Dissertation submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of Doctor of Natural Sciences

Presented by Alex de Marco

Born in Trieste, Italy

Oral examination

Analysis of retroviral assembly and maturation using cryo-electron tomography

Referees: Dr. Anne-Claude Gavin

     Prof. Dr. Hans-Georg Kraeusslich

Since no one that knows me would

ever expect any dedication in this

I dedicate it to my wife

Index

Abstract

Retroviruses are a family of membrane-enveloped RNA viruses that can retro-transcribe and integrate their genome in the host cell chromatin. During the active production of virions the structural polyprotein Gag assembles together with the genomic RNA to form immature particles, which bud out from the host cell. After budding the Gag protein undergoes proteolytic maturation and is cleaved into MA, CA, NC, p6 and two spacer peptides, SP1 and SP2. This leads to dramatic changes in the core morphology and the gain of infectivity. The immature retro-virions are known to have Gag organized into a round, incomplete hexameric lattice with a spacing of ~7.5nm. The core of mature virions is organized into a mixture of hexamers and pentamers which are organized along a lattice with a spacing of ~9.6nm. The shape of the core in the mature virions is genus-dependent, but can be cylindrical, conical or round.

During my PhD I have studied the immature Gag assembly across four retroviral genera in order to understand the structural requirements for the assembly of the immature retroviral lattice, and to shed more light on the principles of HIV-1 maturation.

The major conclusions of my studies are the following: The CA region of Gag is the most structurally conserved across genera. The presence of a domain upstream of CA is not critical for the assembly although it stabilizes the lattice. In order to maintain an immature lattice is important to have a Gag multimerization domain downstream of CA. The region between CA and NC, which is highly variable, is not critical for the assembly but it can stabilise the lattice and therefore affect the structural changes that occur during the maturation.

The maturation in retroviruses is an extremely fast process. In order to investigate the structural changes occurring during the maturation in HIV-1 I analysed the products of partial Gag maturation, which were obtained through selective mutations of the cleavage sites in Gag. This confirmed that the order in which Gag cleavages occur is important for a correct processing. The immature Gag lattice is destabilized only if both sides of the CA-SP1 region are cleaved. Furthermore, it showed that the condensation of the RNP has an effect on the core morphology in the mature virion.

Zusammenfassung

Retroviren sind membranumhüllte RNA-Viren, die ihr Genom revers transkribieren und in die Erbinformation der Wirtszelle einbauen. Während der aktiven Virusproduktion assembliert das virale Strukturprotein Gag zusammen mit der viralen RNA zu unreifen Viruspartikeln, die von der Wirtszelle knospen. Nach der Knospung wird das Gag-Protein proteolytisch in die Untereinheiten MA, CA, NC, p6 und zwei kurze Peptide, SP1 und SP2 gespalten. Diese Reifung führt zu dramatischen Veränderungen der Virusmorphologie, durch die Retroviren infektiös werden. In unreifen Retroviren sind Gag-Moleküle sphärisch in einem hexamerischen, lückenhaften Gitter mit einem Abstand von ~7,5nm angeordnet. In reifen Viruspartikeln ist Gag in Pentamere und Hexamere mit einem Gitterabstand von ~9,6nm angeordnet. Die Form dieses reifen Viruskerns ist gattungsabhängig und kann zylindrisch, konisch oder sphärisch sein.

In meiner Doktorarbeit habe ich die Anordnung von unreifem Gag in vier retroviralen Genera untersucht. Ziel war es, die strukturellen Voraussetzungen für den Aufbau des unreifen retroviralen Gitters zu klären und neue Erkenntnisse über die Grundlagen der Reifung von HIV-1 zu gewinnen.

Die wichtigsten Ergebnisse dieser Arbeit sind folgende: die CA-Region von Gag weist den höchsten Grad struktureller Konservierung auf. Für den Gitteraufbau ist die CA vorgelagerte Domäne nicht essentiell, jedoch stabilisiert sie das Gitter. Eine CA nachgeordnete Multimerisierungsdomäne ist wichtig für die Aufrechterhaltung des unreifen Gitters. Die Region zwischen CA und NC ist hochvariabel und nicht essentiell für den Gitteraufbau, jedoch kann sie das Gitter stabilisieren und somit die strukturellen Veränderungen während der Reifung beeinflussen.

Die Spaltungsprozesse während der Reifung von Gag sind extrem schnell. Um die Strukturveränderungen bei der HIV-1-Reifung zu klären habe ich unvollständig gespaltene Gag-Varianten, die durch selektive Mutation der Spaltstellen hergestellt wurden, untersucht. Ich konnte zeigen dass die Reihenfolge der Spaltvorgänge für die korrekte Reifung von Gag wichtig ist. Nur bei der Spaltung beidseitig der CA-SP1-Region wird das unreife Gag-Gitter destabilisiert. Die Verdichtung des RNP hat zudem Auswirkungen auf die Morphologie des Gag-Gitters in reifen Viruspartikeln.

1 Introduction

This manuscript will introduce and discuss the major findings related to my PhD project. During the last three years I have investigated two key steps in the replication cycle of the retroviruses: the assembly and the maturation.

This is a cumulative thesis; therefore most of the findings are widely discussed in the papers that are included. Here I will give a broad introduction about the retroviruses in order to present the biological problems I have been addressing. Then I will give a brief technical introduction about the methodologies I used. In the chapter 4 I will discuss briefly the results, and the conclusions that have been included in the published papers. A copy of the papers can be found in the appendix after the bibliography. The order in which they appear is not chronological but I chose to organize them in a way that better describes the order of the events in the retroviral replication cycle.

In the sections 5 and 6 I will discuss how the results of my work relate to existing knowledge in the field of retrovirus research, and the possible future experiments to investigate the questions that are still open.
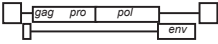
2 Retroviruses

2.1 Taxonomy
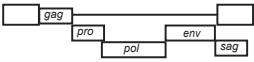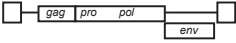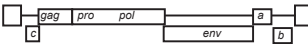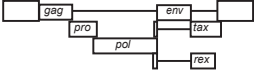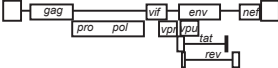
Retroviruses are a family of retro-transcribing RNA viruses. They have an RNA genome, which is transcribed by the viral reverse transcriptase into a dsDNA that integrates into the genome of the infected cell. All retro-virions are enveloped, and their diameter can vary between 80 and 140nm. All retroviral genera display viral glycoproteins incorporated in the lipid bilayer, while the organization of the core is genus-specific (Coffin JM, 1997).

The genome of all retroviruses contains three major coding domains: *gag*, *pol*, and *env*. They code for the main structural elements, the viral protease, the polymerase (reverse transcriptase), the integrase, and the surface glycoproteins. A first and broad division of the retroviruses can be made depending on whether the virions carry only these three genes (simple retroviruses), or whether other viral regulatory and accessory proteins are coded through multiple splicing sites (complex retroviruses). Retroviruses are further divided into seven genera plus the unclassified retroviridae (Table 1).

2.2 Genetic organization

As discussed in the previous paragraph retroviruses can be classified in two main categories based on their genetic composition, simple and complex retroviruses. A common feature is that the viral genes in the provirus (integrated viral DNA) are flanked at the 3' and 5' end by the long terminal repeats (LTR). These sequences are identical at both ends and can be divided into three elements. In the provirus, starting from 5' to 3': U3 corresponds to the 3' end of the viral RNA; R is a repeated sequence that is identical at both ends of the viral RNA; U5 corresponds to the 5' end of the viral RNA. The transcription initiation site sits between U3 and R, while the polyA addition to the transcript is between U5 and R. Proceeding toward the 3' we find a splicing donor site, then the genome-packaging sequence (ψ) that is required for the correct identification of the viral RNA during the assembly of the virions (Jouvenet et al, 2008; Jouvenet et al, 2009; Lu et al, 2011). This sequence is followed by the three genes *gag, pro,* and *pol*. The expression of these genes is regulated differently in the various genera. After *pol* we find the major splicing acceptor

sequence. The last gene before the second LTR is *env*. The complex retroviruses have multiple alternative splicing sites in the region between *pol* and the second LTR. These splicing sites are required for the expression of viral accessory proteins (Table 1).

	Genus	Example of specie	Example of genome
simple / oncoviruses	Alpharetrovirus	Avian leukosis virus (ALV) Rous sarcoma virus (RSV)	
	Betaretrovirus	Mouse mammary tumor virus (MMTV) Mason-Pfizer monkey virus (M-PMV)	
	Gammaretrovirus	Murine Leukaemia virus (MuLV) Feline Leukaemia virus (FeLV)	
	Hepsilonretrovirus	Walleye dermal sarcoma virus Walleye epidermal hyperplasia virus 1	
complex	Deltaretrovirus	Human T-cell leukaemia virus (HTLV) Bovine leukaemia virus (BLV)	
	Lentivirus	Human immunodeficiency virus (HIV) Simian immunodeficiency virus (SIV)	
	Spumavirus	Human Foamy virus	

Table 1

Overview of the taxonomic classification of the retroviruses. The genera are ordered according to their complexity, and according to their oncogenic potential. There is a scheme of the genome for one species per genus to explain the genetic variability between genera.

2.3 Viral proteins

2.3.1 Gag-derived proteins

Gag is a polyprotein (in HIV-1 54kDa) that is composed of three main proteins: membrane-associated or Matrix (MA), Capsid (CA) and Nucleocapsid (NC). Depending on the retroviral genus a variety of peptides and proteins can be present within Gag. See Fig. 1A for an overview of Gag in the alpha-, beta-, gamma-, and lenti-virus. In the next 4 paragraphs I will describe the most conserved parts of Gag and the peptide sequences relevant to my work.

A

Alpha-retroviruses

| MA | | p10 | CA | | NC | PR |

p2   SP

Beta-retroviruses

| MA | p24 | p10 | CA | | NC | p4 |

Gamma-retroviruses

| MA | p12 | CA | NC | |

Lenti-viruses

| MA | CA | | NC | | p6 |

SP1   SP2

B

MA   CA-NTD   CA-CTD   NC
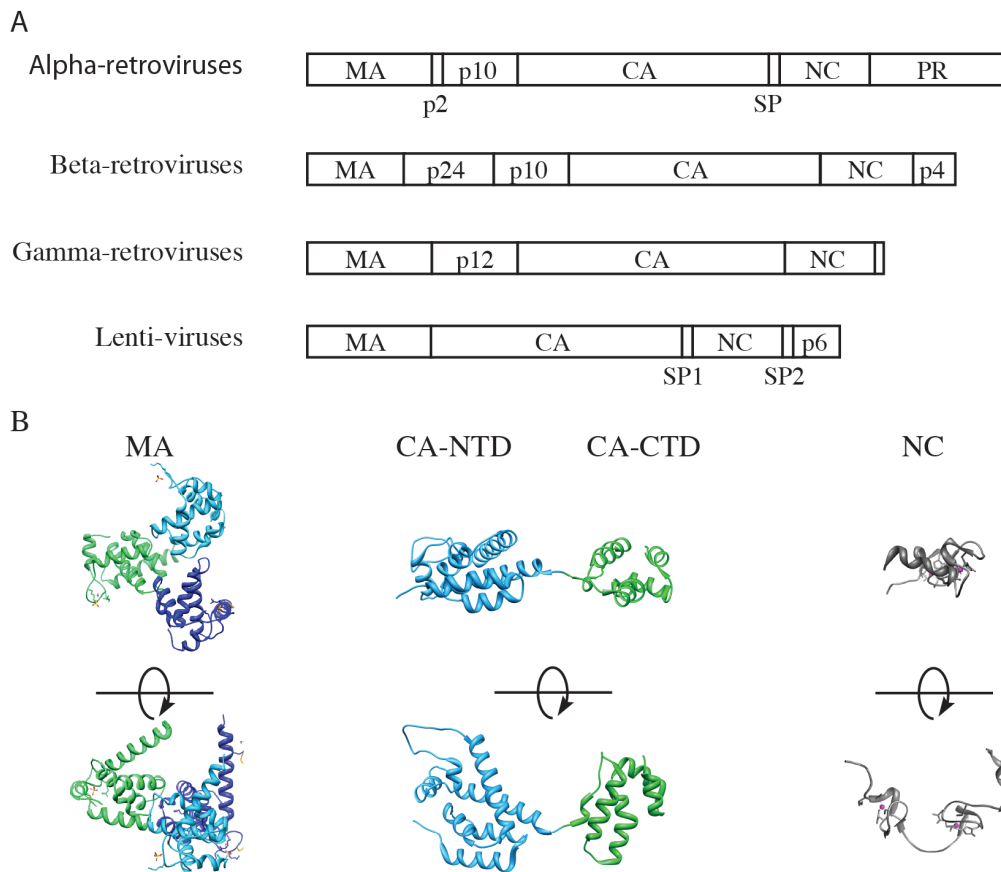
Figure 1
Structure and organization of the structural protein Gag. A) Schematic representation of Gag in the four genera treated in this manuscript. B) Crystal structures of the individual domains of Gag from HIV-1. MA is the only structure represented as a trimer.

2.3.1.1 Matrix

MA is the N-terminal domain of Gag in all retroviruses. This protein associates with the plasma membrane of the host cell. In fact in most of the genera the N-terminal is myristoylated (Shoji & Kubota, 1989). The structure of the MA has been solved in most of the genera by crystallography or NMR. MA from HIV-1 and SIV in the crystal packing showed a trimeric arrangement (see Fig. 1B)(Hill et al, 1996; Rao et al, 1995). The structure presents five helices linked by β-strands or loops. The shape of the trimer suggests that the myristates can insert in the membrane. Even if the crystal structures suggest a trimeric arrangement of the proteins, there is no evidence that in the virions this arrangement is conserved, since all the structural analyses have been made on the non-myristoylated form (Zhou & Resh, 1996).

MA proteins in some genera are capable to bind RNA *in vitro*. There is evidence that in the Bovine Leukaemia Virus (BLV) MA can specifically recognize a sequence located at the 5' of the viral RNA (Katoh et al, 1991). The role of this affinity in the virus replication cycle is not clear yet, and this interaction has never been described in virions (Steeg & Vogt, 1990).

2.3.1.2 Capsid

Among the Gag-derived proteins CA is the most conserved. CA has approximately 230 residues (depending on the species). It contains a highly conserved 20 amino acid sequence called the major homology region (MHR). The MHR is conserved across all genera with the exception of the spumaviruses.

CA consists of two domains that fold independently (CA-NTD and CA-CTD) and that are linked by a disordered region. The structures of both the CA-NTD and CA-CTD have been solved in many retroviral species (Jin et al, 1999; Khorasanizadeh et al, 1999; Kingston et al, 2000; Mortuza et al, 2004; Nandhagopal et al, 2004). In all the cases the CA-NTD and the CA-CTD have been solved independently because a flexible linker separates the two. For the CA-CTD the structure has been solved both in monomeric and dimeric form. The supramolecular organization of CA in the virus has been solved, for both the immature and the mature stages. Although the structure of the immature stage has not yet sufficient resolution to unambiguously define the arrangement (see paragraphs 2.5 and 2.8).

2.3.1.3 Nucleocapsid

NC is a small basic protein that recognizes, binds and organizes the genomic viral RNA. In all retroviruses with the exception of the spumaviruses NC has a characteristic Cys-His motif with this structure: $CX_2CX_4HX_4C$, where X represent a non-conserved residue. This motif is very similar to the Cys-His motif present in the "zinc fingers" of non-viral nucleic acid-binding proteins. NC has been shown to bind Zn^{2+} *in vitro* (Bess et al, 1992; Chance et al, 1992), and any alteration or deletion of its sequence leads to a loss of the viral RNA in the virions (Gorelick et al, 1988). NC recognizes the viral RNA through the ψ element, which is important to distinguish the viral RNA from the pool of cellular RNA (Aldovini & Young, 1990; Gorelick et al, 1990; Meric & Goff, 1989; Meric et al, 1988). NC can specifically bind multiple sites along the RNA due to their secondary structures; some of these interactions are important to discriminate between the spliced and the unspliced RNA (Lu et al, 2011).

2.3.1.4 Peptides

A number of peptide sequences are located between the three conserved main proteins of the Gag polyprotein. These sequences lie between MA and CA, between CA and NC or downstream of NC. The function of most of these peptides is poorly understood.

One of the most interesting sequences sits between CA and NC. This region is poorly conserved across genera and can be flanked by one or two proteolytic sites, depending on the genus. The complete processing of this region is critical for the correct assembly and for the infectivity of the virus (Krausslich et al, 1995; Wiegers et al, 1998).

Downstream of NC in lentiviruses there is an unstructured peptide (p6) whose presence is critical for the virus budding. p6 is rather variable, but it contains a late domain motif, which is required for the recruitment of the cellular machinery ESCRT to promote scission during the budding of the virions (Demirov et al, 2002; Gottlinger et al, 1991; Huang et al, 1995). All known retroviruses have an equivalent motif which is responsible for the recruitment of cellular factors to promote the budding (Puffer et al, 1997). The absence of the late domain leads to an incomplete budding, and the virions remain attached to the plasma membrane (Demirov et al, 2002; Gottlinger et al, 1991).

2.3.2 Proteins derived from *pro* and *pol*

All retroviruses carry three essential enzymes that are encoded by *pro* and *pol*: the reverse transcriptase (RT), the integrase (IN) and the protease (PR) (Coffin JM, 1997).

The *pol* ORF is almost never expressed independently (spumaviruses are an exception): it is always expressed as part of a larger fusion protein with Gag. The relative amount of Gag and Gag-Pro-Pol has to be regulated to be ~10% of the Gag transcript, since Gag-Pro-Pol alone cannot form a stable assembly. This has been tested only *in vitro* (Felsenstein & Goff, 1988; Park & Morrow, 1991), but the ratio appears to be strictly regulated in the virus as well. In alpha-retroviruses and lentiviruses a translational frameshift allows the formation of a Gag-Pro-Pol fusion protein (Jacks & Varmus, 1985). In contrast, in gamma- and epsilon-retroviruses the *pol* ORF is in the same reading frame of *gag*, and the translation happens thanks to a translational read-through that suppresses the stop codon (Yoshinaka et al, 1985).

RT is a RNA dependent DNA polymerase; this enzyme is essential to generate the proviral DNA upon infection. This enzyme has been demonstrated to have also an RNAse H activity, which is important since it degrades the viral RNA when this is hybridized with the DNA strand after the retro-transcription (Molling et al, 1971; Molling et al, 1973; Tanese & Goff, 1988).

IN is an enzyme that promotes the integration of the proviral dsDNA after retro-transcription. IN selectively binds the proviral dsDNA, removes two nucleotides from the 3'-end and joins it to a target DNA. The target DNA is preferably a large intronic region of highly transcribed genes, even though a fixed rule has not yet been found (De Marco et al, 2008; Han et al, 2004; Lewinski et al, 2005).

PR is the viral protease that plays a crucial role to pass from an immature and non-infectious form to a mature and infectious one. PR cleaves the polyprotein Gag in the late assembly phase or after budding (Coffin JM, 1997; Vogt, 1996). The sequence conservation in PR is not very high across genera, but in the active site the sequence D-S / T-Q is preserved identical in all species apart from HFV where the site is D-S / T-G.

2.3.3 *env* derived proteins

The membrane envelope of retroviruses contains viral glycoproteins that are used to identify the possible targets for the infection and to mediate the binding.

All retroviruses produce two proteins from the common Env glycoprotein transcribed from *env*. These are SU and TM. The nascent polypeptide, as all the proteins destined for secretion, associates with the ER membrane, it is transferred to the ER lumen where the glycosylation takes place. Env is cleaved in the Golgi into SU and TM which remain attached to each other through non-covalent interactions and are transported through the cellular secretory pathway to the plasma membrane (Coffin JM, 1997). Env cleavage is essential for its normal function (Brody et al, 1994; Rein et al, 1994). The cytoplasmic tail of TM has been shown to interact with Gag in the immature virions (Wyma et al, 2000). Env is structurally organized as a trimer. Three copies of TM form a stem protruding out from the envelope surrounded by three identical copies of SU (Liu et al, 2008).

2.4 Replication cycle

The retroviral replication cycle is a complex procedure. The following section summarizes only the major steps, and a schematic representation of the replication cycle is shown in Fig. 2.

2.4.1 Host cell recognition, receptor binding and fusion

These are the first steps in the retroviral cycle. A number of events occur at this stage, which involves the binding of Env to the receptors on the target cell surface. The receptors recognized are different across the genera, and their cellular function is not related to their virus receptor activity (e.g. the lentiviruses recognize the antigen CD4; the gamma-retroviruses recognize a variety of membrane transporter molecules) (Miller et al, 1994; O'Hara et al, 1990; Sattentau, 1988; Takeuchi et al, 1992).

The fusion between the viral envelope and the plasma membrane is a temperature-dependent process, in fact although the virus can bind the plasma membrane at temperatures below 20ºC, the fusion occurs at temperatures closer to the physiological ones, i.e. 37-38ºC). This feature was shown to be common to alpha-retroviruses and lentiviruses (Frey et al, 1995; Gilbert et al, 1990).

2.4.2 Entry of the core and uncoating

The entry and the uncoating are the least understood steps in the replication cycle of the retroviruses. After its release in the cytoplasm the core is thought to undergo structural modifications, which lead to the formation of the pre-integration complexes. It has been shown how these can associate to the cytoskeleton to be
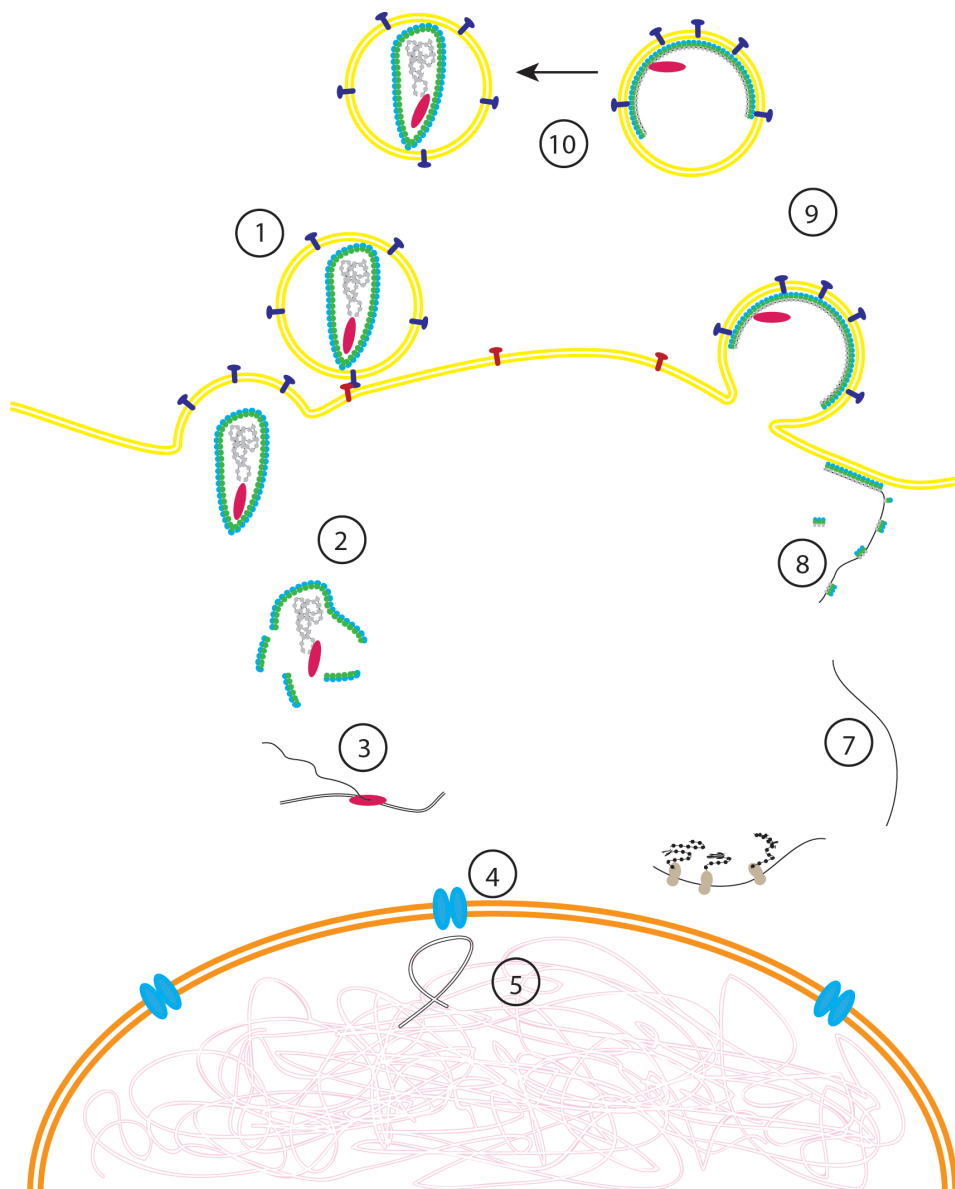
Figure 2

Replication cycle of a retrovirus. In yellow the plasma and viral membranes, in orange is displayed the nuclear envelope. Gag is represented in green/cyan/grey. The viral DNA is represented as a single black line, while the provirus (dsDNA) is represented as a double black line. In pink is represented the cellular chromatin. The blue trans-membrane proteins represent Env, while the red are the cell specific surface proteins that the virus recognizes (e.g. CD4 is recognized by HIV-1). The numbers summarize the steps: 1) host cell recognition and viral entry; 2) uncoating; 3) retro-transcription; 4) import of the provirus in the nucleus and integration; 5) activation of the transcription; 6) translation of the viral proteins; 7) export in the cytoplasm of the full-length transcript; 8) assembly on the plasma membrane; 10) maturation.

transported in proximity of the nuclear membrane (Bukrinskaya et al, 1998; McDonald et al, 2002). In the uncoating process some cellular proteins have been shown to play a decisive role, the most important appears to be the Cyclophilin-A, which destabilizes the core (Luban, 1996).

2.4.3 Reverse transcription of the viral genome to dsDNA

The reverse transcription is a process that produces one copy of the provirus from the genomic RNA and is carried out by the viral enzyme RT. Despite the presence of two copies of the genomic RNA in the core only one copy is retro-transcribed. Reverse transcription can initiate thanks to the binding of a cellular tRNA to a specific primer-binding site in the viral genomic RNA. The retro-transcription can start on one copy of the genome and end on the other, increasing the chances to recover information in case of damage of one of the two copies of the RNA (Coffin JM, 1997).

2.4.4 Nuclear import of the proviral dsDNA and integration in the host genome

The integration of the provirus in the cellular genome is a process promoted by the viral enzyme IN. This enzyme binds both ends of the provirus, selects the region for the integration and promotes the insertion of the provirus. The mechanisms involved in the selection of the region are not clear, although long intronic regions appear to be preferred.

2.4.5 Transcription of the provirus

The transcription is operated by the cellular RNA polymerase II. The full-length transcript will be packaged in the virions as viral genome. The amount of viral RNA produced in the cell is dependent on the position of the integration in the host genome. It is regulated by cellular factors and (in case of the complex retroviruses) by viral accessory proteins. The main factor that is involved in the activation of the transcription in complex retroviruses is the trans-activator protein (Tax in HTLV, Tat in HIV), which promotes the transcription by binding a specific sequence (transactivation-responsive region TAR) in the 5' LTR and recruiting the RNA Pol II machinery.

2.4.6 Translation of the viral proteins

In order to promote the assembly and the generation of new virions the retroviral transcripts are translated by the cellular ribosomes. The production of the individual

proteins is controlled (in complex retroviruses) through the alternative splicing. The most expressed proteins during the active production of new virions are Gag, Gag-Pol and Env.

2.4.7 Assembly of the virion and genome packaging

Gag is the main structural element of the retroviruses. During the assembly procedure it aggregates to form a shell, which binds a cellular membrane on one side and organizes two copies of the genomic RNA on the other. The complex retroviruses encode for accessory proteins that prevent the degradation of the viral RNA (e.g. rex in HTLV and rev in HIV). A more extensive description of the assembly procedure can be found in section 2.5.

2.4.9 Budding

This step is required for virus exit from the cell. Virus budding is promoted by Gag, which is able to bind the plasma membrane and to bend it to form the bud. Membrane scission is mediated by the cellular ESCRT machinery, which is recruited by the late domain contained in the C-terminal portion of Gag.

2.4.10 Proteolytic maturation

Immediately after budding the virions undergo a proteolytic maturation, which leads to dramatic changes in the organization of the core. Gag is cleaved in multiple positions in order to allow the re-arrangement of CA, and the RNP condensation. For a better description of the maturation and the mature stage see sections 2.7 and 2.8.

2.5 Assembly and immature lattice organization

The assembly of the newly formed virions is driven by Gag. There are two major ways for retroviruses to assemble, depending on whether the Gag assembly happens on the plasma membrane or in the cytoplasm.

In most of the retroviruses the assembly of the Gag shell occurs on the plasma membrane. Due to the hydrophobic or basic sequences in the MA and the myristoyl group, Gag targets the plasma membrane (Coffin JM, 1997). It is not yet clear whether Gag is transported as a monomer or in multimeric patches. Gag aggregates underneath the plasma membrane, most likely by side-to-side contacts to form a curved shell. This curved shell induces a curvature change in the plasma membrane, and generates a protruding bud. The bud can grow into a complete sphere, until it is

pinched off, and the virion is released. The structure of the immature virus is roughly
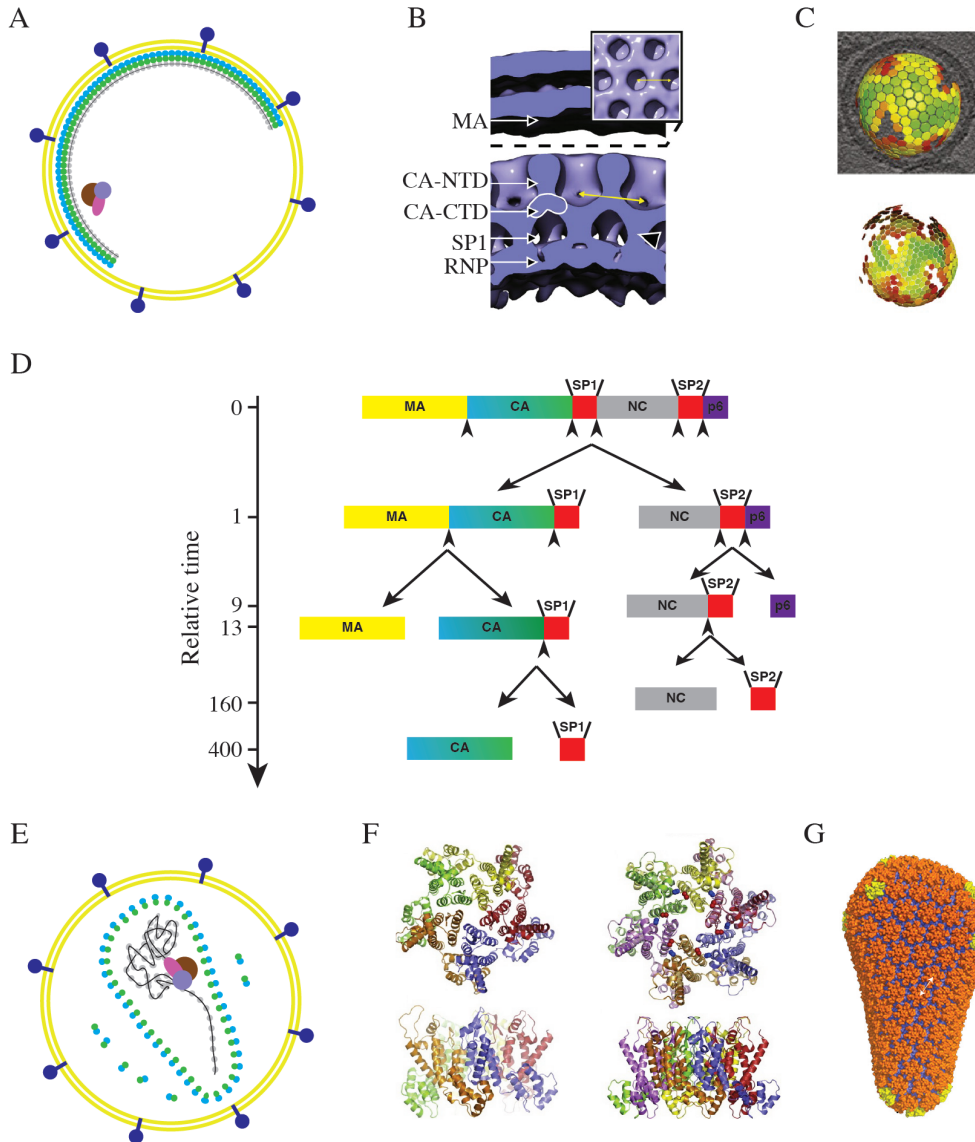
Figure 3

Organization and maturation of Gag in HIV. A) Schematic representation of an immature HIV-1 virion (the colour-code is the same as in Fig.2). B) *In virus* structure of the Gag immature lattice. The yellow double-headed arrow indicates the distance between hexamers, which at the CA-CTD is 7.4nm. C) Macroscopic organization of the Gag lattice in the virus. The colour (red to green) of the hexamers is an indication of the regularity of the lattice. Panels B and C have been adapted from (Briggs & Krausslich, 2011). D) Schematic representation of the proteolytic maturation. The black arrowheads represent the cleavage sites that have not been cleaved. The relative time scale shows when each cleavage occurs (the last cleavage is 400X slower than the first cleavage). The cleavage rates have been deducted from (Pettit et al, 2004). E) Schematic representation of a mature HIV-1 virion (the colour-code is the same as in Fig.2). F) Crystal structure of the pentameric and hexameric organization of CA in the mature arrangement. G) Three dimensional model of a fullerene cone assembled with the twos crystal structures showed in panel F. in orange are colours the hexamers, while in yellow are displayed the pentamers. The double-headed white arrow indicates a distance of 9.3nm. Panels F and G are adapted from (Pornillos et al, 2011).

spherical. Gag is radially arranged and its domains are organized toward the centre from the N- to the C-terminal.

The structure of the Gag shell has been solved for HIV-1 in purified immature viruses and *in vitro* assembled VLPs by cryo-electron tomography (Briggs et al, 2009; Wright et al, 2007). The CA domain is organized into a hexameric lattice with an inter-hexamer distance of 74Å. Toward the centre of the virion there are rod-like structures that connect the hexameric lattice of CA to a disordered layer, which can be identified as the RNP. The rod-like densities have been modelled as a six-helical bundle of the SP1 region (see Fig. 3A-C). It has been shown that the Gag lattice doesn't cover the 100% of the internal surface (Briggs et al, 2009; Carlson et al, 2008). In fact there is a main gap in the shell that most probably corresponds to the scission site after budding. Furthermore, instead of introducing pentamers in the lattice to induce the curvature (as in an icosahedron), the curvature is induced by the presence of irregular cracks and small changes in the lattice regularity (Briggs et al, 2009). At present it is still not clear which is the reason to prefer irregular defects to more a stable pentameric arrangement. The structures available do not allow us to answer these problems in a satisfactory manner.

A similar organization of CA appears to be adopted by MuLV (Yeager et al, 1998). Unfortunately a structure is not available for this virus, instead visual inspection of the virions and Fourier analyses of the Gag shell have pointed out the presence of a hexameric lattice with the same spacing measured in HIV. This type of information has been retrieved as well for RSV, and the outcome was that it also adopts a hexameric lattice with a spacing of 78Å (Briggs et al, 2006).

The assembly procedure of betaretroviruses and the spumaviruses are exceptions, since their assembly occurs in the pericentriolar region of the cytoplasm, from which the viruses are transported to the plasma membrane (Rhee et al, 1990). In order to have a correct transport of the Gag shell to the plasma membrane the myristylation of Gag is required as for the other genera (Rhee & Hunter, 1987; Rhee & Hunter, 1990). Studies from *in vitro* assembled VLPs of M-PMV CA-NC showed a hexameric arrangement, and irregular defects across the particles (Kuznetsov et al, 2007).

The average size of a retrovirus is between 80nm and 140nm and the amount of Gag molecules is related to the size of the particle (Briggs et al, 2004). Considering a

similar immature arrangement of the CA protein across genera, it can be estimated that the amount of Gag protein per virion is conserved.

The structural role of Gag-Pro-Pol protein in the formation of the immature Gag lattice, if any, is not yet clear.

2.6 Budding

For all retroviruses the exit from the host cell does not lead to lysis and cell death. Budding starts with the binding of the myristoylated N-terminal region of MA to the inner layer of the plasma membrane. In HIV-1 this process has been extensively studied, with the conclusion that the Gag shell induces the curvature of the plasma membrane. The recruitment of the ESCRT complex by the late domain induces the closure of the bud and the subsequent scission; the completeness of the Gag lattice is higher in late budding events (Carlson et al, 2008).

2.7 Maturation

All retroviruses require a proteolytic maturation in order to become infectious (the spumaviruses are an exception since Gag in this genus does not contain cleavage sites). The viral protease cleaves Gag in multiple positions (5 for HIV) during or immediately after budding. This set of cleavages leads to a dramatic change in the organization of the CA layer, the condensation of the RNP and the gain of the infectivity (Coffin JM, 1997).

The maturation process is extremely fast; therefore no intermediates can be seen *in vivo*. *In vitro* studies on the HIV-1 protease activity shed light on the timing of this process (Pettit et al, 2004). In Fig. 3D the steps and the relative timing of the single events are explained. Many studies made by mutagenesis and using protease inhibitors showed that not all the cleavage sites have the same importance for the infectivity, and there are indications that the order of the cleavages has a certain importance for the infectivity (Krausslich et al, 1995; Muller et al, 2009; Wiegers et al, 1998).

2.8 Mature virion

Once the maturation is completed, the viral core morphology has dramatically changed.

The shape of the mature core is genus-specific. Alpha- and gamma-retroviruses have a round shell; beta-retroviruses have a cylindrical core; lentiviruses have a conical

core. In all the genera CA is the protein that forms the core. The RNP it is generally organized into a condensed mass within the core (Coffin JM, 1997).

For both gamma-retroviruses and lentiviruses Fourier analyses of the mature virions showed a core where CA is organized in a hexameric lattice with a spacing of ~11nm (Briggs et al, 2003; Yeager et al, 1998). For RSV the structure of the *in vitro* assembled capsid has been solved. This showed to be a hexameric lattice, the particles reconstructed were icosahedral and the size of the hexamers and pentamers were consistent with the expected size in a gamma-retrovirus or lentivirus (Cardone et al, 2009).

The structures of *in vitro* CA assemblies have been solved by EM or X-ray crystallography. This allowed the characterization of the arrangement of CA in the basic units that form the core: the hexamer and the pentamer (Fig. 3E-G) (Ganser-Pornillos et al, 2007; Pornillos et al, 2009; Pornillos et al, 2011). The mature core in all the retroviruses has been modelled to contain both hexamers and pentamers. The relative arrangement of the two basic units can generate all the shapes that are generally seen in the retroviral cores; they all correspond to arrangements allowed for the fullerene (cylinders, icosahedrons, cones). In multiple retroviruses has been shown the presence of a linker between the envelope and one tip of the core (Briggs et al, 2003; Butan et al, 2008).

3 Methodological background

In this section I will give a brief explanation about the techniques I used during my PhD. I have performed structural analysis on retroviruses, which could be done through the use of electron microscopy (EM) and derivate techniques. The reasons to use EM are that the size of the features I have been investigating on was below the diffraction limit of the visible light and the EM techniques allow the *in situ* investigation of the protein arrangement. I have performed all the analyses on vitrified samples in order to not compromise the molecular arrangement.

In the next to sections I will explain the basics of the data collection and the image processing used.

3.1 Cryo-electron microscopy and tomography

Biological samples are extremely dynamic systems. Since they have high water content, in order to get structural information we can vitrify them and collect images using a cryo-electron microscope. Depending on the sample and on the type of information one wants to obtain, one can collect single images in each spot (cryo-electron microscopy (cEM)) or multiple images but at each at a different angle (cryo-electron tomography (cET)).

In both cases the sample is deposited on a copper grid covered with holey carbon. The vitrification is done through fast freezing, by plunging the sample in liquid ethane and then storing it in liquid nitrogen. This process immobilizes the sample, and the embedding vitreous water is transparent to the electron beam. If the sample is thicker than 1μm the vitrification can be obtained increasing the high pressure during the freezing procedure.

During the data collection it is very important to consider that the frozen biological specimens are highly sensitive to the electrons. Therefore a long exposure to the electron beam induces changes and damage of the sample. Images are collected applying the minimal dose possible in order to avoid or minimize the beam damage. This leads to a very low signal to noise ratio in the images, and therefore to a low resolution since the details are lost in the noise. Another consideration is that in order to have a complete three-dimensional map of the specimen we need to investigate its structural organization from as many angles as possible. If the sample is

homogeneous in shape, or if it has a defined geometrical organization, it is possible to apply cEM in combination with image processing techniques (e.g. single particle, icosahedral or helical reconstruction) to get a high-resolution structure, and a complete (or as much as possible complete) set of views.

If the sample has no homogeneous shape or if it is embedded in a complex environment (e.g. the cytoplasm) we can apply cET. This technique is based on the collection of a set of images of the same sample from different orientations (tilt series). These images are then convolved into a volume, which is a 3d computational representation of the sample. The spatial resolution achievable in cET is anisotropic, since it is technically impossible to have a complete angular coverage. The resolution in z is dependent to the maximum angle at which the tilt series are collected (typically 60-75 deg). The signal to noise ration in a cryo-electron tomogram is 20 to 50 times lower that in a typical cEM image. This reduced signal to noise is due to the fact that the maximum electron dose applicable to the sample has to be divided among the images in the tilt series. In order to make the resolution isotropic and improve the signal to noise ratio, if the sample contains a repetitive unit, it is possible to apply subtomogram averaging.

3.2 Subtomogram averaging

As for cEM, also in cET it is possible to improve the spatial resolution and the signal to noise ratio by applying image processing techniques. The image processing allowing this is called subtomogram averaging. The subtomogram averaging consists in the alignment and average of multiple copies of an object present in a tomogram (or multiple tomograms). The immediate advantage is that the signal to noise ratio improves significantly. Furthermore, if the orientation of the individual copies is random the resolution gains isotropy filling the angular gaps in the information. This results in a resolution improvement. The subtomogram averaging has been used to solve structures of purified viruses up to a resolution of ~2nm and for big cellular complexes to ~5nm (Beck et al, 2007; Briggs et al, 2009).

There are two types of information which can be retrieved from the output of a subtomogram averaging reconstruction: the structure of the repetitive unit that has been averaged at a higher resolution than in the raw tomograms; the precise position and orientation of the repetitive unit in the tomographic reconstruction (and therefore in the sample).

In the projects that are part of this thesis I used cET and subtomogram averaging in order to retrieve structural information about the retroviral immature lattice assembly, and the intermediates during the maturation. The choice of this technique was due to the fact that the retroviral immature lattice shows a conserved local arrangement but the global arrangement is variable.

4 Summary of the conclusions

In this section I will summarize the results and the major conclusions drawn from the papers I published during my PhD, plus the result of a collaboration, which has not been published yet. The conclusions do not appear in chronological order of the publications, but rather in an order that reflects the replication cycle of the retroviruses.

The section 4.1 summarizes the conclusions from (de Marco et al, 2010a; Hadravova et al, 2012).

The section 4.2 summarizes the conclusion from (Carlson et al, 2010; de Marco et al, 2010b)

4.1 Structural conservation across retroviruses

A common and critical step in the retroviral replication cycle is the assembly of the immature Gag lattice. The Gag polyprotein presents a high degree of variability across genera, but there are conserved features in CA. This conservation results in a similar local organization of CA in the mature virions (Cardone et al, 2009; Pornillos et al, 2011). At present a high-resolution structure of the immature arrangement from any of the genera is missing, but a structure at a resolution of ~2nm exists for HIV-1 (Briggs et al, 2009). There are various indications of a general conservation of the organization into a hexameric lattice with an inter-hexamer spacing of ~7.5nm in HIV-1, RSV and MuLV (Briggs et al, 2006; Yeager et al, 1998). For this reason we have solved the structure of the immature lattice organization in 4 retroviruses (HIV-1, M-PMV, RSV, XMRV) each from a different genus. This was done in *in vitro* assembled VLPs.

The VLPs from all viruses were variable in size and organized into a hexameric lattice with irregular defects. These defects can have various shapes, and most of the time they look like cracks, even though in the smallest particles they often tend to have a pentameric shape. The inter-hexamer distance in all four viruses was ~7.5nm at CA-CTD level. The intra-hexamer arrangement of the CA-CTD across genera was highly conserved, although the angle between adjacent hexamers was quite variable resulting in a variable particle size. We measured the variability of the particle size and attributed it to be a genus specific feature.

The CA-NTD is generally organized into a hexameric lattice shell outside the CA-CTD. In XMRV the CA-NTD appeared to be disordered, but this is likely to be an assembly defect related specifically to the construct used. While the assembly constructs used for M-PMV and XMRV did not have any portion of the upstream proteins (but it started from the 2nd residue of CA), HIV-1 and RSV needed to have some parts upstream, otherwise the particles either do not assemble properly or they form tubular arrangements.

The region immediately downstream of CA-CTD has no or poor conservation. Although this region is poorly conserved across genera it is critical for the assembly ant it shows a high structural conservation.

4.2 Budding and maturation in HIV-1

In order to better understand the assembly and the budding process of HIV-1 we have solved the structure of the Gag lattice directly on the budding sites. For this purpose we infected human glioblastoma cells with an adenoviral vector expressing Gag and Gag-Pro-Pol, and collected cryo-electron tomograms at the budding sites on the cell surface. From the tomograms the Gag layer appeared to be identical in the purified immature virions, *in vitro* assembled VLPs and in the buds. We applied subtomogram averaging on single buds in order to get more detailed information about the organization of the layer. The results suggested that CA assembles into a hexagonal lattice underneath the membrane and is connected to the RNP via rod-like structures. The CA structure is consistent with the structure of the immature Gag lattice, indicating that there are no major rearrangements after budding and before the maturation takes place. The hexameric lattice shows multiple irregular defects, which are comparable to the defects present in budded immature particles.

Once budded from the cell the virions undergo a series of proteolytic cleavages that dramatically modify the virion organization in order to become infectious. Since these events are extremely fast and happen immediately after the budding, it is very difficult to study the process in detail. In order to better understand it we have analysed a pool of mutants that are not able to complete the maturation due to point mutations at single or multiple the proteolytic sites. These mutants do not match perfectly with all the maturation intermediate states, but they provide information on the role and the importance of each cleavage site.

Our results suggested that the immature organization of CA is stable unless both sides of CA-SP1 are cleaved. In fact if only the site between MA and CA is cleaved the CA-p6 disconnects from the membrane envelope, but its organization is identical to the immature. If only the cleavages downstream of CA are allowed the CA organization does not change, and the RNP condenses into a globular mass connected to the CA lattice. The nature of this connection is not yet clear. Comparing the structure of virions that can cleave everything downstream of CA and virions that can cleave everything downstream of NC we noticed no difference; and comparing the cleavage efficiency of the proteolytic site between CA and SP1 we noticed that it was reduced due to the missing cleavage between MA and CA. We could conclude that the relative timing for the cleavage events is important for an efficient processing.

Interestingly, in some of the buds we observed a single thin and striated layer underneath the membrane. The structure of this layer appeared to be identical to the MA-CA-SP1 region of Gag in a normal immature lattice. Furthermore the lattice appeared to have the same type of defects present in the immature virions. Given the similarity to the phenotype seen in the MA-SP1 mutant the most likely hypothesis is that the proteolytic maturation in these buds took place too early. We also observed that these cases were characterized by the absence of the RNP, suggesting that the premature cleavage leads to the release of the RNP with subsequent production of non-infectious virions. Furthermore the completeness was generally higher, this consideration makes sense since the late domain is not anymore in place to recruit the ESCRT complex. This turned out to be a cell dependent feature.

The dramatic changes involved in the maturation process are not only related to the changes in the organization of the CA lattice: the RNP condenses into a globular mass that has to be included in the conical core. This process starts from the first cleavage (between SP1 and NC) and continues with two more cleavages. In order to shed light on the importance of the cleavages in this region and on the importance of SP2 we analysed a pool of mutants carrying point mutations on the cleavage sites downstream NC or missing SP2.

The results suggested that SP2 is not required for the infectivity, although it plays a role in the viral fitness. The cleavage of the p6 region appeared to be the most critical event for the infectivity, even though the effect appears to be stronger in case both cleavage sites upstream p6 are mutated. Surprisingly interfering with the cleavages in

this region has an effect on the core morphology, impeding in some cases the formation of a conical core. This effect is not directly related to the infectivity; in fact we found a high amount of irregular cores when suppressing the cleavage between NC and SP2, even though this mutant is fully infectious.

4.3 HIV-1 Leucine-Zipper

4.3.1 Background

The importance and the relative contribution of the domains of Gag during the assembly is a key feature that still needs to be investigated. A well-tested approach is to exchange the NC domain of Gag with the leucine zipper domain of the yeast transcription factor GCN4 (Accola et al, 2000). It has been proved that this exchange leads to a normal particle formation in *in vitro* assembled VLPs for RSV, and that the alteration of the Gag sequence upstream the NC protein leads to the formation of tubes independent of whether NC or its substitute leucine zipper is present (Johnson et al, 2002). An intriguing question is whether the Gag arranges to form a typical immature lattice when NC is exchanged with the leucine zipper, or whether it organizes in a different way. We have applied cET and subtomogram averaging on budding and purified HIV-1-Leu-Zip particles in order to characterize the arrangement of the Gag layer in absence of NC.

4.3.2 Methods

In order to apply cET on budding sites, similarly to what was done in (Carlson et al, 2010) we used an adenoviral vector to express HIV-*gag* and HIV-*gag-pol* in a Rev-independent manner. *gag* and *gag-pol* were modified in order to contain MA-CA-SP1-Leu-Zip as described in (Accola et al, 2000). The sample preparation, and the data collection were performed as described in (Briggs et al, 2009; Carlson et al, 2010; de Marco et al, 2010b). Averaging the rotation-autocorrelation functions of the 2D power-spectra computed along the surface of the buds and the viruses we retrieved the lattice information. Through this analysis we generated a Radius-Angle-Frequency (RAF) plot, which allowed us to identify the presence of an ordered lattice without applying subtomogram averaging (de Marco et al, 2010a).

The structure obtained by subtomogram averaging was computed in a reference-free manner in order to be confident that the outcome would not contain any bias. The structure was generates as described in (Carlson et al, 2010).
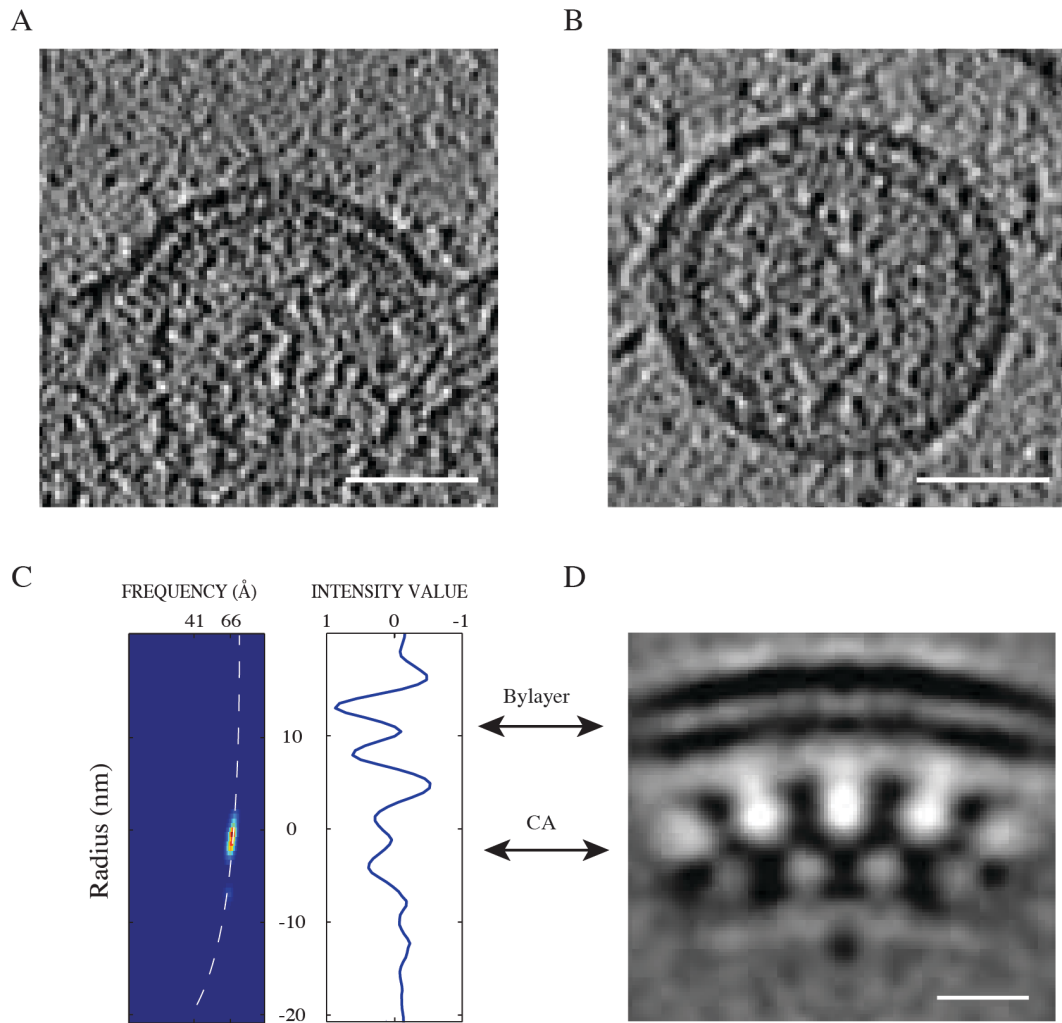
Figure 4

Structure and organization of MA-CA-SP1-Leu-Zip VLPs. A) Computational slice of 0.9nm thickness from a tomographic reconstruction of a budding VLP. B) Computational slice of 0.9nm thickness from a tomographic reconstruction of a purified VLP. The scale bar is 50nm long. C) RAF analysis of the purified particles. The peak shows a 6-fold symmetric arrangement with a spacing of 7.5nm in the striated layer. This can be attributed to the CA domain. D) Radial slice through the subtomogram averaging reconstruction of the purified VLPs. The structure has been aligned to the density and the frequency plots in the panel C. The scale bar is 1nm.

4.3.3 Results

From a visual analysis of the budding sites and the purified particles we could identify a single striated layer underneath the membrane (Fig. 4A-B). In general the particles appeared to be less rigid, and the curvature on the budding sites often was lower if compared to the particles containing NC.

We computed the RAF as done in (de Marco et al, 2010b; Hadravova et al, 2012) to verify the presence of an ordered arrangement of this layer. The peak distribution

suggested the presence of a hexameric lattice with an inter-hexamer distance of ~7.5nm (Fig. 4C). This distribution is typical of a retroviral immature lattice at the CA-CTD level.

The structure obtained by subtomogram averaging (Fig. 4D) showed a typical immature lattice arrangement in the CA-SP1 region. The Leu-Zip region does not form a dense layer (as we see when NC is present instead), while the densities are "hanging" toward the centre of the particle continuing the SP1 bundle.

4.3.4 Discussion

The lower curvature and the apparent loss of rigidity in the Leu-Zip particles if compared to virus particle containing NC can be explained by the absence of the continuous layer generally formed by the RNP.

From the RAF and from the cET structures it is clear that NC is not a structural determinant in the assembly of the immature lattice, although a multimerization domain is needed downstream CA in order to allow the immature assembly. The distance between the multimerization domain (NC or Leu-Zip) and the CA-CTD is crucial for the formation of the assembly (Accola et al, 2000). In order to assemble the concentration of CA needs to be high, but *in vivo* the concentration cannot be increased without being toxic for the host cell. The presence of a multimerization domain downstream CA provides sufficient interaction surfaces to allow the assembly with a low protein concentration.

5 Discussion and future perspectives

In this section I will discuss the structural features for an immature retroviral lattice, I will discuss the role of the various domains of Gag during the assembly, I will describe the structural changes that occur during the maturation process and toward the end I will describe in detail the characteristics of the mature virions. Furthermore I will describe some possible experiments to address the open questions.

5.1 Requirements for the immature assembly and stability of the lattice

In vitro CA is sufficient to arrange into a hexagonal lattice, which has the mature organization and spacing. Independently of the conditions used to promote the assembly, in order to generate an immature lattice the domains surrounding CA are needed. Different viruses have different requirements to generate an immature lattice. Therefore: "Which are the domains surrounding CA that are needed to have an immature lattice organization, and what is their relative importance?".

To form *in vitro* an immature lattice for both HIV-1 and RSV a sequence upstream CA is needed. In M-PMV and in XMRV a CA-NC construct is perfectly capable to assemble a VLP with an immature lattice. The NC region can be exchanged with the leucine zipper domain of the yeast transcription factor GCN4 to supply the interactions needed for the assembly, and this domain has to sit at a specific distance from CA (at least in RSV) (Accola et al, 2000).

The structures of the Gag lattices in 4 retroviruses are available at low resolution (Briggs et al, 2009; de Marco et al, 2010a; Hadravova et al, 2012; Wright et al, 2007). The arrangement of the CA region across genera is well conserved, and the different composition of the constructs analysed would suggest that the regions upstream of CA are not needed for the assembly of the immature lattice, even though these can under some conditions facilitate it. The arrangement of the region between CA and NC, which is not conserved across species, seems to be related to the need (or not) to have a domain upstream CA. In (de Marco et al, 2010a) we have modelled the SP1 region in HIV-1 and the SP region in RSV as a 6-helix bundle. In M-PMV we noticed that the likelihood that a 6-helix bundle is present is extremely low, and the structure confirmed that the region is not ordered into a rod-like structure. To explore the relative importance of upstream regions, CA-CA interactions, CA-SP interactions,

and SP-SP interactions in assembly, it would be interesting to generate chimeric constructs, for example constituted by CA and NC from HIV1 and the region between the last helix of CA and NC from M-PMV. My expectation would be that CA-CA and SP-SP interactions are predominant, and that this construct would assemble into an immature lattice. Along this line, further analyses could be done on most of the constructs that have been published in the last 15 years where an assembly was detected *in vitro* (often tubular), but no information has been retrieved about their lattice organization.

The CA-CA interactions are different between the mature and immature arrangement. The fact that there are no constructs that can assemble into an immature lattice without a multimerization domain downstream CA leads to the conclusion that the CA-CA interactions in the immature form are not sufficient to drive assembly. In (de Marco et al, 2010b) we demonstrated that in HIV-1 both sides of CA-SP1 have to be cleaved in order to promote the destabilization of the immature lattice. This would therefore suggest that although both CA and NC are required to assemble into an immature lattice, the region upstream CA is itself sufficient to stabilize the lattice after it is formed. The region upstream CA is as well not required to stabilize the lattice after it is assembled in HIV-1 since the CA-p6 variant showed to have an immature organization.

Summarizing, CA is not sufficient to form an immature lattice, it requires the presence of a downstream multimerization domain. The region upstream CA can be important during the assembly depending on the structure of the region between CA and NC. The region upstream CA is not required to stabilize the immature lattice although it keeps the lattice stable upon cleavage downstream CA.

5.2 Defects and curvature in the immature lattice

The Gag lattice has many features that are well conserved across genera. The CA region presents a highly conserved hexameric arrangement, the immature lattice is curved and the shells are not closed. The lattice is strewn with irregular defects in the lattice. These defects are extremely variable in size, and shape. Generally there is one extremely large defect or gap (up to ~1/3-1/2 of the virion surface), which corresponds to the point where the membrane scission occurred. The other defects are mostly cracks in the lattice 10-15nm wide. A close look to these lattices makes evident that the hexagonal arrangement is not perfect, in fact the relative orientation

and the inter-hexamer distance are slightly variable. This arrangement is reflected in the structures solved by subtomogram averaging where the central hexamer has a better definition than the side ones. In (de Marco et al, 2010a) we have shown how the curvature of the lattice can change across genera, as well as within a single particle. The particles are almost never spherical, they are generally elongated, and therefore even within a single particle more than one curvature is allowed. The fact that multiple curvatures are allowed by the same proteins means that either the CA-CA interactions can vary, or the flexible linkers present in CA allow different degrees of stretching, or a combination of the two. The features that appear to be genus-dependent are the mean curvature and the degree of variability. Considering the high conservation of CA and the relatively poor conservation of the surrounding regions I would argue that the curvature is a feature that depends on the regions surrounding CA. Since the possibility that the small differences in CA play a role in the curvature cannot be excluded, I would imagine that a screening of the allowed (present) curvatures from chimeric constructs (between HIV and RSV or M-PMV for example) could clarify and attribute the relative importance of each region.

The presence of defects upon bending a hexameric lattice, which is formed by rigid subunits, is expected. Considering that there are also small changes in the lattice spacing I would argue that the hexamers and the relationship between hexamers are not totally rigid. The presence of both, defects and small changes in the lattice spacing induces to think that the flexibility allowed by the CA-CA interactions is not high enough to allow the changes in curvature present in the assemblies, without destabilizing the overall assembly. The flexibility probably allows lower curvatures than the one observed in the particles. The imposition of a higher curvature than allowed by the CA-CA interactions could come from the RNP. In order to allow high curvatures in the CA shell an alternative to the irregular defects could be the introduction of pentamers or pentameric defects, to generate an icosahedral shell. The reason why the observed defects are not replaced by pentamers could be related to an intrinsic need of instability in order to allow the dramatic changes in the maturation.

Another possible reason to have irregular defects could be that these defects are the initiation sites for the proteolytic maturation. The defects in the lattice could be needed in order to allow the protease to access all the cleavage sites and cleave them in the correct order. This hypothesis can be supported by the fact that the cleavage

order is important to have an efficient processing of Gag. Knowing that the first cleavage that occurs is located between SP1 and NC, and knowing the supramolecular organization it is tempting to speculate that the protease can pass through the defects better than through the lattice itself. In this scenario, defects in the lattice would provide multiple initiation sites for processing at the same time, which would speed up the completion of the first cleavage in all the Gag molecules in the virion.

In order to properly address all the above open questions it would be important to obtain the structure of an immature virus at a resolution at which the CA-CA interactions can unambiguously be characterized. This structure will most likely come from *in vitro* assembly systems as it happened for the mature lattice. These systems allow having the lattice repeated on tubes or sheets, which allows the use of crystallographic or single-particle techniques to resolve the structure. Once the structure will be solved *in vitro* the obvious step is to verify it in the virus. In order to solve a high-resolution structure in the virus a major technical development of the subtomogram averaging technique is needed. Subtomogram averaging is generally needed in order to solve structures of repetitive objects whose local arrangement is conserved while the global arrangement is not. The structure *in virus* will not only verify the structure obtained from the *in vitro* assembly, but it will also allow an unambiguous interpretation of the role that the gaps in the lattice have in the organization of the virions. If there is any repetitive pattern in the organization of the edges we will be able to reveal it by applying subtomogram averaging. For example if the gap are initiation sites for the proteolytic processing the edges might have better exposed cleavage sites and a comparison between a high resolution structure of the lattice and of the edges would reveal it.

5.3 Model for the maturation

(Pettit et al, 2002) shed some light on the order of the cleavage events during the maturation. Their experiments were done *in vitro* but gave a good starting point to understand the process. The fastest cleavage separates the CA lattice from the RNP, allowing through the other cleavages two partially independent re-organizations. The scheme in Fig. 5 explains a possible model for the maturation. At present it still needs to be understood how the CA lattice disassembles from the immature organization and reassembles into the mature form. Understanding the structure of the CA-SP1 layer (not always present) in the CA-SP1 mutant would shed some light on this
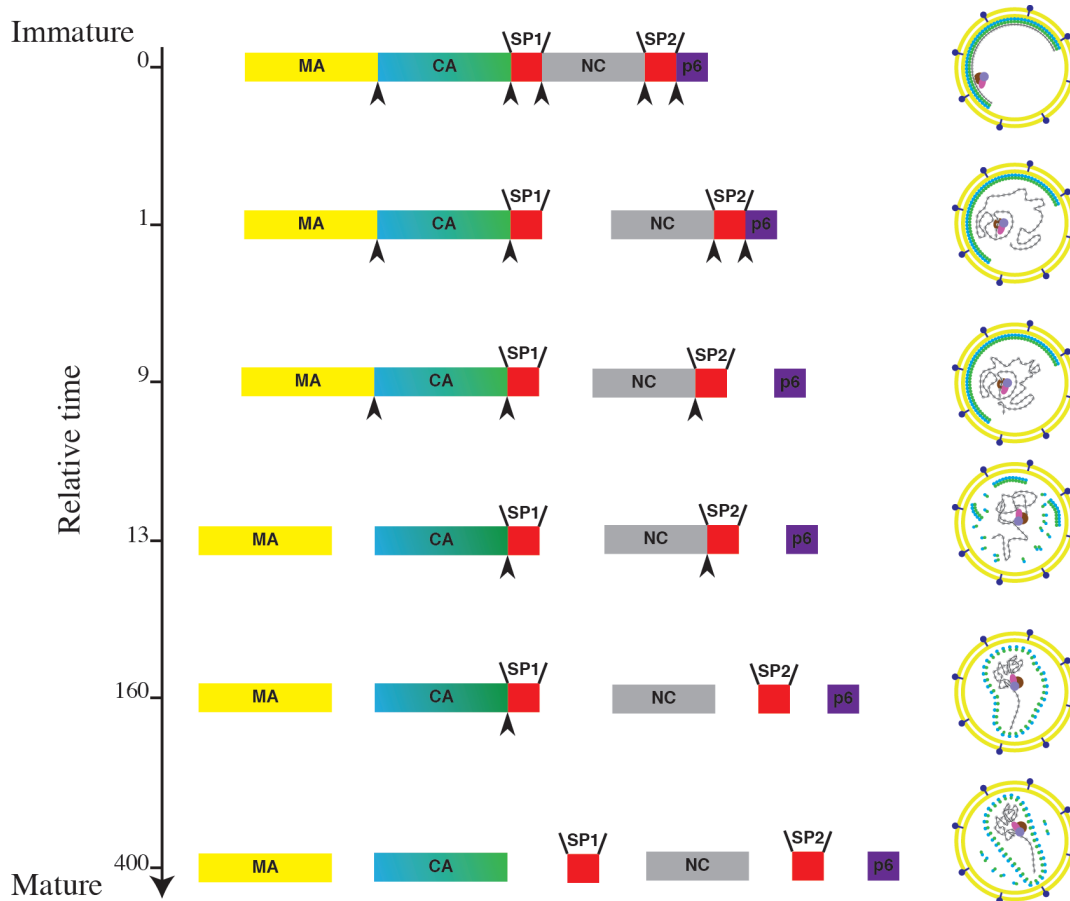
Figure 5
Schematic representation of the changes that occur during the maturation. The model has been made from the results in the (de Marco et al, 2010b) and (de Marco et al, submitted), and it shows which are the changes in the core morphology induced by the cleavages step by step. The colour-code is conserved from Fig.2, and the schematic representation of the immature and mature states have been adapted from (Briggs & Krausslich, 2011)

question. In fact it would clarify whether the re-arrangement occurs before or after the last cleavage, and eventually which is the role of the SP1 region in the maturation (e.g. impede the organization into a mature lattice or slow it down, impede the formation of pentamers or promote it).

In the replication cycle of many viral species there is a step that brings the virions from a non-infectious to an infectious state. This is generally a consequence of a proteolytic cleavage. The maturation in retroviruses occurs upon budding. There are indications that to start the cleavages either a drop in pH or a high concentration of protease is needed. Both are situations likely to occur in the virion after the budding. We have shown that the maturation event can start earlier, but it leads to non-infectious virions (Carlson et al, 2010). A possible reason could be that there are

different requirements for the assembly and for the infectivity. In this case the proteolytic processing allows big structural changes fulfilling all the requirements with one protein. For example the disassembly of the core after the virus entry would not work if Gag were not arranged differently from when it assembled, especially because the cytoplasm of the newly infected cell have the same conditions present during the assembly. Knowing that there are cellular proteins involved in the core disassembly of retroviruses (e.g. Cyclophilin A), a conformational change would prevent the action of these proteins during the assembly.

5.4 Structural changes during the maturation

The maturation is an extremely fast process; at present there are no techniques available for the direct visualization of the intermediate states. The most reliable approach is the one we have been using in the last years, which is based on the generation of artificial intermediate states. This approach allows the identification of the steps responsible for the major changes, but it is missing the dynamics involved in the process. The main limitation is that the generation of artificial intermediate assumes that each cleavage site is cleaved in all the Gag molecules before the next cleavage starts. This is unlikely the case: it is more probable that during the maturation exists a mixture of states, and that the order of the cleavage events is valid only within single Gag molecules or within small patches. At present I cannot imagine a better method to investigate the maturation, but in the next few years we might benefit from the progress in molecular engineering, which could produce, for example, a temperature-dependent protease, allowing the live investigation of the process.

5.5 Structure of the mature virus

Retroviruses are heterogeneous in shape and size at both immature and mature stages. In order to investigate the structure of the assembly a variety of methods have been applied. The mature CA assembly has been extensively studied from *in vitro* systems, using either X-ray crystallography or cEM. The structures solved allowed us to characterize the assembly. In HIV-1 cEM and cET have contributed to generate a model, which describes the cores as fullerene cones.

At present the local organization of the CA lattice in the mature virus has been solved, while the general organization has been indirectly understood measuring the aperture

angle of the conical cores (Ganser et al, 1999). It has been modelled as a fullerene. The solution of the structure of the cores *in virus* might help to determine the reasons for this arrangement, and clarify if these cores are effectively fullerene cones. Given the size and shape variability is not unreasonable to consider the presence of irregular defects in some of the cores. The general architecture of a fullerene cone/cylinder or a fullerite sphere is basically the same, with the only difference that in the fullerite the spacing between pentamers is constant. Understanding the reasons that bring to prefer an irregular shape to a more regular and stable could be crucial to understand the disassembly mechanism in HIV-1. Recently it has been shown that the cellular restriction factor TRIM5α assembles into a rigid hexagonal lattice, which has a spacing of exactly 4 hexamers of the HIV-1 mature lattice (Ganser-Pornillos et al, 2011). If this structure is present also in the cell during the viral infection, the presence of a conical core, which has the same properties of the fullerite but it is less regular might give the virus a better fitness reducing the chances of TRIM5α to bind the core. As supported by our results in (de Marco et al, to be submitted) the genome condensation could regulate the formation of the conical core in HIV-1. This speculation could explain the fact that the mutant NC-SP2 has been shown to be more infectious than the wild type beside it has an increased amount of "irregular" cores (Coren et al, 2007) and (de Marco et al, to be submitted). We classified the cores as "irregular" in any case these were not conical. The non-conical cores could present different degrees of irregularity (which we could not appreciate) and possibly in NC-SP2 these "irregular" cores still fulfil the requirements for the disassembly, but the action of the cellular restriction factor is less effective than in the wild type. At present it is hard to prove whether the shape of the core is defined as a response to the action of the restriction factors, since very little is known about the action of the various orthologs in of TRIM5α. All that is known is that the TRIM family has some highly conserved features needed for the retroviral recognition, which co-evolved across the mammalian species (Sawyer et al, 2007). In the next years, if the structures of the TRIM proteins from various species will be solved, it will be possible to better answer this question.

My hypothesis would be that the organization of the viral genome and its interactions with the NC proteins has evolved in order to promote the formation of a conical core in HIV-1. This could be due to the steric interference when a spherical core is
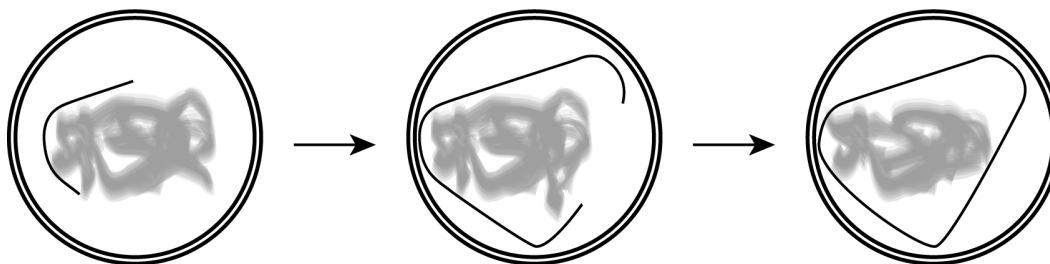
Normal genome condensation at the early stage of the assembly
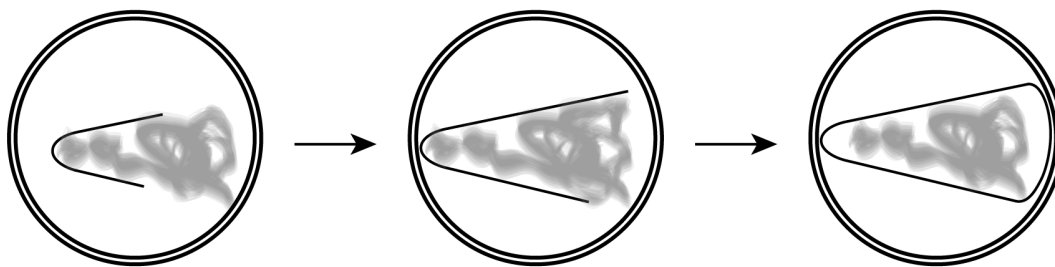
High genome condensation at the early stage of the assembly
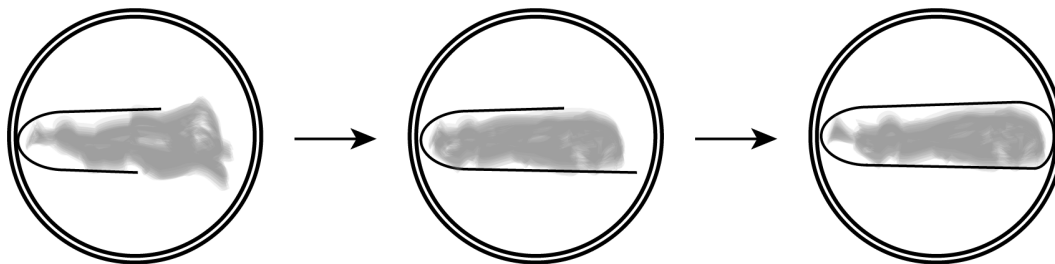
Low genome condensation at the early stage of the assembly

Very low genome condensation at the early stage of the assembly
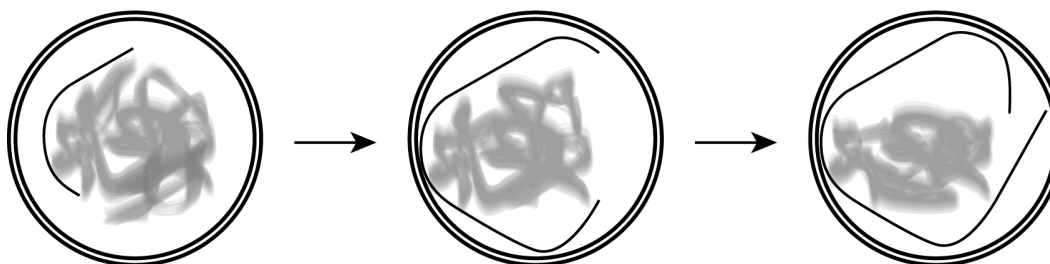
Figure 6
Schematic representation of the possible models for the relationship between the mature core shaping and the organization of the RNP. In case the condensation level is very high CA starts the assembly of the core with 6 pentamers, allowing the formation of a cylinder. With lower levels of condensation the number of pentamers decreases, inducing the formation of conical, or aberrant cores. In case the condensation is too low the core assembly starts with only one pentamer and the core in some cases cannot close. This effect has been described in (de Marco, to be submitted).

forming. The steric interference role played by the RNP as the major determinant in

the core shaping could explain also the formation various shapes of cores. A possible model that is in agreement with the model presented in (Briggs et al, 2006) is explained in Fig. 6.

Another possible reason to prefer a shape that is not as highly regular as the fullerite could be related to the requirements for the disassembly. The disassembly of the core could be regulated by its inherent instability: it has been proved that both destabilization and over-stabilization of the core can lead to a loss of infectivity (Forshey et al, 2002). An irregular cone is a system that has a relatively high energy, while a fullerite sphere is more stable. One could speculate that the core is kept intact in the virion (e.g. high CA concentration, pH, no or very little Cyclophilin-A in the virions), while it is destabilized in the cytoplasm.

5.6 Organization of the MA layer in HIV-1

In the immature virions the MA layer appears as a thickening of the viral envelope. At present there are no structures available to identify the arrangement that this protein has in the virus. The understanding of the arrangement of MA in the virions is a key point to better understand the assembly. MA is known to form trimers in the crystal packing, but from the structures of the immature HIV-1 no order has been seen, even though in two papers a faint peak (never discussed) in Fourier space has been seen for a 6-fold or 3-fold arrangement in that layer (Briggs et al, 2006; de Marco et al, 2010b). As future project I would investigate on the arrangement of the MA layer using cET and subtomogram averaging.

Another question is about the relationship between the MA layer and the cytoplasmic tail of Env. At present we know that Gag assembles underneath the plasma membrane where a high concentration of Env is present. It is not yet clear whether there is a stable interaction between these two proteins or if Env, once it has recruited Gag, can still move in the membrane. One way to address this question would be to get two parallel reconstructions from the same virion: one of the immature Gag lattice and one of the Env spikes. The information retrieved by the relative position of the CA hexamers relative to the spikes would answer this question.

After the cleavage between CA and MA during the maturation, the MA layer is not detectable anymore in the virions (de Marco et al, 2010b). The most likely explanation is that MA diffuses along all the internal face of the viral envelope (being covered before the cleavage 1/2 - 2/3 there is the possibility of a wide rearrangement).

Probably the easiest way to check this would be by immuno-gold labelling of MA and the determination of its position in the virions. If the outcome of this experiment showed that MA is diffused all along the envelope, it would be interesting to see whether the position of MA is in any way related to the position of the spikes. In case a stable binding exists between Env and MA in the immature virions the solution of the structure of Env by subtomogram averaging would allow the recovery of the density of MA. If the relative arrangements are not conserved at least a blurred mass should appear underneath the membrane. This has never been seen in the previously published papers on the structure of Env (Bartesaghi & Subramaniam, 2009; Liu et al, 2008; Zanetti et al, 2006), but no one has been looking in detail at this feature yet, and the published structures where always masked at the membrane.

5.7 Organization of the RNP

As we have shown in (de Marco et al, to be submitted) the correct organization of the RNP is critical in the replication cycle of the virus. But at present it is still not clear what are the determinants defining the correct RNP organization, which is the three dimensional organization of the RNP and which is the exact relationship with the core shape.

Investigating the determinants for the correct RNP condensation is an important step to understand the maturation and the encapsidation processes. (Berkowitz et al, 1995) demonstrated that the core is not anymore conical if the part of the NC is exchanged with the NC orthologs from MuLV. At present the sequences involved in the recognition of the viral RNA, and the ones involved in the binding of the NC proteins are known (Lu et al, 2011), but it is not yet clear which is the role of each sequence in the organization of the RNP during its condensation. A possible experiment to investigate on the importance of the relative position of the NC binding sites could be to change the length of the genome by adding or removing sequences in various points of the genome, and then check the effects that these changes have on the core formation. It would be as well informative to look at the changes in the core morphology across the strains in each species. Another option would be to modify the genome in order to impede the dimerization, or exchange the RNA dimer with a double length RNA. Some of these experiments have already been done in the past, but generally the information sought after was the amount of RNA included in the

genome and the specificity (Lu et al, 2011). No one has yet investigated the changes in core morphology.

Understanding the three dimensional organization of the RNP in both immature and mature virions would be a key step to understand the assembly and the maturation processes. Unfortunately at present the electron microscopy techniques do not allow to discriminate the organization of such a compact mass. A possible approach I can imagine would be single molecule light microscopy based. This would be based on the incorporation of fluorescently labelled RNA in the virions, and subsequent imaging by ground state depletion microscopy (GSD) and Fluorescence Resonance Energy Transfer (FRET) (Folling et al, 2008; Masi et al, 2010). These techniques are both capable of positioning an object or an interaction extremely precisely in space (10nm GSD and 0.1-1nm FRET). Knowing that the cores are all slightly different to each other, and all randomly oriented we might obtain a three dimensional map of the probability of the distance and orientation between various points along the genome. To do this we will need to collect an extremely large number or cores for each label, but coupling the single molecule fluorescence microscopy, the 3D-EM reconstruction methods, a complex probability analysis and the solution of an interesting biological problem could give rise to an exciting and high impact project. Beside the potential outcome of this approach there are some limitations to it: first of all are the actual costs of the *in vitro* synthesis of the RNA. In order to get a precise map of the RNP it is critical to tag no more than 2 residues per RNA, and this is feasible only by *in vitro* synthesis including modified bases in chosen positions. The second is that the map will be possible only if there is a certain degree of regularity in the RNP condensation. A possible option that might give some interesting outcome would be to mark the 5' and the 3' ends of the viral genomic RNA in order to see whether the relative position of the ends is conserved. Another suitable option could be to take advantage of the MS2 system. This system allows the visualization of the position of a specific RNA sequence thanks to the affinity of the Ms2-coat protein. Introducing this sequence in the viral genome and in parallel expressing the Ms2 coat protein tagged with a fluorescent protein suitable for PALM or STORM as could be mEos (Bates et al, 2007; Betzig et al, 2006; Rust et al, 2006), would allow the precise positioning of the two copies of the Ms2 target sequence in the RNP. A further advancement in this sense would be to couple the light microscopy with the EM,

taking advantage of the recently well-developed correlative techniques (Kukulski et al, 2011). This experiment would allow positioning of the sequence of interest in the core.

5.8 Technical development of the subtomogram averaging

The image processing techniques currently used in cEM (e.g. single particle, icosahedral and helical reconstruction) are capable to generate structures at a resolution below 10Å. The resolution obtainable with these techniques is dependent on the sample rigidity and regularity. If the sample has not a conserved global arrangement, or if it is heterogeneous in shape one would apply cET and subtomogram averaging in case a regular unit is present. At present this technique has been pushed to the resolution of ~17Å, and on the theoretical basis there are no reasons to not improve the limit of the subtomogram averaging to a sub-nanometre resolution.

Considering the fact that the samples generally analysed with this technique do not have a conserved global organization the main limitation is probably the flexibility. But the subtomogram averaging can be applied to highly regular structures that are located in a complex environment. Therefore we could be able to obtain the structure of cellular components *in situ*.

Currently the main limitations of the subtomogram averaging are divided across three categories:

1. Data collection
2. Tilt series processing
3. Improvement of the subtomogram averaging

5.8.1 Data collection

The main limiting factor for the resolution improvement I have identified during the data collection is related to the sample movement during the tilt series acquisition.

The sample during the data collection moves due to the interactions with the electron beam. These movements are small (up to 8Å) but enough to smear the information after the reconstruction of the tomogram. At present I have verified only the movements of the gold beads used as fiducials for the alignment of the tilt images. The next step would be to quantify the movement of the biological samples, and to find possible solutions in the grid preparation that interfere with the movement. I have

thought about the use of a continuous layer of gaphene instead the holey carbon, or to use a continuous carbon layer.

5.8.2 Tilt series processing

The signal present in the raw images of the tilt series is too low to precisely qualify the data quality and to allow the correction of the contrast transfer function.

The contras transfer function (CTF) is a function that modulates the signal transfer in the electron microscope. This function has multiple parameters: some depend on the microscope used; others change with the imaging conditions. The only parameter that can vary across images is the defocus at which the image is collected. In order to correct for the CTF one has to know exactly the shape of the function, the best is to extract this information from the data. The images from the tilt series are generally extremely noisy; therefore it is extremely difficult to retrieve this information. I have partially solved the problem changing the data collection scheme: I collect the minimum amount of tilt images to discriminate the repetitive information in the raw tomogram. In this way it is possible to apply up to 4 times the electron dose in each images than in a standard data collection scheme, resulting with a better signal to noise ratio in each image. The next step will be to optimize the fitting procedure, in order to gain precision in the CTF fit.

5.8.3 Improvement of the subtomogram averaging

The subtomogram averaging at present is limited by two issues: the computing speed and the data interpolation.

The computing speed is crucial in order to work with huge datasets, as it is generally done for the single particles reconstruction methods. I worked on the optimization of the existing algorithms and the next step will be to use the Graphic Processor Units (GPU) instead the computer processors. I already started to test the use of the GPU and the expected gain is ~300 fold in the actual processing time.

The data are interpolated twice with the actual procedures. The first interpolation occurs during the tomogram back-projection, and the second occurs after the shifts and the rotations when the subtomograms are averaged. The idea to improve this procedure would be to use the subtomograms to compute the alignments and to generate the final average through the back-projection of the aligned raw tilt images.

At present this procedure has been only theorized, I plan to try it in the next few months.

7 Bibliography

Accola MA, Strack B, Gottlinger HG (2000) Efficient particle production by minimal Gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. *Journal of virology* **74:** 5395-5402

Aldovini A, Young RA (1990) Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *Journal of virology* **64:** 1920-1926

Bartesaghi A, Subramaniam S (2009) Membrane protein structure determination using cryo-electron tomography and 3D image averaging. *Current opinion in structural biology* **19:** 402-407

Bates M, Huang B, Dempsey GT, Zhuang X (2007) Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science* **317:** 1749-1753

Beck M, Lucic V, Forster F, Baumeister W, Medalia O (2007) Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature* **449:** 611-615

Berkowitz RD, Ohagen A, Hoglund S, Goff SP (1995) Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric Gag polyproteins during RNA packaging in vivo. *Journal of virology* **69:** 6445-6456

Bess JW, Jr., Powell PJ, Issaq HJ, Schumack LJ, Grimes MK, Henderson LE, Arthur LO (1992) Tightly bound zinc in human immunodeficiency virus type 1, human T-cell leukemia virus type I, and other retroviruses. *Journal of virology* **66:** 840-847

Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, Davidson MW, Lippincott-Schwartz J, Hess HF (2006) Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313:** 1642-1645

Briggs JA, Johnson MC, Simon MN, Fuller SD, Vogt VM (2006) Cryo-electron microscopy reveals conserved and divergent features of gag packing in immature particles of Rous sarcoma virus and human immunodeficiency virus. *Journal of molecular biology* **355:** 157-168

Briggs JA, Krausslich HG (2011) The molecular architecture of HIV. *Journal of molecular biology* **410:** 491-500

Briggs JA, Riches JD, Glass B, Bartonova V, Zanetti G, Krausslich HG (2009) Structure and assembly of immature HIV. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 11090-11095

Briggs JA, Simon MN, Gross I, Krausslich HG, Fuller SD, Vogt VM, Johnson MC (2004) The stoichiometry of Gag protein in HIV-1. *Nature structural & molecular biology* **11:** 672-675

Briggs JA, Wilk T, Welker R, Krausslich HG, Fuller SD (2003) Structural organization of authentic, mature HIV-1 virions and cores. *The EMBO journal* **22:** 1707-1715

Brody BA, Rhee SS, Hunter E (1994) Postassembly cleavage of a retroviral glycoprotein cytoplasmic domain removes a necessary incorporation signal and activates fusion activity. *Journal of virology* **68:** 4620-4627

Bukrinskaya A, Brichacek B, Mann A, Stevenson M (1998) Establishment of a functional human immunodeficiency virus type 1 (HIV-1) reverse transcription complex involves the cytoskeleton. *The Journal of experimental medicine* **188:** 2113-2125

Butan C, Winkler DC, Heymann JB, Craven RC, Steven AC (2008) RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis. *Journal of molecular biology* **376:** 1168-1181

Cardone G, Purdy JG, Cheng N, Craven RC, Steven AC (2009) Visualization of a missing link in retrovirus capsid assembly. *Nature* **457:** 694-698

Carlson LA, Briggs JA, Glass B, Riches JD, Simon MN, Johnson MC, Muller B, Grunewald K, Krausslich HG (2008) Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis. *Cell host & microbe* **4:** 592-599

Carlson LA, de Marco A, Oberwinkler H, Habermann A, Briggs JA, Krausslich HG, Grunewald K (2010) Cryo electron tomography of native HIV-1 budding sites. *PLoS pathogens* **6:** e1001173

Chance MR, Sagi I, Wirt MD, Frisbie SM, Scheuring E, Chen E, Bess JW, Jr., Henderson LE, Arthur LO, South TL, et al. (1992) Extended x-ray absorption fine structure studies of a retrovirus: equine infectious anemia virus cysteine arrays are coordinated to zinc. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 10041-10045

Coffin JM HS, Varmus HE (1997) In *Retroviruses*, Coffin JM, Hughes SH, Varmus HE (eds). Cold Spring Harbor (NY)

Coren LV, Thomas JA, Chertova E, Sowder RC, 2nd, Gagliardi TD, Gorelick RJ, Ott DE (2007) Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. *Journal of virology* **81:** 10047-10054

De Marco A, Biancotto C, Knezevich A, Maiuri P, Vardabasso C, Marcello A (2008) Intragenic transcriptional cis-activation of the human immunodeficiency virus 1 does not result in allele-specific inhibition of the endogenous gene. *Retrovirology* **5:** 98

de Marco A, Davey NE, Ulbrich P, Phillips JM, Lux V, Riches JD, Fuzik T, Ruml T, Krausslich HG, Vogt VM, Briggs JA (2010a) Conserved and variable features of Gag structure and arrangement in immature retrovirus particles. *Journal of virology* **84:** 11729-11736

de Marco A, Muller B, Glass B, Riches JD, Krausslich HG, Briggs JA (2010b) Structural analysis of HIV-1 maturation using cryo-electron tomography. *PLoS pathogens* **6:** e1001215

Demirov DG, Orenstein JM, Freed EO (2002) The late domain of human immunodeficiency virus type 1 p6 promotes virus release in a cell type-dependent manner. *Journal of virology* **76:** 105-117

Felsenstein KM, Goff SP (1988) Expression of the gag-pol fusion protein of Moloney murine leukemia virus without gag protein does not induce virion formation or proteolytic processing. *Journal of virology* **62:** 2179-2182

Folling J, Bossi M, Bock H, Medda R, Wurm CA, Hein B, Jakobs S, Eggeling C, Hell SW (2008) Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature methods* **5:** 943-945

Forshey BM, von Schwedler U, Sundquist WI, Aiken C (2002) Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *Journal of virology* **76:** 5667-5677

Frey S, Marsh M, Gunther S, Pelchen-Matthews A, Stephens P, Ortlepp S, Stegmann T (1995) Temperature dependence of cell-cell fusion induced by the envelope glycoprotein of human immunodeficiency virus type 1. *Journal of virology* **69:** 1462-1472

Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI (1999) Assembly and analysis of conical models for the HIV-1 core. *Science* **283:** 80-83

Ganser-Pornillos BK, Chandrasekaran V, Pornillos O, Sodroski JG, Sundquist WI, Yeager M (2011) Hexagonal assembly of a restricting TRIM5alpha protein. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 534-539

Ganser-Pornillos BK, Cheng A, Yeager M (2007) Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell* **131:** 70-79

Gilbert JM, Mason D, White JM (1990) Fusion of Rous sarcoma virus with host cells does not require exposure to low pH. *Journal of virology* **64:** 5106-5113

Gorelick RJ, Henderson LE, Hanser JP, Rein A (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. *Proceedings of the National Academy of Sciences of the United States of America* **85:** 8420-8424

Gorelick RJ, Nigida SM, Jr., Bess JW, Jr., Arthur LO, Henderson LE, Rein A (1990) Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *Journal of virology* **64:** 3207-3211

Gottlinger HG, Dorfman T, Sodroski JG, Haseltine WA (1991) Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 3195-3199

Hadravova R, de Marco A, Ulbrich P, Stokrova J, Dolezal M, Pichova I, Ruml T, Briggs JA, Rumlova M (2012) In Vitro Assembly of Virus-Like Particles of a Gammaretrovirus, the Murine Leukemia Virus XMRV. *Journal of virology* **86:** 1297-1306

Han Y, Lassen K, Monie D, Sedaghat AR, Shimoji S, Liu X, Pierson TC, Margolick JB, Siliciano RF, Siliciano JD (2004) Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *Journal of virology* **78:** 6122-6133

Hill CP, Worthylake D, Bancroft DP, Christensen AM, Sundquist WI (1996) Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 3099-3104

Huang M, Orenstein JM, Martin MA, Freed EO (1995) p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *Journal of virology* **69:** 6810-6818

Jacks T, Varmus HE (1985) Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* **230:** 1237-1242

Jin Z, Jin L, Peterson DL, Lawson CL (1999) Model for lentivirus capsid core assembly based on crystal dimers of EIAV p26. *Journal of molecular biology* **286:** 83-93

Johnson MC, Scobie HM, Ma YM, Vogt VM (2002) Nucleic acid-independent retrovirus assembly can be driven by dimerization. *Journal of virology* **76:** 11177-11185

Jouvenet N, Bieniasz PD, Simon SM (2008) Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* **454:** 236-240

Jouvenet N, Simon SM, Bieniasz PD (2009) Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 19114-19119

Katoh I, Kyushiki H, Sakamoto Y, Ikawa Y, Yoshinaka Y (1991) Bovine leukemia virus matrix-associated protein MA(p15): further processing and formation of a specific complex with the dimer of the 5'-terminal genomic RNA fragment. *Journal of virology* **65:** 6845-6855

Khorasanizadeh S, Campos-Olivas R, Clark CA, Summers MF (1999) Sequence-specific 1H, 13C and 15N chemical shift assignment and secondary structure of the HTLV-I capsid protein. *Journal of biomolecular NMR* **14:** 199-200

Kingston RL, Fitzon-Ostendorp T, Eisenmesser EZ, Schatz GW, Vogt VM, Post CB, Rossmann MG (2000) Structure and self-association of the Rous sarcoma virus capsid protein. *Structure* **8:** 617-628

Krausslich HG, Facke M, Heuser AM, Konvalinka J, Zentgraf H (1995) The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *Journal of virology* **69:** 3407-3419

Kukulski W, Schorb M, Welsch S, Picco A, Kaksonen M, Briggs JA (2011) Correlated fluorescence and 3D electron microscopy with high sensitivity and spatial precision. *The Journal of cell biology* **192:** 111-119

Kuznetsov YG, Ulbrich P, Haubova S, Ruml T, McPherson A (2007) Atomic force microscopy investigation of Mason-Pfizer monkey virus and human immunodeficiency virus type 1 reassembled particles. *Virology* **360:** 434-446

Lewinski MK, Bisgrove D, Shinn P, Chen H, Hoffmann C, Hannenhalli S, Verdin E, Berry CC, Ecker JR, Bushman FD (2005) Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *Journal of virology* **79:** 6610-6619

Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455:** 109-113

Lu K, Heng X, Summers MF (2011) Structural determinants and mechanism of HIV-1 genome packaging. *Journal of molecular biology* **410:** 609-633

Luban J (1996) Absconding with the chaperone: essential cyclophilin-Gag interaction in HIV-1 virions. *Cell* **87:** 1157-1159

Masi A, Cicchi R, Carloni A, Pavone FS, Arcangeli A (2010) Optical methods in the study of protein-protein interactions. *Advances in experimental medicine and biology* **674:** 33-42

McDonald D, Vodicka MA, Lucero G, Svitkina TM, Borisy GG, Emerman M, Hope TJ (2002) Visualization of the intracellular behavior of HIV in living cells. *The Journal of cell biology* **159:** 441-452

Meric C, Goff SP (1989) Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. *Journal of virology* **63:** 1558-1568

Meric C, Gouilloud E, Spahr PF (1988) Mutations in Rous sarcoma virus nucleocapsid protein p12 (NC): deletions of Cys-His boxes. *Journal of virology* **62:** 3328-3333

Miller DG, Edwards RH, Miller AD (1994) Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 78-82

Molling K, Bolognesi DP, Bauer H, Busen W, Plassmann HW, Hausen P (1971) Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nature: New biology* **234:** 240-243

Molling K, Bolognesi DP, Bauer H, Busen W, Plassmann HW, Hausen P (1973) Association of the viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Bibliotheca haematologica* **39:** 536-550

Mortuza GB, Haire LF, Stevens A, Smerdon SJ, Stoye JP, Taylor IA (2004) High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature* **431:** 481-485

Muller B, Anders M, Akiyama H, Welsch S, Glass B, Nikovics K, Clavel F, Tervo HM, Keppler OT, Krausslich HG (2009) HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. *The Journal of biological chemistry* **284:** 29692-29703

Nandhagopal N, Simpson AA, Johnson MC, Francisco AB, Schatz GW, Rossmann MG, Vogt VM (2004) Dimeric rous sarcoma virus capsid protein structure relevant to immature Gag assembly. *Journal of molecular biology* **335:** 275-282

O'Hara B, Johann SV, Klinger HP, Blair DG, Rubinson H, Dunn KJ, Sass P, Vitek SM, Robins T (1990) Characterization of a human gene conferring sensitivity to infection by gibbon ape leukemia virus. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **1:** 119-127

Park J, Morrow CD (1991) Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. *Journal of virology* **65:** 5111-5117

Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH (2004) Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. *Journal of virology* **78:** 8477-8485

Pettit SC, Henderson GJ, Schiffer CA, Swanstrom R (2002) Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. *Journal of virology* **76:** 10226-10233

Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI, Hill CP, Yeager M (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137:** 1282-1292

Pornillos O, Ganser-Pornillos BK, Yeager M (2011) Atomic-level modelling of the HIV capsid. *Nature* **469:** 424-427

Puffer BA, Parent LJ, Wills JW, Montelaro RC (1997) Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein. *Journal of virology* **71:** 6541-6546

Rao Z, Belyaev AS, Fry E, Roy P, Jones IM, Stuart DI (1995) Crystal structure of SIV matrix antigen and implications for virus assembly. *Nature* **378:** 743-747

Rein A, Mirro J, Haynes JG, Ernst SM, Nagashima K (1994) Function of the cytoplasmic domain of a retroviral transmembrane protein: p15E-p2E cleavage activates the membrane fusion capability of the murine leukemia virus Env protein. *Journal of virology* **68:** 1773-1781

Rhee SS, Hui HX, Hunter E (1990) Preassembled capsids of type D retroviruses contain a signal sufficient for targeting specifically to the plasma membrane. *Journal of virology* **64:** 3844-3852

Rhee SS, Hunter E (1987) Myristylation is required for intracellular transport but not for assembly of D-type retrovirus capsids. *Journal of virology* **61:** 1045-1053

Rhee SS, Hunter E (1990) A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a type C retrovirus. *Cell* **63:** 77-86

Rust MJ, Bates M, Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods* **3:** 793-795

Sattentau QJ (1988) The role of the CD4 antigen in HIV infection and immune pathogenesis. *AIDS* **2 Suppl 1:** S11-16

Sawyer SL, Emerman M, Malik HS (2007) Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. *PLoS pathogens* **3:** e197

Shoji S, Kubota Y (1989) [Function of protein myristoylation in cellular regulation and viral proliferation]. *Yakugaku zasshi : Journal of the Pharmaceutical Society of Japan* **109:** 71-85

Steeg CM, Vogt VM (1990) RNA-binding properties of the matrix protein (p19gag) of avian sarcoma and leukemia viruses. *Journal of virology* **64:** 847-855

Takeuchi Y, Vile RG, Simpson G, O'Hara B, Collins MK, Weiss RA (1992) Feline leukemia virus subgroup B uses the same cell surface receptor as gibbon ape leukemia virus. *Journal of virology* **66:** 1219-1222

Tanese N, Goff SP (1988) Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proceedings of the National Academy of Sciences of the United States of America* **85:** 1777-1781

Vogt VM (1996) Proteolytic processing and particle maturation. *Current topics in microbiology and immunology* **214:** 95-131

Wiegers K, Rutter G, Kottler H, Tessmer U, Hohenberg H, Krausslich HG (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *Journal of virology* **72:** 2846-2854

Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, Sundquist WI, Jensen GJ (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *The EMBO journal* **26:** 2218-2226

Wyma DJ, Kotov A, Aiken C (2000) Evidence for a stable interaction of gp41 with Pr55(Gag) in immature human immunodeficiency virus type 1 particles. *Journal of virology* **74:** 9381-9387

Yeager M, Wilson-Kubalek EM, Weiner SG, Brown PO, Rein A (1998) Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 7299-7304

Yoshinaka Y, Katoh I, Copeland TD, Oroszlan S (1985) Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *Proceedings of the National Academy of Sciences of the United States of America* **82:** 1618-1622

Zanetti G, Briggs JA, Grunewald K, Sattentau QJ, Fuller SD (2006) Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS pathogens* **2:** e83

Zhou W, Resh MD (1996) Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. *Journal of virology* **70:** 8540-8548

Acknowledgements

As first I thank John Briggs for the supervision, the teaching, the helpful and interesting discussions.

A special acknowledgement to every collaborator, follow in detail and in a more or less random order:

- Hans-Georg Kraeusslich for the collaboration in most of the projects, and for the supervision.

- Baerbel Glass and Anke-Mareil Hauser for the high quality virus preparations.

- Barbara Mueller for the collaboration, and the discussions on the maturation of HIV-1.

- Volker Vogt for the collaboration and the helpful discussions.

- Kay Gruenewald, Lars-Anders Carlson and Sarah Stauffer for the collaboration and for allowing me to process their own tomograms while studying the budding sites and the Leu-Zip.

- Pawel Ulbrich, Tomas Ruml, Misa Rumlova, for the collaborations in the projects related to M-PMV and XMRV.

I acknowledge anyone that gave me helpful comments during all PhD, my colleagues, my TAC members and anyone else I got feedback from.

Thanks to Rodolfo Ciuffa and Sonja Welsch for critical reading of this manuscript.

A big thank to Daniel Castaño-Diez and Sebastian Streichan for teaching me most of the programming I know, and for the helpful discussions.

Appendix 1: Papers

Conserved and Variable Features of Gag Structure and Arrangement in Immature Retrovirus Particles[▽][†]

Alex de Marco,[1] Norman E. Davey,[1] Pavel Ulbrich,[2] Judith M. Phillips,[3] Vanda Lux,[4]
James D. Riches,[1] Tibor Fuzik,[2] Tomas Ruml,[2] Hans-Georg Kräusslich,[4]
Volker M. Vogt,[3] and John A. G. Briggs[1]*

*Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg,
Germany[1]; Department of Biochemistry and Microbiology and Center of Applied Genomics, Institute of
Chemical Technology, Prague, Technicka 3, 166 28, Prague, Czech Republic[2]; Department of
Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853[3]; and
Department of Infectious Diseases, Virology, University Heidelberg,
Im Neuenheimer Feld 324, 69120 Heidelberg, Germany[4]*

The assembly of retroviruses is driven by oligomerization of the Gag polyprotein. We have used cryo-electron tomography together with subtomogram averaging to describe the three-dimensional structure of *in vitro*-assembled Gag particles from human immunodeficiency virus, Mason-Pfizer monkey virus, and Rous sarcoma virus. These represent three different retroviral genera: the lentiviruses, betaretroviruses and alpharetroviruses. Comparison of the three structures reveals the features of the supramolecular organization of Gag that are conserved between genera and therefore reflect general principles of Gag-Gag interactions and the features that are specific to certain genera. All three Gag proteins assemble to form approximately spherical hexameric lattices with irregular defects. In all three genera, the N-terminal domain of CA is arranged in hexameric rings around large holes. Where the rings meet, 2-fold densities, assigned to the C-terminal domain of CA, extend between adjacent rings, and link together at the 6-fold symmetry axis with a density, which extends toward the center of the particle into the nucleic acid layer. Although this general arrangement is conserved, differences can be seen throughout the CA and spacer peptide regions. These differences can be related to sequence differences among the genera. We conclude that the arrangement of the structural domains of CA is well conserved across genera, whereas the relationship between CA, the spacer peptide region, and the nucleic acid is more specific to each genus.

Retrovirus assembly is driven by the oligomerization of Gag, a multidomain protein, including an N-terminal membrane binding domain (MA), a two-domain structural component (CA), and an RNA binding domain (NC). The Gag proteins of all orthoretroviruses, including the alpha-, beta-, and lentiretroviruses discussed here, share this conserved modular architecture (Fig. 1). Despite very weak sequence conservation, the tertiary structures of MA, CA, and NC are conserved among retroviruses. Outside these conserved domains the Gag proteins of different retroviruses exhibit substantial variability. Other domains may be present or absent, and the length and sequence of linker peptides may also vary (12) (Fig. 1).

Oligomerization of Gag in an infected cell leads to the formation of roughly spherical immature virus particles, where Gag is arranged in a radial fashion with the N-terminal MA domain associated with a surrounding lipid bilayer, and the more C-terminal NC pointing toward the center of the particle (15, 44, 46). Subsequent multiple cleavages of Gag by the viral protease lead to a rearrangement of the virus. NC and the RNA condense in the center of the particle, CA assembles into a capsid or shell around the nucleoprotein, and MA remains associated with the viral membrane. This proteolytic maturation is required to generate an infectious virion (2). In contrast to the mature CA lattice, which has been extensively studied (11, 16, 36), the Gag lattice in immature particles is incompletely understood.

Gag itself contains all of the necessary determinants for particle assembly. For example, the expression of Gag alone in an insect cell expression system is sufficient to generate virus-like particles (3, 17, 22, 38). Retroviral Gag proteins also can be assembled *in vitro* in the presence of nucleic acids to form spherical particles (9, 19, 39, 43, 47). The arrangement of Gag within these *in vitro*-assembled Gag particles is indistinguishable from that found in immature virus particles (6), and the *in vitro* assembly systems have proved valuable for unraveling the principles of virus assembly (18, 28, 29, 39). Multiple layers of interaction promote the assembly of Gag *in vivo*, including MA-membrane-MA interactions, CA-CA interactions, and NC-RNA-NC interactions. An extensive body of literature has explored which regions of Gag are required for assembly and which can be replaced or deleted without compromising assembly. MA-membrane-MA interactions contribute but are not essential. NC-RNA-NC interactions appear to function to nonspecifically link Gag molecules together and can be replaced both *in vivo* and *in vitro* by other interaction domains such as leucine zippers (4, 13, 20, 32, 48). The C-terminal
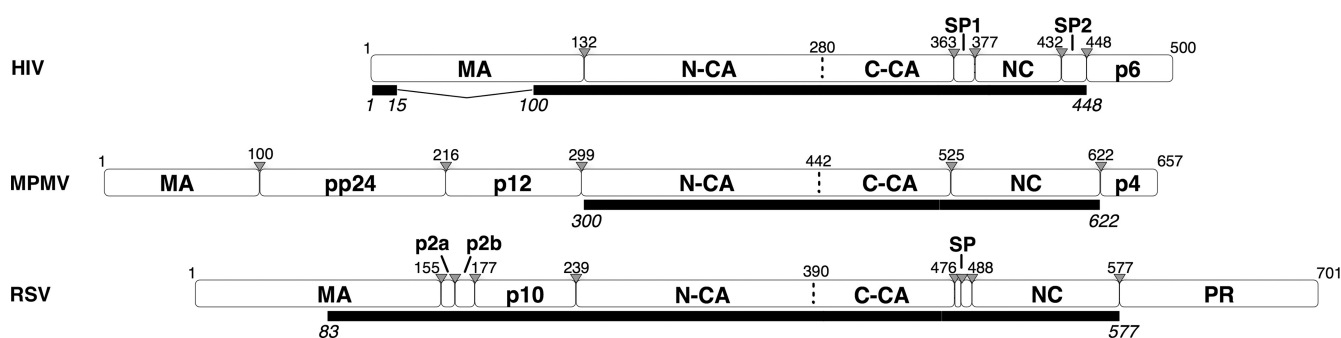
FIG. 1. Modular architecture of the full-length Gag proteins of HIV, M-PMV, and RSV. White rectangles illustrate Gag polyprotein cleavage products. The extent of the constructs used in the electron microscopic analysis is specified under each protein as a black rectangle. Gray triangles specify cleavage sites. Residue numbers are counted from the beginning of Gag.

domain of CA (referred to here as C-CA) and the stretch of amino acids immediately following this domain (termed the spacer peptide [SP] region) are critical for assembly and sensitive to mutation (1, 22, 27, 30).

We set out to understand how the substantial sequence variation among Gag proteins in different retroviruses is manifested in structural differences in the immature Gag lattice. To do this, we studied three retroviruses from different genera: the lentivirus human immunodeficiency virus type 1 (HIV-1), the betaretrovirus Mason-Pfizer monkey virus (M-PMV), and the alpharetrovirus Rous sarcoma virus (RSV). These retroviruses are those for which *in vitro* assembly was first established and has been most extensively studied (6, 19, 24, 28, 29, 35, 43, 47).

The domain structures of the three retroviruses differ most substantially upstream of CA. Both M-PMV and RSV have domains located between MA and CA that are absent in HIV (Fig. 1). In M-PMV there are 198 residues forming the pp24 and p12 domains; in RSV there are 84 residues forming the p2a, p2b, and p10 domains. The three retroviruses have different requirements for regions upstream of CA during assembly. The C-terminal 25 residues of p10 are essential for proper immature RSV assembly, both *in vitro* and *in vivo*, and these residues are inferred to interact directly with N-CA to stabilize the hexamer by forming contacts between adjacent N-CA domains (35). An equivalent assembly domain has not been described for other retroviruses. Within M-PMV p12 is the so-called internal scaffolding domain that is not essential for assembly *in vitro* (43) but is required for particle assembly when the precursor is expressed under the control of the M-PMV promoter (41). It is a key domain for the membrane-independent assembly of immature capsids (40).

In HIV, five residues upstream of CA must be present for assembly of immature virus-like spherical particles *in vitro*, although larger upstream extensions, including part of MA, are required for efficient assembly of regular particles, both for HIV and RSV. For HIV, if the entire MA domain is included, *in vitro* assembly requires the presence of inositol penta- or hexakis phosphate (8). If no sequences upstream of CA are present, the *in vitro* particles in both of these viruses adopt a mature-type tubular morphology (10, 18). It has been hypothesized that cleavage at the N terminus of N-CA during maturation leads to the N-terminal residues of CA folding back into the N-CA structure to form a β-hairpin. The β-hairpin is im-

portant for assembly of the mature CA lattice, whereas its absence is important for immature assembly (23, 42). These requirements explain why, in HIV and RSV, immature Gag lattice-like structures are formed only if regions upstream of CA are present (18). In M-PMV, an immature Gag lattice can be produced when the regions upstream of CA are deleted if this is combined with mutations (such as deleting the initial proline of CA), which prevent β-hairpin formation (43).

During maturation, HIV and RSV Gag proteins are cleaved twice between CA and NC to release a small peptide called SP1 or SP. In RSV the most N-terminal of these two cleavages can occur at one of two possible positions such that the released peptide is either 9 or 12 amino acids long (33). In M-PMV only one cleavage occurs between CA and NC, and no short peptide is produced. The region between the final helix of CA and the Zn fingers has been proposed to adopt a helical bundle architecture in HIV and RSV based on bioinformatic prediction, on mutational analysis, and on structural studies (1, 22, 27, 45). In all three viruses, C-CA and the residues immediately downstream are critical for assembly and are sensitive to mutation. C-CA contains the major homology region, a group of residues that are highly conserved across the retroviruses.

Cryo-electron tomography (cET) studies of immature virus particles (6, 45) have resolved the electron density of the HIV Gag lattice in three dimensions at low resolution. Using these methods, we have also described the three-dimensional architecture of *in vitro*-assembled HIV Gag particles (6). In immature viruses and *in vitro*-assembled particles, Gag is seen to adopt an 8 nm hexameric lattice, as was predicted from previous Fourier analysis of two-dimensional images (7, 46). The hexameric lattice is interrupted by irregularly shaped holes and cracks in the lattice (6, 45). A similar observation has been made using AFM of *in vitro*-assembled particles of M-PMV Gag (26). These holes and cracks allow an otherwise planar hexameric lattice to form the surface of an approximately spherical particle.

The radial positions of the MA, CA, and NC domains had been assigned previously from cryo-electron micrographs (44, 46). Based on these assignments and the shape of the density, the position and relative orientations of CA domains can be modeled into the low-resolution structure of the HIV lattice (6, 45). Density ascribed to the N-terminal domain of CA (N-CA) forms rings around large holes at the 6-fold symmetry

TABLE 1. Subtomogram data sets used in this study

Virus	Defocus (μm)	No. of subtomograms		Resolution (Å)
		Total	In final reconstruction	
HIV	−4.5	12,220	9,698	28
	−2.1	10,231	5,438	25.5
M-PMV	−3.9	18,821	13,209	28
	−2.2	16,453	10,858	26.5
RSV	−3.8	9,802	4,853	28
	−2	11,001	6,732	23

positions in the lattice. Below this layer, at the expected radius of the C-CA, are 2-fold densities, interpreted as corresponding to dimers of C-CA. These densities are linked by rodlike densities, which descend into the NC-nucleic acid layer.

HIV is the only retrovirus for which the arrangement of Gag in the immature particle has been described in three dimensions. Prior to this work, important open questions were therefore: which features of the arrangement of Gag are conserved between genera and therefore reflect general principles of Gag-Gag interactions, and which features are specific to certain genera? We have applied subtomogram averaging of cryo-electron tomograms to generate reconstructions of *in vitro*-assembled Gag particles from HIV, M-PMV, and RSV. These allow identification of the general and variable features of the arrangement of Gag and the architecture of immature retroviruses.

MATERIALS AND METHODS

Sample preparation. Gag proteins were expressed in *Escherichia coli* and purified. Immature virus-like Gag particles were prepared by *in vitro* assembly of proteins with nucleic acid as described previously for HIV (19), M-PMV (26), or RSV (35). The constructs used are illustrated in Fig. 1. The nucleic acid was a 50-nucleotide (nt) single-stranded DNA (RSV), a 73-nt single-stranded DNA (HIV), or MS2 phage RNA (M-MPV). The assembled particles were kept at 4°C for at most 6 days until flash frozen in liquid ethane for analysis.

Electron microscopy and image processing. *In vitro*-assembled particles of HIV, M-PMV, and RSV Gag were mixed with 10-nm gold beads, deposited on C-flat Holey carbon grids, and vitrified by plunge freezing in liquid ethane. Tilt series were collected on an FEI Tecnai F30 Polara transmission electron microscope with Gatan GIF 2002 post column energy filter and a 2kx2k Multiscan charge-coupled device camera. The data collection was performed at 300 kV using the University of California at San Francisco (UCSF) software package and tomograms reconstructed using the IMOD software package (25) as described previously (6). Tilt series were typically collected between −60° and +60° with an angular increment of 3° and a total electron dose of ~70 e/Å². Subtomogram averaging was carried out by using a six-dimensional search (14) in a reference-free manner as described previously (6). Independent reconstructions of half datasets were used to verify the reliability of the structures (see Results) and combined to produce the final reconstruction. Only subtomograms giving a cross correlation with the reference of greater than the mean cross correlation were included in the final reconstruction. The final reconstruction was filtered to the resolution to which the two half datasets are comparable, according to the Fourier shell correlation (see Fig. S3 in the supplemental material). The datasets are summarized in Table 1.

Lattice map representations were generated by using Amira (Visage Imaging), together with the EM Package (37). A hexagon is placed at the final aligned position of each tomogram and colored according to the cross-correlation value between the subtomogram and the average. Tomograms aligned to an inappropriate radial position were excluded based on the radii distribution of the set of 20 nearest subtomograms surrounding each subtomogram. If the radius of the selected subtomogram was in the first or in the fourth quartile of the distribution,

the subtomogram was excluded. Only hexagons with a cross-correlation over a defined threshold are displayed. Structure interpretation and fitting were carried out using UCSF Chimera (34), which was also used to produce structure figures.

The angles plotted in Fig. 4B for each hexamer are 180° minus the mean angle between the 6-fold axis of a hexamer and the 6-fold axes of its immediate neighbors.

Bioinformatics. To establish conserved retroviral features present or absent in our proteins of interest, 12 Gag protein sequences of the 27 International Committee on Taxonomy of Viruses (ICTV) reference species from these genera were chosen as follows. From the 27 ICTV reference species, those with incomplete Gag sequences or from defective viruses were not considered. To avoid overweighting the importance of very similar viruses, only one representative sequence was chosen from groups of viruses with sequence similarity in Gag greater than 70%, giving 12 sequences. These sequences were aligned by using MAAFT (21). Conserved features are those identified in >75% of the sequences.

The SP1 region is hypothesized to form an alpha-helix based on bioinformatic prediction, mutational analysis, and structural studies (1, 22, 27). A helical wheel representation of the 18 residues after the CA adjacent cleavage site was created to examine the helical hydrophobic moment of the region for the three viruses of interest. The complete region between the final helix of CA and the first zinc finger of NC was modeled in three dimensions as an alpha-helix using chimera (see Fig. S4 in the supplemental material).

RESULTS AND DISCUSSION

We describe here the local and global architecture of the Gag lattice in *in vitro*-assembled, immature Gag particles for three retroviral genera, with the aim of identifying general features of Gag assembly and those features which vary among genera. *In vitro*-assembled Gag particles were produced using expression vectors and assembly conditions that have previously been described (18, 35, 43). Each Gag protein represents approximately the minimal size protein for which *in vitro* assembly is efficient and correct and can occur without other additions (see the introduction). As a consequence, domains upstream of the conserved CA domains were variably truncated or absent (constructs are illustrated in Fig. 1).

The structure of the Gag lattice in *in vitro*-assembled immature particles. Vitrified, *in vitro*-assembled Gag particles were imaged by using cryo-electron tomography. Initial data collection was carried out at a defocus of approximately −4 μm. In order to assess the reliability of the reconstructions, for each type of *in vitro*-assembled Gag particle, the data set was split into two halves and two completely independent reference-free subtomogram averaging reconstructions were carried out. A second round of data collection was carried out with a defocus of approximately −2 μm. Again, the data set for each type of particle was split into two halves, and each half was reconstructed using a different reconstruction from the −4 μm data set as a starting model. Sections through these independent reconstructions are shown in Fig. S1 in the supplemental material. No significant differences were observed between independent reconstructions of particles assembled from the same Gag protein. The independent reconstructions were compared to one another by using Fourier Shell Correlation and were comparable to a resolution of 26.5 Å (see Fig. S2 in the supplemental material). This comparison of two independent reconstructions verified the robustness of the reconstruction process and allowed us to assess the resolution to which the reconstructions could be reliably interpreted. The two independent reconstructions were therefore aligned, averaged, and filtered to 26.5 Å to generate final reconstructions (Fig. 2A). The same process was carried out with *in vitro*-assembled Gag
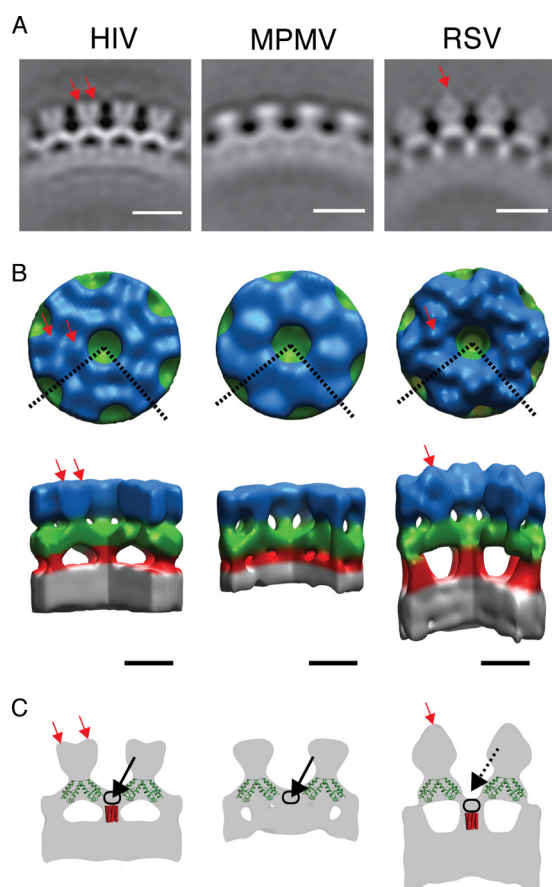
FIG. 2. Structure of the *in vitro*-assembled Gag particles. (A) Radial section through the average structures. The density is white. The scale bar is 10 nm. (B) Surface rendering of the average structures. The color scheme follows an approximate radial division of the structure according to the different domains of Gag as previously assigned (6, 44, 45, 46): gray is the NC-nucleic acid layer, red is the SP1 region, green is the C-CA region, and blue is the N-CA. The dashed lines on the top view represent the section cut out in the side view. Red arrows indicate densities protruding from the N-CA density (see the text). Scale bar, 5 nm. (C) Sections through the isosurface representations (gray), illustrating features referred to in the text. The crystal structure of the CA dimer (PDB: 3ds2) (green) is superimposed on the sections indicated previously (6). Red rods indicate possible positions of a hypothetical 24-Å six-helix bundle. Red arrows indicate densities protruding from the N-CA density. Black arrows indicate density forming the floor of the hole in the C-CA ring. Black rings indicate possible positions of the residues between the final helix in the C-CA and the six-helix bundle. For explanations, see the text.

particles from each of the three different retroviruses. Isosurface representations of the reconstructions were generated (Fig. 2B) and colored for illustrative purposes.

The three different reconstructions exhibit substantial similarity. An outer density (blue in Fig. 2B) is arranged in hexameric rings around large holes. Where these rings meet, 2-fold densities (green in Fig. 2B) extend between adjacent 6-fold rings. Six of these 2-fold densities converge at each 6-fold axis, beneath the holes in the hexameric rings. Here they connect with a density (red in Fig. 2B), which descends further toward the center of the particle. This density meets the NC-nucleic acid layer (gray in Fig. 2B), which is not ordered in a

hexameric manner. Despite these similarities, structural differences are seen at all radii of the Gag layer.

Outermost in the particles (blue in Fig. 2B) are the N termini of the Gag-derived proteins, where the three constructs used differ substantially (Fig. 1). The M-PMV construct begins directly with the N-CA domain, lacking the N-terminal proline residue. Accordingly, the M-PMV reconstruction shows the least featured outer density (blue in Fig. 2B), which can be presumed to represent N-CA.

The HIV construct includes the first 15 and last 33 residues of MA. This protein has been found to efficiently assemble *in vitro* (18). The reconstruction shows small protrusions around the holes at the 6-fold axis (red arrows in Fig. 2B and C). The extra MA residues may form these densities. Alternatively, the presence of part of MA may lead to a more rigid N-CA lattice than present in M-PMV, allowing individual N-CA domains to be more clearly resolved. In this case, the small protrusions could also represent part of N-CA.

The RSV construct includes p10, p2, and about one-half of MA, which together total 156 amino acid residues upstream of the CA domain. The reconstruction shows large extra densities above the 2-fold axes of N-CA (red arrow in Fig. 2A to C, visible as prominent caplike densities in Fig. 2A). The densities are positioned above the 2-fold axes of the lattice, where they link between adjacent hexamers. These densities could include the N-terminal part of p10 and perhaps sequences further upstream in the p2 and/or MA domains. In this model, each of the caplike densities would be formed from two p10 domains, one from each hexamer, which by linking adjacent hexamers would stabilize the hexameric lattice.

Moving toward the center of the reconstructions, the next layer can be assigned to the C-CA domain as described previously (6, 44–46) (green in Fig. 2B). If the crystal structure of the C-CA dimer (PDB: 3ds2) is superimposed onto the HIV density (as in reference 6 or Fig. 2C), the density occupied by the crystal structure can be assigned to the ordered part of the C-CA domain. The shape of this density is similar in all three reconstructions and, considering the high level of conservation of retroviral CA domains, the same crystal structure has also been superimposed on the RSV and M-PMV reconstructions for illustrative purposes. The density around the 6-fold axis that is not occupied by the crystal structure can be assigned to residues downstream of the final CA helix. In HIV and M-PMV the top of this region forms a density below the hexagonal hole created by the N-CA domains, at the center of the C-CA ring, forming the floor of the hole in the ring (black arrows in Fig. 2C). The position of this floor makes it most likely that it is formed from the residues immediately downstream of the final helix of C-CA. In RSV the floor is farther toward the center of the particle, making the hexagonal hole deeper (compare the positions of the solid and dashed black arrows for HIV and RSV in Fig. 2C).

The lower floor in the center of the C-CA hexamer hole in the RSV reconstruction suggests a difference in the arrangement of the residues immediately downstream of C-CA and in their interactions with the structured domain of C-CA. This suggests that the arrangements of C-CA, and the dimensions of the hexamer, which are well conserved across the retroviruses, do not require a specific conserved arrangement of the amino acid sequences that comprise the floor or a specific conserved
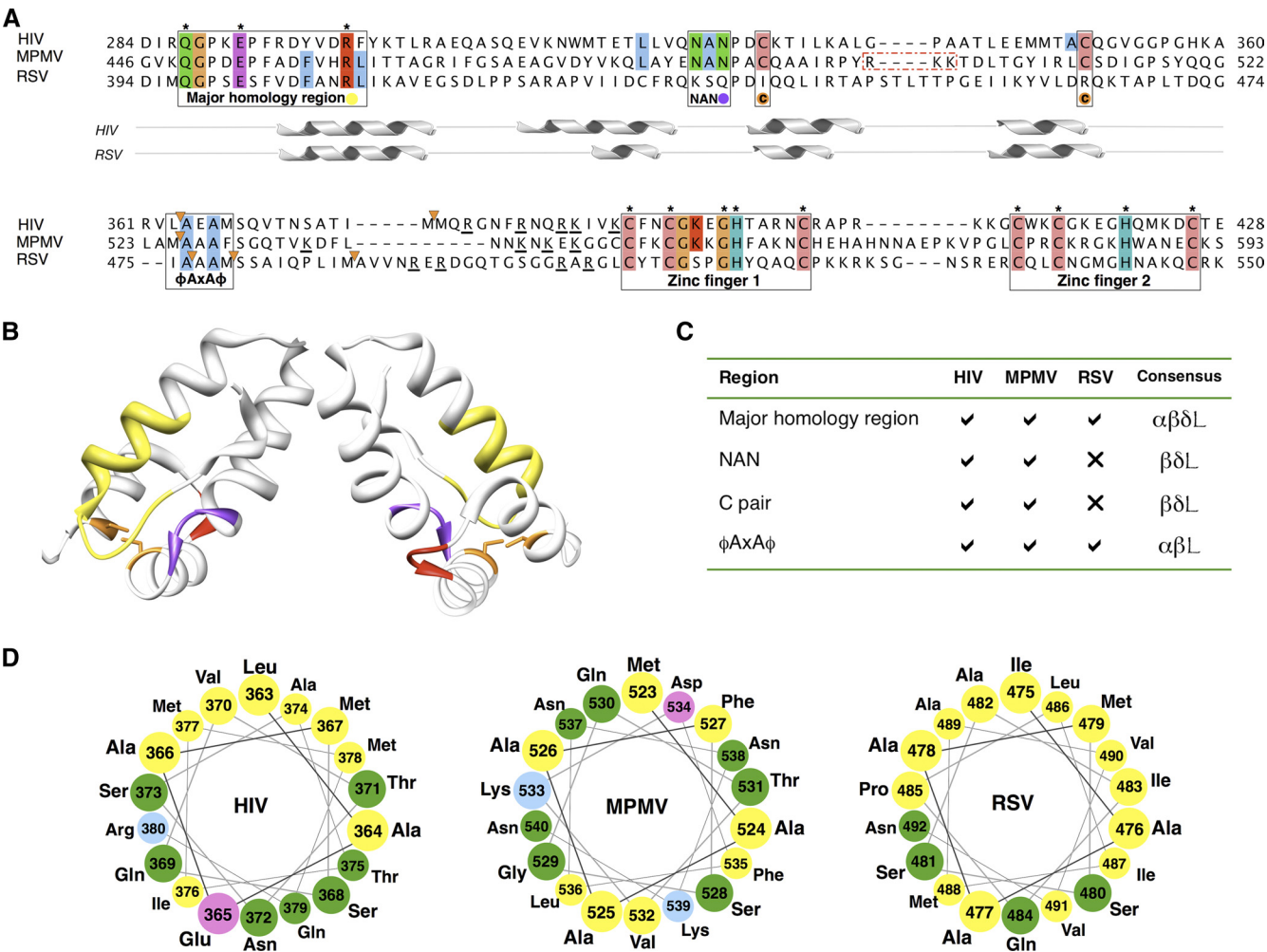
FIG. 3. Comparisons of the Gag sequences from HIV, M-PMV, and RSV. (A) Multiple alignment of Gag C-CA-NC for RSV, M-PMV, and HIV. Regions of interest are boxed, and colored circles indicate the coloring of these regions in panels B and C. Residues conserved across >75% of viruses in an extended alignment, including 12 viral species (see Fig. S3 in the supplemental material), are colored using ClustalX coloring; residues conserved across all species are specified by an asterisk. Orange triangles denote cleavage sites. Basic residues adjacent to the NC are underlined. A dashed red box surrounds an RKK motif unique to M-PMV C-CA. The position of helices from the solved structures for RSV (PDB: 1d1d) and HIV (PDB: 3ds2) are shown under the alignment for illustrative purposes. Residues are numbered from the beginning of Gag. (B) HIV C-CA (PDB: 3ds2) with residues colored to indicate the positions of regions of interest: yellow, MHR; purple, NAN; orange, paired cysteine residues; and red, RKK motif (see the text and panels A and C). (C) Table of regions of interest defined in panel A; the colors correspond to the coloring in the multiple alignment (A) and HIV structure (A). Consensus signifies genera in which the feature is found (α signifies alpha, β signifies beta, δ signifies delta, and L signifies lentiretroviruses). There is currently no solved structure for the CA region in M-PMV, the final helix of M-PMV is defined assuming structural homology to HIV (PDB: 3ds2). (D) Helical wheel representation of the 18 residues after the CA adjacent cleavage site for the proteins of interest.

interaction with the downstream structure present on the 6-fold axis. Instead, the conserved arrangement of C-CA is presumably predominantly mediated by interactions between CA dimers around the 6-fold axis.

The region between C-CA and the NC-nucleic acid layer is strikingly different in the three viruses. In RSV a strong rodlike feature descends from below C-CA to the NC-nucleic acid layer along the 6-fold axis (red in Fig. 2B). In HIV a similar feature is seen, but as described above the density begins within the C-CA layer, where it forms the floor of the hole at the 6-fold axis in the C-CA lattice. In M-PMV the floor is present, but there is no clear rodlike density between C-CA and the NC-nucleic acid layer.

Comparing structural differences with sequence differences. A sequence alignment of the C-CA-NC region highlights features conserved across multiple orthoretroviruses (see Materials and Methods and see Fig. S3 in the supplemental material). The alignment for the three viruses of interest is shown in Fig. 3A. This alignment identifies a number of previously described conserved features, the major homology region (MHR), nucleotide-binding Zn finger motifs, a CA-adjacent cleavage site, and a conserved cysteine pair. Most strikingly conserved is the MHR, which is known to be present in all orthoretroviruses. A pair of well-conserved cysteine residues, which approach close to one another between helices 3 and 4 of C-CA in the HIV CA crystal structure, is present in betaretroviruses, lentivi-

ruses, and deltaretroviruses but absent in the alpharetroviruses. Not previously described is a conserved NAN motif in the turn between helices 2 and 3, which is also absent in the alpharetroviruses. These regions are highlighted on the structure of the C-CA dimer from HIV-1 in Fig. 3B, and summarized in Fig. 3C.

The conserved cysteine pair and NAN turn motif within C-CA, which are absent in RSV but present in HIV and M-PMV, are located away from the dimer axis, in a position that would lie close to the hole at the 6-fold axis (Fig. 3B). This is the region where the position of the floor differs between the retroviruses. We speculate that the cysteine pair and NAN sequence motifs may contribute to the organization of this region in HIV and M-PMV.

A stretch of amino acids roughly corresponding to the HIV-1 SP1 has been suggested to have helical propensity based upon computational prediction, mutational analysis, and nuclear magnetic resonance (NMR) (1, 31). It has been proposed that this region assembles into a six-helix bundle in the immature HIV-1 Gag lattice (45). We examined this proposal by considering the helical hydrophobic moment of this region (Fig. 3D and see Fig. S4 in the supplemental material). Both HIV and RSV show a 16-residue stretch, starting from the residue immediately upstream of the cleavage site and running into SP1 or SP, with a strong helical hydrophobic moment, a finding consistent with the assembly of this region into a six-helix bundle held together by hydrophobic side chains in its center. M-PMV does not show a strong helical hydrophobic moment. The calculated hydrophobic face in RSV is slightly larger than that in HIV. In both HIV and RSV, the 16-residue alpha-helical bundle would be a maximum of ~24 Å long depending on the angle between the helices and the axis of the bundle. A 24-Å bundle is superimposed onto the structures in Fig. 2C for illustrative purposes.

The resolution of the reconstructions is not sufficient to resolve the secondary structural details, which would be necessary to describe the structure of the CA-NC linker region unambiguously, and multiple models are therefore consistent with the observed structures. In one such model, the region between the final helix of C-CA and the beginning of the proposed helical bundle is arranged to form the floor density in HIV (outlined by black ring in Fig. 2C), while the rodlike structure is formed by the proposed six-helix bundle. Below this region, basic residues close to the Zn fingers are involved in binding nucleic acid. In RSV, the region between the final helix of C-CA and the start of the helical bundle is not held in the floor position. It instead forms the top part of the rodlike structure (outlined by black ring in Fig. 2C). In the model, the lower part of the rod would then be formed by the six-helix bundle. In M-PMV the residues after the final helix of CA also form a floorlike structure as for HIV. The rodlike structure is not visible in M-PMV, and the nucleic acid directly abuts C-CA. This observation can be interpreted in two ways. Either the helical bundle is absent or the helical bundle is present despite the absence of a strong helical hydrophobic moment, but the rodlike density is obscured because it is buried in the nucleic acid layer. Presently available data do not allow these two possibilities to be distinguished.

Although we attribute the differences in tomographic density features in the NC region to differences in amino acid sequences among the three viruses studied, we cannot entirely exclude a role for the type of nucleic acid used to promote assembly. Whereas M-PMV particles were formed with RNA, the HIV and RSV particles were formed with DNA oligonucleotides. These respective nucleic acids were found previously to promote assembly in the most robust manner. We think it is unlikely that the HIV and RSV lattices are affected by the use of DNA since reconstructions of immature HIV particles (6, 45), which contain RNA, show the same arrangement of the NC layer.

The observation that the density interpreted to be the NC-nucleic acid layer directly abuts the base of the CA dimer in M-PMV suggests a possible direct interaction between the two. If the sequence of the C-CA domain of M-PMV is mapped onto the crystal structure of the HIV dimer, a basic RKK motif present in M-PMV but absent in RSV and HIV (see the dashed red boxes in Fig. 3A and the red area in Fig. 3B) is exposed on the underside of each monomer. These basic residues provide a potentially potent binding site for nucleic acid, and may explain the proximity of the nucleic acid layer to C-CA in the M-PMV structure. We hypothesize that such an interaction may be sufficient to pull the nucleic acid layer toward the CA layer, even if the Zn fingers are held at a distance from the CA layer by a rodlike structure, leading to the rod becoming buried within the nucleic acid layer.

Global arrangement of Gag in the *in vitro*-assembled Gag particles. The global arrangement of the Gag shell within an individual *in vitro*-assembled particle can be visualized by placing a hexamer at the position and orientation to which each subtomogram has converged during the alignment and reconstruction procedure (6). Two example particles for each genus are illustrated in Fig. 4A. The relative positions and orientations of the placed hexamers are not random but instead are consistent with a hexameric lattice. All three types of particles are assembled from a single assembled hexameric lattice with irregularly shaped defects, indicating that this pattern is a conserved outcome of retroviral Gag protein assembly.

The mean curvature of the particles is not the same in all cases. Similar to previous observations using conventional cryo-EM (5, 47) *in vitro*-assembled HIV Gag particles had a diameter of (94 ± 4 nm, $n = 12$), measured to the C-CA domain, which is similar to the diameter of M-PMV particles (90 ± 7 nm, $n = 12$) and significantly larger than RSV particles (60 ± 3 nm, $n = 12$). The differing diameters are reflected in the different curvatures of the lattice visible in Fig. 4A. This change in size reflects changes in the angle between adjacent hexamers, which in HIV and M-PMV are approximately 170° and 169°, compared to 163° in RSV (Fig. 4B). It is tempting to speculate that the tighter curvature in RSV may result from a difference in the relative orientations of the C-CA domains around the 6-fold axis, facilitated by the structural differences observed in the CA-SP1 region. This speculation is supported by the observation that, upon *in vitro* assembly of chimeric HIV-RSV Gag-derived proteins, the main determinant for particle size was CA (3). RSV CA was found to induce the formation of smaller, more tightly curved particles than HIV-CA in these chimeric proteins.

Although less dramatic than the differences between the genera, there is also variability in the angle between hexamers within each genus, when different particles are compared and
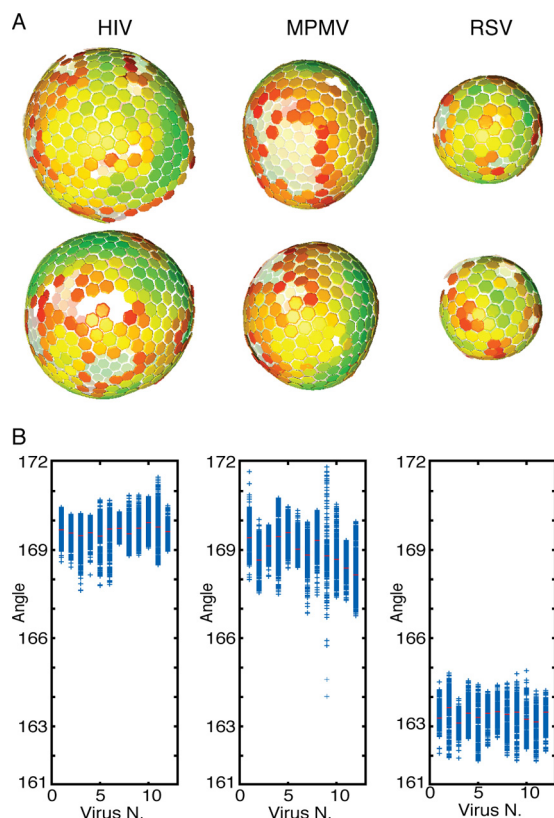
FIG. 4. Global arrangement of Gag lattice in the particles. (A) Global lattice maps of *in vitro*-assembled Gag particles. The centers of each hexameric unit cell are marked with hexamers, which are colored according to cross correlation on a scale from low (red) to high (green). Higher cross-correlation values indicate that the subtomogram is more similar to the average structure. The cross-correlation range in the each map has been set between the minimum and the maximum cross-correlation value present in the map. Maps are shown in perspective, such that hexamers on the rear surface of the particle appear smaller and fainter. Two representative particles are shown for each genus. (B) Scatter plot showing the distribution of the angles between each hexamer and its neighbors (see Materials and Methods). The mean value for each particle is marked in red. For each construct, 12 particles have been analyzed.

also within individual particles. This variability is manifested in the formation of particles with slightly varying diameters and with deviations from a perfect spherical shape. Flexibility in the angle between the hexamers is a necessary basis for the formation of heterogeneous particles, a known feature of all retroviruses.

Summary. The comparison of the immature Gag lattice from three retroviral genera highlights the general features of particle architecture. All three of the particles assemble from a single hexameric lattice with irregular defects. The arrangement of the two CA domains is strikingly conserved between different genera, whereas the angles between adjacent hexamers appear to vary, changing the degree of curvature and thus the particle size. Outside of the CA domains, sequence differences among the three genera are manifested as structural differences. Based on mutational data and chimeric Gag proteins containing foreign domains, the C-CA-SP region seems to contain the critical assembly determinants for the immature

Gag shell. Nevertheless, the inferred structure of this region, and its relationship to CA and the NC-nucleic acid layer, appear to be specific to each genus.

REFERENCES

1. **Accola, M. A., S. Hoglund, and H. G. Gottlinger.** 1998. A putative alpha-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:**2072–2078.
2. **Adamson, C. S., and E. O. Freed.** 2007. Human immunodeficiency virus type 1 assembly, release, and maturation. Adv. Pharmacol. **55:**347–387.
3. **Ako-Adjei, D., M. C. Johnson, and V. M. Vogt.** 2005. The retroviral capsid domain dictates virion size, morphology, and coassembly of gag into virus-like particles. J. Virol. **79:**13463–13472.
4. **Borsetti, A., A. 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:**9313–9317.
5. **Briggs, J. A., M. C. Johnson, M. N. Simon, S. D. Fuller, and V. M. Vogt.** 2006. Cryo-electron microscopy reveals conserved and divergent features of gag packing in immature particles of Rous sarcoma virus and human immunodeficiency virus. J. Mol. Biol. **355:**157–168.
6. **Briggs, J. A., J. D. Riches, B. Glass, V. Bartonova, G. Zanetti, and H. G. Krausslich.** 2009. Structure and assembly of immature HIV. Proc. Natl. Acad. Sci. U. S. A. **106:**11090–11095.
7. **Briggs, J. A., M. N. Simon, I. Gross, H. G. Krausslich, S. D. Fuller, V. M. Vogt, and M. C. Johnson.** 2004. The stoichiometry of Gag protein in HIV-1. Nat. Struct. Mol. Biol. **11:**672–675.
8. **Campbell, S., R. J. Fisher, E. M. Towler, S. Fox, H. J. Issaq, T. Wolfe, L. R. Phillips, and A. Rein.** 2001. Modulation of HIV-like particle assembly in vitro by inositol phosphates. Proc. Natl. Acad. Sci. U. S. A. **98:**10875–10879.
9. **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:**4425–4435.
10. **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:**6487–6497.
11. **Cardone, G., J. G. Purdy, N. Cheng, R. C. Craven, and A. C. Steven.** 2009. Visualization of a missing link in retrovirus capsid assembly. Nature **457:**694–698.
12. **Coffin, J. M., S. H. Huges, and H. E. Varmus.** 1997. Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. **Crist, R. M., S. A. Datta, A. G. Stephen, F. Soheilian, J. Mirro, R. J. Fisher, K. Nagashima, and A. Rein.** 2009. Assembly properties of human immunodeficiency virus type 1 Gag-leucine zipper chimeras: implications for retrovirus assembly. J. Virol. **83:**2216–2225.
14. **Forster, F., O. Medalia, N. Zauberman, W. Baumeister, and D. Fass.** 2005. Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. Proc. Natl. Acad. Sci. U. S. A. **102:**4729–4734.
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:**729–738.
16. **Ganser-Pornillos, B. K., A. Cheng, and M. Yeager.** 2007. Structure of full-length HIV-1 CA: a model for the mature capsid lattice. Cell **131:**70–79.
17. **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 Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. Cell **59:**103–112.
18. **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:**4798–4810.

19. **Gross, I., H. Hohenberg, and H. G. Krausslich.** 1997. In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. Eur. J. Biochem. **249:**592–600.

20. **Johnson, M. C., H. M. Scobie, Y. M. Ma, and V. M. Vogt.** 2002. Nucleic acid-independent retrovirus assembly can be driven by dimerization. J. Virol. **76:**11177–11185.

21. **Katoh, K., K. Misawa, K. Kuma, and T. Miyata.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. **30:**3059–3066.

22. **Keller, P. W., M. C. Johnson, and V. M. Vogt.** 2008. Mutations in the spacer peptide and adjoining sequences in Rous sarcoma virus Gag lead to tubular budding. J. Virol. **82:**6788–6797.

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

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

25. **Kremer, J. R., D. N. Mastronarde, and J. R. McIntosh.** 1996. Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. **116:**71–76.

26. **Kuznetsov, Y. G., P. Ulbrich, S. Haubova, T. Ruml, and A. McPherson.** 2007. Atomic force microscopy investigation of Mason-Pfizer monkey virus and human immunodeficiency virus type 1 reassembled particles. Virology **360:**434–446.

27. **Liang, C., J. Hu, R. S. Russell, A. Roldan, L. Kleiman, and M. A. Wainberg.** 2002. Characterization of a putative alpha-helix across the capsid-SP1 boundary that is critical for the multimerization of human immunodeficiency virus type 1 gag. J. Virol. **76:**11729–11737.

28. **Ma, Y. M., and V. M. Vogt.** 2004. Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles in vitro. J. Virol. **78:**52–60.

29. **Ma, Y. M., and V. M. Vogt.** 2002. Rous sarcoma virus Gag protein-oligonucleotide interaction suggests a critical role for protein dimer formation in assembly. J. Virol. **76:**5452–5462.

30. **Melamed, D., M. Mark-Danieli, M. Kenan-Eichler, O. Kraus, A. Castiel, N. Laham, T. Pupko, F. Glaser, N. Ben-Tal, and E. Bacharach.** 2004. The conserved carboxy terminus of the capsid domain of human immunodeficiency virus type 1 Gag protein is important for virion assembly and release. J. Virol. **78:**9675–9688.

31. **Morellet, N., S. Druillennec, C. Lenoir, S. Bouaziz, and B. P. Roques.** 2005. Helical structure determined by NMR of the HIV-1 (345-392) Gag sequence, surrounding p2: implications for particle assembly and RNA packaging. Protein Sci. **14:**375–386.

32. **Muriaux, D., S. Costes, K. Nagashima, J. Mirro, E. Cho, S. Lockett, and A. Rein.** 2004. Role of murine leukemia virus nucleocapsid protein in virus assembly. J. Virol. **78:**12378–12385.

33. **Pepinsky, R. B., R. J. Mattaliano, and V. M. Vogt.** 1986. Structure and processing of the p2 region of avian sarcoma and leukemia virus gag precursor polyproteins. J. Virol. **58:**50–58.

34. **Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin.** 2004. UCSF Chimera: a visualization system for exploratory research and analysis. J. Comput. Chem. **25:**1605–1612.

35. **Phillips, J. M., P. S. Murray, D. Murray, and V. M. Vogt.** 2008. A molecular switch required for retrovirus assembly participates in the hexagonal immature lattice. EMBO J. **27:**1411–1420.

36. **Pornillos, O., B. K. Ganser-Pornillos, B. N. Kelly, Y. Hua, F. G. Whitby, C. D. Stout, W. I. Sundquist, C. P. Hill, and M. Yeager.** 2009. X-ray structures of the hexameric building block of the HIV capsid. Cell **137:**1282–1292.

37. **Pruggnaller, S., M. Mayr, and A. S. Frangakis.** 2008. A visualization and segmentation toolbox for electron microscopy. J. Struct. Biol. **164:**161–165.

38. **Royer, M., S. S. Hong, B. Gay, M. Cerutti, and P. Boulanger.** 1992. Expression and extracellular release of human immunodeficiency virus type 1 Gag precursors by recombinant baculovirus-infected cells. J. Virol. **66:**3230–3235.

39. **Rumlova-Klikova, M., E. Hunter, M. V. Nermut, I. Pichova, and T. Ruml.** 2000. Analysis of Mason-Pfizer monkey virus Gag domains required for capsid assembly in bacteria: role of the N-terminal proline residue of CA in directing particle shape. J. Virol. **74:**8452–8459.

40. **Sakalian, M., and E. Hunter.** 1999. Separate assembly and transport domains within the Gag precursor of Mason-Pfizer monkey virus. J. Virol. **73:**8073–8082.

41. **Sommerfelt, M. A., S. S. Rhee, and E. Hunter.** 1992. Importance of p12 protein in Mason-Pfizer monkey virus assembly and infectivity. J. Virol. **66:**7005–7011.

42. **Tang, C., Y. Ndassa, and M. F. Summers.** 2002. Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. Nat. Struct. Biol. **9:**537–543.

43. **Ulbrich, P., S. Haubova, M. V. Nermut, E. Hunter, M. Rumlova, and T. Ruml.** 2006. Distinct roles for nucleic acid in in vitro assembly of purified Mason-Pfizer monkey virus CANC proteins. J. Virol. **80:**7089–7099.

44. **Wilk, T., I. Gross, B. E. Gowen, T. Rutten, F. de Haas, R. Welker, H. G. Krausslich, P. Boulanger, and S. D. Fuller.** 2001. Organization of immature human immunodeficiency virus type 1. J. Virol. **75:**759–771.

45. **Wright, E. R., J. B. Schooler, H. J. Ding, C. Kieffer, C. Fillmore, W. I. Sundquist, and G. J. Jensen.** 2007. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J. **26:**2218–2226.

46. **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. U. S. A. **95:**7299–7304.

47. **Yu, F., S. M. Joshi, Y. M. Ma, R. L. Kingston, M. N. Simon, and V. M. Vogt.** 2001. Characterization of Rous sarcoma virus Gag particles assembled in vitro. J. Virol. **75:**2753–2764.

48. **Zhang, Y., and E. Barklis.** 1997. Effects of nucleocapsid mutations on human immunodeficiency virus assembly and RNA encapsidation. J. Virol. **71:**6765–6776.
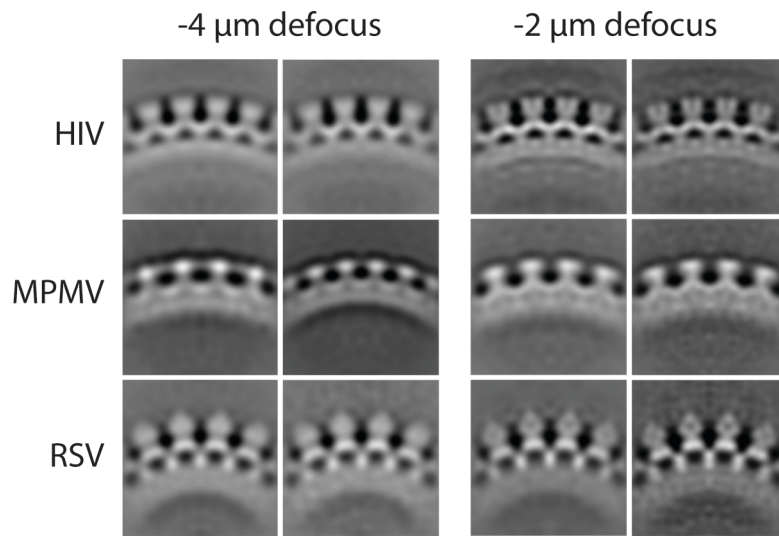
Supplementary figure 1

Orthoslices through reconstructions produced using each half of the datasets independently at (A) -4 µm defocus and (B) -2 µm defocus. For further explanation see main text.
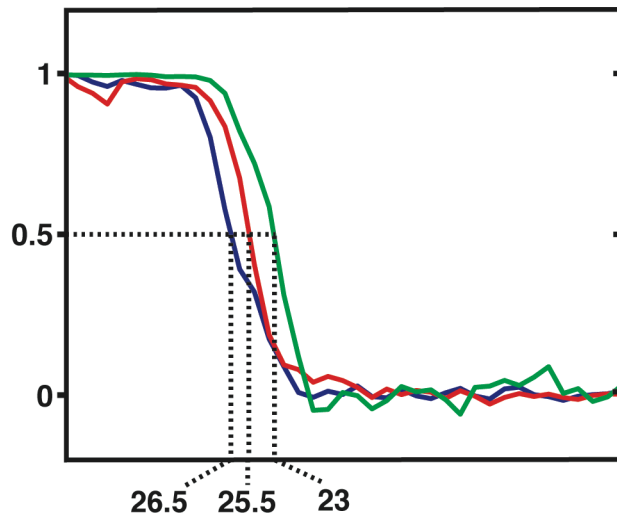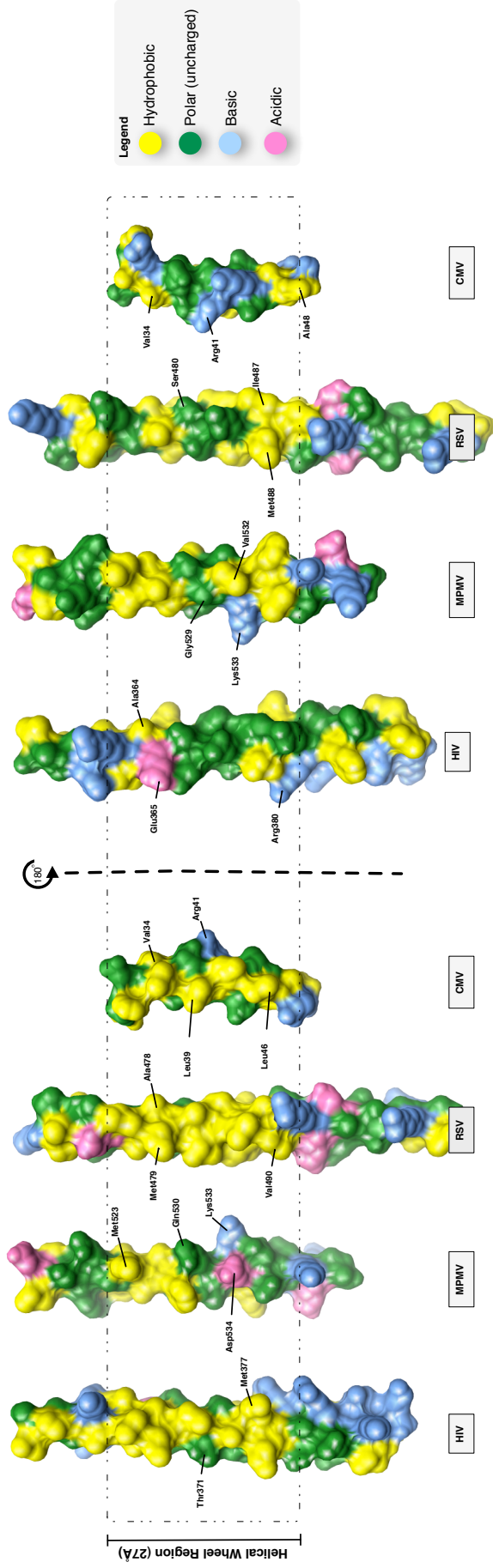
Supplementary figure 2

Fourier Shell Correlations between the CA-SP1 region from independent half reconstructions at approximately 2 μm defocus for HIV (red), M-PMV (blue) and RSV (green). The NC-nucleic acid layer was excluded from the comparison since it is not symmetrically ordered.

sp|P04591|GAG_HV1H2/284-428   284 DIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALG----PAATLEEMMTACQGVGGPGHKA 360
sp|P07567|GAG_MPMV/446-593   446 GVKQGPDEPFADFVHRLITTAGRIFGSAEAGVDYVKQLAYENANPACQAAIRPYR---KKTDLTGYIRLCSDIGPSYQQG 522
sp|P03322|GAG_RSVP/394-550   394 DIMQGPSESFVDFANRLIKAVEGSDLPPSARAPVIIDCFRQKSQPDIQQLIRTAPSTLTTPGEIIKYVLDRQKTAPLTDQG 474
sp|P17756|GAG_HV2D1/1-142   1 DVKQGPKESFQSYVDRFYKSLRAEQTDPAVKNWMTQTLIQNANPDCKLVLKGLG----MNPTLEEMLTACQGVGGPSQKA 77
sp|P19027|GAG_FIVSD/279-411   279 QLRQGAKEDYSSFIDRLFAQIDQEQNTAEVKLYLKQSLSIANANAECKKAMSHLK----PESTLEEKLRACQEIGSPGYKM 355
sp|P69731|GAG_EIAVC/276-417   276 NIRQGAKEPYPEFVDRLLSQIKSEGHPQEISKFLTDTLTIQNANEECRNAMRHLR---PEDTLEEKMYACRDIGTTKQKM 352
sp|P33458|GAG_CAEVC/280-414   280 LVKQKTNEPYEDFAARLLEAIDAEPVTQPIKDYLKLTLSYTNASADCQKQMDRTLGQRVQQASVEEKMQACRDVGSEGFKM 360
sp|P19558|GAG_BIV29/289-435   289 NIHQGPKEPYTDFINRLVAALEGMAAPETTKEYLLQHLSIDHANEDCQSILRPLG---PNTPMEKKLEACRVVGSQKSKM 365
sp|P14076|GAG_HTL1C/265-395   265 SILQGLEEPYHAFVERLNIALDNGLPEGTPKDPILRSLAYSNANKECQKLLQARGH---TNSPLGDMLRACQAWTPKDKTK 342
sp|P25058|GAG_BLVAU/243-387   243 TIVQGPAESYVEFVNRLQISLADNLPDGVPKEPIIDSLSYANANKECQQILQGRGL---VAAPVGQKLQACAHWAPKTKQP 320
sp|P31622|GAG_JSRV/398-551   398 KVRQGPDEPYQDFVARLLDTIGKIMSDEKAGMVLAKQLAFENANSACQAALRPYR---KKGDLSDFIRICADIGPSYMQG 474
sp|P11284|GAG_MMTVC/417-569   417 GLKQGNEESYETFISRLEEAVYRVMPRGEGSDILIKQLAWENANSLCQDLIRPMR---KTGTMQDYIRACLDASPAVVQG 493

sp|P04591|GAG_HV1H2/284-428   361 RVLAEAMSQVTNSATI-----MMQRGNFRNQRKIVKCFNCGKEGHTARNCRAPR------KKGCWKCGKEGHQMKDCTE 428
sp|P07567|GAG_MPMV/446-593   523 LAMAAAFSGQTVKDFL------NNKNKEKGGCCFKCGKKGHCHFAKNCHEHAHNNAEPKVPGLCPRCKRGKHWANECKS 593
sp|P03322|GAG_RSVP/394-550   475 --IAAAMSSAIQPLIMAVVNRERDGQTGSGGRARGLCYTCGSPGHYQAQCPKKRKSG--NSRERCQLCNGMGHNAKQCRK 550
sp|P17756|GAG_HV2D1/1-142   78 RLMAEALKEALTKVQVVQ--------PFAAAQQ--RRAIRCWNCGKEGHSAKQCRAPR------RQGCWKCGKPGHIMANCPE 142
sp|P19027|GAG_FIVSD/279-411   356 QLLAEALTKVQVVQ-------SKGSGP------VCFNCKKPGHLARQCRDVK------KCNKCGKPGHLAAKCWQ 411
sp|P69731|GAG_EIAVC/276-417   353 MLLAKALQTGLAGPF------KGGALKGGPLKAAQTCYNCGKPGHLSSQCRAPK------VCFKCKQPGHFSKQCRS 417
sp|P33458|GAG_CAEVC/280-414   361 QLLAQALRP------GKGKGNQP-QRCYNCGKPGHQARQCRQGI----ICHNCGKRGHMQKECRG 414
sp|P19558|GAG_BIV29/289-435   366 QFLVAAMKEMGIQSPIPAHTPEAYASQTSGPEDGRRCYGCGKTGHLKRNCKQQK-----CYHCGKPGHQARNCRS 435
sp|P14076|GAG_HTL1C/265-395   343 --VLVVQPKK-P-PPNQP-------CFRCGKAGHWSRDCTQPR------PPPGPCPLCQDPTHWKRDCPR 395
sp|P25058|GAG_BLVAU/243-387   321 AILVHTPGPKMPGPR-------QPAPKRPPPGPCYRCLKEGHWARDCPTKTTG--PPPGPCPICKDPSHWKRDCPT 387
sp|P31622|GAG_JSRV/398-551   475 IAMAAALQGKSIKEVL--FQQQARNKKGLQKSGNSGCFVCGQPGHRAAVCPQKHQTSV--NTPNLCPRCKKGKHWARDCRS 551
sp|P11284|GAG_MMTVC/417-569   494 MAYAAAMRGQKYSTFV--KQTYGGGKGGQGSKGPVCFSCGKTGHIKRDCKEEKGSKR--APPGLCPRCKKGYHWKSECKS 569

Supplementary figure 3

Alignment of the C-CA-NC region of Gag from 12 viral species (see methods). Residues conserved in more than 75% of species are coloured using ClustalX colouring, numbered from the first residue of Gag.

Supplementary figure 4

The complete region between the final helix of CA and the first Zn finger of NC modelled as an alpha helix using Chimera compared to the solved structure of the cucumber mosaic virus coat protein helix for illustrative purposes (colouring as in Figure 3D). The left hand panel shows the most likely hydrophobic face the right hand panel shows the left hand panel rotated 180°. The 18-residue region considered in the helical wheel corresponds to a 24 Å alpha-helix (Figure 3C).

In Vitro Assembly of Virus-Like Particles of a Gammaretrovirus, the Murine Leukemia Virus XMRV

Romana Hadravová,[a] Alex de Marco,[b] Pavel Ulbrich,[c] Jitka Štokrová,[a] Michal Doležal,[a,c] Iva Pichová,[a] Tomáš Ruml,[c] John A. G. Briggs,[b] and Michaela Rumlová[a]

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., IOCB and Gilead Research Center, Prague, Czech Republic[a]; Structural and Computational Unit, European Molecular Biology Laboratory, Heidelberg, Germany[b]; and Department of Biochemistry and Microbiology, Institute of Chemical Technology, Prague, Czech Republic[c]

Immature retroviral particles are assembled by self-association of the structural polyprotein precursor Gag. During maturation the Gag polyprotein is proteolytically cleaved, yielding mature structural proteins, matrix (MA), capsid (CA), and nucleocapsid (NC), that reassemble into a mature viral particle. Proteolytic cleavage causes the N terminus of CA to fold back to form a β-hairpin, anchored by an internal salt bridge between the N-terminal proline and the inner aspartate. Using an *in vitro* assembly system of capsid-nucleocapsid protein (CANC), we studied the formation of virus-like particles (VLP) of a gammaretrovirus, the xenotropic murine leukemia virus (MLV)-related virus (XMRV). We show here that, unlike other retroviruses, XMRV CA and CANC do not assemble tubular particles characteristic of mature assembly. The prevention of β-hairpin formation by the deletion of either the N-terminal proline or 10 initial amino acids enabled the assembly of ΔProCANC or Δ10CANC into immature-like spherical particles. Detailed three-dimensional (3D) structural analysis of these particles revealed that below a disordered N-terminal CA layer, the C terminus of CA assembles a typical immature lattice, which is linked by rod-like densities with the RNP.

T he assembly of fully infectious retroviruses comprises two different phases: immature- and mature-particle formation. Initially, the virus is formed in immature form by self-association of the major structural polyprotein Gag into spherical particles either at the plasma membrane (alpharetroviruses, gammaretroviruses, or lentiviruses) or at a distinct site within the infected cell (betaretroviruses). The particles are organized by lattices of Gag and Gag-derived polyproteins that associate with the plasma membrane that becomes the outer layer of the immature particle released during budding. Upon release from the host cell, the Gag precursor is proteolytically processed into N-terminal membrane binding matrix protein (MA), two-domain capsid protein (CA), and RNA-binding nucleocapsid protein (NC). This cleavage leads to a dramatic morphological change and reorganization of the virus particle from its immature form into a mature, infectious virion.

Gag itself is sufficient for immature-particle assembly. The interactions critical for immature-particle assembly are mediated by CA-CA and NC-RNA interfaces. CA consists of two subdomains, the CA–N-terminal domain (NTD) and the CA–C-terminal domain (CTD), connected by a short sequence. Despite a low level of sequence homology, all available retroviral CA-NTD structures show conserved structural motifs consisting of six or seven α-helices and an N-terminal β-hairpin anchored by a salt bridge between the N-terminal proline and an aspartate in helix 3 (10, 20, 29, 30, 42–44). CA-CTD contains a dimerization domain and is essential for the assembly of Gag. In the alpharetroviruses (e.g., Rous sarcoma virus [RSV]) and lentiviruses (e.g., HIV-1), a short spacer peptide (SP) (12 or 14 amino acids, respectively) separates CA from NC. This SP domain plays a critical role in assembly, as mutations or deletions in the SP influence immature-spherical-particle assembly (1, 28, 33, 37). Both HIV-1 and RSV SP sequences were also suggested to function as molecular switches, because the deletion of SP1 of HIV-1 Gag led to the formation of

tubular instead of spherical particles (23). Similarly, mutations within the SP sequence of RSV Gag change spherical-particle formation to tubular-particle formation (28).

The appearances of all immature retroviral particles observed by transmission electron microscopy (TEM) are similar; they are roughly spherical, with an inner electron-dense ring. Based on cryo-electron microscopy (cEM) studies, the Gag polyprotein is arranged radially and is associated with the plasma membrane through its MA domain, with the NC domain oriented toward the center of the particle (65, 67). Three-dimensional (3D) reconstructions of immature HIV-1 and RSV particles obtained by cryo-electron tomography have shown that the immature Gag lattice adopts a hexagonal arrangement (5, 66). To form a closed sphere from a hexagonal lattice, the incorporation of pentamers or irregular defects is required. Pentamers have not been observed in the immature Gag lattice; instead, the Gag hexameric lattice closes through the incorporation of heterogeneously shaped defects (6, 35). Upon maturation, NC with bound RNA condenses in the center of the particle, surrounded by a core shell assembled from released CA. The shapes of the mature cores differ according to the retrovirus genus: in gamma-, delta-, and alpharetroviruses, they are spherical; in betaretroviruses, cylindrical; and in lentiviruses, conical. These assemblies are composed of a curved hexagonal lattice closed by incorporation of CA pentamers. The final
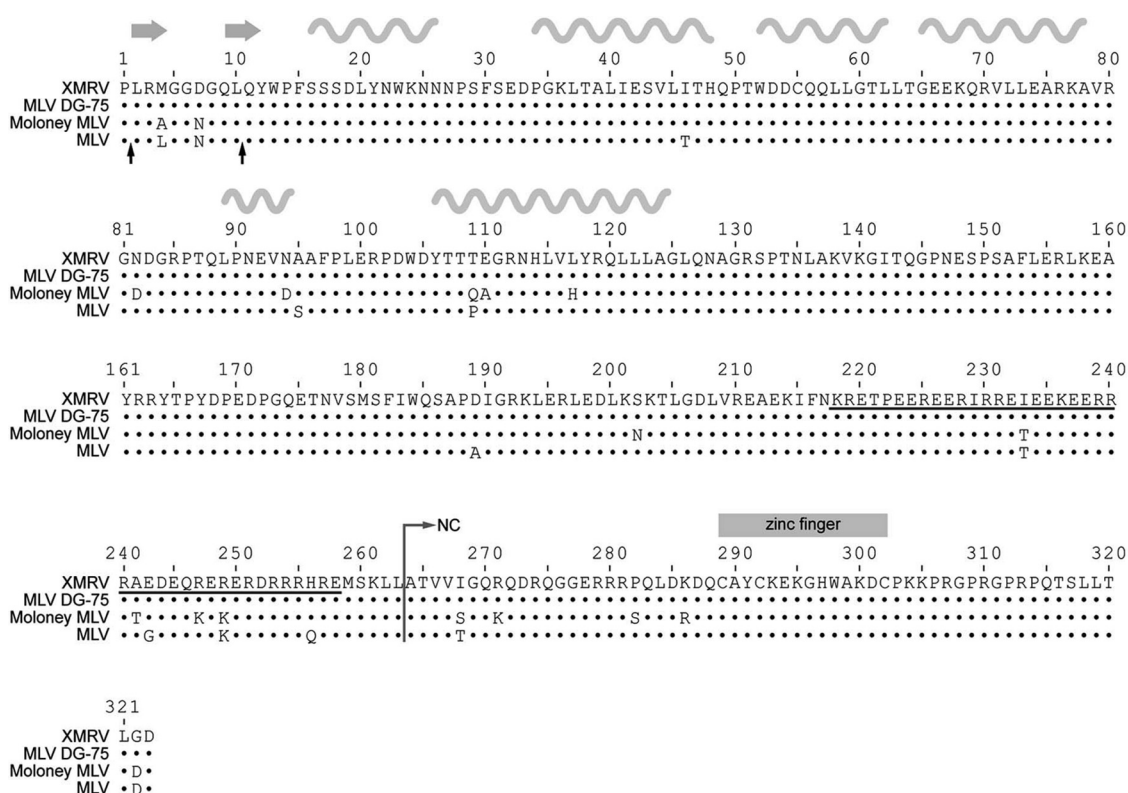
FIG 1 Amino acid alignment of XMRV and MLV DG-75, Moloney MLV, and MLV (strain BM5) CANC proteins. The dots represent identical amino acid residues, the black arrows under the alignment show the start sites of XMRV ΔProCANC and Δ10CANC truncated proteins, and the underlined sequence shows the charged residues at the C termini of MLV and XMRV CA. The positions of β-strands and α-helices from the solved MLV CA-NTD structure (pdb 3bp9) are shown by the gray arrows and wavy lines, respectively, above the alignment.

shapes of the retroviral cores are then determined by the locations of the pentamers. The core of the most studied retrovirus, HIV-1, is a fullerene-like cone composed of a hexameric lattice with 12 pentamers, 5 at the narrow end of the cone and 7 at the broad end of the cone (18, 36, 47). Although the CA domain drives both mature- and immature-particle assembly, the lattices of immature particles differ from those of the mature particles (5–7, 66). The spacing of the immature HIV-1 Gag lattice is smaller than that in the mature particle, and the positions of the CA-NTD and CA-CTD also differ in the immature and mature lattices (6).

Gammaretroviruses, like other type C retroviruses, assemble their immature particles at the plasma membrane. Similarly to other retroviruses, their genome contains three open reading frames: *gag*, *pol*, and *env*. One member of the gammaretroviral family, the xenotropic murine leukemia virus-related virus (XMRV), was recently discovered in association with human prostate tumors (58) and chronic fatigue syndrome (CFS) (38). However, the association has not been confirmed in other studies (15–17, 25, 26, 51–53, 56, 60), thus the relevance of XMRV to human disease remains unclear. Recent publications have provided strong evidence that XMRV does not cause CFS (32, 45, 49, 54). Its genome is most similar to that of exogenous murine leukemia virus (MLV) DG-75, isolated from a human B-lymphoblastoid cell line (48), with which it shares 94% sequence identity (58). The XMRV Gag polyprotein consists of 536 amino acids and is separated from the Pro-Pol sequence by a UAG stop codon. Based on the sequence similarity with the closely re-

lated MLV, it is expected that XMRV Gag also comprises four structural proteins: MA, p12, CA, and NC. The MLV capsid protein, like all retroviral CAs, contains an N-terminal β-hairpin that is stabilized by Pro1-Asp54 interaction (43, 44). This hairpin is formed only upon maturation, and prevention of its maturation markedly reduces viral infectivity (61, 62, 64).

Similarly to betaretroviruses, gammaretroviruses lack a spacer peptide between CA and NC. However, gammaretroviruses possess a unique feature, that is, the presence of a 41-amino-acid stretch rich in both positively and negatively charged residues at the C terminus of CA (Fig. 1, underlined sequence). This region, due to its possible α-helical structure, was named the "charged assembly helix" and is important for the assembly and production of infectious MLV (9, 63).

Studies on *in vitro*-assembled virus-like particles (VLPs) have been proven to be a valuable tool for gaining basic information on retroviral assembly. A number of studies have mimicked the assembly of both immature virus particles and mature cores *in vitro* using purified recombinant Gag-derived capsid-nucleocapsid (CANC) proteins (8, 21, 22, 31, 41, 57, 68). Although the *in vitro* assembly of retroviral genera, including the lentivirus HIV-1 (8, 12, 21, 22), the betaretrovirus Mason-Pfizer monkey virus (M-PMV) (31, 57), and the alpharetrovirus RSV (8, 40, 41, 68), has been intensively studied, there is a lack of information on the *in vitro* assembly of gammaretroviruses. *In vitro* assembly was achieved by using His-tagged MLV CA molecules anchored to Ni^{2+}-chelating lipid nanotubes as an assembly substrate (24). Re-

cently, a comparison of MLV and HIV-1 monomeric Gag molecules revealed their different behaviors in solution. In contrast to HIV Gag, MLV Gag has a much weaker propensity for interprotein interactions and has a rigid and extended structure, likely caused by a proline-rich region between MA and p12 (11). In this work, Datta et al. also demonstrated that MLV Gag could be assembled *in vitro* into fragile VLPs with an irregular appearance.

Here, we present the assembly of XMRV CANC protein into immature virus-like spherical VLPs by using an *in vitro* expression/assembly system. We prepared three N-terminal variants of XMRV CANC and compared their ability to assemble. We found that formation of the N-terminal β-hairpin must be prevented to obtain immature-like particles from XMRV CANC. Using electron microscopy and cryo-electron tomography, we have characterized the structure of assembled immature XMRV VLPs and compared it with that of other retroviruses.

MATERIALS AND METHODS

Cloning. All expression vectors were prepared by standard cloning techniques, propagated in *Escherichia coli* DH5α, and verified by sequencing. The fragment encoding XMRV CANC was obtained by PCR (using primers 5′CAT ATG CCA CTC CGC ATG GGG and 3′CTC GAG CTA GTC ACC TAA GGT CAG G) of cDNA from human prostate carcinoma 22Rv1 cells infected with XMRV (kindly provided by M. Hajduch). The PCR fragment was digested with NdeI and XhoI and ligated into an expression vector, pET-22b, under the T7 promoter. The CANC-derived constructs were prepared analogously, with the 5′ primers CAT ATG CAG TAC TGG CCG TTT TCC TC for Δ10CANC and CAT ATG CTC CGC ATG GGG GGA G for ΔProCANC.

Bacterial expression. Luria-Bertani medium containing ampicillin (100 μg/ml) was inoculated with *E. coli* BL21(DE3) cells carrying the appropriate construct. At the optical density at 590 nm of ~0.8, expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. The cells were harvested 4 h postinduction.

Protein purification. Bacterial cells were resuspended in lysis buffer A (50 mM Tri-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0), disrupted with lysozyme, sonicated, and incubated with sodium deoxycholate. After centrifugation at 10,000 × *g*, proteins were solubilized from pellets by the addition of 0.5% Triton X-100 and 1 to 1.5 M NaCl in lysis buffer A. The XMRV protein was precipitated from the supernatant by ammonium sulfate (final concentration, 25% [wt/vol]). After low-speed centrifugation, the pellet was resuspended in buffer E (20 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 50 μM ZnCl₂, 10 mM dithiothreitol [DTT], 1 mM phenylmethylsulfonyl fluoride [PMSF]), dialyzed overnight against the same buffer at 4°C, and loaded on the top of a DEAE cellulose column. The flowthrough and washing fractions were pooled and loaded on the top of a phosphocellulose column. The bound protein was eluted by an NaCl gradient from 0.1 M to 1 M NaCl in buffer E. The fractions containing the desired proteins were dialyzed overnight against a storage buffer (buffer E with 0.5 M NaCl), concentrated, and loaded on the top of a Sephadex G-100 column. The purified proteins were concentrated (CANC to 5 mg/ml, ΔProCANC to 2 mg/ml, and Δ10CANC to 4 mg/ml) and stored at −70°C. The purity of each protein was analyzed by SDS-PAGE (see Fig. 3), and the protein identities were confirmed by N-terminal sequencing.

***In vitro* assembly.** An aliquot of 60 μg of purified protein was mixed with 6 μg of DNA (oligonucleotide 30-mer, 40-mer, λDNA, or phage M13 single-stranded DNA [ssDNA]) or RNA (MS2) in a total volume of 100 μl of the storage buffer (20 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 50 μM ZnCl₂, 10 mM DTT, 1 mM PMSF). The mixture was dialyzed from 2 h to overnight against the assembly buffer (50 mM Tris buffer, pH 8.0, containing 100 mM NaCl and 1 μM ZnCl₂) at 4°C. One hundred microliters of the assembled material (in the presence or absence of oligonucleotide) was loaded on the top of a 20 to 65% (wt/wt) sucrose gradient and cen-

trifuged at 35,000 rpm in a Beckman SW41Ti for 16 h at 4°C. Individual fractions were analyzed by Western blot analysis using rabbit anti-XMRV CANC antibodies. The sucrose density in the fractions was determined using a refractometer.

Electron microscopy. Particles formed during the assembly reaction were negatively stained with 4% sodium silicotungstate (pH 7.4) on carbon-coated grids and studied by transmission electron microscopy (JEOL JEM-1200EX) with a microscope operated at 60 kV. For thin-section electron microscopic analysis, the bacterial pellets of the induced samples were fixed in 2.5% glutaraldehyde, postfixed in 1% osmium tetroxide, dehydrated by applying an ethanol series, and embedded in Agar 100 epoxy resin. Ultrathin sections (70 nm) were counterstained with saturated uranyl acetate and lead citrate.

Cryo-electron microscopy. *In vitro*-assembled particles were deposited on C-flat holey carbon grids and vitrified by plunge freezing in liquid ethane. Tilt series were collected on an FEI Tecnai F30 "Polara" transmission electron microscope with a Gatan GIF 2002 postcolumn energy filter and a 2,000 by 2,000 Multiscan charge-coupled device (CCD) camera. Data collection was performed at 300 kV with an electron dose of approximately 30 e⁻/Å². The nominal defocus was 2.0 μm, with a magnification of ×34,000, resulting in a pixel size at the specimen level of 4 Å.

Cryo-electron tomography. *In vitro*-assembled particles were mixed with 10-nm gold beads, deposited on C-flat holey carbon grids, and vitrified by plunge freezing in liquid ethane. Tilt series were collected on an FEI Tecnai F30 "Polara" transmission electron microscope with a Gatan GIF 2002 postcolumn energy filter and a 2,000 by 2,000 Multiscan CCD camera. Data collection was performed at 300 kV using the SerialEM software package, and tomograms were reconstructed using the IMOD software package (34). Tilt series were collected between −60° and +60° with an angular increment of 3° and a total electron dose of approximately 70 e⁻/Å². The defocus range was between 3.0 and 4.0 μm, with a magnification of ×34,000, resulting in a pixel size at the specimen level of 4 Å.

2D Fourier analysis. Seventy individual particles were boxed out using an EMAN boxer (39) and padded to a final size of 1,024 by 1,024 pixels. A soft spherical mask was applied with a radius of 22 nm for the rotationally averaged power spectrum or 15 nm for the rotationally aligned power spectrum. (The larger mask includes more data, giving the stronger peaks necessary for the alignment. The narrower mask minimizes the smearing of the peak due to particle curvature.) Calculation of power spectra and their rotational average and rotational alignment were performed in IMAGIC (59).

RAF analysis. Radius-angle-frequency (RAF) plots were generated as described previously (13). Briefly, subtomograms of 384 nm³ were extracted from tomograms of 15 VLPs. The extraction was done along the surface of a sphere centered in the center of each VLP and with a radius equal to the mean radius at CA-CTD level. The subtomograms underwent an orthographic radial projection. The focus of the projection was determined as the geometric center of the particle from which the subtomogram was extracted. The resulting volumes corresponded to flattened subtomograms. The volumes were then padded to 1,024 by 1,024 pixels in *xy*, and 2D power spectra were calculated at each radius. The rotation-autocorrelation function with a rotation range between 0° and 180° was calculated for each power spectrum. For each subtomogram, this generates a 3D plot with radius on one axis and the other two axes representing those of the rotational autocorrelation function of the power spectrum, namely, angle and frequency. The presence of a peak at a particular point in radius, frequency, and angle indicates that at that radius in the virus, the 2D power spectrum of the protein layer has peaks at that frequency which are arranged rotationally, symmetrically repeating at that angle.

Subtomogram averaging. The extracted subtomograms were iteratively aligned and averaged using a six-dimensional search as described previously (6). The initial reference used for the alignment was the average of the subtomograms in the extraction position. Sixfold symmetry was applied to the average at each iteration. The threshold for the subtomo-
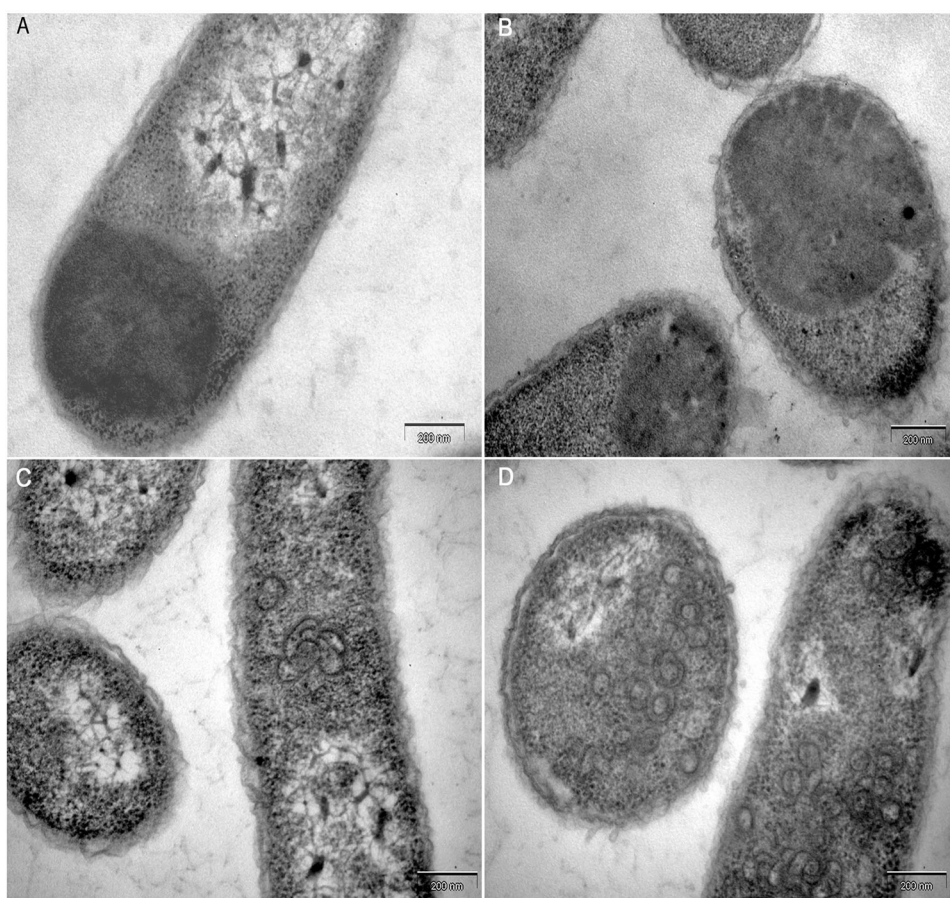
FIG 2 Electron micrographs of thin sections of *E. coli* cells expressing XMRV CA (A), CANC (B), ΔProCANC (C), and Δ10CANC (D).

grams that had to be averaged was set to the mean cross-correlation value of all subtomograms to the reference.

RESULTS

Bacterial and *in vitro* assembly. To obtain the XMRV expression vector, the CA and CANC regions were amplified using cDNA of human prostate carcinoma 22Rv1 cells infected with XMRV. DNA sequencing confirmed 100% identity of the CANC with the CANC of XMRV clone VP62 (Fig. 1). To study the ability of XMRV CA and CANC to assemble into VLPs, we used both bacterial and *in vitro* assembly systems (50, 57). TEM of thin-sectioned *E. coli* BL21(DE3) expressing XMRV CA and CANC showed amorphous inclusion bodies instead of assembled mature-core-like structures (Fig. 2A and B). This finding was surprising, because several studies have demonstrated that retroviral CANCs of different species (HIV-1 and RSV), with Pro1 at their N termini, assemble into tubular, mature-core-like structures (8, 21, 22, 41).

It is known that the prevention of β-hairpin formation (by either the deletion or substitution of Pro1 or by N-terminal extension of CA) can shift the assembly mode to formation of immature-virus-like structures (21, 46, 50, 57, 61). We therefore decided to truncate the XMRV CANC construct by the deletion of the first proline residue (ΔProCANC) with the aim of assembling immature-like structures. TEM analysis of the cells expressing ΔProCANC protein indicated its tendency to assemble into nu-

merous curved structures (Fig. 2C), together with a small number of approximately spherical particles. In order to delete a substantial portion of the N-terminal β-hairpin, we next generated a further N-terminal truncation of CANC, locating the start of the XMRV CA protein 10 amino acids downstream from Pro1, at Gln11 (Fig. 1). TEM of the thin sections of *E. coli* expressing the resulting Δ10CANC construct showed the presence of roughly spherical particles (Fig. 2D).

To further analyze the ability of XMRV CANC-derived proteins to assemble *in vitro*, all studied proteins were purified using procedures similar to those described earlier for M-PMV and HIV-1 CANC (8, 41, 57) (Fig. 3A). An *in vitro* assembly reaction was performed by overnight dialysis of a mixture of purified XMRV CANC-derived proteins with DNA (oligonucleotide 30-mer, 40-mer, λ DNA, and phage M13 ssDNA) or RNA (MS2). To investigate the homogeneity of the assembled particles, the material was centrifuged through a linear 20 to 65% sucrose gradient, and individual fractions were analyzed by Western blotting using rabbit anti-XMRV CANC antibody. In the presence of the oligonucleotide (40-mer), the majority of the assembled particles were found in the fractions with sucrose densities of 1.19 to 1.23 g/ml, corresponding to the properly assembled spherical particles (Fig. 3B, top). In the absence of the oligonucleotides, the bulk of unassembled material remained at the top of the sucrose gradient (Fig. 3B, bottom).
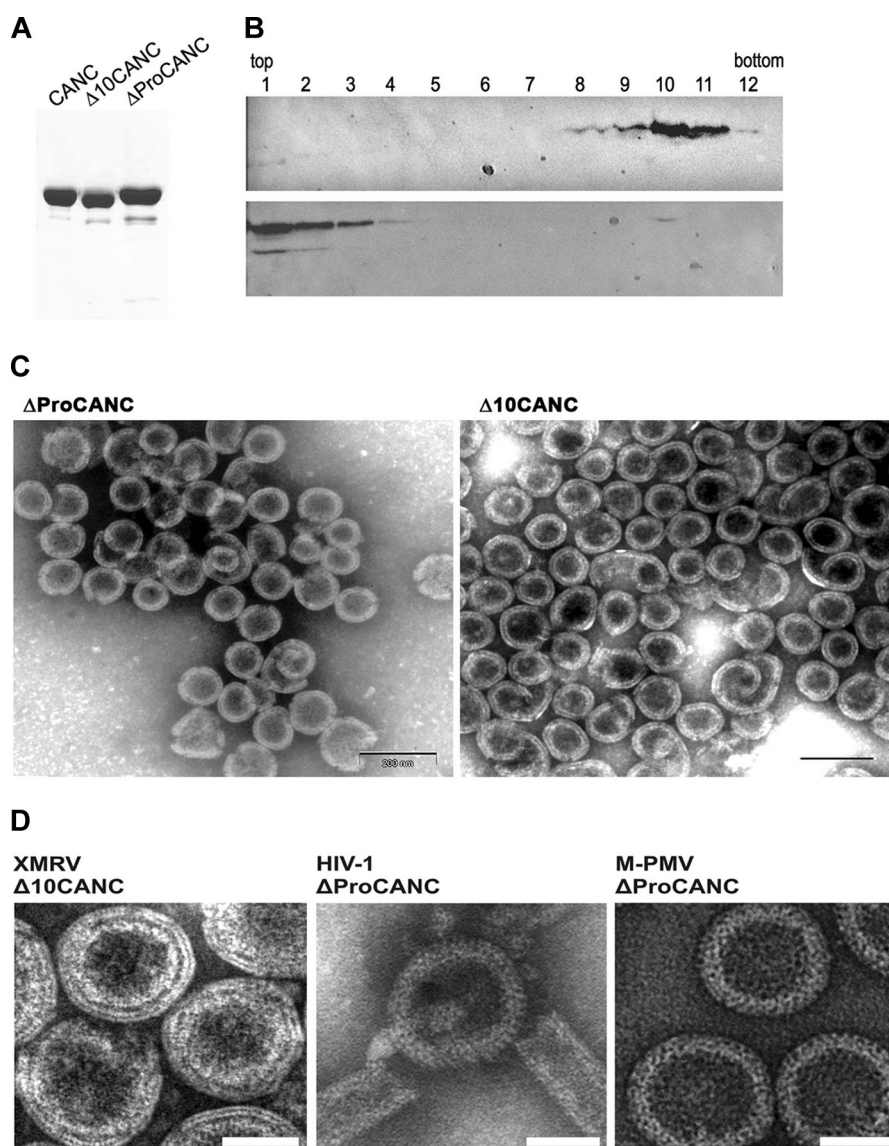
FIG 3 (A) SDS-polyacrylamide analysis of purified XMRV proteins (Coomassie blue-stained gel). (B) Western blot analysis of *in vitro*-assembled XMRV particles in sucrose gradient fractions in the presence (upper blot) or in the absence (lower blot) of oligonucleotides. Δ10CANC protein (60 μg) was mixed with 6 μg of oligonucleotide 40-mer, dialyzed overnight against the assembly buffer, and centrifuged to equilibrium through a linear 20 to 65% sucrose gradient. The fractions were analyzed by Western blotting with rabbit anti-XMRV CANC antibodies. The sucrose densities in fractions 8 to 12 were as follows: 8, 1.19 g/ml; 9, 1.21 g/ml; 10, 1.22 g/ml; 11, 1.23 g/ml; 12, 1.25 g/ml. (C) TEM images of negatively stained *in vitro*-assembled material from XMRV ΔProCANC (left) and XMRV Δ10CANC (right). (D) TEM images of negatively stained VLPs assembled *in vitro* from XMRV Δ10CANC, HIV-1 ΔProCANC, and M-PMV ΔProCANC. Scale bars,e 50 nm.

The gradient profiles were similar for both the ΔProCANC and Δ10CANC proteins. The samples of the *in vitro*-assembled particles were negatively stained with 4% sodium silicotungstate (pH 7.4) and studied by TEM. In contrast to the XMRV CANC protein, where only free or aggregated proteins were observed, the ΔProCANC and Δ10CANC proteins assembled into spherical particles (Fig. 3C). Particles of XMRV ΔProCANC and Δ10CANC were not detected if the *in vitro* assembly was carried out in the absence of nucleic acid or at an NaCl concentration higher than 0.5 M (data not shown), suggesting that assembly must be initiated by the addition of nucleic acids under suitable conditions. Similar assembly requirements have been observed for the *in vitro*-assembled particles of RSV

CANC (8), HIV-1 CANC (8, 21, 22), and M-PMV ΔProCANC (57).

TEM images of negatively stained *in vitro*-assembled XMRV Δ10CANC particles were then compared to those of HIV-1 and M-PMV ΔProCANC (Fig. 3D). The major difference between the particles was the appearance of particle protein layers. While the HIV-1 and M-PMV ΔProCANC particles appeared as a single-layer protein shell with visible rod-like projections (Fig. 3D), the XMRV Δ10CANC particle's shell appeared as three smoother protein layers separated by distinct rings of lower electron density (Fig. 3D).

To study the fine structure of the *in vitro*-assembled XMRV particles in more detail and to avoid artifacts associated with neg-
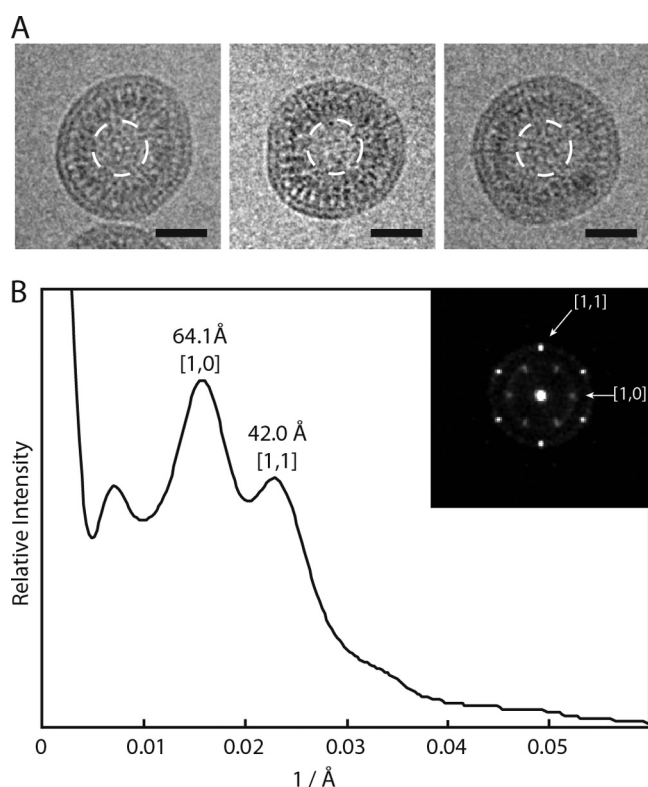
FIG 4 Cryo-electron microscopy of XMRV Δ10CANC particles. (A) Cryo-electron micrographs of particles assembled *in vitro* from XMRV Δ10CANC. Scale bars, 50 nm. The white dashed circles show the region used for Fourier analysis. (B) Rotationally averaged power spectra of the particle centers, showing the positions of the [1,0] and [1,1] reflections from a hexameric lattice. The inset displays the sum of the rotational aligned power spectra, indicating the presence of a hexagonal lattice.

ative staining, we carried out cEM of particles made of Δ10CANC protein. The sample was prepared by the dialysis of Δ10CANC against pH 8.0 buffer in the presence of lambda double-stranded DNA (dsDNA). The CEM images showed three density layers, with the outer two layers close together (Fig. 4A). The mean size of the particles was 82 nm ± 11 nm.

We compared our images of the *in vitro*-assembled XMRV particles (Fig. 4A) with published images of PR⁻ MLV (Fig. 3A and 5 in Yeager et al. [67]). The numbers of layers and the appearances of the layers are extremely similar. For an optimal comparison, we applied the same image analysis procedure that was used for PR⁻ MLV. We calculated the rotationally averaged power spectra of the particle centers, which showed peaks corresponding to repeating elements in the protein lattice (Fig. 4B). Two peaks were seen, centered at 64 ± 9 Å and 42 ± 3 Å. If these are indexed as [1,0] and [1,1] reflections of a hexagonal lattice, they describe a unit spacing of 74 Å or 73 Å. The equivalent analysis for PR⁻ MLV is shown in Fig. 6 in Yeager et al. (67). The pattern of peaks is the same. For PR⁻ MLV, the peaks are centered at 67 ± 11 Å and 45 ± 10 Å, describing a unit cell spacing of 77 Å and 78 Å. We next carried out a rotational alignment and averaging of the power spectra, as described by Briggs et al. (5), revealing that a hexagonal pattern is present in the XMRV particles (Fig. 4B) and confirming the assignment of the [1,0] and [1,1] reflections. Similar patterns were described for individual particles of PR⁻ MLV (Fig. 3 in Yeager et al. [67]).

To further understand the structural organization of the XMRV particles, we collected cryo-electron tomograms (Fig. 5A) and directly compared the radial density profile of XMRV Δ10CANC with those of HIV-1 (using data sets from de Marco et al. [12]). The radial density profile of XMRV Δ10CANC is similar to that of HIV-1 (Fig. 5B). By comparison with HIV-1, the two outer density layers correspond to the NTD and CTD of CA and the inner layer to NC with bound nucleic acid.

In order to identify which layers show 6-fold symmetry, we used the 3D information present in the cryo-electron tomograms and performed RAF analysis (13) (Fig. 5C). Peaks in the RAF plot indicate the presence of symmetrically ordered density at a particular radius and with a particular spacing. We identified two clear peaks: one corresponds to the [1,0] reflection of the expected hexameric lattice in the region between NC and the CA-CTD, the other to the [1,1] reflection of the hexameric lattice in the CA-CTD region. *In vitro*-assembled HIV particles show the same two peaks, as well as two further peaks in the region between CA-CTD and in the CA-NTD region. These peaks are missing in XMRV, indicating that, unlike in HIV, the CA-NTD region of the XMRV particles does not show any detectable hexameric lattice.

Knowing that the region between the RNP and the CA-CTD adopts a hexagonal lattice, we performed subtomogram averaging in order to get a 3D representation of the structure (Fig. 6A). Consistent with the RAF analysis, outermost in the structure is a poorly featured layer that corresponds to the CA-NTD (blue in Fig. 6B). In contrast, the CA-CTD domain (green in Fig. 6B) forms an arrangement similar to that seen in other retroviral immature lattices (12). As in HIV and RSV, the CA-CTD of XMRV appears to be connected to the RNP (gray in Fig. 6B) via rod-like structures (red in Fig. 6B).

DISCUSSION

The release of the N and C termini of CA upon maturation leads to local conformational transitions that influence the overall structure of the protein and thus creates novel protein-protein interfaces (for a review, see reference 19). The model for subunit arrangement in immature particles implies that the hexagonal lattice is held together by intra- and interhexamer interactions within SP and CA-CTD (66). The CA-NTD is not absolutely required for immature-particle assembly (2, 4). Cleavage of Gag by viral protease allows the formation of a β-hairpin at the N terminus of CA-NTD and the separation of CA from the putative SP1 helical bundle. This induces structural rearrangement of CA and the formation of mature hexameric CA-NTD rings that hold together the mature lattice. CA-CTD interactions also mediate interhexamer interactions in the mature lattice.

Structural and biochemical data suggest that the ordered arrays of CA are regulated by the presence (β^+) or absence (β^-) of the β-hairpin at the N terminus and the presence (SP⁺) or absence (SP⁻) of the SP sequence at the CA C terminus, thus determining the particle morphology (23, 28). By this convention, within the immature virus, CA is in the form β^-/SP⁺, and within the mature virus, it is in the form β^+/SP⁻. *In vitro*, the HIV-1 ΔMACANC and RSV ΔPRGag proteins, where β-hairpin formation is prevented by downstream N-terminal sequences of CA but SP is present (β^-/SP⁺), assembled into immature spherical particles (21, 27). *In vitro* and *in vivo* assembly studies on HIV CANC and RSV CANC show that the presence of both N- and C-terminal deter-

A

B

C
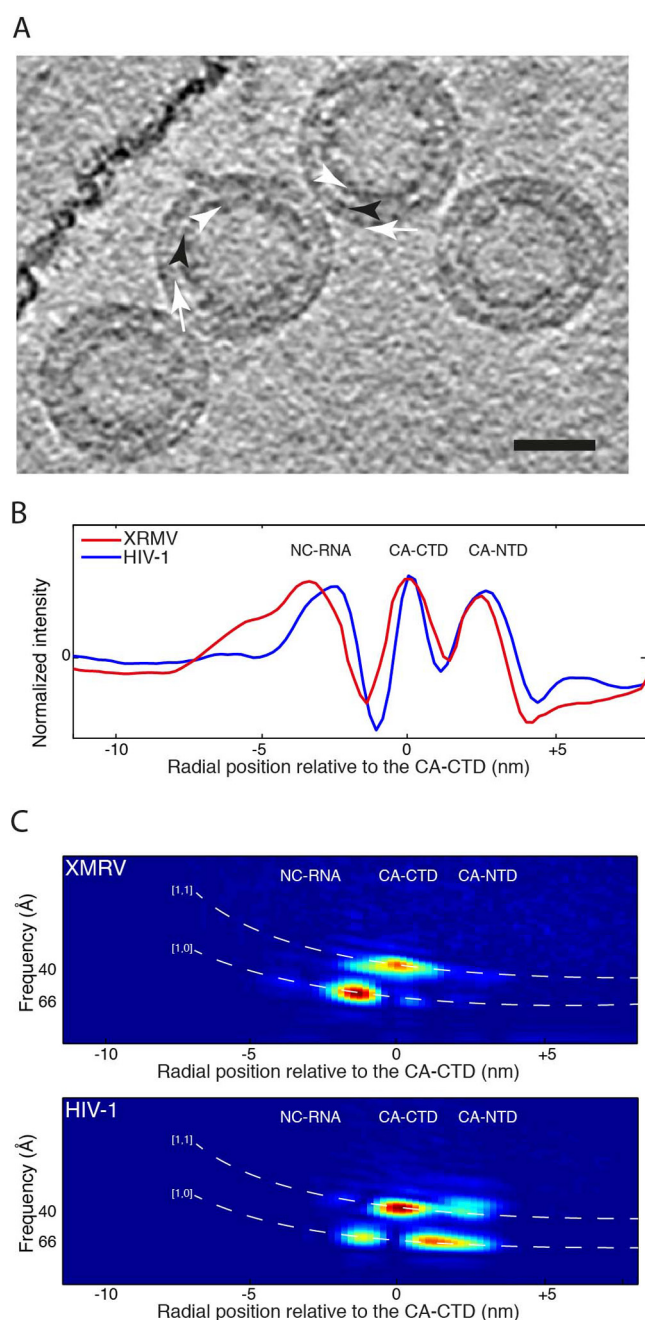
FIG 5 Cryo-electron tomography of XMRV Δ10CANC particles. (A) Computational slice 0.8 nm thick through a cryo-electron tomogram of particles assembled *in vitro* from XMRV Δ10CANC. The white arrows point to the CA-NTD, the black arrowhead points to the CA-CTD, and the white arrowheads point to the NC region. The scale bar is 50 nm. (B) Comparison between the radial density profiles of XMRV Δ10CANC (red) and HIV-1 (blue). The two profiles have been aligned on the CA-CTD, and the *x* axis shows the radial distance from the CA-CTD region. The peaks corresponding to NC and to the NTD and CTD of CA are marked. (C) The two images show slices through the RAF plot at a 60° angle for XMRV Δ10CANC (top) and HIV-1 (bottom). The *x* axes are aligned to the *x* axis of the radial density plot in panel B. The expected positions of the [1,0] and [1,1] reflections for a hexameric lattice with a spacing of 7.4 nm at the CA-CTD are indicated by the white dashed lines. Strong peaks indicating a hexagonal order are seen for both XMRV Δ10CANC and HIV-1 in the CA-CTD and downstream regions. Only HIV shows strong peaks in the CA-NTD region. The radial positions of the RNP and CA domains are indicated on the plots.

A

B
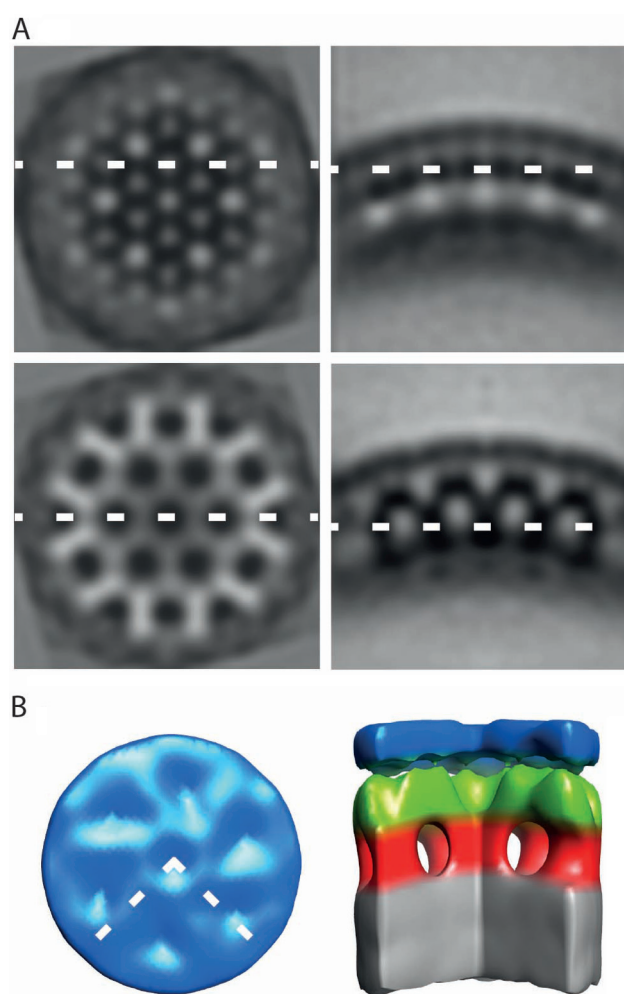
FIG 6 Structure of *in vitro*-assembled particles of XMRV Δ10CANC. (A) Orthoslices through the average structure. On the right are radial sections, while tangential sections are on the left. The dashed lines show the relative position of the orthoslices between the panels. Density is black. (B) Surface rendering of the average structure. The color scheme follows an approximate radial division of the structure according to the different domains of Gag as previously assigned (12). Gray is the NC-nucleic acid layer, red is the region between NC and the CA-CTD, green is the CA-CTD region, and blue is the CA-NTD. The dashed lines in the top view represent the section cut out in the side view. The CA-NTD region does not show an ordered hexameric lattice.

minants (β^+/SP$^+$) of CA results in mature-like tubular (and, in the case of HIV-1, also conical) particle assembly (8, 18, 22). The HIV-1 ΔMACAΔSP1NC and RSV ΔPRGag proteins, which do not form the β-hairpin and in which SP is mutated or deleted (β^-/SP$^-$), assemble into tubular (and, in the case of HIV-1, also conical) particles (23, 28). The mature CA proteins (β^+/SP$^-$) of HIV and RSV assemble *in vitro* into tubular structures; however, the conditions under which these structures are formed (~1 M NaCl) suggest different types of CA-CA interactions (14, 22, 30).

M-PMV, which lacks the SP sequence between CA and NC, appears to behave differently *in vitro*. In the absence of the β-hairpin and in the presence of downstream NC, M-PMV forms spherical particles. In the presence of the β-hairpin, M-PMV cannot be assembled *in vitro*. Here, we show that XMRV behaves in the same way: it is unable to form mature-like structures *in vitro*

from CA or CANC but can form immature-like ΔProCANC and Δ10CANC particles. In contrast to HIV-1 and RSV, XMRV and M-PMV do not have SP separating the CA and NC domains. This begs the question whether some part of XMRV CA or NC might substitute for the SP-like domain. Based on recent data showing that MLV, in contrast to other retroviruses, does not have strict structural requirements for the N-terminal portion of NC (55), it is likely that the SP-like domain occurs within the CA CTD. Here, we show that the region between CA and NC forms a rod-like structure similar to that previously described for HIV and RSV and consistent with the presence of a structured region between the two folded domains. A possible candidate is the "charged assembly helix" at the C terminus of MLV CA (9). This region was suggested to form an amphipathic α-helix, and mutations within the region showed its necessity for MLV assembly (9). The SP-like domain in M-PMV was localized at the CANC junction (3); however, its precise length is currently being studied. It is possible that such sequences at the C terminus of CA in XMRV and M-PMV are what prevents the *in vitro* assembly of mature virus-like assemblies.

Given the high degree of sequence similarity to MLV Gag, the arrangement in the immature-like *in vitro*-assembled XMRV particles should be comparable to that in immature MLV particles, as described by Yeager et al. (67). Indeed, we found that the two particles have similar appearances and adopt a hexameric lattice with the same spacing. The CA and NC domains formed noticeable layers in both CEM and TEM images of the *in vitro*-assembled XMRV Δ10CANC, which correspond well to the features in the immature MLV particles shown by Yeager et al. (67).

3D structural analysis of the particles revealed that they form an ordered hexameric lattice with structural features mostly similar to those previously described for HIV, M-PMV, and RSV particles. Unexpectedly, in contrast to the other viruses, we found that the outermost layer in the XMRV Δ10CANC particles, corresponding to the CA-NTD region, does not form an ordered hexameric lattice. Two formal explanations can be considered: the CA-NTD may be laterally disordered in the *in vitro*-assembled XMRV particles, or gammaretroviruses may adopt a structural arrangement different from that of other retroviruses, in which, for example, part of CA-NTD may fill the typical holes in the CA-NTD layer, giving a smooth appearance. Given the high degree of secondary-structure similarity between the CA-NTD of MLV and those of other retroviruses, we consider the second possibility highly unlikely. It is sensible to conclude that the *in vitro*-assembled XMRV Δ10CANC particles faithfully mimic the structure of the immature XMRV Gag lattice in the CA-CTD and regions immediately downstream but show some lateral disorder in the CA-NTD region. We hypothesize that the difference in the order of CA-NTD (as shown in the RAF plots and in the 3D structure [Fig. 5C and 6B, respectively]) could explain the different appearances of the TEM images of negatively stained XMRV particles compared to HIV-1 and M-PMV (Fig. 3D). Although the cEM radial densities of XMRV and HIV-1 are very similar (Fig. 5B), showing the same radial positions of the CA-NTD, CA-CTD, and RNP domains, the cryo-electron tomography clearly shows that the degree of order in the XMRV CA-NTD (as shown in the RAF plots and in the 3D structure [Fig. 5C and 6B, respectively]) is very different.

Many previous studies indicated the importance of *in vitro* assembly studies of different retroviruses to obtain valuable infor-

mation on the principles of retroviral assembly. A system for the *in vitro* assembly of MLV was described for CA protein (24) recruited to lipid nanotubes as an assembly template. Another *in vitro* assembly was reported recently for MLV Gag (11). However, in this system, the MLV Gag VLPs were fragile and of "irregular appearance." The protocol presented here for the efficient *in vitro* assembly of virus-like Gag-derived particles in solution provides the basis for future studies of detailed protein packing and arrangement in gammaretroviral particles. Importantly, although the region downstream of XMRV CA does not contain a spacer peptide and has an unusual run of charged residues (9), we found that it assembles into a typical immature retroviral lattice with rod-like densities linking CA to the RNP. Further investigation of this system could therefore help us to better understand the critical role of the region in retrovirus assembly.

ACKNOWLEDGMENTS

REFERENCES

1. **Accola MA, Hoglund S, Gottlinger HG.** 1998. A putative alpha-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**:2072–2078.
2. **Accola MA, Strack B, Gottlinger HG.** 2000. Efficient particle production by minimal gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. J. Virol. **74**:5395–5402.
3. **Bohmova K, et al.** 2010. Effect of dimerizing domains and basic residues on in vitro and in vivo assembly of Mason-Pfizer monkey virus and human immunodeficiency virus. J. Virol. **84**:1977–1988.
4. **Borsetti A, Ohagen A, Gottlinger HG.** 1998. The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. J. Virol. **72**:9313–9317.
5. **Briggs JA, Johnson MC, Simon MN, Fuller SD, Vogt VM.** 2006. Cryoelectron microscopy reveals conserved and divergent features of gag packing in immature particles of Rous sarcoma virus and human immunodeficiency virus. J. Mol. Biol. **355**:157–168.
6. **Briggs JA, et al.** 2009. Structure and assembly of immature HIV. Proc. Natl. Acad. Sci. U. S. A. **106**:11090–11095.
7. **Briggs JA, et al.** 2004. The stoichiometry of Gag protein in HIV-1. Nat. Struct. Mol. Biol. **11**:672–675.
8. **Campbell S, Vogt VM.** 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.
9. **Cheslock SR, et al.** 2003. Charged assembly helix motif in murine leukemia virus capsid: an important region for virus assembly and particle size determination. J. Virol. **77**:7058–7066.
10. **Cornilescu CC, Bouamr F, Yao X, Carter C, Tjandra N.** 2001. Structural analysis of the N-terminal domain of the human T-cell leukemia virus capsid protein. J. Mol. Biol. **306**:783–797.
11. **Datta SA, et al.** 2011. Solution properties of murine leukemia virus Gag protein: differences from HIV-1 Gag. J. Virol. **85**:12733–12741.
12. **de Marco A, et al.** 2010. Conserved and variable features of Gag structure and arrangement in immature retrovirus particles. J. Virol. **84**:11729–11736.
13. **de Marco A, et al.** 2010. Structural analysis of HIV-1 maturation using cryo-electron tomography. PLoS Pathog. **6**:e1001215.
14. **Ehrlich LS, Agresta BE, Carter CA.** 1992. Assembly of recombinant human immunodeficiency virus type 1 capsid protein in vitro. J. Virol. **66**:4874–4883.

15. **Erlwein O, et al.** 2011. Investigation into the presence of and serological response to XMRV in CFS patients. PLoS One **6**:e17592.
16. **Fischer N, et al.** 2008. Prevalence of human gammaretrovirus XMRV in sporadic prostate cancer. J. Clin. Virol. **43**:277–283.
17. **Furuta RA, et al.** 2011. No association of xenotropic murine leukemia virus-related virus with prostate cancer or chronic fatigue syndrome in Japan. Retrovirology **8**:20.
18. **Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI.** 1999. Assembly and analysis of conical models for the HIV-1 core. Science **283**:80–83.
19. **Ganser-Pornillos BK, Yeager M, Sundquist WI.** 2008. The structural biology of HIV assembly. Curr. Opin. Struct. Biol. **18**:203–217.
20. **Gitti RK, et al.** 1996. Structure of the amino-terminal core domain of the HIV-1 capsid protein. Science **273**:231–235.
21. **Gross I, Hohenberg H, Huckhagel C, Krausslich HG.** 1998. N-terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. J. Virol. **72**:4798–4810.
22. **Gross I, Hohenberg H, Krausslich HG.** 1997. In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. Eur. J. Biochem. **249**:592–600.
23. **Gross I, et al.** 2000. A conformational switch controlling HIV-1 morphogenesis. EMBO J. **19**:103–113.
24. **Hilditch L, et al.** 2011. Ordered assembly of murine leukemia virus capsid protein on lipid nanotubes directs specific binding by the restriction factor, Fv1. Proc. Natl. Acad. Sci. U. S. A. **108**:5771–5776.
25. **Hohn O, et al.** 2009. Lack of evidence for xenotropic murine leukemia virus-related virus(XMRV) in German prostate cancer patients. Retrovirology **6**:92.
26. **Hong P, Li J, Li Y.** 2010. Failure to detect xenotropic murine leukaemia virus-related virus in Chinese patients with chronic fatigue syndrome. Virol. J. **7**:224.
27. **Johnson MC, Scobie HM, Vogt VM.** 2001. PR domain of rous sarcoma virus Gag causes an assembly/budding defect in insect cells. J. Virol. **75**:4407–4412.
28. **Keller PW, Johnson MC, Vogt VM.** 2008. Mutations in the spacer peptide and adjoining sequences in Rous sarcoma virus Gag lead to tubular budding. J. Virol. **82**:6788–6797.
29. **Khorasanizadeh S, Campos-Olivas R, Summers MF.** 1999. Solution structure of the capsid protein from the human T-cell leukemia virus type-I. J. Mol. Biol. **291**:491–505.
30. **Kingston RL, et al.** 2000. Structure and self-association of the Rous sarcoma virus capsid protein. Structure **8**:617–628.
31. **Klikova M, Rhee SS, Hunter E, Ruml T.** 1995. Efficient in-vivo and in-vitro assembly of retroviral capsids from Gag precursor proteins expressed in bacteria. J. Virol. **69**:1093–1098.
32. **Knox K, et al.** 2011. No evidence of murine-like gammaretroviruses in CFS patients previously identified as XMRV-infected. Science **333**:94–97.
33. **Krausslich HG, Facke M, Heuser AM, Konvalinka J, Zentgraf H.** 1995. The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. J. Virol. **69**:3407–3419.
34. **Kremer JR, Mastronarde DN, McIntosh JR.** 1996. Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. **116**:71–76.
35. **Kuznetsov YG, Ulbrich P, Haubova S, Ruml T, McPherson A.** 2007. Atomic force microscopy investigation of Mason-Pfizer monkey virus and human immunodeficiency virus type 1 reassembled particles. Virology **360**:434–446.
36. **Li S, Hill CP, Sundquist WI, Finch JT.** 2000. Image reconstructions of helical assemblies of the HIV-1 CA protein. Nature **407**:409–413.
37. **Liang C, Hu J, Whitney JB, Kleiman L, Wainberg MA.** 2003. A structurally disordered region at the C terminus of capsid plays essential roles in multimerization and membrane binding of the Gag protein of human immunodeficiency virus type 1. J. Virol. **77**:1772–1783.
38. **Lombardi VC, et al.** 2009. Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. Science **326**:585–589.
39. **Ludtke SJ, Baldwin PR, Chiu W.** 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. **128**:82–97.
40. **Ma YM, Vogt VM.** 2002. Rous sarcoma virus Gag protein-oligonucleotide interaction suggests a critical role for protein dimer formation in assembly. J. Virol. **76**:5452–5462.
41. **Ma YM, Vogt VM.** 2004. Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles in vitro. J. Virol. **78**:52–60.
42. **Macek P, et al.** 2009. NMR structure of the N-terminal domain of capsid protein from the Mason-Pfizer monkey virus. J. Mol. Biol. **392**:100–114.
43. **Mortuza GB, et al.** 2008. Structure of B-MLV capsid amino-terminal domain reveals key features of viral tropism, gag assembly and core formation. J. Mol. Biol. **376**:1493–1508.
44. **Mortuza GB, et al.** 2004. High-resolution structure of a retroviral capsid hexameric amino-terminal domain. Nature **431**:481–485.
45. **Paprotka T, et al.** 2011. Recombinant origin of the retrovirus XMRV. Science **333**:97–101.
46. **Phillips JM, Murray PS, Murray D, Vogt VM.** 2008. A molecular switch required for retrovirus assembly participates in the hexagonal immature lattice. EMBO J. **27**:1411–1420.
47. **Pornillos O, Ganser-Pornillos BK, Yeager M.** 2011. Atomic-level modelling of the HIV capsid. Nature **469**:424–427.
48. **Raisch KP, et al.** 2003. Molecular cloning, complete sequence, and biological characterization of a xenotropic murine leukemia virus constitutively released from the human B-lymphoblastoid cell line DG-75. Virology **308**:83–91.
49. **Robinson MJ, Erlwein O, McClure MO.** 2011. Xenotropic murine leukaemia virus-related virus (XMRV) does not cause chronic fatigue. Trends Microbiol. **19**:525–529.
50. **Rumlova-Klikova M, Hunter E, Nermut MV, Pichova I, Ruml T.** 2000. Analysis of Mason-Pfizer monkey virus Gag domains required for capsid assembly in bacteria: role of the N-terminal proline residue of CA in directing particle shape. J. Virol. **74**:8452–8459.
51. **Satterfield BC, et al.** 2011. Serologic and PCR testing of persons with chronic fatigue syndrome in the United States shows no association with xenotropic or polytropic murine leukemia virus-related viruses. Retrovirology **8**:12.
52. **Schutzer SE, Rounds MA, Natelson BH, Ecker DJ, Eshoo MW.** 2011. Analysis of cerebrospinal fluid from chronic fatigue syndrome patients for multiple human ubiquitous viruses and xenotropic murine leukemia-related virus. Ann. Neurol. **69**:735–738.
53. **Sfanos KS, et al.** 2008. A molecular analysis of prokaryotic and viral DNA sequences in prostate tissue from patients with prostate cancer indicates the presence of multiple and diverse microorganisms. Prostate **68**:306–320.
54. **Shin CH, et al.** 2011. Absence of XMRV retrovirus and other murine leukemia virus-related viruses in patients with chronic fatigue syndrome. J. Virol. **85**:7195–7202.
55. **Still A, Huseby D, Barklis E.** 2011. Analysis of the N-terminal region of the murine leukemia virus nucleocapsid protein. Virus Res. **155**:181–188.
56. **Switzer WM, et al.** 2010. Absence of evidence of xenotropic murine leukemia virus-related virus infection in persons with chronic fatigue syndrome and healthy controls in the United States. Retrovirology **7**:57.
57. **Ulbrich P, et al.** 2006. Distinct roles for nucleic acid in in vitro assembly of purified Mason-Pfizer monkey virus CANC proteins. J. Virol. **80**:7089–7099.
58. **Urisman A, et al.** 2006. Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. PLoS Pathog. **2**:e25.
59. **van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M.** 1996. A new generation of the IMAGIC image processing system. J. Struct. Biol. **116**:17–24.
60. **Verhaegh GW, et al.** 2011. Prevalence of human xenotropic murine leukemia virus-related gammaretrovirus (XMRV) in Dutch prostate cancer patients. Prostate **71**:415–420.
61. **von Schwedler UK, et al.** 1998. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J. **17**:1555–1568.
62. **von Schwedler UK, Stray KM, Garrus JE, Sundquist WI.** 2003. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J. Virol. **77**:5439–5450.
63. **Wang MQ, Goff SP.** 2003. Defects in virion production caused by mutations affecting the C-terminal portion of the Moloney murine leukemia virus capsid protein. J. Virol. **77**:3339–3344.
64. **Wildova M, et al.** 2008. The effect of point mutations within the N-terminal domain of Mason-Pfizer monkey virus capsid protein on virus core assembly and infectivity. Virology **380**:157–163.

65. **Wilk T, et al.** 2001. Organization of immature human immunodeficiency virus type 1. J. Virol. **75**:759–771.
66. **Wright ER, et al.** 2007. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J. **26**:2218–2226.
67. **Yeager M, Wilson-Kubalek EM, Weiner SG, Brown PO, 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. U. S. A. **95**:7299–7304.
68. **Yu F, et al.** 2001. Characterization of Rous sarcoma virus Gag particles assembled in vitro. J. Virol. **75**:2753–2764.

Cryo Electron Tomography of Native HIV-1 Budding Sites

Lars-Anders Carlson[1,2], Alex de Marco[3], Heike Oberwinkler[1], Anja Habermann[1], John A. G. Briggs[3], Hans-Georg Kräusslich[1]*, Kay Grünewald[2,4]*

1 Department of Infectious Diseases, Virology, Universitätsklinikum Heidelberg, Heidelberg, Germany, 2 Department of Molecular Structural Biology, Max-Planck-Institute of Biochemistry, Martinsried, Germany, 3 Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany, 4 Oxford Particle Imaging Centre, Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

Abstract

The structure of immature and mature HIV-1 particles has been analyzed in detail by cryo electron microscopy, while no such studies have been reported for cellular HIV-1 budding sites. Here, we established a system for studying HIV-1 virus-like particle assembly and release by cryo electron tomography of intact human cells. The lattice of the structural Gag protein in budding sites was indistinguishable from that of the released immature virion, suggesting that its organization is determined at the assembly site without major subsequent rearrangements. Besides the immature lattice, a previously not described Gag lattice was detected in some budding sites and released particles; this lattice was found at high frequencies in a subset of infected T-cells. It displays the same hexagonal symmetry and spacing in the MA-CA layer as the immature lattice, but lacks density corresponding to NC-RNA-p6. Buds and released particles carrying this lattice consistently lacked the viral ribonucleoprotein complex, suggesting that they correspond to aberrant products due to premature proteolytic activation. We hypothesize that cellular and/or viral factors normally control the onset of proteolytic maturation during assembly and release, and that this control has been lost in a subset of infected T-cells leading to formation of aberrant particles.

Introduction

HIV-1 particles are assembled at the cell membrane, as the 55 kDa viral polyprotein Gag multimerizes on its inner face [1]. Gag recruits other viral components such as the RNA genome and the surface spike proteins, as well as cellular proteins of the ESCRT machinery required for virus release [2,3,4]. The viral protease (PR) is essential to convert the immature form of the virion into an infectious mature particle. Both forms of the virion are pleiomorphic structures, with the repetitive structural elements of the virus arranged non-symmetrically and variably from one particle to the other.

In the immature virion, uncleaved Gag is anchored to the plasma membrane via a charged surface and a myristoyl tail in its N-terminal matrix (MA) domain [1]. As shown by cryo electron microscopy (cEM), Gag arranges in a regular manner, with its internal capsid (CA) domain forming a hexameric lattice with a spacing of 8.0 nm [5]. C-terminally of CA, the nucleocapsid (NC) domain binds the RNA genome, and the p6 domain recruits the ESCRT machinery to facilitate particle release [6,7]. CA and NC, as well as NC and p6, are separated by short spacer peptides (SP1 and SP2, respectively) which are processed during maturation.

Proteolytic maturation of HIV-1 has been proposed to initiate at or shortly after assembly and release [8]. The active dimeric form of the viral PR cleaves Gag and GagPol at multiple sites, leading to the structural transition from the immature particle with its Gag shell forming a truncated sphere to the mature particle with its

cone-shaped CA core encasing the condensed nucleoprotein complex in the interior of the virion. In an *in vitro* study Pettit *et al.* detected large differences in processing kinetics at the five cleavage sites in Gag, the cleavage at the site between SP1 and NC being an order of magnitude faster than the second fastest cleavage event [9]. These results suggested ordered processing during virion maturation, and this was supported by mutagenesis studies of individual cleavage sites [10]. Processing at all sites except NC-SP1 [11] appears to be important for infectivity and small amounts of partially processed Gag products have been shown to exhibit a strong *trans*-dominant negative effect on viral infectivity [12,13,14]. Furthermore, premature processing [15] as well as low concentrations of protease inhibitors, insufficient to significantly affect proteolytic maturation [14,16,17], were both shown to efficiently block viral infectivity, indicating an intricate interplay between proteolytic activation and virus formation. Morphological maturation is believed to occur rapidly following release of the immature virion as no intermediates of maturation have been observed so far and all cellular budding sites appear to carry an immature Gag shell.

The last years have seen increasingly detailed structural studies of released HIV-1 particles in their immature and mature forms [18,19,20,21,22,23,24,25]. Analysis of immature virions by cryo electron tomography (cET) revealed that the immature lattice covers only part of the viral membrane [22], arranging as an incomplete "~2/3" sphere [26]. These studies suggested that ESCRT involvement in HIV-1 release occurs earlier than

Author Summary

The production of new HIV-1 particles is initiated at the plasma membrane where the viral polyprotein Gag assembles into a budding site, and proceeds through release of an immature virion which is subsequently transformed to the infectious virion by proteolytic cleavage of Gag. Here, we established experimental systems to study HIV-1 budding sites by cryo electron tomography. This technique allows three-dimensional structure determination of single objects at macromolecular resolution, thus being uniquely suited to study variable structures such as HIV-1 particles and budding sites. Using cryo electron tomography, we obtained three-dimensional images with unprecedented detail of the formation of HIV-1 particles. By analyzing these images we show that the organization of released immature HIV-1 is determined at its intracellular assembly without major subsequent rearrangements. We further identify a lattice structure of the viral protein Gag present in budding sites that seem to lack the viral genome and thus cannot be precursors of infectious viruses. We show that some HIV-1 infected T-cells preferentially carry these budding sites, suggesting that they have lost a crucial control of the proteolytic maturation of the virus.

previously thought and that so-called late budding sites carrying an almost complete Gag lattice are likely to be dead-end products rather than intermediates in the release pathway [26]. A later analysis by cET and subtomogram averaging provided a 17 Å structure of the immature Gag lattice and showed how this incomplete hexagonal lattice attains its curvature through inclusion of symmetry defects of irregular shape and size [19]. In the mature capsid, CA also forms a hexameric lattice, but with a different lattice constant and arrangement than in the immature lattice [27]. Curvature of the mature capsid is achieved by asymmetric incorporation of pentameric defects at the narrow and wide end of the cone [28]. While cET studies of released virions have significantly advanced our understanding of HIV-1 morphogenesis, we are currently lacking any three-dimensional information regarding native intracellular particle assembly and budding. These processes have hitherto been inaccessible for detailed structural studies due to the lack of an experimental system enabling their visualization by cET. Here, we present a structural study of HIV-1 assembly sites by cET of intact, plunge-frozen human cells. These snapshots allow detailed structural interpretations of the Gag lattice structure at its assembly site *in situ*, as well as the arrangement of the cortical actin in its immediate vicinity. We further determined the structure of a previously not reported Gag lattice type lacking the NC-RNA layer which we propose to be associated with loss of control of PR activation.

Results

Visualization of HIV-1 assembly by cryo electron tomography

To enable structural analysis of HIV-1 assembly in its cellular context, we set out to establish a system for studying this process by cET on intact human cells. Since cET is limited to samples thinner than ~500 nm, and thus to thin peripheral areas of adherent cells, the natural HIV-1 host cell types (T-cells and macrophages) cannot be used for cET. This led us to use two human glioblastoma cell lines, U-87 MG and U-373 MG. These cells grow adherently, have extended, thin peripheral areas, and are amenable to transfection or transduction with adenoviral vectors. Transfection efficiency proved insufficient for structural analyses, however, and we therefore constructed adenoviral vectors expressing HIV-1 Gag (AdGag) and HIV-1 Gag-Pol (AdGagPol), respectively. Both adenoviral vectors were based on Rev-independent versions of the respective genes [29], allowing expression of either Gag or Gag-Pol without co-transduction of Rev. U-87 MG or U-373 MG cells were seeded on EM sample grids concomitant with transduction using the respective adenoviral vector. Immunofluorescence staining for HIV-1 CA at 24–48 h post transduction revealed that almost 100% of the cells expressed HIV-1 Gag, while retaining normal adherent morphology (Fig. 1A–B). Furthermore, when U-373 MG cells were transduced with AdGagPol, specific cleavage products of HIV-1 Gag were found in the cell lysate and in the particle fraction purified from the culture medium (Fig. 1C), while particle release but no processing was observed for AdGag (data not shown). Cells transduced with AdGag or AdGagPol were vitrified by plunge-freezing and subjected to cEM. Upon inspection of the vitrified grids, HIV-1 assembly sites and released virions were present at low but reproducible frequencies at the periphery of the cells (Fig. 1D) in regions accessible to cET. A total of 40 tilt series were recorded on such positions of AdGag and AdGagPol infected cells, respectively (Table 1).

In cryo electron tomograms of AdGag or AdGagPol transduced cells (Fig. 2) structural features of the cytoplasm were preserved at the high level previously reported for cellular cET [30], with single ribosomes, actin filaments and microtubules being clearly resolved. Both viral budding sites (Fig. 2A, C–F) and released virions (Fig. 2B–C, E) were present in the recorded tomograms. The budding sites were found at the plasma membrane on lamellipodia–like areas (Fig. 2C, E), occasionally in membrane invaginations (Fig. 2B, D) or at the tips (Fig. 2F, 3B) or sides (Fig. 2A, 3A) of filopodial structures. In the immature Gag lattice of these budding sites and released particles, the two density layers corresponding to CA and NC-RNA could be resolved, and occasionally the hexagonal symmetry in the CA layer was apparent by eye in more tangential tomographic slices. Released immature particles (ip) were found adjacent to cells expressing Gag (Fig. 2B, D) or GagPol (Fig. 2E), whereas released particles with mature morphology (mp) were only found adjacent to cells expressing GagPol (Fig. 2E). No density attributable to components of the ESCRT complex was detected at budding sites at the current resolution.

The preservation of cytoplasmic structure in cET allowed for three-dimensional snapshots of filamentous actin associated with HIV-1 budding sites. The presence of actin filaments was generally high at budding sites (Fig. 2A, C, F and Fig.3A–D), as judged by visual inspection where F-actin is clearly resolvable from *e.g.* intermediate filaments. Furthermore, we often observed what appeared to be a direct interaction of actin filaments with the budding site (Fig. 2A, F; Fig. 3D). To categorize the budding sites with respect to their actin context, they were sorted into 5 classes according to the type of actin structures they were associated with (Fig.3, top panel). This classification revealed that 34 of the 39 budding sites analyzed were found adjacent to filamentous actin (Fig. 3A–D), with half of the buds (20 of 39) appearing on the sides or tips of filopodia-like structures characterized by a parallel actin organization (Fig. 3A–B).

In situ organization of the assembling Gag lattice

Cryo electron tomograms containing budding sites on intact cells with apparently high signal-to-noise ratio were selected for
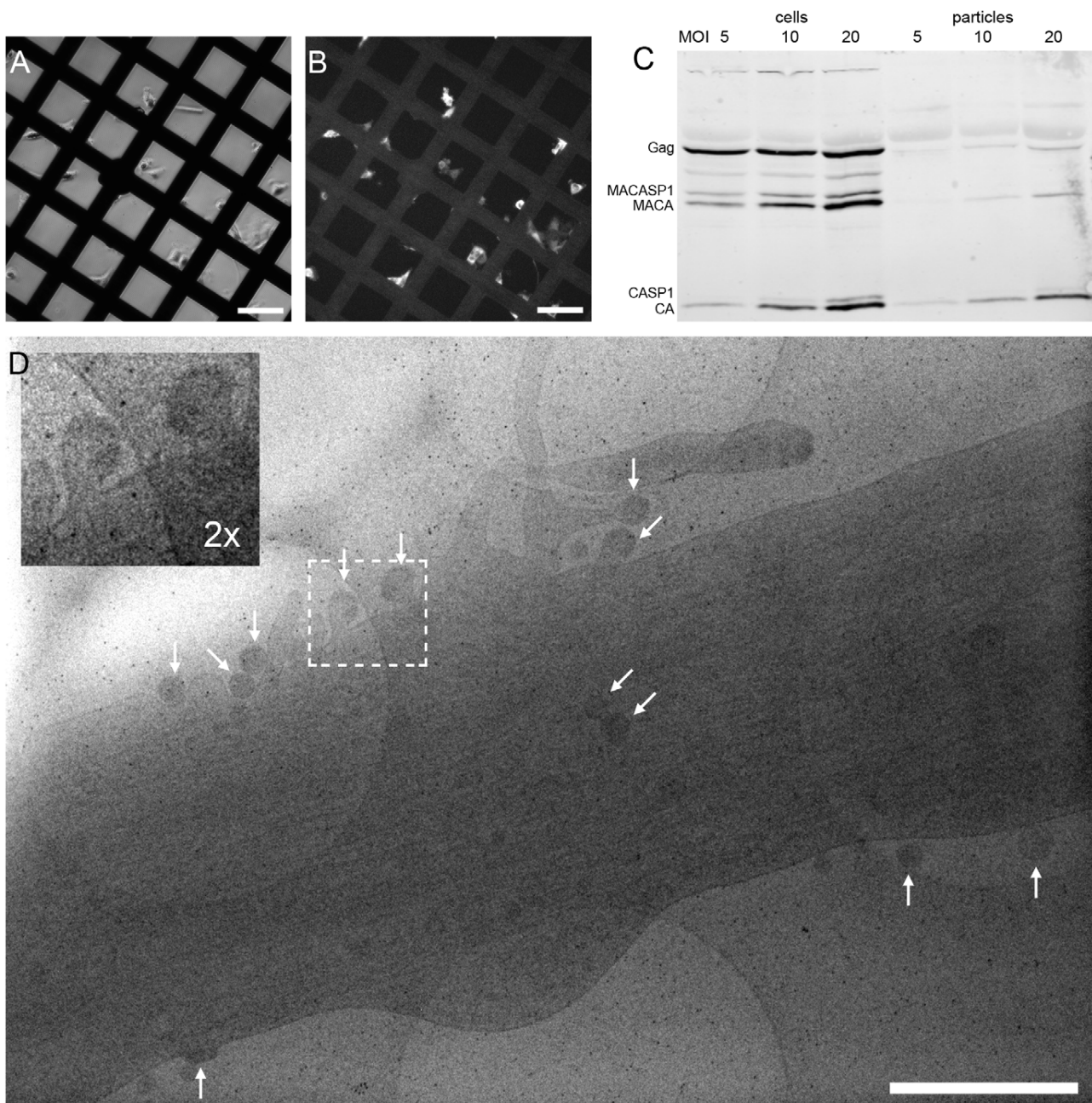
Figure 1. HIV-1 assembly sites imaged by cryo electron microscopy of intact human cells. (A) Light microscopy phase contrast and (B) immunofluorescence image of human glioblastoma cells (U-373 MG) grown adherently on carbon coated EM grids. Cells were transduced with AdGag concomitant with seeding on grids, and subjected to immunofluorescence staining for Gag two days post seeding to estimate the relative number of Gag expressing cells. (C) U373-MG cells were transduced with AdGagPol at various MOI as indicated above the panel. Cell and particle lysates were recovered 2 days after transduction and subjected to immunoblot analysis using polyclonal antiserum raised against CA. The positions of the Gag and GagPol polyproteins and of several processing products are indicated on the left. (D) Cryo electron micrograph (projection image) of a U-373 MG cell plunge-frozen two days after transduction with AdGagPol. The image shows a cell protrusion containing viral assembly sites (marked by arrows) accessible for cET (The large light circles are 2 μm diameter holes in the carbon support films that the cells grow on, the small dark spots are 10 nm colloidal gold beads). The inset in the upper left corner is a magnified view of the boxed area in the large image. Scale bars are 100 μm (A–B) and 1 μm (D).

doi:10.1371/journal.ppat.1001173.g001

further analysis by sub-tomogram averaging as previously performed for isolated immature HIV-1 [19]. This analysis provides two kinds of information about the Gag protein lattice: the global arrangement of the repetitive elements into an ordered lattice structure ("lattice map") and the local structure of these repetitive elements ("unit cell structure").

The immature Gag lattice parameters in the budding sites were the same as previously determined for released immature virions

Table 1. Data statistics for cET of HIV-1 budding.

Data set number	Sample	Number of tilt series	Number of budding sites	Number of released particles
1	U-87 MG transduced with AdGag	6	4	5
2	U-373 MG transduced with AdGag	9	9	26
3	U-373 MG transduced with AdGagPol	25	26	39

For each data set, the number of recorded tilt series is stated along with the number of budding sites and released particles contained in the data set.
doi:10.1371/journal.ppat.1001173.t001

[5,19]: a hexagonal lattice with a lattice constant of 8.0 nm. The lattice maps derived from the budding sites (Fig. 4A) revealed several sites of symmetry breakage, similarly heterogeneous in size and shape as those described for the released virions [19].

The unit cell structure derived from the cellular tomograms contained density corresponding to the N-terminal and C-terminal domains of CA, respectively, as well as density for the membrane and the NC-RNA layer (Fig. 5A). At the present resolution (roughly 40 Å by the 0.5 Fourier shell correlation criterion), this structure was indistinguishable from that previously reported for released particles [19].

A previously undescribed Gag lattice form is present in tomograms of GagPol-expressing cells

Immunoblot analysis of AdGagPol-transduced cells and the particle fraction recovered from the culture medium showed proteolytic processing of Gag (Fig. 1C). Furthermore, particles containing mature-looking cores were found in cryo electron tomograms in immediate proximity of cellular budding sites and next to immature particles (Fig. 2E). Of note, structures resembling the mature capsid were never found in budding sites still connected to the plasma membrane. Besides the well-described mature and immature particles, we observed a third, previously not described form of particles in the tomograms of AdGagPol-transduced cells (Fig. 6A). This apparently "intermediate morphology" exhibiting a thinner Gag lattice than the immature structure was observed both in extracellular particles (Fig. 6A, Fig. S2 in Supporting Information S1) and in HIV-1 budding sites connected to the cell (Fig. 6A–B). Budding sites and extracellular particles with the previously unrecognized Gag lattice were characterized by a single density layer bound to the inner face of the membrane (Figure 6B), as opposed to the two layers (CA and NC-RNA) seen in the immature lattice (Fig. 6C). Density plots orthogonal to the cell membrane confirmed that the density layer found in these novel structures was located at the same position relative to the plasma membrane (peak at −12 nm) as the CA layer of the complete immature lattice (Fig. S1 in Supporting Information S1).

Strikingly, budding sites and released particles with this novel Gag lattice displayed the same hexagonal symmetry and lattice spacing as the immature lattice (Fig. 4B). The apparently intermediate type lattices were typically more complete with smaller lattice defects than the immature lattices, but the lattice defects were similarly heterogeneous. Furthermore, the unit cell structure of the new lattice revealed an arrangement identical to the CA region of the immature lattice, with density corresponding to the N- and C-terminal domains of CA (Fig. 5B), but lacking any density corresponding to the NC-RNA layer. The cellular budding sites and extracellular particles containing this lattice also lacked the internal density corresponding to a condensed NC-RNA complex. Such densities were readily observed within mature capsids and in virions carrying partially processed Gag proteins (de Marco *et al.*, accompanying paper).

Budding sites with the novel Gag lattice are also observed in infected T-cell lines

To determine whether the newly described lattice also occurs in other HIV-1 producing cells including T-cell lines, we recorded electron tomograms on sections of resin-embedded cells producing HIV-1. Although resin-embedded material is not preserved at the molecular level, it might still be possible to distinguish the two lattice types in these data, and this would allow recording sufficiently large data sets for statistical analyses. Visual inspection of resin-embedded sections from infected MT4 cells suggested the presence of a thinner density layer in some budding sites (Fig. 7A, B). This interpretation was validated based on the following criteria: (i) different types of membrane-bound densities in budding sites in resin embeddings could be classified by an unbiased scheme, and (ii) the presence of structures corresponding to the novel Gag lattice was dependent on an active viral PR.

Classification of budding sites and particles was based on linear density plots of the membrane-bound Gag density. Fig. 7C shows a typical example, demonstrating that enough structure is retained to resolve the plasma membrane and distinct regions in the Gag layer. The classification was based on the parts of the plots containing the density bound to the inner face of the membrane (shaded area in Fig. 7C). Data were recorded on resin embeddings of HIV-1 infected MT4 cells, HeLa cells transfected with a proviral HIV-1 plasmid as well as positive and negative controls for the presence of the two lattice types (Table 2). In order to avoid overrepresentation of single cells, two tilt series were recorded per cell, and they were recorded as far apart as possible in the section plane (usually >5 μm). The same cell was never studied in consecutive sections. Our published data on infected MT4 cells [26] were also included in the analysis (Table 2, data set 8).

Principal component analysis of the variance in the data set revealed the presence of one major cluster with residual variance and a smaller, more homogeneous one (Fig. S3A–D in Supporting Information S1). Hierarchical clustering of the cropped line plots identified the two main classes as the two clusters observed in the PCA representation (Fig. S3E,F in Supporting Information S1). The less populated class (class 1, dashed line in Fig. 7D) lacked the innermost density peak corresponding to NC-RNA, whereas this peak was present in the more populated class (class 2, solid line in Fig. 7D). Further division of class 2 into subclasses mainly revealed a slight variation in the position and relative intensity of the NC-RNA peak (Fig. S4 in Supporting Information S1).

Table 3 shows the distribution of budding sites into the two classes for each sample. Class 1 contained 18% of the budding sites on HIV-1 infected MT4 cells and 2.9% of the budding sites on transfected HeLa cells. For infected MT4 cells treated with the PR inhibitor lopinavir, no class 1 budding sites were observed. As a control for the thinner Gag lattice, released virions carrying uncleaved MA-SP1 were studied. In this case, Gag processing is stalled at the MACASP1 fragment due to inactivating mutations at cleavage sites between MA and CA and CA and SP1, respectively
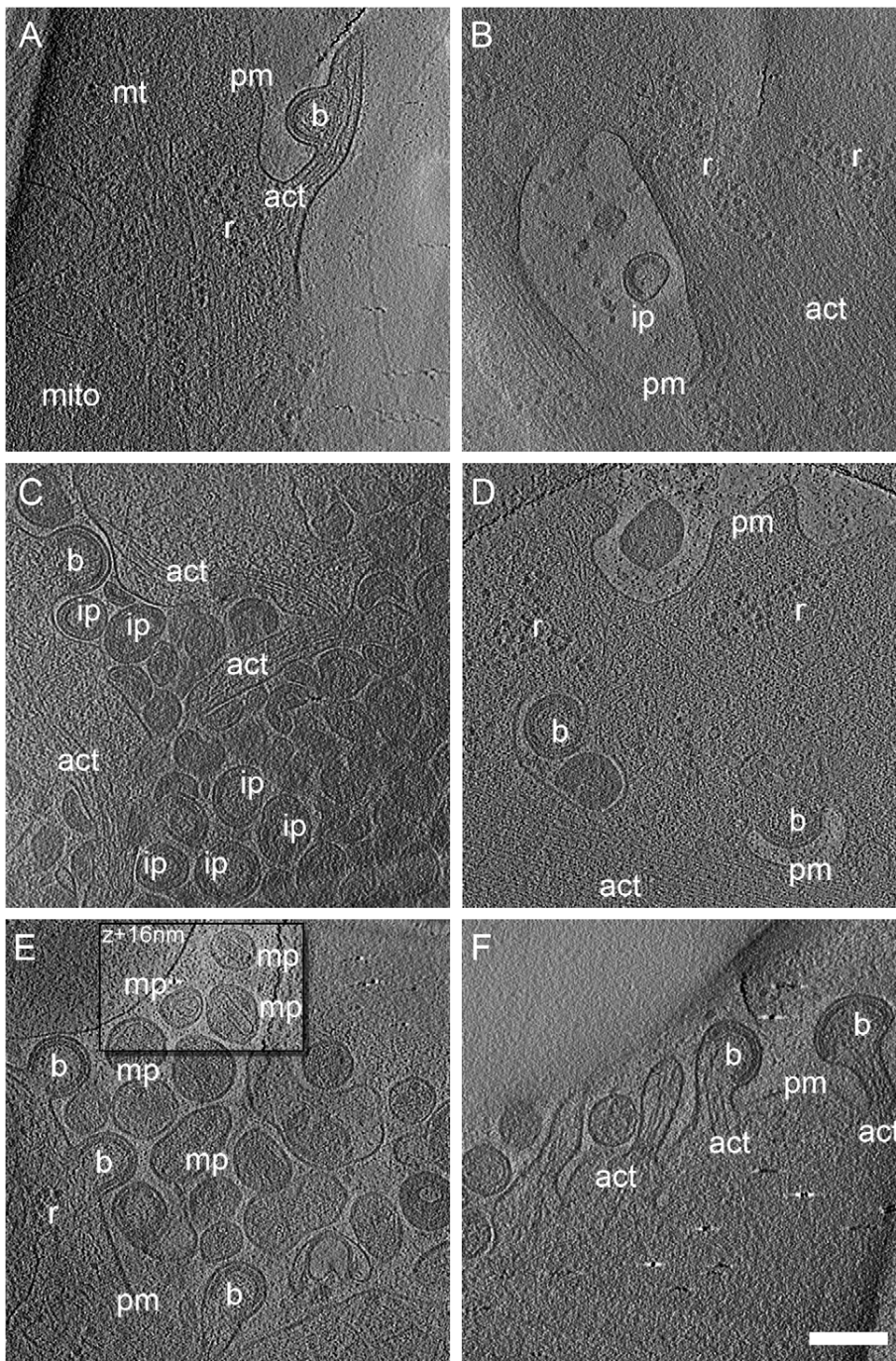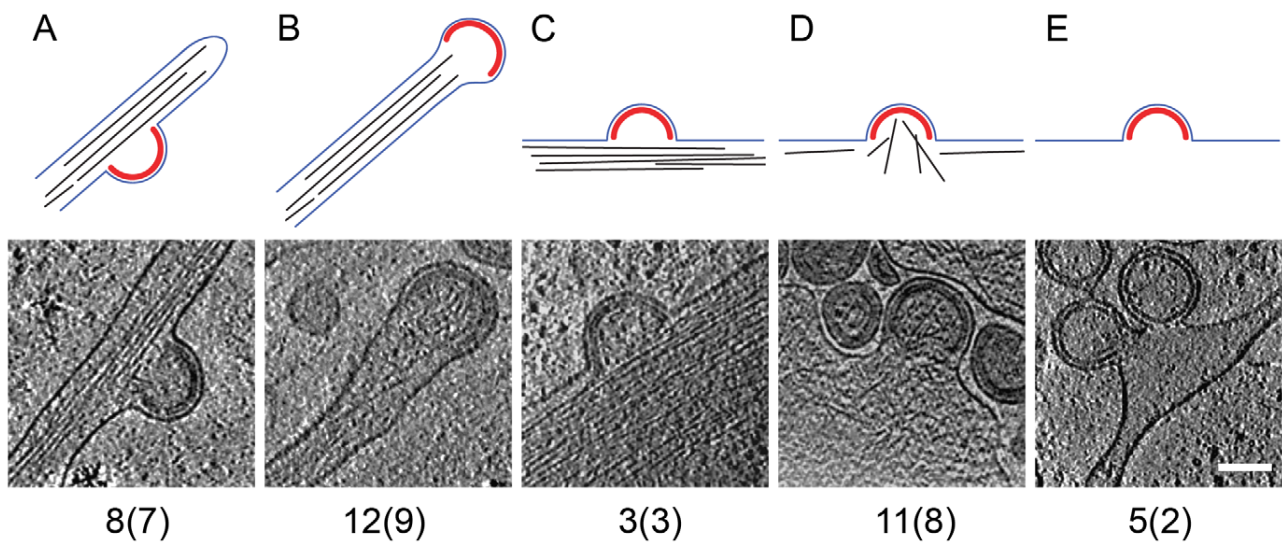
Figure 2. Cryo electron tomograms of HIV-1 budding sites and released particles. Computational slices of 1.6 nm thickness through cryo electron tomograms of U-87 MG (A–B) and U-373 MG (C–F) cells transduced with AdGag (A–C) or AdGagPol (D–F). Density is black. *Bona fide* designations of structures: act, actin; b, budding sites; ip, immature particles; mp, mature particles; mt, microtubule; r, ribosome; pm, plasma membrane. The inset in (E) is shifted by 16 nm perpendicular to the image plane, at which location the morphology of the mature particles appears clearer. Scale bar is 200 nm.
doi:10.1371/journal.ppat.1001173.g002

(see [10] and accompanying paper by de Marco *et al.*). These particles generally fell into class 1. Taken together, these results suggest that class 1 corresponds to the newly described lattice detected by cET, while class 2 corresponds to the immature lattice.

Radius and closure of the Gag layer was calculated separately for the two classes (Table 4, Fig. 7F). Whereas the average radii of the two classes were similar, the closure of the Gag layer was significantly higher for the class 1 buds (Wilcoxon rank sum test, $p = 0.0021$). Assuming the same lattice packing in both classes, class 1 budding sites in this data set contain $5,000 \pm 1,300$ Gag molecules while class 2 budding sites in this data set contain $3,300 \pm 1,500$ Gag molecules.

Figure 3. Filamentous actin at HIV-1 budding sites. The 39 budding sites reconstructed by cryo electron tomography were sorted into five categories, based on the type of filamentous actin structures (if any) present in their vicinity. Top panels are a sketch of the category, showing the Gag lattice in red, the plasma membrane in blue and filamentous actin in black. The lower panels show a computational slice through one budding site from the category, and the numbers below state the number of budding sites (and the number of tomograms they were obtained from) in the category. (A) Budding site at the side of actin-filled filopodium; (B) budding site at the tip of actin-filled filopodium; (C) budding site with cortical actin parallel to the plasma membrane; (D) budding site with cortical actin directed towards or protruding into the budding site; (E) budding site at the plasma membrane, not adjacent to filamentous actin. Scale bar is 100 nm.
doi:10.1371/journal.ppat.1001173.g003

Finally, we asked whether individual cells carried only one class of budding sites or both. In the case of HIV-1 infected MT4 cells, 47 cells exhibited more than one budding site in the tomogram (Table 2, data set 4). Fig. 7E (black bars) shows the relative frequency of class 1 budding sites on these 47 cells, overlaid with the theoretical distribution (grey circles) that would be expected if all cells had the average of 18% class 1 budding sites. Clearly, most cells had either only class 1 or only class 2 budding sites, indicating that the budding phenotype is determined at the single cell level.

Discussion

Here, we established a system to visualize HIV-1 virus-like particle assembly and release *in situ* at macromolecular resolution in three dimensions using cET. This method permits structural interpretation of assembly sites and their cytoplasmic and membrane surroundings without the limitations imposed by the fixation, staining and dehydration methods commonly used in cellular electron microscopy. It allowed a detailed analysis of the Gag lattice in cellular budding sites, suggesting that the organization of the immature HIV-1 particle is already determined at the point of its intracellular assembly, and further identified a hitherto undescribed Gag lattice form. We show that this Gag lattice form is predominant in a subset of infected T-cells, and hypothesize that it is the result of premature PR activation in these cells.

The cET analysis of HIV-1 Gag and Gag-Pol expressing cells also provided unprecedented snapshots of cortical actin filaments in the vicinity of viral budding sites (Fig. 2, 3). It is well established that HIV-1 particles contain substantial amounts of actin [31], but the interplay between cortical actin and retrovirus assembly is still unknown. A recent study indicated a role for star-shaped actin filament arrangements emanating from the viral budding sites in HIV-1 assembly [32]. Our cET data provide 3D structures suggesting directed arrangement of actin filaments towards some

budding sites, particularly in filopodia-associated buds. Strikingly, half of the budding sites in the current data set were present on actin-filled filopodia. To what extent this is a consequence of the particular region of the cell accessible to cET remains to be determined. Future cET studies using altered Gag constructs and selective manipulation of the actin dynamics will provide the necessary tools to untangle the molecular details of the involvement of actin at HIV-1 budding sites.
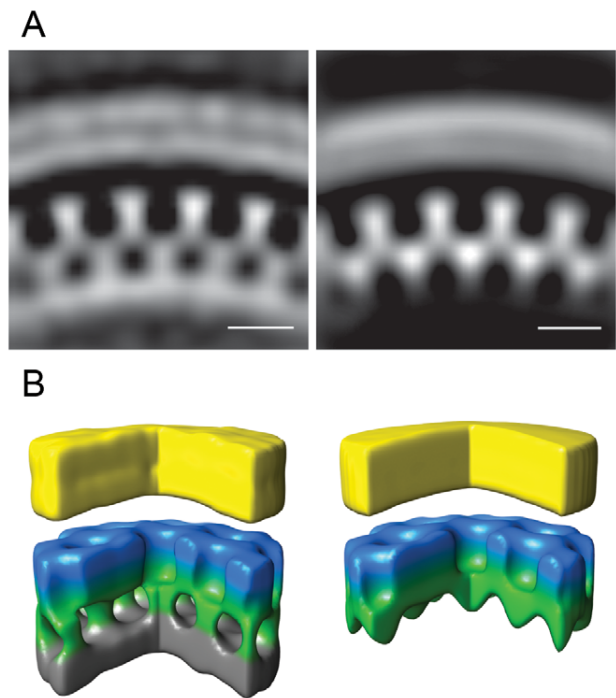
Extracellular immature HIV-1 particles are the direct products of viral budding at the plasma membrane, making it likely that their Gag protein lattice is determined during assembly. The observation that the Gag lattice is incomplete in the immature virus [22,26] and contains defects [19,22] raised the alternative hypothesis that Gag rearrangements occur following release. This could involve post-release association of smaller patches of Gag hexamers, or post-release symmetry breakage and dissociation of a hexagonal lattice containing evenly spaced pentameric defects [22]. Analysis of the lattice in extracellular immature virions revealed continuous hexameric symmetry over most of the lattice with curvature induced by defects of irregular size and shape [19]. These studies were performed on purified released particles, however, and the degree of order attained during assembly at the plasma membrane may have been higher (and broken after release) or lower (and subsequently assembled into a more ordered lattice). Our cET data of budding sites revealed structures of the immature Gag lattice in intact unperturbed cells prior to release, allowing us to draw conclusions on the early steps of the HIV-1 assembly process. There was no indication of separated "islands" of hexameric Gag in the *in situ* budding sites; upon visual inspection they consistently contained one continuous layer of Gag protein. A more detailed analysis of the lattice in the immature budding sites revealed that its unit cell structure was highly similar to that previously described for released immature virions [19], and included irregular lattice defects of similar size and distribution. This suggests that the organization of the immature
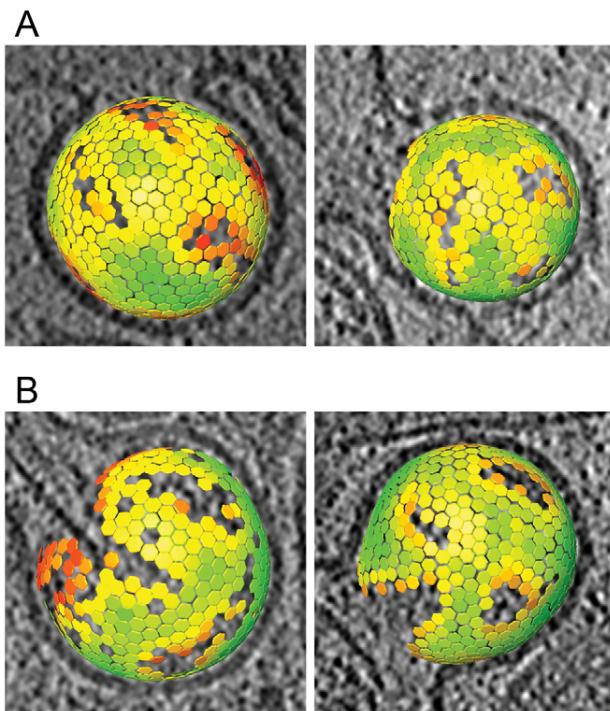
Figure 4. Lattice maps of immature and intermediate HIV-1 budding sites. Lattice maps of budding sites. The center and orientation of each aligned subtomogram is marked with a hexagon and is colored according to the cross correlation on a scale from low (red) to high (green). The lattice maps have the threshold set to the same value used for the subtomogram averaging procedure. The lattice map has been superimposed on a slice of 8 nm thickness through the budding viruses. (A) Lattice maps showing two budding particles with the newly reported lattice. (B) Lattice maps showing two budding particles with an immature lattice.
doi:10.1371/journal.ppat.1001173.g004

Figure 5. Gag unit cell structure in HIV-1 budding sites. Left panels show the average of the aligned subtomograms extracted from an individual budding site with the immature lattice. The right panels show the average from an individual budding site with the newly described lattice. (A) Central radial sections from the structures. Density is white. (B) Isosurface rendering of the structures. A segment of 90° has been cut out to better reveal the internal organization of the lattice. The surfaces have been colored radially to illustrate different domains in Gag: Yellow - membrane + MA; blue/green - CA; grey - NC + RNA. Scale bars 10 nm.
doi:10.1371/journal.ppat.1001173.g005

HIV-1 particle is indeed determined at the point of its intracellular assembly, and not as a result of large scale post-release ordering or disordering.

Besides the described immature Gag lattice, we observed a previously undescribed lattice type in extracellular particles and in budding sites still connected to the cell. Formation of this lattice was dependent on an active HIV-1 PR. The newly discovered lattice exhibited hexagonal symmetry with a CA layer identical to the immature lattice, but lacked the second layer of density corresponding to the NC-RNA complex. The particles and budding sites also contained no visible NC-RNA condensate in the center of the particles or adjacent to the buds, suggesting that the viral genome is absent. A comparison with the analysis performed in the accompanying paper by de Marco *et al.* on isolated virions with specific Gag cleavage site mutations showed that the newly described lattice is most likely composed of the MA-SP1 region of HIV-1 Gag. Although the tomographic average at current resolution does not allow identification of the density corresponding to the SP1 peptide, de Marco *et al.* showed that this peptide is present, independent of whether the cleavage site between CA and SP1 is intact or not. Notably, the particles produced from the MA-CA and MA-SP1 cleavage site mutants which exhibit an MA-SP1 type lattice do contain a clearly distinguishable density corresponding to condensed NC-RNA (de Marco *et al.*), which distinguishes them from the particles with this lattice type in the current study.

As further described in the accompanying paper by de Marco *et al.*, the MACASP1 lattice is likely to constitute a natural intermediate in the proteolytic maturation of HIV-1. Cleavage between SP1 and NC is the fastest processing step *in vitro* and abscission of the C-terminal NC-p6 regions from the immature lattice is required for condensation of the inner ribonucleoprotein complex of the virion prior to formation of the capsid. Thus, extracellular particles containing the MACASP1 lattice and inner density corresponding to NC-RNA are suggested to represent a "frozen" intermediate in HIV-1 maturation (de Marco *et al.*). Such particles were not observed for wild-type HIV-1 constructs, however, indicating that maturation intermediates are normally short-lived. An MACASP1 lattice without adjacent NC-RNA density was observed in 18% of HIV-1 budding sites on infected MT-4 cells, on the other hand, and its frequent presence suggests that it is a metastable structure with a significant lifetime in this case.

We previously reported that budding sites on infected MT4 cells have a broad distribution of Gag layer morphologies. Some resemble actual released immature virions, but some have a more complete Gag layer, containing more Gag molecules than the released virions and thus rather resemble arrested "late" budding sites not thought to be virus precursors [26]. When re-evaluating these data with respect to the type of Gag lattice, the immature lattice was predominantly found in budding sites resembling released virions, whereas the budding sites with the novel lattice had a significantly more closed Gag layer than released virions
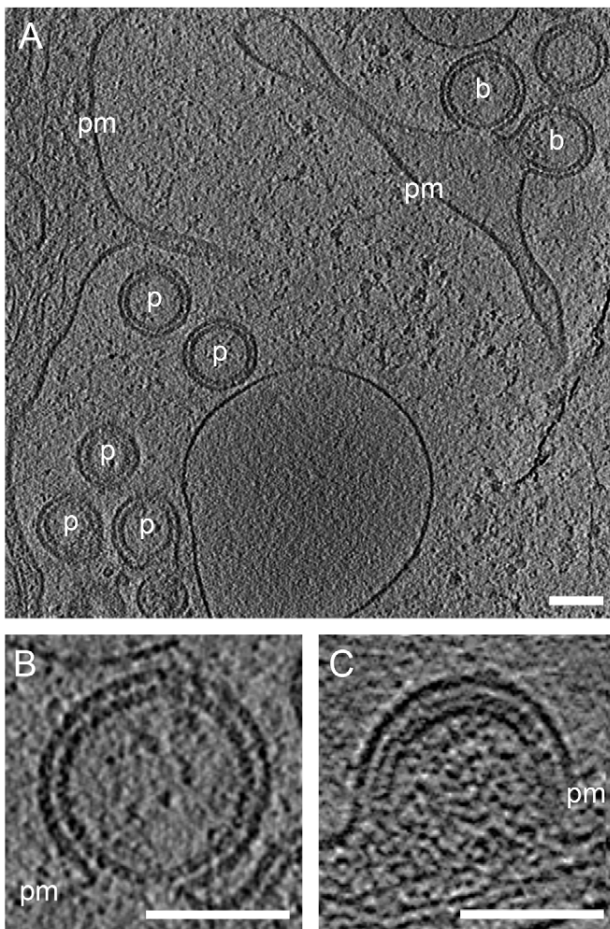
Figure 6. A previously undescribed Gag lattice in budding sites and released particles. (A) Computational slice through cryo electron tomograms of U-373 MG cells transduced with AdGagPol. Budding sites (b) and released particles (p) exhibit a Gag lattice different from both the immature and the mature lattice. (B) Enlarged slice through the top budding site from (A). Density is observed at the expected position of the CA domain, but no prominent density corresponding to NC-RNA is detected. (C) For comparison, a slice through a budding site with the immature Gag lattice showing the characteristic two density layers underneath the plasma membrane. Thickness of computational slices through cryo electron tomograms is 1.6 nm. pm: plasma membrane. Scale bars are 100 nm.
doi:10.1371/journal.ppat.1001173.g006

(Table 4, Fig. 7F). In the case of late-domain defective variants, formation of a closed layer is thought to be caused by a failure to recruit the ESCRT machinery and thus to drive the release process. Consequently, Gag assembly continues until reaching an equilibrium structure, but the resulting late budding sites are mostly dead-end products. It appears likely that a similar process is also relevant for the budding sites with the novel lattice. We hypothesize that premature proteolytic activation prior to confinement of virion constituents in the budding virion leads to removal of NC-p6 from Gag and of PR and downstream *pol* products from Gag-Pol. The cleaved C-terminal fragments are lost from the assembly site, while the MACASP1 lattice is stably retained. No ESCRT recruitment occurs since the p6 domain has been removed and no further processing occurs because PR, not being confined in a viral particle, has been lost. Accordingly, no budding structures with mature cores were observed. Further

assembly of the truncated lattice may involve additional Gag molecules undergoing partial processing or addition of cytoplasmic MACASP1 molecules since a truncated MACASP1 construct has been shown to form budlike structures when overexpressed [33,34]. The presence of extracellular particles with the MA-CASP1 lattice and lacking NC-RNA may indicate some ESCRT-independent release. Release of virions lacking most of the NC domain (but in this case carrying the p6 domain) has also been observed in previous studies [35,36].

The results of the current study clearly show that morphological maturation can be initiated inside the producer cell, but indicate that it is largely non-productive if the virion constituents are not confined in a structure sequestered from the cytoplasm. Assembly, release and maturation would therefore require the sequential processes of Gag association, ESCRT recruitment (prior to closure of the Gag shell), and PR activation (after or concomitant with ESCRT-mediated sequestration of the bud). The observation of aberrant, apparently dead-end structures in a subset of cells in HIV-1 infected T-cell lines suggests that the control of these processes can be lost, however. Infected MT-4 cells generally contained only one type of budding sites indicating that this loss of control may be caused by the host cell environment. Conceivably, host cell factors may be involved in regulating the kinetics of the consecutive stages of viral assembly, budding and release, and the subset of T-cells may have lost this regulation, although a general cytopathic effect cannot be ruled out at present. Identifying viral or host factors involved in this regulation will be of great importance for our understanding of HIV-1 morphogenesis, and may also lead to new approaches to render HIV-1 infection non-productive.

Materials and Methods

Construction of adenoviral vectors for HIV-1 Gag and GagPol

Adenoviral vectors for HIV-1 Gag (AdGag) and GagPol (AdGagPol) were constructed from the Rev independent versions of these genes [29], using the BD Adeno-X Expression System 1 (BD Biosciences Clontech). Briefly, the 1539 nt *gag*-gene was amplified by PCR adding flanking *NheI/XbaI* sites and ligated into the transfer vector pShuttle2. The *gag-pol* gene was excised with *NdeI/XbaI* from pcDNA3.1syngagpol [29] and directly cloned into pShuttle2. After verification of the insert by sequence analysis, these plasmids were digested with *PI-Sce/I-Ceu*. The inserts were ligated into the pre-digested Adeno.X viral DNA. Resulting plasmids were verified by restriction digest and DNA sequencing. Recombinant adenoviruses were generated by transfection of the AdGag or Ad GagPol plasmids into HEK 293 cells and amplified according to the protocol supplied with the kit. Titers of adenoviral vectors were determined using the Adeno-X rapid Titer Kit (Clontech); titers of both AdGag and Ad GagPol typically reached $1-2 \times 10^9$/ml. Immunofluorescence staining of AdGag and AdGagPol transduced cells for HIV-1 Gag was performed using a rabbit polyclonal anti-CA antibody and a FITC-conjugated goat-anti-rabbit secondary antibody. Western blots of cell lysates and particles purified from the culture media of AdGag and AdGagPol transduced cells were performed using the same polyclonal anti-CA antibody, with an IRDye 800CW goat anti-rabbit secondary antibody (LI-COR Biosciences) as recommended by the manufacturer.

Cell culture and adenoviral vector transductions for cryo electron tomography

Human glioblastoma cell lines U-87 MG and U-373 MG were maintained in DMEM and MEM, respectively, supplemented
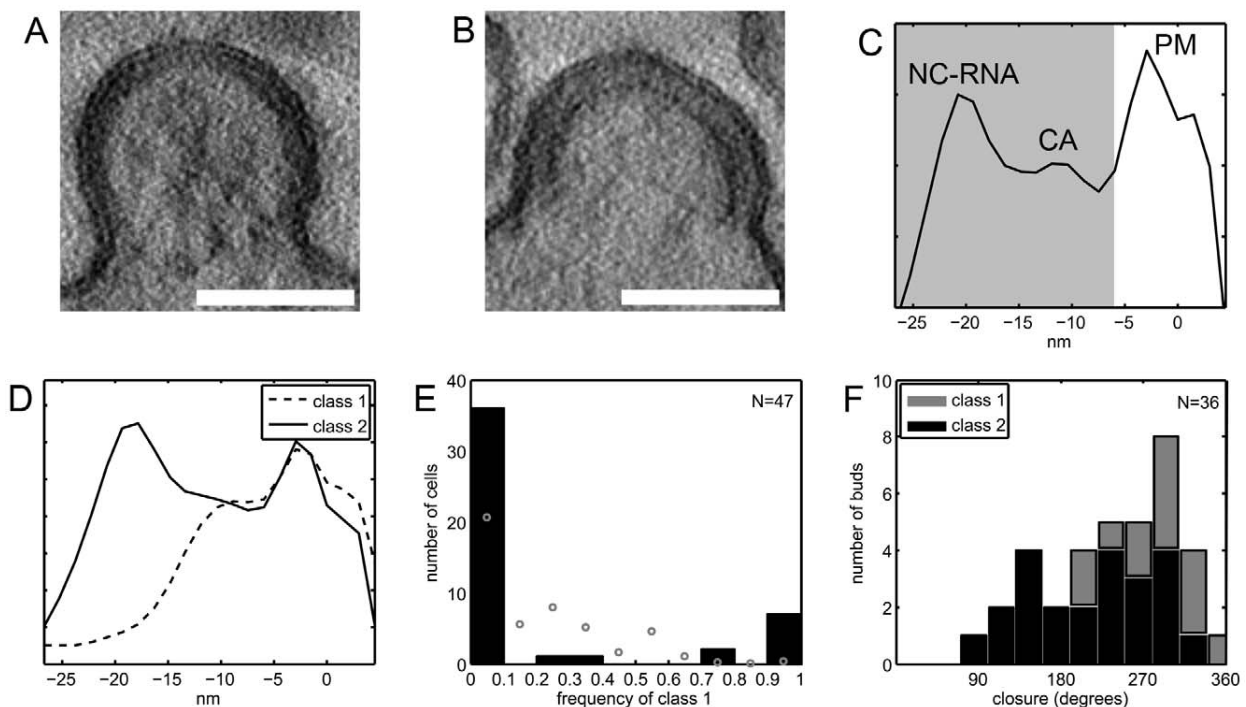
Figure 7. Budding sites with the novel lattice in HIV-1 infected T cells. (A–B) Computational slices of 1.5 nm thickness through tomograms of resin-embedded HIV-1 infected MT4 cells. By eye, two types of budding sites are discernible, one type having a thinner Gag density (A) and the other (B) the typical thickness of the immature Gag lattice. (C) Linear density profile of the Gag layer in one representative resin-embedded immature budding site. The profile was calculated orthogonally to the membrane in the central part of the budding site, as described in Materials and methods. Membrane center at 0 nm, negative direction corresponds to cell interior. Peaks are discernible at the position of the membrane leaflets (PM), the NC-RNA layer, and (less strong) the CA layer of the immature lattice. The shaded area indicates the part of the plots used for classification of bud types. (D) Budding sites in resin embeddings of infected T cells were classified based on line plots like the one shown in (C). The average plots for the two major classes differ at the position of the NC/RNA layer, class 1 (dashed line) resembling the novel lattice and class 2 (solid line) resembling the immature lattice. (E) Frequency of class 1 budding sites on 47 infected MT-4 T cells. Black bars indicate the number of cells carrying either no (frequency 0) or exclusively (frequency 1) class 1 budding sites or a mixture of class 1 and class 2 budding sites. The theoretical frequency distribution which would be expected if all cells exhibited the average frequency of 18% buds in class 1 is given by grey circles. This distribution has local maxima around 25% and 50% due to the limited number of budding sites sampled per cell, as explained further in the Supplementary Information. (F) Gag layer closure in budding sites. The previously published data on the closure of the Gag layer in budding sites on MT-4 cells show a broad distribution of morphologies [26]. Separating these data into the two classes of Gag densities shows that class 2 budding sites (black bars) are more similar to released immature HIV-1 particles, whereas class 1 budding sites (grey bars) exhibit a more closed Gag layer. Scale bars are 100 nm.
doi:10.1371/journal.ppat.1001173.g007

with Penicillin-Streptomycin and 10% foetal bovine serum. When grown on EM grids (CF-2/1-2AU, Protochips, inc.), cells were seeded at low density and kept in medium with 0.5% serum which increased the extent of flat areas of the cells. Transduction with

AdGag or AdGagPol was performed concomitant with or one day after seeding cells on EM grids. A multiplicity of 'infection' (MOI) of 10 was typically used, calculated based on cell culture dish area and assuming that added vector will adhere to any substrate

Table 2. Data sets of resin embedded samples evaluated in the classification of Gag layer types.

Data set number	Sample	Number of tilt series	Number of cells in the data set	Number of budding sites or virions fully contained within sections
4	MT4 cells infected with HIV-1	85	50	224
5	MT4 cells infected with HIV-1 and cultured in presence of lopinavir	29	18	53
6	HEK293T cells transfected with pNL4-3-MACASP1 (only released virions)	7	-	16
7	HeLa cells transfected with wt pNL4-3	33	20	69
8	MT4 cells infected with HIV-1	19	-	36

The number of tilt series recorded, and containing budding sites or released particles, are stated for each data set. Data set 8 was previously evaluated for Gag layer closure [26], whereas data sets 4–7 were recorded in a manner to allow for statistical evaluation of frequencies of Gag lattice types (see text).
doi:10.1371/journal.ppat.1001173.t002

Table 3. Number of budding sites or virions assigned to class 1 and class 2 Gag layer morphology, respectively, for the different sample groups.

Sample (data set number)	Number of budding sites or virions	Class 1	Class 2	Frequency of class 1 (%)	Frequency of class 1, 95% confidence interval (%)
MT4 cells infected with HIV-1 (4)	224	41	183	18	13.5–24.0
MT4 cells infected with HIV-1 in the presence of lopinavir (5)	53	0	53	0	0.0–6.7
Particles containing uncleaved MA-SP1 (6)	16	15	1	-	-
HeLa cells transfected with wt pNL4-3 (7)	69	2	67	2.9	0.35–10
MT4 cells infected with HIV-1 (8)	36	13	23	-	-

Where applicable (data acquisition scheme allowing for statistical interpretation), a frequency and its 95% confidence interval was calculated using the MATLAB function *binofit*, which uses an F-function approximation of the exact Clopper-Pearson formula [41].
doi:10.1371/journal.ppat.1001173.t003

surface. One to three days after seeding cells on the grids, they were plunge-frozen into liquid ethane after application of 10 nm colloidal gold.

Sample preparation for electron tomography of sections of resin-embedded cells

Cell culture and preparation for sections of resin embedded cells were performed as previously described [26]. For production of immature virus, cells were grown in the presence of 1 μM lopinavir. For production of particles containing an uncleaved MA-SP1 layer, HEK293T cells were transfected with a proviral HIV-1$_{NL4-3}$ plasmid carrying mutations in the cleavage sites between MA and CA and CA and SP1, respectively ([14] and accompanying paper by de Marco *et al.*). Cells were fixed, stained, dehydrated and EPON-embedded, and 300 nm sections of the embeddings were cut and transferred to EM grids after deposition of colloidal gold onto the EM grid support film.

Electron tomography

Cryo electron tomography was performed with a Philips CM300 or a FEI Tecnai F30 Polara transmission electron microscope (FEI; Eindhoven, The Netherlands), both equipped with 300 kV field emission guns, Gatan GIF 2002 post-column energy filters and 2048×2048 Multiscan CCD cameras (Gatan; Pleasanton, CA). All data collection was performed at 300 kV, with the energy filter operated in zero loss mode. Tilt series of cryo specimens were typically recorded from −60° to +60° with an angular increment of 1.5°, a total electron dose of 80 electrons Å^{-2}, and a defocus of −6.0 μm to −8.0 μm. The pixel sizes at the specimen level were 8.21 Å at the CM300, and 8.05 Å or 7.13 Å

at the FEI Tecnai F30 Polara, respectively. Electron tomography of resin sections was performed at an FEI Tecnai F20 transmission electron microscope (FEI; Eindhoven, The Netherlands), equipped with 200 kV field emission gun and a 4096×4096 Eagle CCD camera (FEI; Eindhoven, The Netherlands). Tilt series on resin sections were recorded at room temperature from −60° to +60° with an angular increment of 1.0°, with a one time binned pixel size of 2*3.72 Å and a defocus of -4.0 μm. For both cryo and room temperature tilt series, the applied electron dose for a given tilt angle α was proportional to $1/\cos(\alpha)$ to compensate for the higher effective specimen thickness at higher tilts. At the Tecnai microscopes, the SerialEM acquisition software [37] was used for tilt series acquisition. Three-dimensional reconstructions from tilt series were calculated using the weighted back-projection method [38], as implemented in the TOM toolbox [39] for MATLAB (Mathworks, Natick, Massachusetts, United States) with the "exact weighting" option, or in IMOD [40].

Subtomogram averaging

Subtomogram averaging was performed using MATLAB. Sub-volumes of $(38.6 \text{ nm})^3$ were extracted along the surface of the budding particles. The subtomograms were then iteratively aligned against the average in a reference free manner, while applying 6-fold symmetry as described in the accompanying article (de Marco *et al*). For individual particles, the distribution of the cross correlation values between the aligned subtomograms and the reference was bimodal. The threshold of the subtomograms that went into the reconstruction was set as minimum between the two peaks.

Classification of Gag lattice types in tomograms of resin-embedded sections

The classification of Gag densities in resin embeddings of HIV-1 budding sites was performed as described in the Supplementary Information. In brief, linear density profiles were calculated orthogonally to the plasma membrane in the budding sites. These one-dimensional density profiles were cropped to contain only the membrane-apposed density, and subjected to hierarchical clustering.

Table 4. Properties of budding sites on MT4 cells according to Gag layer class.

	Class 1	Class 2	Total
Number of buds	13	23	36
Radius (nm)	76±5	74±8	75±7
Closure (°)	289±45	218±70	244±70
Number of Gag molecules per bud	5,000±1,300	3,300±1,500	3,900±1,600

The calculation is based on the published evaluation of data set 8 (Table 2), assuming the same Gag packing density in both class 1 and class 2, which is valid if they represent the NC-RNA-less and immature lattice, respectively.
doi:10.1371/journal.ppat.1001173.t004

Supporting Information

Supporting Information S1 Supporting image analysis procedures and additional cryo electron tomography images.
Found at: doi:10.1371/journal.ppat.1001173.s001 (1.45 MB PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: LAC JAGB HGK KG. Performed the experiments: LAC AdM HO AH. Analyzed the data: LAC AdM JAGB HGK KG. Wrote the paper: LAC JAGB HGK KG.

References

1. Freed EO (1998) HIV-1 gag proteins: diverse functions in the virus life cycle. Virology 251: 1–15.
2. Hurley JH (2008) ESCRT complexes and the biogenesis of multivesicular bodies. Curr Opin Cell Biol 20: 4–11.
3. Morita E, Sundquist WI (2004) Retrovirus budding. Ann Rev Cell Dev Biol 20: 395–425.
4. Williams RL, Urbe S (2007) The emerging shape of the ESCRT machinery. Nat Rev Mol Cell Biol 8: 355–368.
5. Briggs JAG, Simon MN, Gross I, Kräusslich HG, Fuller SD, et al. (2004) The stoichiometry of Gag protein in HIV-1. Nat Struct Mol Biol 11: 672–675.
6. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, et al. (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell 107: 55–65.
7. VerPlank L, Bouamr F, LaGrassa TJ, Agresta B, Kikonyogo A, et al. (2001) Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). Proc Natl Acad Sci U S A 98: 7724–7729.
8. Coffin JM, Hughes SH, Varmus HE (1997) Retroviruses: Cold Spring Harbor Laboratory Press. .
9. Pettit SC, Moody MD, Wehbie RS, Kaplan AH, Nantermet PV, et al. (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: 8017–8027.
10. Wiegers K, Rutter G, Kottler H, Tessmer U, Hohenberg H, et al. (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J Virol 72: 2846–2854.
11. Coren LV, Thomas JA, Chertova E, Sowder RC, Gagliardi TD, et al. (2007) Mutational analysis of the C-terminal Gag cleavage sites in human immuno-deficiency virus type 1. J Virol 81: 10047–10054.
12. Checkley MA, Luttge BG, Gross I, Soheilian F, Nagashima K, Freed EO (2010) The capsid-spacer peptide 1 Gag processing intermediate is a dominant-negative inhibitor of HIV-1 maturation. Virology 400: 137–144.
13. Lee SK, Harris J, Swanstrom R (2009) A strongly transdominant mutation in the human immunodeficiency virus type 1 gag gene defines an Achilles heel in the virus life cycle. J Virol 83: 8536–8543.
14. Müller B, Anders M, Akiyama H, Welsch S, Glass B, et al. (2009) HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. J Biol Chem 284: 29692–29703.
15. Kräusslich HG (1991) Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. Proc Natl Acad Sci U S A 88: 3213–3217.
16. Kaplan AH, Zack JA, Knigge M, Paul DA, Kempf DJ, et al. (1993) Partial inhibition of the human immunodeficiency virus type 1 protease results in aberrant virus assembly and the formation of noninfectious particles. J Virol 67: 4050–4055.
17. Moore MD, Fu W, Soheilian F, Nagashima K, Ptak RG, et al. (2008) Suboptimal inhibition of protease activity in human immunodeficiency virus type 1: effects on virion morphogenesis and RNA maturation. Virology 379: 152–160.
18. Benjamin J, Ganser-Pornillos BK, Tivol WF, Sundquist WI, Jensen GJ (2005) Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography. J Mol Biol 346: 577–588.
19. Briggs JA, Riches JD, Glass B, Bartonova V, Zanetti G, et al. (2009) Structure and assembly of immature HIV. Proc Natl Acad Sci U S A 106: 11090–11095.
20. Briggs JAG, Grünewald K, Glass B, Förster F, Kräusslich HG, et al. (2006) The mechanism of HIV-1 core assembly: Insights from three-dimensional recon-structions of authentic virions. Structure 14: 15–20.
21. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. Nature 455: 109–113.
22. Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, et al. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J 26: 2218–2226.
23. Zhu P, Liu J, Bess J, Jr., Chertova E, Lifson JD, et al. (2006) Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature 441: 847–852.
24. Zhu P, Winkler H, Chertova E, Taylor KA, Roux KH (2008) Cryoelectron tomography of HIV-1 envelope spikes: further evidence for tripod-like legs. PLoS Pathog 4: e1000203.
25. Briggs JAG, Wilk T, Welker R, Kräusslich HG, Fuller SD (2003) Structural organization of authentic, mature HIV-1 virions and cores. EMBO J 22: 1707–1715.
26. Carlson LA, Briggs JA, Glass B, Riches JD, Simon MN, et al. (2008) Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis. Cell Host Microbe 4: 592–599.
27. Li S, Hill CP, Sundquist WI, Finch JT (2000) Image reconstructions of helical assemblies of the HIV-1CA protein. Nature 407: 409–413.
28. Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI (1999) Assembly and analysis of conical models for the HIV-1 core. Science 283: 80–83.
29. Wagner R, Graf M, Bieler K, Wolf H, Grunwald T, et al. (2000) Rev-independent expression of synthetic gag-pol genes of human immunodeficiency virus type 1 and simian immunodeficiency virus: implications for the safety of lentiviral vectors. Hum Gene Ther 11: 2403–2413.
30. Medalia O, Weber I, Frangakis AS, Nicastro D, Gerisch G, et al. (2002) Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. Science 298: 1209–1213.
31. Ott DE, Coren LV, Kane BP, Busch LK, Johnson DG, et al. (1996) Cytoskeletal proteins inside human immunodeficiency virus type 1 virions. J Virol 70: 7734–7743.
32. Gladnikoff M, Shimoni E, Gov NS, Rousso I (2009) Retroviral assembly and budding occur through an actin-driven mechanism. Biophys J 97: 2419–2428.
33. Crist RM, Datta SA, Stephen AG, Soheilian F, Mirro J, et al. (2009) Assembly properties of human immunodeficiency virus type 1 Gag-leucine zipper chimeras: implications for retrovirus assembly. J Virol 83: 2216–2225.
34. Hockley DJ, Nermut MV, Grief C, Jowett JB, Jones IM (1994) Comparative morphology of Gag protein structures produced by mutants of the gag gene of human immunodeficiency virus type 1. J Gen Virol 75(Pt 11): 2985–2997.
35. Ott DE, Coren LV, Chertova EN, Gagliardi TD, Nagashima K, et al. (2003) Elimination of protease activity restores efficient virion production to a human immunodeficiency virus type 1 nucleocapsid deletion mutant. J Virol 77: 5547–5556.
36. Popova E, Popov S, Gottlinger HG (2010) Human Immunodeficiency Virus Type 1 Nucleocapsid p1 Confers ESCRT Pathway Dependence. J Virol 84: 6590–6597.
37. Mastronarde DN (2005) Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol 152: 36–51.
38. Frank J (2006) Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell Berlin: Springer.
39. Nickell S, Förster F, Linaroudis A, Del Net W, Beek F, et al. (2005) TOM software toolbox: acquisition and analysis for electron tomography. Journal of Structural Biology 149: 227–234.
40. Kremer JR, Mastronarde DN, McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. J Struct Biol 116: 71–76.
41. Westfall PH, Young SS (1993) Resampling-based multiple testing: examples and methods for P-value adjustment: Wiley-Interscience. 360 p.

Supplementary Information

Cryo Electron Tomography of native HIV-1 Budding Sites

Lars-Anders Carlson[1,2], Alex de Marco[3], Heike Oberwinkler[1], Anja Habermann[1], John A.G. Briggs[3], Hans-Georg Kräusslich[1,]*, Kay Grünewald[2,4]*

[1]Department of Infectious Diseases, Virology, Universitätsklinikum Heidelberg, Im Neuenheimer Feld 324, D-69120 Heidelberg, Germany

[2]Dept. Molecular Structural Biology, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany

[3]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

[4]Oxford Particle Imaging Centre, Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

*joint corresponding authors

HGK: Tel +496221 565001; Fax +496221 565003;

Email: hans-georg.kraeusslich@med.uni-heidelberg.de

KG: Tel +441865 287817; Fax +441865 287547;

Email kay@strubi.ox.ac.uk

Classification of Gag lattice types in tomograms of resin-embedded sections

The position of the plasma membrane in the budding sites was manually traced using IMOD [1] on every third xy-slice in the central 30 slices of the budding sites (where the membrane was cut in an angle of 90±~20° by the tomographic slice). From this set of linearly connected points, a triangular mesh surface of the plasma membrane was interpolated and smoothed using the IMOD functions smoothsurf, imodmesh and imodfillin, combined with custom-made scripts for MATLAB (Mathworks, Natick, Massachusetts, United States). The points in this surface model have an approximate spacing of one voxel, and the surface normal at every point can be calculated using the connection to the neighboring points. Using this, the tomogram grey values along a line orthogonal to the membrane were calculated for every point, and the average of these one-dimensional plots over the membrane surface was calculated.

When necessary, the resulting line plots were scaled to achieve the same object pixel size. They were cropped to either [-27 nm, -6 nm] or [-27 nm, 4.5 nm] (where 0 nm is the center of the membrane and the negative direction is the cell/virion interior) to contain either the membrane-bound density or the membrane-bound density and the membrane. After normalization of the scaled and cropped line plots to average 0 and standard deviation 1, they were subjected to further analysis.

Classification of the line plots was performed in MATLAB using functions in the statistics toolbox. Principal component analysis [2] was performed with the MATLAB function *princomp*, hierarchical clustering with the MATLAB functions *pdist*, *linkage* and *cluster*. Briefly, the distance between the 15 element vectors (containing the membrane-bound density) was calculated as the Euclidean distance and was used for linkage analysis with *linkage*, using

Ward's linkage criterion of increase in intragroup variance. Linkage analysis using the criterion of mean intercluster distance gave indistinguishable results.

Calculation of the theoretical frequency distribution in Figure 7E

The gray circles in Figure 7E represent the theoretical frequency distribution expected if all cells had 18% budding sites in class 1 (thin lattice). This distribution was calculated by taking into account the small number of budding sites sampled per cell, in the following way:

(i) If the same number n of budding sites had been recorded on every cell, the probability of having k (k≤n) budding sites in class 1 (thin lattice) is given by the Binomial distribution

$$P_n(k) = (n!/(k!(n-k)!)) \times 0.18^k \times (1-0.18)^{(n-k)}.$$

The frequency (x-axis in Figure 7E) is given by k/n and its probability (y-axis in Figure 7E) is given by $P_n(k)$.

(ii) The actual experimental data contain a varying number of budding sites per cell. We denote the probability of n budding sites on a cell in the data set by W_n. The theoretical frequency distribution displayed in Figure 7E is then given as the weighted average of the $P_n(k)$ by W_n, where the frequencies k/n for each individual $P_n(k)$ was rounded to the closest 0.1 for the histogram representation. This combined frequency distribution has local maxima around 0.25 and 0.5 since these values are among the few that can be formed for certain values of n (2,4,8) that occur often in the data.
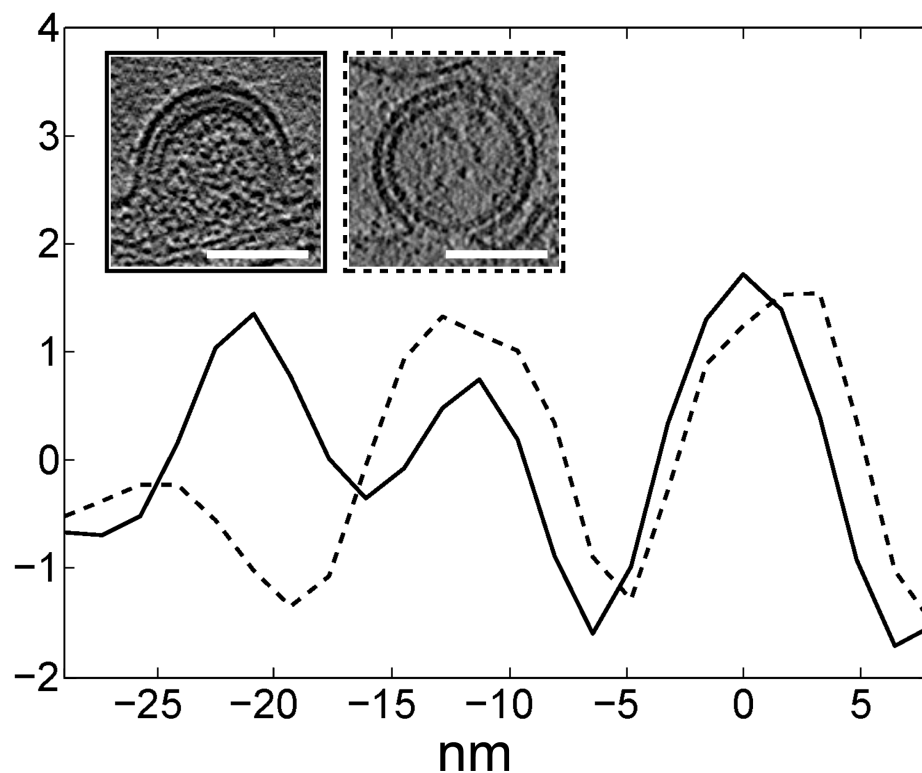
Figure S1. Comparison of budding sites with the immature or the newly described thinner lattice in cryo electron tomograms. Linear density profiles were produced by integration of the tomogram density perpendicular to the plasma membrane in one budding site with the immature (solid line) and one with the newly described (dashed line) lattice. Membrane at 0 nm, negative direction corresponds to cell interior. The three peaks in the immature budding site correspond to the plasma membrane and the known positions of the CA and NC/RNA layers of the immature Gag lattice. The budding site with the newly described lattice has the CA peak at the same distance as the immature budding site, but lacks the NC/RNA peak. The inset show computational slices of 1.6 nm thickness through the cryo electron tomograms of the analyzed budding sites with the immature (left) and newly described (right) lattice. Scale bars are 100 nm.
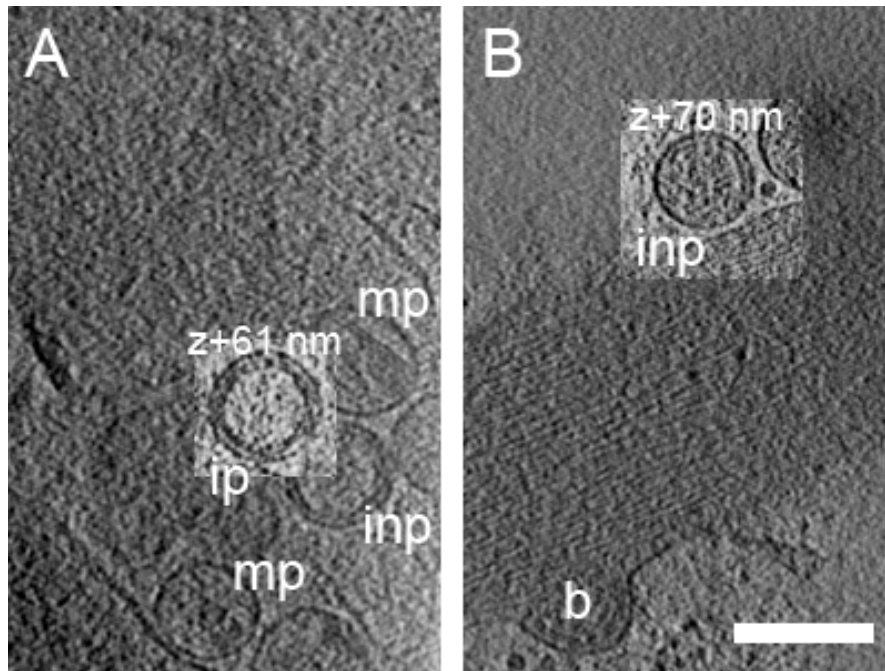
Figure S2. The newly described Gag lattice in released particles. Computational slices of 3.2 nm thickness through cryo electron tomograms of cells infected with AdGagPol. Released particles with the newly described lattice were found next to released particles with immature and mature morphology (A), as well as adjacent to budding sites with an immature Gag layer (B). Structure designation: b, budding site, ip, immature particle, inp= particle with an apparently intermediate lattice, mp=mature particle. Scale bar is 200 nm.
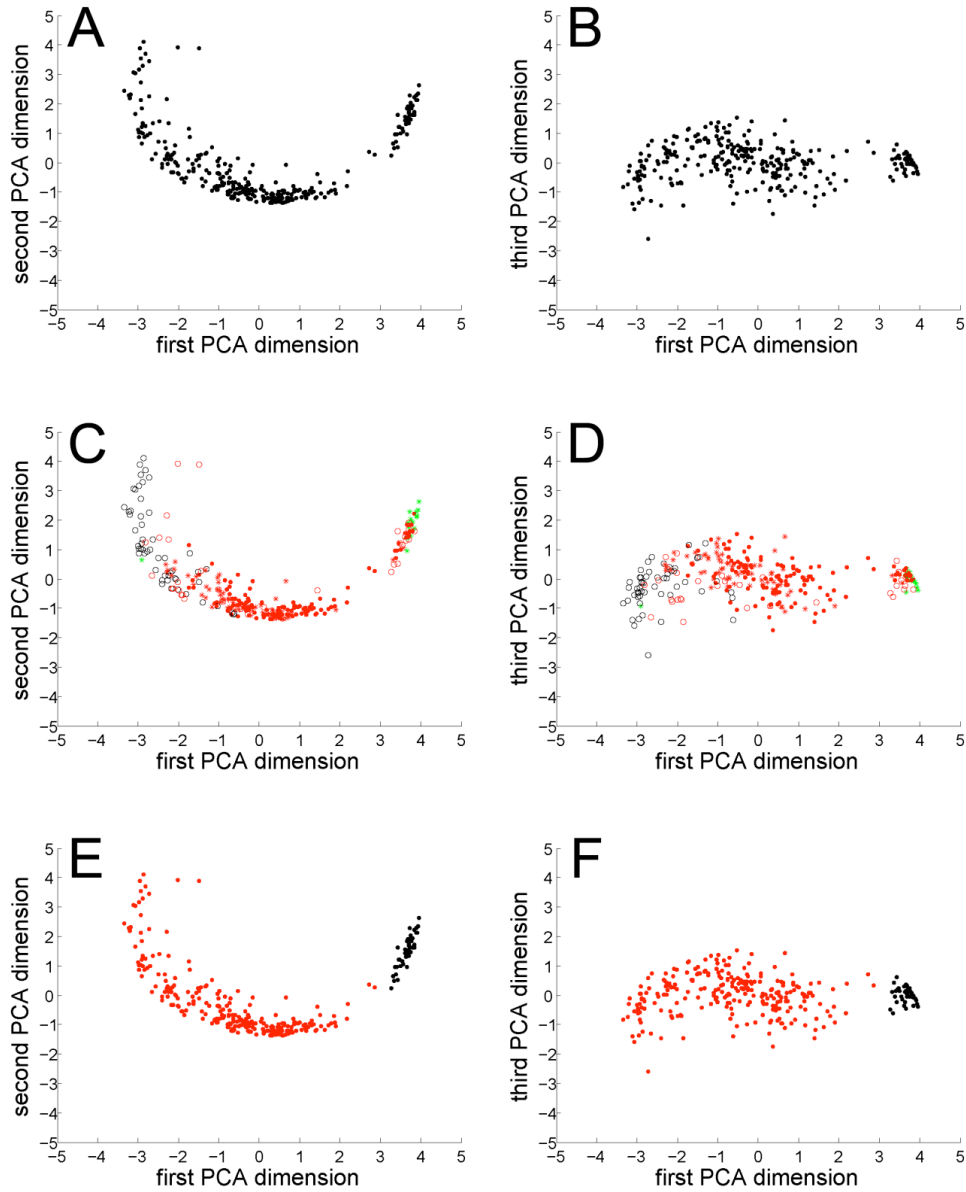
Figure S3. Visualization of variance between budding sites by principal component analysis (PCA). The cropped density profiles were subjected to PCA, and projected on the first and second (A, C, E) or the first and third (B, D, F) base vectors, respectively. (A-B) All data points. (C-D) All data points, colored according to sample group listed in Table 2: MT4 co-culture (data set 4), filled red circles, MT4 co-culture (data set 8), open red circles; HeLa wt, red stars; MT4 co-culture+lopinavir, open black circles; and released MACASP1 virions, filled green circles. (E-F) All data points, colored according to the class assignment in the hierarchical classification, class 1 in black, class 2 in red.
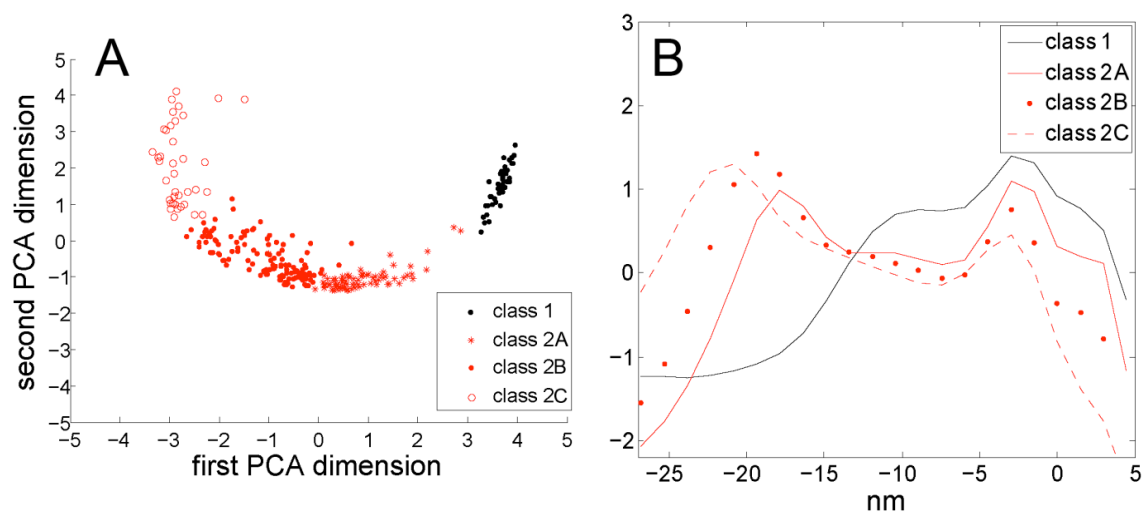
Figure S4. Variance within class two is mainly NC-RNA position. When four instead of two classes are separated from the hierarchical clustering (A), the main variance in the more populated class 2 of Figure S3 is seen to be slight variations in the position of the NC-RNA layer (B). There are also variations in the membrane region between the four classes, but since this part of the density profile is not used in the PCA or clustering (but only the shaded part as indicated in Fig. 6C), it will not contribute to the variation in class 2 found there.

References

1. Kremer JR, Mastronarde DN, McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. Journal of Structural Biology 116: 71-76.

2. Duda RO, Hart PE, Stork DG (2001) Pattern Recognition. New York: John Wiley & Sons, Inc.

Structural Analysis of HIV-1 Maturation Using Cryo-Electron Tomography

Alex de Marco[1], Barbara Müller[2], Bärbel Glass[2], James D. Riches[1], Hans-Georg Kräusslich[2], John A. G. Briggs[1]*

1 Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany, 2 Department of Infectious Diseases, Virology, Universitätsklinikum Heidelberg, Heidelberg, Germany

Abstract

HIV-1 buds form infected cells in an immature, non-infectious form. Maturation into an infectious virion requires proteolytic cleavage of the Gag polyprotein at five positions, leading to a dramatic change in virus morphology. Immature virions contain an incomplete spherical shell where Gag is arranged with the N-terminal MA domain adjacent to the membrane, the CA domain adopting a hexameric lattice below the membrane, and beneath this, the NC domain and viral RNA forming a disordered layer. After maturation, NC and RNA are condensed within the particle surrounded by a conical CA core. Little is known about the sequence of structural changes that take place during maturation, however. Here we have used cryo-electron tomography and subtomogram averaging to resolve the structure of the Gag lattice in a panel of viruses containing point mutations abolishing cleavage at individual or multiple Gag cleavage sites. These studies describe the structural intermediates correlating with the ordered processing events that occur during the HIV-1 maturation process. After the first cleavage between SP1 and NC, the condensed NC-RNA may retain a link to the remaining Gag lattice. Initiation of disassembly of the immature Gag lattice requires cleavage to occur on both sides of CA-SP1, while assembly of the mature core also requires cleavage of SP1 from CA.

Introduction

The assembly of HIV-1 occurs at the plasma membrane of infected cells. The primary structural component of the assembling virus is the 55-kDa polyprotein Gag. Multiple copies of Gag assemble to form an incomplete sphere underneath the plasma membrane, which recruits components of the cellular ESCRT machinery to mediate membrane scission and release of the budding virus from the cell [1,2,3]. Gag consists of three major structural components: MA, the membrane binding domain, CA, the capsid domain, and NC, the nucleocapsid domain, which interacts with the viral RNA. CA and NC are separated by a short linker peptide SP1, and downstream of NC are two further peptide domains, SP2 and p6 (Figure 1A) [4]. p6 contains the short linear motifs which are responsible for ESCRT recruitment [2,5,6]. The viral genome is recruited to the assembling particle from the pool of cellular mRNA via interactions between a packaging signal (Ψ) present at the 5′ end of the genome [7] and zinc-fingers present in NC [4].

During or concomitant with budding, the viral protease (PR) cleaves Gag in five positions into its component domains. These cleavages lead to the dramatic morphological changes required for creation of a mature, infectious virion. Proteolytic maturation is essential for infectivity, and PR inhibitors are a key element of current antiretroviral therapies [8]. Maturation appears to be a rapid process, and viruses imaged by electron microscopy (EM) in the vicinity of infected cells are predominantly mature, with only occasional immature particles. Intermediate maturation states have not been detected so far. For this reason, the majority of morphological and structural studies have focused either on the mature virus, or on the immature virus that can be produced in the presence of a PR inhibitor or by mutation of the PR active site.

In the immature virus, Gag is arranged in a radial fashion, with the N-terminal MA domain associated with the viral membrane, and the C-terminus of the protein pointed towards the centre of the particle. Gag adopts a curved hexameric lattice with an 8 nm unit cell [9], which closes through the incorporation of heterogeneously shaped defects [10]. After proteolytic cleavage, the MA layer is thought to remain associated with the viral membrane, whereas NC and the RNA are condensed into a ribonucleoprotein complex layer surrounded by the viral CA capsid core. The core is most often cone-shaped, though other shapes including capped tubes are seen. The core is thought to have a fullerene cone geometry formed from a hexameric lattice with twelve pentamers incorporated to allow closure [11], five at the narrow end of the cone and seven at the broad end of the cone shaped structure. CA adopts a larger 10 nm unit cell in the mature core [12,13].

The structure of the immature lattice has been explored by cryo-electron tomography combined with sub-tomogram averag-
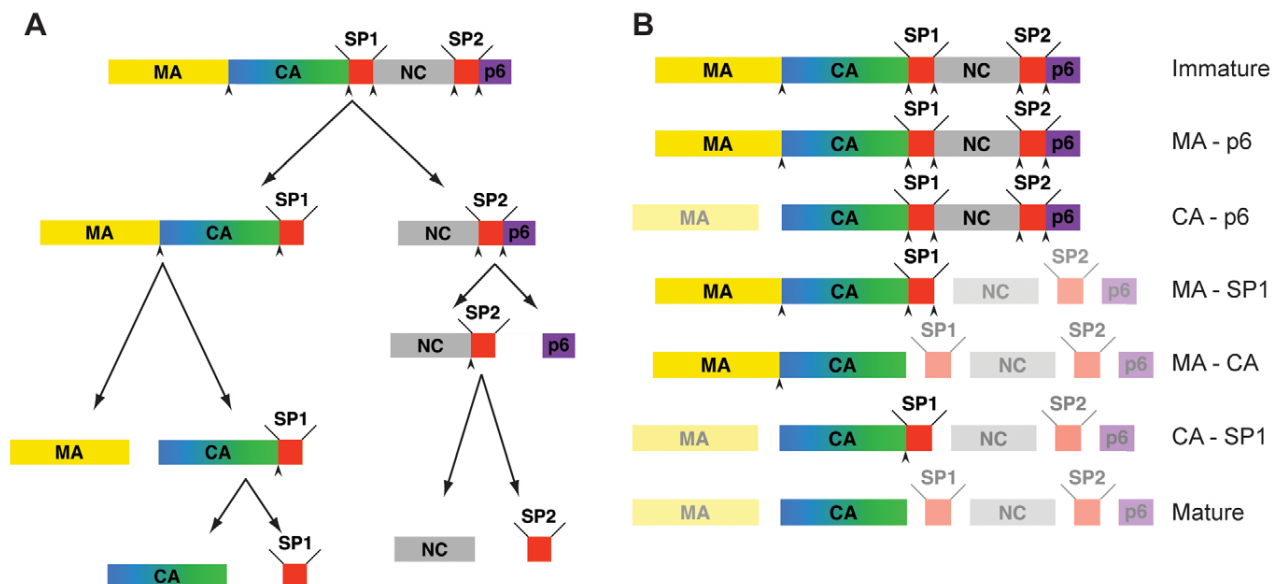
Author Summary

HIV-1 buds from the plasma membrane of infected cells in an immature form with the polyprotein Gag as its major component. Maturation into an infectious form requires cleavage of Gag in five positions. This process is an important target for antiretroviral drugs. Here we studied changes in the structure of the virus that occur during maturation, making use of virus variants in which different combinations of cleavage sites were mutated to prevent cleavage at those sites. We used cryo-electron tomography and sub-tomogram averaging to visualise the arrangement of Gag in 3D. We show that the fastest cleavage event leads to condensation of the RNA genome complexed with viral proteins. This inner RNA/protein structure appears to maintain a link with the remaining Gag lattice. Processing on both sides of CA-SP1, the main structural module of Gag, is required for disassembly of the immature Gag lattice, while removal of SP1 is needed in addition for mature core formation. The results provide structural correlates of the ordered processing events during HIV-1 maturation and shed light on the mechanism of action of bevirimat, an inhibitor of CA-SP1 cleavage in clinical trials.

Studies of the arrangement of CA in the mature core are much more advanced through the use of in vitro assembly of purified proteins to form regular arrays appropriate for high-resolution structural study [15,16]. In the mature core, the N-terminal residues of CA are folded back into the structure to form a β-hairpin [17]. The formation of this structure appears to be conserved in all retroviruses with the exception of the spumaviruses [17,18,19,20]. Mutagenesis studies suggest that the formation of the β-hairpin is important for the formation of the mature core [17,21]. This structure cannot be present in the immature virus since the N-terminal residues are covalently linked to the MA domain.

The five proteolytic cleavage sites in Gag are cleaved at very different rates in vitro. The fastest cleavage is that between SP1 and NC. Relative to this cleavage, SP2-p6 is cleaved 9x slower, MA-CA 14x slower, NC-SP2 350x slower, and CA-SP1 400x slower [22]. These differences, together with the observation of distinct and reproducible processing intermediates in lysates from infected cells or upon partial inhibition of PR [23], suggest that cleavage is a stepwise process in vivo. The order of cleavage based on these rates determined in vitro is illustrated in Figure 1A with initial processing separating NC-p6 from the membrane-bound N-terminal part of Gag, secondary cleavages separating MA from CA-SP1 and p6 from NC-SP2 and final processing removing the two spacer peptides from the C-termini of CA and NC, respectively. Mutational studies revealed that the efficacy and order of these cleavage events is highly relevant for viral infectivity: blocking cleavage at any individual processing site - with the exception of NC-SP2 - severely diminished or abolished virion infectivity [24,25,26,27], mostly affecting fusion or early post-fusion events [24,25,28].

Due to the fast time scale of maturation and the absence of any system for synchronizing the process, studies of maturation intermediates have relied upon viruses in which individual

ing [10,14]. The N-terminal CA domain forms hexameric rings surrounding central holes. Below these rings, 2-fold symmetric densities link adjacent hexamers. These densities have been assigned to the dimeric C-terminal CA domain. Six such dimers converge below the holes in the N-terminal CA domain into a rod-like structure suggested to represent a six-helix bundle formed by SP1 [14]. This rod-like structure descends towards the centre of the virus into the NC-RNA layer, which is not hexagonally ordered [10].

Figure 1. Steps in HIV-1 proteolytic maturation, and variants analysed. A) Schematic outline of the proteolytic cleavages which take place in Gag during the HIV-1 maturation process. Arrowheads indicate proteolytic sites before cleavage. The order of cleavage events shown is based on the rates of cleavage in vitro as described in [22]. B) Schematic representation of Gag after completion of cleavage for each variant analyzed. The non-cleaved products due to the inactivation of the proteolytic sites are highlighted. Mutated, and therefore uncleaved processing sites are indicated by an arrowhead.
doi:10.1371/journal.ppat.1001215.g001

cleavage sites have been mutated [22,27,29,30]. These studies showed that mutation of the cleavage site between MA and CA led to normal condensation of the NC-RNA complex without formation of the cone shaped capsid [27,29]. Inhibition of both cleavages flanking SP1 led to viruses with immature-like morphology, whereas inhibition of the CA-SP1 cleavage allowed NC-RNA condensation, but prevented formation of a mature-like core [27]. A thin electron-dense layer separated from the virion membrane was observed in this case and suggested to correspond to CA. A similar phenotype was observed when virus was produced in the presence of bevirimat [31], which specifically inhibits cleavage between CA and SP1. Although the inhibition of CA-SP1 processing by bevirimat is incomplete [31,32], the infectivity of virions raised in its presence is severely diminished. Similarly, partial inhibition of polyprotein processing by co-expression and co-assembly of wild-type (wt) and cleavage-site mutated Gag polyprotein showed a strong trans-dominant effect of the mutated proteins, which appeared strongest for mutations affecting the CA-NC border [25,26,33]. Mutations spanning the CA-SP1 regions also had a trans-dominant negative effect on processing of the respective wt protein when co-assembled into particles, suggesting that the CA-SP1 region forms a stable multimeric structure [26].

Here we set out to elucidate and describe the stepwise transitions occurring during HIV-1 maturation on a structural level. By doing so we also aimed to shed light on the structure and integrity of the Gag lattice by understanding structural changes associated with specific cleavage events. To do this we have applied cryo-electron tomography and sub-tomogram averaging to describe the morphology and structure of an extended pool of cleavage site mutants.

Results

Virus production and cleavage efficiency

We analyzed a panel of viruses derived from HIV-1 NL43, which carry point mutations at single or multiple of the cleavage sites in Gag [24,27,28] to prevent proteolytic cleavage at the targeted sites. In parallel we transfected a wt proviral plasmid in the presence or absence of the PR inhibitor Lopinavir to generate immature and mature viruses, respectively. The constructs used are shown in Figure 1B. Based on the scheme shown in Figure 1A [22], MA-p6 (together with PR inhibited virus), MA-SP1, CA-SP1 (CA5 in [27]) and the wt virus represent true maturation products. The other constructs CA-p6, and MA-CA were selected to shed light on the importance of individual cleavage sites upstream and downstream of CA.

HIV-1 wt and all variant viruses were produced by transfection of 293T cells, and purified by gradient centrifugation as described previously [10]. We assessed the degree of proteolytic processing via SDS PAGE followed by silver staining or immunoblotting (Figure 2A and data not shown). In most cases all predicted cleavage products were seen, indicating that proteolytic cleavage at non-mutated sites proceeded to completion, while mutated sites were blocked. The exception was MA-CA carrying a single mutation at the cleavage site between MA and CA, but exhibiting strongly reduced cleavage efficiency at the CA-SP1 boundary of only 20%–50% depending on the virus preparation. Failure to cleave at MA-CA therefore leads to a reduction in the efficiency of cleavage at the CA-SP1 boundary.

3D morphology of the viruses

In order to characterize the effect of cleavage site mutations on structure and morphology of the viruses, purified viruses were
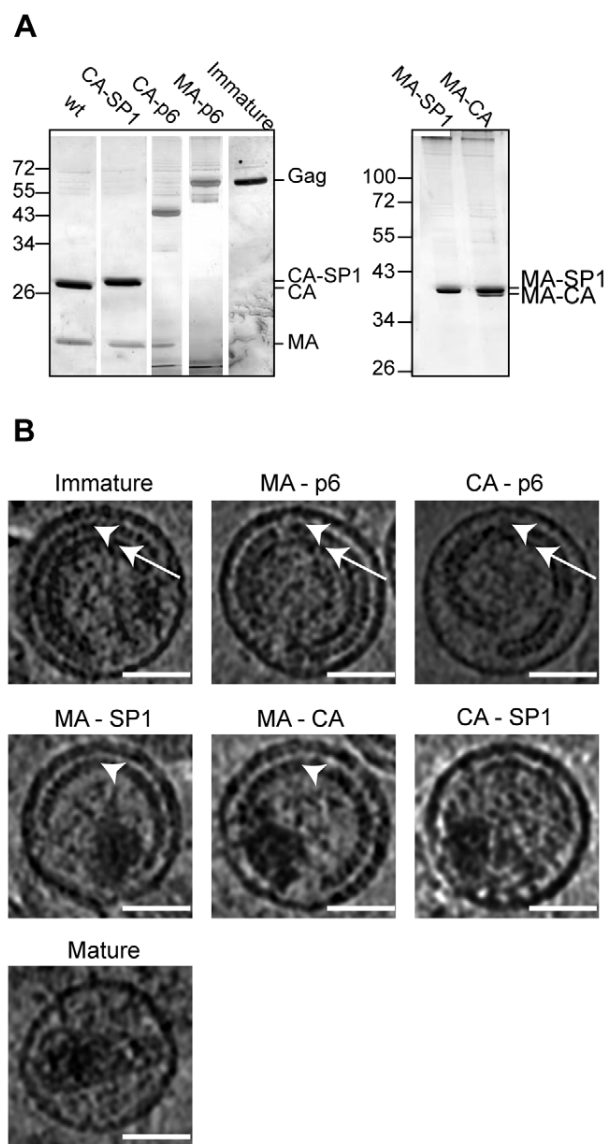
Figure 2. Characterisation of HIV-1 variants. A) Partially processed Gag-derived products detected by SDS-PAGE. Iodixanol gradient purified particle preparations of the indicated HIV-1 variants were separated on 12.5% SDS gels. Virion-associated proteins were visualized by silver staining according to standard procedures, showing MA and CA containing Gag derivatives as the most prominent bands. Lanes in the left panel were combined from two independent gels and were size adjusted for comparison. The right panel illustrates incomplete processing at the CA-SP1 boundary for variant MA-CA. Positions of molecular mass standards (in kDa) are indicated to the left. B) Central sections of 0.8 nm thickness from 6 µm defocus tomographic reconstructions of the HIV-1 variants analyzed. White arrowheads point to the immature CA layer, whereas white arrows point to the NC-RNA layer. Immature, MA-p6 and CA-p6 variants show both CA and NC-RNA layers. In CA-p6 the distance between the CA layer and the membrane is more variable. In MA-SP1 and MA-CA, the CA layer is seen, but the NC-RNA has condensed. CA-SP1 and the mature virus do not show the characteristic striated CA layer. The scale bar is 50 nm.
doi:10.1371/journal.ppat.1001215.g002

plunge-frozen in liquid ethane and imaged in the electron microscope. To eliminate morphological variability resulting from transfection or purification, a minimum of 3 independent

preparations of each virus were imaged. With the exception of the CA-SP1 construct (see below), independent preparations of the same variant showed no significant differences in morphology. In order to obtain three-dimensional structural information, cryo-electron tomograms were acquired. A minimum of 46 virus particles were reconstructed in 3D for each variant, in each case distributed across all 3 preparations.

Figure 2B shows central sections through representative tomograms of viruses from each of the variants (see also Figure S1). Both the PR inactivated wt virus and the MA-p6 variant show the characteristic morphology, which has previously been described for immature HIV imaged by cryo-electron microscopy and cryo-electron tomography [10,14,34]. These virus particles are approximately spherical and display an incomplete protein lattice underneath the viral membrane. The membrane appears thicker in regions where Gag is present, due to the presence of the membrane associated MA domain in these regions [10]. The protein lattice shows a striated pattern in the outer ring of density, previously assigned to the CA domain (Figure 2B, arrowheads), and a smoother inner ring of density, previously assigned to the NC-RNA complex (Figure 2B, arrows) [34]. The similarity of the PR-inhibited virus and the MA-p6 variant indicates that there are no substantial morphological differences between a virus which has an active PR, but mutations in all proteolytic cleavage sites in Gag, and a virus with intact cleavage sites but an inhibited PR.

The characteristic two layers of density corresponding to the CA and NC-RNA domains were also seen in the CA-p6 virus; however, the distance between these layers and the membrane was significantly more variable than in the immature virus (compare CA-p6 with immature in Figure 2B), leading to greater variability in the curvature of the layer.

MA-SP1 and MA-CA showed a thinner Gag layer. The striated CA layer of density was present, but the inner NC-RNA layer of density that is seen in the immature virus was missing. Cleavage between CA and NC therefore leads to loss of the NC-RNA layer. Instead we observed a globular dense structure within the virus that is likely to represent condensed NC-RNA. The volume of the condensed NC-RNA was measured by visual assessment of the boundaries of the condensed region. In MA-SP1 it was $4.4 \pm 1.1 *$ 10^4 nm^3 (n = 46), and in MA-CA was $4.3 \pm 1.2 * 10^4$ nm^3 (n = 46), corresponding to a ratio between the volume of the virus particle and the volume of the condensed NC-RNA of 26.9 ± 8.5 for MA-SP1 and 28.0 ± 8.2 for MA-CA and thus indicating no detectable difference in the degree of condensation.

The CA-SP1 viruses presented high variability in the morphology depending on the preparation. Of the 5 independent preparations, 2 showed no internal structures, whereas the other 3 showed features, and were further analysed. In contrast to what was observed for the other mutated variants, the striated CA layer was not present in any of the particles, while a discontinuous and thin layer was visible in ~45% of the viruses (29 of 64 particles). The NC-RNA was condensed to the same degree as in MA-SP1 and MA-CA (volume $4.3 \pm 1.0 * 10^4$ nm^3 (n = 46), ratio of virus volume to NC-RNA volume 26.4 ± 7.2). The membrane appears thinner, similar to the regions of the immature virus where Gag is absent.

The mature virus showed the same morphology as previously described by tomography of equivalent preparations [35]. Virus particles contained predominantly cone shaped cores consisting of a thin CA lattice containing density which we attribute to the NC-RNA. The NC-RNA volume was $3.5 \pm 0.8 * 10^4$ nm^3 (n = 46), corresponding to a ratio between the volume of the virus particle and the volume of the condensed NC-RNA of 35.0 ± 8.9. The slightly lower NC-RNA volume in the mature virus when compared to MA-SP1, MA-CA and CA-SP1 probably reflects the better definition of the NC-RNA edges when confined within the core, rather than a change in the degree of condensation. The membrane thickness appears similar to that in the CA-SP1 viruses.

Global arrangement of the Gag lattice

To describe the changes in the global arrangement of the Gag lattices in greater detail, we applied subtomogram averaging to the cryo-electron tomography data for those variants which showed a striated lattice (the thinner CA layer in CA-SP1 and wt proved insufficiently featured for subtomogram averaging). By plotting in 3D the position of the subtomograms after the alignments, we were able to define the position and the orientation of the centre of each identified hexagon of the lattice. All the lattice maps showed the same general arrangement of a continuous hexameric lattice with irregularly shaped defects, partially covering the inner surface of the viral membrane, but leaving one large gap (Figure 3). There was no substantial difference in the degree of shell completeness between the different variants.

Structure of the Gag lattice by sub-tomogram averaging

To describe the changes in the Gag lattice structures during the maturation process in greater detail, we analysed the average structures generated during subtomogram averaging. These represent a 3D reconstruction of the local structure of the Gag lattice in each variant.
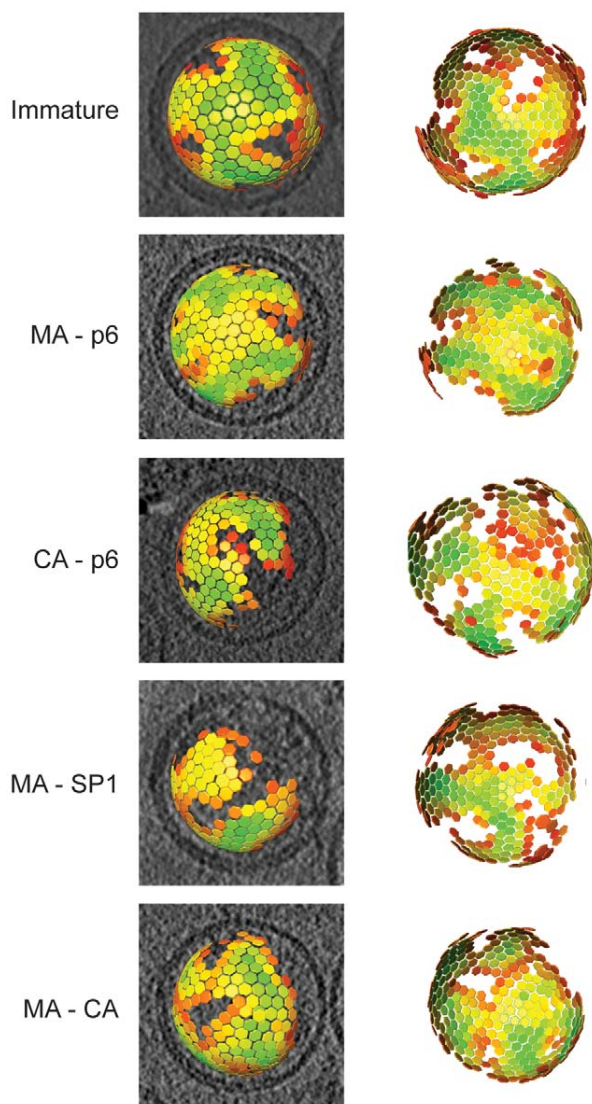
The structure of the Gag lattice in the PR inhibited virus (Figure 4) was consistent with that previously published for immature HIV using these methods [10]. The lattice had a six-fold symmetric structure with an inter-hexagon distance of 8 nm measured at the C-terminal CA domain. The outermost density, previously assigned to the N-terminal domain of CA, formed hexameric rings around large holes. The C-terminal domain of CA sat at a lower radius at the 2-fold symmetrical position between the hexameric rings. These rings were linked at the 6-fold positions, beneath the holes in the N-terminal CA layer, where they joined a rod-like density, previously proposed to represent SP1 descending into the NC-RNA layer, which was not hexagonally ordered. At this resolution, the structure of MA–p6 was indistinguishable from the PR-inhibited wt virus.

The structure of the CA–p6 lattice showed that the outer hexameric rings corresponding to N-CA, the 2-fold densities corresponding to C-CA and the rod-like densities were all unchanged. The inter-hexagon distance remained 8 nm. These observations imply that the CA, SP1 and NC-RNA layers in the CA-p6 variant adopt a structure that is similar to that in the immature virus.

The N-CA and C-CA features of the MA-SP1 structure were similar to the PR-inhibited virus. The inter-hexagon distance remained 8 nm. The rod-like densities that descend from the CA layer towards the NC-RNA could also still be seen in this structure. Below the rod-like densities, the reconstruction of the lattice of the MA-SP1 variant lacked the inner disordered layer that in the immature virus has been assigned to the NC-RNA. The structure of the lattice in MA-CA was identical to the one in MA-SP1, despite the partial cleavage of the SP1 region.

A link between partially cleaved Gag and the NC-RNA

All preparations contained some partially broken virus particles. Condensed NC-RNA, and sections of Gag lattice, which were not part of intact virus particles, could be visualised in the tomograms. We noticed that sections of MA-SP1 lattice from broken particles were commonly associated with condensed NC-RNA (data not

92

particles or the orientation made interpretation difficult, but in most of the cases, the size and shape of the density suggested that it represented the condensed NC-RNA.

Discussion

Here we have studied HIV-1 maturation by using cryo-electron tomography of a panel of viruses with mutations in proteolytic cleavage sites to describe the structures of potential maturation intermediates at macromolecular resolution. Analysis of these viruses reveals how individual cleavage events influence the structure and disassembly of the immature lattice as well as the formation of the mature lattice.
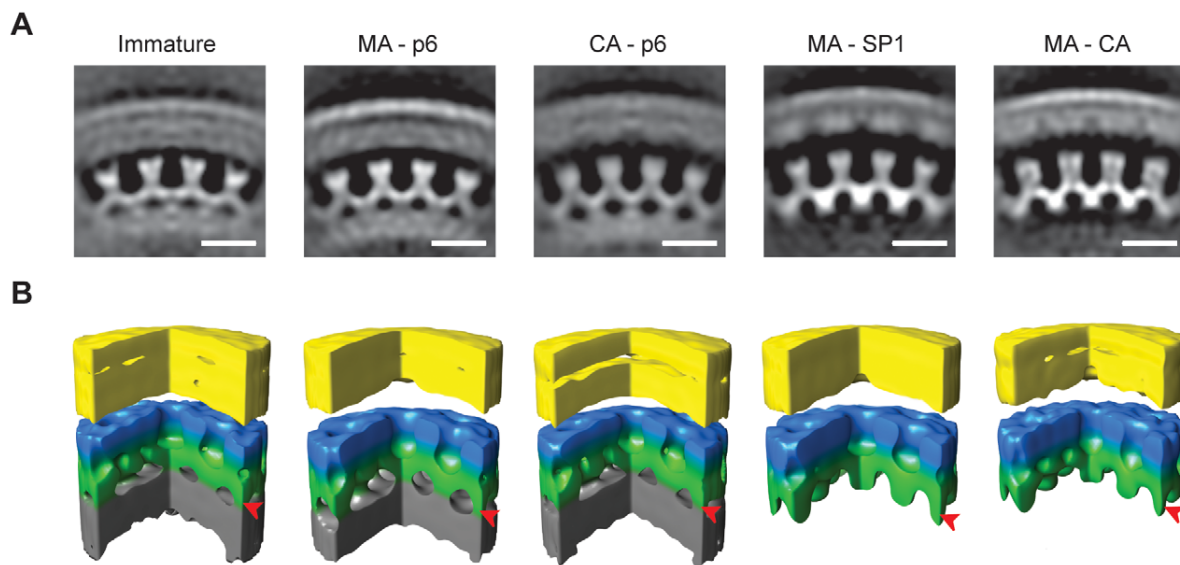
MA-p6 viruses lacking functional cleavage sites in Gag, but exhibiting wt PR activity and regular processing of Pol, present a lattice structure that is indistinguishable from immature HIV-1 produced in the presence of a PR inhibitor. The completeness of the Gag shell and the size of the viruses is also the same as for protease inhibited wt viruses. This suggests that (i) processing of the Pol domain does not detectably influence the lattice structure if Gag remains uncleaved, (ii) the alteration of cleavage sites did not have a significant effect on the structure and assembly of the Gag layer. Further analysis of variants with mutated cleavage sites revealed that processing at both sides of CA-SP1 is needed for dissolution of the immature lattice, while release of SP1 is needed, in addition, for formation of the regular mature core.

MA-SP1 forms a structure which is stable in the absence of NC

The structure of the MA-SP1 variant shows that cleavage between SP1 and NC leads to removal of the NC-RNA layer from Gag, and to condensation of the NC-RNA core. On the other hand, the structure of the MA-SP1 region of the lattice remained unchanged compared to the immature virus. NC-RNA interactions or other non-specific interaction domains downstream of SP1 have been shown to be essential for virus-like particle assembly both in vitro and in tissue culture, however [36]. The stability of the lattice in MA-SP1 indicates that these interactions concentrate and align Gag at the membrane during assembly, but are dispensable for maintaining the structure or stability of the Gag lattice once it is formed.

The same structure as is observed here for MA-CA and MA-SP1 was also seen in a subset of cellular budding sites and released particles in vivo, (see accompanying manuscript by Carlson *et al*). This structure can therefore also be formed by wt Gag if premature SP1-NC cleavage occurs and leads to loss of PR from the budding particle. Observing the same structure also in particles carrying wt Gag polyproteins provides further support for it representing a true maturation intermediate.

In the MA-CA construct, all cleavage sites downstream of CA remain in their wt form. Nevertheless, cleavage efficiency for the CA-SP1 site is reduced to between 20% and 50%, due to the failure of cleavage between MA and CA. These observations suggest that failure to cleave upstream of CA inhibits access of PR to the CA-SP1 cleavage site as has also been reported for mutation of the downstream SP1-NC cleavage site [37]. It has been shown that cleavage at the wt CA-SP1 site is also inhibited in a *trans*-dominant way if small amounts of MA-SP1 are incorporated into HIV-1 particles [26]. Our data indicate that in all of these situations there is a failure to sufficiently disassemble the immature CA lattice (see below), and that disassembly of the immature CA layer is required for efficient access of PR to the CA-SP1 cleavage site.

Figure 3. The global arrangement of the Gag layer. Global lattice maps for each variant superimposed on central sections of the tomographic reconstruction of the virus (left panels), or viewed from the direction of the largest gap in the lattice (right panels). The centres of each hexameric unit cell are marked with hexamers, which are coloured according to cross correlation on a scale from low (red) to high (green). Higher cross correlation values indicate that the subtomogram is more similar to the average structure. The cross-correlation range in each map has been set between the minimum and the maximum cross-correlation value present in the map. Maps are shown in perspective such that hexamers on the rear surface of the particle appear smaller.
doi:10.1371/journal.ppat.1001215.g003

shown), suggesting the possibility of an interaction between the condensed NC-RNA and the MA-SP1 lattice.

To further explore the possibility of a direct link between the condensed NC-RNA and the remaining Gag lattice, we gently disrupted MA-SP1 and MA-CA virus particles by vortexing for 15 s in the presence of 10 nm gold beads. Cryo-electron micrographs of these preparations showed large numbers of virus particles with disrupted membranes. All of the broken particles had condensed densities associated with the remaining MA-SP1 lattice (Figure 5). In some cases the presence of neighbouring

Figure 4. The local structure of the Gag layer. A) Radial sections from the subtomogram average reconstructions coming from each variant. Density is white. The scale bar is 10 nm. B) Surface rendering of the subtomogram average reconstructions. A sector of 90° has been cut out to better show the organization of the lattice. The colours on the surfaces are for illustrative purposes. The membrane and the MA layer, which cannot be delineated at this resolution, are yellow. N-terminal CA domain is blue. C-terminal CA domain and SP1 spacer are green. NC-RNA layer is grey. In MA-SP1 and MA-CA the grey NC-RNA layer is missing because condensation has already taken place, and rod-like protrusions are seen descending from the C-terminal CA domain towards the centre of the virus particle. Arrowheads mark the radial position of the rod-like densities.
doi:10.1371/journal.ppat.1001215.g004

Both MA-CA and MA-SP1 variants showed rod-like structures descending from the CA layer towards the centre of the virus. These rod-like structures were not disrupted by cleavage between SP1 and NC, or by subsequent partial cleavage between CA and SP1. This argues that SP1-SP1 interactions together with CA-SP1 interactions are sufficient to maintain the integrity of the rod-like structures even in the presence of partial CA-SP1 cleavage. SP1 contains important assembly determinants, and this observation

Figure 5. Analysis of disrupted virus particles. Micrographs of MA – SP1 (A) and MA – CA (B) virus particles after disruption by vortexing in the presence of gold beads show patches of viral membrane with underlying CA lattice. The condensed NC – RNA is visible as an irregular, globular density similar to that seen within intact virus particles (compare with figure 2B). Rather than diffusing away from the broken particles, the NC-RNA appears to remain associated with the CA lattice. The scale bar is 50 nm.
doi:10.1371/journal.ppat.1001215.g005

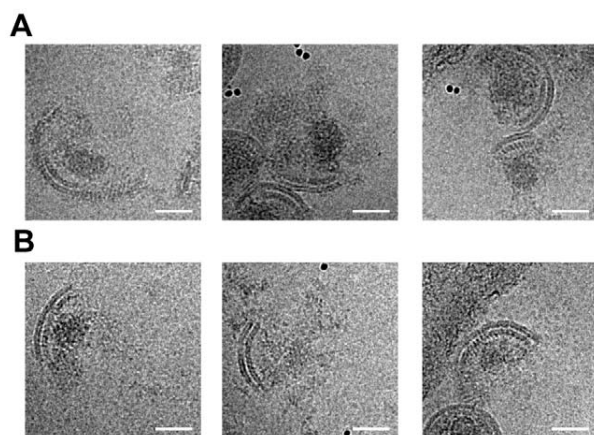suggests that these direct formation of a specific structure, which does not require NC tethering for its stability.

CA-NC forms a structure which is stable in the absence of MA

In solution, cleavage between MA and CA has been demonstrated to cause rearrangement in the N-terminal domain of CA, leading to part of the MA-CA flexible-linker folding back into the N-terminal domain of CA to form a β-hairpin structure [20]. However in the CA-p6 variant, where this cleavage has occurred, the CA lattice was identical to the one present in the immature virus, at the resolution of the reconstructions. It also had the same overall completeness. This suggests that formation of the β-hairpin structure either requires downstream processing in addition, or that its formation is insufficient to cause structural changes in the immature lattice.

Although the structure between CA and p6 in the CA-p6 variant was the same as that in the immature virus, the distance between CA and the membrane became more variable, indicating the absence of a MA-CA link. The curvature of the CA layer was more variable than in the immature virus, suggesting that interactions with the membrane contribute to the uniformity of the lattice curvature.

Cleavage upstream and downstream of CA-SP1 is required for immature lattice disassembly: cleavage within CA-SP1 for mature lattice assembly

As discussed above, both MA-SP1 and CA-p6 show CA lattices with an immature structure. We can therefore conclude that cleavage either only upstream or only downstream of the CA-SP1 domain does not affect the organization of the immature lattice. The immature lattice is disassembled, however, when cleavage has occurred both upstream of CA and downstream of SP1 in the case of the CA-SP1 variant, which contains no characteristic, striated

CA lattice. In thin section electron micrographs, an irregular electron-dense structure separated from the virion membrane was commonly observed for this virus [27]. Such irregular layers were occasionally seen in our analyses as well, but were too thin to represent an immature CA lattice. They are more comparable in thickness to the mature CA core, which shows a thinner CA layer than the immature virus due to the closer proximity of the N- and C- terminal CA domains. Inhibition of cleavage at the CA-SP1 cleavage site therefore does not appear to prevent disassembly of the immature lattice, but to affect formation of the regular mature core.

CA-SP1 and even CA-NC constructs are able to assemble in vitro into mature-like lattices [13,21,36], indicating that there is no absolute requirement for cleavage at the CA-SP1 junction for mature CA lattice assembly. Indeed the thin layers sometimes seen in CA-SP1 particles may represent patches of mature CA lattice, but formation of the closed mature core around the NC-RNA complex does not occur unless SP1 is removed from CA. It is currently not clear why SP1 cleavage is needed for formation of the mature core inside the virion, while being dispensable for in vitro assembly of the mature-like lattice. We speculate that this may be due to a defect in nucleating formation of the cone-shaped mature core in the virion if SP1 is not cleaved. It is interesting, in this context, that in vitro assembly of CA or C-terminally extended versions of CA generally leads to hollow tubes, and formation of a closed cone is observed only very rarely. The in vitro system may thus not completely mimic this aspect of virion maturation. Cleavage between CA and SP1 is the site of action of the maturation inhibitor bevirimat, suggesting that this drug has no effect on disassembly of the immature lattice but acts by preventing formation of the closed cone-shaped core of the mature virion.

A link between the condensed NC-RNA and the immature CA lattice

In the immature virus the NC-RNA is organized in a disordered but compact layer below the CA lattice. When Gag is cleaved downstream of SP1, the NC region condenses. Approximately the same degree of condensation was seen in the MA-SP1, MA-CA, CA-SP1 and wt variants, suggesting that there is no role for free CA in NC-RNA condensation. The high variability in shape and volume (if related to the size of the virus) indicates that the condensed NC-RNA does not form a well-defined structure.

MA-SP1 and MA-CA exhibited a condensed NC-RNA, as well as an intact immature CA layer. Surprisingly, broken MA-SP1 and MA-CA particles exhibited associated density consistent with NC-RNA. The size and shape of this density and the absence of any other strong candidates in the purified virus preparation suggests to us that this density is indeed NC-RNA. We expect that diffusion of the NC-RNA away from the protein shell would have occurred during the vortexing procedure used to disrupt the particles, or within the time taken to transfer the sample to the grid and prepare it for electron microscopy, if no physical linkage was present. These observations therefore suggest that the NC-RNA remains attached to the CA lattice, suggesting the presence of an interaction or link between the partially cleaved lattice and the condensed NC-RNA in these variants.

If a non-covalent interaction is present, it could represent a non-specific interaction between the NC-RNA complex and the CA lattice. The interaction could also be specific: a cellular or viral protein present within the CA lattice may expose a defined binding site for a component of the NC-RNA complex. A more attractive hypothesis is that there is a direct link between NC-RNA and CA in the form of uncleaved Gag at this position. Such a link

could be locally maintained if, for example, binding of the RNA packaging signal to NC protects the SP1-NC cleavage site such that it is not accessible to the viral PR. It is tempting to speculate that maintenance of such a link during maturation may allow the NC-RNA to function as a template for formation of the conical core, and/or to ensure that the NC-RNA is packaged within the core.

In summary, our data suggest that cleavages at both sides of the CA-SP1 module are required to disassemble the immature Gag lattice with removal of the NC-p6 region initiating the independent condensation of the NC-RNA core, still connected with the membrane bound lattice. Subsequent cleavage between MA and CA causes dissociation of the immature lattice, but the resulting CA-SP1 module requires final cleavage of the SP1 peptide to permit assembly of the mature cone.

Methods

Constructs and virus preparation

Derivatives of proviral plasmid pNL4-3 carrying mutations at specific PR cleavage sites within Gag, as well as the PR defective variant NL4-3D25A have been described previously [26,38].

293T cells were maintained in DMEM with 10% fetal calf serum and antibiotics. Transfections with the indicated proviral derivatives were performed using the calcium phosphate method. Culture media were harvested at 42 h post transfection cleared by low speed centrifugation (5 min, 1500 g) followed by filtration through 0.45 µM nitrocellulose filters. Particles were purified by centrifugation through a 20% (w/w) sucrose cushion and subsequent centrifugation on an Iodixanol gradient as described [39]. Purified virus was inactivated with 1% paraformaldehyde for 1 h on ice. Successful inactivation was confirmed by infection of C8166 cells and scoring for syncytia formation up to 10 days after inoculation. At least 3 independent particle preparations were analyzed for each variant shown with no variability in the structures with the exception of the CA-SP1.

Sample preparation and data acquisition

Purified viruses were mixed with 10 nm colloidal gold particles, deposited on C-flat holey Carbon grids, and vitrified by plunge-freezing in liquid ethane. Tilt series were collected on an FEI Tecnai F30 "Polara" transmission electron microscope with Gatan GIF 2002 post column energy filter and 2 k×2 k Multiscan CCD camera. Data collection was performed at 300 kV using the SerialEM Software. Tilt series were collected between 60° and −60° with 3° angular increment; the total electron dose applied to the tomograms was approximately 90 e/Å^2. Tomograms were acquired at defocuses between 2.6 and 6.0 µm, with a magnification of 34000X resulting in a pixel size at the specimen level of 4.0 Å.

Image processing

Tomograms were reconstructed using IMOD [40]. Subtomogram averaging was carried out as described in [10] and below using MATLAB (Mathworks). Lattice map representations were generated using Amira (Visage Imaging), together with the EM Package [41]. A hexagon is placed at the final aligned position of each tomogram, and coloured according to the cross-correlation value between the sub-tomogram and the average. Tomograms aligned to an inappropriate radial position were excluded based on the radii distribution of the set of 20 nearest subtomograms surrounding each subtomogram. If the radius of the selected subtomogram was in the first or in the fourth quartile of the distribution the subtomogram was excluded. Only hexagons with a

cross-correlation over a defined threshold are displayed. The threshold was appropriately set such that hexagons were not displayed where density corresponding to Gag was absent in the tomogram. Density maps were displayed using Amira, surface rendering was done using UCSF Chimera [42].

Subtomogram averaging

For subtomogram averaging, tomograms collected at defocuses between 2.6 and 2.9 μm were used. Sub-volumes of $(38.3 \text{ nm})^3$ were extracted from tomograms along the surface of a sphere centred in the centre of the virus and with a radius equal to the mean radius at CA level. The sub-volumes were iteratively aligned. The initial reference used for the alignment was the average of the subtomograms in the extraction position. All variants subjected to sub-tomogram averaging show 6-fold symmetry, as evidenced using radius-angle-frequency plots (see Text S1 and Figure S2). 6-fold symmetry was therefore applied to the average after all iterations. The threshold for the subtomograms to be averaged was set to the mean cross-correlation value between all subtomograms and the reference. The final reconstructions had a resolution according to the Fourier shell correlation with a 0.5 criterion of approximately 26.5 Å in the CA region (Figure S3A) and are filtered to this resolution. The resolution varies with radius, with the highest resolution in the CA region (Figure S3B).

Supporting Information

Text S1 Supplementary methods.
Found at: doi:10.1371/journal.ppat.1001215.s001 (0.05 MB PDF)

Figure S1 The morphology of virus variants. Central sections of tomographic reconstructions of the HIV-1 variants analyzed acquired at different defocuses (df), and coming from 3 different preparations, to illustrate consistency of virus morphology between preparations. A gaussian filter was applied the tomograms (8 kernel, 0.4 sigma). The scale bar is 50 nm.
Found at: doi:10.1371/journal.ppat.1001215.s002 (5.19 MB TIF)

Figure S2 Radius-angle-frequency plots. A) 3D view of the radius-angle-frequency plot calculated from the immature HIV data, illustrating the relationship between the three axes. Three perpendicular sections are shown intersecting at the point: 53 nm

radius, 60° angle and 7 nm frequency. The colour bar represents the value of autocorrelation and is common to all the panels in the figure. The presence of a peak at a particular point in radius, frequency and angle indicates that at that radius in the virus, the 2D power spectrum of the protein layer has peaks at that frequency which are arranged rotationally symmetrically repeating at that angle (see supplementary methods). B–F) Data from 5 variants. The top panel is a section at the radius where the C-CA domain is found showing two peaks at 60° and 120°, with 7 nm frequency, as expected from a hexagonal unit cell with 8 nm spacing (see supplementary methods). The peaks at 0 and 180° are seen in all 2D power spectra since power spectra have intrinsic 2-fold symmetry. The middle panel is a section at 60° angle that shows that the 7 nm peak is extended across the CA region. The third panel is the radial density profile of the virus in the regions containing Gag. Starting outside the virus (high radius) the first two peaks, typically between 55 and 65 nm, represent the two leaflets of the bilayer and the associated MA, the next two peaks, typically between 45 and 55 nm, represent CA, and the peak below 45 nm, where present, represents the NC - RNA.
Found at: doi:10.1371/journal.ppat.1001215.s003 (2.11 MB TIF)

Figure S3 Fourier shell correlation. A) Fourier shell correlation plots for all the variants. The resolution was determined as the frequency at which the FSC curve drops below 0.5 correlation which is highlighted with the dashed line. B) Plot showing the variation in resolution according to radius. At each radius the resolution was determined by Fourier shell correlation at 0.5, with a mask centred at that radius (see supplementary methods). The positions of NC -RNA, CA, MA and membrane are indicated.
Found at: doi:10.1371/journal.ppat.1001215.s004 (0.89 MB TIF)

Author Contributions

Conceived and designed the experiments: AdM HGK JAGB. Performed the experiments: AdM BM BG JDR. Analyzed the data: AdM JAGB. Wrote the paper: AdM HGK JAGB.

References

1. Demirov DG, Freed EO (2004) Retrovirus budding. Virus Res 106: 87–102.
2. Morita E, Sundquist WI (2004) Retrovirus budding. Annu Rev Cell Dev Biol 20: 395–425.
3. Ganser-Pornillos BK, Yeager M, Sundquist W (2008) The structural biology of HIV assembly. Curr Opin Struct Biol 18: 203–217.
4. Coffin JM, Hughes SH, Varmus HE (1998) Retroviruses.
5. Hurley JH (2008) ESCRT complexes and the biogenesis of multivesicular bodies. Curr Opin Cell Biol 20: 4–11.
6. Williams RL, Urbe S (2007) The emerging shape of the ESCRT machinery. Nat Rev Mol Cell Biol 8: 355–368.
7. Aldovini A, Young RA (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.
8. Anderson J, Schiffer C, Lee SK, Swanstrom R (2009) Viral protease inhibitors. Handb Exp Pharmacol. pp 85–110.
9. Briggs JAG, Simon M, Gross I, Krausslich HG, Fuller S, et al. (2004) The stoichiometry of Gag protein in HIV-1. Nat Struct Mol Biol 11: 672–675.
10. Briggs JAG, Riches JD, Glass B, Bartonova V, Zanetti G, et al. (2009) Structure and assembly of immature HIV. Proc Natl Acad Sci USA 106: 11090–11095.
11. Ganser BK, Li S, Klishko V, Finch J, Sundquist W (1999) Assembly and analysis of conical models for the HIV-1 core. Science 283: 80–83.
12. Briggs JAG, Wilk T, Welker R, Krausslich HG, Fuller SD (2003) Structural organization of authentic, mature HIV-1 virions and cores. EMBO J 22: 1707–1715.
13. Li S, Hill C, Sundquist W, Finch J (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein. Nature 407: 409–413.
14. Wright E, Schooler J, Ding H, Kieffer C, Fillmore C, et al. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J 26: 2218–2226.
15. Ganser-Pornillos BK, Cheng A, Yeager M (2007) Structure of full-length HIV-1 CA: a model for the mature capsid lattice. Cell 131: 70–79.
16. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, et al. (2009) X-ray structures of the hexameric building block of the HIV capsid. Cell 137: 1282–1292.
17. von Schwedler UK, Stemmler TL, Klishko VY, Li S, Albertine KH, et al. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J 17: 1555–1568.
18. Kingston RL, Fitzon-Ostendorp T, Eisenmesser EZ, Schatz GW, Vogt VM, et al. (2000) Structure and self-association of the Rous sarcoma virus capsid protein. Structure 8: 617–628.
19. Cornilescu CC, Bouamr F, Yao X, Carter C, Tjandra N (2001) Structural analysis of the N-terminal domain of the human T-cell leukemia virus capsid protein. J Mol Biol 306: 783–797.
20. Tang C, Ndassa Y, Summers MF (2002) Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. Nat Struct Biol 9: 537–543.
21. Gross I, Hohenberg H, Huckhagel C, Kräusslich HG (1998) N-Terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. J Virol 72: 4798–4810.
22. Pettit SC, Moody MD, Wehbie RS, Kaplan AH, Nantermet PV, et al. (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: 8017–8027.

23. Kaplan AH, Manchester M, Swanstrom R (1994) The activity of the protease of human immunodeficiency virus type 1 is initiated at the membrane of infected cells before the release of viral proteins and is required for release to occur with maximum efficiency. J Virol 68: 6782–6786.

24. Coren LV, Thomas JA, Chertova E, Sowder RC, 2nd, Gagliardi TD, et al. (2007) Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. J Virol 81: 10047–10054.

25. Lee SK, Harris J, Swanstrom R (2009) A strongly transdominant mutation in the human immunodeficiency virus type 1 gag gene defines an Achilles heel in the virus life cycle. J Virol 83: 8536–8543.

26. Müller B, Anders M, Akiyama H, Welsch S, Glass B, et al. (2009) HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. J Biol Chem 284: 29692–29703.

27. Wiegers K, Rutter G, Kottler H, Tessmer U, Hohenberg H, et al. (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J Virol 72: 2846–2854.

28. Wyma DJ, Jiang J, Shi J, Zhou J, Lineberger JE, et al. (2004) Coupling of human immunodeficiency virus type 1 fusion to virion maturation: a novel role of the gp41 cytoplasmic tail. J Virol 78: 3429–3435.

29. Gottlinger HG, Sodroski JG, Haseltine WA (1989) Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A 86: 5781–5785.

30. Kräusslich HG, Fäcke M, Heuser AM, Konvalinka J, Zentgraf H (1995) The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. J Virol 69: 3407–3419.

31. Li F, Goila-Gaur R, Salzwedel K, Kilgore NR, Reddick M, et al. (2003) PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. Proc Natl Acad Sci USA 100: 13555–13560.

32. Zhou J, Yuan X, Dismuke D, Forshey BM, Lundquist C, et al. (2004) Small-molecule inhibition of human immunodeficiency virus type 1 replication by specific targeting of the final step of virion maturation. J Virol 78: 922–929.

33. Checkley MA, Luttge BG, Soheilian F, Nagashima K, Freed EO. The capsid-spacer peptide 1 Gag processing intermediate is a dominant-negative inhibitor of HIV-1 maturation. Virology 400: 137–144.

34. Wilk T, Gross I, Gowen B, Rutten T, de Haas F, et al. (2001) Organization of immature human immunodeficiency virus type 1. J Virol 75: 759–771.

35. Briggs JAG, Grunewald K, Glass B, Forster F, Krausslich HG, et al. (2006) The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. Structure 14: 15–20.

36. Campbell S, Vogt VM (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.

37. Pettit SC, Henderson GJ, Schiffer CA, Swanstrom R (2002) Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. J Virol 76: 10226–10233.

38. Konvalinka J, Litterst MA, Welker R, Kottler H, Rippmann F, et al. (1995) An active-site mutation in the human immunodeficiency virus type 1 proteinase (PR) causes reduced PR activity and loss of PR-mediated cytotoxicity without apparent effect on virus maturation and infectivity. J Virol 69: 7180–7186.

39. Dettenhofer M, Yu XF (1999) Highly purified human immunodeficiency virus type 1 reveals a virtual absence of Vif in virions. J Virol 73: 1460–1467.

40. Kremer JR, Mastronarde DN, McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. J Struct Biol 116: 71–76.

41. Pruggnaller S, Mayr M, Frangakis AS (2008) A visualization and segmentation toolbox for electron microscopy. Journal of Struct Biol 164: 161–165.

42. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF chimera - A visualization system for exploratory research and analysis. J Comp Chem 25: 1605–1612.

Text S1:

Radius-angle-frequency plots

The same subtomograms used in the subtomogram averaging were iteratively aligned allowing only translations in Z (no rotations or translations in other directions were allowed, no symmetry was applied). The starting reference used was the average of the subtomograms in the extraction position. A mask that included the CA layer and the membrane was applied during alignment. Only the subtomograms that contained Gag (to exclude those from regions of the surface where no Gag layer was present) were used for the analysis; this was done by considering only the subtomograms selected for the lattice maps.

Aligned subtomograms underwent a radial orthographic projection. The focus of the projection was determined as the geometric centre of the virus from which the subtomogram was extracted.

The resulting volumes corresponded to flattened subtomograms. 2D power spectra were calculated at each radius. The rotation-autocorrelation function with a rotation range between 0° and 180° was calculated for each power spectrum. For each subtomogram this generates a 3D plot with radius on one axis, and the other two axes representing those of the rotational autocorrelation function of the power spectrum, namely angle and frequency. This analysis is a 3D implementation and combination of the analyses presented in figures 3c and 5 of [1]. The 3D plots for all aligned, selected subtomograms were averaged to generate the final plot.

Radial density profiles

The flattened subtomograms used for the radius-angle-frequency plots were averaged, and the mean density value at each radius was plotted against radius.
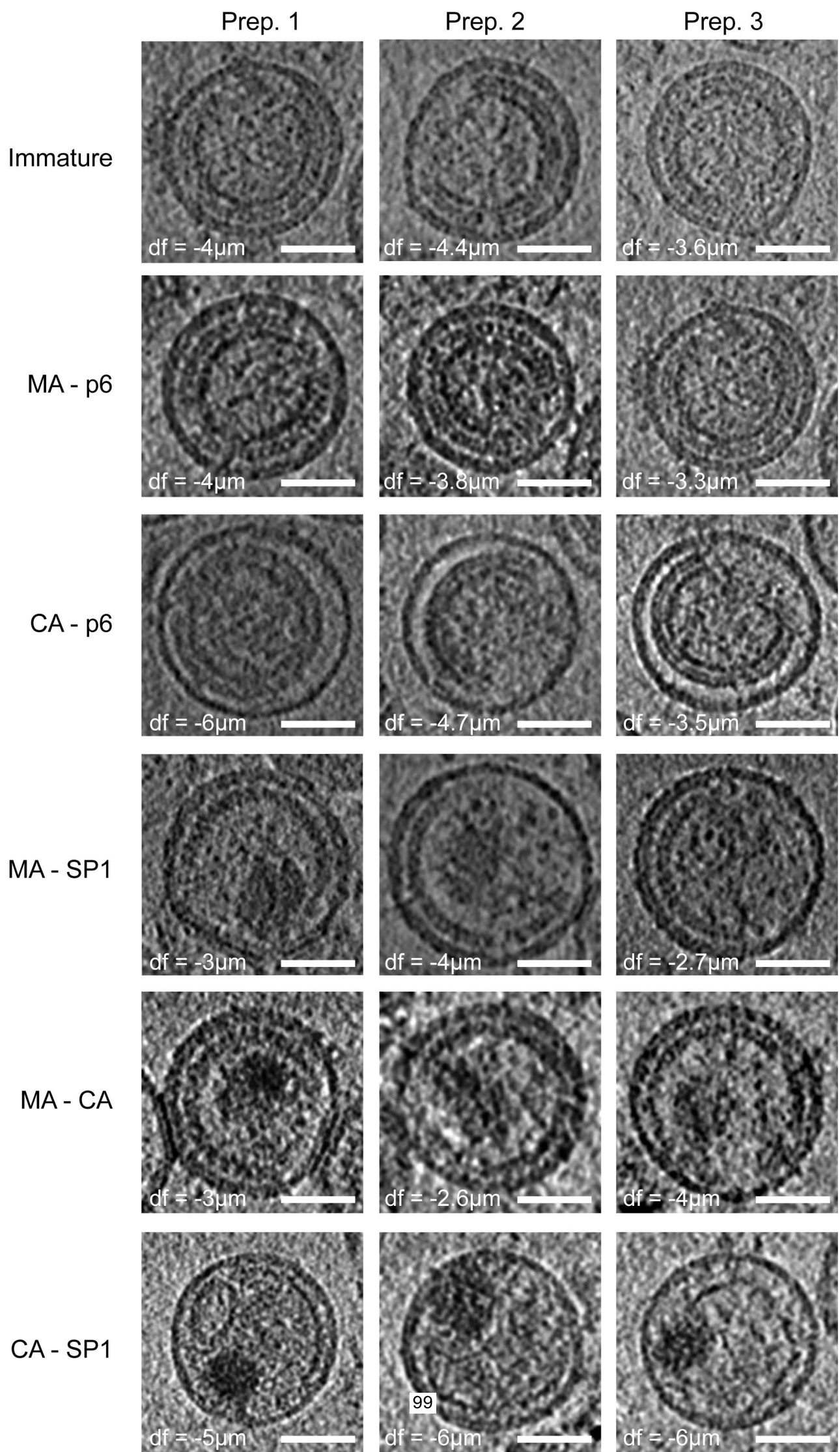
Fourier shell correlation (FSC)

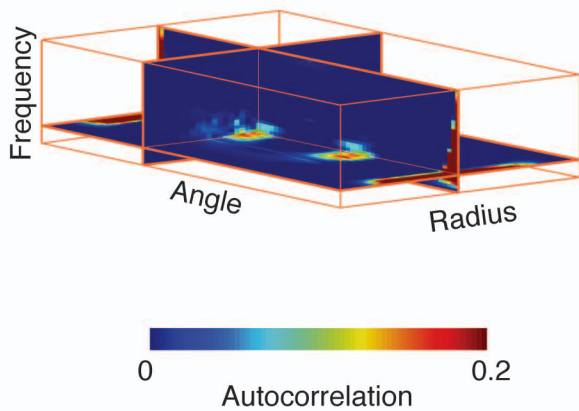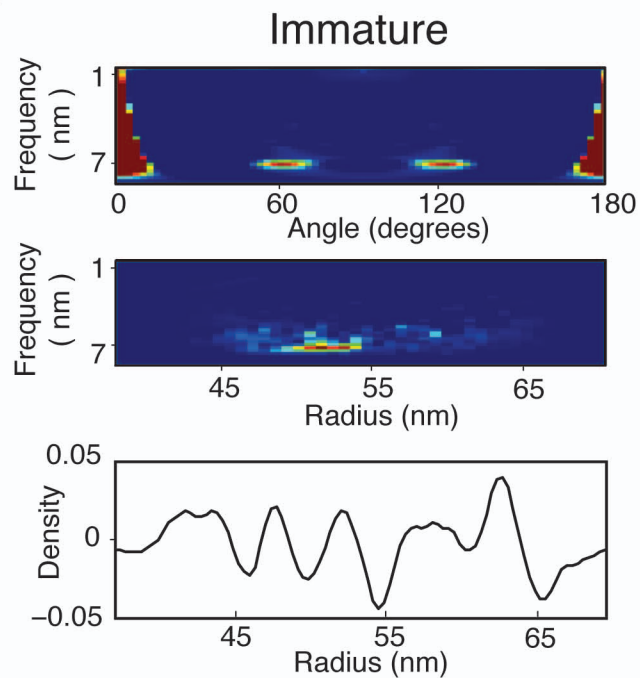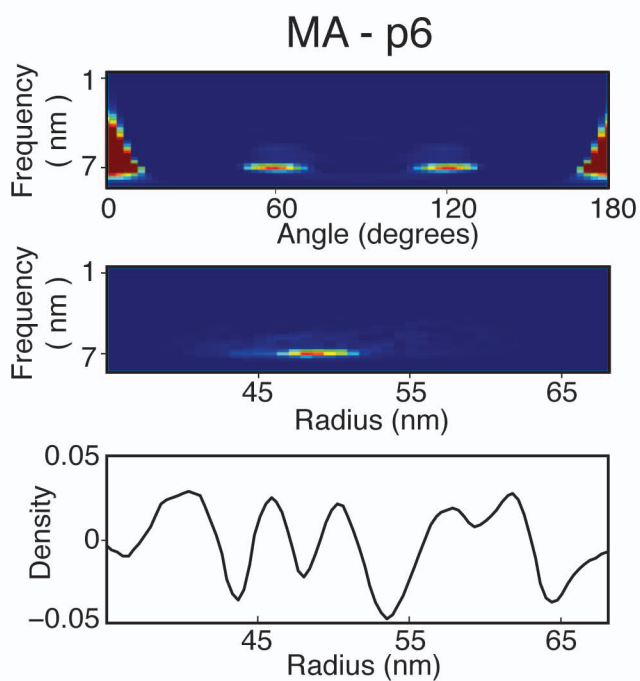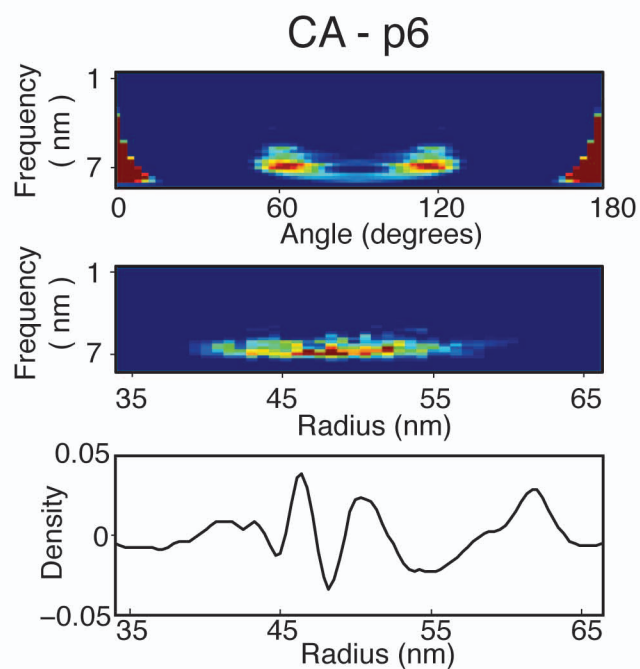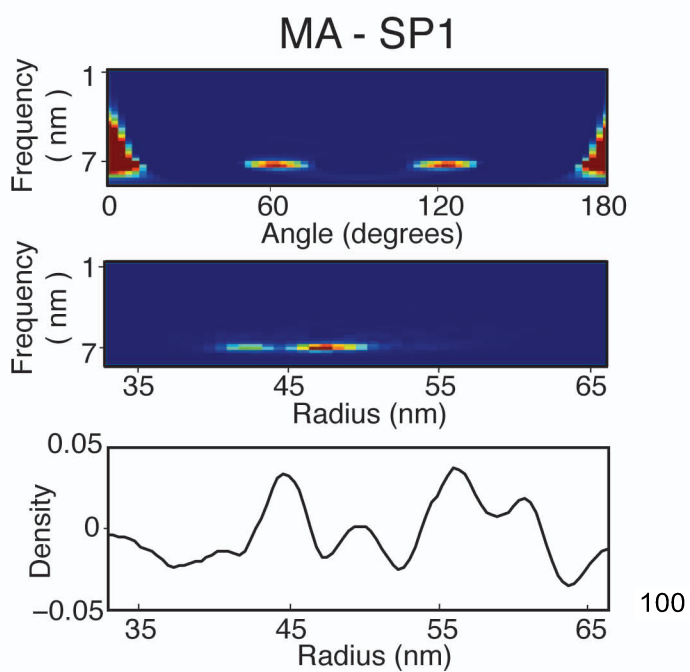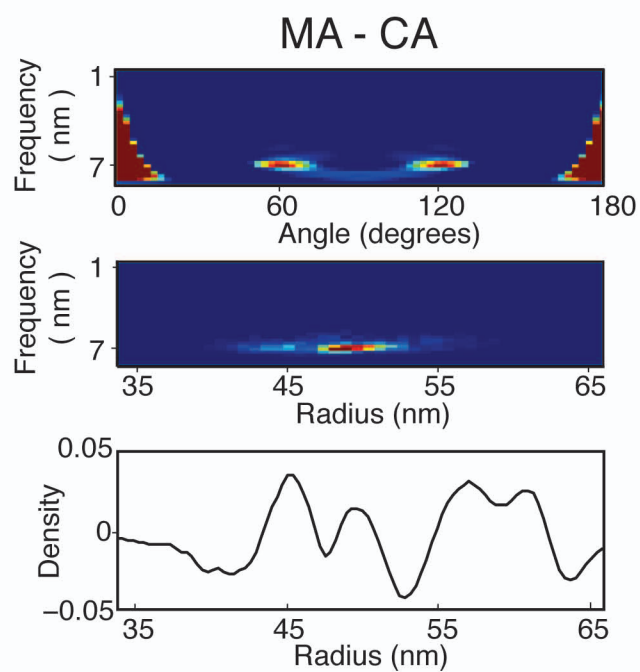The dataset was split two halves, and two averages where generated. To avoid the possibility of an artificial resolution measurement, where multiple subtomograms had shifted and rotated during the alignment to bring their overlapping regions together, all but one were discarded. Since the resolution has variability across the lattice the averages were masked with a soft mask that included only the CA region, and the FSC was calculated.

In order to define the variation of the resolution according to radius, the same mask used for the calculation of the resolution in CA was shifted by increments of 1 pixel, and the resolution at which the FSC curve dropped below 0.5 for each radius was plotted.

References

1. Briggs JA, Johnson MC, Simon MN, Fuller SD, Vogt VM (2006) Cryo-electron microscopy reveals conserved and divergent features of gag packing in immature particles of Rous sarcoma virus and human immunodeficiency virus. J Mol Biol 355: 157-168.

	Prep. 1	Prep. 2	Prep. 3
Immature	df = -4µm	df = -4.4µm	df = -3.6µm
MA - p6	df = -4µm	df = -3.8µm	df = -3.3µm
CA - p6	df = -6µm	df = -4.7µm	df = -3.5µm
MA - SP1	df = -3µm	df = -4µm	df = -2.7µm
MA - CA	df = -3µm	df = -2.6µm	df = -4µm
CA - SP1	df = -5µm	df = -6µm	df = -6µm

A

B
Immature

C
MA - p6

D
CA - p6

E
MA - SP1

F
MA - CA

100
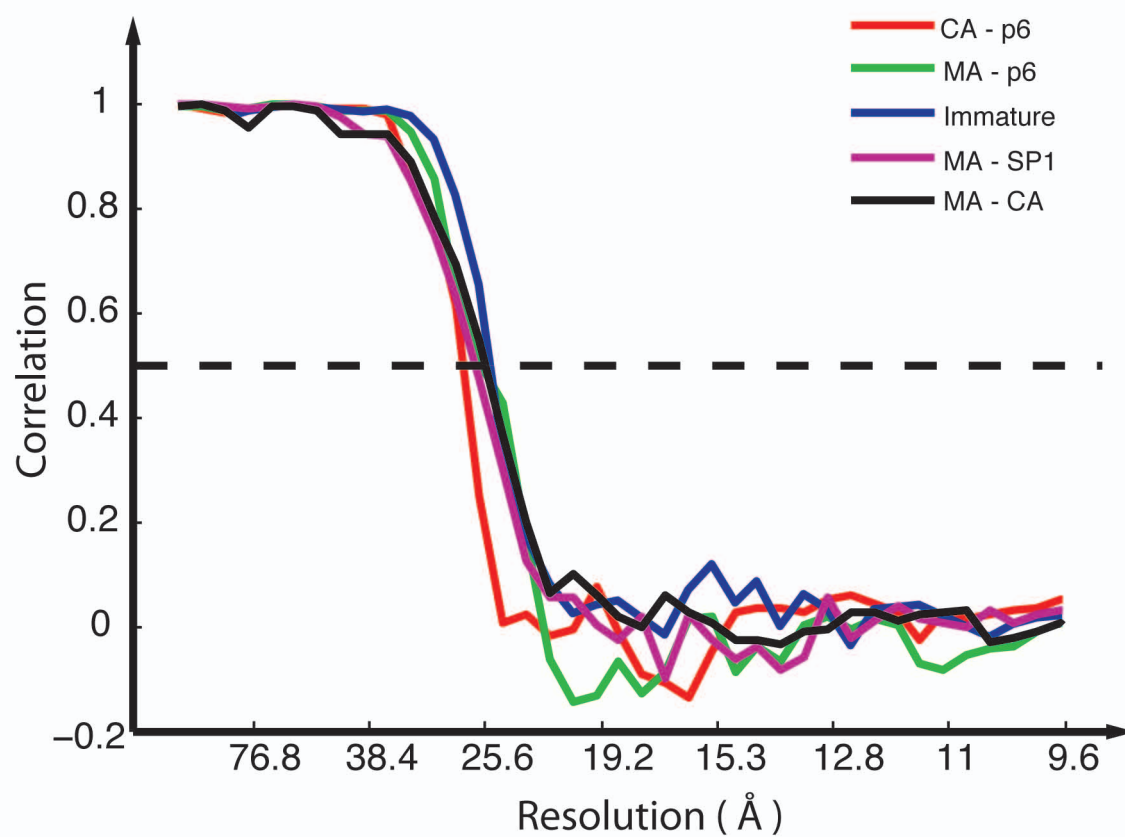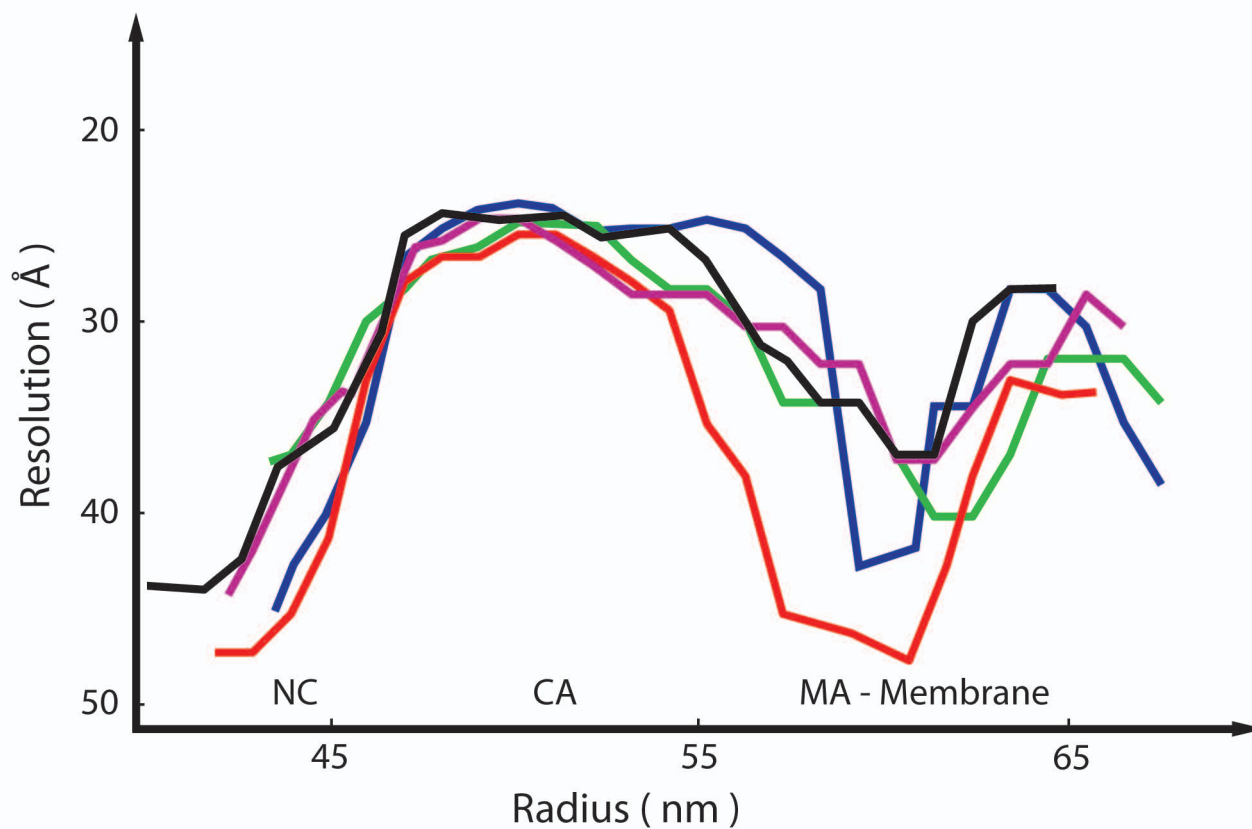
Title

The role of proteolytic cleavage around SP2 in the structural maturation of HIV-1

Authors

Alex de Marco[1], Barbara Mueller[2], Anke-Mareil Heuser[2], Baerbel Glass[2], Hans-Georg Kraeusslich[2]*, John A.G. Briggs[1]*

1- Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

2- Department of Infectious Diseases, Virology, Universitätsklinikum Heidelberg, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany

*Correspondence to:

JAGB, briggs@embl.de, Tel. +49 6221 3878482

HGK, Hans-Georg.Kraeusslich@med.uni-heidelberg.de, Tel. +49 6221 565001

Abstract

The main structural element in HIV-1 is Gag, a polyprotein that from N- to C-terminus comprises a membrane binding MA domain, structural CA domain, spacer peptide 1, genome binding NC domain, spacer peptide 2, and the C-terminal peptide p6. In the immature virus particle Gag is arranged radially, with its N-terminus at the membrane and its C-terminus pointed towards the centre of the particle. During the maturation Gag is cleaved in five positions by the viral protease to separate the domains and peptides, leading to a dramatic change in the internal organization of the virus particle. The NC-RNA complex forms a condensed globular structure in the centre of the mature, where a cone shaped CA capsid core surrounds it. Here we have used biochemistry and cryo-electron tomography to describe the functional role of the proteolytic cleavages between NC and p6 for maturation and infectivity, as well as the role of spacer peptide 2. Surprisingly we find that correct proteolytic processing between NC and p6 is required for proper assembly of the conical CA core, and that the spacer peptide 2 is not required for a proper core formation.

Keywords:

HIV-1, RNP condensation, NC p6, cryo-ET, infectivity, core morphology

Introduction

The main structural component of retroviruses, like human immunodeficiency virus (HIV-1) is Gag, a 50-75-kDa polyprotein. All retroviral Gag proteins contain three structurally conserved functional domains: matrix (MA), which contains a membrane binding domain; capsid (CA), a protein composed of two domains which forms the protective capsid encasing the genome; and nucleocapsid (NC), responsible for packaging and condensing the viral RNA genome. The 55 kDa HIV-1 Gag polyprotein includes in addition the C-terminal p6 protein, which acts as an adapter to recruit the cellular ESCRT machinery responsible for virus budding (Weiss & Gottlinger, 2011)and two short spacer peptides: SP1 and SP2, separating CA and NC, and NC and p6, respectively. SP1 plays an important role in regulating the formation of the mature HIV capsid essential for virus infectivity. The function of SP2 in virus morphogenesis and infectivity remains unclear.

The process of HIV morphogenesis involves the assembly of a variable number of Gag molecules at the cytoplasmic face of the plasma membrane of the virus-producing cell (Briggs & Krausslich, 2011). Approximately 2500 copies of Gag form a spherical virus particle which buds from the plasma membrane in an immature, non-infectious form. Cryo-electron microscopy (cEM) and tomography (cET) analyses of immature HIV-1 revealed a radial organization of Gag molecules, with the MA domain adjacent to the membrane (Wilk et al, 2001). Towards the interior of the particle CA domains are organized in a hexameric lattice with a unit cell of ~8nm (Briggs et al, 2009; Briggs et al, 2004). Rod-like structures presumably formed by SP1 connect CA to the NC-RNA and p6 layer, which is not hexagonally ordered and not well resolved in these analyses (Briggs et al, 2009; de Marco et al, 2010a; de Marco et al, 2010b; Wright et al, 2007).

Concomitant with HIV budding, the viral protease (PR) cleaves Gag in five positions (Fig. 1A). This process, called maturation, is accompanied by dramatic structural and morphological changes within the virion, resulting in the formation of mature, infectious HIV (Coffin JM, 1997). In the mature virus MA remains associated with the viral envelope, the CA layer is detached from MA and forms a fullerene cone with a hexagonal lattice spacing of ~10nm (Briggs et al, 2003; Ganser et al, 1999; Li et al, 2000). This wider spacing of the hexagonal lattice, together with the fact that only a part of all CA molecules are required to form the mature core structure, indicates that formation of the mature capsid involves disassembly of the immature lattice followed by assembly of a new lattice structure within the confined space of the virion. The RNP is detected in electron microscopic analyses as a condensed mass within the cone-shaped capsid core.

The complex process of morphological maturation is apparently tightly controlled through sequential proteolysis of Gag. Accordingly, the five main PR recognition sites within Gag are cleaved at different rates in vitro, with a ~400 fold difference in rates between the fastest cleavage (separating SP1 and NC) and the slowest cleavage (separating CA and SP1) (Pettit et al, 2004). Proteolytic maturation is an important target for antiretroviral drugs, and understanding the intermediate stages remains an important research goal (Anderson et al, 2009). The process of HIV maturation is not synchronized and distinct morphological intermediates of maturation intermediates are not detected in the vicinity of infected cells, suggesting that the overall process

proceeds rapidly. A possible approach to to investigate on the intermediate states during the maturation is to impede the processing of specific cleavage sites introducing mutations at individual PR recognition sites (Krausslich et al, 1995), or by using the drug bevirimat, which selectively inhibits processing between CA and SP1 (Fig 1B). Using these approaches, the effect of cleavage at distinct sites between MA and NC on HIV has been recently characterized (de Marco et al, 2010b; Keller et al, 2011). The cleavage between MA and CA is essential for the detachment of the CA lattice from the viral envelope, but is not sufficient to disrupt the immature organization in of CA. The cleavage between CA and NC induces condensation of the RNP in the centre of the virus particle, but again by itself not sufficient for immature CA lattice assembly. Cleavage at both sides of the CA-SP1 module must occur in order to induce a structural change in the CA region, which leads to dissociation of the immature lattice and allows CA to assemble into a mature lattice. Inhibition of the cleavage between CA and SP1 lead to formation of aberrant cores and a loss in infectivity (de Marco et al, 2010b; Wiegers et al, 1998).

The role of the cleavages in the C-terminal region of Gag and the importance of SP2 for virion morphogenesis not yet defined. Viruses carrying mutations that prevent cleavage at the site between NC and SP2 have been reported to show almost wild type infectivity in tissue culture (Coren et al, 2007). Processing at the faster cleavage site between SP2 and p6 appears to be more critical for virus replication, since prevention of proteolysis at this site leads to a significant drop in infectivity. Preventing all cleavages between NC and p6 cleavages leads a severe reduction in infectivity of between 3 and 5 logs in tissue culture systems (Coren et al, 2007; Muller et al, 2009), although the mature MA and CA proteins are still efficiently produced

Here we studied the effect of preventing Gag processing between NC and p6 as well as of the presence of SP2 on structural maturation and infectivity using a combination of cryo-electron tomography and biochemical assays. Our observations identify a role for the correct processing between NC and p6 during capsid core assembly, which is not mediated by SP2.

Material and Methods

Cell lines and plasmids

293T cells and HeLa TZM-bl cells were grown in Dulbecco's modifies Eagle's medium (DMEM) supplemented with and 10% foetal calf serum (FCS). Proviral plasmids used for virus production were constructed based on pNLC4-3 (Bohne & Krausslich, 2004) or pNL4-3uncoupled, respectively. pNL4-3uncoupled was constructed by inserting a synthetic DNA fragment whose sequence was based on (Leiherer et al, 2009). Plasmids pNLC4-3NC-SP2 and NC-p6 have been described previously (Muller et al, 2009). Plasmid pNLC4-3SP2-p6 was generated by introducing mutation F16S (Coren et al, 2007) into pNLC4-3 by overlap. pNL4-3uncoupledΔSP2 was also created by overlap PCR.

Virus preparation

293T cells were transfected with the respective proviral plasmids using a standard calcium phosphate transfection procedure. At 44 h post transfection, tissue culture supernatant was harvested and passed through a 0.45μm filter. Virions were purified from the cleared supernatant by ultracentrifugation through a 20% (w/w) sucrose cushion and ultracentrifugation through an iodixanol gradient as described previously (Briggs et al, 2009). Particles were re-suspended in phosphate buffered saline (PBS) and stored at -20°C. For cryo-ET analyses, virions were subjected to mild inactivation (1% paraformaldehyde, 1 h on ice). Purity of samples was assessed via SDS-PAGE and silver staining. For immunoblotting, samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Membranes were incubated with antisera raised against recombinant HIV-1 CA, MA and NC, respectively, and bound antibody was detected by quantitative immunoblotting using a LiCor Odyssey system, as well as secondary antibodies and protocols according to the manufacturer's instructions.

Infectivity assay

Tissue culture supernatant from 293T cells transfected with the respective proviral plasmid was harvested and passed through a 0.45μm filter. Virus concentrations in the sample were determined by quantitative immunoblotting (LiCior), using antiserum raised against recombinant HIV-1 CA and purified recombinant CA as a standard.

Cryo-electron tomography

Virions were mixed with 10nm gold beads, deposited on C-flat holey carbon grids, blotted and vitrified by plunge freezing in liquid ethane. Tilt series were collected on an FEI Tecnai F30 "Polara" transmission EM with Gatan GIF 2002 post column energy filter and 2kx2k Multiscan CCD camera. The data collection was performed at 300kV using the SerialEM software package and tomograms were reconstructed using the IMOD software package (Kremer et al, 1996). Tilt series were typically collected between -60 and +60 degrees with a tilt increment of 3 degrees and a total electron dose of approximately 100e/Å^2. The defocus range was between 6.0 and 8.0μm, with a magnification of 27,500X resulting in a pixel size at the specimen level of 4.7Å.

Segmentation and volume measurement

To measure the volume of the virion cores, cores were manually segmented using Amira (Visage) from cryo-electron tomograms low pass filtered to 80Å. To refine the position of the CA layer, we extracted subtomograms centred on the surface of the segmented volume with an even spacing of 6 pixels, pre-oriented to the normal of the surface of the segmented object. The subtomograms underwent iterative translational alignment along the normal vectors. We then spline fitted a closed surface in 3D based on the final coordinate of the subtomograms to define the refined surface. We then calculated the surface area and the volume for the cores. For each virion we measured three orthogonal radii and calculated the volume at the inner side of the envelope. All calculations and image processing were done using Matlab (Mathworks).

Results

Virus production and functional characterization

In order to shed light on the relative importance of the individual processing sites between NC and p6 for the structural reorganization of HIV and to understand the role of SP2 during the maturation and in the infectivity we generated and analysed a panel of mutant HIV-1 derivatives as summarized in Fig. 1B-C. The constructs NC-SP2, SP2-p6 and NC-p6 carried point mutations at the NC-SP2 or/and the SP2-p6 cleavage site, which have been described to inhibit processing by HIV-1 PR (Coren et al, 2007; Muller et al, 2009). Mutational analysis of the SP2 region in the viral context is hindered by the fact that changes in SP2 also affect the overlapping gag-pol frame-shift signal and trans-frame region of the pol open reading frame (ORF); thus effects caused by mutagenesis of SP2 cannot be separated from functional effects on pol. To avoid this problem, we generated the ΔSP2 variant in the context of a pNL4-3 derivative in which gag and pol are uncoupled due to mutational inactivation of the original frame-shift signal and a duplication of the frame-shift region downstream of the gag stop codon (Figure 1C. This modification has been reported to be fully compatible with Gag and Pol production, Gag processing and virus infectivity (Leiherer et al, 2009). NL4-3uncoupledΔSP2 (ΔSP2) lacked the first 10 amino acids of SP2, retaining only the fast cleavage site between SP2 and p6 in order to allow release of NC from p6.

All viruses were purified by gradient centrifugation from the tissue culture supernatant of 293T cells transfected with the respective proviral plasmids as previously described (Briggs et al, 2009). The relative amounts of released viruses were comparable for all variants; the purity of the samples was assessed by SDS-PAGE and silver staining (Fig. 2A). Immunoblot analyses were performed to confirm the presence of the expected partially cleaved Gag products. These analyses also verified that the cleavage at the non-mutated sites proceeded normally in all cases (Fig. 2A).

The infectivity of all variants was compared by titration on HeLa TZM-bl indicator cells as well as by endpoint titration of C8166 T-cells (Fig 2B). In accordance with a previous report (Leiherer et al, 2009), both wild-type viruses, NL4-3 (wt) and NL4-3uncoupled, displayed very similar titers on both cell lines.

Consistent with previous results HIV infectivity in tissue culture was not affected by blocking cleavage at the cleavage site between NC and SP2 (Coren et al, 2007; Muller et al, 2009), while blocking the cleavage at the faster site (between SP2 and p6) causes a detectable reduction in infectivity. The infectivity of the variant in which both cleavage sites were mutated was severely impaired. Deletion of most of SP2, while still allowing a rapid processing step between NC and p6, did not abrogate HIV infectivity. With a 2-3 fold reduction in relative infectivity compared to its wt counterpart, ΔSP2 was only modestly impaired in TZM-bl indicator cells. While the titration in C8166 cells, where virus spread in the culture is required for a positive readout, showed a decrease of 2 orders of magnitude.

Effects on morphology

The morphology of the mutated HIV derivatives was analysed in detail by cryo-ET. This technique allows investigation of the three-dimensional (3D) structure of HI virions in unprecedented detail and has been used to characterize the architecture of mature and immature HIV particles as well as of HIV budding sites. Since it is known from such studies that preparations of HIV present some morphological heterogeneity, the visualization of a few individual particles is not sufficient to draw conclusions. We therefore performed statistical analyses in the case of all mutated viruses in parallel to wild-type controls. In order to control for any variability resulting from the transfection and purification procedure, two independent preparations were analysed for each variant, with a minimum of 50 viruses analysed per preparation. Based on the 3D structures derived from cryo-ET, all individual particles were manually classified into one of four structural groups: (I) mature, conical core, (II) striated, spherical immature Gag shell, (III) irregular, sometimes open, core structure and (IV) containing a condensed RNP, but no discernible Gag layer. Since we did not observe significant differences between independent preparations of the same variant (Table 1), data from both preparations were pooled for analysis.

The NL4-3 preparations showed a range of morphologies comparable with previously published data (Briggs et al, 2003). The vast majority (85%) of viruses contained a conical core; some of these particles contained more than one core. A small proportion (3%) of virus particles had failed to undergo proteolytic maturation and showed the striated Gag density characteristic of immature virions. Besides that, 8% of the virions displayed irregularly shaped, in some cases open, cores of various sizes, while 4% of viruses contained no discernible capsid layer but only contained condensed mass resembling the RNP detected inside mature cores (Fig. 3). Consistent with its observed wild-type infectivity, NL4-3uncoupled showed a very similar distribution of morphologies to the wt (Table 1).

An analogous statistical analysis of particle morphologies for the variants containing cleavage site mutations revealed that the degree of morphological defects paralleled the effects on infectivity. In the case of the NC-SP2 variant, the majority of particles (61%) still carried cone-shaped cores albeit the percentage of disordered cores was increased as compared to wt. Preventing cleavage between SP2 and p6 resulted in a more dramatic phenotype. In this case, only 35% of the particles displayed the characteristic mature core, while in a large proportion (54%) the core was irregular or missing. When both cleavage sites were blocked, the majority of particles (80%) contained irregular cores, while only a minor part (12%) appeared morphologically mature. The amount of immature particles detected did not change systematically; between 2 and 10% were observed in all preparations. In contrast to the clearly detectable differences in case of the cleavage site mutants, the ΔSP2 variant displayed a distribution of particle morphologies that was almost indistinguishable from the wt distribution indicating that the sequential maturation from NC-SP2 to NC does not play a critical role for the formation of the mature HIV capsid.

The variants mutated at cleavage sites between NC and p6 showed impairment of mature capsid formation, although mature CA was produced with normal efficiency. Since one important functional role of NC is the condensation of the viral genome

into a dense RNP complex, a failure to form this properly condensed RNP may be the reason for impaired capsid assembly in the studied cleavage site mutants. Therefore we examined the particles for the degree of condensation of the dense structures representing the RNP. Visual inspection did not show an evident difference in the size of the presumed RNP structures correlated to cleavage site mutations. Unfortunately, variation in density within the RNP-like structure and its smooth shape did not allow a sharp definition of structure borders and thus prevented a precise measurement of the volumes.

Instead we asked whether the volume encapsidated by irregular HIV cores differed from that encapsidated by conical cores. We plotted core volume against virion volume for conical and irregular cores from the wt and the NC-p6 particles (Fig. 4). For both variants, the volume of the conical cores was found to be approximately proportional to the diameter of the virion, in agreement with what has been previously described for wt HIV (Briggs et al, 2003). Irregular cores in both variants had a higher volume in relation to virions size than conical cores, and appeared to be more variable in size.

Discussion

Here we have explored the functional importance of the proteolytic cleavages between NC and p6 for HIV-1 infectivity and for structural maturation. We have also explored the function of the SP2 peptide in infectivity and structural maturation.

Prevention of proteolytic cleavage in the region between NC and p6 did not only affect infectivity, but also affect the correct assembly of the conical CA core. It is informative to compare infectivity and core morphology across the different mutants analysed. While we did observe an overall correlation between the degree of impairment of core assembly as well as of infectivity in the case of the cleavage site mutants, the magnitude of the effects were not directly proportional. Furthermore, the ΔSP2 variant did show reduced infectivity in tissue culture, although conical core formation was not detectably impaired.

Importance of SP2

The infectivity and morphology of *wt* viruses with the uncoupled pol expression was directly comparable with the *wt,* indicating that uncoupling of Gag and Pol does not interfere with infectivity or maturation in this system.

SP2 itself is not essential for formation of infectious virus particles, since the ΔSP2 variant is still infectious. Deletion of SP2, however, does cause a 2 orders of magnitude reduction in infectivity indicating some loss of virus fitness. Deletion of SP2 has no effect on the morphology of the mature virus particles indicating that it plays no role in structural maturation of the capsid core of the virion. Further, these properties of ΔSP2 also imply that when maturation occurs via a single cleavage event between NC and p6 core formation can proceed normally. Since infectivity drops by 2 logs without a change in core morphology, the role of SP2 in virus fitness is not associated with structural maturation.

Importance of the cleavages, effects on core morphology and infectivity

The NC-SP2 variant has wild-type infectivity. To play its role in virus infectivity SP2 does not require to be released from NC. In this variant, cleavage between NC and SP2 is abolished, whereas cleavage at the SP2-p6 site proceeds normally, as is also the case in the ΔSP2 variant. The reduction in infectivity in the ΔSP2 variant is therefore not the result of a slight reduction in the total cleavage rate between NC and p6 as a result of removing one of the protease sites. The NC-SP2 variant produces approximately 25% fewer virions with conical cores. A directly proportional reduction in infectivity would not be detected in the infectivity assays.

We observed an unexpected direct effect of cleavages between NC and p6 on capsid core morphology. The wt and ΔSP2 variants show normal virion morphology; here a single rapid cleavage removes p6. In the NC-SP2 variant, a rapid cleavage separates NC from p6, but release of the peptide SP2 does not occur. This has a small effect on core morphology. In the SP2-p6 variant, only slow cleavage removes SP2 and p6 from NC, and the effect is more severe. In the NC-p6 variant, neither SP2 nor p6 is removed from NC, and only 11% of cores assembly correctly. We speculate that failure to quickly remove p6, and failure to subsequently remove SP2, may have cumulative affects on the rate of RNP condensation, which may in turn hinder proper assembly of the cone shaped core. In SP2-p6 and in NC-p6 the reduction in infectivity is many fold greater than the reduction in the number of conical cores, indicating that the abrogation in infectivity is not exclusively due to impairment in conical core formation.

In summary, the SP2 peptide is not required for infectivity, but it does play a role in viral fitness. Inhibition of proteolytic cleavages between NC and p6 leads to an impairment in structural maturation and more dramatic reductions in viral infectivity. These observations are consistent with a model in which failure to efficiently separate SP2 and p6 from NC alter the kinetics of the RNP condensation, and in which proper RNP condensation is an essential step in proper assembly of a mature conical capsid core.

Author contributions

Conceived and designed the experiments: AdM, BM, HGK, JAGB
Performed the experiments: AdM, AMH, BG, BM
Analyzed the data: AdM, BM, HGK, JAGB
Wrote the paper: AdM, JAGB, BM, HGK

Table 1

Construct	Mean virion Size (nm)	Preparation number	Conical core (%)	Irregular core (%)	Immature (%)	No capsid layer (%)	Number of virions measured
pNL4-3	129 ± 23	1	84	10	4	1	68
		2	87	6	1	6	67
ΔSP2	135 ± 24	1	97	0	2	1	93
		2	70	15	2	13	86
NC-SP2	132 ± 22	1	68	13	6	13	63
		2	58	31	5	7	98
SP2-p6	127 ± 17	1	39	46	11	3	61
		2	30	61	6	2	87
NC-p6	136 ± 28	1	12	75	12	1	69
		2	12	84	2	0	89
pNL4-3 uncoupled	137 ± 22	1	86	5	8	2	65
		2	78	13	2	7	54

Morphological analysis of the variants. The mean size has been measured for each variant, and the classification in the four phenotypes is here showed as percentage. A minimum of 119 virions in total per variant, and 54 per preparation were analysed.

References

Anderson J, Schiffer C, Lee SK, Swanstrom R (2009) Viral protease inhibitors. *Handbook of experimental pharmacology***:** 85-110

Bohne J, Krausslich HG (2004) Mutation of the major 5' splice site renders a CMV-driven HIV-1 proviral clone Tat-dependent: connections between transcription and splicing. *FEBS letters* **563:** 113-118

Briggs JA, Krausslich HG (2011) The molecular architecture of HIV. *Journal of molecular biology* **410:** 491-500

Briggs JA, Riches JD, Glass B, Bartonova V, Zanetti G, Krausslich HG (2009) Structure and assembly of immature HIV. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 11090-11095

Briggs JA, Simon MN, Gross I, Krausslich HG, Fuller SD, Vogt VM, Johnson MC (2004) The stoichiometry of Gag protein in HIV-1. *Nature structural & molecular biology* **11:** 672-675

Briggs JA, Wilk T, Welker R, Krausslich HG, Fuller SD (2003) Structural organization of authentic, mature HIV-1 virions and cores. *The EMBO journal* **22:** 1707-1715

Coffin JM HS, Varmus HE (1997) In *Retroviruses*, Coffin JM, Hughes SH, Varmus HE (eds). Cold Spring Harbor (NY)

Coren LV, Thomas JA, Chertova E, Sowder RC, 2nd, Gagliardi TD, Gorelick RJ, Ott DE (2007) Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. *Journal of virology* **81:** 10047-10054

de Marco A, Davey NE, Ulbrich P, Phillips JM, Lux V, Riches JD, Fuzik T, Ruml T, Krausslich HG, Vogt VM, Briggs JA (2010a) Conserved and variable features of Gag structure and arrangement in immature retrovirus particles. *Journal of virology* **84:** 11729-11736

de Marco A, Muller B, Glass B, Riches JD, Krausslich HG, Briggs JA (2010b) Structural analysis of HIV-1 maturation using cryo-electron tomography. *PLoS pathogens* **6:** e1001215

Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI (1999) Assembly and analysis of conical models for the HIV-1 core. *Science* **283:** 80-83

Keller PW, Adamson CS, Heymann JB, Freed EO, Steven AC (2011) HIV-1 maturation inhibitor bevirimat stabilizes the immature Gag lattice. *Journal of virology* **85:** 1420-1428

Krausslich HG, Facke M, Heuser AM, Konvalinka J, Zentgraf H (1995) The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *Journal of virology* **69:** 3407-3419

Kremer JR, Mastronarde DN, McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. *Journal of structural biology* **116:** 71-76

Leiherer A, Ludwig C, Wagner R (2009) Uncoupling human immunodeficiency virus type 1 Gag and Pol reading frames: role of the transframe protein p6* in viral replication. *Journal of virology* **83:** 7210-7220

Li S, Hill CP, Sundquist WI, Finch JT (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407:** 409-413

Muller B, Anders M, Akiyama H, Welsch S, Glass B, Nikovics K, Clavel F, Tervo HM, Keppler OT, Krausslich HG (2009) HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. *The Journal of biological chemistry* **284:** 29692-29703

Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH (2004) Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. *Journal of virology* **78:** 8477-8485

Weiss ER, Gottlinger H (2011) The role of cellular factors in promoting HIV budding. *Journal of molecular biology* **410:** 525-533

Wiegers K, Rutter G, Kottler H, Tessmer U, Hohenberg H, Krausslich HG (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *Journal of virology* **72:** 2846-2854

Wilk T, Gross I, Gowen BE, Rutten T, de Haas F, Welker R, Krausslich HG, Boulanger P, Fuller SD (2001) Organization of immature human immunodeficiency virus type 1. *Journal of virology* **75:** 759-771

Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, Sundquist WI, Jensen GJ (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *The EMBO journal* **26:** 2218-2226

Figure 1

Schematic overview of the constructs analysed in this study. A) Outline of the proteolytic cleavages within Gag that occur during the maturation of HIV-1.
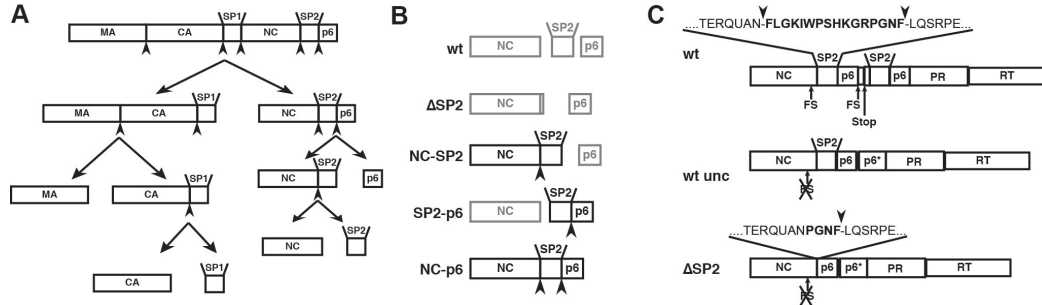

Arrowheads indicate the position of PR recognition sites. The displayed order of cleavage events is based on cleavage rates determined for peptides in vitro and in vitro studies of Gag proteolysis (Pettit et al). B) Schematic representation of Gag after completion of proteolytic maturation for each variant analysed. The uncleaved products remaining due to mutation of the proteolytic site are highlighted. Mutated uncleaved proteolytic sites are marked with arrowheads. C) Schematic representation of the uncoupled gag and pol reading frames in pNL4-3uncoupled and the deletion introduced in the ΔSP2 variant.

Figure 2
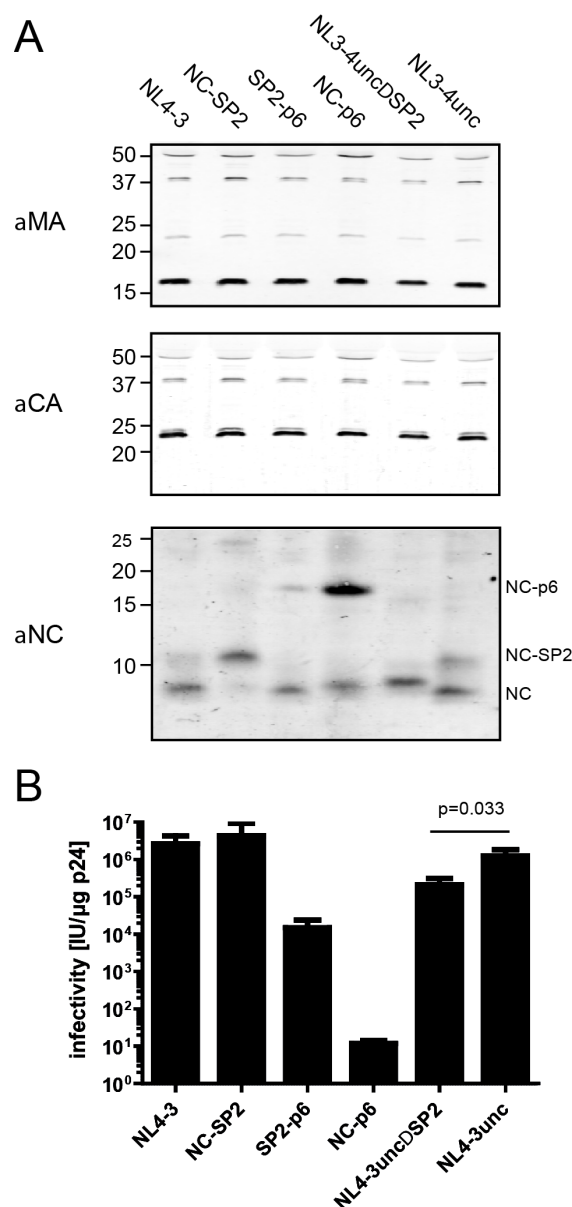
Infectivity of the HIV-1 derivatives. A) Immunoblot analysis of virus particles. Virions were prepared from the tissue culture supernatants of 293T cells transfected with the respective plasmids by ultracentrifugation through a 20% (w/w) sucrose cushion. Samples were separated by SDS-PAGE and Gag derived proteins were detected by quantitative immunoblot (LiCor) using polyclonal antisera raised against the indicated proteins. Positions of molecular mass standards are indicated to the left, positions of NC containing products to the right. B) Relative infectivities were determined by end-point titration of tissue culture supernatants from transfected 293T cells on C8166 T-cells. Values were normalized for the amount of CA p24 detected in the supernatants by quantitative immunoblot. Bars represent mean values and standard deviations from triplicate transfections.

Figure 3
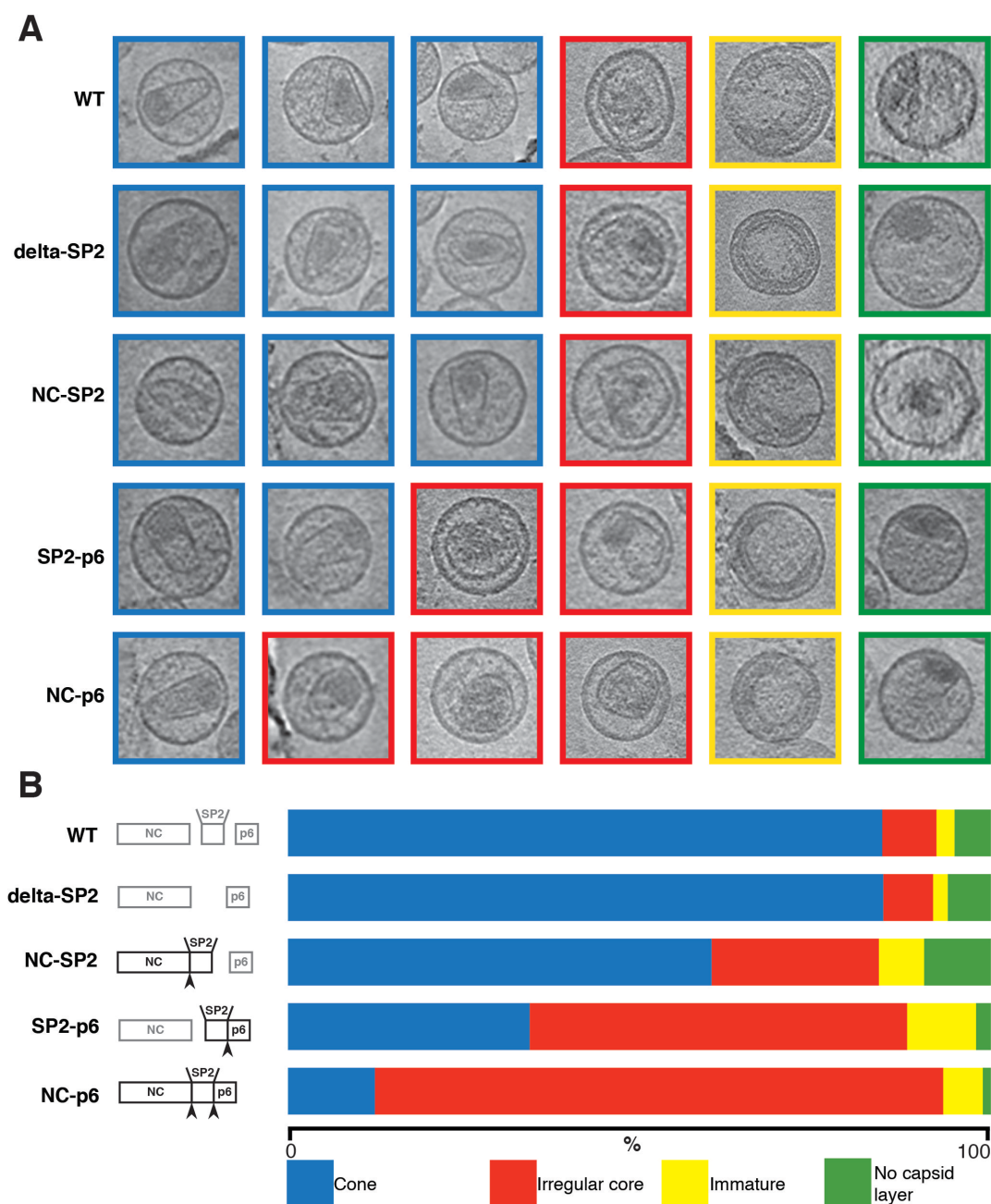
Morphological analysis of the variants. Here choose 4 categories upon the core shape: conical (blue), irregular (red), immature (yellow), and no lattice detected (green). A) Overview of the different morphological phenotypes represented in the statistical evaluation, and the relative abundance. The images are computational slices of 18Å through the cryo-tomograms. The boxes have a size of 150x150nm. B) Bar plot showing the relative abundance of the phenotypes determined for the different virus variants. The color-code used to identify the phenotypes is the same in both panels and is described in the key at the bottom.

Figure 4
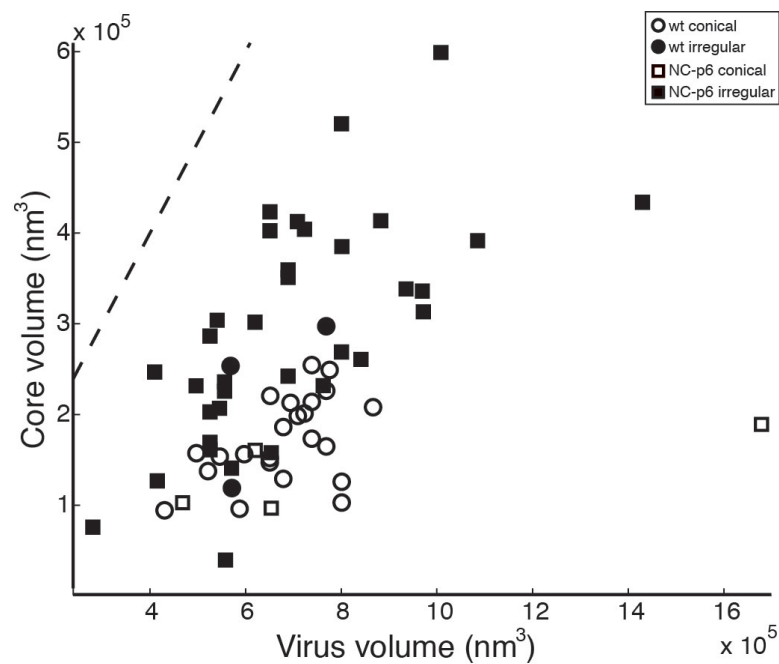
Relation between core and virus volume. The scatterplot shows the relation between the volume of each core in relation to the volume of the respective virion. Circles represent data from wt HIV-1, squares data from NC-p6. Open data points represent conical cores; filled data points identify irregular cores. The dotted line indicates the 1:1 ratio between volumes of the core and the virion.