Dissertation
submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of

Doctor of Natural Sciences

Presented by
M.Sc. Erica Margiotta
Born in Rome, Italy
Oral examination: 26 July 2021

Nuclear entry of HIV-1 capsids through nuclear pore complexes explored *in cellulo* by cryo-electron tomography

Referees:

Dr. Julia Mahamid

Prof. Dr. Hans-Georg Kräusslich

For

Maria, Michele and Gianluca

This study was carried out in the Structural and Computational Biology Unit (SCB) at the European Molecular Biology Laboratory (EMBL) Heidelberg from October 2017 to January 2021 supervised by Dr. Martin Beck.

Summary

Human Immunodeficiency virus type 1 (HIV-1) is a lentivirus that infects non-dividing cells of the immune system. In non-dividing cells, nuclear import of the viral genome occurs through the nuclear pore complex (NPC), a large macromolecular assembly that forms a channel of ~40 nm across the nuclear envelope. This process requires interactions between NPC components known as nucleoporins and CA (capsid) proteins of HIV-1 post-fusion complexes. In mature HIV-1 virions, multiple copies of CA assemble into a lattice of hexamers and pentamers, forming a cone-shaped capsid core of ~60 nm in width and ~110 nm in length encasing the viral genome. Because of its large size, it was generally believed that the capsid core entirely or partially disassembled prior to its translocation through the NPC. However, nuclear entry and uncoating are rare events and challenging to characterize and therefore subject to a long-standing debate.

Here, cryo-electron tomography (cryo-ET) on cryo-focused ion beam (FIB) milled T-cells was combined with subtomogram averaging (SA) to study the ultrastructure of HIV-1 capsids in the process of nuclear entry. Using a HIV-1 variant arrested at the nuclear pore due to a defect in binding of the CA lattice to the host cell protein CPSF6, snapshots of HIV-1 nuclear entry at multiple stages were captured. Reverse transcription-competent HIV-1 complexes were identified in the cytosol, docking to and within the NPC, and in the nucleoplasm of T cells. Surprisingly, in the cytosol and at the NPC, the viral complexes retained cone-shaped capsids highly resembling in size and geometry the mature capsid cores within intact virions. The density at their interior suggested that they were associated with the viral genome. The cone-shape capsids deeply entered into the NPC central channel with their narrow ends. The hexameric lattices of these capsids were intact or nearly intact. These findings argue against the current models of uncoating in the cytosol or at the NPC and rather support translocation of intact HIV-1 capsids through the NPC. Instead, HIV-1 capsids uncoat in the nucleus of T cells. Inside the nucleus, tube-shape fragments still containing few lattice elements were observed. The lack of density at their inside suggested that the viral genome was released from these complexes. Uncoating may thus occur through the partial opening and remodeling of the hexameric lattices of HIV-1 capsids, rather than a step-wise disassembly.

To address how an intact HIV-1 capsid can enter the central channel of NPC, the human NPC from T cells was structurally analyzed *in cellulo* by SA. The NPC overall structure was not altered upon infection but rather dilated in comparison to the structure previously obtained from purified envelopes. The diameter of the central channel of NPC in T cells corresponded to ~64 nm, suggesting that the translocation of intact HIV-1 capsid through the NPC is geometrically possible. The dilated conformation of NPCs in T cells is independent of HIV-1 infection, but represents a physiological condition of actively transporting NPCs.

Zusammenfassung

Das Humane Immundefizienz-Virus Typ 1 (HIV-1) gehört zur Klasse der Lentiviren und infiziert nicht mitotische Zellen des Immunsystems. In diesen Zellen wird das virale Genom durch Kernporenkomplexe (NPCs) in den Zellkern transportiert. NPCs sind große makromolekulare Proteinkomplexe, die einen etwa 40 nm breiten Kanal in der Kernhülle bilden. Notwendig für den Kerntransport ist die Interaktion von Nucleoporinen, den Bausteinen des NPCs, mit viralen Capsid-Proteinen (CA). In reifen HIV-1 Virionen formen zahlreiche CA Kopien ein aus Hexameren und Pentameren bestehendes Gerüst. Dieses bildet ein konusförmiges, etwa 60nm breites und ~110 nm langes Capsid, welches das virale Genom umschließt. Aufgrund der Dimensionen des Capsids nahm man an, dass dieses vor dem Transport in den Kern partiell oder zur Gänze zerlegt wird um die Kernpore passieren zu können. Da der Kernimport bzw die Freisetzung des viralen Genoms sehr selten sind, blieb eine Charakterisierung dieses Prozesses sehr schwierig und ein möglicher Mechanismus war daher über lange Zeit Gegenstand von Diskussionen.

Um die offene Frage der HIV-1 Capsid-Ultrastruktur beim Kernimport zu beantworten analysierten wir in der vorliegenden Arbeit mittels FIBSEM-Technologie geschnittene T-Zellen mit Methoden der Kryo Elektronentomographie (ET) und des Sub-Tomogramm Averaging. Durch die Verwendung einer HIV-1 Variante die aufgrund einer gestörten Interaktion des Capsids mit dem Zellprotein CPSF6 an der Kernpore arretiert, gelang es Momentaufnahmen des Virusimports in den Zellkern aufzunehmen. Diese detektierten Replikations-kompetente HIV-1 Komplexe im Cytosol, an und in der Kernpore, sowie im Nukleoplasma von infizierten T-Zellen. Überraschenderweise behielten die Capside sowohl im Cytoplasma als auch an der Kernpore ihre konusförmige Architektur und entsprachen in ihrer Form und ihren Dimensionen dem reifen Capsid in intakten Virionen. Elekronendichtes Material im Inneren des Capsids zeigte dass dieses mit dem viralen Genom assoziiert ist. An der Kernpore tauchten die konusförmigen Capside mit ihrer schmalen Seite tief in das Poreninnere, wobei die hexamere Gerüststruktur des Capsids intakt blieb. Unsere Resultate argumentieren gegen einen Mechanismus, bei dem das Capsid vor dem Kernimport zerlegt wird. Vielmehr sprechen unsere Ergebnisse für

einen Transport des intakten HIV-1 durch die Kernpore mit anschliessender Freisetzung des viralen Genoms im Kern. Im Nucleoplasma ließen sich röhrenförmige Viruscapsid-Fragmente detektieren die zwar hexamere Gerüstelemente enthielten jedoch vom viralen Genom bereits dissoziiert waren.

Um zu untersuchen, wie ein intaktes HIV-1 Capsid durch den zentralen Kanal der Kernpore gelangt, analysierten wir weiters die Struktur der Kernpore in intakten humanen T Zellen mittels Sub-Tomogramm Averaging. Zwar blieb die allgemeine Struktur der Kernpore durch die Infektion unverändert, jedoch erschien die Pore im Vergleich mit zuvor ermittelten Strukturdaten von NPCs aus aufgereinigten Kernhüllen erweitert. Der Durchmesser des zentralen NPC-Kanals in T Zellen betrug ~64nm, groß genug um den Durchtritt eines intakten HIV-1 Capsids geometrisch zu ermöglichen. Der vergrößerte Durchmesser war nicht Resultat der Infektion sondern entspricht vielmehr der physiologischen Situation von aktiv transportierenden Kernporen.

Table of contents

List of Figures

Contributions and publications

Cone-shaped HIV-1 capsids are transported through intact nuclear pores

Vojtech Zila*, **Erica Margiotta***, Beata Turoňová, Thorsten G. Müller, Christian E. Zimmerli, Simone Mattei, Matteo Allegretti, Kathleen Börner, Jona Rada, Barbara Müller, Marina Lusic, Hans-Georg Kräusslich and Martin Beck

***Contributed equally**

Cell 184, 1032-1046, doi: https://doi.org/10.1016/j.cell.2021.01.025

Results, corresponding figures and methods which were not generated by myself or contain parts of data generated by someone else are marked appropriately throughout the thesis.

Chapter 1

Introduction

1.1. The Human Immunodeficiency Virus (HIV)

The human immunodeficiency virus type-I (HIV-1) is the virus responsible for a widely known and still ongoing pandemic originated 40 years ago with >75 million people infected worldwide, including ~33 millions deaths and ~38 million people still living with the virus (UNAIDS 2019 report). HIV-1 interferes with the immune system by infecting and progressively causing the death of blood cells such as $CD4^+$ T cells, macrophages and dendritic cells. As a result, HIV-1 causes a spectrum of symptoms referred to as acquired immune deficiency syndrome (AIDS), which includes an increased risk of developing common infections and tumors, such as tuberculosis and Kaposi's sarcoma, which are otherwise rare in people who have a normal immune function.

Classified as lentivirus, a subgroup of retroviruses, HIV-1 is not a single virus but comprises four distinct groups of viruses, termed as M, N, O and P (Hemelaar et al., 2006; Robertson, 2000). Among them, group M is predominant and responsible for the great majority of all HIV infection worldwide. It is subclassified into nine clades (A-K) with ~15% of nucleotide sequence variation among them and different prevalence: the most prevalent (52%) clade C is found in Eastern and Southern Africa, while clade B (10%) is present mostly in Europe and the America continent. The concurrent circulation of distinct HIV-1 clades increases the chances of coinfection and recombination of the genetic material, with the generation of hybrid viruses known as circulating recombinant forms (CRFs) and unique recombinant forms (URFs) (Jetzt et al., 2000). Nowadays, 89 CRFs and several URFs already exist. High source of viral diversity is also provided by mutations that arise in a single virus during its life cycle (Coffin, 1995). All these high sequence diversities of HIV-1 constitute a major challenge to HIV-1 vaccine development. Available treatments known as antiretroviral therapy attenuate HIV-1-

related morbidity and mortality rate. However, HIV-1 remains difficult to eradicate and incurable.

1.1.1 Genome and viral proteins

HIV is an enveloped retrovirus with a diameter of 120-140 nm that acquires its lipid membrane by budding from the host cell plasma membrane (Briggs et al., 2009). Apart from the envelope, all the other HIV components are encoded by the viral genome. The HIV genome contains around 10,000 nucleotides and consists of two identical positive single-stranded RNA molecules (gRNA), found as a non-covalent dimer inside the virions (Watts et al., 2010). Each strand has a characteristic organization: the middle RNA-coding region contains nine open reading frames (ORFs), which encode for 15 distinct proteins, and two flanking non-coding regulatory elements termed as 5'- and 3'-long terminal repeats (LTR) (**Figure 1.1**). Both LTRs are identical in sequence but differ in functions (Hu and Hughes, 2012). The 5' LTR works as a promoter *via* the recruitment of the host transcription machinery and the regulatory viral protein Tat (Roebuck and Saifuddin, 1999). The 3' LTR acts in the termination of transcription and polyadenylation of the nascent viral RNA and encodes for the accessory protein Nef (Guntaka, 1993).

In the 5' to 3' direction, three genes, *gag*, *pol*, *env* encode the Gag (also known as $Pr55^{Gag}$), Pol and Env polyproteins, respectively, that represent the main components of the immature HIV-1 (**Figure 1.2 A and B**). Upon maturation of the virion, the polyproteins are subsequently proteolyzed into the individual components of the mature HIV-1 virion described in 1.1.1.1, 1.1.1.2 and 1.1.1.3.

Four genes, *vif*, *vpr*, *vpu*, and *nef*, encode accessory proteins: Vif (virion infectivity factor), Vpr (virion associated protein R), Vpu (viral protein U) and Nef (negative factor), respectively. These proteins are involved in the viral pathogenesis *in vivo* and they work as adapter molecules by interacting with cellular factors to hijack the host cell machinery for the advantage of the virus during different steps of its life cycle. Finally, two genes *tat* and *rev* encode the Tat (trans-activator protein) and Rev (RNA splicing-regulator) proteins, respectively, which both provide essential gene regulatory functions. Specifically, Tat is essential for viral replication as increases the production of viral mRNAs more than 100-fold and transactivates additional HIV-1 genomes present in the cell (Frankel and Young, 1998). Rev mediates the nuclear export of unspliced viral

mRNAs and prevents their splicing in order to package them as viral genomic RNAs into assembling virions (Blissenbach et al., 2010).

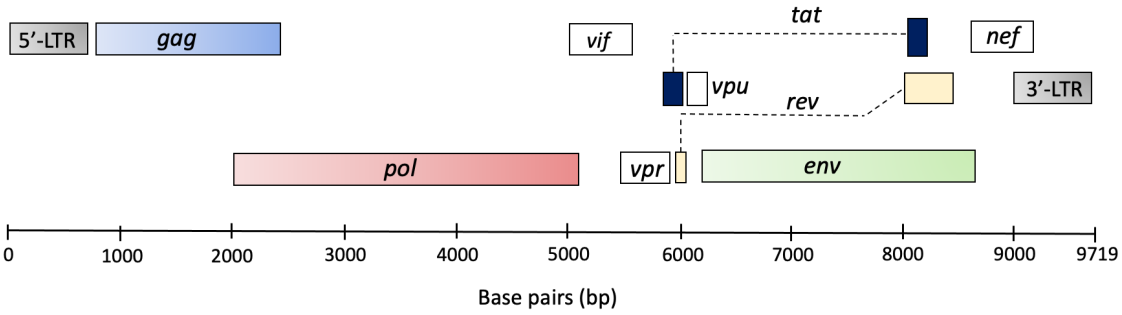


Figure 1.1. Schematic organization of HIV-1 genome. The viral genome contains nine open reading frames flanked at both 5'- and 3'- ends by long terminal repeats (LTR). Three genes gag, pol and env encode for the main viral components: structural proteins (gag and env) and viral enzyme (pol). Four genes, vif, vpr, vpu and nef (white boxes) express the accessory proteins. The last two genes, tat (blue box) and rev (yellow box), have two exons and produce the essential regulatory proteins.

1.1.1.1 Gag polyprotein and structural proteins

The Gag polyprotein (also known as $Pr55^{Gag}$) in the immature HIV-1 virion is the precursor of four structural proteins that compose the mature HIV-1 virion: matrix (MA, p17), capsid (CA, p24), nucleocapsid (NC, p7) and p6 (**Figure 1.2**).

MA are myristoylated membrane-binding proteins that constitute the N-terminal of the Gag polyprotein (Frankel and Young, 1998). They anchor the Gag precursor to the host plasma membrane during virion assembly and facilitate the incorporation of the Env glycoproteins into the viral particles (Mammano et al., 1995). In the mature virion, MA proteins form a layer underneath the envelope and provide stability to the overall virus particle.

CA proteins are the second component of the Gag polyprotein. In the immature Gag shell, CA mediate the main Gag-Gag contacts, whereas in the mature virion, CA form the

conical capsid core that represents a shell for the viral genome, as further described in paragraph 1.1.3 entirely dedicated to the role of CA in HIV-1 infection.
NC proteins are the third domain of the Gag polyprotein with the function to bind the genomic RNA and compact it within the core. The selection and incorporation of the full-length RNA inside the virion occurs during the assembly by the specific binding of NC to a packaging signal also known as ψ localized on the viral RNA sequence (nucleotides number (nts) 240-350) within the 5' end (Clever and Parslow, 1997). Finally, P6 are the C-terminal of the Gag polyproteins that allow the incorporation of the Vpr accessory protein during viral assembly (Kondo and Göttlinger, 1996; Checroune et al., 1995; Dubois et al., 2018) and mediate the budding and release of the viral particle from the host plasma membrane by the interaction with cellular factors (Friedrich et al., 2016). HIV-1 Gag also contains two spacer peptides, named SP1 and SP2, that localize in between the CA/NC and NC/p6 interface domains, respectively.

1.1.1.2 Pol polyprotein and viral enzymes

The Pol polyprotein is incorporated into the virion as part of a Gag-Pol polyprotein precursor which is packaged at a ratio of ~1:20 to Gag (**Figure 1.2**). It provides three viral enzymes essential for viral maturation and replication: protease (PR), reverse transcriptase (RT) and integrase (IN). Because of their importance in the HIV-1 replication cycle, the viral enzymes are the main target in the antiretroviral therapy, which currently represents the only available treatment against AIDS.
HIV-1 PR is a member of the family of aspartic proteases and is a symmetrically assembled homodimer of 22 kDa (Fun et al., 2012). PR triggers the maturation of a noninfectious particle into a mature infectious virus by the cleavage of the Gag and Pol precursors in the individual proteins (Gunik et al., 2000; Fun et al., 2012). Kinetics, order of cleavage and the extent of precursor processing seem to be critical factors for the generation of fully infectious viral particles. Mutations that affect these parameters result in particles with reduced infectivity (Kräusslich et al., 1995; Wiegers et al., 1998). Although the mechanism that activates the viral PR is not yet fully understood, it was reported that PR is responsible for its own release from the precursor polyprotein (autoprocessing) upon dimerization (Pettit et al., 2004). Dimerization might occur when

two adjacent Gag-Pol precursors come together during virus assembly (Navia and Mckeever, 1990; Pettit et al., 2004).

RT and IN mediate two steps in the viral replication cycle which are exclusive for retroviruses, such as the reverse transcription of the viral RNA into DNA and the integration of the newly synthetized DNA into the host chromosomes (Engelman et al., 1993; Craigie and Bushman, 2012). Mature infectious virions contain approximately 150 copies of RT and IN (Christensen et al., 2020).

The mature form of HIV-1 RT is a heterodimer composed of two related subunits: the larger p66, which contains the polymerase and the RNAase H domains and thus provides the main enzymatic activity, and the smaller p51 subunit with a structural role (Jacobo-Molina et al., 1993; Kohlstaedt et al., 1992). In the p66 subunit, the polymerase domain can copy either an RNA or a DNA template but requires a host tRNA primer to initiate the reverse transcription. Different retroviruses use different host tRNA as primers; HIV-1 uses the $tRNA^{Lys,3}$ which is packed inside the virion (Hu and Hughes, 2012). The polymerase activity of the RT, together with that of the host RNA Pol II, contributes to the genetic diversity of HIV-1: RT is error prone due to its lack of proofreading activity and has a high error rate of around 3×10^{-5} (Mansky and Temin, 1995; Preston et al., 1988).

The HIV-1 IN can be structurally divided into three domains: the amino-terminal zinc-binding domain, the central catalytic core domain and the C-terminal domain. The central catalytic core domain is highly conserved in all the retroviral integrase proteins and it has a catalytic activity. The poor solubility of the protein and the flexibility between domains have limited the structural studies of the full-length IN. The partial structures of the IN domains were instead resolved either by crystallization (Dyda et al., 1994) or by NMR spectroscopy (Eijkelenboom et al., 1997), and they contributed in the formulation of different models of the IN in complex with the viral DNA (Savarino, 2007). However, the molecular mechanism of integration mediated by IN *in vivo* became clear only when the crystal structure of the integrase from another retrovirus, the prototype foamy virus (PFV), was solved in complex with the viral DNA (Hare et al., 2010).

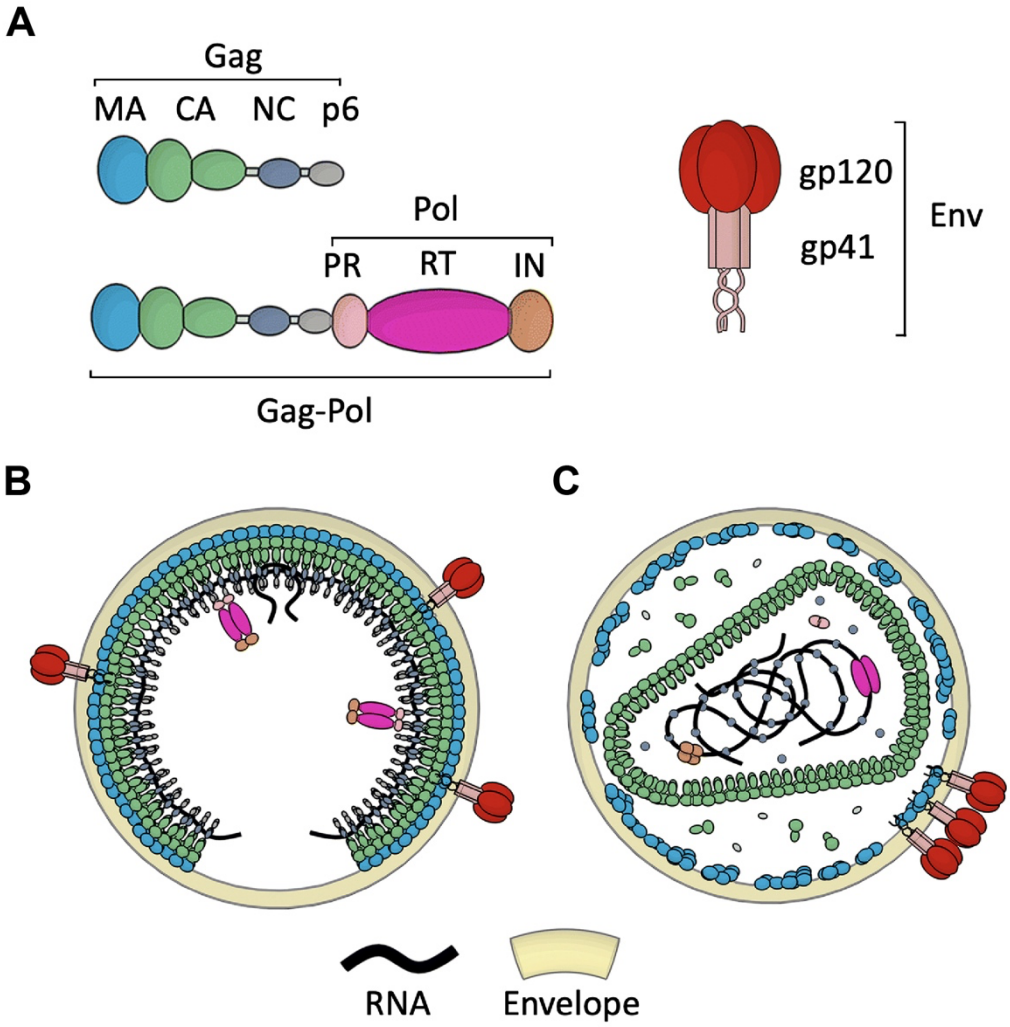


Figure 1.2. Schematic representation of HIV-1 structural components and their organization into the immature and mature virion. (A) Gag, Gag-Pol and Env polyproteins with their respective domains. Gag is the precursor of structural proteins: matrix (MA), capsid (CA), nucleocapsid (NC) and p6 protein. Pol is part of the Gag-Pol polyprotein and precursor of the viral proteins: protease (PR), reverse transcriptase (RT) and integrase (IN). Env contains the transmembrane domain gp41 and the surface protein gp120. (B) Structure of the enveloped immature (left) and mature (right) virion. In the immature virion, Gag and Gag-Pol polyprotein are radially assembled underneath the envelope at the ratio of 1:20. The two copies of full-length RNAs localized at the periphery in association with NC. (C) In the mature virion, MA localize below the envelope and CA assembled into the conical CA core. Inside the core, the genome is packed and associated with RT, IN and NC.

Figure modified from (Chojnacki and Eggeling, 2018).

1.1.1.3 Env polyprotein and trimeric Env spike

The Env polyprotein, also designated as gp160 based on the molecular mass, is translated from a single spliced, bicistronic *vpu/env* mRNA and co-translationally glycosylated with *N*-linked and *O*-linked oligosaccharide side chains on the rough endoplasmic reticulum (RER) membrane to enter the secretory pathway (Checkley et al., 2011; Freed and Martin, 1995). On the RER, monomers of gp160 oligomerize mostly into trimers and to a lesser extent into dimers and tetramers. Upon oligomerization, gp160 traffic from the RER to the Golgi apparatus, where the high mannose ER-acquired N-linked oligosaccharide side chains are further modified. Proteolytic processing of gp160 trimers in the *trans*-Golgi network produces the highly glycosylated surface gp120 (SU) and the transmembrane gp41 (TM) subunits (Checkley et al., 2011). The gp120 subunit is responsible for host-receptor binding; it contains five relatively conserved domains (C1-C5) and five variable loops (V1-V5) that predominantly lie at the surface of gp120 (Wilen et al., 2012). The gp41 subunit is a complex transmembrane protein with various well-defined domains involved at different extent in the fusion of viral and cellular membranes during cytoplasmic entry. The two subunits remain associated by non-covalent bonds and form the trimeric envelope (Env) glycoprotein spike (**Figure 1.2 A**), which localizes on the surface of the viral membrane, tethered by the long cytoplasmic tail of gp41 (**Figure 1.2 B and C**). Env spikes have been visualized by cryo-electron tomography (cryo-ET) but only at low resolution (~20Å) due to high flexibility of the variable loops in gp120 (Chen, 2019; Liu et al., 2008). The Env spikes determine the cell-type specificity or tissue tropism of HIV-1 infection and allow the mature virus to enter the targeted cells by binding their receptors and inducing the fusion of viral and cellular membranes. Each virion incorporates ~7-10 spikes (Zhu et al., 2003). This number of spikes is unusually low for enveloped viruses: for example, influenza virus contains ~400-500 spikes, vesicular stomatitis virus ~1200, Rous sarcoma virus up to 118, and simian immunodeficiency virus ~70. The reason for such a low copy number of Env spikes is still unknown, however it was suggested that this might help HIV-1 to evade the host humoral immune response, such as B cell activation and antibody neutralization, and decrease the virus-induced cytopathicity (Checkley et al., 2011; Egan et al., 1996). To control and keep this number steady, it was suggested that Env spikes are rapidly recycled by endocytosis after their trafficking to the plasma membrane of the host cell (Egan et

al., 1996). How Env is incorporated from the plasma membrane of the host cell to the viral envelope is not fully understood yet. Four general models were put forward, ranging from a passive incorporation of Env simply as a result of gp120 and gp41 expression on the cell surface to a selective incorporation mediated by direct or indirect interaction of gp41 with the Gag polyprotein (Checkley et al., 2011). The lipid membrane composition seems to also play a role in this mechanism, with the preferential Env-Gag targeting on the region of plasma membrane (PM) enriched with lipid rafts or other PM microdomains. Since these models are not mutually exclusive, it is possible that each contributes to Env incorporation to varying degrees in different cellular contexts.

1.1.2 Replication cycle

The HIV-1 life cycle is divided into an early phase including cell binding and entry, reverse transcription, nuclear entry, uncoating and integration of the viral cDNA into the cell genome, and a late phase that includes the expression of viral genes and the assembly, release and maturation of progeny virions (**Figure 1.3**). Each step of infection is introduced in this paragraph, whereas capsid uncoating and nuclear entry are described in more details in sections 1.1.3.3 and 1.1.3.4.

The HIV-1 life cycle begins with the binding of the viral Env glycoprotein trimers to CD4 primary receptors on the target cells. For this to happen, viral particles are initially brought close to the cell surface either by non-specific or specific interactions of Env with negatively charged cell-surface adhesion molecules, such as heparan sulfate proteoglycans, or cell lectins (Wilen et al., 2012). Binding to CD4 induces conformational changes in the Env gp120 which in turn promote its association with a chemokine receptor CXCR4 or CCR5, acting as a coreceptor molecule. The type of coreceptor selected depends on the tropism of the Env gp120. HIV-1 is classified as X4-tropic or R5-tropic depending on its recruitment of either CXCR4 or CCR5 coreceptor, respectively. Some HIV-1 strains can use both coreceptors to enter the host cell; these dual-tropic viruses are called R5X4 HIVs (Berger et al., 1998). Different levels of expression of the coreceptor on a targeted cell determine its permissiveness to X4- or R5-HIV-1 strains. In both cases, the coreceptor engagement induces further conformational changes in the transmembrane domain of Env gp41 which inserts into the host cell membrane and forms a six-helix bundle (Chan et al., 1997). The latter is the driving force that mediates fusion of the viral and cell membranes and the consecutive release of the

gRNA-carrying capsid into the cytoplasm of the target cell. The minimum number of Env, CD4 and coreceptor per Env trimer required for membrane fusion is not determined yet. Several lines of evidence suggested that HIV-1 entry can partially occur *via* endosomal pathway in several cell types, through the incorporation of virus particles into endosomes and micropinosomes (Miyauchi et al., 2009; Stein et al. 1987). This alternative way might offer several advantages to the virus, such as its sheltering from neutralizing antibodies or other inhibitors. However, only a small number of virus particles can make it through the cytoplasm, whereas most of them are degraded when delivered to early and late endosomes and lysosomes (Wilen et al., 2012).

In the cytoplasm of the newly infected cell, the viral genome starts to be converted into double stranded DNA by the viral RT and becomes part of a larger reverse transcription complex (RTC) that includes viral proteins MA, CA, NC, IN and Vpr (Fassati and Goff, 2001; Iordanskiy et al., 2006; Nermut and Fassati, 2003). Reverse transcription proceeds with synthesis of plus-strand DNA, involves a second strand transfer event and terminates at a central termination sequence in the center of the genome. The final product is a linear double-stranded DNA with a central DNA Flap (Charneau, 1991; Charneau and Alizon, 1992; Hu and Hughes, 2012). Reverse transcription was generally believed to entirely happen in the cytosol as viral DNA can be detected within few hours of infection. However, according to recent findings, reverse transcription seems to complete only in the nucleus, where it was proposed to trigger the disassembly or uncoating of the HIV-1 capsid (Dharan et al., 2020; Müller et al. 2020; Christensen et al., 2020). The rate of DNA synthesis was reported to be different between the plus- and the minus-strand, with the plus- being more rapidly synthetized, and to variate among cell types (Hu and Hughes, 2012). Specifically, in non-dividing cells, a lower level of dNTP consequently implies a reduced rate of synthesis in comparison to mitotic cells.

At some point, late in reverse transcription, the RTC is converted into the preintegration complex (PIC) through the association of the viral DNA and IN. Transported along microtubules (MTs) in direction of the nuclear periphery (Arhel et al., 2006; Malikov et al., 2015), the PIC crosses the NPCs to access the nucleus and covalently integrate the viral DNA into the host genome. Nuclear import of PIC and integration of HIV-1 genome are considered as interconnected processes; several factors involved in nuclear localization, such as the nucleoporin Nup358 (RanBP2) and the transportin-3 (TNPO3), appear to also be important for efficient integration (Brass et al., 2008; König et al., 2008).

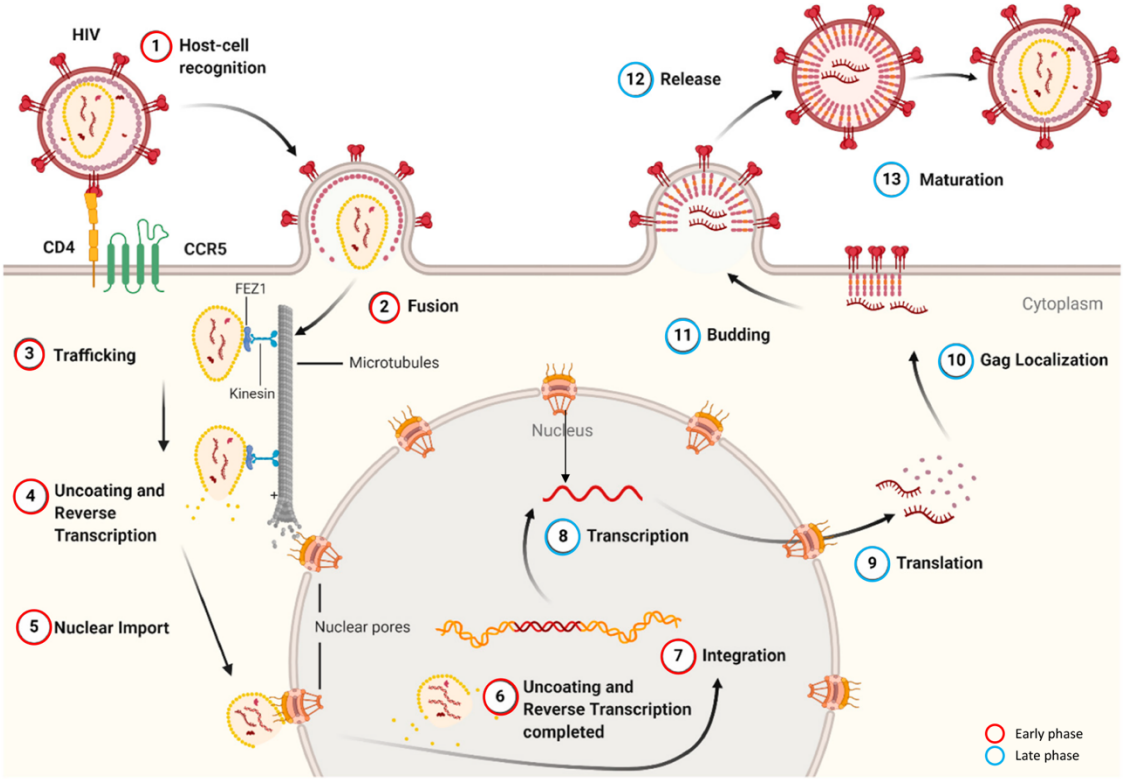


Figure 1.3. Schematic representation of HIV-1 life cycle. The viral life cycle can be mostly divided into thirteen steps: 1) the virus binds to the entry receptor CD4 and co-receptor (CXCR4 or CCR5). The conformational changes of the envelope glycoprotein induce the 2) fusion between the viral and cell membranes, followed by the entry of the viral capsid. 3) In the cytosol, HIV-1 travels along microtubules and 4) reverse transcription induces the formation of the RTC. It is unclear whether the disassembly of the viral core (uncoating) occurs at this stage. 5) Upon reaching the nuclear membrane, RTC is converted into PIC which is imported into the nucleus through the NPC. 6) In the nucleus, the double stranded pro-viral DNA is integrated within the cell genome; this can lead to either viral latency or productive infection. 7) Productive infection consists in the viral gene transcription, mRNAs export and 9) synthesis of the viral proteins. 10) The structural proteins Gag and Gag-Pol are assembled underneath the plasma membrane around the two viral RNA to produce new virions. 11) Upon assembly, the immature particle is budding off from the host cell membrane and 12) is released into the blood stream. 13) Once matured, this virion can start a new replication cycle.
Figure modified from (Rossi et al., 2021).

Viral integration can take place at various locations in the host cell genome, however HIV-1 preferentially targets a subset of genomic sequences that are localized in the heterochromatin-free zone underneath the NPCs, and strongly disfavors centrally located transcriptionally active regions and heterochromatic condensed regions in lamina-associated domains (LADs) (Lusic and Siliciano, 2017; Marini et al., 2015). Efficient integration and integration target site selection are mediated by interactions between host cell factors and viral proteins. Particularly important is the recruitment of the viral IN and CA by the host cell protein lens epithelium derived-grown factor (LEDGF/p75) and cleavage and polyadenylation specificity factor (CPSF6), respectively. Both factors drive integration into transcriptionally active spliced genes and region of open chromatin (Singh et al., 2015; Sowd et al., 2016): CA-CPSF6 interaction recruits PIC into nuclear speckles and gene bodies (Francis et al., 2020), whereas IN-LEDGF/p75 tethers it to active regions of chromosomes (Singh et al., 2015).

Once the provirus is stably inserted into the host genome, the genetic information of HIV is permanently archived in the host; viral DNA can be expressed to enable productive infection and viral replication but also be silenced and determine the viral latency. Viral latency establishes a silent HIV-1 reservoir which persist in the blood stream and has the potential to revert to an active state. Productive infection requires active transcription of the viral genome, nuclear export and processing of viral transcripts. Viral genes are transcribed along with the host DNA by the host RNA polymerase II machinery (Kinoshita et al., 1998) which is recruited at the viral promoter within the 5'-LTR by cellular transcription factors (Garcia et al., 1987). Nuclear export of viral transcripts is mediated by the nuclear export protein CRM1. In the cytosol, a single RNA can yield nearly 40 different transcript upon multiple events of splicing. In the early phase of transcription, mRNA splicing lead to the synthesis of the regulatory proteins Tat and Rev. Tat enhances the level of transcription, whereas Rev controls the late phase of gene expression by the recruitment of CRM1 and the induction of the export of late mRNAs that are un- or partially spliced. The un-spliced mRNAs correspond to the full-length genome packed inside the new virions, whereas the partially spliced mRNAs are translated into the remaining Gag, Gag-Pol and Env polyproteins. Beneath the plasma membrane, Gag and Gag-Pol polyproteins assemble into the nascent, immature virus particle that also incorporates two copies of full-length gRNA at the interior and the trimeric Env spikes at the exterior (see also 1.1.1.1-1.1.1.3 and 1.1.3.1). Membrane fission follows in order to release the nascent particle from the cell surface. Viral budding

is mediated by the constriction of the membrane induced by the viral Gag and Gag-Pol molecules, cellular proteins and membrane lipids. Final release of nascent virions is facilitated by the endosomal sorting complex required for transport (ESCRT) machinery recruited by the p6 domain of Gag (Ganser-Pornillos et al., 2019). During or shortly after budding off of the particle from the cell surface, the viral protease cleaves the Gag polyprotein precursor to trigger HIV-1 maturation (see also 1.1.1.2.).

1.1.3 Role of capsid proteins in HIV-1 infection

1.1.3.1 Gag lattice assembly in the immature virion

The immature HIV-1 virion contains between 1,200 and 2,500 Gag molecules (Carlson et al., 2008; Briggs et al., 2004) within the Gag-Pol fusion protein that arrange radially into a spherical shell beneath the viral envelope with the N-termini facing outward and the C-termini facing inward (**Figure 1.2 B**; **Figure 1.4**). The Gag-Gag molecules interact laterally *via* their CA proteins. Underneath the membrane and the MA proteins, both terminal domains of CA (CA^{NTD} and CA^{CTD}) form a spherical lattice of close-packed hexamers with an inter-hexamers distance of 8.0 nm (Briggs et al., 2004; Ganser-Pornillos et al., 2019). The lattice consists of two hexameric rings (**Figure 1.4 A-C**): one external ring directly connected to the MA proteins formed by the CA^{NTD}, and a second ring directly below formed by the CA^{CTD}. A third layer beneath the CA^{CTD} hexamers was also observed extending toward the NC/RNA region. This is composed by the residues at the CA/SP1 boundary (Ganser-Pornillos et al., 2019). Interestingly, as also known for most of the immature retroviruses, HIV-1 possesses an unusual degree of irregularity as determined by cryo-ET and subtomogram averaging (SA) of either purified immature HIV-1 or *in vitro* assembled Gag particles (**Figure 1.4 D**) (Briggs et al., 2009; Schur et al., 2015). Firstly, the CA lattice is incomplete and covers only around 60% of the spherical membrane of the virus. It is hypothesized that the Gag assembly might be competing with the budding event, which can lead to release of particles with an incomplete Gag shell. Secondly, the CA lattice contains several differently shaped defects that permit the spherical shell of the lattice curvature (Briggs et al., 2009).

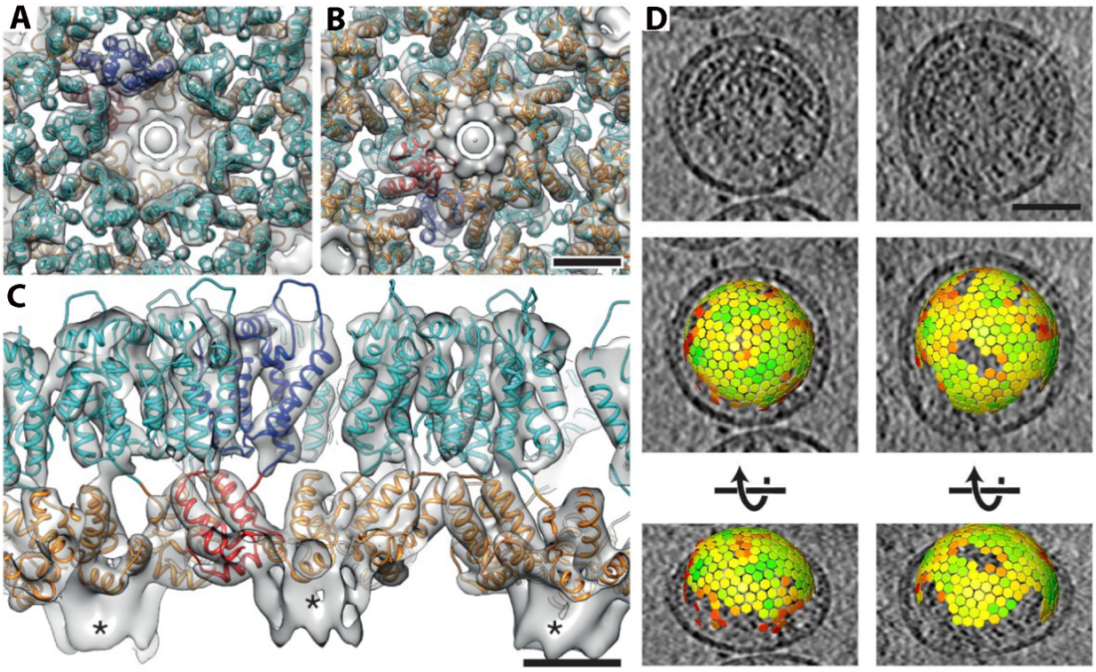


Figure 1.4. Gag lattice in the immature virion. (A-C) Isosurface representation of the Gag lattice structure with high-resolution structures of CA^{NTD} (cyan) and CA^{CTD} (orange) fitted into the density, shown from outside (A), inside (B) and in an orthogonal view (C). One CA monomer is highlighted in blue (CA^{NTD}) and red (CA^{CTD}). Asterixis in (C) show unfilled densities corresponding to the SP1 region. (D) Cryo-electron tomographic slices of immature virions and subtomogram averaging reconstruction of their immature gag lattices. The gag lattice, that localizes underneath the envelope and matrix layer, is spherical and incomplete. Scale bar: 25 Å.
Figure modified from (Schur et al., 2015).

1.1.3.2 Structure of HIV-1 capsid in the mature virion

Maturation implies significant reorganization of the HIV-1 particle, particularly the disassembly of the Gag lattice and formation of the capsid. The major role here is given to the CA protein. The PR protein cleaves the Gag polyprotein at MA/CA and CA/SP1 junctions, thus liberating the CA. Full length CA is composed by independently folded

CA^{NTD} and CA^{CTD}, connected by a flexible linker (**Figure 1.5A**) (Gitti et al., 1996). Both domains are mostly constituted of α-helices: the CA^{NTD} is shaped as an arrowhead with seven α-helices (numbered 1-7) and an amino-terminal β-hairpin, while the CA^{CTD} is globular, with a single-turn 3_{10}-helix and four short α-helices (numbered 8–11) (Gamble, 1997; Worthylake et al., 1999). After cleavage of Gag, the first 13 residues of CA refold from an extended conformation into a β-hairpin and are stabilized by a salt bridge, while the last 11 residues remain disordered (Ganser-Pornillos et al., 2007; Pornillos et al., 2010).

In order to build a capsid that will encase the viral genome, CA proteins form oligomers of 5 or 6 protein copies that spontaneously assemble into a lattice of ~250 hexamers and 12 pentamers (**Figure 1.5**). This organization promotes a cone-shaped structure with a fullerene geometry. Seven pentamers localize at the broad end of the cone and five at narrow end. Their function is to close the capsid and induce the characteristic curvature of the cone.

The CA hexamer in the mature capsid differs from the immature Gag hexamer due to the rearrangements in the intersubunit interactions mediated by the CA^{NTD} and CA^{CTD} that occur during maturation. When assembled into the mature CA hexamer, the CA^{NTD} form a six-fold symmetric inner ring located on the outer surface of HIV-1 capsid, whereas the CA^{CTD} form a "belt" surrounding the central CA^{NTD} ring, oriented towards the interior of the structure (**Figure 1.5 B and C**) (Campbell and Hope, 2015). CA^{NTD}-CA^{NTD} interactions within CA monomers are mediated through helices 1,2 and 3, which associate as an 18-helix bundle in the center of the CA hexamer (Ganser-Pornillos et al., 2007), and are facilitated by the β-hairpin at the CA^{NTD} (von Schwedler et al., 1998). The β-hairpin can adopt alternative conformations that differ by up to 15 Å as measured by comparing different available CA crystal structures (Jacques et al., 2016). In the context of the hexamer, the β-hairpins create an aperture or 'pore' on the outer surface of the capsid, that exhibit an open and close state. In the open state, the pore diameter reaches a maximum of 8 Å, sufficient to allow the transit of small molecules such as dNTPs.

The organization of the CA^{NTD} pentamers **(Figure 1.5 B)** was suggested by X-ray crystallography to be quasi-equivalent to the CA^{NTD} hexamer, with the three-helices in the center stabilized by hydrophobic interaction, but with the exception that the subunits form a five-fold symmetric ring organized around a 15-helix barrel (Pornillos et al., 2011). However, recent cryo-EM and SA analysis reported that the CA^{NTD} undergo a rotation of ~19º relative to the hexameric conformation, that excludes helix 3 from the

interprotomer interface, resulting in a central 10-helix bundle (Mattei et al., 2016). Both the crystal structure and EM map of the CA pentamer confirm the presence of the central pore formed by β-hairpins (Mattei et al., 2016; Pornillos et al., 2011).

Finally, in order to generate a continuous lattice (**Figure 1.5 C and D**), hexamers and pentamers are linked by symmetric CA^{CTD} dimers and trimers that interact side-by-side across local 2-fold axes and 3-fold axes, respectively (Gamble, 1997; Worthylake et al., 1999). This means that the assembly of the mature capsid requires two or three CA subunits linked by CA^{CTD}. The CA^{CTD} dimer interface is mediated by symmetric packing of helix 9 across the dimer interface and by packing of the 3_{10} helix of one subunit and helix 9 of the partner molecule (Gamble, 1997; Worthylake et al., 1999). On the other hand, the interactions that stabilize the CA^{CTD} trimers are not yet know in atomic details (Byeon et al., 2010).

1.1.3.2.1 Curvature of lattice

The assembly of CA into a fullerene cone requires the CA lattice to adopt different local curvatures, with a higher curvature at the tip and at one side of the cone. This is possible due to the flexibility of the CA hexamers and pentamers and depends on their localization along the surface. NMR analyses of full-length CA proteins in solution demonstrate that the CA^{NTD} and CA^{CTD} can rotate almost independently, due to the flexible linker that separates the two domains. In the lattice, however, these movements are more restricted. Cryo-ET and SA of inactivated intact HIV-1 virions (Mattei et al., 2016) allow to reconstruct the entire lattice in the capsid by the repositioning of each CA hexamer and pentamer along the surface (**Figure 1.5 D**) and to determine the tilt and twist between all pairs of neighboring hexamers. In this way, 18 conformations with different tilts from $-1 \pm 3°$ to $29 \pm 3°$, and twist angles from $-12 \pm 3°$ to $12 \pm 3°$ were identified and shown to be accommodated by small movements of CA^{NTD} relative to CA^{CTD} around the flexible linker, as well as by relative rotations around the CA^{CTD} dimer and trimer interfaces.

As already mentioned, the pentamers localize at the broad and narrow end of the cone and introduce a higher local curvature in this area of the cone. This is possible because the pentamer-hexamer CA^{CTD} dimer interface is moved closer together than the hexamer-hexamer CA^{CTD} dimer interface. Changes in local curvature do not only determine the shape of HIV-1 capsid but have also an important impact on the capacity of capsid to

bind different cellular factors that might enhance or inhibit HIV-1 infection, such as cyclophilin A (CypA), tripartite motif-containing proteins (TRIMs), and myxovirus resistance protein B (MxB) (Ni et al., 2020).

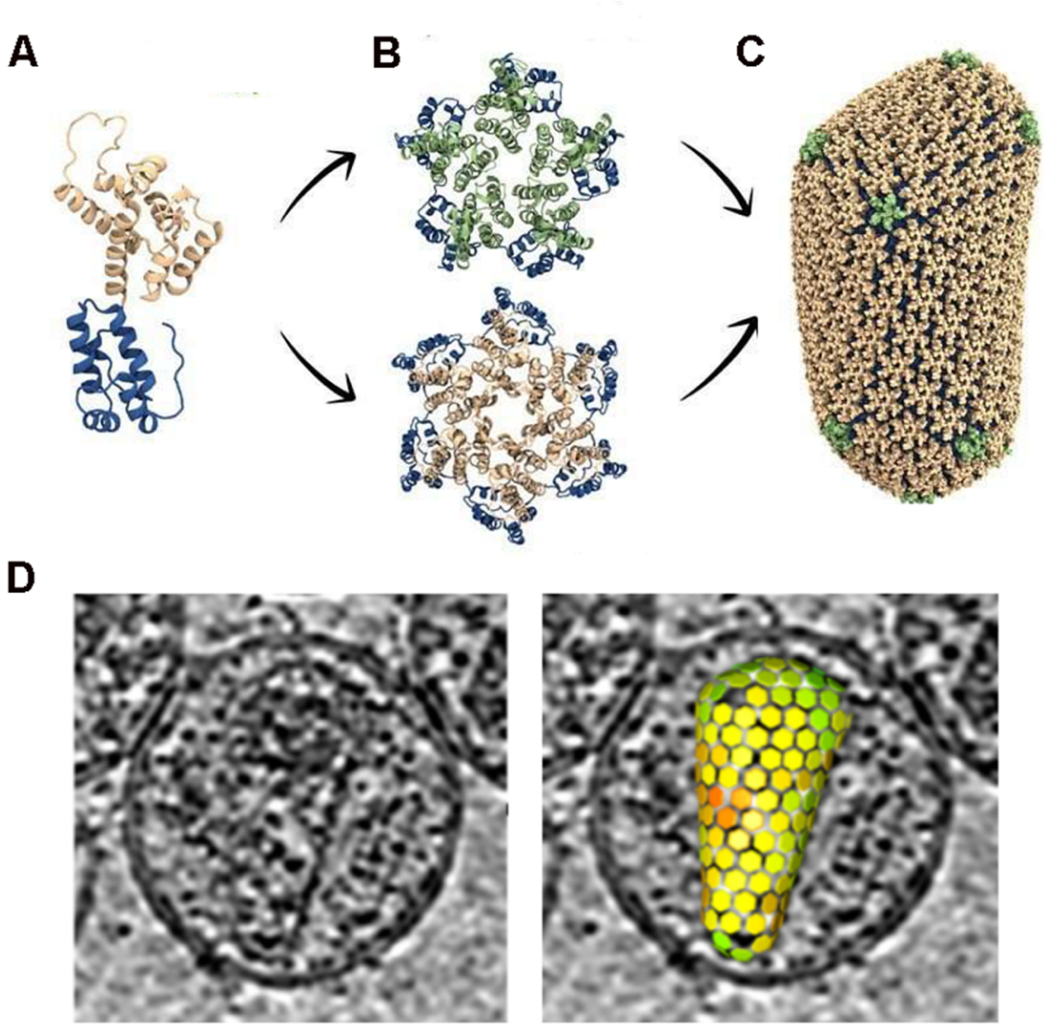


Figure 1.5. Structure of the HIV-1 conical capsid and its components. (A) CA monomer, composed of the CA^{NTD} (beige) and the CA^{CTD} (blue) domains. (B) CA pentamer (top) and CA hexamer (bottom). CA^{NTD} pentamer is shown in green, CA^{NTD} hexamer in beige, CA^{CTD} in blue. (C) CA hexamer and CA pentamers arrangement into the HIV-1 capsid core. (D) Mature virion imaged on EM grids by cryo-ET and its lattice reconstructed by subtomogram averaging.

Figure 1.5 A-C modified from (Quinn et al., 2018); Figure 1.5 D modified from (Mattei et al., 2016).

1.1.3.3 Capsid uncoating

During the HIV-1 life cycle, the capsid shell that encases the viral genome has to disassemble, at least partially, in order to guarantee the release and integration of the viral genome (Ambrose and Aiken, 2014). This process is known as uncoating. Timing, cellular location and extent of CA uncoating during infection are still not clear and might be cell type specific.

Early studies on uncoating were mostly based on biochemical and capsid assembly assays, in which the replication complexes were purified, subjected to ultracentrifugation and detected for the presence of CA proteins (Fassati and Goff, 2001; Campbell and Hope, 2015). More recently, different imaging strategies were developed to track replication complexes inside the infected cell. Multiple viral proteins, including IN and CA proteins, but also CA-binding host factors such as CypA, were targeted for labeling of early replication complexes (Bejarano et al., 2019; Burdick et al., 2020; Francis and Melikyan, 2018; Mamede et al. 2017; Novikova et al., 2019; Peng et al., 2014; Stultz et al., 2017).

In both cases, investigation of early post-entry events is very challenging: first, capsids might not be stable enough during *in vitro* purification steps; second, only a minority of the incoming particles undergo productive infection, whereas the large amount of abortive infection events may cause noise signal that compromises the final readout *in vivo* (Novikova et al., 2019). Consequently, these studies showed high discrepancies that led to the formulation of three main models of uncoating (Campbell and Hope, 2015): i.) rapid core uncoating in the cytoplasm prior to reverse transcription and nuclear import of viral DNA, (ii) partial uncoating in the cytoplasm and co-import of a considerable part of CA proteins and (iii) partial uncoating at the NPC with subsequent nuclear import of viral DNA. A common note of the three uncoating models is to consider it as very unlikely that the capsid core remains intact during its translocation through an intact nuclear pore due to the differences in size between the broad end of capsid core and the nuclear pore inner diameter, which are reported to be ~60 nm and ~40 nm, respectively (von Appen et al., 2015; Mattei et al., 2016). Accumulating evidences in the last years, such as detections of variable amounts of CA in the nucleus of different cell types (Bejarano et al., 2019; Burdick et al., 2020), suggested that CA constitutes a component of the RTC/PIC complexes and excluded the first model of uncoating once per all. In addition, the identification of multiple host factors that CA hijacks in the HIV-1 targeted

cells, indicated that CA play a crucial role during early stages of infection, including cytoplasmic trafficking, protection of RTC/PIC against innate immune sensing, nuclear import of viral genome and integration (Ambrose and Aiken, 2014; Campbell and Hope, 2015; Hilditch and Towers, 2014; Yamashita and Engelman, 2017). However, the extent of assembly and organization of CA proteins in RTC/PIC compared to the CA lattice in mature virion (see 1.1.3.2) remains unknown. A complete understanding of HIV-1 capsid disassembly might be critical for therapeutics development as an intact capsid core may represent a fundamentally different therapeutic target than a small amount of CA that remain associated with the RTC/PIC during infection, due to higher affinity of host proteins and antiviral drugs to the assembled form of CA.

1.1.3.3.1 Determinants of uncoating

Although the mechanism of uncoating remains unclear, several interactions or events within the viral capsid core are reported to inhibit or enhance uncoating.

The host peptidyl prolyl cis-trans isomerase (cyclophilin) CypA is one of the first identified and most characterized interactors of CA (Bossolt et al., 1993). It can be incorporated into virions during assembly, but mostly interacts with HIV-1 CA core upon cytoplasmic entry (Campbell and Hope, 2015; Franke and Yuan, 1994). Although its role was ambiguously interpreted upon infection of different cell types, it is now accepted that CypA is responsible for stabilizing the capsid core and thereby facilitates HIV-1 infection (Liu et al., 2016). CypA binding to CA is enabled by the intrinsic curvature of CA hexamers and occurs at the flexible CA^{NTD} CypA-binding loop that protrudes from the capsid surface (Ni et al., 2020). Mutations in the CA at the CypA-binding loop (P90A) prevent CypA-CA interactions, which alternatively can also be affected by treatment of infected cells with ciclosporin A (CsA) (Braaten and Luban, 2001; Hatziioannou et al., 2005; Sokolskaja et al., 2004; Towers et al., 2003). The stabilization of the CA core mediated by CypA may be dependent on its ability to prevent the binding of some restriction factors to CA, such as TRIM5α (Kim et al., 2019; Selyutina et al., 2020), which in lymphocytes is known to inhibit HIV-1 infection by accelerating core dissociation and preventing reverse transcription (Wagner et al., 2020).

Contrary to CypA, other factors involved in PIC nuclear import, such as Nup358 and the β-importin transportin-1 (TRN-1), are suggested to work as enhancers of CA uncoating.

Nup358-mediated uncoating might be dependent on its cyclophilin-homology domain. This domain is suggested to catalyze the isomerization of CA proteins to destabilize the CA core and generate a complex with dimensions capable of translocating though the nuclear pore (Dharan et al., 2016). TRN-1 is proposed to trigger uncoating based on the observations that tubular capsids are structurally damaged upon TRN-1 binding *in vitro* and that depletion of TRN-1 impairs nuclear entry and causes accumulation of CA proteins *in vivo* (Fernandez et al., 2019). However, the mechanism through which this is happening remains unclear.

Finally, physical forces rather than CA-host factors interactions are also suggested to affect core stability and facilitate uncoating. The complete synthesis of double-stranded DNA upon reverse transcription, for example, seems to induce stress on the capsid structure from within the core. This is supported by the finding that inhibitors of reverse transcription lead to capsid stabilization (Yang et al., 2013) and is in line with a very recent study *in vitro* where the authors observed partially broken HIV-1 capsids with polynucleotide loop emanating from the capsid lattice when cDNA synthesis was largely completed (Christensen et al., 2020). Initially thought to occur in the cytoplasm upon viral entry, reverse transcription is recently reconsidered to happen in the nucleus of the infected cell (Dharan et al., 2020), suggesting a spatiotemporal link between reverse transcription and uncoating. In conclusion, uncoating might be considered as a complex time-regulated process that relies on different cellular factors rather than a single event.

1.1.3.4 Capsid-mediated nuclear import

Nuclear entry of PIC follows cytoplasmic trafficking and occurs through the NPC in non-dividing cells. Considered as a very rare event with only a small fraction of incoming virions (1:200 pseudo typed viral particles) that can enter the nucleus (Yamashita and Engelman, 2017), HIV-1 nuclear entry is particularly challenging to study. Although several viral proteins and DNA flap were reported to facilitate nuclear import, a central role in this process is assigned to CA proteins in the PIC due to their ability to hijack several host proteins and the nucleocytoplasmic transport machinery (**Figure 1.6**) (Campbell and Hope, 2015). Among these, several nucleoporins (Nups) composing the NPCs, such as Nup62, Nup85, Nup98, Nup107, Nup133, Nup153, Nup214 and Nup358/RanBP2, the nuclear receptor factors TRN-1 and TNPO-3, and CPSF6, were

identified as CA interactors (Brass et al., 2008). Notably, the redundancy of Nups found to bind CA and possibly assisting PIC translocation through the nuclear pores led to further exploration of their role in HIV-1 infection. This downsized the number of Nups involved in nuclear entry to only two. Nup358 and Nup153 were found to be essential for HIV-1 nuclear import (Di Nunzio et al., 2012). Nup358 and Nup153 are components of the nuclear pore filaments, which project towards the cytoplasm and nucleoplasm from the NPC, respectively. TRN-1 and TNPO-3 are β-importins that normally bind cargoes without adaptors in the cytoplasm, target them to the nucleus through the NPC and release them in the nucleus following RanGTP binding. CPSF6 is a mRNA processing protein that shuttles among the cytoplasm and the nucleus but has a predominant localization in the nucleus (Rüegsegger et al., 1998). Recruitment of these factors mostly occurs at two binding interfaces in CA: i.) the Cyp-binding loop localized at the CA^{NTD}, where Nup358 and TRN-1 bind to, and ii.) the CA^{NTD}-CA^{CTD} interface of the hexamer, used by Nup153 and CPSF6 (**Figure 1.7**). In particular, TRN-1 recognizes the nuclear localization signal at the CypA-binding loop in CA (Fernandez et al., 2019). Mutations in the conserved Cyp binding loop (P90A) or in the binding pocket of assembled CA (N74D), and/or depletion of these factors cause a disruption of HIV-1 nuclear import.

Nuclear import of PIC relies on active transport (Bukrinsky et al., 1992) and requires three major steps to cross the NPC: docking to the NPC, translocation through the central channel of NPC and release into the nucleoplasm. Nup358 may mediate docking of PIC to the cytoplasmic face of NPC through a mechanism that is not fully understood (Dharan et al., 2016; Di Nunzio et al., 2012). According to one hypothesis, Nup358 migrates on microtubules anterogradely from the NPC to the cytoplasm in a kinesin-1 KIF5B motor dependent mechanism. Upon recruitment of CA, Nup358 moves retrogradely along microtubules to the NPC (Dharan et al., 2016). Translocation of HIV-1 complexes through the central channel remains unclear from a mechanistic and logistic points of view. First, the two nucleoporins of NPCs central channel identified in the genome-wide screening, Nup98 and Nup62, are reported to be mostly involved in the integration of viral genome rather than nuclear translocation (Nunzio, 2013), leaving open some questions on which nucleoporins mediate the actual HIV-1 import and which mechanisms are used in the process. Second, as previously mentioned, the CA core has a diameter of ~60 nm that exceeds in size the central channel of NPCs (~40 nm). This suggests that CA has to at least partially disassemble to enter the central channel of NPCs or alternatively NPC may disrupt by Nup358 displacement to increase the ability of the

viral core to transverse the nuclear pore (Dharan et al., 2016). Once viral complexes have reached the nuclear basket of the NPC, Nup153 may promote the final steps of translocation in the nucleus (Campbell and Hope, 2015; Matreyek et al., 2013; Price et al., 2014).

CPSF6, found highly enriched at the nuclear periphery and in nuclear speckles in different cell types (Bejarano et al., 2019; Francis et al., 2020), is suggested to release PIC in the nucleus by competing for the same binding site of Nup153 in CA. Nuclear localization of CPSF6 seems to be dependent on TNPO3, which may impact PIC nuclear transport directly through its binding to CA, or indirectly by binding CPSF6 at the C-terminal nuclear targeting arginine/serine-rich (RS-) domain (Price et al., 2012). CPSF6 lacking the RS-domain (CPSF6-358) loses its dependency on TNPO3 and is no longer confined in the nucleus; in the cytoplasm CPSF6-358 restricts HIV-1 before nuclear entry. Mutations in CA proteins (A77V) or depletion of CPSF6 cause accumulation of viral complexes at the nuclear side of nuclear pore complexes and integration into transcriptionally repressed, lamina associated heterochromatin (Achuthan et al., 2018; Bejarano et al., 2019; Zila et al., 2019).

In contrast, several host factors do not promote but rather inhibit nuclear import by interacting with CA. For example, MxB, a dynamin-like GTPase, concentrates at the cytoplasmic side of the NPCs and acts as CA-targeting HIV-1 restriction factor that blocks nuclear entry and post-entry events such as integration (Haller, 2015). Interestingly, the N74D CA mutation does not reduce MxB binding to CA assemblies, as reported for CPSF6 and Nup358, but renders HIV insensible to MxB restriction. This happens because MxB mediates its restriction on nuclear import by recruiting and displacing Nup358 from CA binding in a mechanism that uses CPSF6 as co-factor (Xie et al., 2020). Two other inner nuclear membrane proteins SUN1 and SUN2 are reported to inhibit nuclear import and integration (Luo et al., 2018). Thus, nuclear entry is a complex series of events mediated by multiple viral and host factors, and its completion may depend on the balance between enhancing and inhibitory interactions.

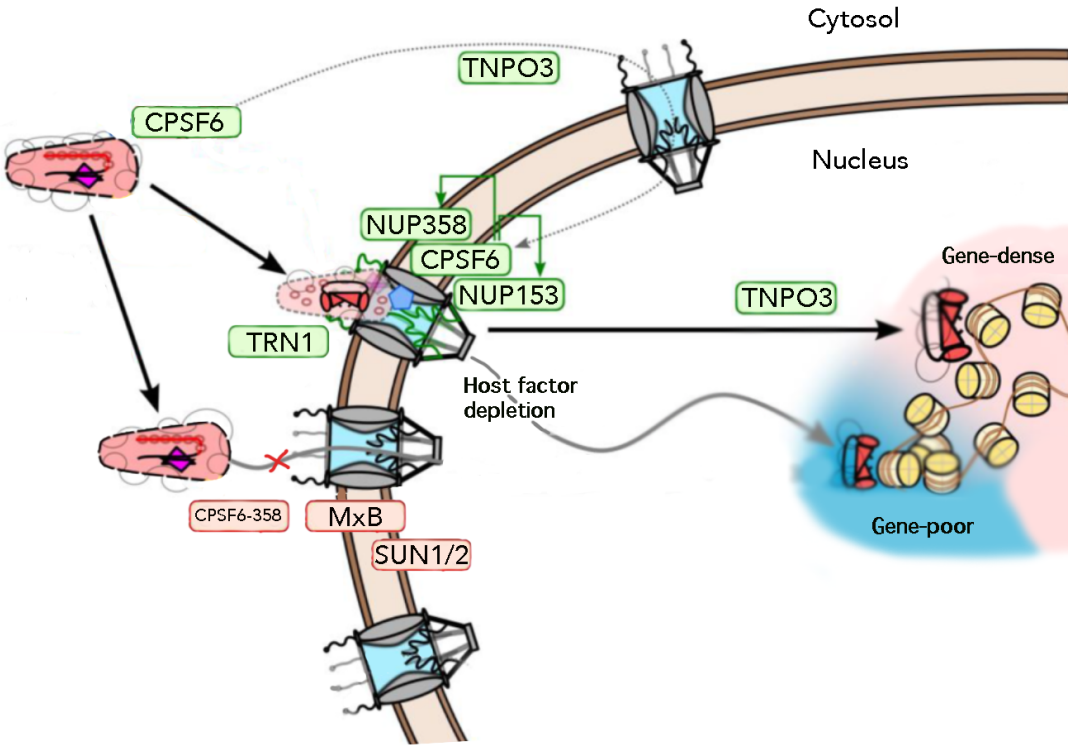


Figure 1.6. Schematic representation of CA-host factors interactions during PIC nuclear entry. NUP358, CPSF6, NUP153, TRN-1 and TNPO3 (green rectangles) bind CA in PIC and facilitate PIC nuclear import across the NPC and viral genome integration into gene-dense chromatin. TNPO3 is also required to localize CPSF6 to the nucleus. Depletion of these factors, such as CPSF6, causes impairment of nuclear entry and viral integration into gene-poor regions of the host genome. Truncated CPSF6-358, MxB and SUN1 and SUN2 proteins are inhibitors of PIC nuclear entry (red rectangles).
Figure modified from (Matreyek and Engelman, 2013).

1.1.3.5 Capsid as therapeutical target for antivirals

HIV-1 CA may be considered as a suitable therapeutical target due to its role in HIV-1 infection. Hence, the importance of understanding the uncoating process and the role of assembled CA during several steps of infection, such as nuclear entry, is connected to the development of new inhibitors of HIV-1 infections that may affect the core stability and the ability of CA to interact with critical host factors during infection. Multiple CA inhibitors

affecting different stages of HIV replication cycle were developed and became commercially available (Carnes et al., 2018). Some examples of compounds that interfere with uncoating and nuclear entry are briefly described here. The first CA-targeting inhibitor, commonly known as CAP-1, was identified in 2003. CAP-1 binds to CA at the base of the CA^{NTD} near the linker region and is reported to alter its conformation. As result, CAP-1 inhibits the ability of CA to self-assemble *in vitro* and causes the production of morphologically altered cores in the viral particles. *In vivo*, CAP-1 inhibits viral maturation and reduces HIV-1 replication by 95%. Another example is Pfizer compound PF-3450074, generally known as PF74. It was originally identified in 2010 in a high-throughput screen as a small molecule inhibitor of HIV-1 infection that destabilizes the capsid core *in vitro*. This compound targets the same binding pocket of CPSF6 and Nup153 to assembled CA (FG-binding pocket in **Figure 1.7**). At high concentrations (10 μM), PF74 interferes with early and late events in the virus life cycle, leading to premature uncoating, inability of reverse transcription and disruption of particle formation. Interestingly, PF74 has an opposite effect at low concentration (~2 μM) as it stimulates the rate of CA self-assembly *in vitro* and stabilizes preassembled CA-NC tubes. BI-2 compound produced by Boehringer Ingelheim binds to the same site in CA as PF74. Contrary to PF74, BI-2 does not block reverse transcription but rather inhibits HIV-1 nuclear entry by interfering with CA binding to CPSF6 and Nup153. Most recently, a new CA-targeting compound, named GS-CA1 (Link et al., 2020), was described and shown to exhibit high antiviral potency in human peripheral blood mononuclear cells with broad spectrum inhibition across all HIV clades. GS-CA1 exploits the same CA binding site as described for PF74 and BI-2, but exhibits a greater potency in comparison to the other two compounds. Accumulation of knowledge on crucial steps in HIV infection might strongly impact the identification of new CA-targeting HIV inhibitors for AIDS therapy.

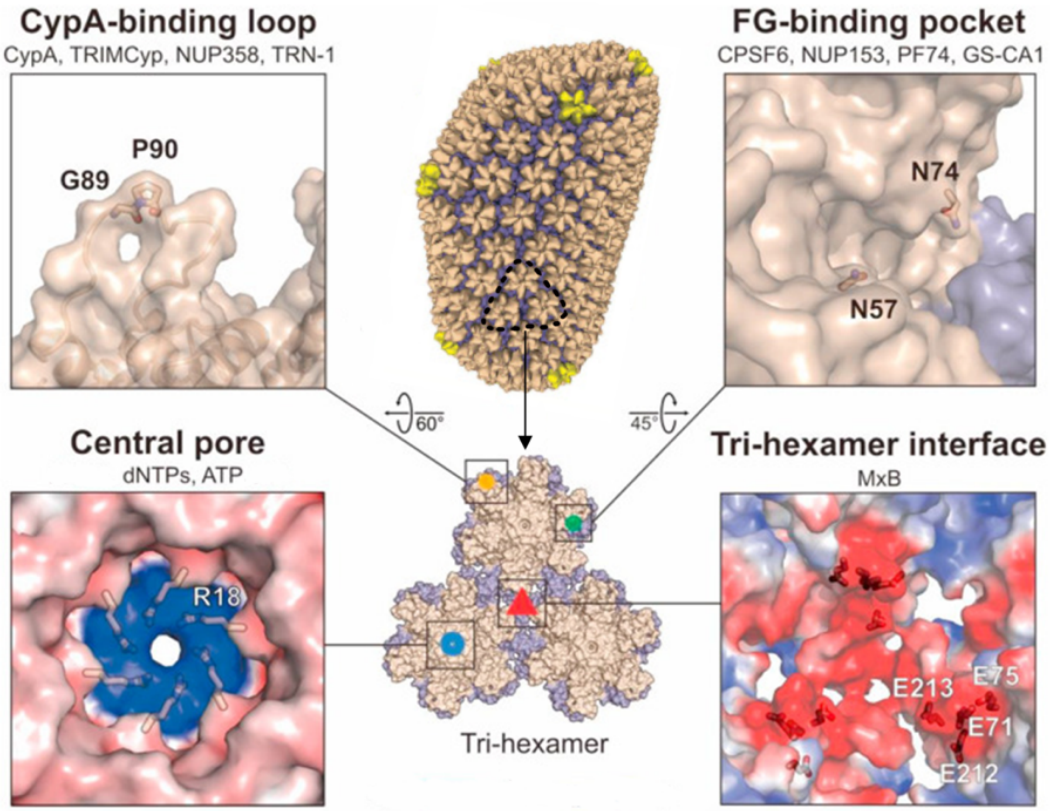


Figure 1.7. Binding interfaces of HIV-1 capsid with host factors and inhibitors.

The tri-hexamer is highlighted from the initial structure of the HIV-1 lattice and the localization of four different binding sites for several host factors are shown. The Cyp-binding loop, a binding platform for CypA, TRIM proteins, Nup358 and TRN-1, is localized at the CA^{NTD} of the hexamer. The FG-binding pockets at the CA^{NTD}-CA^{CTD} interface of the hexamer is shared by the host cellular factors Nup153 and CPSF6, and the CA inhibitors PF74 and GS-CA1. The central pore is formed by the β-hairpins in the CA^{NTD} of CA hexamer; dNTPs and ATP can bind to the central pore in its open conformation. MxB binds at the trimer interface and acts as inhibitor of HIV-1 nuclear entry. Key residues are shown for each binding site.

Figure modified from (Temple, 2020).

1.2 The nuclear pore complexes

NPCs are large macromolecular assemblies embedded into the nuclear envelope (NE), that fuse the outer- and inner nuclear membranes (ONM, INM) and regulate bidirectional nucleocytoplasmic transport. They represent the largest protein complexes occurring in eukaryotic cells: in humans, they have a molecular mass of ~120 MDa.

Each NPC is composed of ~1000 copies from ~30 different Nups, most of which are conserved across eukaryotes (Beck and Hurt, 2017). Their sizes range from a molecular mass of ~30 to 358 kDa. In absence of a unifying nomenclature, the molecular mass is often used to name each Nup (Lin and Hoelz, 2019). Depending on their structural properties, they can be categorized into scaffold Nups, which primarily contain folded protein domains and constitute the stable architectural scaffold, and phenylalanine-glycine-Nups (FG-Nups), which are anchored to the scaffold Nups and carry intrinsically disordered domains rich in FG-repeats. Several scaffold Nups also contains FG-repeats, suggesting that the subdivision of Nups in these two classes is not always clear-cut (Beck and Hurt, 2017). Multiple Nups assemble into a number of distinct substructures, defined as subcomplexes or modules which in turn compose the overall NPC architecture (**Figure 1.8**). The NPC structure, also conserved, displays an eightfold rotational symmetry (Gall, 1967) and includes the following principal structural elements: the inner pore ring (IR), which resides at the fused ONM-INM, the cytoplasmic and nuclear ring (CR, NR) or outer rings, which are anchored to the inner pore ring. The nuclear basket is bound to the NR and projects into the nucleoplasm. Unstructured FG rich repeat domains of the FG-Nups form the cytoplasmic filaments, which are peripheral elements emanating from the nuclear and cytoplasmic rings and projecting towards the central channel. The central channel is rich in FG-Nups and creates a unique milieu with similar biophysical properties of phase-separated, membraneless organelles. Within this environment, cargo proteins, such as transcription factors, histones, polymerases, chromatin binding proteins, assembled ribosomal subunits, mRNA and also viruses or viral proteins, continuously shuttle between the nuclear and the cytoplasmic compartments.

1.2.1 Modular architecture of the human NPCs

The modularity of the NPCs is referred to the ability of several Nups to build multiple subcomplexes that in turn are repeatedly arranged into CR, IN and NR to compose the basic architecture of NPCs. The structure of the human NPC was solved at the resolution of ~20-Å using cryo-ET of intact NPCs in NEs (von Appen et al., 2015). Details in the composite structure of the human NPC were obtained through fitting of human homology models or structures into the human NPC cryo-ET density map (von Appen et al., 2015; Kosinski et al., 2016). This was consistent with the high-resolution biochemical nucleoporin-nucleoporin interaction mapping of fungal nucleoporins (Lin et al., 2016) and with the cross-linking mass spectrometry analysis of human and fungal nucleoporin interactomes (von Appen et al., 2015; Kosinski et al., 2016). The different subcomplexes composing each ring are briefly described in the following paragraphs.

1.2.1.1 Inner ring scaffold: Nup93 and Nup62 subcomplexes

The IR of the human NPC has been largely characterized through a combination of cryo-ET, cross-link mass spectrometry and molecular modelling (von Appen et al., 2015; Bui et al., 2013; Kosinski et al., 2016). It consists of two subcomplexes, namely the Nup93 and Nup62 subcomplexes. The Nup93 subcomplex comprises several Nups, such as Nup53, Nup93, Nup155, Nup188/Nup205 and Nup98, whereas the Nup62 subcomplex, also referred to as channel nucleoporin heterotrimer (CNT), includes three FG-Nups, Nup54, Nup58 and Nup62.

Both IR subcomplexes are repeated in four copies (core modules) in each asymmetric unit (spoke ring) of the NPC (**Figure 1.9A**), whereby binding of Nup188 and 205 is mutually exclusive resulting in 16 copies, for the others in a total of 32 copies in each IR. Further 16 copies of Nup155 connect the IR to the outer rings.

The four core modules symmetrically organize across the NE plane in two cytoplasmic and two nuclear core modules (C2 symmetry). Each of them is in turn arranged in an outer and an inner copy parallelly oriented but rotationally shifted to each other (top of **Figure 1.9B**). Within each core module, Nup155 is anchored to the NE via the N-terminal β-propeller. Similarly oriented, Nup93 interacts with Nup155 and recruits the Nup205/Nup188 complex, which is placed in a horizontally stacked position on top of

Nup155 and Nup93. This arrangement works as platform for the recruitment of the Nup62 subcomplex. The Nups within the Nup62 subcomplex assemble in triangular triple coiled-coil arrangement and project their FG-repeats toward the central channel (bottom of **Figure 1.9B**).

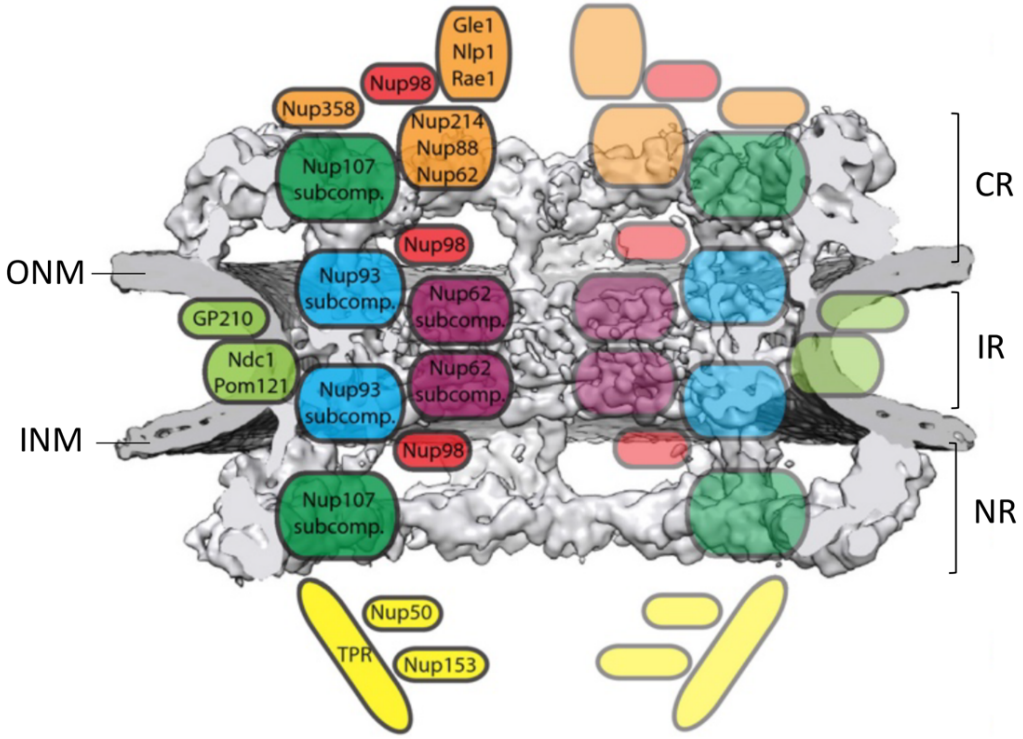


Figure 1.8. Modular architecture of the human NPC. Scheme showing the approximate localization of the subcomplexes with respect to the overall NPC structure. The NPC is embedded into the nuclear envelope and fuses the outer (ONM) and inner nuclear membranes (INM). The overall architecture is composed of three rings: the inner ring (IR), consisting of the Nup93 and Nup62 subcomplexes, and the cytoplasmic ring (CR) and nuclear ring (NR), comprising the Nup107 subcomplex. Nuclear basket shown in yellow.

Figure modified from (Mackmull et al., 2017).

1.2.1.2 Outer rings scaffold: Y-complexes

The outer rings are primarily composed of several nucleoporins that arrange in a head-to-tail fashion to form a Y-shaped structure with a long and short arm followed by a tail. The derived complex is therefore mostly known as Y-complex, but also referred to as the Nup107-Nup160 complex or coat nucleoporin complex (CNC). It consists of 10 different Nups such as Sech13, Seh1, Nup96, Nup75, Nup107, Nup160, Nup133, Nup37, Nup43 and ELYS (Cronshaw et al., 2002; Lin and Hoelz, 2019). The short arm of the Y-complex is composed of Nup85 and is consequently known as Nup85-arm. Nup85 is the binding partner of two other Nups, Seh1 and Nup43. The Seh1 structure consists of an incomplete β-propeller containing six blades out of seven. Nup85 provides a sixth β-propeller blade and complete Seh1 β-propeller. In the assembled NPC, the Nup85-arm points toward the central channel NPC. The large arm of the Y-complex, also known as Nup160-arm, is composed of Nup160 and the two binding partners Nup37 and ELYS. ELYS is localized exclusively in the Y-complexes of NR and in metazoan is believed to recruit the Y-complexes to chromatin to initiate postmitotic NPC assembly.

In the assembled NPC, the long arm points toward the periphery in opposite direction of the short arm and it binds the NE via the β-propeller of Nup160. Attached to Nup160, Nup96 is located at the vertex of the Y-complex and interacts with Sec13 in a similar insertion blade mechanism reported for Nup85 and Seh1. Finally, the Y-complex tail follows Nup96 and consists of Nup107 and Nup133. Nup133 binds the Y-complex tail to the NE via its N-terminal β-propeller.

Both cytoplasmic and nuclear rings in the human NPC count in total 32 Y-complexes, divided in a 16 Y-complexes CR and NR (Bui et al., 2013). Each spoke of NPC thus contains four Y-complexes with two of them belonging to the cytoplasmic ring and the other two to the nuclear ring. Nups in the adjacent spoke connect in a head-to-tail fashion and generate a continuous ring through the interactions mediated by Nup133 and Nup160 at the inter spoke interfaces.

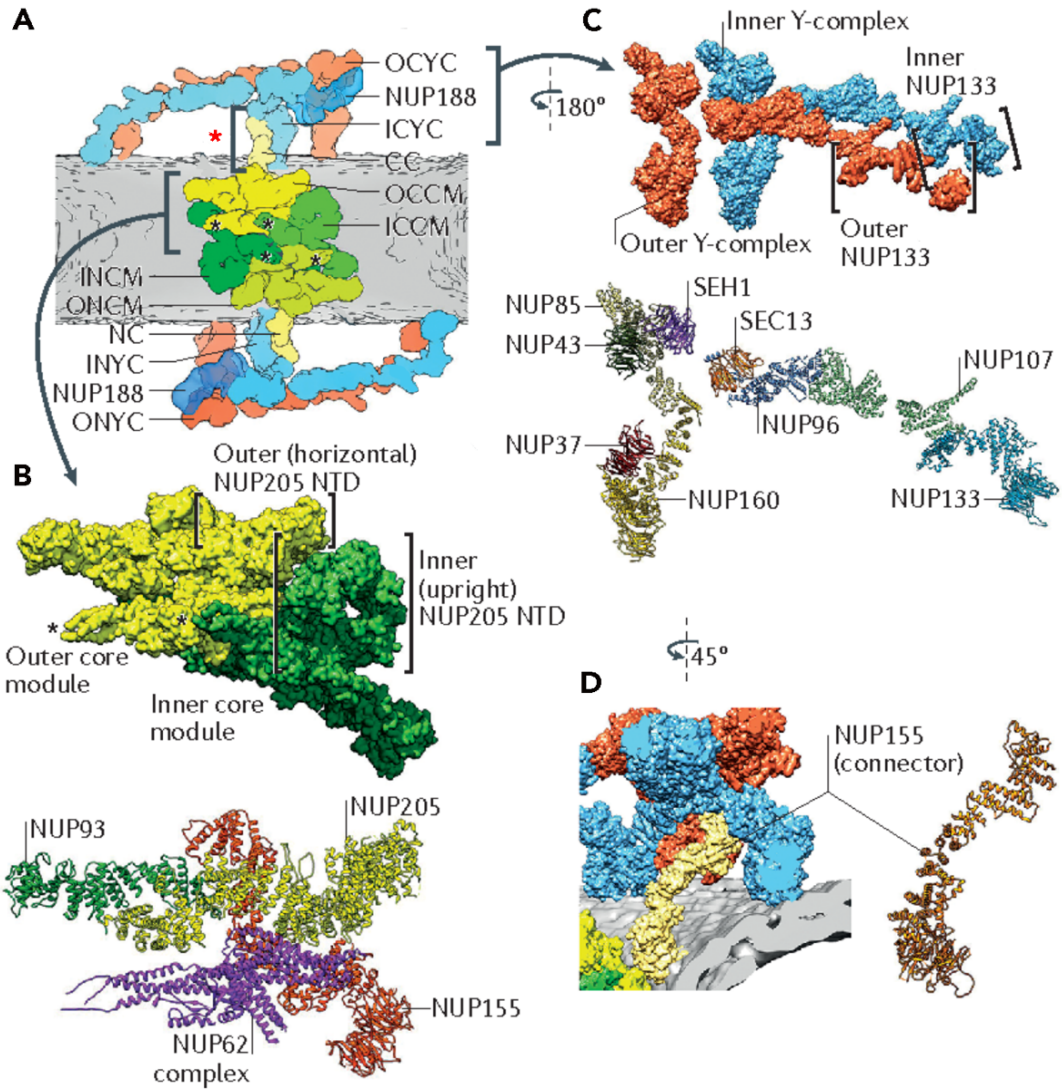


Figure 1.9. Scaffold subcomplexes in the asymmetric unit of NPC. (A) Asymmetric unit of the NPC with (B-D) molecular details of each subcomplex and arrangement of their components: inner and outer inner ring core module (B), inner and outer Y-complex (C), Nup155 connector (D) pointed by the red asterisk in (A).
Figure modified from (Beck and Hurt, 2017).

1.2.1.3 Cytoplasmic filaments

Beside the Y-complexes that build the main structural scaffold, the cytoplasmic ring of the NPC is further composed of two complexes that are considered part of cytoplasmic filaments and are absent in the nuclear ring. These are the Nup358 and Nup214 complexes (von Appen et al., 2015). The cytoplasmic filaments are constituted by the FG-repeat of several FG-Nups and play an important role in the nucleocytoplasmic transport through their interaction with multiple NTFs.

Nup358 is the largest Nup in the NPC (358 kDa). Purified Nup358 forms ~36 nm long, 5 nm thick filaments (Delphin et al., 1997). In humans it comprises multiple domains connected by flexible linkers: a N-terminal α-helical domain, four Ran-binding domain (RanBDs), eight zinc-finger motifs, several FG-repeats, a sumo E3 ligase domain and a C-terminal CypA-homologous domain (Wu et al., 1995; Yokoyama et al., 1995). Except for the N-terminal α-helical domain, all the others have been structurally characterized. Interestingly, the C-terminal end of Nup358, which in the context of HIV-1 infection is responsible for CA binding, can extent towards the cytoplasm up to ~50 nm from the NPC. This suggests that Nup358 is anchored to the CR via its N-terminal end, where it stabilizes the interactions between the inner and outer Y complexes. Depletion of Nup358 induces loss of the outer but not inner Y-complexes in CR (von Appen et al., 2015). Beside its structural role in the organization of CR and the stabilization of the Y-complex dimers, Nup358 provides binding sites for RanGDP/GTP in the cytosol, participates in protein import and export, and regulates mRNA export (Lin and Hoelz, 2019). Nup214 together with Nup88, Nup62, Rae1 and Nup98 forms the Nup214 complex, also known as P-complex, which represents the mRNA export platform (von Appen et al., 2015). Nup214 anchors the complex to the CR scaffold by interacting with the small arm of the Y-complex. In contrast to Nup358, gene silencing of Nup214 does not affect Y-complexes assembly (Bui et al., 2013). Nup88 works as a bridge between Nup214 and Nup98. Nup98 is one of the few Nups that has a stoichiometry of 48 copies instead of 32 in the human NPC. It contains a small structured domain at the C-terminus, that anchors it to Nup88, and a larger N-terminus that harbors the FG-repeats and a highly conserved Gle2-binding sequence that serve as docking for Rae1 and mRNA export factors. Disruption of the interaction between Rae1 and Nup98 leads to mRNA export defects.

1.2.1.4 Nuclear basket and nucleoplasmic filaments

The nuclear basket is located on the nuclear side of the NPC. It interacts with the nucleocytoplasmic transport machinery and provides an anchoring site for the mRNA export and quality control factors. Beside the nucleocytoplasmic transport, nuclear basket is also involved in other non-canonical functions such as chromatin organization, gene expression regulation and DNA repair (Burns and Wente, 2014; Köhler and Hurt, 2007; Krull et al., 2010). Despite its importance, a proper characterization of its structure remains elusive due to the high flexibility of its components. The nuclear basket is composed of three Nups: Nup50, Nup153 and TPR. Nup153, similar to Nup358, contains a tandem array of four Ran-binding zinc-finger domains and a series of FG-repeats in the C-terminal end (Partridge and Schwartz, 2009). TPR consists of a massive coiled-coils region at the N-terminus that interacts with Nup153 and a disordered region at the C-terminus (Hase et al., 2001). Last, Nup50 contains an N-terminal domain that binds Nup153, FG-repeats in the middle and a C-terminal Ran binding-domain (Matsuura and Stewart, 2005). Hence, Nup153 is responsible for the recruitment of both Nup50 and TPR to the nuclear basket and the anchoring of nuclear basket to the nuclear Y-complexes (Walther, 2001). TPR, on the other hand, extends toward the nucleus through its disordered region and represents the main component of the nucleoplasmic filaments.

1.2.2 Nucleocytoplasmic transport

In this paragraph, the nuclear transport system and its molecular mechanism is briefly described, according to the following reviews (Cook et al., 2007; Görlich, 1999; Hampoelz et al., 2019; Mattaj and Englmeier, 1998; Nakielny and Dreyfuss, 1999; Strambio-De-Castillia, 2010).

It was roughly estimated that more than 1 million macromolecules are transferred between the nuclear and the cytoplasmic compartments each minute in a growing mammalian cell. Small molecules < 30-60 kDa (9 nm in diameter) can freely diffuse through the NPC by passive transport, whereas larger complexes are actively transported through the NPC by the energy-dependent nucleocytoplasmic transport system. The nucleocytoplasmic transport system can be divided into a stationary phase and a soluble phase. The first is represented by Nups in the central channel, whereas the second consists of ~30 different mobile nuclear transport factors (NTFs) that transiently interact with the

FG-Nups and move between the nuclear and the cytoplasmic compartments. The last category comprises the nuclear transport receptors (NTRs) or karyopherins, that bind to cargos and mediate their interaction with FG-Nups, and the auxiliary NTFs such as the small GTPase Ran. Depending on the directionality of transport the NTRs are named as importins, exportins, and transportins. Most of importins and exportins belong to the family of karyopherin β also known as importin β. Importin β recognizes specific sequences on the cargo proteins known as nuclear localization signals (NLSs) or nuclear export signals (NESs) in order to facilitate their import into or export from the nucleus, respectively.

Other NTRs belong to the importin α family and often function as adaptor proteins and form a ternary import complex together with importin β and the cargo (Görlich et al., 1995). The small GTPase Ran fuels the active nuclear transport cycle as it establishes and maintains a GTP-GDP gradient across the NE by actively shuttling through the NPC. In the cytoplasm, the Ran GTPase-activating protein (RanGAP) promotes RanGTP hydrolysis leading to a high enrichment of RanGDP in this compartment. Association with the RanGDP recycling factor (NTF2) induces import into the nucleus. In the nucleus, the chromatin-bound Ran guanine nucleotide exchange factor (RanGEF) called RCC1 (regulator of chromatin condensation 1) loads Ran with GTP and thereby significantly increases RanGTP concentration (Recycling in **Figure 1.10**). In the RanGDP environment of the cytoplasm, importins bind the NLSs of their cargo proteins and travel across the central channel of the NPC towards the nucleoplasm. Association of the importins with RanGTP triggers dissociation of the importin-cargo complex and release of cargo in the nucleus (Protein import in **Figure 1.10**). In the opposite direction, exportins recognize and bind NESs on cargo in presence of RanGTP in the nucleus, shuttle towards the NPC central channel and release the cargo in the cytoplasm due to GTP hydrolysis (Protein export in **Figure 1.10**). In both cases, importins and exportins need to be recycled to guarantee the continuity of the nucleocytoplasmic transport. While importing β exits the nucleus on its own, importin is actively recycled by the NTR named CAS (cellular apoptosis susceptibility).

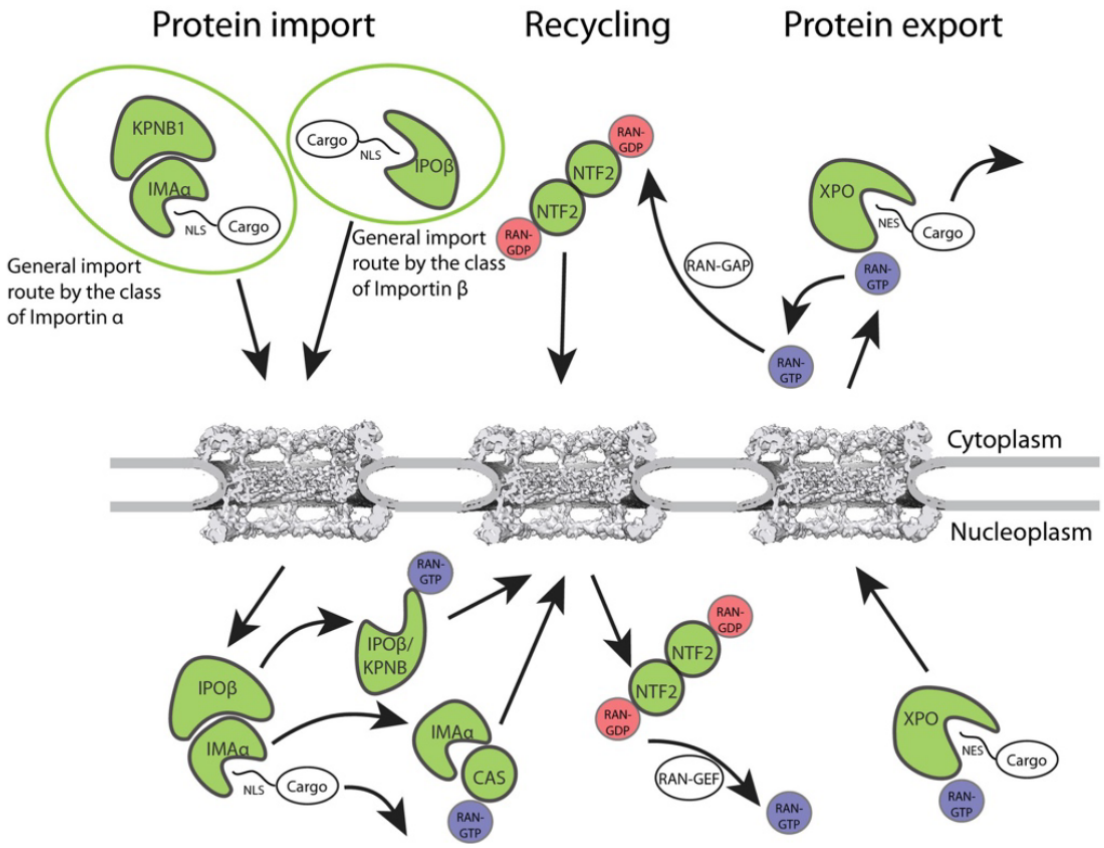


Figure 1.10. Nucleocytoplasmic transport across the NPC. Cargos that contain NLS are bound in the cytoplasm to either importin βs (IPOβ) or importin α that serve as an adaptor for KPNB1. The importin-cargo complex is transported across the NPC and disassembled in the nucleus due to RanGTP high concentration (Protein import). In the cytosol, RanGDP associates with NTF2 and is imported into the nucleus where Ran is further loaded with GTP by RanGEF (Recycling). In the opposite direction, cargos than contain NES are bound in the nucleus by exportins and transported across the NE to the cytoplasm. Their release is dependent on the hydrolysis of GTP bound to Ran in GDP by RanGEF, which highly concentrates RanGDP in the cytoplasm (Protein export).
Figure modified from (Mackmull et al., 2017).

1.2.3 Dilation: NPC diameter *in situ* and *in cellulo*

Due to their large size, modularity and embedding in the nuclear membrane, NPCs have been very challenging to study in their integrity. Purification of NPCs by using detergents has failed due to loss of several components and deformation of the remaining complexes. Approximately 50 atomic structures of individual Nups domains and complexes were solved by X-ray crystallography. Later, the answer on how these pieces assemble together in the large NPC macromolecule was provided by cryo-ET (von Appen et al., 2015). To determine the structure of an entire NPC, nuclei from HeLa cells were purified and NEs obtained by combining DNase and RNase treatments to remove the chromatin. Hence, the NPC structure was determined in the context of NEs but outside of cells. Under this conditions, the central channel of the human NPC was reported to be ~40 nm.

Recent advances in cryo-ET, such as the usage of the dual focused ion beam scanning electron microscope (FIB-SEM), allowed to analyze the NPC structure within cells (*in cellulo*), such as in intact U2OS (Maimon et al., 2012) and in cryo-FIB milled HeLa cells (Mahamid et al., 2016), and visualize the so-called molecular sociology of NPCs not only in the context of nuclear membranes, but also the lamin meshwork, cytoskeletal structure and all other cellular components. Interestingly, the structure of the NPC *in cellulo* is conformationally dilated as compared to previous analyses *in situ* (**Figure 1.11**). In U2OS, the NPC diameter was reported to be ~50 nm at the central plane of the inner ring and ~90 nm at the level of the outer rings (Maimon et al., 2012). Additionally, NPCs showed a high heterogeneity in diameters within the same cell (Mahamid et al., 2016). This variation in pore dimensions between the *in situ* and *in cellulo* structure and either between pores within the same cell are suggested to be dependent on physical forces such as mechanical tension on NE, and related to specific physiological states of the pore, such as the energetic state of the cell (Zimmerli et al., 2020). In a recent study conducted in *Schizosaccharomyces pombe*, starvation or energy depletion of cells seems to recapitulate the conditions occurring upon remotion of NPCs from the cellular context: both passive and active transport are blocked and NPCs constrict up to ~23 nm. This suggests that the integrity of the nucleus has to be retained in order to reconstruct an active NPC and constitutes a prerequisite for studying the functional pore complex.

In conclusion, the term “dilated” can be used to describe the structure of the NPC *in cellulo* in comparison to that *in situ*, and “dilation” can be considered a physiological state of actively transporting NPCs within intact cells. Whether the diameter of NPCs can be increased above this physiological state is not entirely clear yet. Nevertheless, NPCs have been long considered to have an intrinsic plasticity that might allow them to increase the diameter to accommodate for large cargos across the NE.

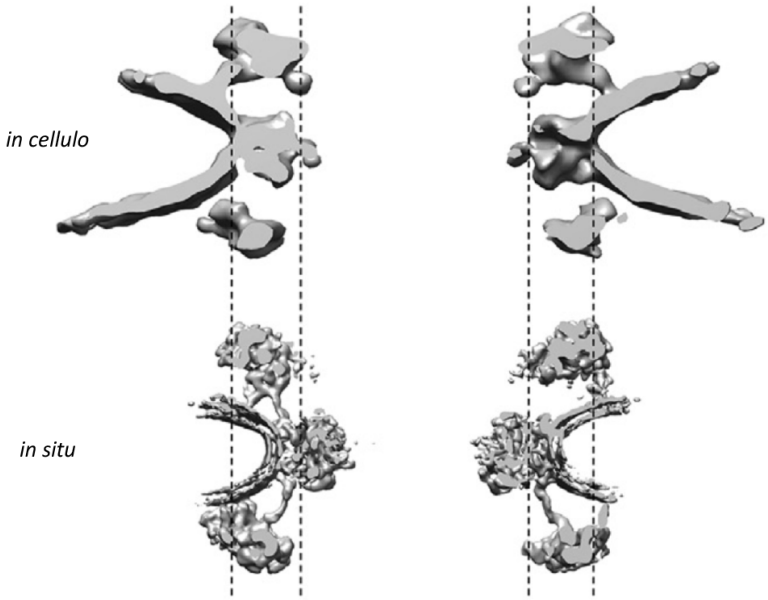


Figure 1.11. Differences in diameters between the *in cellulo* and *in situ* structure of human NPC. The structure of the NPC obtained *in cellulo* from FIB-milled HeLa cells is in a dilated conformation in comparison to the structure of NPC determined *in situ* from purified envelopes. This is evident at the level of the inner ring and at the fusion point between outer and inner nuclear membranes.
Figure modified from (Beck and Baumeister, 2016).

1.3 Cryo-electron tomography *in cellulo*

Cellular tomography reveals the molecular organization of fully hydrated cells in a physiological relevant state and thus provides a 3D reconstruction of multiple macromolecules in their native environment. Structural preservation of cells and their organelles and macromolecules are achieved by their rapid freezing by a method developed in 1982 by Jacques Dubochet and colleagues referred to as cell vitrification (Dubochet et al., 1988). Three-dimensionality is obtained through the acquisition of a set of projections at different angular orientation within the transmission electron microscope (TEM) and the computational alignment and reconstruction of a virtual volume referred to as tomogram. Portions of tomograms that contain repetitive molecular features can be extracted computationally and subjected to SA to determine their structure. Hence, cryo-ET allows to investigate different macromolecular machineries and organelles in a wide range of prokaryotic and eukaryotic cells and elucidate the cellular processes they mediate. The cytoskeleton network, the coat protein complex I, ribosomes, proteasomes, cilia, flagella, are examples of macromolecules targeted by cryo-ET *in cellulo* as reviewed in (Beck and Baumeister, 2016). Although the number of structures deposited in EM databases is progressively increasing, cryo-ET remains very challenging in the aspects of sample preparation, data acquisition and analyses. Among these, sample thickness, radiation damage, low contrast and signal-to-noise ratio (SNR), missing wedge and identification of rare objects inside the cellular compartments, are discussed in the next paragraphs in the context of sample preparation and data collection.

1.3.1 Sample preparation

Sample preparation for cryo-ET mostly consists of two steps: sample vitrification and thinning. Different vitrification and thinning methods have been developed to deal with different types and thicknesses of specimen. In the next paragraphs, some of these methods for cellular vitrification, such as plunge-freezing or high-pressure freezing, and cellular thinning, such as cryo-ultramicrotomy and cryo-focused ion beam milling, are briefly described.

1.3.1.1 Vitrification by plunge- or high-pressure freezing

Vitrification renders biological samples suitable for TEM as it leads to their embedding and preservation in vitreous ice, an amorphous state of ice that closely resembles that of liquid water (Dubochet et al., 1988). This state can be achieved by cooling water faster than the rate of water crystallization. However, the vitrification process of thick biological samples is limited by the fact that water is a poor heat conductor. Generally, the high cooling rate at the cell surface rapidly decays within the sample, and reaches a value that allows water to crystallize at a certain distance from the surface (Dubochet, 2007; Shimoni and Muller, 1998). Consequently, thin samples (<10 μm in thickness) such as viruses, bacteria or thin cells that have a small water content, can generally be vitrified using a method known as plunge freezing, which implies direct plunging of the sample into a cryogen with high heat capacity and thermal conductivity, such as liquid ethane at its melting temperature of ~90 K (-184 ^{o}C). During this process, most of the liquid is removed by blotting with a filter paper and a thin layer of liquid is generated over the carbon-film support in EM grids. On the other hand, samples thicker than 10 μm, such as large mammalian cells, do not properly vitrify using this method, thus water inside of cells forms ice crystals which destroy membranes and organelles. To achieve complete vitrification of thicker samples, the cooling rate of water required for its vitrification has to be reduced. One possibility is to add cryoprotectant that reduce the concentration of free water, such as sucrose, glycerol or DMSO; this however works to the disadvantage of the structural preservation: artefacts such as shrinkage can be introduced due to dehydrating action of these reagents (Meryman, 2007; Studer et al., 2008). The most efficient way to vitrify thick samples (up to 200 μm) is to use high-pressure freezing (HPF). During this process, the expansion of water that occurs during phase transition into ice and the water melting temperature are reduced due to pressure, inhibiting the formation of ice crystal and promoting that of vitrified ice. Also, whereas a cooling rate of ~10^5 K/s is required to vitrify a cell at ambient pressure, a few 10^3 K/s are sufficient at 210 Mpa (Studer et al., 2001; Vanhecke et al., 2008). HPF is often combined with freeze substitution and resin embedding, which generate a frozen crosslinked framework that however limits resolution (Schwarz and Humbel, 2007) and cannot be considered cryo-ET.

1.3.1.2 Thinning of frozen-hydrated biological sample

Only small objects and cells, including viruses, many bacteria and archaea or very thin edges of larger eukaryotic cells (Ben-Harush et al., 2010; Cyrklaff et al., 2007; Hu, 2014; Patla et al., 2010) can be imaged intactly by cryo-ET. Cells above 1 μm in thickness become nontransparent to electron beams at 300 kV acceleration voltage and thus need to be thinned prior to their examination in a TEM. To avoid devitrification of vitreous ice, thinning is performed below ~140K (-130 ^{o}C); sectioning of samples into slices of 50-150 nm in thickness can be done by using a diamond knife in a cryo-microtome (cryo-ultramicrotomy). The sections can then be transferred on EM grids and imaged by cryo-ET (Al-Amoudi et al., 2004). This method offers several advantages: it can be applied to thick specimens vitrified by HPF and allows to generate multiple sections per time from the same sample. However, it is very challenging and produces vitreous sections which are rarely flat and often contain strong cutting artefacts, such as anisotropic compression and crevasses (Al-Amoudi et al., 2005). An alternative to cryo-ultramicrotomy is the focused-ion beam (FIB) milling (Marko et al. 2007; Rigort et al., 2012; Villa, 2013). Initially used in material science to shape specimens for multiple applications, the FIB instrument was lately applied in biology for thinning frozen-hydrated biological samples. In this technique, the milling is performed under vacuum by focusing a beam of accelerated ions, typically gallium (Ga^{+}), at different angles onto the unwanted region of the sample to ablate the material in excess and generate sections of 100-250 nm. These sections are generally referred to as lamellae. The combination of FIB with an electron beam in the same apparatus results in an instrument referred to as dual beam FIB-scanning electron microscope (SEM) or FIB-SEM (Wirth, 2009). The electron beam located at 52^{o} in respect to the FIB allows to target the area of interest and monitor the progress of milling through the SEM image (Wagner et al., 2020).

Major concerns in applying FIB-milling on biological frozen samples for cryo-ET were initially related to a possible transfer of thermal energy to the sample or its damage upon direct interaction with ions. As proof of concept, pure water was FIB-milled and transferred to a TEM; images acquisition showed that water remains vitreous through the entire process (Marko et al., 2006). Subsequent milling of plunge-frozen *E. coli* cells confirmed that FIB-milling of biological sample is compatible with cryo-ET imaging (Marko et al. 2007). Depending on voltage, ion current and angle of incidence, the ions can penetrate ~5-20 nm into the FIB-milled surface, limiting the damage of sample

interior (Matteson et al., 2002). Protection of sample with a thin organometallic layer deposited on the specimen with a gas system injection (GIS) and final sputter coating of the lamellae with a thin conductive layer of platinum further improved the cryo-FIB milling technology for cryo-ET imaging (Hayles et al., 2007; Mahamid et al., 2016). Currently, this approach can be applied to a variety of samples from small bacteria, yeast, algae, to large mammalian cells or even entire multicellular organisms such as *C. elegans* worm (Harapin et al., 2015; Mahamid et al., 2016; Rigort et al., 2012).

1.3.2 Tomographic data collection

Accelerated electrons have shorter wavelengths in comparison to the ~300-800 nm of visible light. As consequence, imaging with electrons allows to resolve molecules relevant for life at near-atomic resolution.

In a TEM, the electrons produced by an electric source (generally a tungsten filament at 2000-3000 ^{o}C) are accelerated by an electric field of typically 200 keV or 300 keV and shaped into an electron beam with high spatial and temporal coherence by a set of electromagnetic condenser lenses. Spatial coherence refers to the uniform spreading direction of the beam, while the term temporal coherence indicates that electrons from the primary electron beam have comparable energies. In a TEM, the column is maintained under high vacuum to avoid electrons interaction with any other type of matter than the specimen inserted in the middle of the objective lens. The nature of this interaction depends on the sample composition. As reviewed in (Orlova and Saibil, 2011), some electron pass through the sample without any interactions, while others interact with the sample and are scattered at certain angles (**Figure 1.12**). The scattered electrons can interact with the sample without losing their energy (elastically scattered electrons), or with partial energy loss through its transfer to the specimen (inelastically scattered electrons). The thickness of the sample is one important factor that determines the type of scattering events that occur. The thicker the sample, the more inelastically scattered events take place. Elastically scattered electrons contribute to image contrast, while the inelastically scattered add noise to the final image and lead to multiple dose-dependent damages of the sample by ionization of atoms, X-ray emission, chemical bonds rearrangement, or formation of free radicals (Orlova and Saibil, 2011). For this reason, both sample thickness and electron exposure have to be limited.

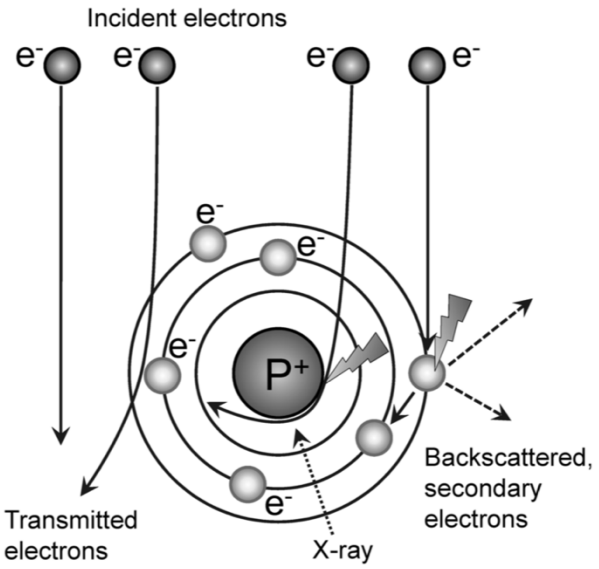


Figure 1.12. Schematic representation of electrons interactions with the sample. Incident electrons from a primary electron beam interact with the sample in different ways. Some electrons can pass through the sample without any interactions. Some electrons are elastically scattered by the sample, meaning that they are deflected at low angles without losing their energy; others are inelastically scattered by the atomic nuclei or the electrons of the sample and this leads to the loss of a part of their energy in form of X-rays or in the emission of secondary backscattered electrons.
Figure modified from (Orlova and Saibil, 2011).

1.3.2.1 Image formation and contrast

Upon interaction with the sample, electrons are focused on the image plane by a set of lenses in the electron microscope and detected by an image detection system, such a digital camera, to form an image. The visualization of the feature of interest in the resulting image highly depends on the contrast of the image. Multiple factors can affect the appearance and the contrast of the image (Orlova and Saibil, 2011).
Some of these are related to the nature of the electron interactions with the specimen. In a TEM, there are essentially two ways of how the electrons that interact with the specimen can generate contrast (Orlova and Saibil, 2011). First, in the so-called

amplitude contrast, some electrons can be physically removed from the primary beam by being absorbed by the specimen. In this ideal condition, few electrons can reach the detector in specific areas and contribute to generate differences in contrast depending on their level of interaction with the specimen. Although very relevant in TEM of heavy-metal stained, plastic embedded samples, in cryo-EM amplitude contrast only partially contributes to the total contrast of an image. This is because the light atoms that compose the specimen (H, O, N, and C) mostly deflect electrons from the incident beam, rather than absorbing them. Consequently, the total number of electrons remains the same after their interaction with the specimen; what changes is the path length of the electrons due to their scattering at varying angles, which generates the so-called phase contrast, the primary image formation mechanism in cryo-EM. Thin transparent samples scatter electrons through small angles and are described as weak phase objects. To improve the contrast of biological samples, images can be collected at a certain defocus, which allows to introduce a phase shift between the scattered and the unscattered electrons; the largest is this shift, the stronger the contrast.

Images can also be collected in focus with the usage of phase plates, such as the Volta potential phase plate (VPP) (Danev et al., 2017). VPP consists of a thin heated amorphous carbon film that applied at the back focal plane of the objective lens can induce an additional shift to the scattered wave when illuminated by the electron beam, thereby increasing phase-contrast.

1.3.2.2 Electron dose exposure

To limit the damage introduced by inelastically scattered events (Glaeser, 1971), the electron dose to which the sample is exposed has to be minimized. This can be achieved by applying the so-called low-dose scheme for the search of an area of interest in the sample and the actual data collection. In this scheme, the desired imaging location is identified at low magnification with low electron dose, while focusing is performed at a location adjacent to the area of interest prior to image acquisition. During tomographic acquisition, the area of interest is imaged from different angles by tilting the stage within the range of typically +/-60^{o} with a tilt increment of 1^{o}-3^{o}. The total dose, generally in the range of 70-140 $e^{-}/Å^{2}$, is distributed (fractionated) over the number of images used to record an entire tilt series, resulting in a low SNR of individual tilt images. Consequently, using a lower electron dose reduces radiation damage, but also introduces a major

limiting factor for cryo-ET, which is a low SNR in individual images (Glaeser, 2016; Grimm et al., 1996).

1.3.2.3 Acquisition schemes

Depending on the purpose of the tomographic acquisition, the distribution of the total electron dose on the sample can be optimized by choosing a proper tilt acquisition scheme. The following general principles have to be considered: i.) when the sample is exposed to the electron beam at different angles, the images collected later in the tilt series lose higher resolution information due to accumulated electron dose (Grant and Grigorieff, 2015); ii.) when a sample with a slab geometry is tilted, the effective thickness of the sample increases with $1/\cos(\alpha)$, where α is the tilt angle. The increasing effective thickness at higher tilts causes further reduction in the SNR. It also limits the range of tilt that can be collected to $\sim$+/-60° and leads to a so called "missing wedge" of information in Fourier space (Schmid and Booth, 2008). The central slice theorem postulates that the Fourier transform of each projection image corresponds to a planar slice through the Fourier transform of the 3D volume (Rosier and Klug, 1968); the resulting tomographic reconstruction is therefore an incomplete 3D representation of the specimen, constituted by a series of planar slices oriented about the tilt axis. Consequently, tomograms have anisotropic resolution and the missing wedge manifests in form of elongated artifacts along the direction of the electron beam.

Different acquisition schemes have been proposed: i.) the continuous, ii.) the bi-directional, and iii.) the dose-symmetric tilt scheme (Hagen et al., 2017). In the continuous tilt scheme, often used for resin embedded samples, the specimen is rotated in one direction and continuous series are acquired, e.g., from +60° to -60°. Although this is a fast way of collection, it is not optimal to achieve high resolution in cryo-ET. The first electrons are irradiated at high tilts, thus the images that less suffer from electron radiation are those that have the highest thickness and lowest SNR. On the other hand, images at lower tilts with the highest SNR are collected only when the sample has been largely exposed to the electron beam.

In the bi-directional tilt scheme, the tilt series acquisition is divided in two separate branches, e.g. with a continuous branch from 0° to +60° or only +30-45° with a 3° increment acquired as first, and another from -3° to -60° as second. This way the images

at low tilts in the first branch are acquired at lower accumulative dose as compared to the continuous scheme. However, the low tilts from the second branch that have higher SNR contain less high-resolution information due to the accumulated dose.

Finally, the dose-symmetric scheme allows to achieve a better dose distribution across the tilt series: it starts the collection at 0^{o} and then alternatively increases the tilt range for both positive and negative branch by a given tilt increment (Hagen et al., 2017). This acquisition scheme was shown to be superior to the others for subsequent high-resolution SA (Turoňová et al., 2020).

1.3.3 Tomographic data analyses

1.3.3.1 General principles of subtomogram averaging

While the anisotropic resolution caused by the missing wedge is difficult to overcome for a tomogram, isotropy of a structure of interest can be partially recovered and its resolution improved by averaging multiple sub-volumes of tomograms that repetitively contain the same structure. This process is known as SA, and, in brief, consists in a series of iterative alignment and averaging steps (**Figure 1.13**). During the SA procedure, a reference is rotated within a defined angular range and at each rotation it is compared to a subtomogram with the aim to identify their mutual orientation. Their mutual spatial shift and an overall similarity score are typically evaluated using constrained cross correlation (CC) functions (Förster and Hegerl, 2007). Once this is repeated for all subtomograms, the found orientations and shifts are used to align and average the subtomograms into a new reference. The progress of the alignment can be evaluated by examining the average, which should improve in contrast and visibility of details, and the shifts and orientations which should decrease during refinement. The procedure is thus repeated until the successive average does not contain any finer details and the shift and orientations reach a plateau, suggesting that the procedure cannot be furtherly improved.

The reference for the initial alignment can be a priori structure, or alternatively when this is not known a suitable reference can be generated from the data by using a free-reference alignment approach (Orlova and Saibil, 2011). During the alignment, the reference is masked to deal with the low SNR of the tomographic data and the missing wedge. A mask is applied in real space around the particle to reduce noise by excluding the

background regions outside the particle and thus maximize the contribution of the signal to the average. Another mask referred to as wedge-mask is applied in reciprocal space to both reference and sub-volumes to avoid alignment of the missing wedges (Bartesaghi et al., 2008; Forster et al., 2005; Förster and Hegerl, 2007; Schmid and Booth, 2008). Additionally, a band-pass filter can be applied in reciprocal space to focus the alignment to frequencies that correspond to the structural features of interest.

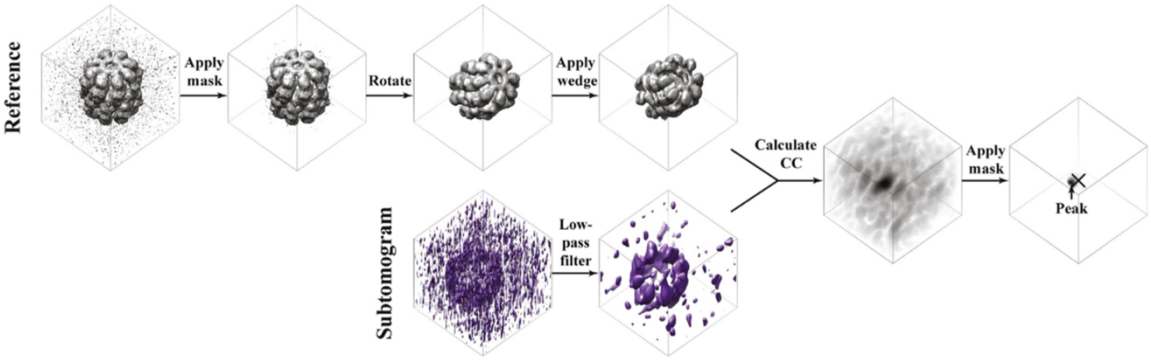


Figure 1.13. Schematic representation of the iterative subtomogram averaging workflow. During subtomogram averaging a series of image processing routines are applied to the reference (in gray) and to each subtomogram volume (in purple). The reference is rotated into defined orientations with defined angular steps in comparison to each subtomogram. Cross-correlation scores (CC) are calculated for each orientation and shift of the reference in relation to the particle. In a final step, all particles are oriented according to the orientation and shift with highest CC score and averaged to generate a new reference.
Figure modified from (Wan and Briggs, 2016).

Many structures have already been analyzed by SA, some of them at a sub-nanometer resolution, others even at a near-atomic resolution (Bharat et al., 2012, 2017; Hutchings et al., 2018; Kovtun et al., 2018; Wan et al., 2017). Viruses are considered particularly optimal targets for SA because of the high repetition of subunits, such as capsid and matrix proteins, into relatively regular lattice-like arrangements (Obr and Schur, 2019);

this is the case of HIV-1, for which either the gag structure in the immature virion (as shown in **Figure 1.4**) (Schur et al., 2015) and the structure of the hexamer and the pentamer in the lattice of the mature HIV-1 capsid core (**Figure 1.5**) (Mattei et al., 2016) were determined at near-atomic resolution. In this context, SA can also provide important information on the global arrangement of the individual proteins into a structure.

Nevertheless, the identification of features of interest *in cellulo*, even if these are highly organized structures, and their picking out of a tomogram remains challenging due to the low contrast of images and the cellular molecular crowding. For this reason, very often the localization of objects inside the cell has been driven by correlative light and electron microscopy (CLEM).

Chapter 2

Aim of this study

Human immunodeficiency virus (HIV-1) remains a major health threat. To prevent viral capsid uncoating with the following release and nuclear entry of viral genome would limit viral replication and latency. Towards that goal, a better understanding of these processes is required. Here, I propose to develop *in situ* structural biology methods based on cryo-ET to structurally analyze HIV-1 capsids and NPCs during the nuclear entry process in a close to native state. Specifically, I want to address the following questions:

- Where inside of the cell is the hexameric lattice of the HIV-1 capsid disassembled?
- Can intact HIV-1 capsids be transported through intact NPCs?
- Is the NPC architecture and diameter altered upon nuclear entry of individual HIV-1 particles?
- Is the NPC architecture and diameter different in infected as compared to non-infected cells?

Chapter 3

Materials and Methods

3.1 Sample cultivation and preparation

3.1.1 Cell culture

Primary human T lymphoblast CD4$^+$ T cells were isolated from the blood of healthy, anonymous volunteer donors, provided by the Heidelberg Hospital Blood Bank, and activated by IL-2 by Ines de Castro in the group of Marina Lusic at the Department of Molecular Virology, Infectious Disease at the University Hospital of Heidelberg.

Primary human monocyte-derived macrophages were isolated from the blood of healthy, anonymous volunteer donors, provided by the Heidelberg Hospital Blood Bank, and differentiated by Hanke-Mareil Heuser in the group of Hans-Georg Kräusslich at the Department of Molecular Virology, Infectious Disease at the University Hospital of Heidelberg.

Human T lymphoblast cells SupT1-R5, stably expressing exogenous CCR5 under puromycin selection (gift from Robert Doms, University of Pennsylvania, USA; certified by Eurofins according to DAkkS ISO 9001:2008), primary human T lymphoblast CD4$^+$ T cells and monocyte-derived macrophages (MDMs) were cultivated at 37°C in a humidified incubator with a 5% CO_2 atmosphere, using RPMI 1640 medium with GlutaMAX (ThermoFisher Scientific) supplemented with 10% fetal bovine serum (FBS; Merck), 50 U/ml of penicillin, 50 µg/ml of streptomycin (ThermoFisher Scientific) and 0.3 µg/ml puromycin (Merck).

3.1.2 Viral production and purification

The viral production and purification was carried out by Tina Holdt and Vojtech Zila in the group of Hans-Georg Kräusslich, at the Department of Molecular virology,

Infectious disease at the University Hospital Heidelberg. I have participated in the production and characterization of a single batch of viral particles. Afterwards, I obtained the viruses and used it for the infection of cells.

To produce infectious RT-competent virus-like particles (NNHIV), HEK 293T cells grown on 175 cm^2 tissue culture flasks were transfected with 70 µg of pNNHIV, using calcium phosphate precipitation according to standard procedures (Bejarano et al., 2019; Zila et al., 2019). For production of infectious virions labeled with IN.mScarlet, cells were co-transfected with pVpr.IN.mScarlet at a molar ratio of 4.5:1. To produce NNHIV particles carrying A77V CA mutation and the labeled IN.mScarlet, cells were co-transfected with pNNHIV-A77V and pVpr.IN_{NN}.mScarlet at a molar ratio of 4.5:1. Culture media from virus-producing cells were harvested at 44–48 h post-transfection, cleared by filtration through a 0.45 µm nitrocellulose filter, and particles from media were concentrated by ultracentrifugation through a 20% (w/w) sucrose cushion for 90 min at 27,000 rpm (at 4°C) in a Beckman SW32 rotor (Beckman Coulter Life Sciences). Particles were resuspended in PBS containing 10% FBS and 10 mM HEPES (pH 7.2) and stored in aliquots at −80°C. For detection of HIV-1 RT products by ddPCR, virus-containing medium from producing cells was treated with 15 U/ml DNase I (Merck) and 10 mM $MgCl_2$ for 5 h at 37°C prior to ultracentrifugation. Particles were then aliquoted and stored as above. Particle-associated RT activity was determined by SYBR Green-based Product-Enhanced Reverse Transcription assay (SG-PERT) (Pizzato et al., 2009).

3.1.3 NNHIV-1 infection

Infection with NNHIV-1 A77V CA INmScarlet or NNHIV WT CA INmScarlet respectively of wild-type or CPSF6 knockdown SupT1R5 and $CD4^+$ T cells was performed as follows: cells were distributed in well of 96-well plate (4×10^5 cells/well), pelleted (3-5 min / 200 × g) and resuspended in ~100 complete RPMI medium containing IN.mScarlet-labeled NNHIV or NNHIV-A77V particles (25 µUnits of RT/cell, corresponding to MOI of 2.5–5 using the same µUnits of RT/cell by infectivity assay). Cells were incubated with particles for 90 min at 16°C to synchronize virus entry and incubated at 37°C for 15 hours.

3.1.4 CPSF6 knock-down

The CPSF6 knock-down in SupT1-R5 T cells and the cellular infection were carried out by Kathleen Börner and Vojtech Zila in the group of Hans-Georg Kräusslich. I performed plunge-freezing, cryo-FIB milling and cryo-ET of the CPSF6 knockdown sample as described in 3.2.1.

CPSF6 knock-down was performed using adeno-associated virus (AAV) vectors. SupT1-R5 cells were distributed into 96-well plates (5×10^4 cells/well; Flat bottom; Greiner Bio-One) and transduced once with equal amounts of purified AAV (2.5×10^6 vector genomes/cell) expressing three shRNAs against CPSF6 or a non-targeted shRNA (NS control). Mock-transduced cells were used as an additional control. At 72 h after transduction, viability of cells was assessed by trypan blue exclusion using the TC20 Automated Cell Counter (Bio-Rad). To determine CPSF6 knock-down efficiency, cells were fixed with 4% formaldehyde (FA) in PBS and immunostained with anti-CPSF6 antibody, followed by staining with secondary antibody conjugated with Alexa Fluor 488. Proportion of CPSF6-positive cells was scored by flow cytometry, using a BD FACSCelesta flow cytometer (BD Biosciences) and data were processed using Flowjo software (FlowJo LLC, BD Biosciences).

3.1.5 Quantification of INmScarlet in $CD4^+$ T cells

Primary $CD4^+$ T cells were infected with NNHIV-1 A77V CA INmScarlet particles as described in 3.1.3.

14 hours post infection (hpi), they were let to adhere on PEI-coated glass bottom microwell dishes (MatTek dishes; 35 mm petri dish, 14 mm Microwell; No. 1.5 coverglass 0.16-0.19 mm) for 1 hour at 37°C. 15 hours post infection, cells were fixed with 4% PFA in PBS (10 minutes at room temperature). The nucleus was stained with Hoechst33258 (Merck). Image stacks were acquired on a Leica SP8 confocal microscope using a 63x/1.4 NA oil immersion objective and excitation with the 405 and 561 laser line. Brightfield was acquired to visualize plasma membrane. The acquired z-stacks were visually examined using Fiji software (Schindelin et al., 2012). Counting of individual INmScarlet particles localized at PMs, NEs and nuclei was done manually.

3.2 Sample preparation for cryo-ET

3.2.1 EM grids preparation of suspension cells

Vitrification of not-infected $CD4^+$ T and infected SupT1-R5 cells (15 hpi) was performed as follows: 3.5 µl of suspension cells from 80-100 µl of final volume were applied on 200-mesh copper grids coated with R 2/1 holey SiO_2 films (Quantifoil Micro Tool GmbH) glow discharged on both sides and plunge frozen in liquid ethane at ~-184°C using a Leica EM GP grid plunger. In the blotting chamber, maintained at 37°C temperature and 90% humidity, the grids were blotted for 2-3 sec with a filter blotting paper (Whatman 597) applied on the reverse side. All plunge-frozen grids were stored in liquid nitrogen until further use.

3.2.2 EM grids preparation of adherent cells

200-mesh gold grids coated with R 2/1 holey SiO_2 films (Quantifoil Micro Tool GmbH) were glow discharged on both sides for 45 sec and sterilized under UV light for 30 min. The grids were then transferred to sterile glass bottom microwell dishes (MatTek dishes; 35 mm petri dish, 14 mm Microwell; No. 1.5 coverglass 0.16-0.19 mm), coated with 50 µg/ml of fibronectin for 60-90 minutes and washed twice with PBS.
Meanwhile, monocytes-derived macrophages were incubated with 1 ml of accutase (Sigma) for 30-60 minutes at 37°C on a moving shaker to let them detach from petri dishes. Around 150,000 monocyte-derived microphages per well were pipetted onto the grids. The cells were allowed to settle and attach to the grids for 24 h at 37°C in 5% CO_2. Plunge-freezing was done as described in 3.2.1.

3.2.3 Cryo-FIB milling

Modified auto-grids (or cartridges) containing a cut out were used to allow milling at a shallow angle. Before starting, the auto-grids were marked at each side perpendicularly to the cut with a permanent marker in order to facilitate the loading into the Krios cassette after the cryo-FIB milling. The frozen grids were clipped into the auto-grids and transferred to the Aquilos FIB-SEM (Thermo Fisher) at liquid nitrogen temperature. Samples were sputter-coated with inorganic platinum (10 mA, 10-20 sec) and coated with organometallic protective platinum layer using the in-situ gas injection system (GIS)

(Hayles et al., 2007) for 7-12 sec. Ablation of undesired cellular material was performed at stage tilt angles of 18°-20° by focusing Gallium ion beam at 30 kV on parallel rectangular patterns above and below the area of interest. Lamella preparation is conducted in a stepwise milling to produce 150-250 nm sections of the biological sample. 0.3 µm gaps at 4-5 µm distance alongside the lamellae were often generated on grids to reduce lamella bending happening during the final milling step (Wolff et al., 2019). Before unloading, grids were sputter-coated again for 2 sec (10mA) to improve conductivity of final lamellae. The final prepared samples were stored under liquid nitrogen condition until further use.

3.3 Cryo-electron tomography data collection

Grids with lamellae were loaded into the Krios cassette. During sample loading, the marks on the auto-grids were disposed in the cassette in a way to orient the FIB-milled lamellae perpendicularly to the tilt axis of the microscope. Cryo-electron tomographic tilt series were acquired on a Titan Krios (Thermofisher FEI) operating at 300 kV equipped with a Gatan Quantum post-column energy filter and a Gatan K2 4k x 4k direct electron detector.

Previous to tilt series acquisition, an initial overview of the entire grid was acquired at 1500x magnification for the lamellae identification. Subsequently each lamella on grid was mapped at 6500x magnification and the resulting lamella map was used to navigate inside the cell and identify positions of interest for data collection. HIV-1 particles are not visible at this magnification, thus double membranes of NEs or NPCs were targeted for setting up positions for the tilt series acquisition, when possible.

Tilt series were collected at a nominal 42000x magnification resulting in a calibrated pixel size of 3.37Å or 3.45Å, over a tilt range +64 to -40 for a positive pre-tilt with 3° tilt increment corresponding to ~35 images per tilt series, a total dose of ~140 e-/A^2 and a nominal defocus range of -2 to -4.5 µm. The data acquisition was automated using a modified version of the dose-symmetric scheme (Hagen et al., 2017) taking the lamella pre-tilt into account (Allegretti et al., 2020) and controlled by using Serial EM (Mastronarde, 2005).

3.4 Data processing

3.4.1 Tomographic reconstruction in IMOD

All the tilt series acquired at the Titan Krios using the dose-symmetric scheme (Hagen et al., 2017) were pre-processed and dose-filtered before alignment.

Initially, the tilt series were sorted from the minimal to the maximal tilt using Matlab scripts from William Wan and Beata Turonová. Each image 4× binned was manually inspected for abnormalities (such as drift-ice crystal contamination, presence of the edge of lamella or grid bar) using IMOD software package 4.9.2. (Kremer et al., 1996). Since each of these factors can negatively affect the tomographic reconstruction, images containing any of those were removed from the tilt series.

Dose-filtering of each tilt series (Grant and Grigorieff, 2015) was performed in MATLAB by using scripts from William Wan and Beata Turoňová (Wan et al., 2017).

Dose-filtered tilt series were aligned with IMOD software package 4.9.2. (Kremer, Mastronarde and McIntosh, 1996) through the following briefly described main steps:

i.) Pre-processing: the IMOD command ccderaser was used to remove extreme pixels values due to X-rays or damaged pixels from each tilt and generate a new stack sequence;

ii.) Coarse alignment: the command titlxcorr was run to automatically align the tilt series using CC. The aligned stacks were then binned 4x and the tilt series rotated around the tilt axis;

iii.) Fiducial model generation: fiducial markers that spontaneously generated during the final sputtering in the Aquilos were manually picked at their center to create an initial seed model. Generally, up to five fiducials that were visible across the entire tilt series and appeared approximately spherical were chosen. By using the tracking beads option, the chosen beads were automatically tracked across the entire tilt series;

iv.) Final alignment: the resulting fiducial model was inspected for missing and uncentered points and manually fixed; the residual error of fiducials was used as measure to identify the misaligned fiducials. The final alignment was computed applying the options *no rotation*, *fixed magnification at 1.0* and *fixed tilt angle* in the *Global Variable* window;

v.) Final aligned stack: the final alignment transformation was used to generate a final, aligned image stack 4x binned;

vi.) Tomogram generation: each tomogram was reconstructed at bin4, with a specified thickness of 1600 pixels, and applying the SIRT-like filter to increase the contrast;

vii.) Trimming: as last step, the tomogram was trimmed to the desired dimensions and rotated 90° along the x-axis to match the original orientation.

Around 250 tomograms from nnHIV-1-A77V infected SupT1-R5 cells, 150 from nnHIV-1-WT infected CPSF6 knockdown SupT1-R5 cells, and 85 tomograms from non-infected SupT1-R5 cells were reconstructed from 4× binned tilt-series using the SIRT-like filter according to the described workflow. These tomograms were only used to perform visual analyses as described further in the text.

3.4.2 Segmentation of tomograms

To generate 3D models of HIV-1 capsids in proximity to the NPCs, features of interest within the SIRT-like filtered and 4× binned tomograms were segmented in Amira (Thermofisher Scientific). In particular, the NE and the capsid-like structures were manually segmented in every 2-3 slices of the tomogram in the XY view using the *brush tool*. The segmented slices were then summed and interpolated to obtain a uniform volume, which was represented in 3D using the *Volren* module. The final NPC structure in HIV-1 infected T cells obtained by SA was loaded in Amira and inserted in correspondence of the empty holes on the NE, where the NPCs were localized in the tomogram. To do so, the modules *Resample Transformed Image* and *Volren* were used in tandem.

3.4.3 CTF correction and tomogram reconstruction

For each tilt-series the contrast transfer function (CTF) was estimated using gCTF (Zhang, 2016) prior the dose-filtering. 3dCTF correction and tomogram reconstruction was performed with simple radial filter using novaCTF (Turoňová et al., 2017). These tomograms were used to perform SA.

3.5 Structural analyses of the HIV-1 capsid

3.5.1 Viral particles segmentation and motive list generation

Around 250 tomograms were reconstructed 4× binned in IMOD and manually inspected for the presence of viral particles. The SIRT-like filter was used in order to generate sufficient contrast for the identification and the manual picking of viral structures. When found in the tomographic reconstruction, the capsid-like structures were manually segmented using the IMOD drawing tool. The segmentations were used for two purposes: i.) to determine the surface area of each capsid and compare the structures found in the cytoplasm to those localized in the nucleoplasm; ii.) to generate an initial motive list containing the coordinates of the subtomograms to extract from the capsid volumes for SA of the CA hexameric unit.

In the first case, a triangular representation of the segmented surface was created in MATLAB and the surfaces area was computed as a sum of areas of individual triangles. Values were plotted in Prism8 (GraphPad Software Inc., USA) and statistical significance was assessed by unpaired two-tailed Student's t test.

In the second case, binary masks were generated from the segmentation to define the core volume using the *imodmop* command in IMOD. These volumes were filtered with a circular Gaussian kernel of 20 pixels with 3 pixels radius using MATLAB scripts provided by William Wan. The generated filtered volumes were opened in USCF Chimera together with the respective tomograms in order to identify the contour level of the volume that provided the best overlap with the contour of the capsid.

The gaussian filtered volume and its contour level were used as inputs in the MATLAB scripts provided by William Wan to generate starting positions with a 2-pixels spacing for SA. Per capsid-like structure, approximately 1000 starting positions were extracted, corresponding to a 4× oversampling.

3.5.2 Subtomogram averaging of CA hexamer and analysis of their distribution in the 3D volume

SA described in this section was performed by Beata Turoňová in our group.

For SA, a new set of 3D-CTF-corrected tomograms was reconstructed using novaCTF and binned 2 times. SA followed the protocol described in (Mattei et al., 2016). The

EMD-3465 map was resampled to match the data pixel size and used as an initial reference. Three iterations of alignment were run to allow the oversampled starting positions to shift to the positions corresponding to the hexamers in the CA lattice. Misaligned subtomograms were removed using two-step cleaning procedure. First, for each subtomogram position the closest point on the corresponding segmentation surface was found to compute an angle between a surface normal vector at this point and the normal vector of a subtomogram. All subtomograms with angular distance larger than 30° were removed. In the second step, for each subtomogram distances to all other subtomograms were computed and the overlapping subtomograms were removed, leaving only a subtomogram with the best cross-correlation to the reference. The cleaned set of subtomograms was used to generate a new reference. The whole SA procedure was run again from the oversampled starting positions using the new reference. After the removal of misaligned particles, the recovered lattice contained more CA hexamers. The visual representation of the lattice was created in UCSF Chimera (Pettersen et al., 2004) using Place Object plug-in (Qu et al., 2018).

3.6 Structural analyses of the NPC

3.6.1 Manual picking of NPCs

NPC particle picking was performed manually from the 4× binned tomograms reconstructed with a SIRT-like filter, which have an improved contrast in comparison to the CTF-corrected tomograms. In total 99 and 42 NPCs were picked from infected and non-infected SupT1-R5 T cells, respectively. Other 47 NPCs were picked from non-infected primary $CD4^+$ T cells.

Each tomogram was opened in IMOD and a slicer view of the NPC in side view was used to orient the NPC particle in a way that it looked roughly rounded and its eight subunits were visible, if possible. Four points in total were picked per NPC in side view. Two points were picked at the top and bottom of the NPC, where the two NE double membrane fuses; a third point was picked at the membrane fusion point approximately at the center slice through the NPC. The three points picked are used to calculate a plane and its normal vector. The normal vector indicates the initial orientation of the NPC, while the center between the top and bottom points marks the NPC center coordinates. To avoid a flipped initial orientation assignment, a fourth point localized in the cytoplasm is picked to assign directionality to the normal vector in direction of the cytoplasm. The resulting model

with the four points was saved and used to generate a motive list in MATLAB where all the initial orientations of the NPCs were annotated.

3.6.2 Lamella model generation

Due to FIB-milling of cells, some of the NPCs contained in the final lamella might be sliced by the ion-beam and consequently only a subset of the eight subunits (SUs) might be located inside the lamella.

Taken this into account for SA, a lamella model was generated from each 4× binned SIRT-like tomogram to define the lamellae boundaries. This was done by picking the 8 corners of the lamella slab along the y axis of each tomogram in IMOD. This model was later on used during the SA procedure to extract only NPC SUs which were fully contained within the lamella slab.

3.6.3 Subtomogram averaging of NPCs

SA was performed with slight modifications from a previously described workflow (Beck et al., 2007) using novaSTA package developed by Beata Turoňová in our group (https://github.com/turonova/novaSTA).

Subtomograms were extracted from 3D-CTF corrected tomograms and SA was performed using 8-fold symmetry on subtomograms which were binned 8×, followed by the alignment on 4× binned subtomograms. A spherical mask was generated with the software Dynamo and used for the alignment in order to remove the background noise from the average. Another mask of few pixels was generated for CC.

Before extracting individual SUs, each NPC particle was manually investigated and the particles that were misaligned based on their CC score were removed from the motive list using a MATLAB script from Christian Zimmerli in our group. Another round of alignment was run.

Subsequently, the positions and orientations of the asymmetric SUs were computed based on the previous alignment of the whole NPC. Precisely, 792, 336, 376 SUs were extracted from infected SupT1-R5, non-infected SupT1-R5 and non-infected primary $CD4^+$ T cells, respectively. The subunits outside the lamellae boundaries were removed from the subsequent processing by using a MATLAB script that took in account the previously generated lamella boundary model.

For SA of the SUs, an ellipsoidal mask was applied on one half of the pore and the remaining 656, 299, 351 SUs from infected SupT1-R5 cell, non-infected SupT1-R5 and $CD4^+$ T cells, respectively, were aligned for ~20 iterations. SUs were manually inspected to identify and remove the misaligned particles as previously described for the whole aligned pore. The final SUs average from infected SupT1-R5 cells was used as reference to perform SA independently for the CR, NR and IR. For the separate SA of each NPC ring, smaller ellipsoidal masks were generated with Dynamo and placed around each ring using UCSF Chimera (Pettersen et al., 2004) to focus the alignment on each ring separately. For the final alignment of each ring, 15-20 iterations were run.
The aligned SUs from non-infected T cells were not further processed but used for the NPC diameters measurement.

3.6.4 FSC and whole pore assembly

The Fourier shell correlation (FSC) at threshold 0.143 was computed for each ring, using as inputs the even and odd half-maps automatically generated after each iteration.
The whole NPC structure from HIV-1 infected T cells was assembled as follows: first, the final averages for each ring were masked and filtered to the resolution indicated by the 0.143 criterion of the FSC using the EMAN2 (Tang et al., 2007) e2proc3d command. Second, the resulting filtered rings were fitted on top of the SU derived by SA in UCSF Chimera (Pettersen et al., 2004) and a high-resolution SU was generate. Third, the new assembled SU was masked with a tight, soft-edged mask generated in RELION (Scheres, 2012). Finally, an in-house MATLAB script developed by Tanmay Bharat was used to assemble the eight masked SUs into a ring like fashion and generate the final eightfold symmetric NPC.
The nuclear membrane was manually masked and cleaned from the NPC SU average and added to the whole NPC assembly in UCSF Chimera (Pettersen et al., 2004).
The final cryo-EM map was deposited in the Electron Microscopy Data Bank (EMDB) under the accession code EMD-11967.

3.7 NPC diameter measurements

Measurements of NPCs diameter of opposing asymmetric units were done using in-house MATLAB scripts written by Christian Zimmerli in our group and deposited in (http://doi.org/10.5281/zenodo.4266419) (Zimmerli et al., 2020).

To do so, a 1-pixel sphere mask was dragged in UCSF Chimera (Pettersen et al., 2004) on the feature of interest in the SU, such as inner ring and outer-inner nuclear membrane fusion point in this analysis. The mask was resampled on the SU average and exported with the new coordinates with respect to the SU.

The area of interest within each SU composing each NPC were re-centered in the particle box in respect to the offset of the 1-pixel sphere mask using MATLAB scripts. The new coordinates, orientations and shifts for each SU were annotated in a new particle list. To calculate the diameter of each individual NPC, only NPCs with a SU-occupancy of five ore more were considered. Vectors from at least two or more opposing SUs were derived from these coordinates to calculate the center of each NPC. The average distance between the center of the NPC and each respective SU was then chosen as a representative NPC radius for the given pore. The values for individual NPCs from HIV-1 infected and non-infected T cells were plotted with Prism8 and statistical significance was assessed by unpaired two-tailed Student's t test.

Chapter 4

Results

4.1 HIV-1 targeted CD4$^+$ T lymphocytes are optimal samples for cryo-ET *in cellulo*

To work under relevant physiological conditions for HIV-1 infection, I initially explored nuclear entry and capsid uncoating in the primary targets of HIV-1 *in vivo*, such as primary CD4$^+$ T lymphocytes and monocytes-derived macrophages (MDMs). Primary cells were isolated from the blood of healthy volunteer donors (done in Marina Lusic and Hans-Georg Kräusslich group) and subjected to plunge-freezing, cryo-FIB milling and cryo-ET to assess the extent of vitrification due to cellular thickness.

In order to generate reliable cryo-ET workflows (**Figure 4.1**), I optimized the seeding, plunge-freezing and cryo-FIB milling conditions for mammalian cells (Mahamid et al., 2016) and adapted them to both CD4$^+$ T lymphocytes and MDMs. I observed that primary activated CD4$^+$ T cells were particularly suitable for cryo-ET studies for several reasons. CD4$^+$ T cells are suspension cells that can be directly applied on EM grids for plunge-freezing, without an extra seeding step in the dish (**Figure 4.1 A**). Geometrically, they are small rounded cells of ~10-15 μm in diameter. Their volume is dominated by a large nucleus with a very thin cytoplasm layer underneath the plasma membrane (**Figure 4.2 A**), providing advantages for both vitrification and targeting of the nucleus.

MDMs, on the other hand, revealed to be less handling and slightly more challenging than CD4$^+$ T cells. They are adherent cells and therefore required to be seeded on EM grids in cell culture prior to plunge-freezing (**Figure 4.1 E**); they are double the size of activated CD4$^+$ T cells with a diameter of ~20 μm and a comparably larger cytoplasmic volume (**Figure 4.1 F; Figure 4.2 C**).

As a common principle for cryo-FIB milling, I targeted single cells localizing approximately in the center of the grid squares to generate cryo-lamellae of both cell

types (**Figure 4.1 B and F**). Single cells are more likely to vitrify in comparison to clusters of two or more cells, and their localization in the center of the grid square prevents the neighboring grid bars from blocking the electron beam during data acquisition at high tilts.

In total, I FIB-milled 36 $CD4^+$ T lymphocytes and 25 MDMs and acquired approximately 60 and 20 cryo-electron tomograms from the respective cryo-lamellae.

The resulting tomograms allowed to properly distinguish features of interest in the cellular landscape of both primary activated $CD4^+$ T cells and MDMs, such as NPCs on the NE, ribosomes in the cytosol, and chromatin in the nucleus (**Figure 4.2 B and D**). No artefacts due to ice crystals formation were observed. Therefore, I concluded that both $CD4^+$ T and MDMs properly vitrify upon plunge-freezing.

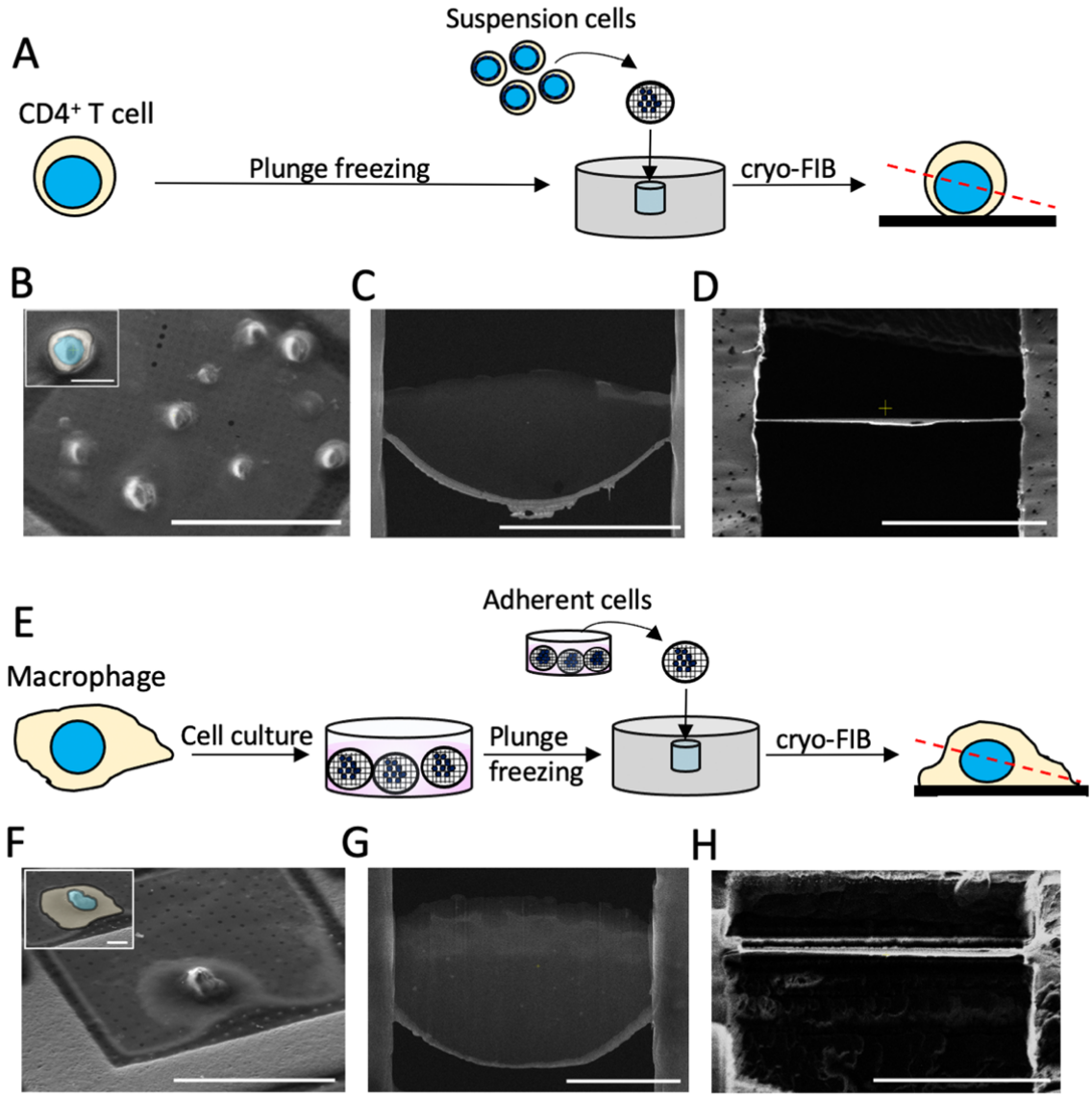


Figure 4.1. Sample preparation varies between suspension and adherent cells.

(A-E) Schematic representation of the experimental workflow optimized for primary $CD4^+$ T cells (A-D) and MDMs (E-H). Nuclei of both cell type are shown in blue, cytoplasm in yellow. **(A)** $CD4^+$ T cells in suspension (400.000 cells in 80-100 µl of medium) are directly applied on EM grid and subjected to plunge-freezing and cryo-FIB milling. **(B)** SEM image shows the geometry of primary $CD4^+$ T cells on grid (scale bar: 10 µm). A single cell is targeted for milling. In the insert, the approximate localization of the cytoplasm (yellow) and nucleus (blue) in $CD4^+$ T cell is shown. Scale bar: 10 µm. **(C-D)** Cryo-lamella of a $CD4^+$ T cell from top view as imaged by the electron beam (C), and side view as imaged by the ion beam (D). Scale bar: 10 µm. **(E)** MDMs cells are cultured adherent on EM grids prior to plunge-freezing and cryo-FIB milling. **(F)** SEM

image shows a macrophage spread on grid. (scale bar: 10 µm). Insert same in (B) but for MDM. Scale bar: 10 µm. **(G-H)** Cryo-lamella of a macrophage shown in top view through the electron beam (C), and side view through the ion beam (D). Scale bar: 10 µm.

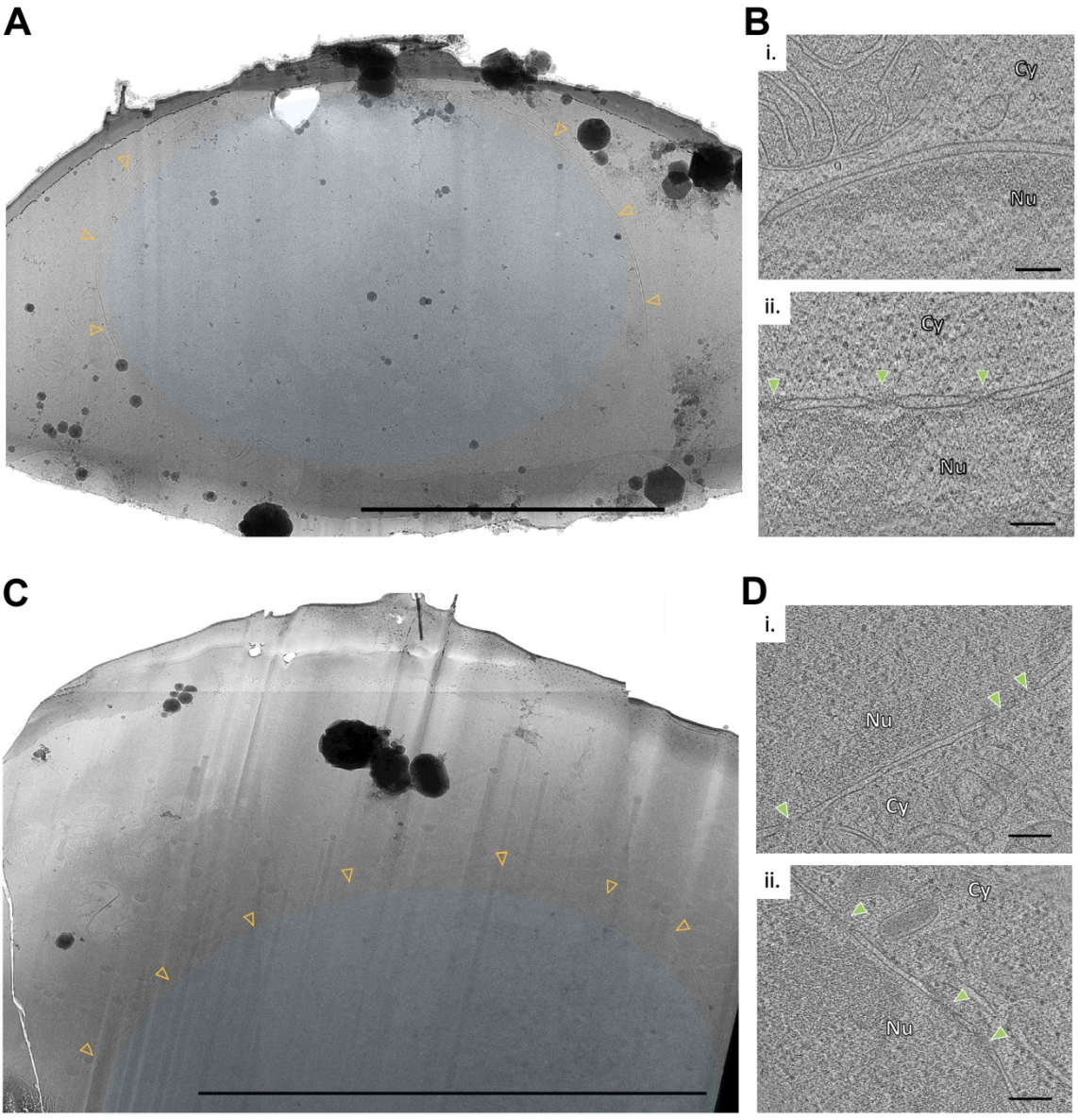


Figure 4.2. Cryo-electron tomography of primary CD4$^+$ T cell and MDM.

(A and C) TEM overview of FIB-milled lamellae of a CD4$^+$ T cell (A) and a macrophage (C) at medium magnification (6500x). Nuclei highlighted in blue and nuclear envelope pointed by yellow arrowheads. Scale bar: 10 µm. **(B and D)** Tomographic slices acquired

at the nuclear envelopes of $CD4^+$ T cell (B) and macrophage (D) in two different positions (i. and ii.). NPCs are pointed by green arrowheads. Cy, cytosol; Nu, nucleus. Scale bar: 200 nm.

4.2 Development of an experimental approach to capture HIV-1 nuclear entry by cryo-ET

4.2.1 Generation of a BSL-1 compatible HIV-1 strain as a prerequisite for structural analysis *in cellulo*

HIV-1 is an infectious agent that was classified into the Risk Group 3 by the World Health Organization. For biosafety reasons, previous tomographic studies of purified HIV-1 virions relied on inactivation of the infectious particles by chemical fixatives (Mattei et al., 2016; Schur et al., 2015; Schur et al., 2016). Alternatively, non-infectious virus-like particles devoid of their genome were assembled *in vitro* (Benjamin et al., 2005). Neither of these procedures, however, constituted a valid option for my studies *in cellulo*. On one hand, chemical fixation negatively affects the structural preservation of the sample; on the other hand, nuclear entry and capsids uncoating were lately suggested to be connected to reverse transcription (Dharan et al., 2020).

To preserve the ultrastructure of HIV-1 replication complexes in infected cells, I therefore performed cryo-ET at the biosafety level 1 (BSL-1) using a non-infectious RT-competent HIV-1 derivative NNHIV that was produced and characterized in the Kräusslich group.

The RT-competent HIV-1 derivative NNHIV contained an inactive IN and a truncated viral accessory protein Tat that impeded integration of the proviral genome and block transactivation of HIV-1 transcription, respectively. Reverse transcription and nuclear import of NNHIV derivative was shown to occur with dynamics similar to wild-type HIV-1 (Zila et al., 2021).

4.2.2 A77V CA mutation accumulates viruses at the nuclear envelope in MDMs but not in CD4$^+$ T cells

Studying HIV-1 nuclear entry and uncoating by cryo-ET revealed to be very challenging for biological and technical reasons. First, HIV-1 nuclear entry is a rare event with few viruses entering the nucleus upon infection (Yamashita and Engelman, 2017). Second, only a thin section of ~150-200 nm is retained from the entire volume of a cell of interest upon cryo-FIB milling (Mahamid et al., 2016), thus further decreasing the chances to capture a *per se* rare event by cryo-ET. As thinning of the sample is unavoidable to generate an electron transparent sample for cryo-ET, I attempted to limit the biological challenge through the enrichment of viral complexes at nuclear pores. For this, I employed an NNHIV-derivative carrying a mutation in CA (A77V) previously reported to prevent the interaction with CPSF6 (Saito et al., 2016). As described in the section 1.1.3.4 of introduction, CPSF6 is suggested to release PIC in the nucleus by competing for the same binding site of Nup153 in CA during nuclear import (Bejarano et al., 2019). In primary MDMs, impairing CA-CPSF6 interaction led to a steady-state accumulation of PICs at nuclear pores up to 2.5-fold more than the wild-type. Under this condition, no major effects on virus infectivity were reported, however a significant donor-to-donor variation was observed (Bejarano et al., 2019).

To test whether A77V-CA mutation induces accumulations of viral complexes also in primary activated CD4$^+$ T lymphocytes as in MDMs, I applied a well-established workflow used for HIV-1 infection of a cell line of CD4$^+$ T lymphocytes, named SupT1-R5 (Zila et al., 2019). CD4$^+$ T cells were incubated with A77V NNHIV particles carrying IN.mScarlet. After low temperature adsorption for 90 minutes at 16°C to synchronize viral entry, cells were incubated for 15 at 37°C to allow for accumulation of viral complexes at NPCs. Viral localizations in the cellular compartments were detected for INmScarlet fluorescence in two different donors (**Figure 4.3**). Most of the A77V CA viral particles did not enter the cells, but rather accumulated at the plasma membrane, whereas very few complexes reached the NE or entered the nucleus. As additional drawback, it was previously shown that HIV-1 infection of primary CD4$^+$ T cells varies greatly between individuals, as also suggested for MDMs, introducing a donor-dependency variability of the cellular infection (Anzinger et al., 2008). Taken together, these results turned out to be incompatible with this study.

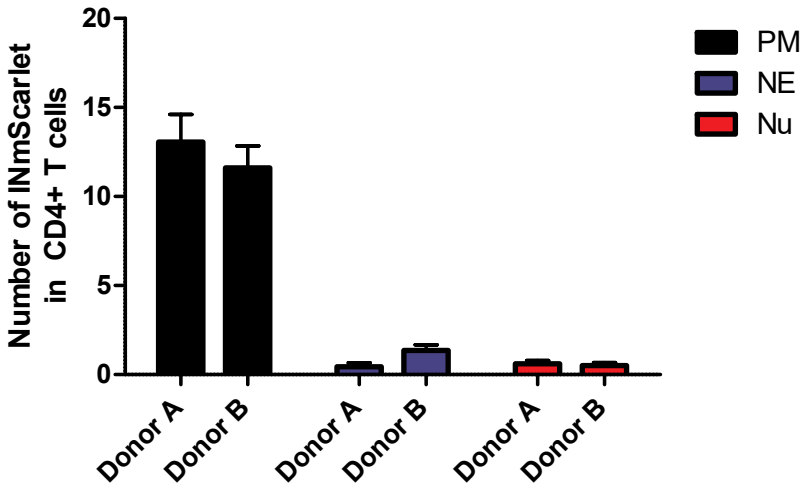


Figure 4.3. Viral complexes carrying the A77V CA mutation do not accumulate at the nuclear envelopes or enter the nuclei of CD4$^+$ T cells. Quantification of INmScarlet positive complexes 15 hours post infection at the plasma membrane (PM), nuclear envelope (NE) and nucleus (Nu) in CD4$^+$ T cells (n = 20) from two different donors.

4.2.3 SupT1-R5 cell line of CD4$^+$ T lymphocytes as optimal sample for cryo-ET studies

Contrary to what I observed in CD4$^+$ T cells, a large number of subviral complexes carrying the A77V CA mutation were arrested at the NPCs in SupT1-R5 cell line between 8 and 15 hpi with few complexes entering the nucleus (Zila et al., 2019). Mildly reduced infectivity of the A77V CA variant was shown in SupT1-R5 T cells, as previously reported for MDMs (Bejarano et al., 2019). Contrary to primary cells, however, cellular infection did not suffer from donor-dependency variability and SupT1-R5 cells were uniformly infected. Based on these observations, I decided to work with SupT1-R5 T cells and postpone the use of primary cells, MDMs mostly, for future studies.

4.3 Visualization of HIV-1 particles in SupT1-R5 by cryo-ET

To assess whether viral complexes can be detected *in cellulo* and investigated in the process of nuclear entry, I combined the cellular HIV-1 infection with the previously optimized cryo-ET workflow for $CD4^+$ T lymphocytes. In order to maximize the number of particles localized at the nuclear pores within the cryo-lamellae, SupT1-R5 cells were infected with NNHIV A77V CA at a boosted concentration of 25 µUnits of RT/cell corresponding to MOI of 2.5–5, approximately 10-fold higher than previously reported (Zila et al., 2019). SupT1-R5 T cells were plunge-frozen 15 hpi and subjected to cryo-FIB milling (**Figure 4.4 A and B**). Criteria for targeting a cell of interest were exclusively based on cellular confluency and localization on the EM grid, as previously described for uninfected primary cells. As SupT1-R5 uniformly absorb viral particles, I initially sought to perform our experiments without recurring to cryo-light microscopy. A cryo-light microscope was running at EMBL at the time I started my experiments, however, not yet at optimal conditions as the samples usually resulted contaminated after imaging. Therefore, I chose to use a brute force approach, consisting in the milling of a large number of cells and the acquisition of a big data set. Around 100 cells were cryo-FIB milled and approximately 250 tomograms of nuclear envelopes in cryo-lamellae were acquired for this study (**Figure 4.4 B and C**).

The resulting tomographic reconstructions contained ~100 NPCs and 9 viral particles. Although my analyses were clearly very low-throughput, it opened the possibility for studying HIV-1 nuclear entry and uncoating *in cellulo* at high resolution and in a near-native condition.

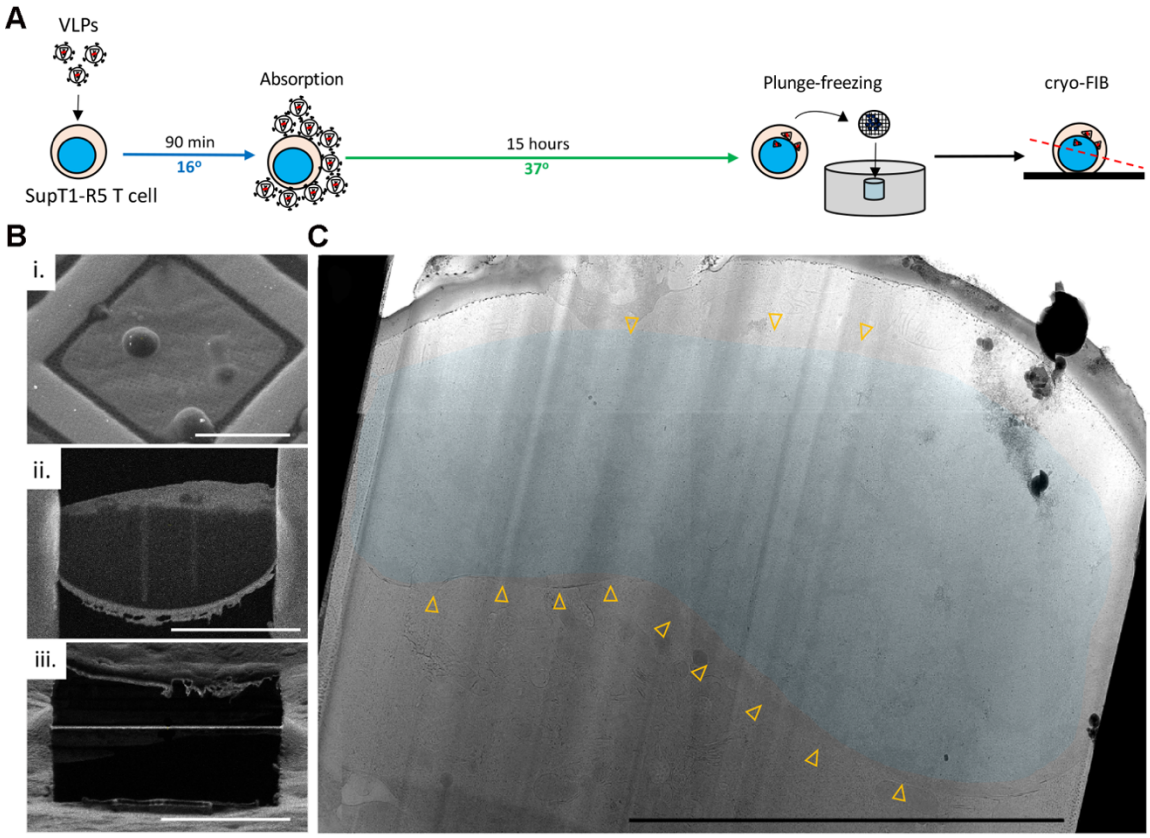


Figure 4.4. Experimental workflow: from cellular infection to cryo-ET.

(A) SupT1-R5 cells are incubated with IN.mScarlet-carrying NNHIV particles for 90 min at 16°C in a 96-well dish. The low temperature allows to accumulate particles at the cell surface (absorption) as prevents both HIV-1 membrane fusion (Henderson and Hope, 2006; Melikyan et al., 2000) and endocytosis (Punnonen et al., 1998; Weigel and Oka, 1981). The sample is shifted to 37°C to initiate virus entry (Zila et al., 2019) and incubated for 15 hours. HIV-1 infected SupT1-R5 T cells are plunge frozen in liquid ethane and transferred to the FIB-SEM for cryo-FIB milling. **(B)** cryo-FIB milling of a single SupT1-R5 T cells, as shown in a high magnification SEM image (scale bar: 50 µm) (i.), leads to the production of a thin cryo-lamella, as shown in top view though the electron beam (ii), and in side view through the ion beam (iii.); scale bars: 10 µm. TEM overview of the cryo-FIB milled lamella at 6500x magnification. The nucleus is highlighted in blue; the yellow arrowheads point to the visible NE and correspond to the regions targeted for the tomographic acquisition. Scale bar: 10 µm.

4.3.1 Cone-shaped HIV-1 capsids penetrate into nuclear pores

In the cryo-electron tomograms, four out of nine viral structures were detected inside the cytoplasm of SupT1-R5 T cells, in close proximity of or even within an NPC (**Figure 4.5 and Figure 4.6**). Without any correlation with a fluorescence signal, the viral complexes stood out from the crowded cellular environment due to their geometry.

Interestingly, they retained capsids with a typical conical shape as known in mature virions (Briggs et al., 2004; Mattei et al., 2016). Their average size of ~112 nm in length and ~55 nm in width of the broad end of the cone also matched previous knowledges of HIV-1 capsids. Inside, they contained dense material, presumably corresponding to the viral nucleoprotein complex.

One of these cone-shaped capsid entered into and penetrated with its narrow end beyond the central channel of the NPC (**Figures 4.6**). The tip of the cone reached to the level of the nuclear ring, the region where Nup153 resides (**Figure 4.6 B-C'**). In tomographic slices at the level of the NPC central channel (**Figure 4.6 D**), the individual spokes of the inner ring were resolved and comfortably accommodated the capsid in between them.

The observed overall morphology and size of the capsids suggested that HIV-1 capsids travel through the cytoplasm and accommodate into the central channel of the NPC still intactly.

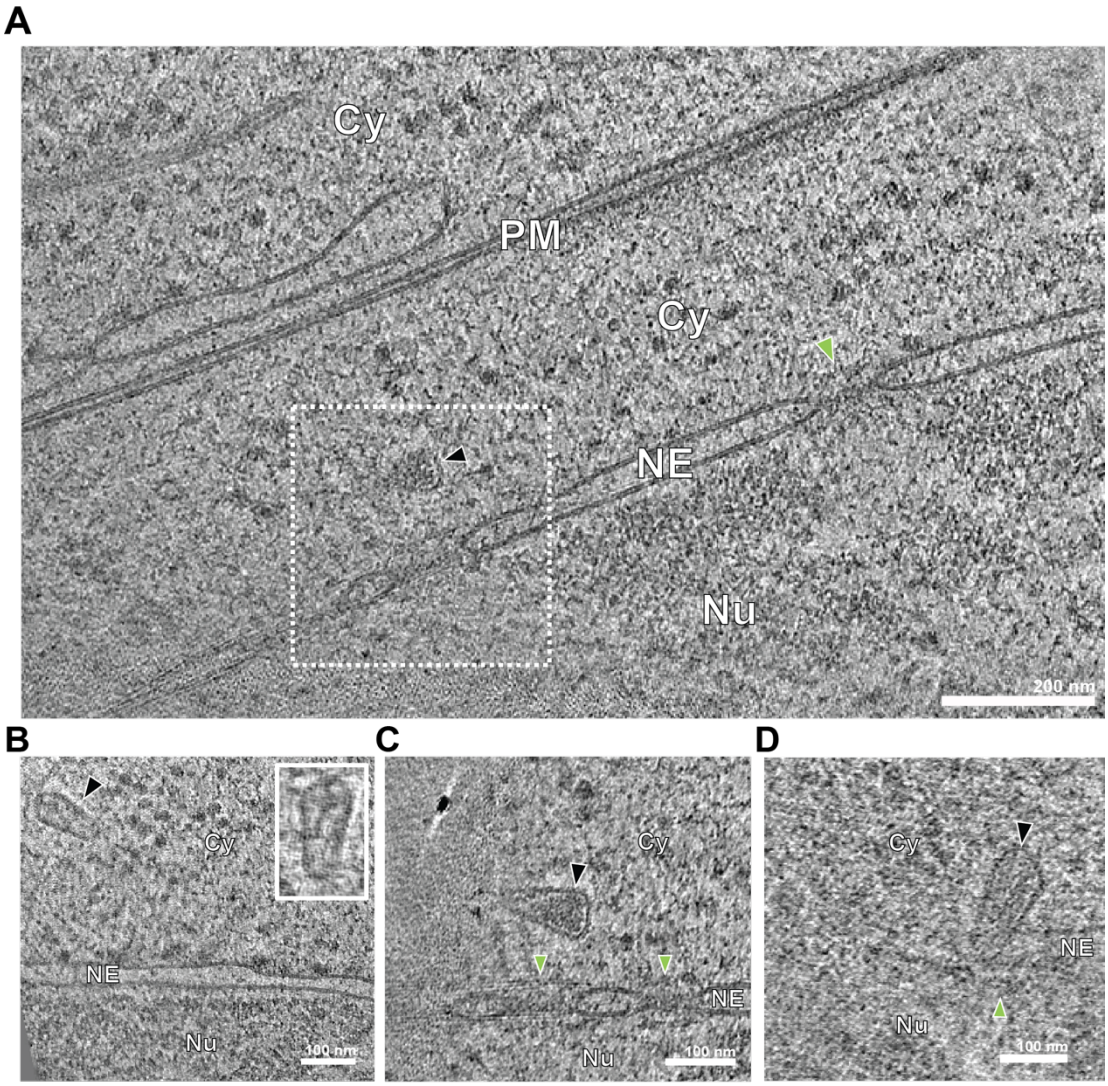


Figure 4.5. **Cone-shaped A77V capsids travel from the cytosol to the nuclear envelope and dock to the NPCs. (A-D)** Slices through tomographic reconstruction showing HIV-1 capsids (black arrowhead) localized in the cytosol (A-C) and in docking position to the NPC (green arrowhead) (C-D) in SupT1-R5 T cells. The area highlighted with the white dashed box in (A) in enlarged in (C). Panels in (B-D) are positioned to simulate the sequence of HIV-1 nuclear entry. The capsids have a conical shape and contain dense material at their inside. Cy, cytosol; Nu, nucleus; NE, nuclear envelope; PM, plasma membrane.

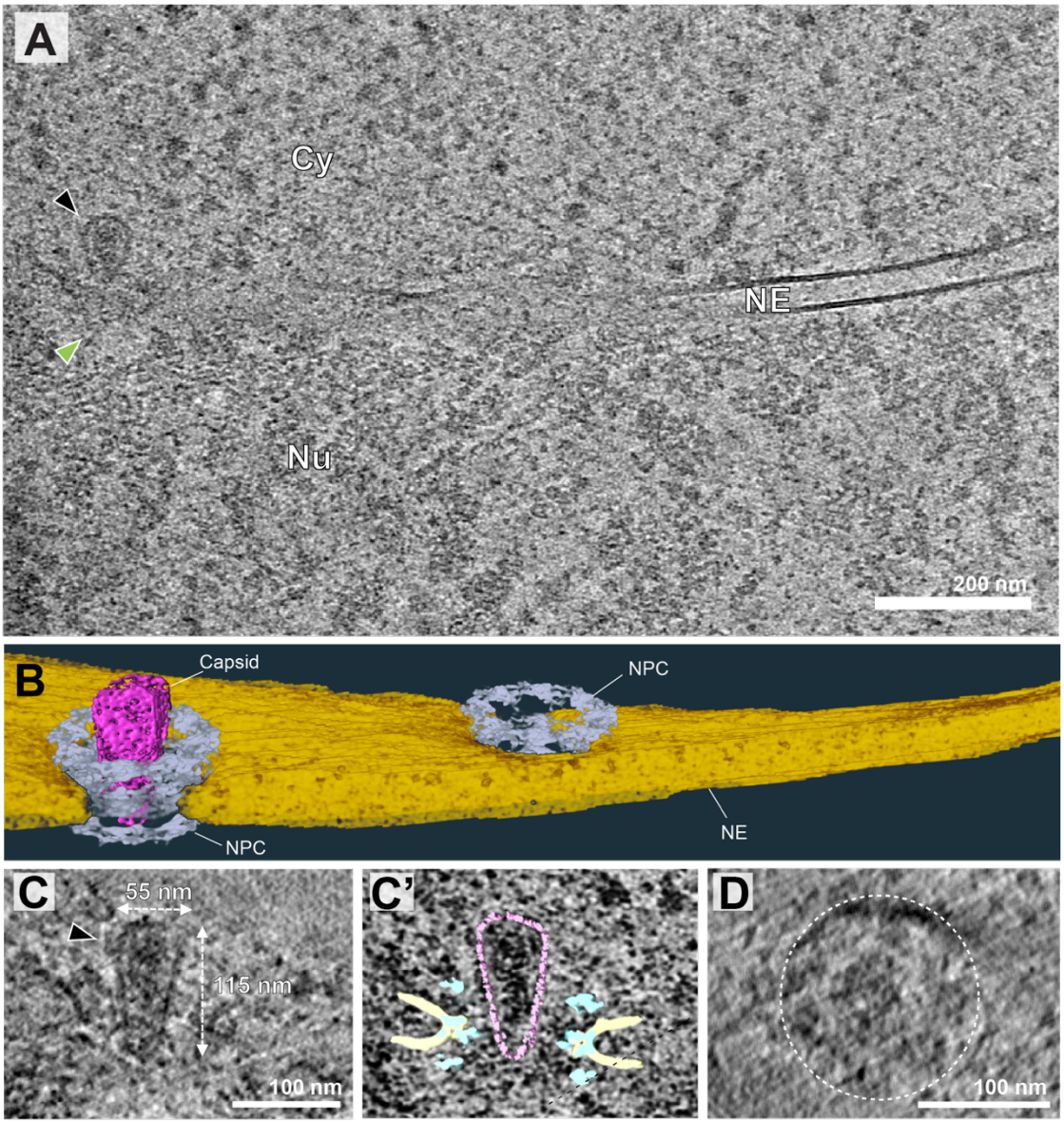


Figure 4.6. Cone-shaped A77V capsids can penetrate into the central channel of the NPC. (A) Slice through a tomographic reconstruction showing a capsid (black arrowhead) localized inside of the NPC (green arrowhead; black box). **(B)** Same as in (B) but displayed segmented and isosurface rendered. Capsid, magenta; NE, yellow; NPC, cyan (cryo-EM map of NPC: this study). **(C)** Enlarged and rotated view of the HIV-1 capsid (black arrowhead) in (A), dimensions are indicated. **(C')** Same as (C) but superimposed with structural models of capsid (magenta; cryo-EM density map of hexameric unit: EMD-3465 (Mattei et al., 2016)) and the NPC (cyan; cryo-EM map of NPC: this study). NE in yellow. **(D)** Same as (C), but displayed in top view. The broad

end of the capsid cone is indicated by the black arrowhead. The NPC is highlighted by dashed circle. Figure modified from (Zila et al., 2021).

4.3.2 HIV-1 capsids are disrupted upon nuclear entry

Five out of the nine viral particles found in the cryo-tomograms localized in the nucleoplasm just underneath the NPC they had passed through. Previous studies revealed that in the absence of CPSF6 binding, PICs are targeted to transcriptionally repressed, lamina-associated heterochromatin (Achuthan et al., 2018). At the same time, the perturbation of CPSF6 binding to the CA hexamer by either the A77V mutation or CPSF6 depletion, resulted in partial co-localization of CA with the basket nucleoporin NUP153 (Bejarano et al., 2019), suggesting that viral complexes reach the nucleoplasm where they remain engaged in interactions with the nuclear pore basket.

The identified structures identified were morphologically altered (**Figure 4.7**) in comparison to the structures in the cytosol. Most of the visualized structures had lost their cone shape, appeared partially open and were devoid of dense material presumed to correspond to the viral genome (**Figure 4.7 A-E**).

While viral particles observed in the cytosol displayed a consistent morphology corresponding to cone-shaped capsids, those inside the nucleus were somewhat more heterogeneous. CLEM and ET performed by Vojtech Zila allowed to identified resembling capsid-like structures at the location of INmScarlet fluorescence, that were morphologically consistent with those observed by cryo-ET (**Figure 4.7 F-H'**).

To properly represent their change in shape and size I segmented all the nine particles identified in the cryo-electron tomograms and measured their surface area. While in the cytosol the capsids were entirely cone-shaped, in the nucleoplasm two were tube-shaped structures and three were fragments (**Figure 4.8 A**); four of these had only ~1/3 of the surface area compared to the conical capsids (**Figure 4.8 B**). Taken together, these results suggested that the conical capsid that enter the central channel of NPC are remodeled in the nucleoplasm upon nuclear entry.

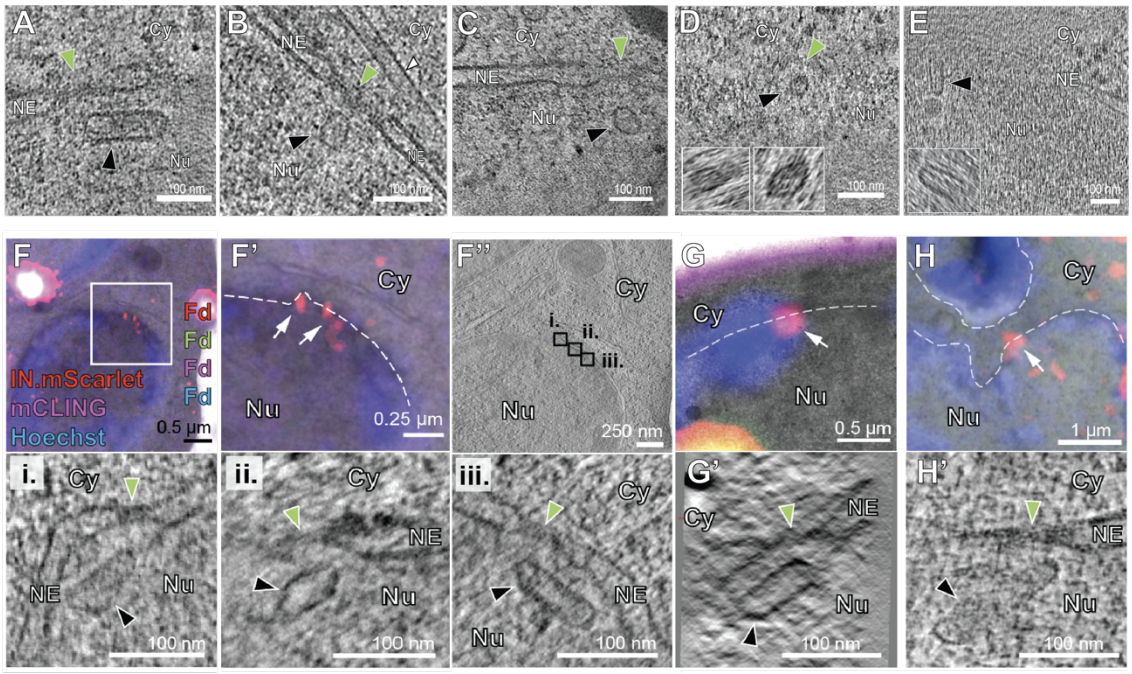
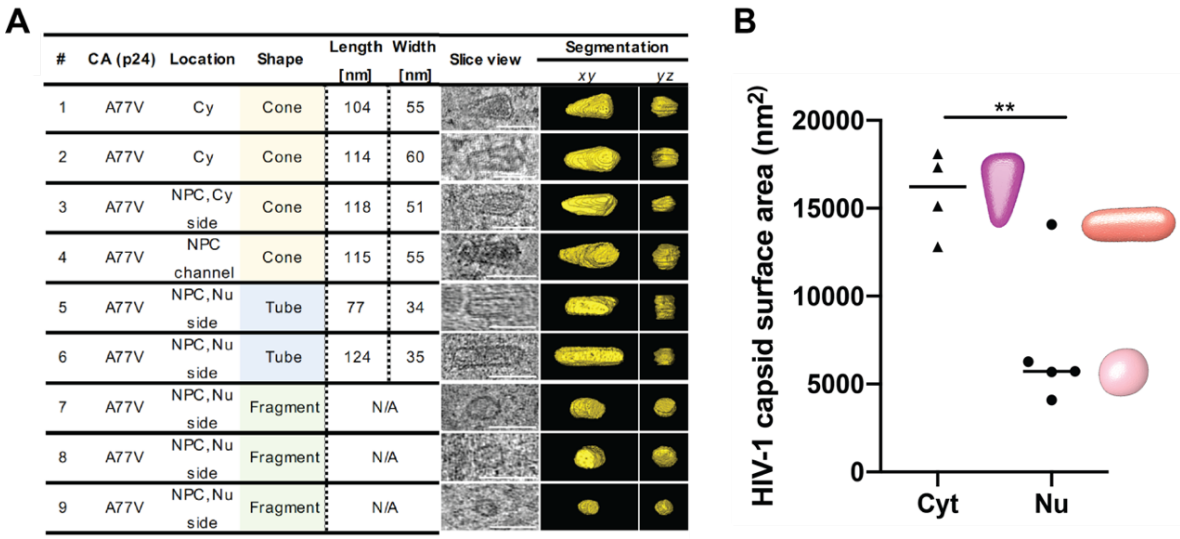


Figure 4.7. CA-A77V structures accumulated at the nuclear basket are morphologically altered. (A-H) Capsid-related structures at the nuclear basket region visualized by cryo-ET (A-E) and CLEM/ET (G-H). Capsids appear as empty tube-shaped or fragments. (F, F') CLEM overlay (F) with enlargement (F') showing the position of IN.mScarlet signal (red; white arrow) at the nuclear envelope (empty white arrowheads) in an EM section stained with mCLING.Atto647N (magenta), post-stained with Hoechst (blue) and decorated with multi-fluorescent fiducials (Fd) for correlation. (F'') Slice through a tomographic reconstruction at the correlated position shown in (F, F'). The features i.–iii. that are shown enlarged in the bottom panel are framed in black and contain three different capsid-related structures at the nuclear basket region. (G–H') Same as (F–F'') showing two capsid-related structures from two different cells.

Figure modified from (Zila et al., 2021). Experiment in (G-H) performed and analyzed by Vojtech Zila.



#	CA (p24)	Location	Shape	Length [nm]	Width [nm]	Slice view	Segmentation xy	Segmentation yz
1	A77V	Cy	Cone	104	55			
2	A77V	Cy	Cone	114	60			
3	A77V	NPC,Cy side	Cone	118	51			
4	A77V	NPC channel	Cone	115	55			
5	A77V	NPC,Nu side	Tube	77	34			
6	A77V	NPC,Nu side	Tube	124	35			
7	A77V	NPC,Nu side	Fragment	N/A				
8	A77V	NPC,Nu side	Fragment	N/A				
9	A77V	NPC,Nu side	Fragment	N/A				

Figure 4.8. Differences in shape and size between capsids in the cytoplasm and nucleoplasm. (A) Summary and segmentations of cones, tubes and fragments captured by cryo-ET in the cytosol, inside NPCs or within the nucleus. **(B)** Surface area of the capsids in the cytosol (Cyt) (n = 4) and inside the nucleus (Nu) (n = 5). Statistical significance was assessed by unpaired two-tailed Student's t test; **p=0.0064. Representative models for three shape-based classes of nuclear structures, such as conical (magenta), tubular (salmon) and spherical (pink) are included in the graph.
Figure modified from (Zila et al., 2021). Segmentations in (A) done by Vojtech Zila.

4.3.3 CPSF6 silencing as complementary strategy to visualize wild-type CA during nuclear entry

To prove that the phenotype observed is not due to the A77V mutation, I used a complementary strategy to enrich nuclear envelopes with viral particles carrying wild-type CA. Without introducing mutations in the CA, CA-CPSF6 interactions were impaired through the knock-down of the cellular CPSF6, as previously reported in both MDMs and SupT1-R5 T cells (Bejarano et al., 2019; Zila et al., 2019).

To achieve CPSF6 silencing, our collaborators in the Kräusslich group transduced SupT1-R5 T cells with AAV vectors expressing a combination of three shRNAs targeting CPSF6, or a non-silencing shRNA as control. CPSF6 was shown to be downregulated with efficiency of 95% without affecting cell viability (Zila et al., 2021). Additionally, intracellular localization, efficiency of nuclear import and infectivity of HIV-1 upon CPSF6 knock-down in SupT1-R5 cells (Zila et al., 2021) were found to be comparable to that of the A77V mutant without knock-down (Zila et al., 2019).

In order to perform cryo-ET of CPSF6 knock-down SupT1-R5 T cells, AAV-transduced cells were infected with the IN.mScarlet labeled wild-type CA-carrying NNHIV for 15 hours prior to plunge-freezing. Approximately 35 lamellae were prepared by cryo-FIB milling and ~120 tomograms reconstructed upon tilt-series acquisition by cryo-ET.

The tomograms were carefully inspected for the capsid-like structures as previously done for the A77V HIV-1 variant. This search was very often challenged by the abundance of vesicles in the cytoplasm and also in proximity to the nuclear envelopes of CPSF6 knockdown SupT1-R5 T cells, which I did not previously observed at the same extent in SupT1-R5 cells.

Apart from one viral particle in the endosome (not shown), only a clearly visible cone-shaped capsid was found in proximity to the nuclear envelope in a docking position to the NPC (**Figure 4.9 A-A''**). The NPC cytoplasmic filaments interacting with the capsids were also visible in the tomographic slice (**Figure 4.9 A'**). As a matter of fact, C-terminal ends of Nup358 were reported to project upward of ~50 nm into the cytoplasm (Walther et al., 2002).

Another capsid-like structure appeared to be localized in the nucleoplasm at the nuclear basket of the NPC (not shown). However, the quality of the cryo-tomogram prevented conclusions about the exact morphology.

Although more data was required to compare the behavior of wild-type HIV-1 CA to that of A77V CA variant during nuclear entry, the upcoming SARS-Cov2 pandemic prevented to perform those experiments.

Nevertheless, CLEM/ET experiments performed in parallel by Vojtech Zila allowed to identified a total of 19 structures in the cellular landscape under CPSF6 silencing.

Similar to the results observed for the A77V mutant, capsids that localized in the cytosol and within the NPC channel were cone-shaped (n = 13), whereas those at the nuclear basket had a tubular shape (n = 4) or were fragments (n = 1) (**Figure 4.9 B-D**).

These data reinforced the notion that conical HIV-1 capsids travel intactly in the cytosol and cross the NPCs unchanged. After passage through the NPC, they do not disassemble into individual subunits, but rather are disrupted into morphologically altered residual capsid structure. Additionally, the results suggested a similar behavior between NNHIV-carrying the A77V CA mutation and the WT CA.

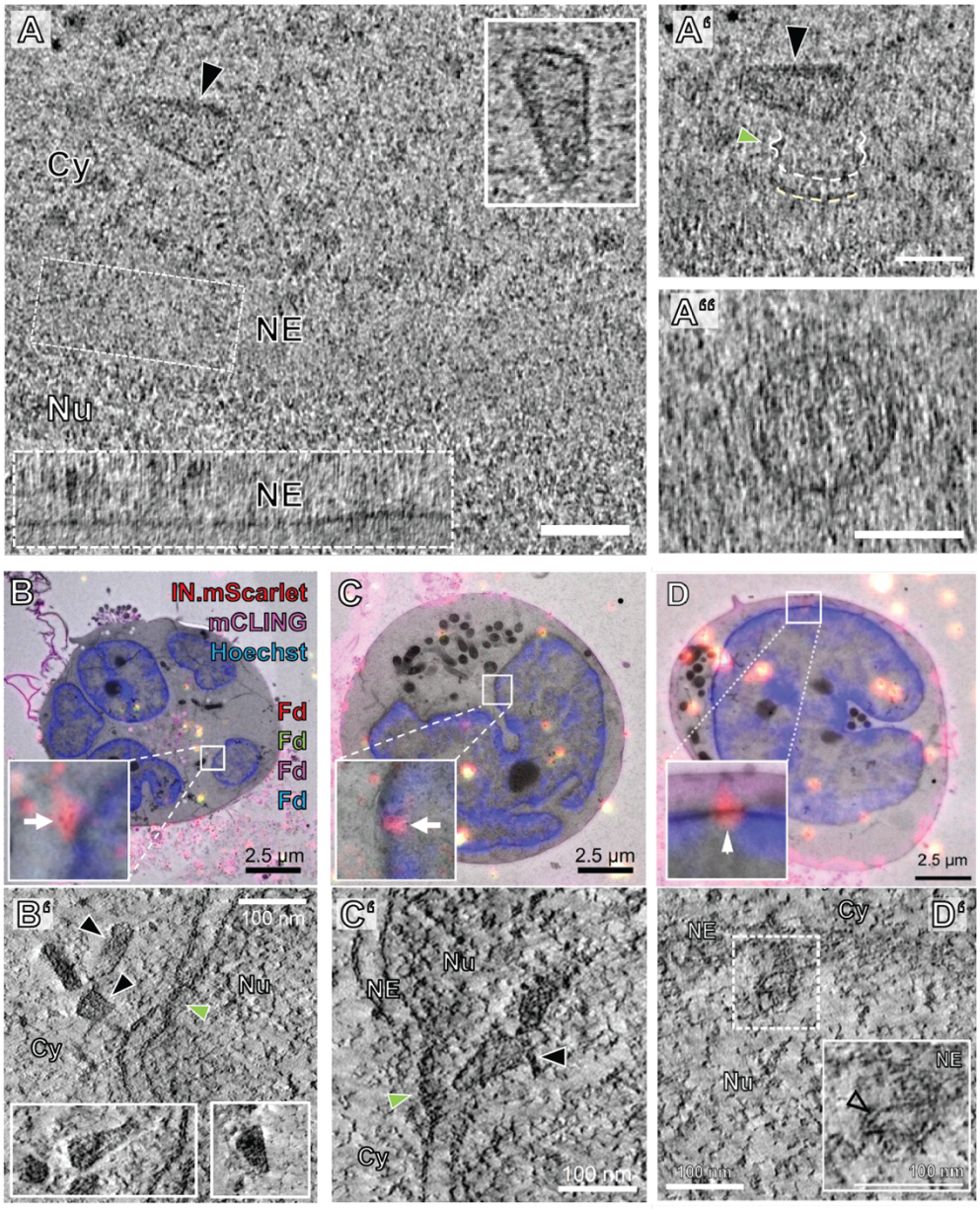


Figure 4.9. Electron tomography of NPC associated HIV-1 complexes upon CPSF6 knock-down. (A-A'') Cryo-electron tomography of cryo-FIB milled CPSF6 k.d. SupT1-R5 cells infected with NNHIV WT-CA. Scale bar: 100 nm. (A) Slice through a

tomographic reconstruction showing a capsid (black arrowhead and insert) localized in proximity to the nuclear envelope, positioned in top view. The insert on top shows the conical capsid with a different orientation. The insert at the bottom left shows the portion of nuclear envelope highlighted in the dashed white rectangle with a different orientation to make double membranes visible. (A') Same as in (A) but in a different orientation. The interpretation on top shows a portion of NPC contour and cytoplasmic filaments and the nuclear envelope below. Green and black arrowheads point to the NPC and the capsid, respectively. (A'') NPC in (A and A') oriented in top view. **(B-D)** CLEM overlays with enlarged regions indicate the positions of IN.mScarlet signal (red; white arrows) at the nuclear envelope in resin sections of cells stained with mCLING.Atto647N (magenta), post-stained with Hoechst (blue) and decorated with multi-fluorescent fiducials (Fd) for correlation. (B') Slices through tomographic reconstructions of three different HIV-1 capsids proximate to the same NPC. Different orientations of the capsid structures are shown in the insets. Capsids at the cytoplasmic face of the NPC display a conical shape and a dense interior. (C', D') Same as (B') but showing examples of empty capsid-related structures (empty black arrowheads) at the nucleoplasmic side of the NPC (C') or in nucleoplasm beneath the nuclear envelope (D'). Cy, cytosol; Nu, nucleus; NE, nuclear envelope.

Figure modified from (Zila et al., 2021). Experiments in panel (B-D) performed and analyzed by Vojtech Zila.

4.4. The hexameric lattice is nearly intact in the cone-shape capsids but barely detectable in the nucleoplasm

To correlate the morphology of the A77V capsids found in the cytosol, within the NPC, and in the nucleoplasm with the status of assembly of their hexagonal lattice, Beata Turoňová and I performed SA as previously described for capsids in intact virions (Mattei et al., 2016) and in methods 3.5 (**Figure 4.11**). Our expectation was to detect a hexagonal lattice nearly comparable to that in mature intact virions for the cone-shaped capsids but not for the tube-shaped or fragments.

During iterative averaging, we observed that the subtomograms of the cone-shaped capsids in the cytosol and within the NPC converged into regular hexameric lattices in which six adjacent CA hexamers surround one central CA hexamer in a regular fashion (**Figure 4.11 A-B'**).

In comparison to the previous data obtained from isolated virions (Mattei et al., 2016), however, the lattice information was recovered for less of the capsid surface. This finding might be interpreted as a partial perturbation of the hexagonal lattice. However, the clearly defined capsid edge and the well-preserved overall cone shape visible in the tomograms suggest that rather technical parameters, such as the reduced SNR ratio due to specimen thickness of the FIB-lamellae and the crowded cellular environment, have resulted in an incomplete lattice recovery during SA.

On the other hand, when SA was applied to the capsid-like structures in the nucleoplasm, only few subtomograms showed high cross-correlation in respect to the reference and, in most of the cases, CA hexamers did not converge into an overall hexagonal lattice.

Exceptionally, lattice elements were observed in the tube-shaped structure. However, only 27 CA hexamers on the surface properly converged, much less than what was detected on the cytoplasmic structures (**Figure 4.10 C**; **Figure 4.11 C and C'**), possibly suggesting higher disorder of the lattice architecture.

We concluded that cone-shaped HIV-1 capsids containing the genomic material, with an either completely or largely intact lattice, can enter the central channel of the NPC. The data also suggest that capsid disassembly should not be conceived as immediate dissolution of the lattice into individual subunits after nuclear entry, but rather as partial disruption of the capsid that allows for the release and dissociation of the viral genome from capsid remnants, possibly due to mechanical strain. These findings support a model in which disruption of the capsid lattice occurs subsequent to translocation through or upon departure from the central channel of the NPC.

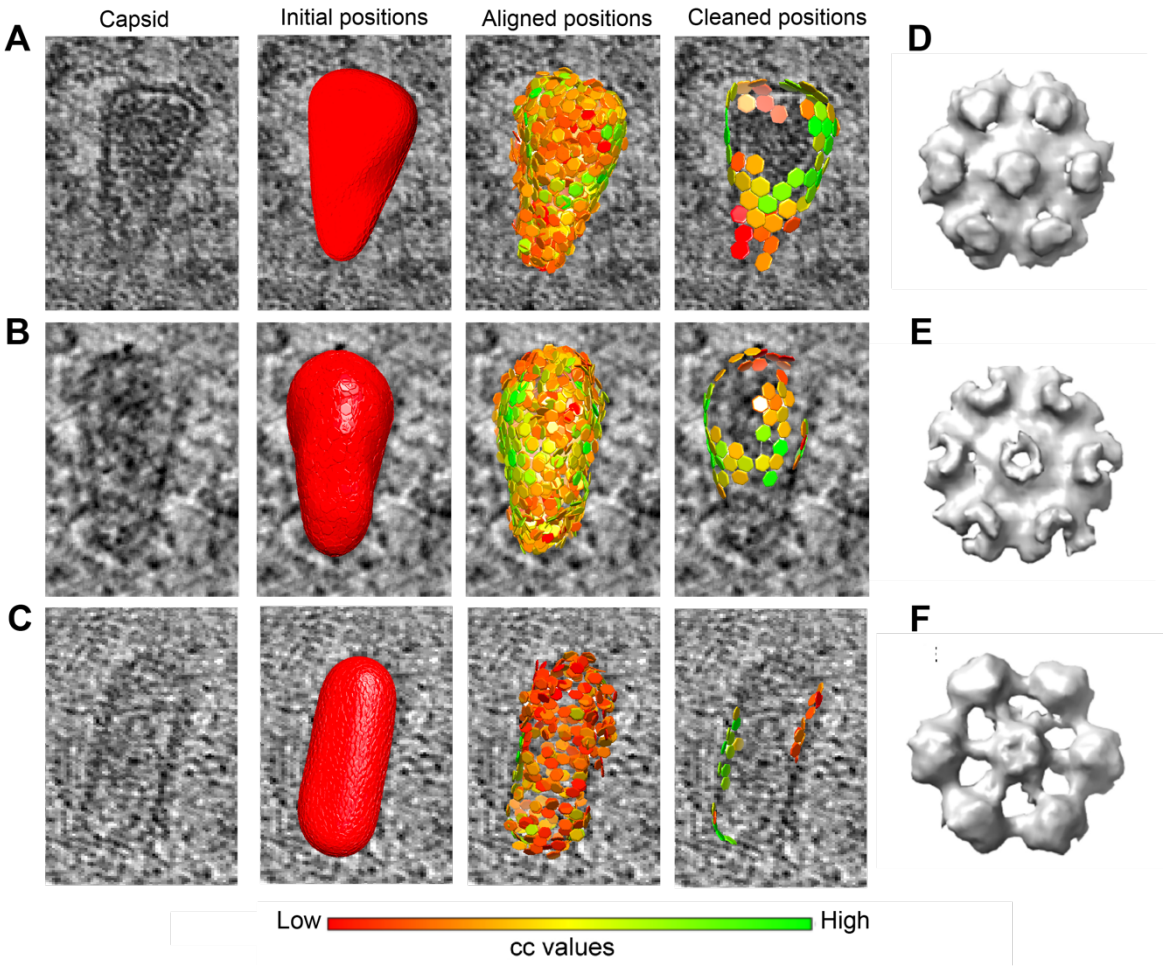


Figure 4.10. Subtomogram averaging of CA hexamers and reconstruction of the capsid lattice. (A-C) For the identified capsids in the cytosol (A), at the NPC (B), and in the nucleus (C), ~1000 initial positions (4× oversampling) were extracted from their manually segmented volumes and subjected to subtomogram averaging. During SA, the oversampled initial positions shifted to the positions corresponding to the hexamers in the CA lattice (aligned positions). Misaligned and overlapping subtomograms were removed, leaving only a subtomogram with the best cross-correlation to the reference (cleaned positions). **(D-F)** Hexagonal subtomogram averages obtained from capsids in the cytosol (D), within the NPC central channel (E) and in the nucleus (F).

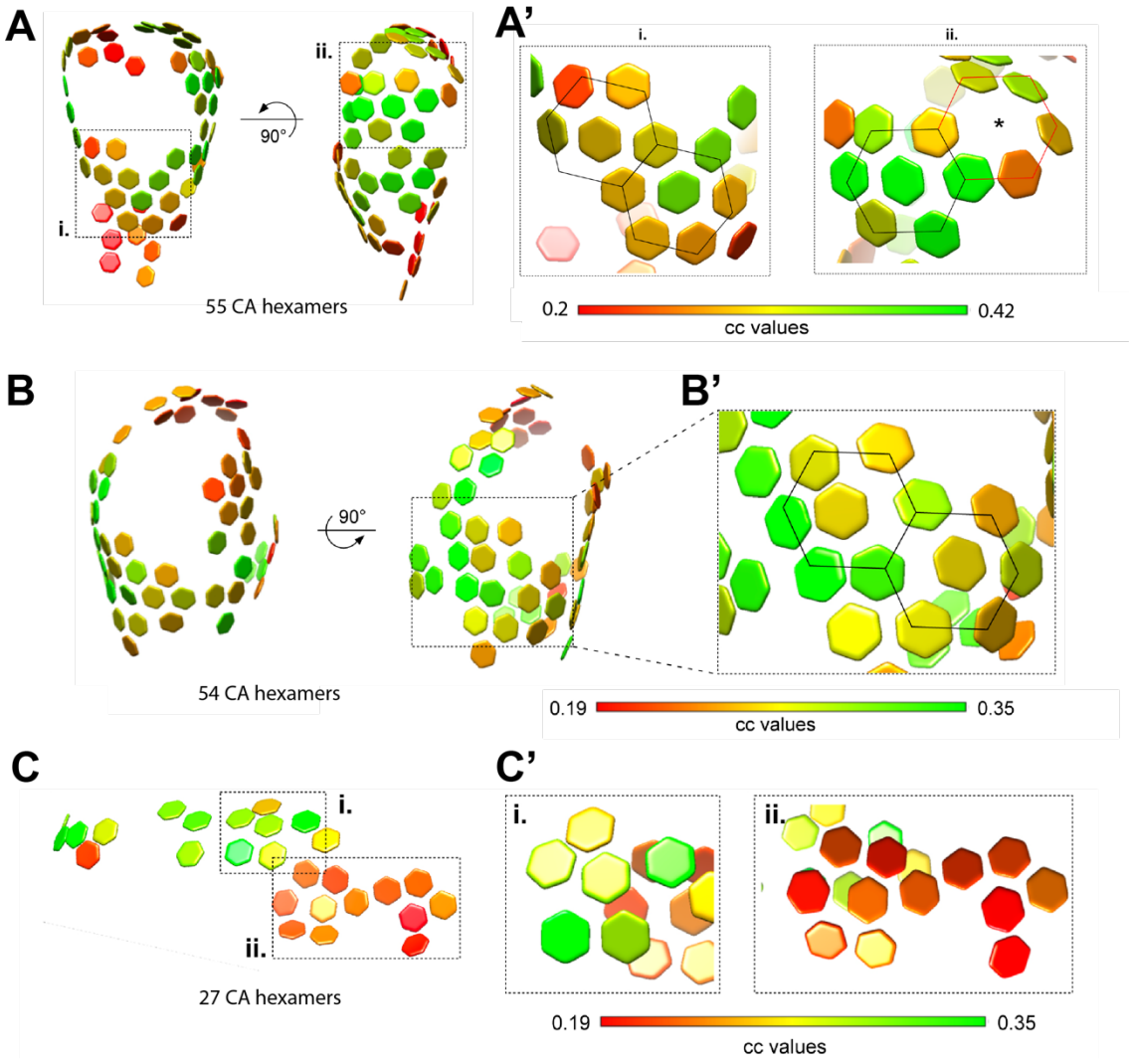


Figure 4.11. Lattice elements are detected more extensively in the cone-shaped than in the tube-shaped capsid. (A) Lattice of CA hexamers obtained after performing subtomogram averaging and cleaning the overlapping aligned subtomograms by cross-correlation (cc) values as in Figure 4.10. Cc values for each CA hexamer in the volume are shown color-coded as indicated in the colormap. **(A')** The areas highlighted as i. and ii. of the lattice shown in (A), contain consecutive regular arrangements of six CA hexamers (highlighted by black lines) surrounding a seventh, central CA hexamer; in ii. six CA hexamers are localized to a highly curved region (red lines). The position of central CA hexamer was not detected (star). **(B)** Same as in (A) but for the capsid within the NPC. (B') Black lines highlight the regular arrangement of six CA hexamers

surrounding a seventh, central CA hexamer. **(C)** Same as in (A) but for the tube-shape structure in the nucleoplasm. **(C')** Incomplete CA hexamers are shown in i. and ii. Figure modified from (Zila et al., 2021).

4.5 The structure of NPC in human T lymphocyte upon HIV-1 infection

The data described above revealed that largely intact HIV-1 capsids can penetrate into and pass through the NPC central channel. However, the central channel diameter of the human NPC was reported to be ~40 nm (von Appen et al., 2015), which is sufficiently wide to allow transport of basically any large cellular cargo, but too narrow to allow passage of intact HIV-1 capsids. This might sound counterintuitive and would require a remodeling or partial disruption of the NPC upon HIV-1 infection as previously suggested (Dharan et al., 2016; Monette et al., 2011). In this regard, displacement of Nup358 from NPCs was proposed as mechanism that might alter NPC structure and increase the chances of viral core to transverse the nuclear pore. Alternatively, the diameter of NPC of T cells might be wider *in cellulo* than reported for NPC *in situ*, as previously shown for HeLa cells *in cellulo* (Mahamid et al., 2016).
To explore these possibilities, I revisited the architecture of the human NPC *in cellulo* under conditions relevant to infection.

4.5.1 NPCs remain intact upon HIV-1 infection

In order to analyze the architecture of the NPC in HIV-1 infected T cells, I manually picked 99 NPCs from the cryo-electron tomograms, as described in methods 3.6.1.
In human T lymphocytes, ~2.5 NPCs can be found on average per tomogram, with the best possible scenario pointing to 3 NPCs per tomogram (**Figure 4.12 A**). After initial alignment of the whole pore, 792 SUs were extracted and subjected to SA (see methods 3.6.3). The resulting cryo-EM map with a resolution of ~37 Å captures the native conformation of actively transporting NPCs in HIV-1 infected T cells (**Figure 4.12 B and C**).

The overall NPC architecture appeared to be organized as previously described (Beck and Hurt, 2017) in three rings, CR, IR and NR, following a 8-fold rotational symmetry (**Figure 4.12**; **Figure 4.13**). Both the CR and the NR showed a dimerized Y-complex arrangement with a total of 32 copies of Y-complexes per ring (**Figure 4.13**). Nup358 density was observed at the CR and connected the outer and inner Y-complexes (**Figure 4.13**), in line with previous analyses of NPCs from isolated NEs (von Appen et al., 2015; Beck and Hurt, 2017; Bui et al., 2013). Although I cannot formulate conclusions for the components of NPC filaments which are excluded from the NPC average due to their flexibility, these results showed that the scaffold of NPC in T cells is generally preserved upon HIV-1 infection. Therefore, I concluded that cone-shaped HIV-1 capsids are translocated through intact NPCs.

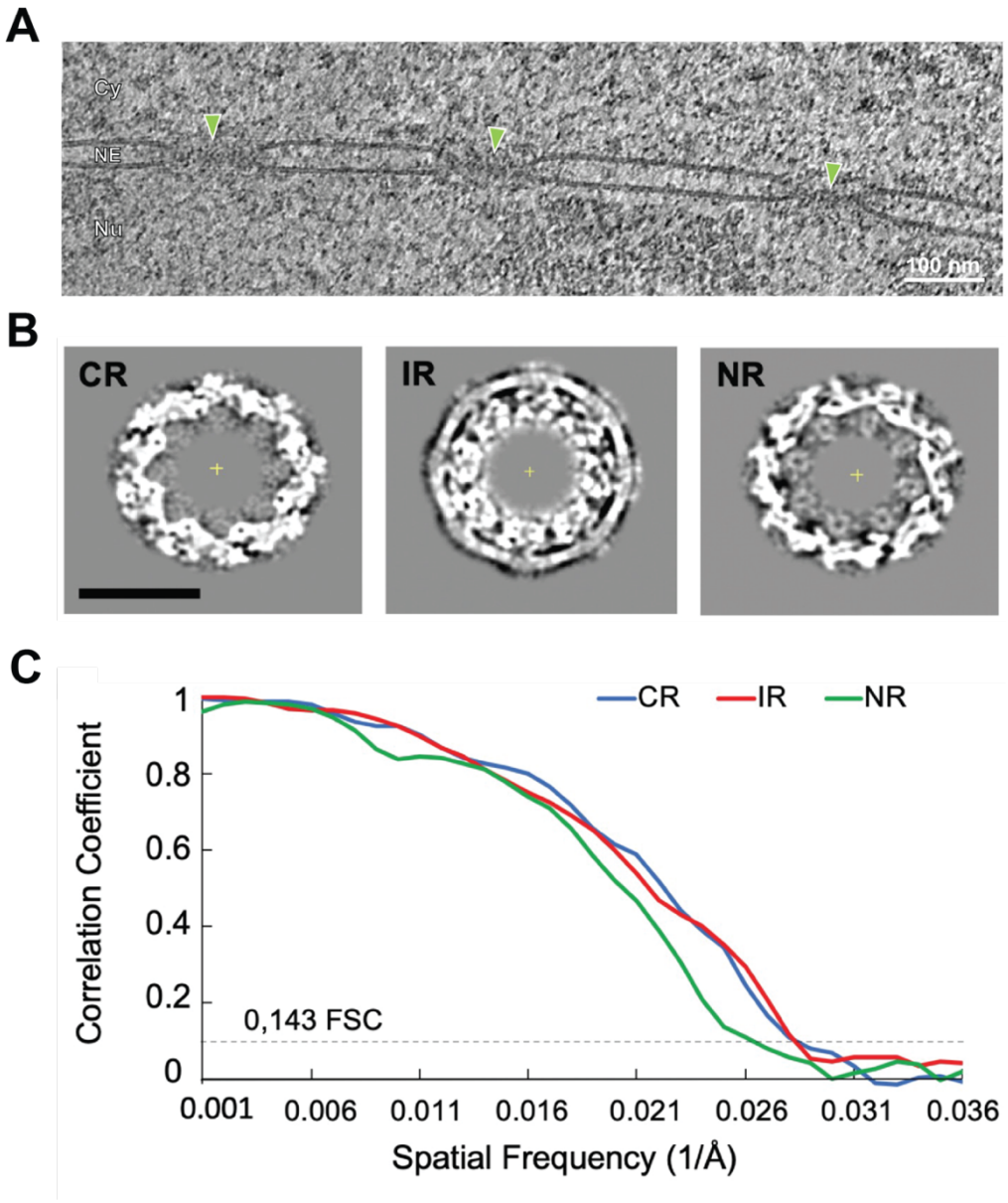


Figure 4.12. *in cellulo* subtomogram average of human NPC in HIV-1 infected T cells. (A) Slice through a representative cryo-electron tomogram as used for structural analysis. NPCs indicated by green arrowheads; Cy, cytosol; Nu, nucleus; NE, nuclear envelope. **(B)** Slices through the cryo-EM map of the NPC at the level of the CR, IR and NR. Scale bar: 50 nm. **(C)** Resolution estimation by gold standard Fourier shell correlation (FSC). After alignment of the whole pore, each NPC was splitted into the even and odd subunits. The cytoplasmic (CR), nuclear (NR) and inner ring (IR) from each subunit were independently subjected to subtomogram averaging. The gold standard FSC was calculated with FSC server – EMBL-EBI with half maps as input. The curves for CR, NR and IR intersect the 0.143 criterium respectively at the resolution of 36.3 Å, 35.8 Å, and 39.9 Å. Figure modified from (Zila et al., 2021).

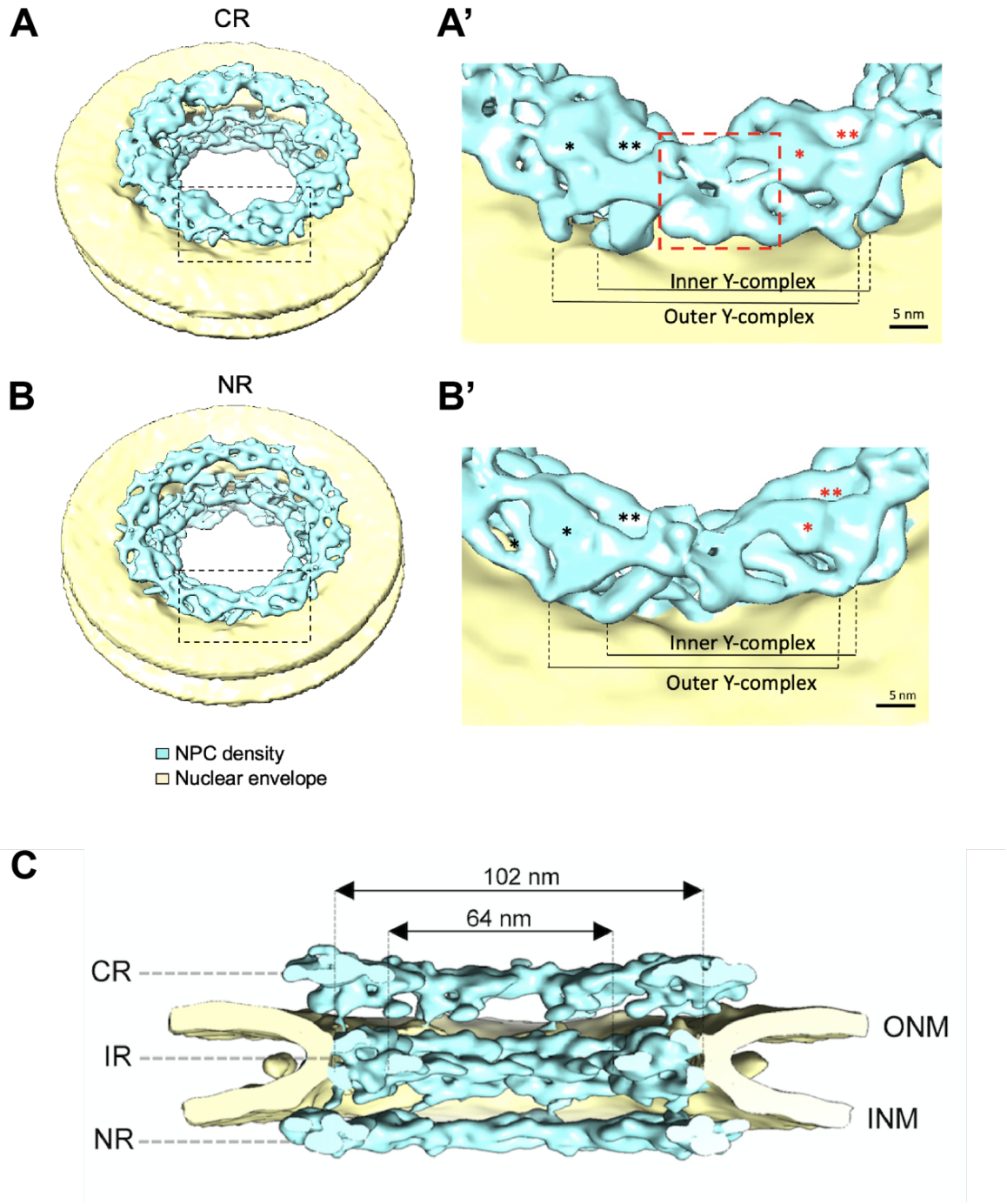


Figure 4.13. The architecture of the NPC in HIV-1 infected SupT1-R5 T cells.

(A-C) Isosurface rendering of the cryo-EM map of the NPC (cyan) within the NE (yellow). (A-A') Top view of the cytoplasmic ring (CR) with black dashed box highlighting the Y-complex (A), enlarged in (A'). Inner and outer Y-complex are labeled. Two consecutive Y-complexes are pointed with black and red asterixis. Single and

double asterixis mark the outer and inner Y-complexes, respectively. Density for Nup358 is highlighted with the red dashed box. (B-B') Same as (A-A') but for nuclear ring (NR) (B) and the Y-complex enlarged in (B'). (C) Same as (A and B) but shown as cut-open view. CR, NR, inner ring (IR), outer nuclear membrane (ONM) and inner nuclear membrane (INM) are labelled.
Figure modified from (Zila et al., 2021).

4.5.2 NPC central channel is sufficiently wide to translocate intact capsids

I next measured the diameter of individual NPCs from the average of the asymmetric unit as described in methods 3.1.6.5. In line with other studies conducted in intact human cells (Beck and Baumeister, 2016; Mahamid et al., 2016), NPCs in T cells appeared dilated in comparison to the NPCs in isolated nuclear envelopes (**Figure 4.14 A**), with 85% of them (77/90) displaying a central channel diameter ≥60 nm. The average NPC diameter corresponded to ~64 nm (**Figure 4.14 B**). This discrepancy between the diameter of NPC *in cellulo* and *in situ* can be attributed to the sample preparation. Previous structural analyses (von Appen et al., 2015) were performed using nuclear envelopes purified from human cells in which mechanical tension is relieved and RanGDP/GTP gradient at the two sides of the nuclear membrane lost. It was recently shown that the energetic status of the cell affects the NPC diameter and that energy depletion causes NPC constriction (Zimmerli et al., 2020).
These findings indicate that the NPC structure observed under the relevant conditions is representative of the transporting state, while the constricted state observed in isolated nuclear envelopes (von Appen et al., 2015) may be more relevant to stress conditions (Zimmerli et al., 2020). Most importantly, my data show that the inner diameter of the central channel exceeds the dimensions of the broad end of intact HIV-1 capsids (~55–60 nm) by ~4–9 nm, rendering the nuclear entry of intact capsids geometrically possible.

4.5.3 HIV-1 infection is not a trigger of NPC dilation

It has been long suggested that NPCs might adopt an increase in diameter to accommodate for large cargoes across the NE or compensate for mechanical stress (Beck et al., 2007; Lin and Hoelz, 2019). Viral infections are also reported to work as triggers of NPC dilation. As examples, herpes simplex virus (HSV-1) and bovine herpesvirus type 1 (BSH-1) are shown to significantly increase the NPC size up to ~6-fold during late infection (Leuzinger et al., 2005).

To address if HIV-1 infection may promote NPC dilation, I collected ~90 cryo-electron tomograms of non-infected SupT1-R5 cells and measured the diameter of 39 NPCs.

As previously done for the reconstruction of the NPC from HIV-1 infected SupT1-R5 cells, only high-quality NPCs within the cryo-lamellae were manually picked and used in the SA. The final average of the SU derived from 336 SUs was used to measure the diameter of NPCs at the IR and at the ONM-INM fusion point in non-infected cells, as described in methods 3.7. The NPC diameter at the IR in non-infected SupT1-R5 T cells corresponded to 62 nm, in comparison to that of 64 nm in HIV-1 infected cells (**Figure 4.14 A**). This difference of 2 nm was significant but very small (**p=0.0055; by unpaired two-tailed Student's t test). On the other hand, measurements of diameters at the ONM-INM fusion point did not show any significant difference.

In conclusion, my results showed similar diameters between HIV-1 infected and non-infected T cells, suggesting that NPC dilates *in cellulo* in respect to the NPC *in situ* but not upon HIV-1 infection.

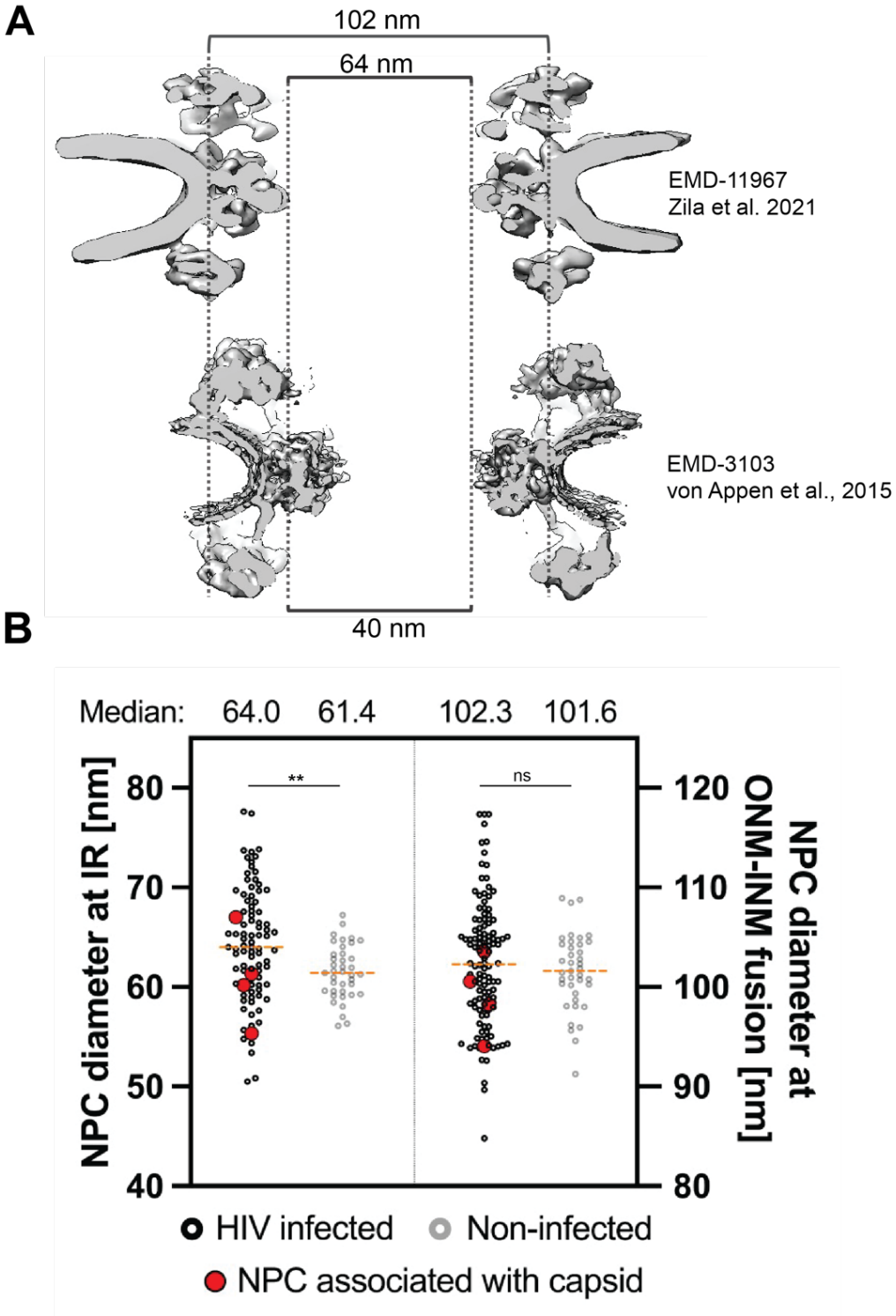


Figure 4.14. The NPC scaffold is dilated in HIV-1 infected cells in comparison to the NPC structure *in situ*, independently by infection. (A) Comparison of the structure of the NPC from this study and the previous human NPC from isolated NEs. Broken lines indicate the diameter at IR or at ONM-INM fusion point and highlight a dilated

conformation of the *in cellulo* structure in comparison to that *in situ*. **(B)** Box plot showing the distribution of diameters of individual NPCs measured either at the relevant opening of central channel at the IR or at the ONM-INM fusion point in HIV-1 infected (n = 90) or non-infected (n = 39) cells. Values of NPCs with associated capsid (n = 4) in red; median values as orange dashed lines. Statistical significance was assessed by unpaired two-tailed Student's t test: **p=0.0055, ns = not significant.
Figure modified from (Zila et al., 2021).

4.5.4 The diameter of the NPC is equivalent between primary CD4+ T cells and SupT1-R5 cell line

As previously mentioned, my original intention was to study HIV-1 nuclear entry in primary $CD4^+$ T in order to work with the most relevant conditions for HIV-1 infection. Although infectivity experiments were not strong enough to continue working with this cell type in cryo-conditions, cryo-electron tomograms of non-infected primary $CD4^+$ that I collected for initial feasibility tests revealed to be useful to reconstruct the NPC structure from a primary cell type. This was not previously done and allowed to determine whether the NPCs diameters varies between a primary cell type and its cell line.

47 NPC were manually picked and subjected to SA. 376 SUs were aligned and final average was used to measure the diameter of individual NPCs as previously done for infected and non-infected SupT1-R5 T cells. I found that the NPCs in primary $CD4^+$ cells have comparable diameters of NPCs in the SupT1-R5 cell line, with an average diameter of ~61 nm at the IR and ~98 nm at the ONM-INM (**Figure 4.14**).

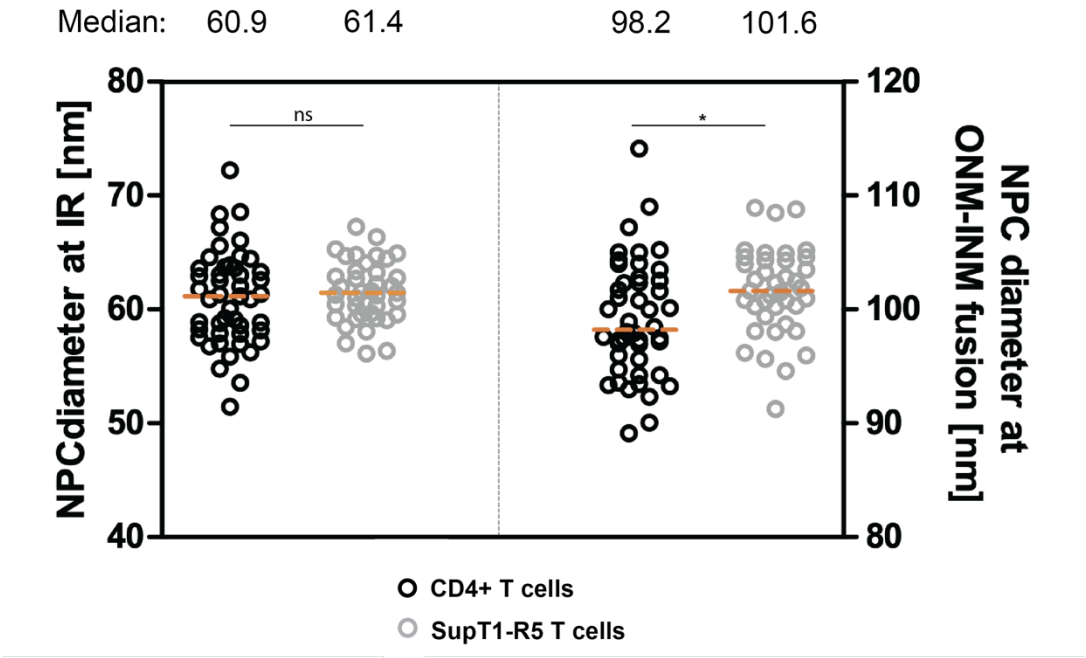


Figure 4.15. Differences in NPCs diameters between primary T cells and the relative cell line. Box plot showing the distribution of diameters of individual NPCs measured either at the relevant opening of central channel at the IR or at the ONM-INM fusion point in not-infected primary $CD4^+$ (n = 47) or non-infected SupT1-R5 (n = 39) T cells. Median values as orange dashed lines. Statistical significance was assessed by unpaired two-tailed Student's t test: *p=0.0172, ns = not significant.

Chapter 5

Discussion

5.1 Mechanism of HIV-1 capsid uncoating

5.1.1 HIV-1 capsids uncoat in the nucleus of infected cells

Nuclear entry of HIV-1 is a rare event as a consequence of the inefficient productive infection. Approximately ~50 particles were reported to enter each cell, of which only ~2-3 enter the nucleus and ~1 forms a provirus (Burdick et al., 2017). Although HIV-1 viral cores require a residence time of ~1.5 ± 1.6 hours at the NEs prior to import (Burdick et al., 2017), actual transport of individual molecules across the NPC is very fast and reported to occur in ≤10 ms (Ribbeck, 2001).
Nevertheless, in this study, I was able to visualize non-infectious RT-competent HIV-1 particles in the process of nuclear entry by cryo-ET *in cellulo*. Precisely, viral particles were localized in the cytosol, docking to or within the central channel of NPCs (**Figure 4.5** and **4.6**), and in the nucleoplasm of SupT1-R5 T cells (**Figure 4.7**).

In the cytosol, they retained conical capsids with a size of ~112 nm in length and ~55 nm in width, nearly identical to those within mature virions (Mattei et al., 2016). The cone-shaped capsids appeared dense at their interior, suggesting they were engaged with the viral genome. This is in line with the capsid primary function in maintaining a close environment where the genome can be protected from the host innate immune responses initiated by cytosolic DNA sensors (Campbell and Hope, 2015). Among these, an interferon-induced DNA sensor named cGAS was described to inhibit HIV-1 infection exclusively upon capsid destabilization in MDMs (Rasaiyaah et al., 2013).

Moreover, our results provided a significant contribution to an open debate related to time and cellular location of uncoating. Several studies pointed towards an early cytoplasmic uncoating within ~30-40 minutes upon HIV-1 entry (Jun et al., 2011; Mamede et al., 2017). Among these, a recent work relied on *in situ* fluorescence imaging of fluorescently labeled replication complexes in live cells, whereas another applied correlative live FM with cryo-ET to directly visualize HIV-1 *in cellulo* (Jun et al., 2011). In the latter, wild-type and hyperstable E45A HIV-1 cores were tracked inside cells by live-imaging and subsequently investigated by cryo-ET upon plunge-freezing. Whereas cone-shaped hyperstable capsids were visualized intracellularly, wild-type capsids were undetectable and therefore suggested to disassemble immediately after membranes fusion.

Our data argue against this model. One cytosolic capsid was found at ~200 nm from the NE (**Figure 4.5 B**), whereas other two were located in very close proximity to the NEs, docking to the NPCs (**Figure 4.5 C** and **4.5 D**). In this regard, one might argue that the A77V CA mutation could affect the stability of the cores as shown for the E45A mutation in (Jun et al., 2011). However, our data of wild-type capsids under CPSF6 silencing strongly disagree (**Figure 4.8**). Although morphology, size and genomic content suggested an integrity of these structures, the assembly of the CA hexamers into a lattice arrangement was confirmed by SA (**Figure 4.7A**; **Figure 4.8A**).

Another study based on light microscopy suggested that intact cores arrive at the NPCs, where they entirely or considerably disassemble prior to nuclear import (Arhel et al., 2007; Burdick et al., 2017; Dharan et al., 2016; Fernandez et al., 2019; Francis and Melikyan, 2018; Novikova et al., 2019; Peng et al., 2014). Our data also challenge this model of uncoating at the NPC.

First, a cone-shaped capsid encasing the viral genome was captured in the process of nuclear import through the NPC: part of the cone was deeply located within the NPC central channel, whereas its narrow end already reached the nucleoplasm. (**Figure 4.6**). This observation was confirmed by CLEM/ET data from Vojtech Zila (Zila et al., 2021). Second, the extent of lattice assembly was comparable to that of the capsid docking to the NPC, as shown by our SA analyses (**Figure 4.11 B** in comparison to **Figure 4.11 A)**, suggesting that HIV-1 capsids are intact or nearly intact during NPCs translocation.

Third, capsids-like structures were found in the nucleoplasm still in close proximity to the NPCs (**Figure 4.7**). Extensive disassembly of capsids at the NPCs would have rather impede the visualization of capsids in the nucleus of infected cells.

In comparison to the capsids in the cytosol, these particles lost their conical shape and differed in size: some of these were large tube-shaped structures, others were small rounded fragments or remnants (**Figure 4.8**). Interestingly, we found a correlation between the morphological change and the rearrangement of the hexameric lattice. As suggested by SA, very few elements of the hexameric lattice were retained within the tube-shape structure (**Figure 4.11 C** in comparison to **Figure 4.11 A** and **B**), whereas they were completely lost within the fragments. These results well complemented other studies that observed variable amounts of CA inside the nucleus of different cell types (Bejarano et al., 2018, 2019; Chen et al., 2016; Chin et al., 2015; Hulme et al., 2015; Peng et al., 2014; Stultz et al., 2017; Zhou et al., 2011). Interestingly, the nuclear structures appeared empty, suggesting that the shape remodeling and the lattice loss cause the release of the viral genome.

Although I cannot entirely exclude that some HIV-1 particles undergo uncoating during NPC translocation (Novikova et al., 2019), two major observations argue against it. First, CLEM/ET identified few capsids inside the nucleus that were still conical rather than tubular (Zila et al., 2021). Second, cryo-ET and SA allowed to demonstrate that conical capsids with intact or nearly intact lattices penetrate the central channel of the NPC. Therefore, I propose a model in which uncoating occurs inside of the nucleus of infected cells.

5.1.2 Capsid uncoating might be triggered by the completion of reverse transcription within intact cones

Beside determining the location of uncoating, the visualization of empty tube-shape or fragmented capsids within the nucleus revealed an interesting insight into the mechanism of uncoating. Previously conceived as a total disassembly of CA proteins from the viral genome, I rather showed that HIV-1 uncoating consists in a partial opening of capsids due to shape remodeling and lattice disruption that allows release of viral genome. This

model of uncoating is in line with a very recent study that visualized newly synthetized cDNA directly emanating from broken capsids by combining the direct labeling of viral cDNA with either immunofluorescence or CLEM in $CD4^+$ T lymphocytes and MDMs (Müller et al., 2020). Separation of the cDNA from the capsids and IN was subsequently observed in both cell types.

On a speculative note, I believe that opening of the conical capsid might potentially be caused by relieving the strain imposed by the CA pentamers. In intact capsids, the closure of the shell encasing the viral genome into a cone depends on the CA pentamers that localize at the broad and narrow ends of the cone. In comparison to the cones, CA tubes are rather open structures that lack pentamers and regions of high curvature. Loss of the curvature imposed by the CA pentamers and lattice rearrangement might be responsible for the remodeling of the capsid. For this to happen, a driving force is further required. Completion of reverse transcription represents one valid possibility. The generation of double-stranded DNA within intact capsid may impose mechanical strain from the inside and therefore induce uncoating.

Although reverse transcription, as well as capsid uncoating, was previously believed to occur in the cytoplasm, there are growing evidences suggesting that this process is only initiated in the cytoplasm but completed in the nucleus (Burdick et al., 2020; Dharan et al., 2020; Müller et al., 2020; Selyutina et al., 2020). In line, blockage of the nuclear pore is shown to prevent completion of reverse transcription (Dharan et al., 2020).

Additionally, reverse transcription in an intact or nearly intact viral core is possible: CA hexamer contains a positively charged channel that allows for transport of deoxynucleotide triphosphates into viral cores, providing the substrates necessary for reverse transcription within intact viral cores (Jacques et al., 2016). This interpretation is also supported by many other finding from previous studies (Rankovic et al., 2017; Rouzina and Bruinsma, 2014). For example, inhibition of reverse transcription was found to prevent uncoating even if completed ~2 to 3 h earlier, suggesting that uncoating is not necessarily initiated upon completion of reverse transcription but requires the completion of viral DNA synthesis (Burdick et al., 2020). Moreover, reverse transcription within purified HIV-1 capsids (Christensen et al., 2020) revealed partially broken capsids with polynucleotide loops emanating from the capsid lattice when cDNA synthesis was largely completed.

Nevertheless, it might be that not a single force but rather a combination of events, such as reverse transcription combined with the binding of several host factors to the CA, induces uncoating. Many host factors have been proposed in this regard, such as Nup358 (Dharan et al., 2016), CPSF6, Nup153, TNR-1 (Fernandez et al., 2019). In order to study their role in this process, one might establish an *in vitro* system where cone-shaped IP6-stable capsids (Christensen et al., 2020) are exposed to isolated nuclei supplemented with cytosolic extract and transport buffers (Adam et al., 1990; Cassany and Gerace, 2008). The effect on capsid shape and assembly of individual or multiple wild-type or mutated proteins at different concentrations can be subsequently explored by negative staining or cryo-ET.

5.1.3 Uncoating in the nucleus occurs near the integration sites in a CPSF6-dependent mechanism

In this study, the HIV-1 viral particles that were used for the infection carried IN fluorescently labeled with mScarlet. As viral INs mediate the integration of the reverse-transcribed cDNA into the host genome, I expected them to remain at least partially associated with the viral genome until completion of integration. Without the correlation of the fluorescence signal, my cryo-ET data did not allow to visualize the viral genome in the dense nuclear environment. In contrast, the INmScarlet signal detected in CLEM experiments led to the identification of the disrupted capsids inside the nucleus (**Figure 4.7 F-H**). Although this system only indirectly labeled the viral genome, our results suggested that the opening of the capsid in the nucleus and the release of the viral genome most luckily occur in close proximity to the integration sites, in a mechanism mediated by the host factor CPSF6. In absence of CPSF6 binding by either the A77V mutation or CPSF6 depletion, tube-shaped or capsids fragments were identified in the nucleoplasm underneath the NPCs, still in association with their nuclear basket (**Figure 4.7**; **Figure 4.9**). On the other hand, in wild-type conditions, HIV-1 capsid-like structure were found deeply inside the nucleus (Zila et al., 2021). Similarly, in Burdick et al., the authors reported that while integration was normally occurring at ~1.5 μm from NE, A77V mutants integrated into within ~0.8 μm from the NE (Burdick et al., 2020). In the last case, the integration sites corresponded to transcriptionally repressed, lamina associated

gene-sparse heterochromatin regions (Achuthan et al., 2018; Burdick et al., 2020). Our data thus support the reported role of CPSF6 in nuclear import (Bejarano et al., 2019; Zila et al., 2019) and in the integration site selection (Achuthan et al., 2018; Sowd et al., 2016) and further indicate that exchange of NUP153 for CPSF6 on hexameric lattice is important for the location of capsid uncoating in the nucleus.

5.2 Mechanism of HIV-1 nuclear entry

The dynamic of processes inside the cells can only partially be reconstructed from cryo-electron tomograms, which represent snapshots of a cell frozen in a certain state. In this sense, my study cannot unravel the dynamics of interactions between the cellular host factors and HIV-1 capsids. However, several observations provided some insights into the mechanism of HIV-1 nuclear entry through the NPC.

5.2.1 The dilated scaffold of NPC allows the translocation of intact or nearly intact HIV-1 capsids

My model of HIV-1 uncoating inside the nucleus is in line with a very recent study from (Burdick et al., 2020). In this work, Burdick et al. tracked GFP-CA labeled virions by live-imaging and compared the average intensities of viral complexes before and after nuclear entry. In support to our SA results, they did not detect loss of GFP-CA upon nuclear entry, suggesting that most of the CA is retained during translocation through the NPCs. How intact capsids with a width of ~61 nm were capable of entering NPCs with an inner diameter of ~39 nm remained unclear from the previous study. My work provided an answer to this.

Remodeling of the NPC or increased NPC permeability by Nups displacement might have constituted a possible explanation, as previously suggested (Monette et al., 2011). Different viruses were indeed reported to alter the NPC structure by displacement of several Nups, such as Nup62 or Nup358 in the case of HIV-1 (Cohen et al., 2012; Dharan et al., 2016; Monette et al., 2011). My SA analyses excluded this hypothesis and revealed

that the overall NPC architecture was not fundamentally altered by HIV-1 infection (**Figure 4.13**). It is worth it to mention however that during SA procedure, multiple NPCs are averaged together, therefore only a persistent change in the population would be detectable in the final average. Additionally, the *in cellulo* human map presented here has a substantially lower resolution in comparison to that obtained from isolated NEs due to a lower number of particles included in my analysis. A cryo-EM map at higher resolution of NPC from HIV-1 infected cells combined with integrative modelling might provide more insights into the arrangement of different subcomplexes, such as Nup93 and Nup62 at the inner ring. Finally, proteomic analyses in the future might be useful to confirm my conclusions.

More interestingly, I revisited the size of the human NPC *in cellulo* and found that the average diameter of the inner ring of NPC is wider than previously reported for the human NPC structure *in situ* (von Appen et al., 2015). Although it has been suspected for a long time that the NPC scaffold architecture might adapt its conformation to e.g. accommodate for the transport of large complexes, I found that this conformation was independent of HIV-1 infection. Non-infected SupT1-R5 also displayed a dilated scaffold of ~62 nm in diameter. Therefore, my explanation of how intact HIV-1 capsids can enter the NPCs is based on their geometry: with an inner ring of ~64 nm, the NPC of SupT1-R5 T cells exceed of ~9 nm the broad end of an intact capsid, and thus is sufficiently large to complete its translocation.

5.2.2 The curvature of the cone facilitates docking and translocation of HIV-1 capsid through the NPC

In this study, I observed that cone-shaped HIV-1 capsids dock to and accommodate into the NPCs with their narrow tips perpendicularly oriented towards the NPCs. On a speculative note, the preferential binding to the NPC with the narrow end of the cone might be dependent on CA pentamers, which are strongly enriched in this area of the capsids. Primary function of CA pentamers is to induce capsid curvature at the extremities of the cone (Ganser et al., 1999; Mattei et al., 2016). Interestingly, changes in local curvature were proposed to influence the capacity of several cellular factors to access their binding sites on the capsid (Ni et al., 2020). For example, Nup153 and CPSF6

bind to a binding pocket formed by the CA_{NTD}-CA_{CTD} interface, which was shown to be opened up in the CA pentamers in comparison to the CA hexamers (Mattei et al., 2016).

The affinity and stoichiometry of CypA binding to *in vitro* self-assembled CA spheres or tubes was recently explored (Lau et al., 2020). CA tubes are open structures that resemble the middle section of the cone as they lack pentamers and regions of high curvature; on the other hand, CA spheres adopt the same pentamer−hexamer interfaces as found in the highly curved ends of the HIV-1 capsids. In line with my hypothesis, CypA was found to bind to highly curved sphere with a higher stoichiometry than to the tubular hexameric lattice.

Similarly, a second study *in vitro* confirmed my observation of preferential binding orientation of the capsid with respect to the NPC. To gain insights into the mechanisms of nuclear entry, the author of this study focused on CA-Nup153 interactions. In this regard, they generated a NuPODs with Nup153 C-terminal domains attached to the inner or outer surface of a 45 nm diameter DNA-origami ring and expose it to *in vitro* assembled CA with different shapes (Shen et al., 2020). In line with the previous study, they showed that CA tubes often bound Nup153 through their flat side surfaces, whereas CA cones bound and inserted into the Nup153 NuPOD with the narrow ends. This strongly support a role of capsids curvature in the interaction with NPCs during HIV-1 nuclear entry.

While the capsid is exposed to Nup153 only once it penetrates deep beyond the central channel and reaches into the nuclear basket region of the NPC, docking to NPCs is reported to be mediated by Nup358 (Burdick et al., 2017; Dharan et al., 2016; Schaller et al., 2011). Similarly to CypA and Nup153, I hypothesize that the Cyp-binding domain of Nup358 binds the narrow end of the capsid with higher stoichiometry than the rest of the cone to mediate docking to the NPC. This might facilitate a capsid orientation that is advantageous for the subsequent penetration through a dense meshwork of FG-Nups located within the central channel. To confirm my hypothesis, one might use a similar *in vitro* systems reported in (Lau et al., 2020; Shen et al., 2020).

The higher affinity and stoichiometry of binding of these factors to regions of high curvature can be speculated to indirectly prove the importance of lattice integrity. In an intact lattice, binding interfaces that are absent in a disassembled lattice would allow to

recruit different host factors to the capsid surface, that can ultimately mediate fundamental processes in HIV-1 life cycle such as HIV-1 nuclear entry and uncoating.

5.2.3 Model of HIV-1 nuclear entry

Taken together, the data in my thesis point to the scenario schematically outlined in **Figure 5.1**. Following cytosolic entry, intact cone-shaped HIV-1 capsids travel towards the nuclear periphery. The subsequent nuclear import of capsids is three-staged. (i) Intact capsids dock to NPCs with the pentamer-rich ends of the capsid, preferably the narrow end. Here they encounter the FG-repeats and Cyp domain of NUP358 bound to the cytoplasmic face of NPCs. (ii) Subsequently, intact capsids penetrate deeply into the central channel of the NPC where they are exposed to a very high local concentration of FG-NUPs of the NUP62 complex. Although this environment is spatially confined, the diameter of the NPC central channel as determined *in cellulo* is physically compatible with translocation of the intact HIV-1 capsid. Up to this stage, the hexagonal lattice and the typical shape of HIV-1 capsids are clearly detectable. (iii) Upon departure from the NPC central channel, capsids encounter NUP153 and CPSF6. At this stage, which can still be conceived as part of the actual nuclear import process, disrupted capsids are detected. The cone shape is lost in many particles and smaller capsid fragments are observed. These structures lack the interior dense material, i.e. the PIC has been released for integration into the host genome.

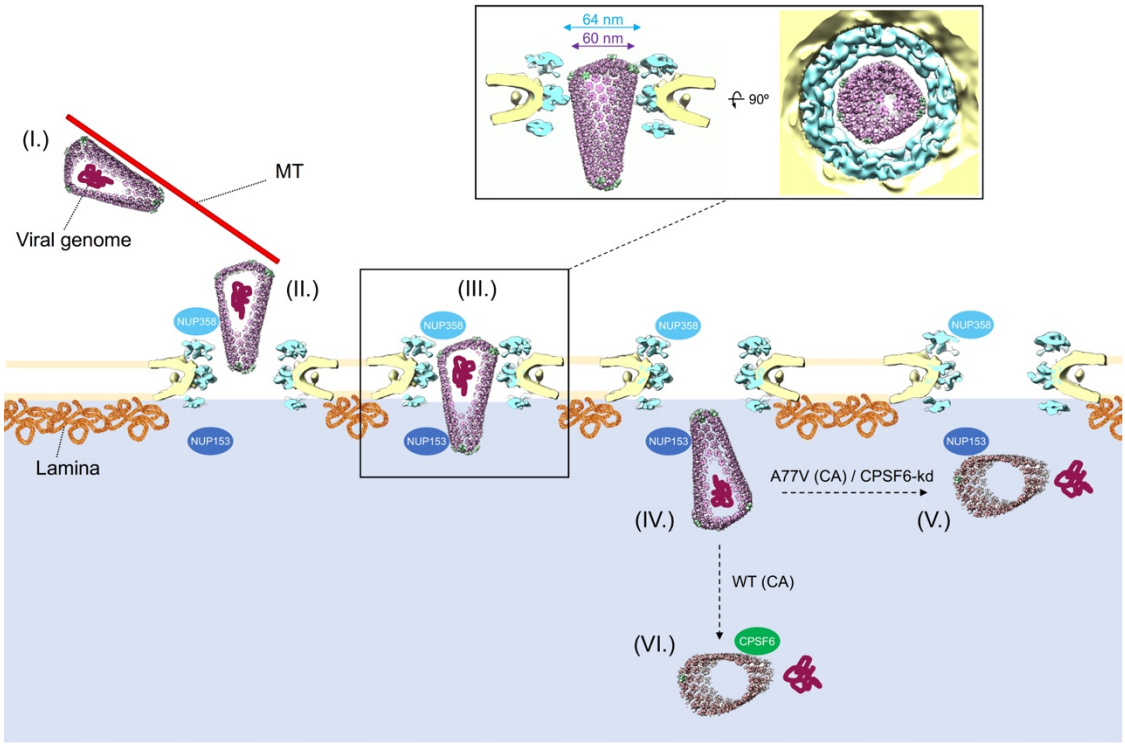


Figure 5.1. Conceptual model of HIV-1 nuclear import in T cells. In the cytosol, the intact cone-shaped HIV-1 capsid encasing the viral genome travels along microtubules (MT) towards the nuclear periphery (I.), where it encounters NUP358 bound to the cytoplasmic face of NPCs and docks to the NPC with the narrow CA pentamer-rich end (II.). Subsequently, the capsid penetrates into the central channel (III.). Superimposition of the mature HIV-1 capsid as determined from intact virions (Mattei et al., 2016) onto the *in cellulo* structural model of the NPC from infected cells (this study) reveals that the diameter of the NPC central channel is sufficiently wide for the transport of intact HIV-1 capsid. Side view (left) and top view (right) are shown as inset (III.). The intact HIV-1 capsid translocates into the nucleus (IV.). Although we cannot formally exclude that the CA lattice is perturbed during transport, all capsids observed inside of the central channel were cone-shaped. Upon departure from the NPC central channel, capsids encounter NUP153 and CPSF6. When CPSF6 binding is perturbed, the capsid uncoats and releases the genome at the NPC nuclear basket region (V.). When CPSF6 is available for interaction with the CA lattice, uncoating occurs deeper in the nucleoplasm (Burdick et al., 2020). In both cases (V. and VI.), the viral genome is released and integration into the host genome occurs close to the site of release. Figure from (Zila et al., 2021).

5.3 Improvement of the low-throughput workflow for the visualization of HIV-1 particles *in cellulo* and the lattice reconstruction

In my work, poorly characterized HIV-1 viral subcomplexes were identified in the crowded cellular environment with the aim to investigate capsid uncoating and nuclear entry. Despite the viral enrichment at the NPCs by A77V CA mutation or CPSF6 silencing, our system suffered from a very low-throughput with only ten viral particles identified in total in approximately 370 cryo-electron tomograms under both conditions. This clearly affected our SA results of the CA hexamer and the reconstruction of the full lattice. Although the visualization of HIV-1 lattice elements was never achieved before *in cellulo*, our results might be further optimized by increasing the number of particles included in the average. This would allow to achieve a better resolution of the CA hexamer in both cytosol and nucleus and consequently fully recover the lattice information at its status of assembly.
Here, several advances in cryo-FIB milling, data collection and cryo-CLEM are discussed in the contest of this study, as methods for improvement of my workflow.

5.3.1 Advances in cryo-FIB milling and data collection will enable further investigations

Persistent efforts were invested into the sample preparation by cryo-FIB milling and the acquisition of a large data set by cryo-ET. As a matter of fact, cryo-FIB milling and cryo-ET are challenging and time-consuming methods and the direct visualization of individual HIV-1 complexes *in cellulo* is further hampered by the infrequency of HIV-1 productive infection (Jun et al., 2011).

Technically, production of cryo-lamellae is a very repetitive task with a *per se* low-throughput when manually performed. Indeed, the high-vacuum environment where the cryo-milling occurs, only partially protects the sample from water condensation (Kuba et al., 2020). Over time during milling, a layer of residual water deposits on the sample

surface and alters sample thickness and vitrification. Size and thickness of the sample also influences the efficiency of FIB-milling, with prokaryotic and small eukaryotic cells that are milled faster than large mammalian cells. $CD4^+$ T lymphocytes used in my study represented an optimal compromise as they are small mammalian cells of ~10-15 µm in diameter (**Figure 4.4**). Nevertheless, the number of final cryo-lamellae with a thickness in the range of 150-200 nm was restricted to 5-8 per day.

Recently, automated sequential FIB-milling methods were developed (Buckley et al., 2020; Kuba et al., 2020; Tacke et al., 2020; Zachs et al., 2020) and proved to notably increase the efficiency of cryo-FIB milling for different type of specimens, reaching up to 27 cryo-lamellae with final thickness of 100-200 nm automatically milled within 16 hours (Tacke et al., 2020). This combined with the installation of a large cryo-shield and a cryo-shutter within the FIB-SEM microscope nearly abrogate the restriction of the FIB-milling to a certain time frame (Tacke et al., 2020).

The data collection is also a time-consuming process. At a Titan Krios microscope equipped with a Gatan K2 direct electron detector, typically ~45 minutes of acquisition are required per tomogram using a dose-symmetric scheme (Hagen et al., 2017). Only recent development of Gatan K3 cameras notably shortened the time of acquisition to approximately 15-20 minutes per tomogram.

Therefore, the combination of the automated cryo-FIB milling with the tomographic acquisition with a faster K3 detector might significantly strengthen my brute force approach by increasing the amount of data in a shorter time frame.

5.3.2 Cryo-CLEM can be used to precisely localize HIV-1 particles *in cellulo*

In comparison to my brute force approach, precise localization of viral particles by CLEM within high-pressure frozen SupT1-R5 T cells, allowed to visualize 16 and 19 viral particles carrying A77V- and WT-CA from few cell sections (Zila et al., 2021). This suggests that an implementation of my system with technological advances in correlative light microscopy (LM) with cryo-ET might also constitute a further improvement.

This approach already revealed to be successful for studying early stages of HIV-1 infection (Jun et al., 2011; Strauss et al., 2016). In one study, fluorescently labeled HIV-1 VLPs tethered to human cells were initially imaged in real time and then cryo-immobilized by plunge-freezing (Strauss et al., 2016). Fluorescent light micrographs were superimposed on selected regions of the cryo-EM montages to drive the acquisition of the area of interest. Cryo-ET was thus used to investigate a host inhibition mechanism which retain particles on the plasma membrane of cells. Although correlation of the fluorescence images in a live-cell series with cryo-EM micrograph was shown to be feasible, changes in the position of the fluorescent particle, cells or even fiducials markers, that occur due to dynamic cell process, blotting and plunge-freezing, might affect the analyses (Jun et al., 2011). Development of cryo-stages to include in confocal microscopes allowed to record fluorescence directly on vitrified samples (Arnold et al, 2016; Jun et al., 2011; Schorb and Briggs, 2014). Among various applications, cryo-CLEM was used to visualize HIV-1 cell attachment and viral entry in HeLa cells (Jun et al., 2011).

While both mentioned studies on HIV-1 entry were performed on thin areas of intact cells, such as filopodia and cortical regions of cells (Jun et al., 2011; Strauss et al., 2016), cryo-FIB milling is unavoidable for targeting HIV-1 nuclear entry through NPCs. In studies where the thickness of the sample exceeds 1 μm, cryo-FM is used to drive the targeted thinning of cell regions that contain fluorescence signal of interest by FIB-milling (Arnold et al., 2016). This step can be speeded up by automatic cryo-FIB milling of lamellae. Subsequent to cryo-FIB milling, the correlation is employed to navigate the data acquisition by cryo-ET within the FIB lamellae (Arnold et al., 2016). Correlation of LM and EM prior to cryo-FIB milling is challenged by the multiple transfers between microscopes, which increases the risk of ice-contamination or devitrification of the sample (Arnold et al., 2016; Rigort, 2010). Dual-purpose cryo-holders that can be used in both instruments were produced to circumvent additional transfer by moving the sample after cryo-LM imaging directly into the FIB-shuttle. Alternatively, photon ion electron microscopes that placed a light fluorescence beam inside the vacuum chamber of a cryo-FIB-SEM microscope were developed and shown to be functional for imaging and milling of different samples (Gorelick et al., 2019).

Another challenge is the high-accuracy targeting of lamella regions based on the cryo-LM images along the *z*-axis, such that the fluorescent targeted structure is contained within the resulting lamella. Nowadays, approaches based on the fiducial correlation are the most applied for achieve targeting accuracies (Arnold et al., 2016; Kukulski et al., 2011; Schorb and Sieckmann, 2017). In this approach, fiducial beads that are visible in both imaging modalities are added to the sample prior to plunge-freezing. At the FIB-SEM microscope, open sources software such as 3dcft (Arnold et al., 2016; https://3dct.semper.space) are used to calculate the image transformation based on the fitting of the center of the beads and thus accurately place the patters for milling.
As validating system of the correlation, a post-correlation on lamellae cryo-CLEM approach was developed to confirm that the targeted structure in the tomograms corresponds to the fluorescent signal (Klein et al., 2021). In this study, cryo-LMs of the lamellae were acquired prior and subsequently to the cryo-TEM imaging. The first cryo-LM map allows to detect the entire fluorescence within the whole cell, whereas the second is used to determine the tilt and position of the lamella. Deconvolution and 3D correlation of the two cryo-LM maps finally permit to exclusively extract the fluorescent signal within the lamellae.

In comparison to cryo-FIB milling automation and the acquisition of data at the Titan Krios with K3 cameras, cryo-CLEM might not necessarily accelerate my workflow but rather increase the number of capsids identified *in cellulo* by their precise localization and targeting for cryo-ET.

Conclusions

- Here, I optimized a workflow for cryo-FIB milling and cryo-ET of primary cells relevant in HIV-1 infection, such as $CD4^+$ T cells and monocyte-derived macrophages.

- More importantly for this study, I combined HIV-1 infection of SupT1-R5 T cells at the BSL-1 with cryo-FIB milling and cryo-ET. This allowed for the first time to visualize HIV-1 particles in the process of nuclear entry at their close-to-native state *in cellulo*.

- When CPSF6 is perturbed, the viral complexes retain cone-shaped HIV-1 capsids enclosing the viral genome in the cytosol and in close proximity to the nuclear envelope. Both shape and size of the capsids *in cellulo* highly resemble the capsids within mature virions *in situ*.

- Interestingly, cone-shape capsids dock to the NPC with their narrow ends perpendicularly oriented toward the NPC. The preferential orientation might be dependent on the high concentrations of CA pentamers in this area of the cone and might facilitate capsid penetration and translocation through the NPCs.

- Strikingly, the cone-shaped capsid still carrying the viral genome enters deeply the central channel of the NPC: while the broad end of the cone is still interacting with the cytoplasmic ring of NPC, the narrow tip already reached the nucleoplasm, possibly interacting with the nuclear basket of NPC.

- In line with the morphology, the lattice of the cone-shaped capsids docking to NPC and localized in the central channel of NPC was intact or nearly intact. In comparison to the study of the hexameric lattice *in situ*, our SA analyses *in cellulo* were affected by the low number of particles included in the average, and a lower SNR of cryo-electron tomograms due to cellular thickness and molecular crowding.

- Nevertheless, translocation of intact HIV-1 capsid through the NPC is geometrically possible: the central channel of NPCs in SupT1-R5 and primary $CD4^+$ T cells is dilated in comparison to the NPC from purified nuclear envelopes.

- The dilated conformation of NPCs in SupT1-R5 is a physiological condition of actively transporting NPCs and is independent of HIV-1 infection.

- Once translocated through the NPCs, HIV-1 capsids uncoat in the nucleus of T cells. In comparison to the cone-shape capsids in the cytosol, the remaining structures are empty tubes or fragments in the nucleus. Few lattice elements within these structures were observed. Therefore, I conceive uncoating as a process where capsids are opened to release the viral genome for its integration into the host chromosomes consists, rather than completely disassembled.

Significance of this study

HIV-1 uncoating is a crucial step in HIV-1 infection as followed by the release of the viral DNA and its integration into the human genome. Due to the irreversibility of this process, the genetic information of HIV-1 remains indelible inside the genome of the patient, making AIDS an incurable disease. It is therefore necessary to fully understand this process in order to potentially prevent it.

In this context, this thesis aimed to study HIV-1 nuclear entry and uncoating. Here, I provided the first evidence of HIV-1 hexameric lattices integrity prior to PICs nuclear entry. Moreover, I determined that capsid uncoating occurs in the nucleus of the infected cell. From a therapeutical perspective, this findings might be relevant for a better understanding of the mechanisms of actions of newly developed targeting-capsid drugs or possibly for the development of other therapeutics.

Chapter 6

Outlook

6.1 The manipulation of NPC diameter might control HIV-1 nuclear entry

In this study, I found that the diameter of NPC in T cells is wider than previously thought for the NPC in human cells. The same result was achieved in many other studies *in cellulo* for different organisms (Allegretti et al., 2020; Eibauer et al., 2015; Mahamid et al., 2016; Maimon et al., 2012; Mosalaganti et al., 2018). Additional variation in diameter between individual NPCs, reflecting NPC plasticity, was previously reported and related to specific physiological states (Mahamid et al., 2016).

It was recently shown that external mechanical forces transmitted to the nuclei that cause their deformation or flattening increase nuclear import of several transcription factors (Elosegui-Artola et al., 2017). This was proposed to happen due to an increased nuclear membrane curvature that in turn causes opening of the pores. In line with the hypothesis of Elosegui-Artola et al. (Elosegui-Artola et al., 2017), Christian Zimmerli et al. in our group recently showed that energy depletion causes constriction of NPC and block of active nuclear transport (Zimmerli et al., 2020). Nuclear size reduction and NE wrinkling connected to a relieve of membrane tension was proposed to be the cause of NPC constriction. Two major evidences became clear from this work: first, the constricted NPC structure *in situ* is related to a stress condition such a low energetic status of the cell; second, the diameter of NPC can be manipulated introducing mechanical stress on nuclei.

Based on these observations, it would be interesting to couple NPC constriction with HIV-1 nuclear entry and establish an on/off system to control the nuclear import of HIV-1 capsids. Here, I demonstrated that intact HIV-1 capsids are translocated through the

NPCs, and the nuclear import of such large cargoes is geometrically possible due to a dilated scaffold of the NPC central channel.

I therefore predict that constriction of NPC would result into a reduction of HIV-1 nuclear entry with a consequent accumulation of HIV-1 capsids at the cytoplasmic side of the NPC. In order to achieve this, one might apply different mechanical stresses that interfere with nuclear shape and membrane tension. Hypo-osmotic shock for example would represents a good candidate: it is simple and fast to perform *in vitro* and decreases flattening of nucleus and membrane tension (Elosegui-Artola et al., 2017).

This type of experiment might provide two different advantages: first, it would represent an alternative method to accumulate viral particles at NPCs and study the ultrastructure of the capsid lattices *in cellulo*, independently from CA mutations; second, it would provide further insights into the importance of the NEs and NPCs biophysical properties in a relevant contest for human health.

Acknowledgments

I would like to express my gratitude to all those people that supported me in the achievement of my doctoral thesis and in my daily life within the past four years.

First of all, I am profoundly grateful to Martin Beck for being such a great and supportive mentor and advisor, and a person who genuinely cares for the people in his lab. You have been a brilliant guidance and example of bright and kind scientist throughout my PhD, and offered me all the tools to grow scientifically and succeed into this experience.

I would like to acknowledge my thesis advisory committee members: Prof. Dr. Hans-Georg Kräusslich, Dr. Julia Mahamid, Dr. Marina Lusic, Dr. Yannick Schwab, for your precious support, advices and feedback. A special thank goes to Hans-Georg and Marina for the opportunity to collaborate and for the interesting discussions and inputs. To Julia, for providing assistance at the microscope and for giving me the chance to conclude my PhD at EMBL. I am also thankful to Dr. Petr Chlanda and PD Dr. Dr. Ada Cavalcanti-Adam for taking part in my PhD defence committee.

Thank you, Christian Zimmerli and Matteo Allegretti for taking me by hand into the field of cryo-EM, for answering my continuous questions and requests for help and for being my first smiling faces in the lab. Thank you, Vojtech Zila for sharing the joey and the pain of our project, for your support, suggestions and continuous discussions. I own you guys, Christian, Matteo, Vojtech a large part of my PhD success.

I would like to acknowledge all the present and former members of the Beck group: Bernhard, Janina, Beata, Iskander, Reiya, Max, Jenny, Patrick, Shyamal, Natalie, Henrik, Anja, Kasia, Marie. Especially, thank you Beata, for your contributions in the project, your efforts in teaching me subtomogram averaging and the proof-reading of my thesis. Bernhard for being my Deutsch summary translator but mostly for your sense of humor, the daily coffee breaks and the fun in the lab, for teaching me how to drive an automatic car and lose flights, sharing fever, happiness, sadness, and for your precious friendship. Thanks to Janina, for being one of the most enjoyable and fun person to work with and for being always available: the fact that you proof-read my thesis during your maternity

leave is a clear example of it. Thank you, Iskander for proof-reading my thesis, and for all the good chats and work breaks we had in the lab.

I would like to thank Ines de Castro for her positivity, suggestions and the fun I had working together.

I am extremely lucky to have met so many great people, who became not only my friends but truly my family in Heidelberg. Thank you, for always been there for me in any situation, for the crazy fun we had, the hikes, the delicious food, the discussions, the parties, the beers and wines, again the food (it was a lot!), the trips around the world, the richness you gave me simply from being from other countries, for filling so much my heart with joy every time. Huge thanks to Carol, Kristina, Ivan, Raj, Luca, Virginia, Veronika, Sara, Natalie, Chris, Alessandra, Cristina, Matteo, Bianca.

A special thank for the constant people in my life, that I always felt so close also in the distance. My ‘gipsies’ Silvia and Emma for sharing a huge history of memories, loughs, fun, support and for being always there in difficult times. Thank you, Chiara, for sharing with me so many experiences and an entire academic path, the fun, the videocalls, the support, for being such a strong person and a great listener, my ‘amichetta’ that, together with Francesco, has been my little piece of home that came along with me in Germany. Thanks to Ilaria, Alessandro, Maria Vittoria, my cousins Rachele and Michela, my aunt Giulia.

Last but not least, my family to whom I dedicate my thesis. To my brother, Gianluca, who I profoundly admire for his intelligence and curiosity, and unconsciously encouraged me in doing always better. My mom Maria and my dad Michele for being my biggest fan and the most supportive people in this world. I will be always grateful to you for supporting me through all my decisions. This means and will always mean a lot to me. Thank you for your teachings and love.

List of abbreviations

AAV : Adeno-associated virus

AIDS : Acquired immune deficiency syndrome

BSH-1 : Bovine herpesivirus type 1

BSL-1 : Biosafety level-1

CA : Capsid protein

CA^{CTD} : Capsid C-terminal domain

CA^{NTD} : Capsid N-terminal domain

CC : Cross correlation

CLEM : Correlative light and electron microscopy

CNC : Coat nucleoporin complex

CPSF6 : Cleavage and polyadenylation specificity factor

CR : Cytoplasmic ring

CRF : Circulating recombinant form

CTF : Contrast transfer function

CypA : Cyclophilin A

EM: Electron Microscopy

EMDB : Electron Microscopy Data Bank

Env : Envelope protein

ESCRT : Endosomal sorting complex required for transport

ET : Electron tomography

FG : Phenylalanine-glycine

FIB : Focused Ion Beam

FIB-SEM : Focused ion beam scanning electron microscope

FSC : Fourier shell correlation

GIS : Gas system injection

gRNA : genomic RNA

HIV : Human immunodeficiency virus

HPF : High-pressure freezing

HSV-1 : Herpes simplex virus

IN : Integrase

INM : Inner nuclear membrane

IR : Inner ring

KIF5B : Kinesin-1

LEDGF : Lens epithelium derived-grown factor

LM : Light microscopy

LTR : Long terminal repeat

MA : Matrix

MDM : Monocytes-derived macrophage

MT : Microtubule

MxB : Myxovirus resistance protein B

NC : Nucleocapsid

NE : Nuclear envelope

NES : Nuclear export signal

NLS : Nuclear localization signal

NPC : Nuclear pore complex

NR : Nuclear ring

NTF : Nuclear transport factor

Nup : Nucleoporin

ONM : Outer nuclear membrane

ORF : Open reading frame

PIC : Pre-integration complex

PM : Plasma membrane

PR : Protease

RanBD : Ran-binding domain

RanGAP : Ran GTPase-activating protein

RanGEF : Ran guanine nucleotide exchange factor

RER : Rough endoplasmic reticulum

RT : Reverse transcriptase

RTC : Reverse transcription complex

SA : Subtomogram averaging

SEM : Scanning electron microscope

SNR : Signal-to-noise ratio

SU : Subunit

TEM : Transmission electron microscope

TNPO3 : Transportin-3

TRIM : Tripartite motif-containing protein

TRN-1 : Transportin-1

URF : Unique recombinant form

VLP : Virus-like particle

VPP : Volta potential phase plate

References

Achuthan, V., Perreira, J.M., Sowd, G.A., Puray-Chavez, M., McDougall, W.M., Paulucci-Holthauzen, A., Wu, X., Fadel, H.J., Poeschla, E.M., Multani, A.S., et al. (2018). Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral DNA Integration. Cell Host & Microbe *24*, 392-404.e8.

Adam, S.A., Marr, R.S., and Gerace, L. (1990). Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. Journal of Cell Biology *111*, 807–816.

Al-Amoudi, A., Chang, J.-J., Leforestier, A., McDowall, A., Salamin, L.M., Norlén, L.P.O., Richter, K., Blanc, N.S., Studer, D., and Dubochet, J. (2004). Cryo-electron microscopy of vitreous sections. EMBO J *23*, 3583–3588.

Al-Amoudi, A., Studer, D., and Dubochet, J. (2005). Cutting artefacts and cutting process in vitreous sections for cryo-electron microscopy. J Struct Biol *150*, 109–121.

Allegretti, M., Zimmerli, C.E., Rantos, V., Wilfling, F., Ronchi, P., Fung, H.K.H., Lee, C.-W., Hagen, W., Turoňová, B., Karius, K., et al. (2020). In-cell architecture of the nuclear pore and snapshots of its turnover. Nature *586*, 796–800.

Ambrose, Z., and Aiken, C. (2014). HIV-1 uncoating: connection to nuclear entry and regulation by host proteins. Virology *454–455*, 371–379.

Anzinger, J.J., Olinger, G.G., and Spear, G.T. (2008). Donor variability in HIV binding to peripheral blood mononuclear cells. Virol J *5*, 95.

von Appen, A., Kosinski, J., Sparks, L., Ori, A., DiGuilio, A.L., Vollmer, B., Mackmull, M.-T., Banterle, N., Parca, L., Kastritis, P., et al. (2015). In situ structural analysis of the human nuclear pore complex. Nature *526*, 140–143.

Arhel, N., Genovesio, A., Kim, K.-A., Miko, S., Perret, E., Olivo-Marin, J.-C., Shorte, S., and Charneau, P. (2006). Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. Nat. Methods *3*, 817–824.

Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prévost, M.-C., Allen, T.D., and Charneau, P. (2007). HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. EMBO J *26*, 3025–3037.

Arnold, J. (2016). Site-Specific Cryo-focused Ion Beam Sample Preparation Guided by 3D Correlative Microscopy. Biophysical Journal 10.

Bartesaghi, A., Sprechmann, P., Liu, J., Randall, G., Sapiro, G., and Subramaniam, S. (2008). Classification and 3D averaging with missing wedge correction in biological electron tomography☆. Journal of Structural Biology *162*, 436–450.

Beck, M., and Baumeister, W. (2016). Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail? Trends in Cell Biology *26*, 825–837.

Beck, M., and Hurt, E. (2017). The nuclear pore complex: understanding its function through structural insight. Nat Rev Mol Cell Biol *18*, 73–89.

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

Bejarano, D.A., Puertas, M.C., Börner, K., Martinez-Picado, J., Müller, B., and Kräusslich, H.-G. (2018). Detailed Characterization of Early HIV-1 Replication Dynamics in Primary Human Macrophages. Viruses *10*.

Bejarano, D.A., Peng, K., Laketa, V., Börner, K., Jost, K.L., Lucic, B., Glass, B., Lusic, M., Müller, B., and Kräusslich, H.-G. (2019). HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. ELife *8*.

Ben-Harush, K., Maimon, T., Patla, I., Villa, E., and Medalia, O. (2010). Visualizing cellular processes at the molecular level by cryo-electron tomography. Journal of Cell Science *123*, 7–12.

Benjamin, J., Ganser-Pornillos, B.K., Tivol, W.F., Sundquist, W.I., and Jensen, G.J. (2005). Three-dimensional Structure of HIV-1 Virus-like Particles by Electron Cryotomography. Journal of Molecular Biology *346*, 577–588.

Berger, E.A., Doms, R.W., Fenyö, E.-M., Korber, B.T.M., Littman, D.R., Moore, J.P., Sattentau, Q.J., Schuitemaker, H., Sodroski, J., and Weiss, R.A. (1998). A new classification for HIV-1. Nature *391*, 240–240.

Bharat, T.A.M., Noda, T., Riches, J.D., Kraehling, V., Kolesnikova, L., Becker, S., Kawaoka, Y., and Briggs, J.A.G. (2012). Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. Proceedings of the National Academy of Sciences *109*, 4275–4280.

Bharat, T.A.M., Kureisaite-Ciziene, D., Hardy, G.G., Yu, E.W., Devant, J.M., Hagen, W.J.H., Brun, Y.V., Briggs, J.A.G., and Löwe, J. (2017). Structure of the hexagonal surface layer on Caulobacter crescentus cells. Nat Microbiol *2*, 17059.

Blissenbach, M., Grewe, B., Hoffmann, B., and Brandt, S. (2010). Nuclear RNA Export and Packaging Functions of HIV-1 Rev Revisited□. J. VIROL. *84*, 7.

Bossolt, K.L., Franke, E.K., and Gofft, P. (1993). Human lmmunodeficiency Virus Type 1 Gag Protein Binds to Cyclophilins A and B. 12.

Braaten, D., and Luban, J. (2001). Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells. EMBO J *20*, 1300–1309.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., and Elledge, S.J. (2008). Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. *319*, 7.

Briggs, J.A.G., Simon, M.N., Gross, I., Kräusslich, H.-G., Fuller, S.D., Vogt, V.M., and Johnson, M.C. (2004). The stoichiometry of Gag protein in HIV-1. Nat Struct Mol Biol *11*, 672–675.

Briggs, J.A.G., Riches, J.D., Glass, B., Bartonova, V., and Zanetti, G. (2009). Structure and assembly of immature HIV. COMPUTATIONAL BIOLOGY 6.

Buckley, G., Gervinskas, G., Taveneau, C., Venugopal, H., Whisstock, J.C., and de Marco, A. (2020). Automated cryo-lamella preparation for high-throughput in-situ structural biology. Journal of Structural Biology *210*, 107488.

Bui, K.H., von Appen, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M.-T., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., et al. (2013). Integrated structural analysis of the human nuclear pore complex scaffold. Cell *155*, 1233–1243.

Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T.L., Bukrinskaya, A.G., Haggerty, S., and Stevenson, M. (1992). Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. Proceedings of the National Academy of Sciences *89*, 6580–6584.

Burdick, R.C., Delviks-Frankenberry, K.A., Chen, J., Janaka, S.K., Sastri, J., Hu, W.-S., and Pathak, V.K. (2017). Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. PLoS Pathog. *13*, e1006570.

Burdick, R.C., Li, C., Munshi, M., Rawson, J.M.O., Nagashima, K., Hu, W.-S., and Pathak, V.K. (2020). HIV-1 uncoats in the nucleus near sites of integration. Proc. Natl. Acad. Sci. U.S.A. *117*, 5486–5493.

Burns, L.T., and Wente, S.R. (2014). From Hypothesis to Mechanism: Uncovering Nuclear Pore Complex Links to Gene Expression. Molecular and Cellular Biology *34*, 2114–2120.

Byeon, I.-J.L., Meng, X., Jung, J., Zhao, G., Yang, R., Shi, J., Concel, J., Aiken, C., and Zhang, P. (2010). Structural Convergence between CryoEM and NMR Reveals Novel Intersubunit Interactions Critical for HIV-1 Capsid Function. 21.

Campbell, E.M., and Hope, T.J. (2015). HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nat Rev Microbiol *13*, 471–483.

Carlson, L.-A., Briggs, J.A.G., Glass, B., Riches, J.D., Simon, M.N., Johnson, M.C., Müller, B., Grünewald, K., and Kräusslich, H.-G. (2008). Three-Dimensional Analysis of Budding Sites and Released Virus Suggests a Revised Model for HIV-1 Morphogenesis. 8.

Carnes, S.K., Sheehan, J.H., and Aiken, C. (2018). Inhibitors of the HIV-1 capsid, a target of opportunity. Current Opinion in HIV and AIDS *13*, 359–365.

Cassany, A., and Gerace, L. (2008). Reconstitution of Nuclear Import in Permeabilized Cells. In The Nucleus, R. Hancock, ed. (Totowa, NJ: Humana Press), pp. 181–205.

Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. (1997). Core Structure of gp41 from the HIV Envelope Glycoprotein. Cell *89*, 263–273.

Charneau, P. (1991). A Single-Stranded Gap in Human Immunodeficiency Virus Unintegrated Linear DNA Defined by a Central Copy of the Polypurine Tract. J. VIROL. *65*, 7.

Charneau, P., and Alizon, M. (1992). A Second Origin of DNA Plus-Strand Synthesis Is Required for Optimal Human Immunodeficiency Virus Replication. J. VIROL. *66*, 7.

Checkley, M.A., Luttge, B.G., and Freed, E.O. (2011). HIV-1 Envelope Glycoprotein Biosynthesis, Trafficking, and Incorporation. Journal of Molecular Biology *410*, 582–608.

Chen, B. (2019). Molecular Mechanism of HIV-1 Entry. Trends Microbiol. *27*, 878–891.

Chen, N.-Y., Zhou, L., Gane, P.J., Opp, S., Ball, N.J., Nicastro, G., Zufferey, M., Buffone, C., Luban, J., Selwood, D., et al. (2016). HIV-1 capsid is involved in post-nuclear entry steps. Retrovirology *13*.

Chin, C.R., Perreira, J.M., Savidis, G., Portmann, J.M., Aker, A.M., Feeley, E.M., Smith, M.C., and Brass, A.L. (2015). Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. Cell Reports *13*, 1717–1731.

Chojnacki, J., and Eggeling, C. (2018). Super-resolution fluorescence microscopy studies of human immunodeficiency virus. Retrovirology *15*, 41.

Christensen, D.E., Ganser-Pornillos, B.K., Johnson, J.S., Pornillos, O., and Sundquist, W.I. (2020). Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. Science *370*.

Clever, J.L., and Parslow, T.G. (1997). Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. Journal of Virology *71*, 3407–3414.

Coffin, J. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science *267*, 483–489.

Cohen, S., Etingov, I., and Panté, N. (2012). Effect of Viral Infection on the Nuclear Envelope and Nuclear Pore Complex. In International Review of Cell and Molecular Biology, (Elsevier), pp. 117–159.

Cook, A., Bono, F., Jinek, M., and Conti, E. (2007). Structural Biology of Nucleocytoplasmic Transport. 28.

Craigie, R., and Bushman, F.D. (2012). HIV DNA Integration. Cold Spring Harbor Perspectives in Medicine *2*, a006890–a006890.

Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., and Matunis, M.J. (2002). Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol *158*, 915–927.

Cyrklaff, M., Linaroudis, A., Boicu, M., Chlanda, P., Baumeister, W., Griffiths, G., and Krijnse-Locker, J. (2007). Whole Cell Cryo-Electron Tomography Reveals Distinct Disassembly Intermediates of Vaccinia Virus. PLoS ONE *2*, e420.

Danev, R., Tegunov, D., and Baumeister, W. (2017). Using the Volta phase plate with defocus for cryo-EM single particle analysis. ELife *6*, e23006.

Delphin, C., Guan, T., Melchior, F., and Gerace, L. (1997). RanGTP Targets p97 to RanBP2, a Filamentous Protein Localized at the Cytoplasmic Periphery of the Nuclear Pore Complex. MBoC *8*, 2379–2390.

Dharan, A., Talley, S., Tripathi, A., Mamede, J.I., Majetschak, M., Hope, T.J., and Campbell, E.M. (2016). KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. PLoS Pathog. *12*, e1005700.

Dharan, A., Bachmann, N., Talley, S., Zwikelmaier, V., and Campbell, E.M. (2020). Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nat Microbiol.

Di Nunzio, F., Danckaert, A., Fricke, T., Perez, P., Fernandez, J., Perret, E., Roux, P., Shorte, S., Charneau, P., Diaz-Griffero, F., et al. (2012). Human Nucleoporins Promote HIV-1 Docking at the Nuclear Pore, Nuclear Import and Integration. PLoS ONE *7*, e46037.

Dubochet, J. (2007). The Physics of Rapid Cooling and Its Implications for Cryoimmobilization of Cells. In Methods in Cell Biology, (Elsevier), pp. 7–21.

Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A.W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. Quart. Rev. Biophys. *21*, 129–228.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R., and Daviest, D.R. (1994). Crystal Structure of the Catalytic Domain of HIV-1 Integras: Similarity to Other Polynucleotidyl Transferases. *266*, 7.

Egan, M.A., Carruth, L.M., Rowell, J.F., Yu, X., and Siliciano, R.F. (1996). Human immunodeficiency virus type 1 envelope protein endocytosis mediated by a highly conserved intrinsic internalization signal in the cytoplasmic domain of gp41 is suppressed in the presence of the Pr55gag precursor protein. Journal of Virology *70*, 6547–6556.

Eibauer, M., Pellanda, M., Turgay, Y., Dubrovsky, A., Wild, A., and Medalia, O. (2015). Structure and gating of the nuclear pore complex. Nat Commun *6*, 7532.

Eijkelenboom, A.P.A.M., van den Ent, F.M.I., Vos, A., Doreleijers, J.F., Hård, K., Tullius, T.D., Plasterk, R.H.A., Kaptein, R., and Boelens, R. (1997). The solution structure of the amino-terminal HHCC domain of HIV-2 integrase: a three-helix bundle stabilized by zinc. Current Biology *7*, 739–746.

Elosegui-Artola, A., Andreu, I., Beedle, A.E.M., Lezamiz, A., Uroz, M., Kosmalska, A.J., Oria, R., Kechagia, J.Z., Rico-Lastres, P., Le Roux, A.-L., et al. (2017). Force Triggers YAP Nuclear Entry by Regulating Transport across Nuclear Pores. Cell *171*, 1397-1410.e14.

Fassati, A., and Goff, S.P. (2001). Characterization of Intracellular Reverse Transcription Complexes of Human Immunodeficiency Virus Type 1. J. Virol. *75*, 3626–3635.

Fernandez, J., Machado, A.K., Lyonnais, S., Chamontin, C., Gärtner, K., Léger, T., Henriquet, C., Garcia, C., Portilho, D.M., Pugnière, M., et al. (2019). Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. Nat Microbiol *4*, 1840–1850.

Förster, F., and Hegerl, R. (2007). Structure Determination In Situ by Averaging of Tomograms. In Methods in Cell Biology, (Elsevier), pp. 741–767.

Forster, F., Medalia, O., Zauberman, N., Baumeister, W., and Fass, D. (2005). Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. Proceedings of the National Academy of Sciences *102*, 4729–4734.

Francis, A.C., and Melikyan, G.B. (2018). Single HIV-1 Imaging Reveals Progression of Infection through CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. Cell Host & Microbe *23*, 536-548.e6.

Francis, A.C., Marin, M., Singh, P.K., Achuthan, V., Prellberg, M.J., Palermino-Rowland, K., Lan, S., Tedbury, P.R., Sarafianos, S.G., Engelman, A.N., et al. (2020). HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. Nat Commun *11*, 3505.

Franke, E.K., and Yuan, H.E.H. (1994). Specific incorporation of cyclophilin A into HIV-1 virions. *372*, 4.

Frankel, A.D., and Young, J.A.T. (1998). HIV-1: Fifteen Proteins and an RNA. Annu. Rev. Biochem. *67*, 1–25.

Freed, E.O., and Martin, M.A. (1995). The Role of Human Immunodeficiency Virus Type 1 Envelope Glycoproteins in Virus Infection. Journal of Biological Chemistry *270*, 23883–23886.

Friedrich, M., Setz, C., Hahn, F., Matthaei, A., Fraedrich, K., Rauch, P., Henklein, P., Traxdorf, M., Fossen, T., and Schubert, U. (2016). Glutamic Acid Residues in HIV-1 p6 Regulate Virus Budding and Membrane Association of Gag. Viruses *8*, 117.

Fun, A., Wensing, A.M., Verheyen, J., and Nijhuis, M. (2012). Human Immunodeficiency Virus gag and protease: partners in resistance. Retrovirology *9*, 63.

Gall, J.G. (1967). OCTAGONAL NUCLEAR PORES. Journal of Cell Biology *32*, 391–399.

Gamble, T.R. (1997). Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein. Science *278*, 849–853.

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

Ganser-Pornillos, B.K., Cheng, A., and Yeager, M. (2007). Structure of Full-Length HIV-1 CA: A Model for the Mature Capsid Lattice. Cell *131*, 70–79.

Ganser-Pornillos, B.K., Yeager, M., and Pornillos, O. (2019). Assembly and Architecture of HIV. 34.

Garcia, J.A., Wu, F.K., Mitsuyasu, R., and Gaynor, R.B. (1987). Interactions of cellular proteins involved in the transcriptional regulation of the human immunodeficiency virus. 10.

Glaeser, R.M. (1971). Limitations to significant information in biological electron microscopy as a result of radiation damage. Journal of Ultrastructure Research *36*, 466–482.

Glaeser, R.M. (2016). Specimen Behavior in the Electron Beam. In Methods in Enzymology, (Elsevier), pp. 19–50.

Gorelick, S., Buckley, G., Gervinskas, G., Johnson, T.K., Handley, A., Caggiano, M.P., Whisstock, J.C., Pocock, R., and de Marco, A. (2019). PIE-scope, integrated cryo-correlative light and FIB/SEM microscopy. 15.

Görlich, D. (1999). Transport Between the Cell Nucleus and the Cytoplasm. 58.

Görlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R.A., Hartmann, E., and Prehn, S. (1995). Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. Curr Biol *5*, 383–392.

Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. Elife *4*, e06980.

Grimm, R., Typke, D., Bärmann, M., and Baumeister, W. (1996). Determination of the inelastic mean free path in ice by examination of tilted vesicles and automated most probable loss imaging. Ultramicroscopy *63*, 169–179.

Gulnik, S., Erickson, J.W., and Xie, D. (2000). HIV protease: Enzyme function and drug resistance. In Vitamins & Hormones, (Elsevier), pp. 213–256.

Guntaka, R.V. (1993). Transcription Termination and Polyadenylation in Retroviruses. MICROBIOL. REV. *57*, 11.

Hagen, W.J.H., Wan, W., and Briggs, J.A.G. (2017). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. Journal of Structural Biology *197*, 191–198.

Haller, O. (2015). Mx GTPases: dynamin-like antiviral machines of innate immunity. *23*, 10.

Hampoelz, B., Andres-Pons, A., Kastritis, P., and Beck, M. (2019). Structure and Assembly of the Nuclear Pore Complex. 24.

Harapin, J., Börmel, M., Sapra, K.T., Brunner, D., Kaech, A., and Medalia, O. (2015). Structural analysis of multicellular organisms with cryo-electron tomography. 5.

Hare, S., Gupta, S.S., Valkov, E., Engelman, A., and Cherepanov, P. (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. Nature *464*, 232–236.

Hase, M.E., Kuznetsov, N.V., and Cordes, V.C. (2001). Amino acid substitutions of coiled-coil protein Tpr abrogate anchorage to the nuclear pore complex but not parallel, in-register homodimerization. Mol Biol Cell *12*, 2433–2452.

Hatziioannou, T., Perez-Caballero, D., Cowan, S., and Bieniasz, P.D. (2005). Cyclophilin Interactions with Incoming Human Immunodeficiency Virus Type 1 Capsids with Opposing Effects on Infectivity in Human Cells. JVI *79*, 176–183.

Hayles, M.F., Stokes, D.J., Phifer, D., and Findlay, K.C. (2007). A technique for improved focused ion beam milling of cryo-prepared life science specimens. J Microsc *226*, 263–269.

Hemelaar, J., Gouws, E., Ghys, P.D., and Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS *20*, W13–W23.

Henderson, H.I., and Hope, T.J. (2006). The temperature arrested intermediate of virus-cell fusion is a functional step in HIV infection. Virol. J. *3*, 36.

Hilditch, L., and Towers, G.J. (2014). A model for cofactor use during HIV-1 reverse transcription and nuclear entry. Current Opinion in Virology *4*, 32–36.

Hu, G.-B. (2014). Whole Cell Cryo-Electron Tomography Suggests Mitochondria Divide by Budding. Microsc Microanal *20*, 1180–1187.

Hu, W.-S., and Hughes, S.H. (2012). HIV-1 Reverse Transcription. Cold Spring Harbor Perspectives in Medicine *2*, a006882–a006882.

Hulme, A.E., Kelley, Z., Foley, D., and Hope, T.J. (2015). Complementary Assays Reveal a Low Level of CA Associated with Viral Complexes in the Nuclei of HIV-1-Infected Cells. Journal of Virology *89*, 5350–5361.

Hutchings, J., Stancheva, V., Miller, E.A., and Zanetti, G. (2018). Subtomogram averaging of COPII assemblies reveals how coat organization dictates membrane shape. Nat Commun *9*, 4154.

Iordanskiy, S., Berro, R., Altieri, M., Kashanchi, F., and Bukrinsky, M. (2006). Intracytoplasmic maturation of the human immunodeficiency virus type 1 reverse transcription complexes determines their capacity to integrate into chromatin. Retrovirology *3*, 4.

Jacobo-Molina, A., Ding, J., Nanni, R.G., Clark, A.D., Lu, X., Tantillo, C., Williams, R.L., Kamer, G., Ferris, A.L., and Clark, P. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proceedings of the National Academy of Sciences *90*, 6320–6324.

Jacques, D.A., McEwan, W.A., Hilditch, L., Price, A.J., Towers, G.J., and James, L.C. (2016). HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. Nature *536*, 349–353.

Jetzt, A.E., Yu, H., Klarmann, G.J., Ron, Y., Preston, B.D., and Dougherty, J.P. (2000). High Rate of Recombination throughout the Human Immunodeficiency Virus Type 1 Genome. J. Virol. *74*, 1234–1240.

Jun, S., Ke, D., Debiec, K., Zhao, G., Meng, X., Ambrose, Z., Gibson, G.A., Watkins, S.C., and Zhang, P. (2011). Direct Visualization of HIV-1 with Correlative Live-Cell Microscopy and Cryo-Electron Tomography. Structure *19*, 1573–1581.

Kim, K., Dauphin, A., Komurlu, S., McCauley, S.M., Yurkovetskiy, L., Carbone, C., Diehl, W.E., Strambio-De-Castillia, C., Campbell, E.M., and Luban, J. (2019). Cyclophilin A protects HIV-1 from restriction by human TRIM5α. Nat Microbiol *4*, 2044–2051.

Kinoshita, S., Chen, B.K., Kaneshima, H., and Nolan, G.P. (1998). Host Control of HIV-1 Parasitism in T Cells by the Nuclear Factor of Activated T Cells. Cell *95*, 595–604.

Klein, S., Wimmer, B.H., Winter, S.L., Kolovou, A., Laketa, V., and Chlanda, P. (2021). Post-correlation on-lamella cryo-CLEM reveals the membrane architecture of lamellar bodies. Commun Biol *4*, 137.

Köhler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol *8*, 761–773.

Kohlstaedt, L., Wang, J., Friedman, J., Rice, P., and Steitz, T. (1992). Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science *256*, 1783–1790.

Kondo, E., and Göttlinger, H.G. (1996). A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr. Journal of Virology *70*, 159–164.

König, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M.C., Irelan, J.T., Chiang, C., Tu, B.P., De Jesus, P.D., Lilley, C.E., et al. (2008). Global Analysis of Host-Pathogen Interactions that Regulate Early-Stage HIV-1 Replication. Cell *135*, 49–60.

Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A.L., Wan, W., Bui, K.H., Hagen, W.J.H., Briggs, J.A.G., Glavy, J.S., et al. (2016). Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science *352*, 363–365.

Kovtun, O., Leneva, N., Bykov, Y.S., Ariotti, N., Teasdale, R.D., Schaffer, M., Engel, B.D., Owen, David.J., Briggs, J.A.G., and Collins, B.M. (2018). Structure of the membrane-assembled retromer coat determined by cryo-electron tomography. Nature *561*, 561–564.

Kräusslich, H.G., Fäcke, M., Heuser, A.M., Konvalinka, J., and 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, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. *116*, 71–76.

Krull, S., Dörries, J., Boysen, B., Reidenbach, S., Magnius, L., Norder, H., Thyberg, J., and Cordes, V.C. (2010). Protein Tpr is required for establishing nuclear pore-associated zones of heterochromatin exclusion. EMBO J *29*, 1659–1673.

Kuba, J., Mitchels, J., Hovorka, M., Erdmann, P., and Berka, L. (2020). Advanced cryo-tomography workflow developments – correlative microscopy, milling automation and cryo-lift-out. 14.

Kukulski, W., Schorb, M., Welsch, S., Picco, A., Kaksonen, M., and Briggs, J.A.G. (2011). Correlated fluorescence and 3D electron microscopy with high sensitivity and spatial precision. Journal of Cell Biology *192*, 111–119.

Lau, D., Walsh, J.C., Mousapasandi, A., Ariotti, N., Shah, V.B., Turville, S., Jacques, D.A., and Böcking, T. (2020). Self-Assembly of Fluorescent HIV Capsid Spheres for Detection of Capsid Binders. Langmuir *36*, 3624–3632.

Leuzinger, H., Ziegler, U., Schraner, E.M., Fraefel, C., Glauser, D.L., Heid, I., Ackermann, M., Mueller, M., and Wild, P. (2005). Herpes Simplex Virus 1 Envelopment Follows Two Diverse Pathways. J. VIROL. *79*, 13.

Lin, D.H., and Hoelz, A. (2019). The Structure of the Nuclear Pore Complex (An Update). 62.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., et al. (2016). Architecture of the symmetric core of the nuclear pore. STRUCTURAL BIOLOGY 20.

Link, J.O., Rhee, M.S., Tse, W.C., Zheng, J., Somoza, J.R., Rowe, W., Begley, R., Chiu, A., Mulato, A., Hansen, D., et al. (2020). Clinical targeting of HIV capsid protein with a long-acting small molecule. Nature.

Liu, C., Perilla, J.R., Ning, J., Lu, M., Hou, G., Ramalho, R., Himes, B.A., Zhao, G., Bedwell, G.J., Byeon, I.-J., et al. (2016). Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. Nat Commun *7*, 10714.

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

Luo, X., Yang, W., and Gao, G. (2018). SUN1 Regulates HIV-1 Nuclear Import in a Manner Dependent on the Interaction between the Viral Capsid and Cellular Cyclophilin A. J Virol *92*, e00229-18.

Lusic, M., and Siliciano, R.F. (2017). Nuclear landscape of HIV-1 infection and integration. Nat Rev Microbiol *15*, 69–82.

Mackmull, M., Klaus, B., Heinze, I., Chokkalingam, M., Beyer, A., Russell, R.B., Ori, A., and Beck, M. (2017). Landscape of nuclear transport receptor cargo specificity. Mol Syst Biol *13*, 962.

Mahamid, J., Pfeffer, S., Schaffer, M., Villa, E., Danev, R., Kuhn Cuellar, L., Forster, F., Hyman, A.A., Plitzko, J.M., and Baumeister, W. (2016). Visualizing the molecular sociology at the HeLa cell nuclear periphery. Science *351*, 969–972.

Maimon, T., Elad, N., Dahan, I., and Medalia, O. (2012). The Human Nuclear Pore Complex as Revealed by Cryo-Electron Tomography. 1.

Malikov, V., da Silva, E.S., Jovasevic, V., Bennett, G., de Souza Aranha Vieira, D.A., Schulte, B., Diaz-Griffero, F., Walsh, D., and Naghavi, M.H. (2015). HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. Nat Commun *6*, 6660.

Mamede, J.I., Cianci, G.C., Anderson, M.R., and Hope, T.J. (2017). Early cytoplasmic uncoating is associated with infectivity of HIV-1. Proceedings of the National Academy of Sciences *114*, E7169–E7178.

Mammano, F., Kondo, E., Sodroski, J., Bukovsky, A., and Göttlinger, H.G. (1995). Rescue of human immunodeficiency virus type 1 matrix protein mutants by envelope glycoproteins with short cytoplasmic domains. Journal of Virology *69*, 3824–3830.

Mansky, L.M., and Temin, H.M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of Virology *69*, 5087–5094.

Marko, M., Hsieh, C., Schalek, R., Frank, J., and Mannella, C. (2007). Focused-ion-beam thinning of frozen- hydrated biological specimens for cryo- electron microscopy. 3.

Marko, M., Hsieh, C., Moberlychan, W., Mannella, C., and Frank, J. (2006). Focused ion beam milling of vitreous water: prospects for an alternative to cryo-ultramicrotomy of frozen-hydrated biological samples. Journal of Microscopy 6.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

Matreyek, K., and Engelman, A. (2013). Viral and Cellular Requirements for the Nuclear Entry of Retroviral Preintegration Nucleoprotein Complexes. Viruses *5*, 2483–2511.

Matreyek, K.A., Yücel, S.S., Li, X., and Engelman, A. (2013). Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. PLoS Pathog. *9*, e1003693.

Matsuura, Y., and Stewart, M. (2005). Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. EMBO J *24*, 3681–3689.

Mattaj, I.W., and Englmeier, L. (1998). NUCLEOCYTOPLASMIC TRANSPORT: The Soluble Phase. Annu. Rev. Biochem. *67*, 265–306.

Mattei, S., Glass, B., Hagen, W.J.H., Kräusslich, H.-G., and Briggs, J.A.G. (2016). The structure and flexibility of conical HIV-1 capsids determined within intact virions. Science *354*, 1434–1437.

Matteson, T.L., Schwarz, S.W., Houge, E.C., Kempshall, B.W., and Giannuzzi, L.A. (2002). Electron backscattering diffraction investigation of focused ion beam surfaces. Journal of Elec Materi *31*, 33–39.

Melikyan, G.B., Markosyan, R.M., Hemmati, H., Delmedico, M.K., Lambert, D.M., and Cohen, F.S. (2000). Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. J. Cell Biol. *151*, 413–423.

Meryman, H.T. (2007). Cryopreservation of living cells: principles and practice. Transfusion *47*, 935–945.

Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G.B. (2009). HIV Enters Cells via Endocytosis and Dynamin-Dependent Fusion with Endosomes. Cell *137*, 433–444.

Monette, A., Panté, N., and Mouland, A.J. (2011). HIV-1 remodels the nuclear pore complex. Journal of Cell Biology *193*, 619–631.

Mosalaganti, S., Kosinski, J., Albert, S., Schaffer, M., Strenkert, D., Salomé, P.A., Merchant, S.S., Plitzko, J.M., Baumeister, W., Engel, B.D., et al. (2018). In situ architecture of the algal nuclear pore complex. Nat Commun *9*, 2361.

Müller, T.G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., Müller, B., and Kräusslich, H.-G. (2020). HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells (Microbiology).

Nakielny, S., and Dreyfuss, G. (1999). Transport of proteins and RNAs in and out of the nucleus. Cell *99*, 677–690.

Navia, M.A., and Mckeever, B.M. (1990). A Role for the Aspartyl Protease from the Human Immunodeficiency Virus Type 1 (HIV-1) in the Orchestration of Virus Assembly. Ann NY Acad Sci *616*, 73–85.

Nermut, M.V., and Fassati, A. (2003). Structural analyses of purified human immunodeficiency virus type 1 intracellular reverse transcription complexes. J Virol *77*, 8196–8206.

Ni, T., Gerard, S., Zhao, G., Dent, K., Ning, J., Zhou, J., Shi, J., Anderson-Daniels, J., Li, W., Jang, S., et al. (2020). Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. Nat Struct Mol Biol *27*, 855–862.

Novikova, M., Zhang, Y., Freed, E.O., and Peng, K. (2019). Multiple Roles of HIV-1 Capsid during the Virus Replication Cycle. Virol Sin *34*, 119–134.

Nunzio, F.D. (2013). Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. 11.

Obr, M., and Schur, F.K.M. (2019). Structural analysis of pleomorphic and asymmetric viruses using cryo-electron tomography and subtomogram averaging. In Advances in Virus Research, (Elsevier), pp. 117–159.

Orlova, E.V., and Saibil, H.R. (2011). Structural Analysis of Macromolecular Assemblies by Electron Microscopy. Chemical Reviews 39.

Partridge, J.R., and Schwartz, T.U. (2009). Crystallographic and biochemical analysis of the Ran-binding zinc finger domain. J Mol Biol *391*, 375–389.

Patla, I., Volberg, T., Elad, N., Hirschfeld-Warneken, V., Grashoff, C., Fässler, R., Spatz, J.P., Geiger, B., and Medalia, O. (2010). Dissecting the molecular architecture of integrin adhesion sites by cryo-electron tomography. Nat Cell Biol *12*, 909–915.

Peng, K., Muranyi, W., Glass, B., Laketa, V., Yant, S.R., Tsai, L., Cihlar, T., Müller, B., and Kräusslich, H.-G. (2014). Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. ELife *3*.

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

Pettit, S.C., Everitt, L.E., Choudhury, S., Dunn, B.M., and Kaplan, A.H. (2004). Initial Cleavage of the Human Immunodeficiency Virus Type 1 GagPol Precursor by Its Activated Protease Occurs by an Intramolecular Mechanism. JVI *78*, 8477–8485.

Pizzato, M., Erlwein, O., Bonsall, D., Kaye, S., Muir, D., and McClure, M.O. (2009). A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. J. Virol. Methods *156*, 1–7.

Pornillos, O., Ganser-Pornillos, B.K., and Yeager, M. (2011). Atomic-level modelling of the HIV capsid. Nature *469*, 424–427.

Pornillos, O., Ganser-Pornillos, B.K., Kelly, B.N., Hua, Y., Whitby, F.G., Stout, C.D., Sundquist, W.I., Hill, C.P., and Yeager, M. (2009). X-Ray Structures of the Hexameric Building Block of the HIV Capsid. 11.

Preston, B., Poiesz, B., and Loeb, L. (1988). Fidelity of HIV-1 reverse transcriptase. Science *242*, 1168–1171.

Price, A.J., Fletcher, A.J., Schaller, T., Elliott, T., Lee, K., KewalRamani, V.N., Chin, J.W., Towers, G.J., and James, L.C. (2012). CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication. PLoS Pathog *8*, e1002896.

Price, A.J., Jacques, D.A., McEwan, W.A., Fletcher, A.J., Essig, S., Chin, J.W., Halambage, U.D., Aiken, C., and James, L.C. (2014). Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. PLoS Pathog. *10*, e1004459.

Punnonen, E.L., Ryhänen, K., and Marjomäki, V.S. (1998). At reduced temperature, endocytic membrane traffic is blocked in multivesicular carrier endosomes in rat cardiac myocytes. Eur. J. Cell Biol. *75*, 344–352.

Qu, K., Glass, B., Doležal, M., Schur, F.K.M., Murciano, B., Rein, A., Rumlová, M., Ruml, T., Kräusslich, H.-G., and Briggs, J.A.G. (2018). Structure and architecture of immature and mature murine leukemia virus capsids. Proc Natl Acad Sci USA *115*, E11751–E11760.

Quinn, C.M., Wang, M., Fritz, M.P., Runge, B., Ahn, J., Xu, C., Perilla, J.R., Gronenborn, A.M., and Polenova, T. (2018). Dynamic regulation of HIV-1 capsid interaction with the restriction factor TRIM5α identified by magic-angle spinning NMR and molecular dynamics simulations. COMPUTATIONAL BIOLOGY 6.

Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C., and Rousso, I. (2017). Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. Journal of Virology *91*.

Rasaiyaah, J. (2013). HIV-1 evades innate immune recognition through specific cofactor recruitment. 16.

Ribbeck, K. (2001) Kinetic analysis of translocation through nuclear pore complexes. 11.

Rigort, A. (2010). Micromachining tools and correlative approaches for cellular cryo-electron tomography. Journal of Structural Biology 11.

Rigort, A., Bauerlein, F.J.B., Villa, E., Eibauer, M., Laugks, T., Baumeister, W., and Plitzko, J.M. (2012). Focused ion beam micromachining of eukaryotic cells for cryoelectron tomography. Proceedings of the National Academy of Sciences *109*, 4449–4454.

Robertson, D.L. (2000). HIV-1 Nomenclature Proposal. Science *288*, 55d–555.

Roebuck, K.A., and Saifuddin, M. (1999). Regulation of HIV-1 Transcription. 18.

Rosier, D.J.D. (1968). Reconstruction of Three Dimensional Structures from Electron Micrographs. 5.

Rossi, E., Meuser, M.E., Cunanan, C.J., and Cocklin, S. (2021). Structure, Function, and Interactions of the HIV-1 Capsid Protein. 25.

Rouzina, I., and Bruinsma, R. (2014). DNA confinement drives uncoating of the HIV Virus. The European Physical Journal Special Topics *223*, 1745–1754.

Rüegsegger, U., Blank, D., and Keller, W. (1998). Human Pre-mRNA Cleavage Factor Im Is Related to Spliceosomal SR Proteins and Can Be Reconstituted In Vitro from Recombinant Subunits. Molecular Cell *1*, 243–253.

Saito, A., Henning, M.S., Serrao, E., Dubose, B.N., Teng, S., Huang, J., Li, X., Saito, N., Roy, S.P., Siddiqui, M.A., et al. (2016). Capsid-CPSF6 Interaction Is Dispensable for HIV-1 Replication in Primary Cells but Is Selected during Virus Passage In Vivo. J. Virol. *90*, 6918–6935.

Savarino, A. (2007). In-Silico docking of HIV-1 integrase inhibitors reveals a novel drug type acting on an enzyme/DNA reaction intermediate. Retrovirology *4*, 21.

Schaller, T., Ocwieja, K.E., Rasaiyaah, J., Price, A.J., Brady, T.L., Roth, S.L., Hué, S., Fletcher, A.J., Lee, K., KewalRamani, V.N., et al. (2011). HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. PLoS Pathog. *7*, e1002439.

Scheres, S.H.W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology *180*, 519–530.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676–682.

Schmid, M.F., and Booth, C.R. (2008). Methods for aligning and for averaging 3D volumes with missing data. J Struct Biol *161*, 243–248.

Schorb, M., and Sieckmann, F. (2017). Matrix MAPS-an intuitive software to acquire, analyze, and annotate light microscopy data for CLEM. 13.

Schorb, M., and Briggs, J.A.G. (2014). Correlated cryo-fluorescence and cryo-electron microscopy with high spatial precision and improved sensitivity. Ultramicroscopy *143*, 24–32.

Schur, F.K.M. (2015). Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution. 16.

Schur, F.K.M., Obr, M., Hagen, W.J.H., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Kräusslich, H.-G., and Briggs, J.A.G. (2016). An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science *353*, 506–508.

Schwarz, H., and Humbel, B.M. (2007). Correlative Light and Electron Microscopy Using Immunolabeled Resin Sections. In Electron Microscopy, J. Kuo, ed. (Totowa, NJ: Humana Press), pp. 229–256.

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

Selyutina, A., Persaud, M., Simons, L.M., Bulnes-Ramos, A., Buffone, C., Martinez-Lopez, A., Scoca, V., Di Nunzio, F., Hiatt, J., Marson, A., et al. (2020). Cyclophilin A Prevents HIV-1 Restriction in Lymphocytes by Blocking Human TRIM5α Binding to the Viral Core. Cell Reports *30*, 3766-3777.e6.

Shen, Q., Xu, C., Jang, S., Xiong, Q., Devarkar, S.C., Tian, T., Bedwell, G.J., Tripler, T.N., Hu, Y., Yuan, S., et al. (2020). A DNA-origami nuclear pore mimic reveals nuclear entry mechanisms of HIV-1 capsid. 31.

Shimoni and Muller (1998). On optimizing high-pressure freezing: from heat transfer theory to a new microbiopsy device. J Microsc *192*, 236–247.

Singh, P.K., Plumb, M.R., Ferris, A.L., Iben, J.R., Wu, X., Fadel, H.J., Luke, B.T., Esnault, C., Poeschla, E.M., Hughes, S.H., et al. (2015). LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. Genes Dev. *29*, 2287–2297.

Sokolskaja, E., Sayah, D.M., and Luban, J. (2004). Target Cell Cyclophilin A Modulates Human Immunodeficiency Virus Type 1 Infectivity. JVI *78*, 12800–12808.

Sowd, G.A., Serrao, E., Wang, H., Wang, W., Fadel, H.J., Poeschla, E.M., and Engelman, A.N. (2016). A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. Proc Natl Acad Sci USA *113*, E1054–E1063.

Stein, B.S., Lifson, D., Bensch, K.G., and Engleman't, E.G. (1987). pH-Independent HIV Entry into CD4-Positive T Cells iia Virus Envelope Fusion to the Plasma Membrane. 10.

Strambio-De-Castillia, C. (2010). The nuclear pore complex: bridging nuclear transport and gene regulation. 12.

Strauss, J.D., Hammonds, J.E., Yi, H., Ding, L., Spearman, P., and Wright, E.R. (2016). Three-Dimensional Structural Characterization of HIV-1 Tethered to Human Cells. J. Virol. *90*, 1507–1521.

Studer, D., Graber, W., Al-Amoudi, A., and Eggli, P. (2001). A new approach for cryofixation by high-pressure freezing. J Microsc *203*, 285–294.

Studer, D., Humbel, B.M., and Chiquet, M. (2008). Electron microscopy of high pressure frozen samples: bridging the gap between cellular ultrastructure and atomic resolution. Histochem Cell Biol *130*, 877–889.

Stultz, R.D., Cenker, J.J., and McDonald, D. (2017). Imaging HIV-1 Genomic DNA from Entry through Productive Infection. Journal of Virology *91*.

Tacke, S., Erdmann, P., Wang, Z., Klumpe, S., Grange, M., Plitzko, J.M., and Raunser, S. (2020). A streamlined workflow for automated cryo focused ion beam milling (Molecular Biology).

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: An extensible image processing suite for electron microscopy. Journal of Structural Biology *157*, 38–46.

Temple, J. (2020). A snapshot of HIV-1 capsid-host interactions. 7.

Towers, G.J., Hatziioannou, T., Cowan, S., Goff, S.P., Luban, J., and Bieniasz, P.D. (2003). Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. Nat Med *9*, 1138–1143.

Turoňová, B., Schur, F.K.M., Wan, W., and Briggs, J.A.G. (2017). Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4 Å. Journal of Structural Biology *199*, 187–195.

Turoňová, B., Hagen, J.W., Obr, M., Mosalaganti S., Beugelink J.W., Zimmerli E.C., Kraüsslich H-G. and Beck M. (2020). Benchmarking tomographic acquisition schemes for high-resolution structural biology. Nature Communications *11*, 876.

Vanhecke, D., Graber, W., and Studer, D. (2008). Chapter 9 Close-to-Native Ultrastructural Preservation by High Pressure Freezing. In Methods in Cell Biology, (Elsevier), pp. 151–164.

Villa, E. (2013). Opening windows into the cell: focused-ion-beam milling for cryo-electron tomography. Current Opinion in Structural Biology 7.

Wagner, F.R., Watanabe, R., Schampers, R., Singh, D., Persoon, H., Schaffer, M., Fruhstorfer, P., Plitzko, J., and Villa, E. (2020). Preparing samples from whole cells using focused-ion-beam milling for cryo-electron tomography. Nat Protoc *15*, 2041–2070.

Walther, T.C. (2001). The nucleoporin Nup153 is required for nuclear pore basket formation, nuclear pore complex anchoring and import of a subset of nuclear proteins. The EMBO Journal *20*, 5703–5714.

Walther, T.C., Pickersgill, H.S., Cordes, V.C., Goldberg, M.W., Allen, T.D., Mattaj, I.W., and Fornerod, M. (2002). The cytoplasmic filaments of the nuclear pore complex are dispensable for selective nuclear protein import. Journal of Cell Biology *158*, 63–77.

Wan, W., and Briggs, J.A.G. (2016). Cryo-Electron Tomography and Subtomogram Averaging. In Methods in Enzymology, (Elsevier), pp. 329–367.

Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., and Briggs, J.A.G. (2017). Structure and assembly of the Ebola virus nucleocapsid. Nature *551*, 394–397.

Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Jr, W.B., Swanstrom, R., Burch, C.L., and Weeks, K.M. (2010). Architecture and Secondary Structure of an Entire HIV-1 RNA Genome. 23.

Weigel, P.H., and Oka, J.A. (1981). Temperature dependence of endocytosis mediated by the asialoglycoprotein receptor in isolated rat hepatocytes. Evidence for two potentially rate-limiting steps. J. Biol. Chem. *256*, 2615–2617.

Wilen, C.B., Tilton, J.C., and Doms, R.W. (2012). HIV: Cell Binding and Entry. Cold Spring Harbor Perspectives in Medicine *2*, a006866–a006866.

Wirth, R. (2009). Focused Ion Beam (FIB) combined with SEM and TEM: Advanced analytical tools for studies of chemical composition, microstructure and crystal structure in geomaterials on a nanometre scale. Chemical Geology 13.

Wolff, G., Limpens, R.W.A.L., Zheng, S., Snijder, E.J., Agard, D.A., Koster, A.J., and Bárcena, M. (2019). Mind the gap: Micro-expansion joints drastically decrease the bending of FIB-milled cryo-lamellae. J. Struct. Biol. *208*, 107389.

Worthylake, D.K., Wang, H., Yoo, S., Sundquist, W.I., and Hill, C.P. (1999). Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. Acta Crystallogr D Biol Crystallogr *55*, 85–92.

Wu, J., Matunis, M.J., Kraemer, D., Blobel, G., and Coutavas, E. (1995). Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. J Biol Chem *270*, 14209–14213.

Xie, L., Chen, L., Zhong, C., Yu, T., Ju, Z., Wang, M., Xiong, H., Zeng, Y., Wang, J., Hu, H., et al. (2020). MxB impedes the NUP358-mediated HIV-1 pre-integration complex nuclear import and viral replication cooperatively with CPSF6. Retrovirology *17*, 16.

Yamashita, M., and Engelman, A.N. (2017). Capsid-Dependent Host Factors in HIV-1 Infection. Trends in Microbiology *25*, 741–755.

Yang, Y., Fricke, T., and Diaz-Griffero, F. (2013). Inhibition of Reverse Transcriptase Activity Increases Stability of the HIV-1 Core. Journal of Virology *87*, 5.

Yokoyama, N., Hayashi, N., Seki, T., Panté, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, T., and Aebi, U. (1995). A giant nucleopore protein that binds Ran/TC4. Nature *376*, 184–188.

Zachs, T., Schertel, A., Hugener, J., Matos, J., and Pilhofer, M. (2020). Fully automated, sequential focused ion beam milling for cryo-electron tomography. 14.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhou, L., Sokolskaja, E., Jolly, C., James, W., Cowley, S.A., and Fassati, A. (2011). Transportin 3 Promotes a Nuclear Maturation Step Required for Efficient HIV-1 Integration. PLoS Pathogens *7*, e1002194.

Zhu, P., Chertova, E., Bess, J., Lifson, J.D., Arthur, L.O., Liu, J., Taylor, K.A., and Roux, K.H. (2003). Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. Proceedings of the National Academy of Sciences *100*, 15812–15817.

Zila, V., Müller, T.G., Laketa, V., Müller, B., and Kräusslich, H.-G. (2019). Analysis of CA Content and CPSF6 Dependence of Early HIV-1 Replication Complexes in SupT1-R5 Cells. *10*, 20.

Zila, V., Margiotta, E., Turoňová, B., Müller, T.G., Zimmerli, C.E., Mattei, S., Allegretti, M., Börner, K., Rada, J., Müller, B., et al. (2021). Cone-shaped HIV-1 capsids are transported through intact nuclear pores. Cell *184*, 1032-1046.e18.

Zimmerli, C.E., Allegretti, M., Rantos, V., Goetz, S.K., Obarska-Kosinska, A., Zagoriy, I., Halavatyi, A., Mahamid, J., Kosinski, J., and Beck, M. (2020). Nuclear pores constrict upon energy depletion (Cell Biology).