-(5-amino-1-carboxypentyl)-iminodiacetic acid) propyl ether)

EIAV -------------------------------- equine infectious anemia virus

EM ----------------------------------- electron microscopy

Env --------------------------------- envelope

FFT --------------------------------- fast-Fourier transform

his-MoCA -------------------------- histidine-tagged M-MuLV CA

his-HIVCA ------------------------- histidine-tagged HIV-1 CA

HIV-1 ------------------------------- human immunodeficiency virus type 1

IN ----------------------------------- integrase

MA ---------------------------------- matrix

MHR -------------------------------- major homology region

M-MuLV ---------------------------- Moloney murine leukemia virus

NC ----------------------------------- nucleocapsid

NTD --------------------------------- (capsid) N-terminal domain

PC ----------------------------------- phosphatidylcholine

PR ----------------------------------- protease

RSV --------------------------------- Rous sarcoma virus

RT ----------------------------------- reverse transcriptase

STATEMENT OF CONTRIBUTION

A statement of my contribution towards each chapter in this Thesis, other contributors are listed as co-authors by chapter, or in the acknowledgements section for that chapter.

Chapter 1   Sole authorship.

Chapter 2   All writing and figure preparation, in addition I was responsible for creation of several expression constructs and performace of duplicate or pertinent experiments not included as figures.

Chapter 3   All writing and figure preparation, in addition I was responsible for creation of many expression constructs, performance of duplicate or pertinent experiments not included as figures, and performance of infectivity experiments.

Chapter 4   In addition to manuscript revision and figure preparation, I was responsible for most EM experiments, including gold labelling experiments and preparation of his-MoCA crystals, and all computer processing steps.

Chapter 5   In addition to manuscript revision and figure preparation, I was responsible for all computer processing steps.

Chapter 6   All writing and figure preparation, in addition I was responsible for gathering duplicate EM data as well as all other experiments outlined in the manuscript.

Chapter 7   Sole authorship.

Chapter 8   Sole authorship, save gathering of raw EM data.

ABSTRACT

Although atomic resolution structures of a number of retroviral capsid proteins have become available recently, relatively little is known about the interactions between capsid proteins that drive virus particle assembly. We have employed genetic, biochemical, and biophysical approaches to investigate the structural interactions of both the human immunodeficiency virus type 1 (HIV-1) and Moloney murine leukemia virus (M-MuLV) capsid proteins (CA).

We used site-directed mutagenesis and biochemical techniques to investigate the ability of cysteine residues introduced into various regions of the CA domain to form intermolecular crosslinks in the presence of cysteine-specific crosslinking reagents. Specific cysteine mutants in both viruses were able to be crosslinked allowing prediction of regions that might interact in virus particles. In M-MuLV we found that the ability of certain mutants to form crosslinks was dependent upon their context in immature or mature virus particles, suggesting that the CA undergoes a significant conformational change during particle maturation. In HIV-1 removal of a naturally occurring cysteine residue from the C-terminal domain of CA abolished particle assembly and release, suggesting that this region is critical for virus particle assembly. In addition, cysteine creations in the major homology region (MHR) of the HIV-1 CA abolished virus infectivity but did not visibly affect assembly, indicating that they were blocked in a post-entry step of infectivity.

To analyze CA-CA interactions further, we established a method for *in vitro* assembly of histidine-tagged (his-tagged) CA proteins on nickel-chelating lipid monolayers. We used this system to obtain projection structures of his-tagged M-MuLV CA (his-MoCA) proteins to 9.5 Å resolution and his-tagged HIV-1 CA (his-HIVCA) proteins to 23 Å resolution. Both structures appeared to have a trigonal or hexagonal organization characterized by hexameric cage-holes. We extended these studies by obtaining three-dimensional (3D) structures of his-MoCA and his-HIVCA proteins to approximately 25Å resolution from images of tilted, negatively-stained 2D crystals. In contrast to expectations from the projection structures, 3D structures displayed a lack of symmetry, appearing to form strands of dimers with the N- and C- terminal domains (NTD and CTD) dimerizing with different partners, and did not exhibit trigonal or hexagonal organization. In both structures the membrane-proximal densities fit expectations for the size of NTD dimers. However, membrane-distal densities appeared to be forming connections with several partners for each NTD, each of which was less dense than would be expected for a CTD. These observations suggest that CTDs adopt alternate conformations which appear as overlapping densities in our structure. Both structures seem to be compatible with helical/spiral models of retovirus assembly but do not agree easily with icosahedral models.

Chapter 1

INTRODUCTION

Summary

Viruses in the family *Retroviridae*, the retroviruses, have been found to play an increasingly important role in human disease. Human Immunodeficiency Virus Types 1 (HIV-1) and 2 (HIV-2), Human T-Cell Lymphoma Virus (HTLV) as well as human endogenous retrovirus (HERV) sequences are the causes of human diseases (31, 102). HIV-1 is the etiological agent of acquired immune deficiency syndrome (AIDS), which has become the fourth leading cause of mortality in the world today (1). Therapeutic drugs against HIV-1 to date have been targeted primarily at the stages of the viral life cycle involved in the reverse transcription of the virus in host cells and most recently at the maturation of infectious virus particles (47). Relatively little information is known about the processes of retroviral assembly that occur in infected cells to produce virus progeny capable of propagation.

Due to the nature of their life cycle and their permanence in the host cell's genome, retroviruses also have become a valuable tool in the development of vectors for use in gene therapy protocols for the treatment of disease. Retroviral vectors can be designed to target specific cell types and a number of approaches have been employed that allow regulation of gene expression. In addition to their clinical importance, retroviruses have been employed for numerous laboratory uses such as transduction of exogenous genes, insertional mutagenesis in eukaryotic models, cell fate determination during organism development and cDNA library construction (108).

Retroviruses are enveloped viruses containing two copies of single-stranded RNA genomes which encode three primary genes, *gag*, *pol* and *env*. The surface of the virus is covered by products of the *env* gene which are involved in binding and entry into the target cell. The inner core structure of the virus is formed by *gag* gene and ranges in shape from spherical to conical (31). The *pol* gene encodes the reverse transcriptase (RT) and an integrase (IN) protein which allow the retrovirus a unique lifestyle by transcribing the RNA genome into DNA followed by incorporation into the genome of the host cell. In most retroviruses the *pol* gene also encodes a protease (PR) responsible for viral maturation. This thesis examines mechanisms of particle assembly for HIV-1, a member of the lentivirus subfamily and Moloney Murine Leukemia Virus (M-MuLV), a type C retrovirus. Specific attention will be given to these two viruses throughout this introduction.

Principles of Virus Structure

All viruses face a similar problem: how to transfer a nucleic acid viral genome from host cell to host cell in the most efficient manner. This consideration imposes a restriction on the size of the genome which is minimized by the use of one or a few kinds of protein structural subunits. Viruses generally accomplish this by enclosing their genomes in a structure with helical symmetry or icosahedral symmetry. In this way, multiple copies of a single structural subunit can participate in the formation of a complex structure with similar repeating protein-protein interactions (71). These structures are capable of

carrying the viral genome as well as infection of the target cell and other functions such as host immune system evasion in the case of animal viruses.

Equivalence

The similarity of protein-protein interactions lies at the heart of formation of both helical and icosahedral viruses from single structural subunits. This quality is referred to as *equivalence*, which simply means that each structural subunit forms identical protein-protein interactions with the neighboring structural subunits in the virus structure. Except in the cases of the most basic manifestations of helical and icosahedral virus structure, maintenance of strict equivalence is not possible for all structural subunits. Instead some or all of the subunits must occupy different environments depending on their locations in the overall virus structure. If the subunits are able to accommodate the different environments with minor adjustments such as slightly altering protein-protein interaction angles, they are described as being *quasi-equivalent* (71, 86, 87). In certain cases, such as in the *polymaviridae*, structural subunits must accommodate large differences in their environments, even taking on different protein conformations (141, 142, 144). This is referred to as *non-equivalence*.

Helical Virus Structure

Helical structures are formed by permutation of the viral structural subunits in a helix to form a tube-like structure (Figure 1d). Helical symmetry can be described by the number of units per turn, u, and the axial rise per unit, p,

with the pitch given by P = u x p. These viruses generally take the form of long tubes whose length is often governed by association of the structural subunits with nucleic acid strand(s) of given length (i.e. the viral genome). Helical structures can be described by a surface lattice, the network of protein-protein interactions between structural subunits on the curved surface of the helix (71). A flat, planar surface lattice can be 'rolled up' to form the helical structure, the number of units that encircle the helix once being u, and the number of units offset for each turn being p, which is not necessarily an integer.

Subunit interactions must take place in the strands that form the helix (side-to-side) as well as between the strands (top-to-bottom). A spherical structure could, in theory, be generated by a permutation of the structural subunits in a spiral instead of a helix, as proposed for retrovirus structures (25). This would involve a continuous variation of the pitch of the helix from one pole of the spheroid to the other and would necessitate formation of quasi-equivalent interactions between strands since the relative locations of the neighboring proteins also would vary continuously.

Icosahedral Virus Structure

Most spherical viruses have been shown to have icosahedral symmetry, characterized by a soccer-ball-like shape having 20 faces which are 3-fold symmetric, 12 vertices which are 5-fold symmetric, and 30 edges which are 2-fold symmetric (see Figure 1a). The icosahedral nature of the virus is defined by these symmetry constraints, not by the physical shape of the particle. Thus an

icosahedral virus may have a spherical outer appearance, but the structural subunits still precisely obey the three symmetry axes. In the simplest case, an icosahedral virus is comprised of 60 structural subunits, each subunit satisfying all three symmetry constraints (see Figure 1b). In this case, each face is composed of three protein subunits, and every protein subunit forms exactly equivalent interactions with its neighbors.

Larger viruses must accommodate more than three subunits per icosahedral face. Some viruses accomplish this with multiple proteins forming one of 60 icosahedral subunits while others have identical protein subunits in some multiple of 60. In these cases the subunits must occupy quasi-equivalent environments depending on their location in the face, or on a vertex or edge. For icosahedral viruses with more than 180 subunits only certain multiples of 60 subunits with very similar interactions are allowed. The triangulation number, T $= h^2 + hk + k^2$ (h and k being integers), gives the allowed multiples of 60 subunits. Common T numbers for icosahedral viruses include; T=3 (many plant and insect RNA viruses), T=4 (togaviruses), and T=7 (heads of bacteriophages P22 and λ) (71).

Figure 1. Models of retrovirus assembly.

(A) **Icosahedral model.** An icosahedron showing 10 of 20 equilateral triangular faces and 9 of 12 vertices. Each face has a three-fold axis of symmetry (**triangle**), each edge has two-fold axis of symmetry (**oval**), and each vertex has a five-fold axis of symmetry (**pentagon**). Each of the symmetry axes is a point of potential quasi- or non- equivalence for protein subunits. Viruses with a triangulation number (T) of 1 have one subunit occupying each face in a universally equivalent manner, viruses of type $T = N$ have $N \times 60$ protein subunits and N units per face.

(B) **T=1 icosahedral model.** An icosahedron with a triangulation number, T=1, schematically showing the location and orientation of protein subunits.

(C) **Hexagonal network.** A hexagonal lattice can be converted into an icosahedron by removing the wedge and folding the remaining lattice down the lines indicated, so that the new edges meet. Adapted from Nermut, *et al.* (41).

(D) **Spiral/helical model of retrovirus assembly.** The spiral/helical model of retrovirus assembly, adapted from Campbell and Vogt (8).

Like helical structures, icosahedral structures can be described with a surface lattice, and flat subunit networks can be folded to form icosahedral structures. Formation of one vertex (and associated five sides) of an icosahedron from a flat lattice can be envisioned as elimination of a 60° wedge from the lattice. The remaining five wedges (each 60°) are then folded to introduce the fivefold symmetric vertex as well as the three dimensional nature of the surface, producing one quarter of an icosahedron (5 of 20 faces). In viruses, this is often accomplished by starting with a hexagonal surface lattice (see Figure 1c). When a 60° wedge is removed from a hexagonal lattice, the fivefold symmetric vertex is formed from a sixfold symmetric hexamer and thus requires minimal differences in structural interactions. It can be seen in Figure 1b, that removal of such a wedge from a hexagonal network also allows the portions of subunits on the edges to match up, giving an integral number of subunits.

Since both helical and icosahedral structures may be formed from surface lattices, it is sometimes the case that these structural forms are combined, such as in the heads of T4 bacteriophages (71). In this case icosahedral 'caps', and tubular sides (exhibiting helical symmetry) are both formed from the same surface lattice. In a similar manner, structural subunits of icosahedral viruses may aberrantly assemble helical structures. In both cases the subunit interactions are sufficiently similar between the icosahedral and helical structures to allow formation of both structural types.

Determination of Virus Structure

Electron microscopy (EM) techniques have been the tools most often used in the determination of general morphology of virus particles. It is often possible to determine whether the virus has helical or icosahedral symmetry, either by simple observation or by using computer image processing techniques. For instance, computational methods exist for the determination of both helical and icosahedral parameters generally using a number of images representing various views of particles. Image reconstruction of particles to moderate resolution from such data is possible. For higher resolution structural information, X-ray crystallography of whole virus particles has been employed. This technique requires large amounts of very homogeneous virus particles that can be crystallized to work effectively. The first X-ray structures of animal viruses were determined for picornaviruses (75, 139).

Results from structural studies need to be tested by genetic and biochemical experiments and reconciled with existing data of this type. For example, initial X-ray structures for the picornaviruses were supported by studies of neutralizing antibodies, some of which reacted to epitopes formed at the junction of two or more structural subunits (114). Other corroborating data has been gained from purely genetic approaches (71, 84) , as well as other biochemical approaches (70, 98).

Overview of the Retroviral Life Cycle

The life cycle of a prototypic retrovirus (see Figure 2) begins when a mature, infectious particle encounters a target cell with the appropriate surface receptor(s). The viral envelope glycoproteins bind to this receptor which causes a conformational change in the envelope to expose a fusion domain that allows fusion of the viral membrane with that of the host cell (82). HIV-1 has two receptors, the CD4 glycoprotein and members of the chemokine receptor family. Binding of the envelope protein to CD4 exposes regions of the envelope which bind to the chemokine receptor and expose the fusion domain (169). The receptor for the ecotropic M-MuLV envelope is a cationic amino acid transporter (82).

After the viral membrane has fused with the host cell membrane, the viral nucleocapsid structure is deposited in the host cell cytoplasm where the particle uncoats to release the reverse transcription complex which includes viral RNA and reverse transcriptase and integrase as well as accessory proteins (165). Retroviral uncoating is one of the most poorly understood steps in the life cycle but *gag* gene products clearly play a significant role in this process (156). Reverse transcription of the viral genomic RNA occurs via a well-defined but complex set of steps involving two distinct functions of the RT, a reverse transcripase and a nuclease (RNase H) which is specific for the RNA strand of RNA:DNA duplexes. The end product of this process is a double stranded proviral DNA form, which differs from the genomic RNA by the presence of a complete copy of the viral long terminal repeat (LTR) at either end of the genome (148, 156).
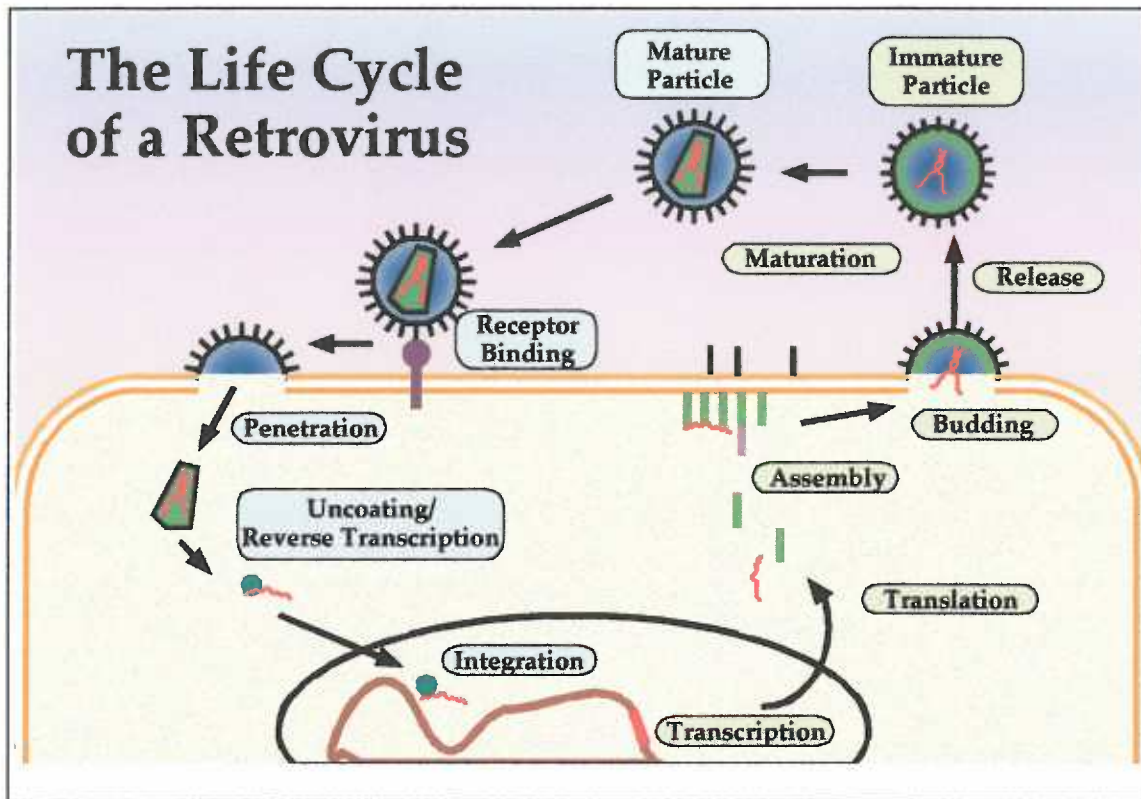
Figure 2. The life cycle of a retrovirus.

The life cycle of the retrovirus begins when a mature virus particle binds to its target receptor(s) on the surface of a host cell and is internalized. Once in the host cytoplasm the viral core uncoats, and reverse transcription of the single-stranded viral RNA to a double-stranded DNA provirus occurs, followed by the transport of the pre-integration complex into the nucleus and integration of the provirus into the cellular genome. Translation of the Gag and Gag-Pol fusion proteins occurs from unspliced viral genomic RNA and translation of *env* gene products as well as other accessory proteins occurs from singly- or multiply- spliced RNA transcripts. Gag and Gag-Pol proteins assemble on the plasma membrane (for lentiviriruses and type C retroviruses), envelope (Env) proteins, viral genomic RNA, as well as viral and/or cellular accessory proteins are incorporated into particles which bud through the membrane and are released from the cell. During or just after particle budding, the viral protease cleaves the Gag and Gag-Pol proteins to form a mature, infectious virus particle.

The complex of viral DNA, IN, and accessory proteins, known as the preintegration complex (PIC) is transported into the nucleus, either by diffusion during mitotic nuclear breakdown or, in the case of HIV and possibly other lentiviruses, active transport through the nuclear pore complex (49). Once in the
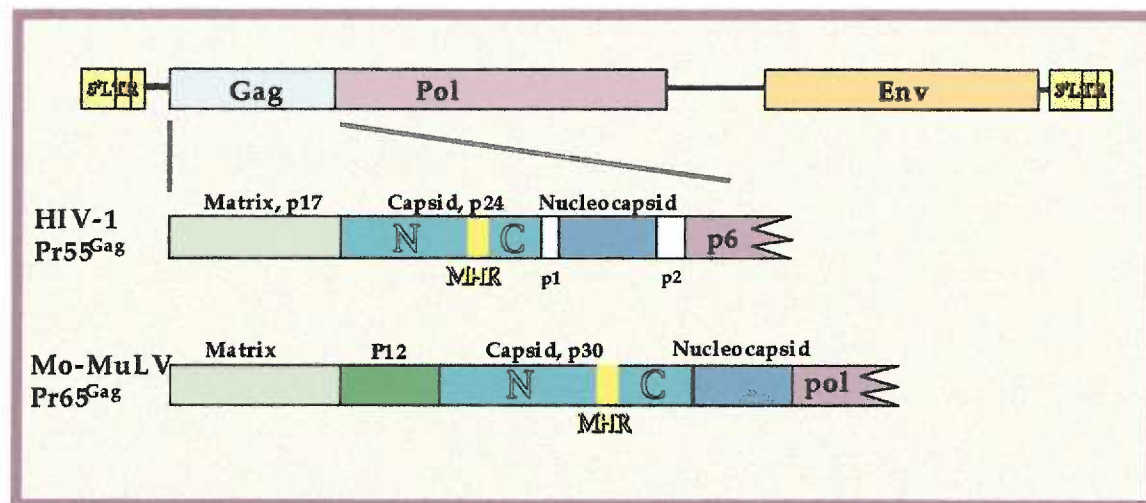
Figure 3. Genomic Organization of M-MuLV and HIV-1.

The genome of a protypic retrovirus (top) showing an integrated proviral form
which includes 5′ and 3′ LTRs, *gag*, *pol* and *env* genes. Organization of Gag proteins
from HIV-1 (middle) and M-MuLV (bottom), showing the Gag domains for each. In
addition the N- and C- terminal domains (NTD and CTD) are noted by N and C
(outlined) and the location of the major homology region (MHR) is noted in yellow).
The HIV-1 Gag protein (pr55; middle) contains the matrix, capsid, nucleocapsid and
p6 domains as well as proteolytic fragments, p1 and p2. The M-MuLV Gag protein
(pr65; bottom) contains the matrix, p12, capsid and nucleocapsid domains.

nucleus, the viral IN directs the looping of the viral DNA and complex formation

with the cellular genomic DNA. The genomic DNA is cleaved by IN via a direct

transesterification reaction and the 3′ ends of the viral DNA are joined to the

genomic DNA. Gap repair, possibly accomplished by RT, and ligation complete

the reaction, yielding a fully integrated provirus (19). Integration sites appear to

be sequence non-specific but show a preference for highly bent regions of DNA

and are influenced by the presence of DNA binding proteins (129).

From the integrated provirus (see Figure 3), RNA transcripts are produced

using the 5′ LTR as a transcriptional promoter. Unspliced, full length, viral RNAs

function as the viral genome as well as coding for the *gag* and *gag-pol* gene

products. A portion of these transcripts are spliced to form subgenomic mRNAs.

In the case of M-MuLV, the one alternatively spliced product encodes the *env* gene, but in HIV and other complex retroviruses, multiple splice pathways are utilized which generate transcripts for multiple accessory proteins as well as Env (130).

Gag proteins are translated as polyprotein precursors (PrGag) from free polyribosomes in the cytoplasm and transported to the plasma membrane by an unknown pathway (155). Approximately 5% of PrGag proteins are produced as Gag-Pol fusion proteins via a frameshift mechanism in the case of HIV or suppressed termination in the case of M-MuLV (72, 83, 157). The Gag proteins of HIV-1 and M-MuLV are 55 kDa and 65 kDa in size, respectively, and thus are referred to as Pr55^{Gag} and Pr65^{Gag} (see Figure 3). PrGag proteins of HIV and M-MuLV, as well as most other retroviral PrGag proteins, are modified by the addition of a myristate group to their N-termini which is essential for targeting to and association with the membrane (20, 73, 97, 135). Upon reaching the plasma membrane Gag proteins oligomerize and direct incorporation of the envelope proteins, RNA and accessory proteins into the assembly complex, or virus particle (8, 32, 51, 97). Concurrent with oligomerization, virus particles begin to bud through the membrane and are released as immature particles, characterized by an electron dense layer just under the viral membrane and an electron lucent center (59, 76, 90, 124). During or just following budding the viral protease is activated, by homodimerization, possibly driven by the oligomerization of the Gag portions of the Gag-Pol proteins (155, 157). The protease cleaves the PrGag and PrGag-Pol proteins into their component parts thus achieving virus maturation and formation of an infectious virus particle, characterized by

condensation of an electron dense core structure, roughly spherical in the case of M-MuLV and conical in shape for HIV-1 (59, 76, 90, 124).

Gag Protein Organization

As "particle-making machines" (42) , Gag proteins are responsible for incorporation of viral RNA, envelope proteins and other viral and/or cellular accessory proteins as well as the generation of the particle structure. The assembly function can be further divided into several steps, membrane targeting and binding (M), formation of structural interactions with other Gag proteins (I), and regulation of functions necessary late in the budding process (L) (36, 155). Retroviral Gag proteins have been shown in several systems to be necessary and sufficient for virus-like particle assembly and release (36, 38, 140, 147, 168). Gag proteins are synthesized as polyprotein precursors (see Figure 3) constituting a minimum of three protein domains, matrix (MA), capsid (CA), and nucleocapsid (NC). Although the presence and relative order of these domains in Gag proteins is invariant between retroviruses (155), some encode other domains in their *gag* genes. M-MuLV has a domain, p12, which lies between the MA and CA domains, and HIV-1 has a domain, p6, which lies at the C-terminus of Gag following the NC domain (see Figure 3). Following are short reviews of what is currently understood about the functions and structures of each of the M-MuLV and HIV-1 Gag domains.

Matrix

The matrix protein is positioned at the N-terminus of the PrGag and has been shown to contain the membrane targeting and association (M) function (97, 155), as well as playing other important roles in the virus life cycle. The MA domains, and thus the PrGag proteins, of both M-MuLV and HIV-1 are modified by the addition of a 14-carbon fatty acid, myristate group at their N-termini (20, 73, 97, 135). Myristate addition has been shown to be essential though not sufficient for the M functionality (20, 136), sequences in the amino-terminal portion of MA are also required (118, 177, 178). These sequences have been proposed to form a positively charged patch, which interacts with negatively charged phopholipids (33, 105, 149, 177). The model that has been proposed to explain these findings is one in which myristoylation provides a weak, and thus reversible, association with the plasma membrane and that the sequences in MA constitute a switch that is turned 'on' during assembly for tight membrane association, or turned 'off', by maturation of MA or other mechanisms, to allow MA to perform other functions (118, 123).

The matrix protein of M-MuLV has not been extensively investigated to date, but several studies have indicated that it contains the M domain for Pr65[Gag] (66, 69, 135, 149). Unmyristoylated Pr65[Gag] proteins are not included in assembling virus particles and are not correctly proteolytically processed (135, 143). Mutations and insertions in the amino-terminal region of MA impaired correct membrane targeting and association of Pr65[Gag] (69, 149), and did not effect on envelope incorporation. M-MuLV MA was shown to be phosphorylated

at tyrosine residues and inhibition of this phosphorylation also inhibited

integration of the provirus in infected cells (146). This and other genetic studies

(37) suggest that MA may play a role in late events of the M-MuLV life cycle

In contrast to M-MuLV MA, the MA of HIV-1 (p17) has been investigated

intensively in recent years and much of what is known about retroviral matrix

proteins comes from this work (33, 155). As outlined above, basic sequences in

the N-terminal portion of HIV-1 MA are essential for membrane targeting and

association (120, 123, 177). In fact, the M domain from HIV-1 MA can replace the

M domain in several other viruses (40, 125, 177). HIV-1 The HIV-1 MA protein

also plays other roles in the life-cycle of the virus. Mutations in HIV-1 MA have

been shown to impair incorporation of envelope proteins into virus particles (34,

45, 119, 162, 171). This effect has been proposed to be through a direct interaction

with the cytoplasmic tails of the viral TM proteins (53, 76, 172). The MA contains

a nuclear localization signal that maps to the same basic sequence important for

membrane binding (22) and has been shown to be a component of the

preintegration complex (23), which allows HIV-1 to infect non-dividing cells, and

may be involved in other processes in the late stage of infection (93). Recent

reports that HIV-1 MA interacts with cellular proteins support these findings (99,

127).

Structures of HIV-1 MA determined by NMR (105, 106) and X-ray

crystallography (74) as well as the related simian immunodeficiency virus (SIV)

MA (131), have been published (see Figure 4). These structures show MA

proteins forming a single, globular domain, composed mainly of alpha-helices,

with flat, basic faces on one side which is thought to be oriented toward the membrane. Trimerization of MA is further supported by a study by Morikawa, et al., which shows that MA preferentially forms trimers in solution and that this phenotype can be blocked by mutations in MA (113). This and other studies (26, 54, 112, 159) suggest a role for the MA protein in the assembly of PrGag, although this notion is not supported by studies showing that HIV-1 MA is dispensable for virus particle assembly (15, 134, 162). The C-terminal tail of MA is disordered in the structures which indicates that the region between MA and CA domains in PrGag may be flexible.

Capsid

The capsid protein forms the viral capsid, a shell which encloses the ribonucleoprotein complex in the mature virus (124, 155, 164). Although large deletions in the capsid domain of the Rous sarcoma virus (RSV) have been found to be compatible with virus-like particle assembly (163), multiple studies have shown that the CA proteins of both HIV-1 and M-MuLV play a significant role in virus particle assembly (52, 59, 155). In addition, *in vitro* assembly studies have demonstrated that CA proteins can assemble into rod-like and spherical structures (46, 67), suggesting that CA is involved in the formation of virion structure. The major homology region (MHR) in CA, is a stretch of 20 residues which are highly conserved among nearly all retroviruses, and forms one end of an assembly domain that stretches over the C-terminal portion of CA and into the spacer region (p2) between CA and NC (2, 15, 43). Mutations in the MHR have been shown to impair assembly (35, 104, 133, 152, 174) and mutations in the

portion of capsid C-terminal to the MHR are also detrimental to assembly (62, 133, 151, 159, 174). In contrast, many mutations in the N-terminal portion of CA are tolerated for assembly (133, 159), and deletion of large portions of the N-terminal portion are compatible with particle assembly and release (15, 159, 160). The organization of CA into functional domains is supported by structural studies which show CA from several retroviruses to be broken into two domains (N and C) with a flexible linker region between them (57, 61, 85, 92, 110). In addition to its role in assembly, CA is responsible for interaction with, and/or incorporation of, viral and/or cellular factors, and for events in the "late" stage of the replicative life cycle (52, 155).

The role of CA in the assembly of M-MuLV has been demonstrated by several groups. The capsid domain was found to be essential for incorporation of mutant Gag-β-galactosidase fusion proteins into wild-type (wt) virus particles (89) and insertions in the N-terminal portion of CA were shown to inhibit mutant Gag incorporation using the same system (69). Results from a two-hybrid type protein interaction screen indicated that the C-terminal regions of CA, coupled to NC, were sufficient to mediate protein dimerization, and that point mutations and deletions in the C-terminal portion of CA were able to abolish these interactions (5). Cysteine residues introduced into the MHR could be crosslinked using cysteine-specific crosslinking agents, indicating that MHR's are close together in immature virus particles (70). Other CA mutants also were shown to abrogate virus particle assembly and release (4, 79).

M-MuLV CA also is involved in the viral life cycle at points after assembly and release of virus particles. M-MuLV CA mutants which were not compromised in assembly, were found to be poorly- or non- infectious, indicating they are blocked in some post-assembly function (4). These mutants were found to contain normal amounts of RT and genomic RNA, but did not synthesize viral DNA, indicating that the replicative block is in viral entry or uncoating of the RT complex (4). Another important role of M-MuLV CA in the late stages of infection is played in the *Fv1* host restriction system. Members of the murine leukemia subfamily can be grouped based on their ability to replicate in the presence of a host restriction system encoded by the *Fv1* locus, which is either N- or B- type (88). A single amino acid in MuLV CA determines whether the virus has N-, B-, or NB- cell tropism (41). *Fv1* restriction appears to act at a stage in infection following reverse transcription but prior to integration (128). This is supported by the finding that CA is present in M-MuLV pre-integration complexes (16). The *Fv1* gene appears to be derived from the *gag* gene of an endogenous retrovirus (14), but the significance of this finding has not been established.
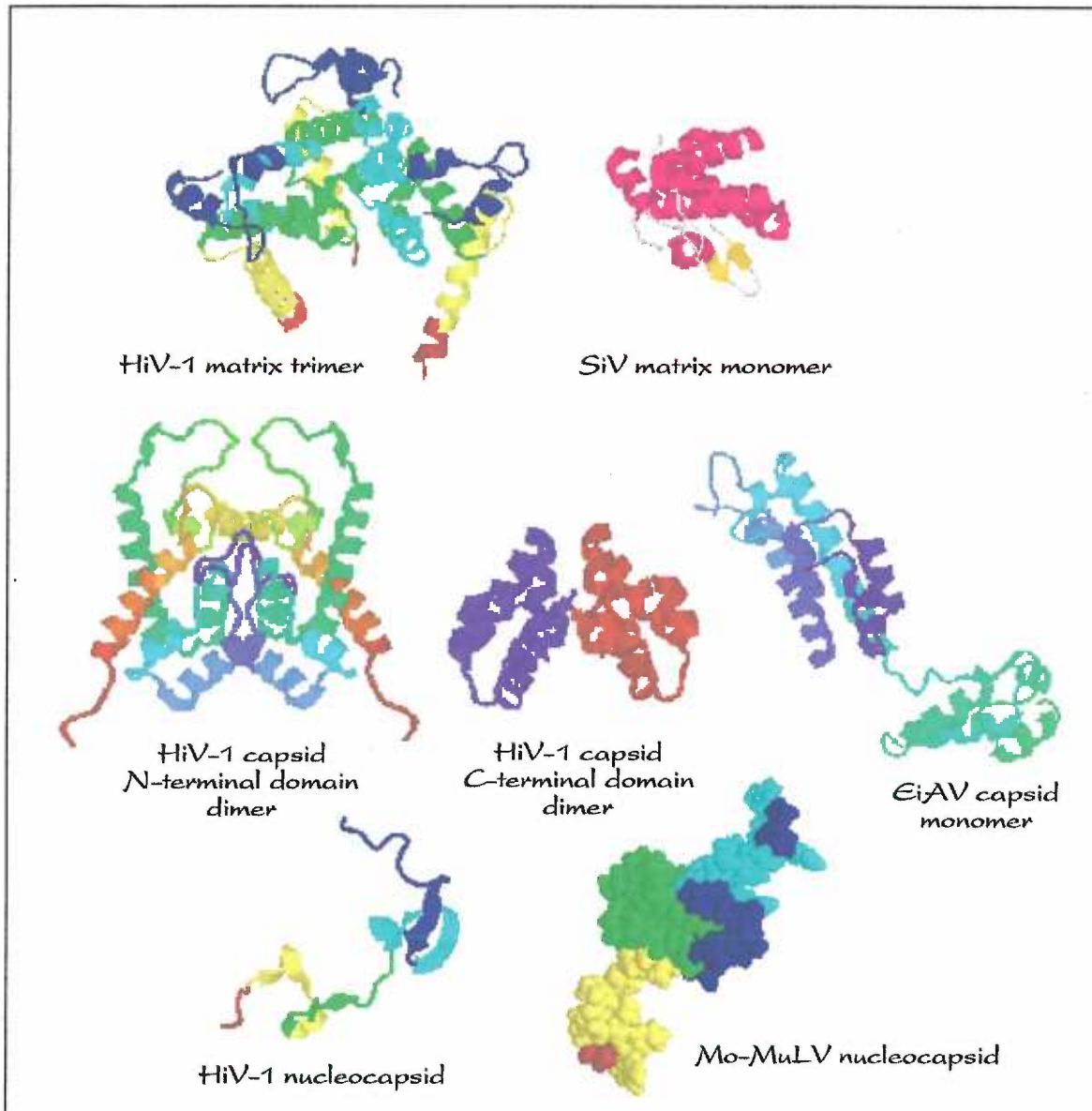
Figure 4. Structures of Gag-derived proteins.

Shown are X-ray and NMR derived structures of Gag domains from several retroviruses. All structures have been rendered with RasMol 2.5 as ribbon diagrams or a space filling model (M-MuLV NC structure) and are shown with N-termini up and C-termini down (approximately), colors are blue (N-termini) to red (C-termini) or are solidly colored (HIV-1 CTD dimer and SIV MA monomer). At top left is the crytal structure of HIV-1 MA protein trimer (PDB #1HIW; 74) followed by the SIV matrix monomer (PDB #1ED1; 131). Middle left, the crystal structure of the HIV-1 capsid N-terminal domain dimer (PDB #1AFV; 110), middle, the HIV-1 capsid C-terminal domain dimer (PDB #1AM3; 57) and, right, a monomer of the EIAV capsid protein (PDB #1EIA; 85). Bottom left, the NMR structure of the HIV-1 nucleocapsid domain (PDB #1MFS; 154) and, right, NMR-derived structure of M-MuLV nucleocapsid domain (PDB #1A6B; 39).

The capsid protein of HIV-1 also plays an important role in virus particle assembly. A study employing Gag-β-galactosidase fusions showed that the CA domain was essential for incorporation of the mutant PrGag into wt particles (161). A similar complementation study further mapped this activity to the C-terminal portion of the CA domain (27), and a short peptide corresponding to a C-terminal sequence was able to inhibit assembly of wt virus particles (117). Deletions, insertions and mutations made in the N-terminal domain of HIV-1 CA had little or no effect on virus particle assembly (133, 151, 159, 174). Conversely, many mutations in the C-terminal domain of CA were found to significantly impair assembly (43, 91, 104, 133, 151, 174). The MHR and C-terminus of CA were also found to be important for the incorporation of Pr160$^{gag-pol}$ into virus particles (80, 151).

Similar to M-MuLV CA, a number of mutations in the CA domain, especially the N-terminus and MHR, which did not affect assembly, were blocked in a post-assembly process (104, 133). Unlike M-MuLV, there is no evidence that HIV-1 CA is associated with pre-integration complexes (49). HIV-1 CA has been shown to specifically direct the incorporation of cyclophilin A (CypA) into virus particles via a proline rich binding loop in the N-terminal domain (18, 50), and abrogating this interaction imposes a block in capsid uncoating before reverse transcription in certain strains of HIV and SIV (18, 103).

A number of high resolution structures of retroviral capsid proteins and fragments have recently been made available (see Figure 4); structures of the HIV-1 CA N- and C- terminal domains individually (56, 57, 61, 110), as well as

the full-length CA protein (11), have been solved by X-ray crystallization

techniques, or multidimensional NMR techniques; structures of HIV-1 and M-

MuLV MHR peptides have been solved by NMR (29, 30); the full-length

structure of the equine infectious anemia virus (EIAV) CA protein has been

determined by X-ray crystallography (85),; and the full length structure of HTLV-

1 CA has been solved by NMR (92). All of these studies show CA as a protein

with two well defined alpha-helical domains joined by a flexible linker. In all the

X-ray structures, except that of the full length HIV-1 CA, the N- and C- terminal

domains form homodimers. In the structures of the C-terminal domains and the

MHR peptide analog structures, the MHR adopts an alpha-helical conformation

which is involved in creating a hydrogen bonding network with the other alpha-

helices in the domain that stabilizes the core of the C-terminal domain (11, 57, 85,

92).

Nucleocapsid

The nucleocapsid domain is responsible for binding and encapsidation of

viral genomic RNA (8, 158). Nucleocapsid proteins from every retrovirus except

those in the spumavirus group have one or two copies of a cysteine-histidine

motif following the pattern, $CX_2CX_4HX_4C$ (158), which are similar to zinc-finger

motifs and have been shown to tightly bind Zn^{++} ions (13, 154). There is one such

motif in the NC of M-MuLV and two in that of HIV-1, and mutations in both

have been shown to impair incorporation of viral genomic RNA (3, 28, 44, 107,

176). Although the mechanism remains unclear, NC seems to provide specificity

for viral RNA (9, 10, 64, 176). Other studies have indicated other roles for NC, by

showing that the NC protein is capable of promoting dimerization of the viral genomic RNAs, can expedite the reverse transcription reaction by assisting in annealing of the primer tRNA to template and promoting strand transfer of RT, and may contribute to genetic variability (6, 52, 63, 132, 137, 138).

The nucleocapsid also is important in virus particle assembly. Deletions in NC, especially deletions in the N-terminal portion, have been shown to significantly impair virus particle assembly (60, 77, 78), indicating the presence of an assembly domain in this region. These results are corroborated by findings that mutations in the N-terminal basic domain of NC (12, 17, 175), as well as in other regions of NC (109, 122), could severely diminish particle assembly. Suprisingly, replacement of the NC with a heterologous protein oligomerization domain was compatible with efficient assembly (175). In addition, nucleocapsid may be involved in the determination of particle size, since mutations in NC produced proteins which assembled particles with a lower-than-normal density (7). Additionally, HIV-1 NC has recently been found to interact with actin (101) which might have implications for transport and assembly of PrGag. Structures of both the HIV-1 (111, 154) and the M-MuLV (39) NC domains have been determined (see Figure 4) and have a compact, folded structure in the presence of zinc ions (154).

Other Gag Proteins

The remaining domains, p12 in M-MuLV, and p6 in HIV-1, both appear to act late in the assembly process during particle budding from the membrane (52,

65, 125, 155, 173). Sequences responsible for this function, the late or L domain, are short, proline-rich, motifs, often containing the amino acid sequence PPPY, can be found in various locations in a number of different retroviruses (155, 173). The functionality of these motifs has been found to be interchangeable between retroviruses and is positionally independent (125). Mutations in the L domain produce a phenotype in which Gag proteins assemble normally and begin to bud through the membrane, but are trapped just prior to particle release from the cell surface (155). In the M-MuLV p12 domain, two copies of a PPPY-type motif are present (173), but HIV-1 contains no such motifs (125). However, deletion of the p6 domain yields a phenotype similar to that found if the L domain is ablated in a number of other retroviruses and this phenotype can be rescued by replacing the p6 domain with an L domain from RSV (125). This function can be mapped to a PTAP sequence at the N-terminus of the p6 domain (65, 80). Similar to other Gag domains, both p12 and p6 have been shown to serve additional functions in the viral life cycle. Mutants with deletions of the p12 domain were shown to be blocked in at a stage in infection just prior to reverse transcription (173). The p6 domain of HIV-1 also is responsible for directing incorporation of the viral Vpr accessory protein (52), and a mutant in the C-terminal portion of p6 was shown to block incorporation of HIV-1 envelope proteins into particles (121).

Virus Particle Assembly

Production of an infectious retroviral particle requires the coordination of a number of different components and their assembly into a complex, macromolecular structure. Viral envelope proteins, genomic RNA, replicative

proteins and other accessory proteins, both viral and cellular, must be incorporated into the structure in a highly regulated fashion and this organization must occur in an appropriate location for the release of the virus particle (52, 158). In many viruses, and certainly in retroviruses, the structure of the assembled virion is not static. In order to generate an infectious particle the assembled, immature virus particle undergoes a maturation process involving a significant rearrangement of protein contacts and a corresponding change in particle morphology from immature to mature (36, 87, 155, 157).

In retroviruses the Gag polyprotein is responsible for the generation of the virus particle structure (36, 155). Formation of this structure requires regular and repeated interprotein contacts between Gag proteins termed assembly or I domains (167). Knowledge about retroviral assembly and the structure of assembled virions comes from a number of sources. Most information comes from genetic studies in which the *gag* gene is modified then used to express viruses which are analyzed for particle release, content, and/or infectivity (36, 155, 167). Similar approaches have been used in biochemical studies of particle structure (70) and for *in vitro* assembly of virus particles (24, 25, 46). In addition electron microscopy (EM) has been employed to examine particle morphology (59, 95, 124) as well as the substructure of the virus (55, 58, 116, 170). Efforts to elucidate high-resolution structural information from EM or X-ray of entire particles, as has been done with other viruses (86), have been hampered by the fact that retroviruses exhibit markedly heterogeneous morphology (158). Finally, high-resolution structures of individual Gag proteins (see Figure 4) have been determined by X-ray and NMR methodologies (11, 29, 39, 48, 56, 57, 61, 74, 85,

105, 106, 110, 131), but as yet have not been well correlated with information from genetic-based assays or with the higher-order structure of the virus.

Assembly Domains

As discussed in previous sections, delineation of assembly domains within Gag has been approached using deletions and mutations in particle release assays as well as complementation-type incorporation assays. In regards to HIV-1, the primary assembly domain begins with the MHR, spans the C-terminal portion of CA into the N-terminal portion of NC and includes the p2 domain (2, 12, 15, 91, 96, 104, 151, 159-161, 174, 175). Evidence from M-MuLV, although not as extensive as that in HIV-1, also points to a similar assembly domain (4, 5, 62, 69, 70, 79, 89). It also has been shown by chemical cross-linking that the MHR's of adjacent M-MuLV capsid proteins are in close proximity to each other in immature virus particles (70), supporting the notion that these regions interact closely inside immature virus particles.

Another assembly domain has been proposed in HIV-1 matrix domain. Mutations and deletions in this region have been shown to impair assembly while retaining appropriate membrane targeting and association properties (26, 54, 112, 159). In addition, mutations in this region have been shown to impair trimerization of MA and MA-CA proteins *in vitro* (113). However, several studies have shown that deletion of the majority of the MA domain is compatible with production of virus particles (15, 134) that maintain infectivity when pseudotyped with a heterologous envelope protein (162). These results suggest

that the MA assembly domain may not be critical for generation of appropriate virion structure.

Maturation and Assembly

Maturation of the virion is a temporally as well as spatially regulated process. The cleavage of PrGag occurs in a temporally ordered fashion regulated by the accessibility of the cleavage site to the protease, sequence of the cleavage site, and other factors (155, 157). In several retroviruses the spacer peptide between the CA and NC domains, SP1, has been shown to be cleaved at the junction adjacent to the NC domain much faster than at the junction adjacent to the CA domain (155, 157). In HIV-1, this regulation is essential for proper particle assembly (68, 96, 126), indicating that the CA-SP1 intermediate may represent a step in assembly distinct from the fully immature (PrGag) or the fully mature (CA) stages. Processing of the N-terminal portion of the HIV-1 CA domain has also been shown to induce a conformational change in structural studies (48, 145).

In HIV-1 the cellular peptidyl-prolyl isomerase cyclophilin A (cypA) is incorporated into virions via interactions with a proline-rich loop in the N-terminal domain of the capsid protein (32). Incorporation of cypA into virions is important for virus infectivity, possibly through the action of the cypA as a destabilization factor for capsid uncoating (32). Blocking incorporation of cypA into virions by treating producer cells with the drug cyclosporine has been

shown in some studies to block virus maturation (153), however other studies report no such block (166).

In Vitro Assembly Systems

A recent advancement in the study of retroviral structure and assembly has been the development of *in vitro* assembly techniques. Capsid-like structures have been generated using PrGag from Mason-Pfizer monkey virus (M-PMV) produced in bacteria (94), HIV-1 PrGag produced in bacteria (25) or by a reticulate lysate system (150), and using CA-NC from RSV produced in bacteria (25). With both RSV and HIV-1, addition of RNA increased the efficiency of assembly and induced the formation of hollow cylinders, the lengths of which were dependent upon the length of the RNA used, indicating a role for RNA in assembly (25). The importance of RNA in assembly was also suggested by studies which showed that the presence of RNA regulated multimerization of HIV-1 Gag proteins lacking the p6 domain as well as size of spherical particles (24), although these results have been difficult to confirm *in vivo*. In HIV-1, the CA domain by itself was shown to form spherical or cylindrical particles and alterations in the ionic strength or pH (46), or extension of the N-terminus of CA (67), promoted formation of spherical particles heterogeneous in size. HIV-1 CA-NC was shown to form only tubular or conical structures when the spacer peptide, p2, was deleted, suggesting proteolytic cleavage of this region from CA during maturation may promote condensation of the core (68).

EM Studies

Electron microscopy (EM) was one of the earliest methods used to study the structure of retroviral particles but recent advances in *in vitro* assembly systems and computational techniques have greatly advanced its power. At the gross level, EM and immuno-gold labeling EM have been used to identify placement of viral antigens in immature and mature virus particles (21, 59, 76, 124), to examine the effects of mutations on the structure of the virus particles (81, 95), and to examine structures formed in the *in vitro* assembly systems discussed above. Some EM studies have suggested that retroviruses display icosahedral symmetry (21, 115, 116) but this assertion has not been supported by evidence showing that retrovirus particles are pleomorphic and seem to lack of icosahedral symmetry (55, 170). Isolated core structures from HIV-1 (164) as well as *in vitro* assembled cores also have been characterized by EM (58).

In other virus systems virion structure was identified at a high enough resolution to be able to identify subunit structure and even atomic level protein-protein contacts that form the structure of the virus itself (86, 87, 100). Due to the caveats listed above, high-resolution structural determination has not been possible with retroviruses. Instead, structural information has been gleaned from very low-resolution studies, i.e. EM of entire virus particles, and very high resolution structures, i.e. X-ray and NMR structures. In HIV-1 Gag proteins assembling on insect cell membranes were analyzed by EM and computer image processing (116). These results showed Gag in a "fullerene-like" arrangement, with protein units forming five or six membered rings that share a unit with the

neighboring ring, spaced at about 68 Å. These results have been corroborated by cryo-EM studies of HIV-1 virus-like particles (55). Cryo-EM and computer image processing also were used to show that immature M-MuLV virions exhibit a paracrystalline, non-icosahedral packing of Gag-RNA complexes, which are spaced at 45 Å intervals and are approximately 200 Å in length (170). Conclusions from these studies indicate that immature retrovirus particles are not icosahedral but do exhibit local, paracrystalline order.

REFERENCES

1. 1998. Report on the global HIV/AIDS epidemic- June 1998. United Nations Joint Programme on HIV/AIDS and World Health Organization.

2. **Accola, M. A., S. Hoglund, and H. G. Gottlinger.** 1998. A putative α-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. J. Virol. **72:**2072-8.

3. **Aldovini, A., and R. A. Young.** 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J. Virol. **64:**1920-6.

4. **Alin, K., and S. P. Goff.** 1996. Amino acid substitutions in the CA protein of Moloney murine leukemia virus that block early events in infection. Virology. **222:**339-51.

5. **Alin, K., and S. P. Goff.** 1996. Mutational analysis of interactions between the Gag precursor proteins of murine leukemia viruses. Virology. **216:**418-24.

6. **Allain, B., J.-B. Rascle, H. d. Rocquigny, B. Roques, and J.-L. Darlix.** 1998. *CIS* elements and *Trans*-acting factors required for minus strand DNA transfer during reverse transcriptioin of the genomic RNA of murine leukemia virus. J. Mol. Biol. **277:**225-35.

7. **Bennett, R. P., T. D. Nelle, and J. W. Wills.** 1993. Functional chimeras of the Rous sarcoma virus and human immunodeficiency virus *gag* proteins. J. Virol. **67:**6487-98.

8. **Berkowitz, R., J. Fisher, and S. P. Goff.** 1996. RNA Packaging. Curr. Top. Microbiol. Immunol. **214:**177-218.

9. **Berkowitz, R. D., and S. P. Goff.** 1994. Analysis of binding elements in the human immunodeficiency virus type 1 genomic RNA and nucleocapsid protein. Virology. **202:**233-46.

10. **Berkowitz, R. D., J. Luban, and S. P. Goff.** 1993. Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid to viral RNAs detected by RNA mobility shift assays. J. Virol. **67:**7190-7200.

11. **Berthet-Colominas, C., S. Monaco, A. Novelli, G. Sibai, F. Mallet, and S. Cusack.** 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:**1124-36.

12. **Berthoux, L., C. Pechoux, M. Ottman, G. Morel, and J.-L. Darlix.** 1997. Mutations in the N-terminal domain of human immunodeficiency virus type 1 nucleocapsid protein affect virion core structure and proviral DNA synthesis. J. Virol. **71:**6973-81.

13. **Bess, J. W. J., P. J. Powell, H. J. Issaq, L. J. Schumack, M. K. Grimes, L. E. Henderson, and L. O. Arthur.** 1992. Tightly bound zinc in human immunodeficiency virus type 1, human T-cell leukemia virus type 1, and other retroviruses. J. Virol. **66:**840-7.

14. **Best, S., P. L. Tissier, G. Towers, and J. P. Stoye.** 1996. Positional cloning of the mouse retrovirus restriction gene *Fv1*. Nature. **382:**826-9.

15. **Borsetti, A., Å. Ohagen, and H. G. Gottlinger.** 1998. The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. J. Virol. **72:**9313-17.

16. **Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus.** 1989. A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev. **3:**469-78.

17. **Bowzard, J. B., R. P. Bennett, N. K. Krishna, S. M. Ernst, A. Rein, and J. W. Wills.** 1998. Importance of basic residues in the nucleocapsid sequence for retrovirus Gag assembly and complementation rescue. J. Virol. **72:**9034-44.

18. **Braaten, D., E. K. Franke, and J. Luban.** 1996. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. J. Virol. **70.**

19. **Brown, P. O.** 1997. Integration, p. 161-204. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

20. **Bryant, M., and L. Ratner.** 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. Proc. Natl. Acad. Sci. USA. **89:**3443-47.

21. **Bugelski, P. J., B. E. Maleeff, A. M. Klinkner, J. Ventre, and T. K. Hart.** 1995. Ultrastructural evidence of an interaction between Env and Gag proteins during assembly of HIV type 1. AIDS Res. Hum. Retroviruses. **11:**55-64.

22. **Bukrinsky, M. I., S. Haggerty, M. P. Dempsey, N. Sharova, A. Adzhubel, L. Spitz, D. Goldfarb, M. Emerman, and M. Stevenson.** 1993. A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. Nature. **365:**666-70.

23. **Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G. Tarpley, and M. Stevenson.** 1993. Association of integrase, matrix and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleaic acids following acute infection. Proc. Natl. Acad. Sci. USA. **90:**6125-29.

24. **Campbell, S., and S. Rein.** 1999. *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J. Virol.

25. **Campbell, S., and V. M. Vogt.** 1995. Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. **69:**6487-97.

26. **Cannon, P. M., S. Matthews, N. Clark, E. D. Byles, O. Iourin, D. J. Hockley, S. M. Kingsman, and A. J. Kingsman.** 1997. Structure-function studies of the human immunodeficiency virus type 1 matrix protein p17. J. Virol. **71:**3474-83.

27. **Carriere, C., B. Gay, N. Chazal, N. Morin, and P. Boulanger.** 1995. Sequence requirements for encapsidation of deletion mutants and chimeras of human immunodeficiency virus type 1 gag precursor into retrovirus-like particles. J. Virol. **69:**2366-77.

28. **Clavel, F., and J. M. Orenstein.** 1990. A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal particle morphology. J. Virol. **64:**5230-4.

29. **Clish, C. B., D. H. Peyton, and E. Barklis.** 1998. Solution structures of human immunodeficiency virus type 1 (HIV-1) and moloney murine leukemia virus (MoMLV) capsid protein major-homology-region peptide analogs by NMR spectroscopy. European Journal of Biochemistry. **257:**69-77.

30. **Clish, C. B., D. H. Peyton, and E. Barklis.** 1996. Spectroscopic study of an HIV-1 capsid protein major homology region peptide analog. FEBS Letters. **378:**43-7.

31. **Coffin, J. M., S. H. Hughes, and H. E. Varmus (ed.).** 1997. Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

32. **Cohen, E. A., R. A. Subbramanian, and H. G. Gottlinger.** 1996. Role of auxiliary proteins in retroviral morphogenesis. Curr. Top. Microbiol. Immunol. **214:**219-235.

33. **Conte, M. R., and S. Matthews.** 1998. Retroviral matrix proteins: a structural perspective. Virology. **246:191-8.**

34. **Cosson, P.** 1996. Direct interaction between the envelope and matrix proteins of HIV-1. EMBO J. **15:**5783-88.

35. **Craven, R. C., A. E. Leure-duPree, R. A. J. Weldon, and J. W. Wills.** 1995. Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. J. Virol. **69:**4213-27.

36. **Craven, R. C., and L. J. Parent.** 1996. Dynamic Interactions of the Gag Polyprotein. Curr. Top. Microbiol. Immunol. **214:**65-94.

37. **Crawford, S., and S. P. Goff.** 1984. Mutations in Gag proteins p12 and p15 of Moloney murine leukemia virus block early stages of infection. J. Virol. **49:**909-17.

38. **Delchambre, M., D. Gheysen, D. Thines, C. Thiriart, E. Jacobs, E. Verdin, M. Horth, A. Burny, and F. Bex.** 1989. The GAG precursor of simian immunodeficiency virus assembles into virus-like particles. EMBO J. **8:**2653-60.

39. **Demene, H., N. Julian, N. Morellet, H. d. Rocquigny, F. Cornille, B. Maigret, and B. P. Roques.** 1994. Three-dimesional 1H-NMR structure of the nucleocapsid protein NCp10 of Moloney murine leukemia virus. J. Biomol. NMR. **4:**153-70.

40. **Deminie, C. A., and M. Emerman.** 1993. Functional exchange of an oncoretrovirus and a lentivirus matrix protein. J. Virol. **68:**4442-9.

41. **DesGroseillers, L., and P. Jolicoeur.** 1983. Physical mapping of the *Fv-1* tropism host range determinant of BALB/c murine leukemia viruses. J. Virol. **48:**685-96.

42. **Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich.** 1984. Protein biosynthesis and assembly, p. 513-648. *In* R. Weiss and e. al. (ed.), Molecular biology of tumor viruses, 2nd ed, vol. RNA Tumor Viruses 1. Cold Spring Harbor Press, Cold Spring Harbor, NY.

43. **Dorfman, T., A. Bukovsky, A. Ohagen, S. Hoglund, and H. G. Gottlinger.** 1994. Functional domains of the capsid protein of human immunodeficiency virus type 1. J. Virol. **68.**

44. **Dorfman, T., J. Luban, S. P. goff, W. A. Haseltine, and H. G. Gottlinger.** 1993. Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. J. Virol. **67:**6159-69.

45. **Dorfman, T., F. Mammano, W. A. Haseltine, and H. G. Gottlinger.** 1994. Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J. Virol. **68:**1689-96.

46. **Ehrlich, L. S., B. E. Agresta, and C. A. Carter.** 1992. Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro*. J. Virol. **66:**4874-83.

47. **Emini, E. A., and H. Y. Fan.** 1997. Immunological and pharmacological approaches to the control of retroviral infections, p. 637-708. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

48. **Endrich, M. M., P. Gehrig, and H. Gehring.** 1999. Maturation-induced conformational changes of HIV-1 capsid protein and identification of two high affinity sites for cyclophilins in the C-terminal domain. J. Biol. Chem. **274:**5326-32.

49. **Fouchier, R. A., and M. H. Malim.** 1999. Nuclear import of human immunodeficiency virus type-1 preintegration complexes. Adv. Virus Res. **52:**275-99.

50. **Franke, E. K., H. E. Yuan, and J. Luban.** 1994. Specific incorporation of cyclophilin A into HIV-1 virions. Nature. **372:**359-62.

51. **Frankel, A. D., and J. A. T. Young.** 1998. HIV-1: Fifteen proteins and an RNA. Annu. Rev. Biochem. **67:**1-25.

52. **Freed, E. O.** 1998. HIV-1 Gag proteins: Diverse functions in the virus life cycle. Virology. **251:**1-15.

53. **Freed, E. O., and M. A. Martin.** 1995. Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tials is blocked by specific, single amino acid substitutionsin the human immunodeficiency virus type 1 matrix protein. J. Virol. **69:**1984-1989.

54. **Freed, E. O., J. M. Orenstein, A. J. Buckler-White, and M. A. Martin.** 1994. Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. J. Virol. **68:**5311-20.

55. **Fuller, S. D., T. Wilk, B. E. Gowen, H. G. Krausslich, and V. M. Vogt.** 1997. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. Curr. Biol. **7:**729-38.

56. **Gamble, T. R., F. F. Vajdos, S. Yoo, D. K. Worthylake, M. Houseweart, W. I. Sundquist, and C. P. Hill.** 1996. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell. **87:**1285-94.

57. **Gamble, T. R., S. Yoo, F. F. Vajdos, U. K. v. Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist, and C. P. Hill.** 1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. Science. **278:**849-53.

58. **Ganser, B. K., S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist.** 1999. Assembly and analysis of conical models for the HIV-1 core. Science. **283:**80-3.

59. **Gelderblom, H. R., E. H. Hausmann, M. Ozel, and G. Pauli.** 1989. Morphogenesis and morphology of HIV. Structure-function relationships. Arch. Virol. **106:**1-13.

60. **Gheysen, D., E. Jacobs, F. d. Foresta, C. Thiriart, M. Francotte, D. Thines, and M. D. Wilde.** 1989. Assembly and release of HIV-1 precursor Pr55gag

virus-like particles from recombinant baculovirus infected cells. Cell.
59:103-12.

61. **Gitti, R. K., B. M. Lee, J. Walker, M. F. Summers, S. Yoo, and W. I.
     Sundquist.** 1996. Structure of the amino-terminal core domain of the HIV-
     1 capsid protein. Science. **273**:231-5.

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

63. **Gorelick, R. J., D. J. Chabot, D. E. Ott, T. D. Gagliardi, A. Rein, L. E.
     Henderson, and L. O. Arthur.** 1996. Genetic analysis of the zinc finger in
     the Moloney murine leukemia virus nucleocapsid domain: Replacement
     of zinc-coordinating residues with other zinc-coordinating residues yields
     noninfectious particles containing genomic RNA. J. Virol. **70**:2593-97.

64. **Gorelick, R. J., L. E. Henderson, J. P. Hanser, and A. Rein.** 1988. Point
     mutants of Moloney murine leukemia virus that fail to package viral
     RNA: Evidence for specific RNA recognition by a "zinc finger-like" protein
     sequence. Proc. Natl. Acad. Sci. USA. **85**:8420-4.

65. **Gottlinger, H. G., T. Dorfman, J. G. Sodroski, and W. A. Haseltine.** 1991.
     Effect of mutations affecting the p6 Gag protein on human
     immunodeficiency virus particle release. Proc. Natl. Acad. Sci. USA.
     88:3195-9.

66. **Granowitz, C., and S. P. Goff.** 1994. Substitution mutations affecting a
     small region of the Moloney murine leukemia virus MA Gag protein block
     assembly and release of virion particles. Virology. **205**:336-44.

67. **Gross, I., H. Hohenberg, C. Huckhagel, and H.-G. Krausslich.** 1998. N-
     terminal extension of human immunodeficiency virus capsid protein
     converts the in vitro assembly phenotype from tubular to spherical
     particles. J. Virol. **72**:4798-810.

68. **Gross, I., H. Hohenberg, T. Wilk, K. Wiegers, M. Grattinger, B. Muller,
     S. Fuller, and H.-G. Krausslich.** 2000. A conformational switch controlling
     HIV-1 morphogenesis. EMBO J. **19**:103-13.

69. **Hansen, M., L. Jelinek, S. Whiting, and E. Barklis.** 1990. Transport and
     assembly of Gag proteins into Moloney murine leukemia virus particles. J.
     Virol. **64**:5306-5316.

70. **Hansen, M. S. T., and E. Barklis.** 1995. Structural interactions between
     retroviral Gag proteins examined by cysteine cross-linking. J. Virol.
     69:1150-9.

71. **Harrison, S. C.** 1991. Principles of Virus Structure, p. 37-62. *In* B. N. Fields
     and D. M. Knipe (ed.), Fundamental Virology. Raven Press, New York.

72. **Hatfield, D. L., J. G. Levin, A. Rein, and S. Oroszlan.** 1992. Translational
     suppresion in retroviral gene expression. Adv. Virus Res. **41**:193-239.

73. **Henderson, L. E., H. C. Krutzsch, and S. Oroszlan.** 1983. Myristyl amino-
     terminal acylation of murine retrovirus proteins: An unusual post-
     translational protein modification. Proc. Natl. Acad. Sci. USA. **80**:339-43.

74. **Hill, C. P., D. Worthylake, D. P. Bancroft, A. M. Christensen, and W. I.
     Sundquist.** 1996. Crystal structures of the trimeric HIV-1 matrix protein:
     Implications for membrane association and assembly. Proc. Natl. Acad.
     Sci. USA. **93**:3099-104.

75. **Hogle, J. M., M. Chow, and D. J. Filman.** 1985. Three-dimensional
     structure of poliovirus at 2.9Å resolution. Science. **229**:1358-65.

76. **Hoglund, S., L. G. Ofverstedt, A. Nilsson, P. Lundquist, H. Gelderblom, M. Ozel, and U. Skoglund.** 1992. Spatial visualization of the maturing HIV-1 core and its linkage to the envelope. AIDS Res. Hum. Retroviruses. **8**:1-7.

77. **Hoshikawa, N., A. Kojima, A. Yasuda, E. Takayashiki, S. Masuko, J. Chiba, T. Sata, and T. Kurata.** 1991. Role of the *gag* and *pol* genes of human immunodeficiency virus in the morphogenesis and maturation of retrovirus-like particles expressed by recombinant vaccinia virus: An ultrastructural study. J. Gen. Virol. **72**:2509-17.

78. **Housset, V., H. d. Rocquigny, B. P. Roques, and J.-L. Darlix.** 1993. Basic amino acids flanking the zinc finger of Moloney murine leukemia virus nucleocapsid protein NCp10 are critical for virus infectivity. J. Virol. **67**:2537-45.

79. **Hsu, H.-W., P. Schwartzberg, and S. P. Goff.** 1985. Point mutations in the p30 domain of the *gag* gene of Moloney murine leukemia virus. Virology. **142**:211-4.

80. **Huang, M., and M. A. Martin.** 1997. Incorporation of Pr160$^{gag-pol}$ into virus particles requires the presence of both the major homology region and adjacent C-terminal capsid sequences within the Gag-Pol polyprotein. J. Virol. **71**:4472-8.

81. **Hughes, B. P., T. F. Booth, A. S. Belyaev, D. McIlroy, J. Jowett, and P. Roy.** 1993. Morphogenic capabilities of human immunodeficiency virus type 1 *gag* and *gag-pol* proteins in insect cells. Virology. **193**:242-55.

82. **Hunter, E.** 1997. Viral entry and receptors, p. 71-120. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

83. **Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus.** 1988. Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. Nature. **331**:280-3.

84. **Jarvik, J., and D. Botstein.** 1975. Conditional-lethal mutations that suppress genetic defects in morphogenesis by altering structural proteins. Proc. Natl. Acad. Sci. USA. **72**:2738-42.

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

86. **Johnson, J. E.** 1996. Functional implications of protein-protein interactions in icosohedral viruses. Proc. Natl. Acad. Sci. USA. **93**:27-33.

87. **Johnson, J. E., and J. A. Speir.** 1997. Quasi-equivalent viruses: A paradigm for protein assemblies. J. Mol. Biol. **269**:66-75.

88. **Jolicoeur, P.** 1979. The Fv-1 gene of the mouse and its control of murine leukemia virus replication. Curr. Top. Microbiol. Immunol. **86**:67-122.

89. **Jones, T. A., G. Blaug, M. Hansen, and E. Barklis.** 1990. Assembly of Gag-β-galactosidase proteins into retrovirus particles. J. Virol. **64**:2265-2279.

90. **Katoh, I., Y. Yoshinaka, A. Rein, M. Shibuya, T. Odaka, and S. Oroszlan.** 1985. Murine leukemia virus maturation: Protease region required for conversion from "immature" to "mature" core form and for virus infectivity. Virology. **145**:280-92.

91. **Kattenbach, B., A. v. Poblotzki, A. Rohrhofer, H. Wolf, and S. Modrow.** 1997. Inhibition of human immunodeficiency virus type 1 particle formation by alterations of defined amino acids within the C terminus of the capsid protein. J. Gen. Virol. **78:**2489-96.

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

93. **Kiernan, R. E., A. Ono, G. Englund, and E. O. Freed.** 1998. Role of matrix in an early postentry step in the human immunodeficiency virus type 1 life cycle. J. Virol. **72:**4116-26.

94. **Klikova, M., S. S. Rhee, E. Hunter, and T. Ruml.** 1995. Efficienct *in vivo* and *in vitro* assembly of retroviral capsids from Gag precursor proteins expressed in bacteria. J. Virol. **69:**1093-98.

95. **Kong, L. B., D. An, B. Ackerson, J. Canon, O. Rey, I. S. Y. Chen, P. Krogstad, and P. L. Stewart.** 1998. Cryoelectron microscopic examination of human immunodeficiency virus type 1 virions with mutations in the cyclophilin A binding loop. J. Virol. **72:**4403-7.

96. **Krausslich, H.-G., M. Facke, A.-M. Heuser, J. Konvalinka, and H. Zentgraf.** 1995. The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. J. Virol. **69:**3407-19.

97. **Krausslich, H.-G., and R. Welker.** 1996. Intracellular transport of retroviral capsid components. Curr. Top. Microbiol. Immunol. **214:**25-63.

98. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London). **227:**680-5.

99. **Lama, J.** 1998. Human immunodeficiency virus type 1 matrix protein interacts with cellular protein HO3. J. Virol. **72:**1671-76.

100. **Liddington, R. C., Y. Yan, J. Moulai, R. Sahli, T. L. Benjamin, and S. C. Harrison.** 1991. Structure of simian virus 40 at 3.8-Å resolution. Nature. **354:**278-84.

101. **Liu, B., R. Dai, C.-J. Tian, L. Dawson, R. Gorelick, and X.-F. Yu.** 1999. Interaction of the human immunodeficiency virus type 1 nucleocapsid with actin. J. Virol. **73:**2901-8.

102. **Lower, R., J. Lower, and R. Kurth.** 1996. The viruses in all of us: Characteristics and biological significance of human endogenous retrovirus sequences. Proc. Natl. Acad. Sci. USA. **93:**5177-84.

103. **Luban, J.** 1996. Absconding with the chaperone: Essential cyclophilin-Gag interaction in HIV-1 virions. Cell. **87:**1157-9.

104. **Mammano, F., A. Ohagen, S. Hoglund, and H. G. Gottlinger.** 1994. Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. J. Virol. **68:**4927-36.

105. **Massiah, M. A., M. R. Starich, C. Paschall, M. F. Summers, A. M. Christensen, and W. I. Sundquist.** 1994. Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. J. Mol. Biol. **244:**198-223.

106. **Matthews, S., P. Barlow, J. Boyd, G. Barton, R. Russell, H. Mills, M. Cunningham, N. Meyers, N. Burns, N. Clark, S. Kingsman, A. Kingsman, and I. Campbell.** 1994. Structural similarity between the p17 matrix protein of HIV-1 and interferon-γ. Nature. **370:**666-68.

107. **Meric, C., and S. P. Goff.** 1989. Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. J. Virol. **63:**1558-68.

108. **Miller, A. D.** 1997. Development and applications of retroviral vectors, p. 437-74. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

109. **Mizuno, A., E. Ido, T. Goto, T. Kuwata, M. Nakai, and M. Hayami.** 1996. Mutational analysis of two zinc finger motifs in HIV Type 1 nucleocapsid proteins: Effects on proteolytic processing of Gag precursors and particle formation. AIDS Res. Hum. Retroviruses. **12:**793-800.

110. **Momany, C., L. C. Kovari, A. J. Prongay, W. Keller, R. K. Gitti, B. M. Lee, A. E. Gorbalenya, L. Tong, J. McClure, L. S. Erlich, M. F. Summers, C. Carter, and M. G. Rossman.** 1996. Crystal structure of dimeric HIV-1 capsid protein. Nature Struct. Biol. **3:**763-70.

111. **Morellet, N., N. Jullian, H. d. Rocquigny, B. Maigret, J.-L. Darlix, and B. P. Roques.** 1992. Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by ^{1}H NMR. EMBO J. **11:**3059-65.

112. **Morikawa, Y., T. Kishi, W. H. Zhang, M. V. Nermut, D. J. Hockley, and I. M. Jones.** 1995. A molecular determinant of human immunodeficiency virus particle assembly located in matrix antigen p17. J. Virol. **69:**4519-23.

113. **Morikawa, Y., W.-H. Zhang, D. J. Hockley, M. V. Nermut, and I. M. Jones.** 1998. Detection of a trimeric human immunodeficiency virus type 1 Gag intermediate is dependent on sequences in the matrix protein, p17. J. Virol. **72:**7659-63.

114. **Mosser, A. G., D. M. Lieppe, and R. R. Rueckert.** 1988. Neutralization of picornanaviruses: support for the pentamer bridging hypothesis, p. 155-67. *In* B. Semler and E. Ehrenfeld (ed.), Molecular aspects of picornavirus infection and detection. ASM Publications, Washington DC.

115. **Nermut, M. V., C. Grief, S. Hashmi, and D. J. Hockley.** 1993. Further evidence of icosahedral symmetry in human and simian immunodeficiency virus. AIDS Res. Hum. Retroviruses. **9:**929-38.

116. **Nermut, M. V., D. J. Hockley, J. B. M. Jowett, I. M. Jones, M. Garreau, and D. Thomas.** 1994. Fullerene-like organization of HIV gag-protein shell in virus-like particles produced by recombinant baculovirus. Virology. **198:**288-96.

117. **Niedrig, M., H. R. Gelderblom, G. Pauli, J. Marz, H. Bickhard, H. Wolf, and S. Modrow.** 1994. Inhibition of infectious human immunodeficiency virus type 1 particle formation by Gag protein-derived peptides. J. Gen. Virol. **75:**1469-74.

118. **Ono, A., and E. O. Freed.** 1999. Binding of human immunodeficiency virus type 1 Gag to membrane: Role of the matrix amino terminus. J. Virol. **73:**4136-44.

119. **Ono, A., M. Huang, and E. O. Freed.** 1997. Characterization of human immunodeficiency virus type 1 matrix revertants: Effects on virus assembly, Gag processing, and Env incorporation into virions. J. Virol. **71:**4409-18.

120.   **Ono, A., J. M. Orenstein, and E. O. Feed.** 2000. Role of Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. J. Virol. **74:**2855-66.

121.   **Ott, D. E., E. N. Chertova, L. K. Busch, L. V. Coren, T. G. Gagliardi, and D. G. Johnson.** 1999. Mutational analysis of the hydrophobic tail of the human immunodeficiency virus type 1 p6^{Gag} protein produces a mutant that fails to package its envelope protein. J. Virol. **73:**19-28.

122.   **Ottmann, M., C. Gabus, and J.-L. Darlix.** 1995. The central globular domain of the nucleocapsid protein of human immunodeficiency virus type 1 is critical for virion structure and infectivity. J. Virol. **69:**1778-84.

123.   **Paillart, J.-C., and H. G. Gottlinger.** 1999. Opposing effects of human immunodeficiency virus type 1 matrix mutations support a myristyl switch model of Gag membrane targeting. J. Virol. **73:**2604-12.

124.   **Palmer, E., C. Sporborg, A. Harrison, M. L. Martin, and P. Feorino.** 1985. Morphology and immunoelectron microscopy of AIDS virus. Arch. Virol. **85:**189-96.

125.   **Parent, L. J., R. P. Bennett, R. C. Craven, T. D. Nelle, N. K. Krishna, J. B. Bowzard, C. B. Wilson, B. A. Puffer, R. C. Montelaro, and J. W. Wills.** 1995. Positionally independent and exchangeable late budding functions of the Rous sarcoma virus and human immunodeficiency virus Gag proteins. J. Virol. **69:**5455-60.

126.   **Petit, S. C., M. D. Moody, R. S. Wehbie, A. H. Kaplan, P. V. Nantermet, C. A. Klein, and R. Swanstrom.** 1994. The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. J. Virol. **68:**8017-27.

127.   **Peytavi, R., S. S. Hong, B. Gay, A. D. d'Angeac, L. Selig, S. Benichou, E. Benarous, and P. Boulanger.** 1999. HEED, the product of the human homolog of the murine eed gene, binds to the matrix protein of HIV-1. J. Biol. Chem. **274:**1635-45.

128.   **Pryciak, P. M., and H. E. Varmus.** 1992. *Fv-1* restriction and its effects on murine leukemia virus integration *in vivo* and *in vitro.* J. Virol. **66:**5959-66.

129.   **Pryciak, P. M., and H. E. Varmus.** 1992. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration *in vivo* and *in vitro.* Cell. **69:**769-80.

130.   **Rabson, A. B., and B. J. Graves.** 1997. Synthesis and processing of viral RNA, p. 205-62. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

131.   **Rao, A., A. S. Belyaev, E. Fry, P. Roy, I. M. Jones, and D. I. Stuart.** 1995. Crystal structure of SIV matrix antigen and implications for virus particle assembly. Nature. **378:**743-7.

132.   **Rascle, J.-B., D. Ficheux, and J.-L. Darlix.** 1998. Possible roles of nucleocapsid protein of MoMuLV in the specificity of proviral DNA synthesis and in the genetic variability of the virus. J. Mol. Biol. **280:**215-25.

133.   **Reicin, A. S., S. Paik, R. D. Berkowitz, J. Luban, I. Lowy, and S. P. Goff.** 1995. Linker insertion mutations in the human immunodeficiency virus

type 1 *gag* gene: Effects on virion particle assembly, release, and infectivity. J. Virol. **69**:642-50.

134. **Reil, H., A. A. Bukovsky, H. R. gelderblom, and H. G. Gottlinger.** 1998. Efficient HIV-1 replication can occur in the absence of the viral matrix protein. EMBO J. **17**:2699-708.

135. **Rein, A., M. R. McClure, N. R. Rice, R. B. Luftig, and A. M. Schultz.** 1986. Myristylation site in Pr65^{Gag} is essential for virus particle formation by Moloney murine leukemia virus. Proc. Natl. Acad. Sci. USA. **83**:7246-50.

136. **Rhee, S. S., and E. Hunter.** 1990. Structural role of the matrix protein of type D retroviruses in Gag polyprotein stability and capsid assembly. J. Virol. **64**:4383-89.

137. **Rocquigny, H. D., C. Gabus, A. Vincent, M. C. Fournie-Zaluski, B. Roques, and J.-L. Darlix.** 1992. Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. Proc. Natl. Acad. Sci. USA. **89**:6472-6.

138. **Rodriguez-Rodriguez, L., Z. Tsuchihashi, G. M. Fuentes, R. A. Bambara, and P. J. Fay.** 1995. Influence of human immunodeficiency virus nucleocapsid protein on synthesis and strand transfer by the reverse transcriptase *in vitro*. J. Biol. Chem. **270**:15005-11.

139. **Rossman, M. G., E. Arnold, and J. W. Erickson.** 1985. The structure of a human common cold virus (rhinovirus 14) and its functional relations to other picornaviruses. Nature (London). **317**:145-53.

140. **Royer, M., M. Cerutti, B. Gay, S. S. Hong, G. Devauchelle, and P. Boulanger.** 1991. Functional domains of HIV-1 Gag-polyprotein expressed in baculovirus-infected cells. Virology. **184**:417-22.

141. **Salunke, D. M., D. L. D. Caspar, and R. L. Garcea.** 1989. Polymorphism in the assembly of polyoma virus capsid protein VP1. Biophys. J. **56**:887-900.

142. **Salunke, D. M., D. L. D. Caspar, and R. L. Garcea.** 1986. Self-assembly of purified polyomavirus capsid protien VP1. Cell. **46**:895-904.

143. **Schultz, A. M., and A. Rein.** 1989. Unmyristylated Moloney murine leukemia virus Pr65Gag is excluded from virus assembly and maturation events. J. Virol. **63**:2370-73.

144. **Schwartz, R., R. L. Garcea, and B. Berger.** 2000. "Local rules" theory applied to polyomavirus polymorphic capsid assemblies. Virology. **268**:461-70.

145. **Schwedler, U. K. v., T. L. Stemmler, V. Y. Klishko, S. Li, K. H. Albertine, D. R. Davis, and W. I. Sundquist.** 1998. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J. **17**:1555-68.

146. **Seri, I., and E. Priel.** 1996. The p15 matrix protein of moloney murine leukemia virus is a phosphotyrosine protein. FEBS Letters. **392**:229-32.

147. **Shioda, T., and H. Shibuta.** 1990. Production of human immunodeficiency virus (HIV)-like particles from cells infected with recombinant vaccinia viruses carrying the *gag* gene of HIV. Virology. **175**:139-48.

148. **Skalka, A. M., and S. P. Goff (ed.).** 1993. Reverse Transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

149.   **Soneoka, Y., S. M. Kingsman, and A. J. Kingsman.** 1997. Mutagenesis analysis of the murine leukemia virus matrix protein: Idenification of regions important for membrane localization and intracellular transport. J. Virol. **71:**5549-59.

150.   **Spearman, P., and L. Ratner.** 1996. Human immunodeficiency cirus type 1 capsid formation in reticulocyte lysates. J. Virol.

151.   **Srinivasakumar, N., M.-L. Hammarskjold, and D. Rekosh.** 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:**6106-14.

152.   **Strambio-de-Castillia, C., and E. Hunter.** 1992. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. J. Virol. **66:**7021-32.

153.   **Streblow, D. N., M. Kitabwalla, M. Malkovsky, and C. D. Pauza.** 1998. Cyclophilin a modulates processing of human immunodeficiency virus type 1 p55Gag: mechanism for antiviral effects of cyclosporin A. Virology. **245:**197-202.

154.   **Summers, M. F., L. E. Henderson, M. R. Chance, J. W. J. Bess, T. L. South, P. R. Blake, I. Sagi, G. Perez-Alvarado, R. C. S. III, D. R. Hare, and L. O. Arthur.** 1992. Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. Protein. Sci. **1:**563-74.

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

156.   **Telesnitsky, A., and S. P. Goff.** 1997. Reverse transcriptase and the generation of retroviral DNA, p. 121-60. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

157.   **Vogt, V. M.** 1996. Proteolytic processing and particle maturation. Curr. Top. Microbiol. Immunol. **214:**95-131.

158.   **Vogt, V. M.** 1997. Retroviral virions and genomes, p. 263-334. *In* J. M. Coffin and S. H. Hughes and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

159.   **Wang, C.-T., and E. Barklis.** 1993. Assembly, processing, and infectivity of human immunodeficiency virus type 1 Gag mutants. J. Virol. **67:**4264-73.

160.   **Wang, C.-T., H.-Y. Lai, and J.-J. Li.** 1998. Analysis of human immunodeficiency virus type 1 gag coding sequences capable of virus-like particle assembly and release. J. Virol. **72:**7950-9.

161.   **Wang, C.-T., J. Stegeman-Olsen, Y. Zhang, and E. Barklis.** 1994. Assembly of HIV GAG-B-Galactosidase fusion proteins into virus particles. Virology. **200:**524-34.

162.   **Wang, C.-T., Y. Zhang, J. McDermott, and E. Barklis.** 1993. Conditional infectivity of a human immunodeficiency virus matrix domain deletion mutant. J. Virol. **67:**7067-76.

163.   **Weldon, R. A. J., and J. W. Wills.** 1993. Characterization of a small (25-kilodalton) derivative of the Rous sarcoma virus Gag protein competent for particle release. J. Virol. **67:**5550-61.

164. **Welker, R., H. Hohenberg, U. Tessmer, C. Huckhagel, and H.-G. Krausslich.** 2000. Biochemical and structural analysis of isolated mature cores of human immunodeficiency virus type 1. J. Virol. **74:**1168-77.

165. **Whitcomb, J. M., and S. H. Hughes.** 1992. Retroviral reverse transcription and integration: Progress and Problems. Annu. Rev. Cell Biol. **8:**275-306.

166. **Wiegers, K., G. Rutter, U. Schubert, M. Grattinger, and H. G. Krausslich.** 1999. Cyclophilin A incorporation is not required for human immunodeficiency virus type 1 particle maturation and does not destabilize the mature capsid. Virology. **257:**261-74.

167. **Wills, J. W., and R. C. Craven.** 1991. Form, function, and use of retroviral Gag proteins. AIDS. **5:**639-54.

168. **Wills, J. W., R. C. Craven, and J. A. Achacoso.** 1989. Creation and expression of myristylated forms of Rous sarcoma virus Gag protein in mammalian cells. J. Virol. **63:**4331-43.

169. **Wyatt, R., and J. Sodroski.** 1998. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science. **280:**1884-8.

170. **Yeager, M., E. M. Wilson-Kubalek, S. G. Weiner, P. O. Brown, and A. Rein.** 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.

171. **Yu, X., X. Yuan, Z. Matsuda, T.-H. Lee, and M. Essex.** 1992. The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. J. Virol. **66:**4966-71.

172. **Yu, X., X. Yuan, M. F. McLane, T.-H. Lee, and M. Essex.** 1993. Mutations in the cytoplasmic domain of human immunodeficiency virus type 1 transmembrane protein impair the incorporation of Env proteins into mature virions. J. Virol. **67:**213-21.

173. **Yuan, B., X. Li, and S. P. Goff.** 1999. Mutations altering the Moloney murine leukemia virus p12 Gag protein affect virion production and early events of the virus life cycle. EMBO J. **18:**4700-10.

174. **Zhang, W., D. J. Hockley, M. V. Nermut, Y. Morikawa, and I. M. Jones.** 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:**743-51.

175. **Zhang, Y., and E. Barklis.** 1997. Effects of nucleocapsid mutations on human immunodeficiency virus assembly and RNA encapsidation. J. Virol. **71:**6765-76.

176. **Zhang, Y., and E. Barklis.** 1995. Nucleocapsid protein effects on the specificity of retrovirus RNA encapsidation. J. Virol. **69:**5716-22.

177. **Zhou, W., L. J. Parents, J. W. Willis, and M. D. Resh.** 1994. Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J. Virol. **68:**2556-69.

178. **Zhou, W., and M. D. Resh.** 1996. Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. J. Virol. **70:**8540-8.

Chapter 2

CROSSLINK ANALYSIS OF N-TERMINAL, C-TERMINAL AND N/B

DETERMINING REGIONS OF THE MOLONEY MURINE LEUKEMIA

VIRUS CAPSID PROTEIN

Jason McDermott[1,2], Sonya Karanjia[1,2], Zachary Love[1], and Eric Barklis[1,3]

[1]Vollum Institute and Department of Microbiology, Oregon Health Sciences University, Portland, Oregon 97201-3098
[2]contributed equally to this work
[3]corresponding author: barklis@ohsu.edu

ABSTRACT

To analyze contacts made by Moloney murine leukemia virus (M-MuLV) capsid
(CA) proteins in immature and mature virus particles, we have employed a
cysteine-specific crosslinking approach which permits the identification of
retroviral Gag protein interactions at particular residues. For analysis, single
cysteine creation mutations were made in the context of protease deficient (PR-),
or PR+ parental constructs. Cysteine creation mutations were chosen near the N-
and C-termini of CA, and at a site adjacent to the M-MuLV glu-ala *Fv1* N/B host
range determination sequence. Analysis of immature virions showed that PrGag
proteins were crosslinked at C-terminal CA residues to form dimers while
crosslinking of particle-associated N-terminal and N/B region mutant proteins
did not yield dimers, but showed evidence of linking to an unknown 140-160
kDa partner. Analysis of mature virions demonstrated that both N- and C-
terminal CA residues participated in dimer formation, suggesting that processed
CA N- and C-termini are free to establish interprotein associations. Interestingly,
N/B region mutant residues in mature virus particles did not crosslink to form
dimers, but showed a novel crosslink band, consistent with an interaction
between the N/B tropism determining region and a cellular protein of 45-55 kDa.

INTRODUCTION

The core (Gag) protein of the Moloney murine leukemia virus (M-MuLV)
is the primary determinant of virus particle structure and expression of Gag has

been shown to be necessary and sufficient for the formation and release of immature virus particles (Craven and Parent, 1996). Gag is synthesized as a myristoylated polyprotein precursor, Pr65^{Gag}, which is cleaved by the viral protease (PR) during or after budding to yield the mature viral proteins matrix (MA), p12, capsid (CA) and nucleocapsid (NC) (Barbacid, Stephenson, and Aaronson, 1976; Bolognesi, Luftig, and Shaper, 1973; Stephenson, Tronick, and Aaronson, 1975). The Gag protein is delivered to the plasma membrane by an undetermined mechanism (Swanstrom and Wills, 1997), assembles there, and buds from the cell to yield a roughly spherical virus particle of approximately 125 nm diameter (Yeager *et al.*, 1998). The MA protein is myristoylated at its N-terminus, associates with membranes, and may interact with the cytoplasmic tail of the viral transmembrane (Env) protein (Swanstrom and Wills, 1997). The role of the p12 domain is not well understood, however deletion of most of the domain does not substantially effect virus particle assembly (Barklis *et al.*, 1997; Crawford and Goff, 1984). The NC protein specifically binds viral genomic RNA and is required for RNA encapsidation as well as serving a function in the viral assembly process (Campbell and Vogt, 1995; Gorelick *et al.*, 1988; Hansen and Barklis, 1995). The CA domain of M-MuLV has been shown to be important in virus particle assembly (Hansen *et al.*, 1990; Jones *et al.*, 1990; Lobel and Goff, 1984) and budding, as well as playing a role early in infection (Alin and Goff, 1996). Contained in the CA domain are the major homology region (MHR), a region of 20 to 30 residues which is the only highly conserved sequence among retroviral capsid proteins (Craven *et al.*, 1995; Mammano *et al.*, 1994; Strambio-de-Castillia and Hunter, 1992), as well as the N/B locus, the determinant of murine C-type NIH/BALB/c (N/B) cell tropism (Brown, 1997; Weiss *et al.*, 1984).

Although it is clear that mutations in the CA domain can effect M-MuLV particle assembly and release from cells (Hansen *et al.*, 1990; Jones *et al.*, 1990; Lobel and Goff, 1984), the exact role of the CA domain in the assembly process remains largely unknown. Studies on HIV-1 have shown that the carboxyl terminus of CA is important for virus particle assembly (McDermott *et al.*, 1996; Reicin *et al.*, 1995) but much of the N-terminal portion is dispensable for virus particle assembly (Borsetti *et al*, 1998; Wang *et al.*, 1994). In Rous sarcoma virus, the essential assembly domains that have been mapped lie outside the CA domain (Swanstrom and Wills, 1997; Weldon and Wills, 1993; Wills *et al.*, 1994). Both the N- and C- terminal portions of the HIV-1 CA have been crystallized as dimers (Gamble *et al.*, 1997; Momany *et al.*, 1996), and recent biophysical studies have revealed some details concerning Gag-Gag interactions of HIV-1 (Barklis *et al.*, 1998; Campbell and Rein, 1999; Fuller *et al.*, 1997; Ganser *et al.*, 1999; Scarlata, Ehrlich, and Carter, 1998), Rous sarcoma virus (Campbell and Vogt, 1995; Campbell and Vogt, 1997; Kovari *et al.*, 1997), and M-MuLV (Barklis *et al.*, 1997; Yeager *et al.*, 1998), but the specific residues involved in interprotein contacts remain to be defined.

As a biochemical approach to examine interactions between CA proteins in M-MuLV, we previously established a system for cysteine crosslink analysis of proteins in M-MuLV virus particles (Hansen and Barklis, 1995). This system, based on the work of others (Pakula and Simon, 1992; Pepinsky, 1983; Pepinsky *et al.*, 1980; Pinter and Fleissner, 1979), allows cysteine residues that are in close proximity on adjacent proteins to be covalently crosslinked through the addition

of cysteine-specific chemical crosslinking reagents such as molecular iodine or *bis*-maleimido hexane (BMH). Cysteines located on the NC domain in immature (Hansen and Barklis, 1995) and mature (Ott *et al.*, 1998; Rein *et al.*, 1996; Rein *et al.*, 1997) M-MuLV particles were efficiently crosslinked by cysteine-specific crosslinking agents. In addition, we found that mutation of all five cysteines present in M-MuLV Gag was compatible with immature virus particle assembly, allowing us to conveniently examine the effects of cysteine creation mutations in immature virions. In our current study, we also have designed an efficiently processed viral construct, containing no CA or MA cysteine residues, so that CA domain interactions can be probed in mature virus particles.

To probe capsid domain contacts, we have made cysteine creation mutations in three different regions of the M-MuLV CA coding region in both PR- and PR+ constructs: the N-terminal region of CA (G220C), the C-terminal region of CA (S475C) and the N/B determining region (T323C; see Figure 1). We observed that in immature virus particles, G220C proteins did not form crosslinked dimers, but S475C proteins did. In contrast, in mature virions, both N- and C-terminally modified proteins crosslinked efficiently to form dimers suggesting that during viral maturation, the N-terminal portion of CA is reoriented so that it can interact with neighbor molecules. Additionally, cellular proteins of 140-160 kDa and 45-55 kDa were identified as CA crosslink partners in immature and mature M-MuLV particles, respectively. These proteins may represent cellular factors which actively participate in the M-MuLV replication cycle.

RESULTS

Assembly of immature and mature virus particles.

To study M-MuLV virus particle assembly and structure, we have
employed a COS cell-based transient transfection system. The vectors used for
transfection were based on the parental construct pXM which carries an SV40
origin of replication that allows replication of the plasmid in COS cells, and the
adenovirus major late promoter, which fosters high levels of transcription
initiation in COS cells. Our parental PR- construct (pXM2453) deletes M-MuLV
pol and *env* regions and directs the expression of immature virus particles
(Hansen and Barklis, 1995). Our wild type (wt) construct (pXMGPE) expresses all
three M-MuLV genes (*gag, pol* and *env*) and directs expression of mature M-
MuLV particles. Two other parental constructs were employed routinely. These
were PR-Cys-, which assembles immature virions in which all five *gag* cysteine
codons were converted to serines (Figure 1; Hansen and Barklis, 1995), and
MA+CACys-, a PR+ (pXMGPE) variant in which the MA and CA cysteine
codons (Figure 1) have been converted to serines. Other pertinent mutations
were cysteine creation mutations near the N-terminus of the CA domain
(G220C), near the CA C-terminus (S475C), at the N/B determining region
(T323C), and a previously studied major homology region (MHR) double
mutation E370C/K373C (Hansen and Barklis, 1995; see Figure 1). The double
mutation was expressed on a PR-Cys- background to yield PR-Cys-
E370C/K373C, while the G220C, S475C and T323C mutations were expressed in
both PR-Cys- and MA+CACys- (but PR+) contexts, giving PR-Cys-G220C, PR-

Cys-S475C, PR-Cys-T323C, MA+CACys-G220C, MA+CACys-S475C, and MA+CACys-T323C.

For assembly analysis, cell lysate and pelleted media supernatant samples were collected three days post-transfection, and *gag*-derived proteins were separated by SDS-PAGE and detected by immunoblotting (see Materials and Methods). As can be seen from Figure 2A, all PR- constructs were released efficiently from cells as indicated by comparison of Pr65^{Gag} levels in the viral supernatants (lanes A-F) versus cell lysates (lanes H-M). The PR+ constructs also directed the efficient release of Gag proteins from cells. As expected (Figure 2B), wt, MA+CACys-S475C, MA+CACys-T323C, and MA+CACys-G220C Pr65^{Gag} and CA proteins were detected readily in media samples (lanes A-D), while only Pr65^{Gag} proteins were observed in cell samples (lanes F-I). Note that for the MA+CACys-S475C mutant, a *gag*-derived 40 kDa media band was observed (lane B) about 80% of the time with the anti-CA antibodies, Hy187 and the polyclonal anti-CA. Since this band was not observed with the anti-p12 antibody Hy548 in virus particles (data not shown) it may represent an incomplete CA-NC cleavage product. That media supernatant Gag proteins were particle-associated was suggested by the fact that they sedimented through sucrose cushions at clearing rates consistent with sedimentation coefficients of >120S and with densities of 1.15-1.18 g/ml (Hansen and Barklis, 1995; data not shown), as expected for virus particles.

Crosslink analysis of immature virus particles.

To test the proximity of CA domain cysteine residues in Pr^{Gag} proteins in immature M-MuLV virus particles we utilized our established crosslinking protocol (Hansen and Barklis, 1995; McDermott *et al.*, 1996). Briefly, unique cysteine residues on adjacent molecules will not form covalent crosslinks if the cysteine residues are separated by too great a distance, when treated with membrane-permeable, cysteine-specific crosslinking agents such as molecular iodine (maximum crosslink distance approximately 5 Å) or BMH (maximum crosslink distance approximately 15 Å). In contrast, adjacent molecules with cysteines in close proximity may form covalent crosslinks that can be detected after gel electrophoresis and immunoblotting. As shown in Figure 3, when PR-virions lacking both *pol* and *env* gene products, were mock-treated, the expected Pr65^{Gag} band was detected with anti-p12 antibodies (lane F). However a parallel crosslinked sample (lane E) gave the 65 kDa band plus a 130 kDa dimer product, consistent with previous observations (Hansen and Barklis, 1995), demonstrating Pr65^{Gag} crosslinking via NC cysteine residues. As expected, crosslink treatment of the PR-Cys- mutant particles showed only a trace of a band at 130 kDa (lane C; discussed below), consistent with the specificity of BMH for preferential crosslinking at cysteine residues. Results from crosslink treatments of PR-Cys-E370C/K373C particles also agreed with previous results, showing no dimer band, but instead a crosslink product at about 190 kDa (Figure 3, lane A). Neither the N-terminal PR-Cys-G220C nor the PR-Cys-T323C mutant particles gave Pr65^{Gag} dimer bands upon crosslinking (lanes I and K), but each showed crosslink bands at 190 kDa, similar to that seen for PR-Cys-E370C/K373C. In contrast,

crosslinking of PR-Cys-S475C particles, which carry single Pr65^{Gag} cysteines near the C-terminus of the CA domain gave a band which comigrates with Pr65^{Gag} dimers (lane G), suggesting that C-terminal CA domains are in close proximity in immature M-MuLV virions.

Although the results shown in Figure 3 were observed in almost all experiments, it is important to point out that experiment-to-experiment variations were observed. While PR- showed Pr65^{Gag} crosslink dimers in 18 of 18 experiments with three different antibodies (Table 1), it also showed evidence of a 190 kDa band at least once. As noted above, a faint dimer band was seen on BMH treatment of PR-Cys- particles (Figure 3, lane C). Faint dimers were observed in 36% of our experiments using two different antibodies: we assume that such dimers resulted from maleimide side reaction with histidines and primary amine groups, which can occur at low levels at pH 7.4-7.5 (Smyth, Blumenfeld, and Konigsberg, 1964). We also observed some crosslinking variability with PR-Cys-G220C, PR-Cys-T323C, PR-Cys-E370C/K373C, and PR-Cys-S475C particles (Table 1). However, the general conclusions of Figure 3 were supported in multiple experiments. Specifically, the PR-Cys-S475C mutant proteins showed 130 kDa bands 91% of the time, while the PR-Cys-G220C, PR-Cys-T323C, and PR-Cys-E370C/K373C mutants gave 190 kDa bands 93%, 82%, and 100% of the time, respectively. A crosslink band this size could represent a Gag protein crosslinked to a cellular protein or, alternatively, a crosslinked Gag trimer which seems unlikely since the band is observed with mutant Gag proteins containing only one cysteine residue.

Crosslink analysis of mature virus particles.

For crosslinking studies on mature virus particles, we employed a parental MA+CA cysteine to serine construct (MA+CACys-), which does not block proteolytic processing. The experimental methods used for crosslink analysis of mature virus particles followed those used for the immature virus particles (see above). As shown in Figure 4, proteins from MA+CACys- particles showed only Pr65^{Gag} and CA bands but no crosslinked bands (Figure 4, lanes A, B) as expected. With mock-treated MA+CACys-G220C a 58 kDa band was observed in addition to the Pr65^{Gag} and CA bands (Figure 4, lane C). Since this band also was detected with our anti-p12 antibody (Hy548; lane E) it is possible that the band represents a processing intermediate, and may comprise the MA-p12-CA domains. Notwithstanding the putative processing intermediate, examination of the crosslinked G220C sample showed an additional 60 kDa band (Figure 4, lane D), consistent with the creation of a CA-CA crosslink dimer. As would be expected for a strictly CA-derived moiety, the 60 kDa band was not detected with the anti-p12 antibody (lane F), supporting the notion that the band corresponds to a CA-CA dimer. In the case of the MA+CACys-S475C mutant, we faintly observed the uncrosslinked 40 kDa band (Figure 4, lane G) which appears to correspond to a CA-NC processing intermediate (as in Figure 2). In lane H, an additional crosslink band at 62 kDa was observed. This 62 kDa band was not detected with the anti-p12 Hy548 antibody (data not shown), suggesting that it also represents a CA-CA dimer, which migrates slightly differently than the apparent G220C CA-CA dimer, possibly because of its different crosslink site. While the MA+CACys-G220C and S475C variants showed putative CA crosslink

dimer bands, the MA+CACys-T323C crosslink pattern was quite different. As shown (Figure 4, lane J) a faint 60 kDa dimer may be evident, but the prominent crosslink bands migrated at 75 and 80 kDa (arrows).

As with the PR- crosslinking results (Figure 3; Table 1), although the crosslink results with mature virus particles were generally reproducible, some variability was observed (Table 2). Nevertheless, with the MA+CACys-G220C mutant, the 60 kDa putative dimer band was observed nearly 90% of the time. While less consistent, the putative MA+CACys-S475C CA dimers were observed nearly two-thirds of the time, and MA+CACys-T323C 75-80 kDa bands were seen in three quarters of the experiments. In contrast, with the wt construct, 60, 62, 75, and 80 kDa bands were observed less than 10% of the time (Table 2).

Both the 75 and 80 kDa crosslink bands were observed after iodine-mediated crosslinking (data not shown), supporting the notion that interacting cysteines were close enough to form disulfides. During follow-up studies with reducible cysteine-specific crosslinking agents, we found that formation of reducible Gag protein crosslinks could be induced simply by particle treatment with DMSO, which often is used to solubilize membrane permeant crosslinking agents. The crosslinking potential of DMSO has been reported previously (Tam *et al.*, 1991), and is consistent with reports of its oxidizing capabilities. To ascertain DMSO crosslink effects on mature virions, particles were treated with DMSO and then either reduced (Figure 5, lanes A to D), or not reduced (lanes E to H), fractionated by SDS-PAGE and immunoblotted to detect Gag proteins. For wt M-MuLV mature virus particles, no crosslinking was detected with DMSO

(compare Figure 5, lanes A and E), as expected from BMH results (Figure 4) and indicating that the natural capsid cysteine residues (C270; Figure 1) are not in close proximity on neighbor molecules. In contrast to the wt results, but in concert with comparable BMH results (Figure 4), MA+CACys-G220C proteins showed the putative CA dimer band at 60 kDa (Figure 5, lane F) and putative 62 kDa MA+CACys-S475C dimer bands were also observed (Figure 5, lane H). Furthermore, one CA crosslink band was present at 75-80 kDa in the treated MA+CACys-T323C mutant lane (Figure 5, lane G), consistent with the appearance of a single 80 kDa crosslink product, or the comigration of the 75 and 80 kDa species seen in Figure 4. From these data it appears that MA+CACys-G220C, T323C and S475C crosslink partners are within 5 Å of each other. Since BMH samples were run under reducing conditions (Figure 3 and 4), it also would appear that BMH either reacts with cysteines more quickly than its solvent, DMSO, or that DMSO-induced disulfides can exchange with BMH to form non-reducible crosslinks.

Based on the size of the MA+CACys-T323C crosslink bands and amounts of Pr65^{Gag} present in virus particle preparations, it was possible that the 75-80 kDa bands could represent fast-migrating crosslinked Pr65^{Gag}-CA species. To test this hypothesis, MA+CACys-T323C controls (Figure 6, lanes B-D, F-G) and BMH crosslink samples (Figure 6, lanes A, E) were electrophoresed and blotted in parallel with anti-CA antibody (Hy187, lanes A-D) or with anti-p12 antibody (Hy548, lanes E-H). The 75 and 80 kDa bands were detected with the anti-CA but not anti-p12 antibody (Figure 6, lane A versus E) indicating that the 75 and 80 kDa bands did not derive from Pr65^{Gag} crosslinks. We also performed parallel

crosslink immunoblots with anti-actin, anti-tubulin, and anti-PDI (protein disulfide isomerase) antibodies. These experiments were performed because HIV-1 Gag proteins have been shown to interact with cytoskeletal proteins (Ott *et al.*, 1996; Perrin-Tricaud *et al*, 1999; Rey *et al*, 1996; Wilk *et al*, 1999) while PDI, which catalyzes protein disulfide formation and exchange, could be postulated to associate with cysteine residues on mutant proteins. Nevertheless, actin, tubulin and PDI antibodies did not bind to the 75-80 kDa MA+CACys-T323C crosslink bands (data not shown), suggesting the existence of alternate crosslink partners in these virions.

DISCUSSION

In this paper, we describe the crosslink analysis of M-MuLV capsid proteins in mature and immature virus particles. These experiments were undertaken to complement structural models derived from *in vitro* studies. Although previous crosslinking studies have been employed successfully to examine viral protein interactions (Pepinsky, 1983; Pepinsky *et al.*, 1980; Pinter and Fleissner, 1979), we chose the cysteine crosslinking approach to avoid complications in interpretation arising from using amine-specific crosslinking agents which can react with a large number of residues present in M-MuLV Gag. Nevertheless, as with other crosslinking methods, the composition of crosslink bands has been inferred from gel mobilities, and interpretations must be considered with this caveat in mind. Because our approach requires making mutations on PR-Cys- and MA+CACys- parental constructs, some results may not be pertinent to wt constructs. Additionally, the PR- construct does not

express *pol* and *env* gene products, an alteration which may disrupt the internal structure of immature virions. Since we have not performed electron microscope analyses of virus particles produced from transiently transfected cells, we do not know whether some mutants might show morphologies which are visibly different from their wt counterparts. However, the PR-Cys- and MA+CACys-Gag proteins directed the efficient assembly of particles with apparently normal masses and densities (Figure 2; Hansen and Barklis, 1995; data not shown), and the approach has been used previously to examine retrovirus particle structures (Hansen and Barklis, 1995; McDermott *et al.*, 1996). Thus, we believe our results are relevant to the analysis of immature and mature M-MuLV particles.

Our studies show that particle-associated wt PR- Gag proteins can be crosslinked to form dimers (Figure 3). This finding is consistent with our previous observations that Pr^{Gag} proteins in immature virions crosslink via the NC zinc finger cysteines (Hansen and Barklis, 1995) and agrees with the previous results showing that NC cysteines are targets for thiol reactive compounds (Ott *et al.*, 1998; Rein *et al.*, 1996; Rein *et al.*, 1997). We also observed that PR-Cys-S475C proteins could be crosslinked to form dimers, suggesting that CA C-terminal domains are closely associated in immature virions. In contrast, $Pr65^{Gag}$ dimers were not observed with the PR-Cys-G220C, T323C, or E370C/K373C mutants (Figure 3). Instead crosslink bands from these mutants migrated with apparent mobilities of 185-190 kDa on 7.5% gels, similar to our previous observations for the PR-Cys-E370C/K373C mutant (Hansen and Barklis, 1995), as well as for the other major homology region mutants PR-Cys-L369C and PR-Cys-E370C (Hansen and Barklis, 1995). The observed 190 kDa band is approximately the size

of three covalently linked Pr^{Gag} monomers, but is seen with mutants containing only one cysteine residue. This consideration is consistent with the crosslinking partner being a cellular protein or with anomalous crosslinking of Pr^{Gag} monomers to form trimers. We are attempting to generate a sufficient quantity of the 190 kDa band to characterize this complex further.

Relative to immature virions, CA proteins in mature M-MuLV particles showed different crosslink reactivities. While the natural CA C270 residues did not mediate dimer formation in mature virions (Figure 4), BMH (Figure 4), iodine (data not shown), and DMSO (Figure 5) treatments both yielded MA+CACys-G220C and S475C 60-62 kDa crosslink bands, which appear to correspond to CA-CA dimers. Assuming that the 60-62 kDa bands represent *bona fide* CA dimers, our results suggest that during maturation, capsid protein interactions reorganize to permit dimer formation at both N- and C- termini. This is consistent with the maturational conformation model proposed for HIV-1 in which constrained Pr^{Gag} CA N-termini refold during proteolytic processing (Schwedler *et al.*, 1998) and previous crosslinking results in M-MuLV (Pepinsky, 1983) and in avian sarcoma and leukemia viruses (Pepinsky *et al.*, 1980). Structural studies on HIV-1 and EIAV have shown that these capsid proteins crystallize as dimers (Gamble *et al.*, 1997; Jin *et al.*, 1999; Momany *et al.*, 1996) with higher order interactions forming the super-molecular particle structure. Our data supports the idea that multiple and distant regions of the M-MuLV capsid domain participate in close interprotein contacts in both the immature and mature virus particle structures.

In contrast to results with MA+CACys-G220C and S475C, the 75-80 kDa T323C crosslink products do not appear to be simple CA dimers. Our evidence suggests that the 75-80 kDa bands contain a CA moiety, but do not appear to be composed of p12-containing Pr^{Gag} molecules (Figure 6). Conceivably, these bands could represent anomalously migrating CA trimer bands, but crosslink trimers are unexpected for monomers which carry only single cysteines. Based on parallel immunoblot analysis, the MA+CACys-T323C crosslink partners do not appear to be actin, tubulin, or PDI. Furthermore the 75-80 kDa products would seem too small to include *env* gene product, gp70, or viral reverse transcriptase proteins. It is intriguing to speculate whether the MA+CACys-T323C crosslink partner could be involved in the *Fv1* host-range restriction phenomenon. In this regard, it is of interest that the predicted size of the *Fv1* N/B-related murine host factor is about 52 kDa, and appears to correspond to a murine homolog of the *gag* gene of the human endogenous retrovirus, HERV-L (Best *et al.*, 1996). However, while the 75-80 kDa MA+CACys-T323C crosslink products suggest that crosslink partners are present at reasonably high levels in virus particles produced from African green monkey (COS) cells, the Fv1 gene product appears to exert its effects in target murine cells, despite very low levels of expression (Best *et al.*, 1996). Consequently, for the present it is not clear how or if the MA+CACys-T323C crosslink partners relate to *Fv1* restriction.

MATERIALS AND METHODS

Constructs

The M-MuLV transient expression vector was derived from pXM (Yang *et al.*, 1986) which uses a simian virus 40 (SV40) origin of replication and adenovirus major late promoter. Construction and sequencing of mutant vectors followed standard protocols (Maniatis, Fritsch, and Sambrook, 1982). M-MuLV was modified by the addition of 5' EcoRI and EcoRV sites (GAATTCGATATCAAGCTT) at the HindIII (M-MuLV nt 563; (Shinnick, Lerner, and Sutcliffe, 1981)) of pMov-Psi (Mann, Mulligan, and Baltimore, 1983) and used to create pXMGPE which spans M-MuLV nt 566 to 7846 and expresses M-MuLV *gag*, *pol* and *env* gene products. The protease deficient (PR-) transient expression vector pXM2453 encodes M-MuLV viral nt 566-2453 and has a termination signal downstream of nt 2453 and thus lacks both the *pol* and *env* genes. Downstream junction sequences for both constructs are as follows: pXMGPE, 5' TTTGGCAAGCTAGA 3', in which nt 7840 from the 3' untranslated sequence is underlined; pXM2453, 5' GGT AAG GTC ACC GCG GAT CCC CCT **TAA** GTT AAC TTA AGG GCT GCA GGA ATT C 3', in which nt 2453 is underlined and the termination codon in bold face. As previously described (Hansen and Barklis, 1995), pXMGPE and pXM2453 were used as PR+ and PR- templates for the introduction of cysteine-to-serine point mutations. All five cysteine-to-serine point mutations, including C504S/C507S, previously described (Gorelick *et al.*, 1988), were cloned into the two parental vectors to yield pXM2453 all cysteine to serine (pXM2453Cys- or PR-Cys-) and pXMGPE all cysteine to serine (pXMGPECys- or wtCys-). The sequences of these cysteine

mutants, with the first nt of the sequence listed, the modified nucleotides in boldface, and the codons where an amino acid was changed underlined, are as follows: C39S (nt 723), 5' TGG GTT ACG TTC TCC TCT GCA 3'; C270S (nt 1419), 5' TGG GAC GAT AGT CAA CAG 3'; C504S/C507S (nt 2123), 5' GAC CAG TCT GCC TAC TCC AAA 3'; C517S (nt 2163), 5' AAA GAT TCT CCC AAG AAG CCT CGA 3' . In addition to the pXMGPE Cys- construct, we constructed a pXMGPE capsid cysteine to serine construct (pXMGPECACys- or CACys-) which contains only the C270S mutation. Additionally, a pXMGPE matrix plus capsid cysteine to serine construct (pXMGPEMA+CACys- or MA+CACys-) which contains the C39S and C270S mutations, also was made.

With regard to cysteine creation mutants, PR-Cys-E370C/K373C, has been described (Hansen and Barklis, 1995). It introduces two cysteine residues in the MHR region (E370C/K373C [nt 1722], 5' TTC CTA TGT CGA CTT TGC GAA GCC TAT CGC ACG TAC ACT CCA TAT 3'). In addition to this cysteine creation mutant, several others were made in the capsid domain: G220C (shown from nt 1266), 5' CCC CTC CGC GCA TGC GGA AAC GGA 3'; T323C (shown from nt 1566), 5' CGC CCA GAC AGG GAT TAC ACA TGT CCC AGG 3'; S475C (shown from nt 2031), 5' CAT AGA GAG ATG TGT AAG CTA TTG GCC 3'. These mutants were introduced into the parental constructs PR-Cys-, CACys-, and MA+CACys-.

Cell culture and viruses

COS1 and COS7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (FCS), penicillin, and streptomycin. Cells were split 1:4 from confluent plates onto 10-cm-diameter plates 24 h prior to transfections and were transfected with 16 _g plasmid DNA by the calcium phosphate precipitation method (Hansen and Barklis, 1995). Three days post-transfection, cell lysates and supernatants were collected as previously described (Hansen and Barklis, 1995). Cell-free supernatants from transfected cells were centrifuged through 4-ml 20% sucrose cushions at 83,000 X g (25,000 rpm in an SW28 rotor) for 2 h at 4° C. Virus pellets were resuspended in phosphate-buffered saline (PBS; 9.5 mM sodium potassium phosphate [pH 7.4], 137 mM NaCl, 2.7 mM KCl). Cells were rinsed twice with PBS, collected in IPB (20 mM Tris-hydrochloride [pH 7.5], 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 0.02% sodium azide), vortexed 1 min, and centrifuged (13,700 X g, 4°C, 10 min) to remove insoluble material. For protein gels, solubilized cellular lysates were mixed with equal volumes of 2X sample buffer (12.5 mM Tris-hydrochloride [pH 6.8], 2% SDS, 20% glycerol, 0.25% bromophenol blue) plus 5% β-mercaptoethanol (β-Me).

Sucrose density gradient fractionation of virus particles was performed as previously described (Hansen and Barklis, 1995; McDermott *et al.*, 1996). Briefly, extracellular virions were isolated by pelleting through sucrose cushions as described above, after which they were applied to sucrose gradients consisting of 1.1 ml layers of 20, 30, 40, and 50% sucrose which had been allowed to mix by

sitting for 1 h at 4°C. Gradients were centrifuged at 300,000 X g (50,000 rpm in an

SW50.1 rotor) overnight at 4°C and 400 µl fractions were collected from top to

bottom for analysis by immunoblotting as described below.

Protein and crosslink analysis

Protein samples were analyzed by SDS-polyacrylamide gel

electrophoresis (PAGE, (Laemmli, 1970)). Unless otherwise stated,

electrophoresis was performed using an 7.5% or 10% SDS-PAGE gels.

Electrophoresis ordinarily was performed under reducing conditions which was

achieved by the addition of an equal volume of 2X sample buffer plus 5% β-Me

(v/vFinal) to the protein samples. For non-reducing conditions β-Me was

omitted. After electrophoresis under standard conditions (Hansen *et al.*, 1993;

Hansen *et al.*, 1990; Hansen and Barklis, 1995; Jones *et al.*, 1990; Laemmli, 1970;

McDermott *et al.*, 1996) gels were electroblotted onto nitrocellulose filters and

proteins were detected using the western procedure previously described

(Hansen and Barklis, 1995). For size estimates, mobilities were plotted versus log

molecular weights of known size standards. For immunoblotting, primary

antibodies used were rat monoclonal antibody hybridoma cell supernatants anti-

M-MuLV p12^{Gag} (Hy548) and anti-M-MuLV p30^{Gag} (Hy187) used at 1:10 dilutions

(Chesebro *et al.*, 1983); polyclonal goat anti-M-MuLV p30^{Gag} (National Cancer

Institute), used at a 1:4,000 dilution; mouse monoclonal anti-actin antibody

(Boehringer Mannheim) used at 1:2,000 dilution; mouse monoclonal anti-protein

disulfide isomerase (PDI; Affinity Bioreagents, Inc.) used at a dilution of 1:2,000,

and mouse monoclonal anti-tubulin (Amersham) used at a dilution of 1:200. Secondary alkaline phosphatase-conjugated antibodies were goat anti-mouse immunoglobulin G (IgG; Promega) used at a 1:20,000 dilution and rabbit anti-goat IgG (Boehringer Mannheim), used at a 1:10,000 dilution.

Iodine crosslinking of cysteine residues was performed as previously described (Hansen *et al.*, 1993; Pakula and Simon, 1992) and analyzed by non-reducing SDS-PAGE. For cross-linking with *bis*-maleimido hexane (BMH; Pierce), BMH was prepared fresh as a 100 mM solution in dimethyl sulfoxide (DMSO, Mallinckrodt). Viral samples were split into equivalent 100-μl fractions, treated with BMH (1 μl of 100 mM BMH in DMSO) or mock treated (1 μl of DMSO), vortexed gently, and incubated at room temperature for 1 h. Reactions were terminated by the addition of 2X sample buffer plus β-Me to 5%, and the samples were boiled for 3-5 min. For DMSO oxidation treatments (Tam *et al.*, 1991), 1 μl of DMSO was added to each 100 μl tube, vortexed gently and incubated at room temperature. Reactions were terminated by the addition of 2X sample buffer without B-Me followed by sample boiling for 3-5 min.

ACKNOWLEDGEMENTS

REFERENCES

Alin, K., and 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.

Barbacid, M. J., Stephenson, J., and Aaronson, S. (1976). Gag gene of mammalian type-C RNA tumor viruses. *Nature* **262,** 554-559.

Barklis, E., McDermott, J., Wilkens, S., Fuller, S., and Thompson, D. (1998). Organization of HIV-1 capsid proteins on a lipid monolayer. *J. Biol. Chem.* **273**(13), 7177-7180.

Barklis, E., McDermott, J., Wilkens, S., Schabtach, E., Schmid, M. F., Fuller, S., Karanjia, S., Love, Z., Jones, R., Rui, Y., Zhao, X., and Thompson, D. (1997). Structural analysis of membrane-bound retrovirus capsid proteins. *EMBO J.* **16**(6), 1199-1213.

Best, S., Tissier, P. L., Towers, G., and Stoye, J. P. (1996). Positional cloning of the mouse retrovirus restriction gene *Fv1*. *Nature* **382**(6594), 826-829.

Bolognesi, R. D., Luftig, R., and Shaper, J. (1973). Localization of RNA tumor virus polypeptides. I. Isolation of further virus substructures. *Virology* **56,** 549-564.

Borsetti, A., Ohagen, A., and 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**(11), 9313-9317.

Brown, P. O. (1997). Integration. *In* "Retroviruses" (J. M. Coffin, S. H. Hughes, and H. E. Varmus, Eds.), pp. 161-203. Cold Spring Harbor Laboratory Press, Plainview.

Campbell, S., and Rein, A. (1999). In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J. Virol.* **73**(3), 2270-2279.

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**(10), 6487-6497.

Campbell, S., and Vogt, V. M. (1997). In vitro assembly of virus-like particles with Rous sarcoma virus Gag deletion mutants: identification of the p10 domain as a morphological determinant in the formation of spherical particles. *J. Virol.* **71**(6), 4425-4435.

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 cirus. *Virology* **127,** 134-148.

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

Craven, R. C., and Parent, L. J. (1996). Dynamic interactions of the Gag polyprotein. *Curr. Top. Microbiol. Immunol.* **214,** 65-94.

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.

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

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

Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999). Assembly and analysis of conical models for the HIV-1 core. *Science* **283**(5398), 80-83.

Gorelick, R., Henderson, L., Hanser, J., and Rein, A. (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for a specific RNA recognition by a "zinc finger-like" protein sequence. *Proc. Natl. Acad. Sci. USA* **85**, 8420-8424.

Hansen, M., Jelinek, L., Jones, R. S., Stegeman-Olsen, J., and Barklis, E. (1993). Assembly and composition of intracellular particles formed by Moloney murine leukemia virus. *J. Virol.* **67**, 5163-5174.

Hansen, M., Jelinek, L., Whiting, S., and Barklis, E. (1990). Transport and assembly of Gag proteins into Moloney murine leukemia virus particles. *J. Virol.* **64**, 5306-5316.

Hansen, M. S. T., and Barklis, E. (1995). Structural interaction between retroviral Gag proteins examined by cysteine cross-linking. *J. Virol.* **69**(2), 1150-1159.

Jin, Z., Jin, L., Peterson, D. L., and Lawson, C. L. (1999). Model for lentivirus capsid core assembly based on crystal dimers of EIAV p26. *J. Mol. Biol.* **286**(1), 83-93.

Jones, T. A., Blaug, G., Hansen, M., and Barklis, E. (1990). Assembly of Gag-b-galactosidase proteins into retrovirus particles. *J. Virol.* **64**, 2265-2279.

Kovari, L. C., Momany, C. A., Miyagi, F., Lee, S., Campbell, S., Vuong, B., Vogt, V. M., and Rossmann, M. G. (1997). Crystals of Rous sarcoma virus capsid protein show a helical arrangement of protein subunits. *Virology* **238**(1), 79-84.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lobel, L. I., and Goff, S. P. (1984). Construction of mutants of Moloney murine leukemia virus by suppressor-linker insertional mutagenesis: Positions of viable insertion mutations. *Proc. Natl. Acad. Sci. USA* **81**, 4149-4153.

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

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). "Molecular cloning: a laboratory manual." Cold Spring Harbor, Cold Spring Harbor, NY.

Mann, R., Mulligan, R., and Baltimore, D. (1983). Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**, 153-159.

McDermott, J., Farrell, L., Ross, R., and Barklis, E. (1996). Structural analysis of human immunodeficiency virus type 1 Gag protein interaction using cysteine-specific reagents. *J. Virol.* **70**(8), 5106-5114.

Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Ehrlich, L. S., Summers, M. F., Carter, C., and Rossman, M. G. (1996). Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.* **3**(9), 763-770.

Ott, D. E., Coren, L. V., Kane, B. P., Busch, L. K., Johnson, D. G., Sowder, R. C. n., Chertova, E. N., Arthur, L. O., and Henderson, L. E. (1996). Cytoskeletal proteins inside human immunodeficiency virus type 1 virions. *J. Virol.* **70**(11), 7734-7743.

Ott, D. E., Hewes, S. M., Alvord, W. G., Henderson, L. E., and Arthur, L. O. (1998). Inhibition of Friend virus replication by a compound that reacts with the nucleocapsid zinc finger: anti-retroviral effect demonstrated in vivo. *Virology* **243**(2), 283-292.

Pakula, A. A., and Simon, M. I. (1992). Determination of transmembrane structure by disufide cross-linking: the Escherichia coli Tar receptor. *Proc. Natl. Acad. Sci. USA* **89**, 4144-4148.

Pepinsky, R. (1983). Localization of lipid-protein and protein-protein interactions within the murine retrovirus gag precursor by a novel peptide-mapping technique. *J. Biol. Chem.* **258**(18), 11229-11235.

Pepinsky, R. B., Cappiello, D., Wilkowski, C., and Vogt, V. M. (1980). Chemical crosslinking of proteins in avian sarcoma and leukemia viruses. *Virology* **102**, 205-210.

Perrin-Tricaud, C., Davoust, J., and Jones, I. M. (1999). Tagging the human immunodeficiency virus gag protein with green fluorescent protein. Minimal evidence for colocalisation with actin. *Virology* **255**(1), 20-25.

Pinter, A., and Fleissner, E. (1979). Structural studies of retroviruses: characterization of oligomeric complexes of murine and feline leukemia virus envelope and core components formed upon cross-linking. *J. Virol.* **30**(1), 157-165.

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

Rein, A., Ott, D. E., Mirro, J., Arthur, L. O., Rice, W., and Henderson, L. E. (1996). Inactivation of murine leukemia virus by compounds that react with the zinc finger in the viral nucleocapsid protein. *J. Virol.* **70**(8), 4966-4972.

Rein, A., Ott, D. E., Mirro, J., Arthur, L. O., Rice, W. G., and Henderson, L. E. (1997). Suppression of retrovirial replication: inactivation of murine leukemia virus by compounds reacting with the zinc finger in the viral nucleocapsid protein. *Leukemia.* **11**(3), 106-108.

Rey, O., Canon, J., and Krogstad, P. (1996). HIV-1 Gag protein associates with F-actin present in microfilaments. *Virology* **220**, 530-534.

Scarlata, S., Ehrlich, L. S., and Carter, C. A. (1998). Membrane-induced alterations in HIV-1 Gag and matrix protein-protein interactions. *J. Mol. Biol.* **277**(2), 161-169.

Schwedler, U. K. v., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H., Davis, D. R., and Sundquist, W. I. (1998). Proteolytic refolding of the HIV-1

capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**(6), 1555-1568.

Shinnick, T., Lerner, R., and Sutcliffe, J. (1981). Nucleotide sequence of Moloney murine leukemia virus. *Nature* **293**, 543-548.

Smyth, D. G., Blumenfeld, O. O., and Konigsberg, W. (1964). Reactions of N-ethylmaleimide with peptides and amino acids. *Biochem. J.* **91**, 589-595.

Stephenson, J., Tronick, S., and Aaronson, S. (1975). Murine leukemia virus mutants with temperature-sensitive defects in precursor processing cleavage. *Cell* **6**, 543-548.

Strambio-de-Castillia, C., and 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.

Swanstrom, R., and Wills, J. W. (1997). Synthesis, Assembly, and Processing of Viral Proteins. *In* "Retroviruses" (J. M. Coffin, S. H. Hughes, and H. E. Varmus, Eds.), pp. 263-334. Cold Spring Harbor Laboratory Press, Plainview.

Tam, J. P., Wu, C.-R., Liu, W., and Zhang, J.-W. (1991). Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. *J. Am. Chem. Soc.* **113**, 6657-6662.

Wang, C.-T., Stegeman-Olsen, J., Zhang, Y., and Barklis, E. (1994). Assembly of HIV GAG-ß-Galactosidase fusion proteins into virus particles. *J. Virol.* **200**, 524-534.

Weiss, R., Teich, N., Varmus, H., and Coffin, J. (1984). "RNA tumor viruses, 2nd ed." Cold Spring Harbor, Cold Spring Harbor, NY.

Weldon, R. A. J., and 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.

Wilk, T., Gowen, B., and Fuller, S. D. (1999). Actin associates with the nucleocapsid domain of the human immunodeficiency virus Gag polyprotein. *J. Virol.* **73**(3), 1931-1940.

Wills, J. W., Cameron, C. E., Wilson, C. B., Xiang, Y., Bennet, R. P., and Leis, J. (1994). An assembly domain of the Rous sarcoma virus Gag protein required late in budding. *J. Virol.* **68**, 6605-6618.

Yang, Y. C., Ciarletta, A. B., Temple, P. A., Chung, M. P., Kovacic, S., Witek-Giannotti, J. S., Leary, A. C., Kriz, R., and Clark, S. C. (1986). Human IL-3 (multi-CSF): identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. *Cell* **47**, 3-10.

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

TABLES AND FIGURES

Table 1. Observed crosslink bands for PR- constructs.

construct	crosslink band	% time observed	total # gels	antibodies
PR-	PrGag dimer	100	18	187,548,CA
	PrGag tetramer	17	18	187,548,CA
	220 kd	6	18	548
	205 kd	6	18	548
	190 kd	6	18	548
PR-Cys-	PrGag dimer (faint)	36	14	548,CA
	190 kd	14	14	548
	160 kd	7	14	CA
PR-Cys-E370C/K373C	PrGag dimer	7	14	CA
	190 kd	93	14	548,CA
PR-Cys-G220C	PrGag dimer	9	11	548
	190 kd	82	11	187,548
PR-Cys-S478C	PrGag dimer	91	11	187,548
	190 kd	9	11	548
PR-Cys-T323C	PrGag dimer	0	8	548
	190 kd	100	8	548

Crosslinking reactions were performed as described in the Materials and Methods on immature virus particles produced on transfection of cells with the indicated constructs. Crosslink bands are designated as putative Pr^{Gag} dimers or tetramers, or as apparent molecular masses of crosslink species. The designation of "faint" for PR-Cys- Pr^{Gag} dimers indicates that bands were detectable, but at levels less than 10-25% of wt Pr^{Gag} dimers. The total number of gels signifies the number of independent crosslinking experiments; the percentage of time observed is based on the total number of gels; and the antibodies column indicates the different antibodies that have been used to detect a specific crosslink band.

Table 2. Observed crosslink bands for PR+ constructs.

Construct	crosslink band	% time observed	total # gels	antibodies
wt	PrGag dimer	40	25	187,548,CA
	60 kd	8	25	187
	62 kd	4	25	187
	37 kd	4	25	187
MA+CACys-G220C	PrGag dimer	15	27	187,548,CA
	60 kd	89	27	187,CA
	75 kd	4	27	187
	40 kd	4	27	187
MA+CACys-S478C	PrGag dimer	9	22	187,548
	62 kd	64	22	187
	60 kd	27	22	187
	45 kd	18	22	187
	75 kd	9	22	187
	40 kd	9	22	187
	37 kd	5	22	187
MA+CACys-T323C	PrGag dimer	4	25	187
	80 kd	76	25	187
	75 kd	76	25	187
	60 kd	24	25	187
	50 kd	4	25	187

Crosslinking reaction were performed as described in Materials and Methods on mature virus particles produced on transfection of cells with the indicated constructs. Crosslink bands are designated as putative PrGag dimers and as apparent molecular masses of crosslink species. The total number of gels signifies the number of independent crosslinking experiments; the percentage time observed is based on the total number of gels; and the antibodies column indicates the different antibodies that have been used to detect a specific crosslink band.
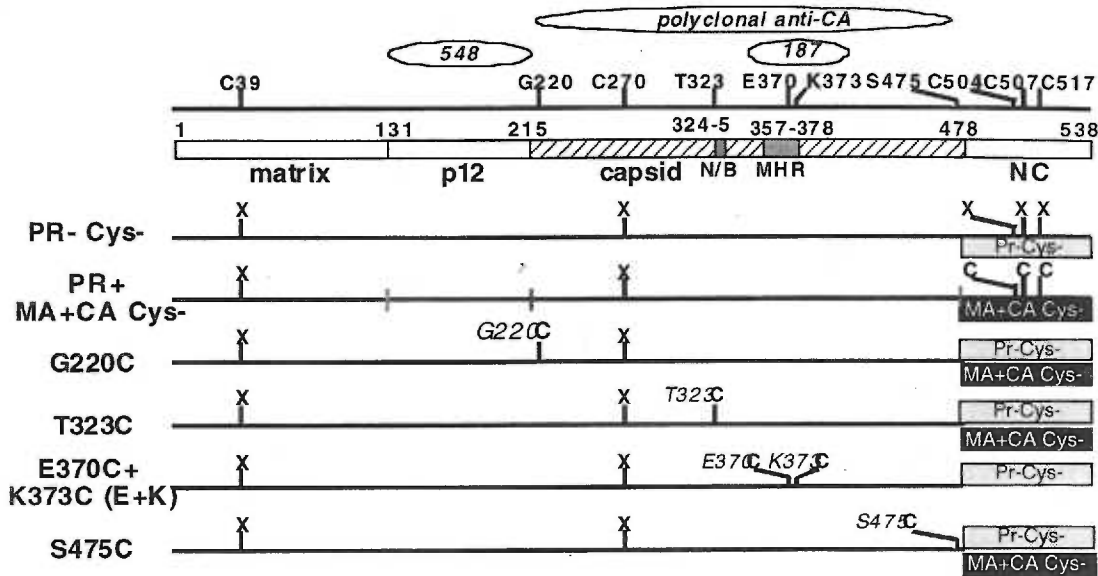
Figure 1. Schematic diagram of M-MuLV Pr65^{Gag} protein. The M-MuLV Pr65^{Gag} protein consists of four domains; matrix (MA), p12, capsid (CA, stippled lines), and nucleocapsid (NC), that are cleaved by the viral protease (PR) during the budding process. The precursor polypeptide is 538 codons in length and the juncture codons are indicated. The five natural cysteines in Pr65^{Gag} and the five introduced cysteine mutations are indicated on the top bar with their respective *gag* codon numbers. Within the capsid domain, the N/B tropism determination site is at codons 324 and 325 and the MHR includes codons 357 to 378. Antibody epitopes for hybridoma 548, which binds the p12 domain, hybridoma 187, which binds near the MHR, and the polyclonal anti-CA antibody, which binds to the capsid domain, are indicated at the top. Note that the parental immature construct is protease-minus (PR-), and that mutations were constructed on a protease-minus, cysteine-minus (PR-Cys-) parent (grey box). In contrast, the mature parent construct eliminated matrix and capsid domain cysteines (MA+CACys-) but retained all NC cysteines (black box). The parental immature and mature constructs as well as the individual cysteine mutations are represented schematically with an X denoting a naturally occurring cysteine that has been removed and a C (preceded by the mutation designation) representing a cysteine creation. Specific cysteine creation mutations use a letter-number-letter nomenclature, where the first letter indicates the wt residue, the number corresponds to the *gag* codon number, and the second letter indicates the newly created amino acid residue. Grey or black boxes associated with each construct denote whether the cysteine creation was expressed in a PR-Cys- or MA+CACys-context, or both.
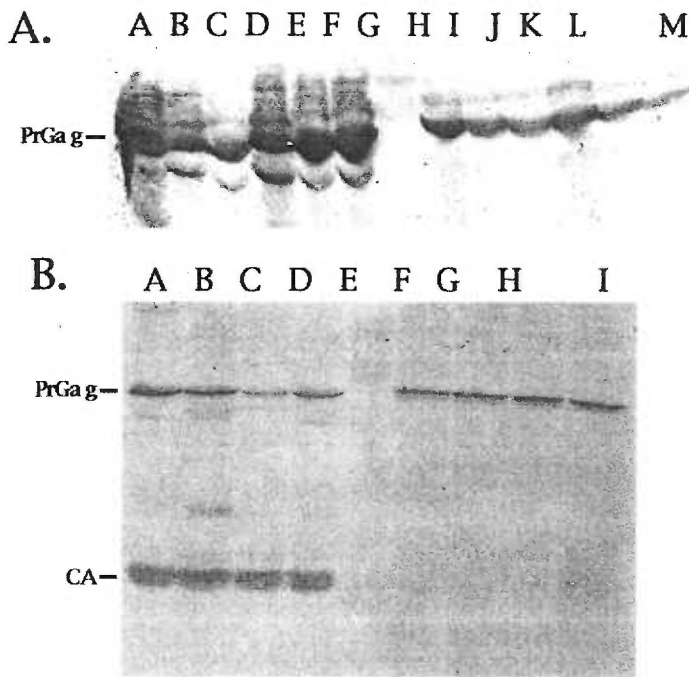
A. A B C D E F G H I J K L M

PrGag—

B. A B C D E F G H I

PrGag—

CA—

Figure 2. Expression and release of mutant Gag proteins. (Panel A) As described in the Materials and Methods, virus particle (lanes A-F) and cell lysate (lanes H-M) samples were collected from cells transfected with the following protease-minus (PR-) pXM2453 vectors: wt PR- (lanes A, H), PR-Cys- (B, I), PR-Cys-E370C/K373C (C,J), PR-Cys-G220C (D,K), PR-Cys-S475C (E,L), PR-Cys-T323C (F,M). Samples were subjected to SDS-PAGE and electroblotted, and Gag proteins were detected with the anti-p12 antibody Hy548. The protein size standard (faint) is in lane G, and the Pr65^{Gag} band is indicated. **(Panel B)** Virus particle (lanes A-D) and cell lysate (lanes F-I) samples were collected from cells transfected with the following pXMGPE (PR+) vectors: wt (lanes A, F), MA+CACys-S475C (B, G), MA+CACys-T323C (C, H), MA+CACys-G220C (D, I). Samples were fractionated and immunoblotted as above, except the anti-CA antibody Hy187 was employed. The protein size standard appears faintly in lane E, and Pr65^{Gag} and CA proteins are as indicated.

Figure 3. Crosslinking of M-MuLV protease-minus Gag proteins. Immature virus particles produced by constructs PR-Cys-E370C/K373C (lanes A, B), PR-Cys- (C, D), PR- (E,F), PR-Cys-S475C (G, H), PR-Cys-G220C (I, J), and PR-Cys-T323C (K, L) were mock treated (lanes B, D, F, H, J, and L) or treated with BMH (lanes A, C, E, G, I, and K), as described in Materials and Methods, to crosslink neighboring cysteine residues. Samples were subjected to SDS-PAGE and electroblotted, and Gag proteins were detected with the anti-p12 antibody Hy548. Pr65^{Gag} proteins and the putative Pr65^{Gag} homodimers are indicated. The arrow denotes the putative heterodimer crosslink band for lanes A, I, and K.

A B C D E F G H I J K L

dimer—

PrGag—
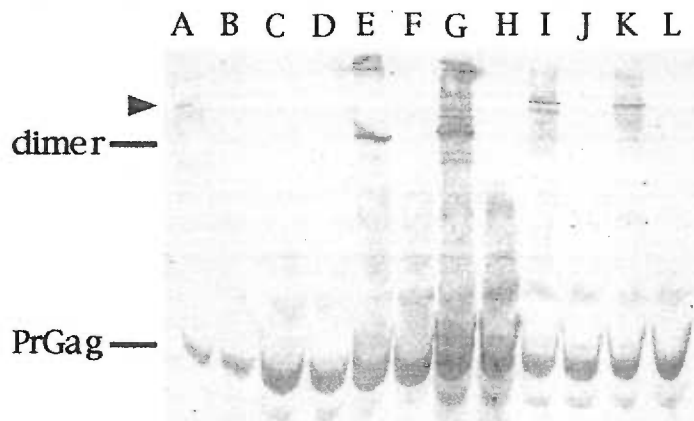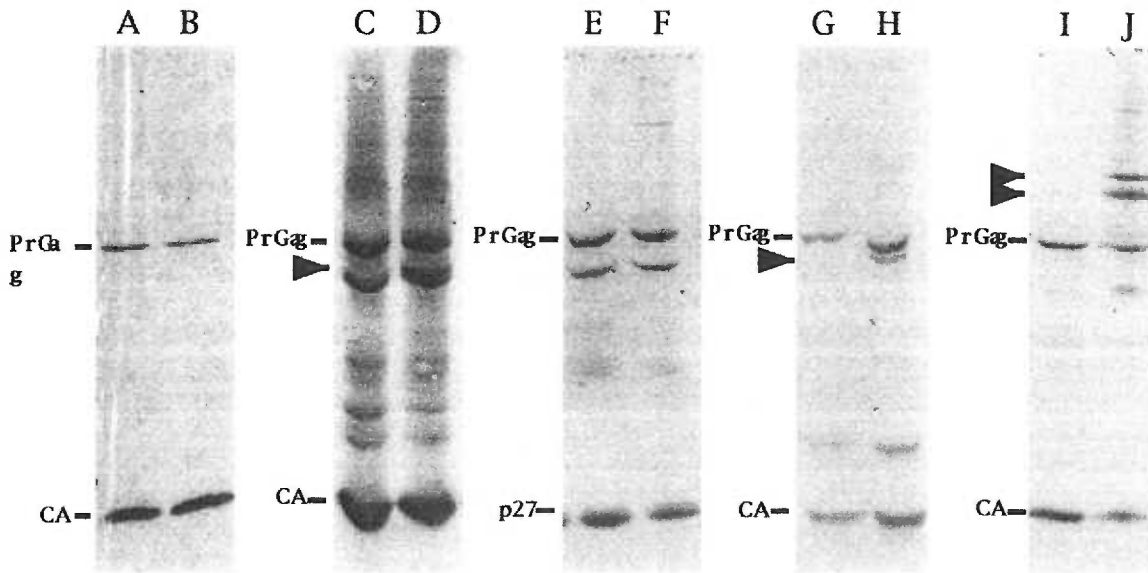
Figure 4. Crosslinking of mature M-MuLV Gag proteins. Virus particles from cells transfected with pXMGPE (wild type; A, B), MA+CACys-G220C (C-F), MA+CACys-S475C (G, H), and MA+CACys-T323C (I, J) were mock treated (lanes A, C, E, G, and I) or treated with BMH (lanes B, D, F, H, and J) as described in Materials and Methods. Protein samples were fractionated by SDS-PAGE and electroblotted then detected with the anti-CA hybridoma Hy187 (lanes A-D, G-J) or the anti-p12 hybridoma Hy548 (lanes E, F). Bands corresponding to Pr65Gag, CA, and the proteolytic MA+p12 fragment, p27, are indicated. Arrows in lanes D, H and J indicate crosslinked bands.

Figure 5. DMSO-induced crosslinking of mature virus particle proteins. Virus particles from cells transfected with pXMGPE (wild type; A, E), MA+CACys-G220C (B, F), MA+CACys-T323C (C, G), and MA+CACys-S475C (D, H) were treated with dimethyl sulfoxide (DMSO) as described in the Materials and Methods. Protein samples were separated on a reducing gel (lanes A-D) or a non-reducing gel (lanes E-H), after which electroblotted Gag proteins were detected using the anti-CA antibody Hy187. Bands representing the Pr65^{Gag} and mature CA proteins are indicated, as are the locations of the crosslink products for each mutant at the right.
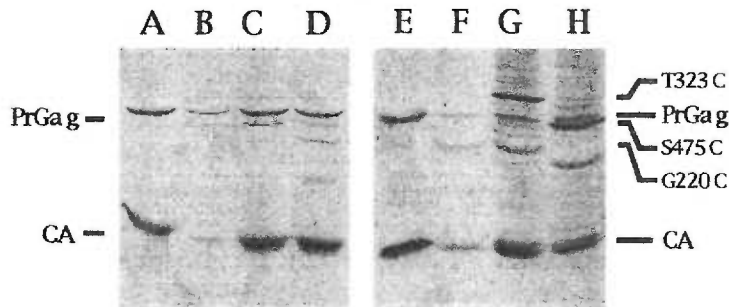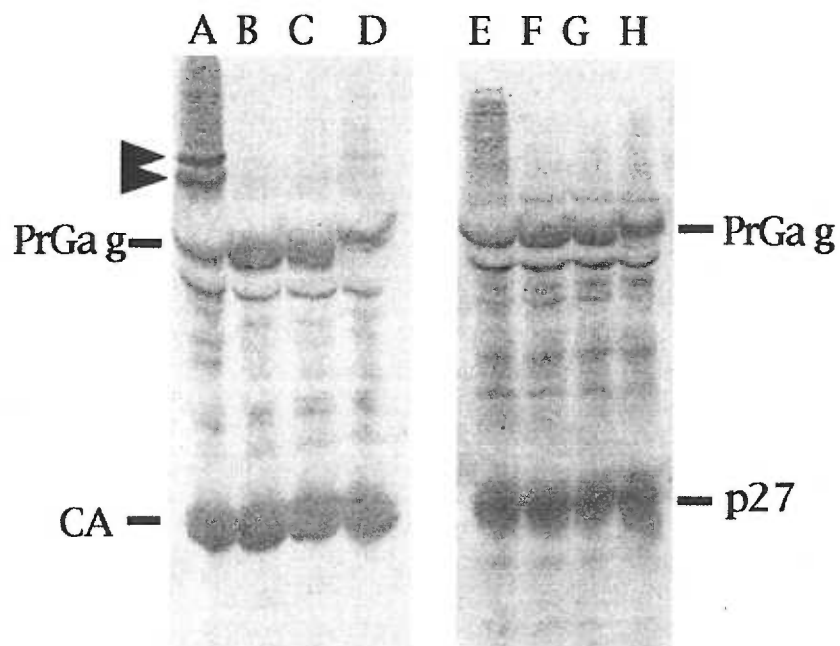
Figure 6. Composition of the MA+CACys-T323C crosslink band. Virus particles from cells transfected with MA+CACys-T323C were either BMH-treated (A, E), or mock treated (B-D, F-H), after which protein samples were fractionated on a reducing SDS-PAGE gel. After electrophoresis, samples were electroblotted and then immunoblotted using either the anti-CA antibody, Hy187 (lanes A-D) or the anti-p12 antibody, Hy548 (lanes E-H). Indicated are $Pr65^{Gag}$, CA and the proteolytic MA+p12 fragment, p27. Arrows indicate the 75 kDa and 80 kDa crosslink bands.

Chapter 3

STRUCTURAL ANALYSIS OF HIV-1 GAG PROTEIN INTERACTIONS

USING CYSTEINE-SPECIFIC REAGENTS

Jason McDermott*, Lori Farrell*, Robert Ross, and Eric Barklis[1]

Vollum Institute for Advanced Biomedical Research and Department of Molecular Microbiology and Immunology, Oregon Health Sciences University, Portland, Oregon, 97201-3098
Phone: (503)-494-8098   FAX: (503)-494-6862

* The two authors contributed equally to this research

[1]Corresponding author

ABSTRACT

We have examined structural interactions of Gag proteins in human immunodeficiency virus type 1 (HIV-1) particles by utilizing cysteine mutagenesis and cysteine-specific modifying reagents. In immature protease-minus (PR-), but otherwise wild-type (wt) particles, precursor Pr55^{Gag} proteins did not form intermolecular cystines naturally, but could be crosslinked at cysteines, and crosslinking appeared to occur across nucleocapsid (NC) domains. Capsid (CA) proteins in wt mature viruses possess cysteines near their carboxy-termini at *gag* codons 330 and 350, but these residues are not involved in natural covalent intermolecular bonds, nor can they be intermolecularly crosslinked using the membrane-permeable crosslinker bis-maleimido hexane (BMH). The cysteine at *gag* codon 350 (C350) is highly reactive to thiol-specific modifying reagents, while the one at codon 330 (C330) appears considerably less reactive, even in the presence of ionic detergent. These results suggest that the HIV-1 CA C-terminus forms an unusually stable conformation. Mutagenesis of C350 to a serine residue in the mutant C350S virtually eliminated particle assembly, attesting to the importance of this region. We also examined a C330S mutant, as well as mutants in which cysteines were created midway through the capsid domain, or in the C-terminal section of the major homology region (MHR). All such mutants appeared wt on the basis of biochemical assays, but showed greatly reduced infectivities, indicative of a post-assembly, post-processing replicative block. Interestingly, capsid proteins of mature MHR mutant particles could be cysteine-crosslinked, implying either that these mutations permit crosslinking of the native C-terminal CA cysteines, or that major homology regions on neighbor capsid proteins are in close proximity in mature virion

INTRODUCTION

The human immunodeficiency virus (HIV-1) Gag proteins are encoded by the viral *gag* gene and are synthesized as a polyprotein precursor, Pr55^{Gag}. Expression of Pr55^{Gag} in eukaryotic cells is necessary and sufficient for the formation of virus-like particles (10, 14, 22, 34, 35). Immature particles are approximately 125 nm in diameter, are surrounded by a lipid bilayer, and show an electron-dense layer juxtaposed to the bilayer (35). In cells infected with wild-type (wt) HIV-1, the Pr55^{Gag} precursor ordinarily is cleaved by the virally encoded protease (PR) to give an occasional processing intermediate (p41), and the mature viral Gag proteins, matrix (MA), capsid (CA), nucleocapsid (NC), and p6 (4, 19, 25, 29). Coincident with processing, HIV particles take on a new morphology: the particles become much more sensitive to disruption by non-ionic detergents (36, 39), and they acquire an electron dense conical or cylindrical core (9).

Most of the knowledge acquired about Gag protein function has come from molecular genetic analysis of mutants. Mutations of MA are characterized by a reduction in membrane binding (33, 39), impaired assembly (8, 33) or assembly on intracellular membranes (8, 17), reduced incorporation of viral envelope products (Env) into virus particles (8, 43), and/or impaired Env function (39, 42). The capsid domain contains a region of Gag protein primary sequence homology shared among almost all primate and avian retroviruses, called the major homology region (MHR,[27, 37, 41]). Mutations in the MHR can impair assembly (27), or can cause assembled particles to be non-infectious (27, 37). Several other mutations in CA block assembly (18, 38), or block virus infection at an undefined late stage of the virus life cycle (39). Mutations in NC

also can cause a reduction in assembly. In addition, many NC mutations impair viral ability to encapsidate viral RNA (1, 11), supporting the notion that the cys-his fingers in NC are required for specific encapsidation of viral RNA. Finally, mutations in p6 cause a reduction in virus budding under some conditions (12).

Although standard molecular genetic approaches have allowed identification of the functions of different Gag proteins, definition of subdomain interactions is in its infancy, and very little is known about how Gag proteins associate with each other to form a virus particle. To begin to identify important Gag-Gag protein contacts in virus particles, we have commenced a series of cysteine mutagenesis experiments aimed at assessing the importance of the Gag protein cysteine residues in the virus life cycle, and at identifying important Gag protein associations by cysteine-specific cross-linking methods (16). Our studies have focused on Pr55^{Gag} in immature virions, and on the CA domain of mature virus particles, as it appears to make essential interprotein contacts in infectious virions (9). In this study, we have examined the status of wild-type (wt) HIV-1 Gag protein cysteines, and of cysteines created in the HIV-1 capsid domain. Our results indicate that intermolecular cystine formation is not essential to particle assembly, but that Pr55^{Gag} proteins can be linked, apparently across their nucleocapsid cysteines. Biochemical and genetic analysis showed that the two HIV-1 CA cysteines at *gag* codons 330 and 350 did not form intramolecular cystines on the majority of capsid proteins, and that mutagenesis of C350 had profound effects on virus particle assembly. Additionally, creation of cysteines in the C-terminal portion of the HIV-1 MHR increased the efficiency of CA-CA crosslink formation, and yielded virus particles that had wt characteristics but were poorly infectious. Taken together, our results indicate that the C-terminal end of HIV-1 CA is extremely important to virus particle assembly, and that

MHR residues appear to mediate Gag protein associations, at least in the mature capsid domains.

MATERIALS AND METHODS

Recombinant Plasmids. The parental wt plasmid used in this study is HIVgpt which was derived from the HXBc2 HIV provirus (32). The plasmid has an active SV40 ori sequence which allows plasmid replication in transiently transfected COS7 cells (30). For infectivity assays, using expression of the bacterial *gpt* gene as a selectable marker, the SV-A-MLVenv plasmid was used as previously described (30). Both the HIVgpt and the SV-A-MLVenv plasmids were generously provided by D. Littman (30). Mutant virus constructs (see Table 1) were based on HIVgpt, and construction and sequencing followed standard protocols (28). One protease-minus version of HIVgpt, HIVgpt 2498T was constructed by the insertion of an exogeneous sequence that provides a stop codon at the N-terminus of the pol gene product. The sequence from nucleotide 2495 is 5' GTC ATC GAA TTC CTG CAG CCC TTA AGT <u>TAA</u> CTT AAG GGG GGA TCA GAT 3'; where the underlined triplet provides a *gag*-frame stop codon. HIVgpt dlMARC is a variation of the HIVgpt dlMA, previously described (40). The notable difference is that the dlMARC construct does not introduce a cysteine: while the sequence of the dlMA construct, starting at nt 827 and ending at nt 1149, is 5' AT CGA TGT **CGA CGA** TCT G 3', the sequence of dlMARC is 5' AT CGA **CGA** TCT G 3'. Note that the dlMARC protease minus construct (dlMARC PR-) is a combination of the dl MARC and 2498T.

Point mutants used in this study were as follows: T242C, where the threonine at *gag* codon 242 was changed to a cysteine giving the sequence, 5' ACT AGT <u>TGC</u> CTG CAG GAA 3'; R294C, where the arginine at codon 294 was

changed to a cysteine (5' CCC TTT <u>TGT</u> GAC TAC GTA 3'); V297C, where the valine at codon 297 was changed to a cysteine (5' GAC TAT <u>TGT</u> GAC CGG 3'); Y301C, where the tyrosine at codon 301 was changed to a cysteine (5' CGG TTC <u>TGT</u> AAA ACT CTG CGC GCC GAG 3'); C330S, in which the cysteine at codon 330 was mutated to a serine (5' CCA GAT <u>TCG</u> AAG ACT 3'); and C350S, replacing the cysteine at codon 350 with a serine (5' ATG ACA GCT <u>AGC</u> CAG 3'). In all cases, mutated codons are underlined and altered nucleotides are in bold. Please note that in some but not all cases, additional conservative nt changes were made to facilitate mutant screening by restriction digest analysis.

As a standard for Gag protein quantitation, we also expressed and purified a histidine-tagged capsid protein. To do so, the HIV-1 capsid coding region was cloned into the plasmid pet15B. The resultant plasmid pet15BHIV1147-1900, has a 5' juncture sequence of 5' G GAT CCC GCT G 3', and a 3' juncture sequence of 5' AAT TCC TGC AGC CCG GGG GAT CC 3'. The expressed recombinant protein has non-CA N- and C-terminal amino acid additions of; MGSSHHHHHHSSGLVPRGSHMLQDP, and AQAMSQVTNSCPGDAANKARKQLAAAT AEQ(Stop).

Cell culture, transfections, infections, and infectivity assays. COS7 and HiL cells (CD4+ HeLa cells) were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (FCS), penicillin, and streptomycin. COS7 cells were split from confluent plates 1:4 onto 10 cm plates 24 hours before each transfection. Plasmid DNAs of HIVgpt wt or *gag* mutant HIVgpt constructs were transfected onto COS7 cells by calcium phosphate precipitation (13), and chloroquine was added (25 mM) to improve efficiency. For transfection of one plasmid, fourteen micrograms of plasmid DNA

were used, but for production of pseudotyped virus, eight micrograms of SV-A-MLVenv plasmid were cotransfected with eight micrograms of wt or mutant HIVgpt plasmid. Three days post-transfection, cell lysates and supernatants were collected as previously described (39). For infections, HiL cells on 10 cm plates were 10% confluent. Adsorption of virus was allowed to take place at 37°C in the presence of 4 µg/ml polybrene. Three days post-infection, plates were split 1:8 onto 10 cm plates containing selection media (standard growth media plus 50 µg/ml of xanthine, 3 µg/ml of hypoxanthine, 4 µg/ml of thymidine, 10 µg/ml of glycine, 150 µg/ml of glutamine, and 25 µg/ml of mycophenolic acid (GIBCO)). Plates were refed with selection media every 3 to 4 days until drug resistant colonies formed. The number of colonies on each plate was used to calculate virus titers (gpt resistant colony forming units per ml), and mutant infectivities were determined by comparison with wt HIVgpt titers derived in parallel experiments.

Hybridoma cells were grown in RPMI media supplemented with 10% heat-inactivated fetal calf serum, penicillin and streptomycin in tissue culture flasks. Antibody-containing media was collected by centrifugation to remove cells then frozen at -80° C with 0.2% sodium azide added to aid in preservation.

Protein analysis. Media supernatants of transfected COS7 cells were collected at 3 days post-transfection and filtered through a 0.45-um-pore-sized filter. The filtered supernatants were centrifuged at 4°C for 45 min at 274,000 X g (40,000 rpm in a SW41 rotor) through 2 ml of 20% sucrose in TSE (10 mM Tris hydrochloride, 100 mM NaCl, 1 mM EDTA, 0.1 mM phenylmethylsulfonylfluoride; PMSF). Pellets were resuspended in 100 ul of IPB (20 mM Tris hydrochloride [pH 7.5], 150 mM NaCl, 1 mM EDTA, 0.1% sodium dodecyl sulfate [SDS], 0.5% sodium deoxycholate, 1% Triton X-100, 0.02%

sodium azide) plus 0.1 mM PMSF for electrophoresis, or in TSE for other procedures. Cells from transfected plates were washed with 10 ml ice cold phosphate buffered saline (PBS) twice and collected in 1 ml PBS for each 10 cm plate. Cells then were gently pelleted and resuspended in 1 ml IPB plus 0.1 mM PMSF and re-pelleted to remove debris. Supernatant and cell samples were prepared for electrophoresis by adding an equal volume of 2X sample buffer (12.5 mM Tris hydrochloride [pH 6.8], 2% SDS, 20% glycerol, 0.25% bromophenol blue) plus B-mercaptoethanol (B-Me) to 5%, followed by boiling for 4 to 5 min (17). Samples were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE [24]), electroblotted onto a nitrocellulose filter and Gag proteins on filters were immunodetected using procedures previously described (21). Two primary antibodies were used for the detection of p24: Hy5001, a mouse anti-p24 monoclonal antibody (the kind gift of Epitope Inc., Beaverton, Or.), used at 1:20,000 dilution; and Hy183, a mouse hybridoma monoclonal antibody (obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr. Bruce Chesebro, clone H12-5C), used at 1:30 dilution from media of cells grown in culture. The primary antibody used for the detection of MA was Hy3H6.D4P (the kind gift of Epitope Inc., Beaverton, Or.). In all cases, the secondary antibody used was a goat anti-mouse immunoglobulin G-alkaline phosphotase conjugate (Promega). Immunodetected bands on filters were quantitated either by scanning densitometry with a Bio-Rad model 620 video densitometer on reflectance mode (39), or by obtaining a computer image file of the filter using a Hewlett Packard color flatbed scanner, and processing the image with the NIH Image software package. Efficiency of proteolytic processing was calculated by dividing the density units of individual proteolytic product bands of each sample by the total Gag density of that sample.

Sucrose density gradient fractionation of virus particles was performed as described (39, 40, 44). Briefly, extracellular virions were isolated by pelleting media supernatant from transfected COS7 cells through a 4 ml, 20% sucrose cushion in TSE. The pellets were resuspended in TSE and applied to sucrose gradients consisting of 1.1 ml layers of 20, 30, 40, and 50% sucrose in TSE which had been allowed to mix by sitting for one hour. Gradients were centrifuged at 300,000 X g (50,000 rpm on an SW50.1 rotor) overnight at 4° C, and 400 ul fractions were collected from top to bottom. Fractions were analyzed for sucrose density and Gag protein levels by densitometric quantitation of immunodetected bands. Gag protein levels were normalized to the band with highest Gag protein level from each gradient.

Crosslinking and chemical reactivity methods. Iodine crosslinking of cysteine residues was performed as described (16, 31). For crosslinking with bis-maleimido hexane (BMH; Pierce) treatments, virus particles or proteins were prepared as described above and resuspended in 200 ul of TSE or PBS. BMH was prepared fresh as a 100 mM solution in dimethyl sulfoxide (DMSO), and viral samples were split into equivalent 100 ul fractions and treated with BMH (1 ul of 100 mM BMH in DMSO), or mock treated (1 ul DMSO), and then vortexed gently and incubated at room temperature for one hour. Reactions were terminated by the addition of 2X sample buffer plus B-Me to 5%, and the samples were boiled for 3-5 min (16). For detergent treatment experiments, Triton X-100 or lithium dodecyl sulfate (LDS) were used at final concentrations of 0.25% in crosslinking reactions.

Biotin maleimide (Sigma) reactions were carried out on proteins prepared as described and resuspended in TSE. Biotin maleimide was prepared as a 100 mM solution in DMSO, and then diluted to 10 mM in 10 mM Tris hydrochloride

(pH 7.4). Reactions were performed in 100 ul total volume with the addition of 10 ul of 10 mM biotin maleimide prepared as described, or 10 μl 10% DMSO in 10 mM Tris hydrochloride (pH 7.4) for the mock. Triton X-100 and LDS were used at 0.25% (final concentration) in specified reactions. Reactions were allowed to proceed in the dark on ice for 1 hour, and were terminated by the addition of 2X sample buffer plus β-Me to 5%. Biotinylated proteins were identified by mobility shift and with alkaline phosphatase-conjugated streptavidin on electroblotted filters.

Chemical cleavage of proteins. Specific protein cleavage at cysteine residues by 2-nitro-5-thiocyanobenzoic acid (NTCB, Sigma) followed previously described procedures (2) with modifications. NTCB was prepared fresh as a 2.24 mg/ml solution (1 M) in buffer bob (6M guanidine hydrochloride, 200 mM Tris hydrochloride [pH 8.0], 0.1 mM EDTA [pH 8.0], 0.1 mM dithiothreitol [DTT]). Cleavage reactions were prepared by mixing 175 μl buffer bob with 25 μl of protein sample (in TSE), and reactions were initiated by addition of 50 μl of NTCB solution, and vortexing. Cleavage reactions took place at 37° C in the dark on a rocking platform for 20-75 h. Reactions were terminated by the addition of β-Me to 1%, after which samples were dialyzed two times two hours in a fume hood against 1 L 50 mM ammonium bicarbonate, pH 8.0. Dialyzed samples were vacuum dried overnight, resuspended in 150 μl IPB, 150 μl 2X sample buffer, 15 μl β-Me, and subjected to SDS-PAGE and immunoblotting. For successive biotin-maleimide then NTCB treatments, samples were processed according to the procedure for biotin-maleimide modification except that the reactions were incubated at room temperature 2 h and no β-Me was added to terminate the reaction. Prior to NTCB cleavage, 10 μl aliquots were set aside for electrophoresis, while the remainder of the biotin-maleimide reactions were

combined with 1.5 ml of buffer bob plus 300 μM DTT, after which 400 μl 10 mM NTCB (in buffer bob plus 300 μM DTT) was added. Cleavage reactions then proceeded as described above.

RNAse protection and reverse transcriptase reactions. RNAse protection assays using the HIV 831-680 riboprobe were performed as previously described (40, 44). The riboprobe spans the HIV major splice donor sequence so it will detect both spliced and unspliced transcripts. Exogeneous reverse transcriptase assays were performed as previously described (39), and results were normalized to densitometrically determined Gag levels from immunoblots. To assay reverse transcriptase activities on native templates, endogeneous reverse transcriptase reactions were performed on permeabilized pelleted particles using previously established reaction incubations. Total nucleic acid was isolated from each reaction mix, and minus strand strong stop DNA was assayed by RNAse protection using a 151 nt sense strand probe.

RESULTS

Cysteine-specific crosslinking of wild type HIV-1 Gag proteins

There are a total of 10 cysteines encoded by the HIV-1 *gag* gene (Figure 2), two in the matrix domain (MA), two in the capsid domain (CA), six in the nucleocapsid domain (NC), and none in p6. The NC cysteines are absolutely conserved among over 70 different HIV-1 *gag* genes that have been sequenced (26), and the two MA cysteines are conserved to a very high degree, one cysteine is invariant, while the other cysteine is absent in only 3 isolates (26). The two CA cysteines also are well conserved. The first cysteine (*gag* codon 330) is absolutely conserved, and the second cysteine (350) is conserved in all but one isolate (26).

The corresponding CA cysteine residues are present and conserved in HIV-2 ROD and SIV isolates (20), but these viruses encode an additional cysteine in the central portion of their capsid domains. Of these cysteine residues, evidence indicates that NC cysteines form two zinc fingers important in viral RNA encapsidation (1). Additionally, MA cysteines have been located in NMR and X-ray structures of the matrix domain. However, it is unclear how CA cysteines contribute to HIV function, and how MA, CA, and NC cysteines interact in Pr55^{Gag}. These studies were performed to clarify the roles of the HIV-1 CA cysteines (C330 and C350) in the viral life cycle and to combine cysteine mutagenesis with crosslinking and chemical reactivity protocols to probe virus particle structure (16).

To determine whether HIV-1 Gag cysteines are in close proximity to cysteines on neighboring Gag proteins, we initially ran proteins from immature and mature virus particles on non-reducing gels and found no evidence for intermolecular Gag protein cystine formation (23). As a second step, we used membrane permeable bis-maleimido hexane (BMH) to crosslink cysteines in intact mature and immature protease minus (PR-) virus particles. When wild type (wt) mature HIV particles were subjected to BMH crosslinking and the protein products were immunodetected with either anti-CA or anti-MA antibodies, no novel crosslinked species were identified (Figure 3, lanes A-B and E-F). These data imply either that BMH cannot penetrate mature virions, or that cysteines on mature MA or CA proteins were not in close enough proximity to cysteines on neighboring proteins to crosslink. In contrast, when immature, PR-particles were subjected to BMH crosslinking, a novel protein band at 110 kD was detected with both the CA (lane D) and the MA (lane H) antibodies, along with an additional less pronounced band of unknown origin at about 100 kD. This result is consistent with the supposition that at least one pair of cysteine

residues on neighboring Pr55^{Gag} molecules can be crosslinked, and is similar to our previous findings with Moloney murine leukemia virus (M-MuLV) PrGag protein crosslinking (16). Note also that the slowest migrating PrGagpol band in lane C disappears on crosslinking (lane D), presumably because PrGagpol proteins crosslink to generate a species which does not enter our running gels readily.

The appearance of a 110 kD band upon crosslinking of immature virus particles was consistent with the presence of a Pr55^{Gag} dimer, but could also be due to the crosslinking of Pr55^{Gag} to an unknown protein of approximately 55 kD. In order to analyze the composition of the putative dimers, we constructed a matrix deletion Pr55^{Gag} protein, dlMARC PR-, which retains N- and C-terminal MA sequences but is deleted from codon 15 to codon 111. The dlMARC mutant is similar to the previously described dlMA mutant (40), which assembles conditionally infectious virus particles, except that dlMARC is two residues shorter, and retains no cysteines in the MA domain. To examine the composition of Gag-reactive crosslinked molecules, we subjected PR- wt and dlMARC proteins in virus particles to BMH treatments (Figure 4). Again, BMH treatment of PR- particles yielded a new band at 110 kd, and a less pronounced 100 kD band of unknown origin (lane B). Crosslinking of dlMARC PR- virus particles produced an additional band at approximately 82 kD, consistent with the predicted size of a dlMARC crosslinked dimer (lane D). When separately isolated PR- and dlMARC PR- particles were mixed prior to crosslinking, the predicted homodimer bands of 82 kd and 110 kD were observed as was the 100 kD band of unknown origin (lane I). In contrast to the mixing experiment, when PR- and dlMARC PR- constructs were cotransfected, crosslinking of the isolated virus particles showed an additional band at approximately 96 kD (lane G). This new band corresponds to the expected size for a mixed dimer of PR- and

dlMARC PR- Gag proteins. However, since the mixed dimer 96 kD band
migrated so closely to the unknown 100 kD band, we repeated the crosslinking
experiment and ran all crosslinked products side by side (Figure 4, lanes J-M).
Again the wt dimer (lane J) and dlMARC (lane K) dimers were observed, and the
mixed particles showed homodimers only (lane M). However in addition to the
homodimers, the cotransfection experiment (lane L) showed a heterodimer at 96
kD. We conclude from these experiments that BMH crosslinks neighboring Gag
proteins at cysteine residues. Furthermore, neither mature MA nor CA crosslink
(Figure 3) and the precursor Gag does not require MA cysteines to crosslink
(Figure 4). Independent experiments showed that a PrGag protein deleted for all
of the cysteine residues in the NC domain was impaired in crosslinking (data not
shown). While we cannot exclude arguments that NC mutant proteins fold
differently form wt PrGag proteins, our results are consistent with the
interpretation that crosslinking of HIV-1 Pr55^{Gag} occurred via the cysteines in
the NC domain. This result is similar to what we have observed with M-MuLV
PrGag proteins (16), and suggests that NC in Gag precursors may play an
important role in multimerization of Gag proteins (3,7).

Analysis of HIV-1 capsid protein cysteines

Since HIV-1 capsid protein cysteines are well conserved (26) and in close
proximity in the primary sequence (only 20 residues apart), we were interested
in examining the status of these two residues. Consequently cysteines at *gag*
codons 330 and 350 were mutated separately to encode serines in the mutants
C330S and C350S (Figure 1). When wt, C330S and C350S HIVgpt constructs were
transfected into COS7 cells, the standard Pr55, p41 and CA proteins were
detected in lysates of wt (Figure 5, lane A) and C330S (lane C) transfected cells.
However, we consistently detected no capsid-related proteins in lysates from

cells transfected with the C350S mutant (lanes B, D), when anti-CA antibody Hy5001 was used. The absence of C350S proteins in cell lysates could be due to something as trivial as failed transfections, so a parallel immunoblot was performed with the second anti-CA monoclonal antibody, Hy183. As illustrated (Figure 5, lanes E-H), the second monoclonal antibody readily detected wt (lane E), C330S (lane G) and C350S (lanes F, H) Pr55, p41 and CA proteins in transfected cell lysates: these results suggest that the C350S mutation altered the Hy5001 epitope so that the mutant protein was no longer recognizable by this antibody on immunoblots. Using the Hy183 monoclonal antibody, which recognizes wt, C330S and C350S proteins, we examined particle-associated Gag protein levels in media supernatants of transfected cells as a measure of the efficiency of wt and mutant proteins in directing the assembly and release of virus particles. As shown in lanes I and K, wt and C330S proteins were released efficiently from transfected cells. In contrast, the total amount of Gag protein released from C350S-transfected cells was 10- to 30-fold lower than that seen for its counterparts. These results indicate that the C350S mutation has a drastic effect on HIV particle assembly. At least two mechanisms might account for the C350S assembly defect: the mutation might cause a mislocalization of the protein within cells, or it could interfere with interactions that are necessary for particle assembly. Because results from subfractionation and immunofluorescence studies appeared indistinguishable for the wt and C350S proteins (data not shown), we believe the C350S mutation interferes with assembly by disrupting Gag protein contacts with either membranes, viral RNA, or other proteins.

The fact that the C330S mutant assembled virus particles, while the C350S mutant was assembly-defective seems to discount the notion that these two residues form an intramolecular cystine. Furthermore, non-reducing gels showed little if any evidence for intermolecular cystine bridge formation between capsid

proteins or Pr55^{Gag} proteins (23). However, we were interested in directly testing the accessibility of the capsid cysteines. To do so, wt HIVgpt virus particles were treated with biotin-maleimide which can biotinylate free thiols. Treatments were performed in the absence of detergent, in the presence of Triton X-100 (which dissembles mature but not immature HIV particles [39]), or in the presence of lithium dodecyl sulfate (LDS; which dissembles both mature and immature HIV particles). Since biotin-maleimide is not membrane permeable in the absence of detergents, it is not surprising that the immunoblot profile of biotin-maleimide-treated samples was comparable to that of the mock (Figure 6; lanes B, E versus A; and lanes I, L versus H). When biotin-maleimide reactions occurred in the presence of either detergent, one slightly lower mobility CA band (biotin-CA) was observed, and 30-70% of the capsid protein was shifted to this mobility when Hy183 was used for detection (lanes J, K, M, N). However, when Hy5001 was utilized as a detection reagent, there was an apparent reduction in the amount of CA, and no biotin-CA was observed (lanes C, D, F, G). We interpret these results to indicate that the biotin-CA band consists of CA biotinylated at C350; at C350 plus C330; or a combination of the mono- and di-biotinylated species.

In contrast to biotinylation results in triton X-100 or in the absence of detergent, when particles were treated with LDS there was evidence of lower mobility of Pr55 and p41 species as detected by Hy183 (lanes J, M). This result is consistent with previous observations that immature HIV particles are resistant to dissembly by non-ionic detergents (17, 36), but also suggests that none of the HIV-1 PrGag cysteine residues were accessible to biotin-maleimide unless immature particles were completely disrupted. Additionally, since neither p41, Pr55, nor their biotinylated counterparts were visible when biotin-maleimide-,

LDS-treated proteins were probed with Hy5001, it would appear that all biotinylated and biotin-Pr55 species were modified at C350.

To extend our observations, wt Gag proteins were mock- or biotin-maleimide-treated, and then subjected to NTCB cleavage at free cysteines. In this case, biotinylated cysteines should be resistant to NTCB cleavage, and cleavage should occur only at residues that were unable to be biotinylated. In this experiment, two different HIVgpt preparations were used, wt and dlMARC, both of which form mature virus particles that appear identical with respect to their capsid proteins. As expected, completely untreated samples yielded only one capsid band (Figure 7, lanes A,D), while biotin-maleimide-reacted samples (which have not been NTCB cleaved) gave the additional lower mobility band seen in Figure 6 (lanes B, C). When unbiotinylated proteins corresponding to lanes A and D were cleaved with NTCB, several bands were seen (lanes E, H), one presumably corresponding to CA which was not cleaved by NTCB (band #2), and bands of higher mobility (bands #3 and #4), cleaved at C330 and/or C350 with NTCB. When samples were reacted first with biotin-maleimide and then cleaved by NTCB, the relative amount of material in bands #1 (biotin-CA) and #2 (CA) increased (lanes F, G), consistent with the notion that biotin-maleimide reactions protected CA from NTCB cleavage. Interestingly, the amount of material in band #3 remained constant, whether or not the samples received biotin-maleimide pretreatment. In contrast, biotin-maleimide pretreatment caused a clear decrease in the relative amount of band #4 (compare band #4 in lanes F versus E, and G versus H). Since C350 preferentially reacts with biotin-maleimide (Figure 6) this suggests that band #4 is the cleavage product at C350, and band #3 is the cleavage product at C330. This result is paradoxical since the C330 cleavage product (band #3) should be smaller than the C350 cleavage product (band #4). However, this type of gel migration

anomaly has been observed in cases where the C-terminus of avian retrovirus proteins are naturally cleaved, or artificially truncated (6, 7). Our results indicate that C350 is highly reactive to biotin maleimide, but that C330 appears considerably less reactive. The lack of C330 reactivity could be due to its inaccessibility, even in the presence of ionic detergent. Alternatively, the thiol at C330 could be blocked by an as yet undetermined low molecular weight factor. However, if this is the case, the factor could not be necessary for virus assembly, since the mutant C330S does assemble virus paricles.

Phenotypes of HIV-1 cysteine substitution mutants

Since we examined the status of the two natural cysteines in the HIV-1 capsid proteins, we thought it might be useful to examine what occured when additional cysteines were introduced into the capsid coding region. In addition to the C330S and C350S mutants (Figure 1) we chose to look at four others. One mutant, T242C (Figure 1), is at a site where a linker insertion mutation killed virus infectivity, and impaired proteolytic processing of the Pr55[Gag] precursor (39). In addition, three mutations were created in the C-terminal portion of the capsid domain major homology region (MHR), R294C, V297C, and Y301C (Figure 1). Previously, others found that mutations in the most conserved, amino-terminal residues of the MHR killed virus infectivity and could block assembly (27, 37). We chose residues 294, 297, and 301 since NMR studies suggested that the HIV MHR forms an amphipathic helix in solution with R294 on the hydrophilic face, and V297 and Y301 on the hydrophobic face (5). We felt it was possible to determine whether cysteines introduced at these positions might interfere with assembly or, if not, might be used to probe structural features of the MHR.

For analysis, mutant and wt HIVgpt constructs first were transfected into COS7 cells and virus particle release levels were assayed as descibed in Figure 5. Similar to the C330S mutant and in contrast to the C350S mutant, media levels of T242C, R294C, V297C and Y301C were all roughly comparable to wt levels (data not shown), suggesting that these mutants were not defective for virus assembly. Furthermore, all well-released mutants made particles of wt density and were appropriately processed (data not shown).

Since particles released by T242C, R294C, V297C, Y301C or C330S were not grossly defective, we tested the abilities of Gag mutant virions to infect target cells. To do so, wt or mutant HIVgpt constructs were cotransfected with an envelope expression construct into COS7 cells, and filtered media supernatants were collected. For infectivity assays, infectious events were monitored by transduction of the proviral gpt gene. Thus, particles were applied to target HiL cells which subsequently were split into gpt selective media, and grown until the appearance of drug resistant colonies. In all cases, mutant and wt HIVgpt infections were performed in parallel. Titers, in colony forming units per ml (cfu/ml) were calculated based on colony numbers, and mutant titers were compared to wt titers obtained in infections performed at the same time. Interestingly, the infectivities of all of the capsid domain single-substitution mutants were reduced considerably from that of wt (Table 1). The near-zero titers for the C350S mutant were expected, based on its defect in release. However, of the other mutants, only T242C showed appreciable titers, 5-20% that of wt. In contrast, C330S titers were 1-5% wt, and all mutations in the MHR (R294C, V297C, Y301C) gave titers of less than 1% of wt.

Why are the infectivities of the capsid domain single-substitution mutants so low? RNAse protection assays showed that the T242C, R294C, V297C, Y301C and C330S mutant particles packaged roughly wt levels of genomic viral (data

not shown), so the mutants do not appear defective for RNA encapsidation. Similarly, efficiently released mutant particles possessed wt levels of exogeneous template reverse transcriptase (RT) activity (39) and endogenous RT minus strand strong stop product levels (data not shown), suggesting that the initial steps of reverse transcription are not impaired. Thus, it seems that our capsid mutants are defective either for virus entry into target cells, or at an as yet undefined post-entry step of infection. These results are reminiscent of those observed for some other HIV capsid mutants (39).

Crosslink analysis of cysteine substitution mutants

To probe the cysteines introduced in the capsid domain, particles were isolated from media supernatants of transfected COS7 cells and either mock treated, or crosslinked at cysteine residues with BMH (Figure 8). As expected (see Figure 3), mature wt particles showed little evidence of capsid protein crosslinking (Figure 8, lane B) when compared to the mock treatment (lane A). Similar results were obtained with T242C (lanes C, D). In contrast, BMH treatment of C330S and MHR mutant particles resulted in the appearance of anti-CA-reactive bands at 46-50 kd, consistent with CA-CA dimers. The 46-50 kD products show slightly different mobilities which vary by mutant. These variations may be due to mobility differences of dimers depending on the exact location of the crosslink. However, an alternative hypothesis, that CA monomers crosslink to unknown 22-26 kD proteins, can not be ruled out.

Given the lack of putative CA-CA dimers for wt HIVgpt, it was surprising that the C330S mutant showed a putative dimer band (Figure 8, lane L). One explanation for this result is that the C330S mutation alters the conformation of CA in virus particles, such that C350 is placed in close proximity to C350 in neighbor molecules, and is able to form a crosslink. An alternative explanation

for this observation is that in wt particles, C330 and C350 preferentially form an intramolecular crosslink, and the absence of a suitable intramolecular partner in the C330S mutant leaves C350 free to form intermolecular crosslinks. While this interpretation might imply that crosslinking to form dimers always will occur when there is an odd number of cysteines in CA, the lack of crosslinking for T242C, which has three cysteines, argues against this implication.

As indicated above, all three MHR cysteine substitution mutants showed putative CA-CA dimer species (Figure 8, lanes F, H J). These crosslinked species disappeared when virus particles were treated with either 0.1% Triton-X 100 or 0.1% SDS (data not shown), indicating that the mature virus particle structure was necessary for crosslink species to appear. Our results with R294C, V297C and Y301C are consistent with the notion that MHR domains are close enough to be crosslinked in neighboring molecules, possibly forming an interface between capsid domains in mature particles. However, one caveat is that MHR cysteine mutations might cause a conformational change such that other CA cysteines (C330 or C350) are able to crosslink, or that mixed crosslinks might form. If this is the case, the conformational change appears to require cysteine substitutions at specific locations, since the T242C did not cause the appearance of a significant crosslink band.

DISCUSSION

There are ten cysteines in the HIV-1 Gag protein, two in the matrix domain, two at the C-terminus of the capsid domain, and six involved in the two zinc-finger motifs in NC (Figure 2). Since we have not found evidence for PrGag or CA dimers in non-reducing gels, it does not appear that intermolecular cystine formation is essential to particle formation. However, PrGag proteins can be

crosslinked at cysteines with BMH to form dimers, and crosslinking appears to occur in the NC domain (Figures 3, 4). In contrast to PrGag, we found no evidence for crosslinking of CA domains in mature wt particles (Figure 4, 8). One trivial explanation for this observation might be that the HIV-1 CA domain cysteines (C330 and C350) naturally form an intramolecular cystine. However, our ability to biotinylate C350 with biotin-maleimide (Figure 6) argues against the existence of a natural intramolecular C330-C350 disulfide on the majority of capsid proteins. Interestingly, C330 appears resistant to biotinylation, even in the presence of ionic detergent (Figures 6, 7). This observation suggests that the C330 thiol could be chemically blocked. However, an alternate explanation is that the CA C-terminus forms a secondary structure that is stable to ionic detergent. Indeed, the anomalous gel migration of cysteine cleavage products (Figure 7) suggests that this is the case, and is reminiscent of results with C-terminally modified avian retrovirus capsid proteins (6, 7). Although we currently have no direct evidence, we also would predict that the C-terminal cysteines of HIV-2 and SIV will possess the same chemical reactivities to those of HIV-1 CA. In this regard, it is noteworthy that HIV-2 and SIV encode CA cysteines besides those near the C-terminus. Based on their non-conservation, it seems likely that these additional cysteines may interact with virus accessory proteins, such as vpx, which are not present in HIV-1.

In our molecular genetic analyses, we found that mutation of C330 to a serine (C330S) gave a phenotype that is similar to many other retrovirus capsid domain mutants. The infectivity of this mutant was greatly reduced relative to that of wt (Table 1), but there was no apparent reduction in particle assembly (Figure 5). Nor were there apparent alterations in virus structure based on proteolytic processing, sucrose gradient profiles, RNA levels, or reverse transcriptase activities (data not shown), although thorough electron microscopy

studies were not done here. These results suggest that the C330S mutation either impairs virus entry or post-penetration processes. In contrast to C330S, the mutant C350S showed the most drastic effect on HIV assembly that we've observed for a single residue substitution (Figure 5). This might seem surprising for a conservative mutation of a cysteine which does not appear to be involved in consistent intra- or intermolecular cystine formation, and could be due either to alteration of important interprotein assembly contacts or to mislocalization of the mutant Gag proteins. In this regard, we have found that the C350S immunofluorescence pattern appeared like that of wt, as did the membrane versus cytosol fractionation pattern (data not shown), suggesting that the mutation does not affect subcellular localization. Consequently, our hypothesis is that the C350S mutation changes an important CA C-terminal conformation that is a prerequisite for Gag protein oligomerization and particle assembly.

Insofar as CA cysteine substitution mutations are concerned, three were created in the C-terminal portion of HIV-1 MHR, since this region has been modeled to form an amphipathic helix (5); and one mutation was created at HIV *gag* codon 242, near a site which, when mutated by linker insertion, had a slight effect on assembly and processing (39). All these mutants had similar phenotypes to the C330S mutant, although T242C was three times more infectious than C330S, and the MHR mutants were considerably less infectious (Table 1). When T242C, R294C, V297C, Y301C and C330S were subjected to BMH-crosslinking, T242C capsid proteins showed little evidence of crosslinking, while the other mutants gave putative CA-CA dimer bands (Figure 8). The results with C330S were surprising given that wt CA proteins did not crosslink (Figure 3, 8). One explanation for this result is that the C330S mutation made C350 more accessible or closer to neighbor C350 residues. However, we favor the hypothesis that BMH preferentially forms intramolecular C330-C350 crosslinks in wt particles, while

blocking the intramolecular reaction in C330S particles permits the intermolecular reaction to occur. Similar to C330S, all the MHR cysteine substitution mutants showed putative CA-CA dimers on BMH-treatment (Figure 8). This could occur if mutations in the MHR affected C350 or C330S so that they now could react with a neighbor molecule. However, assuming that the block to intermolecular crosslinking via residue 350 is due to an intramolecular 330-350 reaction, it is not clear how MHR mutations would affect this interaction. An alternative hypothesis is that cysteines of amphipathic MHR helices (5) on neighboring capsid molecules are in close enough proximity to crosslink. We currently are testing these structural predictions in vitro.

ACKNOWLEDGEMENTS

REFERENCES

1. **Aldovini, A., and R. A.Young.** 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J. Virol. **64:**1920-1926.

2. **Baehler, M., F. Benfenati, F. Valtorta, A.J. Czernik, and P. Greengard.** 1989. Characterization of synapsin I fragments produced by cysteine-specific cleavage: a study of their interactions with F-actin. J. Cell Bio. **108:**1841-1849.

3. **Bennet, R. P., T.D. Nelle, and J.W. Wills.** 1993. Functional chimeras of the Rous Sarcoma Virus and Human Immunodeficiency Virus Gag proteins. J. Virol. **67:**6487-6498.

4. **Cann, A. J., and J. Karn.** 1989. Molecular biology of HIV-1; new insights into the virus life cycle. AIDS **3 (Suppl. 1):**S19-S34.

5. **Clish, C.B., D.H. Peyton, and E. Barklis.** 1996. Spectroscopic study of an HIV-1 capsid protein major homology region peptide analog. FEBS Letters **378:**43-47.

6. **Craven, R.C., A.E. Leure-duPree, C.R. Erdie, C.B. Wilson, and J.W. Wills.** 1993. Necessity of the spacer peptide between CA and NC in the Rous sarcoma virus Gag protein. J. Virol. **67:**6246-6252.

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

8. **Facke, M., A. Janetzko, R.L. Shoeman, and H. Krausslich.** 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.

9. **Gelderblom, H.R., M. Ozel, and G. Pauli.** 1989. Morphogenesis and morphology of HIV structure function relations. Arch. Virol. **106:**1-13.

10. **Gheysen, D., E. Jacobs, F. de Foresta, C. Thiriart, M. Francotte, D. Thines, and M. De Wilde.** 1989. Assembly and release of HIV-1 precursor pr55^{Gag}virus-like particles from recombinant baculovirus-infected insect cells. Cell **59:**103-112.

11. **Gorelick, R. J., S. M.Nigida, Jr, J. W. Bess, Jr, L.O. Arthur, L.E. Henderson, and A. Rein.** 1990. Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. J. Virol. **64:**3207-3211.

12. **Gottlinger, H. G., T. Dorfman, J.G. Sodroski, and W.A. Haseltine.** 1991. Effect of mutations affecting the p6 *gag* protein on human immunodeficiency virus particle release. Proc. Natl. Acad. Sci. USA **88:**3195-3199.

13. **Graham, R.M and A. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456-467.

14. **Haffar, O., J. Garrigues, B. Travis, P. Moran, J. Zarling, and S.-L. Hu.** 1990. Human immunodeficiency virus-like, nonreplicating *gag-env* particles assemble in a recombinant vaccinia virus expression system. J. Virol. **64:**2652659.

15. **Han, K., C. Richard, G. Zhang, and A. Delacourte.** 1986. Sequence homology analysis of proteins by chemical cleavages: using a mono and two

dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Int. J. of Biochem. **18**:1073-1082.

16. **Hansen, M.H., and E. Barklis.** 1995. Structural interactions between retroviral Gag proteins examined by cysteine cross-linking. J. Virol. **69**:1150-1159.

17. **Hansen, M., L. Jelinek, R. S. Jones, J. Stegeman-Olsen, and E. Barklis.** 1993. Assembly and composition of intracellular particles formed by Moloney murine leukemia virus. J. Virol. **67**:5163-5174.

18. **Hansen, M., L. Jelinek, S. Whiting, and E. Barklis.** 1990. Transport and assembly of Gag proteins into Moloney murine leukemia virus. J. Virol. **64**:5306-5316.

19. **Henderson, L.E., M. A. Bowers, R. C. SowderII, S. A. Serabyn, D. G. Johnson, J. W. Bess, Jr., L. O. Arthur, D. K. Bryant, and C. Fenselau.** 1992. Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. J. Virol. **66**:1856-1865.

20. **Henderson, L.E., T.D. Copeland, R.C Sowder, A.M. Schultz, and S. Oroszlan.** 1988. Analysis of proteins and peptides purified from sucrose gradient branded HTLV-III, p135-147. In: D. Bolognesi (ed.) Human Retroviruses and AIDS: Approaches to Prevention and Therapy. Alan R. Liss, Inc., New York.

21. **Jones, T.A., G. Blaug, M. Hansen, and E. Barklis.** 1990. Assembly of *gag*-B-galactosidase proteins into retrovirus particles. J. Virol. **64**:2265-2279.

22. **Karacostas, V., K. Nagashima, M. Gonda, and B. Moss.** 1989. Human immunodeficiency virus-like particles produced by a vaccinia virus expression vector. Proc. Natl. Acad. Sci. USA **86**:8964-8967.

23 **Karanjia, S. and E. Barklis.** Unpublished Results.

24. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680-685.

25. **Leis, J., D. Baltimore, J. B. Bishop, J. Coffin, E. Fleissner, S. P. Goff, S. Oroszlan, H. Robinson, A. M. Skalka, H. M. Temin, and V. Vogt.** 1988. Standardized and simplified nomenclature for proteins common to all retroviruses. J. Virol. **62**:1808-1809.

26. **Louwagie, J., F.E. McCutchan, M. Peeters, T.P. Brennan, E. Sanders-Buell, G.A. Eddy, G. van der Groen, K. Fransen, G.M. Gershy-Damet, R. Deleys, and D.S. Burke.** 1993. Phylogenetic analysis of *gag* genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. AIDS **7**:769-80.

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

28. **Maniatis, T., E.F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

29. **Mervis, R. J., N. Ahmad, E. P. Lillehoj, M.G. Raum, F. H. R. Salazar, H. W. Chan, and S. Venkatesan.** 1988. The *gag* gene products of human immunodeficiency virus type 1: alignment within the *gag* open reading frame, identification of posttranslational modifications, and evidence for altermative *gag* precursors. J. Virol. **62**:3993-4002.

30. **Page, K.A., N.R. Landau, and D.R. Littman.** 1990. Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. J. Virol. **64:**5270-5276.
31. **Pakula, A.A., and M.I. Simon.** 1992. Determination of transmembrane structure by disulfide cross-linking: the Escherichia coli Tar receptor. Proc. Natl. Acad. Sci. USA **89:**4144-4148.
32. **Ratner, L., W. Haseltine, R. Patarca, K.J. Livak, B. Starich, S.F. Josephs, E.R. Doran, J.A. Rafalski, E.A. Whitehorn, K. Baumeisterm L. Ivanoff, S.R. Petteway, Jr., M.L. Pearson, J.A. Lautenberger, T.S. Papas, J. Ghrayeb, N.T. Chang, R.C. Gallo, and F. Wong-Staal.** 1985. Complete nucleotide sequences of the AIDS virus, HTLV-III. Nature (London) **313:**277-284.
33. **Rhee, S. S., and E. Hunter.** 1990. Structural role of the matrix protein of type D retroviruses in Gag polyprotein stability and capsid assembly. J. Virol. **64:**4383-4389.
34. **Shioda, T., and H. Shibuta.** 1992. Production of human immunodeficiency virus (HIV)-like particles from cells infected with recombinant vaccinia viruses carrying the *gag* gene of HIV. Virology **175:**139-148.
35. **Smith, A. J., M.-I. Cho, M.-L. Hammarskjold, and D. Rekosh.** 1990. Human immunodeficiency virus type 1 $Pr55^{gag}$ and $Pr160^{gag-pol}$ expressed from a simian virus 40 late replacement vector are efficiently processed and assembled into viruslike particles. J. Virol. **64:**2743-2750.
36. **Stewart, L., G. Shatz, and V. Vogt.** 1990. Properties of avian retrovirus particles deficient in viral protease. J. Virol. **64:**5076-5092.
37. **Strambio-de-Castillia, C., E. and Hunter.** 1992. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. J. Virol. **66:**7021-7032.
38. **Von Poblotzki, A., R. Wagner, M Niedrig, G. Wanner, H. Wolf, and S. Modrow.** 1993. Identification of a region in the $Pr55^{Gag}$ polyprotein essential for HIV-1 particle formation. Virology **193:**981-985.
39. **Wang, C., and E. Barklis.** 1993. Assembly, processing, and infectivity of human immunodeficiency virus type 1 *gag* mutants. J. Virol. **67:**4264-4273.
40. **Wang, C., Y. Zhang, J. McDermott, and E. Barklis.** 1993. Conditional infectivity of a human immunodeficiency virus matrix domain mutant. J. Virol. **67:**7067-7076.
41. **Wills, J.W. and R.C. Craven.** 1991. Form, function and use of retroviral Gag proteins. AIDS **5:**639-654.
42. **Yu, X, Q. Yu, T. Lee, and M. Essex.** The C terminus of human immunodeficiency virus type 1 matrix protein is involved in early steps of the virus life cycle, (1992) J.Virol. **66:**5667-5670.
43. **Yu, X., X. Yuan, Z. Matsuda, T. Lee, and M. Essex.** 1992. The matrix protein of HIV-1 is required for incorporation of viral envelope protein into mature virions. J.Virol. **66:**4966-4971.
44. **Zhang, Y., and E. Barklis.** 1995. Nucleocapsid protein effects on the specificity of retrovirus RNA encapsidation. J. Virol. **69:**5716-5722.

TABLES AND FIGURES

Table 1. Infectivity of HIV *gag* mutants.

Construct	Titer	HIVgpt titer	% Infectivity
HIVgpt	2,296	2,296	100
T242C	416	2,178	19.1
	34	874	3.9
	102	749	13.6
	192	2,007	9.6
R294C	20	7,847	0.3
	5	1,504	0.3
	7	3,831	0.2
V297C	5	2,451	0.2
	13	3,831	0.3
	21	1504	1.4
	4	2,007	0.2
Y301C	17	2,451	0.7
	12	2,451	0.5
	22	7,831	0.3
	0	1504	0
	3	3,831	0.1
C330S	61	1,400	4.3
	41	2,451	1.7
C350S	0	1,400	0
	1	2,451	0.04

Each HIVgpt-based construct was cotransfected with SV-A-MLVenv into COS7 cells and 3 days later, cell supernatants were used to infect HiL cells. Infections and selections for mycophenolic acid resistant colonies were performed as described in Materials and Methods. The value for the HIVgpt (wt) titer was calculated as an average of 11 independent infectivity experiments. Replicate experiments were performed with different DNA plasmid preparations, at different times, or both. Infectivities for each mutant were determined by the ratio of its titer versus the HIVgpt-wt titer (middle column) in a parallel experiment.

Figure 1. HIV-1 Gag mutants.

2498T	[2495]	5' GTC	**ATC**	**GAA**	**TTC**	**CTG**	**CAG**	**CCC**	**TTA**	**AGT**	<u>**TAA**</u> 3'
	(pol 47)	S	I	E	F	L	Q	P	L	S	Stop

dlMA	[828]	5' GAT	CGA	TGT	**CGA**	**CGA**	TCT	GAC	ACA	3'	[1154]

$$\text{D} \quad \text{R}_{15} \quad \text{C} \quad \text{R} \quad \text{R} \quad \text{S} \quad \text{D}_{111} \quad \text{T}$$

dlMARC [828]   5' GAT CGA **CGA** TCT GAC ACA   3'   [1154]
$$\text{D} \quad \text{R}_{15} \quad \text{R} \quad \text{S} \quad '\text{D}_{111} \quad \text{T}$$

T242C   [1507]   5' ACT AGT **TGC** CTG CAG GAA   3'
$$\text{T}_{240} \quad \text{S} \quad \text{T>C} \quad \text{L} \quad \text{Q} \quad \text{E}$$

R294C   [1662]   5' CCC TTT **TGT** GAC TAC GTA   3'
$$\text{P}_{292} \quad \text{F} \quad \text{R>C} \quad \text{D} \quad \text{Y} \quad \text{V}$$

V297C [1671]   5' GAC TAT **TGT** GAC CGG TTC   3'
$$\text{D}_{295} \quad \text{Y} \quad \text{V>C} \quad \text{D} \quad \text{R} \quad \text{F}$$

Y301C [1683]   5' CGG TTC **TGT** AAA ACT CTG   3'
$$\text{R}_{299} \quad \text{F} \quad \text{Y>C} \quad \text{K} \quad \text{T} \quad \text{L}$$

C330S [1770]   5' CCA GAT **TCG** AAG ACT ATT   3'
$$\text{P}_{328} \quad \text{D} \quad \text{C>S} \quad \text{K} \quad \text{T} \quad \text{I}$$

C350S [1830]   5' ACA GCT **AGC** CAG GGA GTA   3'
$$\text{T}_{348} \quad \text{A} \quad \text{C>S} \quad \text{Q} \quad \text{G} \quad \text{V}$$

Mutant plasmids used here are all HIVgpt vectors and were constructed by standard methods. The mutation in 2498T is an insertion which provides a premature stop codon to the pol gene and inactivates the protease protein. The mutations dlMA and dlMARC are both 316 nt deletions in the MA domain which remove 96 codons. Capsid domain mutations are as follows: T242C changes the threonine at codon 242 to a cysteine; R294C changes the arginine at 294 to a cysteine; V297C changes the valine at 297 to a cysteine; Y301C changes the tyrosine at 301 to a cysteine; C330S changes the cysteine at 330 to a serine; and C350S changes the cysteine at 350 to a serine. In all cases sequences are shown 5' to 3', altered or inserted nucleotides and residues are in bold, nt numbers for all 5' and some 3' nucleotides are in brackets, and codon numbers are subscripted.

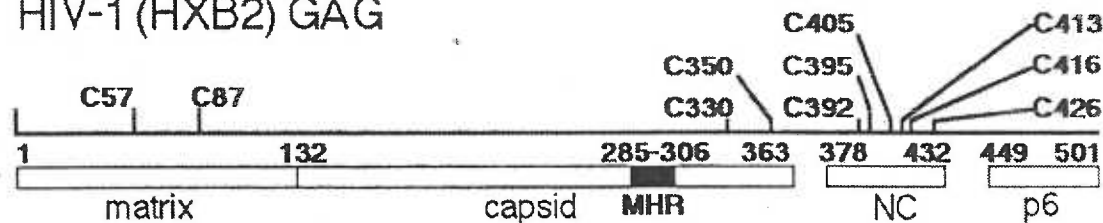Figure 2. **Location of cysteine residues in HIV Pr55^{Gag}.** The HIV-1 (HXBc2) Pr55Gag protein consists of four domains; matrix (MA), capsid (CA), nucleocapsid (NC), and p6, that are cleaved to maturation by the virally encoded protease (PR) during the budding process. The precursor polypeptide is 501 codons in length, and the juncture codons are indicated (note that both the CA/NC and NC/p6 junctures are formed by the excision of small peptides of unknown function, p1 and p2). Pr55Gag contains 10 cysteines which are indicated on the top bar with codon numbers. The HIV MHR (major homology region) is indicated and spans *gag* codons 285 to 306.

Figure 3. Crosslinking of HIV-1 Gag proteins. Wild-type (wt; lanes A,B,E,F) or protease-minus (PR-; lanes C,D,G,H) HIVgpt particles were mock treated (lanes A,C,E,G), or treated with BMH (lanes B,D,F,H) to crosslink neighboring cysteine residues. Samples were separated by SDS-PAGE, electroblotted and Gag proteins were detected by immunoblotting with either anti-CA antibody (lanes A-D) or the anti-MA antibody (lanes E-H). HIV Gag proteins MA, CA, and Pr55^{Gag} are indicated, as are putative 110 kd Pr55^{Gag} dimers. Also evident are PrGagpol proteins (lanes C,G) and 100 kd species in lanes D,H.
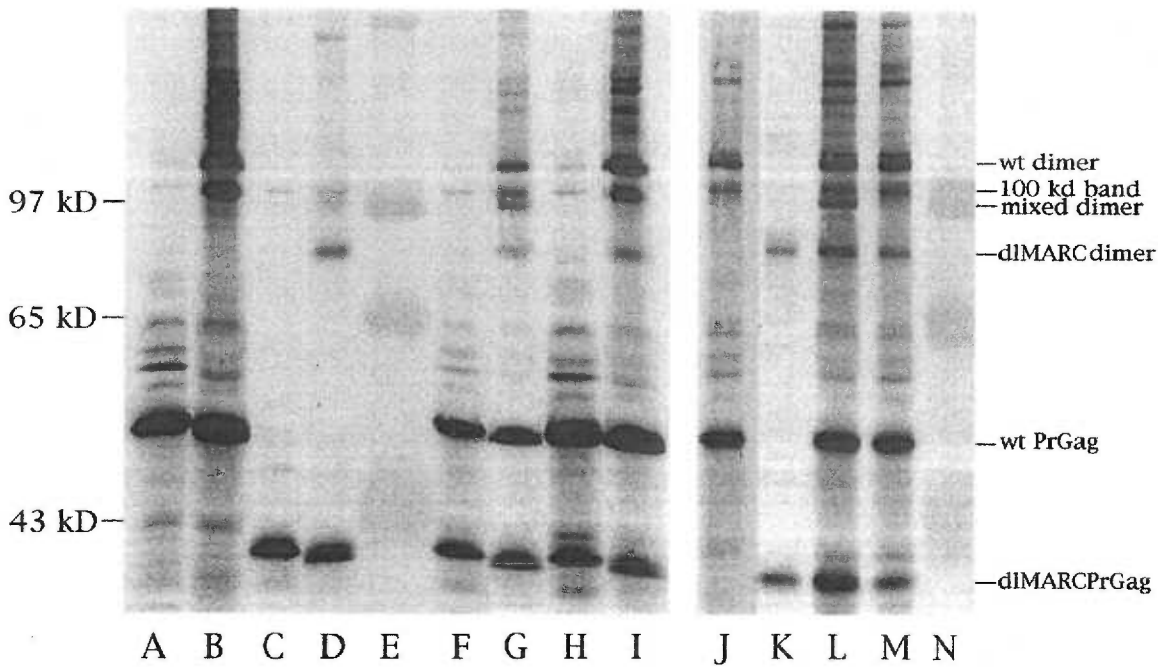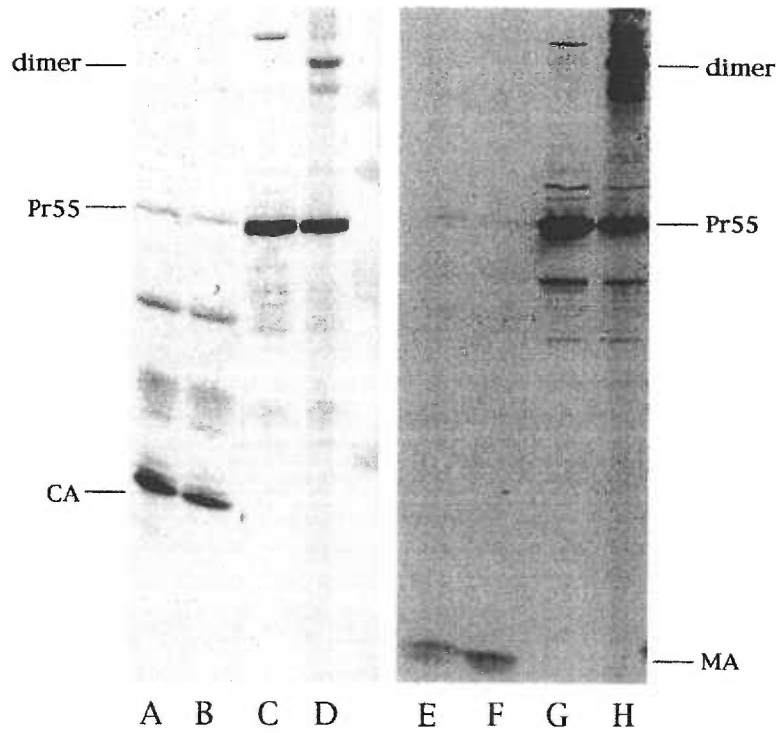

Figure 4. Composition of crosslinked Gag species. Gag proteins from virus particles that were mock treated (lanes A,C,F,H) or BMH-treated (lanes B,D,G,I,J,K,L,M,N) were electrophoresed and immunodetected with an anti-CA antibody as described in Figure 3. Virus particles derived from COS7 cells transfected with the following constructs: PR-, lanes A,B,J; dlMARC PR-, lanes C,D,K; particles from cells co-transfected with PR- and dlMARC PR-, lanes F,G,L; particles mixed from cells singly transfected with PR- and dlMARC PR- constructs, lanes H,I,M. Standard molecular weight sizes are indicated on the left and derive from marker lanes E and N. Wt and dlMARC PrGag proteins, homodimers, mixed dimers and the 100 kd band of unknown origin are indicated on the right.
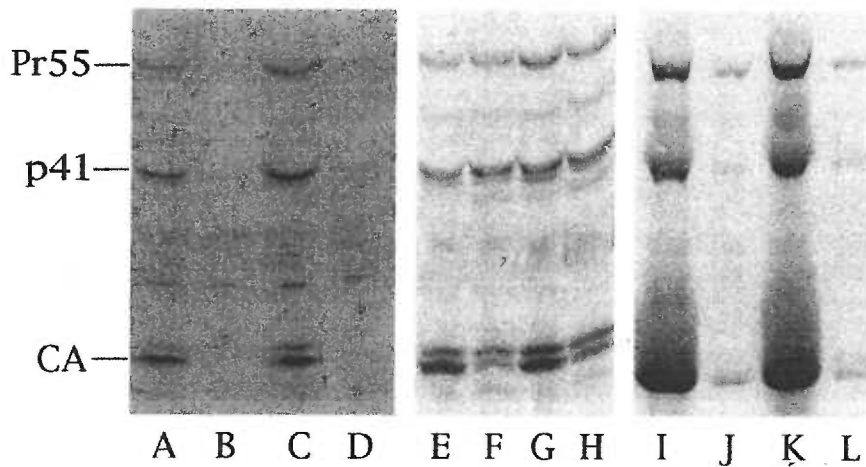
Figure 5. Expression, release, and antibody reactivity of mutant Gag proteins. Cell lysates (lanes A-H) and virus particles from media supernatants (lanes I-L) were collected from cells transfected with wt (lanes A,E,I), C330S (lanes C,G,K) and C350S (lanes B,D,F,H,J,L) HIVgpt constructs as described in the Materials and Methods. Samples were subjected to SDS-PAGE, electroblotted, and Gag proteins were immunodetected with anti-CA antibodies Hy5001 (lanes A-D) or Hy183 (lanes E-L). Pr55, p41 and CA proteins are as indicated.
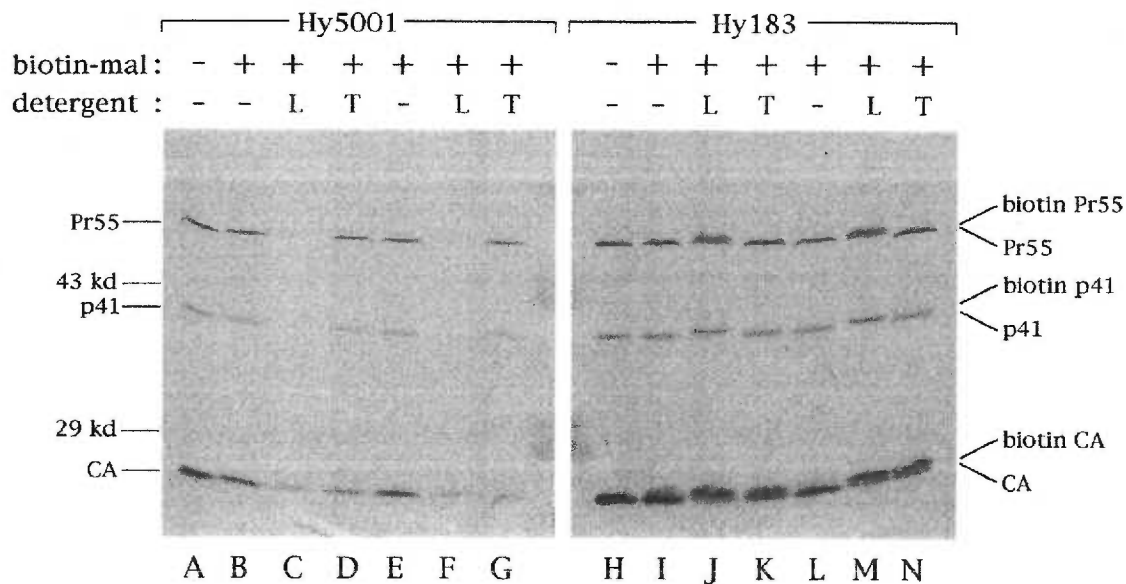

Figure 6. Biotin-maleimide reactivity of HIV-1 capsid proteins. Wild-type (wt) HIVgpt virus particles were mock treated (lanes A,H; indicated by -) or biotin maleimide treated (lanes B-G, I-N; indicated by +) in the absence of detergent (-; lanes B,E,I,L), or in the presence of 0.25% Triton X-100 (lanes D,G,K,N; indicated by T) or 0.25% LDS (lanes C,F,J,M; indicated by L). After treatments, proteins in samples were separated by SDS-PAGE, electroblotted and immunodetected with anti-CA antibodies Hy5001 (lanes A-G) or Hy183 (lanes H-N). Note that Hy5001 does not recognize C350S mutant Gag proteins, while Hy183 does (see Figure 5). Marker protein sizes, Gag proteins Pr55, p41, and CA, and biotinylated species are as indicated.
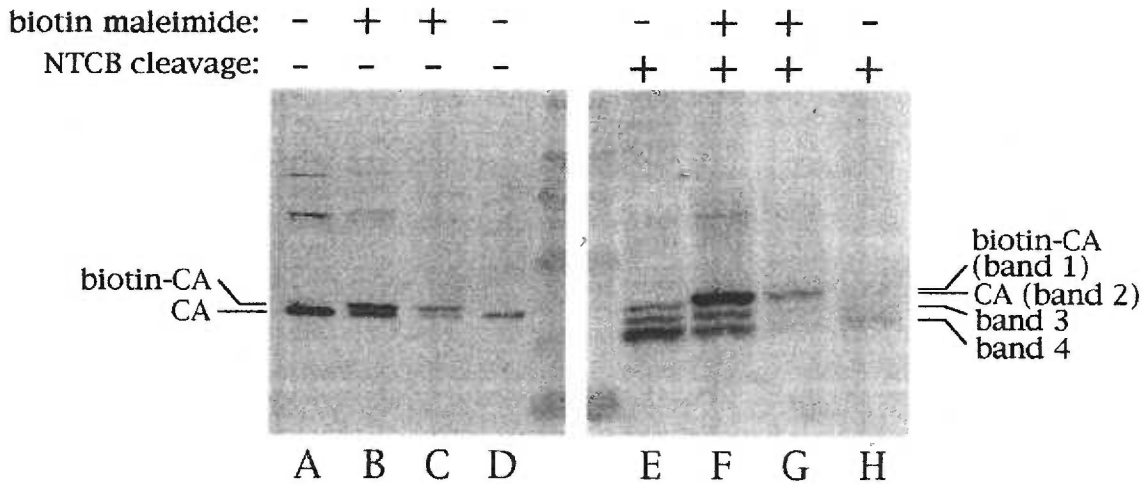
Figure 7. Cleavage of HIV-1 capsid proteins at cysteine residues. Wild-type (wt; lanes A,B,E,F) or dlMARC (lanes C,D,G,H) HIVgpt particles were mock treated (lanes A,D,E,H) or treated with biotin-maleimide (lanes B,C,F,G), after which proteins were processed for electrophoresis (lanes A-D), or NTCB-cleaved at non-biotinylated cysteines and processed for electrophoresis. Electrophoretically separated proteins were electroblotted and immunodetected with anti-CA antibody Hy183. Four different capsid derived proteins are indicated: band 1 (biotin-CA), band 2 (CA), band 3, and band 4.
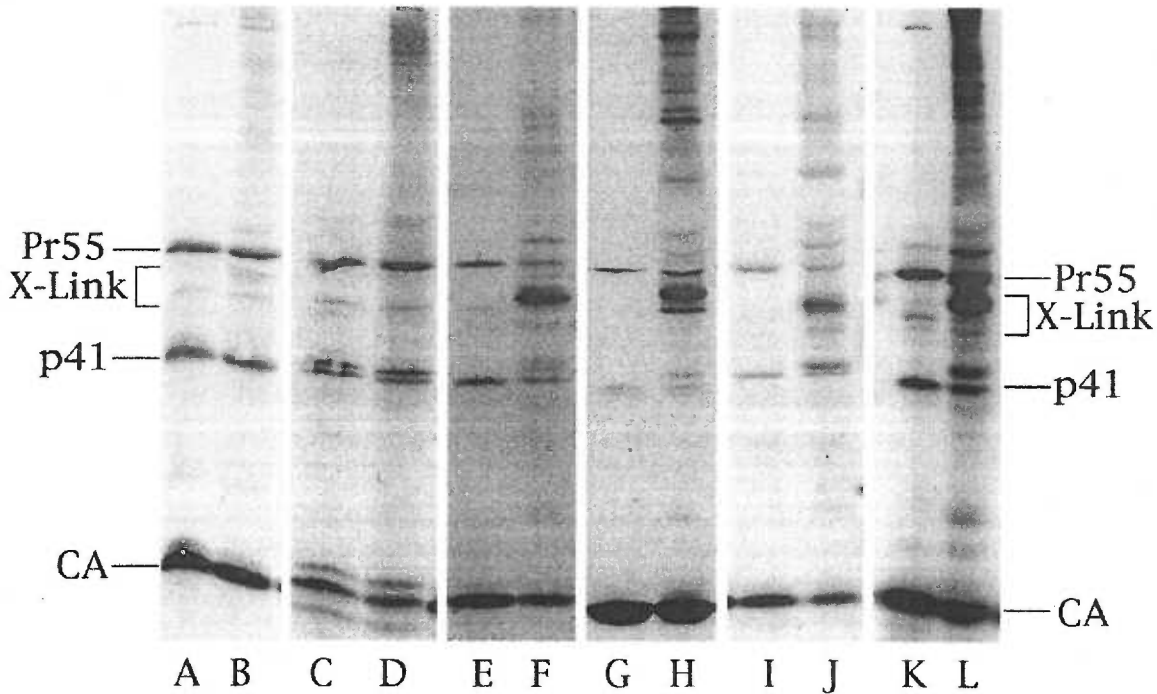
Figure 8. Chemical crosslinking of HIV mutants. Wt (lanes A,B) and mutant T242C (lanes C,D), R294C (lanes E,F), V297C (lanes G,H), Y301C (lanes I,J), and C330S (lanes K,L) HIVgpt particles were mock treated (lanes A,C,E,G,I,K) or treated with BMH (lanes B,D,F,H,J,L) to crosslink neighboring cysteine residues. Samples were subjected to SDS-PAGE, electroblotted and Gag proteins were immunodetected with anti-CA antibody Hy183. HIV Gag proteins Pr55, p41, and CA are indicated, as are putative CA homodimer bands.

Chapter 4

STRUCTURAL ANALYSIS OF MEMBRANE-BOUND RETROVIRUS

CAPSID PROTEINS

Eric Barklis[*+1], Jason McDermott[+1], Stephan Wilkens[+2], Eric Schabtach[3], M.F.

Schmid[4], Stephen Fuller[5], Sonya Karanjia[1], Zachary Love[1], Russell Jones[6],

Yuanjui Rui[7], Xiumin Zhao[7], and David Thompson[+7]

[1]Vollum Institute and Department of Microbiology, Oregon Health Sciences University
[2]Institute of Molecular Biology, University of Oregon, Eugene, OR, 97403 USA
[3]Institute of Neuroscience, University of Oregon, Eugene, OR, 97403, USA
[4]Verna and Marrs McLean Department of Biochemistry and W.M.Keck Center for Computational
   Biology, Baylor College of Medicine, Houston, TX 77030, USA
[5]Structural Biology Programme, European Molecular Biology Laboratory, D-69012 Heidelberg,
   Federal Republic of Germany
[6]Department of Pathology, Oregon Health Sciences University, Portland, Oregon, 97201 USA
[7]Department of Chemistry, Purdue University, West Lafayette, IN, 47907 USA

* Corresponding author
+ Authors contributed equally to the completion of this work

ABSTRACT

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, 1,2-di-O-hexadecyl-*sn*-glycero-3-(1'-(2"-R-hydroxy-3"-N-(5-amino-1-carboxypentyl)-iminodiacetic acid) propyl ether) (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, and the corresponding diffraction patterns could be indexed as either orthorhombic (C222) or hexagonal (p6), with reflections out to 9.5Å 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 which are either circular with a diameter of 19.2Å, or roughly triangular (length=28.0Å, width=23.2Å). 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.

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; Faecke 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 protease (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 Pr8^{ag} 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; Faecke et al., 1993; Wang et al., 1993). However, the central 80-90% of HIV MA can be deleted without blocking assembly of virus particles (Faecke et al., 1993) that are conditionally infectious (Wang et al., 1993). At or near the C-termini of lentivirus or C-type retrovirus Pr8^{ag} 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 Pr8^{ag} proteins, and assembly of Gag-Pol and Gag fusion proteins into virions (Jones et al., 1990; Hansen et al., 1990; Mammano et al., 1994; Chazal et al., 1994; Wang et al., 1994; Hansen and Barklis, 1995; Craven et al., 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; Hansen and Barklis, 1995; Craven et al., 1995; Clish et al., 1996; McDermott et al., 1996). In addition to the three conserved mature Gag proteins (MA, CA, 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 Pr8^{ag}: 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; Griffith 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 icosohedral symmetry, which would be a great help in image reconstruction (Choi et al., 1991; Fuller et al., 1995; Fuller et al., 1996). In the absense of homogeneous, *in vivo*-derived virus preparations, several groups have undertaken the analysis of virus-like particles formed *in vitro* from Prg^{ag} proteins (Nermut et al., 1994; Klikova et al., 1995) or their derivatives (Ehrlich et al., 1992; Campbell and Vogt, 1995; McDermott and 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 Prg^{ag} 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; Kubalek et al., 1995; Ng et al., 1995; Dietrich et al., 1995,1996; 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 moeity of the Gag protein, and have examined crystalline arrays of a his-tagged M-MuLV capsid proteins (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 to 9.5 Å resolution, and can be classified as either orthorhombic (a=79.6 Å, b=137.5 Å, γ=90°) or hexagonal (a=b=79.7 Å, γ=60°). Image reconstruction shows a cage-like array of proteins, similar to that seen for HIV-1 Prg^{ag} 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 1,2-di-O-hexadecyl-*sn*-glycero-3-(1'-(2'-R-hydroxy-3''-N-(5-amino-1-carboxypentyl) iminodiacetic acid) propyl ether) (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 infrared and nuclear magnetic resonance 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 ul 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 used with control lipids OG (octyl glucoside) and DHP (dihexadecyl phosphate), Ni^{+2} 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 Ni^{+2}, inhibiting DHGN-mediated extraction of nickel.

The above experiments indicated that DHGN chelates Ni^{+2}, 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 monolayers 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 the 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 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 lacey 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 1). All conjugates showed total gold counts of 100-300 particles/nl, and PC monolayer binding of 3-32 particles/1000 nm^2. 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 2). As expected, PC only monolayers (Table 2; 0% DHGN) and monolayers of PC plus DHGN that had not been charged with Ni^{+2} (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 (6x10^8 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 2). However, certain divalent cations seemed to interfere with the DHGN-his-tag interaction. In particular, $CaCl_2$ and $FeCl_2$ reduced binding levels 4- to 5-fold, and 5 mM $NiCl_2$ in the subphase eliminated binding in a more dramatic fashion even than EDTA. Finally, as might be expected from their effects on metal chelate chromatography (Hochuli et al., 1987), dithiothreotol (DTT) and B-mercaptoethanol (B-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 (Jones et al., 1990; Hansen et al., 1990; Hansen et al., 1993; Wang et al., 1993,1994; Mammano et al., 1994; Hansen and Barklis, 1995; Craven et al., 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 *E. coli*, his-MoCA is a 35 kd protein, composed of the 30 kd M-MuLV capsid domain, a his-tag-containing 25 residue N-terminal leader sequence, and a vector-derived, 16 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 kd (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 was for the 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 greater than 1

mg/ml, and never formed on PC-only monolayers or on DHGN-containing

monolayers in which the DHGN had not been prebound 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 approximately 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 (see Materials and Methods) retained his-MoCA

array order. Consequently, four negatives (#717, #719, #255, #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 approximately 9.5Å 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$1.7 Å, γ =89.72$\pm$0.89°, or a=79.4$\pm$0.6 Å,

b=80.0$\pm$1.3 Å, γ =59.89$\pm$0.73°, depending on the choice of indexing (Table 3).

After unbending and 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 23Å

resolution, using a primitive (p1) unit cell, while film #719, gave a phase residual of 21.0° at 14Å 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 less than 25° (Table 3), while the best merging for other space groups consistently gave phase residuals of >50°. 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 approximately 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,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 interprotein 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 envision 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 Prg^{ag} 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 wt M-MuLV *gag* and *pol* genes; pXM2453T, which expresses an unprocessed, protease-minus (PR-) Gag protein; pXMGPEΔMA+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, pXMGPEΔMA+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, three 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 Pr8^{ag} 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 approximately 40 kDa and 57 kd respectively (lanes B, F). With the wt virus pellet sample, three reactive bands (Pr8^{ag}, 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, approximately 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 possiblities, we performed immunofluorescent localization of Gag proteins in Cos7 cells transfected with either pXM2453T or pXMpetMoCA, expressing wt and NC-deleted Pr8^{ag} proteins, respectively. Interestingly, the localization patterns of the two proteins

showed significant differences (see Figure 9). The wt Pr8^{ag} proteins

demonstrated a punctate staining pattern over the entire area of staining cells

(Figure 9, A-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 9, I-L 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; Kubalek et al., 1995;

Dietrich et al., 1995,1996; 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 nitrilotriacetic acid (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 2). 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 3). 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 preincubated 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, activating transcription factor), 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 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 can not 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 $Pr8^{ag}$ proteins (Jones et al., 1990; Hansen et al., 1990; Hansen and Barklis, 1995). The his-MoCA protein lacks the M-MuLV NC domain, and consequently can not 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 3). 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 to consist 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=85Å, cross-sectional diameter=34Å, with cage hole-to-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 approximately 19.2 Å and a hole-to-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 cm^3/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: (a) building of hexamer rings from monomers via (5-3) contacts, followed by network assembly through 1-1 contacts; or (b) homodimer formation at the 1-1 faces, followed by network assembly mediated by 5-3 contacts. We prefer the latter (b) pathway, because purified HIV and M-MuLV capsid proteins form stable dimers in solution, and retrovirus *gag* crosslinking 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 (a).

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 presumes 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 icosohedral-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 icosohedral, but do display an ordered arrangement of Gag proteins with unit cell parameters similar to those reported here (Fuller, Wilk, Krausslich, and Vogt; manuscript 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 (Jones et al., 1990; Hansen et al., 1990,1993; 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 NC domains in immature M-MuLV and HIV-1 particles can be crosslinked 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; 1,2-di-O-hexadecyl-*sn*-glycero-3-(1'-(2"-R-hydroxy-3"-N-(5-amino-1-carboxypentyl) iminodiacetic acid) propyl ether)) 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--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 mmole) 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 mmole) 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 infrared (IR) and nuclear magnetic resonance (NMR; chemical shifts reported in ppm using $CDCl_3$ as solvent) spectroscopy with results as follows: IR (compound 5): 2955, 2918, 2850, 1467, 1377, 1259, and 1118 cm^{-1} ; ^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); ^{13}C 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^{-1} ; ^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, dihexadecyl phosphate (DHP; Aldrich), and octyl glucoside (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 micromolar solutions in distilled water, and 20 microliter samples were flash-atomized at 2500 degrees 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 microns. Under these conditions, chloroform, lipid solutions, and 50 mM EDTA (ethylenediamine-tetraacetic acid) 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 overlayered with an equal volume of 20 micromolar nickel sulfate in distilled water. Samples were rocked 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 microliter 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 (Ni^{+2}) phase to a final concentration of 50 mM prior to mixing with the DHGN-chloroform phase.

Protein purification and analysis. The M-MuLV capsid (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 BamHI site of pET15B (Novagen). The cassette was constructed by PCR-generation of a BamHI site at the amino-terminus of the CA coding region and addition of a BamHI linker at the carboxy-terminal MscI 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 nt 1266 (Shinnick et al., 1981), and TTG/GCGGATCC, where the bold G is viral nt 2055. When expressed in bacteria, the his-MoCA protein has an amino-terminal leader sequence of MGSSH HHHHH SSGLV 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 (Jones et al., 1990; Hansen et al., 1990, 1993; 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 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, PMSF (5 mM final concentration) and the samples were mixed with one volume of 2x 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% polyacrylamide gel electrophoresis (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-ClaI pet15B-MoCa fragment (encoding the CA C-terminal coding region) in
place of the M-MuLV-derived XhoI-ClaI 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 nt 645 and nt 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 immunoglobulin G (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 40x 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 SO163 film, the search mode was performed at 1,000-5,000x magnification, focusing was at 100,000-150,000x, and micrographs were taken at 30,000-60,000x, 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 (Uzgiris and Kornberg, 1983; Darst et al., 1991a,b; Celia et al., 1994) with a variety of modifications. For our standard lipid mix, a working (1x) stock of 200 µg/ml phosphatidyl choline (Avanti Polar Lipids) plus 50 µg/ml nickel-charged DHGN was made up in 1:1 hexane:chloroform from 10x stocks. In some control experiments, the final proportion of DHGN in the 1x lipid mix varied from 0-50%, with the total lipid weight remaining constant. The 10x nickel-charged DHGN stock was prepared by mixing 2 µl of 100 mM nickel chloride (1 mM final) with 200 ul 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 30 sec at 20 min intervals. Note that although all lipids were stored at -80°C under nitrogen gas, we found that working 1x stocks went bad for unknown reasons after 3-12 weeks, as observed by poor protein crystallization.

To set up monolayer incubations, 5 ul of 0.2-2.0 mg/ml MoCa protein in distilled water was mixed with an equal volume of 2x subphase buffer to yield a final mix of 0.1-1.0 mg/ml protein in 1x 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 2x buffers (see Table 2). 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 1x lipid mix (pre-loaded with Ni^{+2}) 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 distilled water) in plastic 150 mm petri plates. Lids were placed on plates and the plates were sealed with strips of parafilm prior to overnight incubations, which usually were at 30 degrees C, although room temperature incubations also were successful.

After incubations, arrays were transfered to lacey carbon EM grids (Ted Pella 300 mesh copper "Netmesh," cat. no. 01883; or lacey 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 sec. For samples to be stained, grids with adherent arrays were placed 30-45 sec on 50-100 µl water drops, wicked from the side, stained 30 sec 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, transfered 30-45 sec onto water drops, carefully lifted with forceps, transfered 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), bovine serum albumin (BSA; Sigma), purified histidine-tagged activating transcription factor (his-ATF), and purified histidine -tagged cAMP responsive element mediator protein (his-CREM). Using 10 nm gold colloid (ICN), protein concentration and pH optima for conjugations

were determined in small scale reactions (50 μl colloid, 10 μl protein in distilled water) by observation of color changes after addition of 10 μl 10% NaCl. For large scale conjugations, 5 ml colloid (pH optimized) was vortexed 10 sec with the optimized amount of protein in 0.2 ml of distilled water, and incubated 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; 2,000 x 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 x g (Beckman SW50.1; 25,000 rpm), 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 ul 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 four 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 250,000x (1.06 micron^2) 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 2x subphase buffer (see above), overlayed with 1x lipid mix, incubated overnight at 30°C under conditions described above, lifted onto lacey EM grids, washed once for 30 sec on a 50 ul drop of distilled water, incubated 30 sec 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 91,000x (4500 nm^2) were counted to calculate the average number of particles bound per 1000 nm^2 .

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 717,719), 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 A/pixel for negative 717; 4.13 A/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 FFTRANS (Henderson et al., 1990). Transforms were represented as diffraction patterns using SPECTRA (Schmid et al., 1993), indexed by hand as either orthorhombic, or with $\gamma*=60°$, 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 contrast transfer function (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 ORIGTILTB. 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, FFTRANS, 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

REFERENCES

Avila-Sakar, A.J., and Chiu, W. (1996). Visualization of _ 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 A 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 A resolution.*Cell* **66**, 121-128.

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.

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.

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. *Meth. Enzymology* **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. *Meth. Enzymology* **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., 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 icosohedral particles: The uncommon line. *J. Structural Biology* **116**, 48-55.

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.

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., 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. of Chromatography* **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-B-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 Fujiyhoshi (1994). Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367**, 614-621.

Leis, J., Baltimore, D., Bishop, J., Coffin, J., Fleissner, E., Goff, S., Oroszlan, S., Robinson, H., Skalka, A., Temin, H., and Vogt, V. (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, Cold Spring Harbor, NY.

Matthews, S., Barlow, P., Boyd, J., Barton, G., Russell, R., Mills, H., Cunningham, M., Meyers, N., Burns, N., Clark, N., Kingsman, S., Kingsman, A., and Campbell, I. (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., Kovari, L., Prongay, A., Keller, W., Gitti, R., Lee, B., Gorbalenya, A., Tong, L., McClure, J., Ehrlich, L., Summers, M., Carter, C., and Rossman, M. (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 ^{1}H 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., Luftif, 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. of Microscopy* **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., Vander 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., Hammarskjold, 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 phopholipids. 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-B-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, 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.

TABLES AND FIGURES

Table 1. 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 histiding-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 ul of 1:10 dilutions of gold-conjugate stocks onto carbon-formvar grids, and counting all gold particles in 200 EM fields at 250,000x magnification (1.06 micron^2) 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 phosphatidyl choline (PC) only, or PC plus nickel-charged DHGN(PC+DHGN), as described in the Materials and Methods. After incubations, monolayers were lifted onto lacey carbon grids, rinsed 30 sec in 1% glucose, dried, and viewed and photographed. Binding was quantitated by counting particles on hole-free, carbon support-free monolayer areas in 20 fields at 91,000x magnification to obtain particles bound per 1000 nm^2.

Table 2. Variables affecting histidine-tagged gold particle binding to lipid monolayers.

lipid/salt variations		binding as % standard conditions
0 %	DHGN	0
1 %	DHGN	68
20%	DHGN*	100
50 %	DHGN	16
20 %	DHGN (no nickel)	7
0 mM	NaCl	58
250 mM	NaCl*	100
500 mM	NaCl	119

ion/misc. variations	binding as % standard conditions
No additions*	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 B-Me	0
1 mM B-Me	9

pH/buffer variations	binding as % standard conditions
pH9.0, phosphate	90
pH8.5, phosphate	142
pH7.8, phosphate*	100
pH7.4, phosphate	112
pH7.0, phosphate	76
pH6.5, phosphate	96
pH6.0, phosphate	135
pH5.5, phosphate	133
pH5.0, phosphate	130
pH7.8, phosphate*	100
pH7.8, Tris-HCl	118
pH7.8, Hepes	70

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 the Materials and Methods, and were compared with results from standard condition incubations (incubations marked with asterisks) 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.

TABLE 3. Unit cell parameters for his-MoCA 2D arrays.

FILM	717	719	255	260
ORTHORHOMBIC* UNIT CELL DIMENSIONS	a = 79.6 Å b =137.3 Å a/b= 0.58 γ = 89.26°	a = 78.9 Å b =138.5 Å a/b=0.57 γ = 90.48°	a = 79.1 Å b =138.9 Å a/b=0.57 γ = 90.45°	a = 80.8 Å b =135.1 Å a/b= 0.60 γ = 90.00°
HEXAGONAL* UNIT CELL DIMENSIONS	a = 80.0 Å b = 79.1 Å a/b=1.01 γ = 60.22°	a = 80.0 Å b = 79.1 Å a/b= 1.01 γ = 60.22°	a = 78.6 Å b = 80.3 Å a/b= 0.98 γ = 58.8°	a = 79.0 Å b = 81.8 Å a/b= 0.97 γ = 60.31°
PHASE RESIDUAL FOR P1 MERGE	----	23A= 10.2° 17A= 15.3° 14A= 21.0°	23A= 18.9°	23A= 19.4°
PHASE RESIDUAL FOR P2 MERGE	----	23A= 10.8° 17A= 17.4° 14A= 21.6°	23A= 19.8°	23A= 19.9°
PHASE RESIDUAL FOR C222 MERGE	----	27A= 13.2°	27A= 15.2°	27A= 14.9°
PHASE RESIDUAL FOR P3 MERGE	----	23A= 16.8° 17A= 18.0° 14A= 19.9°	23A= 22.5°	23A= 14.4°
PHASE RESIDUAL FOR P6 MERGE	----	23A= 15.5° 17A= 16.3° 14A= 18.4°	23A= 23.6°	23A= 17.8°

Electron Micrographs of unstained his-MoCA arrays were taken on the University of Oregon CM12 at 60,000x magnification and scanned on the Baylor flat bed scanning densitometer (Films 717,719), or on the EMBL CM200 and CCD-scanned (Films 255,260). 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*; γ approximately 90°) or roughly hexagonal (hexagonal*; γ approximately 60°): 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 the Materials and Methods. Film 717 was chosen as a reference film, phase origin-centered, and merging was performed using ORIGTILT 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 greater than 50° at 23 or 27Å resolution.

Model for forming histidine-tagged protein arrays
on a monolayer of nickel-chelating lipids at the air-buffer interface
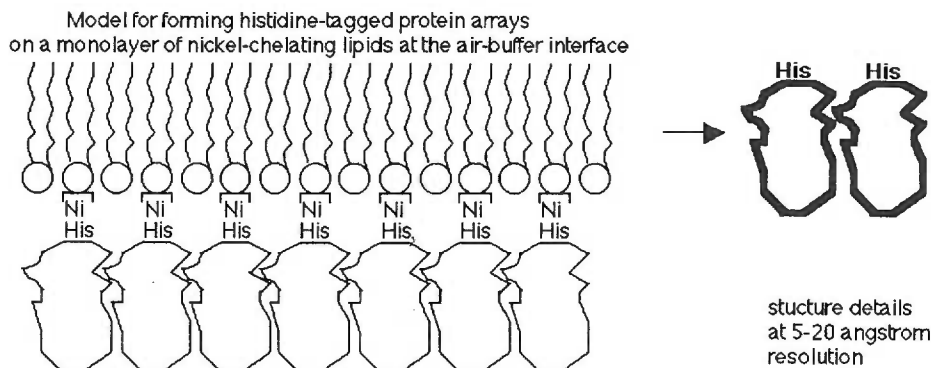
stucture details
at 5-20 angstrom
resolution

Figure 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 histidine-tagged (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 a nitrilotriacetic acid (NTA; Hochuli et al., 1987) headgroup. Analysis of two-dimensional (2D) protein arrays, following previously established methods (Unwin and Henderson, 1975; Fuller et al., 1979; Schmid et al., 1993), involves adhering arrays to electron microscope (EM) grids; photography of stained or unstained protein crystals at 20,000-100,000x 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 angstrom 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).
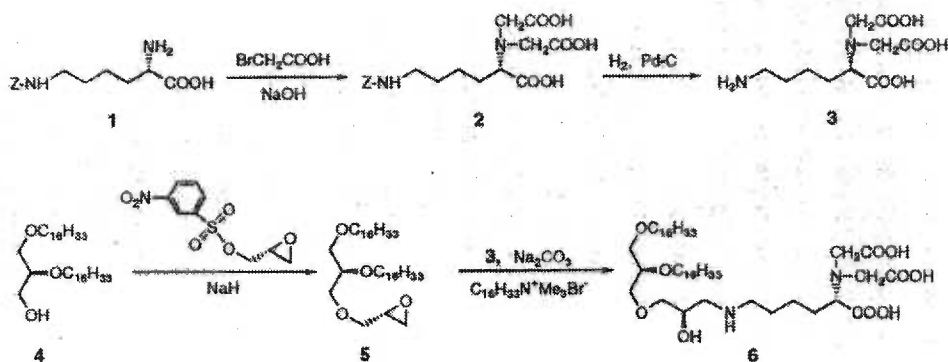

Figure 2. Synthesis of the nickel-chelating lipid, DHGN.

DHGN (compound 6; 1,2-di-O-hexadecyl-sn-glycero-3-(1'-(2"-R-hydroxy-3"-

N-(5-amino-1-carboxypentyl) iminodiacetic acid) propyl ether) was prepared from starting materials N—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-nitrobenzenesufonate. 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.

Figure 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 micromolar lipid solutions in 0.1 ml of chloroform were overlayed with an equal volume of 20 micromolar 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 microliter 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. Lipids were OG (octylglucoside), DHP (dihexadecyl phosphate), and DHGN (1,2-di-O-hexadecyl-sn-glycero-3-(1'-(2"-R-hydroxy-3"-N-(5-amino-1-carboxypentyl) iminodiacetic acid) propyl ether). In one experiment, EDTA (ethylenediamine-tetraacetic acid) was added to the aqueous phase to a final concentration of 50 mM. Neither EDTA, not any of the stock lipid solutions showed any nickel content above backround prior to the phase fractionation experiments.
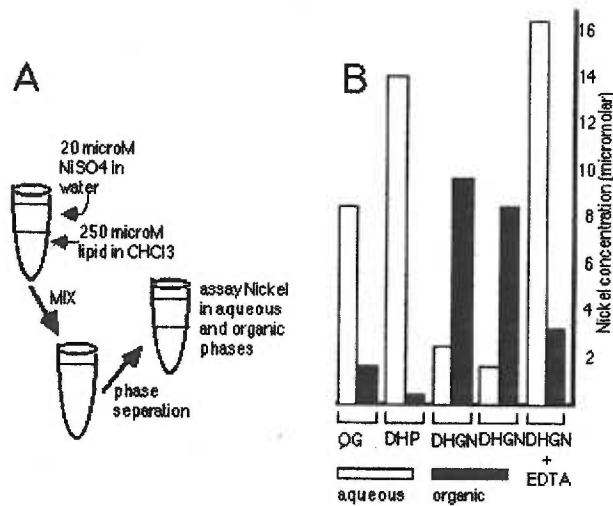
Figure 4. Protein-gold conjugate binding to lipid monolayers.
Histidine-tagged CREM proteins (his-CREM), conjugated to 20 nm gold particles as described in the Materials and Methods, were incubated under standard conditions beneath lipid monolayers composed of PC only (panel A) or PC plus nickel-charged DHGN (panel B). After incubations, monolayers were lifted onto lacey carbon grids, rinsed 30 seconds in 1% glucose, dried, and viewed and photographed on the Phillips 301 TEM. Carbon
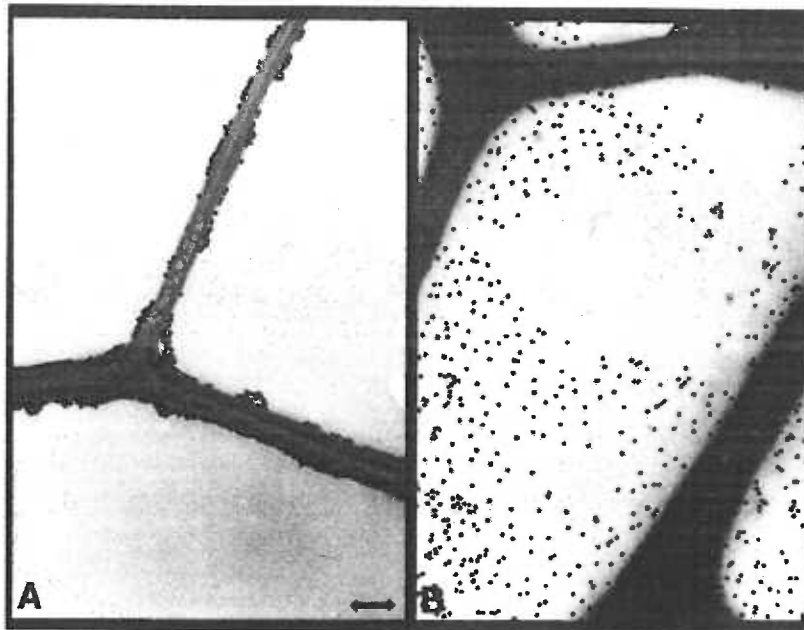
supports are clearly visible as the broad dark lines; monolayers are visible with occassional holes caused by beam damage, while gold particles appear as black dots. Note that with either monolayers, gold particles adhere or are washed to the carbon support regions, however, gold-conjugated his-CREM particles are present on the monolayer only when DHGN is present (panel B). The black bar in panel A indicates 10 nm.

Figure 5. M-MuLV his-MoCA two dimensional arrays.
His-MoCA proteins were incubated beneath a PC+DHGN monolayer, lifted onto a lacey EM grid, washed, uranyl acetate-stained, and photographed on the Philips CM12. The lacey carbon support appears as thick grey 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 (60,000x), and its length corresponds to 417 nm.
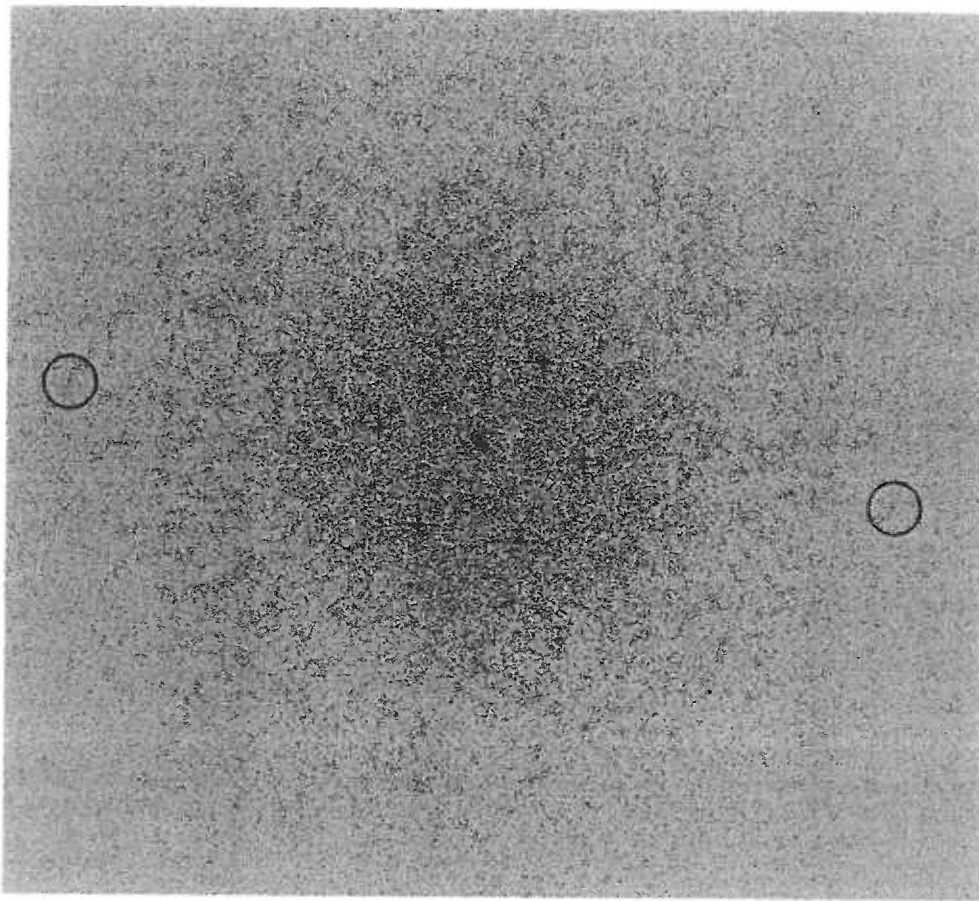
Figure 6. Calculated diffraction pattern for an unstained his-MoCA 2D crystal. An unstained
his-MoCA crystalline area from a cryo sample frozen in vitreous ice (negative #717) was scanned
at 2.67Å/pixel, transformed, and displayed as a diffraction pattern, which appears hexagonal or
orthorhombic, with 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 $\gamma^* = 60°$, or the 2,0; 1,3; -1,3; -2,0; -1,-3; and 1,-3 reflections if $\gamma = 90°$. The two faint
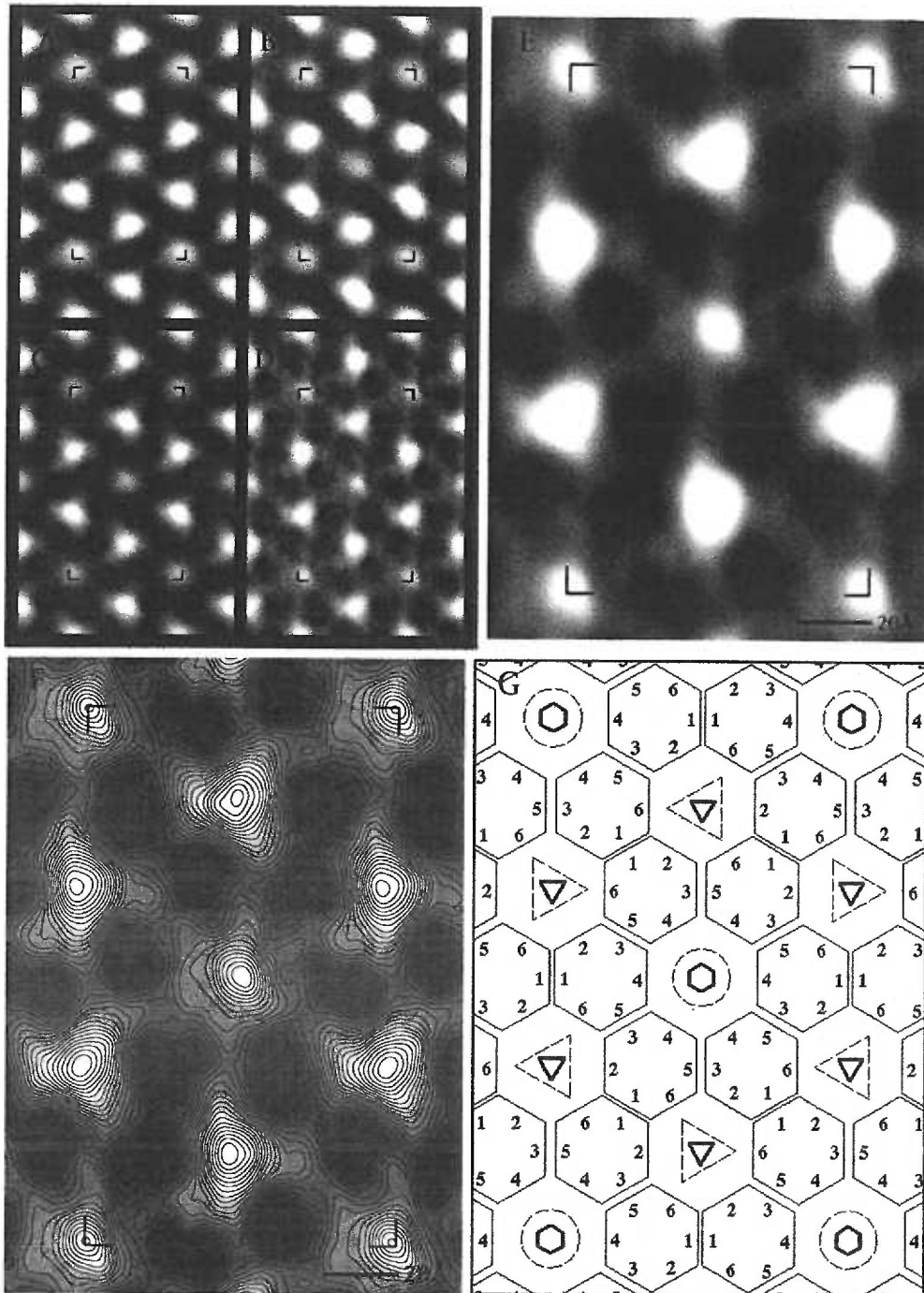circled reflections are at approximately 9.5Å resolution.

Figure 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 79.2Åx137.5Å, and 20 Å marker bars are indicated. Panel E is a larger scale version of panel D, panel F is a contour map of panel E, and panel G shows a schematic interpretation of the projections shown in panels 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 six-fold and three-fold crystallographic axes, respectively. Additional two-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 asymmetry dimer leads to the same symmetry in the final network.
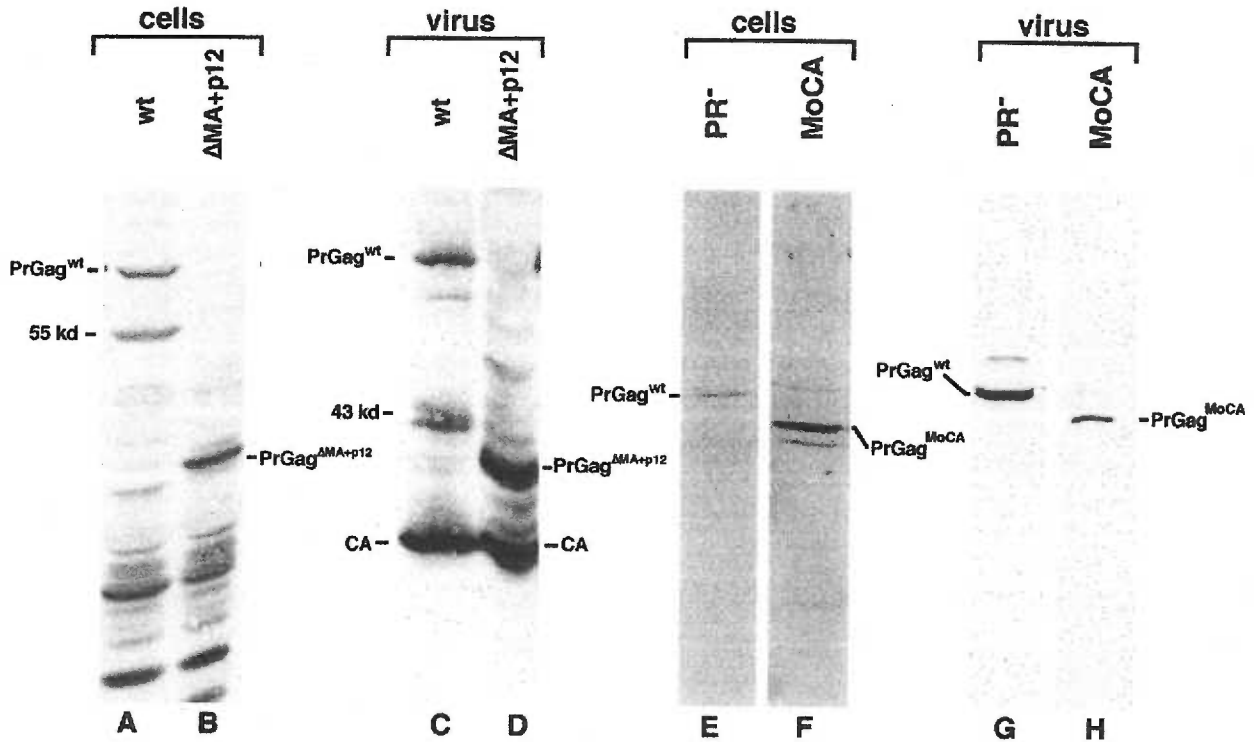
Figure 8. Assembly and release of Gag proteins from transfected mammalian cells. Cos7 tissue culture cells were transfected with the M-MuLV Gag protein expression constructs pXMGPE (lanes A,C), pXMΔMA+p12 (lanes B,D), pXM2453T (lanes E,G), and pXMpetMoCA (lanes F,H). Three days post-transfection, cell lysate (lanes A,B,E,F) and virus pellet (lanes C,D,G,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 kd and 43 kd in lanes A and C also are 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; pXM_MA+p12, 3.3; pXM2453T, 5.0; pXMpetMoCA, 0.2.
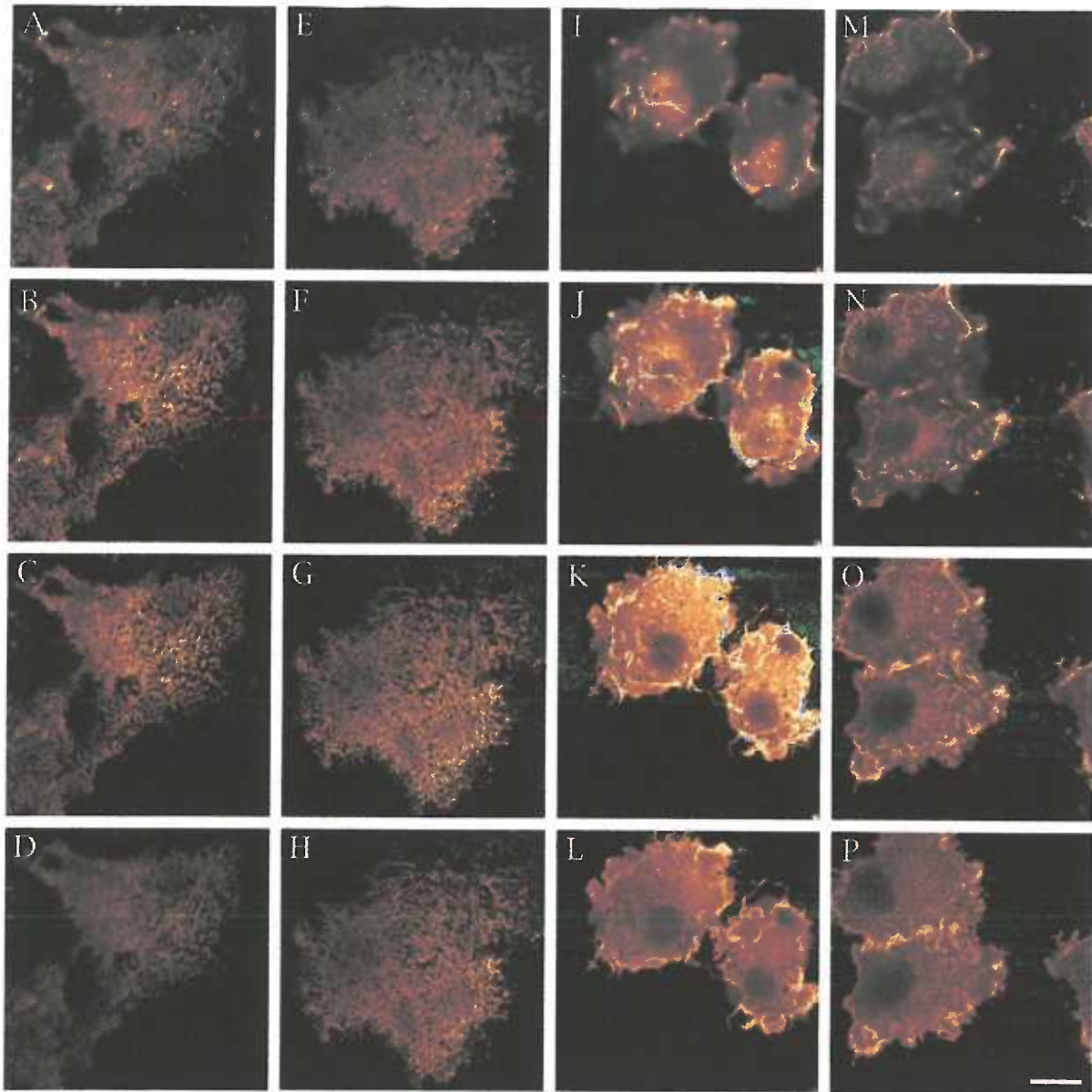
Figure 9. Immunofluorescent localization of Gag proteins in mammalian cells. Cos7 cells on coverslips were transfected with either pXM2453T (panels A-H) or pXMpetMoCA (panels I-P), grown 3 days, and processed for immunofluorescent localization of Pr^{gag} proteins using an anti-p12 monoclonal antibody (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 40x oil immersion objective and rhodamine filters. Each set of images (A-D, E-H, I-L, M-P) was taken in four 1.25 micron slices from the apical (A,E,I,M) to basolateral (D,H,L,P) surfaces. Fluorescence voltage settings for panels A-H were 950, and 760 for panels I-P. The size bar in panel P indicates 50 microns.

Chapter 5

ORGANIZATION OF HIV-1 CAPSID PROTEINS ON A LIPID MONOLAYER

Eric Barklis[1*], Jason McDermott[1], Stephan Wilkens[2], Stephen Fuller[3], and

David Thompson[4]

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

[2] Institute of Molecular Biology, University of Oregon, Eugene, Oregon, 97403, USA

[3] Structural Biology Programme, European Molecular Biology Laboratory, D-69012 Heidelberg, Germany

[4] Department of Chemistry, Purdue University, West Lafayette, Indiana, 47907, USA

SUMMARY

In an *in vitro* system that mimics the assembly of immature HIV particles, ordered arrays of HIV-1 capsid (CA) proteins encoded by the viral *gag* gene have been obtained by incubation of histidine-tagged capsid proteins (his-HIVCA) beneath lipid monolayers containing the nickel-chelating lipid, DHGN. The membrane-bound his-HIVCA proteins formed small crystalline arrays of primitive (p1) unit cells with dimensions of a=74.2Å, b=126.2Å, γ=89.3°. The image-analyzed two-dimensional (2D) projection of his-HIVCA assemblies shows a cage-like lattice, consisting of hexamer and trimer units, surrounding protein-free cage holes. The hexamer coordinated cage holes of 26.3Å diameter are spaced at 74.2Å intervals: these distances, and the hexamer-trimer arrangement are consistent with previous, lower resolution studies on immature HIV-1 virus particles produced *in vivo*. Additionally, HIV-1 matrix (MA) protein trimer unit structures align to the his-HIVCA trimer units such that residues previously shown to interact with the HIV-1 gp120/gp41 Envelope (Env) protein complex are oriented towards the hexamer cage holes. Our results form a bridge between results from conventional methods for the analysis of HIV particle structure.

INTRODUCTION

The products of human immunodeficiency virus (HIV-1) assembly are 125 nm diameter, enveloped, immature and mature virus particles[1]. The principal internal protein component of HIV-1 virions is encoded by the viral *gag* gene. Immature HIV-1 particles are composed of unprocessed 55 kd Gag proteins

(Pr55Gag), and appear to have an electron dense layer of material juxtaposed to the inner faces of their lipid envelopes[1,2]. However, during or just after virus particle release from cells, Pr55Gag proteins are processed by the HIV-1 protease into the major mature Gag proteins matrix (MA), capsid (CA), nucleocapsid (NC) and p6. Proteolytically processed HIV-1 particles adopt a mature morphology, in which electron dense material reorganizes into a central cone- or rod-shaped structure[1].

Complete or partial structures for the MA, NC and CA Gag proteins have been determined in nuclear magnetic resonance spectroscopy (NMR) and X-ray crystallographic studies[3-10]. The matrix protein, which interacts with the HIV-1 envelope (Env) protein complex (SU/TM, or gp120/gp41) and serves a membrane binding function, forms characteristic trimers in three different crystal forms[5]. The NC domain possesses two cys-his finger motifs which coordinate zinc ions, and are important in RNA binding[10]. The capsid domain, which is crucial to virus particle assembly[11], has proven more difficult. However, partial structures of HIV-1 CA have shown that it is composed primarily of alpha helices[6-9], an unexpected contrast with the B-barrel or jellyroll structures which constitute the capsid proteins of a number of other animal viruses.

While NMR and X-ray structures of individual HIV-1 Gag proteins provide clues as to potential interprotein contacts within HIV-1 particles, much remains obscure. Several reports have suggested that mature and immature HIV-1 virions show icosahedral symmetry[1,2], but high resolution support for this suggestion is needed. Analysis of mature virus cores has been hampered by difficulties in their isolation, although the recent demonstration that RNA-Gag (CA plus NC) complexes can be assembled *in vitro*[11] may lead to an accurate

model for mature HIV-1 cores. However, the most developed model for the structure of HIV-1 is based on the immature virus form. In particular, examination of negatively stained membranes and immature HIV particles produced from a baculovirus vector led Nermut *et al.*[2] to propose a "fullerene-like" particle structure. By this model, Pr55Gag proteins at viral or cell membranes appear to form hexamer rings surrounding protein-free "holes." Although first order diffraction reflections at 65 Å^{-1} were barely discernable, results of averaging five subimages, assuming six-fold rotational symmetry, suggested an arrangement in which only one Pr55Gag monomer would be shared by two adjacent hexamer units[2]. More recent cryo-EM analysis of disrupted HIV-1 virus-like particles also showed that Pr55Gag proteins appear to form a cage-like structure, although in this case, cage hole-to-hole distances appeared to be on the order of 48Å[13]. In our current study, we describe a method for the analysis of HIV-1 capsid proteins, assembled in vitro on a lipid monolayer. Our results show that membrane-bound HIV-1 CA proteins organize into a hexamer-trimer cage-like network that explains previous structural results in a consistent model.

EXPERIMENTAL PROCEDURES

Lipid monolayer incubations and electron microscopy. The his-HIVCA protein has been described previously[11] and consists of the HIV-1 coding region with an amino-terminal histidine tag and a 3.3 kDa C-terminal extension. The his-HIVCA protein was purified and analyzed as described previously[11,15]. For monolayer incubations, we followed our standard protocol[14] with the subphase containing 0.5-2.0 mg/ml his-HIVCA in subphase buffer at pH 7.8 or 8.3. Subphase

solutions were overlayered as described[14] with 1:1 hexane:chloroform

containing 200 µg/ml phosphatidylcholine (Avanti Polar Lipids) plus 50 µg/ml

nickel-charged DHGN (1,2-di-O-hexadecyl-*sn* -glycero-3-(1'-2''-R-hydroxy-3'-N-

(5-amino-1-carboxypentyl)-iminodiacetic acid) propyl ether). After overnight

incubation at 25-30° C, arrays were lifted onto lacey grids, washed 30 sec in

distilled water, and either negative stained in 1.33% uranyl acetate or plunge

frozen in liquid ethane. Monolayer arrays were viewed and photographed at the

Portland VA Hospital JEOL JEM12OOEX, the University of Oregon Philips

CM12, or the EMBL-Heidelberg Philips CM200-FEG as described previously[14].

Crystalline areas on micrographs were identified by optical diffraction, and

images were scanned and digitized using either the EMBL-Heidelberg Perkin

Elmer 1010GM flat-bed scanner at 2.63 Å/pixel or an Optronics CCD mounted

on a dissecting microscope at 5.25 Å/pixel. Since CM200-FEG images used in

averages were taken at 900 nm defocus, corresponding to a first CTF zero of

approximately 15.1Å, CTF correction was not applied in our analyses.

Image analysis. Image analysis steps employed both the MRC[16,17] and

SPIDER[18] image analysis packages. Initially, ordered areas, identified by optical

diffraction, were scanned, converted to MRC image files, boxed using BOXMRC,

Fourier transformed, and viewed as diffraction patterns on SPECTRA[19].

Diffraction patterns were indexed by hand, masked, and back-transformed to

yield crude filtered images[14,17]. For more thorough analysis, SPIDER[18] real

space operations were used. Twenty 168.3Å x 168.3Å windows, corresponding

to approximately two p1 unit cells from the initial filtered images, were picked

using the SPIDER operation WI from scanned negative #8347 that had been

Gaussian low-pass filtered to 32.9Å resolution. The filtered windows were

aligned using the AP RA and AP SA operations, and averaged using the AD

command. The averaged image then was used as a reference with the programs CC and PK D to pick three sets of cross-correlation peaks: 100 image windows from negative #8347; the best 280 cross-correlation peaks from five digitized negatives; and 1000 windows from the five negatives. Windows from each data set were aligned to a black and white contrast version of the averaged reference window. Aligned windows in each set were averaged using the SPIDER AD operation. For statistical analysis, data sets were halved, averaged, and compared to obtain Fourier Ring Correlation (FRC) indices at different rings of resolution using RF M. Image averages subsequently were filtered using FQ to 24.0 Å resolution in accordance with FRC results.

For rotational averaging, a rotational series of the averaged, negative #8347 data set was generated using RT, and rotated images were compared with the unrotated original using RF M, identifying rotational correlation peaks. The original plus its 60, 120, 180, 240, and 300° rotations were averaged to give a rotationally averaged image: the FRC of the 0, 120 and 240° average versus the 60, 180 and 300° average was 0.90 at 36.0 Å resolution, and 0.62 at 22.4Å resolution. To simulate the lower resolution obtained in previous studies, the six-fold rotationally averaged image was Gaussian low-pass filtered to 52.6 Å resolution using the FQ operation. For alignment of HIV-1 matrix protein trimers onto his-HIVCA arrays, the MA trimer unit from the Brookhaven Protein Data Bank (accession # 1H1W) was imaged on MIDAS, rotated so that the threefold axis was normal to the projection and the predicted internal face of the virus[4,5], and scaled. The scaled MA trimer image was converted to SPIDER image format, used to generate a 120° rotational series of windows at 1° intervals, and each rotated trimer was cross-correlated to the six-fold rotationally averaged his-HIVCA image. The best alignment had the highest value cross-correlation peak.

For display purposes, all images were converted from either MRC or SPIDER image formats to TIFF files, and displayed using Adobe Photoshop software.

RESULTS AND DISCUSSION

To circumvent difficulties that have limited the structural analysis of HIV-1 particles, we have adapted the lipid monolayer method of two dimensional (2D) protein crystallization[26] for analysis of HIV-1 Gag protein interactions. To do so, lipid monolayers were made from phosphatidyl choline (PC) plus a dialkyl-glycerol derivative, DHGN, which carries a nickel-chelating head-group that binds histidine-tagged (his-tagged) proteins[14]. The HIV-1 Gag protein chosen for analysis, his-HIVCA, was an HIV-1 (HXB2) capsid protein derivative with an amino-terminal his-tag[11]. Our rationale for using the matrix-deleted his-HIVCA protein was that the his-tag-DHGN interaction should substitute for the membrane binding function of MA. Furthermore, studies have shown that large HIV-1 *matrix* deletions are compatible with virus particle assembly[19,20]. Although his-HIVCA also lacks the NC and p6 C-terminal domains of Pr55Gag, this did not prevent his-HIVCA from mimicking interactions observed for the full-length protein (see below).

The recombinant his-HIVCA protein was produced in *E. coli*, and purified to greater than 95% homogeneity by two rounds of non-denaturing nickel-chelate chromatography[15]. As a conformation check, his-HIVCA and HIV-1 CA isolated from virus particles were subjected in parallel to a partial proteolytic treatments, and each protein yielded a characteristic C-terminal 14 kd partial trypsin fragment. When his-HIVCA proteins were used in monolayer

incubations with PC plus DHGN, protein arrays with small patches of crystallinity were observed with either stained samples (Figure 1A), or samples frozen in vitreous water. While such regions could be difficult to detect in cryo-EM micrographs (see Figure 1B), they were evident in low pass filtered cryo-EM images (Figure 1C). In small regions, his-HIVCA proteins were ordered well enough to obtain distinct diffraction patterns, as shown in Figure 1D. Despite sample drift for this negative, the reflections at 26.1 Å^{-1} are clearly evident, and the primitive (p1) unit cell of a=74.8Å, b=126.2Å, γ=89.3° is reminiscent of the a=79.2Å, b=137.5Å, γ=89.7° unit cell observed for capsid proteins of another retrovirus, the Moloney murine leukemia virus (M-MuLV)[14]. Interestingly, back-transformation of masked diffraction patterns from small ordered regions yielded images (Figure 1E) showing a cage-like structure of proteins (light) surrounding protein-free cage holes (dark), similar to those seen for M-MuLV CA[14] and HIV-1 Pr55Gag[2].

Because the sizes of ordered his-HIVCA patches were not large enough to permit a thorough analysis by standard diffraction analysis methods[16,17], we opted to analyze his-HIVCA monolayer structures by a modification of single particle analysis methods[18]. To do so, an image-averaged reference corresponding to approximately two p1 unit cells (168.3 Å x 168.3 Å) was used to locate cross correlation peaks from five separate digitized micrographs. As expected from the low pass filtered image (Figure 1C), peaks tended to be located in small ordered patches throughout the micrographs (Figure 1F). From cross-correlation maps, three data sets were derived. These consisted of 100 image windows from our highest quality micrograph; 280 windows corresponding to the best cross-correlation peaks; and 1000 windows (total) from the five micrographs. All windows were aligned to a black and white averaged

reference window, after which data sets were merged to obtain 100 image, 280 image and 1000 image averages (Figure 2A-C). To evaluate results, each data set was halved, and averages from each half set were compared with their counterparts at different rings of resolution. As shown in Table 1, for each data set, Fourier Ring Correlation values fall off at resolutions better than 24.0 Å. These results are consistent with the observation of reflections at 26.1 Å^{-1} (Figure 1D), and Figures 2D-F have been filtered to this resolution limit.

The most obvious features of the his-HIVCA structures shown in Figure 2 are the large circular holes of 26.3 Å in the protein cage, which are bordered by six rectangularly-shaped protein units. However, each circular hole also is surrounded by six smaller protein-free zones, which in turn are bordered by three rectangular units. Thus, membrane-bound HIV-1 capsid appears to form a cage consisting of hexamer and trimer units. Since previous lower resolution studies on *in vivo*-derived immature HIV-1 particles employed a six-fold rotational averaging scheme in image reconstructions[2], we wished to assess the results of rotational averaging for comparative purposes. As shown in Figure 3A (top panel), the large cage holes in the HIV-1 capsid arrays appeared as six-fold rotational axes, and the averaged image correlated with its 60° rotation to about 21.3Å (Figure 3A, bottom panel). Six-fold averaging increased the distinction of putative monomer units (Figure 3B), and when this image was filtered to 52.6 Å resolution, the four rectangular units separating each trimer hole appeared as single electron-dense units (Figure 3C). The resultant image compares well with HIV-1 Pr55Gag arrays at membranes and in disrupted immature virus particles[2,13]. In particular, the 74.2Å hexamer to hexamer cage hole spacing that we observe (Figures 2,3) is consistent with the spacing observed for Pr55Gag proteins arranged on the plasma membranes of baculovirus vector-infected

cells[2]. Since that study used negatively stained samples and was at low resolution, we believe that trimer cage holes were not resolved, much as they are not resolved in our low pass filtered image (Figure 3C). In contrast, cryo-EM analysis of disrupted immature HIV-1 particles has shown a cage hole center-to-center spacing of approximately 48Å[13], which is in good agreement with our observation of trimer to hexamer cage hole spacing of 44.9Å. Given the above observations, it appears that membrane-bound HIV-1 CA proteins assemble in a fashion similar to the full-length Pr55Gag proteins. This may seem surprising, since Gag protein nucleocapsid (NC) domains, possibly through associations with RNA, exert a strong influence on retrovirus particle assembly[12,14]. However, the similarity of our HIV-1 CA projection structure (Figures 2,3) and that of Pr55Gag proteins in virus-like particles[2], is consistent with indications that MA, CA, and NC domains in Pr55Gag molecules stack roughly as long rods[13]. We thus hypothesize that NC and RNA interactions do not alter the basic arrangements of membrane-tethered capsid proteins, at least at our level of detection.

Although the HIV-1 CA hexamer-trimer arrangement agrees with available data on immature HIV-1 patches, it differs slightly from the hexamer-hexamer arrangement which M-MuLV capsid proteins form on membrane monolayers[14]. In particular, M-MuLV capsid proteins form arrays in which roughly symmetrical hexamer cage holes are surrounded by six skewed hexamer cage holes. Also, in contrast with HIV-1 CA arrays, in which each monomer unit contributes to one hexamer and one trimer; M-MuLV monomer units each contribute to one symmetrical and two skewed cage holes[14]. Furthermore, at hexamer vertices, it appears that putative HIV-1 CA monomers coordinate with three other monomer units, while M-MuLV monomers coordinate only with two

additional monomers. However, both M-MuLV and HIV-1 arrangements suggest that cytoplasmic portions of the retroviral Env proteins may be accomodated in the hexamer cage hole regions. For HIV, if this is the case, then one might expect that the Pr55Gag matrix domains ordinarily should be positioned between the viral membranes and the CA domains. In keeping with this hypothesis, when HIV-1 matrix trimers as observed in 3D crystals were oriented towards viral membranes as previously predicted[3,4,5] and scaled to our projection structure, they overlayed neatly onto the capsid trimer units (Figures 3D). In this structure, matrix residues 11-32 and 87-100 are oriented towards hexamer holes, and could interact with the cytoplasmic tails of HIV-1 gp120/gp41 Env protein trimers[21], which have been postulated to localize to cage holes[2]. At least two studies have mapped Env protein-interaction regions to these residues[22,23], although other investigations have identified additional interacting residues elsewhere on matrix[24,25].

While the extent to which our system mimics authentic HIV assembly is unclear, for comparative purposes, it is of interest to put our results in the context of individual, immature HIV-1 particles. Assuming that electron-dense units surrounding cage holes are capsid monomers, the CA projection dimensions are approximately 18.4 Å by 34.2 Å, consistent with dimensions obtained for the fragments of HIV-1 CA observed in X-ray diffraction studies[6-9]. Based on a spherical 110 nm diameter for immature HIV-1 particles, this corresponds to a maximum estimate of approximately 4200 Gag monomer molecules per virion, which could accomodate up to 700 Env protein trimers[21] in hexamer holes. We believe that our results, combined with results from NMR, X-ray crystallography, and molecular genetic studies will help identify protein-protein interactions in HIV virions that can be targeted for antiviral therapies.

REFERENCES

1. Gelderblom, H.R. *AIDS* **5** , 617-638 (1991).
2. Nermut, M., Hockley, D., Jowett, J., Jones, I., Garreau, M., and Thomas, D. *Virology* **198**, 288-296 (1994).
3. Matthews, S., Barlow, P., Boyd, J., Barton, G., Russell, R., Mills, H., Cunningham, M., Meyers, N., Burns, N., Clark, N., Kingsman, S., Kingsman, A., and Campbell, I. *Nature* **370**, 666-668 (1994).
4. Rao, Z., Belyaev, A., Fry, E., Roy, P., Jones, I., and Stuart, D. *Nature* **378**, 743-747 (1995).
5. Hill, C., Worthylake, D., Bancroft, D., Christensen, A., and Sundquist, W. *Proc. Natl. Acad. Sci. USA* **93**, 3099-3104.
6. Gitti, R., Lee, B., Walker, J., Summers, M., Yoo, S., and Sundquist, W. *Science* **273**, 231-235 (1996).
7. Momany, C., Kovari, L., Prongay, A., Keller, W., Gitti, R., Lee, B., Gorbalenya, A., Tong, L., McClure, J., Ehrlich, L., Summers, M., Carter, C., and Rossmann, M. *Nat. Struct. Biol.* **3**, 763-770 (1996).
8. Gamble, T., Vajados, Yoo, S., Worthylake, D., Houseweart, M., Sundquist, W., and Hill, C. *Cell* **87**, 1285-1294 (1996).
9. Gamble, T., Yoo, S., Vajdos, F., von Schwedler, U., Worthylake, D., Wang, H., McCutcheon, J., Sundquist, W., and Hill, C. *Science* **278**, 849-853 (1997).
10. Morrelet, N., Jullian, N., De Rocquigny, H, Maigret, B., Darlix, J.-L., and Roques, B. *EMBO J.* **11** , 3059-3065 (1992).
11. McDermott, J., Farrell, L., Ross, R., and Barklis, E. *J. Virol.* **70**, 5106-5114 (1996).
12. Campbell, S., and Vogt, V. *J. Virol.* **69** , 6487-6497 (1995).
13. Fuller, S., Wilk, T., Gowen, B., Krausslich, H.-G., and Vogt, V., *Current Biology* **7**, 729-738 (1997).
14. Barklis, E., McDermott, J., Wilkens, S., Schabtach, E., Schmid, M., Fuller, S., Karanjia, S., Love, Z., Jones, R., Rui, Y., Zhao, X., and Thompson, D. *EMBO J.* **16**, 1199-1213 (1997).
15. Hochuli, E., Dobeli, H., and Schacher, A. *J. Chromatog.* **411**, 177-184 (1987).
16. Crowther, R., Henderson, R., and Smith, J. *J. Struct. Biol.* **116**, 9-16 (1996).
17. Hardt, S., Wang, B., and Schmid, M.F. *J. Struct. Biol.* **116**, 68-70 (1996).
18. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. *J. Struct. Biol.* **116**, 190-199 (1996).
19. Faecke, M., Janetzko, A., Shoeman, R., and Krausslich, H.-G. *J. Virol.* **67**, 4972-4980 (1993).
20. Wang, C., Zhang, Y., McDermott, J., and Barklis, E. *J. Virol.* **67**, 5550-5561 (1993).
21. Chan, D., Fass, D., Berger, J., and Kim, P. *Cell* **89**, 263-273 (1997).
22. Freed, E., and Martin, M. *J. Virol.* **69**, 1984-1989 (1995).
23. Yu, X.-F., Yuan, X., Matsuda, Z., Lee, T.-H., and Essex, M. *J. Virol.* **66**, 4966-4971 (1992).
24. Dorfman, T., Mammano, F., Haseltine, W., and Gottlinger, H. *J. Virol.* **68** , 1689-1696 (1994).

25. Lee, Y.-M., Tang, X.-B., Cimakasky, L., Hildreth, J., and Yu, X.-F., *J. Virol.* **71**, 1443-1452 (1997).
26. Uzgiris, E., and Kornberg, R. *Nature* **301**, 125-129 (1983).
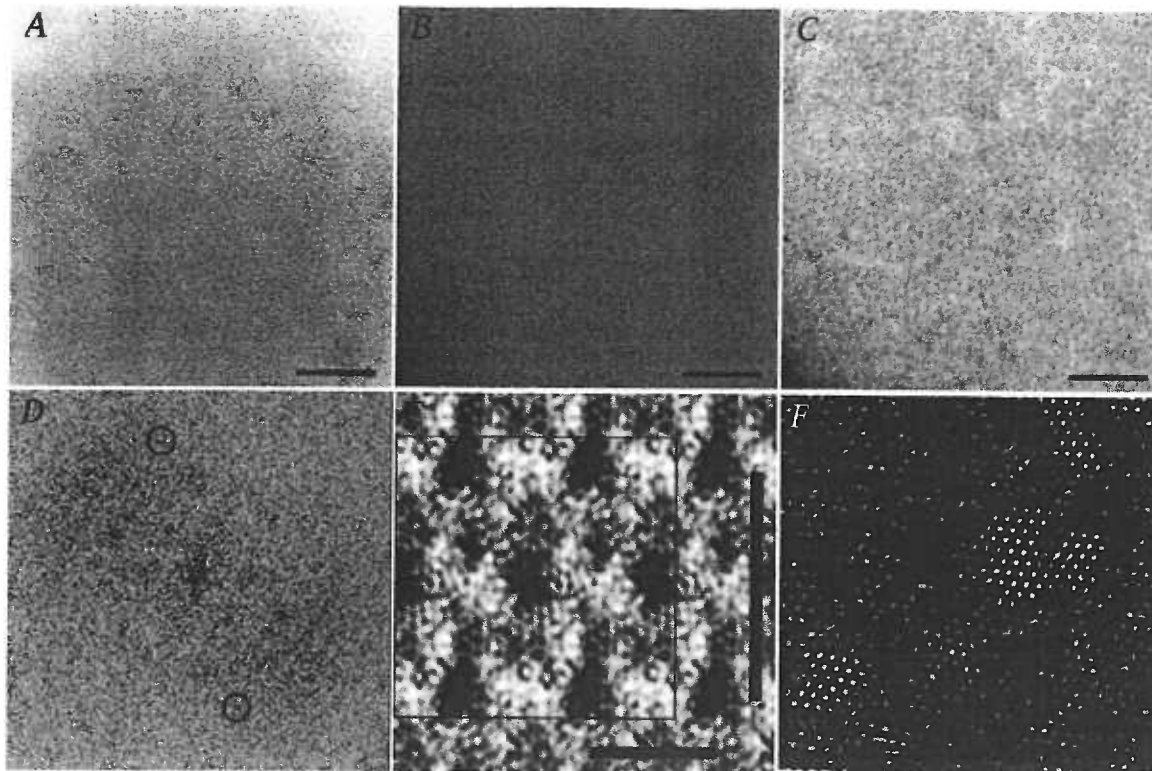
TABLES AND FIGURES

Figure 1. Two dimensional arrays of his-HIVCA on lipid monolayers. A, Negatively stained his-HIVCA array formed on PC plus DHGN. The contrast on the scanned EM negative was inverted so that protein-free regions are dark. The scale bar is 53.6 nm. **B,** Unfiltered cryo-EM image of his-HIVCA arrays. Protein free regions appear dark, and the size bar indicates 50 nm. **C,** The image from panel B was Gaussian low pass filtered to 32.9Å resolution using the SPIDER operation FQ. The size bar indicates 50 nm. **D,** The power spectrum from a 135 nm x 135 nm his-HIVCA array in vitreous ice was calculated from the scanned image and can be indexed as a primitive (p1) unit cell with a=74.8Å, b=126.2Å, γ=89.2°. The circled 2,-4 and 2,4 reflections are at 26.1 Å^{-1}. **E,** The calculated diffraction pattern from an ordered section of negative #8347 was displayed, indexed and masked using SPECTRA[18], and back transformed to yield the filtered image, in which protein-free areas appear darkly, while electron-dense areas are bright. The thick black lines indicate p1 unit cell dimensions of 74.8Å and 126.2Å, while the thin black line shows the box format used to pick windows to average for the cross correlation reference in panel F. **F,** White spots indicate cross correlation peaks from panel B, using a reference averaged from twenty 168.3Å x 168.3Å negative #8347 windows that were boxed using the two unit cell box from panel E as a guide.

Resolution Range (Å)	FRC		
	100-Image	280-Image	1000-Image
≥168.3Å	0.88	0.95	0.99
168.3-56.1Å	0.95	0.99	0.99
56.1-33.7Å	0.92	0.96	0.97
33.7-24.0Å	0.61	0.69	0.73
24.0-18.7Å	0.16	0.06	0.29

Table 1. Fourier Ring Correlation values of averaged data sets at different resolution ranges. Fourier Ring Correlation (FRC) values were determined for each data set by splitting each data set in half, averaging each half set, and comparing half averages at different resolution rings using the SPIDER operation RF M[18]. An FRC of 1.0 indicates complete correlation, while a value of 0.5 may be taken as an operational indicator of the resolution limit.
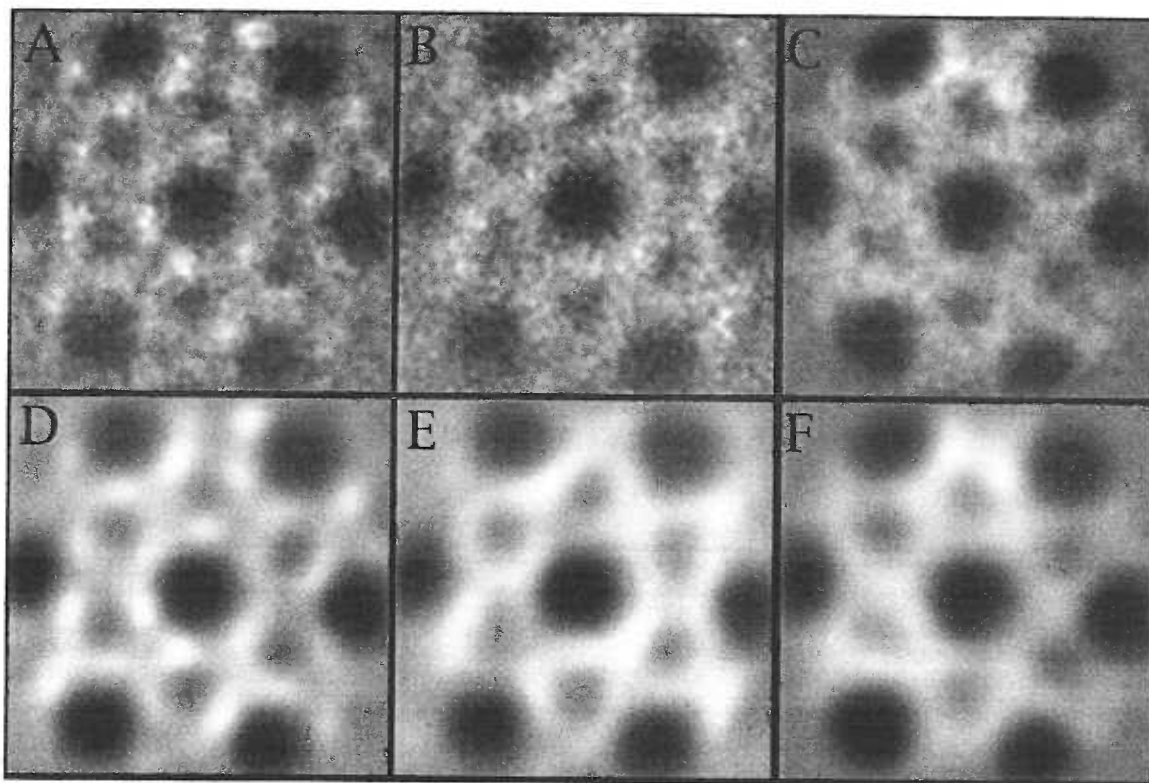
Figure 2. Arrangement of HIV-1 capsid proteins bound to a lipid monolayer. Image windows from cryo-EM micrographs of two dimensional his-HIVCA arrays were aligned and averaged as described in the Methods section. Averages are displayed as 168.3Å x 168.3Å areas with protein-free zones dark, and electron-dense areas light. Averages were either unfiltered, or Gaussian low pass filtered to 24.0Å resolution, in accordance with results from Table 1. A, 100 window average, unfiltered; B, 280 window average, unfiltered; C, 1000 window average, unfiltered; D, 100 window average, filtered; E, 280 window average, filtered; F, 1000 window average, filtered.

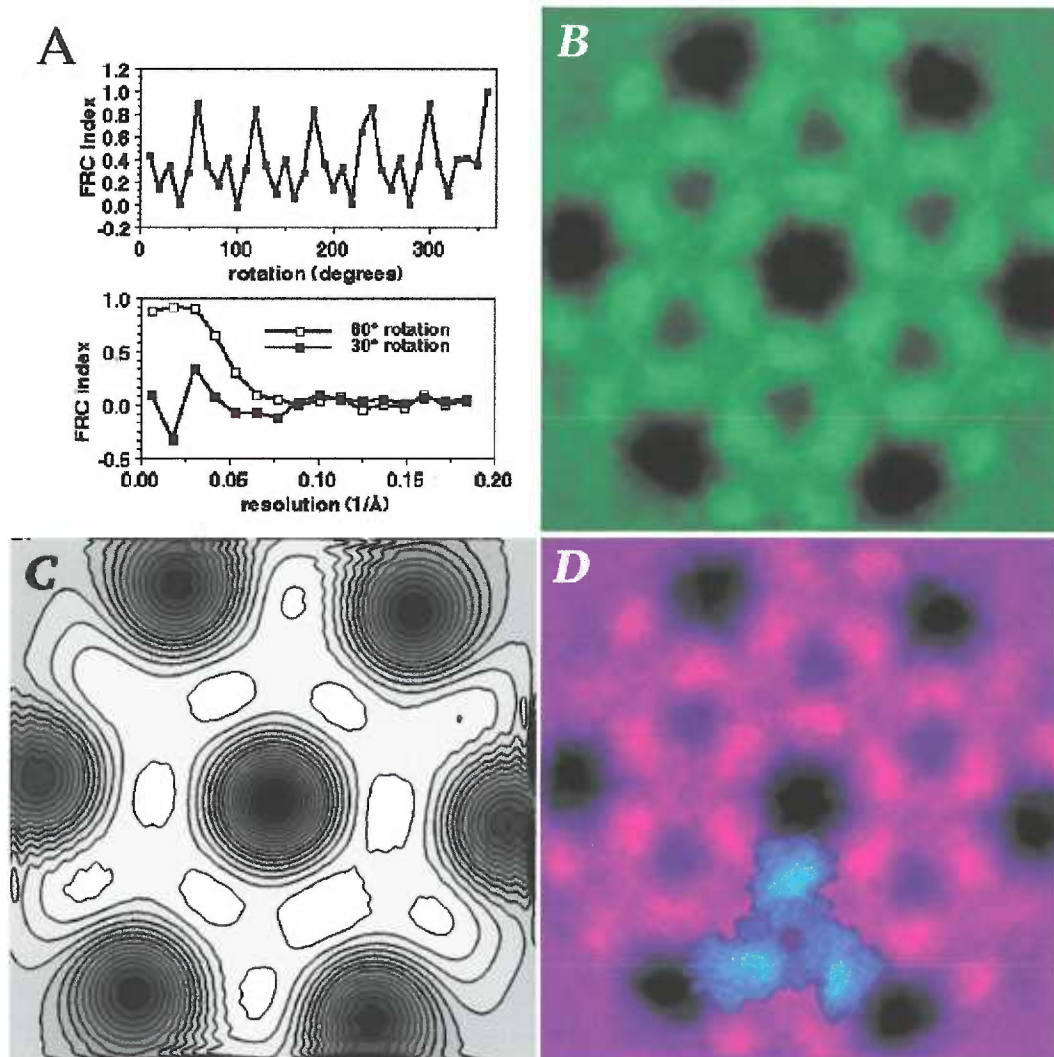Figure 3. Rotational averaging, and comparison with virus particle and atomic structures. **A, Top panel,** The unfiltered average from Figure 2A was used to generate a 360° rotational series of images, which were compared with the original, unrotated average. Fourier Ring Correlation (FRC) values at 33.7Å resolution were plotted versus rotation angles, and show six peaks. **Bottom panel,** The averaged image from Figure 2A was compared with 30° (filled squares) or 60° (open squares) rotated versions of itself, and results of the comparisons were plotted as Fourier Ring Correlation (FRC) indices versus resolution in reciprocal space. For the 60° comparison, the FRC was 0.5 at 21.3Å. **B,** The six-fold rotationally averaged image from the 100 window data set is shown with dark protein-free zones, and electron-dense areas in green. The image size is 168.3Å x 168.3Å. **C,** To simulate lower resolution results obtained with immature HIV-1 virus-like particles produced *in vivo*[2], the rotationally averaged image from panel B was filtered to 52.6Å resolution. **D,** The HIV-1 matrix protein trimer structure[5] (Brookhaven PDB #1H1W) in blue was scaled and fitted with the his-HIVCA image (purple) to maximize projection overlap as described in the Methods. After identification of the best-fit cross-correlation peak, the scaled matrix trimer, depicted as a space-filling model, was oriented so that its predicted membrane-binding face[4,5] is away from the reader. In this representation, matrix residues 11-32 and 87-100 are oriented towards hexamer holes, where HIV-1 gp120/gp41 envelope protein complex cytoplasmic tails have been postulated to reside[2].

Chapter 6

THREE-DIMENSIONAL ORGANIZATION OF CAPSID PROTEINS ON A LIPID MONOLAYER

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

[1]Vollum Institute and Department of Microbiology, Oregon Health Sciences University, Portland, Oregon 97201-3098
[2]corresponding author: barklis@ohsu.edu

SUMMARY

We have used a method for the two-dimensional (2D) crystallization of retroviral structural proteins to obtain a three-dimensional (3D) 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 γ=119.5°, but appeared to have no intrinsic symmetry (p1) in 3D, in contrast to the trigonal or hexagonal appearance of their two-dimensional 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.

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 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; Gitti *et al.*, 1996; Momany *et al.*, 1996; Gamble *et al.*, 1997; 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 alpha-helices in a bundle-like arrangement, and the C-terminal domains (CTDs) adopt a globular organization formed by four alpha-helices. In each of these structures the orientation of the CTD relative to the NTD appears to be flexible (Berthet-Colominas *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 alpha-helix which overlaps the linker region and the CTD and stabilizes the structure of the CTD by forming hydrogen bonds with other CTD alpha 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; Fuller *et al.*, 1997; Nermut *et al.*, 1998) and M-MuLV (Yeager *et al.*, 1998) have shown that PrGag proteins form cage-like structures with a 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; Campbell & Vogt, 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. Previously we obtained two-dimensional (2D) 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; Fuller *et al.*, 1997; Nermut *et al.*, 1998; Yeager *et al.*, 1998). We have extended these studies by obtaining a three-dimensional (3D) 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 must 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 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 1a, which shows a 140.9 nm x 140.9 nm crystalline patch. The crystalline order is apparent in the corresponding diffraction pattern (power spectrum; Figure 1b). 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 by 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 x 79.6Å; 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, sixty-one untilted diffraction patterns were boxed, indexed, unbent, CTF-corrected and merged, assuming no symmetry constraints (p1). Amplitude and phase values for each reflection were vector averaged, and averaged values were used in back-transformations to yield the Figure 1c signal-enhanced 225 Å x 225Å 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 Figures 1b 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 (p3) and hexagonal (p6) space groups, in addition to the primitive (p1)

space group. All other 2D crystal space groups gave significantly higher internal symmetry phase residuals.

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. We chose to do this with phase centered references in all three indicated space groups, p1, p3, and p6, since merging with higher symmetry (p6) gives a greater sampling of the data, while merging with lower symmetry (p1) requires fewer assumptions about the data. However, p1 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 p1 axes in one diffraction pattern correspond to the same p1 set in another. Arbitrarily choosing axes by eye for each image is equivalent to applying six-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 Methods section). A comparison of averaged phase residuals from such an optimization (Figure 2a) showed a high match for one set of axes for each image, with decreasing values for the other five potential axes sets. This result is compatible with p1 space group symmetry, but less clearly so with p3 or p6 symmetry, which would be expected to yield either three or six approximately equally good normalized phase match values. Other analysis also argued against

p3 or p6 space group symmetry assignments for his-MoCA crystals. Comparison of phase residuals versus tilt angles in his-MoCA p1, p3, and p6 merges (Figure 2b) showed a clear distinction. In particular, non-zero tilt angles yielded rising phase residuals for p3 and p6 merges, but relatively constant p1 residuals (Figure 2b). Thus, the p3 and p6 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 1c) 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 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 one. 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 x 72.5 Å p1 unit cell. For comparison with the 2D projection in Figure 1c, examination of negative stain accumulation areas corresponding to protein-free holes in a view perpendicular to the membrane (Figure 3a) was helpful. Three different types of protein-free zones occurred in an apparent hexagonal or trigonal arrangement (Figure 3a), reminiscent of Figure 1c. Looking at the protein areas from the same view gives a different impression (Figure 3b). 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 to the membrane (Figure 4a), the membrane-proximal domains of the C-shaped units from Figure 3b, which we take to be his-MoCA NTDs, appear to form dimers (Figure 4b). The connections between the putative NTD dimers occur in membrane-distal regions, which appear to represent CTDs (Figures 4a, 4b). The overall look of the structure is that of chains of CA proteins, alternately linked by putative NTD and CTD dimers (Figure 4a-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 (Figures 4c, 4d). Viewed from an angle slightly off of perpendicular to the membrane, where the linear strings of proteins are oriented left to right, interconnections between strands mediated by these membrane-distal regions can be seen, using two different density levels (Figure 5a, 5b). 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 5c, **X**); one set appears to connect neighbor strands in one direction (Figure 5c, **Y**); and one set appears to connect neighbor strands in the other direction (Figure 5c, **Z**). Assuming our interpretation is correct, at least six types of CA monomers (Figures 5c-e, #1-6) appear to occupy positions in the crystalline lattice. Because the crystal represents an average, all these alternatives are superimposed in the final reconstruction (Figures 5a, 5b; 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 3a) correspond to 2D projection hole-to-hole distances (Figure 1c) 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 75-100Å, 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). 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 (p1) space group yielded low phase residuals at all tilt angles, and Q values

approaching one (Figure2, 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. Also apparent 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 5c-f). 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). 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 what has been seen in retroviral CA and CA-NC helical rod assemblies (Ehrlich *et al.*, 1992; Campbell & Vogt, 1995; Campbell & Vogt, 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 interconnect 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 previously described (Barklis *et al.*, 1997). Briefly, the his-tagged M-MuLV CA protein (his-MoCA) was expressed in *E. Coli.* strain BL21(DE3)/pLysS (Novagen) from a bacterial expression construct containing the capsid coding region inserted in the BamHI 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 0.2-2 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% glycerol, and 0.02% 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 10X stock solutions in chloroform under nitrogen gas.

Monolayer crystallization incubations were performed as previously 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 NaAc, pH 7.6; 20% glycerol) was overlaid with 1 µl of a 1:1

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° in sealed, distilled water-humidified plates, after which they were subjected to optional 0.5-3h 4° C post-incubations. Crystalline arrays were lifted onto lacey EM grids, and grids were washed 30-60 sec in distilled water, stained 45-60 sec with 1.33% 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 52,000x were taken using the CM120 low-exposure protocol, with the search mode at 2,500-5,000x, and the focus mode at 135,000x. 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 52,000x, the OHSU CM120 TEM tilt to x angle was determined to be $116.5 \pm 7.1°$.

Digitization, transformation, and initial indexing

Negatives were scanned using an Optronics DEI-470 CCD mounted on a Fisher Stereomaster dissecting microscope equipped with a 0.5x 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 previously 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 ($\gamma*$ 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 p1, p3, and p6 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 p1 reference was converted to positive h index form, while p3 and p6 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 p1 and p3 and ORIGTILTB for p6 (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 γ=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 p1 merges, an optimization step was performed to facilitate indexing. Since γ* = 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) which 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 p1, p3, and p6 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 p1

merged set, images from the index-optimized p1 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 35Å 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 figures are as follows: Figures 3a and b, X[-1, 0, 0], Y[0, 1, 0], Z[0, 0, -1]; Figures 4a and b, X[0.5719, 0.8095, -0.1328], Y[0.8203, -0.5664, 0.0797], Z[-0.0107, -0.1545, -0.9879]; Figure 4c, X[0, 1, 0], Y[0, 0, 1], Z[1, 0, 0]; Figure 4d, X[0.866, 0.5, 0], Y[0, 0, -1], Z[-0.5, 0.866, 0]; Figures 5a-e, X[-0.0147, 0.9879, -0.1543], Y[0.9982, 0.0234, 0.0549], Z[0.0579, -0.1532, -0.9865]; Figure 6 a-b, X[0.5031, 0.8624, 0.0558], Y[0.7491, -0.4030, -0.5257], Z[-0.4309, 0.3063, -0.8488]. Figures 6a-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 dimension

ACKNOWLEDGEMENTS

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**, 339-51.

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**, 585-91.

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**(13), 7177-80.

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.

Barth, M., Bryan, R. K., Hegerl, R. & Baumeister, W. (1988). Estimation of missing cone data in three-dimensional electron microscopy. *Scanning Microscopy - Supplement* **2**, 277-84.

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**, 1124-36.

Borsetti, A., Ohagen, Å. & 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**(11), 9313-17.

Campbell, S. & Rein, S. (1999). *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J. Virol.*(2270-9).

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**, 6487-97.

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**, 4425-35.

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**, 633-49.

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**, 4213-27.

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**, 4874-83.

Fuller, S. D., Capaldi, R. A. & Henderson, R. (1979). Structure of cytochrome c oxidase in deoxycholate-derived two-dimensional crystals. *J. Mol. Biol.* **134**(2), 305-27,.

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**(10), 729-38.

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**, 1285-94.

Gamble, T. R., Yoo, S., Vajdos, F. F., Schwedler, U. K. v., 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**, 849-53.

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**, 80-3.

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**, 231-5.

Glaeser, R. M., Tong, L. & Kim, S. H. (1989). Three-dimensional reconstructions from incomplete data: interpretability of density maps at "atomic" resolution. *Ultramicroscopy* **27**(3), 307-18.

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**, 4798-810.

Hahn, T., Ed. (1983). International tables for crystallography: Volume A Space group symmetry. Dordrecht, Holland: D. Reidel Pub. Co.

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 p30 domain of the *gag* gene of Moloney murine leukemia virus. *Virology* **142**, 211-4.

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**, 117-23.

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**, 4927-36.

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**(8), 5106-5114.

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

Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Erlich, L. S., Summers, M. F., Carter, C. & Rossman, M. G. (1996). Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.* **3**, 763-70.

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**, 143-9.

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**, 288-96.

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**, 642-50.

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**, 367-75.

Schmid, M., Dargahi, R. & Tam, M. (1993). SPECTRA: A system for processing electron images of crystals. *Ultramicroscopy* **48**, 251-64.

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**, 1555-68.

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**, 279-82.

Smith, J. L., Hendrickson, W. A., Honzatko, R. B. & Sheriff, S. (1986). Structural heterogeneity in protein crystals. *Biochemistry* **25**, 5018-27.

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**, 6106-14.

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**, 24769-75.

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**(12), 7021-32.

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**, 125-9.

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**, 281-7.

Wang, C.-T. & Barklis, E. (1993). Assembly, processing, and infectivity of human immunodeficiency virus type 1 Gag mutants. *J. Virol.* **67**, 4264-73.

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**, 7950-9.

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**, 5550-61.

Wilson-Kubalek, E. M., Brown, R. E., Celia, H. & Milligan, R. A. (1998). Lipid nanotubes as substrates for helical crystallization of macromolecules. *PNAS* **95**, 8040-5.

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. *PNAS* **95**(13), 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**, 743-51.

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**(373-84).

Zuber, G., McDermott, J., Karanjia, S., Zhao, W., Schmid, M. F. & Barklis, E. (2000). Alternative assembly of retrovirus capsid-nucleocapsid proteins in the presence of membranes or RNA. *Submitted.*

TABLES AND FIGURES

Dimensions	
parameter	measurement
a	72.6±1.7Å
b	72.5±1.7Å
γ	119.5±1.5°

Space group fit	
space group	phase residual
p1	21.6±3.0°
p3	16.7±8.6°
p6	18.9±8.0°

Table 1. Untilted unit cell parameters. Diffraction patterns from untilted crystalline his-MoCA images were indexed, boxed, unbent, and CTF-corrected as described in the Methods section. 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 p1 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 p1, p3, and p6 values derive from an average of 28, 22, and 58 comparisons, respectively.

Phase residuals versus resolution

resolution range	overall phase residual
>27.4Å	20.9°
27.4Å-23.7Å	23.4°
23.7Å-19.4Å	17.5°
19.4Å-15.0Å	42.1°

Vector addition analysis

multiplicity	frequency	Q value
1	117	1.000
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
21-120	15	0.911

Table 2. Analysis of merged three dimensional data. Merging of three-dimensional (3D) data was performed on index-optimized amplitude and phase (APH) files from 218 images using ORIGTILTC, assuming no symmetry constraints (p1). For the purposes of phase residual versus resolution comparisons, merging was performed to 15 Å, using IQ ≤ 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 ≤ 4 reflections, and reflections of IQ ≤ 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.
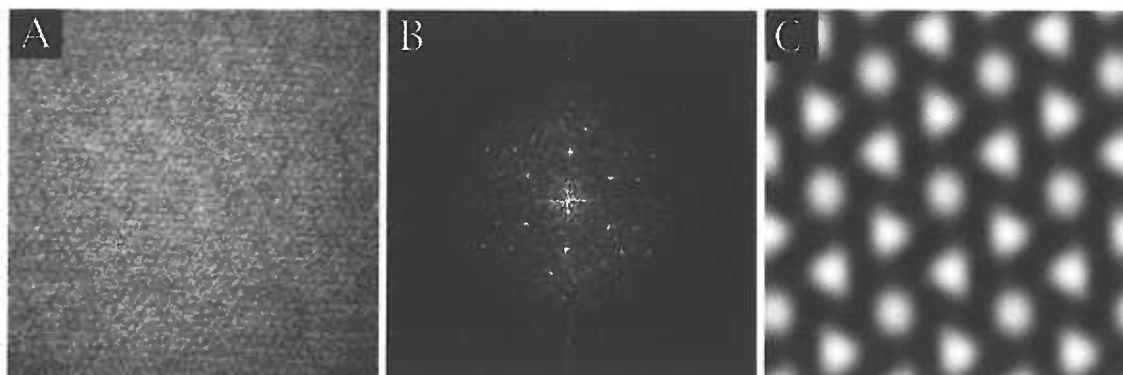
Figure 1. Projection structure of membrane-bound his-MoCA proteins.
A. Shown is a micrograph of a his-MoCA array formed on a PC + DHGN lipid monolayer, lifted onto an electron microscope grid and negatively stained with uranyl acetate as described in the Materials and Methods. The image was taken at 52,000x magnification and represents a 140.9 x 140.9 nm patch. Protein areas appear light and protein-free areas appear dark. **B.** The calculated diffraction pattern from the image in panel 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 224Å x 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 (p1), using high signal to noise reflections (IQ ≤ 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.
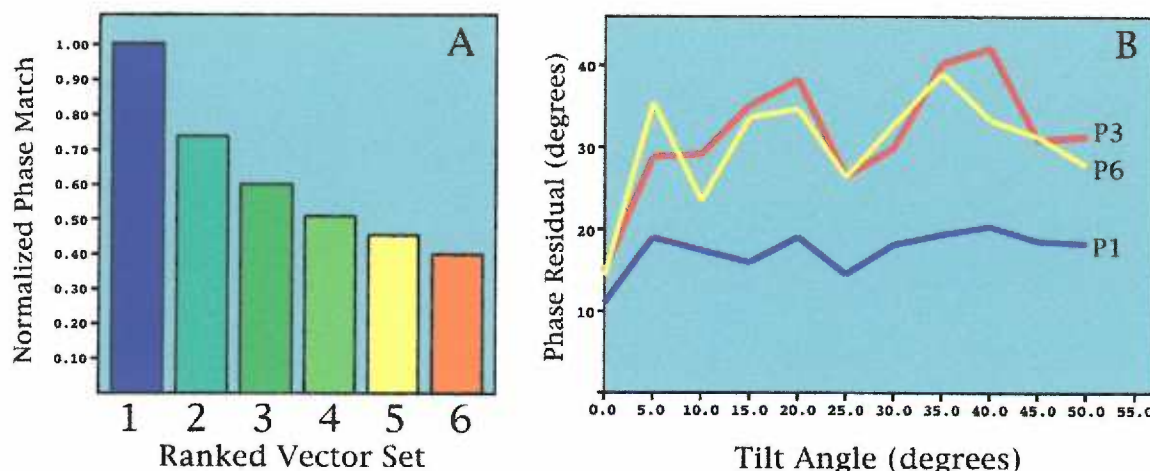
Figure 2. Space group assignment of his-MoCA crystals
A. Shown is an analysis of the consequences of indexing diffraction patterns using the six possible γ*=60° axes sets. Each diffraction pattern was indexed in the six possible indexing schemes, and all indexing schemes for each image were tested in p1 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 (p1) diffraction pattern will show only one high phase match value; a trigonal (p3) pattern will show three high values, and all six vector sets should give equally high values for hexagonal (p6) diffraction patterns. **B.** The full index-optimized MoCA dataset, consisting of 232 tilted and untilted images, was merged assuming p1, p3, and p6 space group symmetry as described in the 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 p1 (blue), p3 (red), and p6 (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.
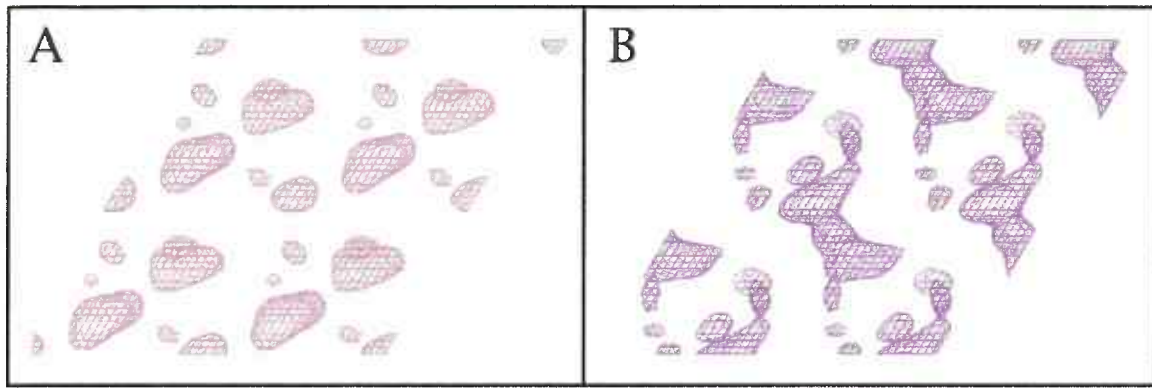
Figure 3. Membrane-bound his-MoCA protein reconstruction viewed perpendicular to the membrane. The his-MoCA dataset was merged into a p1 unit cell, binned, filtered, and back-transformed as described in the Materials and Methods. Protein density was located in the central 100 Å thick Z axis slab, and volumes were viewed using XFIT. Shown in the figure are approximately four unit cells (a=b=144Å) 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 σ= –1.0 to show protein-free zones. (B) The contour level was set at σ=2.5 to show protein regions.
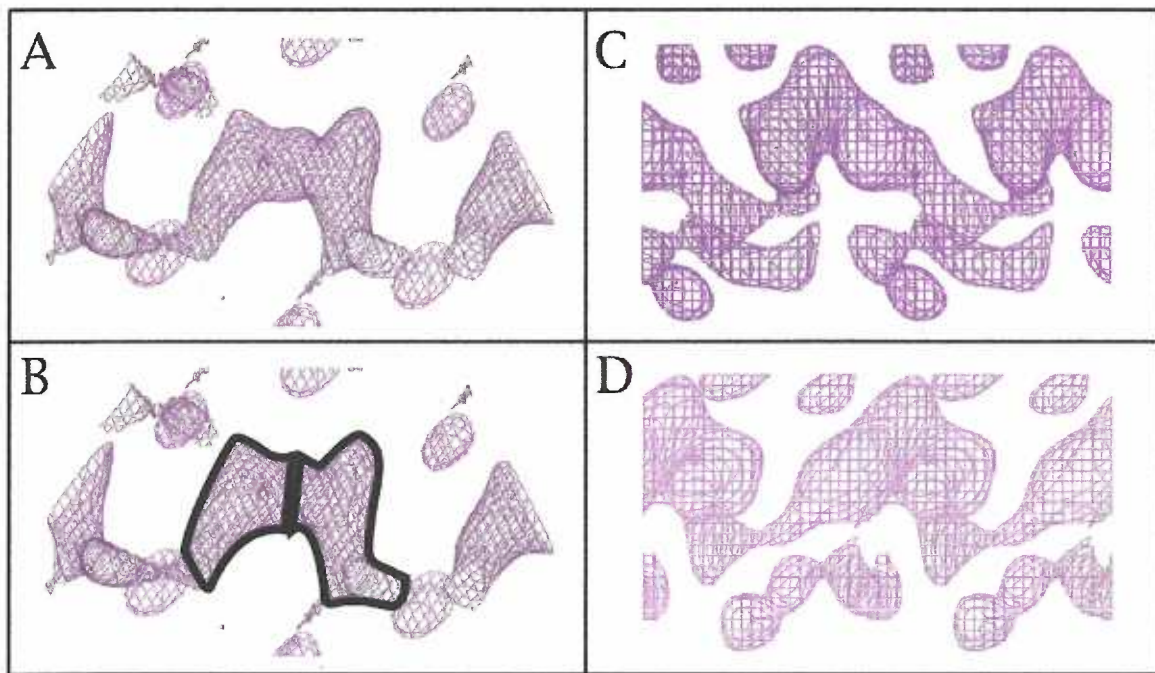
Figure 4. His-MoCA protein-proteins interactions. Shown is the 3D his-MoCA structure from the same 3D density map and using the same a,b,c dimensions as in Figure 3, viewed with σ=2.0, at a slight angle to the Z-axis, by itself (A), and with the members of the putative N-terminal dimer (NTD) outlined (B). In panel (C), the structure is viewed along the X-axis, at a σ=2.0, with the membrane face towards the top of the page , and panel (D) is viewed similarly but rotated 60° clockwise around the Z-axis.
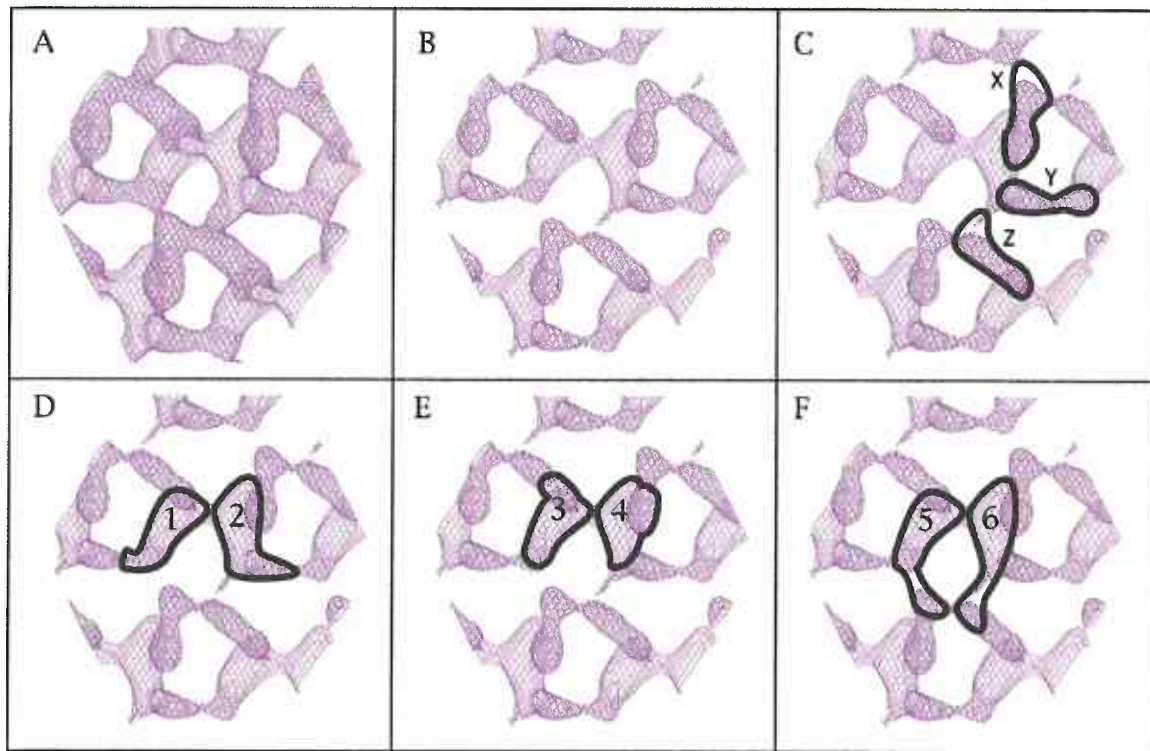
Figure 5. Alternate conformations for the his-MoCA C-terminal domain (CTD). The structure of the membrane-distal portion of the map described in Figure 3 is shown so as to facilitate putative CTD viewing. In all panels, a and b dimensions were 144 Å, and thicknesses were 50 Å. **(A)** Alternate CTD interactions shown with σ=1.5. **(B-F)**Alternate CTD interactions shown with σ=2.0. **(C)** Three electron densities corresponding to putative CTD dimers are labeled X, Y, and Z. **(D-E)**Putative N-terminal domain (NTD) dimers are shown with each of the three alternate locations for the putative CTDs outlined. **(D)** CTDs at right angles to the NTDs. **(E)** CTDs project beneath and backwards from NTDs. **(F)** CTDs extend forward from the NTDs.

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 to the membrane, clipped to a 37.5 Å slab at σ=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.
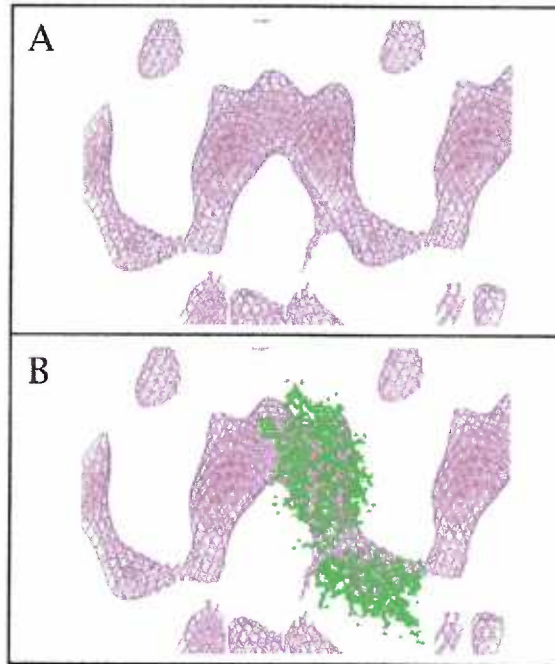
Chapter 7

CONCLUSIONS

Summary

Although multiple atomic resolution X-ray and NMR derived structures of various retrovirus capsid proteins have become available recently (4, 9, 16, 17, 20, 26, 29, 36), the protein-protein interactions which specify the assembly and structure of retrovirus particles have not been elucidated. Information about the macromolecular structure of retrovirus particles has come from genetic and biochemical studies (1, 2, 5, 11, 23, 48), *in vitro* assembly studies (6, 8, 13, 18, 21, 30, 51), and electron microscopy analysis of immature and mature virus particles (15, 25, 39, 41, 49). These studies have shown that the CA domain is an important structural determinant of virus particle structure. In particular the C-terminal portion of CA, which comprises part of an assembly domain that spans the MHR, C-terminal domain of CA and the proteolytic p2 fragment (1, 5, 12) has been shown to be critical for assembly. *In vitro* assembly studies have shown that under certain conditions, the CA domain alone can form ordered structures such as spheres and rods (13, 18). EM studies have suggested that the subdomain structures of retroviruses and *in vitro* derived particles may be based on a hexagonal-like arrangement of protein units, although increasing evidence suggests that retrovirus particle structure is not icosahedral (15, 18, 39, 41, 51). We have employed genetic, biochemical, and biophysical approaches to further investigate the structure of both the HIV-1 and M-MuLV capsid proteins with a specific focus on delineating the interactions which drive virus particle assembly. Firstly, results of the three-dimensional (3D) structures obtained from membrane-bound, his-tagged M-MuLV and HIV-1 capsid proteins (his-MoCA

and his-HIVCA) will be discussed. Next, results from crosslinking studies of both M-MuLV and HIV-1 capsid proteins will be discussed and placed in the context of the 3D structural results. Finally, the two-dimensional (2D) projection structures from his-MoCA and his-HIVCA will be discussed as they relate to the 3D structures and retroviral assembly models.

Three-dimensional structures of retrovirus capsid proteins

A two-dimensional protein crystallization technique was used to gain three-dimensional (3D) structural information from HIV-1 and M-MuLV capsid proteins. His-tagged capsid proteins were crystallized on lipid monolayers containing nickel-chelating lipids which promoted concentration of the proteins in two dimensions at the air-buffer interface, and permitted lateral and rotational mobility during the formation of crystals. Images of crystals were analyzed using image processing techniques, and 3D information was gained by merging images of negatively-stained crystals at various tilt angles, followed by Fourier-reconstruction from 3D datasets. The resulting structures had coherent data to approximately 23 Å (his-MoCA) and 30Å (his-HIVCA) resolution in the X-Y plane, and approximately 35 Å resolution along the Z-axis. Although fine structural detail could not be seen at this resolution, information about the overall shape of the proteins, subdomain shapes and relative orientations, and protein and subdomain oligomerization states was obtained.

The 3D structure of his-MoCA crystals

The 3D structure of his-MoCA was obtained by merging 218 images of negative-stained crystals at tilt angles of 0°-55°. The projection structure obtained by merging data from untilted images had a trigonal or hexagonal appearance with unit cell dimensions a=b=72.5Å, γ=119.7°, similar to the projection structure of unstained cryo-preserved his-MoCA crystals. Contrary to expectations, the 3D dataset was determined to have no intrinsic symmetry and was best described by the primitive p1 space group. The organization of proteins observed also was at odds with 2D projections: instead of being organized as trimers and hexamers, the structure was formed by CA dimers, with the N-terminal domains (NTD) and C-terminal domains (CTD) dimerizing with different partners in a strand-like organization. The NTD of his-MoCA appeared to be localized to membrane-proximal electron density in the form of two lobes, each corresponding to one NTD of a dimer. Electron density connecting NTD dimers into strands was membrane-distal, presumably formed by the CTDs. The structure of the EIAV CA protein (26) showed a reasonable fit to both these regions, both by overlap of electron density and by relative orientation of the NTD to CTD. However, the CTD of the EIAV seemed too large to accommodate fitting two CTDs between each NTD dimer: the electron density for the CTD strand connections seemed lower than expected relative to the NTD. Two regions of additional membrane-distal electron density were present in the structure with approximately the same shapes and densities as the strand connections, suggesting that the CTDs may adopt multiple conformations in the crystal structure. This notion is consistent with the fact that the linker region connecting the NTD and the CTD has been

shown to be flexible in multiple CA structures (4, 9, 26, 29). Interestingly, the MHR, which overlaps this region and the CTD, is one of the most highly conserved region in Gag (47), and this may reflect the need of the MHR to stabilize several conformations of CA. In this regard, other viral structural proteins have been shown to adopt multiple conformations in the setting of the virus particle (27, 28, 31). A model describing this type of virus particle assembly, is the local rules theory of Berger, *et al.* (3, 46) and is discussed further below.

The 3D structure of his-HIVCA crystals

We obtained a 3D structure of membrane-bound his-tagged HIV capsid (his-HIVCA) protein crystals by merging data from 88 images of crystalline areas taken at various tilt angles and reconstructing a 3D density map from the merged dataset. Although data from the merged his-HIVCA dataset was less complete than the data from the his-MoCA dataset (52% vs. 70%), the merged data gave a low phase residual and the reconstructed 3D map showed distinct organization. Firstly, like his-MoCA, the projection structure obtained by merging untilted images from his-HIVCA crystals was similar to that obtained from 2D analysis of ice-embedded crystals. The 3D unit cell was found to be described most accurately by primitive symmetry (p1), and the unit cell size was a=65.3, b=65.8 Å, γ=120.0° with protein density concentrated in a 120 Å slab at the center of the reconstruction. The strand-like arrangement of his-HIVCA was evident by viewing the structure perpendicular to the membrane and, although membrane-proximal electron density was less clearly dimer-like than his-MoCA, the

structure still provided a reasonable fit to the structure of the HIVCA NTD

dimers. Strand connections between NTD dimers were at a level appropriate for

the CTD but seemed of insufficient size to accommodate CTDs. In fact, all

membrane-distal density seemed less well- defined, which could be due to

incomplete or poorly merged data in this region, or to diminished electron

density due to alternate CTD conformations. Merging more, better quality

images into the negative stain 3D dataset, or deriving a structure based on ice-

embedded his-HIVCA crystals should provide more information on this subject.

Biochemical studies of retroviral capsid structure

Crosslinking in immature and mature M-MuLV virus particles

We introduced several cysteine residues into various regions of the capsid

domain of M-MuLV and assayed both immature and mature mutant virus

particles for their abilities to form crosslinked PrGag or CA protein species to

gain structural information about both viral forms. In immature M-MuLV

particles, we found that a cysteine creation in the C-terminal portion of capsid

allowed the formation of crosslinked dimers upon treatment with cysteine

specific crosslinking agents. However, cysteine creations in the N-terminus and

in the N/B determining regions of the CA domain did not facilitate crosslinking

in immature particles. In contrast, the same mutants in mature virus particles

showed different phenotypes. The cysteines in the N- and C- terminal portions of

CA were able to form crosslinks whereas the cysteine in the N/B determining

region was not. From these results, we propose that the CA domain of M-MuLV

undergoes a significant conformational and/or subunit rearrangement during the maturation process.

This finding is supported by studies in HIV-1 showing that the structure of the CA protein undergoes a conformational change upon cleavage at its N-terminus (14, 22). In addition, results from *in vitro* assembly studies of HIV-1 and Rous sarcoma virus (RSV) have shown that N-terminal and C-terminal extensions of the capsid domain can alter the kinds of higher order structures formed by the capsid protein (6, 21). This conformational switch seems to be a conserved feature among many retroviruses (14), but how the switch affects the structure of the virus particle is still unknown. Although circumstantial evidence links the conformational switch in CA with the morphological changes during virus particle maturation, a correlation has not been proven.

Crosslinking in HIV-1 virus particles

In crosslinking studies of the HIV-1 capsid domain, we found that elimination of C350, one of the naturally occurring cysteine residues in the C-terminal portion of CA, abolished virus particle assembly. However, elimination of the other cysteine, C330, had little effect on particle assembly. The cysteines also were differentially reactive to cysteine-specific modifying reagents, suggesting that they were not involved in forming inter- or intra-molecular cystine bonds. Crosslink analysis of introduced cysteine residues showed that cysteine residues introduced into the MHR were able to form intermolecular crosslinks in the context of mature virus particles whereas a cysteine residue

introduced in the N-terminal portion of CA (T242C) did not have this ability. All cysteine mutants showed decreased infectivity, and the MHR mutants as well as the C330S mutant in the C-terminus of CA were nearly completely non-infectious, despite showing little reduction in particle assembly and release.

These observations suggest that HIV-1 CA mutants may be blocked in a post-entry step of infectivity, a phenotype which has been observed in other studies with capsid domain mutants including some in the MHR (32, 43). Other reports have determined that at least one block in the infectious process with MHR mutants occurs after entry into the host cell, but prior to the reverse transcription step (2). This suggests that the block occurs during the ill-defined step of virus uncoating. This process is considered to involve an alteration in subunit interactions of the retrovirus core which blocks reverse transcription complex access to the host cell cytoplasm. In X-ray and NMR structures of CA proteins, the MHR has been shown to form multiple hydrogen bonds with alpha-helices in the C-terminal domain, thus stabilizing it (4, 9, 17, 26, 29). Mutations in the MHR may cause a defect in the assembly process which does not manifest until the uncoating step, and is not observable in assays on particle density, *in vitro* RT function, RNA content or morphological differences by EM (2, 32, 43). Another possibility, as we have proposed, is that mutations in the MHR may compromise the ability of the MHR to stabilizing alternate CTD conformations, and that these alternate conformations may involved in a conformational switch mechanism during the uncoating process.

Two-dimensional projection structures of retroviral capsid proteins

Projection structures of both his-MoCA and his-HIVCA in vitreous ice were determined by Fourier analysis or patch-averaging techniques, respectively. Both structures had trigonal or hexagonal appearances, with cage-holes being surrounded by what appeared to be six protein monomers. The his-MoCA structure showed reflections to 9.5 Å resolution and could be indexed as orthorhombic (C222) or hexagonal (p6), with a hexagonal unit cell size of a=b=79.7 Å, γ=120°. Two to three types of protein-free holes were present in this structure, and each hole appeared to be surrounded by six protein monomers. His-HIVCA proteins formed small arrays which were indexed as primitive (p1) giving orthorhombic unit cell dimensions of a=74.2 Å, b=126.2 , γ=90.7° (corresponding to a hexagonal indexing of a=b=74.2 Å, γ=120.0°). A projection structure based on patch-averaging of 100 crystalline patches showed good data agreement to approximately 24 Å resolution and displayed six-fold symmetry on the basis of Fourier ring correlation values. We proposed a model based on this structure which included head-to-tail hexamers surrounding the large holes and trimers surrounding smaller holes.

The observations from these projection structures are supported by multiple EM studies of assembling Gag and immature virus particles *in vivo* and *in vitro* studies by EM and AFM (15, 18, 41, 49-51). The hole-to-hole spacings seen in these studies and the apparent hexameric arrangements of the Gag proteins agree with our structures (15, 18, 41, 49-51). In addition, projection structures

based on merged untilted images of negatively stained crystals of his-MoCA and his-HIVCA agree very closely with our projection structures in vitreous ice. However, 3D structures from both capsid proteins show a lack of trigonal or hexagonal symmetry or protein arrangement, exhibiting a strand-like arrangement of dimerized proteins. Thus, while 2D projection structures appear trigonal or hexagonal, their 3D counterparts do not display this symmetry.

Two predictions based on the 2D models must be re-examined based on the results of our 3D structures. Firstly, we proposed that the cage-holes formed by these structures were plausible sites for positioning of envelope protein transmembrane or cytoplasmic domains. This proposal is not substantially altered by details of the 3D structure, since the protein-free holes are still present and of similar sizes. Secondly, in the case of the his-HIVCA projection structure, we modeled the trimeric matrix protein (MA) as fitting on top of the proposed CA trimer. The MA proteins of HIV-1 and a number of other retroviruses have been shown to form trimers, both in crystal structures as well as in *in vitro* assays (10, 24, 34, 35, 38, 42) and this trimeric form has been proposed to be important for proper virus particle assembly in the context of PrGag (10, 37, 38). It is difficult to reconcile our dimer-based, stranded structure with the presence of a trimer. Nevertheless, several explanations can be proposed. Our structures may not accurately reflect virus particle structure. Alternately, our structures may reflect the structures of mature virions, and the structure of the immature virus has the potential for a trimer arrangement of CA domains. Another possibility is that CA domains in PrGag are asymmetrically attached via their N-termini to

MA trimers (see Appendix A, Figure 4). Further structural analyses will be necessary to clarify this issue.

Retroviral structural models

There are two primary models for retroviral structure: the icosahedral model and the helical/spiral model (see also Introduction). Retroviruses have long been thought to have some sort of icosahedral symmetry based primarily on observations by EM groups (18, 19, 33, 39-41, 47). HIV-1 has been proposed to have icosahedral symmetry based on observations of apparently hexagonal networks of assembling HIV Gag proteins on membranes (40, 41) and supported by EM studies reporting an icosahedral appearance of retroviral particles, which led to the proposal that the structure of HIV-1 could be assembled by forming an icosahedron from the hexagonal network using a T number of 63 (i.e. 63 monomers per icosahedral face). However, retroviral particles formed *in vivo* as well as *in vitro* show a large variance in size and morphology inconsistent with a simple icosahedral model which would predict more rigidly defined sizes and shapes (6-8, 13, 15, 49). In addition, analysis of virus-like particles by image processing techniques failed to support the presence of icosahedral symmetry (15).

An alternate model for retroviral assembly is the helical/spiral model, which is based partly on observations of tubular assemblies observed in *in vitro* assembly experiments (8). The helical/spiral model attempts to explain the variance in size and morphology of particles *in vivo* and of structures formed *in*

vitro, as well as the apparent dependence of the shape and size of these structures on the presence of RNA in the *in vitro* reactions. However, other studies have found that the CA domain alone as well as other Gag constructs lacking RNA binding function can form similar structures *in vitro*. There also must be cross-strand connections in this sturcture which contribute to the morphology of the products (e.g. cylindrical versus spherical), suggesting the model needs to be elaborated to account for these considerations.

How do the organizations of our his-MoCA and his-HIVCA 3D structures relate to these models? A strand-like organization would be necessary for formation of a spiral/helical model (8), and indeed a model of this type would fit well with the strand organization of the 3D structures we have presented. A simple strand organization, however, fails to explain the formation of a regular, 2-D network, such as ours. Clearly an arrangement which simply describes our NTD dimers as interacting with only two other dimers, like links in a chain, cannot explain the formation of a crystalline array which appears trigonal or hexagonal. Even if different conformations of CTDs form alternate connections, the result is still a chain with only two links per unit. To generate the crystal structure we have observed, the existence of at least one other interaction between dimer units of the strands most be proposed. Since the wrapping of the strands into a spiral to form spherical particles suggests that proteins in the strands will have different environments due to their positioning relative to the positioning of their neighboring strands, the proteins would have to have quasi-equivalent or non-equivalent characteristics. Flexibility of the capsid CTD would

seem to support this notion but it remains unclear as to what the interstrand interaction(s) might be from our structures.

Our 3D models of both his-MoCA and his-HIVCA, with their strand-like appearances would also seem incompatible with icosahedral models. However, the strand-like appearance of our structures does not necessarily imply a strand-like assembly pathway. It is possible that the strand organization is an artifact or merely a reflection of stronger connections with particular partners in the 2D lattice. Additionally, retroviral morphology and the morphology of higher order structures formed *in vitro* from Gag and CA proteins, share features in common with the icosahedral polyomavirus family. Like retroviruses, the polyomaviruses exhibit polymorphic phenotypes, at least in *in vitro* assembly experiments, forming spheres, tubes, spirals and extended sheets (44-46). Also like retroviruses, alterations in experimental conditions, including buffer composition and mutations in the structural proteins, can shift the assembly to favor certain morphologies (44, 45). The polyomaviruses have T=7 icosahedral symmetry but exhibit a large degree of nonequivalence in their structures (27, 28, 44-46). Nonequivalence in the *polyomaviridae* is mediated by the ability of the C-terminus of the structural protein to adopt alternate conformations and form alternate contacts within the context of the virus particle (27, 28, 31, 44-46). The "local rules" model, proposed by Berger, *et al.* (3, 46), establishes a set of rules in which individual structural proteins adopt alternate conformations depending on their local environments, and each conformation contributes in a different way to the formation of the 3D structure of the icosahedron. Retroviral particles may be formed with some sort of local rules organization and that our structures without

apparent 2-fold, 3-fold, or 5-fold axes of symmetry, could be involved in the formation of nonequivalent icosahedral-type structures by virtue of their conformational flexibility.

Although our structural data would seem to support the spiral/helical model of retrovirus assembly, the notion that retroviral structure is based on a loosely icosahedral model with nonequivalent protein subunits, cannot be ruled out. In addition, the spiral/helical and icosahedral models are not necessarily mutually exclusive, so retrovirus structure might be based on a combination of both, for example assembling in a spiral/helical pathway to generate a particle with icosahedral symmetry. Obtaining higher resolution structures of his-MoCA and his-HIVCA crystals, negatively-stained or ice-embedded, may shed light on this subject. However, studies such as these will need to be combined with studies from whole virus particles, isolated cores, and/or *in vitro* assembly products of other types to resolve this issue.

Summary of Structural Findings

The results from the experiments described in this thesis suggest that the structure of retroviral particles is dynamic in nature and that the C-terminal domain, including the MHR, plays a particularly important role in retrovirus particle assembly and generation of retroviral structure. Cysteine-specific crosslinking studies of the MHR showed that MHRs are in close proximity inside virus particles. Other crosslinking experiments showed that only the C-terminal portion of CA was able to be crosslinked in immature virus particles while both

the N- and C- terminal portions of CA could be crosslinked in mature virus particles suggesting that the capsid domain undergoes a conformational change during virus particle maturation. Furthermore, from mutagenesis studies it was determined that the C-terminal portion of the capsid domain is an important determinant of virus particle assembly and/or release.

Studies of capsid proteins assembling on membranes yielded structures which were made up of dimers of capsid proteins in a strand-like arrangement where the NTD and CTD of capsid formed dimers with different partners. The CTD appeared to adopt alternate conformations relative to the NTD in the structures which implies that the region joining the two domains is flexible and suggests a role for the MHR in stabilizing the alternate conformations. The strand-like arrangement of the capsid proteins, as well as the dimeric nature of the assemblies fit with a helical/spiral model of retroviral assembly more easily than with an icosahedral model.

Future Directions

Atomic resolution structure in the X-Y plane and approximately 5Å resolution in the Z dimension can be attained by a 2D crystallography technique such as ours. To obtain higher resolution structures of his-MoCA and his-HIVCA it will be necessary to perform the 3D image reconstructions using well-diffracting images of tilted crystals embedded in vitreous ice. Perhaps with higher resolution structures some of the questions raised by our current structures can be answered. Central issues are the flexible nature of the capsid

CTD and the role of this domain in particle morphogenesis, how our structures are related to structures of assembled full-length PrGag, and how the stranded nature of our structures relates to models for virus structure and assembly. Icosahedral and helical/spiral models of retroviral structure should be more specifically investigated using genetic and biochemical approaches as well as by high-resolution investigation of virus particles and *in vitro* assembly products. Atomic resolution structures available for capsid proteins can be more accurately fit to our EM structures using computational methods and such studies may shed light on capsid-capsid interactions at the molecular level. With such information, molecular docking simulations can be used to generate lead compounds for use as inhibitors of assembly and/or viral uncoating. We have already begun preliminary research in this area. Better understanding of retroviral structure may lead to antiviral drug development, improved use of retroviral vectors as research or gene therapy tools as well as increasing our knowledge of the retroviral life cycle.

REFERENCES

1.   **Accola, M. A., S. Hoglund, and H. G. Gottlinger.** 1998. A putative α-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. J. Virol. **72:**2072-8.

2.   **Alin, K., and S. P. Goff.** 1996. Amino acid substitutions in the CA protein of Moloney murine leukemia virus that block early events in infection. Virology. **222:**339-51.

3.   **Berger, B., P. W. Shor, L. Tucker-Kellog, and J. King.** 1994. Local-rule-based theory of virus shell assembly. Proc. Natl. Acad. Sci. USA. **91:**7732-6.

4.   **Berthet-Colominas, C., S. Monaco, A. Novelli, G. Sibai, F. Mallet, and S. Cusack.** 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:**1124-36.

5.   **Borsetti, A., Å. Ohagen, and H. G. Gottlinger.** 1998. The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. J. Virol. **72:**9313-17.

6.   **Campbell, S., and S. Rein.** 1999. *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J. Virol.

7.   **Campbell, S., and V. M. Vogt.** 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:**4425-35.

8.   **Campbell, S., and V. M. Vogt.** 1995. Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. **69:**6487-97.

9.   **Campos-Olivas, R., J. L. Newman, and M. F. Summers.** 2000. Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. J. Mol. Biol. **296:**633-49.

10.   **Conte, M. R., and S. Matthews.** 1998. Retroviral matrix proteins: a structural perspective. Virology. **246:191-8.**

11.   **Craven, R. C., A. E. Leure-duPree, C. R. Erdie, C. B. Wilson, and J. W. Wills.** 1993. Necessity of the spacer peptide between CA nad NC in the Rous sarcoma virus Gag protein. J. Virol. **67:**6246-52.

12.   **Dorfman, T., A. Bukovsky, A. Ohagen, S. Hoglund, and H. G. Gottlinger.** 1994. Functional domains of the capsid protein of human immunodeficiency virus type 1. J. Virol. **68.**

13.   **Ehrlich, L. S., B. E. Agresta, and C. A. Carter.** 1992. Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro.* J. Virol. **66:**4874-83.

14.   **Endrich, M. M., P. Gehrig, and H. Gehring.** 1999. Maturation-induced conformational changes of HIV-1 capsid protein and identification of two high affinity sites for cyclophilins in the C-terminal domain. J. Biol. Chem. **274:**5326-32.

15.   **Fuller, S. D., T. Wilk, B. E. Gowen, H. G. Krausslich, and V. M. Vogt.**
      1997. Cryo-electron microscopy reveals ordered domains in the immature
      HIV-1 particle. Curr. Biol. **7:**729-38.

16.   **Gamble, T. R., F. F. Vajdos, S. Yoo, D. K. Worthylake, M. Houseweart,
      W. I. Sundquist, and C. P. Hill.** 1996. Crystal structure of human
      cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell.
      87:1285-94.

17.   **Gamble, T. R., S. Yoo, F. F. Vajdos, U. K. v. Schwedler, D. K.
      Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist, and C. P. Hill.**
      1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1
      capsid protein. Science. **278:**849-53.

18.   **Ganser, B. K., S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist.** 1999.
      Assembly and analysis of conical models for the HIV-1 core. Science.
      283:80-3.

19.   **Gelderblom, H. R., P. Bauer, M. Ozel, S. Hoglund, M. Niedrig, H. Renz,
      B. Morath, P. Lundquist, A. Nilsson, J. Mattow, C. Grund, and G. Pauli.**
      1992. Morphogenesis and morphology of human immunodeficiency virus,
      p. 33-54. *In* R. C. Aloia and C. C. Curtin (ed.), Membrane interactions with
      HIV. Wiley-Liss, New York.

20.   **Gitti, R. K., B. M. Lee, J. Walker, M. F. Summers, S. Yoo, and W. I.
      Sundquist.** 1996. Structure of the amino-terminal core domain of the HIV-
      1 capsid protein. Science. **273:**231-5.

21.   **Gross, I., H. Hohenberg, C. Huckhagel, and H.-G. Krausslich.** 1998. N-
      terminal extension of human immunodeficiency virus capsid protein
      converts the in vitro assembly phenotype from tubular to spherical
      particles. J. Virol. **72:**4798-810.

22.   **Gross, I., H. Hohenberg, T. Wilk, K. Wiegers, M. Grattinger, B. Muller,
      S. Fuller, and H.-G. Krausslich.** 2000. A conformational switch controlling
      HIV-1 morphogenesis. EMBO J. **19:**103-13.

23.   **Hansen, M. S. T., and E. Barklis.** 1995. Structural interactions between
      retroviral Gag proteins examined by cysteine cross-linking. J. Virol.
      69:1150-9.

24.   **Hill, C. P., D. Worthylake, D. P. Bancroft, A. M. Christensen, and W. I.
      Sundquist.** 1996. Crystal structures of the trimeric HIV-1 matrix protein:
      Implications for membrane association and assembly. Proc. Natl. Acad.
      Sci. USA. **93:**3099-104.

25.   **Hoglund, S., L. G. Ofverstedt, A. Nilsson, P. Lundquist, H. Gelderblom,
      M. Ozel, and U. Skoglund.** 1992. Spatial visualization of the maturing
      HIV-1 core and its linkage to the envelope. AIDS Res. Hum. Retroviruses.
      8:1-7.

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

27.   **Johnson, J. E.** 1996. Functional implications of protein-protein interactions
      in icosohedral viruses. Proc. Natl. Acad. Sci. USA. **93:**27-33.

28.   **Johnson, J. E., and J. A. Speir.** 1997. Quasi-equivalent viruses: A
      paradigm for protein assemblies. J. Mol. Biol. **269:**66-75.

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

30.   **Klikova, M., S. S. Rhee, E. Hunter, and T. Ruml.** 1995. Efficienct *in vivo*
       and *in vitro* assembly of retroviral capsids from Gag precursor proteins
       expressed in bacteria. J. Virol. **69**:1093-98.

31.   **Liddington, R. C., Y. Yan, J. Moulai, R. Sahli, T. L. Benjamin, and S. C.
       Harrison.** 1991. Structure of simian virus 40 at 3.8-Å resolution. Nature.
       354:278-84.

32.   **Mammano, F., A. Ohagen, S. Hoglund, and H. G. Gottlinger.** 1994. Role
       of the major homology region of human immunodeficiency virus type 1 in
       virion morphogenesis. J. Virol. **68**:4927-36.

33.   **Marx, P. A., R. J. Munn, and K. I. Joy.** 1988. Computer emulation of thin
       section electron microscopy predicts an envelope-associated
       icosadeltahedral capsid for human immunodeficiency virus. Labor.
       Investig. **58**:112-8.

34.   **Massiah, M. A., M. R. Starich, C. Paschall, M. F. Summers, A. M.
       Christensen, and W. I. Sundquist.** 1994. Three-dimensional structure of
       the human immunodeficiency virus type 1 matrix protein. J. Mol. Biol.
       244:198-223.

35.   **Matthews, S., P. Barlow, J. Boyd, G. Barton, R. Russell, H. Mills, M.
       Cunningham, N. Meyers, N. Burns, N. Clark, S. Kingsman, A.
       Kingsman, and I. Campbell.** 1994. Structural similarity between the p17
       matrix protein of HIV-1 and interferon-γ. Nature. **370**:666-68.

36.   **Momany, C., L. C. Kovari, A. J. Prongay, W. Keller, R. K. Gitti, B. M.
       Lee, A. E. Gorbalenya, L. Tong, J. McClure, L. S. Erlich, M. F. Summers,
       C. Carter, and M. G. Rossman.** 1996. Crystal structure of dimeric HIV-1
       capsid protein. Nature Struct. Biol. **3**:763-70.

37.   **Morikawa, Y., D. J. Hockley, M. V. Nermut, and I. M. Jones.** 2000. Roles
       of matrix, p2, and N-terminal myristoylation in human immunodeficiency
       virus type 1 Gag assembly. J. Virol. **74**:16-23.

38.   **Morikawa, Y., W.-H. Zhang, D. J. Hockley, M. V. Nermut, and I. M.
       Jones.** 1998. Detection of a trimeric human immunodeficiency virus type 1
       Gag intermediate is dependent on sequences in the matrix protein, p17. J.
       Virol. **72**:7659-63.

39.   **Nermut, M. V., C. Grief, S. Hashmi, and D. J. Hockley.** 1993. Further
       evidence of icosahedral symmetry in human and simian
       immunodeficiency virus. AIDS Res. Hum. Retroviruses. **9**:929-38.

40.   **Nermut, M. V., D. J. Hockley, P. Bron, D. Thomas, W.-H. Zhang, and I.
       M. Jones.** 1998. Further evidence for hexagonal organization of HIV gag
       protein in prebudding assemblies and immature virus-like particles. J.
       Struct. Biol. **123**:143-9.

41.   **Nermut, M. V., D. J. Hockley, J. B. M. Jowett, I. M. Jones, M. Garreau,
       and D. Thomas.** 1994. Fullerene-like organization of HIV gag-protein shell
       in virus-like particles produced by recombinant baculovirus. Virology.
       198:288-96.

42.   **Rao, A., A. S. Belyaev, E. Fry, P. Roy, I. M. Jones, and D. I. Stuart.** 1995.
       Crystal structure of SIV matrix antigen and implications for virus particle
       assembly. Nature. **378**:743-7.

43. **Reicin, A. S., S. Paik, R. D. Berkowitz, J. Luban, I. Lowy, and S. P. Goff.** 1995. Linker insertion mutations in the human immunodeficiency virus type 1 *gag* gene: Effects on virion particle assembly, release, and infectivity. J. Virol. **69:**642-50.

44. **Salunke, D. M., D. L. D. Caspar, and R. L. Garcea.** 1989. Polymorphism in the assembly of polyoma virus capsid protein VP1. Biophys. J. **56:**887-900.

45. **Salunke, D. M., D. L. D. Caspar, and R. L. Garcea.** 1986. Self-assembly of purified polyomavirus capsid protien VP1. Cell. **46:**895-904.

46. **Schwartz, R., R. L. Garcea, and B. Berger.** 2000. "Local rules" theory applied to polyomavirus polymorphic capsid assemblies. Virology. **268:**461-70.

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

48. **Wang, C.-T., and E. Barklis.** 1993. Assembly, processing, and infectivity of human immunodeficiency virus type 1 Gag mutants. J. Virol. **67:**4264-73.

49. **Yeager, M., E. M. Wilson-Kubalek, S. G. Weiner, P. O. Brown, and A. Rein.** 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.

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

51. **Zuber, G., J. McDermott, S. Karanjia, W. Zhao, M. F. Schmid, and E. Barklis.** 2000. Alternative assembly of retrovirus capsid-nucleocapsid proteins in the presence of membranes or RNA. Submitted.

Appendix A

THREE-DIMENSIONAL STRUCTURE OF MEMBRANE-BOUND HIV-1 CAPSID PROTEIN CRYSTALS

Jason McDermott, Doug Huseby, and Eric Barklis

Summary

We have determined a three-dimensional (3D) structure of membrane-bound, his-tagged HIV-1 capsid protein (his-HIVCA) crystals. Images from negatively-stained his-HIVCA arrays at various tilt angles were merged and reconstructed into a 3D electron density map. Although, two-dimensional (2D) projection structures of his-HIVCA appeared to have trigonal or hexagonal symmetry, the 3D dataset was determined to be best described in the p1 space group (no symmetry) and had a unit cell of a=65.3 Å, b=65.8 Å, γ=120.0°. The organization of the his-HIVCA proteins was reminiscent of his-MoCA protein organization under similar conditions. Membrane-proximal densities appeared to be CA N-terminal dimers, and were connected by membrane-distal protein regions, presumably C-terminal domains (CTDs). Additional membrane-distal connecting regions were present, suggesting that HIVCA CTDs may adopt alternate conformations in the crystal structure. These results support a helical-spiral (3) model for retrovirus particle assembly rather than an icosahedral model (7, 8, 11-13).

Materials and Methods

Sample preparation, data gathering and image processing techniques were as described in the Materials and Methods section for his-MoCA (Chapter 6). Briefly, his-HIVCA was purified by nickel-chelate chromatography and used to form crystals in incubations under DHGN or DOGS at 22°-30° C overnight.

Crystals were photographed on the OHSU Phillips CM120 Biotwin under low-dose conditions at 52,000x, scanned by CCD and converted to the MRC image format (4). Calculated diffraction patterns from 88 images (0-55°) were indexed in SPECTRA (14) as p1 space group symmetry (a=65.3 Å, b=65.8 Å, γ=120.0°), and merged using ORIGTILTC (p1, Z-thickness=150Å, Z-window=0.0002Å^{-1}, using reflections with IQ $\leq$ 5 and a resolution cutoff 20Å for phase origin search) giving an overall phase residual of 18.14°, the dataset was then optimized for p1 sets, improving the phase residual to 9.76°. Then merged output was binned using a Z-thickness of 600Å and including IQ values less than 8, filtered to 32.5 Å in the X-Y plane and 35Å in the Z dimension giving a dataset completeness of 52% (165 reflections from a possible 315). 3D volumes were obtained by back-transformation and viewed with the XTALVIEW suite of crystallography programs (10) as described in Chapter 6.

Results and Discussion

Diffraction patterns from 18 negatively-stained, untilted images of his-HIVCA crystalline areas were indexed with a primitive unit cell (p1; a=65.3, b=65.8, γ=120.0°), and merged in the p1 space group. The merged output file was vector averaged and back-transformed to give a projection structure (Figure 1), showing nine unit cells with protein regions being shown in white and protein-free holes in black. Note that the 2D projection map appears similar to the his-HIVCA projection structure previously obtained from crystals in vitreous ice, but had a smaller unit cell than the ice-embedded structure (a=b=65.5 Å versus

a=b=74.2 Å), similar to the difference observed between negatively-stained and ice-embedded his-MoCA crystal dimensions.

To obtain a 3D structure, 88 images of his-HIVCA crystals at tilt angles between 0° and 55° were merged assuming no symmetry (p1) and then optimized for p1 sets. A phase match comparison of phase residuals from each p1 vector set was performed (see Materials and Methods, Chapter 6) and, like results from the his-MoCA 3D merges, the his-HIVCA phase match distribution (Figure 2a) was indicative of a p1 space group with one vector set ranking best with the others ranking poorly. Trigonal or hexagonal crystals, on the other hand, would be expected to have three or six vector sets with approximately equal higher values. In addition, phase residuals for the p1 merge remained fairly constant versus tilt angle (Figure 2b). In contrast, merges performed in the p3 and p6 space groups showed increasing phase residual values at higher tilt angles, indicating that the apparent p3 or p6 symmetry of the projection structure is not maintained in the third dimension. These results indicate that the his-HIVCA is best described by the p1 space group.

Amplitude and phase information output from the p1, vector set-optimized, merged dataset was binned at 600Å thickness, filtered and used to reconstruct a 3D volume for viewing. Viewed at an angle perpendicular to the membrane, with membrane-proximal regions closest to the viewer and at a density level of σ=2.0 (Figure 3a), electron-dense regions appear to be organized in a strand-like arrangement, with oblong electron-dense protein regions positioned to one side of the large holes. These regions appeared to be

appropriately sized for N-terminal domain (NTD) dimers (data not shown).

Viewed parallel to the membrane (down the a axis) at a density level of $\sigma = 1.8$

(Figure 3b), the profiles of membrane-bound capsid proteins can be observed,

with membrane-proximal putative NTD domains oriented up, and C-terminal

domains (CTDs) down. The lengths of these units are approximately 75 Å,

slightly shorter than the 85 Å length observed in immature virus particle radial

density profiles (6). Viewed at an angle 40° off of perpendicular to the membrane

(membrane face closest) with decreasing density levels (Figure 3c, $\sigma = 2.0$, Figure

3d, $\sigma = 1.8$, Figure 3e, $\sigma = 1.75$), membrane-distal protein regions can be seen

connecting putative NTD dimers. The multiple, relatively weak connections are

reminiscent of the his-MoCA structure, and may be indicative of the his-HIVCA

CTDs ability to adopt alternate conformations to form different connections.

Viewed from the membrane at a 50 Å Z-axis slab ($\sigma = 1.5$; Figure 3f), the putative

NTDs make at least three connections; to the front, to the back, and strand-like

connections to the side.

Although the completeness of the dataset was relatively low (52%), the

dataset merged to yield a low phase residual (9.76°) and density maps obtained

with the merge looked similar to previously obtained structures when viewed as

an untilted projection. The structure of his-HIVCA appears to be similar in

several respects to that of his-MoCA: the 3D structure of his-HIVCA was best

described using a primitive (p1) space group and did not exhibit trigonal or

hexagonal symmetry as predicted from the projection structures; membrane-

proximal densities appeared to be strands of NTD dimers; the membrane-distal

protein regions that form connections between NTD dimers appear to be the CTDs with alternate conformations in the crystal lattice. A model for matching of CA dimers from our structure with MA trimer structures derived from X-ray crystallography is shown in Figure 4. The strand-like nature of the his-HIVCA structure resembles what has been seen with helical *in vitro* assembly products of retroviral CA and CA-NC proteins (1-3, 5, 7, 9, 15) and would seem to agree with helical/spiral models of retrovirus assembly (3). Our structure also suggests that the hexagonal-appearing networks of HIV-1 Gag proteins observed by EM on cell membranes (13) may be actually organized in a strand fashion which appears hexagonal in projection, and also would seem incompatible with icosahedral models of retroviral structure.

REFERENCES

1. **Campbell, S., and S. Rein.** 1999. *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J. Virol.

2. **Campbell, S., and V. M. Vogt.** 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**:4425-35.

3. **Campbell, S., and V. M. Vogt.** 1995. Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. **69**:6487-97.

4. **Crowther, R., R. Henderson, and J. Smith.** 1996. MRC image processing programs. J. Struct. Biol. **116**:9-16.

5. **Ehrlich, L. S., B. E. Agresta, and C. A. Carter.** 1992. Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro*. J. Virol. **66**:4874-83.

6. **Fuller, S. D., T. Wilk, B. E. Gowen, H. G. Krausslich, and V. M. Vogt.** 1997. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. Curr. Biol. **7**:729-38.

7. **Ganser, B. K., S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist.** 1999. Assembly and analysis of conical models for the HIV-1 core. Science. **283**:80-3.

8. **Gelderblom, H. R., E. H. Hausmann, M. Ozel, and G. Pauli.** 1989. Morphogenesis and morphology of HIV. Structure-function relationships. Arch. Virol. **106**:1-13.

9. **Gross, I., H. Hohenberg, C. Huckhagel, and H.-G. Krausslich.** 1998. N-terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. J. Virol. **72**:4798-810.

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

11. **Nermut, M. V., C. Grief, S. Hashmi, and D. J. Hockley.** 1993. Further evidence of icosahedral symmetry in human and simian immunodeficiency virus. AIDS Res. Hum. Retroviruses. **9**:929-38.

12. **Nermut, M. V., D. J. Hockley, P. Bron, D. Thomas, W.-H. Zhang, and I. M. Jones.** 1998. Further evidence for hexagonal organization of HIV gag protein in prebudding assemblies and immature virus-like particles. J. Struct. Biol. **123**:143-9.

13. **Nermut, M. V., D. J. Hockley, J. B. M. Jowett, I. M. Jones, M. Garreau, and D. Thomas.** 1994. Fullerene-like organization of HIV gag-protein shell in virus-like particles produced by recombinant baculovirus. Virology. **198**:288-96.

14. **Schmid, M., R. Dargahi, and M. Tam.** 1993. SPECTRA: A system for processing electron images of crystals. Ultramicroscopy. **48**:251-64.

15. **Zuber, G., J. McDermott, S. Karanjia, W. Zhao, M. F. Schmid, and E. Barklis.** 2000. Alternative assembly of retrovirus capsid-nucleocapsid proteins in the presence of membranes or RNA. Submitted.

FIGURES

Figure 1. Projection structure of membrane-bound his-HIVCA proteins. Approximately nine unit cells of a two-dimensional reconstruction of membrane-bound his-HIVCA proteins are shown as a 202.7 Å x 200.0Å panel in which electron dense (protein) regions are light. Merging was performed with no symmetry assumed (p1), using reflections above a resolution of 20 Å and with high signal to noise ratio (IQ $\leq$ 4) for the phase origin search. Amplitude and phase values for each reflection were vector-averaged and then back-transformed to give the projection structure.
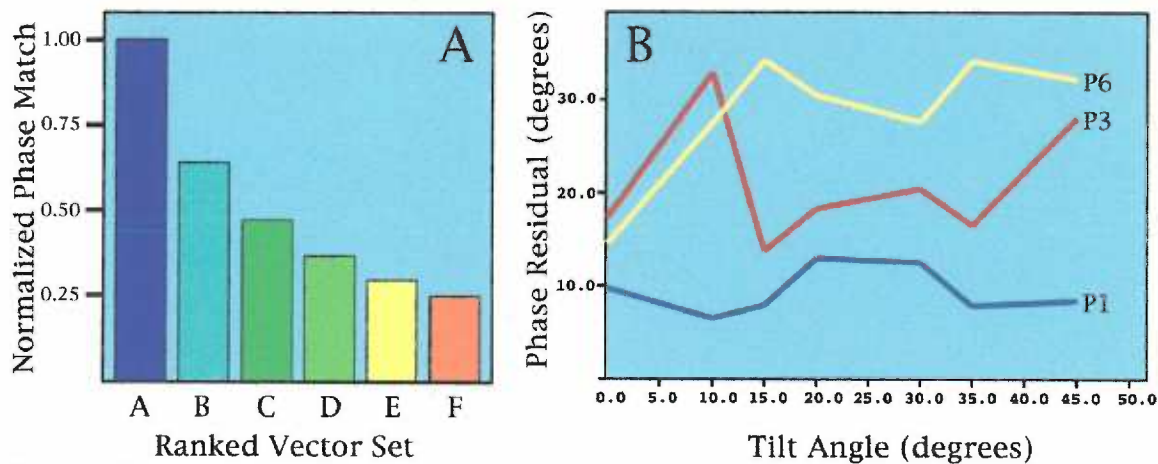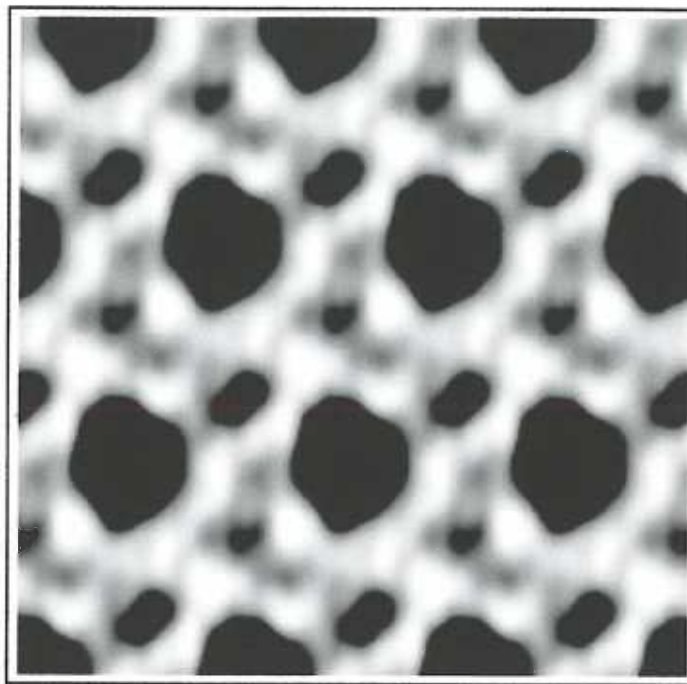

Ranked Vector Set

Tilt Angle (degrees)

Figure 2. Space-group assignment of his-HIVCA crystals.
A. Shown is an analysis of the consequences of indexing diffraction patterns using the six possible $\gamma^*=60°$ axes sets. Each diffraction pattern was indexed in the six possible indexing schemes, and all indexing schemes for each image were tested in p1 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 88 images. For display purposes, all top, second, third, fourth, fifth, and sixth place vector sets were averaged. Ideally a strictly primitive (p1) diffraction pattern will show only one high phase match value; a trigonal (p3) pattern will show three high values, and all six vector sets should give equally high values for hexagonal (p6) diffraction patterns. **B.** The full index-optimized his-HIVCA dataset, consisting of 88 tilted and untilted images, was merged assuming p1, p3, or p6 space group symmetry. For each of the three merges, phase residual values from individual images were grouped according to their tilt angles and averaged. Results for p1 (blue), p3 (red), and p6 (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.
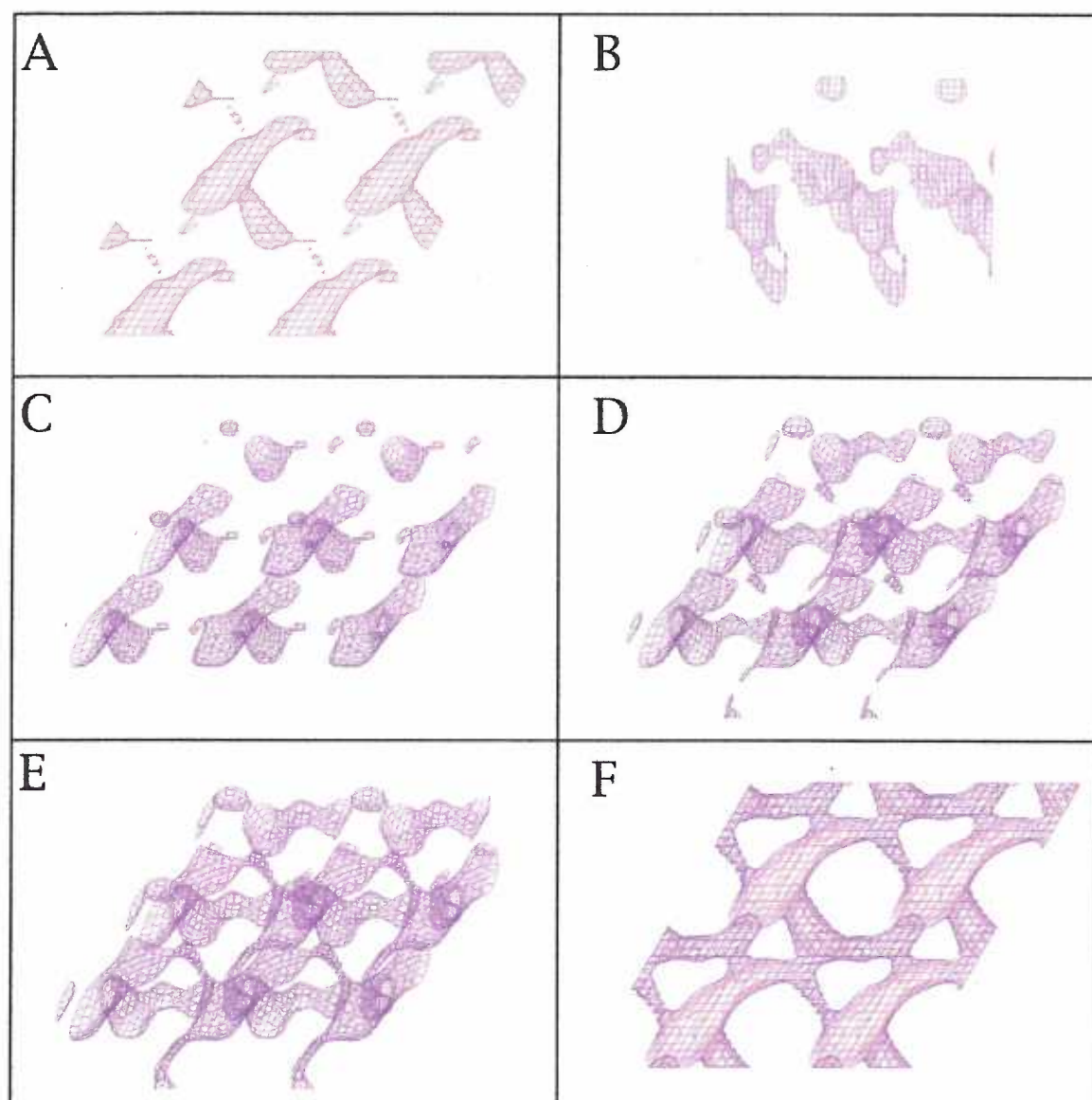
Figure 3. Membrane-bound his-HIVCA protein three-dimensional reconstruction.
The his-HIVCA dataset, consisting of 88 images of negatively-stained his-HIVCA crystals, was merged into a p1 unit cell, binned, filtered, and back-transformed into a 3D electron density map as described in the Materials and Methods. Protein density was localized in the central 120 Å Z-axis slab, and the map was displayed using the XFIT program from XTALVIEW. **(A)** Approximately four unit cells are shown (a=b=130Å) in a view perpendicular to the membrane with the membrane face closest to the viewer at a density level $\sigma = 2.0$. **(B)** The map viewed parallel to the membrane viewed along the a axis, and along the protein strands, at a density level $\sigma = 1.8$. **(C-E)** Viewed at 40° from perpendicular with the membrane face closest to the viewer at the following density levels: **(C)** $\sigma = 2.0$, **(D)** $\sigma = 1.8$, and **(E)** $\sigma = 1.75$. **(F)** Viewed perpendicular to the membrane, membrane face closest, a 50 Å, membrane-distal slab on the Z-axis, showing the putative C-terminal domain connections at a density level of $\sigma = 1.5$.
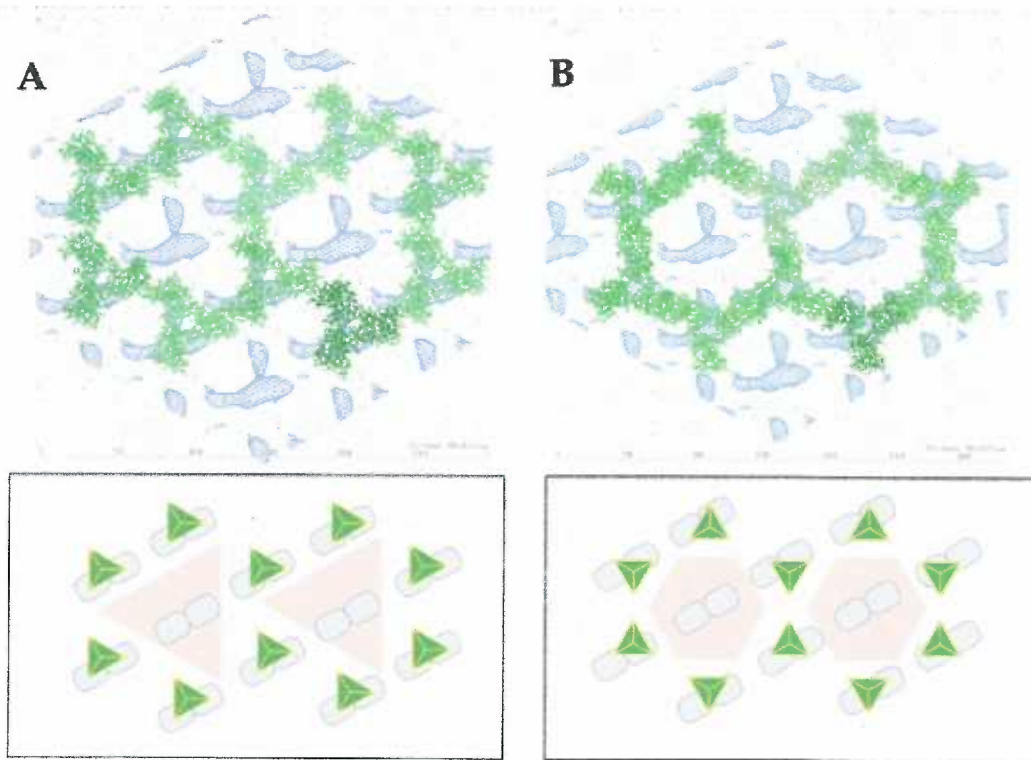
Figure 4. Model of MA trimers fit to CA dimers. Matrix trimer structures (green) were fit to the projection structure (blue) from his-HIVCA proteins assembling on lipid monolayers by matching MA and CA monomers one-to-one in a repeatable fashion. **A.** All trimers are oriented the same direction (top). Schematic representation (bottom) with MA trimers in green, CA dimers in blue and holes in pink. **B.** MA trimers oriented in different directions (top). Schematic representation (bottom).