Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Sandra Schifferdecker

Born in: Mannheim, Germany

Oral examination: 28.03.2022

Direct HIV-1 CA labeling by genetic code expansion using the amber suppression system

Referees:

Prof. Dr. Hans-Georg Kräusslich

Prof. Dr. Friedrich Frischknecht

I. Contents

List of Figures

List of Tables

List of Abbreviations

AAV	adeno-associated virus
a.u.	arbitrary units
AIDS	acquired immunodeficiency syndrome
APC	aphidicolin
cART	combined antiretroviral therapy
BICD2	bicaudal D homolog 2
CA	capsid protein
Capsid	the assembled cone-shaped core consisting of CA
CCR5	C-C chemokine receptor type 5
CLEM	correlative light electron microscopy
CMV	cytomegalovirus
COVID-19	coronavirus disease 2019
CpK	cyclopropene-L-lysine
CPSF6	cleavage and polyadenylation specificity factor 6
Cryo-ET	electron cryotomography
CsA	cyclosporine A
CTD	C-terminal domain
CXCR4	C-X-C Motif Chemokine Receptor 4
CypA	cyclophilin A
CypA-DsRed	fluorescently tagged version of the CA-binding host protein CypA
dsDNA	double-stranded DNA
EM	electron microscopy

Env	envelope
eRF1	eukaryotic release factor 1
ET	electron tomography
FEZ1	fasciculation and elongation protein zeta 1
FG	phenylalanine glycine
FISH	fluorescence in-situ hybridization
GCE	genetic code expansion
h p.i.	hours post-infection
h p.t.	hours post-transfection
H3K36me3	trimethylation of lysine 36 in histone H3
HIV-1	human immunodeficiency virus subtype 1
HIV-1*	plasmid of HIV-1 molecular clone NLC4-3 Vpr_{TGA}
HIV-1*CAxx$^{TGA/CpK/SiR}$	plasmid of HIV-1 molecular clone NLC4-3 Vpr_{TGA} with stop codon insertion within the CA; incorporated ncAA (CpK); coupled to Tet-SiR
IF	immunofluorescence
iGFP	internal fluid-phase GFP marker
IN	integrase
IN.FP	fluorescently tagged IN
IP6	inositol hexakisphosphate
LADs	lamina-associated domains
LEDGF/p75	lens epithelium-derived growth factor p75 splicing variant
LPV	lopinavir
LTR	long-terminal repeat
MA	matrix
mCLING	membrane-binding fluorophore-cysteine-lysine-palmitoyl group
MDM	monocyte-derived macrophages
MFI	mean fluorescence intensity
MOI	multiplicity of infection
NC	nucleocapsid
ncAA	non-canonical amino acid
NE	nuclear envelope
NPC	nuclear pore complex
NTD	N-terminal domain
Nup153	nucleoporin 153
Nup358	nucleoporin 358
ORFs	open reading frames
ParB	chromosome-partitioning protein B
parS	chromosome-partitioning sequence S
PFA	paraformaldehyde
PF74	PF-3450074; HIV-1 inhibitor
PIC	pre-integration complex
PM	plasma membrane
PR	protease
PylRS	pyrrolysine-tRNA synthetase

qPCR	quantitative/real-time polymerase chain reaction
RF1	release factor 1
RT	reverse transcriptase
RTC	reverse transcription complex
SDCM	spinning-disc-confocal microscopy
Sec24C	component of the coat protein complex II (COPII)
SG-PERT	SYBR green I-based real-time enhanced reverse transcriptase
SiR	silicon-rhodamine
smFRET	single-molecule fluorescent resonance energy transfer
sp1	spacer peptide 1
sp2	spacer peptide 2
SPIEDAC	strain-promoted inverse electron-demand Diels-Alder cycloaddition
ssRNA	single-stranded RNA
STED	stimulated emission depletion
T-20	peptide drug Enfuvirtide; fusion inhibitor
TAR	trans-activation response element
Tat	trans-activator of transcription
TC	tetracysteine
Tet-SiR	tetrazine coupled silicon-rhodamine dye
TNPO3	transportin-3
TRN1	transportin-1
Trim5α	tripartite motif-containing protein 5
vDNA	viral DNA
Vpr	viral protein R
Vpu	viral protein U
WT	wild-type

List of publications

Schifferdecker, S., Zila, V., Müller, T.G., Sakin, V., Anders-Össwein, M., Laketa, V., Kräusslich, H.-G., Müller, B. (2021). Direct capsid labeling of infectious HIV-1 by genetic code expansion allows detection of largely complete nuclear capsids and suggests nuclear entry of HIV-1 complexes via common routes. bioRxiv 2021.09.12.460218. doi: https://doi.org/10.1101/2021.09.14.460218

Müller, T. G., Zila, V., Peters, K., **Schifferdecker, S.**, Stanic, M., Lucic, B., Laketa, V., Lusic, M., Müller, B., Kräusslich, H.-G. (2021). HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells', eLife, 10, pp. 1–32. doi: 10.7554/ELIFE.64776.

Ritter, C., Newrly, A., **Schifferdecker, S.,** Roggenbach, I., Müller, B., Rohr, K. (2020). Optimizing particle detection by colocalization analysis in multi-channel fluorescence microscopy Images. IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, pp. 882-885.

II. Summary

HIV-1 infects $CD4^+$ T cells and macrophages as specific target cells where it replicates by hijacking host cell components. This capsid containing the viral RNA genome consists of ~ 1,200-1,500 capsid proteins (CA) and is transported towards the nucleus to integrate into the host genome. Before integration, the capsid needs to be removed in a process termed uncoating. However, the exact nature and intracellular location of uncoating is not known. Microscopy-based analysis would be capable to address this issue by detecting individual particles. The lack of an experimental system for direct fluorescent labeling of the capsid structure impaired microscopic analyses of CA content within subviral complexes. Experiments relying on immunofluorescence staining depend on epitope accessibility, while CA fusions to fluorescent proteins severely affect infectivity. To overcome these issues, a site-specific and minimally invasive labeling strategy of HIV-1 CA based on genetic code expansion *via* amber suppression followed by click-labeling with silicon rhodamine (SiR) was developed. 18 amino acids within the protein were screened to identify positions permissive for modification. Out of these candidates, alanine 14 allowed labeling of > 95 % of particles within a preparation while retaining ~ 50 % infectivity without the need for complementation with the wild-type protein. Therefore, this strategy enabled reliable CA content analysis for subviral complexes in different subcellular compartments and single virus tracking by confocal imaging, super-resolution STED nanoscopy, and correlative light electron microscopy (CLEM). Directly labeled particles were detected inside the nucleus of HeLa-derived cells and primary $CD4^+$ T cells with intensities corresponding to the amount of a completely assembled CA lattice. CLEM revealed several closely adjacent cone-shaped objects inside the nucleus of cell cycle arrested T cells correlating to positions of CA SiR signals. These data strongly argue for the translocation of largely intact capsids through nuclear pore complexes (NPCs) into the nucleus in both cell types. Additionally, the analysis of cytoplasmic and nuclear signals by STED imaging suggested nuclear capsid trafficking on common routes resulting in dose-dependent capsid clustering inside the nucleus. The unique properties of the direct labeling approach allowed to clarify the controversially discussed effect of the CA-targeting inhibitor PF74. Upon adding the drug, the complete CA-content of nuclear complexes was retained, indicating that PF74 has a stabilizing effect on the capsid lattice. Together these data support a model in which intact capsids enter the nucleus through NPCs, where uncoating must occur before productive integration. In conclusion, the establishment and thorough validation of this HIV-1 CA labeling method provides a versatile tool to investigate uncoating dynamics. The approach may be applicable to the capsids of other viruses.

III. Zusammenfassung

HIV-1 infiziert gezielt $CD4^+$ T-Zellen und Makrophagen und nutzt deren eigene Komponenten zur Replikation. Das Kapsid umhüllt die virale RNA und besteht aus 1200 bis 1500 Kapsidproteinen (CA). Dieses wird in Richtung Zellkern transportiert, um sein Erbgut dort in das Wirtsgenom zu integrieren. Vorher muss die Kapsidhülle jedoch in einem als „Uncoating“ bezeichnetem Prozess entfernt werden. Der genaue Mechanismus und der intrazelluläre Ort dieses Uncoatings ist bislang nicht bekannt. Ein besseres Verständnis dieses Prozesses würde man mittels mikroskopischer Analyse gewinnen, da diese eine Detektion einzelner viraler Partikel erlaubt. Durch das Fehlen einer Methode zur direkten Fluoreszenzmarkierung des Kapsids war es bislang nicht möglich subvirale Komplexe mikroskopisch zu analysieren und deren CA-Mengen genau zu messen. Während die bisher verwendeten Immunfluoreszenzfärbungen CA-Epitope nicht ausreichend binden konnten, beeinträchtigte das Anhängen fluoreszierender Proteine die Infektiosität der Viren. Um solche Probleme zu umgehen, wurde eine ortsspezifische und minimal invasive Markierungsstrategie für HIV-1 CA entwickelt. Diese basiert auf der Erweiterung des genetischen Codes unter Verwendung des Amber-Suppressionssystems mit anschließender CA-spezifischer Klickfärbung unter Verwendung von Silizium-Rhodamin (SiR). In der vorliegenden Arbeit wurden 18 verschiedene Proteinstellen getestet, um eine geeignete Aminosäure zu identifizieren, welche für eine Modifikation in Frage kommt. Von diesen Kandidaten ermöglichte Alanine 14 die Markierung von > 95 % der produzierten Partikel, während eine Infektiosität von etwa 50 % erhalten blieb, ohne dass eine Komplementierung mit dem Wildtyp-Protein erforderlich war. Dieser Ansatz ermöglicht eine zuverlässige Analyse des CA-Gehalts in verschiedenen zellulären Kompartimenten. Außerdem können hierdurch einzelne virale Komplexe mit Hilfe konfokaler Bildgebung, superauflösender STED-Analyse und korrelativer Licht- und Elektronenmikroskopie (CLEM) auf ihrem Weg durch die Zelle verfolgt werden. Sowohl in den Zellkernen von HeLa-Zellen als auch primären CD4+ T-Zellen konnten direkt markierte Partikel nachgewiesen werden. Ihre Intensitäten entsprachen dabei denen vollständig zusammengesetzter CA-Gitter. Mittels CLEM wurden mehrere dicht nebeneinander liegende kegelförmige Objekte im Zellkern teilungsunfähiger T-Zellen nachgewiesen, deren Positionen, mit denen der CA-SiR-Signale korrelierten. Diese Daten sprechen dafür, dass in beiden Zelltypen weitgehend intakte Kapside vermutlich durch Kernporenkomplexe (NPCs) in den Zellkern eingetreten sind. Darüber hinaus deutet die STED-Analyse der zytoplasmatischen und nukleären Signale darauf hin, dass die Kapside innerhalb des Zellkerns auf definierten Routen transportiert werden und es innerhalb des Kerns zu einer dosisabhängigen

Anhäufung von Kapsiden kommt. Durch den direkten Markierungsansatz konnte außerdem die kontrovers diskutierte Wirkung des CA-bindenden Inhibitors PF74 untersucht werden. Da auch nach dessen Zugabe der CA-Gehalt erhalten blieb, scheint PF74 eine stabilisierende Wirkung auf das CA-Gitter zu haben. Diese Ergebnisse unterstützen die Theorie, dass intakte Kapside durch NPCs in den Zellkern gelangen, wo schließlich auch das Uncoating stattfindet. Die in dieser Arbeit etablierte und validierte Methode zur Markierung von HIV-1 CA bietet ein vielseitiges Instrument zur Untersuchung des Uncoatings und könnte künftig für weitere Viren angewendet werden.

1 Introduction

1.1 The human immunodeficiency virus (HIV)

Viruses are infectious agents, which aim at transmitting their genomic information by replication within a living organism. According to several estimates, ~ 10^{31} viruses exist on our planet, infecting animals, plants, and microorganisms, but less than 1 % of these species are human pathogens[1]. Interestingly 8 % of the human genome, consisting of roughly 22,000 genes, contains sequences from remnants of ancient viruses. Furthermore, ~ 40 % of the human genome is built of repetitive sequences with a viral origin, which can alter the expression of other genes[2], called transposons. Until this day, the origin of viruses remains unclear and is under constant debate[3]. In order to give a sorted overview of this large variety, different viruses were assigned to different classes. Since we lack detailed information about the origin and virus evolution, the Baltimore scheme is used to classify viruses based on their genome and replication strategy[4]. Within this system, the 1981 isolated human immunodeficiency virus (HIV)[5] was categorized as a lentivirus within the family of *Retroviridae* and assigned to class VI. The name 'retrovirus' stems from the mechanism used by these viruses. They use RNA-dependent DNA polymerase to reverse transcribe their single-stranded RNA (ssRNA) into double-stranded DNA (dsDNA) to incorporate permanently in the host DNA. By identifying this molecule named reverse transcriptase (RT)[6,7], the initial understanding of the central dogma of molecular biology was extended, showing that the flow of genetic information was not limited to 'from DNA to RNA'. Compared to retroviruses from other genera such as γ-retroviruses (e.g., murine leukemia virus), the more complex lentiviruses carry several viral proteins leading to long incubation periods. Besides, lentiviruses do not need nuclear envelope (NE) breakage to enter the nucleus and efficiently infect a cell.

The human immunodeficiency virus 1 (HIV-1), as a representative of lentiviruses, attacks the human immune system mainly by depleting $CD4^+$ T cells. If an HIV-1 infection remains untreated, it causes the acquired immunodeficiency syndrome (AIDS) in the vast majority of patients and eventually death due to opportunistic diseases. Although researchers have studied this virus for more than 40 years, it still constitutes a global public health issue without the possibility to eliminate the virus from an infected person. The virus's capability to successfully evade the immune system building a provirus integrated into the host genome without replicating makes it harder to diagnose as well as to discover a protective vaccine. Nevertheless, the development and constant advancements of

combined antiretroviral therapy (cART), which can control the virus and prevent disease outbreak, has made HIV/AIDS a potentially manageable disease[8]. The main goal of cART, when combined with more routine screening and early diagnosis, is to prevent virus transmission and lower HIV-associated mortality by suppressing several viral replication steps.

Despite large-scale awareness and prevention efforts, a total of 37.7 million people were known to live with HIV[9] in 2020, and 1.5 million new infections and 680.000 deaths were registered that same year. Especially in low-income countries AIDS is one of the most common causes of death[10]. The fear of stigmatization and unequal access to health care are still causing a large number of HIV-cases to remain undetected[11,12]. Recent advances have further been interrupted by the outbreak of the coronavirus disease 2019 (COVID-19) in 2020[13]. Several studies assume that quarantine and social distancing in 2020 and 2021 led to reduced testing, diagnosis, and delayed ART, which hampered the medically necessary care of HIV-infected patients[14,15]. The full impact of COVID-19 and related public health measures on people living with poorly treated HIV needs to be assessed further[16].

Due to issues with undetected cases, unequal access to treatment, and (perinatal) transmission prevention further investigation of HIV as well as extended measures to increase routine testing are still necessary. To develop the best possible treatment options and eventually a cure, a full understanding of the viral replication cycle is required.

1.2 Genome and particle structure of HIV-1

The structure of HIV-1 is well understood and characterized. The ~ 9 kb genome of HIV-1 consists of nine open reading frames (ORFs) encoding for 15 different viral proteins (Figure 1a)[17]. Three central genomic domains, *gag* (group-specific antigen), *pol* (polymerase), and *env* (envelope), encode polyproteins, which are then proteolytically cleaved. Gag supplies proteins for the inner structure of the virion, namely matrix (MA), CA, nucleocapsid (NC). Also Gag-derived are p6, spacer peptide 1 (sp1), and spacer peptide 2 (sp2). Pol is expressed as Gag-Pol by ribosomal frameshifting in a 1:10 ratio to Gag. It is cleaved into the viral enzymes protease (PR) to cleave Gag, RT to transcribe ssRNA into dsDNA, and integrase (IN) for integration into the host DNA. Env provides the two glycoproteins gp120 and gp41 essential for receptor binding and fusion. Apart from structural proteins, HIV-1 encodes several nonstructural proteins to support the viral replication. Accessory proteins viral protein U (Vpu), viral infectivity factor (Vif), viral protein R (Vpr), and negative factor (Nef) regulate the innate and intrinsic immunity. Two essential regulatory RNA elements, Tat (trans-activator of transcription) and Rev (regulator of virion), enhance viral gene expression. The proviral DNA is

flanked by peripheral long-terminal repeats (LTRs), and the 5'-LTR serves as the promoter for expression.

The accessory proteins are required not only for efficient viral RNA synthesis, but also release of assembled virions. Tat activates the RNA-polymerase II by binding to an RNA stem-loop secondary structure, which builds the trans-activation response element (TAR) to recruit a general elongation factor and ensure efficient transcriptional elongation. By differential RNA splicing, all 15 proteins can be expressed. Rev has the major function of exporting unspliced pre-mRNAs out of the nucleus. Rev response elements (RREs) on the mRNA induce Rev binding and multimerization, which disclose a nuclear export signal. By building a complex with other proteins, pre-mRNAs can be saved from splicing and translocated into the cytoplasm of an infected cell. Subsequently, the NC domain guides the exported viral genomic RNAs into particles ready for assembly and budding, driven by Gag, at the plasma membrane (PM). While CA ensures Gag-Gag interactions, the N-terminal MA region interacts with gp120 to incorporate Env into the Gag lattice and anchors into the host PM. After budding at the PM, immature particles with a spherical Gag shell must be processed by the viral PR resulting in infectious and mature virions of ~ 100-200 nm in diameter with a closed fullerene cone (capsid) inside comprising essential proteins in the center (Figure 1b).

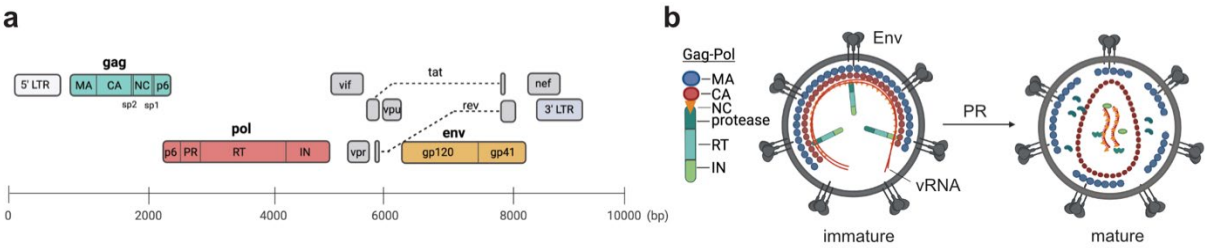


Figure 1. HIV-1 genome and particle structure. a. HIV-1 genome organization. **b.** Scheme of an immature (left) and mature (right) HIV-1 virion. After cleavage by viral PR and structural arrangement of the components, the mature virion is ready to infect new host cells. Created with BioRender.com

1.3 HIV-1 replication cycle

In order to reach the host cell nucleus and productively infect a cell with subsequent release of new infectious particles the HIV-1 virus must pass several barriers (Figure 2)[18]. The Env gp120 protein is necessary for the initial target cell binding and coreceptor interactions inducing fusion of the viral lipid bilayer with the host PM. Upon cytoplasmic entry of the viral core, the capsid containing all

essential viral components must traffic towards the nuclear pore complex (NPC) using a network of microtubules and supporting host proteins[19].

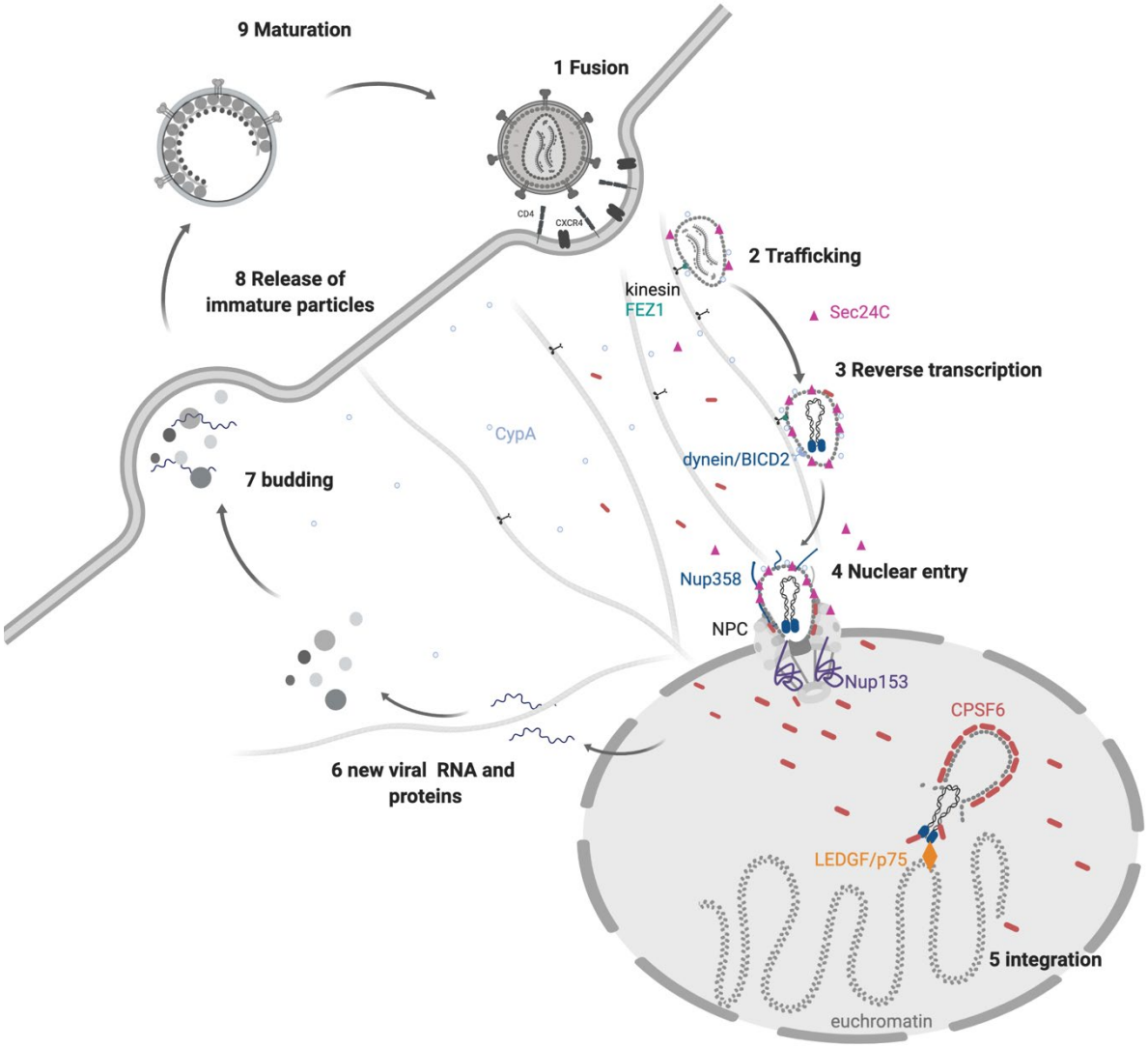


Figure 2. The HIV-1 replication cycle. Cell entry is mediated by the viral Env and receptor binding leading to fusion at the PM and release into the cytoplasm (1). The viral capsid is trafficking towards the nucleus by hijacking different cytoskeleton components (2) while reverse transcription is initiated inside the core (3). Trafficking is supported by several host factors (kinesin/FEZ1, Sec24C, dynein/BICD2, CypA). After docking at the NPC via Nup358 interaction, the viral complex enters the nucleus through Nup153 binding and interacting with different host factors (4). Inside the nucleus, the capsid is surrounded by CPSF6, and the viral genome is released and integrated into the host DNA (5). New viral RNA and proteins are released after transcription and translation by the host machinery (6). Budding of assembled viral components occurs at the PM (7), and new immature particles are released (8). To become fully infectious, the particles must undergo maturation with structural rearrangements (9). The figure was created with BioRender.com

During cytoplasmic trafficking, reverse transcription is initiated within the capsid core to reverse-transcribe ssRNA into dsDNA. The docking of the capsid at the NPC and interactions with host proteins seem to be indispensable for NPC passaging and nuclear entry. Inside the nucleus, the IN

enzyme catalyzes the integration of the viral genome into the host DNA. The integrated provirus then serves as a template for viral RNA synthesis by the host machinery. After producing all viral proteins, the progeny virions assemble at the PM, and the final event of budding releases new immature particles. Finally, PR cleavage and structural rearrangements lead to fully mature virions.

1.3.1 Cell entry, cytoplasmic trafficking, and nuclear entry

The first stage of the replication cycle requires the virus to enter the host cell. Enveloped viruses lead to productive infection upon fusion with the host PM using different molecular mechanisms[20–23]. The HIV-1 Env precursor glycoprotein gp160 is cleaved by the furin protease into gp120 and gp41[24]. The gp120 fragment initiates CD4 receptor binding, which induces conformational rearrangements and exposure of additional coreceptor binding site[25]. The concurrent binding of CD4 and co-receptors (C-X-C motif chemokine receptor 4 (CXCR4) or C-C chemokine receptor type 5 (CCR5)) triggers the gp41 refolding to translocate the N-terminal fusion peptide into the host PM. To cross this first kinetic barrier of the infection, the viral Env proteins undergo different conformational changes[22]. Starting from a high-energy prefusion state, it forms intermediates with lower energy levels to reach the hemifusion state where both bilayers are in close contact and ready to merge. Conformational changes in Env trigger membrane fusion and capsid release released into the host cytosol. Exposed to the host intrinsic immune system, the capsid interacts with protective cyclophilin A (CypA) and Sec24C (Figure 2, light blue and magenta) before exploitation of the cellular cytoskeleton and proteins.

Due to a dense cytoplasmic environment filled with different organelles and proteins, mega-dalton-sized viral complexes cannot reach the nucleus by passive diffusion alone[26]. Therefore, the cytoskeleton network, consisting of actin and microtubules, is indispensable for the cytoplasmic transport of incoming replication complexes towards the nucleus[27,28]. The major viral player facilitating cellular trafficking is the capsid hijacking the cytoskeleton network via interaction with many host factors via CA binding[19]. One such CA-interacting protein is the fasciculation and elongation protein zeta 1 (FEZ1; Figure 2, green), serving as an adaptor for kinesin-1 binding to facilitate directed movement along microtubules[29]. FEZ1 was suggested to interact with multiple capsid subdomains and mutations of distinct FE1 residues resulted in reduced observed cytosolic trafficking underlining the importance of this molecule interacting with capsid[30]. Another adaptor protein, bicaudal D homolog 2 (BICD2; Figure 2, blue), was found in a large RNAi screening to interact with HIV-1 and to be essential for replication[31]. BICD2 is a dynein motor adaptor that binds to the dynein-dynactin complex activating dynein processivity and increasing trafficking of cellular

cargoes[32,33]. BICD2 interacts with HIV-1 particles via its coiled-coiled 3 domain binding *in vitro* assembled CA-NC cores, and depletion decreased cytoplasmic trafficking and nuclear import[34]. In addition, Carnes *et al*.[35] showed that BICD2 associates with capsid-like assemblies in cell extracts and that it is crucial for the capsid-dynein interaction, retrograde movement, and viral infectivity.

While the dependence of HIV-1 on the host cell cytoskeleton, which facilitates the transport towards the nucleus and productive cell infection has been clearly demonstrated, the mechanisms of viral trafficking, including interactions with motor proteins, are not yet fully understood[27,28,36]. Different studies reported that microtubule network disruption has only a slight effect on infectivity and suggested an alternative microtubule-independent trafficking[37,38]. Conceivably, HIV-1 viruses can hijack actin microfilaments instead[39], which are essential for particle fusion[37,40], the association of viral complexes with the cytoskeleton[37], and productive reverse transcription[37]. Accordingly, a distinct role of actin was recently also connected to the initiation of podosome-mediated virus entry in macrophages[41]. Yet, the reason why the virus uses minus- and plus-end-directed motors for microtubular trafficking while reaching a net movement towards the nucleus remains unclear. One hypothesis is the so-called tug-of-war model describing the situation of two types of motors pulling against each other on the same cargo[42,43]. As soon as one motor gains the upper hand, it can pull the cargo in its direction.

After arriving at the nuclear envelope, the viral genome must be transported into the nucleus to integrate and productively infect the cell. The capsid is the primary interaction partner for host proteins and a major key player during cytoplasmic trafficking. Furthermore, serval findings led to the general conclusion that the viral capsid mediates not only cytosolic transport and ongoing reverse transcription but also nuclear translocation[19,44–46]. The NPC is a scaffold-building complex, an octagonal pore, of ~ 30 nucleoporins (Nups) filled with phenylalanine glycine (FG)-containing Nups measuring 95 nm in diameter[47]. This subset of scaffold Nups regulate bidirectional transport through the NPC allowing fast but selective transport through the pores distributed along the NE barrier[48–50]. While small proteins up to ~ 9 nm in diameter can cross the NPC by passive diffusion, larger complexes, such as the HIV-1 capsid, must be transported actively in an energy-dependent process[51]. The stationary phase of nucleocytoplasmic transport starts with the interaction of long cytoplasmic filaments of the nucleoporin 358 (Nup358; Figure 2, dark blue) serving as the first NPC docking platform for viral complexes. Nup358 is interacting with the capsid through FG repeats binding its cyclophilin homology domain to the capsid's CypA loop[52,53]. So-called nuclear transport factors interact with FG-Nups to move between the cytoplasmic and nuclear compartments also supporting the transport through the NPC. Nuclear transport receptors (e.g., transportin-1 (TRN1),

transportin-3 (TNPO3) are known to interact with capsid[54,55]) can recognize nuclear import or export signals on cargos mediating FG-Nup interaction and transport through the channel[51,56]. It is assumed that the inner nuclear basket component nucleoporin 153 (Nup153)[53] (Figure 2, purple) takes over from the central channel and replaces the Sec24C protein, which is known to interact with capsid in the cytoplasm by binding the same capsid motif (Figure 2)[57]. By following this binding cascade, the spatial distance between the core and the nucleoplasmic site can be reduced. The release into the nucleus might be facilitated in turn by cleavage and polyadenylation specificity factor 6 (CPSF6) binding on the nucleoplasmic region of the NPC, pulling viral complexes into the nucleus.

1.3.2 Reverse transcription

The characteristic feature of retroviruses is the RT with a polymerase and RNase H activity, responsible for reverse transcription of the viral RNA into DNA within the capsid core. At this stage, the viral complex is referred to as reverse transcription complex (RTC). The heterodimeric RT architecture comprises a hand-like p66 domain (fingers, palm, thumb) with the DNA-binding groove, active site, RNase H domain, and structural p51 domain (Figure 3a). To initiate reverse transcription (Figure 3b), the RT recruits host $tRNA^{Lys3}$ as a primer to anneal to the pbs (primer binding site) at the 5'-end of the HIV-1 genome (Figure 3b A). A complex reverse transcription process with two crucial strand transfers is required to reconstitute the viral LTR promoter, which is needed for Tat-induced transcription in virus producing cells. After initializing at the tRNA primer, (-) strand synthesis is completed until the 5'-end (Figure 3b B) and the product is subsequently transferred to the 3'-end of the genomic RNA (Figure 3b C; 1. strand transfer). During the generation of the (-) strand DNA, a DNA-RNA intermediate complex is formed, which is recognized by the RNase H, leading to the digestion of the viral RNA (vRNA; Figure 3b D). Polypurine tract (ppt) fragments are resistant against RNase H activity and stay as primers (Figure 3b E) for the following (+) strand cDNA synthesis (Figure 3b F). By synthesizing the 3'-LTR regions (Figure 3b G), the reverse transcription at the 3'-end is completed, and the first dsDNA fragment is formed containing the $tRNA^{Lys3}$ pbs sequence (Figure 3b H). The second strand transfer implies the pbs hybridization of the (+) strand DNA and (-) strand sequence (Figure 3b I). Finally, the whole dsDNA can be completed, and the functional promoter (LTRs) are reinstalled (Figure 3b J). From now on, the viral complex is named pre-integration complex (PIC) since the double-stranded vDNA is prepared for integration into the host genome. After successful integration, the provirus can remain unremarkably in the genome or serve as the template for producing new virions.

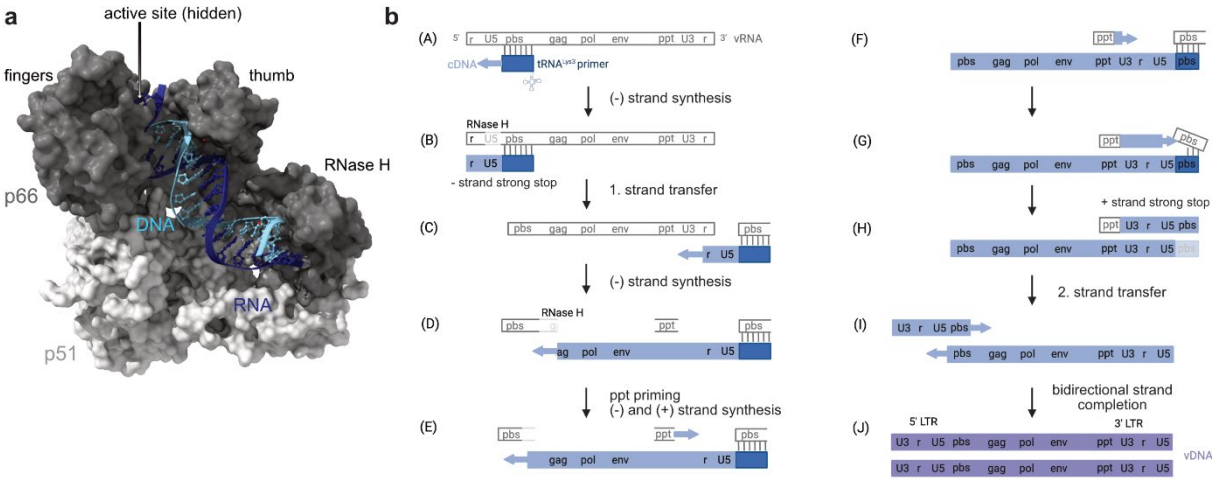


Figure 3. The HIV-1 RT and the reverse transcription process. a. Surface rendered RT-RNA/DNA polymerase complex (pdb: 4pqu) with the p66 (dark grey) and p51 structural domain (light grey). Nucleic acids (blue) are represented in cartoons/atoms mode. The active site of the polymerase, where dATP substrate is bound for elongation, is hidden behind the fingers and not visible. Rendering was done using ChimeraX. **b.** Individual steps of the reverse transcription mechanism to generate full LTR sequences. Viral RNA is represented in grey, the $tRNA^{Lys3}$ primer in dark blue and generated DNA in light blue. r: repetitive region in the LTR; U5: unique region of 5'-LTR; pbs: primer binding site; ppt: polypurine tract; U3: unique region of 3'-LTR. The figure was created with BioRender.com.

While essential RT components and the mechanism behind reverse transcription have been revealed, the exact location of the process is not yet fully identified. A widely accepted model entails early onset of reverse transcription after cellular entry of incoming viral structures and a completion of the process during trafficking of the viral core inside the cytosol. However, recent studies indicate that reverse transcription of the viral genome might only be completed after nuclear entry. These studies reported a maintained sensitivity of viral complexes to RT inhibitors even after nuclear import[58–60] or higher RT products in the nucleus compared to the cytosol measured in cell fractionation experiments[61]. But reverse transcription does not seem to be crucial for the nuclear import process itself[58,61–63].

This model, which described the nucleus as the subcellular location where reverse transcription is finished is also supported by our recently published study showing that dsDNA of viral structures became only accessible after nuclear entry[64]. This study used the ANCHOR system, which was developed based on the ParB-parS chromosomal partitioning system to visualize single genes in mammalian cells[65,66]. The ParB protein (OR3) binds the parS seed sequence (ANCH) specifically. The introduction of the ANCH sequence into the HIV-1 genome and expression of fluorescently tagged OR3 by a stable cell line allowed the detection of vial dsDNA[64,67]. Visualization of reverse transcribed cDNA by this fluorescent marker showed that cDNA was exclusively detected inside the nucleus. By

using the nucleotide analogue 5-ethynyl-2'deoxyuridine (EdU), which is incorporated into nascent vDNA during reverse transcription, the same study showed that EdU signals were higher at nuclear structures arguing together for completion of reverse transcription after nuclear entry. Together with high nuclear CA intensities detected by immunofluorescence (IF) these data suggested the connection between RT and capsid uncoating. It described the nucleus as the site where RT is finished and partial uncoating occurs. However, we do not know how these complex processes are spatio-temporally connected and if reverse transcription dynamics are cell type dependent.

1.3.3 Integration

After reaching the nucleus and opening of the capsid to expose the viral genome, the integration of the viral genome into host chromatin, either to express its genome or latently infect a cell, is primarily mediated by the viral IN and CA[68,69]. The linear viral reverse transcripts are connected by a few IN molecules in a so-called intasome complex[70]. The IN has the two primary tasks of 3' processing and strand transfer. During trafficking before reaching the nucleus, the IN removes terminal nucleotides resulting in 3'-hydroxyl groups to cut chromosomal DNA[71]. Then, the 3'-ends of the vDNA serve as nucleophiles targeting the 5'-phosphate groups of the host DNA, initiating the strand transfer and integration. After the host machinery and proteins repair the generated gaps, the integrated viral genome remains a provirus. Investigation of this integration process and tracking of incoming PICs led to the conclusion that the virus integrates at preferred nuclear sites[72] in a gene-dense region of highly expressed and active genes[73,74].

An important role in this process is ascribed to CPSF6, which is an integration-influencing factor interacting with CA[63,75,76] (Figure 3, red). As CPSF6 is binding the same cleft in the hexameric CA lattice as Nup153 (Figure 2), it is assumed that CPSF6 is pulling the CA-containing PIC from the nuclear basket towards nuclear integration sites transferring the CA directly to the process of integration[77]. Convincing evidence for CPSF6-mediated integration was shown by altered integration sites at heterochromatin domains in so-called lamina-associated domains (LADs)[76,78] upon CPSF6 removal. A second factor affecting integration is the lens epithelium-derived growth factor p75 splicing variant (LEDGF/p75)[75,79,80] (Figure 3, orange) interacting with the viral IN in the nucleus. It was reported to mediate the localization of PICs to mentioned highly expressed gene, whereas LEDG/p75 depletion led to integration at gene-poor regions[81].

LEDGF/p75 and CPSF6 were also reported to mediate the virus into clusters of super-enhancer domains containing high level of specific histones and bind co-activators for transcription[82–84]. In

line with this observation, the virus was reported to be directed by CA-CPSF6 interactions towards nuclear speckles and speckle-associated domains[85,86], which imply gene-dense compartments[87] and frequently contain super enhancers[88].

Currently it can be concluded that the incoming capsids are bound by CPSF6, which accumulate around PICs and direct the capsids towards specific nuclear sites. Since different studies detected huge clusters of several individual viral structures associated with CPSF6[64,85,89], nuclear viral clustering could be a determinant of successful integration[73] by LEDGF/p75-mediated chromatin binding[90]. Via interaction of LEDGF/p75 with the trimethylation of lysine 36 in histone H3 (H3K36me3) transcription and DNA damage response is regulated[91].

1.3.4 Different views of capsid uncoating

Despite the good understanding of the general HIV-1 replication process, CA uncoating remains a controversially discussed aspect of the HIV-1 life cycle[92–94]. The viral nucleic acid must be released from the capsid core in order to integrate into the host genome, thus being exposed to the cellular environment. Although we know that the CA play a central role as interaction partner promoting the viral infection at different stages of the replication cycle, the extent and kinetics of CA dissociation, described in a mechanism called uncoating in order to expose the genome to the cell environment, are still heavily discussed[92,94]. Notably, existing literature does not agree on the exact timing, subcellular location, and uncoating mechanism, offering different capsid uncoating models.

Early investigations of the uncoating process were mainly based on biochemical CA detection after fractionation of cytoplasmic and nuclear extracts, where only traces of CA were found to be associated with subviral complexes within the cytosolic region[70,95,96]. But at the same time, Zennou *et al*.[97] presented evidence arguing for later uncoating. In this study, fluorescence in-situ hybridization (FISH) experiments even suggested a possible uncoating after nuclear entry. Yet, the authors could not explain how the large capsid could be transferred through smaller NPCs. Given the lack of a convincing explanation for nuclear transfer of the core, the capsid was regarded as a shipping container needed only to penetrate the cell and dissociate from the subviral complex immediately after cytosolic entry (Figure 4 A).

The so-called ‘fate of capsid’ assay[98,99] offered the possibility to overcome previous problems with bulk analysis including many endosomal particles, which do not contribute to productive

infection[100]. This assay uses sedimentation to exclude the endosomal and lysosomal particles from the analysis[99] based on different densities of complete virions and capsids[101]. This biochemical approach used different densities also to distinguish soluble monomeric CA from assembled cores in the cytosol by analyzing cell lysates on a sucrose gradient[99]. Several other studies utilizing this assay showed that the amount of assembled capsid in the cytosol was clearly underestimated in previous studies[70,95,96], detecting high amounts even several hours post-infection[99,102]. The assumption of late cytosolic uncoating was further supported by the cyclosporine A (CsA) washout assay, which indirectly monitored uncoating by measuring the sensitivity of HIV-1 to restriction by the TRIM-Cyp (tripartite motif-containing protein-cyclophilin) restriction factor[103], which binds to the capsid and degrades it. This restriction of the virus by TRIM-Cyp can be rescued by CsA. By counting the number of infected cells after removing CsA at different timepoints post-infection, the kinetics of uncoating could be monitored. The observation of a gradual sensitivity loss was equated to gradual uncoating in the cytoplasm[103,104].

Implementation of more advanced imaging techniques allowed measurements of the capsid integrity by following an internal GFP signal produced by the incorporation of the GFP sequence into the Gag polyprotein[105]. This internal GFP was used as a fluid phase content marker trapped within the capsid[106]. In this study, the GFP signal outside the capsid was lost upon fusion at the plasma membrane. In a second step inner GFP was released rapidly inside the cytoplasmic region ~ 30 minutes after membrane fusion indicating the first capsid opening large enough to release GFP. By using low viral doses, the authors concluded that only particles that lost the signal early led to productive infection of a cell. Together with decreasing CA IF intensity measurements of subviral complexes found in the cytoplasm, these data suggested that early uncoating is required for infection. However, this indirect marker was not able to monitor uncoating since it is not suitable to distinguish between capsid integrity loss and complete disassembly of the core. This biphasic signal loss indicating uncoating was already proposed some years earlier by viral RNA labeling and confocal imaging[107]. This study used the RNA signal measurement and CA mutants with different stabilities. Since the staining relies on RNA accessibility to a small dye, they suggested an initial core opening followed by complete disassembly. The striking difference between these imaging techniques following capsids during cell infection and the earlier mentioned *in-vitro* assays was the presence of cellular factors creating a unique capsid stabilizing environment. The failure of previous studies to retain and detect capsid structures from isolated cores can be explained by recent findings that the capsid is a metastable structure, stabilized by cellular cofactors such as inositol hexakisphosphate (IP6) binding to a central pore of the CA lattice[108–110]. Addition of IP6 to isolated

capsids led to a significantly increased half-life of the cores to up to 10 h[108,111]. Therefore, the hypothesis about uncoating immediately after cell entry was not reliable, and only cell-based assays commonly suggested the model of gradual cytosolic uncoating.

Although the model of cytosolic dissociation of the CA lattice was supported, this uncoating model was not consistent with the observation that CA plays a crucial role in the infection of non-dividing cells[112,113] or in the protection of the viral genome during intracellular trafficking[114–116]. This discrepancy seemed to be why several research groups started intensive research on uncoating using mainly imaging approaches to detect CA signals. Using better microscopy techniques, various studies reported high CA amounts associated with subviral complexes in the cytosol and near the NE[117–121].

Accordingly, some of these studies, using new fluorescent markers to detect capsid or electron microscopy (EM), suggested the model of uncoating at the NPC, when significant loss of CA was observed to occur shortly before nuclear entry (Figure 4 B)[58,117,119–121]. In line with this model, the CA mutation N74D provided resistance to an inhibiting effect of different Nup depletions (Nup153, Nup160), which usually resulted in lower infectivity of WT HIV-1[122]. It was also found that Nup153 and Nup98 interact with the CA lattice[123–125] binding CA hexamers or pentamers primarily through the FG repeats[57,125]. These results strongly indicated the presence of at least partially intact capsids at the NE. Further support for this theory came from EM images of a NE-associated cone-shaped capsid[120], crystal structures of CA-Nup153 interaction[126], CA detection by IF in fixed cells[117], or loss of a labeled CA-binding CypA protein[119] all in proximity to the NE. This led to the assumption that the capsid, measuring a ~ 60 nm diameter at the wide end[127], dissociates upon docking at the NE to pass through the ~43 nm-wide NPC[128].

Eventually, the detection of nuclear CA-signals led to a third model. Although the capsid diameter would exceed the NPC's diameter and no explicit explanation for capsid passaging was described, several studies reported various amounts of nuclear CA associated with subviral complexes with strong IF signals in macrophages[62,63,85,89,129] but lower or no signal in HeLa-derived cells[62,130] or T cells[131]. The CA appeared to further impact the integration position in the nucleus[52,132] and interact with the nuclear host factor CPSF6[52,133,134]. Recently published data even indicated the importance of the at least partially intact CA lattice inside the nucleus for successful integration[59–61]. Time-of-addition experiments with CA lattice binding molecules[61], nuclear import blockage[59], and nuclear CA detection by IF in HeLa and T cells[64] or a direct CA labeling approach[60] strongly argued for the presence of CA in the nucleus implying the model of nuclear uncoating[59–61,64,135] (Figure 4 C).

Very recently, Zila *et al.* used 3D correlative light electron microscopy (CLEM) and cryo-electron tomography (cryo-ET) to show that cone-shaped structures were translocated into the nucleus of infected T cells[136]. This strategy revealed that the diameter of the NPC *in situ* is larger than previously assumed[128]. They reported that intact NPCs (~ 64 nm) could fit a cone-shaped capsid (~60 nm at the wide end), which had groundbreaking importance for the field, by explaining the appearance of nuclear CA by passaging through intact NPCs. This finding was supported by another study available at the same time showing that the NPC diameter is regulated by the mechanical status of the nuclear membrane and can be enlarged by pulling forces[137].

Nevertheless, the contentious debate whether the capsid passes into the nucleus or not and to which extent is still ongoing. Although high-resolution imaging provided good evidence for translocation of capsids into the nucleus[60,64], other recently published literature argued against it[58,85,121]. In one study, the authors proposed that the detection of nuclear CA signal in monocyte-derived macrophages (MDM)[62,63,131,138] only arose from nuclear clustering of several capsid fragments but not due to the transfer of intact cores[85]. The same authors reported that the phenotype of uncoating at the NPC observed in HeLa-derived cells[118] was conserved in primary cells[58]. Inconsistent conclusions could, among other reasons, be explained by the limitation of only having indirect techniques to follow incoming capsids. Since the CA is highly fragile and modifications of the protein rarely resulted in normal virus production[139,140], a direct labeling of the protein leading to infectious and morphologically intact particles is urgently needed.

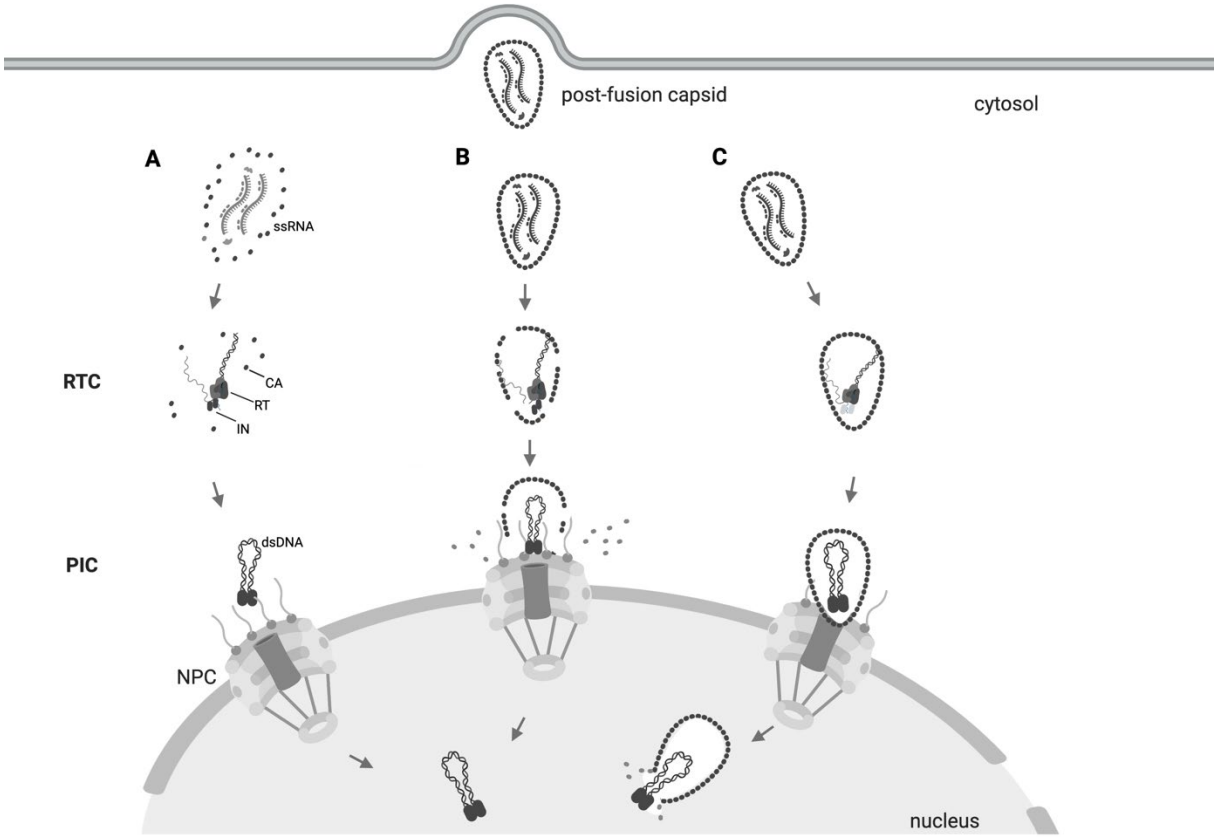


Figure 4. Capsid uncoating models. Different hypotheses about the cellular location of capsid disassembly. **A.** The capsid dissociates early in the cytosol (immediately) after cell penetration and only the PIC without the lattice is transported into the nucleus. **B.** The capsid stays largely intact throughout cytosolic trafficking and dissociates at the NE shortly before transferring the PIC through NPCs into the nucleus. **C.** The capsid core structure is preserved throughout trafficking and nuclear translocation. CA can be detected inside the nucleus of infected cells. ssRNA: single-stranded RNA; RTC: reverse transcription complex; PIC: pre-integration complex; dsDNA: double-stranded DNA; CA: capsid protein; RT: reverse transcriptase; IN: integrase; NPC: nuclear pore complex. The figure was created with BioRender.com.

1.4 Labeling strategies of HIV-1 CA to follow uncoating dynamics

The main obstacles to reaching a standard definition of capsid uncoating are technical limitations, the use of different experimental approaches, and the lack of a directly tagged CA.

Since the 24 kDa HIV-1 CA is building a highly ordered viral capsid structure consisting of hexamers and pentamers in a defined arrangement (Figure 5), it presents a highly challenging protein to target for adding a fluorescent label. It is not only involved during structural rearrangements from the immature to the mature Gag shell, but its outer surface constitutes an interface for protein-protein interactions with host factors essential for various steps of the life cycle[141–144]. Even the exchange of single amino acids in extensive screenings of this highly fragile protein disturbed the viral replication caused by, e.g., aberrant particle maturation and/or a severely compromised ability to

infect cells[139,140]. Accordingly, several attempts to incorporate smaller peptide tags into the CA were only achievable upon complementation with wild-type (WT) CA or resulted in severely impaired infectivity[121,145,146].

Pereira *et al*. used the biarsenical-tetracysteine labeling system[147] incorporating a 6-12 amino acid large tetracysteine motif (TC tag) into the N- and C-terminal CA sequence[145]. This modification, as well as incorporation of this tag into the CypA loop of CA[146], led to diminished reverse transcription and infectivity. Only the complementation with WT virus in equimolar amounts resulted in CA-TC or TC-CA tagged particles with decreased but still detectable infectivity[121,146]. However, this complementation limits the system due to large variations in the amount of incorporated labeled CA making reliable conclusions about CA loss impossible. Due to this technical restriction and the fragility of the CA not tolerating major changes, most studies addressing uncoating dynamics (*cf.* 1.3.3) had to use indirect approaches[59,61,103,106,111,118,119,135,148,149] or IF staining to visualize CA[62–64,117,120,130,150,151].

Various studies used the GFP as a fluid phase content marker to follow incoming particles measuring intensities of capsid encapsulated GFP[106,111,135]. Whereas this approach initially argued for early uncoating[106], a recent publication using GFP in combination with CPSF6 imaging convincingly suggested the opposite by finding capsid inside the nucleus[135]. The imaging approach presented by Francis *et al*. uses a fluorescently tagged version of the CA-binding host protein CypA (CypA-DsRed)[118]. Tracking of the CypA-DsRed signal indicated CA loss at the NPC and no (or only low amounts) CA entering the nucleus[118,119]. The fact that different or even the same approach was used to conclude different uncoating scenarios might raise doubts regarding suitability of these strategies to draw conclusions about the subcellular location of capsid uncoating. One reason is that the GFP fluid phase content marker cannot distinguish between initial capsid opening and complete disassembly. Furthermore, the tagged CypA-DsRed cannot exclude that the signal loss is only caused by CypA replacement from capsid or removal during NPC passaging. In both cases, signal loss would not necessarily describe a major uncoating event.

Later publications using different indirect methods to follow uncoating, e.g., imaging of CA-binding CPSF6[61] or measurement of PF-3450074 (PF74) sensitivity[59], showed increasing evidence for higher than previously assumed amounts of capsid inside the nucleus[59,61,135].

However, consistent detection of CA IF intensities inside the nucleus in all cell types was still lacking. Only strong and clear nuclear CA signals were detected in primary (monocyte-derived) macrophages[58,62,63,85,129,131,150], apparently indicating cell type-dependent differences. Based on the

assumption that the HIV-1 capsid was too large to pass through NPCs and enter the nucleus, it was not considered that IF staining failed to detect CA only in the nuclei of infected HeLa-derived or T cells[62,103,117,118,131]. But first doubts about cell type-dependent differences were already raised after detecting weak nuclear CA IF signals also in other cell types[62,130,150].

Therefore, the development of virions containing directly labeled CA in 2020, which were still functional and infectious[60,121] can help to better understand uncoating dynamics in different cell types. Burdick *et al*. used a GFP-CA fusion protein, which was suitable to produce CA tagged particles with normal morphology and only ~2-fold reduced infectivity[60]. Extensive microscopic imaging suggested the transfer of capsid into the nucleus based on strong nuclear GFP-CA signals. However, this technique comes with a significant drawback of complementation with a large amount of unlabeled WT virus in a 1:15 plasmid ratio. On the other hand, a concurrently published article using the same approach of a directly labeled CA version on the contrary argued for cytoplasmic uncoating as no nuclear CA was detected[121]. Their labeling method required complementation with an even larger excess of WT virus, resulting in only two GFP molecules per virion. Although these two studies offered a new possibility of directly labeled CA, their conclusions about the subcellular site of uncoating was still contrary. On the basis that IF staining and approaches using WT complementation are insufficient to give an unveiled overall concept about uncoating kinetics, further CA label development would be a desirable objective.

1.5 The HIV-1 capsid as the main interaction partner for host proteins

The cone-shaped capsid consists of roughly 1,500 single CA monomers, which assemble into approximately 250 hexamers[152] and 12 pentamers[153,154], forming the ~ 120 nm long and ~ 60 nm wide capsid shell[127] (Figure 5a). Incorporating a specific number of pentamers into the core is crucial to allow bending and curvature at the broad and narrow end of the capsid[141]. Immature virions incorporate ~ 2,400 Gag molecules[155], but not all are needed for capsid formation[152,156]. The remaining CA molecules may be encapsulated into the core or evenly distributed outside the capsid surrounded by the Env lattice.

The conical capsid composed of assembled mature hexameric and pentameric structures (Figure 5b,c) appears to be a main viral element that interacts with multiple host factors[157–159] (*cf*. Figure 2). Assembled capsomers are not only binding partners for infection-promoting proteins (Sec24C[57], Nup153[63,126], CPSF6[126,134]) and host cell restriction factors (e.g., Trim5α [160], MxB[159],

TrimCyp[158,161]) but also for small molecules involved in stability (PF74[126,162], IP6[108,163], GS-CA1[164]). As mentioned already earlier (*cf*. 1.3.4), the capsid was initially considered to be an instable complex. Over the last years, more and more studies attempted to identify the multiple distinct roles of the ordered capsid structures, increasingly arguing for a more stable capsid, and underlining the importance of this evolutionary deliberated and complex viral element. For instance, the striking feature of a central opening in the assembled hexamer surrounded by an unexpected positively charged arginine R18 and lysine K25 ring (Figure 5d) has a distinct function in binding negatively charged small molecules[165], e.g., IP6[110]. IP6 was shown to be incorporated into immature particles *via* R18 and K25 binding to increase the assembly efficiency of immature virus-like particles *in vitro* and have a hexamer stabilizing effect[108,163]. If IP6 was not packaged into newly produced particles, maturation was hampered and resulted in premature uncoating[109,110,166]. It was also reported that IP6 had an even higher stabilizing effect on capsid pentamers[167], which could induce favored pentamer incorporation in the capsid leading to better curvature and stability[141]. Apart from IP6 binding for structural reasons, the arginine ring creates a central pore within the capsid that allows molecule passaging giving an explanation of how dNTPs can enter the still assembled capsid to promote viral DNA (vDNA) synthesis[110,165,168].

Besides the central pore within the capsid, other residues of the CA are also crucial for protein or ligand binding. Every single CA monomer consists of the CA C-terminal domain (CTD) and N-terminal domain (NTD) connected by a flexible linker (Figure 5e). On the outer surface, it exposes the cyclophilin A binging loop (CypA loop; residues 83-100) as a central structure to bind proteins essential for the infection, e.g., CypA[169], TRN1[54], and Nup358 (also known as RanBP2)[52,53,170]. Binding of cytoplasmic CypA to the loop and incorporation into virions[171,172] has been reported to modulate viral replication[52,102,173–175] by blocking tripartite motif-containing protein 5 (Trim5α) binding[61,176] or stabilizing the core[177,178]. Interestingly, it was shown that CypA recognizes a special capsid surface geometry[177] and capsid curvature affects the number of binding CypA molecules[179]. This means that CypA senses capsid curvature and fully assembled capsids are required for efficient CypA binding.

A mass spectrometry-based proteomics approach recently identified Sec24C protein, a component of the COPII complex, to interact with capsid[57]. While accompanying the hexameric capsid lattice on its way towards the nucleus, Sec24C seemed to interact with the virus since Sec24C downregulation affected viral core stability, nuclear import, and infectivity. Furthermore, Rebensburg *et al*.[57] showed that Sec24C contains a FG motif, which was also detected for distinct Nups (*cf*. 1.3.1), binding to a pocket between two adjacent capsid subunits (Figure 6a). The importance of FG motifs was already described earlier in the context of CPSF6[133] and Nup153[125] binding to the CA.

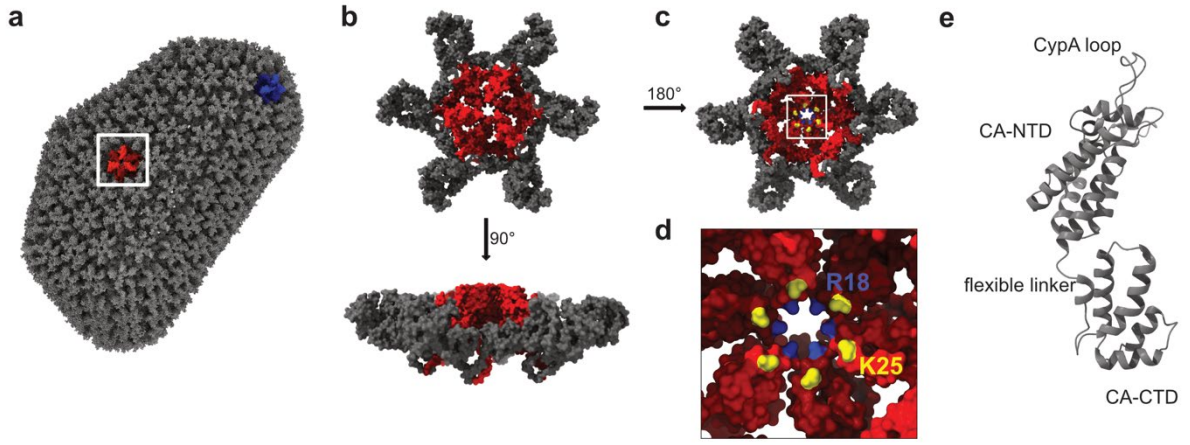


Figure 5. The HIV-1 capsid structure. a. Fullerene capsid core (pdb 3j3q) made of ~250 CA hexamers and 12 pentamers. One representative hexamer is colored in red; the representative pentamer is colored in blue. The structure was rendered in ChimeraX. **b.** Enlargement of the surface-rendered CA hexamer (pdb 5mcx) boxed in (c). upper panel: top view, lower panel: side view. **c.** Top view of the surface-rendered hexamer rotated by 180°. **d.** Enlargement of the central pore within the capsid boxed in (e) with color-coded residues R18 (blue) and K25 (yellow). **e.** CA monomer (pdb 5mcx) consisting of CA-NTD and CA-CTD connected by a flexible linker. The cyclophilin A (CypA) loop is exposed to the outer surface.

Considering former studies, this hydrophobic binding pocket at the CA-NTD involving helices 3 and 4 provides a platform for interactions with other proteins. After entering the NPC, Nup153 binds to the same hydrophobic pocket (Figure 6b) to facilitate core translocation into the nucleus[53,122,125]. An overlap between CA binding interfaces was also found for the CPSF6 binding later at the nucleoplasmic site[126,133,134] (Figure 6c). CPSF6 binds incoming cores to allow entry into the nuclear region[63], directs the capsid inside the nucleus[61,85,180], and induces viral integration into special domains[85,86,180]. Taken together all cellular interactions of this pocket, it can be hypothesized that the hydrophobic pocket in the assembled capsid is a key actor in a relay sequence, with consecutive binding of the same motif by different proteins to mediate transport to and import into the nucleus. Therefore, the capsid constitutes a decisive component during early HIV-1 replication[19,44–46,92,181].

1.6 Capsid-binding small molecules

Due to its multiple host factor interactions, the capsid represents an interesting target for antiviral compounds. Several molecules were identified to bind HIV-1 CA with antiviral activity[182,183] (e.g., BI-2[184,185], CAP-1[186], PF74[187], GS-6207[188,189]) of which most target the exact same motif in the capsid as shown for Sec24C, Nup153, and CPSF6 (*cf.* 1.5). The prototype of this class is the phenylalanine-

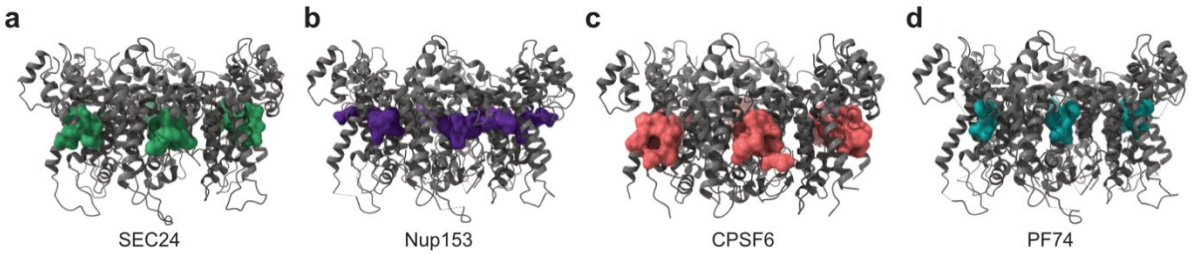


Figure 6. Binding pocket for host proteins and ligands within the hexameric lattice. All proteins bind a hydrophobic cleft between two CA subunits. **a.** Binding of Sec24C (green) during cytosolic trafficking (pdb 6pu1). **b.** Nup153 (purple) binding at the same pocket at the NPC (pdb 4u0c). **c.** CPSF6 (indian red) binding after Nup153 to release capsid into the nucleus (pdb 4u0a). **d.** The CA binding small molecule PF74 (dark cyan) binds to the same site and can replace CPSF6 (pdb 4u0e). All CA hexamer structures are illustrated in cartoon representation (grey) and bound proteins/ligands are surface rendered and color-coded using ChimeraX.

derived peptidomimetic PF74[187], which binds the CA within a binding pocket involving the CA-NTD helices H3, 4, 5 and 7 and CA-CTD helices H8 and 9 from a neighboring capsid subunit[126,134] (Figure 6d). This small molecule was shown to influence viral replication in a bimodal and concentration-dependent mode[190,191]. Low PF74 concentrations (2 µM) inhibited the infection without affecting reverse transcription and formation of the PIC[62], whereas higher concentrations (> 10 µM) blocked reverse transcription and were reported to have destabilizing effects on mature capsids[60,126,187,192]. Decreased infectivity upon PF74 treatment was connected to preventing CA binding to cellular factors supporting nuclear entry[125,193]. Contrary to studies showing a destabilizing effect, no major changes in capsid intensity were reported based on retained CA intensities upon PF74 treatment[64]. Another study observed even increased CA signal intensities upon PF74 treatment[85], while another claimed PF74 had stabilizing effects on the capsid core[194]. Using the internal GFP marker for single-molecule microscopy[111] and atomic force microscopy [195] revealed that PF74 indeed accelerated capsid opening, but the remaining lattice was stabilized. Therefore, further investigation is needed to understand the actual effect of PF74 on the mature capsid and the mechanism by which it affects capsid stability.

1.7 The genetic code expansion

Since the capsid constitutes an important structure of HIV-1 assisting viral infection and existing fluorescence CA labeling techniques suffer from several technical limitations, other CA labeling strategies would be beneficial in order to produce highly labeled and infectious particles. The principle of genetic code expansion (GCE) offers the possibility of producing engineered proteins by site-specific incorporation of hundreds of non-standard amino acids exceeding the naturally

occurring ones, which enables the investigation of biological processes[196–200]. This minimally invasive method could be adapted to HIV-1 CA modifications and additional click-labeling would allow production of fluorescently labeled particles in a proviral context suitable for a wide range of microscopy techniques.

The natural system offers a limited number of 20 naturally encoded proteinogenic amino acids, which the translational machinery can translate from genetic information into a polypeptide (Figure 7a). The enzymatically active aminoacyl tRNA synthetase (aaRS) aminoacylates a specific tRNA with the appropriate amino acid. This tRNA recognizes a particular mRNA sequence and binds to it with its matching anticodon. The mRNA is translated with the help of a moving ribosome forming a polypeptide chain of different amino acids until it reaches a stop codon where the peptide is released. The conserved mechanism of gene expression translating information-carrying nucleic acids by ribosomes can be altered by artificially modifying the system.

Figure 7b shows a scheme of GCE in the context of amber suppression, which describes the modification of a protein after incorporating a amber stop codon[201]. For this purpose, a bioorthogonal set of an aaRS and tRNA is required, which together incorporate a non-canonical amino acid (ncAA) at the site of an artificially introduced stop (UAG). Instead of releasing the polypeptide from the ribosomal complex and protein folding, the ncAA is incorporated and translation proceeds until it reaches one of the other natural opal (UGA) or ochre (UAA) stop codons, which are not recognized by this specialized tRNA. This aaRS/tRNA pair, which does not crosstalk with other host components, allows protein modifications by changing only one amino acid.

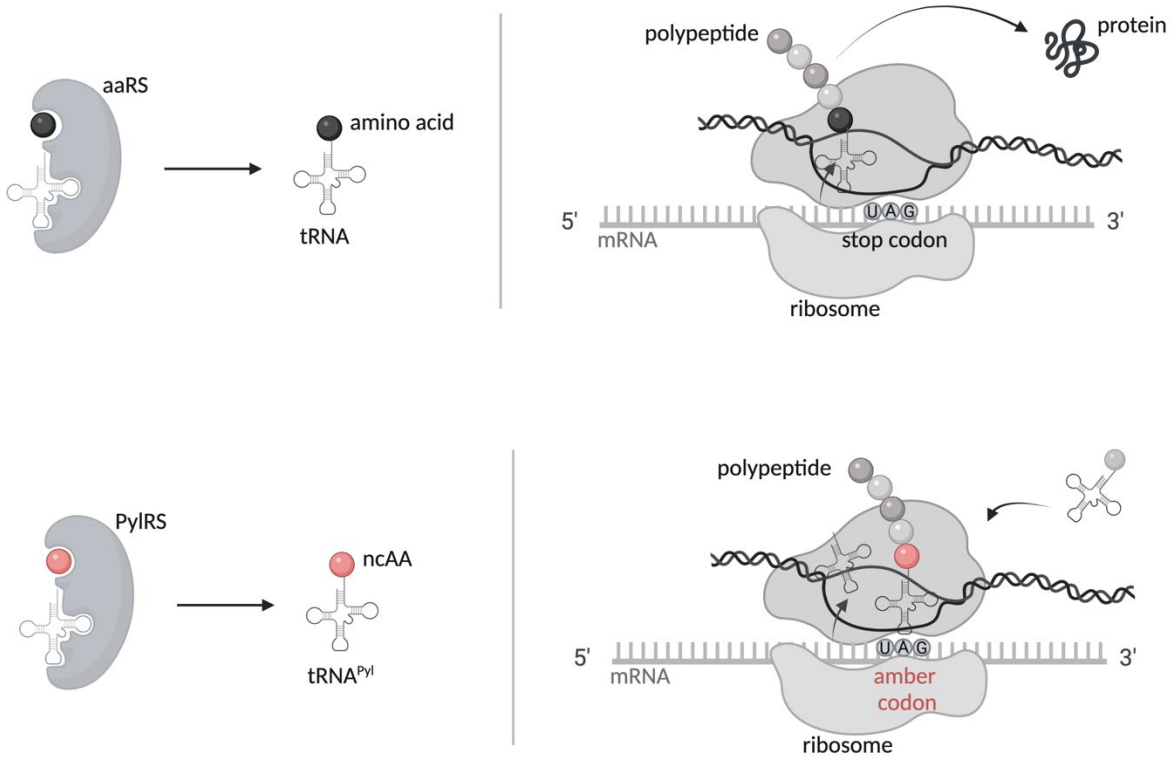


Figure 7. Genetic code expansion. Schematic scheme of translation. **a.** Aminoacyl tRNA synthetase (aaRS) binds tRNA and a natural amino acid. aaRS attaches the amino acid to the tRNA. The amino acid-loaded tRNA binds a specific sequence determined by its anticodon. The ribosome facilitates the attachment of the AA to the polypeptide chain. When the ribosome reaches a stop codon, which is not recognized by any host tRNA, the polypeptide is released and can form a protein. **b.** Ribosomal translation of mRNA containing an amber stop codon (UAG) in the presence of an orthogonal aaRS (pyrrolysine-tRNA synthetase; PylRS)-$tRNA^{Pyl}$-pair. The PylRS recognizes only $tRNA^{Pyl}$ and attaches a specific ncAA (pink). Instead of terminated translation and release of the polypeptide chain, the tRNA binds the amber stop codon and incorporates the ncAA. Figure was created with BioRender.com.

The efficiency of the GCE system on protein yield strongly depends on several factors and amber suppression competes with natural translational termination. Thus, the efficiency, usually reaching 10-100 %, can differ between different proteins, but various sites of the same protein can also differ. Throughout optimization of orthogonality by designing and directed evolution, the read-through of stop codons was significantly increased[202–204]. Accordingly, the selection of the best suitable orthogonal pair for each host is essential. Whereas some aaRS/tRNA pairs are orthogonal only in *E. coli* or eukaryotic cells, the widely used pyrrolysine-tRNA synthetase (PylRS)-$tRNA_{CUA}$ pair from certain methanogens is orthogonal in eukaryotic cells and bacteria[205].

GCE is a versatile tool for modifying expression in different hosts in order to understand biological processes[200]. GCE was already applied for animals such as *C. elegans*, fly cell culture[205], and reversible gene expression in mice[206,207].

1.7.1 Genetic code expansion via amber suppression

The UAG codon is one of the three non-sense codons indicating the termination of translation. By incorporating this amber codon, ncAAs can be incorporated by an altered tRNA and read-through *via* amber suppression[196]. By now, the ncAA incorporation using an amber stop codon is the mostly used technique to expand the genetic code in prokaryotic or eukaryotic cells[197,208–210], especially due to its lower abundance in *E.coli*.

This principle makes advantage of the existence of additional canonical amino acids, such as pyrrolysine as the 22nd natural amino acid[211–213]. This genetically encoded pyrrolysine is rarely used in translation and most likely evolved during adaptation of *Archaea* to new environments[214,215]. It is encoded by the amber stop codon UAG, which can be recognized by an orthogonal tRNA^{Pyl}/PylRS pair. This PylRS cannot recognize any other natural amino acids and by that offers the perfect tool for site-specific GCE. The versatility of this orthogonal system was shown by the successful incorporation of pyrrolysine analogues by amber suppression in *E. coli,* which only requires high affinity for the hydrophobic binding pocket of the PylRS[216].

Improvements of efficiency of the system were achieved by better designing the orthogonal pair sequence and expression[217–219], by predicting ncAA incorporation efficiency based on the target sequence[220], by optimization of elongation factors[221,222] and by altering release factors[223–226]. For example, the release factor 1 (RF1) in bacteria, normally terminating the synthesis at the ribosome, competes with the decoding tRNA at amber stop codons affecting the overall efficiency. Mutations of the 16S ribosomal RNA disturb the interaction of RF1 with the ribosome leading to higher suppression efficiency[227]. The same principle was adapted to eukaryotes providing a dominant negative version of the eukaryotic release factor 1 (eRF1), resulting in less competition with the tRNA and increase the incorporation of ncAAs[217]. Another recently described strategy created an artificial organelle within a cell to achieve a local enrichment of amber suppression components[228]. By building these phase-separated compartments where the ncAA incorporation occurs, the translation level reached almost half of the efficiency compared to the WT protein control due to less interference with the host machinery. A striking improvement was achieved by Nikic *et al*. leading to significantly less fluorescent background for imaging of amber suppressed and click-

labeled proteins in cells[229]. They identified a nuclear localization signal within the sequence of the PylRS, which dramatically decreased GCE efficiency and resulted in very strong nuclear nonspecific background. By adding a nuclear export signal to the PylRS N-terminus (resulting in NESPylRS), a cytoplasmic localization was induced increasing overall efficiency and applicability of the pyrrolysine GCE system for microscopic imaging.

The orthogonal $tRNA^{Pyl}$/PylRS pair allows protein modification in several different ways. A large variety of applications was reported[230] including the incorporation of fluorescent amino acids or environmental sensors and protein conjugation with drugs, PEGs, or other probes. One specific application involves chemically reactive amino acids, which can be incorporated into proteins to serve as reaction partners for electrophiles. A used technique in fixed cells include, e.g., the copper-catalyzed 1,3-dipolar cycloaddition of ncAAs containing an azide or alkynyl[231–233]. Reactions in living cells became possible via copper-free click reactions with strained alkenes[234–236]. The strain-promoted inverse electron-demand Diels-Alder cycloaddition (SPIEDAC) gained more and more importance offering a rapid and selective click reaction of proteins and biorthogonal reactive side chains[237], e.g., alkenes, alkynes, and tetrazines[235,238–242] (Appendix Figure 34). This SPIEDAC reaction in combination with an engineered PylRS variant was effectively used for site-specific biorthogonal fluorescent protein tagging[235,243], *in vivo* labeling[239,241,244,245] or protein cross-linking[209,246].

1.7.2 GCE applications in a viral context

GCE recently became a promising approach to test the development of vaccine candidates, which are conditionally replication-competent[247–249]. In this context, the incorporation of amber stop codons into the influenza hemagglutinin protein or HIV-1 PR and Gag[250] enables replication of viral proteins only upon addition of amber suppression components, which could trigger an immune response. The ability of the virus to reinstall sense codons instead of incorporated stop codons would require the incorporation of several stop codons within the genome. The incorporation of higher numbered stop codons[251], demanding a suitable amino acid and translation system, together with a system allowing multiple-round infection[252] could be a starting point for the successful development of an efficient viral vaccine.

In contrast to the attachment of larger tags, GCE provides the unique opportunity of a minimally invasive modification of viral proteins. Due to the compact genome of viruses and partially overlapping ORFs, fusion proteins rarely result in intact particles or desired protein labeling.

Although ncAAs can be incorporated throughout the whole genome by GCE, only sporadic applications of GCE in a viral context were reported[253].

The first study using GCE and amber suppression to site-specifically label viral proteins generated infectious and intact Hepatitis D virus particles click-labeled at outer surface proteins[254]. Another study targeted the Hepatitis C helicase, which was purified and site-specific labeled to investigate binding motifs and enzyme activity[255]. In the context of influenza, virus-like particles containing a TAG mutant of hemagglutinin were produced and individual filaments were imaged in super-resolution microscopy[256]. Also, amber suppression was used to produce the HIV-1 Env protein carrying a ncAA. Subsequent click-labeling allowed investigation of the Env distribution at the plasma membrane by stimulated emission depletion (STED) nanoscopy[257]. Others introduced the amber stop codon into the *env* sequence to generate click-labeled Env, which was used to observe different conformational states of the viral protein by single molecule fluorescence resonance energy transfer (smFRET) imaging[258].

Others applied the strategy to non-enveloped viruses such as adeno-associated viruses (AAVs) to site-specifically functionalize the capsid by PEGylation[259] or fluorescent labeling[260,261]. Besides monitoring the host cell infection of fluorescently click-labeled AAVs[261], the GCE and AAV capsid functionalization also enabled gene delivery into specific cell types and visualization of selective glioblastoma cell killing[260]. A notable aspect of the AAV CA in contrast to the highly fragile HIV-1 CA is that this protein accepts more extensive modifications and peptide insertions[262–265], making it a suitable candidate to tolerate changes of the protein sequence. So far, however, click-labeling by GCE of the CA in enveloped viruses was not reported.

1.8 Aim of this thesis

The discussion about HIV-1 uncoating dynamics and its subcellular location is influenced by various techniques used to follow incoming capsids. To a large degree, the understanding of the process relies on indirect approaches or bulk analyses, including biochemical CA measurements or imaging of incoming particles. The lack of a method to directly label the CA to allow live-cell microscopy and quantitative assessments about the amount of PIC-associated CA hindered more detailed insights into the uncoating process. To overcome these technical limitations the present thesis focused on two main objectives.

The first main objective was to identify a position within the CA sequence, which allows the incorporation of a ncAA and subsequent fluorescent labeling of infectious and morphologically

intact virions. The goal was to use GCE with an engineered $tRNA^{Pyl}$/$NESPylRS^{AF229}$ combined with a chemically reactive ncAA to attach a fluorescent dye by click-chemistry covalently. In addition to identifying a promising residue, this approach included establishing and optimizing the system to produce high particle yields and determine production and labeling parameters best suitable to follow highly labeled particles without disturbing the particle maturation or viral replication cycle. As a subgoal this objective included the production of a non-infectious CA-labeled construct that still allows for reverse transcription to perform live-cell imaging outside of the BSL-3 area. For this purpose, an integrase-deficient plasmid with a *tat* deletion provided by Thorsten Müller (University Hospital Heidelberg) was used[136]. The established direct CA-labeling was combined with additional staining of the viral IN, which enabled the detection and tracking of double-labeled particles.

The second main objective of this thesis was to validate the directly labeled CA variant and determine infection kinetics in different cell types. By having the first system without WT complementation, this study provides a reliable quantitative analysis of CA content in the cytoplasm and inside the nucleus of infected cells for the first time. The minimally invasive labeling of the fragile HIV-1 CA is a versatile tool for live-cell imaging and tracking of incoming capsids also allowing the application of super-resolution and correlative microscopy. This system was applied to detect CA signals and elucidate the mechanism of capsid uncoating in different cell types.

2 Results

2.1 Establishment of GCE by amber suppression of HIV-1

2.1.1 CA click-labeling using GCE and the amber suppression system

To establish a new direct labeling method for the HIV-1 CA, GCE and site-specific labeling were tested to achieve highly labeled particles suitable for a wide spectrum of microscopy imaging techniques. With the help of GCE, the ncAA cyclopropane-L-lysine (CpK) was site-specifically incorporated at the introduced amber codon in the presence of a bacteria-derived orthogonal tRNA/tRNA synthetase pair (Figure 8a). The CA carrying one single ncAA at a defined position can then be subjected to SPIEDAC click-reaction with a suitable fluorescent dye like silicon rhodamine (SiR) coupled to tetrazine (Tet-SiR; Figure 8b, c).

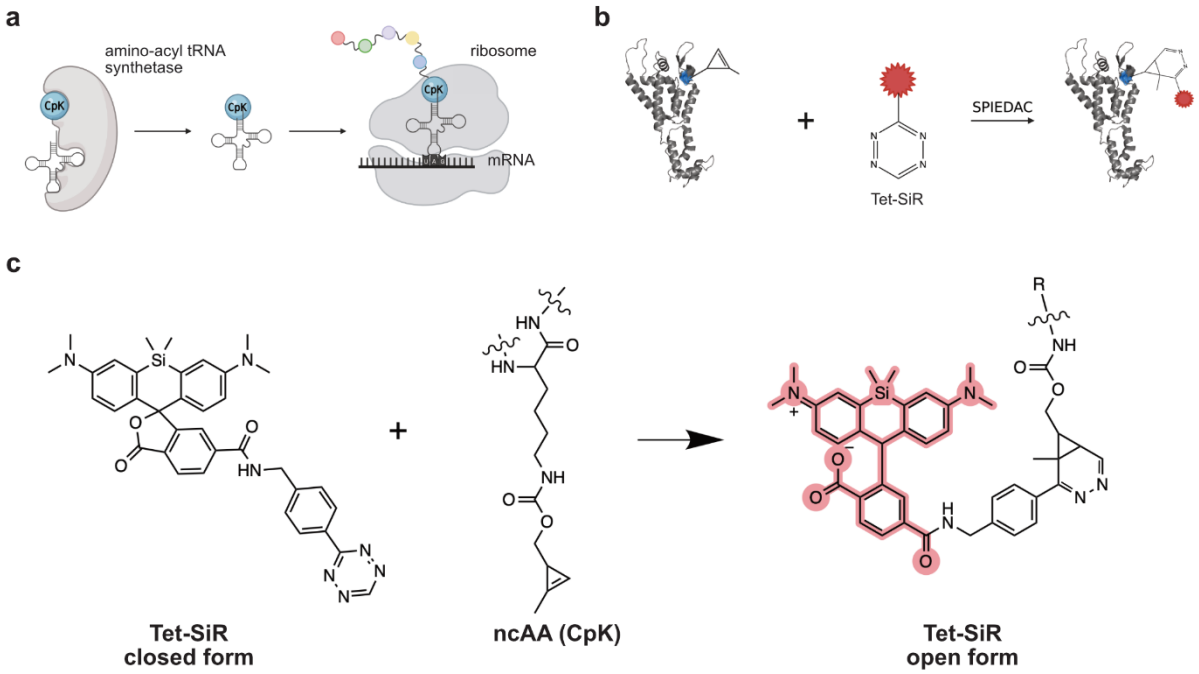


Figure 8. Site-directed CA labeling via amber codon suppression and click-labeling. Schematic representation of the two-step process of GCE followed by click-labeling. **a.** The first step is the incorporation of a ncAA cyclopropene-L-lysine (CpK) at an artificially introduced amber stop codon (UAG) with the help a bacteria-derived bioorthogonal engineered tRNA/aminoacyl tRNA synthetase pair. **b.** The reactive group of CpK is covalently ligated to a fluorophore (e.g., silicon rhodamine, SiR) in a second step. The strain-promoted inverse electron-demand Diels-Alder [4+2] cycloaddition (SPIEDAC) between the cyclopropene group of CpK and a tetrazine-conjugated fluorophore is employed. C. Scheme of SPIEDAC of a tetrazine-derivative of silicon rhodamine (Tet-SiR; electron-poor diene) and the ncAA CpK (electron-rich dienophile). The highly reactive tetrazine group reacts with the strained alkene of CpK via SPIEDAC and the open SiR conformation results in fluorescence. Created with BioRender.com. Modified from Schifferdecker *et al.,* 2021[266].

2.1.2 Generation of a modified proviral HIV-1 plasmid suitable for protein-specific labeling by GCE

To ensure site-specific and unique click-labeling of CA within the virion, the proviral plasmid used for particle production required modification. HIV-1 Vpr is incorporated into virions in a ratio of 1:7 compared to Gag molecules via its interactions with the p6 domain and associates with mature viral cores[152,267,268]. Since the *vpr* ORF of the used isolate HIV-1 NL4-3 naturally ends with an amber stop codon (TAG), the sequence was changed to an opal stop codon (TGA) (resulting in HIV-1*) to avoid unwanted labeling of Vpr by the GCE system. To test if the modification in the *vpr* gene has an impact on particle yield compared to HIV-1 WT, the RT activity of mature particles was measured using SYBR green I-based real-time enhanced reverse transcriptase assay (SG-PERT)[269]. The amount of virus associated RT activity is a surrogate measure of the particle number in the supernatant of virus-

producing cells or a concentrated particle solution. The RT is a product of the Gag-Pol polyprotein resulting from a frameshift readthrough in the p6 region. Since Gag-Pol is taking part in the assembly process, the cleaved RT is a component of completely assembled virions incorporated into virions in a ratio of 1:20 to Gag. In the context of HIV-1 protein modifications by the GCE system, RT can only be produced after successfully suppressing the incorporated amber codon in the upstream CA coding region, indicating the yield of modified particles. Analysis of SG-PERT measurements showed that particle yields of HIV-1 NLC4-3 (8.45 ± 3.84 x10^8 pUnits RT/µl) and HIV-1* (5.89 ± 3.03 x 10^8 pU RT/µl) were comparable (Figure 9a). Importantly, the single round infectivity was not altered for the modified proviral construct with relative infectivity of 94.01 ± 34.81 % compared to NLC4-3 (Figure 9b). These data indicate that the modified *vpr* stop codon has no unintended negative effect on particle assembly, release, or infectivity of the proviral construct and can therefore be used as backbone for the introduction of amber codons for GCE within the CA.

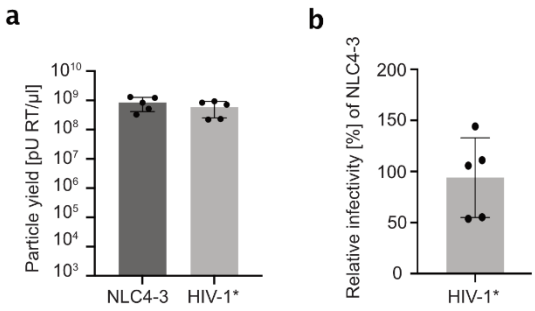


Figure 9. Characterization of the proviral HIV-1* plasmid for site-specific labeling by GCE. To allow specific and unique click-labeling of CA, the *vpr* stop codon TGA in the proviral pNLC4-3 plasmid was exchanged by an opal codon TGA. The resulting plasmid was named HIV-1*. Particles were concentrated from the supernatant of virus-producing HEK293T cells transfected with either pNLC4-3 or HIV-1*. At 48 hours post-transfection (h p.t.), the supernatant of the cells was harvested and filtered. **a.** Particle yield was determined and compared via quantitation of RT activity by the SG-PERT assay[269]. **b.** The relative infectivity of HIV-1* was determined in TZM-bl indicator cells by a luciferase assay and normalized by RT activity. The values obtained for HIV-1* were normalized to HIV-1 NLC4-3. The graphs show mean values and SD from five replicates performed in three independent experiments. Modified from Schifferdecker *et al*. (in preparation)[270].

2.1.3 Selection of candidate residues for genetic code expansion

In order to identify suitable sites for amino acid substitutions, screening experiments were performed. The selection of candidate residues (Figure 10a, b) was mainly based on structural information about the HIV-1 CA monomer and the immature (pdb: 4usn) and mature hexameric/pentameric structure (pdb: 5mcx, 5mcy). The HIV mutation database (https://www.hivmut.org) was used to obtain functional information on existing virus mutants. The CA structure was used to exclude sites, which were part of the immature interface[271] important for

Gag assembly or hexamerization interface in the mature lattice since these residues can be crucial for viability and the fitness of viral particles[272]. The selection was limited to sites on the outer surface of the mature CA hexamer to ensure accessibility for click-labeling (Figure 10c).

Furthermore, only amino acids of the NTD (aa 1-145) were selected in order to avoid a dominant negative effect of truncated proteins on the assembly process[140,273–275]. Additionally, the impact of potentially disruptive truncated CA forms incorporated into the virions was also reduced by targeting the NTD. These considerations resulted in a subset of 18 codons, which were replaced by amber codons in the context of the proviral plasmid pNLC4-3.

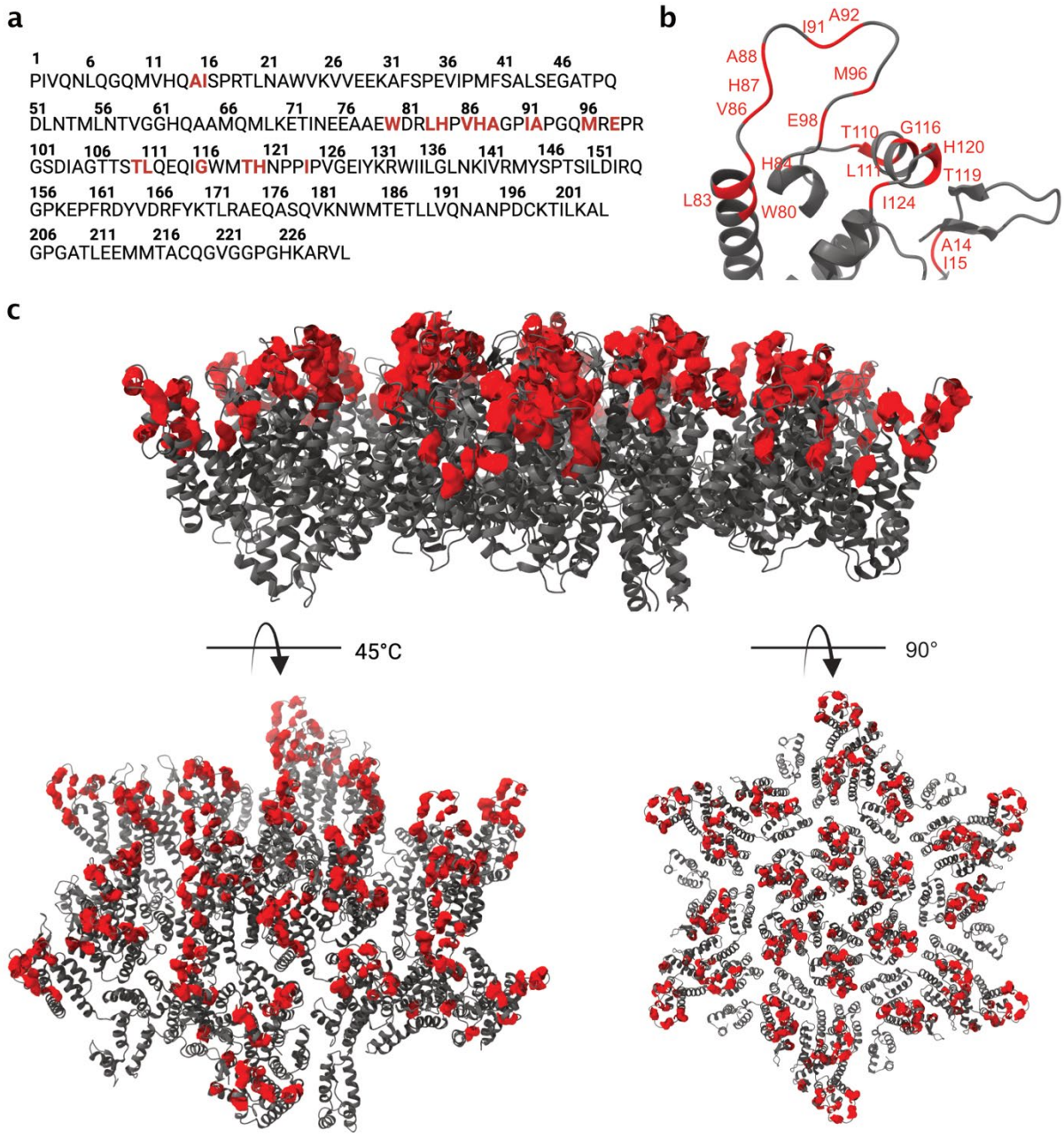


Figure 10. HIV-1 CA structure with selected residues highlighted. 18 different amino acid residues were selected to incorporate a ncAA in the presence of the amber suppression system and subsequent click-labeling. **a.** The CA sequence with amino acids residues 1-231 (uniport: P12493). Residues mutated and targeted for click-labeling in this study were highlighted in red. **b.** NTD of the CA monomer structure (grey) with targeted residues highlighted in red (pdb: 5mcx). **c.** CA hexamer side view, angled (45°), and top view (90°) (pdb: 5mcx) in cartoon representation (grey) with selected residues rendered in surface representation (red). Protein structures were illustrated with UCSF ChimeraX [276,277]. Figure from Schifferdecker *et al*. (in preparation)[270].

2.1.4 Optimization of GCE-altered particle production with a modified expression system

For particle production of GCE click-labeled virus, the transfection of cells required a plasmid encoding the HIV-1 coding sequence and another plasmid delivering the components for the amber suppression system. Initial experiments were performed in a system co-transfecting HIV-1 proviral plasmid pNLC4-3 and two additional plasmids carrying the amber suppression system. The required aminoacyl tRNA synthetase and cognate tRNA were delivered on the plasmid ptRNA^{Pyl}/NESPylRS^{AF229} that was optimized with a nuclear export signal (NES) to increase expression yields and decrease the untargeted background. The efficiency of ncAA incorporation can be improved by co-expression of a dominant-negative form (E55D) of eRF1, which alleviates competition of the amber suppression system with eRF1 and translation termination[217]. In this specific context, protein yields were reported to be enhanced by 17-20-fold, which reflected 43 % efficiency compared to a no amber stop control[217]. Sakin *et al.* had observed a positive effect of this eRF1 mutant in the context of GCE of the HIV-1 Env protein[257], which was therefore included in the GCE system for CA labeling.

To simplify the transfection setup, GCE components were combined into a single plasmid based on the multi-cassette pEA168 plasmid[278]. The final plasmid pNESPylRS-eRF1dn-tRNA (provided by Anna-Lena Schäfer, Bachelor thesis, University Hospital Heidelberg) contained the NESPylRS^{AF} driven by the cytomegalovirus (CMV) promoter, the coding sequence for eRF1 E55D under EF1α control, and four cognate tRNA expression cassettes (Appendix Figure 35). Comparing the particle yield of the new 2-plasmid system with the 3-plasmid system revealed no clear differences using four representative HIV-1*CAxx^{TAG} variants (Figure 11a). Next, the optimal transfection molar ratio of pNESPylRS-eRF1dn-tRNA to the plasmid encoding the HIV-1 coding sequence was identified. For the variant HIV-1*CA14^{TAG}, the highest particle yield was measured at a molar 2.2:1 ratio of HIV-1*CA14^{TAG} and pNESPylRS-eRF1dn-tRNA (Figure 11b). Furthermore, the addition of 100 µM ascorbic acid to transfected cells increased the particle production, presumably due to antioxidant properties and protection of the cells from free radicals caused by the addition of CpK (Figure 11c).

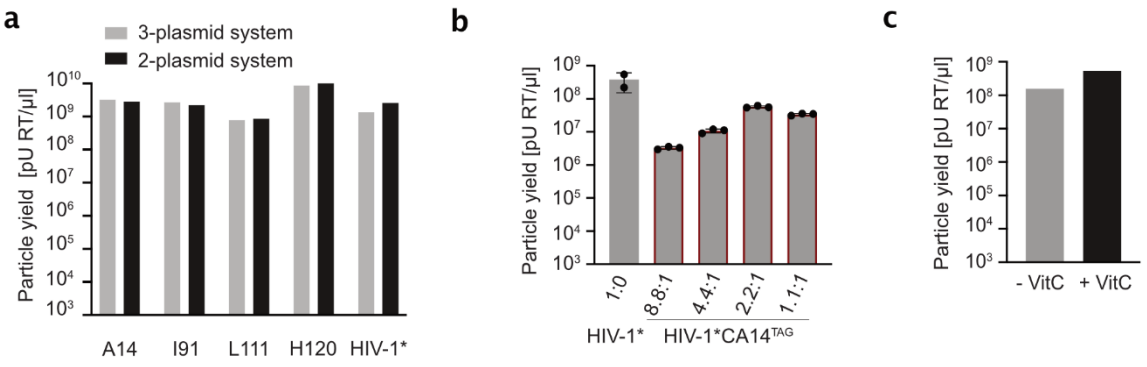


Figure 11. Optimization of the transfection system. Transfections were performed using the proviral plasmid HIV-1* or an HIV-1*CAxx^{TAG} variant (A14, I91, L111, H120) together with either the old system using two plasmids carrying the GCE components (3-plasmid system, light grey) or one single GCE plasmid (2-plasmid system, dark grey). Co-transfecting three plasmids was performed in a 2.4:1:1 ratio of HIV-1*/HIV-1* CAxx^{TAG} variant, ptRNA^{Pyl}/NESPylRS^{AF} and eRF1-E55D. Co-transfecting two plasmids was performed in a 2:1 ratio of HIV-1*/HIV-1* CAxx^{TAG} variant and pNESPylRS-eRF1dn-tRNA. The supernatant was harvested, and particle yield was determined via RT activity measurement by the SG-PERT assay in a single experiment. Experiment performed by Volkan Sakin (University Hospital Heidelberg). **b.** Particle yield determined via SG-PERT analysis of HIV-1* or HIV-1*CA14^{TAG} co-transfected with pNESPylRS-eRF1dn-tRNA at different molar ratios. The graph shows mean values and SD from three technical replicates. **c.** Particle yield of HIV-1*CA14^{CpK} virus produced in cells, with (light grey) or without (dark grey) the addition of 100 µM ascorbic acid (vitamin C) during transfection, was determined via SG-PERT analysis. Experiment was performed by Volkan Sakin (University Hospital Heidelberg). Modified from Schifferdecker *et al*. (in preparation)[270].

2.2 Screening of CA residues targeted for direct click-labeling by GCE

2.2.1 Determination of amber suppression efficiency

The efficiency of amber suppression is not only determined by the nature of the suppression system and the host cell, but also depends on the sequence context of the introduced TAG codon. Furthermore, different sites can have different propensities for incorporation of a natural amino acid instead of a ncAA. Therefore the 18 CAxx^{TAG} variant candidates were screened in presence and absence of CpK.

Transfected cells were lysed at 48 hours post transfection (h p.t.), separated by SDS-PAGE and proteins transferred onto membranes were stained with an antibody raised against HIV-1 MA. Previously reported amber suppression expression systems reported ncAA incorporation efficiencies between 25 - 50 % of WT protein yields in eukaryotic cells[217,257,279]. By intensity measurements of truncated protein and full-length Gag bands of cell lysates in the presence of CpK,

the efficiency of amber suppression in the context of HIV-1 CA was estimated (Figure 12 and Appendix Table 8). The efficiency was defined as the percentage of full-length Gag relative to the amount of full-length and truncated Gag, which was determined in a range of ~ 2-26 %.

In a second step, immunoblot analysis was used to compare the amount of full-length Gag produced after transfecting HIV-1*CAxx^{TAG} variants in the presence or absence of CpK. CpK-independent expression or codon skipping became visible for several variants by detecting full-length Gag at the size of 55 kDa. HIV-1*CAxx^{TAG} variants I15, W80, L83, H84, V86, T110, G116, and T119 expressed Gag at detectable levels in the absence of CpK arguing for site-specific leakiness of the amber suppression system. Note that Gag-derived MA was also detected in samples grown in the absence of CpK, which could be a contamination from immature or mature virus particles remained within the cell lysates.

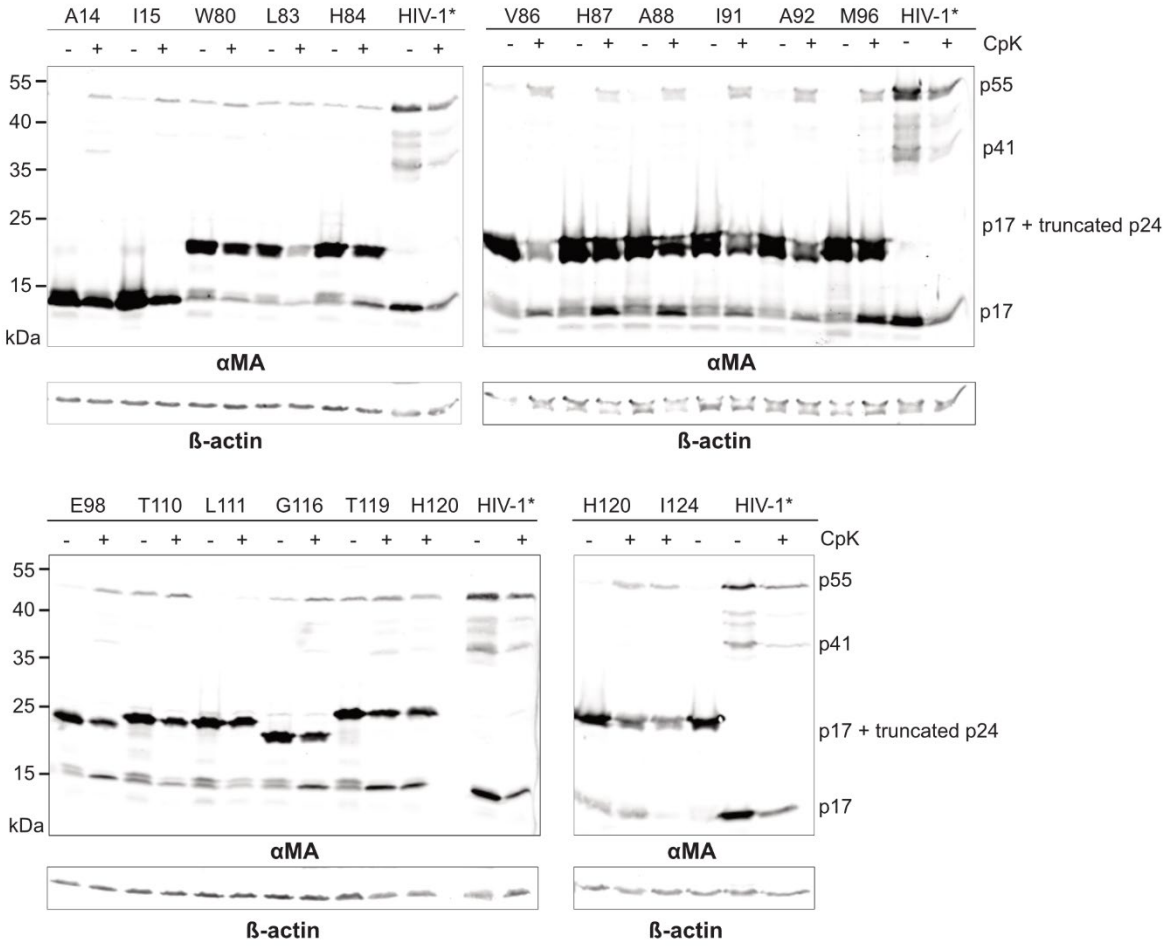


Figure 12. Screening of 18 HIV-1* CA^{TAG} variants to detect Gag and determine amber suppression efficiency. HEK293T cells were co-transfected with pNESPylRS-eRF1dn-tRNA and a plasmid encoding the coding sequence of one of the CA^{TAG} variants HIV-1*CAxx^{TAG} or HIV-1* in the absence or presence of 500 µM CpK. After harvesting the supernatant of transfected cells, cells were lysed and subjected to immunoblot analysis. Proteins were stained with a polyclonal rabbit antiserum raised against recombinant HIV-1 MA and

the housekeeping protein ß-actin. Bound polyclonal Alexa Fluor secondary antibodies were detected and quantified using a Li-COR Odyssey infrared scanner and the Li-COR software Image Studio lite 5.0. Immunoblots show specific bands for Gag-derived proteins MA (p17), truncated Gag (p17+truncated p24), Gag-intermediate product (p41), and full-length Gag protein (p55). Note that truncated MA-CA14 (~16.32 kDa) and MA-I15 (~16.39 kDa) proteins cannot be distinguished from mature MA (~14.8 kDa). The percentage of full-length Gag was calculated relative to the total reactive bands of full-length and truncated Gag (Appendix Table 8). From Schifferdecker *et al*. (in preparation)[270].

2.2.2 Quantification of particle yield and ncAA-independent Gag expression

Next, the proteolytic processing of the ncAA-containing CA variants was analyzed, which can serve as a marker indicating viral particle production. The supernatant of transfected cells, grown in the presence of CpK, was harvested and filtered. Particles were concentrated by ultracentrifugation through a sucrose cushion. Viral particles were then lysed and subjected to immunoblotting. A polyclonal antiserum raised against HIV-1 CA was used to semi-quantitatively analyze processed CA of the amber suppressed constructs. As shown in figure 13a, fully mature CA was released in detectable amounts for most variants with concentrations reaching from ~ 0.1-0.6 ng/µl. HIV-1*CAxx^{CpK} variants I15, W80, L111 and I124 yielded lower CA concentrations below 0.1 ng/µl.

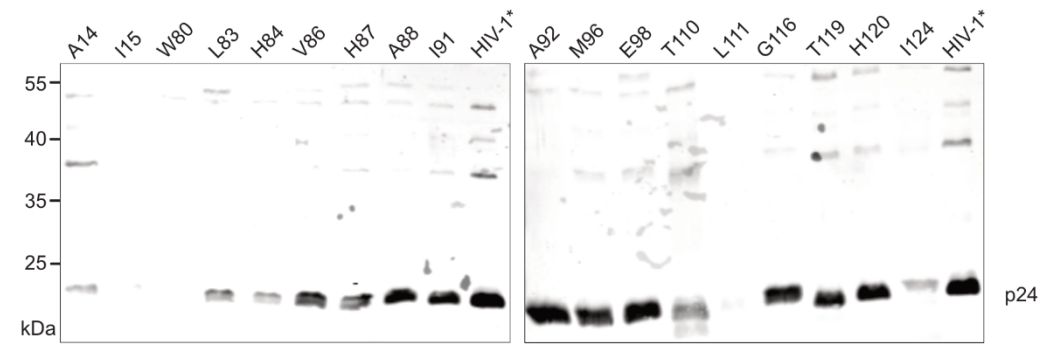
After identifying proteolytically processed CA of viral particles in immunoblot analysis, particle production was additionally quantitated using the SG-PERT assay. Cells supplemented with CpK produced a relatively substantial number of viral particles for all variants in the range of 10^7 – 10^8 pU RT/µl (Figure 13b, dark grey). The reduction due to amber suppression expression was only modest compared to HIV-1* in presence of CpK, whereas in absence of CpK the yield was approximately 10-fold higher indicating a general effect of CpK on expression performance. Obtained SG-PERT values are in accordance with the CA quantification performed in immunoblot analysis (Figure 13a) assessing relatively high yields for most variants compared to HIV-1*, which is illustrated in addition by the calculated ratio of HIV-1*CAxx^{CpK} and HIV-1* based on SG-PERT values (Figure 13c). As expected, for particles produced in the absence of CpK, the RT activity dropped drastically for all variants due to incomplete Gag expression at the amber stop codon (Figure 13b, light grey). However, ratio calculations of particle yield with and without CpK resulted in a wide range of ~ 10-1,000-fold increase upon CpK addition (Figure 13c). Residues that showed only a ~ 10-fold increase (W80, L83, H84, V86, H87, E98, T110, G116, T119) thus argue for site-dependent leakiness of the system, which was already observed in previous immunoblot analysis (*cf*. 2.2.1; Figure 12). This might be explained by aminoacylation and loading of a natural amino acid (most likely lysine) onto the tRNA^{TAG}, near-cognate suppression of endogenous tRNAs via non-Watson-

Crick interaction, or by stop codon skipping resulting in an amino acid deletion due to a +3 frameshift[280–284].

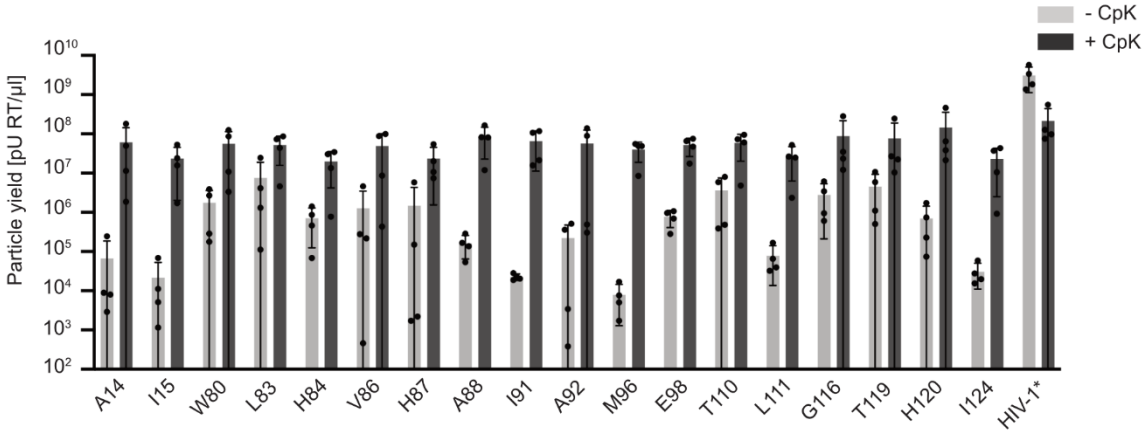
Figure 13d combines results from all screening experiments in order to allow a comparison of the 18 variants in a systematic manner. Apart from examining protein expression and processing of GCE-modified HIV-1 CA variants, relative infectivity of unlabeled viral particles determined in TZM-bl reporter cells using a luciferase assay was included. As expected, the fragile CA was not tolerant of most changes in the amino acid sequence resulting in severely decreased single round infectivity. The majority of HIV-1*CAxx^{CpK} variants lost the ability to detectably infect host cells (Figure 13d and Appendix Table 9). Only HIV-1*CAxx^{CpK} A14, L83, I91, L111, and H120 particles displayed detectable infectivity in TZM-bl indicator cells with relative infectivity between ~ 30 and > 50 % of HIV-1 WT. Thus, several tested CA variants retain infectivity and might be suitable to analyze the early steps of HIV-1 infection.

The wide range of efficiencies and specificities show the importance of screening several CA positions for the ncAA-incorporation. The combination of results obtained from different experiments performed with all HIV-1*CAxx^{TAG} variants indicated a subset of residues for further investigation. Maintained infectivity, high Gag expression and amber suppression efficiencies were crucial for the selection of protein sites A14, L83, I91, L111 and H120.

a



b



c

	A14	I15	W80	L83	H84	V86	H87	A88	I91
$yield_{-CpK}$ [pU RT/µl]	$6.7x10^4$	$2.1x10^4$	$1.8x10^6$	$7.5x10^6$	$7.0x10^5$	$1.3x10^6$	$1.5x10^6$	$1.6x10^5$	$2.3x10^4$
$yield_{+CpK}$ [pU RT/µl]	$6.1x10^7$	$2.4x10^7$	$5.6x10^7$	$5.2x10^7$	$2.0x10^7$	$4.9x10^7$	$2.3x10^7$	$8.5x10^7$	$6.5x10^7$
ratio +CpK/-CpK	913	1100	32	7	28	40	16	528	2883
ratio +CpK/HIV-1*	0.29	0.11	0.26	0.24	0.09	0.23	0.11	0.40	0.31

	A92	M96	E98	T110	L111	G116	T119	H120	I124
$yield_{-CpK}$ [pU RT/µl]	$2.2x10^5$	$7.9x10^3$	$7.5x10^5$	$3.7x10^6$	$7.6x10^4$	$2.8x10^6$	$4.6x10^6$	$7.0x10^5$	$3.1x10^4$
$yield_{+CpK}$ [pU RT/µl]	$5.7x10^7$	$4.0x10^7$	$5.2x10^7$	$5.9x10^7$	$2.7x10^7$	$8.8x10^7$	$7.6x10^7$	$1.5x10^8$	$2.3x10^7$
ratio +CpK/-CpK	261	5117	68	16	348	31	17	210	753
ratio +CpK/HIV-1*	0.27	0.19	0.24	0.27	0.12	0.41	0.36	0.68	0.11

d

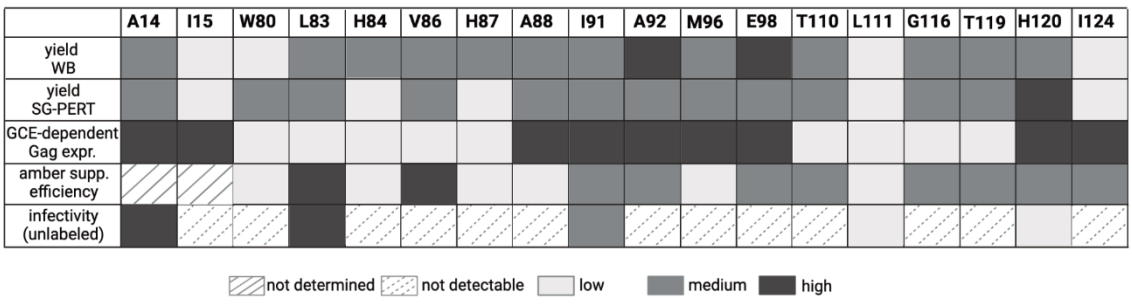


Figure 13. Screening experiments of HIV-1*CAxx^{CpK} variants. a. CA quantification. At 48 hp.t., he supernatant of HEK293T cells co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1*/HIV-1*CAxx^{TAG} variant grown in the presence of 500 µM CpK was harvested, filtered, and concentrated via ultracentrifugation through a sucrose cushion. Viral proteins were separated by SDS-PAGE and transferred onto a nitrocellulose membrane by semi-dry blotting. 10 µl particle lysate for each variant and 5 µl for HIV-1* were loaded. According to the manufacturer's instructions, proteins were quantitatively detected by IF using a polyclonal antibody raised against HIV-1 CA, a polyclonal Alexa Fluor secondary antibody, a Li-COR CLx infrared scanner. The Li-COR software Image Studio lite 5.0 and a CA standard (2.5 ng/µl; in-house) added in serial dilutions were used to quantify CA amounts. **b.** Harvested supernatant was used to determine particle yield in a SG-PERT assay[269] for samples produced in the absence (light grey) or presence of CpK (dark grey). The graphs show mean values and standard deviation of four individual experiments. **c.** Representation of individual SG-PERT values [pU RT/µl] for all variants obtained from (b). The ratio of SG-PERT values from HIV-1*CAxx^{CpK} and HIV-1* (ratio +CpK/HIV-1*) illustrates the relative number of particles produced compared to WT. The ratio of SG-PERT values from HIV-1*CAxx^{TAG} and HIV-1*CAxx^{CpK} (ratio +CpK/-CpK) illustrates the site-specific leakiness of the system. d. Overview of performed screening experiments with results assigned to performance categories (not determined, not detectable, low, medium, high). Table contains the yield determined via CA quantification in immunoblotting and RT activity in SG-PERT measurements, GCE-dependent Gag expression based on comparison of SG-PERT values of particles produced in the presence or absence of CpK, calculated amber suppression efficiency based on immunoblot analysis and Steady-Glo infectivity measurements determined in TZM-bl reporter cells. Modified from Schifferdecker *et al*. (in preparation)[270]

2.2.3 Control experiments to detect truncated protein and CypA incorporation

In the context of amber suppression, transfected cells produce full-length proteins and truncated proteins due to inefficient readthrough of the amber stop codon, which might have dominant-negative effects on the viral particles. Due to this, HEK293T cells were co-transfected with pNESPylRS-eRF1dn-tRNA and a proviral plasmid encoding the coding sequence of HIV-1* or one of the prior selected HIV-1*CAxx^{TAG} variants A14, L83, I91, L111 or H120. Incorporated truncated proteins in mature viral particles were detected by immunoblot analysis with an antibody against HIV-1 MA (Figure 14a). Processed MA was detected for all variants. For L83, I91, and L111, faint bands were visible at the size of truncated MA-CA^{TAG} (~ 20 kDa). The band was absent for H120 particles assuming that readthrough efficiency was higher in this residue context. Truncated protein forms for A14 (~ 16.32 kDA) were not detected either because truncated proteins were not incorporated or because the truncated protein form was not distinguished from mature MA running at the size of ~ 15 kDa. Of note, however, is that all samples show prominent bands at the size of full-length Gag (55 kDa) and intermediate p41, which are not part of released and mature particles. These bands most likely correspond to an adverse contamination of concentrated particles with virus-producing cells. Thus, the truncated protein bands detected in some of the samples might be also caused by cell contaminations instead of truncated protein incorporation.

CypA is one of the first host proteins found to interact with HIV-1 CA[285]. It is incorporated in virions at a 1:10 molar ratio with a significant impact on viral infection[52,174,175]. The binding of CypA in a

proline-rich region of CA (CypA binding loop) can be disrupted by single amino acid mutations, such as residues G89 or P90[52,175]. Thus, it was investigated whether the candidates with residues near the CypA binding loop disturb the binding and incorporation of the protein. Figure 14b shows immunoblot analysis of lysed particles stained against HIV-1 CA and CypA. Three different variants with amber codons at positions L83, H87 and I91 were tested. Mutations of residue G89V and P90A were used as negative controls. All three amber suppressed protein variants showed incorporation of CypA into the particles detected by the specific protein bands at the size of ~ 15 kDa for CypA. The absence of the CypA band confirmed the inability of CA variants G89V and P90A in CypA binding.

Thus, it was assumed that residue candidates near the CypA-loop do not disturb host factor binding and may have no adverse effects on host cell infection.

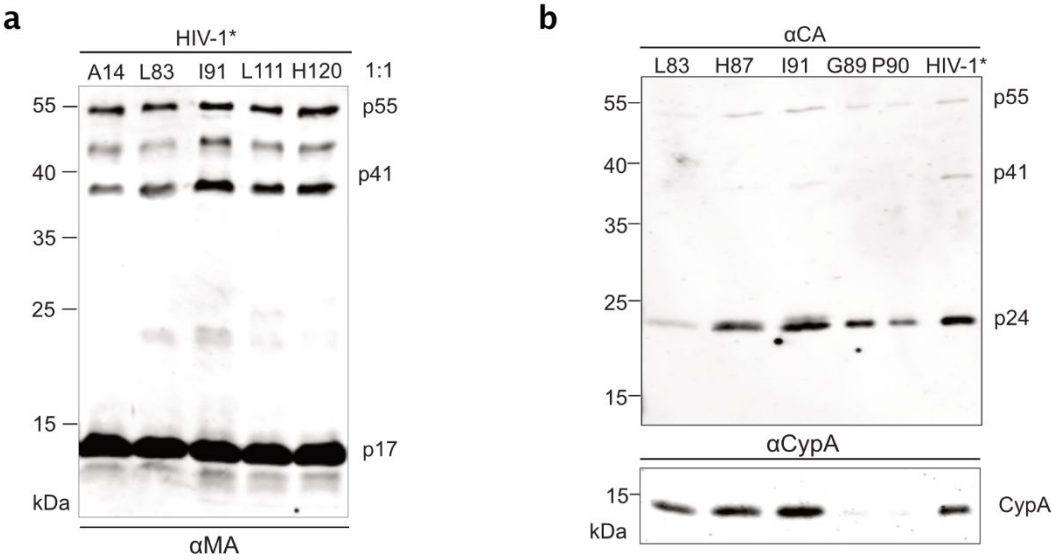


Figure 14. Control experiments for incorporation of truncated proteins into mature virions and binding of CypA. a. HEK293T cells were co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1* plus indicated HIV-1*CAxx^{TAG} variant in a 1:1.8:1.8 µg ratio, grown in the presence of 500 µM CpK. At 48 h p.t., concentrated particles were lysed, proteins were separated by SDS-PAGE, blotted, and analyzed by immunostaining with an antibody raised against HIV-1 MA. Bands were detected for mature MA (p17), Gag intermediate product (p41), and full-length Gag (p55). **b.** HEK293T cells were co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1* or one of the HIV-1*CAxx^{TAG} variants L83, H87, I91. G89V and P90A were used as positive controls for interfering with CypA incorporation. Concentrated particles containing the ncAA bicyclo [6.1.0] nonyne-lysine (BCN) were subjected to immunoblot analysis stained against CA and CypA. Experiment was performed by Volkan Sakin (University Hospital Heidelberg). Modified from Schifferdecker *et al*. (in preparation)[270].

2.2.4 Site-specific click-labeling of HIV-1 CA

Since CA click-labeled HIV particles offer a versatile tool for live-cell imaging and super-resolution microscopy, the fluorescent probe for CA labeling required specific characteristics. To comply with desired properties such as membrane-permeability allowing the labeling of assembled particles, high photostability for intensive imaging, and adaptability for imaging of live cells[286], silicon rhodamine (SiR)[287] was selected. Coupling the SiR-carboxy to a tetrazine, which can react rapidly with strained alkynes/alkenes[288,289] such as CpK (Figure 8c and 15a), allowed imaging under physiological conditions without the need of extensive washing steps and comparably low background[290,291].

Initial screening experiments (*cf.* 2.2.1 - 2.2.3) offered the possibility to decide on a smaller panel of candidate residues of CA to apply GCE and click-labeling with a fluorescent dye. Supernatant of cells transfected with HIV-1*CAxx^{TAG} variants A14, L83, I91, L111, H120, and HIV-1* was harvested and 125, 250, or 500 nM of tetrazine-SiR (Tet-SiR) was added for 30 min at 37° C to stain CA in viral particles. After concentration of click-labeled particles by ultracentrifugation, particle lysates were separated by SDS-PAGE to detect labeled proteins by In-gel fluorescence. The appearance of bands at the size of CA (~ 25 kDa) reflected site-specific and successful protein labeling with Tet-SiR (Figure 15b). Particles stained with 500 nM Tet-SiR showed a specific band of the expected molecular mass for variants A14, L83, I91, and L111. Protein bands were not detectable for HIV-1* particles produced under the same labeling conditions indicating protein-specific click labeling. Although L83 had shown leakiness of amber suppression (*cf*. 2.2.1; Figure 12 and 2.2.2; 13b,c), labeling of GCE-modified L83 variant was possible for all three tested Tet-SiR concentrations. This result indicates that this residue achieved UAG readthrough by using escape mechanisms in the absence of CpK, but incorporated CpK when present in the medium to at least a certain extent.

Subsequently, the functional impact of the addition of Tet-SiR to the CA was evaluated. Therefore, the relative infectivity of labeled particles after adding different dye concentrations was determined (Figure 15c). As already shown for unlabeled particles (*cf*. 2.2.2; Figure 13), the overall infectivity of CA variants decreased compared to HIV-1*. The infectivity of HIV-1*CAxx^{SiR} variants A14, I91, and H120 was only mildly affected with ~ 2-fold drop. However, no significant differences in infectivity of particles labeled at different Tet-SiR concentrations was observed, suggesting that the additional Tet-SiR mass is not the limiting factor to produce infectious CA click-labeled particles.

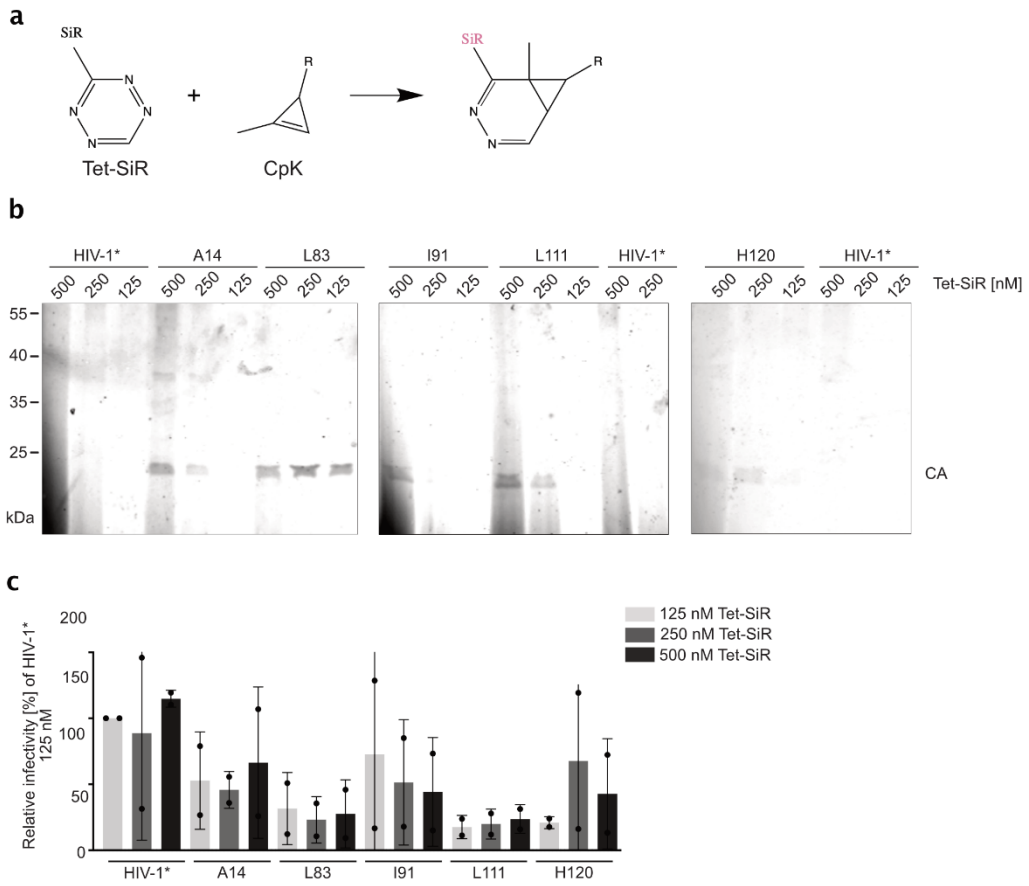


Figure 15. Click-labeled particles retained infectivity. a. Scheme of the strain-promoted inverse electron-demand Diels-Alder [4+2] cycloaddition (SPIEDAC) reaction of a tetrazine-derivative of silicon rhodamine (Tet-SiR; electron-poor diene) and the ncAA cyclopropene-L-lysine (CpK; electron-rich dienophile). The tetrazine group reacts with the strained alkene of CpK, resulting in a CA covalently bound to SiR. **b.** Filtered supernatant of virus producing cells was incubated with 125, 250, or 500 nM Tet-SiR for 30 min at 37° C. Click-labeled particles were concentrated by ultracentrifugation, and particle lysates were separated by SDS-PAGE. In-gel fluorescence was detected using a Li-COR CLx infrared scanner at a wavelength of 700 nm. Particles with successfully labeled CA showed specific bands migrating at the molecular mass of CA (~ 25 kDa). **c.** Relative infectivity was determined by titration on TZM-bl cells and Steady-Glo analysis for different CAxx^{TAG} variants labeled with indicated Tet-SiR concentrations (125, 250, 500 nM) normalized for RT activity measured in SG-PERT analysis. Values obtained for each virus variant were normalized to the values for HIV-1* labeled with 125 nM Tet-SiR. Graphs show mean values and standard deviation from two individual experiments. Modified from Schifferdecker *et al.*(in preparation)[270].

2.2.5 Detection of CA click-labeled particles by confocal microscopy

Based on the different labeling properties and viral fitness of CA click-labeled particles identified in prior experiments (Figure 16), HIV-1*CA14^{SiR} and HIV-1*CA91^{SiR} were considered for further assessment. Thus, click-labeled HIV-1*CA14^{SiR} and HIV-1*CA91^{SiR} particles were imaged by spinning-disc-confocal microscopy (SDCM) (Figure 16). Immobilized HIV-1* particles served as a negative control for the click-reaction (Figure 16a). The positions of particles on the coverslip were identified

by CA IF. Imaging of particles labeled with different Tet-SiR concentrations confirmed the result obtained from In-gel fluorescence analysis (*cf*. 2.2.4; Figure 15) showing that HIV-1* particles (Figure 16a, upper row) lacked detectable signal intensities in the CA SiR (Figure 16a, lower row) channel. HIV-1*CA14^{SiR} and HIV-1*CA91^{SiR} showed specific colocalization of CA IF and CA SiR foci (Figure 16b, c). While only a few CA SiR-positive particles were detected upon staining with 125 nM Tet-SiR, the proportion of click-labeled particles increased with 250 and 500 nM. Since visual inspection showed that 250 nM Tet-SiR was sufficient to label the vast majority of CA IF detected particles, further experiments were performed at this concentration.

Infectivity measurements described earlier were performed in large screening experiments using all (*cf*. 2.2.2; Figure 13d) or a subset of mutants (*cf*. 2.2.4; Figure 15c). However, these experiments were performed without inhibiting a second round of infection making the GCE-modified and click-labeled particles, which are unable to undergo secondary infection in target cells not expressing the GCE system, not comparable to WT virus. Since newly infected cells do not have the GCE system, HIV-1*CAxx^{TAG} variants cannot be translated. Thus, cells cannot produce new amber suppressed particles to infect in a second round. To investigate if a second round of HIV-1* might have influenced the result, the infectious units of HIV-1*CA14^{SiR} and HIV-1*CA91^{SiR} were determined in presence of the fusion inhibitor T-20 using a more precise blue-cell counting assay. Data from three independent experiments showed that click-labeled HIV-1*CA91^{SiR} is ~ 50 % less infectious compared to HIV-1*CA14^{SiR} (Figure 16d).

To determine the labeling efficiency of both variants, CA SiR intensities of detected particles were plotted against CA IF. Every particle with CA SiR intensity above the threshold was considered CA SiR-positive. Imaging HIV-1* particles resulted in 3.17 ± 1.28 % positive particles due to the unspecific background of Tet-SiR binding to the surface (Figure 17a). Both variants counted high efficiencies in labeling with 95.53 ± 5.04 % CA IF/CA SiR colocalization for HIV-1*CA14^{SiR} (Figure 17b) and 99.94 ± 0.08 % for HIV-1*CA91^{SiR} (Figure 16c), which showed successful labeling of almost all particles. Although In-gel fluorescence did not show SiR signals for HIV-1*CA91^{SiR} labeled with 250 and 125 nM (*cf*. 2.2.4; Figure 15b), SDCM revealed CA SiR positive signals for particles labeled with all three concentrations. Measurement of mean fluorescent intensity (MFI) of CA SiR-positive particles showed 3-fold higher intensities for HIV-1*CA14^{SiR} (21,376 a.u.) HIV-1*CA91^{SiR} (21,566 a.u.) compared to HIV-1* (7,008 a.u.) confirming unspecific background detection with HIV-1* particles.

Since click-labeled A14 is more infectious than I91, with comparable labeling efficiency and SiR intensities, the A14 variant was selected for further experiments to investigate early HIV-1 replication.

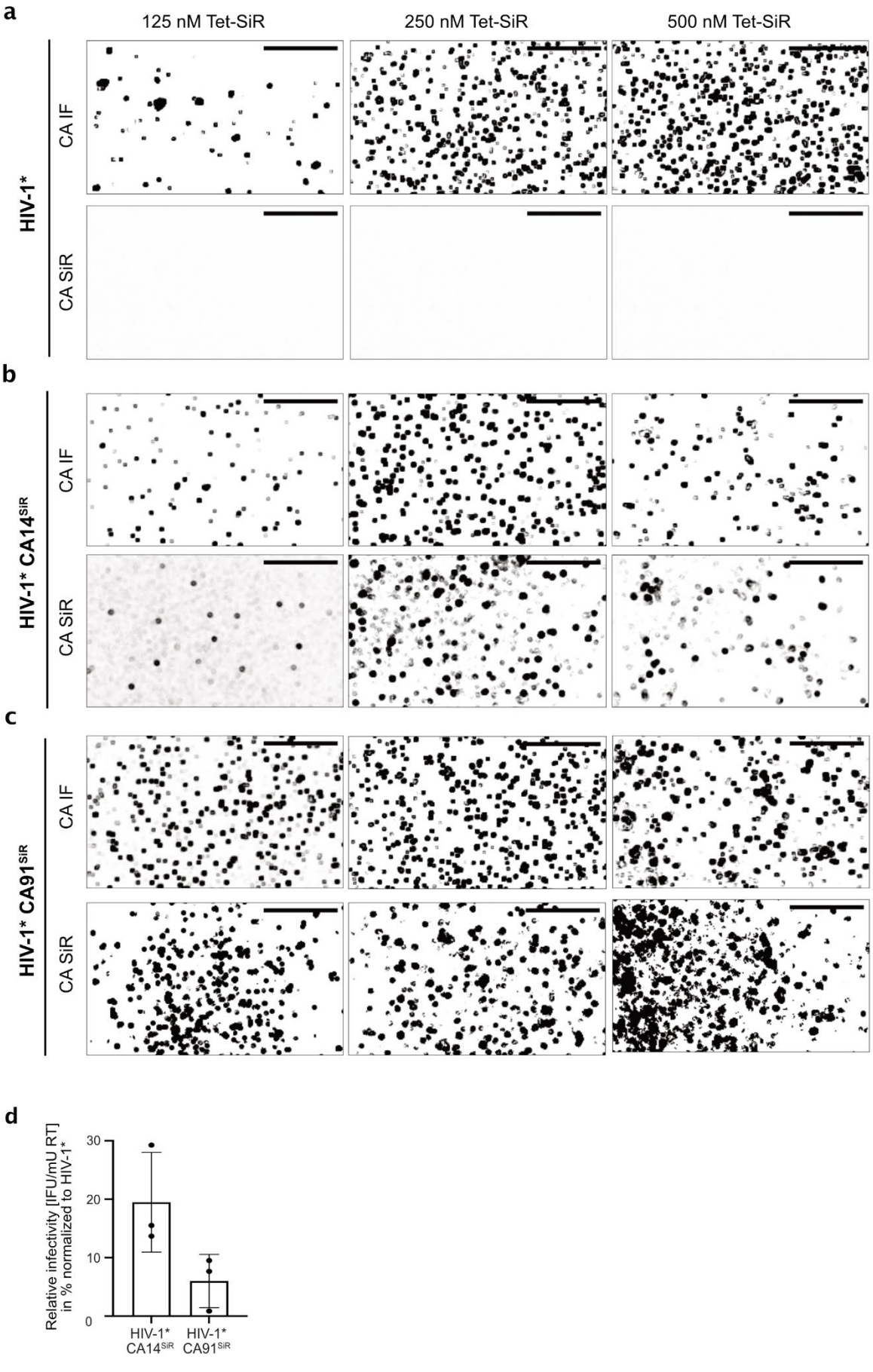

a
125 nM Tet-SiR
250 nM Tet-SiR
500 nM Tet-SiR
HIV-1*
CA IF
CA SiR
b
HIV-1* CA14SiR
CA IF
CA SiR
c
HIV-1* CA91SiR
CA IF
CA SiR
d
Relative infectivity [IFU/mU RT] in % normalized to HIV-1*
30
20
10
0
HIV-1* CA14SiR
HIV-1* CA91SiR

Figure 16. Detection of click-labeled particles in SDCM. a-c. Harvested supernatant of virus-producing HEK293T cells, co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1* (a), HIV-1*CA14^{TAG} (b), or HIV-1*CA91^{TAG} (c) was filtered and labeled with 125 (left panel), 250 (middle panel) or 500 (right panel) nM Tet-SiR for 30 min at 37° C. Click-labeled particles were concentrated via ultracentrifugation through sucrose cushion. Particles were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes, fixed with 4 % PFA, stained by IF using a polyclonal antibody raised against HIV-1 CA and imaged by SDCM. Representative images are shown for HIV-1* and two virus variants. Variance filter and background subtraction were applied for clarity. Scale bars: 10 µm. **d.** TZM-bl cells were seeded into 96-well cell culture plates one day before HIV-1*, HIV-1*CA14^{SiR} and HIV-1*CA91^{SiR} particles (250 nM Tet-SiR) were titrated. At 48 hours post-infection (h p.i.), cells were fixed for 10 min with 100 % methanol, and beta-galactosidase expressing cells were quantitated by microscopy (Blue cell assay). Calculated infectious units (IFU) were normalized for RT activity as a measure of particle amount, and values obtained for HIV-1* were set to 100 %. Bar graphs show mean and standard deviation from three independent experiments performed in technical triplicates. Modified from Schifferdecker *et al.*(in preparation)[270].

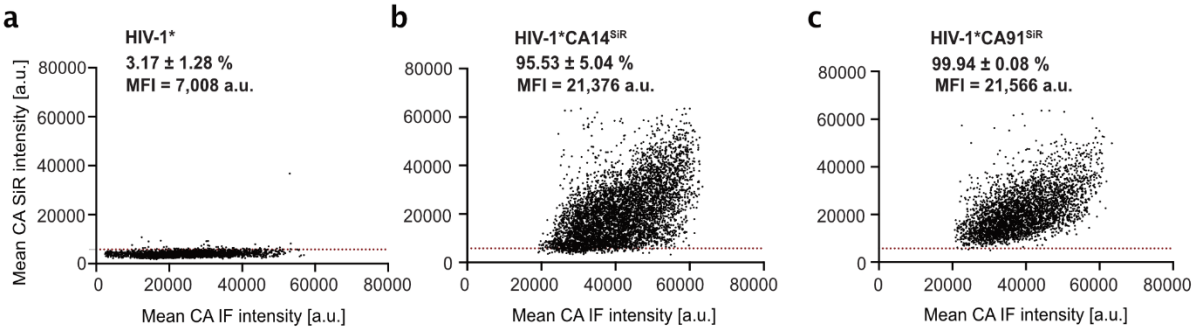


Figure 17. Labeling efficiency of HIV-1*CA14^{SiR} and HIV-1*CA91^{SiR}. The supernatant of cells, co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1*/ HIV-1*CA14^{TAG}/ HIV-1*CA91^{TAG} was harvested, filtered, and labeled with 250 nM Tet-SiR. After ultracentrifugation, viral particles were immobilized, fixed, stained against CA by IF, and imaged by SDCM. Mean CA SiR intensities of detected HIV-1* (**a**, n = 2349), HIV-1*CA14^{SiR} (**b**, n = 5484), and HIV-1*CA91^{SiR} (**c**, n = 3598) particles were plotted against CA IF intensities. Mean intensity of the background measured in Fiji was set as a threshold indicated by the red lines.

2.3 Microscopic imaging of HIV-1*CA14^{SiR} particles to follow incoming capsids

Validation and characterization assays were performed to compare 18 candidate residues in order to find the most promising site in the CA for ncAA incorporation and click-labeling. Summarized, modification of the residue A14 resulted in the most infectious variant showing GCE-dependent Gag expression suggesting a comparably high CpK incorporation with moderate particle yield (*cf.* 2.2.2; Figure 13d). Therefore, A14 was selected to perform intense microscopic analysis in order to fully characterize a CA click-labeled virus variant investigating uncoating in different cell types.

2.3.1 HIV-1*CA14^{SiR} particles show normal morphology during budding and release

The usefulness of modified variants for the planned studies depends on faithful assembly and formation of a mature capsid. After showing that proteolytic processing appears undisturbed by immunoblot (*cf.* 2.2.1; Figure 12), the morphology of HIV-1*CA14^{CpK} particles during and after release was analyzed by thin-section EM of HEK293T cells transfected with the proviral plasmid and GCE system grown in the presence of CpK.

At 44 h p.t., typical HIV-1 budding sites were observed in virus-producing cells (Figure 18a). Released particles were detected in immature but also mature states with densities representing the mature capsid lattice. These data show that ncAA-dependent Gag and particle production led to fully mature particles without apparent inhibition of immature or mature lattice assembly.

2.3.2 Large-scale particle preparation and characterization

For the subsequent microscopic investigation of infected cells, large-scale particle preparations were performed using calcium phosphate-mediated transfection[292–294] and resulting particles were characterized.

CpK-dependent Gag expression and processing of cells transfected with the CA-modified plasmid HIV-1*CA14^{TAG} were analyzed by immunoblot analysis of cell lysates stained against MA (Figure 18b). Whereas HIV-1* transfected cells showed Gag expression in the absence and presence of CpK, the production of HIV-1*CA14^{TAG} transfected cells relied on the addition of CpK to enable readthrough of the amber codon. The overall yield of HIV-1*CA14^{SiR}, determined by quantitation of RT activity (Figure 18c) and quantitation of HIV-1 CA by immunoblot (Figure 18d), was 5-10-fold reduced compared to the yield of HIV-1*, presumably caused by incomplete amber suppression[217,257,279]. In line with the detection of mature particles at budding sites, immunoblot analysis (Figure 18e) confirmed that click-labeled HIV-1*CA14^{SiR} particles contained the mature RT heterodimer (p66, p51) and mature CA (p24) in detectable amounts. The detection of in-gel fluorescence corresponding to the molecular mass of CA (~ 25 kDa) for HIV-1*CA14^{SiR} particles and the absence of the band in the HIV-1* sample verified the specificity of CA labeling under amber suppression conditions (Figure 18f).

EM analyses had revealed immature Gag shells and particles with mature morphology (Figure 18a). Immature and mature click-labeled particles were also analyzed by STED nanoscopy. For this, HIV-1*CA14^{CpK} particles were produced in the absence or presence of the proteasomal inhibitor lopinavir

(LPV), which inhibits proteolytic Gag cleavage and morphological particle maturation[295]. Subsequently, particles were immobilized and click-labeled with a fluorescent dye in order to detect typical morphologies in STED nanoscopy of immature mature virions as described previously in another study from our lab using a Gag label[296]. CA click-labeled particles produced in the presence of LPV showed a typical hemispherical Gag shell on the outer surface of the virion (Figure 18g, left panel), which confirms that CpK clicked to the fluorescent label was introduced specifically into the protein of the particles suitable for super-resolution nanoscopy. Without LPV treatment, STED imaging revealed the expected round shape of a mature CA-labeled virion with a condensed blob-like inner appearance (Figure 18g, right panel), which suggests that redistribution of CA within the particle was possible even after CpK incorporation.

These results verified that GCE-modified CA at residue number 14, produced in a large-scale optimized protocol, was accessible for the click-reaction in the immature but also mature virion, and CpK-incorporation did not harm particle maturation already assumed from prior immunoblot and EM analysis.

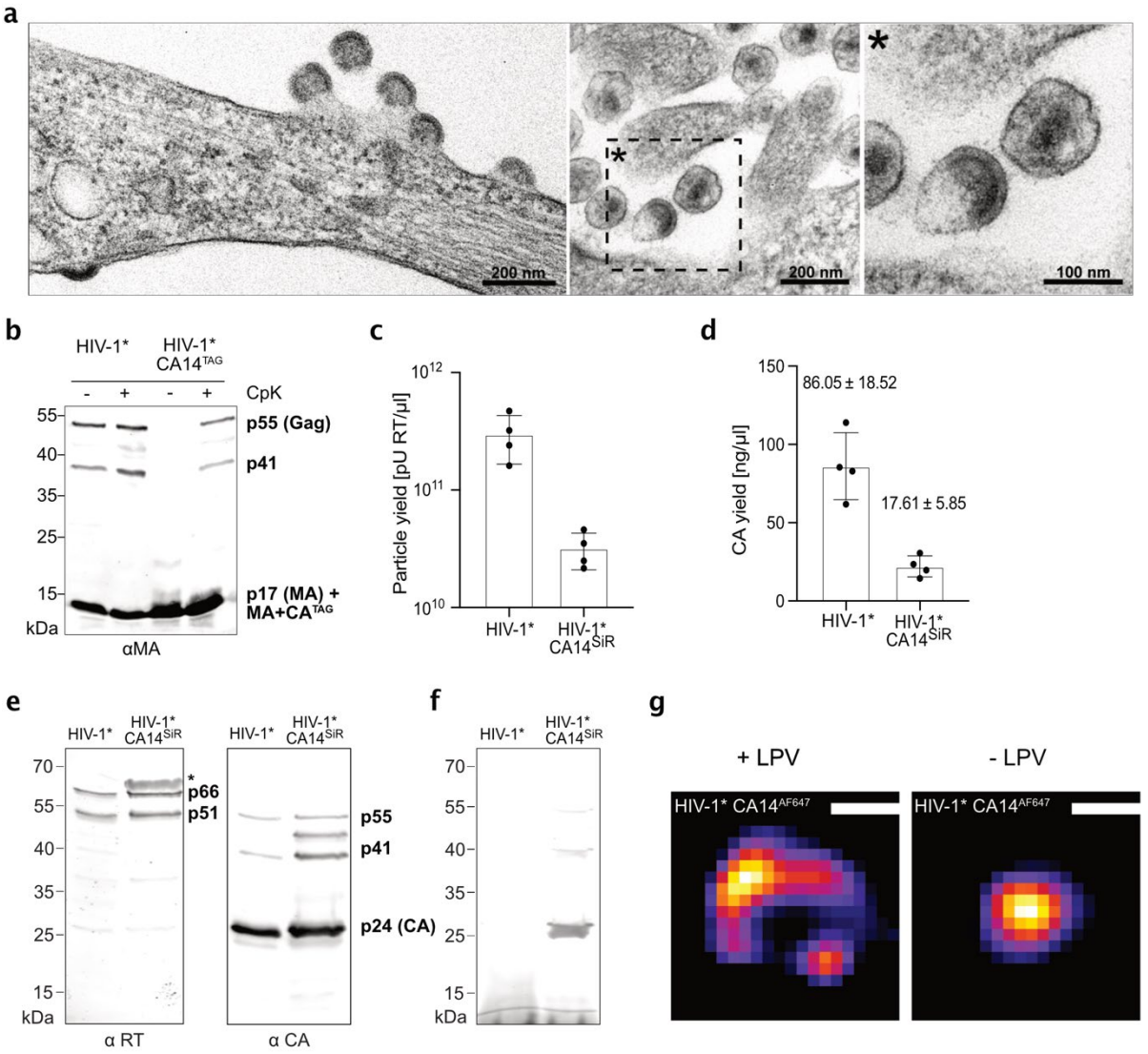


Figure 18. Characterization of HIV-1*CA14^{SiR} click-labeled particles in large-scale production. **a.** Morphology of mature particles and assembly sites in virus-producing cells. HEK293T cells were co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1*CA14^{TAG} and supplemented with 500 µM CpK. At 44 h p.t., cells were fixed, embedded, and analyzed by thin-section EM (*cf.* 5). EM was performed by Vojtech Zila (University Hospital Heidelberg). **b.** GCE-dependent Gag expression. HEK293T cells were co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1* or HIV-1*CA14^{TAG} and grown in the presence or absence of 500 µM CpK. At 48 h p.t., cell lysates were separated by SDS-PAGE, and transferred onto a nitrocellulose membrane by semi-dry blotting. Gag-derived proteins (p55, p41, p17) were detected by IF with an antibody against HIV-1 MA, fluorescent secondary antibody, and a Li-COR CLx infrared scanner. Note that truncated Gag products due to incomplete amber suppression (MA+CA truncated: ~ 16.32 kDa) were not separated from mature MA (~ 14.8 kDa) in these experiments. **c-f.** The supernatant of virus-producing cells was filtered, labeled with 250 nM Tet-SiR and particles were concentrated via sucrose cushion and ultracentrifugation. **c.** Particle yield for HIV-1* and HIV-1*CA14^{SiR} quantified by RT activity (SG-PERT assay[295]). Data show mean values and standard deviation of four independent experiments. **d.** CA yield for HIV-1* and HIV-1*CA14^{SiR} particles determined by quantitative immunoblot analysis as shown in (e). Data show mean values and standard deviation of four independent experiments. **e.** Immunoblot analysis of particle lysates. Particles were separated by SDS-PAGE, blotted onto a nitrocellulose membrane by semi-dry blotting, and stained by IF using antibodies raised against HIV-1 RT protein (RT) and capsid (CA). Secondary antibodies were detected by a Li-COR CLx infrared scanner. The asterisk indicates non-specific binding of the antibody to bovine serum albumin, a

contamination from growth medium. **f.** In-gel fluorescence of click-labeled CA. HIV-1* and HIV-1*CA14^{SiR} particle lysates were separated by SDS-PAGE. The gel was imaged using a Li-COR CLx infrared scanner at a wavelength of 700 nm. **g.** HEK293T cells, co-transfected with pNESPylRS-eRF1dn-tRNA and HIV-1*CA14^{TAG} IN.eGFP, were grown in the presence of 500 µM CpK and with (left panel) or without (right panel) 2 µM LPV. Concentrated particles were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes, fixed with 4 % PFA, permeabilized, stained with 1µM pyrimidyl-tetrazine-AF647 for 60 min at 37° C and imaged by STED nanoscopy. Scale bars: 200 nm. Modified from Schifferdecker *et al*., 2021[266].

2.3.3 Highly CA-labeled particles retained viral infectivity

Mean intensities of HIV-1*CA14^{SiR} particles labeled with different Tet-SiR concentrations were compared to ensure that the optimized large scale particle production did not compromise the SiR intensities of click-labeled particles. Figure 19a shows that increasing Tet-SiR concentration used to click-label HIV-1*CA14^{CpK} particles also increased SiR intensities of individual particles detected by SDCM (SiR mean intensities for 125 nM (3,580.59 ± 913.39 a.u.), 250 nM (5,269.51 ± 1,941.68 a.u.), and 500 nM (6,811.31 ± 2,712.03 a.u.)). In contrast, only unspecific background was detected upon staining of HIV-1* particles. Based on the high SiR intensity obtained, 250 nM Tet-SiR was chosen for further experiments, also in large scale productions.

Figure 19b,c show that the labeling efficiency of HIV-1*CA14^{SiR} was retained reaching more than 95 % efficiency, while HIV-1* particles, produced and stained in parallel, measured only unspecific background intensities. Subsequently, the infectivity of those particles was assessed in a microscopy-based analysis (Figure 19d,e). HIV-1* and HIV-1*CA14^{SiR} labeled particles were titrated on TZM-bl indicator cells, treated with 50 µM T-20, fixed at 48 hours post infection (h p.i.), and stained against HIV-1 MA by IF. Due to the inability of GCE-modified particles after infection to produce full-length Gag, only the truncated MA-CA^{TAG} protein could be used as an indicator for productive infection of click-labeled particles. The number of MA-positive cells upon infection was quantified (Figure 19d) and used to determine the relative infectivity of HIV-1*CA14^{SiR}, resulting in a 2-fold drop compared to HIV-1* (Figure 19e).

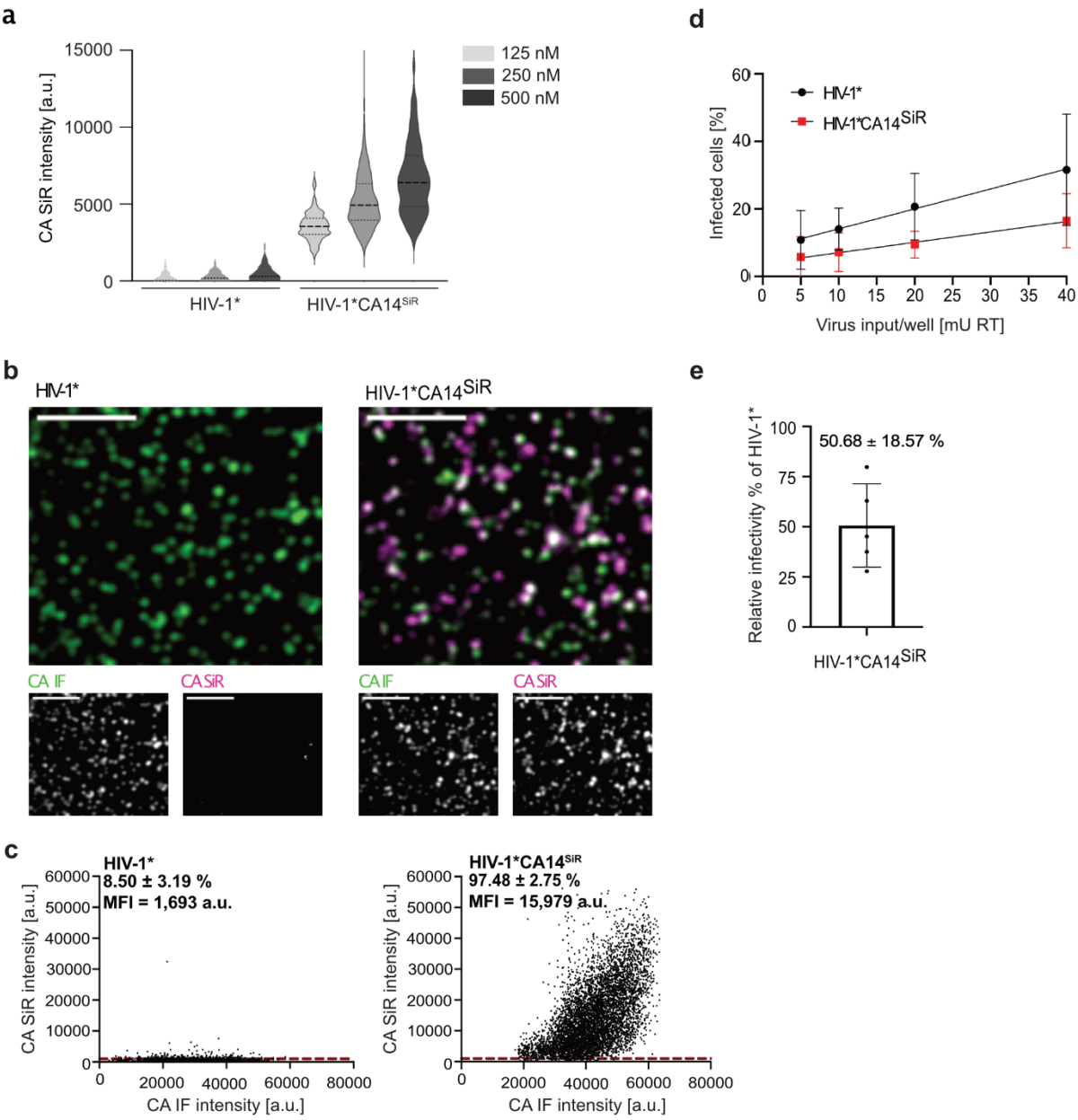


Figure 19. Click-labeled HIV-1*CA14^{SiR} particles retain infectivity. a-c. Purified and labeled HIV-1*CA14^{SiR} particles (125, 250, 500 nM Tet-SiR) were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes, stained against HIV-1 CA, and imaged by SDCM. **a.** Particles were detected by CA IF and CA SiR mean intensities were plotted for HIV-1* and HIV-1*CA14^{SiR} labeled with indicated Tet-SiR concentrations. n > 100 particles per condition. Graph shows mean values and quartile lines of particle intensities detected in one experiment. **b.** Representative single z-slice SDCM images of HIV-1* and HIV-1*CA14^{SiR} particles labeled with 250 nM Tet-SiR. Separated CA IF (green) and CA SiR (magenta) channels are shown below. Mean filter and background subtraction was applied to all images for clarity as described in material and methods. **c.** Mean CA SiR intensities of particles detected in (b) were plotted against mean CA IF intensities using the icy software and spot detector. The threshold was set to t = 1,000 a.u.. All HIV-1* (n = 3,300) and HIV-1*CA14^{SiR} (n = 6,900) particles above the threshold were counted as sufficiently labeled by Tet-SiR. Graph represents data of three independent experiments and virus preparations. For each experiment, 12 FOVs were analyzed. **d.** Infectivity of click-labeled particles in TZM-bl cells. HIV-1* and HIV-1*CA14^{SiR} particles were titrated on TZM-bl cells seeded in 15-well ibidi µ-Slide angiogenesis dishes. At 6 h p.i., 50 µM T-20 was added to prevent second-round infection. At 48 h p.i., cells were fixed, permeabilized, stained against HIV-1 MA by IF, and imaged by SDCM. The proportion of infected cells was quantified based on detected MA signal and total number of nuclei

counted automatically in Fiji. Data show mean values and standard deviation from five independent experiments (n = 5,500 - 7,700 cells counted per condition). Lines indicate linear regression based on the mean values to determine infectious units/mU RT. **e.** Relative infectivity of particles analyzed in (c). Values obtained for HIV-1*CA14^{SiR} were normalized to the values of HIV-1* in each experiment. Bar graph shows mean values and standard deviation of five individual experiments and particle preparations. Modified from Schifferdecker *et al*., 2021[266].

2.3.4 Click-labeled particles can be efficiently detected inside the cytoplasm of infected cells

Subsequent to the detection of free particles on a coverslip, I tested the detection of particles within the environment of an infected cell. In contrast to the PBS-buffered environment used in prior particle detection, the intracellular environment exhibits different physiochemical properties. Therefore, cells were infected with HIV-1*CA14^{SiR} particles and SDCM was used to detect single particles in a cellular environment.

Initial experiments were performed in a widely used HeLa-derived TZM-bl model cell line[297]. Cells infected with HIV-1*CA14^{SiR} revealed highly CA SiR-labeled particles in the cytoplasmic region and near the NE (Figure 20a). Click-labeled particles was not affected by the cell environment and displayed the same high level of CA IF/CA SiR colocalization (labeling efficiency of 99.37 ± 1.43 % (Figure 20b)). The unbound Tet-SiR dye in cells infected with HIV-1* was less attached to the cell surface than the glass surfaces or accumulated less in the cytosol, resulting in even lower detection of unspecific background (0.74 ± 0.94 %). High CA SiR intensities detected in the cellular environment enabled imaging with moderate laser intensities between 7-30%. Enlargements of representative cytoplasmic particles indicated by the white arrowheads show clear colocalization of directly labeled CA (CA SiR) with CA IF (Figure 20c). A closer look at the intensity distribution on a profile plot along the white line drawn in figure 20c confirmed that intensity peaks of CA IF and CA SiR were overlapping at the particle sites (Figure 20d).

To conclude, I have achieved minimally invasive labeling of the HIV-1 CA via amber suppression. Residue A14 was identified to allow the production of viral particles, which retained the ability to productively infect cells after incorporating an amber stop codon and adding a fluorescent dye. Detection of click-labeled HIV-1*CA14^{SiR} particles in infected cells by SDCM and STED indicated the general applicability of this modified virus variant for confocal imaging and nanoscopy.

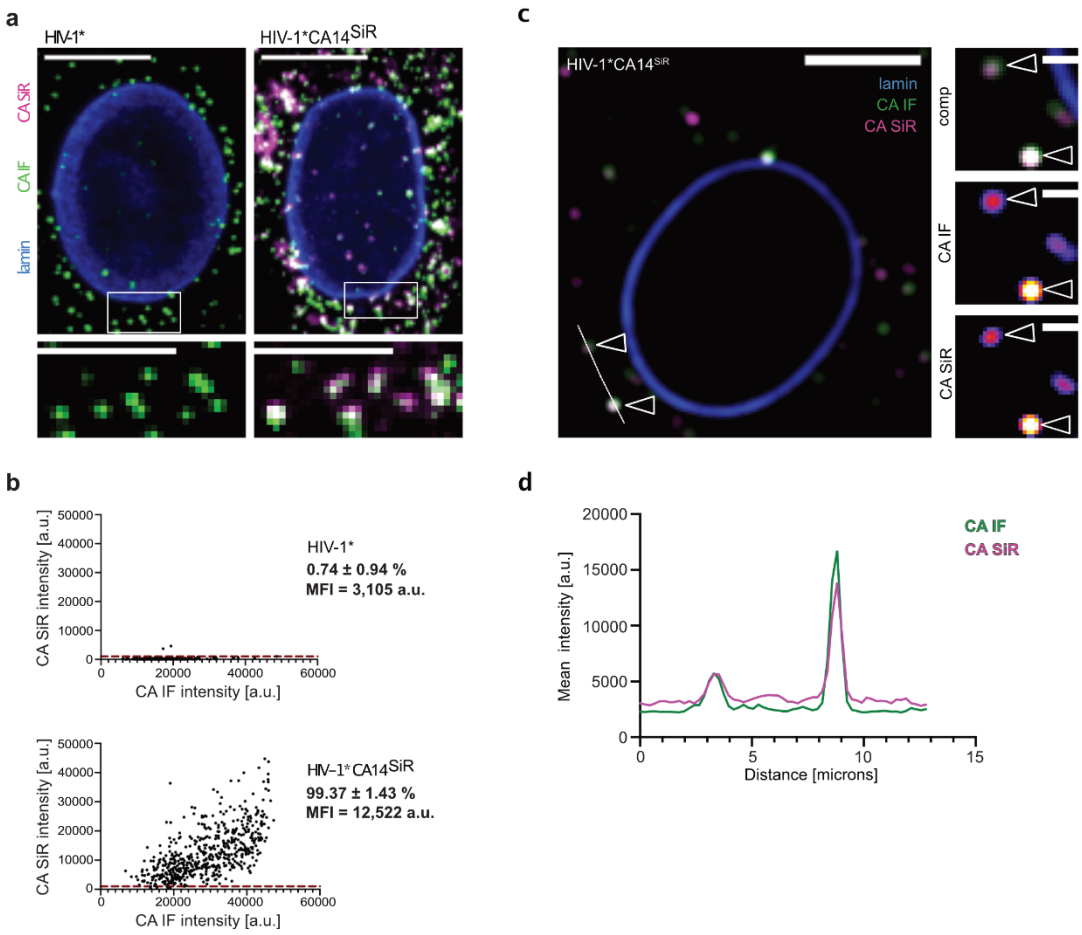


Figure 20. CA click-labeled particles in infected TZM-bl cells. a. TZM-bl cells were infected with HIV-1* or HIV-1*CA14^{SiR}. At 8 h p.i., cells were fixed, permeabilized, and immunostained against HIV-1 CA (green) and lamin A/C (blue). White boxed region is enlarged below each image. Scale bars: 10 µm (cell), 5 µm (enlargement). Mean filter and background subtraction was applied to all images for clarity as described in material and methods. **b.** Fluorescence intensity for HIV-1* (n = 463) and HIV-1*CA14^{SiR} (n = 567) particles was determined using icy software. Graph shows data of one representative experiment. Five cells were analyzed per condition. The threshold was set to t = 1,000 a.u. indicated by the red line. **c.** Single z-slice of a representative image showing a TZM-bl cell infected with HIV-1*CA14^{SiR} click-labeled particles (250 nM Tet-SiR) at 6 h p.i.. Enlargement of double-positive particles (CA IF/CA SiR) detected within the cell (white boxed arrowheads) is shown below with separated channels. Mean filter and background subtraction was applied for clarity as described in material and methods. Scale bars: 10 µm (cell), 2 µm (enlargement). **d.** Quantification of intensity profile measured across the dashed white line in (d) for CA IF (green) and CA SiR (magenta). The two peaks represent the detection of two double-positive particles with colocalization of the signals. Modified from Schifferdecker *et al.*, 2021[266] and Schifferdecker *et al.* (in preparation)[270].

2.3.5 Detection of CA IF inside the nucleus strongly depends on epitope accessibility

As part of the capsid uncoating debate, the question of whether CA enters the nucleus controversially discussed. Several publications reported different results with varying amounts of CA entering the nucleus and large variations depending on the cell type used in the respective

studies[59–61,64,103,104,106,107,114,120,121,151,298]. Detection by IF was the most commonly used approach to image HIV-1 CA, which was shown to be dependent on sample treatment and fixation protocols[64,131,150].

Detection of directly labeled CA was tested in TZM-bl cells combined with different immunostaining conditions to shed more light on nuclear CA detection by IF. Fixation of HIV-1* infected cells with paraformaldehyde (4 % PFA) alone revealed exclusively cytoplasmic CA punctae without detectable CA signal inside the nucleus (Figure 21a). The same was true in cells infected with HIV-1*CA14^{SiR}. However, distinct SiR foci were detected within the nucleus (Figure 21d). Extraction of HIV-1* infected and PFA-fixed cells with 100 % methanol for 10 min allowed detection of cytoplasmic and nuclear CA IF objects (Figure 21b), as previously reported[64]. However, infection with HIV-1*CA14^{SiR} and subsequent methanol extraction did not result in colocalization of nuclear SiR foci with CA IF (Figure 21e).

We have previously shown that nuclear CA-positive structures colocalized with large clusters of CPSF6, which may impair antibody binding to CA by shielding of epitopes[64]. Addition of high concentrations of the small molecule CA inhibitor PF74[187] dissolved these CPSF6 clusters resulting in a strong nuclear CA IF signal even without methanol extraction.

To confirm this in our experimental setup, HIV-1*CA14^{SiR} infected cells were treated additionally with 15 µM PF74 for 1 h, subsequently immunostained for CA and imaged by SDCM. Cells infected with HIV-1* indeed showed clear nuclear CA IF signals after PF74 treatment (Figure 21c). In this case, we could also identify nuclear CA SiR foci co-localizing with CA IF (Figure 21f).

The interaction of HIV-1*CA^{SiR} with CPSF6 was assessed by parallel CA and CPSF6 immunostaining (Figure 21g,h). In line with figure 21b and e, DMSO-treated cells lacked nuclear CA IF positive signals, but showed distinct CA SiR punctae (Figure 21g,i.). Strong CPSF6 signals were located together with CA SiR (Figure 21g ii.). In contrast, cells treated with PF74 were positive for nuclear CA IF/CA SiR (Figure 21h iii.) while no CPSF6 clusters were detected at CA SiR positive punctae (Figure 21h iv.).

Taken together, I can conclude that CPSF6 accumulations at subviral structures inside the nucleus can be displaced by PF74 to allow the binding of antibodies to CA epitopes (Figure 21i). However, PF74 addition alone without extraction by methanol was not sufficient to enable detection of nuclear events by IF for Tet-SiR labeled HIV-1*CA14^{SiR} particles as already observed for nuclear CA-IF detection in T cells[64]. Tet-SiR labeled particles required a combination of PF74 and methanol, while either one of these treatments alone was sufficient for WT. The reason for this difference is currently not understood.

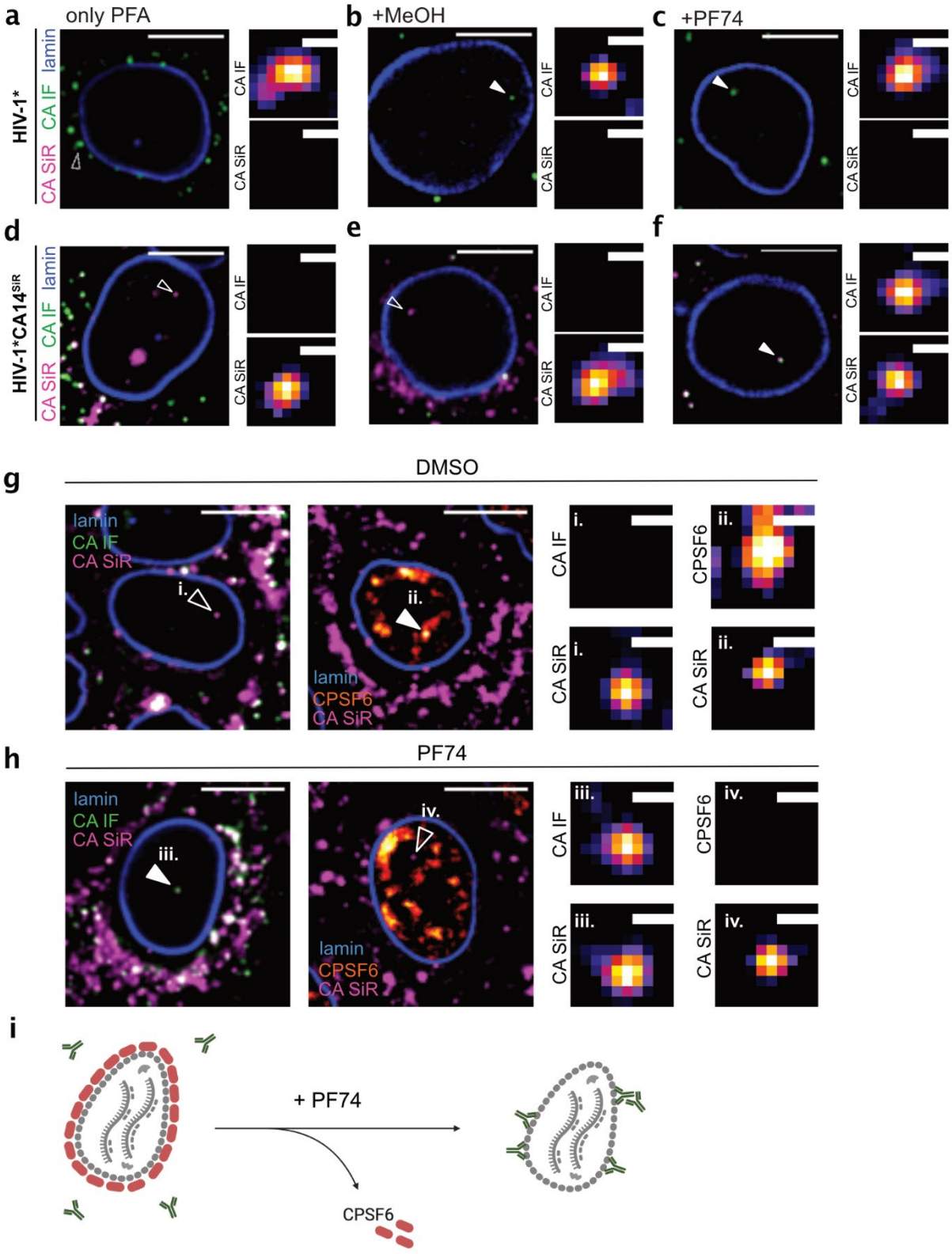


Figure 21. Nuclear CA IF detection depends on the accessibility of CA epitopes upon CPSF6 displacement. a-f. TZM-bl cells were seeded into 15-well ibidi µ-Slide angiogenesis dishes and infected for 18 h with 200 mU RT/well HIV-1* (a-c) or HIV-1*CA14^{SiR} (d-h). See enlargements of particles indicated by arrowheads next to each cell. Scale bars: 10 µm (cell) and 1 µm (enlargement). **a,d.** Cells were fixed, permeabilized, and stained against CA (green) and lamin A/C (blue) by IF. CA IF positive foci can be detected in the cytoplasm but not in the nucleus (empty arrowheads). **d.** HIV-1*CA14^{SiR} shows distinct CA SiR foci (magenta) also inside the nucleus. **b,e.** Cells were fixed, permeabilized, and extracted using 100 % ice-cold

methanol (MeOH) for 10 min. Cells infected with HIV-1* showed nuclear CA IF signals (b, white filled arrowhead) while nuclear SiR foci in cells infected with HIV-1*CA14^{SiR} were not positive for CA IF (e, white-bordered arrowhead). **c,f.** Cells were treated with 15 µM PF74 for 1 h before fixation, permeabilization and methanol extraction. Nuclear CA IF was detectable in cells infected with HIV-1* (c, white filled arrowhead). Nuclear CA SiR was also positive for CA IF (f, white filled arrowhead). **g,h.** TZM-bl cells infected with HIV-1*CA14^{TAG} were treated with DMSO or 15 µM PF74 1 h before fixation and methanol extraction at 18 h p.i. **g.** DMSO-treated cells showed nuclear CA SiR foci but no colocalization with CA IF (i.). Strong CPSF6 positive punctae colocalize with CA SiR (ii.). **h.** PF74 treated cells revealed CA IF/CA SiR double-positive foci in the nucleus (iii.) while PF74 treatment removes CPSF6 accumulations at CA SiR foci (iv.). **i.** Scheme of CPSF6 displacement by PF74. CPSF6 (red) shields the viral capsid inside the nucleus and prevents CA antibodies from binding. After adding 15 µM PF74, CPSF6 is removed from nuclear viral complexes, and antibodies (green) can bind to the CA epitopes resulting in a positive nuclear CA IF signal.

2.3.6 CA click-labeled particles enter the nucleus with a delay at the NE

After establishment of a protocol for reliable CA IF detection of nuclear CA punctae, TZM-bl cells were infected with HIV-1*CA14^{SiR} fixed at different time points under conditions suitable for detecting nuclear CA positive objects. Nuclear SiR foci detected in cells infected with HIV-1*CA14^{SiR} always colocalized with CA-IF (Figure 22a, Appendix Figure 36). Abundant nuclear viral complexes could be detected at 6 h p.i. for HIV-1*, in line with existing literature[60,64]. Quantification of cells displaying nuclear CA objects over time showed a constant high level after infection with HIV-1* already at early time points (Figure 22b, black), while the proportion of positive cells for HIV-1*CA14^{SiR} increased more slowly over time. Comparable numbers of nuclear post-fusion complexes for HIV-1*CA14^{SiR} and HIV-1* were reached at 12 - 18 h p.i. (Figure 22b, red). Neither CA IF nor CA SiR signals were detected in the nucleus of non-infected cells (Figure 22b, grey). The absolute number of subviral particles in the nucleus per cell was lower at earlier time points for the click-labeled virus, reaching comparable levels as HIV-1* at 12 h p.i., with 4.58 ± 4.12 nuclear HIV-1*CA14^{SiR} particles compared to 5.91 ± 4.11 for HIV-1* (Figure 22c).

The number of subviral particles was quantified at different intracellular localizations at the indicated time points to identify the limiting step for the observed difference in nuclear entry kinetics (Figure 22d). Most HIV-1* and HIV-1*CA14^{SiR} particles were detected at the PM or in the cytoplasm, regardless of the time point. HIV-1* and HIV-1*CA14^{SiR} particles reached the NE of cells (NE associated) already after 6 h with similar kinetics. The only apparent difference was shown inside the nucleus where the highest portion with 4.20 ± 1.80 % of nuclear events was seen at 18 h p.i. for HIV-1*CA14^{SiR} compared to 4.80 ± 3.39 % for HIV-1*already at 6 h p.i. This result argued for unaffected cytosolic trafficking but a significant delay at nuclear entry.

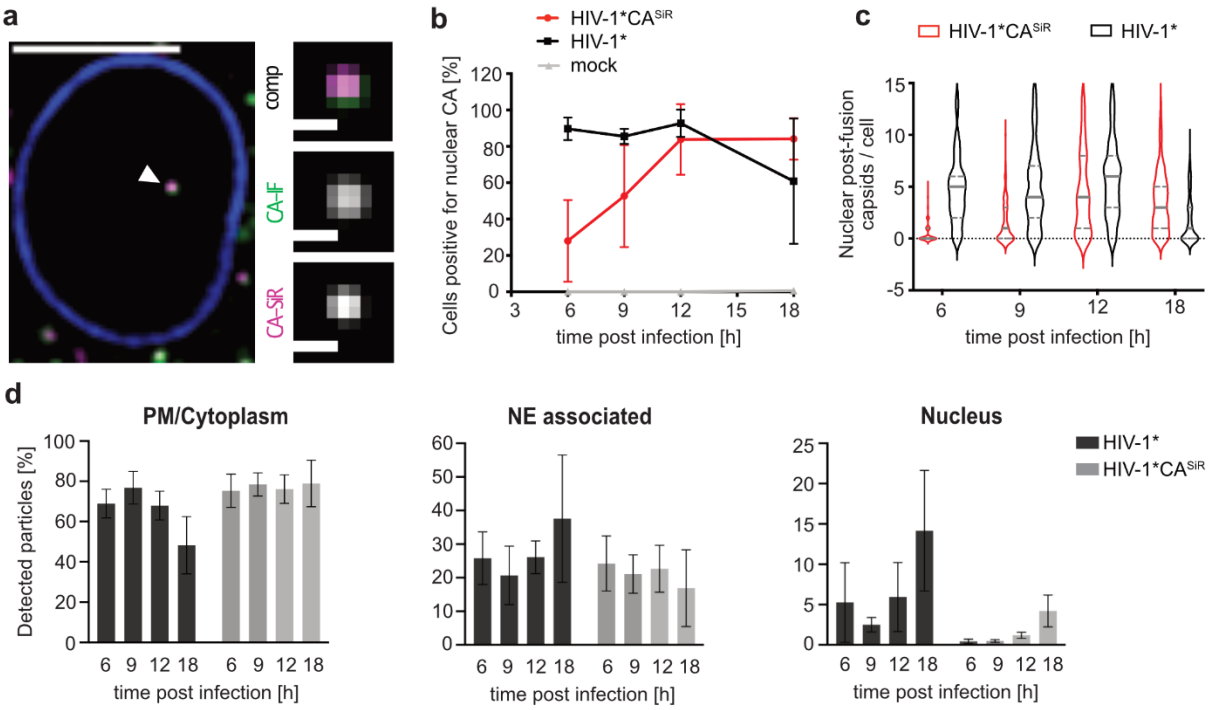


Figure 22. CA click-labeled particles infect cells with a delay in nuclear entry. TZM-bl cells were infected with HIV-1* or HIV-1*CA14^{SiR} (MOI ~0.8), treated for 1 h with PF74, and fixed at indicated time points. After permeabilization and extraction with methanol, cells were stained by IF with antibodies against HIV-1 CA (green) and lamin A/C (blue). **a.** Single z-slice of a TZM-bl cell infected with HIV-1*CA14^{SiR} and fixed at 18 h p.i. Representative image of one out of three independent experiments. Nuclear CA IF foci colocalized with CA SiR intensities. Scale bars: 10 µm (cell) and 1 µm (enlargements). **b-d.** Time-course analysis of nuclear entry. **b.** The number of cells positive for nuclear CA objects at indicated time points for cells infected with HIV-1* (CA IF; black), HIV-1*CA14^{SiR} (CA IF/CA SiR; red), and negative, uninfected control (mock; grey). n > 115 cells per condition were counted. Graph represents mean and standard deviation of three independent experiments. **c.** Absolut numbers per cell of nuclear CA foci at indicated time points for HIV-1* (black), HIV-1*CA14^{SiR} (red). n > 120 cells counted per condition. Graph represents median and quartile lines of three independent experiments. **d.** Quantification of HIV-1* (dark grey) and HIV-1*CA14^{SiR} (light grey) particles at indicated time points detected in different cellular compartments (PM/cytoplasm, NE associated, nucleus). Graph shows mean values and standard deviation of three independent experiments. n > 20 cells were counted per time point. Figure from Schifferdecker *et al.*, 2021[266].

2.3.7 PF74 treatment did not induce capsid lattice dissociation

Recently the effect of PF74 was investigated using GFP as a capsid content marker. The results argued for only partial permeabilization of the capsid core with a subsequent stabilizing effect of PF74 on the remaining lattice[111,195]. Moreover, CA IF detection did not reveal significant loss of CA signal upon PF74 treatment[64]. On the contrary, CA IF at nuclear objects was enhanced by PF74 treatment. However, the IF approach always relies on accessibility of the epitopes and results may thus be influenced by unknown variables.

The directly labeled CA variant developed in this thesis, however, enabled the investigation of the PF74 effect without such confounding variables. Adding PF74 to isolated HIV-1*CA14^{SiR} particles imaged on a coverslip did not change their SiR intensity colocalizing with CA IF positive objects (Figure 23a). However, since the viral membrane is still present, dissociation of the capsid into its subunits would not be visible. Therefore, TZM-bl cells were infected with HIV-1*CA14^{SiR} and treated the cells with PF74 for 1 h prior to fixation. As shown in Figure 23b CA SiR intensities of nuclear subviral structures detected upon PF74 treatment were similar to those in DMSO treated control cells indicating that PF74 treatment does not lead to capsid disassembly. This observation contrasts with results from a recently published study using a directly labeled CA (GFP-CA)[60]. In this study, the authors observed rapid GFP-CA loss from nuclear complexes upon PF74 addition. As this labeling method relies on WT complementation, it may be assumed that the tagged CA molecules were not fully integrated into the mature lattice, so the observed GFP-CA intensity loss may not reflect changes of the capsid.

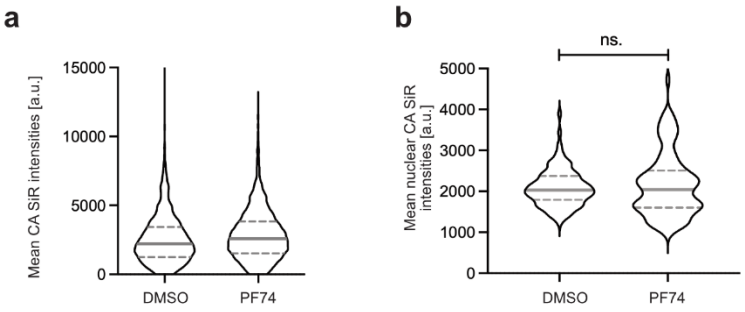


Figure 23. The effect of high PF74 concentrations on capsid integrity. a. Quantification of mean CA SiR intensities of complete virions. HIV-1*CA14^{SiR} particles were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes, treated with DMSO or 15 µM PF74 for 1 h, fixed, permeabilized, stained against HIV-1 CA by IF used to automatically detect viral particles, and imaged by SDCM. **b.** Quantification of mean CA SiR intensities of particles detected in the nucleus of TZM-bl cells. Cells were infected with HIV-1*CA14^{SiR} for 17 h, treated with DMSO or 15 µM PF74 for 1 h, fixed, permeabilized, extracted with methanol, and imaged by SDCM. Graph shows mean values (DMSO: 2,086.83 ± 456.35, n = 100; PF74: 2,164.51 ± 783.45, n = 100) and quartile lines. The difference between the two samples was not significant ($p > 0.05$ (ns), two-tailed Student's t-test). Figure from Schifferdecker *et al*., 2021[266].

2.3.8 Cone-shaped capsids cluster inside the nucleus in a dose-dependent manner

The development of CA directly labeled with synthetic fluorophores suitable for super-resolution imaging provided the possibility for further characterization of CA SiR positive structures by STED nanoscopy. Previous studies using EM and CLEM revealed the accumulation of clusters of several closely adjacent capsid-like structures in the nuclei of infected cells[64]. For a quantitative assessment

of this phenomenon, TZM-bl cells were infected at two different multiplicities of infection (MOIs) (~ 0.8, as usually employed in my experiments, and ~ 8), fixed and stained against CA by IF at 18 h p.i. (Figure 24a). Diffraction-limited nuclear subviral objects were identified under both conditions by CA IF and CA SiR labeling. STED nanoscopy of these punctae, with a lateral resolution of < 50 nm, revealed mostly single CA positive objects at the position of diffraction-limited punctae at an MOI ~ 0.8 (Figure 24b, left panel). In contrast, diffraction-limited clusters often compromised 2 - 4 separated objects in cells infected with a higher MOI of ~ 8 (Figure 24b, right panel). Quantification of diffraction-limited foci showed that 88 % of nuclear structures represented single objects after infection with a low MOI ~ 0.8, but in cells infected with a higher MOI ~ 8 almost half of nuclear punctae corresponded to clusters of CA-containing structures, consistent in size with single capsid cores (Figure 24c). Thus, CA objects clustered inside the nucleus in a dose-dependent manner. Based on these data, the vast majority of analyzed signals in all previous and following experiments, which were performed at an MOI of ~ 0.8, correspond to single capsids.

2.3.9 Most viral particles enter TZM-bl cells *via* endosomal uptake

STED nanoscopy was also applied to CA structures inside the cytoplasmic region of infected cells. First, the PM staining with a membrane-binding fluorophore cysteine lysine palmitoyl group (mCLING)[299] ATTO488, used in previous studies to identify HIV-1 particles contained in endosomal structures[131] was validated in TZM-bl cells. Cell viability measurements showed that more than 90 % of the cells survived staining with 0.25 - 2 µM for 30 min at 16° C (Figure 25a). To ensure strong and sufficient staining of the PM, the cells were stained with 2 µM mCLING ATTO488 before infection with HIV-1*CA14^{SiR} and fixed at 3 h p.i. (Figure 25b). Viral particles in the cytoplasmic region were identified by detecting CA SiR punctae. Colocalization with mCLING488 indicated particles enclosed in endosomal vesicles as a pre-fusion state of the virus. Quantification of CA SiR/mCLING488 colocalization showed that under our conditions > 95 % of incoming viral structures were trapped in endosomes (Figure 25c).

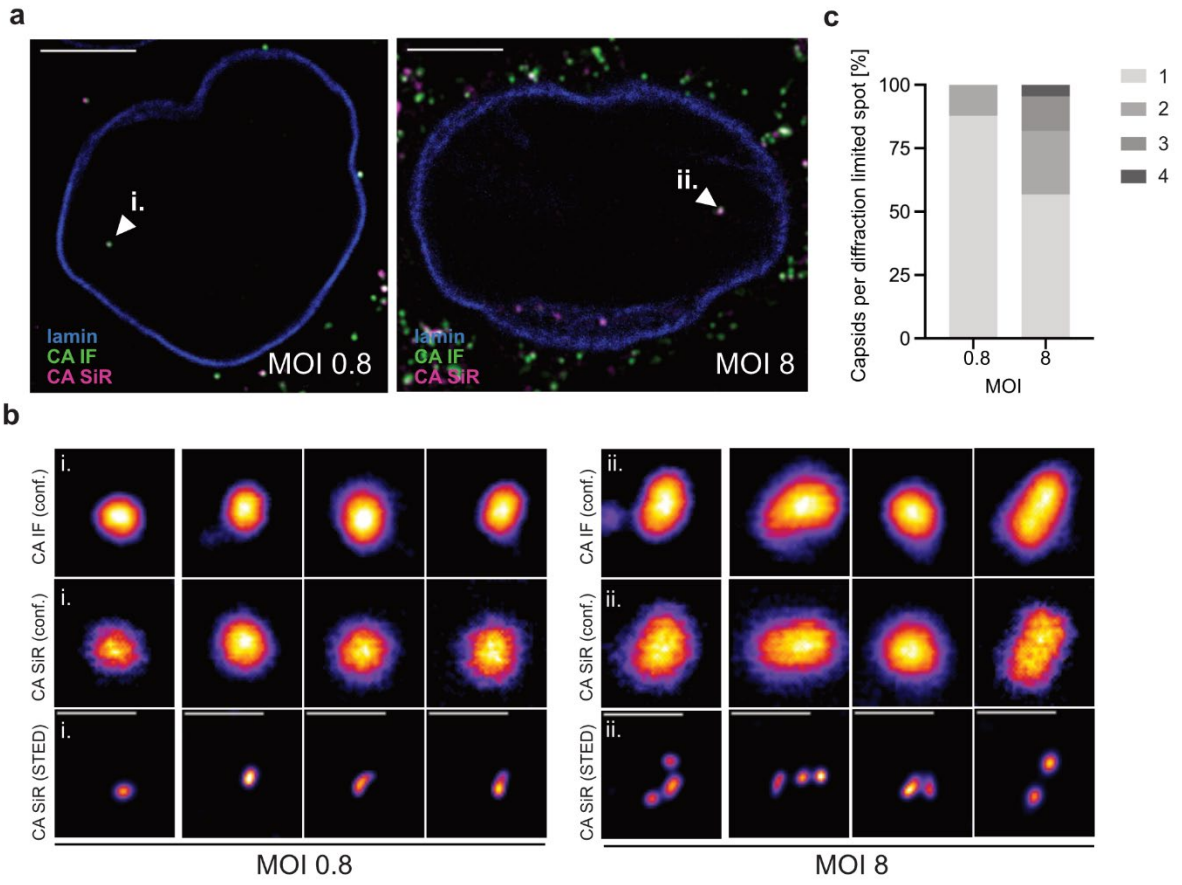


Figure 24. Dose-dependent clustering of nuclear capsids detected in super-resolution. TZM-bl cells, infected with HIV-1*CA14^{SiR} at an MOI ~ 0.8 (left) or MOI ~ 8 (right), were treated with 15 µM PF74 for 1 h, fixed, permeabilized, and extracted with methanol. After IF staining with antibodies against HIV-1 CA (green) and lamin A/C (blue), cells were imaged by STED nanoscopy. **a.** Representative single z-slices of two cells infected at different MOIs. White arrowheads indicate nuclear CA IF/CA SiR positive structures (i.-ii.). Scale bars: 10 µm. **b.** Enlarged nuclear objects in cells such as indicated in (a) by an exemplary cell. Detected particles are shown in CA IF confocal (upper row), CA SiR confocal (middle row), and CA SiR in STED mode (lowest row). Diffraction-limited spots (confocal mode) were resolved into individual capsid structures in STED mode. Scale bars: 500 nm. **c.** Quantification of capsids per diffraction-limited spot for MOI ~ 0.8 (n = 33) and MOI ~ 8 (n = 44) as shown in (b). Graph shows data of two independent experiments. Figure from Schifferdecker *et al*., 2021[266].

STED nanoscopy of cytoplasmic diffraction-limited CA SiR objects in cells infected at an MOI ~ 0.8 (Figure 25d) or MOI ~ 8 (Figure 25e) revealed that most cytoplasmic punctae represented single CA positive objects. Infection with a lower viral dose (MOI ~ 0.8) resulted in detection of individual particles in ~ 95 %, while ~ 5 % were resolved as clusters of not more than two objects (Figure 25f).

Although propagated nuclear clustering was found during infection with a higher MOI (MOI ~ 8), most cytoplasmic diffraction-limited foci in cells infected at the same MOI were also resolved into single capsids (89.02 %) with an only slightly higher number of clusters of two or three closely adjacent objects (10.98 %, Figure 25f).

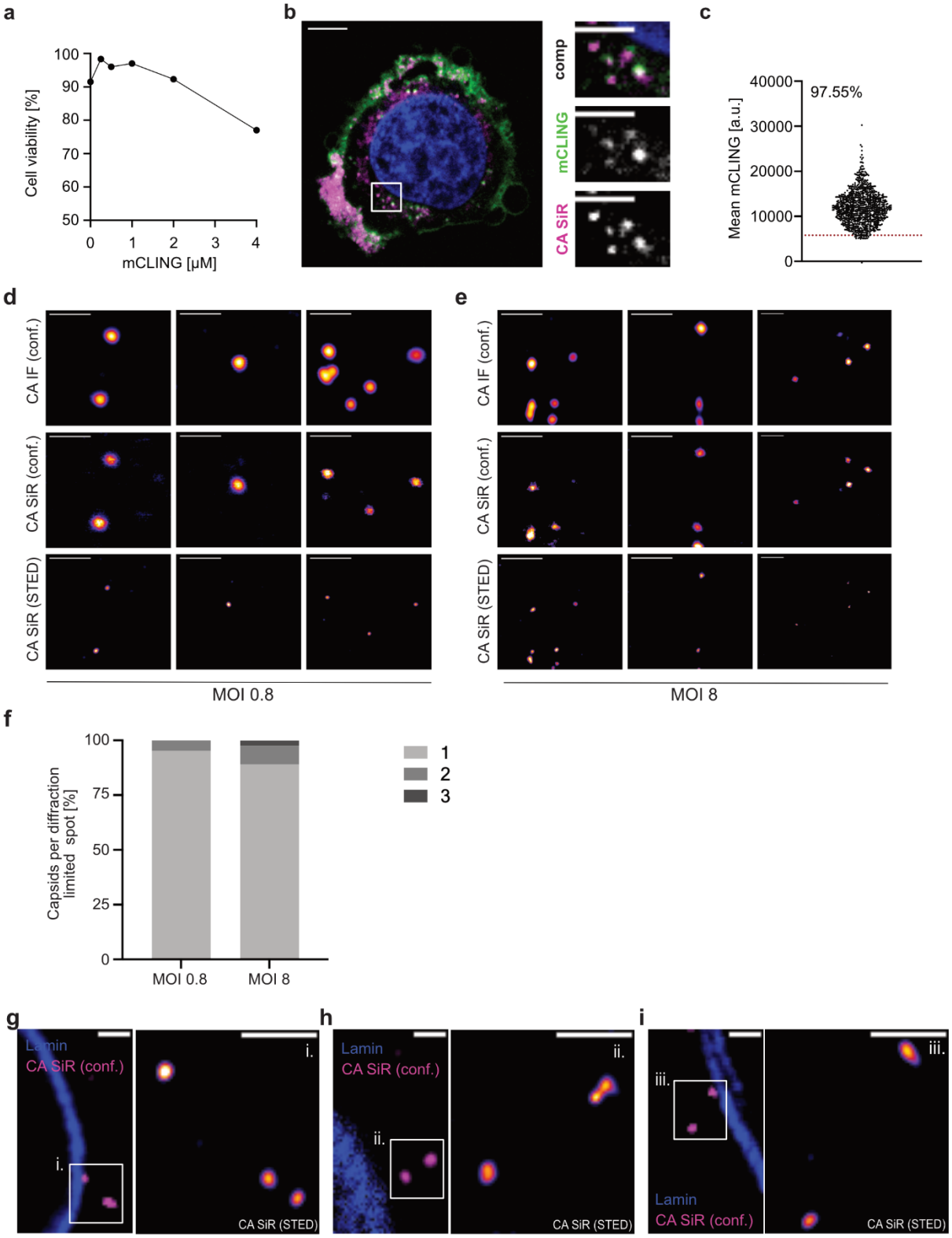


Figure 25. Incoming HIV-1 detected in endosomal vesicles. a. Viability measurement of TZM-bl cells. mCLING ATTO488 was titrated on TZM-bl cells and incubated at 16° C for 30 min. The number of cells was quantified using Trypan blue staining and an TC20 automated cell counter. Graph shows mean values of technical duplicates. **b,c.** TZM-bl cells were incubated with 2 µM mCLING ATTO488 at 16° C for 30 min and

infected with HIV-1*CA14^{SiR} (MOI ~ 0.8) at 37° C. At 3 h p.i., cells were fixed with 4 % PFA + 0.2 % GA and stained with Hoechst. **b.** Representative single z-slice of a cell infected with HIV-1*CA14^{SiR} at 18 h p.i. Boxed area enlarged on the right with colocalization of mCLING ATTO488 (green) and CA SiR (magenta). Scale bars: 10 µm (cell) and 5 µm (enlargements). **c.** Quantification of mean mCLING ATTO488 intensities of CA SiR positive objects in the cytoplasm (n > 900). The threshold was set to t = 8,000 a.u. based on visual inspection using Imaris software. Every detected object above the threshold was counted as mCLING ATTO488-positive. **d-h.** TZM-bl cells, infected with HIV-1*CA14^{SiR} (MOI ~ 0.8 and ~ 8), were fixed at 18 h p.i. and immunostained against HIV-1 CA and lamin A/C. **d,e.** STED nanoscopy of CA IF/CA SiR positive structures in the cytosol. Cytoplasmic structures were identified in confocal mode by CA IF (upper row) and CA SiR (middle row). Diffraction-limited spots in the SiR channel were resolved in STED mode (lower row). Scale bars: 1 µm. **f.** Quantification of capsids per diffraction-limited spot resolved by STED. Graph shows the number of detected objects resolved in single capsid structures (n(MOI ~ 0.8) = 79, n(MOI ~ 8) = 72), two individual structures (n(MOI ~ 0.8) = 4, n(MOI ~ 8) = 7) or three capsids (n(MOI ~ 0.8) = 0, n(MOI ~ 8) = 2) per diffraction-limited foci. **g-i.** Cytoplasmic viral structures were detected by CA SiR (magenta) trafficking towards the NE (lamin, blue). In each image, the CA SiR channel of the boxed regions (i.-iii.) is shown enlarged on the right in STED mode. Scale bars: 1 µm (overviews) and 500 nm (enlargements STED mode). Modified from Schifferdecker *et al*., 2021[266].

HIV-1 particles hijack the cellular cytoskeleton to facilitate trafficking on microtubule and microfilament networks[28]. Imaging cytoplasmic particles on the way towards the nucleus often showed stringing objects in proximity of the NE (Figure 25g-i). Enlargements of detected punctae in STED mode could show such cytoskeleton-directed trafficking of individual particles on common routes [300]. Important to note here that most objects detected in the cytoplasmic region in another experiment with mCLING staining were identified to be endosomal. Thus, we cannot conclude for sure that the imaged punctae represent post-fusion events during cytoskeleton-directed trafficking.

2.3.10 Direct CA labeling enabled relative quantification of CA in different cellular compartments

The previous experiments showed that the developed CA labeling strategy allows to label the vast majority of subviral structures while infectivity is retained. This feature stands in contrast to the necessity of WT complementation used in other, more intrusive labeling methods[60,121]. Therefore, this method permits to draw conclusions about the CA content of subviral structures in different cellular compartments based on intensity measurements. Previous analysis of click-labeled HIV-1*CA14^{SiR} particles in cells (MOI ~ 0.8) ensured that diffraction-limited punctae in the cytoplasm and the nucleus represented single virions or subviral particles.

Complete virions comprise ~ 2,400 CA molecules[155] (Figure 26a). In contrast, post-fusion capsids inside the nucleus represent only ~ 60 % of total CA molecules, namely those incorporated into the

mature core lattice (~ 1,500 molecules)[152,155,156] (Figure 26b). TZM-bl cells were infected at an MOI ~ 0.8 and fixed at 18 h p.i. to measure intensities of individual objects associated with the cell (at the membrane or within the cytosolic region) or inside the nucleus and determine the amount of CA associated with nuclear particles compared to complete virions at the PM or within endosomes. Quantitation of CA SiR intensities (Figure 26c) showed ~ 56 % mean intensity for subviral particles inside the nucleus (9,835 ± 5,708 a.u.) compared to intensities of complete virions in the cytoplasmic region (17,649 ± 11,663 a.u.). Based on the assumption that all CA molecules within the virion were labeled with equal probability, the relatively high SiR intensities of post-fusion nuclear objects indicates that nuclear capsids contain an amount of CA similar to a complete mature capsid structure.

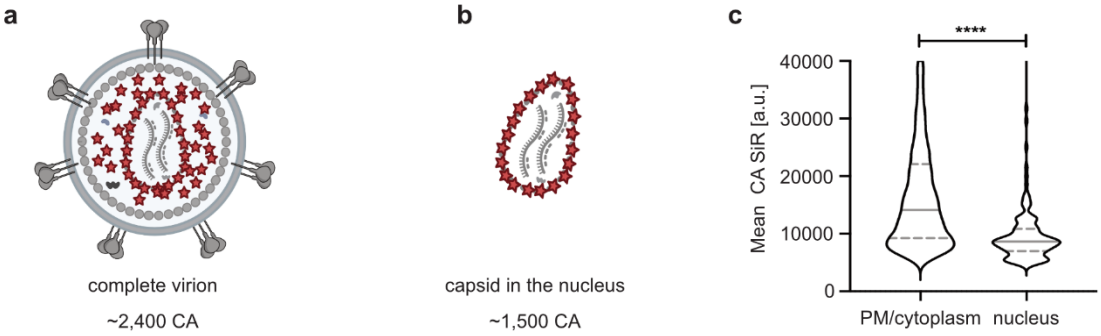


Figure 26. Subviral particles inside the nucleus contain high amounts of CA. a,b. Scheme of a complete virion (a) and the mature capsid (b) comprising different amounts of CA (red stars). **c.** Quantification of mean CA SiR intensities of viral structures detected in the cytoplasm and nucleus of HeLa-derived cells by IF. TZM-bl cells, infected with HIV-1*CA14^{SiR} (MOI ~ 0.8), were treated with 15 µM PF74, fixed, permeabilized, immunostained against HIV-1 CA and lamin A/C, and imaged by SDCM. Mean CA SiR intensity was measured for particles detected at the PM/cytoplasm (mean = 17,649.22 ± 11,663.47 a.u.) and inside the nucleus (mean = 9,835.08 ± 5,708.14 a.u.). Violin plots show median and quartile lines ($n_{PM/cytoplasm}$ = 6,441, $n_{nucleus}$ = 135; combined data from three experiments). Significance was assessed by a two-tailed Student's t-test ($p < 0.001$). Modified from Schifferdecker *et al.*, 2021[266].

2.3.11 Large amounts of CA are associated to subviral complexes detected in the nucleus of primary CD4$^+$ T cells

In contrast to HeLa-derived cell lines, T cell lines and primary T cells demonstrated comparably low nuclear CA IF signals in previous studies arguing for cell-type-specific differences in the site of uncoating. Since we were able to show that nuclear CA IF detection in T cells depends on sample treatment requiring methanol extraction or PF74 treatment to expose epitopes[64], I was interested in the detection of SiR-labeled particles also in primary cells.

To follow directly labeled incoming capsids in a more relevant cell type, primary $CD4^+$ T cells, as a main natural target of HIV-1, were infected with HIV-1*CA14^{SiR}. Infection of activated primary $CD4^+$ T cells with HIV-1*CA14^{SiR} at an MOI ~ 0.8 (determined in TZM-bl cells) for 24 h revealed nuclear CA SiR foci mostly lacking CA IF staining (Figure 27a). As already shown before for HeLa-derived cells (*cf*. 2.3.5; Figure 21), nuclear CA IF detection required the removal of the dense CPSF6 coat by PF74 to allow binding of the antibody and detection of CA SiR/CA IF colocalization (Figure 27b and Appendix Figure 37). Upon treatment of infected cells for 1 h with 15 µM PF74 before fixation, 81.25 % of nuclear SiR punctae were also found to be CA IF positive, while this was only the case for ~ 9.09 % in DMSO treated control cells (Figure 27c). The following experiments were performed under conditions allowing the detection of nuclear CA by IF.

Quantitative analysis in primary cells from six different blood donors showed that HIV-1*CA14^{SiR} and HIV-1* enter the nucleus with similar efficiency at late time points (Figure 27d). CA SiR positive objects in the nucleus of these cells, infected at a low MOI ~ 0.8, displayed similar mean intensities (12,485 a.u.) compared to those measured in the nucleus of HeLa-derived TZM-bl cells (Figure 26c; 9,835 a.u.), suggesting the transfer of nearly complete capsids into the nucleus in primary cells in line with recently published findings of largely complete nuclear capsid structures in TZM-bl cells[60,61,64]. CA SiR positive objects detected in the cytoplasmic region of infected $CD4^+$ T cells were not analyzed due to high SiR background in the smaller cytosol. This high background prevented a precise localization of subviral objects as well as a reliable calculation of SiR intensities. Infections at higher MOI (~ 8) yielded 3-fold increased CA SiR intensities (39,502 a.u.), most likely reflecting capsid clustering, as detected in the previous STED experiments (*cf*. 2.3.8; Figure 24).

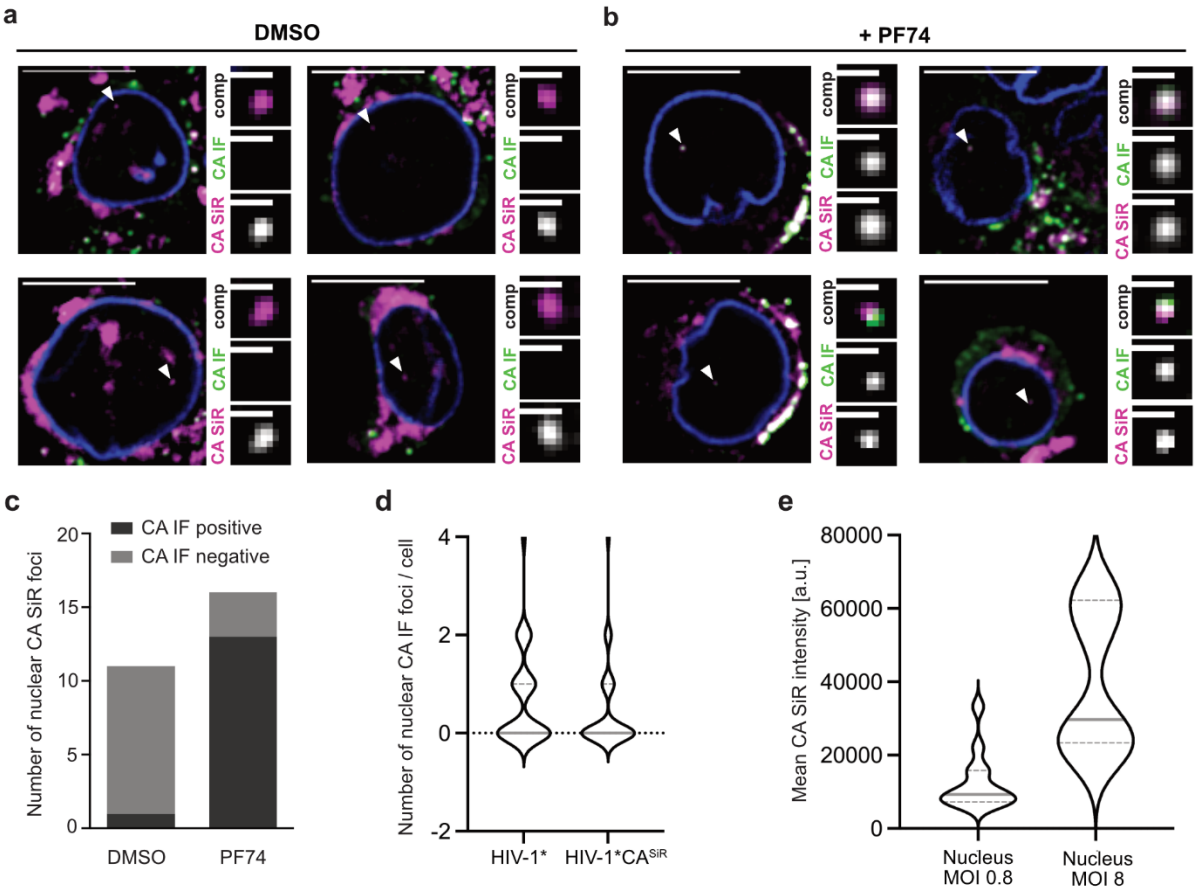


Figure 27. A high amount of CA associated to subviral objects is transmitted into the nucleus of primary CD4⁺ T cells with dose-dependent capsid clustering. a,b. Representative images of single z-slices of three independent experiments. Activated primary human CD4⁺ T cells were isolated and infected with HIV-1*CA14^{SiR} (MOI ~ 0.8, determined in TZM-bl cells). At 23 h p.i., cells were treated with DMSO (a) or 15 µM PF74 (b) for 1 h before fixation, permeabilization, extraction with methanol, and staining against HIV-1 CA and lamin B by IF. White arrowheads indicate nuclear viral structures, which are shown enlarged next to each cell. Nuclear CA SiR foci, detected in the DMSO and PF74 sample, only became CA IF positive after PF74 treatment and CPSF6 removal (b, enlargements). Mean filter and background subtraction was applied for clarity. Scale bars: 10 µm (cells) and 1 µm (enlargements). **c-e.** Quantitative analysis of the experiment explained in (a, b). Graphs show pooled data of six different blood donors. **c.** The number of CA SiR foci detected in the nucleus of activated CD4⁺ T cells. In cells treated with DMSO, 1 of 11 foci were positive (dark grey), while 10 were negative (light grey) for CA IF. After treatment with 15 µM PF74, only 3 of 16 nuclear SiR events were counted as CA IF negative and 13 CA IF positive. **d.** Quantification of nuclear CA IF spots per cell after infection with HIV-1* (mean = 0.85, n = 35 cells) or HIV-1*CA14^{SiR} (mean = 0.51, n = 73 cells). All events counted for HIV-1*CA14^{SiR} were double-positive: CA IF/CA SiR. Violin plots show median and quartile lines. **e.** Quantification of mean CA SiR intensities of viral structures detected in the nucleus by CA IF at an MOI ~ 0.8 (mean = 12,485 ±7,445 a.u., n = 13) and MOI ~ 8 (mean = 39,502 ± 18,025 a.u., n = 7). Violin plots show median and quartile lines. Modified from Schifferdecker *et al.*, 2021[266].

To investigate the ultrastructure of CA SiR containing fluorescent objects CLEM analysis was performed in more detail. SupT1 cells were treated overnight with aphidicolin (APC) to prevent cell division. After incubation cells were infected with HIV-1*CA14^{SiR}. To enhance the chances to detect nuclear events, prebinding was performed for 3 h at 16° C preventing cellular uptake of

particles[301,302] followed by temperature shift to 37° C for synchronized infection[131]. Subsequently, high-pressure freezing and freeze substitution of the specimens was performed. 250 nm-thick resin sections were imaged by SDCM to identify CA SiR positive objects, which were then analyzed by CLEM and electron tomography (CLEM-ET). Figure 28a (left) illustrates that CA SiR intensity is bright enough inside the nucleus in resin embedded sections. Correlative ET analysis of the region of interest indicated by the white box and enlarged next to the cell overview revealed that the CA SiR containing object corresponded to a cluster of electron-dense structures with the morphology of capsid cones identified in different tomographic sections (Figure 28b). 3D rendering of tomograms revealed largely intact and closely adjacent capsid-like objects co-localizing with CA SiR (Figure 28c). These data are consistent with another publication from our lab, which I participated, using the ANCHOR vDNA labeling system showing tomograms of broken cone-shaped structures releasing an elongated chromatin-like electron-dense structure at the site of fluorescently tagged IN (IN.FP) and OR3 signals[64].

However, these data presented in this thesis show for the first time that intact cone-shaped capsids can be found inside the nucleus of APC-treated primary T cells. Therefore, they likely have passed through the NPC, although it cannot be directly proven that this respective cell did not escape the block to cell division. The correlation of conical capsid structures to directly labeled CA particles provides direct evidence for high numbers of full capsid complement in the nucleus of T cell.

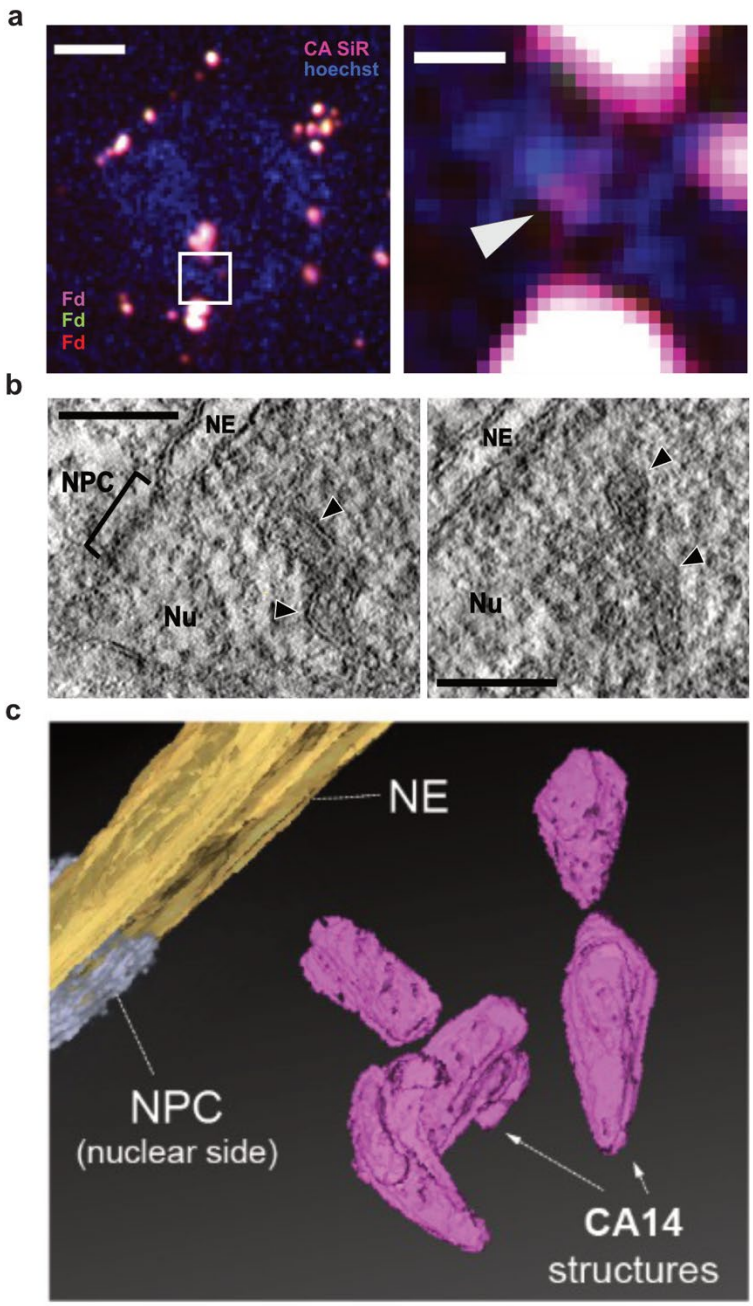


Figure 28. Largely intact capsids enter the intact nucleus of T cells. CLEM-ET analysis of nuclear capsid events in SupT1 cells. Cells were treated with 1 µM aphidicolin (APC) for 16 h to block cell division and infected with HIV-1*CA14^{SiR} (MOI ~ 0.4, determined in TZM-bl cells) for 24 h before cryo-immobilizion by high-pressure freezing. Cells were then freeze-substituted and processed for CLEM and ET (*cf.* 5). **a.** 250-nm thick resin section imaged by SDCM. Cells were stained with Hoechst, and multi-fluorescent fiducials (Fd) were added for correlation. The image shows a nuclear CA SiR positive structure in infected SupT1 cell. The boxed region indicated in the image is enlarged to visualize the CA SiR-dense structure (magenta) within the Hoechst staining (blue) by the white arrowhead. Scale bars: 1 µm (cell) and 200 nm (enlargement). **b.** Two computational slices through the tomographic reconstruction at the correlated region are shown in the enlargement image in (a, right). Tomograms show a cluster of capsid-related objects (black arrowheads). NPC: nuclear pore complex, NE: nuclear envelope, Nu: nucleus. Scale bars: 100 nm. **c.** Segmented and iso-surfaced rendering of the detected capsid structures shown in (b). Cyan: NPC, yellow: NE, magenta: capsids. CLEM-ET

has been performed by Vojtech Zila (University Hospital Heidelberg). Modified from Schifferdecker *et al.*, 2021[266].

2.3.12 Generation of directly labeled CCR5-tropic virus variant allows infection of monocyte-derived macrophages

Besides T cells monocytes and macrophages are primary targets for HIV-1 infection *in vivo* with a considerable role for disease progression[303]. While the main HIV-1 co-receptor in T cells is CXCR4, macrophages express CCR5 as a co-receptor. In order to employ the click-labeled virus in macrophages, a CCR5-tropic CA click-labeled virus variant was generated based on the HIV-1-NL4-3-AD8 strain. Mutation of the *vpr* amber stop codon into an opal stop codon resulted in the proviral plasmid named AD8* and the plasmid with an amber stop codon at residue 14 of the CA was termed AD8*CA14^{TAG}.

Infectivity of this virus variant (suitable to productively infect macrophages) was determined via titration of labeled AD8* and AD8*CA14^{SiR} viral particles on TZM-bl cells (Figure 29a). The proportion of infected cells was quantified and normalized to the control infected with AD8*, resulting in ~ 60 % relative infectivity for AD8*CA14^{SiR}. Infectious R5 tropic particles were immobilized, immunostained against CA by IF, and imaged by SDCM to determine labeling efficiency (Figure 29b,c). Quantitative analysis of CA IF/CA SiR colocalization revealed high labeling of > 80 % for AD8*CA14^{SiR} particles (Figure 29d; MFI = 5,845 a.u.) and 16 % unspecific background labeling with AD8* (Figure 29e; MFI = 3,565 a.u.).

Following characterization of click-labeled virus particles, monocytes were isolated from two different blood donors and differentiated into MDMs. Infection of MDMs showed that labeled particles were efficiently detected in the cellular environment. AD8*CA14^{SiR} and AD8* particles were detected mainly in the cytoplasmic region and near the NE (Figure 30). CA IF positive punctae in cells infected with AD8*CA14^{SiR} were all positive for CA SiR. In contrast to T cells, MDMs revealed low background intensity in the SiR channel (Figure 30a i.-viii.). In 20 imaged cells from a single experiment, only one subviral structure with high CA SiR intensity was detected inside the nucleus (Figure 30a iii.), while nuclear events could not be detected for AD8* (Figure 30b). Uninfected cells showed unspecific PM staining in the CA IF channel and low background in the nucleus and cytoplasm (Figure 30c). A quantitative analysis of CA SiR intensities will require the collection of a large data set. Such analyses will reveal whether the viral dose-dependent capsid-clustering phenotype observed in HeLa-derived cells and primary T cells is also present in MDM.

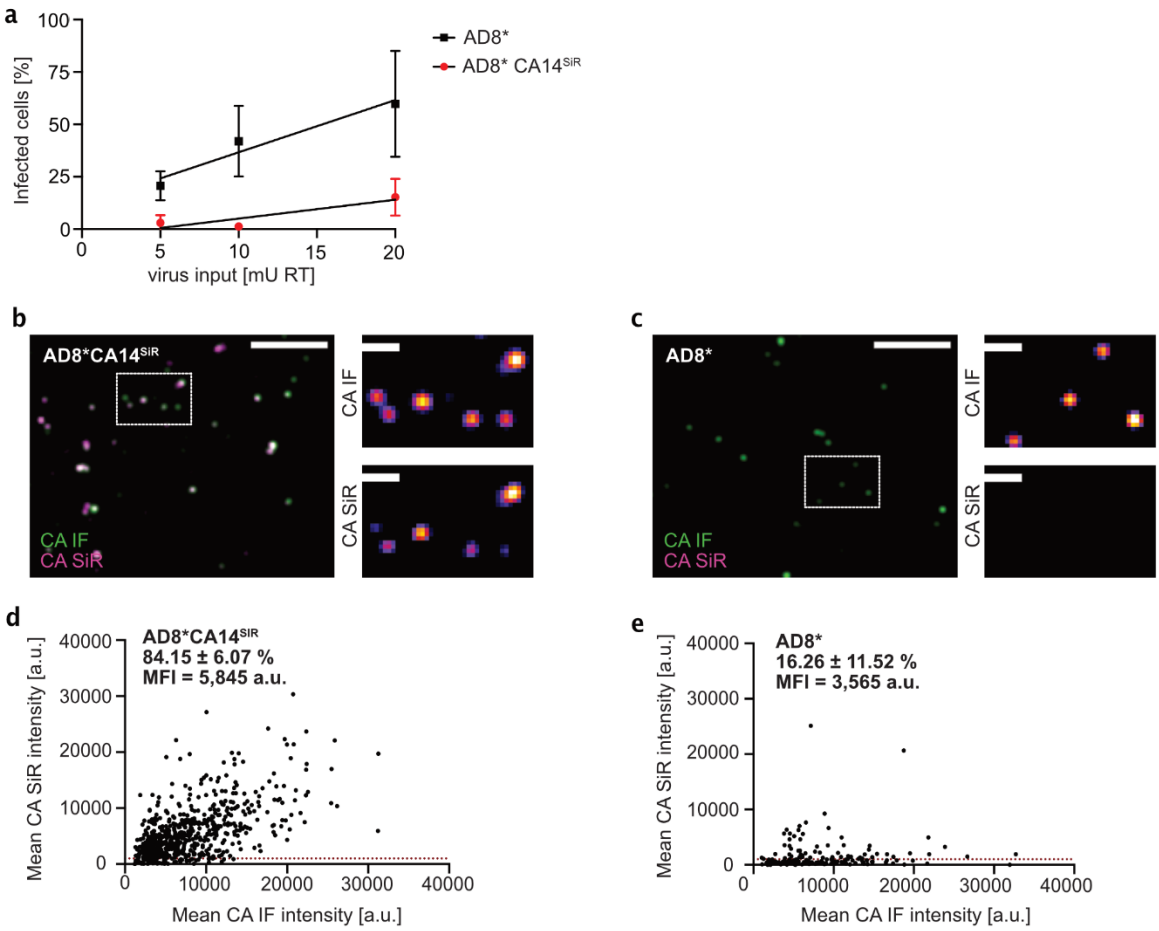


Figure 29. CCR5-tropic CA click-labeled virus variant used to infect MDM. Analysis of infectivity and labeling efficiency of the virus variant AD8*CA14^{SiR}. **a**. AD8* and AD8*CA14^{SiR} were titrated on TZM-bl cells. At 6 h p.i., new media supplemented with 50 µM T-20 (fusion inhibitor) was added. At 48 h p.i., cells were fixed, permeabilized, stained by IF using an antibody raised against HIV-1 MA, and imaged by SDCM. Infected cells were counted based on the MA signal. Graph shows mean values and standard deviation of 10 FOVs per condition (n = 200 - 1,000 cells counted per condition in one experiment). Lines indicate linear regression based on the mean values. **b-e.** Labeling efficiency of AD8*CA14^{SiR} from one experiment. Particles were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes, fixed, permeabilized, stained against HIV-1 CA by IF and imaged by SDCM. **b.** AD8*CA14^{SiR} particles (n = 788) were detected in a single experiment by CA IF, showing specific colocalization with SiR signals. **c.** CA IF detected AD8* particles (n = 306) did not colocalize specifically with the SiR channel (magenta). Scale bars: 10 µm (overview) and 2 µm (enlargements). **d,e.** Quantification of mean CA SiR intensities of particles shown in (b) and (c). Particles were detected by CA IF. Mean CA SiR intensities were plotted against mean CA IF intensities. The threshold was set to t = 1,000 a.u.. All particles with SiR intensities above the background were counted as efficiently labeled by Tet-SiR.

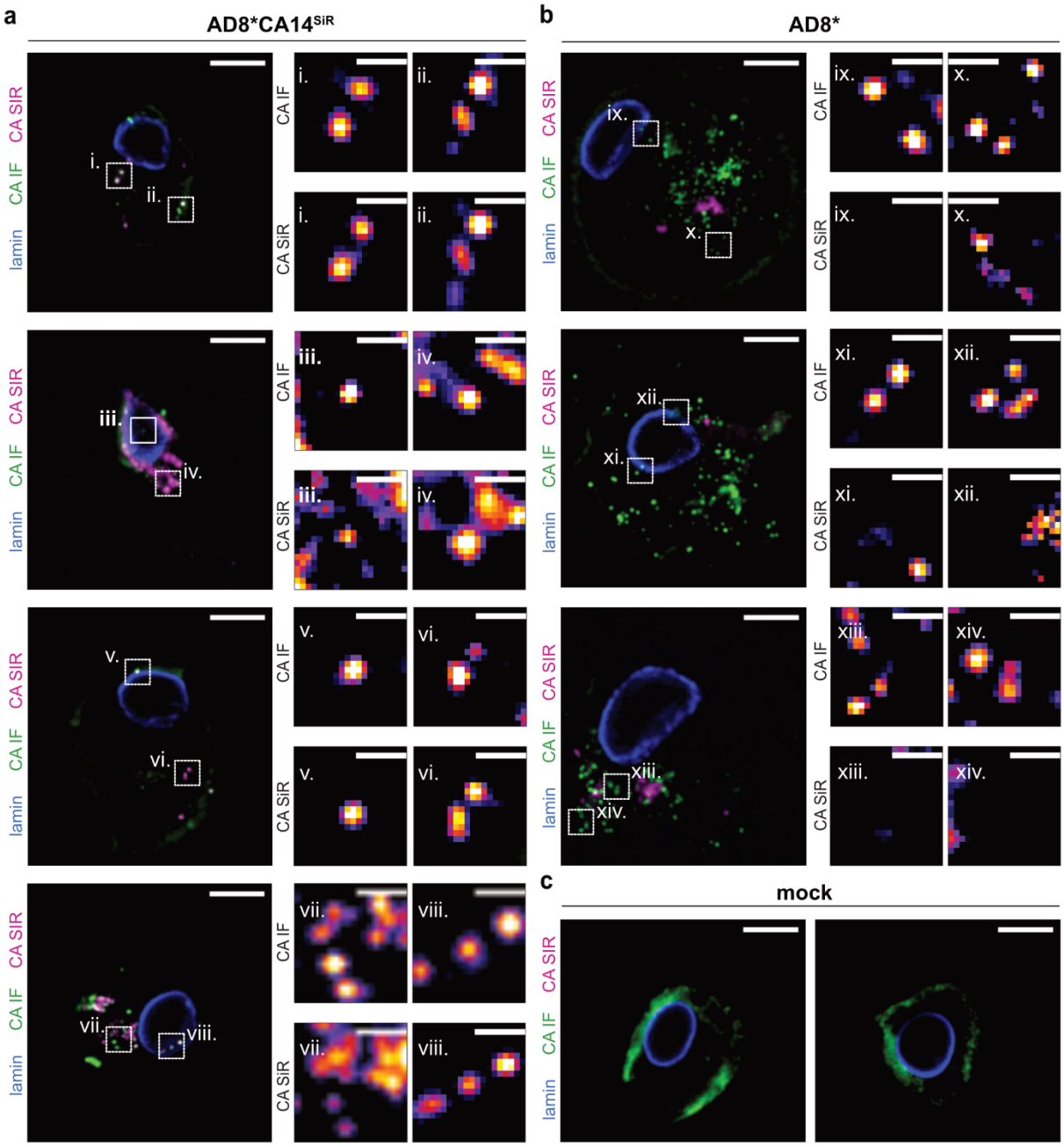


Figure 30. CA click-labeled objects can be identified in MDM. MDMs were infected with AD8*CA14^{SiR} (a) or AD8* (b) at MOI ~ 0.8 (determined in TZM-bl cells). At 72 h p.i., cells were fixed, permeabilized, extracted with methanol, immunostained against HIV-1 CA (green) and lamin A/C (blue), and imaged by SDCM. All white boxed regions in the overview cell images are shown enlarged next to the respective image. Scale bars: 10 µm (cells) and 2 µm (enlargements). **a.** In AD8*CA14^{SiR} infected cells, double-positive particles (CA IF/CA SiR (magenta)) were detected mainly at the PM/in the cytoplasm (i., ii., iv., vi., vii.) or the NE (v., viii.). A CA IF/CA SiR structure was detected inside the nucleus (iii.). **b.** AD8* infected cells showed distinct CA IF structures at the PM/in the cytoplasm (ix.-xiv.) and the NE (ix., xi., xii.) without specific colocalization of the SiR channel. **c.** No CA-positive structures in the non-infected control. Unspecific binding of the CA antibody (CA IF) showed irregular staining of the PM.

2.3.13 Live-cell imaging of directly labeled HIV-1 CA

Apart from quantifying intensities of CA SiR containing objects in fixed cells, directly labeled CA offered the possibility of live-cell imaging and tracking of subviral particles. In order to follow incoming particles and possibly measure CA intensities to conclude uncoating and CA loss during the post-entry phase of viral replication, a second marker is required. For this purpose virus particles were prepared by co-transfecting HEK293T cells with pVpr-IN.mScarlet, HIV-1* or HIV-1*CA14TAG and pNESPylRS-eRF1dn-tRNA in a 1:2.45:1 molar ratio. Infection of TZM-bl cells with a viral particles containing an IN.mScarlet label (HIV-1* IN.mScarlet) showed detection of particles in the cytoplasm (Figure 31a-c) and inside the nucleus with high mScarlet intensities (Figure 31c, white arrowheads). Double-labeled particles, containing directly labeled CA and IN.mScarlet (HIV-1*CA14^{SiR} IN.mScarlet), were exclusively detected in the cytosol of infected cells (Figure 31d) and close to NE (Figure 31e) with clear colocalization of CA SiR and IN.mScarlet (Figure 31f). Infection with HIV-1*CA14^{SiR} did not reveal any objects inside the nucleus. Relative infectivity of IN-only- and double-labeled particles was determined via titration on TZM-bl cells (Figure 31g). In line with our previous results, the incorporation of IN.mScarlet leads to a decrease in infectivity of ~ 50 % compared to HIV-1*^{136} (Figure 31h, black-boxed bar). In this experiment, HIV-1*CA14^{SiR} particles displayed an infectivity of roughly 30 % compared to HIV-1* (Figure 31h, red-boxed bar). Combining both markers (HIV-1*CA14^{SiR} IN.mScarlet) reduced infectivity to only ~ 6 % compared to unlabeled HIV-1* particles (Figure 31h, yellow, red-boxed bar).

To expand investigation of viral kinetics and the fate of capsids with a direct CA label inside the nucleus, the IN.mScarlet labeling was replaced by an IN.eGFP tag. Incorporation of eGFP labeled IN into WT HIV-1 was shown to apparently not alter infection rates tremendously compared to unlabeled WT HIV-1 particles[62].

In preliminary experiments, infection of TZM-bl cells with HIV-1*CA^{SiR} IN.eGFP and fixation at 18 h p.i. revealed several particles inside the nucleus (Figure 32a, white arrowheads) with co-localizing CA SiR and IN.eGFP signals (Figure 32a i., ii.). Due to the triple transfection required to add a labeled IN, the probability of producing double-positive particles was lower to previously used particles containing only the CA SiR label. Particle preparations revealed ~ 75 % labeling efficiency, which means 75 % of CA IF positive particles incorporated both tags displaying positive signals for CA SiR and IN.eGFP (Figure 32b). HIV-1*CA^{SiR} IN.eGFP particles were used for infection and live-cell imaging of TZM-bl cells (Figure 32c). The recording was started at 6 h p.i. and images were taken every 2 min. IN.eGFP and CA SiR (Figure 32d) intensities were bright enough for imaging with low laser power

over an extended period (120 min). Manual tracking of two detected particles in the cytoplasm showed persisting colocalization of CA SiR and IN.eGFP signals during cellular trafficking (Figure 32e,f). Of note, this experimental setup did not include a staining to identify post-fusion events. Thus, it is not possible to distinguish endosomal virions from post-fusion cores.

The combination of directly labeled CA, the possibility to quantify CA content, and live tracking of incoming particles provide a versatile tool to shed more light on capsid uncoating in the nucleus and CA interactions. However, further characterization of double-labeled particles is needed in order to use this virus variant to investigate post-entry processes in live. Optimized transfection ratios could possibly lead to higher proportion of particles positive for CA SiR and IN.eGFP. Using this virus in future imaging would allow tracking of CA SiR intensity measurements in living cells and follow CA interactions and trafficking routes inside the nucleus even in higher resolved STED imaging.

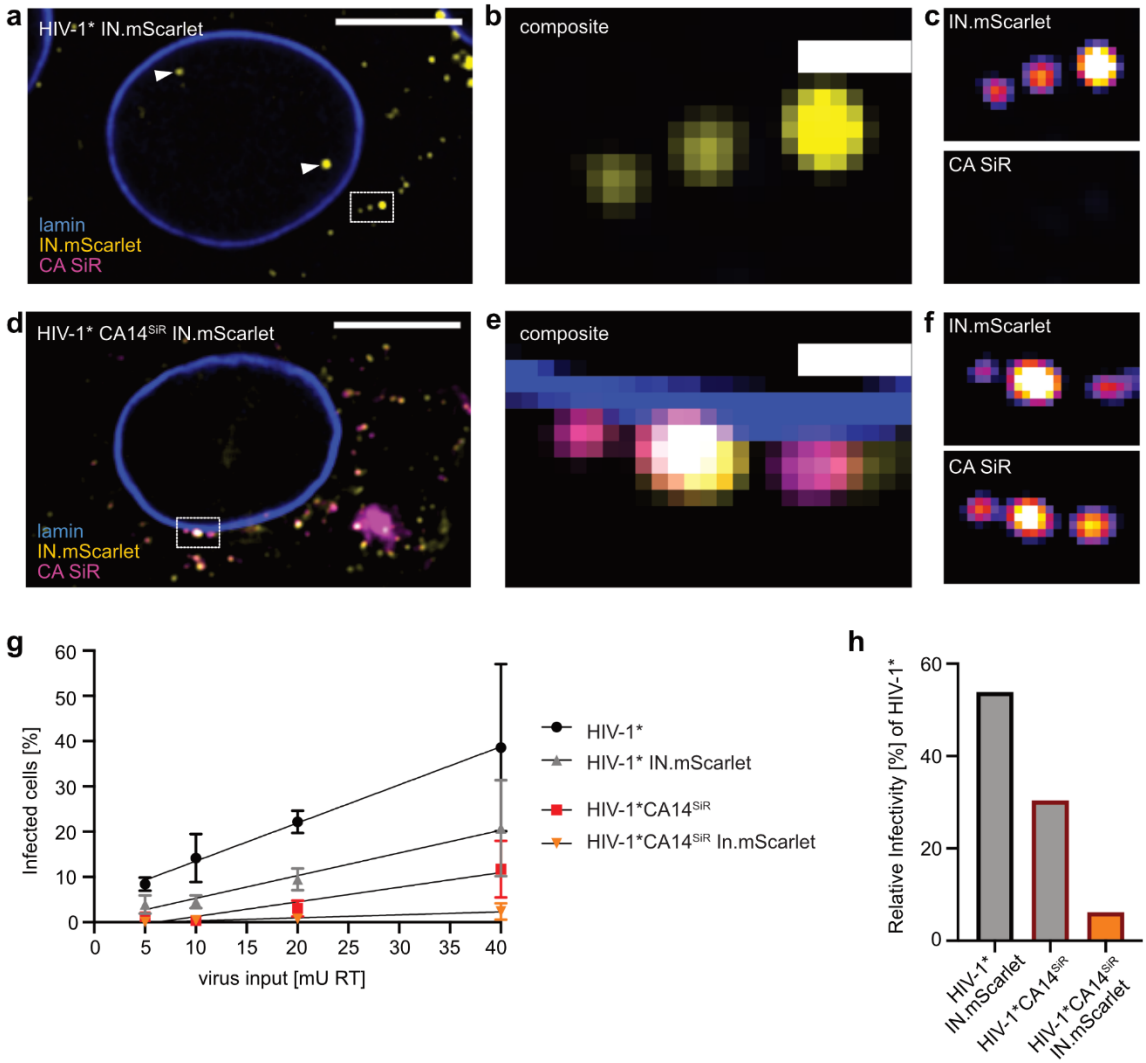


Figure 31. Additional IN.mScarlet labeling of CA click-labeled particles resulted in low infectivity. a-f. TZM-bl cells, infected with 200 mU RT/well HIV-1*IN.mScarlet or HIV-1*CA14^{SiR} IN.mScarlet, were fixed, permeabilized, stained by IF using an antibody raised against lamin A/C, and imaged by SDCM. Scale bars: 10 µm (cells) and 1 µm (enlargements). **a.** Cells infected with HIV-1*IN.mScarlet showed cytoplasmic (boxed region) and nuclear (white arrowheads) IN.mScarlet foci (yellow). **b,c.** Enlargement of the boxed region indicated in (a) showing separated IN.mScarlet and CA SiR channels. **d**. Cells infected with HIV-1*CA14^{SiR} IN.mScarlet showed cytoplasmic/NE associated (boxed region) IN.mScarlet foci (yellow). **e,f.** Enlargement of the boxed region indicated in (d) showing separated IN.mScarlet and CA SiR channels. **g,h.** Infectivity of HIV-1*CA14^{SiR} IN.mScarlet. **g**. HIV-1*(black), HIV-1*IN.mScarlet (grey), HIV-1*CA14^{SiR} (red) and HIV-1*CA14^{SiR} IN.mScarlet (yellow) particles were titrated on TZM-bl cells. At 6 h p.i., fresh medium supplemented with 50 µM T-20 was added to prevent second-round infection. At 48 h p.i., cells were fixed, permeabilized, stained with Hoechst and against HIV-1 MA by IF, and imaged by SDCM. The proportion of infected cells was quantified in Fiji based on the MA IF signal. Graph shows mean values and standard deviation of five FOVs per condition (n > 200 cells counted per condition). Lines indicate linear regression based on the mean to calculate relative infectivity. **h.** Relative infectivity of particles titrated in (g). Values obtained for HIV-1*IN.mScarlet (53.89 %), HIV-1*CA14^{SiR} (30.39 %) and HIV-1*CA14^{SiR} IN.mScarlet (6.16 %) were normalized to the infectivity of HIV-1*.

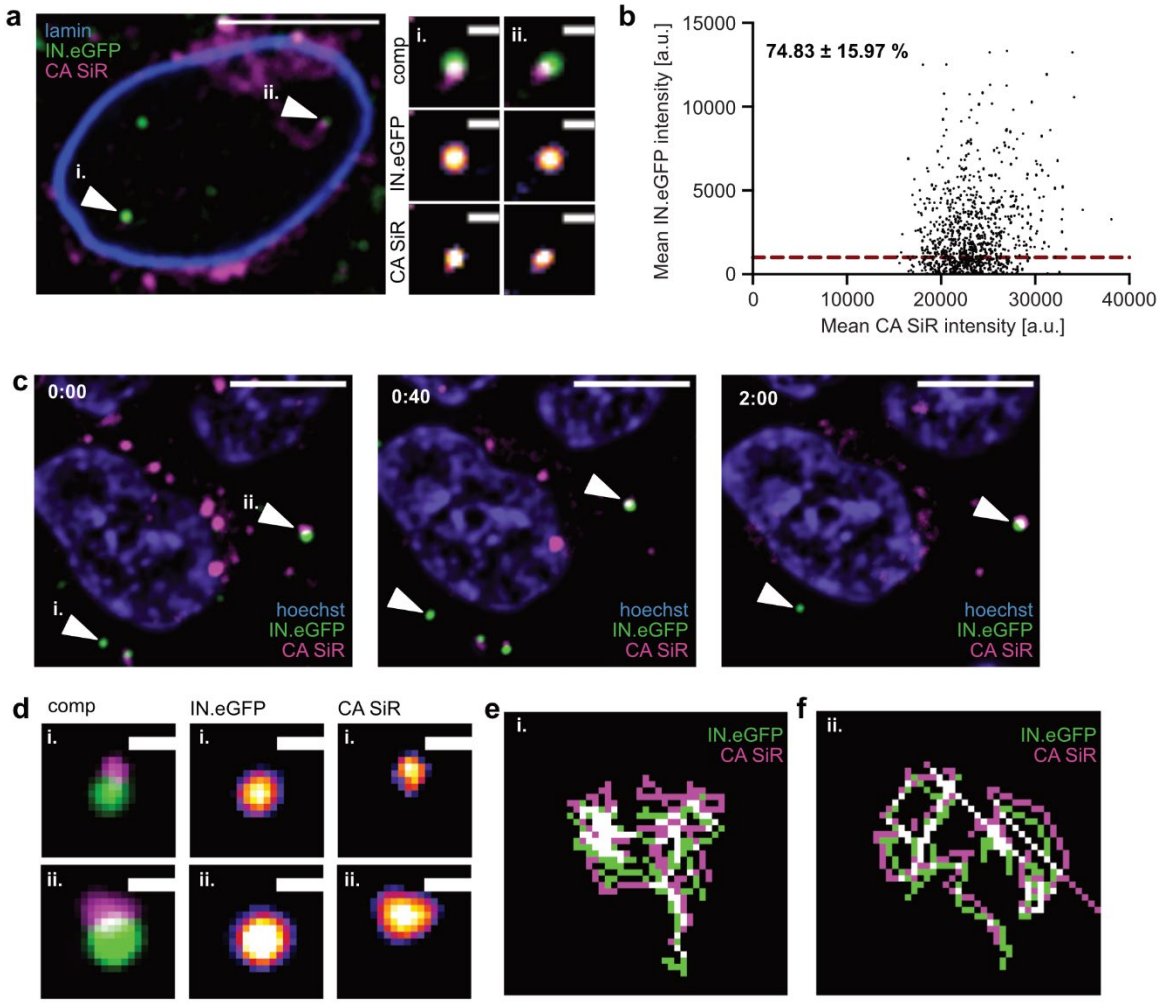


Figure 32. Directly labeled CA can be followed in live-cell microscopy. a,b. TZM-bl cells, infected with 200 mU RT/well HIV-1*CA14^{SiR} IN.eGFP, were fixed at 18 h p.i., permeabilized, and immunostained for lamin A/C. **a.** Single z-slice of a representative cell with nuclear CA SiR/IN.eGFP signals (i., ii.). Two enlarged single z-slices of nuclear particles indicated by the white arrowheads are shown next to the cell image. Mean filter and background subtraction was applied for clarity. Scale bars: 10 µm (cell) and 1 µm (enlargements). **b.** Labeling

efficiency of HIV-1*CA14^{SiR} IN.eGFP. Particles were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes, fixed and imaged by SDCM. CA SiR positive structures were detected in Icy. Mean IN.eGFP intensities were plotted against mean CA SiR intensities (n = 1,204). The threshold for IN.eGFP was set to t = 100 by visual inspection indicated by the red line. All particles above the set threshold were considered as double-labeled (IN.eGFP/CA SiR). **c-f.** Live-cell imaging of HIV-1*CA14^{SiR} IN.eGFP in TZM-bl cells. Imaging was started at 6 h p.i. recording one frame every 2 min. Mean filter and background subtraction was applied for clarity**. c.** Maximum intensity projection of exemplary frames at indicated timepoints. White arrowheads indicate detected IN.eGFP/CA SiR double-labeled particles. Scale bars: 10 µm. **d.** Enlarged single z-slices of two double-labeled particles indicated in (c) (i., ii.). Scale bars: 1 µm. **e,f.** Traces of double-labeled particles (i., ii.) manually tracked (IN.eGFP, CA SiR) in Fiji of all frames (120 min). White lines indicate the overlapping regions of both signals.

3 Discussion

In this thesis, a direct HIV-1 CA labeling approach was developed and validated, with the aim to investigate capsid uncoating in different cell types. The system was applied to determine the amount of nuclear CA of subviral complexes entering the nucleus through intact NPCs. This represents the first application of GCE together with click-chemistry to label the CA of an enveloped virus, leading to mature and infectious particles. Moreover, it is the first successful strategy for direct HIV-1 CA labeling independent of complementation with unlabeled WT virus. After extensive screening to identify a sequence position targeted for ncAA incorporation combined with subsequent copper-free click-labeling and establishing a working system, a virus variant was validated to follow highly labeled particles convincingly during early post-entry infection. This thesis contains quantitative image intensity analysis of directly labeled CA showing large amounts of nuclear CA corresponding to the fullerene core in HeLa-derived and primary activated CD4$^+$ T cells. CLEM and ET analysis provide the first direct evidence that previously observed capsid-like structures in the nucleus correspond to a detectable CA-specific signal. Additional STED imaging suggests nuclear clustering of individually trafficked particles in the cytosol, indicating common routes for nuclear entry. This method also allows the combination with another fluorescent IN staining facilitating live-cell imaging and tracking of CA signals. The new versatile tool, applicable for various proteins of different viruses, allows minimally invasive labeling and advanced super-resolution imaging.

3.1 The site targeted by GCE must be carefully considered

Point mutations of HIV-1 CA were found earlier to be poorly tolerated, shown by scanning mutagenesis[139,140]. Due to the highly ordered capsid structure, fusion of protein tags did previously

not result in functional particles suitable to investigate capsid disassembly[145,146]. The fragility of the CA also limits suitable positions for the incorporation of ncAAs and required extensive screening of several different positions within the sequence. Important intra-and inter-hexameric CA-CA interactions[304], which must be sustained during particle maturation and assembly, the capsid with its distinct structural features, such as the CypA binding loop (*cf*. 1.5; Figure 5), serves as a primary interaction partner for several host factors. Residues P85, V86, H87, A88, and P93 are known to contribute to productive CypA binding[305] together with the main determinants G89 and P90[169]. Although it was suggested that a CypA-independent mechanism exists to infect a cell[306], mutations within this region were connected to decreased infectivity and diminished CypA interaction[139], which was not the case for L83 and I91 tested for CypA incorporation in this thesis (*cf*. 2.2.3; Figure 14). A dramatic drop in infectivity was observed for amino acid substitutions at positions A92 and G94[307] or V83, H87 and R100[139]. The loop mutations V86, H87, A88, and A92 altered the CypA binding capability and, in some cases, infectivity[308,309]. Mutations of H87, A88, or I91 were more recently reported to increase the overall capsid stability[310].

In order to identify a residue suitable for codon exchanges, several different aspects need to be considered. Possible reasons for excluding certain residues could be site-specific low amber suppression efficiency, reduced particle release or assembly defects, disturbed proteolytic maturation and lost infectivity due to altered capsid stability or interaction with cellular factors.

Given this long list of possible reasons why distinct residues are not suitable for the incorporation of a ncAA *via* amber suppression, it is not surprising that also in our hands the majority of the 18 tested CAxx^{CpK} residues resulted in non-infectious virus particles after incorporating CpK (*cf.* 2.2.2; Figure 13d). Almost half of the residues tested in this study were part of the flexible CypA loop region, which is known to be crucial for the virus infection interacting with host proteins. The selection of these residues could have led to a possible disruption of CA-host factor-interactions explaining the severe drop in infectivity for residues adjacent to the loop region.

However, to a certain extent, the candidate I91 in the loop region was found to tolerate GCE modifications incorporating CpK and even additional attachment of Tet-SiR. In line with results of this thesis, previous studies reported that I91 mutations had no notable effect on CypA binding[305] accompanied by mild effects on single-cycle infectivity[139]. A possible explanation for viable virus production upon I91 substitutions was given by Gamble *et al*., showing that the I91 side chain had no intermolecular contact in CypA binding[169], leading to the assumption that amino acid changes at this position could be accepted.

Positions selected within the CA helix 6 were considered as promising candidates to incorporate ncAAs since mutations of G116 and L119 were reported to allow regular capsid assembly and achieve an even increased infectivity[140]. The incorporation of the ncAA, with a two-to-three times larger molecular mass (CpK: 256 Da) compared to natural amino acids, might cause sterical hindrances leading to non-infectious particles. The non-viable variants T110 and I124, tested in this thesis, most likely lost their infectivity due to disturbed Gag processing or assembly defects[140,311], which inhibit intact particle formation.

Although we learned from several studies that every structural component of the CA has its distinct function and CA-CA interactions are crucial for immature and mature particle formation, it was possible to identify an amino acid substitution, which allowed site-specific click-labeling of the CA. Residue A14 of CA is directly located after the ß-hairpin (1-13) structure. Existing structural information reported initially that the ß-hairpin itself was not making essential contacts in the assembled capsid. However, the deletion of 16 amino acids of the ß -hairpin (Δ1-16) impaired the maturation process[312] and the single mutation of I15 within the helix 1 was shown to be crucial for capsid assembly[313].

Nevertheless, we still suggested that the capsid might tolerate amino acid changes at residue A14 and the addition of the SiR dye. This assumption was supported by the fact that amino acid substitutions of A14 with polar serine or threonine were reported to only mildly affect single-round infectivity[139]. It was shown in this thesis that the exchange of the relatively small alanine at position 14 with the larger CpK enabled the production of morphologically mature and infectious particles. GCE-modified particles retained their infectivity even after Tet-SiR (641.8 Da) was coupled to CA monomers (*cf*. 2.3.3; Figure 19d,e).

Although production of infectious CA click-labeled particles was possible with the A14 mutant, particle yield was reduced to WT particles due to inefficient amber suppression, which constitutes the main limiting factor for high yields of modified proteins. Thus, several attempts were reported aiming to increase the overall efficiency. It is known that increasing the tRNA and aaRS expression level by e.g., promoter optimization or introduction of multiple tRNA and aaRS copies into the expression plasmid could be beneficial[314–317]. The same is true for improving the tRNA aminoacylation or its delivery to ribosomes by modifying elongation factors[317,318] and increasing the readthrough via release factor knockouts[223,225,226,319]. In this study, I have employed the dominant negative form of the release factor eRF1, a tRNA synthetase containing a nuclear export signal and

multiple tRNA cassettes on the expression plasmid to optimize modified virus production (Appendix Figure 35).

The sequence context is another factor, which needs to be considered to increase amber suppression efficiency since it strongly depends on nucleotides adjacent to the amber codon[320,321]. Possible reasons for this observation could be the interaction of the release factor[322] or the amber suppressor tRNA[323] with distinct nucleotides downstream the amber codon. Besides, chemical properties of neighboring amino acids, such as charge or polarity as well as secondary structures, can affect tRNA binding and codon suppression[324–326]. At the outset of this work, prediction of suppression efficiency in the eukaryotic context based on the sequence was not possible. The very recently developed and described proteomics-based online tool iPASS (identification of permissive amber sites for suppression) offers the possibility to identify amber suppression sites within a protein sequence to incorporate ncAAs efficiently[220]. The program predicts promising residues based on the sequence context of the input ORF by using a linear regression model. The tools' output is the so-called iPASS score ranging from -1.61 to 3.01 in the context of the HIV-1 CA sequence (Table 1). An iPASS score of ≥ 1 indicates that the relative ncAA incorporation efficiency is above average. Another unique advantage of this tool is the consideration of sequence context optimization resulting in higher iPASS scores. The program proposes silent mutations of neighboring codons flanking the amber stop codon, which generate a favorable sequence context. A predicted high ncAA incorporation efficiency (≥ 1) was stated for HIV-1*CAxx^{TAG} variants W80, L83, V86, H87, T110, G116 and I 124. Extremely low iPASS scores were measured for A88, I91, M96, H120. Residue A14 reached a score of 0.72. In an optimized sequence context, suggested by iPASS, the A14 score could be increased to 1.95. Compared to other CA residues targeted for amber stop codon insertion, iPASS predicated a moderate amber suppression efficiency for residue A14. Having regard to the high infectivity of HIV-1*CA14$^{CpK/SiR}$ particles, A14 could have also been included into a smaller subset of candidate residues selected based on the generated iPASS result. Future additional improvements of particle yield could be achieved by optimizing the amber codon-flanking sequence at this position from *catcagTAGatatca* to *catcaaTAGatctcg* as suggested by iPASS in order to increase ncAA incorporation.

Table 1 iPass score for CA variants predicting ncAA incorporation efficiency

residue	iPASS score [WT]	iPASS score [optimized]
A14	0.72	1.95
I15	0.85	2.53
W80	1.60	2.11
L83	1.75	2.33
H84	0.91	2.60
V86	1.87	2.30
H87	1.64	2.75
A88	-0.23	1.11
I91	0.08	1.73
A92	0.68	2.41
M96	-0.53	1.51
E98	0.64	1.79
T110	1.48	2.04
L111	0.37	0.71
G116	1.91	1.91
T119	0.33	0.62
H120	0.03	0.46
I124	1.04	2.60

3.2 Nuclear detection of directly labeled CA does not depend on sample handling

Imaging of HIV-1 infected cells revealed directly labeled CA SiR signals colocalizing with CA IF inside the nucleus of different cell types. The quantitation of CA SiR signal intensities revealed an average intensity corresponding to an almost intact capsid core (*cf.* 2.3.10; Figure 26c and 2.3.11; Figure 27e). Together with CLEM data (*cf.* 2.3.11; Figure 28), these results strongly argued for the presence of largely intact capsids inside the nucleus of HeLa-derived and primary T cells. The unique possibility to connect fluorescence intensities to a relative CA amount was offered by establishing the first direct CA labeling without WT complementation. It was shown that almost every detected particle contains clicked residues resulting in a high labeling efficiency (*cf.* 2.3.3; Figure 19b,c) making these particles suitable to follow incoming viral structures by confocal imaging. Assuming a labeling of CA molecules with equal probability, it can be suggested that a major CA loss (i.e., uncoating) would have come along with a significant decrease of CA SiR intensities. Thus, the CA SiR

intensity drop of only ~ 50 % inside the nucleus would be consistent with a full capsid lattice complement inside the nucleus, which may then need to uncoat (*cf.* 2.3.10; Figure 26).

Most studies employing CA IF could not detect strong nuclear CA signals in other cell types than macrophages[58,62,117,118,130,131]. This thesis and another publication from our group[64] support the hypothesis that the apparent cell-type-specific differences observed may have resulted from inefficient antibody binding due to the dense nuclear environment or clustering of CA-binding nuclear host proteins. Studies that did report detection of strong nuclear CA IF signal had employed particular experimental conditions, namely partial protease digestion of the samples[150] or harsh conditions required for copper catalyzed click chemistry[63,85,89,129]. The specific labeling conditions in these experiments lead to cell extraction and allow nuclear CA detection by IF. Previous work from our lab revealed that extraction with, e.g., methanol or dispersion of CPSF6 clusters at CA-containing subviral structures exposes the CA epitopes at nuclear complexes[64]. In T cells, on the contrary, a reliable CA IF detection was only possible after CPSF6 displacement and additional methanol extraction. The reason for this difference is currently unclear, but it suggests differences in the nuclear architecture and CA accessibility. The here presented direct CA labeling approach confirmed our results from this previous study by showing distinct nuclear CA-SiR foci independent of sample handling or infected cell type (*cf.* 2.3.11; Figure 27). It was also validated that the labeled CA variant was able to interact with CPSF6 (*cf.* 2.3.5; Figure 21g,h) and that CA-SiR foci became positive for IF signals after removing CPSF6 clusters upon extraction and PF74 treatment (*cf.* 2.3.11; Figure 27a-c).

Surprisingly, treatment of cells by either methanol or PF74 was not enough to detect the CA-labeled HIV-1*CA14^{SiR} variant inside the nucleus; only the combination of both yielded strong CA-IF punctae (*cf.* 2.3.5; Figure 21). This observation might suggest that the detection of nuclear click-labeled particles requires more efficient cell extraction. A possible reason could be that the polyclonal antibody displays reduced binding to modified A14 capsid and therefore additional exposure is needed. However, unidentified nuclear proteins apart from CPSF6 might also influence the nuclear environment and thereby CA accessibility. CPSF6 was reported as part of the cleavage factor I (CFI_m) building a complex with the subunits CPSF5 and CPSF7[327,328]. It is unclear whether the whole CFI_m or CPSF6 alone is needed to mediate HIV-1 infection. However, some references suggest that CPSF5 and CPSF7 were not required for active chromatin targeting and that they have only a minor role in facilitating CPSF6-capsid binding[329]. Investigation of possible CA interactions with CPSF5/7 or so far unidentified nuclear proteins could be part of further research.

3.3 Capsid mediates nuclear entry

The detection of nuclear capsids is in line with recent reports also indicating nearly complete capsids inside the nucleus and nuclear uncoating[59–61,64,135]. The detection of conical capsids in the nucleus of T cells arrested by APC treatment demonstrated that capsids most likely entered through intact NPCs (*cf.* 2.3.11; Figure 28). Together with the finding that the NPC diameter is large enough for the passage of the HIV-1 capsid core[136], the detection of nuclear CA signals strongly supports the model of nuclear uncoating. The observation that click-labeled particles displayed delayed nuclear entry kinetics (*cf.* 2.3.6; Figure 22b-d) is consistent with the notion that the viral capsid influences the mechanism of nuclear translocation. Time-resolved analysis revealed that HIV-1*CA^{SiR} particles were not impaired in cytosolic trafficking but reached WT comparable amounts of nuclear subviral structures only at later time points.

Since nuclear entry is the main bottleneck of the viral infection[330], viruses adopted several strategies to overcome this barrier, when they need to replicate inside the nucleus of the host cell[331]. Some viruses have been reported to employ importins *via* binding of nuclear localization signals incorporated into, e.g., the Hepatitis B capsid core (~ 32-36 nm diameter) or the Influenza A ribonucleoprotein complex to access the nucleus[331,332]. Viruses with a larger genome encapsulating capsid – others than HIV-1 – were reported to also interact with Nups of the NPC. For example, the adenovirus with a 90-100 nm-sized capsid interacts with Nup214 before it disassembles and releases the genome into the nucleus[333]. A similar strategy is followed by Herpes simplex virus 1 capsid (~ 120 nm diameter), which docks at the outer NPC via Nup214-Nup358 interactions[334,335] and injects the pressurized genome into the nucleus.

The HIV-1 capsid constitutes the largest cargo transported through intact NPCs known so far[136]. If the ~ 60 nm of the HIV-1 capsid constitutes the upper limit of NPC transport, this may explain the observed delay of SiR-labeled particles at the NPC. Incorporating the larger ncAA CpK could alter the chemical or structural properties of the CA but does not necessarily affect the diameter of the assembled core. However, additionally ligating the dye to a high number of individual CA monomers of the core could enlarge the capsid size in the range of ~ 1-3 nm (distance measurement of Tet-SiR performed after energy minimization in Chem3D, PerkinElmer resulted in a calculated diameter of ~ 1.5 nm), which conceivably might hamper NPC passaging resulting in slower nuclear import kinetics.

Capsid is the main factor mediating infection of non-dividing cells[112] and therefore is likely involved in NPC passage, however, the exact mechanism of HIV-1 translocation is still unclear. Several studies showed that the CA lattice binds different Nups[52,125,126,336], and CA mutations can alter these interactions[122]. One-third of the Nup proteins contain FG motifs (Nup358, 214, 153, 98, 62, 58, 54, 50 45, POM121, NLP1)[337] leading to directed movement of viral complexes into the nucleus[338]. The mature HIV-1 capsid lattice compromises particular sites to interact with certain FG-repeats[123,125,339] (*cf*. 1.3.1). Thus, a complex capsid-Nup interaction cascade might be speculated to mechanistically mediate NPC translocation involving the two major interfaces of mature hexameric lattice, which serves as an interaction platform for host proteins (*cf*. 1.5 and 1.6). Consecutive binding of host proteins could interact already inside the pore applying forces, which direct or pull the capsid through and out of the NPC channel into the nucleus. Speaking in favor of such a transfer process could be evidence from Bejarano *et al*. suggesting CPSF6 competing Nup153 from the common interface on the viral capsid to facilitate entry[63].

Although residue A14 is not located within the FG-binding cleft, it may interfere with interactions of any of the involved components of the NPC, which may lead to the observed delay in kinetics. This might be direct obstruction or abrogation of interactions but also indirect interference by changing the general chemical properties of the capsid structure.

It is currently not understood if the capsid undergoes remodeling after docking to allow the passage through the NPCs central pore, which is one possible model for nuclear entry. The capsid owes its conical shape to the distinct distribution of the 12 pentamers; five at the narrow and seven at the wide end[340]. The most straightforward explanation for the capsid passing through the NPC could be a substantial deformation with a longitudinal shaped capsid as an intermediate state. This process might require a temporarily uniform distribution of the 12 pentamers along the capsid during nuclear entry[341]. However, it is somewhat hard to believe that the capsid can alter its architecture to this degree and switch back to a cone shape as soon as the NPC is passed. This becomes exceedingly doubtful if we consider that the CTD dimers must bridge different distances between hexamers and pentamers, which would require an even more sophisticated system of the virus to fill the gaps between individual subunits after exchanging hexamers and pentamers during remodeling[144]. Still, the hypothesis of capsid remodeling is part of the current debate about nuclear entry of the virus[46]. The aspect of remodeling came up in recent publications using, e.g., the combination of CA immunogold labeling and CLEM, which observed a changed staining pattern during nuclear entry[67]. But the observed “pearl necklace-like” staining pattern, which was suggested by the authors to indicate a changed capsid shape, does not allow conclusions about the structure of the underlying

capsid, and thus still leaves room for speculation without direct evidence of a remodeled capsid. A clearer illustration of capsid shapes during nuclear entry offered the work from Zila *et al*. capturing intact cone-shaped lattices inside the pore by CLEM and cryo-ET[136]. Since they showed many tubular or broken remnants in the nucleoplasm, the obvious question of whether the capsid undergoes remodeling during or after nuclear entry was also raised. Of note, these data were collected using a CPSF6 binding deficient virus variant. This thesis and another study from our lab to which I contributed using WT virus[64], however, showed largely intact nuclear cone-shaped structures, apparently broken remnants (often at the narrow end) or tubular structures suggesting entry of the complete core, which breaks open in the nucleus (*cf.* 2.3.11; Figure 28). Such a mechanism would not argue for major changes of the capsid during entry. Rather, the very recently discovered flexibility of the NPC[137,342] might allow passage of the capsid structure.

Yet, our findings do not exclude a nuclear entry process requiring capsid flexibility or remodeling at the NPC. However, the favored entry model of intact capsids and nuclear breakage without remodeling or larger disassembly before entry was supported by recent findings of another groups[60,135]. These authors monitored capsid disassembly and integrity simultaneously, and concluded that the core disassembles rapidly in the nucleus after initial integrity loss[135]. This model is consistent with the observation that large amounts of assembled CA were retained at subviral complexes until shortly before integration[60]. The CA click-labeling approach, used in this thesis also detected largely assembled capsids even at late time points post infection (*cf.* 2.3.10, Figure 26; 2.3.11, Figure 27e and Figure 28).

Nonetheless, our current knowledge about HIV-1 nuclear entry through the NPC is insufficient to differentiate between capsid remodeling at the NPC, partial disassembly, and local capsid rupture during passaging. However, it can be stated that at least a part of the capsids that translocated through the NPC appear largely morphologically intact. CLEM data obtained in this thesis (*cf*. 2.3.11; Figure 28) and other studies from our lab[64,136] indicate that a major breakage or disassembly does not occur at the pore. Whether cores that show altered structures inside the nucleus were broken during or after translocation is not known. Thus, further investigation of the entry process including dynamic properties of NPC components is required for a mechanistic understanding.

3.4 An improved CA labeling strategy leads to better understanding of nuclear uncoating

As discussed in one of the previous sections, the debate whether the capsid uncoats before or after nuclear entry and to which extent single monomers dissociate from the CA lattice is still ongoing (*cf.* 1.3.4). This thesis presents striking evidence for translocation of at least almost complete capsids into the nucleus in cell lines and primary cells (*cf.* 2.3). The approach using a directly labeled CA might show a holistic picture of uncoating dynamics since it suffers less from technical drawbacks than IF staining or indirect labeling systems. CLEM data from our lab showing complete capsids passaging the NPC[136] and several closely adjacent nuclear capsid-like objects[64] support the presented model of complete capsids transferred into the nucleus.

However, some of the results presented here are not consistent with conclusions from recent imaging studies. In the following, discrepancies to an indirect labeling method[58,85,118,119] and the direct approach with a GFP fused to CA[41,60,121] will be discussed.

The indirect CA labeling technique, already mentioned in chapter 1.4, is based on CypA-DsRed packaged into virions *via* CA binding. According to the authors, following the DsRed signal allows conclusions about the CA amount associated with RTC/PICs. In contrast to this thesis, none of the publications using this labeling technique could detect nuclear capsids. First, they suggested cytoplasmic uncoating due to a rapid DsRed loss after fusion[118]. In close succession, the same group argued for the presence of a high CA content near the nuclear membrane and uncoating at the NPC before nuclear translocation in HeLa-derived cells[119] and primary macrophages[58,85], whereas this thesis showed strong nuclear CA signals at individual subviral particles in a HeLa cell line, primary $CD4^+$ T cells, and preliminary in MDMs. Since the CypA binding loop of the CA serves as the central interaction platform for CypA, it seems likely that non-covalent CypA attachment could be easily replaced by other molecules binding to an overlapping site. The interaction of TRN1[54] and Nup358[52] during docking and nuclear entry might remove CypA via competitive binding, which would result in DsRed signal loss but not reflect disassembly of the capsid. Consistent with my assumption, a recent study showed nuclear CA signals by IF in MDMs that lacked CypA-DsRed signals[85]. The authors argued that nuclear CA IF signals in macrophages were only detectable because small CA amounts remained associated with viral complexes and cell-type-specific clustering of subviral complexes led to positive CA IF signals. Alternatively, failure to detect nuclear CA by IF in certain cell types could be explained by staining conditions, as outlined above. Apart from that, lacking the nuclear CypA-DsRed signal in contrast to IF could be another indication that the signal was ripped off from the

capsid at the NPC. I conclude that indirect CypA-DsRed labeling of capsids is not suitable to study nuclear translocation of capsids.

Burdick *et al*. incorporated GFP-tagged CA subunits into infectious virions[60]. The authors showed strong nuclear CA signals in infected cells, in line with results presented here. However, shortly before integration the GFP-CA signal was lost, which was interpreted by the authors as uncoating by disassembly of the capsid lattice into the individual CA subunits.

To retain infectivity, the authors were forced to use complementation with WT CA and incorporated the labeled variant at substoichiometric concentrations (1:15). The crucial aspect of using a CA-labeled virus to investigate uncoating and reliably measure CA signals is to ensure that tagged proteins are incorporated into the mature CA lattice rather than only as free GFP-CA encapsulated within the capsid or loosely associated. However, it is not conceivable how CA molecules containing a ~ 3 nm sized GFP[343] could fit into the highly structured capsid lattice without disturbing the complex and its maturation. Additionally, if a large proportion of GFP-CA is incorporated into the lattice the substantially larger diameter would likely be detrimental to the nuclear import step. Thus, as recently mentioned in a review article[344], it might be less likely that most of the CA monomers forming the lattice were labeled.

Furthermore, the addition of PF74 induced rapid loss of GFP-CA inside the nucleus within minutes[60]. In contrast, PF74 addition to the CA14 SiR-labeled nuclear particles did not lead to loss of the strong CA SiR signals in the nucleus. *In vitro*, PF74 leads to an initial permeabilization of the capsid, allowing internal fluid-phase marker GFP to diffuse out[111,135]. Following this event, the remaining lattice was reported to be highly stabilized[111] consistent with high CA SiR intensities observed inside the nucleus of cells in this thesis (*cf.* 2.3.7; Figure 23), and high CA IF signals[64].

This suggests that GFP-CA may not be an integral part of the hexameric capsid lattice, but rather loosely associated or influencing capsid stability and thus not faithfully reporting capsid disassembly.

A similar strategy of GFP fused to CA was also employed by another group[121]. These authors reported, in contrast to this thesis and the work from Burdick *et al*., that capsid uncoating occurs before nuclear translocation[121]. However, since the authors used WT complementation in an extreme excess, reporting only ~ 2 GFP-CA molecules per virion, it seems unlikely that this study quantifies CA content reliably.

Thus, this thesis presents a direct CA labeling strategy offering clear advantages over previously described methods. Without WT complementation and the modification of changing only one amino acid, intact capsid cores containing exclusively ncAA modified CA molecules could be assembled. Since the chemical click reaction by addition of Tet-SiR is highly efficient and happens after complete particle maturation, we were able to show that a high number of CA monomers building the CA lattice was labeled detecting high intensities of post-fusion cores (*cf*. 2.3.10; Figure 26). The flexibility to choose different bright and photostable organic dyes, allows super-resolution STED imaging (*cf.* 2.3.8; Figure 24 and 2.3.9; Figure 25) and CLEM analysis (*cf.* 2.3.11; Figure 28). Preliminary results also indicate that CA click-labeled particles can be used for advanced Minflux[345] imaging to achieve low nanometer resolution.

3.5 Nuclear capsid clustering might be important for viral infection

Incoming HIV-1 particles use dynamic microtubules and host proteins such as FEZ1 or BICD2, in order to deliver their viral genome towards the nucleus efficiently[28,36]. These polarized filaments support an active and directed cargo transport towards their minus-ends anchored at perinuclear sites called microtubule-organizing center (MTOC). The viral CA plays a central role in interactions with the microtubular network[38]. In this thesis, imaging of incoming click-labeled HIV-1*CA^{SiR} particles by STED revealed individual cytoplasmic virions mainly trafficking within endosomes (*cf.* 2.3.9; Figure 25d-i). As this experimental setup was not able to identify a large proportion of post-fusion subviral complexes, I cannot comment on MT-mediated cytoplasmic transport. However, consistent with other studies detecting nuclear accumulation of viral complexes[64,85,89], the here presented STED data identified nuclear clustering of several capsids with clustering frequency depending on the virus amount (*cf.* 2.3.8; Figure 24). Low infection units per cell resulted in the detection of mostly individual structures, whereas infection with a higher viral dose showed clustering of two to four capsid for half of the nuclear events (*cf.* 2.3.8; Figure 24c).

Interestingly, this clustering phenotype was retained to at least some extent in this thesis after unsynchronized infection with even roughly ten-fold lower virus amounts (*cf*. 2.3.8; Figure 24b,c and 2.3.11; Figure 28) than used in other studies to detect nuclear clustering in HeLa-derived cells[64] or macrophages[89]. Due to the observation that nuclear capsids accumulate at the same location after reaching the NE individually (detected for low and high MOI) and broken capsid remnants releasing the viral genome[64], it could be suggested that this is the site of final uncoating. We have shown in this study the separation of viral cDNA from IN.FP and captured nuclear clusters with broken capsids

releasing an electron-dense filling by CLEM. This might show nuclear uncoating by capsid breakup spitting out the viral genome. Whether capsid clustering is needed for efficient infection, is not known. As this phenotype was also observed in T cells (*cf*. 2.3.11; Figures 26e and 27), it can only be hypothesized that this might reflect a preferred mechanism of the virus releasing its genome during or after clustering.

In line with such an assumption, would be the observation of clustering subviral complexes in nuclear speckles[85,89], which possibly integrate into speckle-associated domains containing actively transcribed genes[85,86]. Viral structures are directed by CPSF6 beyond the periphery deeper into the nucleus to reach these gene-rich regions to integrate supported by LEDGF/p75[76]. Accordingly, CA mutants deficient in CPSF6 binding were reported to be localized more at the nuclear periphery[119,150,193,346]. These findings and clustering even with low virus amounts detected with HIV-1*CA^{SiR} labeled particles could argue for intranuclear common routes of CPSF6 directed capsids. It was reported, that addition of PF74 induces relocation of IN.FP signals from nuclear speckles and additionally decreases IN.FP signals of individual subviral complexes[85]. This observation might support the hypothesis that incoming viruses are directed towards these speckles in a CPSF6-dependent manner, which is lost upon PF74 treatment. However, here it was shown, that the clustering phenotype is retained following PF74 addition (*cf*. 2.3.8; Figure 24). This may indicate that once capsids are clustered, the PF74-induced interference with the CPSF6-CA interaction leads to relocation of whole clusters to the space adjacent to nuclear speckles. In future, detection of CA SiR positive subviral complexes inside the nucleus could be combined with a 5-580CP-HOECHST chromatin staining in STED imaging pointing out the interchromatin channels used as common nuclear routes by incoming capsids to reach the gene-dense regions deep inside the nucleus.

This thesis identified nuclear largely intact capsids by EM clustering close to the NPC using click-labeled virus (*cf*. 2.3.11, Figure 28). So far, this observation was made only in one experiment but is not consistent with the assumption that clustering occurs at interior nuclear sites where gene-dense regions are located[85,89] and only CA mutants were reported to integrate on the outer nuclear periphery near the NE[60]. Thus, the observation of capsid clustering close to the NE in this thesis was either only one rare instance where nuclear speckles were just close to the NPC, or it could indicate that capsids accumulate rapidly after NPC passaging and cluster before trafficking towards the speckles rather than individual nuclear trafficking. Of note, we cannot exclude that HIV-1*CA14^{SiR} displays a CA mutant phenotype with altered integration sites. Very recently, Burdick *et al*.[346], however, showed that the location of CA-CPSF6 impaired mutants close to the NE was changed after cell division towards the interior of the nucleus making them comparable to WT CA and that

chromosomal sites of integration are dynamic. Furthermore, authors mentioned that T cells showed lower distances between active proviruses and the NE compared to HeLa cells. This might suggest differences in the positioning of active proviruses in different cell types and might explain the detection of a capsid cluster close to the NE in T cells. Thus, it is a question of future research to investigate if this result is reproducible, if such NPC-close clusters colocalize with CPSF6, and if the distance between the clusters and NPC is substantial. Imaging of CA14^{SiR} click-labeled particles would be also helpful in this context to show if capsids cluster before or after trafficking towards these speckles. As confocal imaging will not be sufficient to resolve capsid clusters, live imaging in STED mode of an integrase-deficient construct (NNHIV*CA14^{SiR}) could be applied to follow individual subviral complexes shortly before and after nuclear entry.

Furthermore, it is unknown if the cytosolic filaments lead to a limited number of NPCs and if this can cause capsid accumulation at a subset of NPC pores in front and therewith nuclear clustering. Live imaging of directly labeled CA14^{SiR} particles could visualize cytoplasmic trafficking routes towards NPCs also addressing the question of whether the virus exhibits nuclear entry through a smaller subset of preferred NPCs and if an NPC can be passed subsequently. Zila *et al*. observed that accumulation happened often at the cytosolic side of NPCs, but whether this reflects the directed cytosolic pathways used by viral structures was not addressed further[136]. In contrast to the currently by us used IN.FP system to detect or follow incoming viral subcomplexes in other studies[64,136], the HIV-1*CA^{SiR} particles achieve higher fluorescence intensities due to more dense labeling of several CA monomers. The IN.FP particles usually exhibit very low staining since only some labeled IN.FP are incorporated into virions, while complemented with unlabeled protein. Thus, click-labeled particles offer better imaging conditions to reliably detect a larger number of viral structures and constitute a reliable marker for the detection of viral capsids. Tracking of SiR intensities in live imaging could help to understand, e.g., if one NPC is passed subsequently by multiple particles, which then follow the same intranuclear pathway, indicating the importance of the NPC Nup-composition for capsid entry[347].

3.6 The spatio-temporal connection between uncoating and reverse transcription remains unclear

The process of HIV-1 reverse transcription is initiated in the cytoplasm after fusion and CA was found to be associated to reverse transcribing viral complexes in the cytoplasm[92,120] and close to the NE[119]. First attempts to connect uncoating and reverse transcription were based on mutational studies

using hypo- or hyperstable CA mutants, which impaired reverse transcription in both cases, suggesting that a balanced stability of the capsid core is essential to allow ongoing of the process[101]. Consistent with that, biochemical assays confirmed that destabilization of the CA lattice leads to reverse transcription defects presumably due to the loss of viral enzymes and the genome targeted by the host cell[348]. Inhibition of reverse transcription delayed uncoating based on indirect uncoating assays[103,349] and it was suggested, that reverse transcription of the viral genome might induce gradual CA loss during trafficking[103,104]. However, more recent studies challenged the connection between gradual CA loss and ongoing reverse transcription mostly happening in the cytoplasmic region[350,351]. In line with the detection of largely intact capsids transferred into the nucleus in this thesis and other recent studies, it was reported that reverse transcription may be finished only after nuclear translocation[58,60,64,86]. Although many reverse transcription products were detected already in the cytosol[59,119,151], *in vitro* endogenous reverse transcription assays detected the peak of late reverse transcription products at 8-10 h p.i.[352,353]. If the *in vitro* measured kinetics fit the time course of ongoing reverse transcription in infected cells, this would be a timepoint indicating a nuclear location of the subviral structures. More evidence for reverse transcription completed inside the nucleus of assembled capsids was provided by cell fractionation showing that the nuclear fraction was enriched in intermediate reverse transcription products[61] and detection of both, (+) and (-) DNA strand-specific probes only in the nucleus[59].

Nevertheless, the exact spatio-temporal link between uncoating and reverse transcription remains unclear[350]. A possible relation could be inferred based on theoretical calculations that reverse transcribed dsDNA requires more volume compared to ssRNA[354]. Together with atomic force microscopy measuring differences in the capsid stiffness during reverse transcription (followed by core rupture)[355,356], the data argue for capsid uncoating due to an increasing inner pressure. The viral NC protein was reported to be involved in reverse transcription[357] and to condense dsDNA[14,358], which could be a mechanism to fit the whole genome and overcome limitations of the capsid volume. If the capsid loses its integrity already before complete rupture shortly before integration, the decreased NC concentration might in turn cause genome decondensation and increased inner pressure. Possibly, the pressure is already high while reverse transcription is still processing, but cytoplasmic stabilizing factors such as CypA[178] or Sec24C[57] prevent core opening. Upon removal of these factors during NPC passaging, this capsid stabilizing effect might get lost and other factors with destabilizing effects (e.g., TRN1[54], TNPO3[359]) at the nuclear side could additionally accelerate core breakage.

The recently described GFP-CA labeling strategy reported rapid and almost complete GFP loss from nuclear complexes at 10.5 h p.i., which was interpreted as complete capsid disassembly shortly before integration[60]. If we assume that this labeling method is rather unlikely able to incorporate GFP-CA into the CA lattice (*cf*. 3.4), this signal loss might be explained by losing free GFP. In contrast, this thesis revealed largely assembled capsids even at late timepoints post-infection (18 - 24 h p.i.). Although the data argue for capsid opening but not complete disassembly, the detected CA SiR signals (*cf*. 2.3.10; Figure 26, 2.2.11; Figure 28) still do not tell whether these structures are the ones productively infect the cell. Whether productive particles must uncoat completely to be productive could be a question of future research. Live imaging will be necessary to better comment on the temporal sequence of the timepoint where reverse transcription is finished and uncoating before integration.

The here presented direct CA labeling with Tet-SiR might offer the possibility to develop a new advanced tool by combining it with our previously described ANCHOR labeling to label reverse transcribed viral cDNA[64] (*cf*. 1.3.2). A proviral plasmid containing an amber codon within the CA and the ANCH sequence within Env could be used to generate a pseudotyped virus, which can be detected in infected cells. If the combination of both systems would allow the production of intact and infectious particles, both fluorescent signals of directly labeled CA and cDNA could be detected by confocal (live) imaging in infected cells. The simultaneous staining of reverse transcription (cDNA) and uncoating (CA) markers could contribute to fill the gap in our understanding of the connection between both processes. One possible observation could be loss of CA intensity still colocalizing with OR3.GFP, which would argue for gradual capsid disassembly, while reverse transcription is still ongoing or already finished. If we would detect separation of CA and cDNA signals of individual subviral complexes, the results from Müller *et al*. would be reproduced and live imaging data of labeled CA and vDNA would be added. This observation would then also suggest the model described in this study suggesting physical disruption of the capsid lattice after reverse transcription is completed rather than progressive disassembly. In line with these observations made in our previous study, this thesis detected largely intact capsid cores at late timepoints post infection (*cf*. 2.3.10; Figure 26 and 2.3.11; Figure 27e and 28). The combination of both protein labels could clarify if detected capsids in this thesis are associated with viral DNA.

Another option to improve our understanding of uncoating dynamics before final integration would be to expand the here presented system allowing to identify productively infected cell by an additional marker such as a GFP-reporter virus containing GFP in the place of a gene region nonessential for the virus infection (nef, env)[60,118,119]. This tool would allow live imaging of

pseudotyped HIV-1*CA14^{SiR} particles, identification of the productively infected cells, and retrospective analysis of CA SiR intensities. This approach might answer if the largely intact capsids detected in this thesis represent viral structures that productively infected the cell.

Although we still lack more knowledge about the exact time course of CA loss connected to individual reverse transcription stages, recent studies offered a wide variety of different methods to follow labeled CA, cDNA and measure reverse transcription products. Combining the results from studies applying different experimental methods in the context of HIV research, such as the ANCHOR system[64], CA labeling methods from us and others[60,266], DNA hybridization probes[59], quantitative PCR together with RT inhibition[61] or CLEM[64,136], enabled us to change the current understanding of early HIV-1 replication dynamics.

3.7 Outlook

As discussed within this thesis, detailed knowledge about uncoating dynamics, NPC passaging or the connection between uncoating and reverse transcription is still lacking. The here presented minimally invasive CA labeling technique could be used as a versatile tool for future experiments together with insights from other recent studies to elucidate different viral replication steps in a time-resolved manner.

One central question that needs to be addressed in future studies is how subviral complexes enter the nucleus through intact NPCs. I already discussed briefly that it is not clear whether the capsid remodels its architecture during nuclear import or not (*cf*. 3.3). Another possibility explaining NPC passaging of the large HIV-1 cargo might be NPC flexibility. Dynamic changes of the NPC diameter during virus entry could have been not excluded so far. Thus, we decided not to focus only on the capsid properties and started to apply the CA click-labeling to Minflux measurements[345]. Minflux nanoscopy is a recently developed super-resolution imaging approach combining the strengths of STED[360,361] and PALM/STORM[362,363] imaging. By localizing individual switchable fluorophores with a donut-shaped excitation beam, this system achieves ~ 1-3 nm resolution in three dimensions. We optimized labeling conditions of HIV-1*CA14^{CpK} particles with a blinking AF647 dye suitable for Minflux to achieve highest resolution images of subviral complexes entering the nucleus. By using a U-2 OS CRISPR-Nup96-eGFP cell line, we will try to capture NPCs displaying CA positive structures stuck inside the pore and empty NPCs. The aim of this project will be to measure NPC diameters in order to reveal morphological changes of NPCs in the context of viral HIV-1 infection.

Furthermore, I started first experiments applying the GCE-modified CA variant in the context of murine leukemia virus (MLV). Based on existing literature and advice from the Briggs ab (Kun Qu, John Briggs Lab), I identified nine candidate residues possibly tolerating the incorporation of ncAAs. MLV constitutes an interesting target showing a multilayered morphology of the capsid in contrast to the cone-shaped HIV-1 core[364]. A CA click-labeled MLV variant, can be used in microscopy analysis to investigate uncoating dynamics of a virus, which requires nuclear envelope breakdown during mitosis in order to productively infect a cell[365]. Comparison of CA intensities of different viruses using the same approach might lead to new insights about uncoating dynamics and the presence of nuclear CA of MLV.

4 Conclusion

Together with insights from other recent studies, results from this thesis indicate that the HIV-1 capsid is indispensable for early viral replication (Figure 33). Besides its shielding function to protect the genome from antiviral sensors and degradation, the CA lattice is needed for protein interactions allowing efficient nuclear trafficking and entry. Also, nucleoplasmic trafficking is a guided process most likely dependent on CPSF6 interactions, which direct capsids towards defined nuclear locations to accumulate. The present findings confirm that the capsid escorts RTC/PICs all along the way to accomplish productive infection. Efficient reverse transcription of the genome occurs within the capsid shell, which contains pores to allow diffusion of dNTPs into the core. Inside the nucleus the capsid is directed by the host factor CPSF6 and potentially other factors towards nuclear speckles, offering a gene-rich environment, to uncoat and integrate into the host genome. These multifaceted tasks of the capsid may be the reason for the extreme genetic fragility of the capsid and make it very difficult to modify without destroying the function. Since its delivery into the nucleus was observed in different cell types, it can be suggested that the capsid constitutes the key player protecting the HIV-1 virus during infection.

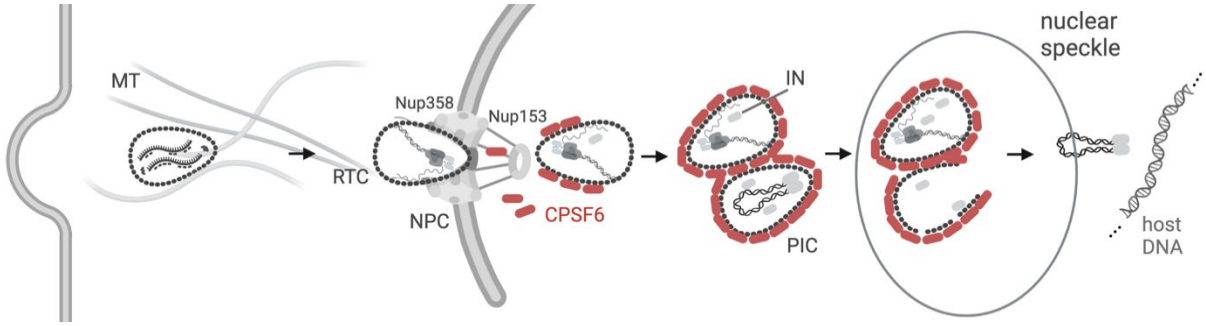


Figure 33. Model of nuclear clustering and uncoating. Individual capsids traffic towards the NE by hijacking cytosolic host microtubule (MT) network. Nuclear entry through NPCs of RTCs is facilitated by different CA-Nup interactions (e.g., Nup358, Nup153) and strong CPSF6 binding on the nuclear side. CPSF6 is directing the capsids *via* common routes towards nuclear gene-rich regions (nuclear speckles), which are the sites of uncoating and release of the viral genome to integrate into the host DNA near these speckles. Figure was created with BioRender.com.

5 Materials and Methods

Table 2 Antibodies

Reagent	Reference	Notes
rabbit polyclonal anti-HIV-1 CA	In-house	1:1000
rabbit polyclonal anti-hCPSF6	Atlas Antibodies; Cat# HPA039973	1:500
mouse monoclonal anti-hLamin A/C	Santa Cruz Biotech.; Cat# sc-7292	1:100; HeLa cells and MDM
mouse monoclonal anti hLamin B1	Santa Cruz Biotech.; Cat# sc-365962	1:200; T cells
sheep polyclonal anti-HIV-1 CA	Müller *et al.,* 2009[366]	1:5000
Anti-Cyp A	Abcam; Cat#: ab137684	1:1000
Anti-beta actin	Abcam; Cat#: ab8226	1:100
rabbit polyclonal anti-HIV-1 RT	Müller *et al.,* 2004[367]	1:1000
rabbit polyclonal anti-HIV-1 MA	In-house	1:100
Goat/donkey polyclonal Alexa Fluor (405, 488, 568 and 647) secondary antibodies	Thermo Fisher Scientific	1:1000
sheep IgG, IRDye 680RD conjugated	Li-COR Biosciences; RRID:AB_10954442	1:10000
rabbit IgG, IRDye 800CW conjugated	Li-COR Biosciences; RRID:AB_621848	1:10000
goat anti-mouse IgG conjugated with STAR Orange	Abberior GmbH; RRID:AB_2847853	1:200

Table 3 Plasmids

Reagent	Reference	Notes
pNLC4-3	Bohne and Kräusslich, 2004[368]	Synthetic CMV-containing pNL4-3 backbone
NNHIV	Zila *et al.,* 2021[136]	pNLC4-3 $IN_{D64N/D116N}$ $tat_{\Delta 33\text{-}64bp}$; Integration and transcription deficient HIV-1 proviral plasmid (provided by Thorsten Müller)
pVpr-$IN_{D64N/D116N}$.eGFP	David Bejarano (University Hospital Heidelberg)	Expression of labeled catalytically dead IN
Vpr-IN.mScarlet	Zila *et al.,* 2021[136]	Expression of labelled IN
pNLC4-3* (HIV-1*)	This thesis	*vpr* stop codon mutation TGA (vpr^{TGA})
HIV-1*AD8	This thesis	*vpr* stop codon mutation TGA (vpr^{TGA}) of CCR5 tropic strain
HIV-1*AD8 $CA14^{TAG}$	This thesis	*vpr* stop codon mutation TGA (vpr^{TGA}) of CCR5 tropic strain; amber stop codon at CA residue 14

eRF1-E55D	Schmied *et al.,* 2014[217]	Dominant-negative version of eRF1
pEA168	Cohen and Arbely, 2016[278]	Eukaryotic vector with expression cassettes for two proteins and 4 tRNA molecules
tRNA^{Pyl}/NESPylRS^{AF}	Nikić *et al*., 2016[229]	Modified pyrrolysine tRNA synthetase
pEA168-CMV-aaRS-4xU6tRNA	Schifferdecker *et al*. (in preparation)[270]	Provided by Anna-Lena Schäfer
pNESPylRS-eRF1dn-tRNA	Schifferdecker *et al*. (in preparation)[270]	Amber suppression components
pcDNA3.1(+) Vpu deltaTat33-64bp Rev Env	Thorsten Müller (University Hospital Heidelberg)	*vpr* and *vpu* Tat$\Delta_{33\text{-}64bp}$, *rev*, *env* subcloned from NNHIV
HIV-1*CA14^{TAG} (A14)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 14
HIV-1*CA15^{TAG} (I15)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 15
HIV-1*CA80^{TAG} (W80)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 80
HIV-1*CA83^{TAG} (L83)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 83
HIV-1*CA84^{TAG} (H84)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 84
HIV-1*CA86^{TAG} (V86)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 86
HIV-1*CA87^{TAG} (H87)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 87
HIV-1*CA88^{TAG} (A88)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 88
HIV-1*CA91^{TAG} (I91)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 91
HIV-1*CA92^{TAG} (A92)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 92
HIV-1*CA96^{TAG} (M96)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 96
HIV-1*CA98^{TAG} (E98)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 98
HIV-1*CA110^{TAG} (T110)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 110
HIV-1*CA111^{TAG} (L111)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 111
HIV-1*CA116^{TAG} (G116)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 116
HIV-1*CA119^{TAG} (T119)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 119
HIV-1*CA120^{TAG} (H120)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 120
HIV-1*CA124^{TAG} (I124)	This thesis	vpr^{TGA}; Amber stop codon at CA residue 124

Table 4 Cell lines

Reagent	**Reference**	**Notes**
HeLa TZM-bl	Wei *et al.,* 2002[297]	-
Embryonic kidney 293 T cells (HEK293T)	Pear *et al.,* 1993[369]	-
Human CD4$^+$ T lymphoblast cells SupT1	Smith *et al.,* 1984[370]	-

Table 5 Chemicals

Reagent	Reference	Notes
Cyclopropene-L-Lysine (CpK)	SiChem; Cat# SC-8017	100 mM in 0.2M NaOH+15% DMSO; -80°C
Bicyclo [6.1.0] nonyne-lysine (Fmoc-protected) (BCN)	SiChem; Cat# SC-8038	100 mM in DMF; -80°C
Tetrazine-silicon rhoadamine (SiR)	Spirochrome; Cat# SC-8017	1 mM stock in DMSO; -80°C
Pyrimidyl-tetrazine-AF647	Jena Biosciences; Cat# CLK-102	1 mM stock in DMSO; -20°C
Paraformaldehyde (PFA)	Electron Microscopy Sciences; Cat #15700	16% aqueous solution
Ascorbic Acid	Sigma Aldrich; Cat# A92902	10 mM in ddH_2O; -20°C
mCLING.ATTO488	Synaptic Systems; Cat# 710006AT3	Membrane probe, 50µM in ddH_2O; -80°C
Hoechst33258	Merck; Cat# 94403	1:1000
Enfuvirtide (T-20)	Roche; NIH AIDS Reagent Program Cat# 4624	Fusion inhibitor, 20mM in ddH_2O; -20°C
PF 3450074 (PF74)	Sigma Aldrich; Cat# SML0835	10 mM in DMSO; -20°C
TetraSpeck beads	Thermo Fisher Scientific; Cat# T7279	Fiducial marker
Aphidicolin (APC)	Merck; Cat# 178273	Cell cycle arrest
Nucleospin Gel and PCR Cleanup Kit	Macherey Nagel; Cat# 740609.50	-
NucleoBond PC 500, Maxi kit	Macherey Nagel; Cat# 740574.50	-
NEBuilder® HiFi DNA Assembly Master Mix	New England Biolabs; Cat# E2621L	-
EasySep™ magnet	Stemcell technologies; Cat# 18000	Magnet for immunomagnetic separation
EasySep™ Direct Human T cell Isolation Kit	Stemcell technologies; Cat# 19661	Immunomagnetic negative T cell selection from whole blood
EasySep™ Direct Monocyte Isolation Kit	Stemcell technologies; Cat# 19669	Immunomagnetic negative monocyte selection from whole blood
TransAct™	Mitenyi Biotec; Cat# 130-111-160	Activate and expand human T cells via CD28 and CD3
IL-2	Sigma Aldrich; Cat# 11011456001	Interleukin-2 human, T cell activation

Table 6 Software

Reagent	Reference	Notes

Fiji 1.53c	Schindelin *et al.,* 2012[371]; RRID:SCR_002285	General image analysis
Icy 2.0.3.0	De Chaumont *et al*., 2012[372]; RRID:SCR_010587	Intensity quantification, Correlation
eC-CLEM (Icy plugin; v 1.0.1.5)	Paul-Gilloteaux *et al.,* 2017[373]; http://icy.bioimageanalysis.org/plugin/ec-clem/	Post-correlation
Imspector 16.1.6905	Abberior Instruments GmbH	STED data acquisition and deconvolution
Prism v9.1.0	GraphPad Software Inc.; RRID:SCR_002798	Visualization and Plotting
Volocity 6.3	Perkin Elmer; RRID:SCR_002668	Data acquisition
Imaris 9.7.2	Bitplane AG; RRID:SCR_007370	Spot detection, Intensity quantification
Huygens Professional Deconvolution	SVI; RRID:SCR_014237	Deconvolution
SerialEM 3.7.9	Mastronarde *et al*., 2005[374]; RRID:SCR_017293	Tomogram acquisition, Pre-correlation
IMOD 4.9.4	Kremer *et al.,* 1996[375]; RRID:SCR_003297	Tomogram reconstruction
Amira 2019.3	Thermo Fisher Scientific; RRID:SCR_007353	Visualization and rendering
Image Studio™ Lite 5.0	Li-COR Biosciences; RRID:SCR_014211	Immunoblot analysis
Tecan i-control 1.10	Tecan	Plate reader software
CFX Manager Software 3.1	Bio-Rad; RRID:SCR_017251	Analysis of real-time PCR data
UCSF ChimeraX	Pettersen *et al.,* 2021[276]; https://www.rbvi.ucsf.edu/chimerax	Molecular graphics of protein structures
BioRender	BioRender.com	Science illustrations
ChemDraw	www.cambridesoft.com	Chemical drawing

Table 7 Primers

primer	**sequence**
Vpr_{TGA} a	ctggcatttgggtcagggagtc
Vpr_{TGA} b	cagggctctagttcaggatctactggctc
Vpr_{TGA} c	gagccagtagatcctgaactagagccctg
Vpr_{TGA} d	gttctcttaatttgctagc
Vpr_{TGA} qc f	ggagccagtagatcctgaactagagccctggtac

Vpr$_{TGA}$ qc r	gtaccagggctctagttcaggatctactggctcc
CA$_{TAG}$ a	cttgctgaagcgcgca
CA$_{TAG}$ d	cctaggggccctgcaatttttggctatgtgc
CA L6$_{TAG}$ c	cctatagtgcagaactagcaggggcaaatggta
CA L6$_{TAG}$ b	taccatttgcccctgctagttctgcactatagg
CA A14$_{TGA}$ b	agttctaggtgatatctactgatgtaccatttg
CA A14$_{TGA}$ c	caaatggtacatcagtagatatcacctagaact
CA I15$_{TAG}$ c	atggtacatcaggcctagtcacctagaacttta
CA I15$_{TAG}$ b	taaagttctaggtgactaggcctgatgtaccat
CA W80$_{TAG}$ c	gaggaagctgcagaataggatagattgcat
CA W80$_{TAG}$ b	tggatgcaatctatcctattctgcagcttcctc
CA L83$_{TAG}$ c	gaatgggatagatagcatccagtgc
CA L83$_{TAG}$ b	gcactggatgctatctatcccattc
CA H84$_{TAG}$ c	gaatgggatagattgtagccagtgcatgc
CA H84$_{TAG}$ b	gcatgcactggctacaatctatcccattc
CA V86$_{TAG}$ c	gattgcatccatagcatgcagggcctattgc
CA V86$_{TAG}$ b	gcaataggccctgcatgctatggatgcaat
CA H87$_{TAG}$ c	gcatccagtgtaggcagggcct
CA H87$_{TAG}$ b	aggccctgcctacactggatgc
CA A88$_{TAG}$ c	cagtgcattaggggcctattgc
CA A88$_{TAG}$ b	gcaataggcccctaatgcact
CA I91$_{TAG}$ c	catgcagggccttaggcaccaggccag
CA I91$_{TAG}$ b	ctggcctggtgcctaaggccctgcatg
CA A92$_{TAG}$ c	catgcagggcctatttagccaggccagatg
CA A92$_{TAG}$ b	catctggcctggctaaataggccctgcatg
CA M96$_{TAG}$ c	accaggccagtagagag
CA M96$_{TAG}$ b	gttctctctactggcct
CA E98$_{TAG}$ c	agatgagatagccaaggg
CA E98$_{TAG}$ b	cccttggctatctcatct
CA T110$_{TAG}$ c	tactagttagcttcagga
CA T110$_{TAG}$ b	tcctgaagctaactagta
CA L111$_{TAG}$ c	ctagtacctagcaggaac
CA L111$_{TAG}$ b	gttcctgctaggtactag
CA G116$_{TAG}$ c	caaatatagtggatgacac
CA G116$_{TAG}$ b	gtgtcatccactatatttg
CA T119$_{TAG}$ c	ggatggatgtagcataat

CA T119$_{TAG}$ b	attatgctacatccatcc
CA H120$_{TAG}$ c	ggatgacatagaatccacc
CA H120$_{TAG}$ b	ggtggattctatgtcatcc
CA I124$_{TAG}$ c	aatccaccttagccagtagg
CA I124$_{TAG}$ b	cctactggctaaggtggatt
AD8 Vpr$_{TGA}$ a	ggaagccataataagaattctgca
AD8 Vpr$_{TGA}$ b	ggctctagtctaggatctact
AD8 Vpr$_{TGA}$ c	agatcctgaactagagccct
AD8 Vpr$_{TGA}$ d	ccctgtaatatttgatgaacatctaatttg
pPylRS fwd	cgggcggccgcgagaacccactgcttactgg
pPylRS rev	cggctcgaggctggatatctgcag aattcc
eRF1 fwd	cggggtacczgacctccatagaag acacc
eRF1 rev	cggacgcgtgctatcattagtaatcatccaaat

Generation of plasmids

Standard molecular biology techniques were used for cloning of plasmids. Plasmid sequences were verified by commercial Sanger Sequencing (Eurofins Genomics, GER; Seqlab, GER). According to manufacturer's instructions, PCR reactions were performed with Q5 High-Fidelity DNA Polymerase (New England Biolabs, GER) or Phusion DNA Polymerase (New England Biolabs). All primers were ordered at Eurofins Genomics. *E. coli* Stbl2 (Thermo Fisher Scientific, USA) were used for plasmid amplification.

pNESPylRS-eRF1dn-tRNA. The plasmid comprising all amber suppression components was provided by Ana-Lena Schäfer (Bachelor Thesis, University Hospital Heidelberg) and based on the eukaryotic expression vector pEA168[278], which was kindly provided by Eyal Arbely (Ben-Gurion University of the Negev, Israel). The sequence of a modified pyrrolysine tRNA synthetase was subcloned from tRNA^{Pyl}/NESPylRS$^{AF\ 229}$ (pPylRS fwd, pPylRS rev) into pEA168 using unique restriction sites HindIII/XbaI. This resulted in the plasmid pEA168-CMV-aaRS-4xUtRNA. Furthermore, the sequence of a dominant version of the eukaryotic release factor 1 (eRF1(E55D)) was PCR amplified from peRF1-E55D[217] (eRF1 fwd, eRF1 rev) and subcloned into a pEA168-CMV-aaRS-4xU6tRNA expression cassette under the EF1 promoter using KpnI/MluI restriction sites. The resulting plasmid was named pNESPylRS-eRF1dn-tRNA.

HIV-1*. A double point mutation was introduced into the *vpr* coding region of the synthetic CMV promoter containing pNLC4-3[368] to avoid ncAA-incorporation by exchanging the TAG amber stop

codon into an opal stop codon TGA. The vpr region was amplified *via* overlap PCR with overlapping forward and reverse primers containing the desired mutations. PCR 1 (Vpr_{TGA} a, Vpr_{TGA} b) and PCR 2 (Vpr_{TGA} c, Vpr_{TGA} d) were performed in parallel with pNLC4-3 as the template and both isolated products were used as templates for PCR 3 using primers Vpr_{TGA} a and Vpr_{TGA} d. The PCR product containing the TGA mutation was subcloned into pNLC4-3 by using unique EcoRI/StuI cutting sites.

HIV-1*CAxx^{TAG}. For construction of 18 plasmids suitable for click-labeling the CA *via* amber suppression, a TAG amber stop codon was incorporated into the protein sequence. To incorporate a ncAA at different sites, the stop codon was introduced at different CA residues. Two overlapping primers containing the TAG codon ($CAxx_{TAG}$ b and c) were designed individually for every selected residue. PCR 1 (CA_{TAG} a, $CAxx_{TAG}$ b) and PCR 2 ($CAxx_{TAG}$ c, CA_{TAG} d) were performed on pNLC4-3 in parallel to use both PCR products in PCR 3 with primers CA_{TAG} a and CA_{TAG} d. The double-stranded PCR 3 product was subcloned into HIV-1* using BssHII/ApaI. See table 6 for primers.

NNHIV*. To generate a pNLC4-3 variant without GCE-modification of Vpr, which is suitable for imaging under BSL-2 conditions, the RT-competent but integration and transcription deficient HIV-1 derivative NNHIV (pNLC4-3 $IN_{D64N/D116N}$ $tat\Delta_{33\text{-}64bp}$)[136] was modified, kindly provided by Thorsten Müller (University Hospital Heidelberg). *vpr* ORF mutation to avoid ncAA incorporation into the Vpr protein was carried out *via* site-directed mutagenesis. pcDNA3.1(+) Vpu $deltaTat_{33\text{-}64bp}$ Rev Env plasmid, containing the *vpr* coding region, *vpu* and $Tat\Delta_{33\text{-}64bp}$ was used for the PCR reaction. Mutagenesis of the *vpr* stop codon was performed using QuikChange Site-Directed Mutagenesis[376] with overlapping primers including the desired mutation (Vpr_{TGA} qc f and r), and the mutated fragment was transferred back into NNHIV *via* EcoRI/StuI, resulting in plasmid NNHIV*.

NNHIV CA14^{TAG}. The integration and transcription deficient variant for CA click-labeling was constructed *via* subcloning of the *capsid* ORF from HIV-1*$CA14^{TAG}$ into NNHIV* using BssHII/ApaI.

HIV-1*AD8 (AD8*). The following reagent was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: Human Immunodeficiency Virus 1 (HIV-1) AD8 Infectious Molecular Clone (pNL(AD8)) , ARP-11346, contributed by Dr. Eric O. Freed. ARP-11346 is a CCR5-utilizing derivative, which contains the NL4-3 backbone with AD8-1 env. It is a full-length replication competent and infectious chimeric molecular clone. This derivative of pNL4-3 (pNL(AD8)) was used for infection of CCR5 exhibiting MDMs. The stop codon of the *vpr* ORF in pNL(AD8) was mutated *via* site-directed mutagenesis into TGA *via* overlap PCR as described for HIV-1* and NNHIV* using primers AD8 Vpr_{TGA} a and AD8 Vpr_{TGA} b for PCR 1 and AD8 Vpr_{TGA} c and AD8 Vpr_{TGA} d for PCR 2. The PCR

3 product was subcloned using EcoRI/BsaBI unique restriction sites into pNL(AD8) resulting in HIV-1*AD8 (AD8*).

HIV-1*AD8 CA14^{TAG} (AD8*CA14^{TAG}). The capsid region of HIV-1*CA14^{TAG} was subcloned into HIV-1*AD8 using BssHII/ApaI restriction sites.

Cell and tissue culture

Human embryonic kidney 293T (HEK293T) cells and HeLa-derived TZM-bl (JC53BL-13) cells were cultured in humidified incubators at 37° C and 5 % CO_2 kept in Dulbecco's Modified Eagle's Medium (DMEM; Thermo Fisher Scientific) supplemented with 10 % fetal calf serum (FCS; Sigma Aldrich), 100 U/ml penicillin and 100 µg/ml streptomycin (PAN Biotech, GER). The human T lymphoblast cell line SupT1, MDM and primary CD4$^+$ T cells were cultured in RPMI 1640 (Thermo Fisher Scientific) supplemented with 10 % heat-inactivated FCS, 100 U/ml penicillin and 100 µg/ml streptomycin. MDM were additionally supplemented with 5 % human AB serum (Sigma Aldrich).

Isolation of primary cells. Primary cells were isolated from buffy coats of healthy blood donors received from Heidelberg University Hospital Blood Bank under the regulations of the local ethics committee. CD4$^+$ T cells were isolated using the EasySep^{TM} Direct Human T Cell Isolation Kit (Stemcell technologies, GER) and an EasySep^{TM} magnet (Stemcell technologies) according to manufacturer's instructions. Cells were activated with 100 U/ml IL-2 (Sigma Aldrich) and 1:100 T cell TransAct^{TM} human (Mitenyi Biotec, GER). Monocytes were isolated with EasySep^{TM} Direct Human Monocyte Isolation Kit (Stemcell technologies) and EasySep^{TM} magnet (Stemcell technologies) according to manufacturer's instructions. Monocytes were differentiated into MDM for 7 - 10 days. Medium was exchanged every third day.

Viral particle production and quantification of released HIV-1 particles

For small-scale particle production, 3 x 10^5 HEK293 T cells were seeded in 6-well plates the day before co-transfection of proviral DNA (HIV-1*/HIV-1*CAxx^{TAG}/NNHIV*/NNHIV*CA14^{TAG}/AD8*/AD8*CA14^{TAG}) and pNESPylRS-eRF1dn-tRNA at a molar 2.22:1 ratio. Initial experiments were using a triple transfection with HIV-1*/ tRNA^{Pyl}/NESPylRS^{AF}/peRF1-E55D in an equal molar ratio (3-plasmid system). Transient transfection was performed with PEI according to standard procedures. Briefly, 3 µg total DNA were mixed with PEI (1:3) in serum-free DMEM and incubated at room temperature for 30 min. The mixture was then added to the cell culture medium in a total volume of 2 ml/well supplemented with a 500 µM final concentration CpK (SiChem, GER; 100 mM prediluted 1:4 in 1 M HEPES shortly before use) and

100 µM ascorbic acid (Sigma Aldrich). Cells were kept at 37° C. At 40 h p.t., supernatant of transfected cells was harvested, filtered through a 0.45 µm nitrocellulose filter (Carl Roth, GER) and layered onto 100 µl of 20 % (w/v) sucrose solution. Particles were concentrated by ultracentrifugation for 45 min at 44,000 rpm using a Beckman TLA-100 fixed angle rotor (Beckman Coulter, GER). Supernatants were carefully removed to solve the particles in phosphate-buffered saline (PBS) with 10 % FCS and 10 mM HEPES (pH 7.5). Particle solution was aliquoted and stored at -80° C.

Large-scale virus production was performed in T175 tissue culture flasks seeded with HEK293T cells (~ 80 % confluency) the day before co-transfection of proviral DNA (HIV-1*/HIV-1*CAxx^{TAG}/NNHIV*/NNHIV*CA14^{TAG}/AD8*/AD8*CA14^{TAG}) and pNESPylRS-eRF1dn-tRNA at a molar 2.22:1 ratio. For production of double labeled NNHIV*CA14^{SiR} IN.eGFP or HIV-1*CA14^{SiR} IN.mScarlet, cells were transfected with NNHIV*/NNHIV*CA14^{TAG}/HIV-1*/HIV-1*CA14^{TAG}, pNESPylRS-eRF1dn-tRNA and pVpr-IN$_{D64N/D116N}$.eGFP/pVpr-IN.mScarlet in a molar 2.45:1:1.1 ratio. Cells were transfected with calcium phosphate precipitation according to standard protocols. Briefly, 50 µg total DNA was used per flask. At 6 h p.t., transfection medium was exchanged by fresh DMEM containing a final concentration of 500 µM CpK (100 mM prediluted 1:4 in 1 M HEPES shortly before use) and 100 µM ascorbic acid. At 44 h p.t., supernatant of virus-producing cells was harvested and filtered through a 0.45 µm nitrocellulose filter (Carl Roth) and layered onto 6 ml of a 20 % (w/v) sucrose solution. Particles were concentrated by ultracentrifugation for 90 min at 28,000 rpm using a Beckman TLA-100 fixed angle rotor. Supernatants were carefully removed to solve the particles in phosphate-buffered saline (PBS) with 10 % FCS and 10 mM HEPES (pH 7.5). Particle solution was aliquoted and stored at -80° C.

Quantitation of released viral particles was performed in a one-step SYBR green I-based product-enhanced reverse transcriptase assay (SG-PERT)[269]. 5 µl of supernatant or concentrated particles was lysed in 2x lysis buffer for 10 min and mixed with 90 µl PCR buffer. 10 µl of this solution was used for RT-PCR measurements using a CFX Touch 96-well Real-Time PCR Detection System (Bio-Rad, USA) and CFX Manager Software 3.1 (Bio-Rad). By adding a standard dilution series of a characterized virus stock, virus concentrations were determined based on a generated standard curve.

Immunoblotting

Transfected cells or concentrated particles were lysed in 3x SDS buffer (150 mM Tris-HCl, 30 % glycerin, 0.06 % bromophenol blue, 20 % ß-mercaptoethanol and 6 % SDS) for 10 min at 95° C.

Proteins were separated on a 15 % SDS-PAGE gel in Laemmli buffer and blotted for 1 h to nitrocellulose membranes by semi-dry blotting (0.8 mA/cm^2). Membranes were blocked in 1:4 pre-diluted Odyssey Blocking buffer (Li-COR Bioscience, GER) for 15 min at room temperature, followed by incubation with primary antibodies against protein of interests at 4° C overnight. After washing with PBS, fluorophore-coupled secondary antibodies IRDye™ anti-sheep 680CW (1:10,000; Li-COR Biosciences) and anti-rabbit 800CW (1:10,000; Li-COR Biosciences) were incubated for 45 min at room temperature, kept in the dark. Bound antibodies were detected and quantified by a Li-COR Odyssey CLx infrared scanner using Image Studio lite software 5.0 (Li-COR Biosciences) according to manufacturer's instructions. To quantify absolute CA amounts, a CA standard (2.5 ng/µl; in-house) was added in a serial twofold dilution.

In-gel fluorescence

Harvested supernatant of virus-producing cells was incubated with Tet-SiR (125-500 nM; Spirochrome, GER), filtered and concentrated *via* ultracentrifugation through a sucrose cushion. Labeled particles were lysed and separated on SDS-PAGE gel. Acrylamide gels were scanned using a Li-COR Odyssey CLx infrared scanner (Li-Cor Biosciences) at an emission wavelength of 700 nm to detect fluorescent bands.

Labeling efficiency of click-labeled particles

1-5 µl concentrated and Tet-SiR labeled particles were incubated on PEI-coated (1 mg/ml) 15-well ibidi µ-Slide angiogenesis dishes in a total volume of 40 µl (in PBS) for 60 min at 37° C. Immobilized particles were washed with PBS, fixed using 4 % PFA (Electron Microscopy Sciences, USA) for 15 min and permeabilized using 0.05 % Triton-X 100 for 15 min at room temperature. After blocking of the slides with 2.5 % BSA/PBS for 15 min, the primary antibody rαCA was diluted 1:1,000 in 0.5 % BSA/PBS and incubated for 1 h at room temperature. Slides were washed three times with PBS and Dαrabbit 488 secondary antibody was then added in 0.5 % BSA/PBS (1:1,000) for 1 h at room temperature. After washing three times in PBS, slides were kept at 4° C until imaging.

Infectivity assays

All assays to determine infectivity of a virus particle preparation were performed in HeLa-derived TZM-bl cells. TZM-bl cells were derived from HeLa cells engineered to express CD4, CXCR4 and CCR5. They contain reporter genes for *E. coli* ß-galactosidase and firefly Luc under the control of HIV-1 long terminal repeats[297], which enables detection of HIV-1 infection.

Luciferase based (Steady Glo). TZM-bl cells (5x10^3 cells/well) were seeded in 96-well plates the day before infection with purified particles or virus-containing supernatant in serial twofold dilutions. At 48 h p.i., medium was removed and 100 µl Steady Glo luciferase assay (Promega, GER) reagent was added to the cells for 10 min at room temperature to lyse the infected cells and viral particles. 80 µl of the cell solution was transferred into 96-well plates suitable for luciferase activity measurement. Relative light units (RLU) were detected using an Infinite 200 PRO plate reader (Tecan Life Sciences, SUI) and the Tecan i-Control software 1.10 (Tecan Life Sciences, SUI). RLUs were plotted against the virus/supernatant volume to perform linear regression analysis. Relative infectivity was calculated by normalization of RLUs to RT activity of each sample determined by SG-PERT assay.

Blue-cell assay. TZM-bl cells (5 x 10^3 cells/well) were seeded in flat-bottom 96-well plates the day before infection with particles diluted 1:200 in DMEM and serial hundredfold dilutions. 100 µl of each dilution was added to the cells in triplicates. At 6 h p.i., medium was exchanged by fresh DMEM containing 50 µM T-20 (Enfuvirtide; Roche, GER) to prevent second-round infection. At 48 h p.i., cells were fixed with ice-cold methanol/acetone (1:1) for 10 min at room temperature. 100 µl of a 1x staining solution (500x stock: 1.5M Potassium-ferri-cyanide, 1.5M Potassium-ferro-cyanide, 0.5M $MgCl_2$) containing 200 µg/ml 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal) in DMF, was added for 2-3 h at 37° C. After removing the staining solution, cells were kept in PBS to manually quantify blue cell colonies per well and calculate infectious units/ml.

MA IF staining. 15-well ibidi µ-Slide angiogenesis dishes were seeded with 3.5 x 10^3 TZM-bl cells. The next day 60, 40, 20, 10 and 5 mU RT/well (determined previously in SG-PERT assay) were used for cell infections. At 6 h p.i., medium was exchanged by fresh DMEM containing 50 µM T-20 (Enfuvirtide) to prevent second-round infection. At 48 h p.i., cells were washed with PBS, fixed and permeabilized using 4 % PFA (15 min) and 0.5 % Triton-X 100 (15 min). Cell were stained with Hoechst and against HIV-1 MA by IF using a polyclonal antibody (in-house). Cells infected with different virus amounts were imaged by SDCM (20x objective). The amount of infected cells was quantified in Fiji[371] in 12 FOVs per condition using the cell counter plugin. Non-infected cells were imaged as a negative control and for setting a threshold for infected cells based on MA intensities. Number of infected cells [%] was plotted against RT values [mU/well] to perform linear regression analysis.

Based on the number of infected cells and the volume used for infection, infection units/ml (IF/ml) of a virus preparation could be determined. Using the formula below, the multiplicity of infection (MOI) was calculated:

$$\text{MOI} = \frac{\text{IF/ml} \times \text{virus volume}}{\text{total cell number}}$$

Cell infection, treatment, and staining

15-well ibidi µ-Slide angiogenesis dishes were seeded with TZM-bl cells (3.5 x 10^3 cells/well) the day before infection. Cells were infected (MOI ~ 0.8 or ~ 8) in 50 µl total volume as indicated in the figure legends. Suspension cells (SupT1 and primary $CD4^+$ T cells) were infected in v-bottom 96-well plates in a total volume of 40 µl medium for 24 h. After infection, the whole volume containing ~2 x 10^4 cells/well were immobilized on PEI-coated 15-well ibidi µ-Slide angiogenesis dishes.

PF74 treatment. Before fixation, growth medium was exchanged at indicated time points with fresh medium containing 15 µM PF-3450074 (PF74; Sigma Aldrich). After 1 h incubation at 37° C, cells were washed with PBS and fixed as described above. For suspension $CD4^+$ T cells, PF74 was added to the inoculum in v-bottom 96-well plates to a final volume of 50 µl with desired final concentration before transfer to PEI-coated 15-well ibidi µ-Slide angiogenesis dishes.

IF staining and fixation. If not indicated differently in the figure legends, infected cells were fixed using 4 % PFA for 15 min, permeabilized using 0.5 % Triton-X100 for 15 min and extracted with ice-cold 100 % methanol for 10 min. Fixed samples were incubated for blocking in 2.5 % BSA/PBS for 15 min at room temperature before 3x washing in PBS. Primary antibodies, diluted in 0.5 % BSA/PBS, were incubated for 1 h at room temperature. Then, slides were washed 3x in PBS and fluorescent secondary antibodies diluted in 0.5 % BSA/PBS were added for 45 min at room temperature. After washing again 3x in PBS, samples were kept at 4° C until imaging. See table 1 for dilutions of the antibodies.

mCLING staining. Seeded TZM-bl cells were stained in DMEM containing 2 µM mCLING ATTO488 (Synaptic Systems, GER) for 30 min at 16° C. Afterwards, infection with viral particles was performed for 3-4 h at 37° C. Cells were washed once in fresh DMEM before fixation with 4 % PFA + 0.2 % glutaraldehyde (GA) for 90 min at room temperature. Nuclei were stained with 5 µg/ml Hoechst (Merck, GER) for 15 min. Cells were washed 3x in PBS and kept in PBS at 4° C until imaging.

Cell viability assay. TZM-bl cells (9 x 10^3 cells/well) were seeded into 96-well flat bottom plates (Greiner Bio-one, GER). The next day cells were incubated in technical duplicates with 100 µl DMEM containing 0, 0.5, 1, 2, 4 µM mCLING ATTO488 for 30 min at 16° C. Subsequently, cells were

trypsinized and stained with Trypan blue[377] to automatically count the cell number using a TC20™ Automated Cell Counter (BioRad, GER).

Imaging

Spinning disc confocal microscopy (SDCM). All images were acquired using a Perkin Elmer Ultra VIEW VoX 3D SDCM (Perkin Elmer, USA). Imaging of infected TZM-bl cells was performed with a 60x immersion oil objective (numeric aperture [NA] 1.49; Perkin Elmer) generating multichannel z-series with a spacing of 200 nm to allow imaging of the whole cell volume (3D). Immobilized particles, CD4+ T cells and MDM were imaged using a 100x immersion oil objective (NA 1.49; Perkin Elmer). Laser and filter configurations were as follows for the different channels: 405/420 - 460 nm (Alexa405, Hoechst), 488/510 - 540 nm (Alexa488, eGFP), 568/589 - 625 nm (Alexa568, mScarlet) and 647/665 - 705 nm (SiR).

Live imaging. Cells were cultured in FluoroBrite DMEM Imaging medium (Thermo Fisher Scientific) containing 10 % FCS, 100 U/ml penicillin, 100 µg/ml streptomycin, 4 mM GlutaMAX (Gibco Life technologies, USA), 20 mM HEPES (pH 7.4) and 2 mM sodium pyruvate (Gibco Life technologies). Time-resolved imaging of living TZM-bl cells was performed with an inverted Nikon Eclipse Ti2 (Nikon, JPN) with a spinning disc unit (Yokogawa CSU-W1, Nikon) and an emission light detector with two EMCCD cameras (iXon 888, Andor, Oxford Instruments, GBR) used in single mode. Cells were kept in a humidified incubation chamber at 37° C and 5 % CO_2 and imaged using a 100x oil immersion objective (Nikon CFI Apochromat TIRF, NA 1.49). Multichannel (405/420 - 460 nm; 488/510 - 540 nm; 647/665 - 705 nm) 3D stacks of several positions were recorded using the automated stage and Nikon Imaging Software Elements v.5.02

Stimulated emission depletion (STED) nanoscopy. Super-resolution microscopy was performed using an Abberior STED System (Abberior Instruments GmbH, GER) with a 775 nm STED laser and 100x oil-immersion objective (Olympus UPlanSApo, NA 1.4; Olympus, GER). The 640 nm laser line was recorded in confocal and STED mode, 488 and 590 nm only in confocal. The STED laser power was set to 70 – 100 % of the maximum power (1250 mW) with a pixel dwell time of 10 µs and 15 nm pixel size. Recorded images were deconvolved using the Imspector Software (Richardson-Lucy algorithm; Abberior Instruments GmbH) and Huygens Professional Deconvolution (Scientific Volume Imaging, GER).

Electron microscopy (EM). 4 x 10^5 HEK293T cells were seeded on glass coverslip-bottom petri dishes (MatTek, USA) the day before co-transfection with HIV-1*$CA14^{TAG}$ and pNESPylRS-eRF1dn-tRNA using calcium phosphate precipitation according to standard procedures. At 6 h p.t., 500 µM

CpK and 100 µM ascorbic acid in DMEM were added. At 44 h p.t., cells were fixed for 90 min at room temperature with pre-warmed 2 % formaldehyde + 2.5 % GA in 0.1 M cacodylate buffer, pH 7.4. After washing once with 0.1 M cacodylate buffer, cells were post-fixed for 1 h on ice with 2 % osmium tetroxide (Electron Microscopy Sciences). Dehydrating of the cells was performed with an increasing cold ethanol series (30, 50, 70, 80, 90, and 100 %; on ice) and two anhydrous acetone series at room temperature. The coverslip with attached cells was flat embedded in Epon resin and 70 nm-thick sections were cut with an ultramicrotome (Leica EM UC6), collected on formvar-coated 100-mesh copper EM grids (Electron Microscopy Sciences) and stained with a 3 % uranyl acetate in 70 % MeOH for 10 min, and lead citrate for 7 min. JEOL JEM-1400 electron microscope operating at 80 kV (Jeol Ltd., JPN), equipped with a bottom-mounted 4K by 4K pixel digital camera (TemCam F416; TVIPS GmbH, GER) was used for imaging of the cells.

CLEM and electron tomography. SupT1 cells were pre-incubated for 16 h with 1 µM APC (Merck) in a 96-well plate (2 x 10^5 cells/well; U-bottom; Greiner Bio-one, 650180). After centrifugation of the cells (200 x g, 3 min), the pellet was resuspended in RPMI medium with HIV-1*CA14^{SiR} particles (MOI ~ 0.4). Cells were incubated with viral particles for 120 min at 16° C and subsequently processed for CLEM and ET as described previously[136]. In brief, cells were transferred to glass-bottomed 'microwell' dishes (MatTek) containing carbon-coated and retronectin-coated sapphire discs (Engineering Office M. Wohlwend, SUI). After high pressure freezing, sapphire discs were transferred from liquid nitrogen to the freeze-substitution (FS) medium (0.1 % uranyl acetate, 2.3 % methanol and 2 % H_2O in acetone) tempered at -90° C. Samples were FS-processed and embedded in Lowicryl HM20 resin (Polysciences, USA) according to a modified protocol of Kukulski *et al.*[378]. For CLEM-ET, 250 nm-thick resin sections were placed on a slot (1 × 2 mm) EM copper grids covered with a formvar film (Electron Microscopy Sciences, FF2010-Cu). Grids were decorated with fiducial marker and stained with Hoechst to visualize the nuclei. z stacks of sections were acquired using the PerkinElmer UltraVIEW VoX 3D Spinning-disc Confocal Microscope (Perkin Elmer) and a 100 × oil immersion objective (NA 1.49; Nikon), with a *z*-spacing of 200 nm and excitation with the 405-, 488-, 568- and 647-nm laser line. Acquired z stacks were visually examined using Fiji software[371] and intracellular CA SiR signals were identified. EM grids were decorated with 15 nm protein-A gold particles for tomogram alignment and stained with uranyl acetate and lead citrate. Grids were loaded to a Tecnai TF20 (FEI, USA) electron microscope (200 kV) with a field emission gun and a 4K-by-4K pixel Eagle CCD camera (FEI). Positions of CA SiR signals were pre-correlated with imported SDCM images in SerialEM[374] as described previously[379]. Single-axis electron tomograms were carried out. Tomographic tilt ranges were typically from –60° to 60° with an angular increment of 1° and

1.13 nm pixel size. Alignments and 3D-reconstructions of tomograms were done with IMOD software[375]. Post-correlation was performed using the eC-CLEM plugin[373] in Icy software[372].

Image analysis

Images were screened in Fiji and filtered by applying a mean filter (kernel size: 0.5-2 pixels). Uneven illuminated background was corrected by the "rolling ball" algorithm.

Fiji standard greyscale or fire lookup table (LUT) were used to visualize single channel images in enlargements.

Semi-automated infectivity assay. Recorded images were segmented in Fiji for automated nuclei counting and infected cells were counted manually using the cell counter plugin.

Labeling efficiency. CA IF particles were detected using the Icy software[372] and spot detector. The mean SiR intensity of 5 randomly drawn rectangles (regions without particles) in each individual image served as a background, which was subtracted from SiR values of detected particles. Mean CA SiR of these foci were plotted against CA IF mean intensities. The threshold was set manually t = 1,000 based on visual inspection after background subtraction and every particle with higher intensity was considered as CA SiR positive.

Particle detection in 3D. Particle distribution and SiR intensity measurement throughout the entire cell volume of infected cells was performed *via* reconstruction of z-series in 3D with Imaris 9.2 software (Bitplane AG, SUI). 3D ellipsoid objects with 300 nm diameter in x-y dimension were automatically generated based on detected CA IF signal with the spot detector module. Threshold was set at t = 7,000 a.u. and adjusted individually for each cell. SiR intensities were quantitated for every object with SiR intensities above the set threshold.

To determine particle distribution within a cell, Imaris and lamin A/C staining were used to identify and classify nuclear objects manually. Lamin A/C intensities were used to identify NE-associated particles (~ 6,300 - 9,100 a.u.), which were manually controlled by visual inspection in 3D for each cell. Remaining particles were classified as PM/cytoplasmic. Clusters of the SiR-dye were visually identified and detected particles in this area were excluded.

To identify endosomal particles the threshold for mCLING ATTO488 was set to t = 5,900 a.u. The threshold was set based on the lowest mCLING ATTO488 intensity of particles in a sample treated with T-20 (fusion inhibitor).

Data visualization and statistical analysis

Data was visualized and statistical significance was determined using Prism v9.1.0 (GraphPad Software Inc, USA). A two-tailed non-paired Mann-Whitney test ($\alpha = 0.05$) was used for statistical analysis of non-parametric data. Individual datapoints, mean/median values and error were defined in the figure legends.

6 Acknowledgements

First, I would like to thank Prof. Dr. Barbara Müller for offering me the possibility of starting in her lab as a master's student and continuing working on this exciting topic during my Ph.D. I highly appreciate your effort to discuss data and proofread my thesis, drafts, and manuscripts. I enjoy working in your group and the feeling that someone cares about my project and progress. I want to emphasize that you encouraged me to work independently and take responsibility for teaching students and presenting our data. Due to all your guidance, input, and criticism, I improved a lot in writing, presenting, and discussing scientific data!

Additionally, I would like to thank Prof. Dr. Hans-Georg Kräusslich for his work as my first reviewer and for listening to all my presentations during our lab meetings and TACs. I am thankful for all the scientific advice and critical questions fundamental for this project and thesis. I am fascinated by his incredible memory, remembering almost every mean fluorescence intensity I presented somewhen in the last years. I would also like to thank Prof. Dr. Friedrich Frischknecht for being my second reviewer, accepting me as a rotation student at the beginning of my master's, and his supervision. Thanks to Dr. Steeve Boulant as a member of my TAC, who initially agreed on being a referee before he moved to the USA. I appreciated your positive attitude, reflected ideas, and interesting discussions about my project or working in academia. This thesis significantly gained value by collaborating with Vojtech Zila, who performed the CLEM experiments presented here. Thank you for commenting on my experimental plans, discussing existing literature, and my confusing results. Special thanks to Thorsten Müller for contributing a lot to this project by proofreading my thesis, discussing my results, and improving experimental setups. I am very lucky that I had the chance to work with you, and I want to thank you for your patience while answering all my (sometimes) naive, stupid, and senseless questions.

One of the most important persons in our lab is Maria Anders-Össwein! Thank you, Maria, for your great support and for performing thousands of particle preparations with hundred different conditions. You saved my (Ph.D.) life by running all these centrifuge rotors in the BSL-3 lab, pipetting all these infectivity assays, and performing tonnes of western blots.

I also want to thank the entire Müller group for being supportive and making lab work less miserable. Special thanks to Maike for proofreading my thesis, finding every typing error, and being the best 'Monk' I could have wished for! Thank you, Tamara Naumoska and Djordje Salai, for your support, proofreading, and every coffee break: A sorrow shared is a sorrow halved! I am really sorry that I did not learn to speak 'gypsy language' to be a proper member of your gypsy family or a replacement for Mia.

I am super happy and a bit surprised that all the last years' effort is now leading to a doctoral thesis and degree. But I am even more grateful for the time I spent with all the great people I met, especially Thorsten, Tamara, Maike, and Djordje. I am not only richer in knowledge but also richer in new friends.

Finally, I would like to thank my parents and friends for their mental support showing me that science and this project are not the only things, which matter. Many thanks to Tess, La, Jenny, Julia, Steffen, Hagen, Miri, and Fabi for being the best friends and having the required humor and understanding for fine wines and good food, making life much more enjoyable.

7 References

1. Microbiology by numbers. *Nat. Rev. Microbiol.* **9**, 628–628 (2011).
2. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci.* **36**, 344–355 (1950).
3. Krupovic, M., Dolja, V. V. & Koonin, E. V. Origin of viruses: primordial replicators recruiting capsids from hosts. *Nat. Rev. Microbiol.* **17**, 449–458 (2019).
4. Baltimore, D. Expression of animal virus genomes. *Bacteriol. Rev.* **35**, 235–241 (1971).
5. Barre-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868–871 (1983).
6. TEMIN, H. M. & MIZUTANI, S. Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of Rous Sarcoma Virus. *Nature* **226**, 1211–1213 (1970).
7. BALTIMORE, D. Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses. *Nature* **226**, 1209–1211 (1970).
8. Jung, I. Y. *et al.* Trends in mortality among ART-treated HIV-infected adults in the Asia-Pacific region between 1999 and 2017: results from the TREAT Asia HIV Observational Database (TAHOD) and Australian HIV Observational Database (AHOD) of IeDEA Asia-Pacific. *J. Int. AIDS Soc.* **22**, e25219 (2019).
9. WHO. HIV/AIDS Key Facts. (2021). Available at: https://www.who.int/news-room/fact-sheets/detail/hiv-aids.
10. WHO. The top 10 causes of death. (2020).
11. Skinner, D. & Mfecane, S. Stigma, discrimination and the implications for people living with HIV/AIDS in South Africa. *SAHARA-J J. Soc. Asp. HIV/AIDS* **1**, 157–164 (2004).
12. Rankin, W. W., Brennan, S., Schell, E., Laviwa, J. & Rankin, S. H. The Stigma of Being HIV-Positive in Africa. *PLoS Med.* **2**, e247 (2005).
13. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
14. Jiang, H., Zhou, Y. & Tang, W. Maintaining HIV care during the COVID-19 pandemic. *Lancet HIV* **7**, e308–e309 (2020).
15. Shiau, S., Krause, K. D., Valera, P., Swaminathan, S. & Halkitis, P. N. The Burden of COVID-19 in People Living with HIV: A Syndemic Perspective. *AIDS Behav.* **24**, 2244–2249 (2020).
16. Cooper, T., Woodward, B., Alom, S. & Harky, A. Coronavirus disease 2019 (COVID-19) outcomes in HIV/AIDS patients: a systematic review. *HIV Med.* **21**, 567–577 (2020).
17. Frankel, A. D. & Young, J. A. T. HIV-1: Fifteen Proteins and an RNA. *Annu. Rev. Biochem.* **67**, 1–25 (1998).
18. Freed, E. O. HIV-1 Replication. *Somat Cell Mol Genet* 13–33 (2001).
19. Novikova, M., Zhang, Y., Freed, E. O. & Peng, K. Multiple Roles of HIV-1 Capsid during the Virus Replication Cycle. *Virol. Sin.* **34**, 119–134 (2019).

20. Kielian, M. Mechanisms of Virus Membrane Fusion Proteins. *Annu. Rev. Virol.* **1**, 171–189 (2014).

21. Harrison, S. C. Viral membrane fusion. *Virology* **479**–**480**, 498–507 (2015).

22. Harrison, S. C. Viral membrane fusion. *Nat. Struct. Mol. Biol.* **15**, 690–698 (2008).

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

24. Allan, J. S. *et al.* Major Glycoprotein Antigens That Induce Antibodies in AIDS Patients Are Encoded by HTLV-III. *Science (80-.).* **228**, 1091–1094 (1985).

25. Dalgleish, A. G. *et al.* The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763–767 (1984).

26. Luby-Phelps, K. Cytoarchitecture and Physical Properties of Cytoplasm: Volume, Viscosity, Diffusion, Intracellular Surface Area. in 189–221 (1999). doi:10.1016/S0074-7696(08)60527-6

27. Walsh, D. & Naghavi, M. H. Exploitation of Cytoskeletal Networks during Early Viral Infection. *Trends Microbiol.* **27**, 39–50 (2019).

28. Naghavi, M. H. HIV-1 capsid exploitation of the host microtubule cytoskeleton during early infection. *Retrovirology* **18**, 19 (2021).

29. Malikov, V. *et al.* HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. *Nat. Commun.* **6**, 1–13 (2015).

30. Huang, P.-T. *et al.* FEZ1 Is Recruited to a Conserved Cofactor Site on Capsid to Promote HIV-1 Trafficking. *Cell Rep.* **28**, 2373-2385.e7 (2019).

31. Zhou, H. *et al.* Genome-Scale RNAi Screen for Host Factors Required for HIV Replication. *Cell Host Microbe* **4**, 495–504 (2008).

32. Carter, A. P., Diamant, A. G. & Urnavicius, L. How dynein and dynactin transport cargos: a structural perspective. *Curr. Opin. Struct. Biol.* **37**, 62–70 (2016).

33. Matanis, T. *et al.* Bicaudal-D regulates COPI-independent Golgi–ER transport by recruiting the dynein–dynactin motor complex. *Nat. Cell Biol.* **4**, 986–992 (2002).

34. Dharan, A. *et al.* Bicaudal D2 facilitates the cytoplasmic trafficking and nuclear import of HIV-1 genomes during infection. *Proc. Natl. Acad. Sci.* 201712033 (2017). doi:10.1073/pnas.1712033114

35. Carnes, S. K., Zhou, J. & Aiken, C. HIV-1 Engages a Dynein-Dynactin-BICD2 Complex for Infection and Transport to the Nucleus. *J. Virol.* JVI.00358-18 (2018). doi:10.1128/JVI.00358-18

36. Gaudin, R., Alencar, B. C. de, Arhel, N. & Benaroch, P. HIV trafficking in host cells: motors wanted! *Trends Cell Biol.* **23**, 652–662 (2013).

37. Bukrinskaya, A., Brichacek, B., Mann, A. & Stevenson, M. Establishment of a Functional Human Immunodeficiency Virus Type 1 (HIV-1) Reverse Transcription Complex Involves the Cytoskeleton. *J. Exp. Med.* **188**, 2113–2125 (1998).

38. McDonald, D. *et al.* Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* **159**, 441–452 (2002).

39. Turville, S. All-Round Manipulation of the Actin Cytoskeleton by HIV. *Viruses* **10**, 63 (2018).

40. Iyengar, S., Hildreth, J. E. K. & Schwartz, D. H. Actin-Dependent Receptor Colocalization Required for Human Immunodeficiency Virus Entry into Host Cells. *J. Virol.* **72**, 5251–5255 (1998).

41. Li, W. *et al.* HIV-1 Uses Dynamic Podosomes for Entry into Macrophages. *J. Virol.* **95**, (2021).

42. Welte, M. A. Bidirectional Transport along Microtubules. *Curr. Biol.* **14**, R525–R537 (2004).

43. Pietrantoni, G., Ibarra-Karmy, R. & Arriagada, G. Microtubule Retrograde Motors and Their Role in Retroviral Transport. *Viruses* **12**, 483 (2020).

44. Scoca, V. & Di Nunzio, F. The HIV-1 Capsid: From Structural Component to Key Factor for Host Nuclear Invasion. *Viruses* **13**, 273 (2021).

45. Ingram, Z., Fischer, D. K. & Ambrose, Z. Disassembling the Nature of Capsid: Biochemical, Genetic, and Imaging Approaches to Assess HIV-1 Capsid Functions. *Viruses* **13**, 2237 (2021).

46. Guedán, A., Caroe, E. R., Barr, G. C. R. & Bishop, K. N. The Role of Capsid in HIV-1 Nuclear Entry. 1–14 (2021).

47. Yang, Q., Rout, M. P. & Akey, C. W. Three-Dimensional Architecture of the Isolated Yeast Nuclear Pore Complex: Functional and Evolutionary Implications. *Mol. Cell* **1**, 223–234 (1998).

48. Hampoelz, B., Andres-Pons, A., Kastritis, P. & Beck, M. Structure and Assembly of the Nuclear Pore Complex. *Annu. Rev. Biophys.* **48**, 515–536 (2019).

49. Knockenhauer, K. E. & Schwartz, T. U. The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell* **164**, 1162–1171 (2016).

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

51. Görlich, D. & Kutay, U. Transport Between the Cell Nucleus and the Cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–660 (1999).

52. Schaller, T. *et al.* HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathog.* **7**, e1002439 (2011).

53. Di Nunzio, F. *et al.* Human Nucleoporins Promote HIV-1 Docking at the Nuclear Pore, Nuclear Import and Integration. *PLoS One* **7**, e46037 (2012).

54. Fernandez, J. *et al.* Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. *Nat. Microbiol.* **4**, 1840–1850 (2019).

55. Diaz-Griffero, F. The Role of TNPO3 in HIV-1 Replication. *Mol. Biol. Int.* **2012**, 1–6 (2012).

56. Strambio-De-Castillia, C., Niepel, M. & Rout, M. P. The nuclear pore complex: bridging nuclear transport and gene regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 490–501 (2010).

57. Rebensburg, S. V. *et al.* Sec24C is an HIV-1 host dependency factor crucial for virus replication. *Nat. Microbiol.* **6**, 435–444 (2021).

58. Francis, A. C., Marin, M., Prellberg, M. J., Palermino-Rowland, K. & Melikyan, G. B. Hiv-1 uncoating and nuclear import precede the completion of reverse transcription in cell lines and in primary macrophages. *Viruses* **12**, 1–22 (2020).

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

Microbiol. **5**, 1088–1095 (2020).

60. Burdick, R. C. *et al.* HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 5486–5493 (2020).

61. Selyutina, A., Persaud, M., Lee, K., KewalRamani, V. & Diaz-Griffero, F. Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating. *Cell Rep.* **32**, 108201 (2020).

62. Peng, K. *et al.* Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *Elife* **3**, e04114 (2014).

63. Bejarano, D. A. *et al.* HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *Elife* **8**, 1–31 (2019).

64. Müller, T. G. *et al.* Hiv-1 uncoating by release of viral cdna from capsid-like structures in the nucleus of infected cells. *Elife* **10**, 1–3 (2021).

65. Germier, T. *et al.* Real-Time Imaging of a Single Gene Reveals Transcription-Initiated Local Confinement. *Biophys. J.* **113**, 1383–1394 (2017).

66. Saad, H. *et al.* DNA Dynamics during Early Double-Strand Break Processing Revealed by Non-Intrusive Imaging of Living Cells. *PLoS Genet.* **10**, e1004187 (2014).

67. Blanco-Rodriguez, G. *et al.* Remodeling of the Core Leads HIV-1 Preintegration Complex into the Nucleus of Human Lymphocytes. *J. Virol.* **94**, (2020).

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

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

70. Miller, M. D., Farnet, C. M. & Bushman, F. D. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J. Virol.* **71**, 5382–5390 (1997).

71. Craigie, R. & Bushman, F. D. HIV DNA Integration. *Cold Spring Harb. Perspect. Med.* **2**, a006890–a006890 (2012).

72. Bedwell, G. J. & Engelman, A. N. Factors that mold the nuclear landscape of HIV-1 integration. *Nucleic Acids Res.* **49**, 621–635 (2021).

73. Lucic, B. *et al.* Spatially clustered loci with multiple enhancers are frequent targets of HIV-1 integration. *Nat. Commun.* **10**, 4059 (2019).

74. Schröder, A. R. W. *et al.* HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. *Cell* **110**, 521–529 (2002).

75. Singh, P. K. *et al.* LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes Dev.* **29**, 2287–2297 (2015).

76. Achuthan, V. *et al.* Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral DNA Integration. *Cell Host Microbe* **24**, 392-404.e8 (2018).

77. Sowd, G. A. *et al.* A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1054–E1063 (2016).

78. Briand, N. & Collas, P. Lamina-associated domains: peripheral matters and internal affairs. *Genome Biol.* **21**, 85 (2020).

79. Ciuffi, A. *et al.* A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**, 1287–1289 (2005).

80. Cherepanov, P. *et al.* HIV-1 Integrase Forms Stable Tetramers and Associates with LEDGF/p75 Protein in Human Cells. *J. Biol. Chem.* **278**, 372–381 (2003).

81. Vranckx, L. S. *et al.* LEDGIN-mediated Inhibition of Integrase–LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV. *EBioMedicine* **8**, 248–264 (2016).

82. Whyte, W. A. *et al.* Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **153**, 307–319 (2013).

83. Lovén, J. *et al.* Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. *Cell* **153**, 320–334 (2013).

84. Hnisz, D. *et al.* Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **155**, 934–947 (2013).

85. Francis, A. C. *et al.* HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat. Commun.* **11**, (2020).

86. Li, W. *et al.* CPSF6-Dependent Targeting of Speckle-Associated Domains Distinguishes Primate from Nonprimate Lentiviral Integration. *MBio* **11**, (2020).

87. Chen, Y. & Belmont, A. S. Genome organization around nuclear speckles. *Curr. Opin. Genet. Dev.* **55**, 91–99 (2019).

88. Ishov, A. M., Gurumurthy, A. & Bungert, J. Coordination of transcription, processing, and export of highly expressed RNAs by distinct biomolecular condensates. *Emerg. Top. Life Sci.* **4**, 281–291 (2020).

89. Rensen, E. *et al.* Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. *EMBO J.* **40**, 1–16 (2021).

90. Lapaillerie, D. *et al.* Modulation of the intrinsic chromatin binding property of HIV-1 integrase by LEDGF/p75. *Nucleic Acids Res.* **49**, 11241–11256 (2021).

91. Li, L. & Wang, Y. Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair. *J. Biol. Chem.* **292**, 11951–11959 (2017).

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

93. Toccafondi, E., Lener, D. & Negroni, M. HIV-1 Capsid Core: A Bullet to the Heart of the Target Cell. *Front. Microbiol.* **12**, 1–17 (2021).

94. Zhang, M. J., Stear, J. H., Jacques, D. A. & Böcking, T. Insights into HIV uncoating from single-particle imaging techniques. *Biophys. Rev.* (2022). doi:10.1007/s12551-021-00922-6

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

96. Bukrinsky, M. I. *et al.* Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6125–6129 (1993).

97. Zennou, V. *et al.* HIV-1 Genome Nuclear Import Is Mediated by a Central DNA Flap. *Cell* **101**, 173–185 (2000).

98. Yang, Y., Luban, J. & Diaz-Griffero, F. The fate of HIV-1 capsid: A biochemical assay for HIV-1 uncoating. *Methods Mol. Biol.* **1087**, 29–36 (2014).

99. Stremlau, M. *et al.* Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5α restriction factor. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5514–5519 (2006).

100. Maréchal, V., Clavel, F., Heard, J. M. & Schwartz, O. Cytosolic Gag p24 as an Index of Productive Entry of Human Immunodeficiency Virus Type 1. *J. Virol.* **72**, 2208–2212 (1998).

101. Forshey, B. M., von Schwedler, U., Sundquist, W. I. & Aiken, C. Formation of a Human Immunodeficiency Virus Type 1 Core of Optimal Stability Is Crucial for Viral Replication. *J. Virol.* **76**, 5667–5677 (2002).

102. Li, Y., Kar, A. K. & Sodroski, J. Target Cell Type-Dependent Modulation of Human Immunodeficiency Virus Type 1 Capsid Disassembly by Cyclophilin A. *J. Virol.* **83**, 10951–10962 (2009).

103. Hulme, A. E., Perez, O. & Hope, T. J. Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9975–9980 (2011).

104. Perez-Caballero, D., Hatziioannou, T., Zhang, F., Cowan, S. & Bieniasz, P. D. Restriction of Human Immunodeficiency Virus Type 1 by TRIM-CypA Occurs with Rapid Kinetics and Independently of Cytoplasmic Bodies, Ubiquitin, and Proteasome Activity. *J. Virol.* **79**, 15567–15572 (2005).

105. Hübner, W. *et al.* Sequence of Human Immunodeficiency Virus Type 1 (HIV-1) Gag Localization and Oligomerization Monitored with Live Confocal Imaging of a Replication-Competent, Fluorescently Tagged HIV-1. *J. Virol.* **81**, 12596–12607 (2007).

106. Mamede, J. I., Cianci, G. C., Anderson, M. R. & Hope, T. J. Early cytoplasmic uncoating is associated with infectivity of HIV-1. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E7169–E7178 (2017).

107. Xu, H. *et al.* Evidence for biphasic uncoating during HIV-1 infection from a novel imaging assay. *Retrovirology* **10**, 1–16 (2013).

108. Mallery, D. L. *et al.* IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *Elife* **7**, (2018).

109. Mallery, D. L. *et al.* Cellular IP6 Levels Limit HIV Production while Viruses that Cannot Efficiently Package IP6 Are Attenuated for Infection and Replication. *Cell Rep.* **29**, 3983-3996.e4 (2019).

110. Renner, N. *et al.* A lysine ring in HIV capsid pores coordinates IP6 to drive mature capsid assembly. *PLOS Pathog.* **17**, e1009164 (2021).

111. Márquez, C. L. *et al.* Kinetics of HIV-1 capsid uncoating revealed by single-molecule analysis. *Elife* **7**, e34772 (2018).

112. Yamashita, M. & Emerman, M. Capsid Is a Dominant Determinant of Retrovirus Infectivity in Nondividing Cells. *J. Virol.* **78**, 5670–5678 (2004).

113. Yamashita, M., Perez, O., Hope, T. J. & Emerman, M. Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells. *PLoS Pathog.* **3**, 1502–1510 (2007).

114. Rasaiyaah, J. *et al.* HIV-1 evades innate immune recognition through specific cofactor recruitment. *Nature* **503**, 402–405 (2013).

115. Lahaye, X. *et al.* The Capsids of HIV-1 and HIV-2 Determine Immune Detection of the Viral cDNA by the Innate Sensor cGAS in Dendritic Cells. *Immunity* **39**, 1132–1142 (2013).

116. Sumner, R. P. *et al.* Disrupting HIV -1 capsid formation causes cGAS sensing of viral DNA . *EMBO J.* **39**, (2020).

117. Burdick, R. C. *et al. Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes*. *PLoS Pathogens* **13**, (2017).

118. Francis, A. C., Marin, M., Shi, J., Aiken, C. & Melikyan, G. B. Time-Resolved Imaging of Single HIV-1 Uncoating In Vitro and in Living Cells. *PLoS Pathog.* **12**, 1–28 (2016).

119. Francis, A. C. & Melikyan, G. B. 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 (2018).

120. Arhel, N. J. *et al.* HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J.* **26**, 3025–3037 (2007).

121. Zurnic Bönisch, I. *et al.* Capsid-Labelled HIV To Investigate the Role of Capsid during. *J. Virol.* **94**, 1–29 (2020).

122. Lee, K. *et al.* Flexible Use of Nuclear Import Pathways by HIV-1. *Cell Host Microbe* **7**, 221–233 (2010).

123. Engelman, A. N. Hiv capsid and integration targeting. *Viruses* **13**, 1–14 (2021).

124. Di Nunzio, F. *et al.* Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. *Virology* **440**, 8–18 (2013).

125. Matreyek, K. A., Yücel, S. S., Li, X. & Engelman, A. Nucleoporin NUP153 Phenylalanine-Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to Mediate Lentiviral Infectivity. *PLoS Pathog.* **9**, (2013).

126. Price, A. J. *et al.* Host Cofactors and Pharmacologic Ligands Share an Essential Interface in HIV-1 Capsid That Is Lost upon Disassembly. *PLoS Pathog.* **10**, (2014).

127. Briggs, J. A. G., Wilk, T., Welker, R., Kräusslich, H. G. & Fuller, S. D. Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* **22**, 1707–1715 (2003).

128. Von Appen, A. *et al.* In situ structural analysis of the human nuclear pore complex. *Nature* **526**, 140–143 (2015).

129. Stultz, R. D., Cenker, J. J. & McDonald, D. Imaging HIV-1 Genomic DNA from Entry through Productive Infection. *J. Virol.* **91**, (2017).

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

131. Zila, V., Müller, T. G., Laketa, V., Müller, B. & Kräusslich, H.-G. Analysis of CA Content and CPSF6 Dependence of Early HIV-1. **10**, 1–20 (2019).

132. Koh, Y. *et al.* Differential Effects of Human Immunodeficiency Virus Type 1 Capsid and Cellular Factors Nucleoporin 153 and LEDGF/p75 on the Efficiency and Specificity of Viral DNA

Integration. *J. Virol.* **87**, 648–658 (2013).

133. Price, A. J. *et al.* CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication. *PLoS Pathog.* **8**, (2012).

134. Bhattacharya, A. *et al.* Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18625–18630 (2014).

135. Li, C., Burdick, R. C., Nagashima, K., Hu, W.-S. & Pathak, V. K. HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc. Natl. Acad. Sci.* **118**, e2019467118 (2021).

136. Zila, V. *et al.* Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell* **184**, 1032-1046.e18 (2021).

137. Zimmerli, C. E. *et al.* Nuclear pores dilate and constrict in cellulo. *Science (80-.).* **374**, (2021).

138. Zuliani-Alvarez, L. & Towers, G. J. Identifying a nuclear passport for HIV. *Elife* **8**, (2019).

139. Rihn, S. J. *et al.* Extreme Genetic Fragility of the HIV-1 Capsid. *PLoS Pathog.* **9**, (2013).

140. von Schwedler, U. K., Stray, K. M., Garrus, J. E. & Sundquist, W. I. Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein. *J. Virol.* **77**, 5439–5450 (2003).

141. Mattei, S., Glass, B., Hagen, W. J. H., Kräusslich, H. G. & Briggs, J. A. G. The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science (80-.).* **354**, 1434–1437 (2016).

142. Zhao, G. *et al.* Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643–646 (2013).

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

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

145. Pereira, C. F. *et al.* Labeling of multiple HIV-1 proteins with the biarsenical-tetracysteine system. *PLoS One* **6**, (2011).

146. Campbell, E. M., Perez, O., Anderson, J. L. & Hope, T. J. Visualization of a proteasome-independent intermediate during restriction of HIV-1 by rhesus TRIM5α. *J. Cell Biol.* **180**, 549–561 (2008).

147. Griffin, B. A., Adams, S. R. & Tsien, R. Y. Specific Covalent Labeling of\nRecombinant Protein Molecules\nInside Live Cells. *Science (80-.).* **281**, 269 (1998).

148. Hulme, A. E. & Hope, T. J. The Cyclosporin A Washout Assay to Detect HIV-1 Uncoating in Infected Cells. in 37–46 (2014). doi:10.1007/978-1-62703-670-2_4

149. Zhou, L. *et al.* Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS Pathog.* **7**, (2011).

150. Chin, C. R. *et al.* Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep.* **13**, 1717–1731 (2015).

151. Cosnefroy, O., Murray, P. J. & Bishop, K. N. HIV-1 capsid uncoating initiates after the first strand transfer of reverse transcription. *Retrovirology* **13**, 1–17 (2016).

152. Briggs, J. A. G. G. *et al.* The stoichiometry of Gag protein in HIV-1. *Nat. Struct. Mol. Biol.* **11**, 672–675 (2004).

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

154. Ganser-Pornillos, B. K., Yeager, M. & Sundquist, W. I. The structural biology of HIV assembly. *Curr. Opin. Struct. Biol.* **18**, 203–217 (2008).

155. Carlson, L. A. *et al.* Three-Dimensional Analysis of Budding Sites and Released Virus Suggests a Revised Model for HIV-1 Morphogenesis. *Cell Host Microbe* **4**, 592–599 (2008).

156. Lanman, J. *et al.* Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nat. Struct. Mol. Biol.* **11**, 676–677 (2004).

157. James, L. C. & Jacques, D. A. The Human Immunodeficiency Virus Capsid Is More Than Just a Genome Package. *Annu. Rev. Virol.* **5**, 209–225 (2018).

158. Li, Y.-L. *et al.* Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. *Elife* **5**, (2016).

159. Smaga, S. S. *et al.* MxB Restricts HIV-1 by Targeting the Tri-hexamer Interface of the Viral Capsid. *Structure* **27**, 1234-1245.e5 (2019).

160. Pertel, T. *et al.* TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* **472**, 361–365 (2011).

161. Sayah, D. M., Sokolskaja, E., Berthoux, L. & Luban, J. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* **430**, 569–573 (2004).

162. Gres, A. T. *et al.* X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science (80-.).* **349**, 99–103 (2015).

163. Dick, R. A. *et al.* Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509–512 (2018).

164. Yant, S. R. *et al.* A highly potent long-acting small-molecule HIV-1 capsid inhibitor with efficacy in a humanized mouse model. *Nat. Med.* **25**, 1377–1384 (2019).

165. Jacques, D. A. *et al.* HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349–353 (2016).

166. Mallery, D. L. *et al.* A stable immature lattice packages IP 6 for HIV capsid maturation. *Sci. Adv.* **7**, (2021).

167. Yu, A., Lee, E. M. Y., Jin, J. & Voth, G. A. Atomic-scale characterization of mature HIV-1 capsid stabilization by inositol hexakisphosphate (IP 6). *Sci. Adv.* **6**, (2020).

168. Xu, C. *et al.* Permeability of the HIV-1 capsid to metabolites modulates viral DNA synthesis. *PLOS Biol.* **18**, e3001015 (2020).

169. Gamble, T. R. *et al.* Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **87**, 1285–1294 (1996).

170. Bichel, K. *et al.* HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. *Retrovirology* **10**, 1 (2013).

171. Franke, E. K., Yuan, H. E. H. & Luban, J. Specific incorporation of cyclophilin a into HIV-1 virions. *Nature* **372**, 359–362 (1994).

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

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

174. Braaten, D., Franke, E. K. & Luban, J. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *J. Virol.* **70**, 3551–3560 (1996).

175. Towers, G. J. *et al.* Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. *Nat. Med.* **9**, 1138–1143 (2003).

176. Jimenez-Guardeño, J. M., Apolonia, L., Betancor, G. & Malim, M. H. Immunoproteasome activation enables human TRIM5α restriction of HIV-1. *Nat. Microbiol.* **4**, 933–940 (2019).

177. Ni, T. *et al.* Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. *Nat. Struct. Mol. Biol.* **27**, 855–862 (2020).

178. Liu, C. *et al.* Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nat. Commun.* **7**, 10714 (2016).

179. Lau, D. *et al.* Self-Assembly of Fluorescent HIV Capsid Spheres for Detection of Capsid Binders. *Langmuir* **36**, 3624–3632 (2020).

180. Achuthan, V. *et al.* Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral DNA Integration. *Cell Host Microbe* **24**, 392-404.e8 (2018).

181. Temple, J., Tripler, T. N., Shen, Q. & Xiong, Y. A snapshot of HIV-1 capsid–host interactions. *Curr. Res. Struct. Biol.* **2**, 222–228 (2020).

182. Kleinpeter, A. B. & Freed, E. O. HIV-1 Maturation: Lessons Learned from Inhibitors. *Viruses* **12**, 940 (2020).

183. Dick, A. & Cocklin, S. Recent Advances in HIV-1 Gag Inhibitor Design and Development. *Molecules* **25**, 1687 (2020).

184. Lamorte, L. *et al.* Discovery of Novel Small-Molecule HIV-1 Replication Inhibitors That Stabilize Capsid Complexes. *Antimicrob. Agents Chemother.* **57**, 4622–4631 (2013).

185. Fricke, T., Buffone, C., Opp, S., Valle-Casuso, J. & Diaz-Griffero, F. BI-2 destabilizes HIV-1 cores during infection and Prevents Binding of CPSF6 to the HIV-1 Capsid. *Retrovirology* **11**, 120 (2014).

186. Kelly, B. N. *et al.* Structure of the Antiviral Assembly Inhibitor CAP-1 Complex with the HIV-1 CA Protein. *J. Mol. Biol.* **373**, 355–366 (2007).

187. Blair, W. S. *et al.* HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog.* **6**, (2010).

188. Link, J. O. *et al.* Clinical targeting of HIV capsid protein with a long-acting small molecule. *Nature* **584**, 614–618 (2020).

189. Bester, S. M. *et al.* Structural and mechanistic bases for a potent HIV-1 capsid inhibitor. *Science (80-.).* **370**, 360–364 (2020).

190. McArthur, C., Gallazzi, F., Quinn, T. P. & Singh, K. HIV Capsid Inhibitors Beyond PF74. *Diseases* **7**, 56 (2019).

191. Thenin-Houssier, S. & Valente, S. T. HIV-1 capsid inhibitors as antiretroviral agents HHS Public Access. *Curr HIV Res* **14**, 270–282 (2016).

192. Shi, J., Zhou, J., Shah, V. B., Aiken, C. & Whitby, K. Small-Molecule Inhibition of Human Immunodeficiency Virus Type 1 Infection by Virus Capsid Destabilization. *J. Virol.* **85**, 542–549 (2011).

193. Saito, A. *et al.* Roles of Capsid-Interacting Host Factors in Multimodal Inhibition of HIV-1 by PF74. *J. Virol.* **90**, 5808–5823 (2016).

194. Dostálková, A. *et al.* PF74 and Its Novel Derivatives Stabilize Hexameric Lattice of HIV-1 Mature-Like Particles. *Molecules* **25**, 1895 (2020).

195. Rankovic, S., Ramalho, R., Aiken, C. & Rousso, I. PF74 Reinforces the HIV-1 Capsid To Impair Reverse Transcription-Induced Uncoating. *J. Virol.* **92**, (2018).

196. Mendel, D., Cornish, V. W. & Schultz, P. G. Site-Directed Mutagenesis with an Expanded Genetic Code. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 435–462 (1995).

197. Wang, L. & Schultz, P. G. Expanding the genetic code. *Chem. Commun.* 1–11 (2002). doi:10.1039/b108185n

198. Ambrogelly, A., Palioura, S. & Söll, D. Natural expansion of the genetic code. *Nat. Chem. Biol.* **3**, 29–35 (2007).

199. Davis, L. & Chin, J. W. Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell Biol.* **13**, 168–182 (2012).

200. Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).

201. GOODMAN, H. M., ABELSON, J., LANDY, A., BRENNER, S. & SMITH, J. D. Amber Suppression: a Nucleotide Change in the Anticodon of a Tyrosine Transfer RNA. *Nature* **217**, 1019–1024 (1968).

202. Liu, C. C. & Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **79**, 413–444 (2010).

203. Zhang, M. S. *et al.* Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat. Methods* **14**, 729–736 (2017).

204. Ellefson, J. W. *et al.* Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol.* **32**, 97–101 (2014).

205. Chin, J. W. Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annu. Rev. Biochem.* **83**, 379–408 (2014).

206. Ernst, R. J. *et al.* Genetic code expansion in the mouse brain. *Nat. Chem. Biol.* **12**, 776–778 (2016).

207. Han, S. *et al.* Expanding the genetic code of Mus musculus. *Nat. Commun.* **8**, 14568 (2017).

208. Chin, J. W. *et al.* Addition of p -Azido- l -phenylalanine to the Genetic Code of Escherichia c oli. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).

209. Chin, J. W. & Schultz, P. G. In vivo photocrosslinking with unnatural amino acid mutagenesis.

ChemBioChem **3**, 1135–1137 (2002).

210. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci.* **99**, 11020–11024 (2002).

211. Polycarpo, C. *et al.* An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proc. Natl. Acad. Sci.* **101**, 12450–12454 (2004).

212. Srinivasan, G., James, C. M. & Krzycki, J. A. Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA. *Science (80-.).* **296**, 1459–1462 (2002).

213. Atkins, J. F. & Gesteland, R. The 22nd Amino Acid. *Science (80-.).* **296**, 1409–1410 (2002).

214. Borrel, G. *et al.* Unique Characteristics of the Pyrrolysine System in the 7th Order of Methanogens: Implications for the Evolution of a Genetic Code Expansion Cassette. *Archaea* **2014**, 1–11 (2014).

215. Prat, L. *et al.* Carbon source-dependent expansion of the genetic code in bacteria. *Proc. Natl. Acad. Sci.* **109**, 21070–21075 (2012).

216. Polycarpo, C. R. *et al.* Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. *FEBS Lett.* **580**, 6695–6700 (2006).

217. Schmied, W. H., Elsässer, S. J., Uttamapinant, C. & Chin, J. W. Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J. Am. Chem. Soc.* **136**, 15577–15583 (2014).

218. Amiram, M. *et al.* Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol.* **33**, 1272–1279 (2015).

219. Zheng, Y., Lewis, T. L., Igo, P., Polleux, F. & Chatterjee, A. Virus-Enabled Optimization and Delivery of the Genetic Machinery for Efficient Unnatural Amino Acid Mutagenesis in Mammalian Cells and Tissues. *ACS Synth. Biol.* **6**, 13–18 (2017).

220. Bartoschek, M. D. *et al.* Identification of permissive amber suppression sites for efficient non-canonical amino acid incorporation in mammalian cells. *Nucleic Acids Res.* **49**, e62–e62 (2021).

221. Park, H.-S. *et al.* Expanding the Genetic Code of Escherichia coli with Phosphoserine. *Science (80-.).* **333**, 1151–1154 (2011).

222. Fan, C., Ip, K. & Söll, D. Expanding the genetic code of Escherichia coli with phosphotyrosine. *FEBS Lett.* **590**, 3040–3047 (2016).

223. Mukai, T. *et al.* Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Res.* **38**, 8188–8195 (2010).

224. Wu, I.-L. *et al.* Multiple Site-Selective Insertions of Noncanonical Amino Acids into Sequence-Repetitive Polypeptides. *ChemBioChem* **14**, 968–978 (2013).

225. Lajoie, M. J. *et al.* Genomically Recoded Organisms Expand Biological Functions. *Science (80-.).* **342**, 357–360 (2013).

226. Johnson, D. B. F. *et al.* RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.* **7**, 779–786 (2011).

227. Wang, K., Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Evolved orthogonal ribosomes

enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol.* **25**, 770–777 (2007).

228. Reinkemeier, C. D., Girona, G. E. & Lemke, E. A. Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes. *Science (80-.).* **363**, (2019).

229. Nikić, I. *et al.* Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy. *Angew. Chemie - Int. Ed.* **55**, 16172–16176 (2016).

230. Young, D. D. & Schultz, P. G. Playing with the Molecules of Life. *ACS Chem. Biol.* **13**, 854–870 (2018).

231. Lang, K. & Chin, J. W. Bioorthogonal reactions for labeling proteins. *ACS Chem. Biol.* **9**, 16–20 (2014).

232. Wang, Q. *et al.* Bioconjugation by Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J. Am. Chem. Soc.* **125**, 3192–3193 (2003).

233. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie - Int. Ed.* **40**, 2004–2021 (2001).

234. Plass, T., Milles, S., Koehler, C., Schultz, C. & Lemke, E. A. Genetically Encoded Copper-Free Click Chemistry. *Angew. Chemie Int. Ed.* **50**, 3878–3881 (2011).

235. Lang, K. *et al.* Genetic Encoding of Bicyclononynes and trans -Cyclooctenes for Site-Specific Protein Labeling in Vitro and in Live Mammalian Cells via Rapid Fluorogenic Diels–Alder Reactions. *J. Am. Chem. Soc.* **134**, 10317–10320 (2012).

236. Baskin, J. M. *et al.* Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci.* **104**, 16793–16797 (2007).

237. Knall, A.-C. & Slugovc, C. Inverse electron demand Diels–Alder (iEDDA)-initiated conjugation: a (high) potential click chemistry scheme. *Chem. Soc. Rev.* **42**, 5131 (2013).

238. Borrmann, A. *et al.* Genetic Encoding of a Bicyclo[6.1.0]nonyne-Charged Amino Acid Enables Fast Cellular Protein Imaging by Metal-Free Ligation. *ChemBioChem* **13**, 2094–2099 (2012).

239. Plass, T. *et al.* Amino Acids for Diels-Alder Reactions in Living Cells. *Angew. Chemie Int. Ed.* **51**, 4166–4170 (2012).

240. Seitchik, J. L. *et al.* Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific in Vivo Bioorthogonal Ligation with trans -Cyclooctenes. *J. Am. Chem. Soc.* **134**, 2898–2901 (2012).

241. Lang, K. *et al.* Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* **4**, 298–304 (2012).

242. Nikić, I., Kang, J. H., Girona, G. E., Aramburu, I. V. & Lemke, E. A. Labeling proteins on live mammalian cells using click chemistry. *Nat. Protoc.* **10**, 780–791 (2015).

243. Kipper, K. *et al.* Application of Noncanonical Amino Acids for Protein Labeling in a Genomically Recoded Escherichia coli. *ACS Synth. Biol.* **6**, 233–255 (2017).

244. Hoffmann, J.-E. *et al.* Highly Stable trans -Cyclooctene Amino Acids for Live-Cell Labeling. *Chem. - A Eur. J.* **21**, 12266–12270 (2015).

245. Peng, T. & Hang, H. C. Site-Specific Bioorthogonal Labeling for Fluorescence Imaging of Intracellular Proteins in Living Cells. *J. Am. Chem. Soc.* **138**, 14423–14433 (2016).

246. Rutkowska, A. *et al.* T-CrAsH: A Heterologous Chemical Crosslinker. *ChemBioChem* **15**, 1765–1768 (2014).

247. Wang, N. *et al.* Construction of a Live-Attenuated HIV-1 Vaccine through Genetic Code Expansion. *Angew. Chemie Int. Ed.* **53**, 4867–4871 (2014).

248. Yuan, Z. *et al.* Controlling Multicycle Replication of Live-Attenuated HIV-1 Using an Unnatural Genetic Switch. *ACS Synth. Biol.* **6**, 721–731 (2017).

249. Si, L. *et al.* Generation of influenza A viruses as live but replication-incompetent virus vaccines. *Science (80-.).* **354**, 1170–1173 (2016).

250. Wang, N. *et al.* Construction of a Live-Attenuated HIV-1 Vaccine through Genetic Code Expansion. *Angew. Chemie Int. Ed.* **53**, 4867–4871 (2014).

251. Chen, Y. *et al.* Controlling the Replication of a Genomically Recoded HIV-1 with a Functional Quadruplet Codon in Mammalian Cells. *ACS Synth. Biol.* **7**, 1612–1617 (2018).

252. Yuan, Z. *et al.* Controlling Multicycle Replication of Live-Attenuated HIV-1 Using an Unnatural Genetic Switch. *ACS Synth. Biol.* **6**, 721–731 (2017).

253. Kelemen, R. E., Erickson, S. B. & Chatterjee, A. Synthesis at the interface of virology and genetic code expansion. *Curr. Opin. Chem. Biol.* **46**, 164–171 (2018).

254. Lin, S. *et al.* Site-specific engineering of chemical functionalities on the surface of live hepatitis D virus. *Angew. Chemie - Int. Ed.* **52**, 13970–13974 (2013).

255. Ablenas, C. J. *et al.* Hepatitis C Virus Helicase Binding Activity Monitored through Site-Specific Labeling Using an Expanded Genetic Code. *ACS Infect. Dis.* **5**, 2118–2126 (2019).

256. Nikič, I. *et al.* Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. *Angew. Chemie - Int. Ed.* **53**, 2245–2249 (2014).

257. Sakin, V. *et al.* A Versatile Tool for Live-Cell Imaging and Super-Resolution Nanoscopy Studies of HIV-1 Env Distribution and Mobility. *Cell Chem. Biol.* **24**, 635-645.e5 (2017).

258. Lu, M. *et al.* Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET. *Nature* **568**, 415–419 (2019).

259. Yao, T. *et al.* Site-Specific PEGylated Adeno-Associated Viruses with Increased Serum Stability and Reduced Immunogenicity. *Molecules* **22**, 1155 (2017).

260. Zhang, C. *et al.* Development of next generation adeno-associated viral vectors capable of selective tropism and efficient gene delivery. *Biomaterials* **80**, 134–145 (2016).

261. Zhang, C., Zhou, X., Yao, T., Tian, Z. & Zhou, D. Precision Fluorescent Labeling of an Adeno-Associated Virus Vector to Monitor the Viral Infection Pathway. *Biotechnol. J.* **13**, 1–11 (2018).

262. Börner, K. *et al.* Pre-arrayed Pan-AAV Peptide Display Libraries for Rapid Single-Round Screening. *Mol. Ther.* **28**, 1016–1032 (2020).

263. Chandran, J. S. *et al.* Site Specific Modification of Adeno-Associated Virus Enables Both Fluorescent Imaging of Viral Particles and Characterization of the Capsid Interactome. *Sci. Rep.* **7**, 14766 (2017).

264. Feiner, R. C. *et al.* rAAV Engineering for Capsid-Protein Enzyme Insertions and Mosaicism Reveals Resilience to Mutational, Structural and Thermal Perturbations. *Int. J. Mol. Sci.* **20**, 5702 (2019).

265. Varadi, K. *et al.* Novel random peptide libraries displayed on AAV serotype 9 for selection of endothelial cell-directed gene transfer vectors. *Gene Ther.* **19**, 800–809 (2012).

266. Schifferdecker, S. *et al.* Direct capsid labeling of infectious HIV-1 by genetic code expansion allows detection of largely complete nuclear capsids and suggests nuclear entry of HIV-1 complexes via common routes. *bioRxiv* 2021.09.14.460218 (2021). doi:10.1101/2021.09.14.460218

267. Cohen, E. A., Dehni, G., Sodroski, J. G. & Haseltine, W. A. Human immunodeficiency virus vpr product is a virion-associated regulatory protein. *J. Virol.* **64**, 3097–3099 (1990).

268. Müller, B., Tessmer, U., Schubert, U. & Kräusslich, H.-G. Human Immunodeficiency Virus Type 1 Vpr Protein Is Incorporated into the Virion in Significantly Smaller Amounts than Gag and Is Phosphorylated in Infected Cells. *J. Virol.* **74**, 9727–9731 (2000).

269. Pizzato, M. *et al.* 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 (2009).

270. Schifferdecker, S. *et al. Minimally invasive chemical functionalization of the HIV-1 capsid protein by genetic code expansion*.

271. Schur, F. K. M. *et al.* Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution. *Nature* **517**, 505–508 (2015).

272. Manocheewa, S., Swain, J. V., Lanxon-Cookson, E., Rolland, M. & Mullins, J. I. Fitness Costs of Mutations at the HIV-1 Capsid Hexamerization Interface. *PLoS One* **8**, 1–10 (2013).

273. Borsetti, A., Öhagen, Å. & Göttlinger, H. G. The C-Terminal Half of the Human Immunodeficiency Virus Type 1 Gag Precursor Is Sufficient for Efficient Particle Assembly. *J. Virol.* **72**, 9313–9317 (1998).

274. Bharata, T. A. M. *et al.* Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8233–8238 (2014).

275. von Schwedler, U. K. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555–1568 (1998).

276. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

277. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

278. Cohen, S. & Arbely, E. Single-Plasmid-Based System for Efficient Noncanonical Amino Acid Mutagenesis in Cultured Mammalian Cells. *ChemBioChem* **17**, 1008–1011 (2016).

279. Rodriguez, E. A., Lester, H. A. & Dougherty, D. A. Improved amber and opal suppressor tRNAs for incorporation of unnatural amino acids in vivo. Part 1: Minimizing misacylation. *Rna* **13**, 1703–1714 (2007).

280. Ma, N. J., Hemez, C. F., Barber, K. W., Rinehart, J. & Isaacs, F. J. Organisms with alternative genetic codes resolve unassigned codons via mistranslation and ribosomal rescue. *Elife* **7**, (2018).

281. Eggertsson, G. & Söll, D. Transfer ribonucleic acid-mediated suppression of termination

codons in Escherichia coli. *Microbiol. Rev.* **52**, 354–374 (1988).

282. O'Donoghue, P. *et al.* Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNA Pyl for genetic code expansion. *FEBS Lett.* **586**, 3931–3937 (2012).

283. Aerni, H. R., Shifman, M. A., Rogulina, S., O'Donoghue, P. & Rinehart, J. Revealing the amino acid composition of proteins within an expanded genetic code. *Nucleic Acids Res.* **43**, e8–e8 (2015).

284. George, S. *et al.* Generation of phospho-ubiquitin variants by orthogonal translation reveals codon skipping. *FEBS Lett.* **590**, 1530–1542 (2016).

285. Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V. & Goff, S. P. Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell* **73**, 1067–1078 (1993).

286. Lukinavičius, G. *et al.* A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat. Chem.* **5**, 132–139 (2013).

287. Koide, Y., Urano, Y., Hanaoka, K., Terai, T. & Nagano, T. Evolution of Group 14 Rhodamines as Platforms for Near-Infrared Fluorescence Probes Utilizing Photoinduced Electron Transfer. *ACS Chem. Biol.* **6**, 600–608 (2011).

288. Blackman, M. L., Royzen, M. & Fox, J. M. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels−Alder Reactivity. *J. Am. Chem. Soc.* **130**, 13518–13519 (2008).

289. Devaraj, N. K., Weissleder, R. & Hilderbrand, S. A. Tetrazine-Based Cycloadditions: Application to Pretargeted Live Cell Imaging. *Bioconjug. Chem.* **19**, 2297–2299 (2008).

290. Nadler, A. & Schultz, C. The Power of Fluorogenic Probes. *Angew. Chemie Int. Ed.* **52**, 2408–2410 (2013).

291. Shieh, P. & Bertozzi, C. R. Design strategies for bioorthogonal smart probes. *Org. Biomol. Chem.* **12**, 9307–9320 (2014).

292. Graham, F. L. & van der Eb, A. J. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456–467 (1973).

293. Dudek, H., Ghosh, A. & Greenberg, M. E. Calcium Phosphate Transfection of DNA into Neurons in Primary Culture. *Curr. Protoc. Neurosci.* **3**, (1998).

294. Kwon, M. & Firestein, B. L. DNA Transfection: Calcium Phosphate Method. in 107–110 (2013). doi:10.1007/978-1-62703-444-9_10

295. Flexner, C. HIV-Protease Inhibitors. *N. Engl. J. Med.* **338**, 1281–1293 (1998).

296. Hanne, J. *et al.* Stimulated Emission Depletion Nanoscopy Reveals Time-Course of Human Immunodeficiency Virus Proteolytic Maturation. *ACS Nano* **10**, 8215–8222 (2016).

297. Wei, X. *et al.* Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* **46**, 1896–1905 (2002).

298. Scoca, V., Louveaux, M., Morin, R. & Ershov, D. HIV-induced membraneless organelles orchestrate post- nuclear entry steps . (2021).

299. Revelo, N. H. *et al.* A new probe for super-resolution imaging of membranes elucidates trafficking pathways. *J. Cell Biol.* **205**, 591–606 (2014).

300. Brandenburg, B. & Zhuang, X. Virus trafficking – learning from single-virus tracking. *Nat. Rev. Microbiol.* **5**, 197–208 (2007).

301. Weigel, P. H. & Oka, J. A. 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 (1981).

302. Melikyan, G. B. *et al.* 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 (2000).

303. Merino, K. M., Allers, C., Didier, E. S. & Kuroda, M. J. Role of Monocyte/Macrophages during HIV/SIV Infection in Adult and Pediatric Acquired Immune Deficiency Syndrome. *Front. Immunol.* **8**, (2017).

304. Pornillos, O. & Ganser-Pornillos, B. K. Maturation of retroviruses. *Curr. Opin. Virol.* **36**, 47–55 (2019).

305. Yoo, S. *et al.* Molecular recognition in the HIV-1 capsid/cyclophilin A complex 1 1Edited by J. A. Wells. *J. Mol. Biol.* **269**, 780–795 (1997).

306. Ptak, R. G. *et al.* Inhibition of Human Immunodeficiency Virus Type 1 Replication in Human Cells by Debio-025, a Novel Cyclophilin Binding Agent. *Antimicrob. Agents Chemother.* **52**, 1302–1317 (2008).

307. Ylinen, L. M. J. *et al.* Cyclophilin A Levels Dictate Infection Efficiency of Human Immunodeficiency Virus Type 1 Capsid Escape Mutants A92E and G94D. *J. Virol.* **83**, 2044–2047 (2009).

308. Veillette, M. *et al.* The V86M mutation in HIV-1 capsid confers resistance to TRIM5α by abrogation of cyclophilin A-dependent restriction and enhancement of viral nuclear import. *Retrovirology* **10**, 25 (2013).

309. Liu, Z. *et al.* The highly polymorphic cyclophilin A-binding loop in HIV-1 capsid modulates viral resistance to MxB. *Retrovirology* **12**, 1 (2015).

310. Setiawan, L. C., van Dort, K. A., Rits, M. A. N. & Kootstra, N. A. Mutations in CypA Binding Region of HIV-1 Capsid Affect Capsid Stability and Viral Replication in Primary Macrophages. *AIDS Res. Hum. Retroviruses* **32**, 390–398 (2016).

311. Novikova, M. *et al.* Identification of a Structural Element in HIV-1 Gag Required for Virus Particle Assembly and Maturation. *MBio* **9**, (2018).

312. López, C. S. *et al.* Determinants of the HIV-1 core assembly pathway. *Virology* **417**, 137–146 (2011).

313. López, C. S., Tsagli, S. M., Sloan, R., Eccles, J. & Barklis, E. Second site reversion of a mutation near the amino terminus of the HIV-1 capsid protein. *Virology* **447**, 95–103 (2013).

314. Ryu, Y. & Schultz, P. G. Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Methods* **3**, 263–265 (2006).

315. Cellitti, S. E. *et al.* In vivo Incorporation of Unnatural Amino Acids to Probe Structure, Dynamics, and Ligand Binding in a Large Protein by Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **130**, 9268–9281 (2008).

316. Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An Enhanced System for Unnatural Amino Acid Mutagenesis in E. coli. *J. Mol. Biol.* **395**, 361–374 (2010).

317. Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H. & Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli. *Biochemistry* **52**, 1828–1837 (2013).

318. Guo, J., Melançon, C. E., Lee, H. S., Groff, D. & Schultz, P. G. Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Proteins Containing Nonnatural Amino Acids. *Angew. Chemie* **121**, 9312–9315 (2009).

319. Johnson, D. B. F. *et al.* Release Factor One Is Nonessential in Escherichia coli. *ACS Chem. Biol.* **7**, 1337–1344 (2012).

320. Bossi, L. & Roth, J. R. The influence of codon context on genetic code translation. *Nature* **286**, 123–127 (1980).

321. Ayer, D. & Yarus, M. The Context Effect Does Not Require a Fourth Base Pair. *Science (80-.).* **231**, 393–395 (1986).

322. Pedersen, W. T. & Curran, J. F. Effects of the nucleotide 3′ to an amber codon on ribosomal selection rates of suppressor tRNA and release factor-1. *J. Mol. Biol.* **219**, 231–241 (1991).

323. Stormo, G. D., Schneider, T. D. & Gold, L. Quantitative analysis of the relationship between nucleotide sequence and functional activity. *Nucleic Acids Res.* **14**, 6661–6679 (1986).

324. Björnsson, A., Mottagui-Tabar, S. & Isaksson, L. A. Structure of the C-terminal end of the nascent peptide influences translation termination. *EMBO J.* **15**, 1696–1704 (1996).

325. Mottagui-Tabar, S., Björnsson, A. & Isaksson, L. A. The second to last amino acid in the nascent peptide as a codon context determinant. *EMBO J.* **13**, 249–257 (1994).

326. Zhang, S., Rydén-Aulin, M. & Isaksson, L. A. Functional Interaction Between Release Factor One and P-site Peptidyl-tRNA on the Ribosome. *J. Mol. Biol.* **261**, 98–107 (1996).

327. Rüegsegger, U., Beyer, K. & Keller, W. Purification and Characterization of Human Cleavage Factor Im Involved in the 3′ End Processing of Messenger RNA Precursors. *J. Biol. Chem.* **271**, 6107–6113 (1996).

328. Gruber, A. R., Martin, G., Keller, W. & Zavolan, M. Cleavage factor I m is a key regulator of 3′ UTR length. *RNA Biol.* **9**, 1405–1412 (2012).

329. Rasheedi, S. *et al.* The Cleavage and polyadenylation specificity factor 6 (CPSF6) subunit of the capsid-recruited pre-messenger RNA cleavage factor I (CFIm) complex mediates HIV-1 integration into genes. *J. Biol. Chem.* **291**, 11809–11819 (2016).

330. Flatt, J. & Greber, U. Misdelivery at the Nuclear Pore Complex—Stopping a Virus Dead in Its Tracks. *Cells* **4**, 277–296 (2015).

331. Paci, G., Caria, J. & Lemke, E. A. Cargo transport through the nuclear pore complex at a glance. *J. Cell Sci.* **134**, (2021).

332. Fay, N. & PantÃ©, N. Nuclear entry of DNA viruses. *Front. Microbiol.* **6**, (2015).

333. Trotman, L. C., Mosberger, N., Fornerod, M., Stidwill, R. P. & Greber, U. F. Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1. *Nat. Cell Biol.* **3**, 1092–1100 (2001).

334. Liashkovich, I., Hafezi, W., Kühn, J. M., Oberleithner, H. & Shahin, V. Nuclear delivery mechanism of herpes simplex virus type 1 genome. *J. Mol. Recognit.* **24**, 414–421 (2011).

335. Ojala, P. M., Sodeik, B., Ebersold, M. W., Kutay, U. & Helenius, A. Herpes Simplex Virus Type 1 Entry into Host Cells: Reconstitution of Capsid Binding and Uncoating at the Nuclear Pore Complex In Vitro. *Mol. Cell. Biol.* **20**, 4922–4931 (2000).

336. Meehan, A. M. *et al.* A Cyclophilin Homology Domain-Independent Role for Nup358 in HIV-1 Infection. *PLoS Pathog.* **10**, 1–17 (2014).

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

338. Terry, L. J., Shows, E. B. & Wente, S. R. Crossing the Nuclear Envelope: Hierarchical Regulation of Nucleocytoplasmic Transport. *Science (80-.).* **318**, 1412–1416 (2007).

339. Buffone, C. *et al.* Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. *J. Virol.* **92**, (2018).

340. Perilla, J. R. & Gronenborn, A. M. Molecular Architecture of the Retroviral Capsid. *Trends Biochem. Sci.* **41**, 410–420 (2016).

341. Nguyen, T. T., Bruinsma, R. F. & Gelbart, W. M. Continuum Theory of Retroviral Capsids. *Phys. Rev. Lett.* **96**, 078102 (2006).

342. Schuller, A. P. *et al.* The cellular environment shapes the nuclear pore complex architecture. *Nature* **598**, 667–671 (2021).

343. Hink, M. A. *et al.* Structural Dynamics of Green Fluorescent Protein Alone and Fused with a Single Chain Fv Protein. *J. Biol. Chem.* **275**, 17556–17560 (2000).

344. Mukherjee, S., Boutant, E., Réal, E., Mély, Y. & Anton, H. Imaging viral infection by fluorescence microscopy: Focus on hiv-1 early stage. *Viruses* **13**, (2021).

345. Balzarotti, F. *et al.* With Minimal Photon Fluxes. *Science (80-.).* **355**, 606–612 (2017).

346. Burdick, R. C. *et al.* Intranuclear Positions of HIV-1 Proviruses Are Dynamic and Do Not Correlate with Transcriptional Activity. *MBio* (2022). doi:10.1128/mbio.03256-21

347. Kane, M. *et al.* Nuclear pore heterogeneity influences HIV-1 infection and the antiviral activity of MX2. *Elife* **7**, 1–44 (2018).

348. Eschbach, J. E. *et al.* Capsid Lattice Destabilization Leads to Premature Loss of the Viral Genome and Integrase Enzyme during HIV-1 Infection. *J. Virol.* **95**, (2020).

349. Yang, Y., Fricke, T. & Diaz-Griffero, F. Inhibition of Reverse Transcriptase Activity Increases Stability of the HIV-1 Core. *J. Virol.* **87**, 683–687 (2013).

350. Shen, Q., Wu, C., Freniere, C., Tripler, T. N. & Xiong, Y. Nuclear Import of HIV-1. *Viruses* **13**, 2242 (2021).

351. AlBurtamani, N., Paul, A. & Fassati, A. The Role of Capsid in the Early Steps of HIV-1 Infection: New Insights into the Core of the Matter. *Viruses* **13**, 1161 (2021).

352. Jennings, J., Shi, J., Varadarajan, J., Jamieson, P. J. & Aiken, C. The Host Cell Metabolite Inositol Hexakisphosphate Promotes Efficient Endogenous HIV-1 Reverse Transcription by Stabilizing the Viral Capsid. *MBio* **11**, (2020).

353. Christensen, D. E., Ganser-Pornillos, B. K., Johnson, J. S., Pornillos, O. & Sundquist, W. I. Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. *Science (80-.).* **370**, (2020).

354. Rouzina, I. & Bruinsma, R. DNA confinement drives uncoating of the HIV Virus. *Eur. Phys. J. Spec. Top.* **223**, 1745–1754 (2014).

355. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C. & Rousso, I. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J. Virol.* **91**, (2017).

356. Rankovic, S., Deshpande, A., Harel, S., Aiken, C. & Rousso, I. HIV-1 Uncoating Occurs via a Series of Rapid Biomechanical Changes in the Core Related to Individual Stages of Reverse Transcription. *J. Virol.* **95**, (2021).

357. Levin, J. G., Guo, J., Rouzina, I. & Musier-Forsyth, K. Nucleic Acid Chaperone Activity of HIV-1 Nucleocapsid Protein: Critical Role in Reverse Transcription and Molecular Mechanism. in 217–286 (2005). doi:10.1016/S0079-6603(05)80006-6

358. Jiang, K. *et al.* Annealing of ssDNA and compaction of dsDNA by the HIV-1 nucleocapsid and Gag proteins visualized using nanofluidic channels. *Q. Rev. Biophys.* **52**, e2 (2019).

359. Shah, V. B. *et al.* The Host Proteins Transportin SR2/TNPO3 and Cyclophilin A Exert Opposing Effects on HIV-1 Uncoating. *J. Virol.* **87**, 422–432 (2013).

360. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780 (1994).

361. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl. Acad. Sci.* **97**, 8206–8210 (2000).

362. Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science (80-.).* **313**, 1642–1645 (2006).

363. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).

364. Qu, K. *et al.* Structure and architecture of immature and mature murine leukemia virus capsids. *Proc. Natl. Acad. Sci.* **115**, E11751–E11760 (2018).

365. Roe, T., Reynolds, T. C., Yu, G. & Brown, P. O. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* **12**, 2099–2108 (1993).

366. Müller, B. *et al.* HIV-1 gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. *J. Biol. Chem.* **284**, 29692–29703 (2009).

367. Müller, B. *et al.* Construction and Characterization of a Fluorescently Labeled Infectious Human Immunodeficiency Virus Type 1 Derivative. *J. Virol.* **78**, 10803–10813 (2004).

368. Bohne, J. & Kräusslich, H. G. Mutation of the major 5′ splice site renders a CMV-driven HIV-1 proviral clone Tat-dependent: Connections between transcription and splicing. *FEBS Lett.* **563**, 113–118 (2004).

369. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci.* **90**, 8392–8396 (1993).

370. Smith, S. D. *et al.* Monoclonal Antibody and Enzymatic Profiles of Human Malignant T-Lymphoid Cells and Derived Cell Lines. *Cancer Res.* **44**, 5657–5660 (1984).

371. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

372. De Chaumont, F. *et al.* Icy: An open bioimage informatics platform for extended reproducible research. *Nat. Methods* **9**, 690–696 (2012).

373. Paul-Gilloteaux, P. *et al.* EC-CLEM: Flexible multidimensional registration software for correlative microscopies. *Nat. Methods* **14**, 102–103 (2017).

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

375. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).

376. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).

377. Strober, W. Trypan blue exclusion test of cell viability. *Curr. Protoc. Immunol.* **Appendix 3**, 2–3 (2001).

378. Kukulski, W. *et al.* Correlated fluorescence and 3D electron microscopy with high sensitivity and spatial precision. *J. Cell Biol.* **192**, 111–119 (2011).

379. Schorb, M. *et al.* New hardware and workflows for semi-automated correlative cryo-fluorescence and cryo-electron microscopy/tomography. *J. Struct. Biol.* **197**, 83–93 (2017).

8 Appendix

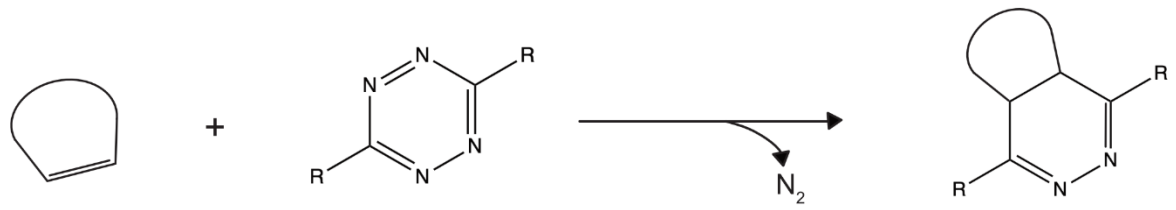


Figure 34. Strain-promoted inverse electron demand Diels-Alder cycloaddition. Strained alkene reacts with a tetrazine *via* strain-promoted inverse demand Diels-Alder [4+2] cycloaddition (SPIEDAC). Structures were created with ChemDraw.

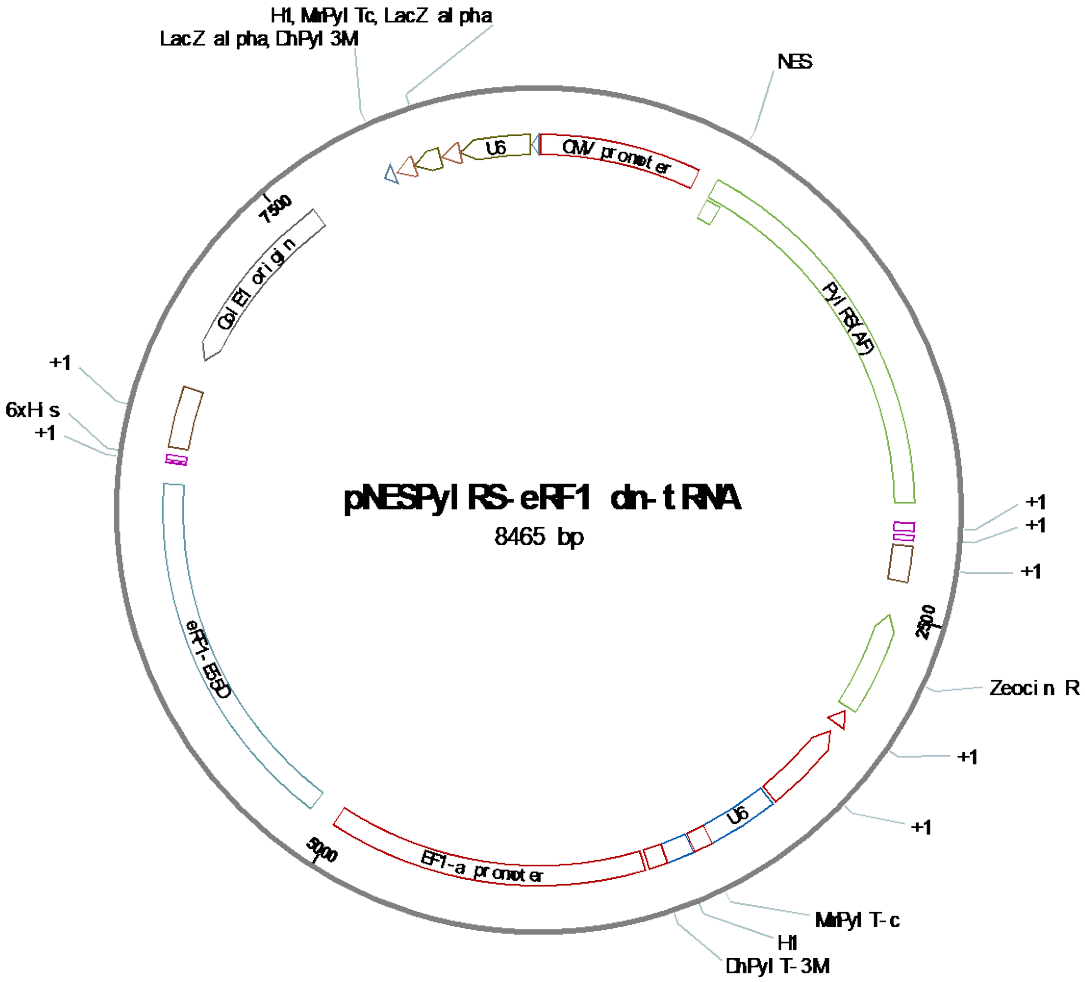


Figure 35. Plasmid map of pNESPylRS-eRF1dn-tRNA plasmid.
Plasmid is based on pEA168 containing two U6-promotor-driven expression modules for tRNA, eRF1 E55D coding sequence under EF1 control and a CMV-driven expression cassette for an optimized pyrrolysine tRNA synthetase.

Table 8 Analysis of amber suppression efficiency of GCE-modified CA variants

Virus variant	amber suppression efficiency [%]
HIV-1*CA A14TAG	not determined
HIV-1*CA I15TAG	not determined
HIV-1*CA W80TAG	8.45
HIV-1*CA L83TAG	24.33
HIV-1*CA H84TAG	4.86
HIV-1*CA V86TAG	25.94
HIV-1*CA H87TAG	3.64
HIV-1*CA A88TAG	6.49

HIV-1*CA I91TAG	12.04
HIV-1*CA A92TAG	15.04
HIV-1*CA M96TAG	6.84
HIV-1*CA E98TAG	16.04
HIV-1*CA T110TAG	17.49
HIV-1*CA L111TAG	2.01
HIV-1*CA G116TAG	16.18
HIV-1*CA T119TAG	18.09
HIV-1*CA H120TAG	13.84
HIV-1*CA I124TAG	16.51

Table 9 Relative infectivity of unlabeled particle variants

Virus variant	Relative Infectivity[a] in % [median + SD] normalized to HIV-1*
HIV-1*CA A14TAG	63.59 ± 82.87
HIV-1*CA I15TAG	-
HIV-1*CA W80TAG	-
HIV-1*CA L83TAG	82.15 ± 44.24
HIV-1*CA H84TAG	-
HIV-1*CA V86TAG	-
HIV-1*CA H87TAG	-
HIV-1*CA A88TAG	-
HIV-1*CA I91TAG	51.05 ± 17.01
HIV-1*CA A92TAG	-
HIV-1*CA M96TAG	-
HIV-1*CA E98TAG	-
HIV-1*CA T110TAG	-
HIV-1*CA L111TAG	41.69 ± 9.44
HIV-1*CA G116TAG	-
HIV-1*CA T119TAG	-
HIV-1*CA H120TAG	31.89 ± 10.75
HIV-1*CA I124TAG	-

[a] determined in Steady Glo assay

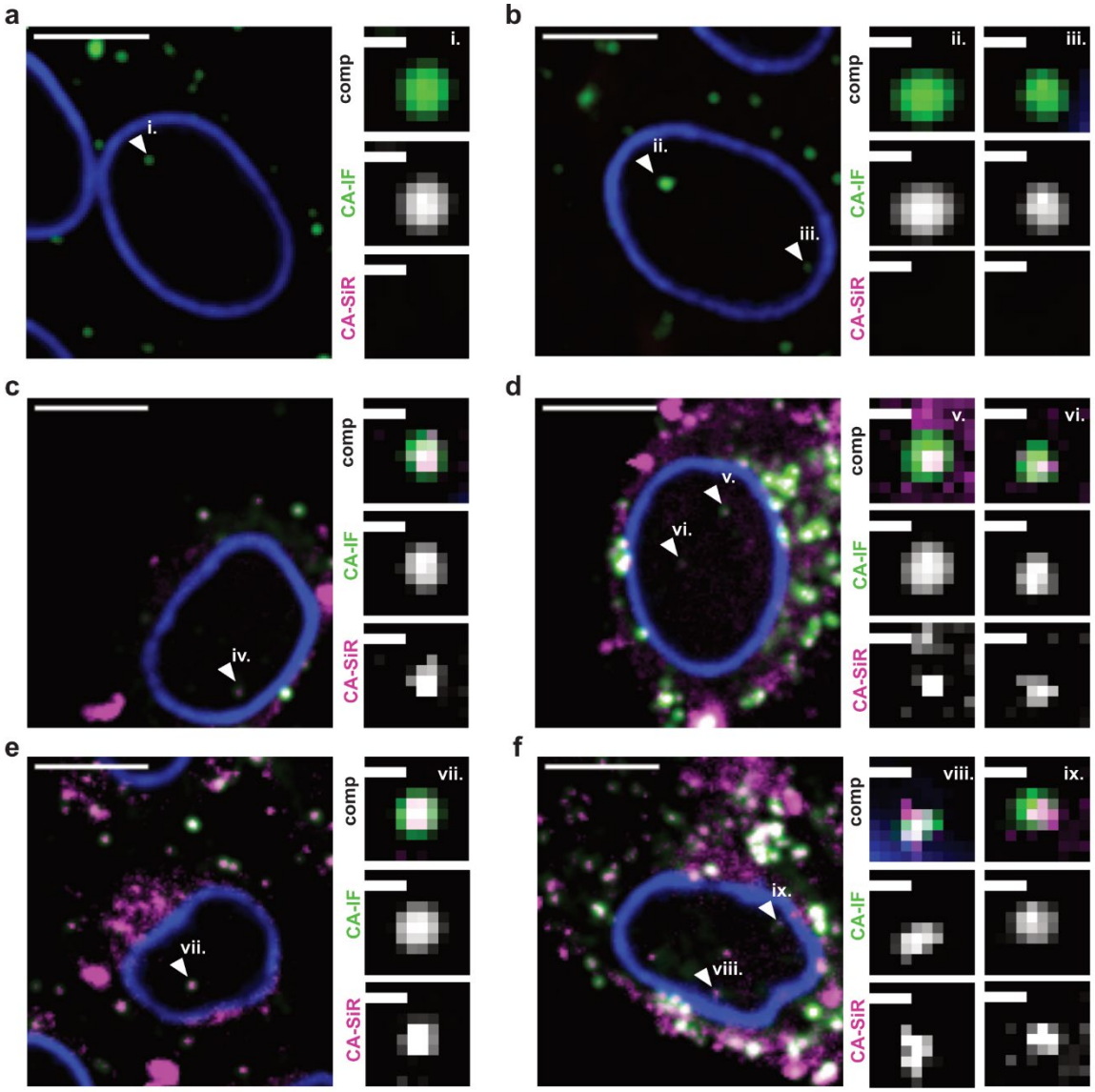


Figure 36. More images of nuclear CA in TZM-bl cells. TZM-bl cells were infected with HIV-1* or HIV-1*CA14^{SiR} (MOI ~ 0.8), treated with 15 µM PF74 at 17 h p.i. for 1 h, fixed, extracted with methanol and permeabilized. Samples were stained by IF using antibodies raised against HIV-1 CA (green) and lamin A/C (blue) and imaged by SDCM. **a,b.** Detection of nuclear CA IF foci (i. - iii.) in cells infected with HIV-1*. Enlargements of indicated spots (white arrowheads) are shown next to each cell. **c-f.** Detection of nuclear CA IF/CA SiR foci (iv. - ix.) in cells infected with HIV-1*CA14^{SiR}. Enlargements of indicated spots (white arrowheads) are shown next to each cell. Scale bars: 10 µm (cell) and 1 µm (enlargements). Figure is published in Schifferdecker *et al.,* 2021[266].

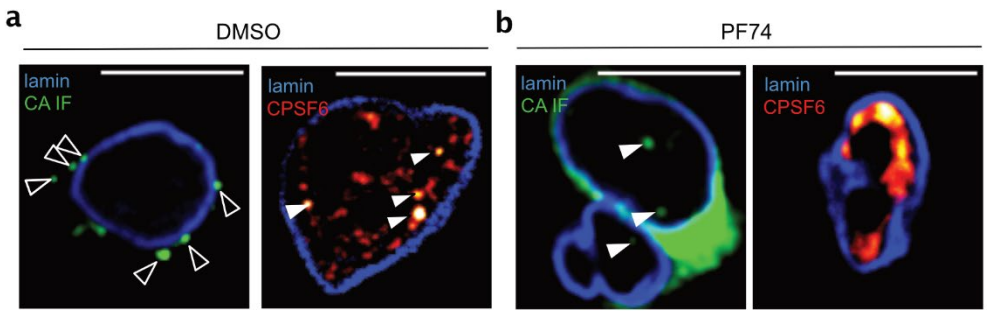


Figure 37. CPSF6 clusters were dissolved upon PF74 treatment. Primary CD4+ T cells were infected with HIV-1*. At 23 h p.i., cells were treated with DMSO (a) or 15 µM PF74 (b) for 1 h, fixed, permeabilized, extracted with methanol and stained against HIV-1 CA (green) or CPSF6 (red) and lamin B (blue) by IF. Scale bars: 10 µm. **a.** Cells treated with DMSO show CA IF positive structures only in the cytoplasm (white bordered arrowhead) and strong nuclear CPSF6 clusters (white filled arrowhead). **b.** Cells treated with 15 µM PF74 show nuclear CA IF signals upon displacement of CPSF6 clusters.