HIV needs the small molecule IP6 to build its capsid

Nadine Renner

Pembroke College

University of Cambridge

MRC Laboratory of Molecular Biology

This dissertation is submitted for the degree of

Doctor of Philosophy

February 2022

Summary

The retrovirus Human Immunodeficiency Virus Type-1 (HIV-1) infects cells of the immune system causing their depletion and resulting in the Acquired Immunodeficiency Syndrome (AIDS). The viral capsid protein (CA) is essential for replication. Upon budding of a new HIV particle, CA is part of the Gag polyprotein which assembles into an immature virus lattice. During maturation, the viral protease cleaves Gag into single components and CA assembles into a mature capsid. The capsid protects the viral RNA genome which is reverse-transcribed into DNA, trafficked to the nucleus, and integrated into the host genome.

Both immature and mature HIV capsid lattices need to maintain a delicate balance of stability to allow efficient assembly without compromising maturation or disassembly. The work described in this thesis shows that coordination of the small molecule inositol hexakisphosphate (IP6), which can bind to both the immature (via residues K158 and K227) and mature (R18 and K25) CA lattice is crucial. The data herein demonstrate the importance of these charged rings for viral replication and infection.

In the mature capsid, R18 and K25 are thought to have a dual role. Both residues are important IP6 binders, and their mutation interferes with mature capsid formation. In addition, they are proposed to recruit dNTPs to facilitate encapsidated DNA synthesis. Here I show that infection and conical capsid formation of K25A can be rescued by compensating mutants without restoring the charged K25 ring. This suggests that K25 is necessary for assembly but not for nucleotide import. In contrast, maintaining the positively charged R18 appears to be indispensable.

In the immature lattice, lack of either of the positively charged rings is strongly destabilizing and results in aberrant assembly. However, simultaneous mutation of both lysine residues (K158 and K227) allows immature lattice formation, indicating that HIV immature assembly can become independent of IP6. Moreover, the immature lattice is more tolerant to changes in inositol phosphate levels compared to the mature lattice. Nevertheless, HIV-1 with a disabled IP6-enrichment mechanism is non-infectious and my data suggest that this is because IP6 is important in the formation of mature cores that support productive infection.

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed the prescribed word limit for the Biology Degree Committee of 60,000 words.

Nadine Renner

Pembroke College

University of Cambridge

February 2022

Acknowledgements

First, I would like to thank my supervisor Leo James for welcoming me to his group and for being so encouraging and supportive. Leo's knowledge, creativity, drive and expertise were crucial for all the exciting and challenging projects I had the pleasure of working on.

I would also like to thank my second supervisor John Briggs and my university supervisor Stephen Graham for their time and valuable suggestions.

During my PhD, I have been lucky to work with many outstanding scientists. In particular, I want to thank Donna Mallery, with whom I worked very closely throughout all HIV projects. Furthermore, I want to thank Alex Kleinpeter, David Jacques, Eric Freed and Till Böcking for the great collaboration resulting in all the exciting experiments and insightful discussions that were invaluable for the work described in this thesis.

I want to thank past and present members of the James group for making my PhD such a special time: Anna Albecka-Moreau, Alex Jonsson, Dean Clift, Guido Papa, Jakub Luptak, Jingwei Zeng, Larissa Labzin, Maria Botterman, Sarah Caddy, Shannon Smyly and Tyler Rhinesmith. Special thanks to Aaron Tan, Jonna Hakulinen and Zunlong Ke and all members of the MRC EM-Facility who helped me with all sorts of EM/ET problems. In addition, I am grateful to Stephen McLaughlin of the Biophysics Facility and all other support services at the MRC-LMB for their assistance during my PhD.

Most importantly, thank you Leo Kiss for your constant emotional and scientific support and contagious excitement. You always helped me with troubleshooting, told me more about ubiquitin than I ever wanted to know, and your support with my TRIM21 projects was crucial. You are the best! Finally, I would like to thank my family and friends for all the support during my PhD.

Table of contents

Abbreviations

AA	Amino acid
Ab	Antibody
AIDS	Acquired Immunodeficiency Syndrome
ALIX	Apoptosis-Linked gene 2 Interaction protein X
AP-1	Activator protein 1
APOBEC3	Apolipoprotein B mRNA-editing catalytic polypeptide-like 3
ART	Antiretroviral therapy
B	B-box
Bp	Base pairs
CA	Capsid protein
CCV	Clathrin-coated vesicle pathway
COX-	Cytochrome C oxidase subunit 4
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
EIAV	Equine infectious anaemia virus
ERT	Endogenous reverse transcription
CPSF6	Cleavage and polyadenylation specific factor 6
Cryo-ET	Cryo-electron Tomography
CTD	C-terminal Domaine
CypA	Cleavage and polyadenylation specific factor 6
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
ER	Endoplasmic reticulum
ESCRT	Host endosomal Sorting Complex Required for Transport
EV	Empty vector
FBS	Foetal bovine serum
Gag	Group antigen specific
HIV	Human immunodeficiency virus
HAART	Highly active retroviral therapy
HRP	Horseradish peroxidase

IB	Immunoblot
Ig	Immunoglobulin
IL	Interleukin
INT	Integrase
IP	Inositol phosphate
IP6	Inositol hexakisphosphate
IPMK	Inositol-polyphosphate multikinase
IPPK	Inositol-pentakisphosphate-2-kinase
IPTG	Isopropyl B-D-1-Thiogalactopyranoside
IPTK	Inositol-tetrakisphosphatase
IRF	Interferon regulatory factor
Kb	Kilobase
KO	Knock-out
LDS	Lithium Dodecylsulfate
LTR	Long Terminal Repeat
MA	Matrix Protein
MES	2-(N-morpholino)-Ethanesulfonic Acid
MHR	Major homology region
MLV	Murine leukemia virus
MOI	Multiplicity of infection
mRNA	messenger RNA
NC	Nucleocapsid protein
NEF	Negative factor
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B-cells
NMR	Nuclear magnetic resonance
NNRTI	Non-nucleoside Inhibitor
NPC	Nuclear pore complex
NRTI	Nucleoside RT Inhibitor
NTD	N-Terminal Domaine
NtRTI	Nucleotide RT inhibitor
NUPs	Nucleoporins

PAGE	Polyacrylamide gel electrophoresis
PBS	Primer binding site
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline with 0.1 % Tween
PCR	Polymerase Chain Reaction
PEI	Polyethylenimine
PIC	Pre-Integration Complex
PIP2	Phosphatidylinositol-(4,5)-bisphosphate
PLC	Phospholipase C
PM	Plasma Membrane
POL	Polymerase
PPT	Polypurine Tracts
R	RING
REV	Regulator of Virion
RING	Really Interesting New Gene
RSV	Rous sarcoma virus
RRE	Rev Response Element
RT	Reverse Transcription
SD	Standard Deviation
SEM	Standard Error of the Mean
SIV	Simian Immunodeficiency Virus
SOC	Super Optimal broth with Catabolite repression
SP	Spacer Peptide
T5	TRIM5
TAR	Trans-activation response
TAT	Transactivator of Transcription
TBE	Tris-borate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus
Tsg101	Tumour Susceptibility Gene 101
TRIM	Tripartite-motif containing Protein
TY	Tryptone Yeast

TYE	Trypticase Yeast Extract
UTR	Untranslated region
VIF	Viral infectivity factor
VLP	Virus-like particle
VPR	Viral protein R
VPU	Viral protein U
WT	Wild type

Chapter 1 General Introduction

1.1 HIV and AIDS

The Acquired Immune Deficiency Syndrome (AIDS) was first described in 1981. It was characterised by people dying of infections that should have been cleared by the immune system[1]. Two years later, the human immunodeficiency virus 1 (HIV-1) was identified to cause AIDS[2,3]. Since then, HIV-1 has developed into one of the most serious global health challenges, and an estimate of 34.6 million people have died since the pandemic started[4]. Major progress has been made in fighting the disease. In 1996, the highly active antiretroviral therapy (HAART or cART) was introduced. HAART is a combination of three antiretrovirals that suppress viral replication[5,6]. Thenceforth, HIV-1 is no longer a death sentence, and a normal life expectancy can be reached if treated swiftly[7]. In 2020, 73% of the 37.6 million people living with HIV-1 received antiretroviral therapy (ART). Nonetheless, this still led to 690 000 HIV-related deaths. Moreover, 1.5 million people were newly infected with HIV[4]. Despite many efforts, an HIV-1 vaccine is still not available. Thus, further HIV research is essential to develop new drugs and vaccines.

1.2 HIV genus, species and origin

HIV-1 belongs to the *Retroviridae* family. All retroviruses have a single-stranded viral genomic RNA (ssRNA, vRNA) that is reverse transcribed into a double-stranded viral DNA (dsDNA or vDNA) intermediate. This vDNA is integrated into the host cell genome. Retroviruses are divided into seven groups: the alpha-, beta-, gamma-, delta-, epsilonretrovirus, lentivirus, and spumavirus (**Figure 1**). HIV belongs to the lentiviruses, which are characterised by a long latency period and their ability to replicate in non-dividing cells[8,9].

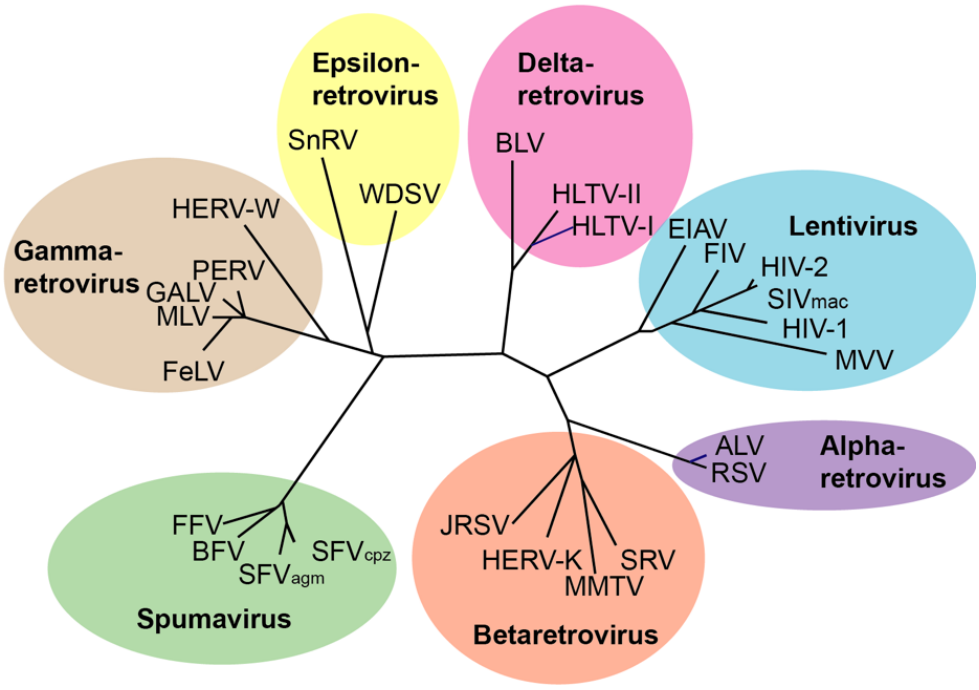


Figure 1: Phylogeny of retroviruses. adapted from[10]

HIV-1 and the later discovered HIV-2 both lead to AIDS, with the latter being less infectious and pathogenic[11-13]. Transmissions of the simian immunodeficiency virus (SIV) from African primates to humans through primate blood or other fluids are thought to be the origin of the viruses[14]. Multiple transmission events seemed to have occurred and gave rise to various HIV subtypes. HIV-1 group M (main) is responsible for the HIV-1 pandemic and was most likely caused by a transmission event of chimpanzee SIV (SIV_{cpz}) to humans[14,15]. Three other separate transmissions from chimpanzees or gorillas resulted in the subgroups N, O, and P. These subgroups appear to be restricted to West Africa[14]. Likewise, the subtypes (A-H) of HIV-2 originate from transmissions from SIV_{smm} from the sooty mangabey monkeys[14].

1.3 HIV and AIDS

The most common route of HIV infection is by sexual transmission, however, HIV is also transmitted through transfusion, needle sharing and maternal-infant exposure[16].

HIV primarily infects activated $CD4^+$ T-cells in mucosal tissues. HIV can also infect myeloid cells e.g. macrophages and monocytes (including dendritic cells)[17]. Initially, the virus replicates at low levels in T-cells followed by a rapid replication and a

corresponding burst of viremia. This allows the virus to spread to lymphoid organs and other tissues[16,18]. Replication usually peaks after 3-5 weeks and is accompanied by a drastic loss of $CD4^+$ T-cells and a short period called the acute HIV syndrome. This phase can be symptomatic, and some people experience flu-like symptoms. Around six weeks after infection, a neutralising antibody response is detected, and $CD8^+$ cytotoxic T-lymphocytes are developing. Subsequently, the immune system gains control over viral replication that remains stable at the so-called viral set-point. The viral load at the set-point varies between individuals and determines the speed of disease progression[16,19-21].

At this stage of infection, the virus establishes a persistent viral reservoir even if the individual is treated with cART. Even though cART can nearly completely inhibit replication, it cannot clear HIV genomes integrated into the host cell genome. Therefore, the virus can replicate as long as the infected cells persist. Especially latently infected resting or memory $CD4^+$ T-cells are maintained indefinitely and are thus a major problem for cART[22,23]. Once cART is stopped, the viral load usually rebounds in a few weeks which is why HIV positive people require cART for the rest of their lifes[24]. Importantly, patients diagnosed early and adhering to cART treatment for the rest of their lives will not develop AIDS and can live a normal life[4,7].

In untreated individuals, the virus can usually only be controlled for a limited amount of time because mutants that can evade the immune system arise quickly[25-27]. Increased activation of the immune system, accompanied by an increased turnover of $CD4^+$ cells and other immune cells such as $CD8^+$ T-cells, B-cells, NK cells, can initially counteract the viral escape variant and the corresponding progressive loss of $CD4^+$-cells. However, after several years at later stages of chronic immune activation, the immune system progressively loses its ability to regenerate immune cells to control viral replication[16,28]. With immune system exhaustion the final stage of the disease, AIDS, commences, and opportunistic infections can no longer be controlled[16]. Most people survive less than two years without treatment[29].

1.4 HIV structure and life cycle overview

1.4.1 HIV genome organisation

The HIV genome has three open reading frames, gag (group-specific antigen), pol (polymerase) and env (envelope). HIV Gag consists of the matrix protein (MA), capsid protein (CA), spacer peptide 1 (SP1), nucleocapsid (NC), spacer peptide 2 (SP2), and the C-terminal peptide p6. During translation of gag in around 5-10% of cases, a frameshift leads to translation of the Gag-Pol fusion protein, which includes the protease (PR), integrase (IN) and reverse transcriptase (RT)[30] (**Figure 2**). In addition, the virus encodes six regulatory proteins: negative factor (Nef), regulator of expression of virion proteins (Rev), trans-activator of transcription (tat), viral infectivity factor (Vif), viral protein R (Vpr) and viral protein U (Vpu)[31,32].

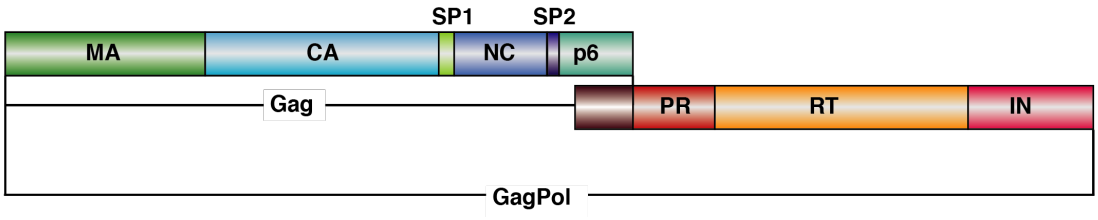


Figure 2: Map of HIV Gag-Pol precursor polyprotein. The polyprotein is processed into matrix (MA), capsid (CA), spacer region 1 (SP1), nucleoprotein (NC), spacer region 2 (SP2), p6, protease (PR), reverse transcriptase (RT) and integrase (IN).

1.4.2 HIV virion

HIV is surrounded by a lipid bilayer that is taken from the plasma membrane of an infected cell during viral budding. The structure of an immature HIV particle is shown in **Figure 3** on the left and of the mature on the right. A newly released immature virion contains around 2500 Gag and 125 Gag-Pol molecules. Gag polymerises, the MA interacts with the membrane, and the NC interacts with two copies of the viral RNA genome. The immature Gag and Gag-Pol polyproteins are cleaved into the single protein components by the protease (PR) to form the infectious mature virus[33]. In the mature virion, the capsid proteins (CA) form a closed conical shell called the capsid (or core). For simplicity, the capsid protein will always be abbreviated as CA, whereas capsid always means the capsid core from here on. The capsid encloses two copies

of the viral RNA genome bound by the viral nucleocapsid (NC), reverse transcriptase (RT) and Integrase (IN)[34,35].

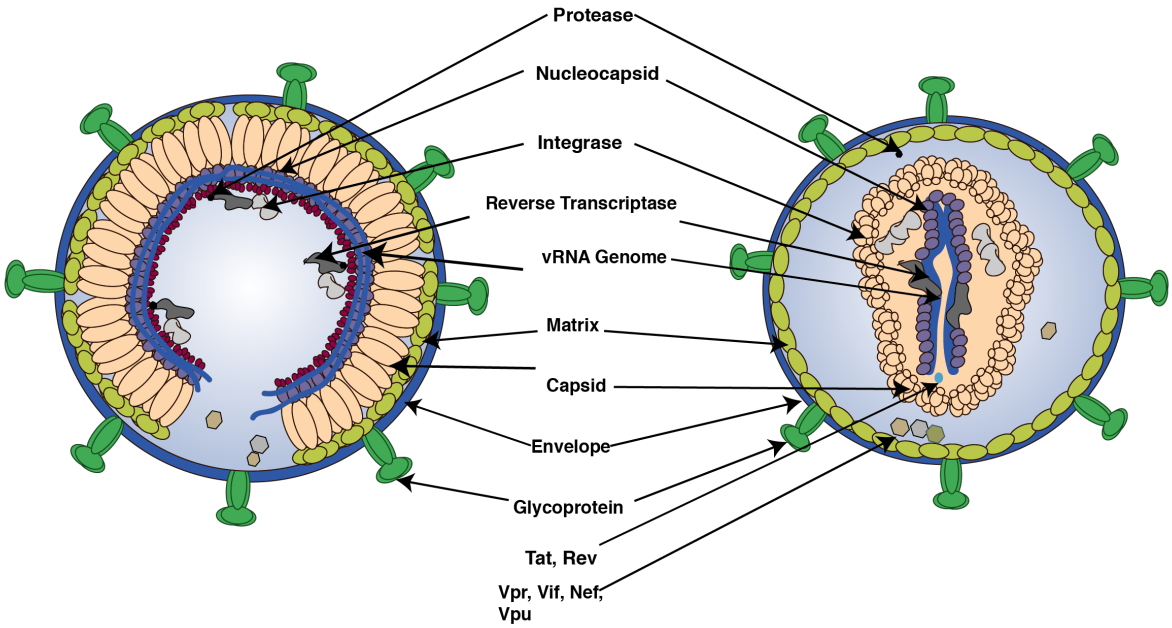


Figure 3: Organisation of immature and mature virus. The virion is enclosed by a lipid bilayer that contains the envelope glycoprotein gp120 and gp41. In the immature virus Gag which includes the structural proteins MA, CA, NC and Gag-Pol (also includes RT, IN, Pol) polymerise. The protease cleaves Gag and Gag-Pol, leading to major rearrangements in the virion. CA forms a conical capsid which encloses the viral RNA which is coated by NC, IN, and Pol.

1.4.3 HIV life cycle overview

During infection, viral envelope (Env) glycoproteins bind to the CD4-receptor and coreceptor CXC chemokine receptor type 4 (CXCR4) or CC chemokine receptor type 5 (CCR5) and binding leads to fusion of the viral and the cellular membrane[36] (**Figure 4**). This results in release of the capsid containing the viral genome, RT and IN into the cytoplasm. The ssRNA genome is reverse transcribed into dsDNA by the RT whilst the capsid is transported to the nucleus using the cellular microtubule network[37]. The timing of capsid disassembly, called uncoating, is still highly debated. Most recent evidence suggests that the capsid remains largely intact until it enters the nucleus and possibly through the nuclear pore itself allowing infection of nondividing cells[38]. Reverse transcription may then be completed in the nucleus prior to uncoating[39,40]. In the nucleus, the vDNA is integrated into actively transcribed regions of the host cell genome by the IN. This integrated provirus either stays latent or is transcribed into mRNAs and translated to produce the viral proteins and full-length RNA. Viral

envelope glycoproteins traffic via the secretory pathway to the plasma membrane, whereas the structural proteins are synthesised in the cytosol. Gag and Gag-Pol recruit the unspliced non-translating vRNA, multimerize, and are targeted to the plasma membrane by a still-undefined pathway[41,42]. The immature virus buds by membrane scission using the host endosomal sorting complex required for transport (ESCRT) machinery. Gag and Gag-Pol are cleaved into the single protein components by the protease (PR) to form the infectious mature virus[33], in which the CA assembles into the conical capsid[43,44].

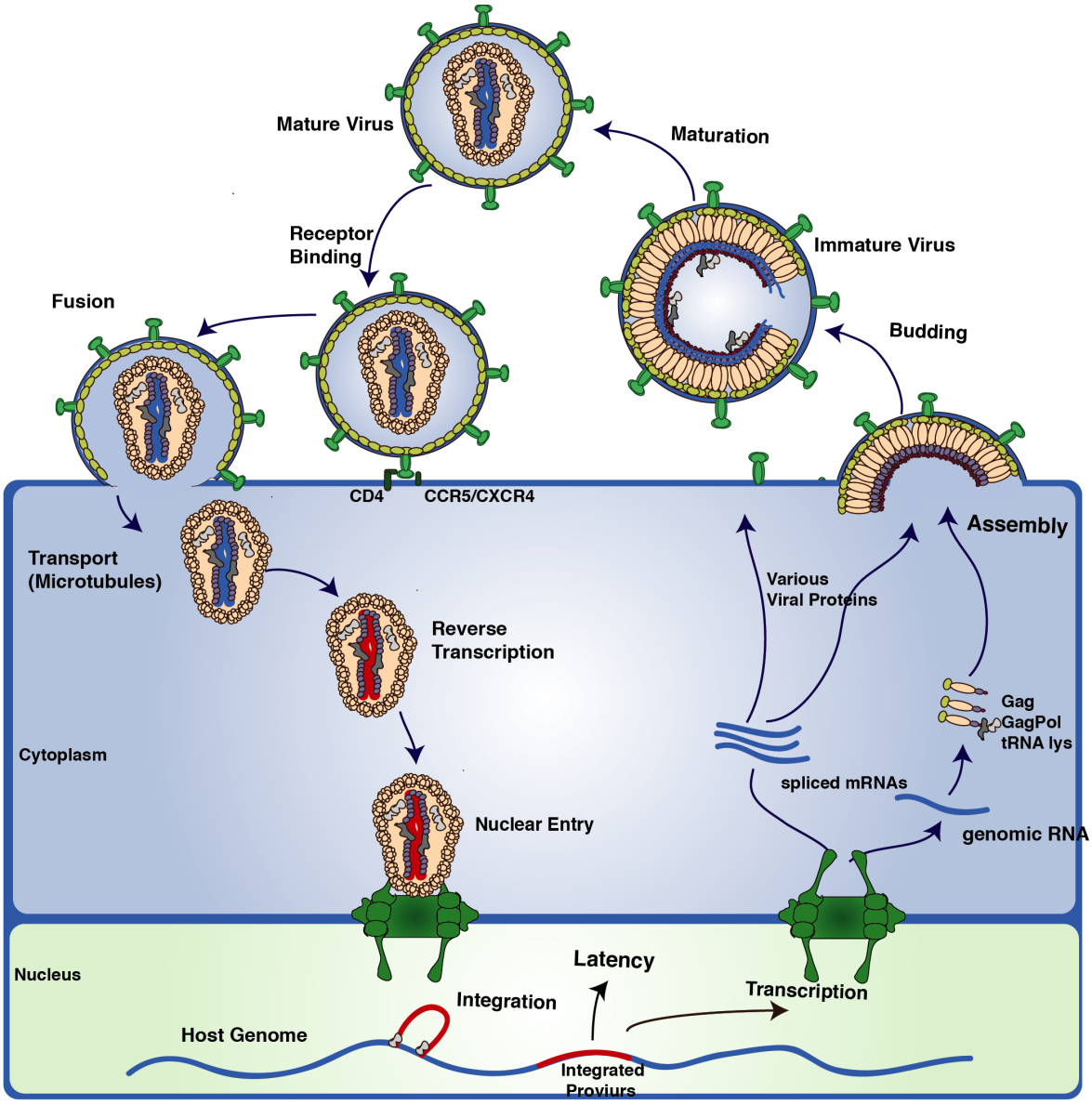


Figure 4: Overview of the HIV infection cycle. The HIV envelope glycoproteins bind to receptors on the cell surface. The viral capsid core is released into the cytoplasm and trafficked towards the nucleus. During this time, the virus initiates reverse transcription. The completely transcribed viral genome is

then inserted into the host cell chromosome. This integrated provirus either stays latent or gives rise to viral proteins necessary for the generation of progeny virions. The virus then buds from the plasma membrane as an immature virus. Maturation leads to the formation of the conical capsid core.

1.5 HIV viral proteins

1.5.1 Matrix (MA)

The MA is a 17 kDa protein that is essential to target Gag to the cell membrane and to allow proper incorporation of envelope glycoproteins. Near its N-terminus, the matrix protein has a highly basic region (HBR) which allows binding to the negatively charged phospholipid PI(4,5)P_2 groups at the inner leaflet of the plasma membrane[45-47]. In addition, the MA carries a myristyl group (14-carbon fatty acid) at its N-terminus, which is a key interactor with the plasma membrane[46,48].

1.5.2 Capsid (CA)

The conical capsid of the mature virus is formed by ~1500 24 kDa CA proteins, which are assembled into ~250 hexamers and exactly 12 pentamers (**Figure 5**). The latter provides the curvature necessary for a closed fullerene capsid core[43,44]. Each CA consists of an N-terminal domain (CA_{NTD}) and a C-terminal domain (CA_{CTD}) that are connected by a flexible linker. CA_{NTD} consist of a β-hairpin, five long and two short α-helices. α-Helices 1, 2, and 3 of CA_{NTD} build an 18-helix inner-hexamer ring through intermolecular NTD-NTD interactions (**Figure 5**). The β-hairpin of CA_{NTD} can isomerise between an open and a closed conformation. An open state gives access to a positively charged pore in the centre of the hexamer[49]. CA_{CTD} is comprised of a 3_{10}-helix and four α-helices. CA_{CTD} forms a belt around the inner hexamer ring and connects the hexamers[44,50,51]. Interactions between the CA_{NTD}-CA_{CTD} domains stabilise the capsid[51,52]. In the mature capsid, seven CA pentamers are located at the wide and five at the narrow end of the core[43]. The NTD of pentamers is organised around the fivefold axis. The pentamer-hexamer CA_{CTD} dimer interface is similar to the interhexamer interface but with more rotations between individual CAs because of higher curvature[53]. The immature CA is described in Chapter 1.6.2.

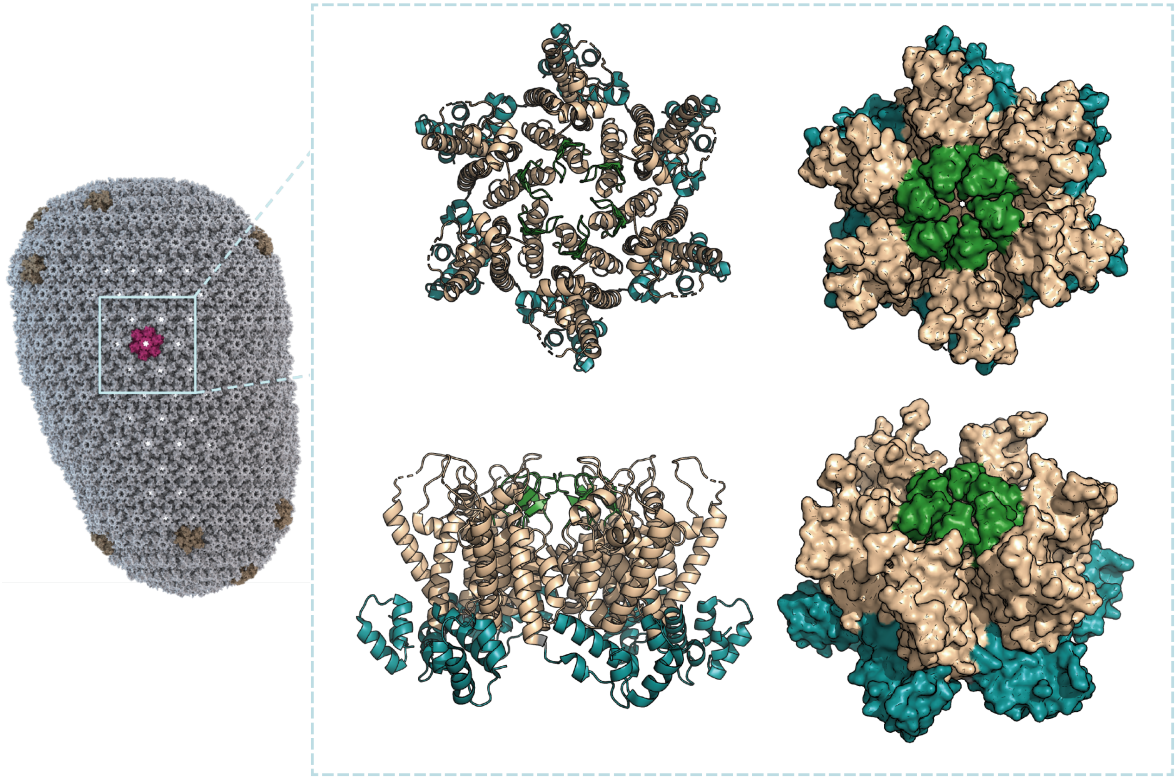


Figure 5: The mature capsid. A model of the assembled capsid is shown on the left (taken from [54]). Top and side views of structures the CA hexamer (PDB: 5HGN[49]) are shown on the right. The CA_{NTD} is coloured in beige, the CA_{CTD} in cyan and the β-hairpin in green.

1.5.3 Spacer region 1 (SP1)

SP1 is a stretch of 14 amino acids (aa) between the C-terminus of the CA and the N-terminus of the NC. SP1 forms a 6-helix bundle with the C-terminal end of CA in the immature Gag-lattice[55]. The helix has to be unwound to allow proteolytic SP1 cleavage from the CA[56,57]. This is exploited by a class of drugs called maturation inhibitors which bind to the 6-helix bundle, thereby stabilising and rearranging it. This impairs proper maturation resulting in less infectious virus[58-60].

1.5.4 Nucleocapsid (NC)

The 6.4 kDa nucleocapsid is required for RNA packaging. The NC has two zinc finger domains connected by a short basic linker. This allows specific binding of the conserved 5' leader of the viral genome (Ψ-sequence)[61-65]. This NC-RNA interaction then serves as a platform for Gag polymerisation and assembly in the immature virus[66]. In the mature virus, the cleaved NC condenses the RNA genome into the ribonucleoprotein complex (RNP) necessary for capsid stability and regulation of

reverse transcription[67-69]. Moreover, NC is a cofactor for the RT during viral DNA synthesis and is important for the completion of reverse transcription and DNA maintenance[70-73]. Furthermore, NC has a role in gRNA recombination and thus increases the rate resistance to antivirals emerges[74,75].

1.5.5 Spacer region 2 (SP2)

The SP2 region does not seem to be necessary for virus maturation and infectivity, but processing and processing kinetics of NC-SP2-p6 appear to be important[76].

1.5.6 p6

p6 is a 52 aa peptide. It regulates the interaction of Gag with the endosomal sorting complex required for transport (ESCRT). p6 has two motifs, the PT/SPA motif that binds to the tumour susceptibility gene 101 (Tsg 101)[77-79] and the YPX(n)L late domain motif that recruits the apoptosis-linked gene 2 interaction protein x (ALIX)[80,81]. Both lead to the recruitment of the ESCRT machinery components, with ALIX seemingly playing a secondary role. However, a consensus model of ESCRT mediated virus abscission remains elusive to date (reviewed in [82]).

1.5.7 Protease (PR)

The HIV protease is an 11 kDa homodimeric aspartic acid protease and a part of Gag-Pol. It cleaves the precursor polyproteins Gag-Pol and Gag into the mature proteins. At the dimer interface, one catalytic aspartic acid residue (Asp25) from each of the protomers form the catalytic site[83,84]. Initial intramolecular PR cleavage within the Gag-Pol precursors occurs upstream of the PR[85]. This is followed by PR auto-processing and results in a stable dimer with improved cleavage efficiency[86,87]. The timing of proteolytic cleavage is critical for the viral life cycle, and variations result in aberrant virus morphologies[88]. The order of proteolytic cleavage of Gag is described in chapter 1.6.3 and **Figure 9**. The protease is a common target in ART since the mid-90s when several HIV-1 protease inhibitors were approved[89].

1.5.8 Reverse transcriptase

The reverse transcriptase is a heterodimer consisting of the subunits p66 and p51. The p66 subunit is catalytically active, and its polymerase domain can carry out RNA-

and DNA-dependent DNA synthesis, whilst its RNaseH domain digests RNA. The p51 unit plays a structural role[90,91].

RT recruits host $tRNA^{Lys3}$ as a primer for the first DNA (-)-strand synthesis, which binds to the primer binding site (PBS) close to the 5' end of genomic vRNA (**Figure 6,** step 1). This initiates reverse transcription until the end of the ssRNA, and an RNA-DNA duplex is formed (**Figure 6,** step 2). The RNA is degraded by RNAse H, the ssDNA repetitive (R) sequence is revealed and can hybridise with a complementary repetitive sequence at the 3' end of one of the ssRNAs, leading to (-) strand-transfer from the 5' to the 3' end (**Figure 6,** step 3). Then DNA elongation continues, and RNaseH continues to degrade RNA. Two purine-rich sequences called polypurine tracts (PPTs) at the central part of the HIV genome are resistant to RNaseH cleavage (**Figure 6,** step 4). The 3' end PPT serves as a primer for (+)-DNA-strand synthesis (**Figure 6,** step 5). During elongation of the (+)-strand, the PBS of the tRNA is also used until the methylA of the tRNA blocks reverse transcription and generates the (+)-strand strong stop. The RNaseH degrades the $tRNA^{Lys3}$ and the PPT, allowing the PBS sequence of the (+)-strand to anneal to the complementary sequence on the 3'end of the (-)-DNA (**Figure 6,** step 6). Finally, both strands are extended, resulting in linear dsDNA with long terminal repeats at both ends (**Figure 6,** step 7). The (+)-strand displaces the DNA that was initiated at the cPPT and a single-strand cFlap is generated that might be removed by host proteins[92-94].

Similar to the PR, the HIV RT is one of the major targets in therapy. The following drug classes are used: Nucleoside RT inhibitors (NRTIs), Nucleotide RT inhibitors (NtRTIS) and non-nucleoside RT inhibitors (NNRTIs). NRTIs and NtRTIs compete with dNTPs and can be incorporated into the newly synthesised DNA strand, where they terminate DNA chain synthesis. In contrast, NNRTIs bind to a hydrophobic pocket close to the active site of the RT, leading to allosteric changes and thus RT inactivation[95].

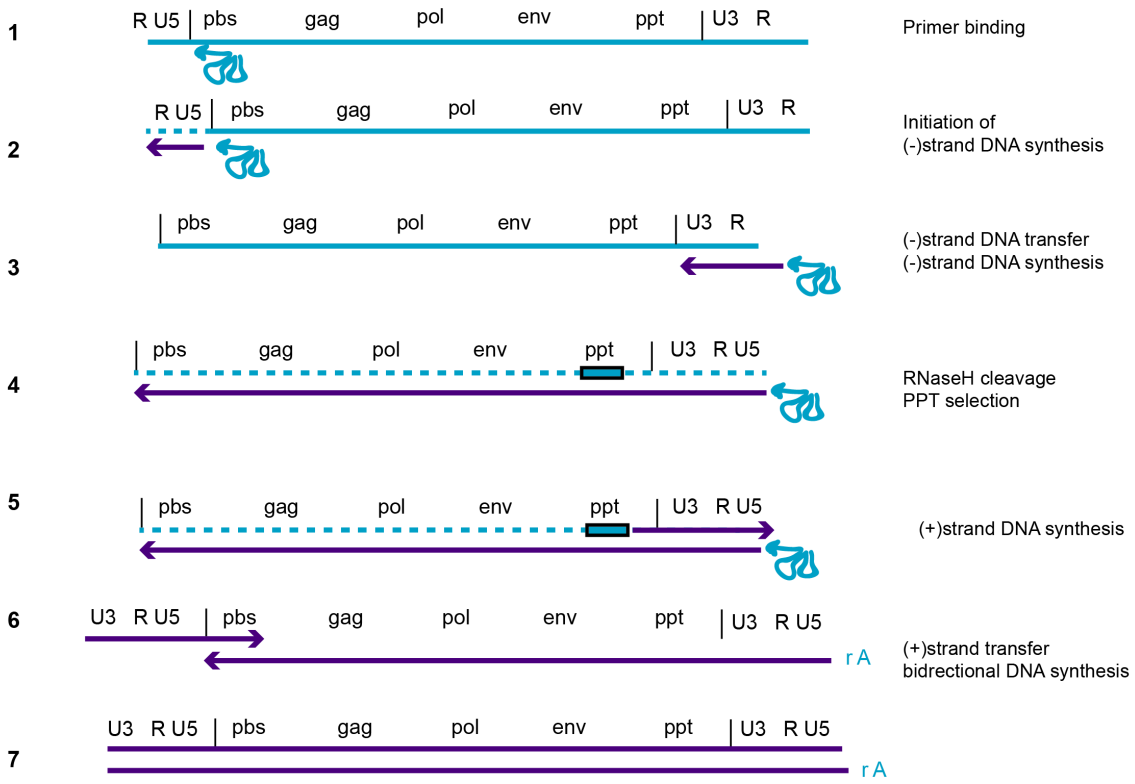


Figure 6: Scheme of reverse transcription. (1) Host $tRNA^{lys3}$ binds to the primer binding site (pbs) near the 5' end of viral RNA where it serves as a primer for the synthesis of the first DNA strand ((-)-strand). (2) RT of the (-)-strand is initiated in 3'-direction, and an RNA-DNA duplex is formed. The RNA is degraded by RNAseH. (3) The (-)-strand is transferred to the repetitive (R) sequence at the 3'-end of the RNA genome, and the (-)-DNA continues to be synthesised. (4) RNA is degraded with exception of a purine-rich sequence (ppt) resistant to RNAse H. (5) The PPTs serve as a primer for (+)-DNA strand synthesis and the RT also copies 18 nts of the tRNA. Then the tRNA and the ppts are removed by RNAseH, and a riboA is left at the 5'-end. (6) The strand is transferred via the pbs (leading to circularisation), and both strands are extended to generate dsDNA (7) with long terminal repeats (U3 – R-U5) on both ends. The (+)-strand displaces the DNA initiated at the cPPT and the cFlap is generated. Figure adapted from [92,93].

1.5.9 Integrase (INT)

The vDNA must be integrated into the host cell genome, and HIV prefers to integrate into gene-rich, transcriptionally active regions[96]. Control of integration targeting is not completely understood but is known to be modulated by several factors by host cell factors and access to chromatin[97]. Retroviral integration is guided by the intasome nucleoprotein complex. This complex consists of multimeric assemblies of integrase at the ends of the viral DNA transcript. The complex initially catalyses processing of the 3' viral DNA ends. Next, it uses the resulting 3' hydroxyl group to cut the

chromosomal DNA and join the viral 3' ends to the chromosomal DNA. Finally, the intasome disassembles, and the 5' viral DNA ends are ligated to the chromosomal DNA with help of host cell factors. This process generates short duplications at both ends of the provirus[97,98]. Moreover, IN binding is important to localise the vRNA into the interior of the mature capsid and thus allows proper mature capsid formation[99]. Integrase strand transfer inhibitors (INSTI) are often used in modern ART treatments[100].

1.5.10 Envelope Glycoproteins (Env)

Envelope glycoproteins enable fusion of the viral and the cellular membrane during infection. The precursor protein gp160 is expressed in the ER and glycosylated in the Golgi[101]. Next, the cellular endoprotease furin cleaves gp160 into gp120 and gp41 in the trans-Golgi. gp120 and gp41 then form non-covalently linked heterodimers. Three such heterodimers form the trimeric envelope glycoprotein on the viral surface[102]. Within the trimer, gp41 serves mainly as a transmembrane protein, whereas gp120 is on the surface. Viral entry is induced by gp120 binding to the CD4-receptor. Subsequent conformational changes within the trimer allow binding to the coreceptors CXCR4 or CCR5[103,104]. This induces conformational changes in the trimer triggering exposure of the gp41 N-terminal fusion peptide that is then inserted into the plasma membrane. Further gp41 rearrangement and refolding bring the viral and plasma membrane closer together, and a fusion pore is formed[36,105]. Fusion is thought to mainly occur at the plasma membrane but can also occur in endosomes via clathrin-mediated endocytosis with the endosomal membrane when CXCR4 is present as a coreceptor[106-109].

1.5.11 Trans-Activator of Transcription (TAT) and Regulator of Expression of Viral proteins (REV)

At the beginning of the HIV life cycle, transcription of the RNA genome is abortive, and the resulting RNAs are fully spliced. These initial transcripts are transported into the cytoplasm, where they give rise to TAT and REV proteins[110].

Trans-activator of transcription (TAT) is important for controlling HIV transcription[111,112]. TAT enters the nucleus where it binds to the TAR sequence at the

5' end of viral transcripts. It then recruits several factors, including the RNA Polymerase II (RNAPII) elongation factor P-TEFb, into the super elongation complex (SEC). The kinase CDK9 is part of P-TEFb and phosphorylates the CTD of RNA pol II and other associated complexes. This is necessary to overcome the inhibitory state of RNA pol II and to allow efficient translational elongation[113-116]. TAT is also involved in other processes, and its uptake by neighbouring cells can cause inflammation and cytotoxicity[117].

Regulator of expression of viral proteins (REV) is necessary to transport unspliced or partially spliced vRNA into the cytoplasm. It enters the nucleus via its nuclear localisation signal and binds to the REV response element (RRE), a highly structured element in the env gene of unspliced or partially spliced viral RNA transcripts[118,119]. REV oligomerises and folds the RRE into the RNA stem-loop[120]. Cellular components are recruited, and the vRNAs are transported into the cytoplasm where the REV-RNA complex dissociates[121,122]. The unspliced mRNAs can then be packaged into virions to provide full-length genomes or are translated into viral proteins[110].

1.5.12 Negative Factor (NEF)

NEF is a nonenzymatic accessory protein of HIV. It is N-terminally myristoylated and mainly associated with the plasma membrane and perinuclear membrane complexes[123-125]. NEF hijacks host proteins necessary for protein trafficking in endocytic and late secretory pathways to favour virus expression. NEF can, for example, hijack the clathrin-coated vesicle pathway (CCV) pathway leading to the downregulation of host proteins. It interacts with adaptors (AP-1, AP-2) and accessory proteins of the CCV to traffic the cellular proteins away from the cell surface. The proteins can then be rerouted to the endolysosomal system by NEF binding to components of the ESCRT machinery. This leads to downregulation of host cell factors such as SERINC3/5 or MHC-I and this downregulation interferes with virus recognition and killing of infected cells. NEF is also involved in other pathways such as COPI vesicles, where it appears to increase components of the endosomal/ exosomal pathway to increase its spread and influence autophagy. In addition, it can also increase loading of proteins into vesicles by influencing the affinity of sorting adaptors, and thus leads to upregulation of surface proteins (reviewed in [126,127]).

1.5.13 Virion Infectivity Factor (VIF)

VIF prevents Apolipoprotein B mRNA-editing catalytic polypeptide-like 3 (APOBEC3) incorporation into virions. The APOBEC3 family is an important host restriction factor against many viruses, with the subtypes s3G and 3F being the most important HIV-1 restriction factors[128]. APOBEC3 is a cytidine deaminase that can be incorporated into virions. During reverse transcription, it turns cytosine into uracil at the (-)-strand of ssDNA[129,130]. VIF hijacks the host ubiquitin-proteasome pathway by acting as a substrate receptor for APOBEC in the Cullin-Ring E3 ubiquitin ligase complex, thereby inducing the degradation of APOBEC by the ubiquitin-proteasome pathway[131-134].

1.5.14 Viral Protein R (VPR) and Viral Protein U (VPU)

VPR is packaged into virions by interacting with the p6 region of Gag[135,136] and has many roles during viral infection. Proteomic studies have shown that VPR changes the proteome of the cell significantly by targeting multiple proteins for degradation by hijacking the ubiquitin-proteasomal system and other pathways[137]. This allows VPR to inhibit activation of the innate immune system[138]. It can furthermore modulate cell cycle arrest and apoptosis. It also reactivates HIV-1 from latency through transactivation of HIV LTRs. In addition, seems to involved in integration and reverse transcription and in many more processes, most of these are not fully understood yet (reviewed in [139,140]).

VPU is a transmembrane protein[141], that downregulates host proteins by targeting the newly synthesised proteins for degradation via the proteasomal or lysosomal pathway[142-145]. In addition, it enhances the release of virions by antagonising the restriction factor tetherin that otherwise restricts the release of enveloped viruses from the membranes of infected cells[143,146]. In addition, it also plays a role in many other pathways including apoptosis, it influences the transport of host cell proteins from ER to Golgi, and it downregulates MHCII molecules (reviewed in [147]).

1.6 The HIV capsid and IP6 during HIV infection

1.6.1 Inositol-hexakisphosphate (IP6)

Inositol phosphates (IPs) consist of a hexagonal carbon ring with phosphates attached. IPs are involved in many physiological functions such as signal transduction, trafficking, and gene expression[148-151]. IP6 is the most abundant inositol phosphate and is involved in trafficking and various nuclear processes[152-154]. Cellular levels have been measured between 24 – 47 μM in mammalian cells lines and tissues[155]. Although most IP6 is taken up via diet, IP6 can also be produced de novo in a complex and not fully resolved process[156,157].

The main pathway of IP6 synthesis is shown in **Figure 7**. Synthesis starts with Inositol a component of phosphatidylinositols in the cell membrane[158]. Hydrolysis of the inositol head group of the membrane enriched PI(4,5)P2 by phospholipase C (PLC) results in IP3 The inositol-polyphosphate multikinase (IPMK) adds phosphate molecules to IP3 to yield IP4 species I(1,4,5,6)P4, I(1,3,4,6)P4, I(1,3,4,5)P4 and finally (I1,3,4,5,6)P5. IP5 can also be generated in a secondary not fully resolved pathway by Inositol-tetrakisphosphatase (IPTK). IPTK can add 1- and 6- phosphate leading to I(3,4,5,6)P4 and I(1,3,4,5)P4 resulting in (I1,3,4,5,6)P5[159-163]. Addition of 2-phosphate to IP5 (I1,3,4,5,6)P5 by inositol-pentakisphosphate-2-kinase (IPPK) yields IP6[164,165]. Additional phosphates can be added to IP6, thereby generating pyrophosphates (PP-IPs) involved in various signalling pathways[166,167].

IP levels are also regulated by the multiple inositol polyphosphatase 1 (MINPP1)[168]. MINPP1 removes 3- phosphate from I(1,2,3,4,5,6)P6 resulting in I(1,2,4,5,6)P5. I(1,3,4,5,6)P5 is catalyse into I(1,4,5,6)P4 and I(1,3,45)P4 dephosphorylation results in I(1,4,5)P3[161,169,170].

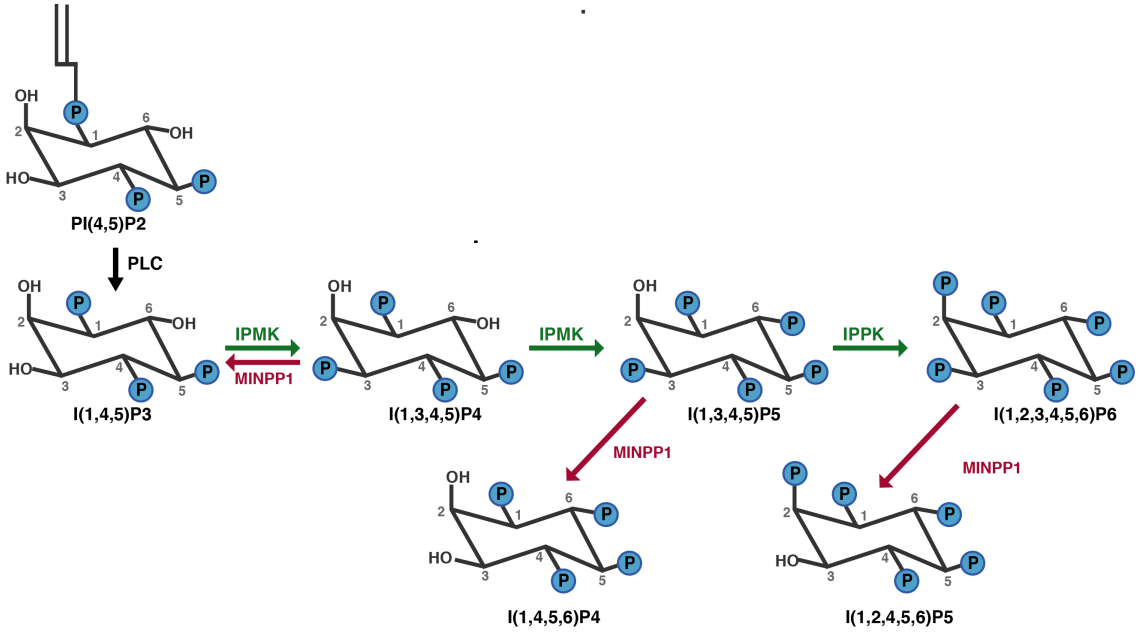


Figure 7: Main pathway of IP6 biosynthesis. Phospholipase C (PLC) removes headgroup of PI(4,5)P2 to generate IP3. Phosphates are added by inositol-polyphosphate multikinase (IPMK) to generate IP5. The final 2-phosphate is added by inositol-pentakisphosphate 2-kinase (IPPK) to generate IP6. Multiple inositol polyphosphatase 1 (MINPP1) removes 3-phosphates of IP6, IP5 and IP4.

Importantly, IPMK, IPPK and MINPP1 are found in all cells including $CD4^+$ T-cells [171-175]. IPPK is present in the nucleus and cytoplasm, whereas IPMK is confined to the nucleus[176,177]. MINPP1 is restricted to the ER lumen and thus has limited access to higher-order IPs[169,178-180].

1.6.2 Assembly of the Immature Capsid

HIV-Gag is trafficked to the plasma membrane where the MA domain interacts with the phospholipids at the inner leaflet[181-183]. The NC engages unspliced vRNA to package viral genomes into the viral particle that nucleates Gag polymerisation[41,184]. The immature Gag assembles into a hexameric lattice via inter-and intramolecular CA interactions (**Figure 8** A). A stretch of 20 amino acids in Gag is conserved in retroviruses and is referred to as the major homology region (MHR) which is important for Gag hexamer assembly[185,186]. In addition, CA_{CTD} and SP1 form a six-helix bundle that stabilises the immature hexamer[187,188]

Inositol hexakisphosphate (IP6) promotes efficient Gag assembly in vitro[189]. One IP6 interacts with two positive lysine rings (K158 and K227) in the CA_{CTD} part of Gag (**Figure 8** B)[190]. Mutation of the IP6 binding sites decreases HIV infection[190] by

decreasing both particle production and virion infectivity[191]. Likewise, reduction of IP6 levels by knocking out the IPMK or IPPK decreases HIV production[190-192]. Infectious virions in IPMK KOs package the same proportion of IP6 molecules suggesting that HIV can enrich IP6[191]. In IPPK KO, loss of the kinase results in increased IP5 levels[191]. Interestingly, IP5 was shown to partially substitute for IP6 in vivo and in vitro [190-192].

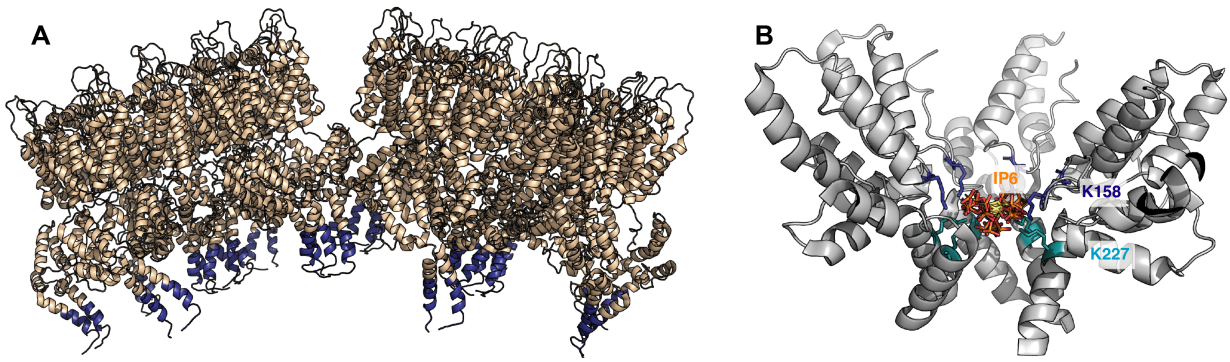


Figure 8: Structure of the immature Gag. (A) Cryo-ET structure of the GagSp1 lattice. CA is colourd in beige and SP1 is coloured blue (PDB: 7ASL[193]) (B) CA_{CTD}-SP1 in complex with IP6 (PDB: 6BHR[190]).

1.6.3 Virus maturation

The released immature virus particles undergo maturation to become infectious. Gag and Gag-Pol proteolytic processing by the PR proceeds in a highly ordered, processive cascade (**Figure 9**). First, the SP1 and NC junctions are cleaved, followed by cleavage between MA-CA, SP2-p6 and NC-SP2[194-196]. After MA-CA cleavage, the 12 N-terminal residues of CA form the β-hairpin, which is a key structural feature of the mature lattice (salt bridge between Pro-1 and Aps51)[197,198]. However, the formation of the hairpin is not necessary for the assembly of the mature capsid [53]. The NC condenses with the RNA and forms the ribonucleoprotein complex (RNP) within the capsid[199]. The final and slowest cleavage is between CA-SP1 and shifts the lattice to a mature form since it disrupts the 6-helix bundle formed by CA_{CTD}-SP1 that stabilises the immature lattice. Interfering with protease cleavage leads to aberrant, non-functional cores[53,200]. Finally, CA assembly into hexamers and pentamers to form a capsid that encloses the vRNA, RT and IN, however this transition is not fully understood.[188,201].

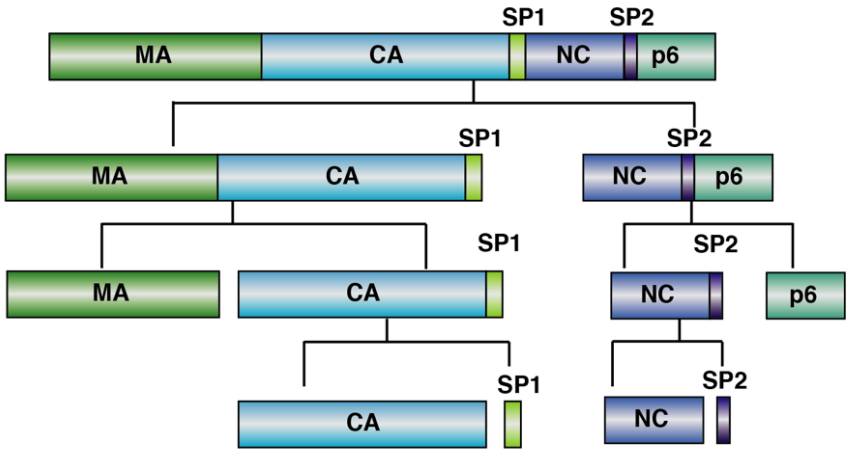


Figure 9: Schematic of Gag processing by HIV-1 protease. The sequential order of proteolytic cleavage is shown. The first cleavage occurs between SP1 and NC, followed by cleavage between MA-CA, SP2-p6 and NC-SP2. Final cleavage occurs between CA-SP1.

However, it is unknown how the mature capsid assembles, and different models were proposed. The disassembly/reassembly model suggests that mature assembly only starts once Gag has been fully cleaved[202-204]. The displacive model claims that the mature lattice forms without liberating the CA protein[205]. A partially displacive model combines elements of both and hypothesises that assembly begins as soon as CA protein is freed at the edge of the immature lattice[206].

1.6.4 Infection and transport to the nucleus

After membrane fusion, the capsid containing the viral genome is released into the cytoplasm. Capsid stability is crucial for productive infection. The HIV capsid is very fragile, and most point mutations lead to virions with a strongly decreased infectivity[207]. This was also observed with mutations that form viral particles with normal shape and amount but have a reduced or increased capsid stability suggesting that the kinetics of uncoating are tightly regulated[208,209].

Within the cytoplasm, an intact capsid core is crucial. It protects the viral DNA genome from DNA sensing host factors that could induce antiviral pathways[210-212]. In addition, reverse transcription primarily happens in the cytoplasm most likely in an at least largely intact capsid. An enclosed environment keeps the RT at a locally high concentration to ensure efficient reverse transcription which is beneficial because the RT dissociates and reassociates during reverse transcription[208,213]. In in vitro endogenous reverse transcription assay (ERT), dNTPs are supplemented to purified

capsids or permeabilised virions. This allows detection of reverse transcription within the capsids during the ERT experiment[49,214]. Reverse transcription within an intact capsid requires import of dNTPs into the interior of the capsid. It was therefore proposed that a positively charged pore in the centre of hexamers is important to recruit dNTPs and transport them into the capsid[49].

However, CA does not only protect the HIV genome and reverse transcription from the hostile host cytoplasm but also interacts with several host proteins in the cytoplasm, which are crucial for infection. The capsid hijacks the microtubule-depended transport to traffic to the nucleus[215]. The capsid was shown first shown to be moved along microtubules by the motor protein dynein[216]. Later more interactions with microtubule-associated factors were identified. It binds to the dynein adaptor protein Bicaudal D2 (BICD2) that activates dynein mediated transport[217]. CA interacts with microtubule-associated proteins 1(MAP1) that promote efficient trafficking[218]. It associates with the fasciculation and elongation factor zeta 1 (FEZ1) to exploit the motor protein kinesin and promote kinesin trafficking towards to nucleus[219,220].

CypA is a host-peptidyl cis-trans isomerase that interacts with a conserved proline-rich loop of CA exposed in mature and immature viruses (called CypA binding loop)[221,222]. It preferentially binds to curved lattice and interacts with CA of three different hexamers[223]. Cytoplasmatic CypA has been shown to be essential for infection in certain cell lines e.g. in Jurkat T-cells and primary $CD4^+$ cells[224-226]. It catalyses cis-trans isomerisation within the binding loop and has been postulated to induce conformational changes that could influence capsid stability. However, data supporting this hypothesis are contradictory[37]. CypA interaction may be involved in trafficking and nuclear entry, but the impact of the CA-CypA interaction seems to be dependent on the cell type[227,228].

IP6 is crucial to stabilise capsid cores by binding into the positively charged pore comprised of six arginine rings in the centre of hexamers[229]. The addition of IP6 to recombinant CA leads to the assembly of CA to cone-shaped capsids suggesting that IP6 promotes cone assembly[190]. In addition, IP6 was shown to influence the efficiency of reverse transcription in vitro in permeabilised viruses and purified capsid cores due to core stabilisation. In line with this, adding capsid components that destabilise the

capsid also inhibit ERT[214,229,230]. Depleting IP6 in target cells had little to no effect on HIV infection, probably because most of the bound IP6 will stay associated with the virus[191,192]

The capsid is also a target for restriction factors. Tripartite motif 5 (TRIM)α can bind to the HIV capsid and form hexagonally patterned higher-order structures[231,232]. This targets the capsids for degradation and thus inhibits reverse transcription[233,234]. In addition, TRIM5α promotes signalling pathways to induce the immune system[235,236]. However, CypA can mask TRIM5α binding sites[237,238]. In addition, the myxovirus resistance protein 2 can bind to the capsid and restrict HIV infection by inhibiting CA dependent nuclear import[239-241].

1.6.5 Nuclear import and Uncoating

Once a capsid reaches the nuclear membrane, it has to pass through the nuclear pore complex (NPC). The NPC is composed of around 30 nucleoporins (NUPs) and mediates transport across the nuclear membrane allowing passive diffusion for small molecules and active transport of larger molecules via transport factors[242]. Nuclear transport of HIV is necessary to allow infection of non-dividing cells[243]. HIV-CA was first shown to bind to phenylalanine-glycine (FG) rich repeats in CPSF6 and nuclear pore proteins such as NUP153. As FG repeats are present in around ¼ of NUPs[244], determining which proteins HIV-1 maintains FG binding for is complex. Nup358 and Nup153 are amongst the most well characterised nuclear entry cofactors for HIV-1 infection. Nup358 is a multidomain protein that forms cytoplasmic filaments and is part of the RanGTPase for nuclear import[242]. It possesses a duplicated CypA domain at its extreme C-terminus, which binds to the CypA binding loop of CA and captures cores at the NPC[227,245]. Within the pore the capsid was show in interact with the FG repeats of various other NUPs and these interactions might facilitate translocation into the pore[246,247]. When the cores reach the nuclear basket, they are thought to interact with NUP153 located at the nuclear site of the NPC where it is integral for controlling transport for example quality control and retention of unspliced RNAs[248,249]. NUP153 binds across a junction of neighbouring hexamers and promotes translocation[250,251].

The cleavage and polyadenylation factor CPSF6 is an mRNA processing protein that is primarily located in the nucleus but can shuffle between the nucleus and cytoplasm[252]. It recognises the same site as NUP153 and might compete for the CA binding site[250,251]. This may allow cores to be released from the NPC into the nucleus[253]. However, certain mutants that do not interact with NUPs or CPFS6 can still enter the nucleus suggesting that alternative import pathways are used[243,246,254].

Transportin 1 (TRN-1) and Transportin 3 (TNPO3) are karyopherins believed to be important for HIV nuclear entry and translocation. It has been reported that they may bind CA directly[255-257], but it is more likely that they interact indirectly through binding of FG-containing proteins or are required for transporting factors such as CSPF6 to the nucleus[258,259]. In addition, TPNO3 was shown to be involved in HIV integration[260-262].

The timing of uncoating is highly debated. Recent data suggest that mostly intact cores are trafficked through the nuclear pore, followed by disassembly within the nucleus[38-40,263]. Whether these cores are 100% intact cannot be resolved in these studies, however, it is now clear that nuclear pores are large enough to allow capsids to pass without disassembly, meaning that prior uncoating is not necessary. In the nucleus, intact capsids were rarely observed[38]. Furthermore, fully transcribed dsDNA was only observed in the nucleus thus nuclear import happens prior to RT completion[39,40,263]. This is line with previous suggestions, that uncoating is dependent on reverse transcription[264-266]. Atomic force microscopy (AFM) studies indicate that a pressure increase due to the increased size and stiffness of the dsDNA triggers progressive uncoating[266,267]. This was further supported by vitro RT assays, where some cores that completed reverse transcription were partially broken with DNA loops protruding[214]. Similar partially broken core structures were also observed in the nucleus, suggesting that local rupture by reverse transcription of the capsid might trigger uncoating[38,263]. This may happen close to the integration sites[40]. The fully reverse-transcribed HIV genome is then integrated into actively transcribed region gene-rich regions (euchromatin). This is directed by CPSF6-CA interaction[264,268,269].

1.7 Aims of this thesis

In 2001, the first report of IP6 induced assembly of Gag was published[189], but the mode of IP6 interaction could not be determined[270]. Recent research gave valuable insights into IP6 binding to the mature and immature capsid; however, fundamental questions remain. Whilst current data suggests that IP6 is an important cofactor for HIV[190-192], it remains unclear if it is essential for HIV-1 replication, and if so, whether it is required for immature or mature assembly or both. In the mature capsid, residues necessary for IP6 binding are also implicated in dNTP recruitment to allow reverse transcription, thus it is difficult to dissect these as separate functions and assess their relative importance[49,271]. In addition, the relationship between capsid binding drugs and IP6 has yet to be understood. The work in this thesis tackles these questions using a combination of structural, biochemical, and cellular methods.

Chapter 2 Materials and Methods

2.1 Bacterial transformation

Between 10 and 200 ng plasmid DNA was mixed with 50 µL chemically competent *Escherichia coli* DH10B, DH5α, XL-1 blue, BL21or C41 on ice. The mixture was incubated on ice for 15-30 min before heat shock at 42 ˚C for ~45 s, followed by incubation on ice for 3 min. Cells were then suspended in SOC or TY-media and plated on TYE agar containing the appropriate antibiotic(s) (100 µg/mL Ampicillin, 34 µg/mL Chloramphenicol, 50 µg/mL Kanamycin). If protein expression was planned for the following day, the cells were added directly into liquid 2xTY media containing the appropriate concentration of antibiotics. Plates and pre-cultures were incubated at 37 ˚C overnight (the pre-cultures were shaking at 200 rpm).

2.2 Molecular cloning

2.2.1 Gibson Assembly

Cloning was mostly performed using Gibson Assembly[272]. Primers were designed using SnapGene (version 5.3.1) with primer overhangs of 15-25 bp and a melting temperature greater than 50˚C. Both vector backbone and insert were amplified by PCR using Q5 high-fidelity polymerase (New England Biolabs) and the PCR program shown in **Table 1**.

Table 1 General PCR program for cloning.

PCR Step	Temperature (˚C)	Time
Initial denaturation (1x)	98	30 s
Denaturation (35x)	98	10 s
Annealing (35x)	72	30 s
Extension (35x)	72	30 s kb^{-1}
Final extension (1x)	72	2 min

After PCR, the products were treated with Dpn1 (New England Biolabs) for 3 h at 37°C to remove template DNA. They were then purified using the QIAquick PCR purification

Kit (QIAGEN). For Gibson Assembly, ~70 ng of the linearised vector was mixed with insert in a molar ratio of 1:3 to 1:7 and mixed with Gibson Assembly Master Mix (1x) (New England Biolabs) on ice. Reactions were incubated at 50 ˚C for 0.5-1h before transformation (of 5 µL maximal volume) into 50 µL *E. coli* DH5α or XL-1 blue cells. On rare occasions, plasmid was linearized using restriction digestion, followed by agarose gel purification using a DNA gel purification kit (QIAGEN).

2.2.2 Site directed mutagenesis

Point mutations and small deletions or insertions were introduced via QuickChange. Primers for most HIV CA mutants were designed by adding 20 – 25 up- and downstream of the mutation site and forward and reverse primers were fully complementary. Primers for pCRV Gag-Pol were generated using the site-directed mutagenesis protocol[273] with a mutation site in the overlapping primer sequence. The overlapping primer sequence had a Tm of ~ 46 ˚C and the non-overlapping sequences were 5-10 ˚C higher.

Mutagenesis PCR was initially performed with Pfu Turbo polymerase (Agilent) using Pfu buffer and 10 mM dNTPs in a 50 µL format using 50 ng of plasmid DNA and 0.5 µM of the forward and reverse primers each and the PCR program given in **Table 2**.

Table 2 General PCR program for quick change mutagenesis.

PCR Step	Temperature (˚C)	Time
Initial denaturation (1x)	94	5 min
Denaturation (35x)	94	1 min
Annealing (35x)	56	1 min
Extension (35x)	72	1 min kb^{-1}
Final extension (1x)	72	10 min

For some plasmids, Pfu did not work efficiently. In those cases, mutagenesis was performed using the Q5® Hot Start High-Fidelity Master Mix (New England Biolabs) with 25 µl Master Mix, 0.5 µM forward and reverse primer and 50 ng DNA in 50 µL.

Table 3 General PCR program for quick change mutagenesis.

PCR Step	Temperature (˚C)	Time
Initial denaturation (1x)	98	3 min
Denaturation (35x)	98	30 min
Annealing (35x)	72	1 min
Extension (35x)	72	30 min kb^{-1}
Final extension (1x)	72	10 min

After PCR, the samples were treated with Dpn1 (New England Biolabs) to remove input DNA, before 2-5 μL were transformed.

2.3 Plasmid purification

Plasmid DNA was purified from 2 mL cultures of *E. coli* cloning cells in 2xTY using the appropriate antibiotic. Cultures were grown over night at 37˚C and 200 rpm. Plasmid purification was performed using the QIAprep Spin Miniprep Kit (QIAGEN). Plasmids were validated by Sanger sequencing,

2.4 Agarose Gels

DNA samples were mixed with Purple Gel Loading Dye (New England Biolabs) to a final concentration of 1 X. Samples were loaded onto 1% agarose gels made with 1 X TBE buffer supplemented with 1:50,000 v/v ethidium bromide or SYBR-safe. Agarose gels were run in 1 X TBE buffer at 85 V for 45 min and visualised by ultra-violet transillumination or blue light transillumination for gel extraction.

2.5 LDS-PAGE

Protein samples in 1x LDS buffer (either containing 50 mM DTT or not) were loaded onto NuPAGE gels (Invitrogen, 10-, 12-, 15- or 17-well) after boiling for 2-10 min at >95 ˚C. Gels were run in MES buffer at 180 V for 40 min (in rare cases 60 min). Proteins were visualized using Instant Blue (Abcam).

2.6 Plasmids

For protein production: Mutants of HIV CA and HIV-Hex in pOPT vectors were generated using QuickChange. HIV ΔMA-CANC was cloned into pET17b and mutations were generated using QuickChange. Ca-SP1 was cloned into pET17b with a SUMO-Tag via restriction digest and Gibson Assembly by Donna Mallery. The protease was ordered codon optimized as synthetic DNA (Integrated DNA technologies) and cloned into pET17b by Gibson assembly.

For viral vector production: packaging plasmid pMDG2, which encodes VSV-G envelope (Addgene plasmid # 12259), pNL4-3-derived pCRV GagPol (HIV-1 clade B)[274], and pCSGW[275] were used. Mutagenesis of CA was performed using the QuickChange method (Stratagene) against pCRV GagPol. These are referred to as viruses in the main text.

2.7 Protein expression and purification

All cells were grown in 2xTY media supplemented with 2 mM $MgSO_4$, 0.5 % glucose and 100 μg mL^{-1} ampicillin unless otherwise indicated.

Capsid proteins were expressed in *E. coli* C41 cells. At an OD600= 0.7-1, expression was induced with 0.5 mM IPTG and incubated for 4 h at 37°C. The cells were harvested by centrifugation (4000 rpm, 20min). Cell lysis was done via sonication in lysis buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20% BugBuster, Protease inhibitor tablets, 1 mM DTT) and centrifuged (24 000 rpm, 1 h). The supernatant was precipitated with 25% ammonium-sulphate (wt/vol) for 15-30 min at 4°C followed by centrifugation (13 000 rpm, 20 min, 4°C). The precipitated CA was resuspended and dialysed against dialysis buffer (50 mM MES (pH 6.0), 0-20 mM NaCl, 1mM DTT). The CA protein was further purified via a cation-exchange column (5ml HiTrap® SP Fast Flow, GE Healthcare) with a gradient from 20mM -1M NaCl followed by size exclusion chromatography on a HiLoad 16/600 Superdex 75 prep grade column (GE Healthcare) with Tris pH 8.0, 20 mM NaCl, 1mM DTT. The protein was concentrated using a Vivaspin Turbo 15 (Sartorius) centrifugal concentrator with a 10-kDa cut-off.

The disulfide-stabilised CA hexamer was purified as previously described with minor modifications[276,277]. Briefly, CA protein with mutations (A14C, E45C, W184A, M185A) was expressed in *E. coli* C41 cells, lysed and cleared by centrifugation. The supernatant was precipitated in 25% ammonium-sulphate, and the pelleted material was resuspended in 50 mM citric acid (pH 4.5), followed by dialysis against the same buffer with 20 mM 2-mercaptoethanol. Afterwards, the protein was dialysed into 50 mM Tris-HCl (pH 8.0), 1 M NaCl, 20 mM 2-mercaptoethanol. The reducing agent was removed by dialyzing against 50 mM Tris (pH 8.0), 1 M NaCl, and then finally into 20 mM Tris (pH 8.0), 40 mM NaCl. Reassembled hexamers were identified by non-reducing SDS–PAGE and were further purified via anion-exchange (2X 5ml HiTrap® Q Fast Flow, GE Healthcare) followed by size exclusion chromatography (Superdex 200).

CA-SP1 proteins were expressed in *E. coli*. expression was induced with 0.4 mM IPTG at 37°C for 4 h. Bacterial pellets were resuspended in lysis buffer (100 mM Tris pH 8.0, 200 mM NaCl, 1 mM DTT, 1 tablet/L protease inhibitor, bug buster, 20 mM Imidazole), lysed by sonication and the cellular debris was removed by centrifugation (1 h, 20000 rpm). The supernatant was filtered through a 0.2-µm filter, applied to a Ni^{2+} affinity resin, and eluted with imidazole. The eluted protein was dialysed overnight in the presence of ULP1 protease and applied to reverse Ni^{2+}-chromatography to remove the SUMO tag and ULP1 protease. The protein was applied to S75 column in 20 mM Tris-HCl, 20 mM Tris-HCl, pH 8, 50 mM NaCl, 1 mM DTT.

The ΔMA-CANC protein was expressed as described previously with a few alterations[278]. ΔMA-CANC expression was induced with 0.5 mM IPTG and 10 μM $ZnCl_2$ in *E. coli* C41 cells for 4 h at 25°C, lysed via sonication in lysis buffer (50 mM Tris-HCl pH 7.4, 500 mM NaCl, 1 mM DTT, 10 μM $ZnCl_2$, 20% BugBuster, Protease Inhibitor Tablets). 0.1% PEI (v/v) was added, and the lysate was stirred for 10 min. The lysate then was centrifuged (24 000 rpm, 1h) and the supernatant was precipitated in 25% ammonium sulphate (wt/vol), followed by centrifugation (13 000 rpm, 20 min). The precipitated protein was resuspended in dialysis buffer (50 mM Tris-HCl pH 7.4, 40 mM NaCl, 1mM DTT, 10 μM $ZnCl_2$), dialysed into the same buffer and applied to a cation-exchange column (5ml HiTrap® Q Fast Flow, GE Healthcare). The flow-

through was precipitated with 0.1% PEI, followed by centrifugation and precipitation with 25% ammonium sulphate. The precipitate was resuspended in dialysis buffer and applied to a cation-exchange column (5ml HiTrap® SP Fast Flow, GE Healthcare) and eluted with a gradient of 40 mM -1M NaCl.

The HIV protease was purified as previously described[279]. Briefly, protease was expressed in *E. coli* BL21(DE3) cells for 4 h at 37°C. Cells were harvested by centrifugation, resuspended in PR buffer (20 mM Tris [pH 8.0], 0.1 M NaCl, 5 mM imidazole, 1 mM β-mercaptoethanol) supplemented with protease inhibitor tablets (Roche). After centrifugation, the pellet containing inclusion bodies was washed twice with 30 ml of PR buffer supplemented with 2 M urea and 1% Triton X-100 and once with water with subsequent centrifugation. The pellet was resuspended in PR buffer supplemented with 8 M urea and incubated overnight. The protein was purified on Ni (NTA) resin (Qiagen) and diluted with 20 mM Tris-HCl (pH 7.9), 100 mM NaCl, 5 mM imidazole, and 8 M urea. PR was refolded by stepwise dialysis against a solution containing 20 mM Tris-HCl (pH 7.9), 100 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 10% glycerol, and 1 mM DTT, with gradually decreasing urea concentrations from 8 to 0 M.

All proteins were snap-frozen as 100 µl aliquots and stored at -80°C.

2.8 Western Blotting

Proteins were transferred onto nitrocellulose membrane using an iBlot Gel Transfer Device (Invitrogen) using programme P1 for 7 min. Blots were blocked in either 5% (w/v) non-fat milk (Marvel) in PBS-T (PBS with 0.01% Tween 20) or 3 % BSA (Sigma Aldrich) in PBS-T. Antibodies were diluted in milk or BSA solutions given above and were incubated for at least 1 h. The membrane was washed with PBS-T. For chemo-luminescence detection, Amersham ECL, ECL Prime or ECL Select detection reagent (GE Healthcare Life Sciences) were used. Fluorescent blot detection was performed using the LI-COR Odyssey CLx imaging system. Band intensities were quantified using Image Studio Lite (LI-COR) software. If required, blot membranes were stripped using ReBlot Plus Strong Antibody Stripping Solution (Millipore).

2.8.1 Antibodies

Anti-HIV-1 p24 (183-H12-5C) was obtained from the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: Anti-HIV-1 p24 Monoclonal (183-H12-5C) (Cat# 3537) from Dr. Bruce Chesebro and Kathy Wehrly[280,281], loading control COX IV (P/N 926-42214) was obtained from Li-Cor Biosciences.

2.9 Cell lines

293T CRL-3216 cells were purchased from ATCC. All cells are regularly tested and are mycoplasma free. HEK293T and HeLa cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin (GIBCO) at 37°C with 5% CO_2). The MT4 and SupT1 T-cell lines were maintained in RPMI 1640 with L-glutamine (Corning) and supplemented with 10% FBS (GenClone), penicillin (100U/ml), and streptomycin (100 mg/ml).

2.10 Viral vector production

Replication deficient VSV-G pseudotyped HIV-1 vector particles were produced in HEK293T cells using the packaging plasmid pMDG2, which encodes VSV-G envelope (Addgene plasmid # 12259), pNL4-3-derived pCRV GagPol (HIV-1 clade B)[274], and pCSGW[275] as described previously[282].

2.11 Infection experiments

For infection experiments with 293T, cells were seeded at $1x10^4$ cells per well into 96-well plates and left to adhere overnight. The media was replaced with FluoroBrite-DMEM (GIBCO) with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin and 5 mg/ml polybrene. Indicated amounts of particles were added, and the plates were scanned every 8 h for up to 72 h in an IncuCyte® (Satorius) to identify GFP-expressing cells. Transduction of HeLa cells was performed in presence of 5 µg/ml polybrene in 6-well plates seeded with 10^5 cells per well. The plates were scanned at the indicated time points in an IncuCyte® (Satorius) to identify GFP-expressing cells.

2.11.1 IPMK and IPPK KO cells

KO cells lines were produced by Donna Mallery (as described here [191]). MINPP1 was located to the PM by inserting a sequence encoding the first 15 amino acids of Gnai2 (MGCTVSAEDKAAAER) containing N-myristol and S-palmitoyl motifs on the N-terminus of MINPP1.

2.11.2 HIV quantification

Vector genomes were quantified using the Cell-to-C_T Kit (Invitrogen). 100 µl viral supernatant was centrifuged to pellet the viral particles, resuspended with 20 µl lysis buffer with DNaseI 1/100 and incubated at room temp for 10 mins. 2 µl Stop solution was added to terminate the reaction. 10 µl RT Buffer (2x) was mixed with 0.5 µl RT enzyme mix and 9.5 µl lysate and incubated at 37˚C for 1 h, followed by 95˚C for 5 min. run on an ABI StepOnePlus Real-Time PCR System (Life Technologies) (37˚C 1 hr, 95˚C, 5 min). The qPCR was performed on an ABI StepOnePlus Real-Time PCR System (Life Technologies) using 2 µl RT product, 5 µl TaqMan Fast Universal PCR Mix (ABI) and 0.5 µl GFP primer-probe (GFP) in a 10 µl reaction ((GFPF (CAACAGCCACAACGTCTATATCAT), GFPR (ATGTTGTGGCGGATCTTGAAG) and probe GFPP (FAM-CCGACAAGCAGAAGAACGGCATCAA-TAMRA). CSGW plasmid with calculated copies was used as a standard.

The level of RT enzyme was quantified using either a colourimetric RT assay kit (Roche) according to manufacturer's instructions or qRT-PCR as described previously with slight alterations[283]. In brief, 5 µl of viral supernatant was mixed with 5 µl lysis buffer (0.25% Triton X-100, 50 mM KCl, 100 mM Tris-HCl (pH 7.4), 40% glycerol) and 0.1 µl RNase Inhibitor and incubated for 10 min at room temperature before diluting to 100 µl with nuclease-free water. 2 µl of lysate was added to 5 µl TaqMan Fast Universal PCR Mix, 0.1 µl MS2 RNA, 0.05 µl RNase Inhibitor and 0.5 µl MS2 primer mix, to a final volume of 10 µl. The reaction was run on an ABI StepOnePlus Real-Time PCR System (Life Technologies) with an additional reverse transcription step (42°C, 20 min) and followed by amplification.

2.11.3 qPCR for in-cell reverse transcription products

Viral supernatants were treated with 250 U/ml DNase (Millipore) for 2 h to remove contaminating DNA. DNase treated viruses were added to cells as per infection protocol and incubated at 37°C. Cells were harvested at indicated time points and the DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions.

Reverse transcription products were detected by qPCR from 2 μl DNA sample using TaqMan Fast Universal PCR Mix (ABI) and RU5 primers to detect strong-stop DNA40 (RU5 forward: 5' CTGGCTAACTAGGGAACCCA-3'; RU5 reverse: 5'-CTGACTAAAAGGGTCTGAGG-3'; and RU5 probe 5'-(FAM) TTAAGCCTCAATAAAGCTTGCCTTGAGTGC(TAMRA)–3') and GFP primers to detect first-strand transfer products (see above). Reverse transcription measurements are representative of 3 experiments with each point measured in triplicate. Results are represented as mean ± standard deviation.

2.12 Preparation of HIV cores and encapsidated reverse transcription assay

Purification of HIV cores and the encapsidated reverse transcription assay were performed as described previously[229]. Briefly, the supernatant containing VSV-G pseudotyped HIV-1 GFP was passed through a 0.45 μm nitrocellulose filter and pelleted through a 20% sucrose cushion (w/v) (28,000 rpm at 4°C, 2 h, Beckman SW32Ti rotor; Beckman Coulter Life Sciences). All following solutions were prepared in CPB (20 mM Tris (pH 7.4), 20 mM NaCl, 1 mM $MgCl_2$). The pellets were resuspended in CPB and treated with DNase I for 2 h (Sigma Aldrich) at room temp. The resuspended virus was layered on top of a gradient of 80–30% sucrose gradient (5% steps) overlayed with 1% Triton X-100 in 15% sucrose and spun at 32,500 rpm at 4°C for 16hr with a Beckman SW40Ti rotor (Beckman Coulter Life Sciences). The gradient was fractionated, and the location of cores was determined via a RT activity assay (see above). Core-containing fractions were pooled, aliquoted and snap-frozen before storage at –80°C. For the ERT assay, viral cores were equalized (as determined by RT) by diluting them in 60% sucrose. 10μl diluted cores were mixed

with 100 mM Tris pH 8, 100 mg/ml DNase I or BSA (as a negative control), dNTPs and or IP6 (were added at indicated concentrations) in a final reaction volume of 20 µl. Reactions were incubated at room temperature for 16 h. Reverse transcriptase products were detected using TaqMan Fast Universal PCR Mix (ABI) with RU5 primers to detect strong-stop DNA (described above).

2.13 TRIM5 abrogation assay

The assay was performed by Donna Mallery as described [49]. GFP or RFP VSV-G pseudotyped HIV-1 virus was produced as described above and concentrated by ultracentrifugation. Viruses were titrated on FRhK-4 cells at the indicated ratios and infection monitored by measuring GFP and RFP expression in cells using an IncuCyte® (Sartorius).

2.14 Assembly Assays

CA proteins were dialysed against 50 mM MES (pH 6.0), 40 mM NaCl, 1mM DTT over night at 4 °C. CA proteins at a final concentration of 25-200 µM were mixed with NaCl (final concentration 2.5 M) or IP6 (final concentration 25 µM-10 mM, Sigma) at 25°C.

250 µM CA-SP1 proteins and the indicated IP6 concentrations and/or 0.5 mM or 1mM PF96 were mixed at 25 °C.

ΔMA-CANC proteins were diluted into assembly buffer (20 mM Tris [pH 7.5], 140 mM KCl, 10 mM NaCl, 5 mM MgCl2, 10 mM TCEP) to 50-200 µM. Assembly of ΔMA-CANC was initiated by adding 1:10 ssRNA GT_{25} and IP6, IP5, IP4, IP3 or IP2 (SiChem) at the indicated concentration.

The increase in Abs_{350} was measured using a PHERAstar FSX Plate reader (BMG Labtech) in 384-well plate with shaking between each measurement at 25°C or 37°C.

2.15 ΔMA-CANC Cleavage Assay

100 µM ΔMA-CANC was assembled with ssRNA and 50 µM IP6 at 37°C for 2 h. The protein was diluted 1:2 with cleavage buffer (20 mM MES (pH 6.0), 140 mM KCl, 10 mM NaCl, 5 mM $MgCl_2$, 10 mM TCEP) and incubated for 1h at 37°C. Protease was

added to the assembly mixture at a 1:50 ratio to Gag and incubated at 25°C. Samples were taken at the indicated time points the reaction was stopped with NuPAGE® LDS Sample Buffer (Invitrogen) to stop the reaction, and then subjected to NuPAGE Novex 4–12% Bis-Tris gel (Invitrogen) for cleavage products analysis and visualized by Coomassie blue staining.

2.16 Differential Scanning Fluorimetry

DSF measurements were performed using a Prometheus NT.48 (NanoTemper Technologies) over a temperature range of 20–95°C using a ramp rate of 2.0 or 2.5°C/min. CA hexamer samples were prepared at a final concentration of 200 μM monomer in PBS in the presence or absence of 4 mM DTT. 200 μM ligands were added 1:1. 75 μM ΔMA-CANC was assembled with 1:10 ssRNA at 37°C for 2 h with and without 50 μM IP6. DSF measurements were performed at least three times.

2.17 Negative Stain

For mature capsid assembly, 4-10 μl of sample was put onto a glow discharged carbon-coated grid (Cu, 400 mesh, Electron Microscopy Services) incubated for 2-5min, washed once with water, and twice with 2% Uranyl-acetate and stained with 2% Uranyl-acetate for 3-5min.

For immature assembly, 4-10 μl of sample was put onto a glow discharged carbon coated grid (Cu, 400 mesh, Electron Microscopy Services) incubated for 2-5min, washed twice with 2% Uranyl-acetate and stained with 2% Uranyl-acetate for 3-5min.

Micrographs were taken at room temperature on a Tencai Spirit (FEI) operated at an accelerated voltage of 120 keV and Gatan 2k × 2 k CCD camera. Images were collected with a total dose of ~30 $e^-/A^{\circ 2}$ and a defocus of 1–3 μm.

2.18 Virus particle production for Tomography

Virus-like particles were produced in HEK293T as described above. Supernatants were harvested and passed through a 0.45 μm filter followed by a 0.22-µm filter. The particles were concentrated by ultracentrifugation over a 20% (wt/vol) sucrose cushion (2 h at 28,000 rpm in a Beckman SW32 rotor; Beckman Coulter Life Sciences). The

pellet was resuspended in PBS. WT, K25A, K158A, T8I, K158/T8I mutant viruses were further purified by applying the resuspended particles to a 6-18% iodixanol gradient (1.2% increment steps and centrifuged for 1.5 h at 250,000 × g in a Beckman SW40 rotor (Beckman Coulter Life Sciences)[284]. The virus-containing fraction was diluted in 1:10 PBS and concentrated by ultracentrifugation (45 min,38,500 rpm in a Beckman SW40 rotor, Beckman Coulter Life Sciences). The pellet was resuspended in PBS and incubated at 4°C overnight to allow full resuspension.

2.19 Cryo-Electron Tomography

10-nm-diameter colloidal gold beads were added to the purified HIV-1 mutants. 4 μl sample-gold suspension was applied to a glow discharged C-Flat 2/2 3C (20 mA, 40 s). Grids were blotted and plunge-frozen in liquid ethane with a FEI Vitrobot Mark II at 15 °C and 100% humidity. Tomographic tilt series of K25A, K25A/T216I, K25A/N21S, KAKA, KAKA/T8I and WT were acquired between –40° and +40°with increments of 3°, on a TF2 Tecnai F20 transmission electron microscope equipped with a Falcon III Direct Electron detector at 200 kV using Serial-EM under low-dose conditions at a magnification of 50000x and a defocus between -3 μm and -6 μm. Tomography of the R18G, K158A, K158A/T8I, T8I, IPMK and IPPK mutants was performed on a FEI Titan Krios transmission electron microscope at 300 kV equipped with a Gatan K2 summit direct electron detector and a Gatan Quantum energy-filter (GIF). Tilt series were acquired between –60° and +60° with increments of 3° using a dose symmetric scheme using Serial-EM[3]. Images were collected at a magnification of 33000x with 10 frames per tilt and a total dose of ~120 e–/Å2 across all of the tilts. Frames were aligned in SerialEM with a final pixel size of 3.667 Å per pixel in the unbinned image stacks. Tomograms were reconstructed using IMOD (4.9)[285]. The alignment of 2D projection images of the tilt series was performed using gold beads as fiducial markers, tomograms were reconstructed by back projection. Tilt-series of K158A, K158A/T8I, T8I were collected by Zunlong Ke.

2.20 Crystallisation

CA hexamer protein was prepared exactly as described previously[229]. Crystals were grown at 17 °C by sitting-drop vapour diffusion in which 100 nl or 200 nl protein was mixed with 100 nl or 200 nl precipitant and suspended above 80 μl precipitant in the MORPHEUS I screen[286].The K25A structure was crystallised and collected by Donna Mallery and solved by Leo James. 12 mg/ml hexamer were mixed with 1mM of myo-IP_5 and cryoprotected with precipitant supplemented with 20% MPD. Crystals were flash-cooled in liquid nitrogen and data was collected at beamline I24 at Diamond Light Source. The data sets were processed using the CCP4 Program suite[287]. Data were indexed and integrated with iMOSFLM and scaled and merged with AIMLESS[288]. Structures were solved by molecular replacement using the model 6ES8 in PHASER[289] and refined using REFMAC5[290]. Between rounds of refinement, the model was manually checked and corrected against the corresponding electron-density maps in COOT[291]. The model and data were deposited in the PDB database with code 6R6Q.

The K25A/N21S and K25T/N21S hexamers at concentrations of 10-15 mg/ml were mixed with 1mM IP_6 or ATP and cryoprotected with precipitant supplemented with 20% MPD. The structures were solved by molecular replacement using PHASER-MR implemented in the Phenix suite[292] with the search model 6ES8. Model building and real-space refinement was carried out in COOT[291] and refinement was performed by using phenix-refine toll in PHENIX and REFMAC5 iteratively[292,293]. Final figures were rendered in The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

2.21 TIRF Assays

The affinity of Atto488-ATP (NU-805-488, Jena Bioscience) binding to CA lattices was measured using a capsid biosensor based on TIRF microscopy[294,295]. Briefly, CA structures were self-assembled using a mixture of CA A204C/A92E (for cross-linking and increased solubility, respectively) and AF647-labelled CA K158C and then captured on the surface of a coverslip. Binding of Atto488-ATP (10–500 nM) to surface-immobilized fluorescent CA assemblies was then imaged by TIRF microscopy

and quantified to obtain an equilibrium binding curve. The K_D of the interaction was obtained by fitting the curve with an equilibrium binding model.

Replication deficient HIV-1 virions without envelope protein were produced in HEK293T cells using pCRV-1 GagPol and pCSGW, biotinylated using EZ-Link Sulfo-NHS-LC-LC-Biotin (Thermo Scientific, 21338) and purified as described [296,297].

TIRF microscopy was carried out following the published method of Marquez et al[296,297]. Briefly, biotinylated viral particles were captured onto coverslips attached to microfluidic flow cells and imaged using a custom built TIRF microscope with an ASI-RAMM frame (Applied Scientific Instrumentation), a Nikon 100 x CFI Apochromat TIRF (1.49 NA) oil immersion objective and NicoLase laser system. Immobilised virions were treated with imaging buffer containing 200 nM PFO, to permeabilize the lipid envelope, and Alexa Fluor 568-labelled CypA (0.5-1 μM), to detect the capsid. TIRF images were then acquired with a frequency of 1 frame/6 s using a 561 nm laser with a 20 ms exposure time for excitation and an Andor iXon 888 EMCCD camera for detection. Single-virion fluorescence traces were extracted from the TIRF image stacks using the JIM Immobilized Microscopy analysis package (https://github.com/lilbutsa/JIM-Immobilized-Microscopy-Suite) and further analysed in MATLAB (The MathWorks, Inc) using software adapted from previous work [298]. Briefly, the duration and intensity of the CypA signal was extracted from fluorescence traces by step-fitting using change point analysis. The lifetime of each capsid was determined as the time difference between acquisition of Alexa Fluor 568-CypA upon permeabilization and loss of fluorescence upon capsid uncoating.

TIRF assays were performed by K.M. Faysal from Till Böckings Lab.

Chapter 3 Role of the R18 and K25 ring in the mature capsid

3.1 Introduction

CA proteins form a conical capsid made up of hexamers and 12 pentamers. At the centre of each hexamer is a positively charged pore. A ring of arginines at position 18 (R18) within this pore, recruits, and binds to polyanions like IP6 and dNTPs[49,229]. IP6 promotes the assembly of mature VLPs in vitro and stabilises mature hexamers[190,229]. Recruitment of dNTPs and their import into the interior of the capsid could be crucial to fuel DNA synthesis[49]. However, the importance of IP6 and nucleotide-binding on mature hexamers and their interplay is not understood. Moreover, it is difficult to pick these functions apart since mutating R18 in the mature capsid destroys both IP6 and dNTP binding[49,229].

A possible candidate to understand more about the role of the pore is the lysine ring K25. Mutating K25 leads to a significant reduction in HIV production and infection, but WT-like virions can still be observed[299]. Molecular dynamics studies suggested that K25 is involved in nucleotide import. It might facilitate nucleotide translocation from R18 into the capsid by creating a lower energy electrostatic potential barrier to the interior of the capsid. It was thus proposed that mutating K25 to alanine increases this energy barrier thereby reducing dNTP movement into the capsid[300]. Along these lines, another study suggested that K25 creates a second energy minimum in the pore if a second polyanion is present and it might be important to coordinate solvent molecules to nucleotides after dewetting of dNTPs at R18[301].

In addition, K25 could also be involved in IP6 binding. When crystal structures of the capsid hexamer crystal structures in complex with IP6 were solved, an additional electron density for a potential second molecule was observed in several data sets. However, the electron density in these data sets was too weak to build this additional ligand with confidence. This is due to the 6-fold symmetry within the pore leading to averaging of the data. As a result, the orientation of the axial phosphate in IP6 is ambiguous. To circumvent this problem, hexamers were crystallised in complex with

IP5, which only contains equatorial phosphates. In these structures, clear density for two molecules of IP5 was visible. Both IP5s were located at the centre of the pore in a planar conformation separated by ~10-12 Å. One molecule was co-ordinated above the R18 and a second one between the R18 and the K25 (**Figure 10**).

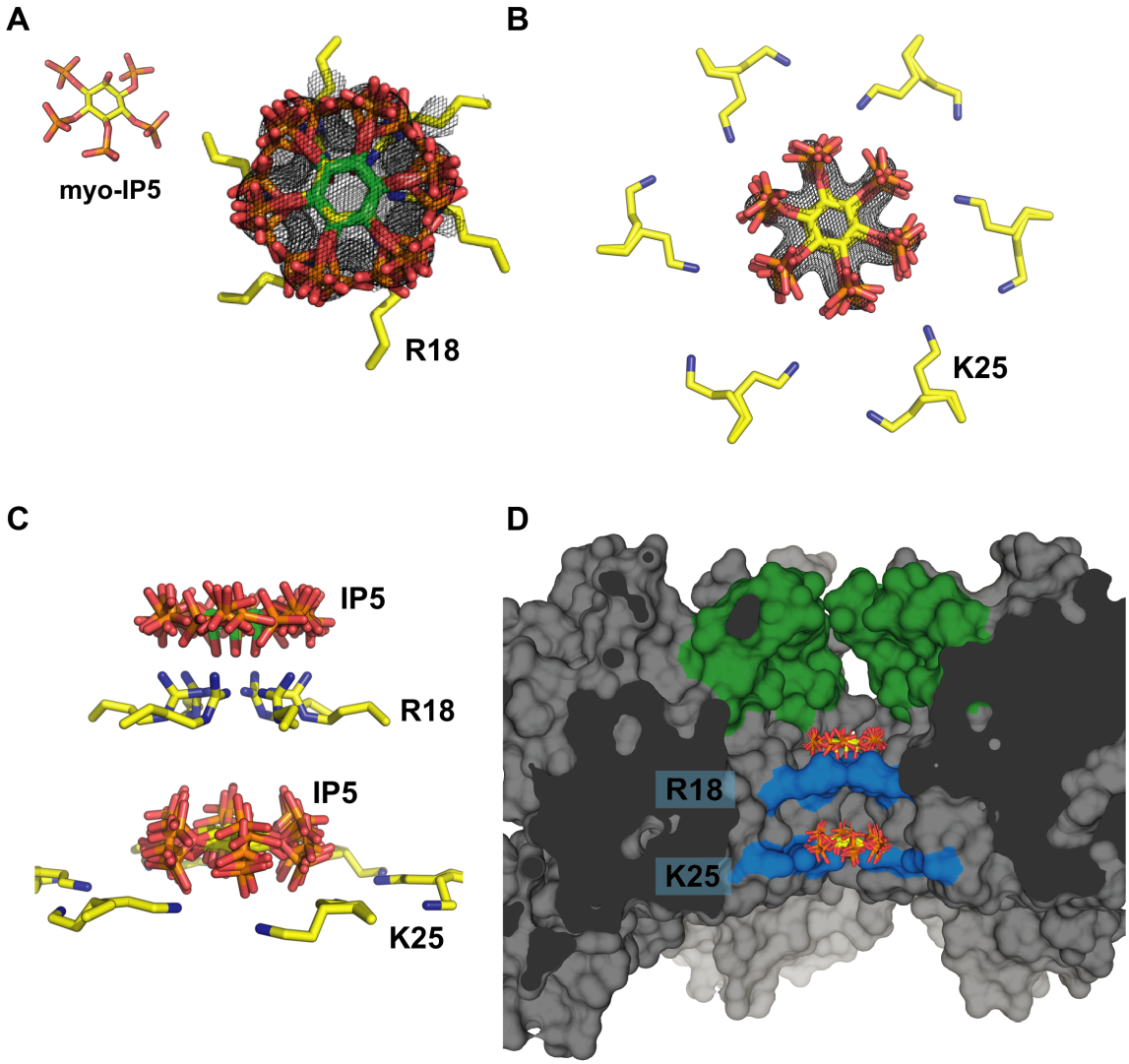


Figure 10: Structure of IP5 bound to HIV CA hexamer. (A) (left) Structure of myo-IP5 ligand. Fo – Fc omit density (mesh) contoured at 2.0σ centred on IP5 bound to R18 and viewed down the 6-fold axis (right) and from the side (below). (B) Fo – Fc omit density (mesh) contoured at 2.0σ centred on a second IP5 molecule next to K25 (two rotamer side chains are shown). All six-symmetry equivalent IP5 molecules are shown. (C) View showing the two IP5 molecules, one binding above R18 and one above K25. Note the second IP5 molecule is located closer to the K25 ring than R18. (D) Cross section through the hexamer, showing the central chamber where the IP5 molecules are bound. The β-hairpin is shown in green and the location of R18 and K25 in blue. The structure in this figure was determined by Donna Mallery and Leo James.

3.1.1 Aims

K25 is implicated in virion production, infection and nucleotide import but there is little consensus on its role or mechanism of action. Similar to R18, K25 was suggested to be involved in dNTP and IP6 binding. Thus, it represents an ideal target for understanding more about the role of binding of both molecules in the mature capsid. In this project I used infection experiments, capsid assembly assays and cryo-ET to determine the role of K25 in the pore.

Some figures in this section were published in[302] and adapted for this thesis.

3.2 Results

3.2.1 K25 mutation impairs infection and reverse transcription

To investigate the importance of the lysine ring, K25 was mutated to either arginine, alanine, asparagine, isoleucine, or serine. K25A, K25N K25I, and K25S mutants were significantly less infectious than WT (**Figure 11** A). The infection-deficit was similar to R18G, a previously characterised IP6- and nucleotide-binding deficient mutant[49,229] (**Figure 11** B). K25R was substantially more infectious than the other K25 mutants (**Figure 11** A, B), highlighting the importance of a positive charge at this position for productive infection. Infectivity of K25R was similar to P38A, a known CA instability mutant[208,209].

To investigate whether the decreased infectivity resulted from impaired reverse transcription, varying amounts of mutant and WT Gag plasmid were transfected to create chimeric virions with a different number of mutant to WT CA per hexamer. It was shown that sequential removal of arginine at position 18 results in a dose-dependent loss of infectivity and a corresponding loss of DNA synthesis[49]. Likewise, the decrease in the infectivity of K25A chimeric viral particles was dose-dependent (**Figure 11** C). Viruses were infectious if at least 4 out of 6 lysines were predicted per hexamer, but with < 3 the infectivity decreased. Importantly, the loss of infectivity was closely paralleled with a reduction in DNA synthesis. Next, I performed a series of encapsidated reverse transcription (ERT) experiments. Fort this, I isolated viral cores and dNTPs to initiate reverse transcription in vitro. The input of cores was normalised by RT (**Appendix Figure 1** B), and DNase was added to the experiments to degrade

all DNA that is not protected within cores. DNA synthesis in wild-type cores increased with increasing dNTPs concentrations (**Figure 11** D). In contrast, K25A showed very little ability to carry out reverse transcription even at high dNTP concentrations. I then tested if introducing additional stability by titrating IP6 at a low dNTP concentration can rescue reverse transcription. Increased IP6 concentrations increased DNA products in WT cores up until 10 μM IP6. At higher concentrations, reverse transcription plateaued, similar to what was observed previously[229]. In contrast, the addition of IP6 did not affect reverse transcription in K25A cores (**Figure 11** E). ERT experiments with K25S mutant were like K25A, although a slight increase in RT products was seen upon adding increasing dNTP concentrations (**Appendix Figure 2** A). IP6 did not increase K25S reverse transcription efficiency (**Appendix Figure 2** B). This suggests that K25 mutants are defective in either dNTP import or the ability to recruit IP6 to stabilise the capsid.

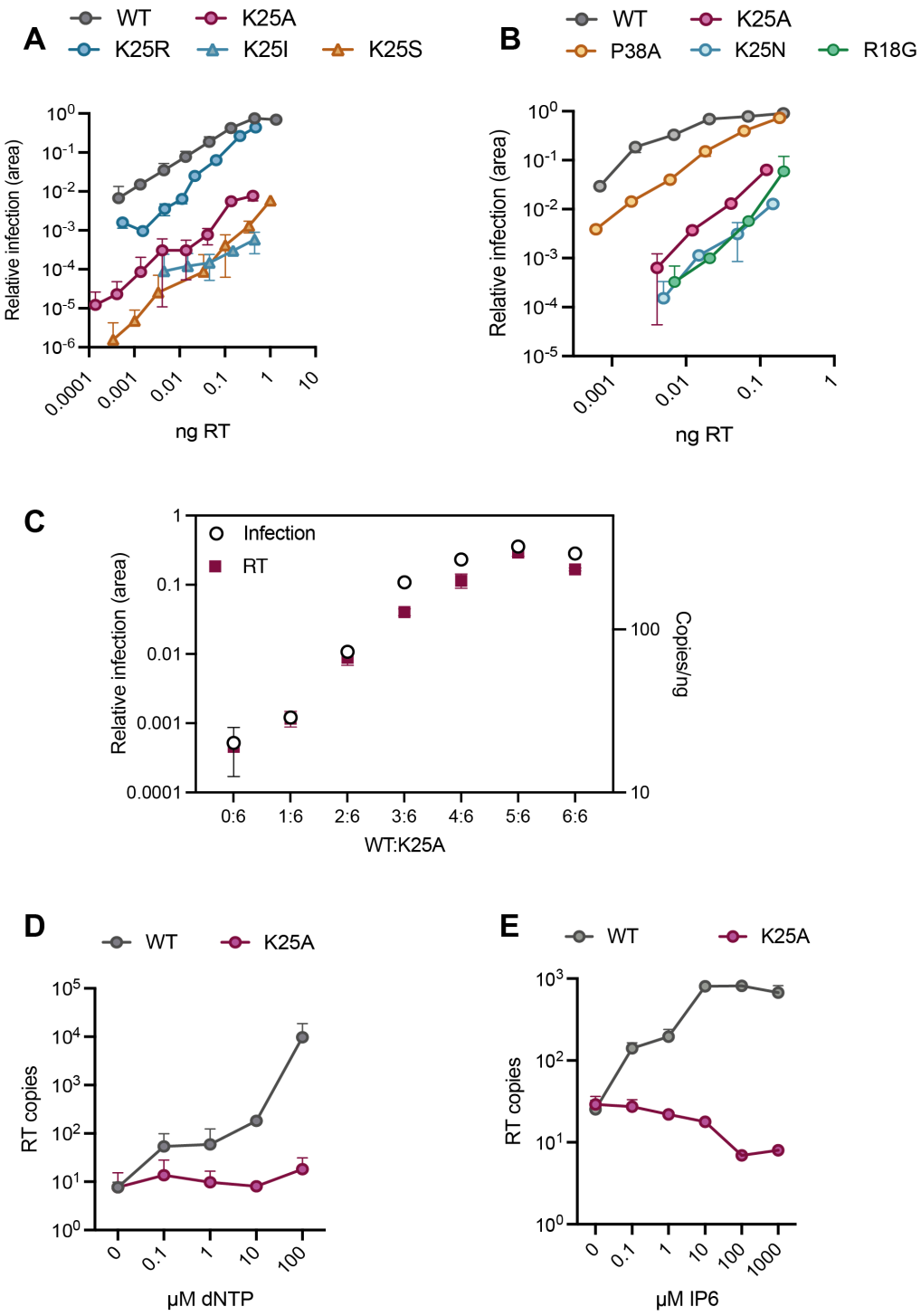


Figure 11: K25 is essential for reverse transcription and infection. (A, B) Infectivity of WT HIV and selected mutants in HEK293T cells. Infectivity is measured using an Incucyte and determined as the proportion of cell area that is GFP +ve at a given viral dose (in ng RT). Error bars in infection experiments depict mean ± SD of three replicates from one experiment representative of three independent experiments. (C) Matching infectivity (GFP +ve cell area) and reverse transcription (strong-stop, RU5, at 4 hours post-infection) of chimeric viruses produced with an increasing ratio of WT:K25A Gag in Hela cells. Error bars in infection experiments depict mean ± SD of three replicates from one experiment representative of three independent experiments. (D) ERT assay measuring the synthesis of strong-stop DNA in the presence of DNase of WT and K25A cores at increasing concentrations of dNTPs in Hela cells. Error bars in infection experiments depict mean ± SD of three replicates from one

experiment representative of three independent experiments. (E) ERT assay measuring the synthesis of strong-stop DNA in the presence of DNase at 1 μM dNTPs of WT and K25A cores at increasing concentrations of IP6 in Hela cells. Error bars in infection experiments depict mean ± SD of three replicates from one experiment representative of three independent experiments. The experiments shown in panel (B) and (C) were performed by Donna Mallery.

Next, I determined whether K25 mutants can bind to the ligands dATP, ATP, IP6, and IP5 by measuring the thermal stability of disulfide-linked capsid hexamers (**Figure 12** A, B). K25A was able to bind to all ligands, and a similar degree of stabilisation was observed as with WT protein (**Figure 12** A). This stabilisation was only seen in hexamers, and when the disulfide bonds that stabilise the hexamers were reduced by DTT, no increase in melting temperature was measured. As expected, mutation of R18G leads to a loss of ligand binding since R18 is the main polyanion binder (**Figure 12** B). Thus, K25 does not significantly contribute to IP6-mediated stabilisation in pre-formed hexamers.

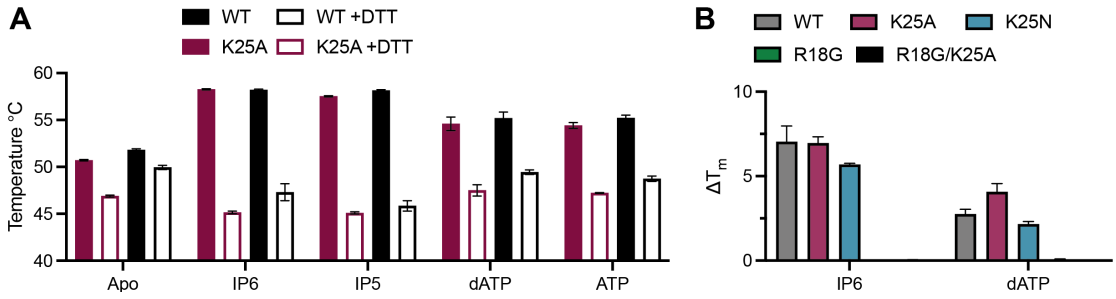


Figure 12: K25 is required for the production of HIV virions with mature capsid cores. (A) Changes of NanoDSF-derived melting temperature (ΔTm's) of K25A, K25N, R18G and R18G/K25A in the presence of different polyanions compared to the respective disulfide-stabilised hexamers in absence of ligands. Error bars depict mean ± SD of technical replicates. (B) NanoDSF-derived melting temperatures (Tm's) of WT or K25A disulfide-stabilised hexamers ± DTT and in the presence of different polyanions. Error bars depict mean ± SD of at least three technical replicates.

3.2.2 K25 and R18 are necessary to form stable capsids

I then tested the impact of K25 and R18 mutation viruses by electron cryotomography (cryo-ET). I purified WT and mutant viruses, prepared grids and collected tilt-series to generate tomograms. The majority of wild-type virions possessed visible mature capsids (**Figure 13**, **Appendix Figure 3**). Notably, I could not observe any cores for K25A. Most particles appeared to contain unstructured densities, and were, therefore, classified as ambiguous (**Figure 13**, **Appendix Figure 4**). In contrast to K25A, the majority of R18G virions contained cores. However, most were irregularly shaped, and

only very few were conical or tubular. Closer analysis revealed that compared to WT, R18G virions contained more additional higher-order structures besides single capsids (**Appendix Figure 5**). Especially multiple cores or cores with attached open or closed structures were present more frequently. Immature virions were similar in WT, K25A and R18G virions suggesting that K25A and R18G only affect the mature stage of the virus. However, this defect does not seem to be explained by impaired proteolytic processing since neither K25A nor R18G mutants show any obvious cleavage deficits (**Appendix Figure 1** D, E).

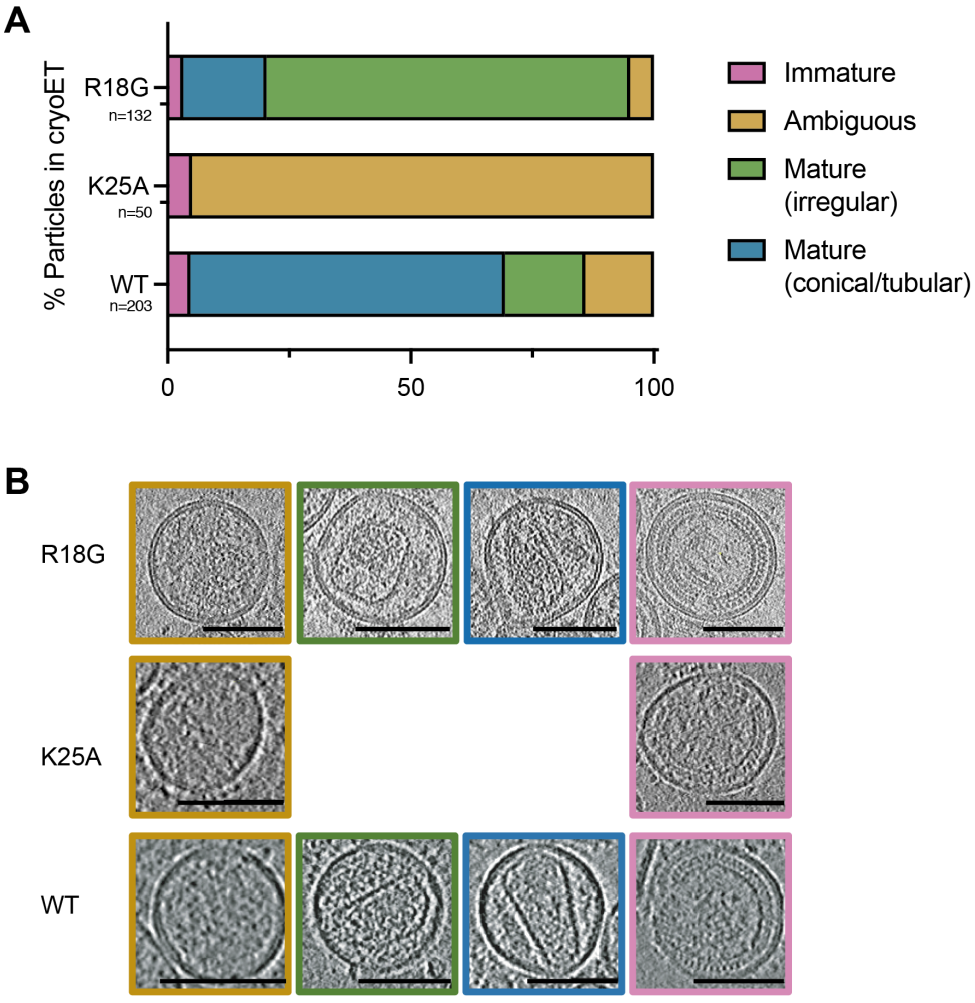


Figure 13: K25A virions cannot form capsids, whereas R18G can form irregular capsids. (A) Cryo-ET analysis of infectious WT, K25A and R18G HIV virions. A total of 203 WT, 50 K25A and 132 R18G viral particles were analysed, and the frequency of each phenotype was plotted as a percentage. Virions were classified as either immature (pink), mature with a conical or tubular capsid (blue), mature with an irregular capsid (green) or ambiguous (yellow) and slices through tomograms (B) show representative examples of the viral morphologies observed in the different datasets. Scale bars: 100 nm.

Although cryo-ET provides insight into capsid morphology, it does not reveal whether assembled cores are stable. A series of TRIM5 abrogation assays were carried out by infecting cells expressing rhesus macaque TRIM5. This assay uses the fact that TRIM5 binds to the capsid and the resulting TRIM5 restriction becomes saturated at a high multiplicity of infection. In the experiment, cells were infected with a consistent dose of WT virus that encodes a GFP reporter gene and different doses of mutant viruses encoding an RFP reporter gene. When increasing doses of WT RFP virus were added (**Figure 14** B), a corresponding increase in GFP signal from the co-infecting WT GFP virus was measured (**Figure 14** A). This occurs as TRIM5 becomes saturated by the RFP virus and no longer efficiently blocks GFP virus infection. In contrast, the K25A RFP virus gave poor RFP expression, and no increase in GFP expression from the co-infecting WT GFP virus was measured (**Figure 14** A, B). Thus, K25A virions are not only poorly infectious but also have too few stable capsids to efficiently saturate TRIM5. The experiment was repeated with K25R. This mutant could saturate TRIM5 in a dose-depended manner corresponding to an increase in the GFP signal. The saturation ability of the K25R RFP virus was ~ 1-log less efficient than WT virus. This closely correlates with its 1-log infectivity decrease (**Figure 11**) and suggests that K25R virus is less infectious because it has fewer stable capsids. R18G-RFP virus had a similar ability to saturate TRIM5 as K25R, but R18G-RFP infectivity was drastically lower compared to K25R (**Figure 14** B). This agrees well with the cryo-ET data that showed the ability of R18G viruses to form capsid structures (**Figure 13**).

To further confirm this theory, single-molecule TIRF microscopy was used to analyse the stability of the K25 mutant[296,297]. Virions are captured on a slide and permeabilised by a pore-forming toxin. Fluorescent-labelled CypA can then enter through the membrane pore and bind to capsid lattices. The fluorescence signal at locations of the viral particles through time is measured and allows estimation of the lifetime of each capsid (**Figure 14** C, D). WT and K25A viruses were compared in the presence and absence of IP6. WT capsids had an intrinsic half-life of ~5-10 minutes, but the addition of IP6 stabilised the capsids (**Figure 14** E). In contrast, most K25A capsids were short-lived with a half-life of less than 1 minute. This was not changed by adding IP6. K25R was slightly more stable than K25A in the presence of IP6 but significantly less stable

than WT (**Figure 14** E). R18G TIRF showed that 95% of the capsids were unstable capsids with a half-life of 1-2 min (**Figure 14** F). In addition, IP6 had no effect on increasing the fraction of stable capsids and probably also hardly any effect on half-life.

Taken together, this data suggests that the ability of K25A mutants to from stable cores is strongly impaired. Although R18G mutants can form mature capsids, these were mainly irregular in shape and highly unstable.

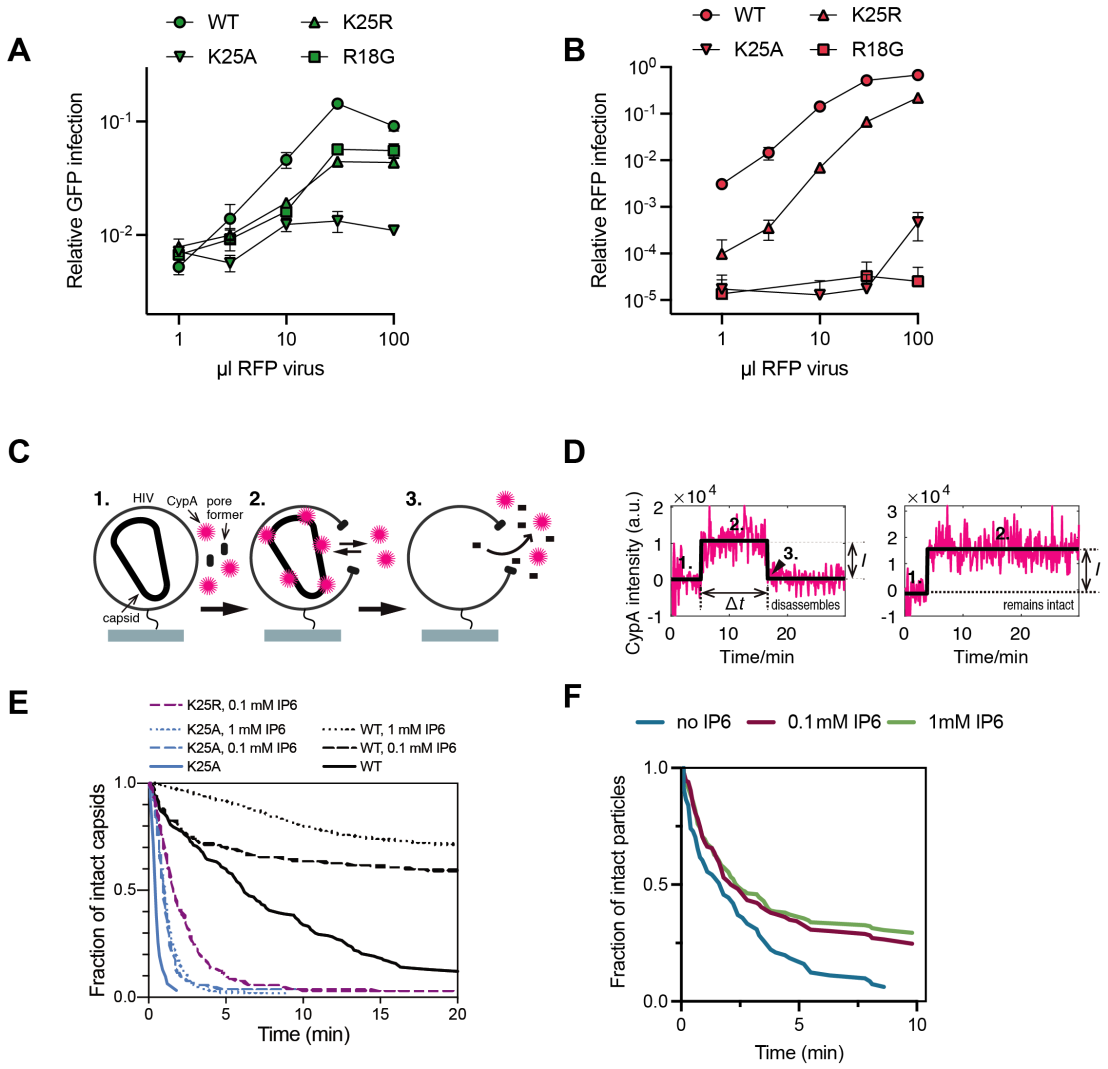


Figure 14: K25 and R18 mutant capsids are unstable. (A, B) TRIM5 abrogation assay in FRhK-4 cells. Cells expressing rhesus TRIM5 were infected with a consistent dose of WT GFP-expressing virus and increasing doses of RFP-expressing WT, K25A and K25R viruses. The level of infection of each virus (green symbols for GFP (A), red symbols for RFP (B)) was determined as the proportion of cell area that was GFP or RFP positive. Increasing doses of WT RFP virus results in an increase in RFP positive cells but also in GFP positive cells, because TRIM5 becomes saturated with RFP virus and no longer restricts infection by the GFP virus. Viruses that do not contain capsids capable of recruiting and maintaining TRIM5 binding will not saturate. Error bars depict mean ± SD of three replicates from one experiment representative of two independent experiments. (C) Schematic of the capsid uncoating

assay using TIRF microscopy and the CypA paint method. Immobilized viral particles are permeabilized in the presence of fluorescently labelled CypA while recording fluorescence traces at the locations of single virions (1). CypA binds to the capsid, resulting in the appearance of a stable fluorescence signal (2). Capsid disassembly results in the disappearance of the CypA signal (3). (D) Example traces for a capsid that disassembles (left) and a capsid that remains intact (right) during the acquisition time. The fluorescence traces (magenta) are analysed by step fitting (black) to extract lifetime (Δt) and intensity (I) for each capsid. (E, F) Capsid survival curves generated from single-virion uncoating traces. CA K25A mutant capsids fall apart rapidly and IP6 does not stabilise CA K25A or K25R capsids, unlike wild type capsids (E) and R18G (F). The experiments shown in panel (A, B) were performed by Donna Mallery and in (C-F) by K.M. Rifat Faysal.

3.2.3 Mutation of K25 and R18 impairs IP6 dependent mature capsid assembly

To determine if the specific loss of IP6 coordination is the reason for reduced assembly, I carried out in vitro assembly reactions using recombinant CA protein. I first induced assembly of WT, K25A, K25R, K25N, P38A, and R18G CA by adding 2.5 mM NaCl. High salt was previously shown to induce the formation of capsid tubes[303]. I then monitored the kinetics of assembly by measuring the absorbance at 350 nm over time and collected samples at the end of each experiment for analysis via negative stain electron microscopy (**Figure 15**). K25A, K25R, and K25N mutant viruses were capable of assembling capsid tubes like WT (**Figure 15**). Assembly kinetics were similar, but K25A showed an initial lag phase. This suggests that the mutation of K25 does not impair the CA structure to an extent where it is no longer able to form higher-order assemblies. In line with previous work, the instability mutant P38A assembled slightly faster than WT[209], and R18G mainly formed small spheres[51].

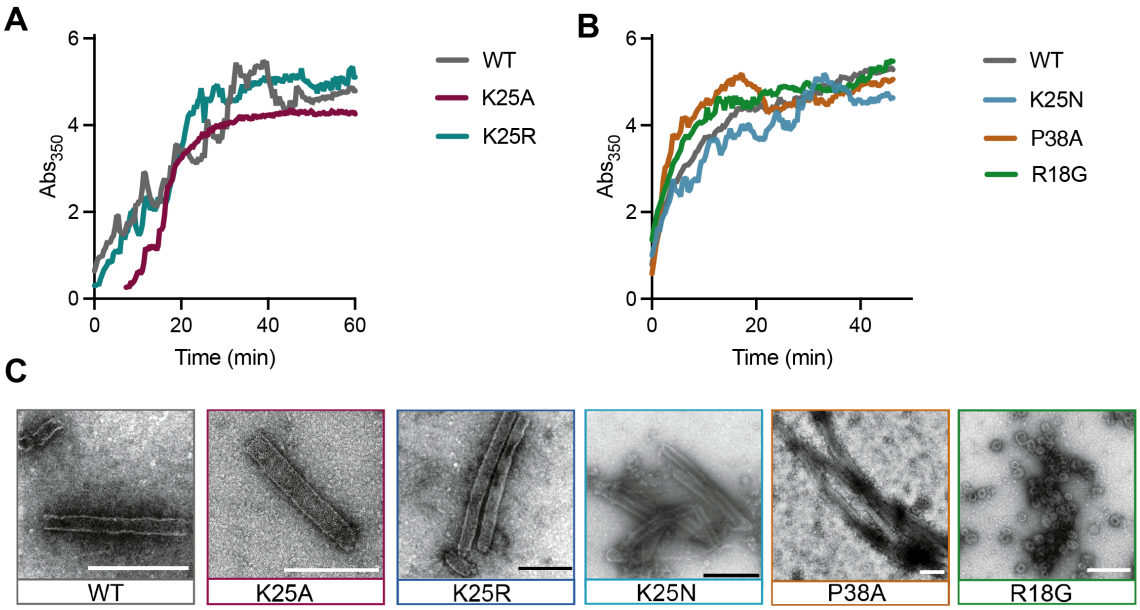


Figure 15: K25 mutants can assembly in high salt conditions. In vitro assembly reactions were monitored in real-time by measuring the absorbance at 350 nm. (A, B) Assembly reactions using 75 μM capsid protein and 2.5 M NaCl in 50 mM MES pH 6.0 (C) Negative stain EM images of material taken from (A, B). Scale bars: 200 nm.

Next, I assembled WT at low salt conditions and added increasing amounts of IP6 (**Figure 16** A). I observed a striking dose-dependent increase of assembly upon titration of IP6. In negative stain micrographs, WT CA formed conical and tubular capsids reminiscent of mature HIV cores under IP6 conditions (**Figure 16** B), similar to a previous report[190]. I then chose the highest IP6 concentration to compare the assembly of WT CA to the mutants K25A, K25N and K25R. K25A and K25N were not able to assemble, whereas K25R assembled, albeit slower (**Figure 16** C). K25R was able to form conical capsids, but in line with the decreased absorbance, I observed a reduced yield of cores in micrographs (**Figure 16** E). K25R cores also appeared to be larger than WT cores. Like R18G, K25A and K25N could not form any capsid structures. However, P38A formed cores with similar kinetics and morphology as WT. Since K25A and K25N were able to assemble under high salt conditions, I increased CA and IP6 concentrations (**Figure 16** D). At 200 μM capsid and 6 mM IP6, K25N formed tubes and cores (**Figure 16** D, E). In contrast, only a slight increase in absorbance was measured for K25A CA, but no tubes or cores could be found in negative stain EM. Finally, the K25A assembly was measured for 24 h (**Figure 16** F). At 10 mM IP6, the WT reaction was complete within 10 minutes, while the K25A absorption reached a plateau after ~ 10 hours. Negative stain electron microscopy

showed tubes similar to CA assembled at high salt conditions (**Figure 16** G). This suggests that K25A mutants are severely impaired for IP6-driven assembly of capsid tubes, while capsid cones do not form or do so too rarely to be detected by in vitro methods.

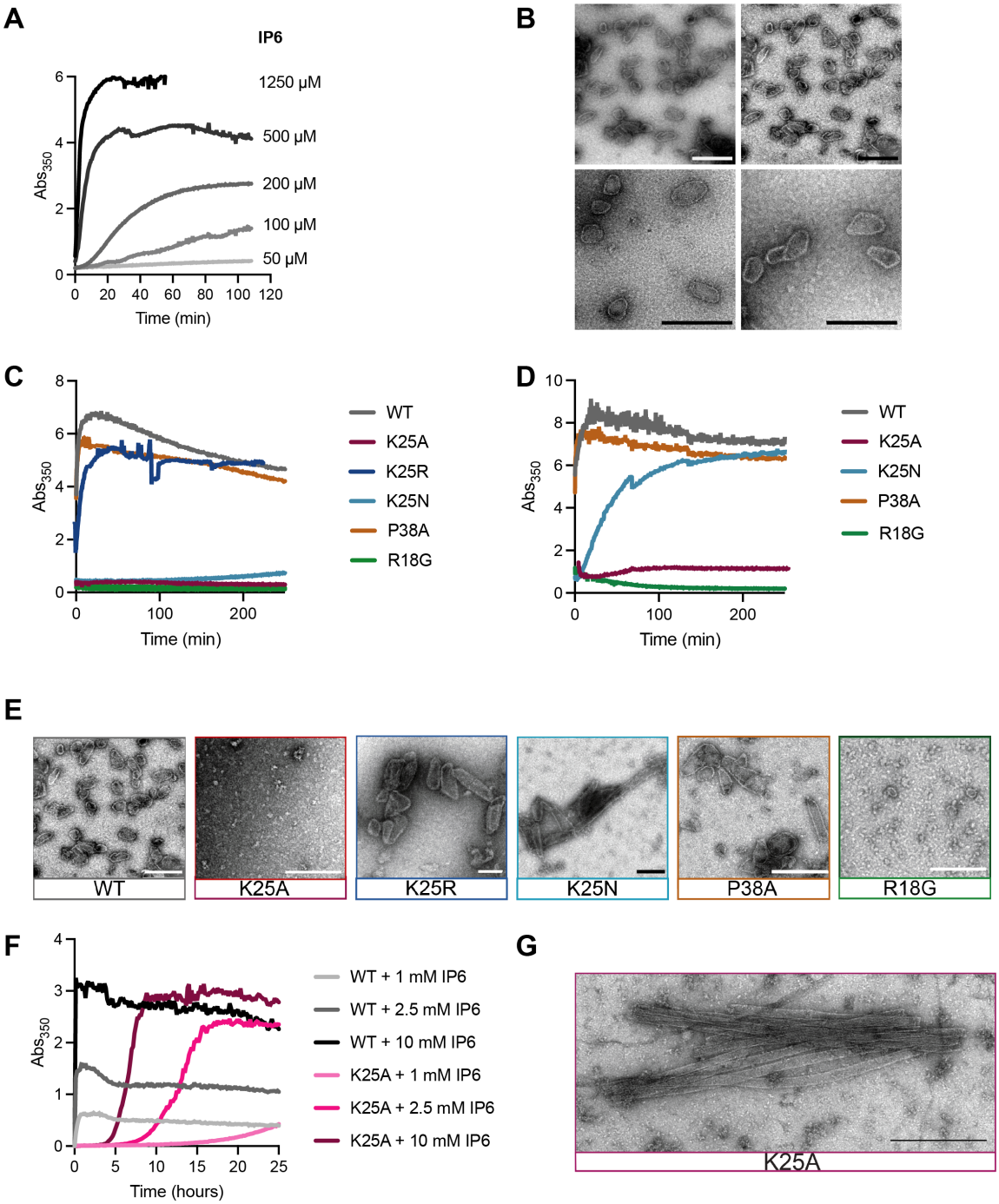


Figure 16: K25 is required for IP6-mediated assembly of HIV CA. In vitro assembly reactions were monitored in real-time by measuring the absorbance at 350 nm. (A) Assembly reactions using 50 µM WT and mutant capsid protein and increased IP6 concentrations (50-1250 µM). (B) Negative stain EM images of material taken from (A), Scale bars: 200 nm. (C) Assembly reactions using 100 µM WT and

mutant capsid protein with 1.25 mM IP6 and (D) 200 μM WT and mutant capsid protein with 6 mM IP6 in 50 mM MES pH 6.0, 40 mM NaCl (E) Negative stain EM images of material taken from (C, D). (F) Assembly reactions of 250 μM WT or K25A CA in 50 mM MES pH 6, 100 mM NaCl and a range of IP6 concentrations up to 10 mM and corresponding negative stain EM images (G), Scale bars: 200 nm. The experiments shown in panel F and G were performed by Wang Peng.

Finally, I mutated both K25 and R18. The double mutant R18G/K25A assembled into spheres in high salt conditions similar to R18G (**Figure 17**) suggesting that R18G has a more dominant effect in high salt conditions. Interestingly, at very high CA and IP6 concentrations, only R18G can be induced to form spheres and a few tubes and cones. In contrast, R18G/K25A cannot assemble into stable higher-order structures. This suggests that the presence of K25 might allow some IP6 binding because if it only were a crowding effect, the double mutant should assemble more efficiently as less repulsive charges are present in the hexamer.

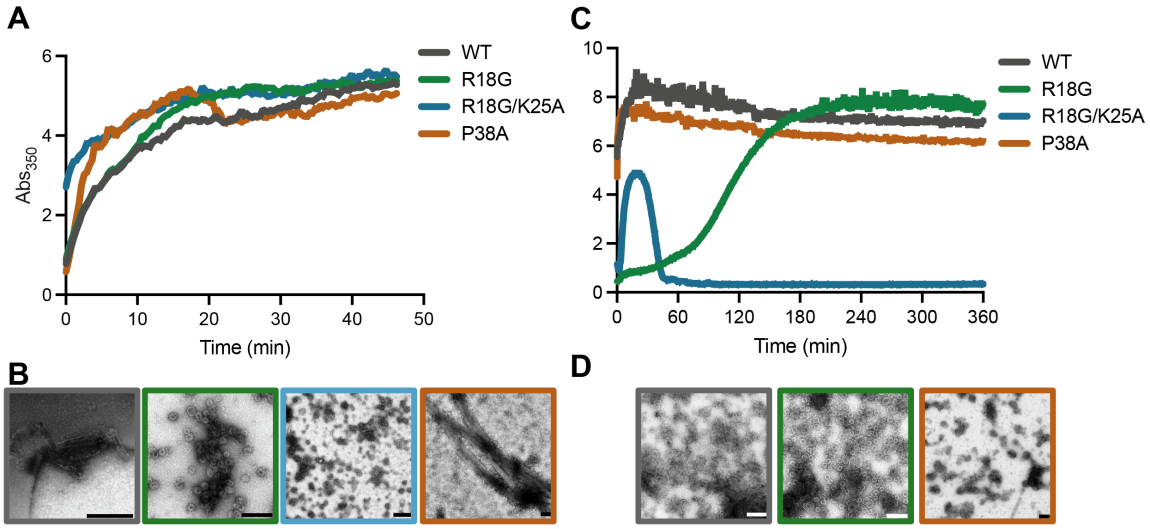


Figure 17: R18G has a more dominant effect on mature assembly. In vitro assembly reactions were monitored in real-time by measuring the absorbance at 350 nm. (A) Assembly reactions using 75 μM WT and mutant capsid protein and 2.5 M NaCl (B) Negative stain EM images of material taken from (A). (C) Assembly reactions using 400 μM capsid protein with 9 mM IP6 (D) Negative stain EM images of material taken from (C), Scale bars: 200 nm.

3.2.4 Introducing K25 compensatory mutants

To dissect K25's involvement in reverse transcription from its impact on capsid assembly, passaging experiments of K25A virions were performed by Alex Kleinpeter. The compensatory mutant T216I was discovered at the CTD-CTD interhexamer interface (**Figure 18** A). In addition, an N21S mutant was detected which is located within the hexamer pore in-between the R18 and the K25 ring (**Figure 18** B). Both T216I and N21S could rescue the K25A infection (**Figure 18** C, D). Interestingly,

mutation of K25A to threonine was also observed, and K25T was also significantly more infectious than K25A. This was surprising given that K25S is infection deficient (**Figure 11** A). If N21S and T216I were added to the K25T background, infectivity was close to WT levels. Interestingly, none of these mutations compensates for the loss of the positive K25 charge. Changing residue N21 to lysine and thus re-introducing a second positively charged ring in the pore could not rescue infection to the same extent as the N21S mutant but increased infection by 2-log compared to K25A (**Appendix Figure 6** A).

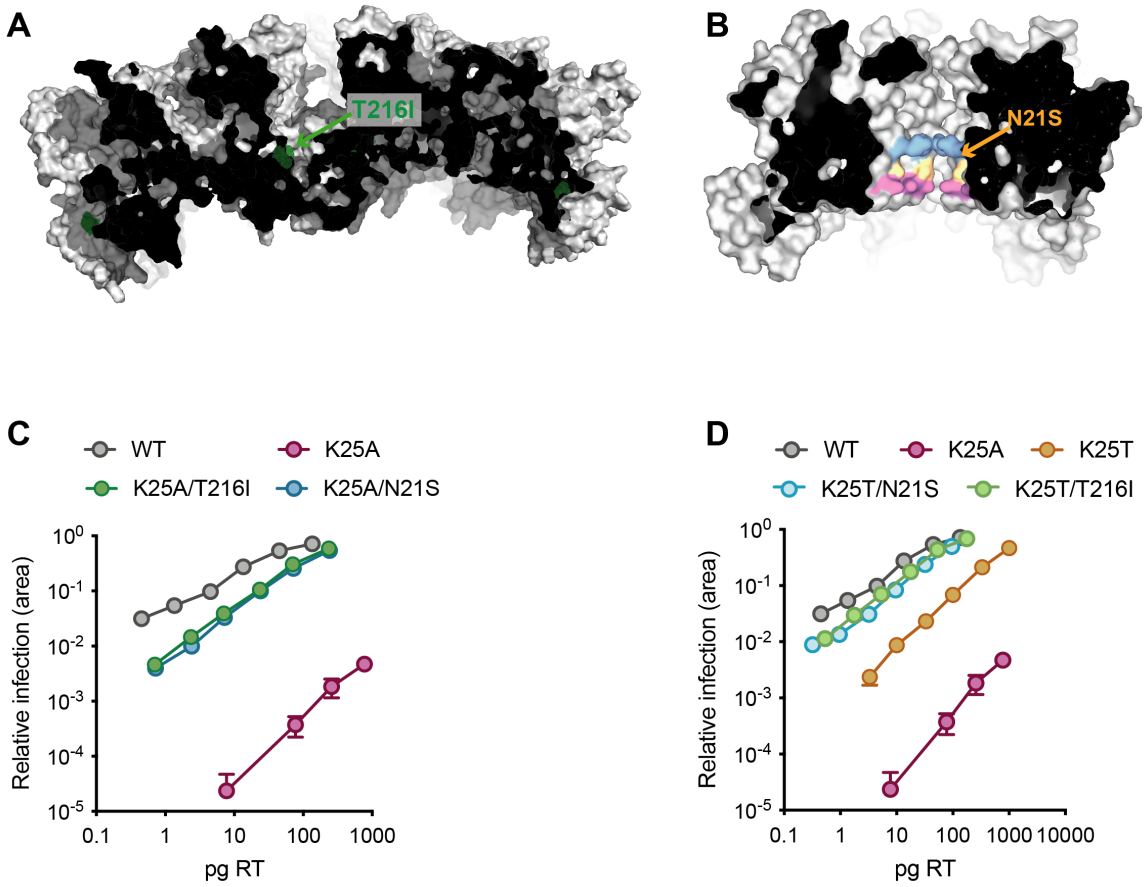


Figure 18: K25 Escape mutants. (A, B) Crystal structures with position of T216 highlighted in green (A) and N21 in orange (B) (PDB: 5HGN[49]). (C, D) Infectivity of WT HIV and selected mutants in HEK293T cells. Infectivity is measured using an Incucyte and determined as the proportion of cell area that is GFP +ve at a given viral dose (in ng RT) after 72h. Error bars in infection experiments depict mean ± SD of three replicates from one experiment representative of three independent experiments. Experiments for panel (C) and (D) were performed by Donna Mallery.

3.2.5 K25 compensatory mutants allow mature core formation

To test the influence of the compensating mutants T216I and N21S on K25A core formation, I purified K25A/N21S and K25A/T216I virions and performed cryo-ET. I

classified the virions as either immature, mature with conical, tubular or irregular capsids or as ambiguous when I could not identify any capsid (**Figure 19**). As described above, most mature WT virions contained conical capsid (46%), whereas for the K25A mutant no mature cores could be observed. Importantly, the addition of either N21S or T216I allowed the formation of conical and tubulure capsids. However, more ambiguous particles (N21S/K25A: 29% K25A/T216I: 21%, WT: 7%) and a higher amount of irregularly shaped cores were present (31% or 29% compared to 21% for WT). This shows that the compensatory mutants N21S and T216I can largely rescue the K25A mature assembly defect, with T216I being slightly more efficient.

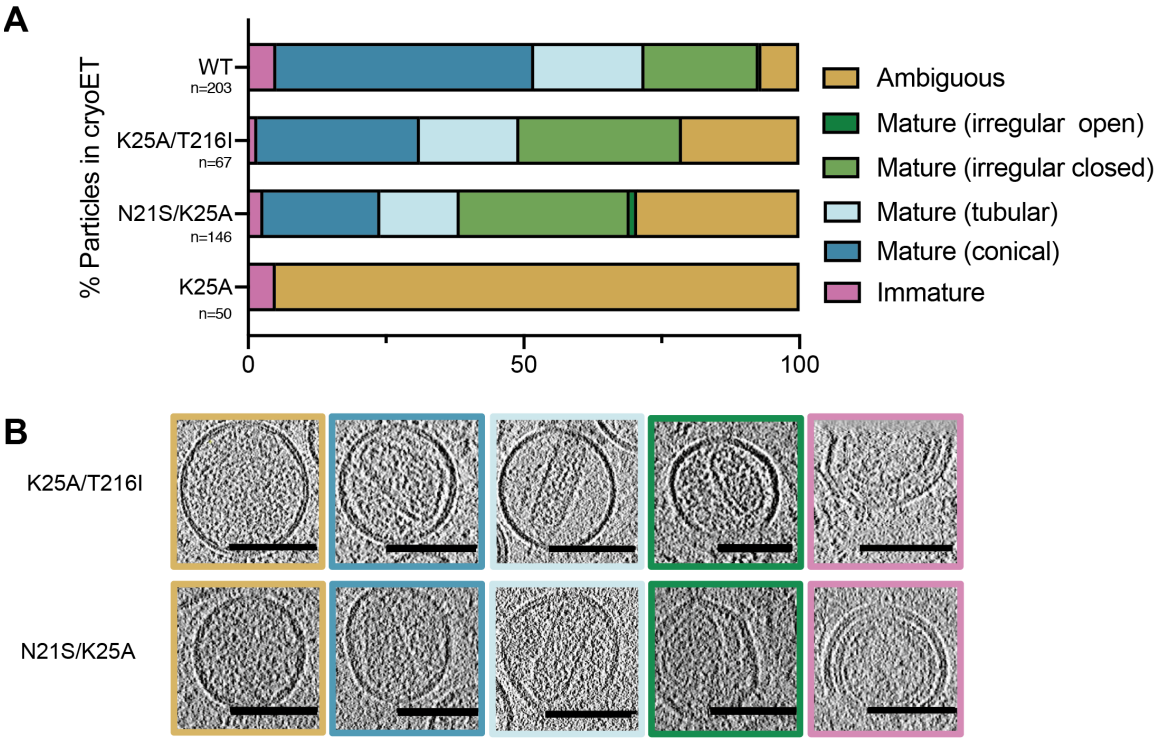


Figure 19: Escape mutants can rescue mature core formation in virions. Cryo-ET analysis of a total of 203 WT, 67 K25A/T216I, 146 N21S/K25A, and 50 K25A virions. (A) Virions were classified as immature (pink), mature with conical (dark blue) or tubular cores (light blue), mature with an irregular but closed structure (green), or ambiguous (partial or no lattices; yellow), and their frequency was plotted as a percentage of all viruses. Representative slices through tomograms of mutant viruses for each category are shown as a gallery in (B), Scale bars: 100 nm.

To test how the compensatory mutants impact assembly, I carried out in vitro capsid assembly experiments with recombinant CA protein and measured the absorbance at 350 nm over time. First, I added 2.5 mM NaCl to WT, K25A, K25A/T216I and N21S/K25A CA protein. K25A/T216I assembled slightly faster than WT, whereas N21S/K25A assembled marginally slower (**Figure 20** A). Most reactions resulted in

the formation of tubes of varying lengths (**Figure 20** C) with exception of N21S/K25A CA that mainly formed spheres. K25T assembled into tubes quicker than WT (**Figure 20** B, C). Similar to the K25A background, K25T/T216I assembled the fastest, whereas N21S/K25T was the slowest. In negative stain micrographs, cone-shaped capsids and a few spheres were visible for K25T/T216I, whereas N21S/K25T mainly formed spheres. This suggested, that N21S promotes pentamer formation at high salt. Indeed, CA carrying N21S alone assembled into spherical particles quickly (**Figure 20** B, C).

Next, I investigated whether the escape mutants could rescue IP6-induced assembly of K25A cores. I chose conditions where WT assembled quickly. As expected K25A did not assemble and neither did N21S/K25A (**Figure 20** D). However, K25A/T216I CA formed cores, albeit with slower kinetics than WT. Increasing capsid and IP6 concentrations allowed N21S/K25A assembly into conical capsids (**Figure 20** E, F). A slight increase in absorbance was visible with K25A, but no proper higher-order structures could be observed via negative stain EM. This shows that T216I and N21S escape mutants rescue IP6 induced formation of K25A capsids, albeit not to WT levels WT. In contrast to K25A, K25T was able to assemble albeit significantly less efficient than WT (**Figure 20** G, H). Like K25T, CA carrying K25T/T216I or N21S/K25T assembled into cores and tubes, with K25T/T216I showing the fastest assembly kinetics (**Figure 20** G, H). N21S assembled efficiently into mostly spheres and occasionally small tubes, further suggesting that N21S enhances pentamer assembly.

Finally, I measured the assembly of N21K mutants (**Appendix Figure 6** B, D). Interestingly, lower capsid concentrations were necessary for IP6 induced assembly of N21K/K25A compared to the N21S and T216I escape mutants, suggesting that a 2nd positively charged ring in the pore increases assembly efficiency in vitro. Summing up, the assembly and cryo-ET data suggest that both T216I and N21S can compensate for the K25A assembly defect and T216I seems to be more efficient.

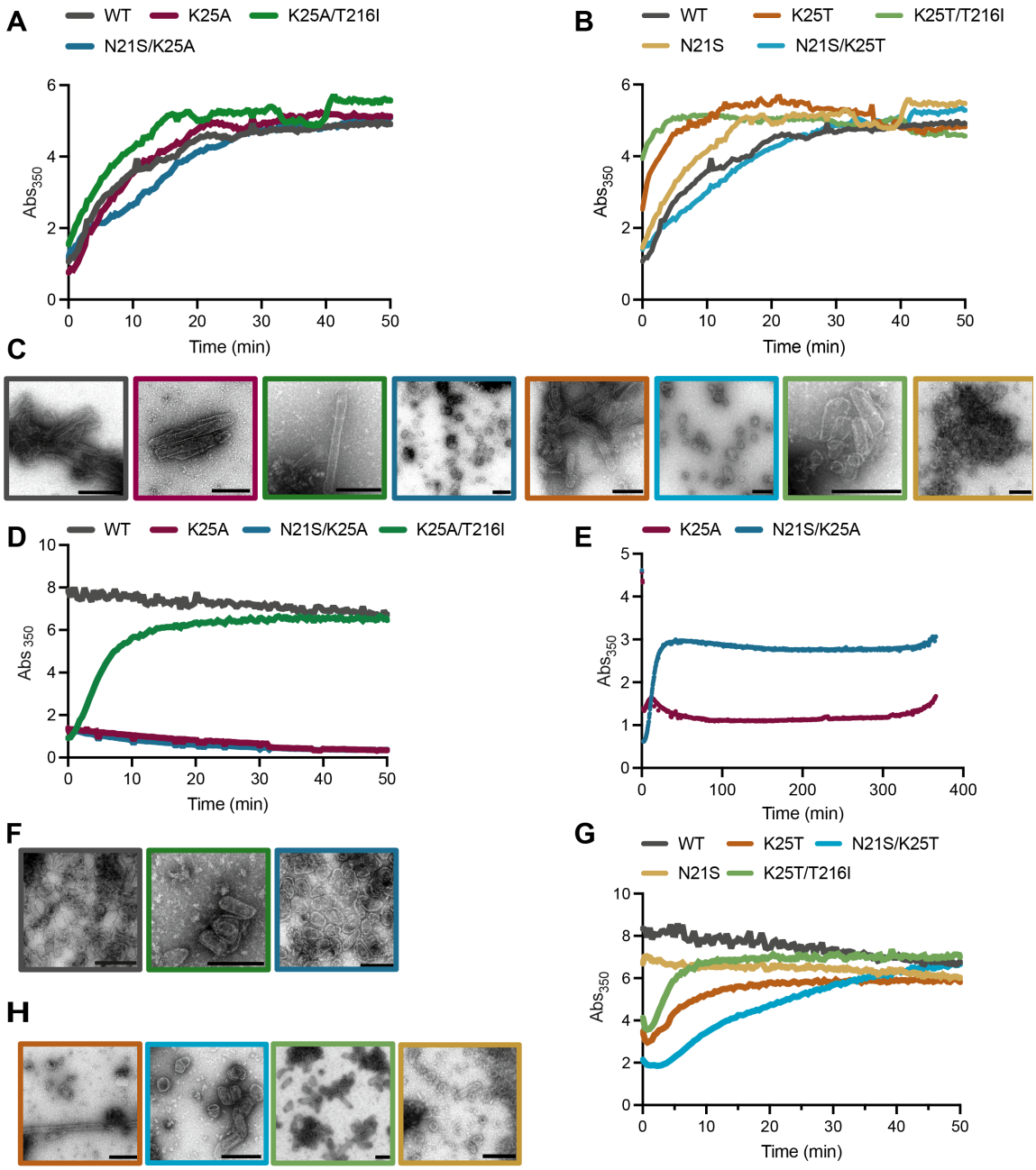


Figure 20: Escape mutants can rescue mature core formation in vitro. In vitro assembly reactions were monitored in real-time by measuring the absorbance at 350 nm. (A, B) Assembly reactions using 75 μM capsid protein and 2.5 M NaCl in 50 mM MES, pH 6.0. (C) Negative stain EM images of material taken from (A, B), Scale bars: 200 nm. (D) Assembly reactions using 200 μM WT and mutant capsid protein with 1.25 mM IP6 and (E) 400 μM capsid protein with 10 mM IP6 in 50 mM MES pH 6.0, 40 mM NaCl. (F) Negative stain EM images of material taken from (D, E). (G) Assembly reactions using 200 μM WT and mutant capsid protein with 1.25 mM IP6 in 50 mM MES pH 6.0, 40 mM NaCl. (H) Negative stain EM images of material taken from (G), Scale bars: 200 nm.

3.2.6 Crystal structures of N21S hexamers

To investigate how compensating mutants may be rescuing assembly by altering packaging within hexamers, I solved crystal structures of N21S/K25T and N21S/K25A

hexamers. N21S/K25T and N21S/K25A hexamers were crystallised in presence of ATP and IP6 **Figure 21**, **Table 3**. The densities corresponding to the ligands are shown in **Figure 21**. As mentioned in the introduction, the six-fold symmetry makes ligand modelling challenging because the ligands can occupy six different positions which are averaged. However, it was apparent that the ligands only seemed to interact with the R18 ring. In addition, no obvious interaction of the serine at position 21 was observed, suggesting that it might not be involved in hexamer formation. Thermal stability measurements of the hexamers showed that the K25A/N21S and K25A/K25T hexamers had a similar melting temperature to WT and were all able to bind to the ligand ATP, IP6 and IP5, and a similar degree of stabilisation was observed as with WT protein (**Appendix Figure 7**).

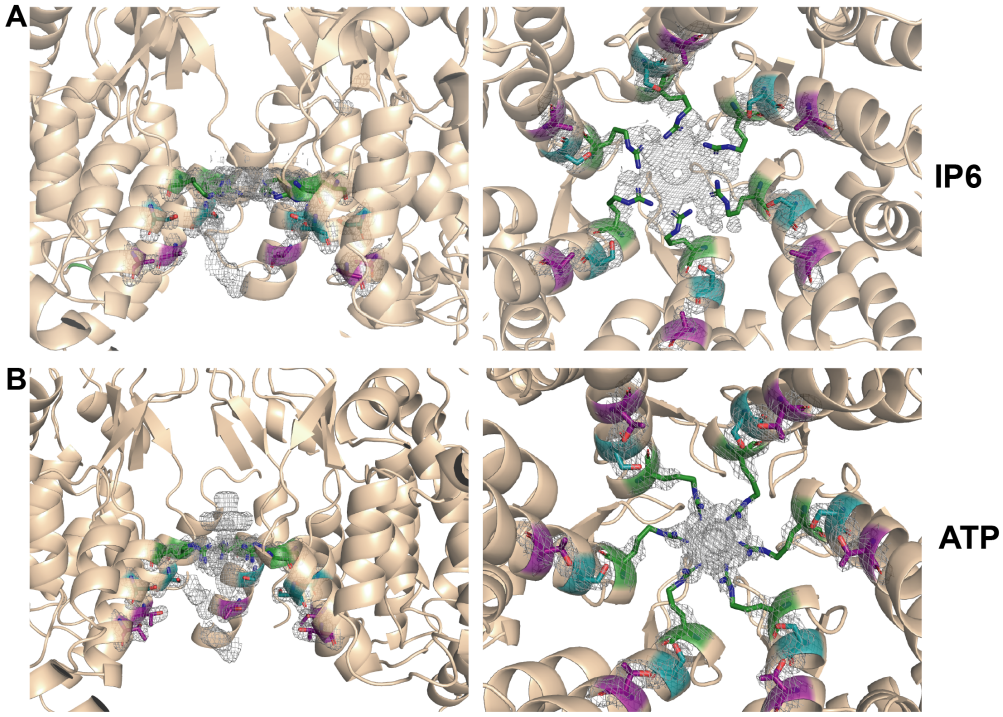


Figure 21: Crystal structures of hexamers with ligands. (A) Crystal structure of N21S/K25T hexamer in complex with IP6. Fo – Fc omit density (mesh) contoured at 1.0σ around the pore show ligand binding. (B) Crystal structure of N21S/K25A hexamer in complex with ATP. Fo – Fc omit density (mesh) contoured at 1.0σ. Residue R18 is coloured green, N21S cyan and K25T/K25A purple.

3.2.7 K25 compensatory mutant capsids are stable and allow reverse transcription

To assess capsid stability, TRIM5 abrogation assays were carried out by infecting cells expressing rhesus macaque TRIM5. Co-infection of increasing K25A/N21S or

K25A/T216I RFP virus with a single dose of WT GFP virus resulted in a corresponding increase in GFP signal, suggesting that these virions can form stable capsid in cells (**Figure 22** A). Finally, the ability of the K25A escape mutants N21S and T216I to reverse transcribe was investigated. In cells infected with K25A virions very little reverse transcription were measured after 8h. In contrast, K25A/N21S and K25A/T216I virions were able to reverse transcribe, albeit less efficiently than WT (**Figure 22** B).

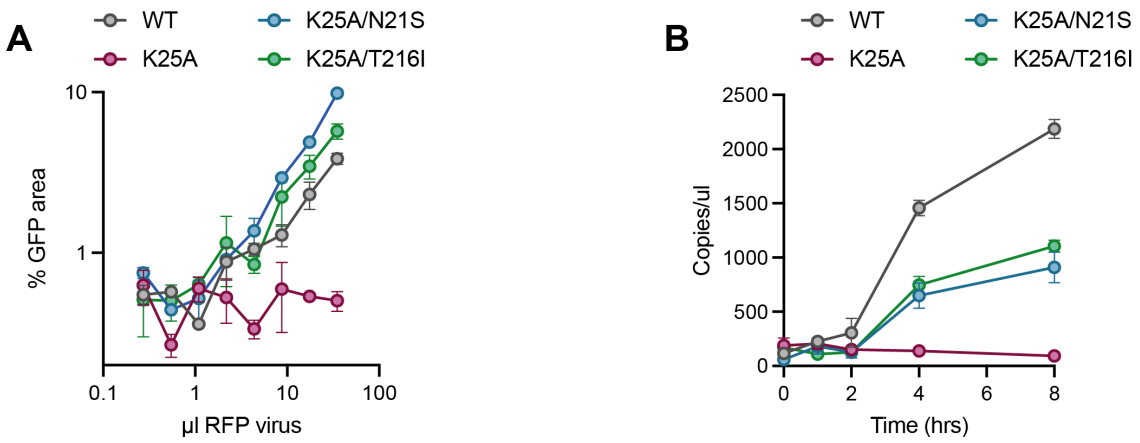


Figure 22: Escape mutants form stable capsids and can reverse transcribe. (A) TRIM5 abrogation assay in FRhK-4 cells. Cells expressing rhesus TRIM5 were infected with a consistent dose of WT GFP-expressing virus and increasing doses of RFP-expressing WT and mutant viruses. The level of infection of each virus was determined as the proportion of cell area that was GFP positive. Increasing doses of WT RFP virus results in an increased GFP positive cells if TRIM5 becomes saturated with RFP virus and no longer restricts infection by the GFP virus. Error bars depict mean ± SD of three replicates from one experiment representative of two independent experiments. (B) Amount of reverse transcription products (GFP products) in Hela cells at the indicated time points. Error bars depict mean ± SD of three replicates from one experiment representative of two independent experiments Experiments for this Figure were performed by Donna Mallery.

3.3 Discussion

3.3.1 Role R18 in the mature capsid

HIV capsids contain positively charged pores that can bind multiple polyanions, including IP6 and dNTPs[49,190,191,229]. Mutation of either K25 or R18 results in a dramatic loss in infectivity. R18 was identified as the primary IP6 and dNTP binding site, and mutation of R18 was accompanied by strongly decreased reverse transcription[49,190,191,229]. Consistent with previous reports, my cryo-ET data showed that mutation of R18 allows the formation of mature capsids in virions albeit with irregular shapes[304,305]. Removal of the positive R18 charges promoted sphere

formation in vitro, possibly due to an increased frequency of pentamer incorporation[304]. Packing in the pentamer pore is tighter, thus, electrostatic repulsion of the R18s is likely more destabilising in pentamers. Indeed, a recent cryo-ET study suggested that the IP6 density above the R18 ring seemed to be more pronounced in pentamers compared to hexamers, which was attributed to tighter binding[306]. As a result, formation of hexameric tubes is preferred in the absence of IP6, whereas removal of R18 favours pentamer formation. The altered incorporation of pentamers into forming cores can also explain the irregularly shaped capsids observed in tomograms of R18G virions. Moreover, these capsids are highly unstable and likely fall apart quickly once the capsid is released into the cytosol. The remaining R18G capsid fragments still appear to saturate TRIM5 to some extent, but they cannot carry out efficient reverse transcription resulting in unproductive infection. This highlights the importance of IP6 binding for capsid stability, but it makes it impossible to determine whether R18 is involved in dNTP recruitment and import. However, the fact that no compensatory R18G mutants were discovered so far suggests that the decreased infection might not only be caused by reduced stability. More research into possible R18G rescue mutants is necessary to answer these questions.

3.3.2 Role of K25 in the mature capsid

Mutating residue K25 provides another avenue to investigate the interplay between IP6 and dNTPs. It was proposed that K25 is necessary for the directional import of nucleotides after they are recruited by R18[300,301]. In line with this, mutation of K25 results in loss of reverse transcription both in purified capsids and in cells. However, cryo-tomography on HIV virions showed that mature capsid formation is strongly impaired in K25A mutants as none could be found in my data. K25A capsids were highly unstable in TRIM5 abrogation studies and single-molecule TIRF assays, and the addition of IP6 failed to rescue stability in the latter. In vitro assembly experiments demonstrated that K25 is involved in IP6 induced capsid assembly. K25A, K25R and K25N were capable of forming capsid tubes in the presence of high salt with similar kinetics and morphology as WT but IP6-induced formation of cores was lost in K25A and significantly reduced in K25R and K25N. Thus, K25 mutants appear to be distinct from other mutations affecting capsid stability. For example, the instability mutant

P38A had a similarly reduced infectivity as K25R but was able to form cores in IP6 conditions efficiently. Although IP6 binding to K25 was observed in crystal structures, no loss of IP6-mediated stabilisation was observed upon mutating K25 in NanoDSF experiments with cross-linked hexamers, in contrast to R18G mutation. This confirms that R18 is the primary nucleotide-binding site and suggest that IP6 does not bind K25 in cross-linked hexamers or that binding does not stabilise them. Taken together, these data corroborate that K25 is involved in the IP6-driven assembly of mature capsid cores either by recruiting a 2nd IP6 or by maintaining IP6 in the capsids, possibly by holding to or passing the molecule between the rings. This likely explains why no K25A mature capsids were observed in virions.

The K25A compensatory mutants N21S, N21K, K25T and T216I could largely rescue infection and core formation in cells and in vitro. Importantly, these mutants also allowed reverse transcription, suggesting that K25 does not seem to be essential for nucleotide import. However, this does not explain why K25 mutation interferes with assembly. K25:IP6 interactions might be necessary to stabilise pentamers and thus allow the formation of conical cores rather than just tubes. If K25 were only a destabilising force then K25A should form pentamers more efficiently similar to R18G, nevertheless, the fact that K25A assembles tubes but not cones in vitro indicates that K25 is participating in an interaction that actively stabilises pentamers. Interestingly, MD simulations also suggested a possible second IP6 binding site comprised of the K25/E28/E29/K30 and R162 in pentamers[306]. However, E29A is as infectious as WT virus[302] and a positive charge at position K25 does not appear to be required for core formation since the mutants N21S, K25T and T216I do not restore the K25 charge but nonetheless allow capsid assembly. Moreover, although N21K mutants assembled most efficiently in vitro, they did not rescue infection to the same extent as the T216I and N21S mutants. The latter compensatory mutants appear to allow core formation through different mechanisms. N21S sphere formation in vitro indicates an increased propensity to form pentamers. Since sidechains cannot be identified in cryo-EM pentamer structures, it is not obvious how this stabilisation occurs. Therefore, I modelled the possible side-chain arrangement within a pentamer based on an EM structure of HIV-1 capsid pentamers in intact virions (5MCY[307]) Within the pentamer N21 is in close proximity to M39 (**Figure 23**) and steric repulsions might disfavour

pentamer formation. However, mutation to the smaller serine (N21S) might allow easier accommodation of the residues in the pentamer. This could explain, why N21L was shown to have an assembly defect in turbidity assays, but N21A appeared to assemble with similar efficiency as WT[304,308]. Thus, adding the pentamer-favouring N21S mutant to the hexamer-favouring K25A could restore conical core formation. Interestingly, the feline immunodeficiency virus (FIV) capsid looks similar to the N21S mutant as it contains a K18, S20 and E25 motif in the pore (**Appendix Figure 8**) and also forms spheres in high salt conditions but conical cores in virions[309,310].

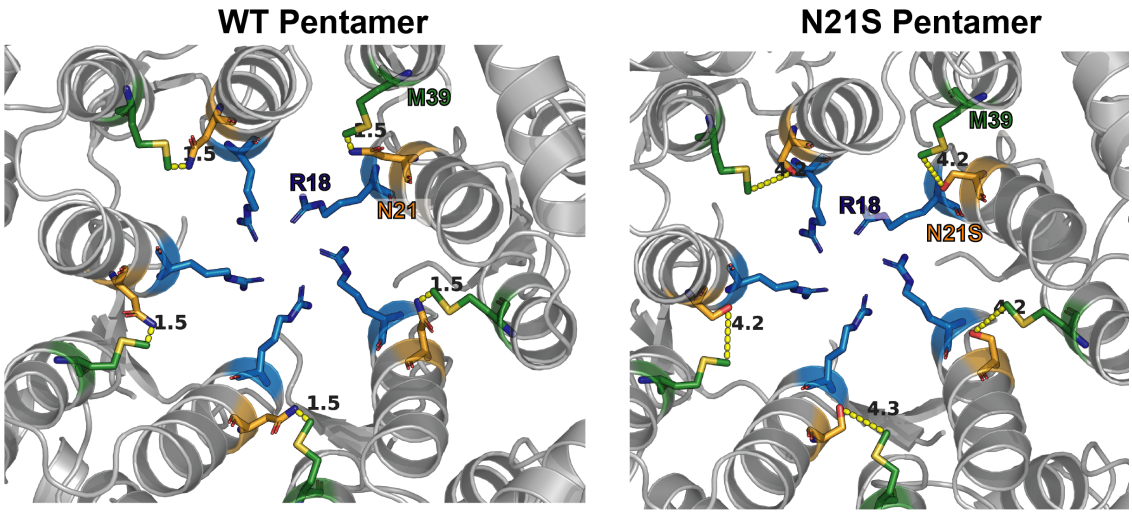


Figure 23: Model of WT and N21S pentamer. Model of a CA pentamer, generated using monomers from the hexamer in **Figure 10** A and the pentamer cryo-EM structure 5MCY[311].

T216I was previously shown to partially restore infection, replication, reverse transcription and stability of the P38A mutant. It was speculated that T216I might stabilise the CTD-CTD interface to compensate for structural changes induced by P38A[209]. In addition, CA tube stability of P38A/T216I was increased compared to WT in vitro and seemed to slow down core disassembly of P38A[312]. Since IP6 binding and hexamer stability appear to be interdependent, the increased stability introduced by T216I will most likely allow IP6 binding to R18 and thus allow capsid formation.

While the details of how K25 participates in assembly are of interest, perhaps the more important question is why HIV has evolved this particular mechanism. The data above suggest that IP6-stabilisation at position K25, whether in hexamers or pentamers, is not a requirement for assembly. In addition, it does not seem to be necessary for reverse transcription. However, the escape mutants are not as infectious as WT,

indicating that K25 plays a role in allowing optimal capsid stability to allow ordered uncoating and reverse transcription. Interestingly, FIV appears to be less IP6 dependent, suggesting that changes in the pore influence IP6 dependency[313]. Thus, as discussed in the outlook, the IP6 dependency of different retroviruses might help to answer these questions.

Chapter 4 Maturation Inhibitors

4.1 Introduction

Full proteolytic cleavage of the immature lattice is crucial to produce a mature infectious virion[53,200]. The final cleavage occurs between CA and SP1 and shifts the viral lattice from immature to mature (**Figure 9**)[188,201]. The CA-SP1 cleavage site is buried within the six-helix bundle (called 6HB) and unfolding of the 6HB is necessary for PR-cleavage[314]. This is exploited by a class of drugs called Maturation inhibitors (MIs). MIs stabilise the immature lattice and slow down CA-SP1 proteolytic cleavage. The resulting altered Gag processing was shown to reduce infectivity[315-317].

HIV can acquire resistance mutations for MIs that increase CA-SP1 processing by decreasing the stability of the immature lattice[318,319]. These escape mutants differ for different MIs. For example, BVM selects for mutations around K227[316,318], whereas for the structurally distinct MI, PF96, mutations occur around K158[320,321]. The BVM analogues 7m and 7r select for mutations around both the K158 and K227 lysine rings[319]. Interestingly, these are the residues that package IP6 in the immature lattice: K158 engages the 5-equatorial phosphates while K227 interacts with the axial phosphate[190]. Removal of the lysine rings decreased virus production and infectivity and mutation of K158 has a more pronounced effect than K227[229]. In line with that, MI escape mutants around K158 are generally less replication-competent than their K227 counterparts[318,321].

The correlation between mutations that confer escape from MIs and their proximity to IP6 binding residues is striking. A structure of the immature hexamer in complex with BVM obtained by microelectron diffraction suggested that the IP6-coordinating residue K227 may be directly involved in BVM binding (**Figure 24** A)[58]. This predicted that IP6 and MIs would bind competitively. Cellular data showed the opposite, and when viruses were produced in cells produced in presence of the MIs PF96, BVM, and the BVM analogue 7m, no reduction in IP6 incorporation was observed (**Figure 24** B). Moreover, the MI PF96 restored the infectivity of the IP6 deficient mutant K158A

(**Figure 24** C) and increased IP6 incorporation (**Figure 24** D). Thus, the MI seems to be needed to compensate for the lattice-destabilizing effect of K158A.

Upon passaging IP6-binding-deficient mutants, compensatory mutations that stabilise the lattice were obtained. The mutant T8I is located in the SP1 region and was previously shown to stabilise the CA-SP1 six-helix bundle[318,321]. T8I on top of the escape mutants was able to restore virus production and infection (**Figure 24** E and F) as well as IP6 incorporation (**Figure 24** G).

Figures in this section were published in [322] and adapted for this thesis.

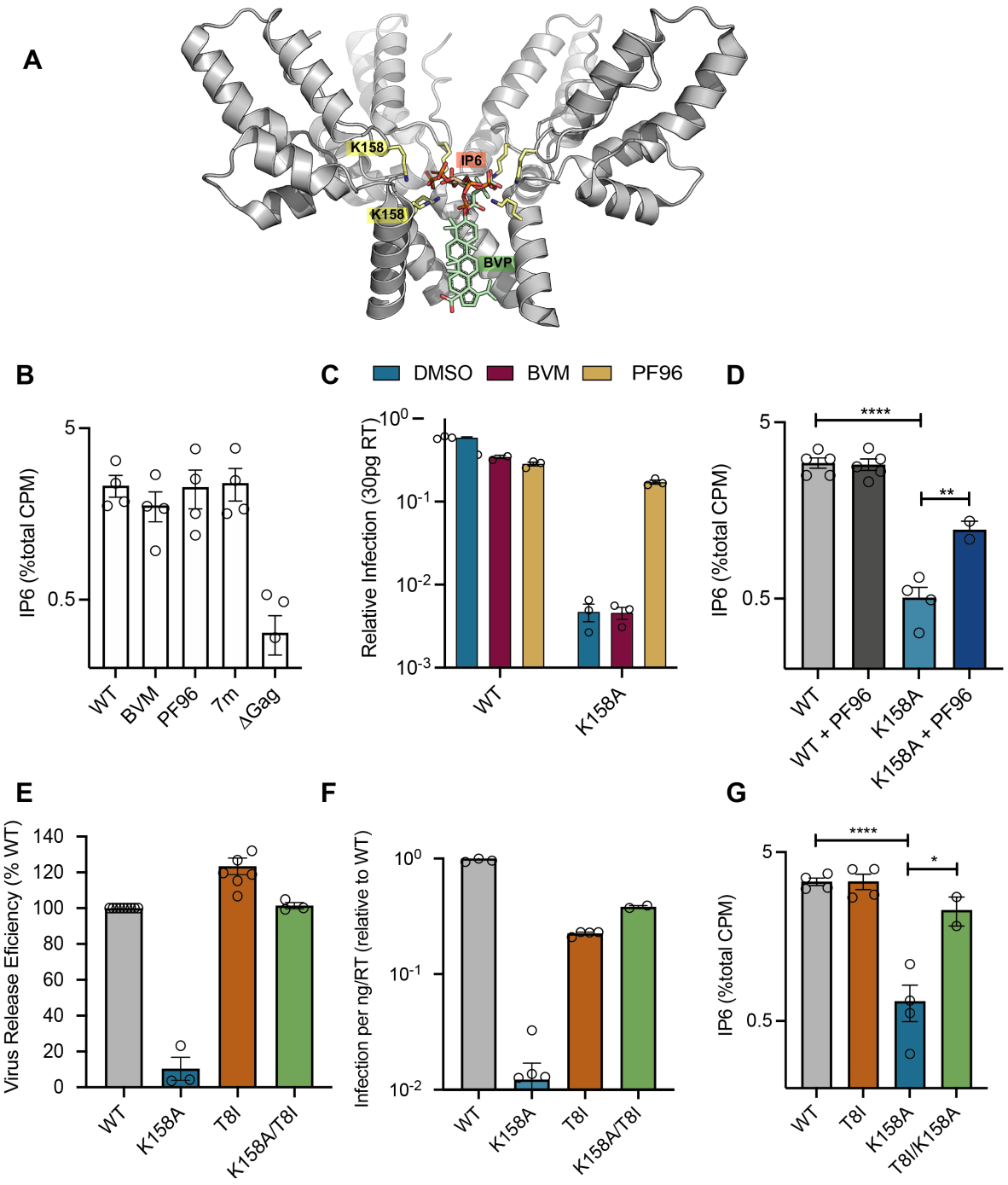


Figure 24: Relationship between immature Gag lattice stability, MI sensitivity and IP6 incorporation. (A) Model of an immature hexamer showing the binding sites of IP6 and BVM. Based on 6BHR[190] and 6N3U[323]. (B) IP6 incorporation of purified virions produced in the presence of the indicate MIs and 3H inositol expressed as % total counts per minutes (CPM) of ^{3}H inositol–containing species quantified by scintillation counting. Data are presented as mean ± standard error of the mean from a minimum of three independent experiments. (C) Infectivity of WT and K158A viruses produced in presence of 5μM BVM or PF96 in HEK293T cells. Data are normalized to quantity of input virus (30pg RT). Error bars depict mean ± SEM of at least two replicates from three independent experiments. (D) IP6 in cooperation in WT and K158A virions as a % of total CPM in the presence or absence of MI PF96. Error bars depict mean %CPM ± SEM from at least two independent experiments. Statistics were performed using the students t-test (WT versus K158A, ****P < 0.0001; K158A versus K158A/T8I, *P =

0.011). (E) Virus release efficiency of the K158A and K158A/T8I Gag mutants calculated based on the level of p24 in purified virions expressed as a percentage of p24 and Pr55 expression in HEK293T producer cells and p24 expression in purified virions. Data are normalized to give the efficiency relative to WT virus. Error bars depict the SEM from three independent experiments. (E) Infectivity of the mutants shown in (D). (G) IP6 incorporation as % total CPM in WT, T8I, K158A or K158A/T8I viruses. Error bars depict IP6 as % total CPM ± SEM from at least two independent experiments. Statistical analysis was performed using the students t-test. Experiments shown in panel (B-G) were performed by Donna Mallery and Alex Kleinpeter (E-F).

4.1.1 Aims

The aim of this chapter was to investigate the impact of the escape mutants and the T8I compensatory mutants on virus morphology and assembly behaviour. I used cryo-ET to assess the phenotype of K158A virions alone and in combination with the compensatory mutant T8I. In addition, I used an in vitro assembly assay of recombinant CA-SP1 protein, to investigate the differences in assembly and IP6 binding of those mutants.

Tomogram collection was performed by Zunlong Ke and CA-SP1 protein purification was performed by Leo Kiss according to my instructions.

4.2 Results

4.2.1 Assembly of K158A and K158A/T8I VLPs

It was previously shown that IP6 cannot promote the immature lattice formation of lysine mutant K158A in vitro[190]. Therefore, I wanted to test if adding the second site mutant T8I restores assembly of recombinant CA-SP1 protein to VLPs in vitro. First, I compared assembly kinetics of WT and T8I in presence of IP6. T8I assembled significantly faster than WT CA-SP1 (**Figure 25** A), while both formed immature VLPs with similar morphology as assessed by negative stain EM (**Figure 25** E). Titration of IP6 with WT and T8I CA-SP1 revealed that T8I CA-SP1 assembled efficiently even at low IP6 concentrations (125 μM). In contrast, WT CA-SP1 assembly was reliant on higher IP6 concentrations (**Figure 25** B). Strikingly, T8I assembled faster at 125μM IP6 than WT at 1 mM. Next, I tested whether T8I could restore K158A CA-SP1 VLP assembly (**Figure 25** C). However, no assembly was observed, suggesting that in the in vitro system, T8I can only enhance assembly if the IP6 binding site K158A is present. Since T8I and PF96 both lead to enrichment of IP6 in K158A mutant virions in cell experiments, I tested whether adding all these components would allow K158A

assembly (**Figure 25** D). Remarkably, the addition of both PF96 and IP6 allowed K158A/T8I to assemble. Negative stain EM confirmed that immature VLPs were formed (**Figure 25** F). This supports that IP6 binding to K158A can be restored indirectly by increasing the stability of the immature lattice. It also shows that PF96 and IP6 can act synergistically to increase immature VLP assembly of WT CA-SP1, consistent with the IP6 incorporation data suggesting that IP6 and PF96 are not competitive (**Figure 24** D).

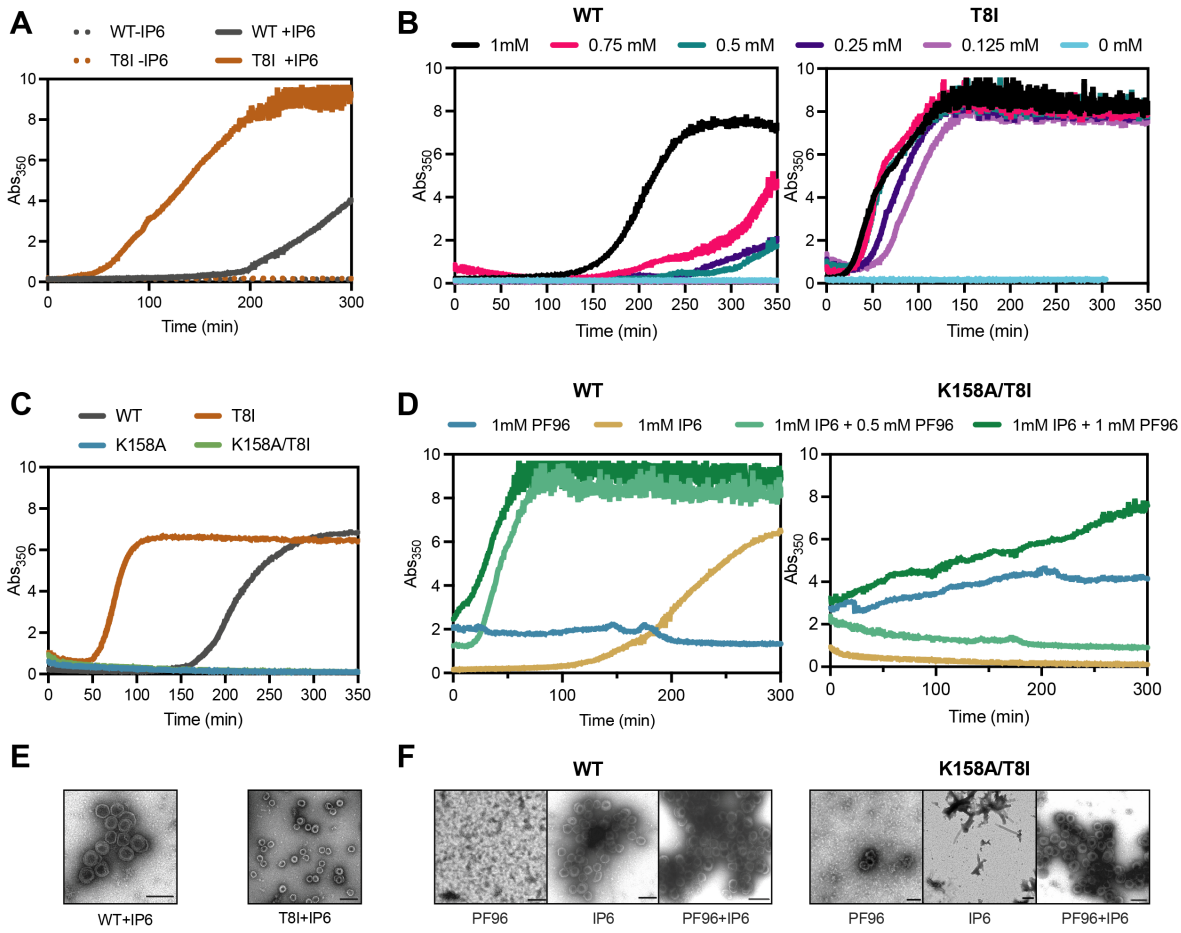


Figure 25: Rescue of IP6 binding–deficient mutant K158A restores the assembly of single capsids. In vitro assembly reactions using 250 µM WT and mutant CA-SP1 recombinant protein in 50 mM TRIS-HCl (pH 8.0), 50 mM NaCl, 1 mM dithiothreitol (DTT) were monitored in real time by measuring the absorbance at 350 nm (Abs350). (A) Assembly of WT and T8I CA-SP1 into VLPs in the presence and absence of 1 mM IP6. (B) Assembly of WT and T8I with indicated concentration of IP6. (C) Assembly of WT, K158A, T8I and T8I/K158A with 1mM IP6. (D) Assembly of WT and K158A/T8I with 1 mM IP6 and/or 0.5 or 1 mM PF96. (E) Negative stain images of material taken from (A), Scale bars: 200 nm. (F) Negative stain EM images of material taken from (D), Scale bars: 200 nm.

4.2.2 Second site compensatory mutant rescues mature core formation

To test the influence of T8I on mature capsid formation, I carried out a series of cryo-electron tomography (Cryo-ET) experiments. I purified WT, T8I, K158A, and K158A/T8I viruses and prepared grids. Tomograms were collected from multiple grids per sample. I used this data set for 3D-tomographic reconstructions on 60-130 virions per mutant. Then, I divided the virions into the categories: immature, mature with a conical or tubular capsid, mature with an irregular capsid, and ambiguous, when no clear lattice was observable (**Figure 26** A, **Appendix Figure 9**, **Appendix Figure 10**, **Appendix Figure 11**). Most WT virions were mature, with only ~ 5 % displaying an immature lattice. Most WT cores were conical or tubular, while in a minor fraction (28%) cores had an irregular shape. A larger fraction of immature virions (13%) was observed for the T8I mutant. This was expected because T8I was previously shown to stabilise the six-helix bundle and thus slows down CA-SP1 proteolytic processing[321,324,325]. K158A/T8I immature virions were also increased compared to K158A alone (7% and 2% respectively). Interestingly, K158A and the double mutant K158A/T8I showed fewer conical or tubular capsids than WT (43% and 34% respectively, vs 52% for WT). However, this does not quantitively account for the infection defect in K158A or the rescue observed with K158A/T8I. To gain deeper insight into the mature virions observed by cryo-ET, I subdivided them into five categories – single, multiple and multi-layered cores, and cores with additional structures that were either open or closed (**Figure 26** B). Mature WT viruses contained a single capsid core in most cases (88%), and only a small fraction (11%) of additional structures were observed. Strikingly, most K158A virions contained other structures such as multiple cores, multi-layered cores, or cores with additional open or closed structures (**Figure 26**, **Appendix Figure 9**). This suggests that K158A leads to a capsid over-assembly phenotype. Importantly, when T8I was added to K158A the number of virions with additional structures decreased, and the majority of virions (64%) had a single capsid (**Figure 26**, **Appendix Figure 10**). Thus, adding T8I seems to counteract the K158A over-assembly phenotype. This might be due to increased IP6 levels in the double mutant virions since T8I rescues IP6 incorporation into the immature lattice. IP6 may be required to ensure proper conical core formation, as has been demonstrated in vitro[190]. In addition, many of the K158A virions showed a density

outside of cores that may represent unpackaged RNA genome (**Appendix Figure 12, Appendix Figure 2**). A similar phenomenon of condensed RNP located outside empty core structures has been reported for HIV virions that have undergone complete but delayed CA processing[88]. Taken together, the data presented here show that K158A mutation impacts mature core formation in virions but can be rescued by the addition of T8I.

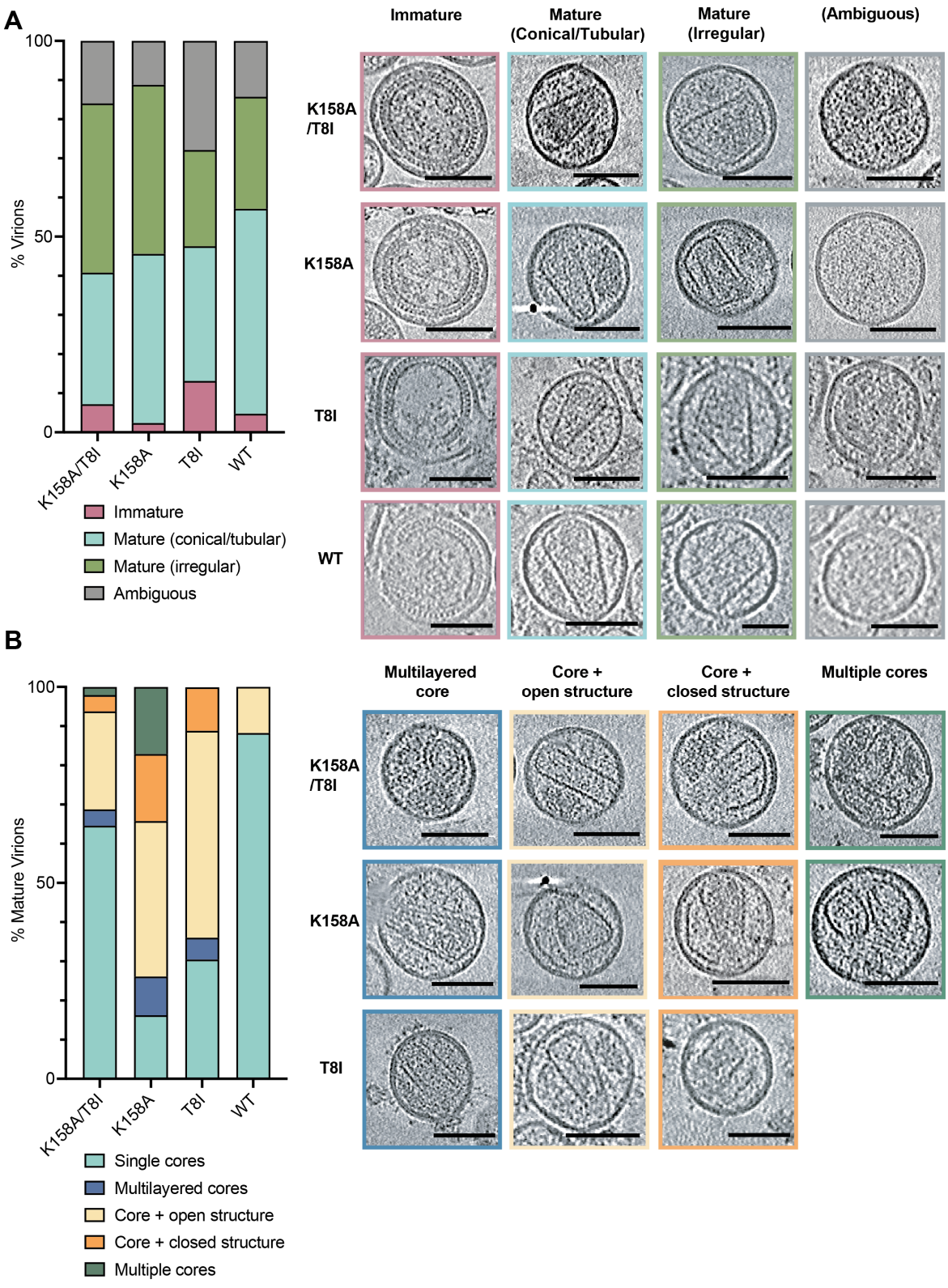


Figure 26: T8I restores the K158A deficiency in assembling single capsids. (A) Cryo-ET analysis of WT, K158A, K158A/T8i and T8I virion morphologies. Capsids were classified into the indicated categories. A total of 64 WT, 69 T8I, 127 K158A, and 119 K158A/T8I tomograms was analysed, and the percentage of each category was calculated. Virions were classified as immature (pink), mature with conical or tubular cores (cyan), mature with an irregular but closed structure (green), or ambiguous (partial or no lattices; grey), and their frequency was plotted as a percentage of all viruses.

Representative slices through tomograms of WT and mutant viruses for each category are shown as a gallery. (B) Virions with mature lattices were further subdivided into single cores (light blue), multi-layered cores (dark blue), cores with additional open structures (light orange), cores with additional attached closed structures (dark orange), or multiple cores (dark green), and their frequency was plotted as a percentage of mature viruses. Representative tomograms of mutant viruses for each category are shown as a gallery, Scale bars: 100 nm.

4.3 Discussion

IP6 binding is important for immature hexamer formation, and thus for efficient HIV-1 infection. IP6 is recruited into assembling virions via two lysine rings (K158 and K227) within immature Gag hexamers, and mutation of the K158 ring abolishes immature VLP formation in vitro. Introducing additional stability in the immature lattice, for instance by the stability mutant T8I or by adding MIs like PF96, can restore infectivity of K158A mutant virions. In vitro assembly showed that T8I assembled remarkably faster in vitro than WT (**Figure 25**), suggesting it is more efficient in enriching IP6. However, K158A/T8I did not assemble in vitro, probably because the assay used does not recapitulate all components of immature assembly, for instance, the interaction between the nucleocapsid (NC) domain of Gag and the viral RNA. Remarkably, the presence of PF96 allowed K158A/T8I assembly. This confirms that K158A assembly can be restored by increasing the stability of the immature lattice. Moreover, it also shows that PF96 and IP6 can act cooperatively to increase immature VLP assembly of WT CA-SP1, similar to the reported effect of BVM[326]. It also explains why the MI can restore both infectivity and IP6 incorporation into the K158A virus.

Analysis of cores via cryo-ET revealed many additional structures in K158A virions that were largely missing in K158A/T8I virions. This is in line with previously published thin-section EM data where K158A virions contained irregularly shaped cores and multiple cores[305]. This suggests that IP6 may be required to ensure the formation of a single conical core, as has been demonstrated in vitro[190]. In line with this, K158A cores were shown to be highly unstable but the addition of T8I increased capsid stability most likely by restoring IP6 binding[322]. However, it is worth mentioning that K158A capsid protein assembled into cylinders in high salt, but in contrast to WT, the K158A cylinders tended to aggregate in vitro and it was speculated that the removal of K158

might introduce a new hydrophobic surface enabling aggregation[304]. This may contribute to the over-assembly phenotype observed in my tomograms.

Taken together, the data presented in this chapter suggests that because the K158 mutant packages less IP6 and is therefore less stable than wild-type, it can escape MIs, which can otherwise make immature hexamers hyperstable and defective for PR-mediated cleavage. The reduced capsid stability of the lysine mutants also explains why PF96 promotes their infectivity; it restores hexamer stability and IP6 binding. The second-site mutations behave like MIs, stabilising the immature hexamers and restoring infectivity and IP6 incorporation.

Chapter 5 IP6 dependency of immature and mature capsid

5.1 Introduction

In the previous chapters, I have established the importance of IP6 for producing infectious viral particles. IP6 is recruited into virions during immature assembly by two lysine rings (K158 and K227) and promotes VLP assembly[190,191,322]. Mutation of either K158 or K227 prevents IP6 enrichment, decreases virion production and results in a profound loss of infectivity[191,322]. In the mature capsid, one or two IP6s are coordinated by two charged rings (R18 & K25), and these residues were shown to be essential for productive infection[229,271,327]. HIV's dependence on IP6 can also be investigated by interfering with the inositol phosphate metabolism for instance by knocking out the IP6 generating kinases IPMK or IPPK or overexpressing the phosphatase MINPP1 (the inositol phosphate metabolism pathway is shown in **Figure 7**). To understand the effect of these manipulations, the cellular IP profile in KOs +/- MINPP1 overexpression were measured. KO of IPMK reduced IP6 levels by ~ 10-fold compared to WT cells, whereas KO of IPPK reduced IP6 levels > 50-fold, while IP5 was unaffected (**Figure 27** A). Overexpression of full-length MINPP1 (MINPP1 FM1) resulted in further reduction of IP5 and IP6 in both KOs. Interestingly, the amount of IP4 was significantly higher in IPPK KO + MINPP1 than in the IPMK KO cells, while IP6 levels were > 5-times lower (**Figure 27** A). In line with previous publications, KO of IPMK or IPPK decreased viral production[191,192,313] and further reduction of IP6 levels by overexpressing MINPP1 nearly abolished viral production[313], especially when MINPP1 was localised to the plasma membrane (MINPP1 PM1) (**Figure 27** B). Virions produced in IPPK or IPMK KOs showed no or only a modestly reduced infectivity. However, infectivity was reduced substantially when virions were produced in KOs plus MINPP1 overexpression. Similar to what has been observed for viral production, this effect was exacerbated by targeting MINPP1 to the plasma membrane in the IPPK KO background (**Figure 27** C, D). The production and infectivity data suggest that HIV-1 actively recruits IP6 at the plasma membrane during immature Gag lattice assembly

and viral budding. Reduced IP6 incorporation, as found in the IPMK KO knockout, combined with MINPP1 has a dramatic effect on Gag processing leading to an accumulation of uncleaved Pr55 Gag (**Figure 27** E).

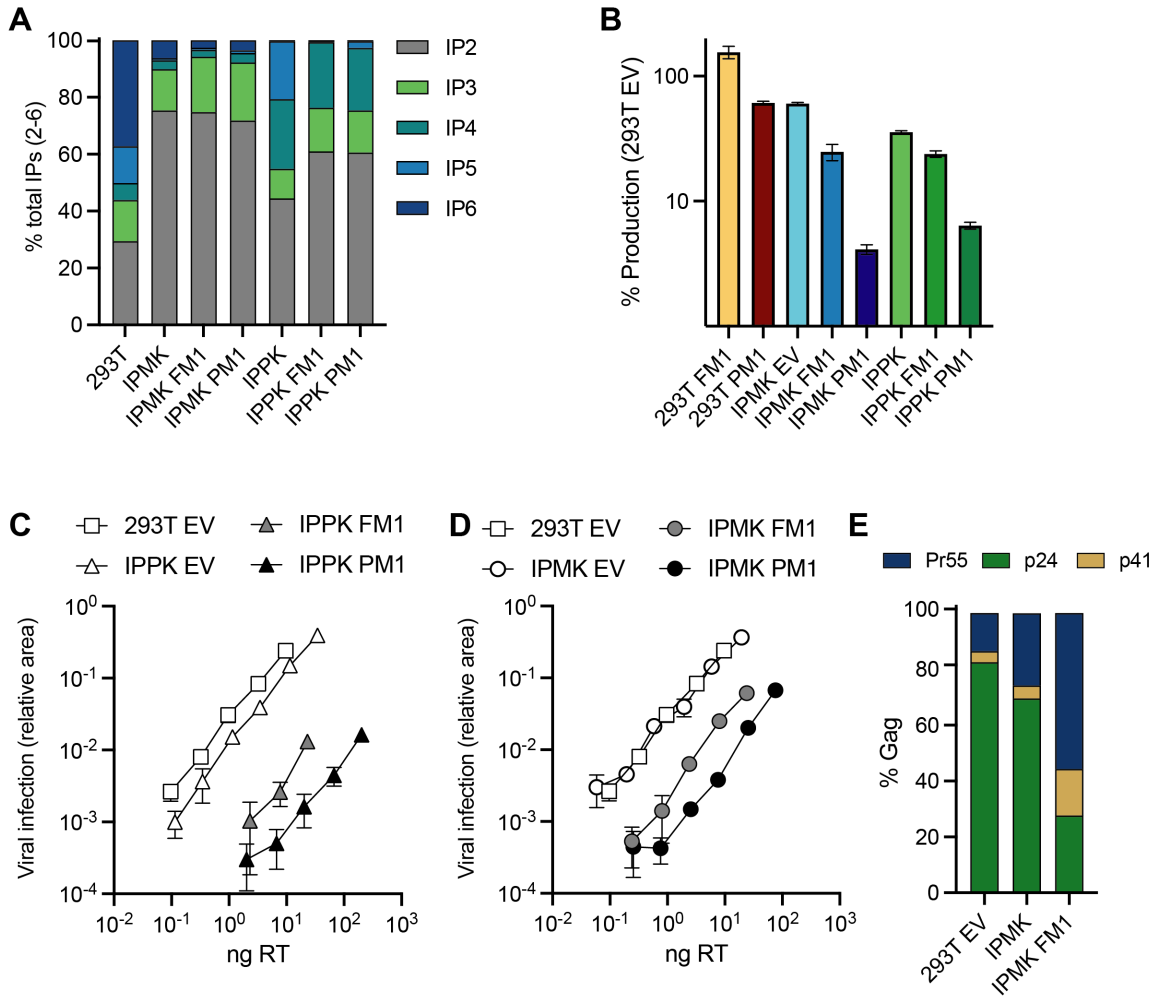


Figure 27: Altering IP composition in producer cells changes HIV-1 particle production, Gag processing and infectivity. (A) 293Ts, CRISPR KOs of IPMK or IPPK[191], or KOs over-expressing either full-length (FM1) or plasma-membrane targeted (PM1) MINPP1 were grown in tritiated inositol. IP species were extracted and separated by SAX-HPLC in HEK293T cells. The proportion of each IP species as a fraction of the total IP2-IP6 concentration. (B) HIV-1 production, as measured by RT activity, expressed as a percentage of virus production in HEK293T cells. (C-D) Infectivity of viruses from (B) plotted against quantity of input virus. (E) Gag cleavage efficiency of WT purified virions, calculated as the relative amount of p24 (CA), p41 and Pr55Gag as a percentage of total Gag. Error bars depict SEM from three experiments. Experiments in this figure were performed by Donna Mallery.

IP6 incorporation into the virion can also be reduced by mutating both lysine rings in the immature lattice simultaneously (K158A/K227A or KAKA). This has a smaller impact on production than mutation of the individual lysine residues (**Figure 28** A). However, KAKA mutants displayed markedly reduced infectivity, with a ~100-fold lower infectivity than WT and therefore lie between K158A and K227A. Interestingly,

in contrast to the single mutants, addition of the CA-SP1 stability mutant T8I does not rescue infectivity (**Figure 28** B) and the KAKA mutant failed to incorporate IP6 (**Figure 28** C).

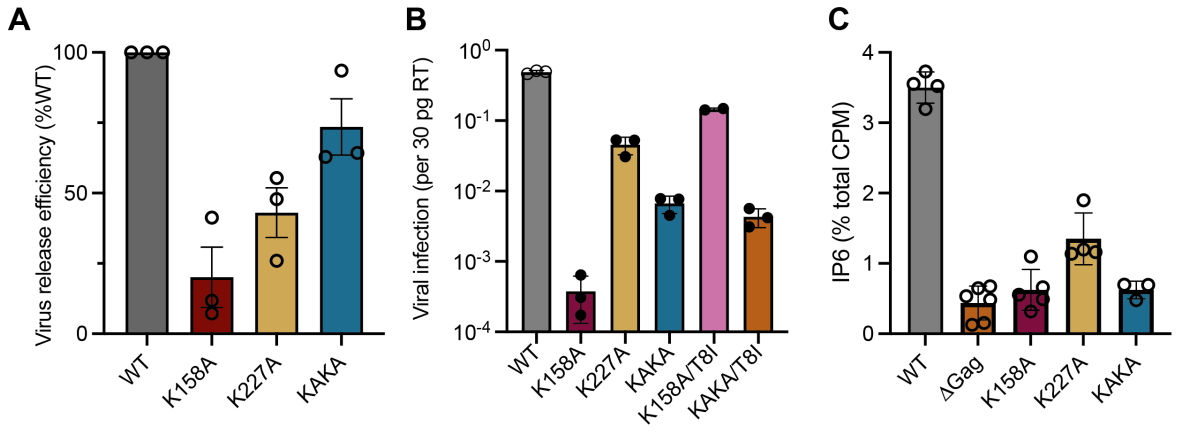


Figure 28: HIV-1 can become independent of IP6 for immature particle assembly. (A) Virus release efficiency of Gag mutants, calculated as the percentage of particle-associated p24 (CA) as a fraction of total (cell- + particle-associated Gag) normalised to WT virus in HEK293T cells. Error bars depict the SEM from at least three independent experiments. WT vs K158A p=0.0017(**), WT vs K227A p=0.003(**), WT vs KAKA p=0.058(ns). (B) Infectivity of Gag mutants normalised to the quantity of input virus [per 30 pg of reverse transcriptase (RT) in HEK293T cells. Error bars depict the SEM from three independent experiments. C) Viruses produced in cells supplemented with tritiated inositol were purified and inositol phosphate species extracted and fractioned by SAX-HPLC in HEK293Ts. The counts per minute (CPM) of IP6 shown as a fraction of total CPM in the sample. Error bars depict mean CPM ± SEM from at least two independent experiments. Experiments in this figure were performed by Donna Mallery and Alex Kleinpeter (A).

5.1.1 Aims

The aim of this chapter was to investigate whether IP6 is essential for immature or mature assembly or both.

First, I investigated the impact of significant IP6 reduction in virions by analysing the morphology of viruses produced in IP6 depleted cells via cryo-ET. I then determined if other IPs can functionally replace IP6 in the immature and the mature capsid by measuring capsid assembly in presence of IP5, IP4, IP3 and IP2. In addition, I measured the stability and cleavage behaviour of immature VLPs assembled with different IPs.

Lastly, I determined whether the KAKA mutant assembles independent of IP6 in vitro and what impact the KAKA mutation has on mature capsid formation by cryo-ET.

5.2 Results

5.2.1 Virions produced in IP5/IP6-depleted cells show aberrant morphology

To investigate the impact of IP6 depletion on particle production, I used cryo-ET to analyse particles produced in IPPK and IPMK KOs cells over-expressing full-length MINPP1. These cells contain high IP3 and IP4 but low IP5 and IP6 levels (**Figure 27** A). I classified all particles that possessed Gag structures into immature and mature with conical, tubular, or irregular cores. Surprisingly, the proportion of mature conical capsids of virions produced in IPPK/FM1 cells was only slightly reduced, and no reduction was seen in IPMK/FM1 cells compared to WT (**Figure 29** A, B). However, more irregular cores were present in both KOs. I further subclassified the mature particles into single cores or multiple cores, multi-layered cores and cores with additional open or closed structures (**Figure 29** C, D). Mature capsid cores formed by virion produced in the IPPK or IPMK KO cells had more additional structures for example multiple or multi-layered cores in virions, when compared to WT. However, the magnitude of the morphology defects revealed by cryo-ET does not explain the severe infectivity defect of virions produced in the KO/FM1 cells. During analysis, I realised that the number of virions on KO/FM1 grids appeared to be significantly lower compared to what was expected by RT quantification. This would suggest that particles contain viral proteins but are not able to from visible Gag assemblies. Therefore, I quantified all membrane-enclosed particles that had identifiable VSV-G surface protein. I divided them into clearly identifiable virions, vesicles with structured or unstructured density, filamentous structures, and partial densities (**Figure 29** E, F). VSV-G positive particles from KO/FM1 cells were more likely to appear ‘empty’ or contain only partial or unstructured density.

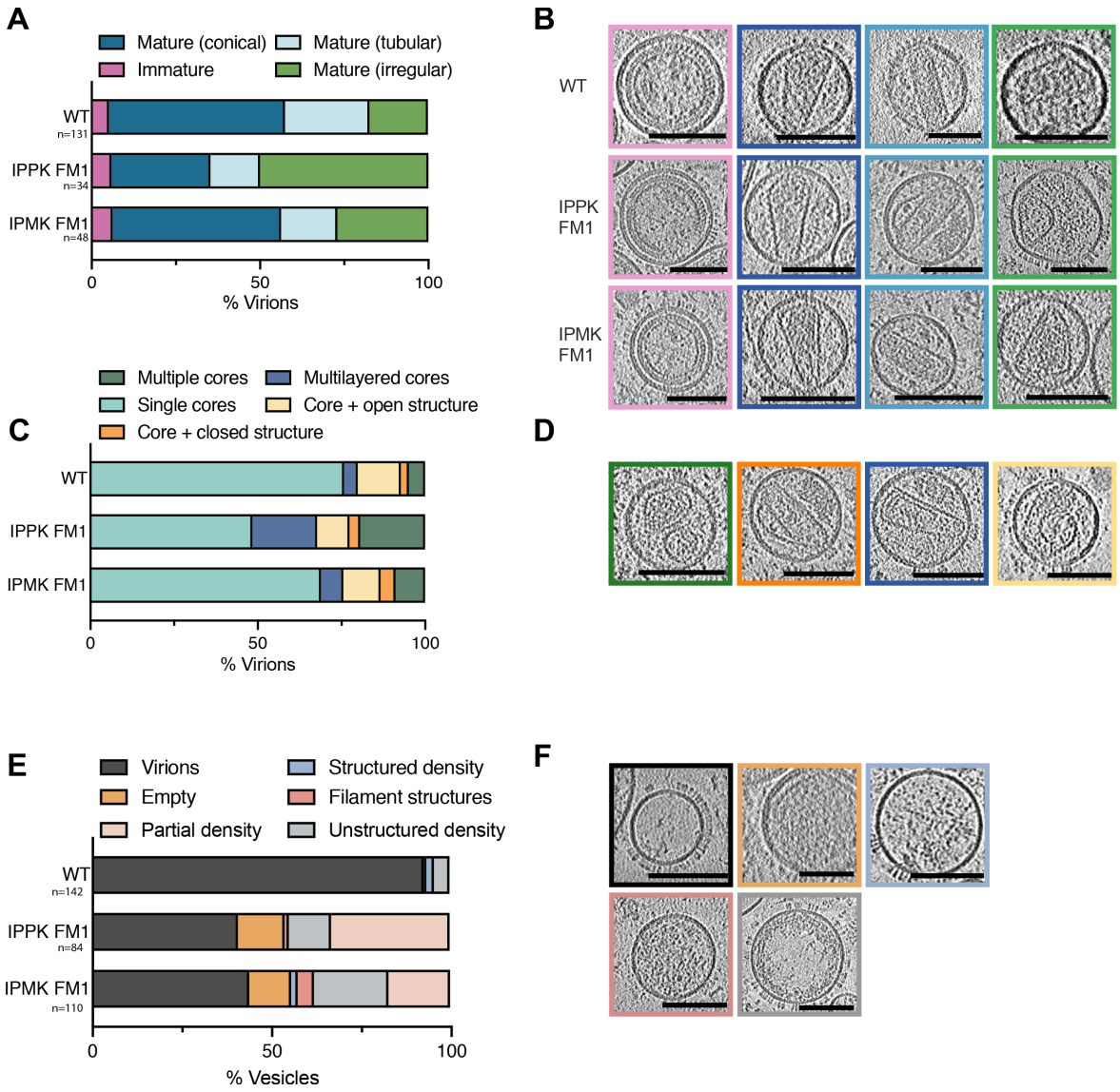


Figure 29: HIV-1 particles produced in IP5/IP6-depleted cells display aberrant morphology and are often empty. (A-D) Cryo-ET on HIV-1 virions produced in IPMK and IPPK KO cells while MINPP1 was overexpressed. Tilt series were collected, and reconstructions performed to assess capsid morphology. A total of 131 WT, 34 IPPK FM1, 48 IPMK FM1 and 163 KAKA virions were analysed. (A) Virions were classified into the categories: Immature (pink), Mature Conical (dark blue), Mature Tubular (light blue), Mature Irregular (green). (B) Slices through tomograms show representative examples of the viral morphologies, Scale bars: 100 nm. (C) Virions with mature lattices were further subdivided into: Multiple Cores (dark green), Single Cores (cyan), Cores with additional closed structure (orange), Cores with additional open structure (light orange), Multi-layered Cores (blue). (D) Slices through tomograms show representative examples of the viral morphologies, Scale bars: 100 nm. (E) All vesicles with visible VSV-G surface protein on the grid were categorized into: Virions (black), Partial density (gray), Empty (orange), Structured density (blue), Filament structures (dark pink), Partial density (light pink). (F) Slices through tomograms show representative examples of the viral morphologies, Scale bars: 100 nm.

This suggests that lack of IP6 and/or the incorporation of smaller IPs results in the production of aberrant vesicular particles without an organised capsid structure.

Aberrant particles may form when immature or mature Gag does not polymerize efficiently. This might be due to incomplete or altered cleavage or insufficient IP6 levels to allow the formation of cores. Thus, part of the infectivity defect of KO/FM1-produced virus may be due to an increase in the number of aberrant particles resulting in an over-estimate of the number of true virions.

5.2.2 Lower IPs can promote immature but not mature capsid assembly

Thus, I compared the ability of different IPs to promote in vitro assembly of immature and mature VLPs to understand how different IPs could give rise to defects in HIV-1 particle production and infection. I started by looking at immature assembly with a ΔMA-CANC construct, comprising part of the matrix (MA) domain and all of the capsid (CA) and nucleocapsid (NC) domains[326]. In presence of RNA and an IP concentration of 50 μM, IP5 and IP6, but not IP2-4, promoted immature assembly (**Figure 30** A, D). IP6 promoted assembly substantially faster than IP5 with a shorter lag time. IP4 assembly only occurred at elevated concentrations (**Figure 30** B, C, D). Neither IP2 nor IP3 were able to stimulate immature VLP formation under these conditions. To test the importance of Gag concentration, I increased the amount of ΔMA-CANC in the reactions while maintaining a stoichiometry of 1 IP molecule per 1 immature Gag hexamer. At 50 μM ΔMA-CANC, only IP6 stimulated immature assembly, whereas at 100 μM IP5 and at 150 μM IP4 were also assembly competent (**Appendix Figure 14**). There were clear differences in assembly efficiency and kinetics, with reduced particle formation by smaller IPs at equivalent Gag and IP concentrations.

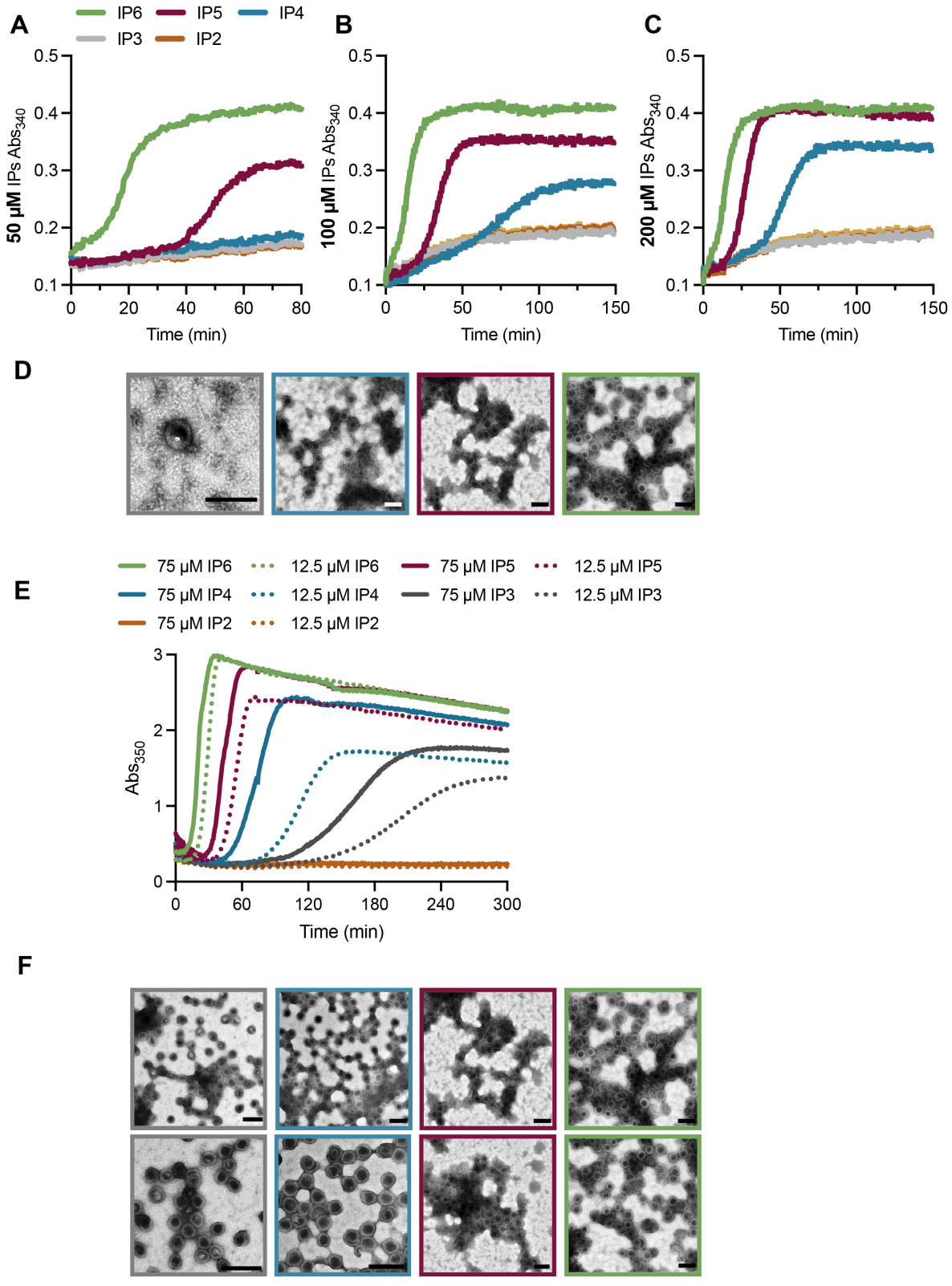


Figure 30: Immature particles assemble more slowly with smaller IPs. In vitro assembly of immature VLPs using recombinant ΔMA-CANC protein. IPs at the indicated concentrations and 7.5 μM RNA were added to 75 μM ΔMA-CANC and assembly monitored through absorbance changes at 350 nm at 25°C. IPs at 50 μM (A), 100 μM (B) or 200 μM (C) were added. (E) Negative stain images of the final assembly reactions in (C), Scale bars: 200 nm. (E) In vitro assembly at 37°C. Indicated IPs were

added at equimolar ΔMA-CANC concentration or 12.5 μM, to achieve a stoichiometry with immature hexamers of 6:1 or 1:1 respectively. (F) Negative stain images of the final assembly reactions shown in (E), Scale bars: 200 nm.

Subsequently, I tested the temperature-dependency of immature VLP assembly kinetics and efficiency. Whilst IP3 was not able to assemble VLPs at 25°C (**Figure 30** A-C), at 37 °C with 75 μM ΔMA-CANC, IPs from IP3 to IP6 were all able to stimulate the formation of immature VLPs, albeit with decreasing kinetics (IP6>IP5>IP4>IP3). IP2 was not able to assemble at these conditions (**Figure 30** E, F). In addition, lower concentrations of IPs were necessary to promote efficient assembly at 37°C. IP6 assembled efficiently at a stoichiometry of 1 molecule IP6 per 1 immature hexamer and increasing the IP concentration further did not enhance assembly (**Figure 30** E). This is consistent with there being only one assembly-promoting IP bxinding site per immature hexamer. Next, I added sub-stoichiometric amounts of IP6 to 75 μM ΔMA-CANC and monitored its assembly (**Figure 31** A). Assembly plateaued at a lower absorbance when less IP6 was added and at a ratio of 0.25:1 (IP6: hexamer) about half the signal of a ratio of 0.5:1 was measured. Moreover, adding more IP6 to either condition to achieve a stoichiometric ratio of 1:1 immediately restarted assembly and a similar final yield was reached. As expected, addition of more IP6 to the 1:1 assembly condition did not restart assembly since all available CANC was already assembled. In contrast, increasing the concentration of IP3 and IP4 beyond a 1:1 stoichiometry increased both the rate and efficiency of assembly. This might be due to a lower affinity because lower IPs have fewer phosphates to engage with the lysine rings in immature hexamers. Thus, a higher concentration might be required to drive assembly to completion. To test this, I assembled ΔMA-CANC in the presence of 12.5 μM IP3 and IP6 (**Figure 31** B). When the assembly plateaued, I added a further 12.5 μM of either IP. Additional IP3 did not have an effect but the addition of IP6 restarted assembly of plateaued IP3. However, upon addition of an excess 500 μM IP3 to the plateaued IP3 assembly, assembly resumed to the same level as the IP6 condition (**Figure 31** C). Taken together, immature assembly kinetics are fastest with IP6, followed by IP5, IP4 and IP3 most likely due to decreasing ΔMA-CANC affinities.

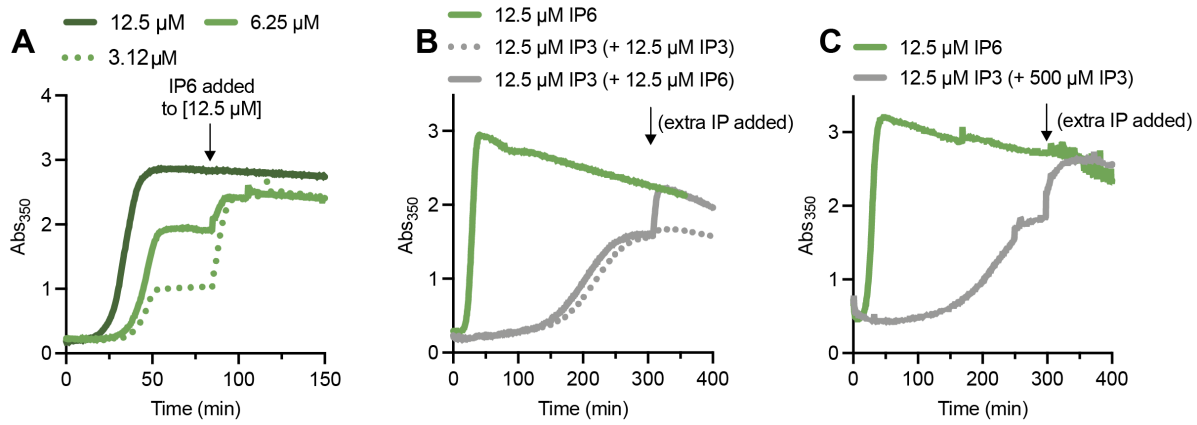


Figure 31: Stoichiometry of immature assembly. In vitro assembly reaction of 75 μM ΔMA-CANC and RNA with indicated IP concentrations at 37°C. (A) In vitro assembly reaction with an IP6 to immature hexamers ratio of 1:1 (12.5 μM), 0.5:1 (6.25 μM) or 0.25:1 (3.12 μM). Once assembly had plateaued, additional IP6 was added to achieve 1:1 stoichiometry, resulting in similar final yields by absorbance. (B) In vitro assembly with 75 μM ΔMA-CANC and 12.5 μM (1:1 stoichiometry) IP3 or IP6. At the indicated time point, an additional 12.5 μM of IP3 or IP6 was added. When additional IP6 but not IP3 is added to the reaction there is a renewed increase in absorbance indicative of further assembly. (C) As with (B), but at the indicated time point excess IP3 is added (500 μM) leading to a resumption in assembly to IP6-stimulated levels.

Since virions produced in IPMK and IPPK KOs overexpressing MINPP1 cells showed aberrant cleavage (**Figure 27** E), I reasoned that different IP molecules could influence the stability of the immature lattice. Thus, I measured the thermal stability of immature VLPs assembled with different IPs. WT VLPs assembled in presence of IP6 showed a > 15 °C increase in melting temperature (Tm) relative to unassembled protein. The stability of VLPs decreased with lower IPs (IP6 >IP5>IP4 >IP3) (**Figure 32** A, **Appendix Figure 15**). I then tested whether this reduced stability contributes to altered Gag processing in an in vitro cleavage assay. I added purified HIV-1 protease to immature VLPs that had been assembled *in vitro* with different IPs and assayed the cleavage reaction over time (**Figure 32** C). **Figure 32** B shows a scheme of proteolytic cleavage and the molecular weights of the cleavage products. I also used tartrate as a control, since it was previously shown to induce VLP assembly[314]. As expected, cleavage of IP6-assembled VLPs started with NC cleavage and ended with SP1 liberation from CA. Processing of VLPs assembled in the presence of IP5 was very similar. However, VLPs formed with smaller IPs, most noticeably with IP3 or IP2 or tartrate, displayed differences in the processing sequence. In particular, SP1 cleavage began much earlier, immediately after liberation of NC, and proceeded more quickly

than in the presence of IP6. As a result, multiple cleavage products were observed. On example was a band possibly corresponding to ΔMA-CA, a cleavage intermediate rarely produced under normal processing conditions, that only occurs due to premature SP1 liberation. This is consistent with the reduced thermal stability of VLPs assembled with smaller IPs and supports the importance of IP in stabilising the six-helix bundle within SP1 and hence the hexameric lattice. Taken together, the data presented here suggest that usage of IP3 or IP4 by virions produced in cells lacking IP5 and IP6 will result in a less stable immature lattice and altered Gag processing.

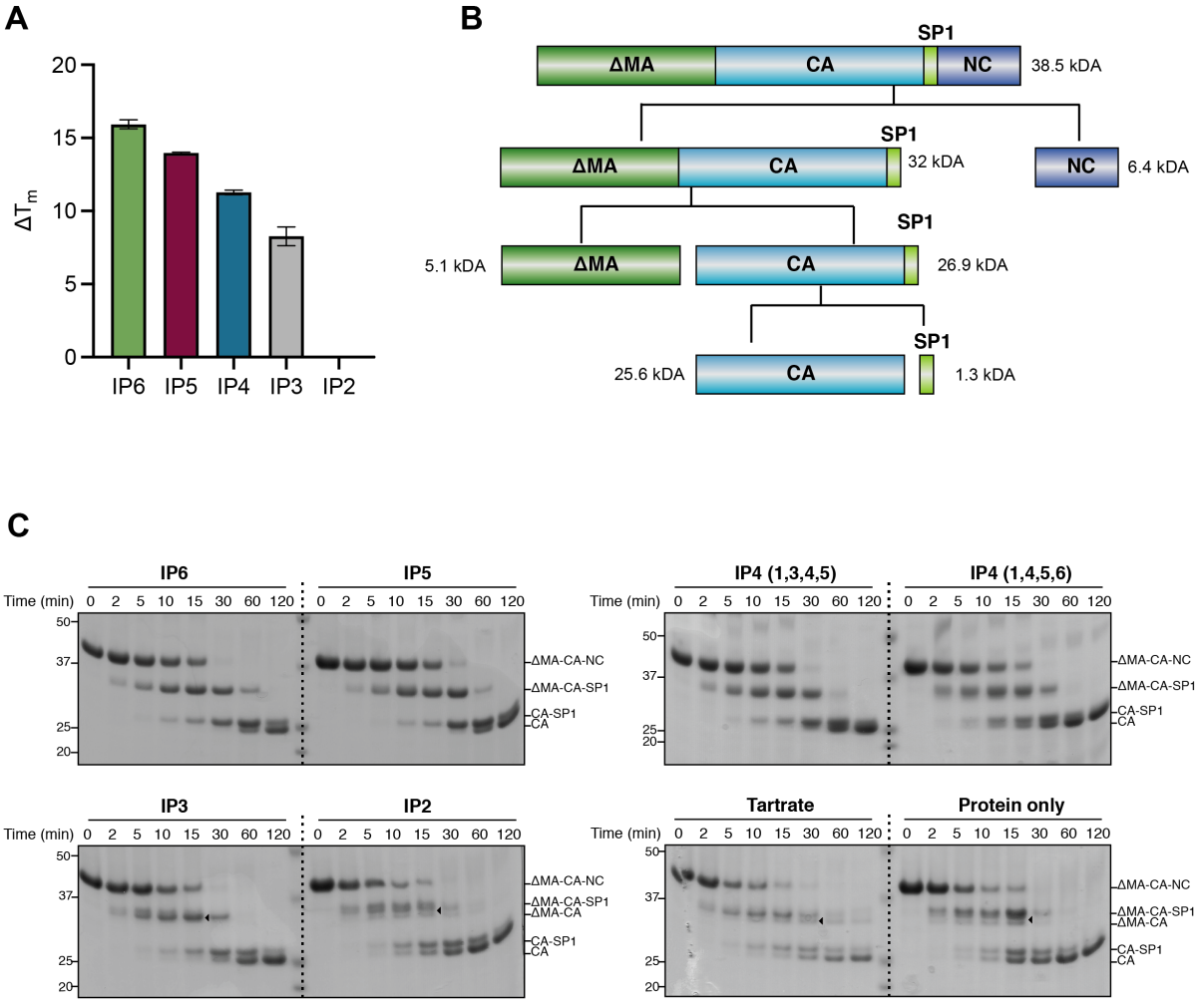


Figure 32: Immature particles assembled in presence of smaller IPs are less stable and have altered Gag processing relative to VLPs assembled in the presence of IP5 or IP6. (A) Thermostability of in vitro assembled VLPs with 7.5 μM μM RNA, 75 μM ΔMA-CANC and 12.5 μM IP was measured by differential scanning fluorimetry (DSF). The change in melting temperature (ΔTm) was calculated with respect to the thermostability of unassembled ΔMA-CANC protein. Curves are shown in **Appendix Figure 15**)(B) Schematic of Gag processing by HIV-1 protease: The ΔMA-CANC construct is the starting material used for in vitro assembly and proteolysis experiments throughout this work. The sequential order of proteolytic cleavage and the MW of the cleavage products are shown.

(C) In vitro assembled VLPs but with an additional condition including 375 μM Tartrate rather than IP and unassembled ΔMA-CANC were incubated with recombinant HIV-1 protease for the indicated times and analysed by SDS PAGE. The probable cleavage products, based on size, are indicated. Triangles point to additional cleavage products that are not present in particles assembled with IP6.

Based on the immature assembly experiments presented above, I wondered whether lower IPs could also promote mature VLP assembly. Although the lower IPs increased the melting temperature and thus bind to pre-formed hexamers (**Appendix Figure 13**), only IP6 was able to promote efficient mature capsid assembly even at very high IP concentration (2.5 mM) (**Figure 33**). IP5 was only capable of promoting assembly to a small extent as was IP4 although slower. No assembly was observed for IP3. IP5 and IP4 cores were rare in negative stain and often small and non-conical. This establishes that depletion of IP6 impacts mature assembly significantly more.

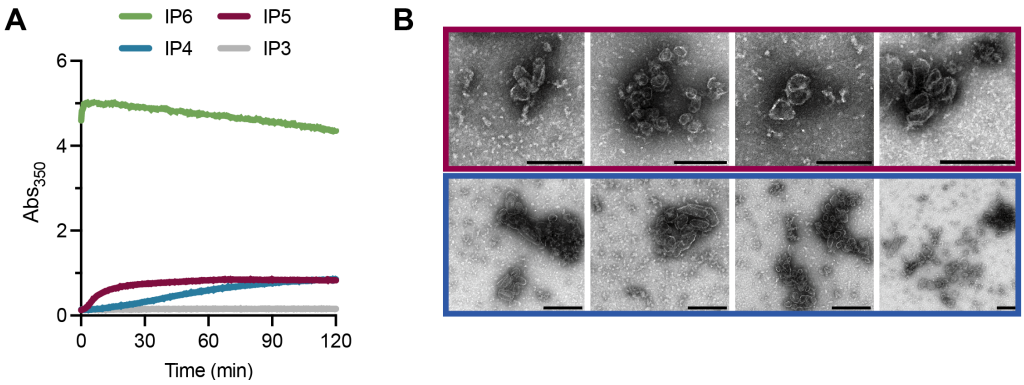


Figure 33: Mature cores only assemble efficiently with IP6 (A) In vitro assembly of mature VLPs using 150μM recombinant CA protein and 2.5 mM of the indicated IP. (B) Negative stain EM images of the final assembly reactions shown in (A), Scale bars: 200 nm.

5.2.3 Immature VLP assembly can be IP6 independent

IP6 levels within virions can also be manipulated by removing residues in the immature lattice responsible for its enrichment - K158 and K227. The double lysine mutant, KAKA, produces more virion than either single lysine mutant alone, to approximately 75% WT, but it remains significantly impaired for infection (**Figure 28**). The increased production of the KAKA mutant suggests its Gag lattice forms more readily than to the single mutants.

I investigated the assembly of WT and KAKA ΔMA-CANC protein proteins in vitro. WT immature particles assembled only in presence of both IP6 and RNA, whereas RNA alone was sufficient to drive assembly of the KAKA mutant (**Figure 34** A) as described

previously[328]. As expected, IP6 did not enhance KAKA assembly as this mutant does not contain the lysine rings required to bind IP6. However, assembly of KAKA was slower compared to WT with IP6. This observation contrasts with previous work suggesting KAKA assembly is more efficient than WT[328]. Negative stain analysis showed that KAKA VLPs were similar in size and shape to WT assembled with IP6 (**Figure 34** B). I then tested the effect of the SP-1 mutant T8I on assembly (**Figure 34** C, D). T8I assembled efficiently with IP6 and RNA, and like WT it did not form VLPs without IP6. KAKA/T8I ΔMA-CANC assembled in the absence of IP6 with increased kinetics compared to KAKA. As expected, the addition of IP6 did not enhance KAKA/T8I VLP assembly. This data suggests that immature hexamer stability is crucial factor in VLP assembly velocity.

To assess the stability of KAKA VLPs, I measured the increase in thermal stability upon VLP assembly (**Figure 34** E). WT VLPs assembled in the presence of IP6 and RNA showed a ~12 °C increase in melting temperature (Tm) relative to unassembled protein. No increase in melting temperature was observed in the absence of IP6 as no WT assembly occurred under these conditions. Mutant SP1-T8I VLPs assembled with IP6 and RNA achieved a higher melting temperature than WT. KAKA VLPs assembled with RNA only, displayed an increased melting temperature of ~8 °C. As expected, this was not further increased in presence of IP6. The Tm of KAKA/T8I VLPs assembled with RNA only was slightly increased compared to KAKA. No further effect could be measured in the presence of both RNA and IP6. This shows that IP6 is required for optimal assembly and stability of immature HIV-1 VLPs. However, if IP6 recruitment is impaired, e.g., by mutating one lysine ring, then an uncharged immature hexamer core lacking both lysine rings is preferred.

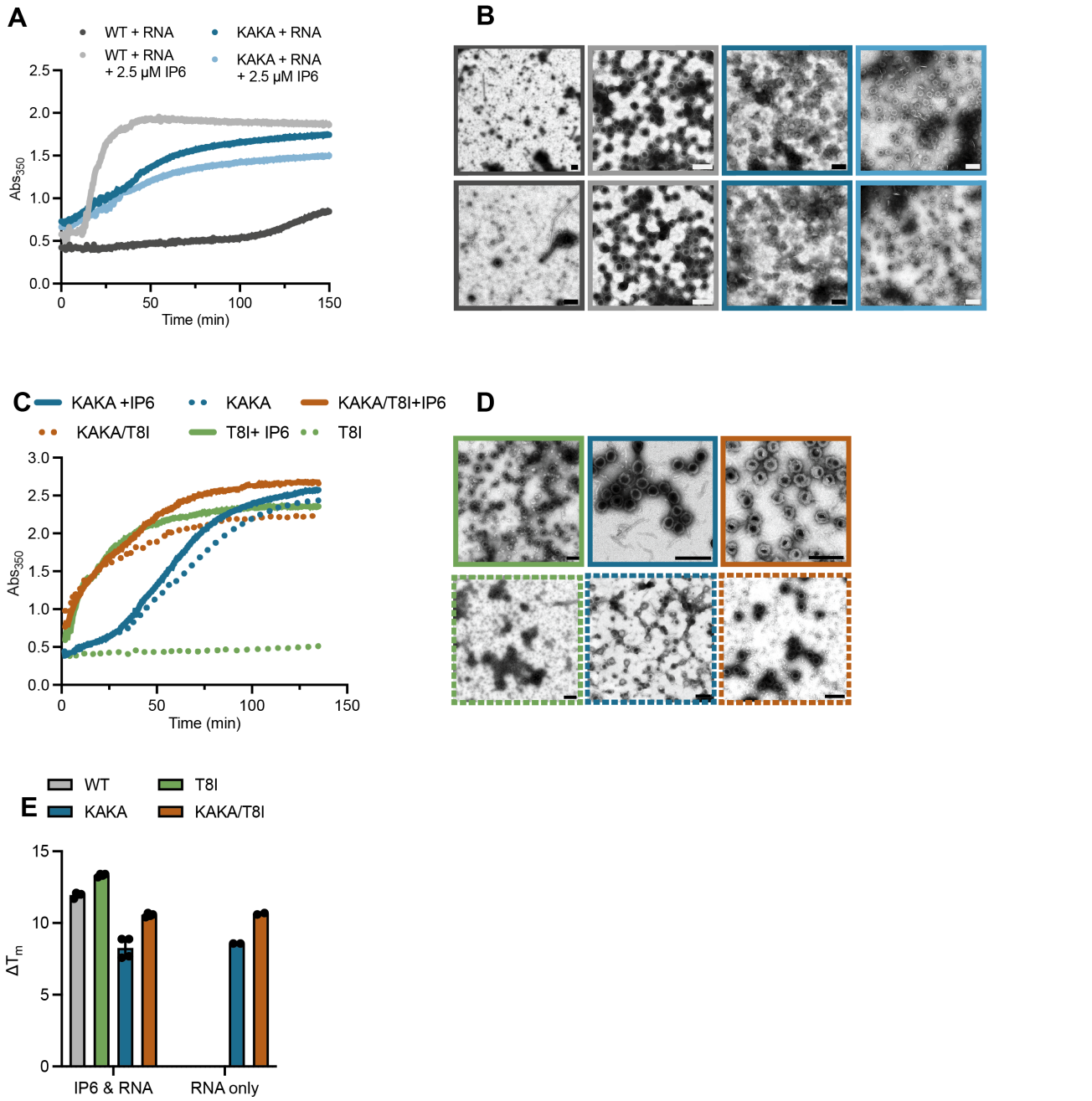


Figure 34: HIV-1 immature assembly can become independent of IP6. (A) In vitro assembly of immature VLPs using recombinant ΔMA-CANC protein 7.5 μM RNA was added to 75 μM ΔMA-CANC and assembly monitored through absorbance changes at 350 nm. (B) Negative stain EM of samples of the final assembly reactions shown in (A), Scale bars: 200 nm. (C) Kinetics of immature VLP assembly of ΔMA-CANC mutants using 10 μM RNA and 100 μM ΔMA-CANC. (D) Negative stain EM of samples of the final assembly reactions shown in (C), Scale bars: 200 nm. (E) The thermostability of in vitro assembled VLPs of (C) was measured by differential scanning fluorimetry and expressed as a change in melting temperature (Tm) compared to unassembled ΔMA-CANC.

The ability of the KAKA mutant to undergo immature particle assembly in vitro suggested that its reduced infectivity may be due to defective mature core formation as a consequence of reduced IP6 levels within the virion. I, therefore, used cryo-

electron tomography (cryo-ET) to examine the structures of capsids in purified virions (**Figure 35**). WT, KAKA and KAKA/T8I virions were classified into immature or mature virions. The latter was subdivided into conical, tubular or irregular capsid or ambiguous virions with no proper Gag structure (**Figure 35** A, B). KAKA virions had fewer conical capsids than WT and an increased proportion of irregular or tubular mature capsids. In addition, more ambiguous virions were observed. However, this does not explain the infectivity defect. I further analysed the mature virions and classified possible substructures (**Figure 35** C). This showed that KAKA mostly contained single cores similar to WT. This is in contrast to virions of the single mutant K158A where additional structures were observed often (**Figure 35** C, **Figure 26**). The K158A infection and processing defects were previously shown to be rescued by the second-site compensating mutation SP1-T8I, which stabilises the six-helix bundle at the centre of the immature Gag hexamers and restores IP6 enrichment into virions[322]. Addition of T8I to K158A was also shown to reduce the presence of additional structures (**Figure 35** C, **Figure 26**). In contrast, addition of SP1-T8I had little effect on KAKA (**Figure 35** A-C). This is in line with the in vitro assembly data and suggests that the KAKA mutant is already capable of forming a stable immature hexameric Gag lattice.

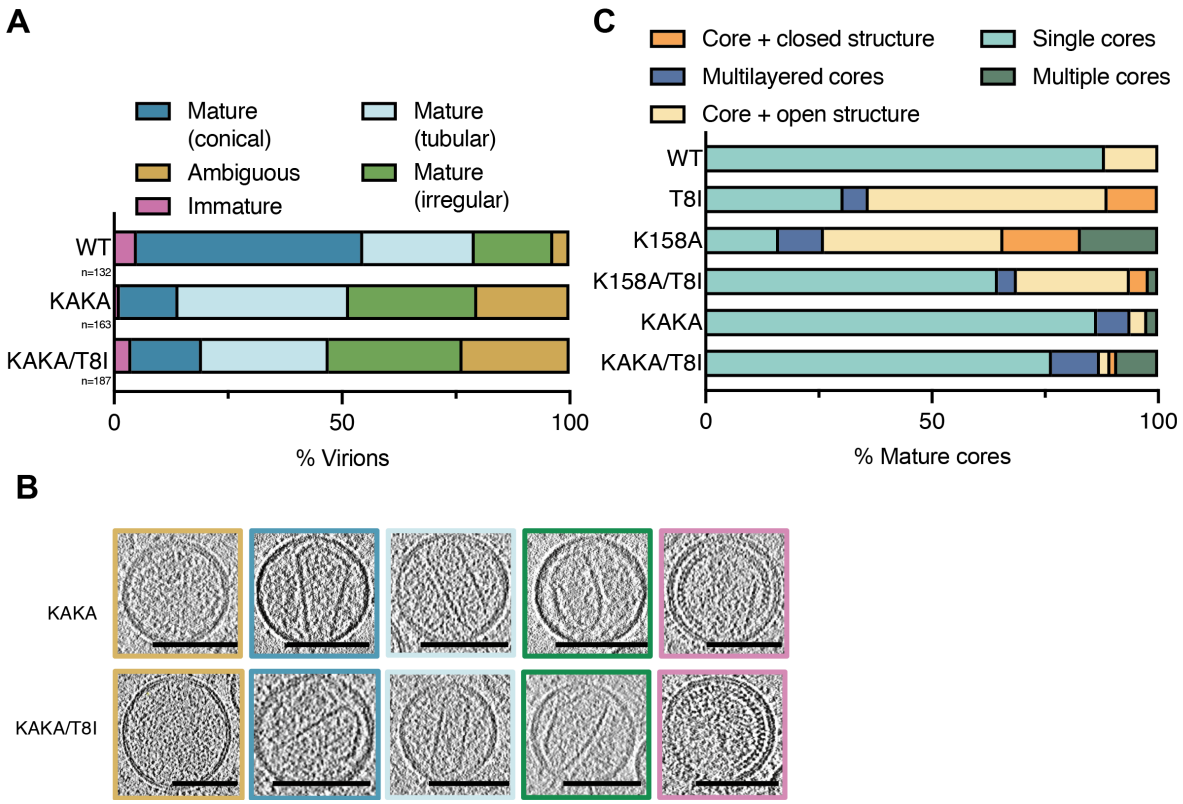


Figure 35: KAKA can form immature and mature capsids. (A-B) Cryo-ET of indicated HIV-1 mutants. Tilt series were collected, and reconstructions performed to assess capsid morphology. A total of 163 KAKA, 187 KAKA/T8I, 69 T8I, 127 K158A and 119 K158A/T8I virions were analysed. (A) Virions of the indicated Gag mutants were classified into the indicated categories: Ambiguous (orange), Immature (pink), Mature Conical (dark blue), Mature Tubular (light blue), Mature Irregular (green). (B) Example tomograms of the virions quantified in (A), Scale bars: 100 nm. (C) Virions with mature lattices were further subdivided into: Multiple Cores (green), Single Cores (cyan), Cores with additional closed structure (orange), Cores with additional open structure (light orange), Multi-layered Cores (blue).

The KAKA mutation also rescued Gag processing in virions produced in IP6-depleted cells; irrespective of kinase KO or MINPP1over-expression, KAKA achieved similar Gag cleavage efficiency as virions produced in WT cells (**Appendix Figure 16** C). I, therefore, used a proteolytic cleavage assay to monitor KAKA cleavage efficiency (**Figure 36**). WT and mutant ΔMA-CANC was assembled in the presence of IP6, RNA and tartrate to promote VLP formation of all constructs (**Appendix Figure 17**). Processing of the mutant VLPs proceeded slightly differently compared to WT. SP1 liberation from K158A VLPs occurred earlier, and multiple cleavage products were visible. K227A and KAKA cleavage was more similar to WT but, SP1 appeared to be cleaved slightly faster than WT. The CA-SP1 band was more dominant in the WT reaction and persisted longer, whereas for the mutants CA-SP1 was cleaved earlier. Processive proteolytic cleavage is depended on immature lattice stability, and

cleavage of SP1 is crucial to shift the lattice to a mature form to allow infectious particle formation[329]. Thus KAKA, but not K158A, appears to be able to assemble a stable immature lattice allowing ordered WT-like proteolytic cleavage.

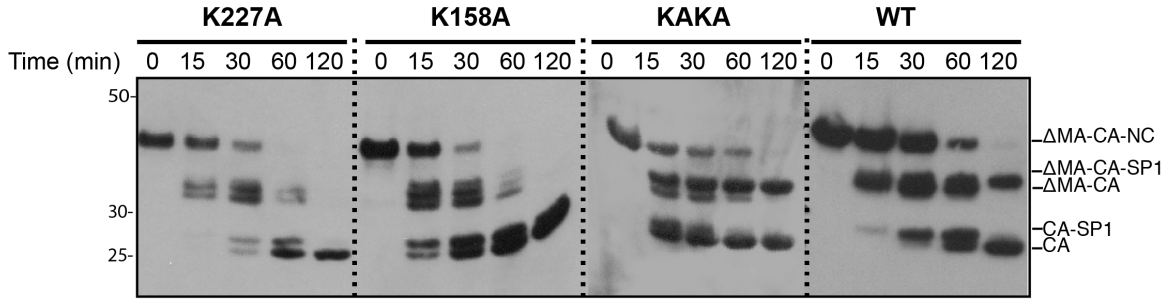


Figure 36: KAKA rescues cleavage defect in vitro. In vitro assembled VLPs assembled in 375 μM Tartrate, IP6 and RNA were incubated with recombinant HIV-1 protease for the indicated times and analysed by SDS PAGE. The probable cleavage products, based on size, are indicated.

Making the immature lattice IP6-independent by mutating the lysine residues also made viral production independent of IP6 levels in producer cells (**Appendix Figure 16** A). However, IP6 depletion dose-dependently decreases the infectivity of all lysine mutants (**Appendix Figure 16** B). Despite KAKA virions having a reduced ability to enrich IP6 into virions, IP6 can still be incorporated passively into virus particles in WT cells. However, in IP6 depleted cells the pool of IP6 available for passive incorporation into KAKA virions is drastically reduced and thus, mature capsid assembly might be impaired.

5.2.4 Mature and Immature Assembly require a different concentration of IP6

To understand how immature and mature capsid formation is impacted by IP6 levels, I carried out in vitro mature assembly reactions with recombinant CA. I measured the maximal absorbance values (A_{max}) corresponding to maximal assembly. A range of different CA and IP6 concentrations were measured to calculate the concentration necessary for half-maximal absorbance for mature assembly (Abs_{MA}). An Abs_{MA} of 224 ± 17 μM was calculated for mature assembly (**Figure 37** A). Likewise, the half-maximal absorbance for immature assembly (Abs_{IA}) was measured using ΔMA-CANC in the presence of RNA. The calculated Abs_{IA} constant was 2.8 ± 1.4 μM IP6 (**Figure 37** B). This data shows that immature VLP assembly occurs at IP6 concentrations ~ 100-fold below what is required for mature VLP assembly. As cellular concentrations of IP6 are 22-44 μM[155], immature assembly can occur efficiently in producer cells. In stark

contrast, mature capsids formation cannot. Moreover, this strongly implies that mature capsid formation must be more sensitive than immature assembly to any reduction in cellular IP6 levels.

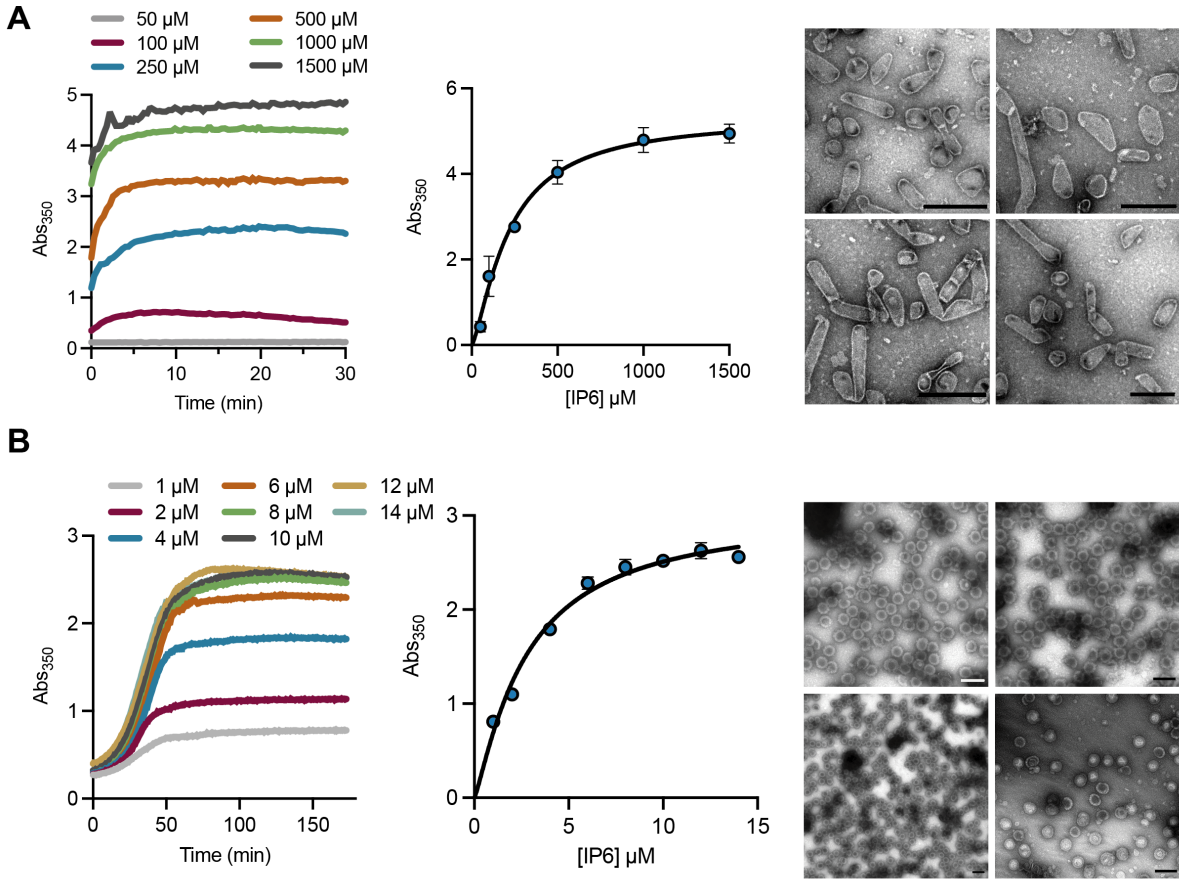


Figure 37: Half-maximal assembly for mature and immature capsid assembly. (A) In vitro mature assembly kinetics with 75 μM CA and 50-1500 μM IP6 are shown on the left. The maximum assembly at different IP6 concentrations was fitted to equation I: $Y=A_{max}*X^h/(Abs_{MA}{}^h + X^h)$; where A_{max} is the maximum assembly, Abs_{MA} is the half maximal assembly and h is the Hill slope. Fitting gave a Abs_{MA} of 217 ± 16 μM. Error bars depict mean Amax ± SD of at least three independent measurements. EM images of negatively stained samples of the final assembly reactions are shown on the right, Scale bars: 200 nm. (B) In vitro immature VLP assembly with 75 μM CANC and a range of IP6 concentrations. Data) were fitted as to give a Abs_{IA} of 2.8 ± 1.4 μM. Error bars depict mean Amax ± SD of at least three independent measurements. EM images of negatively stained samples of the final assembly reactions are shown on the right, Scale bars: 200 nm.

5.3 Discussion

The data shown here highlights the importance of IP6 in the HIV-1 replication cycle. IP6 binds to and promotes assembly of both immature VLPs and mature capsids. In vitro assembly kinetics showed that IP6 concentrations have to be significantly higher to promote mature assembly compared to immature assembly. This might be

necessary, because the immature CANC protein has the ability to form mature particles in vitro[190]. However, the cellular IP6 concentration of 24-47 μM measured in mammalian cell lines[155] is too low to support mature assembly. Thus, mature assembly can only occur within virions where IP6 has been enriched. Tritiated inositol measurements suggest there over 300 IP6 molecules per virion[229]. If 300 IP6 molecules were present in a 120 nm virion, the concentration would be around 0.55 mM IP6 suggesting a 10-fold enrichment within the virion. Thus, if 1500 CA monomers assemble into a capsid, 1–2 IP6 molecules per hexamer would be present in a fully assembled capsid. This is consistent with there being two known binding sites per hexamer provided by R18[190,229] and K25[327].

Previous experiments have tested the dependence of HIV-1 on IP6 by knocking out biosynthetic IP kinases[191,192,313]. It was shown that IPMK or IPPK KO results in mild infection and production defects. IP6 appears to be efficiently recruited by the lysine rings into the immature Gag lattice even at low cellular concentration, thereby enriching IP6 into virions. This would allow mature assembly even if cellular IP6 levels are reduced. Over-expression of the phosphatase MINPP1 reduces the levels of IP5 in IPPK KOs and IP6 in IPMK KOs such that they are at level well below the concentration required for immature assembly, while IP3 and IP4 levels remain unchanged or even increased. This leads to a severe impact on viral production, Gag processing and particle infectivity. In vitro assembly data showed that immature VLPs can assemble with IP5 and to a lesser extent IP4 at close to WT kinetics or slower with IP3. In contrast, mature assembly is only efficient in presence of IP6 and strongly decreased with IP5 or IP4. This highlights that mature assembly is far more dependent upon IP6 than immature assembly. If IP6 is not enriched into virions then there will only be enough molecules to stabilise 10-30 % of hexamers in the capsid, thus risking post-entry premature capsid collapse. Taken together, these results strongly support the critical importance of IP6 in HIV-1 replication.

IP6 incorporation can also be depleted by mutating one or both of the lysine rings of the immature hexamer. Unlike the single mutants, mutation of both rings (KAKA) does not have a substantial defect in particle production or Gag processing. In addition, KAKA can assemble into immature VLPs in the absence of IP6. This demonstrates that there is no intrinsic structural requirement for IP6 in building an immature HIV-1

Gag lattice. Meanwhile, cryo-ET of KAKA virions shows that mature capsids can also form in absence of IP6 enrichment, albeit in ¼ of cases compared to WT. Capsid formation within KAKA virions may be driven by the extremely high local CA concentration. However, these capsids are most likely very unstable once this concentration drops. Thus, as soon as virions release their capsid into the cytoplasm, the capsid will likely disassemble immediately. As a result, KAKA shows a strong reduction in infectivity compared to WT. Interestingly, infectivity of KAKA lies between K158A and K227A. Whilst K227A can enrich some IP6, KAKA and K158A have lost this ability, explaining their reduced infectivity. KAKA is probably more infectious than K158A because it does not show the same over-assembly phenotype as K158A (see chapter 4) which most likely has a negative impact on infection.

However, the fact that there are no natural KAKA isolates[330] indicates that enriching IP6 into virions provides an important selective advantage for HIV-1 transmission. KAKA virus is essentially independent of IP6 at the level of immature assembly and also rescues the cleavage defects observed with WT Gag expressed in IP6-deficient cells. Nevertheless, it has a profound infectivity defect that strongly suggests that IP6 is required for mature capsid assembly and/or stability. This is further supported by the fact that the infectivity defect of KAKA is exacerbated by IP6-depletion in producer cells. Interestingly, the KAKA mutant seems to be very similar to Rous sarcoma virus (RSV). The immature capsid of the RSV does not possess the basic rings found in HIV and RSV immature assembly is consequently IP6 independent. However, IP6 is essential for mature capsid formation and is coordinated by a ring of lysine and arginine residues in mature hexamers. Viral infectivity and release is even more sensitive to IP6 depletion than WT HIV-1[331]. This is most likely because like in the KAKA mutant IP6 cannot be enriched but is only passively incorporated into the budding virus. Whether there are mutations equivalent to KAKA in the mature capsid that make mature assembly IP6 independent remains to be determined.

Chapter 6 Conclusions and outlook

6.1 Conclusions

In this thesis, I aimed to investigate the role IP6 plays in mature and immature HIV-1 assembly and whether it is essential for HIV-1 replication.

Chapter 3 describes the role of two positively charged rings, R18 and K25, in the mature hexamer. Both residues were proposed to recruit dNTPs to facilitate encapsidated DNA synthesis and removal of these rings strongly reduces infectivity[49,271]. Since clustering of charges is destabilising, it was also proposed that R18 recruits the metabolite IP6 to allow capsid assembly[190,229]. Whilst R18 engages IP6 at the top of the pore, K25 can simultaneously bind a second IP6 molecule. Mutation of R18 leads to irregularly shaped and unstable mature capsid in virions and IP6 cannot drive assembly of capsid cones in vitro. Likewise, mutation of K25 results in a severe defect in capsid assembly and stability. Monitoring both the kinetics and morphology of capsids assembled in vitro revealed that while mutant K25A can still form tubes in presence of IP6, the ability of IP6 to drive assembly of capsid cones has been lost. This suggests that recruitment of a second IP6 by K25 drives the assembly of conical capsids or that K25 can help to maintain IP6 in capsomers. Due to the strongly decreased capsid stability, it cannot be directly tested whether R18 and K25 are involved in nucleotide import because the reverse transcription defect that results from their removal could be explained by unstable cores that disassemble rapidly during infection. In contrast to R18G, capsid formation of K25A can be rescued by compensating mutants without restoring the charged K25 ring. These compensatory mutants also largely rescue infection and allow reverse transcription. This suggests that K25 is not essential for capsid assembly and nucleotide import, but rather enhances capsid assembly. Why HIV-1 has evolved to utilise two charged rings, whereas other lentiviruses have just one, is both an interesting question itself and may yet help to address exactly what roles the rings play.

Chapter 4 describes the role of IP6 in the context of immature lattice stability. The immature Gag hexamer recruits IP6 to promote immature lattice assembly and provide stability. Lattice stability has to be carefully modulated to allow processive proteolytic

cleavage. This is exploited by a class of drugs called maturation inhibitors which stabilise the immature lattice and thus interfere with cleavage to reduce infectivity. It was shown that IP6-binding deficient virions are resistant to the maturation inhibitor PF96. Importantly, PF96 and IP6 do not compete for binding to immature hexamers, but PF96 can act with positive co-operativity to restore IP6 incorporation into virions with impaired IP6 binding. Likewise, passaging experiments yielded compensatory mutants that stabilise the immature lattice and can also rescue IP6 incorporation, assembly and infectivity of IP6-binding-deficient mutants. This indicates that HIV can maintain IP6 binding by altering the stability of the immature Gag lattice. Thus, designing MIs that mimic IP6 but block its incorporation may provide more potent antiretrovirals that are more challenging for the virus to escape from.

Chapter 5 shows that the immature lattice is required to recruit IP6 into immature virions to catalyse mature capsid assembly. An HIV-1 Gag mutant, which lacks both the positively charged amino acid residues responsible for binding IP6, does not enrich IP6 but allows immature lattice formation and viral production. However, without sufficient IP6 molecules inside each virion, HIV-1 can no longer build a stable mature capsid and becomes non-infectious. Moreover, manipulating the availability of cellular inositol phosphates reveals that HIV-1 cannot functionally replace IP6, as substitution with other IPs profoundly slows mature assembly kinetics and results in virions with gross morphological defects. These results suggest that HIV-1 enriches IP6 into virions not because it is essential for the assembly of an immature lattice but because it is needed to form and stabilise a mature capsid.

6.2 Outlook

As established in this thesis, IP6 might play several roles in HIV replication. It may be necessary to maintain ideal capsid stability in the cytoplasm, withstand nuclear import and ensure ordered uncoating in the nucleus. Another possibility is that IP6 binding is required to build a mature capsid that contains a charged pore capable of importing dNTPs for encapsidated reverse transcription. The positively charged pores necessary for dNTP import and translocation are a destabilising feature if they are not masked by IP6. While this thesis shows that K25 does not seem to be essential for

efficient reverse transcription, it cannot prove or disprove if the pore is necessary for dNTP import.

One approach to gain further insights into the necessity of IP6 is the investigation of other retroviruses. It is becoming increasingly evident that IP6 binding is evolutionarily conserved in lentiviruses and possibly in other retroviruses (phylogeny of retroviruses is shown in **Figure 1**). The Gag proteins of the lentiviruses HIV-1, HIV-2, simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anaemia virus (EIAV), and bovine immune deficiency virus (BIV) all have conserved lysine residues (equivalents to K158 and K227 in HIV-1) in their immature lattice. In vitro assembly assays have shown that the immature assembly of these lentiviruses can be induced by IP6[313,332]. In line with this, the production of HIV-1, HIV-2, SIV, and FIV was reduced in IPMK or IPPK KO cells[191]. Another study confirmed that for SIV_{mac} production and infectivity, IP6 or IP5 are essential in producer but not target cells[313]. However, IP6 did not seem to be necessary for FIV or EIAV production as no effect was observed when IPPK or MINPP1 were added[313,332]. Nevertheless, it should be noted that in neither study the abundance of inositol phosphate species was quantified. Although IP6 binding to immature EIAV Gag is identical to HIV, EIAV can form immature VLPs in absence of IP6 in vitro, because unlike in HIV-1, its 6 -helix bundle is stable in the absence of IP6[332]. However, the formation of immature EIAV VLPs was substantially increased by IP6 addition, suggesting that whilst IP6 may not be critical for non-primate lentiviruses, it significantly enhances assembly and production. So far, the IP6 dependence of the mature capsids has not been investigated for other lentiviruses than HIV-1. They all have at least one positively charged ring in the hexamer, but it is likely that the IP6-dependence for mature assembly might differ. All lentiviruses can infect non-dividing cells, so they can pass through an intact nuclear envelope[333,334] suggesting that they either have similar capsid stability to HIV-1 to be imported via the same pathways or they use different pathways. Thus, investigating their stability could provide valuable insights into different infection pathways which might also potentially be used by HIV.

Interestingly, although the immature capsid of the alpharetrovirus rouse sarcomavirus (RSV) does not possess the basic rings found in lentiviruses, IP6 depletion in cells reduces infectivity and release. Although, RSV immature assembly was shown to be

independent of IP6 in vitro, infectious particle production was sensitive to IP6 depletion. This was partly explained by a reduced particle release and the observation that in the mature capsid IP6 is coordinated by a ring of lysine and arginine residues and is involved in mature capsid formation [331]. Beta- and gamma-retroviruses do not have lysine residues in the immature that could bind IP6. Not surprisingly, IP6 does not seem to be required for the production or assembly of the Gammaretrovirus Murine Leukemia Virus (MLV) or the Betaretrovirus Mason-Pfizer Monkey Virus (MPMV)[313,335]. It is worth mentioning that the mature capsid of lentiviruses and alpharetroviruses are presumed to be closed[43,50,331]. In contrast, the IP6 independent gamma retrovirus MLV does not form a closed capsid shell, thus nucleotides might be able to pass through gaps into the capsid[336].

It would be of great interest to investigate why certain viruses have evolved to use IP6 and, conversely, why certain viruses appear to be IP6-independent. This thesis showed that the immature HIV-1 lattice can become IP6 independent, but the mature cannot. Investigation of whether that holds true for more retroviruses would provide valuable insights into mature capsid assembly. In addition, it might be possible to generate compensatory mutants of R18 mutants based on structural insights gained by other retroviruses. Such data could reveal whether mature HIV assembly can become IP6-independent and if so, whether this would still allow encapsidated reverse transcription. In addition, it could give insights into the physical properties of mature capsid assembly e.g. HIV-1 mainly forms conical particles, whereas RSV forms highly pleomorphic particles[331]. In addition, it might be worth investigating these observations and IP6 levels in general in primary T-cell macrophages and dendritic cells.

Understanding why viruses use IP6 will improve our knowledge of how infectious retroviral particles are formed, which co-factors they use and when they disassemble. Importantly, this could lead to the development of new antiviral drugs. Since the HIV capsid is very conserved this could be a useful option for the long-term antiviral treatment.

References

1 Greene, W. C. A history of AIDS: looking back to see ahead. *Eur J Immunol* **37 Suppl 1**, S94–102, doi:10.1002/eji.200737441 (2007).

2 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, doi:10.1126/science.6189183 (1983).

3 Gallo, R. C. *et al.* Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *science* **224**, 500–503 (1984).

4 UNAIDS. GLOBAL AIDS UPDATE (2021).

5 Hammer, S. M. *et al.* A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. *New England Journal of Medicine* **337**, 725–733 (1997).

6 Gulick, R. M. *et al.* Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *New England Journal of Medicine* **337**, 734–739 (1997).

7 May, M. T. *et al.* Impact on life expectancy of HIV-1 positive individuals of CD4+ cell count and viral load response to antiretroviral therapy. *AIDS* **28**, 1193–1202, doi:10.1097/QAD.0000000000000243 (2014).

8 Burrell, C. J., Howard, C. R. & Murphy, F. A. *Fenner and Whites medical virology*. (Academic Press, 2016).

9 Narayan, O. & Clements, J. E. Biology and pathogenesis of lentiviruses. *J Gen Virol* **70 (Pt 7)**, 1617–1639, doi:10.1099/0022-1317-70-7-1617 (1989).

10 Weiss, R. A. The discovery of endogenous retroviruses. *Retrovirology* **3**, 1–11 (2006).

11 Clavel, F. *et al.* Isolation of a new human retrovirus from West African patients with AIDS. *Science* **233**, 343–346, doi:10.1126/science.2425430 (1986).

12 Campbell-Yesufu, O. T. & Gandhi, R. T. Update on human immunodeficiency virus (HIV)-2 infection. *Clin Infect Dis* **52**, 780–787, doi:10.1093/cid/ciq248 (2011).

13 Reeves, J. D. & Doms, R. W. Human immunodeficiency virus type 2. *J Gen Virol* **83**, 1253–1265, doi:10.1099/0022-1317-83-6-1253 (2002).

14 Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* **1**, a006841, doi:10.1101/cshperspect.a006841 (2011).

15 Keele, B. F. *et al.* Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**, 523–526, doi:10.1126/science.1126531 (2006).

16 Moir, S., Chun, T. W. & Fauci, A. S. Pathogenic mechanisms of HIV disease. *Annu Rev Pathol* **6**, 223–248, doi:10.1146/annurev–pathol–011110–130254 (2011).

17 Deeks, S. G., Overbaugh, J., Phillips, A. & Buchbinder, S. HIV infection. *Nature Reviews Disease Primers* **1**, 15035, doi:10.1038/nrdp.2015.35 (2015).

18 Haase, A. T. Perils at mucosal front lines for HIV and SIV and their hosts. *Nat Rev Immunol* **5**, 783–792, doi:10.1038/nri1706 (2005).

19 Piatak, M. *et al.* High levels of HIV–1 in plasma during all stages of infection determined by competitive PCR. *Science* **259**, 1749–1754 (1993).

20 Little, S. J., McLean, A. R., Spina, C. A., Richman, D. D. & Havlir, D. V. Viral dynamics of acute HIV–1 infection. *The Journal of experimental medicine* **190**, 841–850 (1999).

21 Mellors, J. W. *et al.* Prognosis in HIV–1 infection predicted by the quantity of virus in plasma. *Science* **272**, 1167–1170, doi:10.1126/science.272.5265.1167 (1996).

22 Buzon, M. J. *et al.* HIV–1 persistence in CD4+ T cells with stem cell–like properties. *Nature medicine* **20**, 139–142 (2014).

23 Chomont, N. *et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med* **15**, 893–900, doi:10.1038/nm.1972 (2009).

24 Davey, R. T. *et al.* HIV–1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proceedings of the National Academy of Sciences* **96**, 15109–15114 (1999).

25 Wei, X. *et al.* Antibody neutralization and escape by HIV–1. *Nature* **422**, 307–312 (2003).

26 Richman, D. D., Wrin, T., Little, S. J. & Petropoulos, C. J. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proceedings of the National Academy of Sciences* **100**, 4144–4149 (2003).

27 Deng, K. *et al.* Broad CTL response is required to clear latent HIV–1 due to dominance of escape mutations. *Nature* **517**, 381–385 (2015).

28 Deeks, S. G. *et al.* Immune activation set point during early HIV infection predicts subsequent CD4+ T–cell changes independent of viral load. *Blood* **104**, 942–947 (2004).

29 Deeks, S. G., Lewin, S. R. & Havlir, D. V. The end of AIDS: HIV infection as a chronic disease. *The Lancet* **382**, 1525–1533 (2013).

30 Jacks, T. *et al.* Characterization of Ribosomal Frameshifting in Hiv–1 Gag–Pol Expression. *Nature* **331**, 280–283, doi:DOI 10.1038/331280a0 (1988).

31 Frankel, A. D. & Young, J. A. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* **67**, 1-25, doi:10.1146/annurev.biochem.67.1.1 (1998).

32 Swanson, C. M. & Malim, M. H. SnapShot: HIV-1 proteins. *Cell* **133**, 742, 742 e741, doi:10.1016/j.cell.2008.05.005 (2008).

33 Sundquist, W. I. & Krausslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med* **2**, a006924, doi:10.1101/cshperspect.a006924 (2012).

34 Kotov, A., Zhou, J., Flicker, P. & Aiken, C. Association of Nef with the human immunodeficiency virus type 1 core. *J Virol* **73**, 8824-8830 (1999).

35 Accola, M. A., Ohagen, A. & Gottlinger, H. G. Isolation of human immunodeficiency virus type 1 cores: retention of Vpr in the absence of p6(gag). *J Virol* **74**, 6198-6202, doi:10.1128/jvi.74.13.6198-6202.2000 (2000).

36 Chen, B. Molecular Mechanism of HIV-1 Entry. *Trends in Microbiology* **27**, 878-891, doi:10.1016/j.tim.2019.06.002 (2019).

37 Campbell, E. M. & Hope, T. J. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Microbiol* **13**, 471-483, doi:10.1038/nrmicro3503 (2015).

38 Zila, V. *et al.* Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell* **184**, 1032-1046 e1018, doi:10.1016/j.cell.2021.01.025 (2021).

39 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, doi:10.1038/s41564-020-0735-8 (2020).

40 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, doi:10.1073/pnas.1920631117 (2020).

41 Freed, E. O. HIV-1 assembly, release and maturation. *Nature Reviews Microbiology* **13**, 484-496 (2015).

42 Chen, J. *et al.* Visualizing the translation and packaging of HIV-1 full-length RNA. *Proceedings of the National Academy of Sciences* **117**, 6145-6155 (2020).

43 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* **283**, 80-83, doi:10.1126/science.283.5398.80 (1999).

44 Li, S., Hill, C. P., Sundquist, W. I. & Finch, J. T. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409-413, doi:10.1038/35030177 (2000).

45 Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K. & Freed, E. O. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 gag targeting to the plasma membrane. *Mol Biol Cell* **15**, 122a-123a (2004).

46 Saad, J. S. *et al.* Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *P Natl Acad Sci USA* **103**, 11364-11369, doi:10.1073/pnas.0602818103 (2006).

47 Freed, E. O., Englund, G. & Martin, M. A. Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. *J Virol* **69**, 3949-3954, doi:10.1128/JVI.69.6.3949-3954.1995 (1995).

48 Tang, C. *et al.* Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proc Natl Acad Sci U S A* **101**, 517-522, doi:10.1073/pnas.0305665101 (2004).

49 Jacques, D. A. *et al.* HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349-353, doi:10.1038/nature19098 (2016).

50 Briggs, J. A., 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, doi:10.1093/emboj/cdg143 (2003).

51 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).

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

53 Mattei, S. *et al.* High-resolution structures of HIV-1 Gag cleavage mutants determine structural switch for virus maturation. *Proc Natl Acad Sci U S A* **115**, E9401-E9410, doi:10.1073/pnas.1811237115 (2018).

54 James, L. C. & Jacques, D. A. The human immunodeficiency virus capsid is more than just a genome package. *Annual review of virology* **5**, 209-225 (2018).

55 Schur, F. K. M. *et al.* Structure of the immature HIV-1 capsid in intact virus particles at 8.8 angstrom resolution. *Nature* **517**, 505-508, doi:10.1038/nature13838 (2015).

56 Wiegers, K. *et al.* Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J Virol* **72**, 2846-2854, doi:10.1128/JVI.72.4.2846-2854.1998 (1998).

57 Pettit, S. C., Lindquist, J. N., Kaplan, A. H. & Swanstrom, R. Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. *Retrovirology* **2**, 66, doi:10.1186/1742-4690-2-66 (2005).

58 Purdy, M. D. *et al.* MicroED structures of HIV-1 Gag CTD-SP1 reveal binding interactions with the maturation inhibitor bevirimat. *Proc Natl Acad Sci U S A* **115**, 13258-13263, doi:10.1073/pnas.1806806115 (2018).

59 Keller, P. W., Adamson, C. S., Heymann, J. B., Freed, E. O. & Steven, A. C. HIV-1 maturation inhibitor bevirimat stabilizes the immature Gag lattice. *J Virol* **85**, 1420-1428, doi:10.1128/JVI.01926-10 (2011).

60 Nguyen, A. T. *et al.* The prototype HIV-1 maturation inhibitor, bevirimat, binds to the CA-SP1 cleavage site in immature Gag particles. *Retrovirology* **8**, 101, doi:10.1186/1742-4690-8-101 (2011).

61 De Guzman, R. N. *et al.* Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **279**, 384-388, doi:10.1126/science.279.5349.384 (1998).

62 Gorelick, R. J. *et al.* Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J Virol* **64**, 3207-3211, doi:10.1128/JVI.64.7.3207-3211.1990 (1990).

63 Berkowitz, R. D., Luban, J. & Goff, S. P. Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. *J Virol* **67**, 7190-7200, doi:10.1128/JVI.67.12.7190-7200.1993 (1993).

64 Lu, K. *et al.* NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. *Science* **334**, 242-245, doi:10.1126/science.1210460 (2011).

65 DSouza, V. & Summers, M. F. How retroviruses select their genomes. *Nat Rev Microbiol* **3**, 643-655, doi:10.1038/nrmicro1210 (2005).

66 Pak, A. J. *et al.* Immature HIV-1 lattice assembly dynamics are regulated by scaffolding from nucleic acid and the plasma membrane. *Proc Natl Acad Sci U S A* **114**, E10056-E10065, doi:10.1073/pnas.1706600114 (2017).

67 Mirambeau, G. *et al.* Transmission electron microscopy reveals an optimal HIV-1 nucleocapsid aggregation with single-stranded nucleic acids and the mature HIV-1 nucleocapsid protein. *J Mol Biol* **364**, 496-511, doi:10.1016/j.jmb.2006.08.065 (2006).

68 Mougel, M., Houzet, L. & Darlix, J. L. When is it time for reverse transcription to start and go? *Retrovirology* **6**, 24, doi:10.1186/1742-4690-6-24 (2009).

69 Briggs, J. A. & Krausslich, H. G. The molecular architecture of HIV. *J Mol Biol* **410**, 491-500, doi:10.1016/j.jmb.2011.04.021 (2011).

70 Barat, C. *et al.* HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO J* **8**, 3279-3285 (1989).

71 Wu, T. *et al.* Fundamental differences between the nucleic acid chaperone activities of HIV-1 nucleocapsid protein and Gag or Gag-derived proteins: biological implications. *Virology* **405**, 556-567, doi:10.1016/j.virol.2010.06.042 (2010).

72 Thomas, J. A., Bosche, W. J., Shatzer, T. L., Johnson, D. G. & Gorelick, R. J. Mutations in human immunodeficiency virus type 1 nucleocapsid protein zinc fingers cause premature reverse transcription. *J Virol* **82**, 9318–9328, doi:10.1128/JVI.00583-08 (2008).

73 Allain, B., Lapadat-Tapolsky, M., Berlioz, C. & Darlix, J. L. Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *EMBO J* **13**, 973–981 (1994).

74 Onafuwa-Nuga, A. & Telesnitsky, A. The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiol Mol Biol Rev* **73**, 451–480, Table of Contents, doi:10.1128/MMBR.00012-09 (2009).

75 Negroni, M. & Buc, H. Mechanisms of retroviral recombination. *Annu Rev Genet* **35**, 275–302, doi:10.1146/annurev.genet.35.102401.090551 (2001).

76 de Marco, A. *et al.* Role of the SP2 domain and its proteolytic cleavage in HIV-1 structural maturation and infectivity. *J Virol* **86**, 13708–13716, doi:10.1128/JVI.01704-12 (2012).

77 Gottlinger, H. G., Dorfman, T., Sodroski, J. G. & Haseltine, W. A. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc Natl Acad Sci U S A* **88**, 3195–3199, doi:10.1073/pnas.88.8.3195 (1991).

78 Garrus, J. E. *et al.* Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**, 55–65, doi:10.1016/s0092-8674(01)00506-2 (2001).

79 Demirov, D. G., Ono, A., Orenstein, J. M. & Freed, E. O. Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. *Proc Natl Acad Sci U S A* **99**, 955–960, doi:10.1073/pnas.032511899 (2002).

80 Strack, B., Calistri, A., Craig, S., Popova, E. & Gottlinger, H. G. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell* **114**, 689–699, doi:10.1016/s0092-8674(03)00653-6 (2003).

81 Fisher, R. D. *et al.* Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding. *Cell* **128**, 841–852, doi:10.1016/j.cell.2007.01.035 (2007).

82 Lippincott-Schwartz, J., Freed, E. O. & van Engelenburg, S. B. A Consensus View of ESCRT-Mediated Human Immunodeficiency Virus Type 1 Abscission. *Annu Rev Virol* **4**, 309–325, doi:10.1146/annurev-virology-101416-041840 (2017).

83 Graves, M. C., Lim, J. J., Heimer, E. P. & Kramer, R. A. An 11-kDa form of human immunodeficiency virus protease expressed in Escherichia coli is sufficient for enzymatic activity. *Proc Natl Acad Sci U S A* **85**, 2449–2453, doi:10.1073/pnas.85.8.2449 (1988).

84 Brik, A. & Wong, C. H. HIV-1 protease: mechanism and drug discovery. *Org Biomol Chem* **1**, 5-14, doi:10.1039/b208248a (2003).

85 Pettit, S. C., Everitt, L. E., Choudhury, S., Dunn, B. M. & Kaplan, A. H. Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. *J Virol* **78**, 8477-8485, doi:10.1128/JVI.78.16.8477-8485.2004 (2004).

86 Agniswamy, J., Sayer, J. M., Weber, I. T. & Louis, J. M. Terminal interface conformations modulate dimer stability prior to amino terminal autoprocessing of HIV-1 protease. *Biochemistry* **51**, 1041-1050, doi:10.1021/bi201809s (2012).

87 Tang, C., Louis, J. M., Aniana, A., Suh, J. Y. & Clore, G. M. Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature* **455**, 693-696, doi:10.1038/nature07342 (2008).

88 Mattei, S. *et al.* Induced maturation of human immunodeficiency virus. *J Virol* **88**, 13722-13731, doi:10.1128/JVI.02271-14 (2014).

89 Ghosh, A. K., Osswald, H. L. & Prato, G. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *J Med Chem* **59**, 5172-5208, doi:10.1021/acs.jmedchem.5b01697 (2016).

90 Kohlstaedt, L. A. & Steitz, T. A. Reverse transcriptase of human immunodeficiency virus can use either human tRNA(3Lys) or Escherichia coli tRNA(2Gln) as a primer in an in vitro primer-utilization assay. *Proc Natl Acad Sci U S A* **89**, 9652-9656, doi:10.1073/pnas.89.20.9652 (1992).

91 Jacobo-Molina, A. *et al.* Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc Natl Acad Sci U S A* **90**, 6320-6324, doi:10.1073/pnas.90.13.6320 (1993).

92 Hu, W.-S. & Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harbor perspectives in medicine* **2**, a006882 (2012).

93 Hughes, S. H. Reverse transcription of retroviruses and LTR retrotransposons. *Mobile DNA III*, 1051-1077 (2015).

94 Ilina, T., LaBarge, K., Sarafianos, S. G., Ishima, R. & Parniak, M. A. Inhibitors of HIV-1 reverse transcriptase—associated ribonuclease H activity. *Biology* **1**, 521-541 (2012).

95 Esposito, F., Corona, A. & Tramontano, E. HIV-1 Reverse Transcriptase Still Remains a New Drug Target: Structure, Function, Classical Inhibitors, and New Inhibitors with Innovative Mechanisms of Actions. *Mol Biol Int* **2012**, 586401, doi:10.1155/2012/586401 (2012).

96 Schroder, A. R. *et al.* HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521-529, doi:10.1016/s0092-8674(02)00864-4 (2002).

97 Maertens, G. N., Engelman, A. N. & Cherepanov, P. Structure and function of retroviral integrase. *Nat Rev Microbiol*, doi:10.1038/s41579-021-00586-9 (2021).

98 Lesbats, P., Engelman, A. N. & Cherepanov, P. Retroviral DNA Integration. *Chem Rev* **116**, 12730-12757, doi:10.1021/acs.chemrev.6b00125 (2016).

99 Kessl, J. J. *et al.* HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. *Cell* **166**, 1257-1268 e1212, doi:10.1016/j.cell.2016.07.044 (2016).

100 Scarsi, K. K., Havens, J. P., Podany, A. T., Avedissian, S. N. & Fletcher, C. V. HIV-1 Integrase Inhibitors: A Comparative Review of Efficacy and Safety. *Drugs* **80**, 1649-1676, doi:10.1007/s40265-020-01379-9 (2020).

101 Willey, R. L., Bonifacino, J. S., Potts, B. J., Martin, M. A. & Klausner, R. D. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. *Proc Natl Acad Sci U S A* **85**, 9580-9584, doi:10.1073/pnas.85.24.9580 (1988).

102 Hallenberger, S. *et al.* Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature* **360**, 358-361, doi:10.1038/360358a0 (1992).

103 Kwong, P. D. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648-659, doi:10.1038/31405 (1998).

104 Berger, E. A. *et al.* A new classification for HIV-1. *Nature* **391**, 240, doi:10.1038/34571 (1998).

105 Chan, D. C. & Kim, P. S. HIV entry and its inhibition. *Cell* **93**, 681-684, doi:10.1016/s0092-8674(00)81430-0 (1998).

106 Berger, E. A., Murphy, P. M. & Farber, J. M. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol* **17**, 657-700, doi:10.1146/annurev.immunol.17.1.657 (1999).

107 Daecke, J., Fackler, O. T., Dittmar, M. T. & Krausslich, H. G. Involvement of clathrin-mediated endocytosis in human immunodeficiency virus type 1 entry. *J Virol* **79**, 1581-1594, doi:10.1128/JVI.79.3.1581-1594.2005 (2005).

108 Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V. & Melikyan, G. B. HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. *Cell* **137**, 433-444, doi:10.1016/j.cell.2009.02.046 (2009).

109 Foster, T. L. *et al.* Resistance of Transmitted Founder HIV-1 to IFITM-Mediated Restriction. *Cell Host Microbe* **20**, 429-442, doi:10.1016/j.chom.2016.08.006 (2016).

110 Karn, J. & Stoltzfus, C. M. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med* **2**, a006916, doi:10.1101/cshperspect.a006916 (2012).

111 Dayton, A. I., Sodroski, J. G., Rosen, C. A., Goh, W. C. & Haseltine, W. A. The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* **44**, 941-947, doi:10.1016/0092-8674(86)90017-6 (1986).

112 Kao, S. Y., Calman, A. F., Luciw, P. A. & Peterlin, B. M. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* **330**, 489-493, doi:10.1038/330489a0 (1987).

113 Parada, C. A. & Roeder, R. G. Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature* **384**, 375-378, doi:10.1038/384375a0 (1996).

114 Yamada, T. *et al.* P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* **21**, 227-237, doi:10.1016/j.molcel.2005.11.024 (2006).

115 Fujinaga, K. *et al.* Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol* **24**, 787-795, doi:10.1128/MCB.24.2.787-795.2004 (2004).

116 Rice, A. P. The HIV-1 Tat Protein: Mechanism of Action and Target for HIV-1 Cure Strategies. *Curr Pharm Des* **23**, 4098-4102, doi:10.2174/1381612823666170704130635 (2017).

117 Ajasin, D. & Eugenin, E. A. HIV-1 Tat: Role in Bystander Toxicity. *Front Cell Infect Microbiol* **10**, 61, doi:10.3389/fcimb.2020.00061 (2020).

118 Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V. & Cullen, B. R. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**, 254-257, doi:10.1038/338254a0 (1989).

119 Heaphy, S. *et al.* HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. *Cell* **60**, 685-693, doi:10.1016/0092-8674(90)90671-z (1990).

120 Mann, D. A. *et al.* A molecular rheostat. Co-operative rev binding to stem I of the rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J Mol Biol* **241**, 193-207, doi:10.1006/jmbi.1994.1488 (1994).

121 Daugherty, M. D., Booth, D. S., Jayaraman, B., Cheng, Y. & Frankel, A. D. HIV Rev response element (RRE) directs assembly of the Rev homooligomer into discrete asymmetric complexes. *Proc Natl Acad Sci U S A* **107**, 12481-12486, doi:10.1073/pnas.1007022107 (2010).

122 Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. & Luhrmann, R. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used

by specific cellular RNAs. *Cell* **82**, 475–483, doi:10.1016/0092-8674(95)90436-0 (1995).

123 Fackler, O. T. *et al.* Association of human immunodeficiency virus Nef protein with actin is myristoylation dependent and influences its subcellular localization. *Eur J Biochem* **247**, 843–851, doi:10.1111/j.1432-1033.1997.00843.x (1997).

124 Kohleisen, B. *et al.* Cellular localization of Nef expressed in persistently HIV-1-infected low-producer astrocytes. *AIDS* **6**, 1427–1436, doi:10.1097/00002030-199212000-00002 (1992).

125 Greenberg, M. E. *et al.* Co-localization of HIV-1 Nef with the AP-2 adaptor protein complex correlates with Nef-induced CD4 down-regulation. *EMBO J* **16**, 6964-6976, doi:10.1093/emboj/16.23.6964 (1997).

126 Pereira, E. A. & daSilva, L. L. HIV-1 Nef: Taking Control of Protein Trafficking. *Traffic* **17**, 976–996, doi:10.1111/tra.12412 (2016).

127 Buffalo, C. Z., Iwamoto, Y., Hurley, J. H. & Ren, X. How HIV Nef Proteins Hijack Membrane Traffic To Promote Infection. *J Virol* **93**, doi:10.1128/JVI.01322-19 (2019).

128 Chaipan, C., Smith, J. L., Hu, W. S. & Pathak, V. K. APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages. *J Virol* **87**, 444–453, doi:10.1128/JVI.00676-12 (2013).

129 Lecossier, D., Bouchonnet, F., Clavel, F. & Hance, A. J. Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* **300**, 1112, doi:10.1126/science.1083338 (2003).

130 Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98, doi:10.1038/nature01707 (2003).

131 Marin, M., Rose, K. M., Kozak, S. L. & Kabat, D. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat Med* **9**, 1398–1403, doi:10.1038/nm946 (2003).

132 Yu, X. *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056–1060, doi:10.1126/science.1089591 (2003).

133 Sheehy, A. M., Gaddis, N. C. & Malim, M. H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med* **9**, 1404–1407, doi:10.1038/nm945 (2003).

134 Guo, Y. *et al.* Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* **505**, 229–233, doi:10.1038/nature12884 (2014).

135 Selig, L. *et al.* Interaction with the p6 domain of the gag precursor mediates incorporation into virions of Vpr and Vpx proteins from primate lentiviruses. *J Virol* **73**, 592–600, doi:10.1128/JVI.73.1.592-600.1999 (1999).

136 Paxton, W., Connor, R. I. & Landau, N. R. Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. *J Virol* **67**, 7229–7237, doi:10.1128/JVI.67.12.7229-7237.1993 (1993).

137 Greenwood, E. J. D. *et al.* Promiscuous targeting of cellular proteins by Vpr drives systems-level proteomic remodeling in HIV-1 infection. *Cell reports* **27**, 1579-1596 (2019).

138 Khan, H. *et al.* HIV-1 Vpr antagonizes innate immune activation by targeting karyopherin-mediated NF-κB/IRF3 nuclear transport. *Elife* **9**, e60821 (2020).

139 Guenzel, C. A., Herate, C. & Benichou, S. HIV-1 Vpr-a still "enigmatic multitasker". *Front Microbiol* **5**, 127, doi:10.3389/fmicb.2014.00127 (2014).

140 Nodder, S. B. & Gummuluru, S. Illuminating the Role of Vpr in HIV Infection of Myeloid Cells. *Front Immunol* **10**, 1606, doi:10.3389/fimmu.2019.01606 (2019).

141 Cohen, E. A., Terwilliger, E. F., Sodroski, J. G. & Haseltine, W. A. Identification of a protein encoded by the vpu gene of HIV-1. *Nature* **334**, 532-534, doi:10.1038/334532a0 (1988).

142 Willey, R. L., Maldarelli, F., Martin, M. A. & Strebel, K. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J Virol* **66**, 7193-7200, doi:10.1128/JVI.66.12.7193-7200.1992 (1992).

143 Strebel, K., Klimkait, T., Maldarelli, F. & Martin, M. A. Molecular and biochemical analyses of human immunodeficiency virus type 1 vpu protein. *J Virol* **63**, 3784-3791, doi:10.1128/JVI.63.9.3784-3791.1989 (1989).

144 Margottin, F. *et al.* A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol Cell* **1**, 565-574, doi:10.1016/s1097-2765(00)80056-8 (1998).

145 Douglas, J. L. *et al.* Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta}TrCP-dependent mechanism. *J Virol* **83**, 7931-7947, doi:10.1128/JVI.00242-09 (2009).

146 Neil, S. J., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425-430, doi:10.1038/nature06553 (2008).

147 Khan, N. & Geiger, J. D. Role of Viral Protein U (Vpu) in HIV-1 Infection and Pathogenesis. *Viruses* **13**, doi:10.3390/v13081466 (2021).

148 Kumar, V., Sinha, A. K., Makkar, H. P. S. & Becker, K. Dietary roles of phytate and phytase in human nutrition: A review. *Food chemistry* **120**, 945-959 (2010).

149 Irvine, R. F. & Schell, M. J. Back in the water: the return of the inositol phosphates. *Nature reviews Molecular cell biology* **2**, 327-338 (2001).

150 Sauer, K. & Cooke, M. P. Regulation of immune cell development through soluble inositol–1, 3, 4, 5–tetrakisphosphate. *Nature Reviews Immunology* **10**, 257–271 (2010).

151 Shears, S. B. The versatility of inositol phosphates as cellular signals. *Biochimica et Biophysica Acta (BBA)–Molecular and Cell Biology of Lipids* **1436**, 49–67 (1998).

152 Shears, S. B. Assessing the omnipotence of inositol hexakisphosphate. *Cellular signalling* **13**, 151–158 (2001).

153 Macbeth, M. R. *et al.* Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**, 1534–1539 (2005).

154 Hanakahi, L. A., Bartlet–Jones, M., Chappell, C., Pappin, D. & West, S. C. Binding of inositol phosphate to DNA–PK and stimulation of double–strand break repair. *Cell* **102**, 721–729 (2000).

155 Qiu, D. *et al.* Analysis of inositol phosphate metabolism by capillary electrophoresis electrospray ionization mass spectrometry. *Nat Commun* **11**, 6035, doi:10.1038/s41467–020–19928–x (2020).

156 Chakraborty, A., Kim, S. & Snyder, S. H. Inositol pyrophosphates as mammalian cell signals. *Sci Signal* **4**, re1, doi:10.1126/scisignal.2001958 (2011).

157 Raboy, V. Approaches and challenges to engineering seed phytate and total phosphorus. *Plant Science* **177**, 281–296 (2009).

158 Holub, B. J. Metabolism and function of myo–inositol and inositol phospholipids. *Annu Rev Nutr* **6**, 563–597, doi:10.1146/annurev.nu.06.070186.003023 (1986).

159 Malabanan, M. M. & Blind, R. D. Inositol polyphosphate multikinase (IPMK) in transcriptional regulation and nuclear inositide metabolism. *Biochemical Society Transactions* **44**, 279–285 (2016).

160 Kim, E., Ahn, H., Kim, M. G., Lee, H. & Kim, S. The expanding significance of inositol polyphosphate multikinase as a signaling hub. *Molecules and cells* **40**, 315 (2017).

161 Yang, X. & Shears, S. B. Multitasking in signal transduction by a promiscuous human Ins(3, 4, 5, 6)P4 1–kinase/Ins(1, 3, 4)P3 5/6–kinase. *Biochemical Journal* **351**, 551–555 (2000).

162 Wilson, M. P. & Majerus, P. W. Isolation of Inositol 1, 3, 4–Trisphosphate 5/6–Kinase, cDNA Cloning, and Expression of the Recombinant Enzyme (*). *Journal of Biological Chemistry* **271**, 11904–11910 (1996).

163 Saiardi, A. & Cockcroft, S. Human ITPK1: a reversible inositol phosphate kinase/phosphatase that links receptor–dependent phospholipase C to Ca2+–activated chloride channels. *Science signaling* **1**, pe5–pe5 (2008).

164 Verbsky, J. W., Wilson, M. P., Kisseleva, M. V., Majerus, P. W. & Wente, S. R. The synthesis of inositol hexakisphosphate: characterization of human inositol 1, 3, 4,

5, 6–pentakisphosphate 2–kinase. *Journal of Biological Chemistry* **277**, 31857–31862 (2002).

165 Monserrate, J. P. & York, J. D. Inositol phosphate synthesis and the nuclear processes they affect. *Current opinion in cell biology* **22**, 365–373 (2010).

166 Stephens, L. *et al.* The detection, purification, structural characterization, and metabolism of diphosphoinositol pentakisphosphate (s) and bisdiphosphoinositol tetrakisphosphate (s). *Journal of Biological Chemistry* **268**, 4009–4015 (1993).

167 Menniti, F. S., Miller, R. N., Putney Jr, J. W. & Shears, S. B. Turnover of inositol polyphosphate pyrophosphates in pancreatoma cells. *Journal of Biological Chemistry* **268**, 3850–3856 (1993).

168 Chi, H. *et al.* Targeted deletion of Minpp1 provides new insight into the activity of multiple inositol polyphosphate phosphatase in vivo. *Mol Cell Biol* **20**, 6496–6507, doi:10.1128/MCB.20.17.6496–6507.2000 (2000).

169 Windhorst, S. *et al.* Tumour cells can employ extracellular Ins (1, 2, 3, 4, 5, 6) P 6 and multiple inositol–polyphosphate phosphatase 1 (MINPP1) dephosphorylation to improve their proliferation. *Biochemical Journal* **450**, 115–125 (2013).

170 Kilaparty, S. P., Agarwal, R., Singh, P., Kannan, K. & Ali, N. Endoplasmic reticulum stress–induced apoptosis accompanies enhanced expression of multiple inositol polyphosphate phosphatase 1 (Minpp1): a possible role for Minpp1 in cellular stress response. *Cell Stress and Chaperones* **21**, 593–608 (2016).

171 Karlsson, M. *et al.* A single–cell type transcriptomics map of human tissues. *Science Advances* **7**, eabh2169 (2021).

172 Thul, P. J. *et al.* A subcellular map of the human proteome. *Science* **356** (2017).

173 Atlas, H. P. *The Human Protein Atlas –IPPK*, <https://www.proteinatlas.org/ENSG00000127080–IPPK/single+cell+type/lymph+node> (

174 Atlas, H. P. The Human Protein Atlas –IPMK.

175 Atlas, H. P. The Human Protein Atlas –MINPP1.

176 Brehm, M. A. *et al.* Intracellular localization of human Ins (1, 3, 4, 5, 6) P 5 2–kinase. *Biochemical Journal* **408**, 335–345 (2007).

177 Nalaskowski, M. M., Deschermeier, C., Fanick, W. & Mayr, G. W. The human homologue of yeast ArgRIII protein is an inositol phosphate multikinase with predominantly nuclear localization. *Biochemical Journal* **366**, 549–556 (2002).

178 Ali, N., Craxton, A. & Shears, S. B. Hepatic Ins (1, 3, 4, 5) P4 3–phosphatase is compartmentalized inside endoplasmic reticulum. *Journal of Biological Chemistry* **268**, 6161–6167 (1993).

179 Caffrey, J. J., Hidaka, K., Matsuda, M., Hirata, M. & Shears, S. B. The human and rat forms of multiple inositol polyphosphate phosphatase: functional homology with

a histidine acid phosphatase up-regulated during endochondral ossification. *FEBS letters* **442**, 99-104 (1999).

180 Chi, H. *et al.* Targeted deletion of Minpp1 provides new insight into the activity of multiple inositol polyphosphate phosphatase in vivo. *Molecular and cellular biology* **20**, 6496-6507 (2000).

181 Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K. & Freed, E. O. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc Natl Acad Sci U S A* **101**, 14889-14894, doi:10.1073/pnas.0405596101 (2004).

182 Chukkapalli, V., Hogue, I. B., Boyko, V., Hu, W. S. & Ono, A. Interaction between the human immunodeficiency virus type 1 Gag matrix domain and phosphatidylinositol-(4,5)-bisphosphate is essential for efficient gag membrane binding. *J Virol* **82**, 2405-2417, doi:10.1128/JVI.01614-07 (2008).

183 Mucksch, F., Laketa, V., Muller, B., Schultz, C. & Krausslich, H. G. Synchronized HIV assembly by tunable PIP2 changes reveals PIP2 requirement for stable Gag anchoring. *Elife* **6**, doi:10.7554/eLife.25287 (2017).

184 DSouza, V. & Summers, M. F. How retroviruses select their genomes. *Nat Rev Microbiol* **3**, 643-655, doi:10.1038/nrmicro1210 (2005).

185 Strambio-de-Castillia, C. & Hunter, E. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. *Journal of Virology* **66**, 7021-7032 (1992).

186 Mammano, F., Ohagen, A., Höglund, S. & Göttlinger, H. G. Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *Journal of virology* **68**, 4927-4936 (1994).

187 Schur, F. K. *et al.* Structure of the immature HIV-1 capsid in intact virus particles at 8.8 A resolution. *Nature* **517**, 505-508, doi:10.1038/nature13838 (2015).

188 Schur, F. K. *et al.* An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science* **353**, 506-508, doi:10.1126/science.aaf9620 (2016).

189 Campbell, S. *et al.* Modulation of HIV-like particle assembly in vitro by inositol phosphates. *Proceedings of the National Academy of Sciences* **98**, 10875-10879 (2001).

190 Dick, R. A. *et al.* Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509-512, doi:10.1038/s41586-018-0396-4 (2018).

191 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 e3984, doi:10.1016/j.celrep.2019.11.050 (2019).

192 Sowd, G. A. & Aiken, C. Inositol phosphates promote HIV-1 assembly and maturation to facilitate viral spread in human CD4+ T cells. *PLoS Pathog* **17**, e1009190, doi:10.1371/journal.ppat.1009190 (2021).

193 Mendonca, L. *et al.* CryoET structures of immature HIV Gag reveal six-helix bundle. *Commun Biol* **4**, 481, doi:10.1038/s42003-021-01999-1 (2021).

194 Pettit, S. C. *et al.* The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J Virol* **68**, 8017-8027, doi:10.1128/JVI.68.12.8017-8027.1994 (1994).

195 Pettit, S. C., Sheng, N., Tritch, R., Erickson-Viitanen, S. & Swanstrom, R. The regulation of sequential processing of HIV-1 Gag by the viral protease. *Adv Exp Med Biol* **436**, 15-25, doi:10.1007/978-1-4615-5373-1_2 (1998).

196 Pettit, S. C., Henderson, G. J., Schiffer, C. A. & Swanstrom, R. Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. *J Virol* **76**, 10226-10233, doi:10.1128/jvi.76.20.10226-10233.2002 (2002).

197 Gitti, R. K. *et al.* Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**, 231-235, doi:10.1126/science.273.5272.231 (1996).

198 von Schwedler, U. K. *et al.* Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J* **17**, 1555-1568, doi:10.1093/emboj/17.6.1555 (1998).

199 Benjamin, J., Ganser-Pornillos, B. K., Tivol, W. F., Sundquist, W. I. & Jensen, G. J. Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography. *J Mol Biol* **346**, 577-588, doi:10.1016/j.jmb.2004.11.064 (2005).

200 Kaplan, A. H. *et al.* Partial inhibition of the human immunodeficiency virus type 1 protease results in aberrant virus assembly and the formation of noninfectious particles. *J Virol* **67**, 4050-4055, doi:10.1128/JVI.67.7.4050-4055.1993 (1993).

201 Checkley, M. A., Luttge, B. G., Soheilian, F., Nagashima, K. & Freed, E. O. The capsid-spacer peptide 1 Gag processing intermediate is a dominant-negative inhibitor of HIV-1 maturation. *Virology* **400**, 137-144, doi:10.1016/j.virol.2010.01.028 (2010).

202 Briggs, J. A. G. *et al.* The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. *Structure* **14**, 15-20 (2006).

203 Keller, P. W. *et al.* A two-pronged structural analysis of retroviral maturation indicates that core formation proceeds by a disassembly-reassembly pathway rather than a displacive transition. *Journal of virology* **87**, 13655-13664 (2013).

204 Woodward, C. L., Cheng, S. N. & Jensen, G. J. Electron cryotomography studies of maturing HIV-1 particles reveal the assembly pathway of the viral core. *Journal of virology* **89**, 1267-1277 (2015).

205 Frank, G. A. *et al.* Maturation of the HIV-1 core by a non-diffusional phase transition. *Nature communications* **6**, 1-9 (2015).

206 Ning, J. *et al.* In vitro protease cleavage and computer simulations reveal the HIV-1 capsid maturation pathway. *Nature communications* **7**, 1-12 (2016).

207 Rihn, S. J. *et al.* Extreme genetic fragility of the HIV-1 capsid. *PLoS pathogens* **9** (2013).

208 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, doi:10.1128/jvi.76.11.5667-5677.2002 (2002).

209 Yang, R. *et al.* Second-site suppressors of HIV-1 capsid mutations: restoration of intracellular activities without correction of intrinsic capsid stability defects. *Retrovirology* **9**, 30, doi:10.1186/1742-4690-9-30 (2012).

210 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, doi:10.1016/j.immuni.2013.11.002 (2013).

211 Rasaiyaah, J. *et al.* HIV-1 evades innate immune recognition through specific cofactor recruitment. *Nature* **503**, 402-405, doi:10.1038/nature12769 (2013).

212 Gao, D. *et al.* Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science* **341**, 903-906, doi:10.1126/science.1240933 (2013).

213 Huber, H. E., McCoy, J. M., Seehra, J. S. & Richardson, C. C. Human immunodeficiency virus 1 reverse transcriptase. Template binding, processivity, strand displacement synthesis, and template switching. *J Biol Chem* **264**, 4669-4678 (1989).

214 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* **370**, doi:10.1126/science.abc8420 (2020).

215 Dharan, A. & Campbell, E. M. Role of Microtubules and Microtubule-Associated Proteins in HIV-1 Infection. *J Virol* **92**, doi:10.1128/JVI.00085-18 (2018).

216 McDonald, D. *et al.* Visualization of the intracellular behavior of HIV in living cells. *The Journal of cell biology* **159**, 441-452 (2002).

217 Dharan, A. *et al.* Bicaudal D2 facilitates the cytoplasmic trafficking and nuclear import of HIV-1 genomes during infection. *Proc Natl Acad Sci U S A* **114**, E10707-E10716, doi:10.1073/pnas.1712033114 (2017).

218 Fernandez, J. *et al.* Microtubule-associated proteins 1 (MAP1) promote human immunodeficiency virus type I (HIV-1) intracytoplasmic routing to the nucleus. *J Biol Chem* **290**, 4631-4646, doi:10.1074/jbc.M114.613133 (2015).

219 Malikov, V. *et al.* HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. *Nat Commun* **6**, 6660, doi:10.1038/ncomms7660 (2015).

220 Huang, P.-T. *et al.* FEZ1 is recruited to a conserved cofactor site on capsid to promote HIV-1 trafficking. *Cell reports* **28**, 2373-2385 (2019).

221 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, doi:10.1016/s0092-8674(00)81823-1 (1996).

222 Braaten, D., Franke, E. K. & Luban, J. Cyclophilin A is required for the replication of group M human immunodeficiency virus type 1 (HIV-1) and simian immunodeficiency virus SIV(CPZ)GAB but not group O HIV-1 or other primate immunodeficiency viruses. *J Virol* **70**, 4220-4227, doi:10.1128/JVI.70.7.4220-4227.1996 (1996).

223 Ni, T. *et al.* Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. *Nature structural & molecular biology* **27**, 855-862, doi:10.1038/s41594-020-0467-8 (2020).

224 Hatziioannou, T., Perez-Caballero, D., Cowan, S. & Bieniasz, P. D. Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. *J Virol* **79**, 176-183, doi:10.1128/JVI.79.1.176-183.2005 (2005).

225 De Iaco, A. & Luban, J. Cyclophilin A promotes HIV-1 reverse transcription but its effect on transduction correlates best with its effect on nuclear entry of viral cDNA. *Retrovirology* **11**, 11, doi:10.1186/1742-4690-11-11 (2014).

226 Braaten, D. & Luban, J. Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells. *EMBO J* **20**, 1300-1309, doi:10.1093/emboj/20.6.1300 (2001).

227 Schaller, T. *et al.* HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog* **7**, e1002439, doi:10.1371/journal.ppat.1002439 (2011).

228 Lee, K. *et al.* Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7**, 221-233, doi:10.1016/j.chom.2010.02.007 (2010).

229 Mallery, D. L. *et al.* IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *Elife* **7**, doi:10.7554/eLife.35335 (2018).

230 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**, doi:10.1128/mBio.02820-20 (2020).

231 Ganser-Pornillos, B. K. *et al.* Hexagonal assembly of a restricting TRIM5alpha protein. *Proc Natl Acad Sci U S A* **108**, 534-539, doi:10.1073/pnas.1013426108 (2011).

232 Li, Y. L. *et al.* Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. *Elife* **5**, doi:10.7554/eLife.16269 (2016).

233 Stremlau, M. *et al.* Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci U S A* **103**, 5514-5519, doi:10.1073/pnas.0509996103 (2006).

234 Grutter, M. G. & Luban, J. TRIM5 structure, HIV-1 capsid recognition, and innate immune signaling. *Curr Opin Virol* **2**, 142-150, doi:10.1016/j.coviro.2012.02.003 (2012).

235 Pertel, T. *et al.* TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* **472**, 361-365, doi:10.1038/nature09976 (2011).

236 Fletcher, A. J. *et al.* Trivalent RING Assembly on Retroviral Capsids Activates TRIM5 Ubiquitination and Innate Immune Signaling. *Cell Host Microbe* **24**, 761-775 e766, doi:10.1016/j.chom.2018.10.007 (2018).

237 Kim, K. *et al.* Cyclophilin A protects HIV-1 from restriction by human TRIM5alpha. *Nat Microbiol* **4**, 2044-2051, doi:10.1038/s41564-019-0592-5 (2019).

238 Selyutina, A. *et al.* Cyclophilin A prevents HIV-1 restriction in lymphocytes by blocking human TRIM5α binding to the viral core. *Cell reports* **30**, 3766-3777 (2020).

239 Goujon, C. *et al.* Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **502**, 559-562, doi:10.1038/nature12542 (2013).

240 Liu, Z. *et al.* The interferon-inducible MxB protein inhibits HIV-1 infection. *Cell Host Microbe* **14**, 398-410, doi:10.1016/j.chom.2013.08.015 (2013).

241 Kane, M. *et al.* MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature* **502**, 563-566, doi:10.1038/nature12653 (2013).

242 Lin, D. H. & Hoelz, A. The Structure of the Nuclear Pore Complex (An Update). *Annu Rev Biochem* **88**, 725-783, doi:10.1146/annurev-biochem-062917-011901 (2019).

243 Lee, K. *et al.* Flexible use of nuclear import pathways by HIV-1. *Cell host & microbe* **7**, 221-233 (2010).

244 Kane, M. *et al.* Nuclear pore heterogeneity influences HIV-1 infection and the antiviral activity of MX2. *Elife* **7**, doi:10.7554/eLife.35738 (2018).

245 Bichel, K. *et al.* HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. *Retrovirology* **10**, 81, doi:10.1186/1742-4690-10-81 (2013).

246 Engelman, A. N. HIV Capsid and Integration Targeting. *Viruses* **13**, 125 (2021).

247 Zila, V., Muller, T. G., Muller, B. & Krausslich, H. G. HIV-1 capsid is the key orchestrator of early viral replication. *PLoS Pathog* **17**, e1010109, doi:10.1371/journal.ppat.1010109 (2021).

248 Sukegawa, J. & Blobel, G. A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and faces the nucleoplasm. *Cell* **72**, 29-38, doi:10.1016/0092-8674(93)90047-t (1993).

249 Zwerger, M., Eibauer, M. & Medalia, O. Insights into the gate of the nuclear pore complex. *Nucleus* **7**, 1-7, doi:10.1080/19491034.2015.1130197 (2016).

250 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**, e1004459, doi:10.1371/journal.ppat.1004459 (2014).

251 Matreyek, K. A., Yucel, 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**, e1003693, doi:10.1371/journal.ppat.1003693 (2013).

252 Ruepp, M. D. *et al.* Mammalian pre-mRNA 3' end processing factor CF I m 68 functions in mRNA export. *Mol Biol Cell* **20**, 5211-5223, doi:10.1091/mbc.E09-05-0389 (2009).

253 Bejarano, D. A. *et al.* HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *Elife* **8**, doi:10.7554/eLife.41800 (2019).

254 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. *Nature Microbiology* **5**, 1088-1095 (2020).

255 Krishnan, L. *et al.* The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase. *J Virol* **84**, 397-406, doi:10.1128/JVI.01899-09 (2010).

256 Zhou, L. *et al.* Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS Pathog* **7**, e1002194, doi:10.1371/journal.ppat.1002194 (2011).

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

258 De Iaco, A. *et al.* TNPO3 protects HIV-1 replication from CPSF6-mediated capsid stabilization in the host cell cytoplasm. *Retrovirology* **10**, 20 (2013).

259 Maertens, G. N. *et al.* Structural basis for nuclear import of splicing factors by human Transportin 3. *Proceedings of the National Academy of Sciences* **111**, 2728-2733 (2014).

260 Rodríguez-Mora, S. *et al.* The mutation of Transportin 3 gene that causes limb girdle muscular dystrophy 1F induces protection against HIV-1 infection. *PLoS pathogens* **15**, e1007958 (2019).

261 Zhou, L. *et al.* Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS pathogens* **7**, e1002194 (2011).

262 Valle-Casuso, J. C. *et al.* TNPO3 is required for HIV-1 replication after nuclear import but prior to integration and binds the HIV-1 core. *Journal of virology* **86**, 5931-5936 (2012).

263 Muller, T. G. *et al.* HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *Elife* **10**, doi:10.7554/eLife.64776 (2021).

264 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**, E1054-1063, doi:10.1073/pnas.1524213113 (2016).

265 Yang, Y., Fricke, T. & Diaz-Griffero, F. Inhibition of reverse transcriptase activity increases stability of the HIV-1 core. *J Virol* **87**, 683-687, doi:10.1128/JVI.01228-12 (2013).

266 Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C. & Rousso, I. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J Virol* **91**, doi:10.1128/JVI.00289-17 (2017).

267 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*, doi:10.1128/JVI.00166-21 (2021).

268 Francis, A. C. *et al.* HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat Commun* **11**, 3505, doi:10.1038/s41467-020-17256-8 (2020).

269 Achuthan, V. *et al.* Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral DNA Integration. *Cell Host Microbe* **24**, 392-404 e398, doi:10.1016/j.chom.2018.08.002 (2018).

270 Datta, S. A. K. *et al.* Interactions between HIV-1 Gag molecules in solution: an inositol phosphate-mediated switch. *Journal of molecular biology* **365**, 799-811 (2007).

271 Xu, C. *et al.* Permeability of the HIV-1 capsid to metabolites modulates viral DNA synthesis. *PLoS Biol* **18**, e3001015, doi:10.1371/journal.pbio.3001015 (2020).

272 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).

273 Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* **8**, 91, doi:10.1186/1472-6750-8-91 (2008).

274 Zennou, V., Perez-Caballero, D., Gottlinger, H. & Bieniasz, P. D. APOBEC3G incorporation into human immunodeficiency virus type 1 particles. *J Virol* **78**, 12058-12061, doi:10.1128/JVI.78.21.12058-12061.2004 (2004).

275 Naldini, L. *et al.* In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263-267 (1996).

276 Pornillos, O. *et al.* X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**, 1282-1292 (2009).

277 Price, A. J. *et al.* Active site remodeling switches HIV specificity of antiretroviral TRIMCyp. *Nature structural & molecular biology* **16**, 1036-1042 (2009).

278 Kucharska, I. *et al.* Biochemical reconstitution of HIV-1 assembly and maturation. *Journal of virology* **94**, e01844-01819 (2020).

279 Nguyen, H.-L. T. *et al.* An efficient procedure for the expression and purification of HIV-1 protease from inclusion bodies. *Protein Expression and Purification* **116**, 59-65 (2015).

280 Toohey, K., Wehrly, K., Nishio, J., Perryman, S. & Chesebro, B. Human immunodeficiency virus envelope V1 and V2 regions influence replication efficiency in macrophages by affecting virus spread. *Virology* **213**, 70-79, doi:10.1006/viro.1995.1547 (1995).

281 Wehrly, K. & Chesebro, B. p24 antigen capture assay for quantification of human immunodeficiency virus using readily available inexpensive reagents. *Methods* **12**, 288-293, doi:10.1006/meth.1997.0481 (1997).

282 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**, e1004459, doi:10.1371/journal.ppat.1004459 (2014).

283 Vermeire, J. & Verhasselt, B. Quantification of Retro- and Lentiviral Reverse Transcriptase Activity by Real-time PCR. *Bio-protocol* **3**, e780, doi:10.21769/BioProtoc.780 (2013).

284 Dettenhofer, M. & Yu, X.-F. Highly purified human immunodeficiency virus type 1 reveals a virtual absence of Vif in virions. *Journal of virology* **73**, 1460-1467 (1999).

285 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76, doi:10.1006/jsbi.1996.0013 (1996).

286 Gorrec, F. The MORPHEUS protein crystallization screen. *J Appl Crystallogr* **42**, 1035-1042, doi:10.1107/S0021889809042022 (2009).

287 Winn, M. D. An overview of the CCP4 project in protein crystallography: an example of a collaborative project. *J Synchrotron Radiat* **10**, 23-25 (2003).

288 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214, doi:10.1107/S0907444913000061 (2013).

289 McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta crystallographica. Section D, Biological crystallography* **63**, 32-41, doi:10.1107/S0907444906045975 (2007).

290 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255, doi:10.1107/S0907444996012255 (1997).

291 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

292 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

293 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367, doi:10.1107/S0907444911001314 (2011).

294 Lau, D. *et al.* Fluorescence Biosensor for Real-Time Interaction Dynamics of Host Proteins with HIV-1 Capsid Tubes. *ACS Appl Mater Interfaces* **11**, 34586-34594, doi:10.1021/acsami.9b08521 (2019).

295 Lau, D. *et al.* Self-Assembly of Fluorescent HIV Capsid Spheres for Detection of Capsid Binders. *Langmuir* **36**, 3624-3632, doi:10.1021/acs.langmuir.0c00103 (2020).

296 Márquez, C. L., Lau, D., Walsh, J., Rifat Faysal, K., Parker, M. W., Turville, S. G. and Böcking, T. Fluorescence Microscopy Assay to Measure HIV-1 Capsid Uncoating Kinetics in vitro. *Bio-protocol* **9**, e3297, doi:DOI: 10.21769/BioProtoc.3297 (2019).

297 Marquez, C. L. *et al.* Kinetics of HIV-1 capsid uncoating revealed by single-molecule analysis. *Elife* **7**, doi:10.7554/eLife.34772 (2018).

298 Bocking, T., Aguet, F., Harrison, S. C. & Kirchhausen, T. Single-molecule analysis of a molecular disassemblase reveals the mechanism of Hsc70-driven clathrin uncoating. *Nat Struct Mol Biol* **18**, 295-301, doi:10.1038/nsmb.1985 (2011).

299 Rihn, S. J. *et al.* Extreme genetic fragility of the HIV-1 capsid. *PLoS Pathog* **9**, e1003461, doi:10.1371/journal.ppat.1003461 (2013).

300 Song, G. Structure-based insights into the mechanism of nucleotide import by HIV-1 capsid. *J Struct Biol* **207**, 123-135, doi:10.1016/j.jsb.2019.05.001 (2019).

301 Xu, C. *et al.* Permeability of the HIV-1 capsid to metabolites modulates viral DNA synthesis. *bioRxiv*, 2020.2004.2030.071217, doi:10.1101/2020.04.30.071217 (2020).

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

303 Bush, D. L. & Vogt, V. M. In Vitro Assembly of Retroviruses. *Annu Rev Virol* **1**, 561-580, doi:10.1146/annurev-virology-031413-085427 (2014).

304 Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C. & Sundquist, W. I. Assembly properties of the human immunodeficiency virus type 1 CA protein. *J Virol* **78**, 2545-2552, doi:10.1128/jvi.78.5.2545-2552.2004 (2004).

305 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, doi:10.1128/jvi.77.9.5439-5450.2003 (2003).

306 Ni, T. *et al.* Structure of native HIV-1 cores and their interactions with IP6 and CypA. *Sci Adv* **7**, eabj5715, doi:10.1126/sciadv.abj5715 (2021).

307 Mattei, S., Glass, B., Hagen, W. J., Krausslich, H. G. & Briggs, J. A. The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science* **354**, 1434-1437, doi:10.1126/science.aah4972 (2016).

308 Lopez, C. S. *et al.* Determinants of the HIV-1 core assembly pathway. *Virology* **417**, 137-146, doi:10.1016/j.virol.2011.05.011 (2011).

309 Serrière, J., Fenel, D., Schoehn, G., Gouet, P. & Guillon, C. Biophysical characterization of the feline immunodeficiency virus p24 capsid protein conformation and in vitro capsid assembly. *PLoS One* **8**, e56424 (2013).

310 Luttge, B. G. *et al.* Molecular characterization of feline immunodeficiency virus budding. *Journal of virology* **82**, 2106-2119 (2008).

311 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* **354**, 1434-1437 (2016).

312 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**, doi:10.1128/JVI.00984-20 (2020).

313 Ricana, C. L., Lyddon, T. D., Dick, R. A. & Johnson, M. C. Primate lentiviruses require Inositol hexakisphosphate (IP6) or inositol pentakisphosphate (IP5) for the production of viral particles. *PLoS Pathog* **16**, e1008646, doi:10.1371/journal.ppat.1008646 (2020).

314 Wagner, J. M. *et al.* Crystal structure of an HIV assembly and maturation switch. *Elife* **5**, doi:10.7554/eLife.17063 (2016).

315 Kleinpeter, A. B. & Freed, E. O. HIV-1 Maturation: Lessons Learned from Inhibitors. *Viruses* **12**, doi:10.3390/v12090940 (2020).

316 Li, F. *et al.* PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *Proc Natl Acad Sci U S A* **100**, 13555-13560, doi:10.1073/pnas.2234683100 (2003).

317 Zhou, J. *et al.* Small-molecule inhibition of human immunodeficiency virus type 1 replication by specific targeting of the final step of virion maturation. *J Virol* **78**, 922-929, doi:10.1128/jvi.78.2.922-929.2004 (2004).

318 Adamson, C. S. *et al.* In vitro resistance to the human immunodeficiency virus type 1 maturation inhibitor PA-457 (Bevirimat). *J Virol* **80**, 10957-10971, doi:10.1128/JVI.01369-06 (2006).

319 Urano, E. *et al.* Resistance to Second-Generation HIV-1 Maturation Inhibitors. *J Virol* **93**, doi:10.1128/JVI.02017-18 (2019).

320 Blair, W. S. *et al.* New small-molecule inhibitor class targeting human immunodeficiency virus type 1 virion maturation. *Antimicrob Agents Chemother* **53**, 5080-5087, doi:10.1128/AAC.00759-09 (2009).

321 Waki, K. *et al.* Structural and functional insights into the HIV-1 maturation inhibitor binding pocket. *PLoS Pathog* **8**, e1002997, doi:10.1371/journal.ppat.1002997 (2012).

322 Mallery, D. L. *et al.* A stable immature lattice packages IP6 for HIV capsid maturation. *Sci Adv* **7**, doi:10.1126/sciadv.abe4716 (2021).

323 Purdy, M. D. *et al.* MicroED structures of HIV-1 Gag CTD-SP1 reveal binding interactions with the maturation inhibitor bevirimat. *Proceedings of the National Academy of Sciences* **115**, 13258-13263 (2018).

324 Fontana, J. *et al.* Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. *J Virol* **89**, 9765-9780, doi:10.1128/JVI.01522-15 (2015).

325 Fontana, J. *et al.* Identification of an HIV-1 Mutation in Spacer Peptide 1 That Stabilizes the Immature CA-SP1 Lattice. *J Virol* **90**, 972-978, doi:10.1128/JVI.02204-15 (2016).

326 Kucharska, I. *et al.* Biochemical Reconstitution of HIV-1 Assembly and Maturation. *J Virol* **94**, doi:10.1128/JVI.01844-19 (2020).

327 Renner, N. *et al.* A lysine ring in HIV capsid pores coordinates IP6 to drive mature capsid assembly. *PLoS Pathog* **17**, e1009164, doi:10.1371/journal.ppat.1009164 (2021).

328 Dostalkova, A. *et al.* In Vitro Quantification of the Effects of IP6 and Other Small Polyanions on Immature HIV-1 Particle Assembly and Core Stability. *J Virol* **94**, doi:10.1128/JVI.00991-20 (2020).

329 Mattei, S. *et al.* High-resolution structures of HIV-1 Gag cleavage mutants determine structural switch for virus maturation. *Proceedings of the National Academy of Sciences* **115**, E9401-E9410 (2018).

330 Kuiken, C., Korber, B. & Shafer, R. W. HIV sequence databases. *AIDS Rev* **5**, 52-61 (2003).

331 Obr, M. *et al.* Structure of the mature Rous sarcoma virus lattice reveals a role for IP6 in the formation of the capsid hexamer. *Nature Communications* **12**, 3226, doi:10.1038/s41467-021-23506-0 (2021).

332 Dick, R. A. *et al.* Structures of immature EIAV Gag lattices reveal a conserved role for IP6 in lentivirus assembly. *PLoS pathogens* **16**, e1008277 (2020).

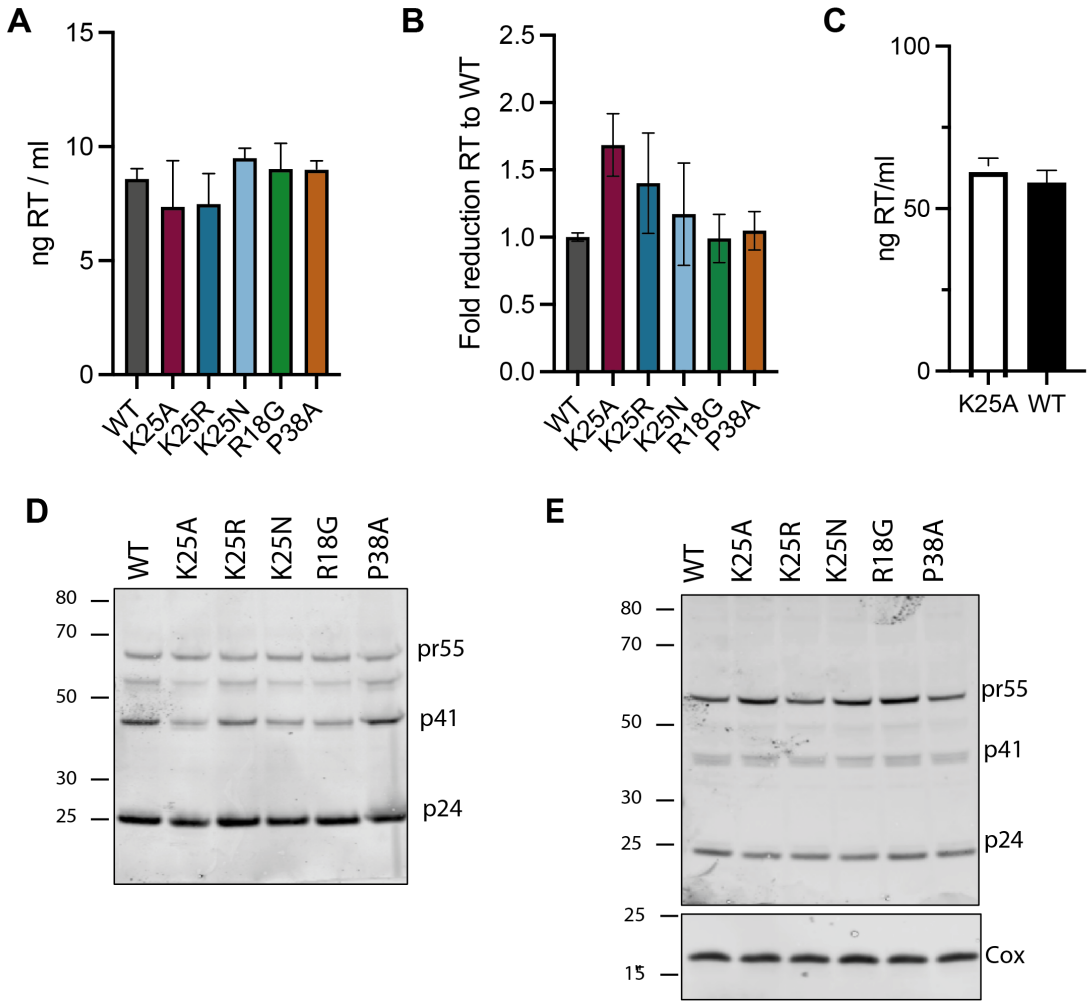
333 Naldini, L. Lentiviruses as gene transfer agents for delivery to non-dividing cells. *Current opinion in biotechnology* **9**, 457-463 (1998).

334 Vodicka, M. A. Determinants for lentiviral infection of non-dividing cells. *Somatic cell and molecular genetics* **26**, 35-49 (2001).

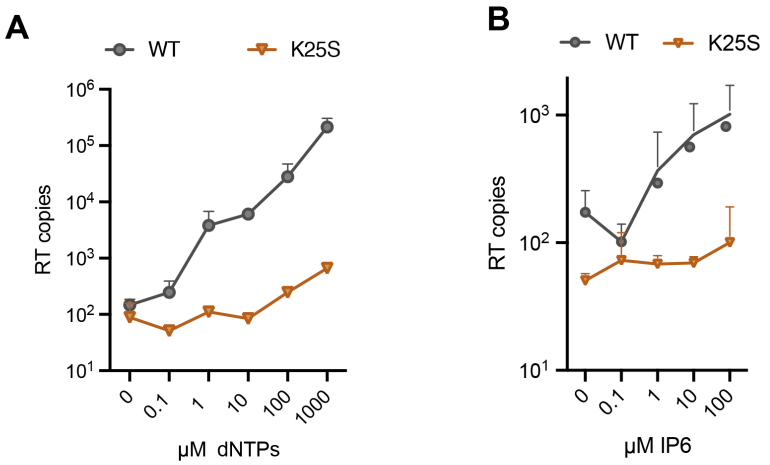
335 Dostálková, A. *et al.* Effect of Small Polyanions on In Vitro Assembly of Selected Members of Alpha-, Beta- and Gammaretroviruses. *Viruses* **13**, doi:10.3390/v13010129 (2021).

336 Qu, K. *et al.* Structure and architecture of immature and mature murine leukemia virus capsids. *Proceedings of the National Academy of Sciences* **115**, E11751-E11760 (2018).

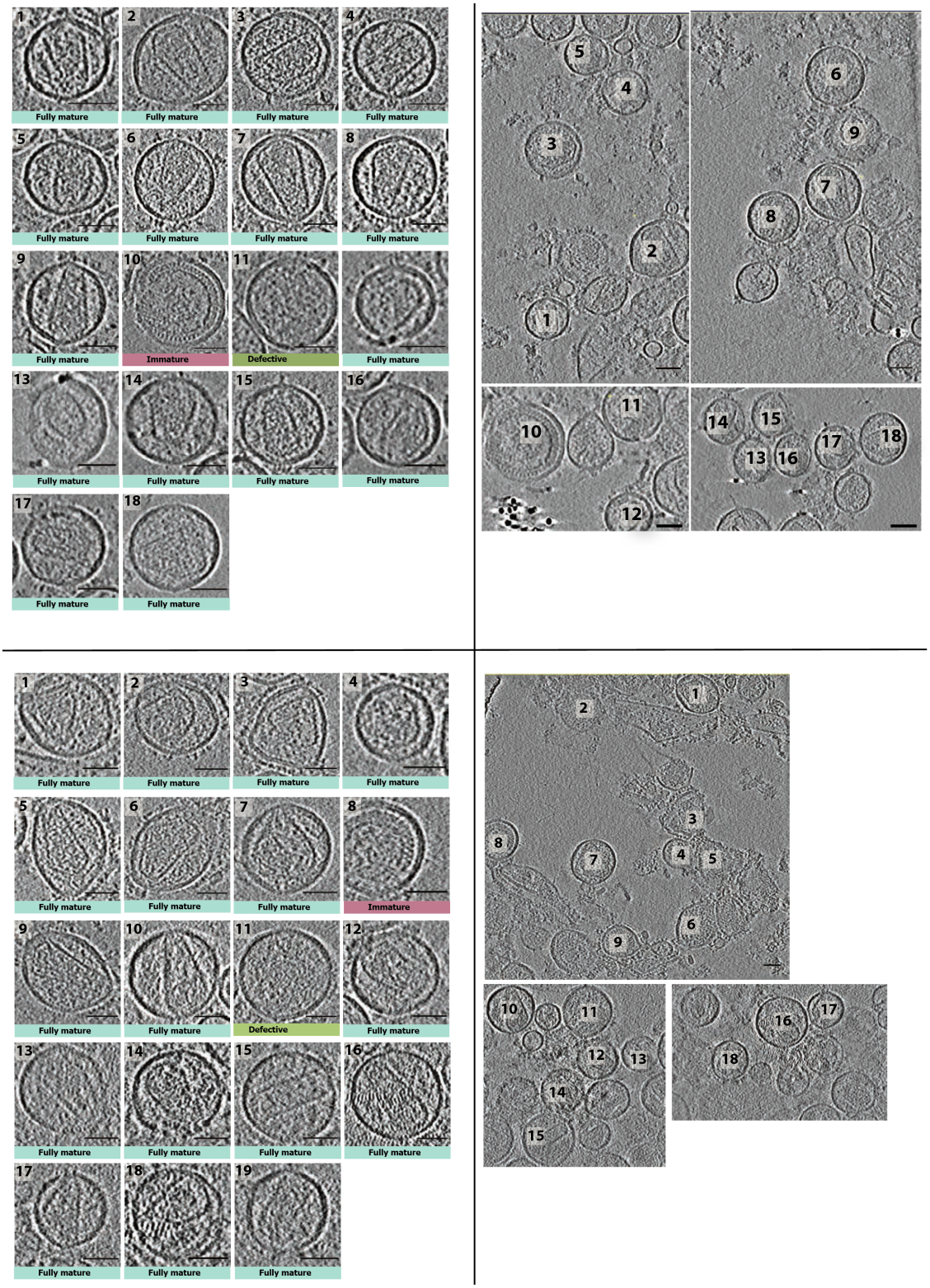
337 Folio, C., Sierra, N., Dujardin, M., Alvarez, G. & Guillon, C. Crystal structure of the full-length feline immunodeficiency virus capsid protein shows an N-terminal β-hairpin in the absence of N-terminal proline. *Viruses* **9**, 335 (2017).



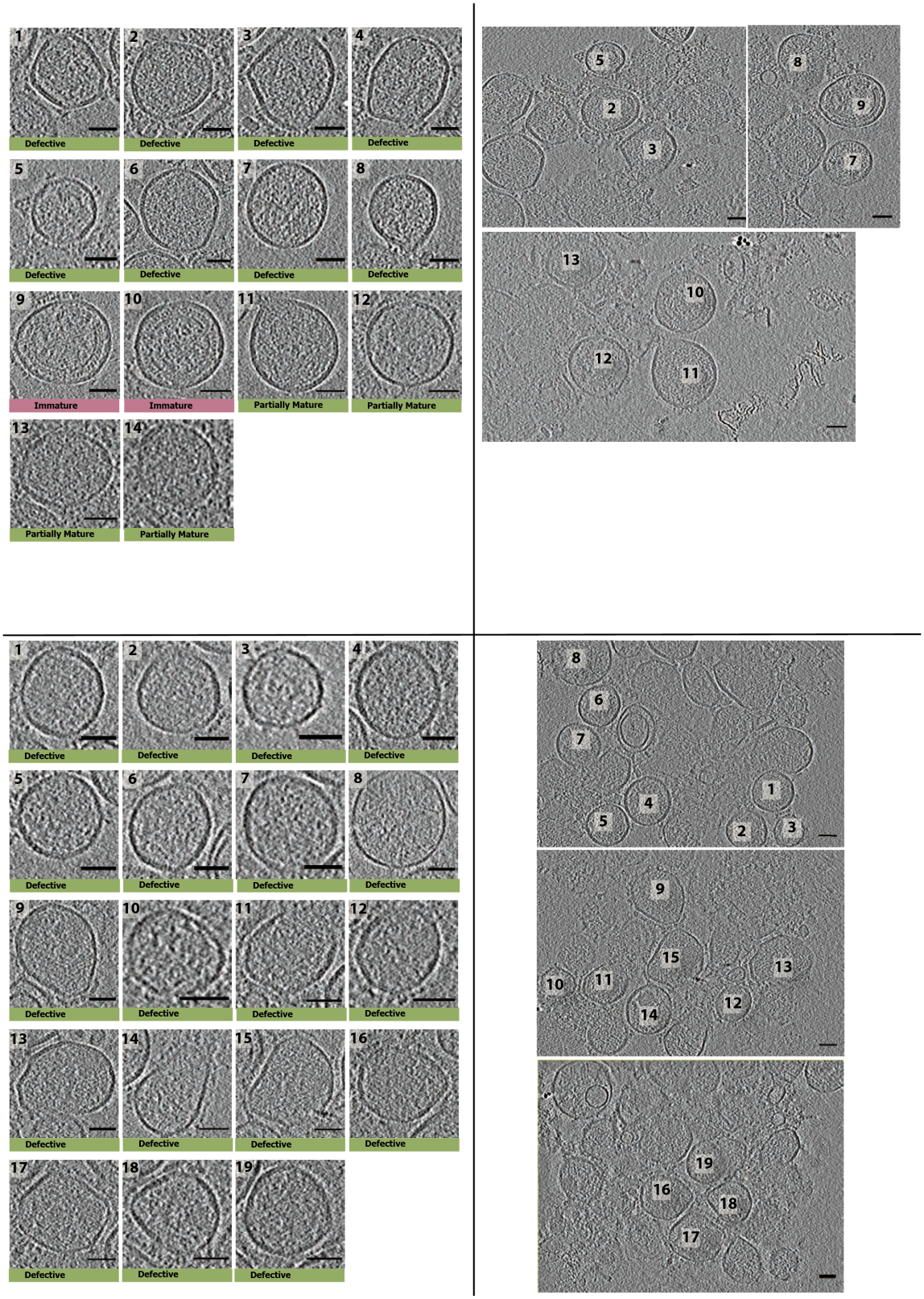
Appendix Figure 1: (A) RT assay of viral supernatant. Error bars depict mean ± SD of three replicates from one experiment representative of three independent preps. (B) Fold reduction of RT present in viral supernatant of mutants compared to WT in 2 biological replicates. (C) RT assay to show similar incorporation of the enzyme in WT and K25A cores for ERT experiments. (D) Western-blot of purified virions showing that WT and mutants have similar levels of p24 processing, consistent with normal maturation. (E) Western-blot of infected cells showing that WT and mutants have similar levels of p24. Error bars depict mean ± SD of three replicates from one experiment representative of three independent core preps. All experiments of panels except (C) were done by Donna Mallery.



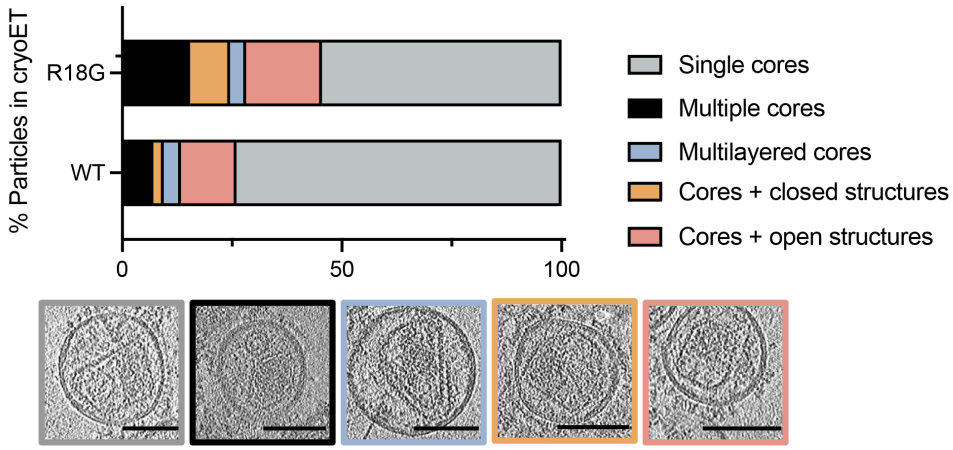
Appendix Figure 2: (A) ERT assay measuring the synthesis of strong-stop DNA in the presence of DNase of WT and K25S cores at increasing concentrations of dNTPs. (B) ERT assay measuring the synthesis of strong-stop DNA in the presence of DNase at 1μM dNTPs of WT and K25S cores at increasing concentrations of IP6. Error bars in infection experiments depict mean ± SD of three replicates from one experiment and representative of three independent experiments.



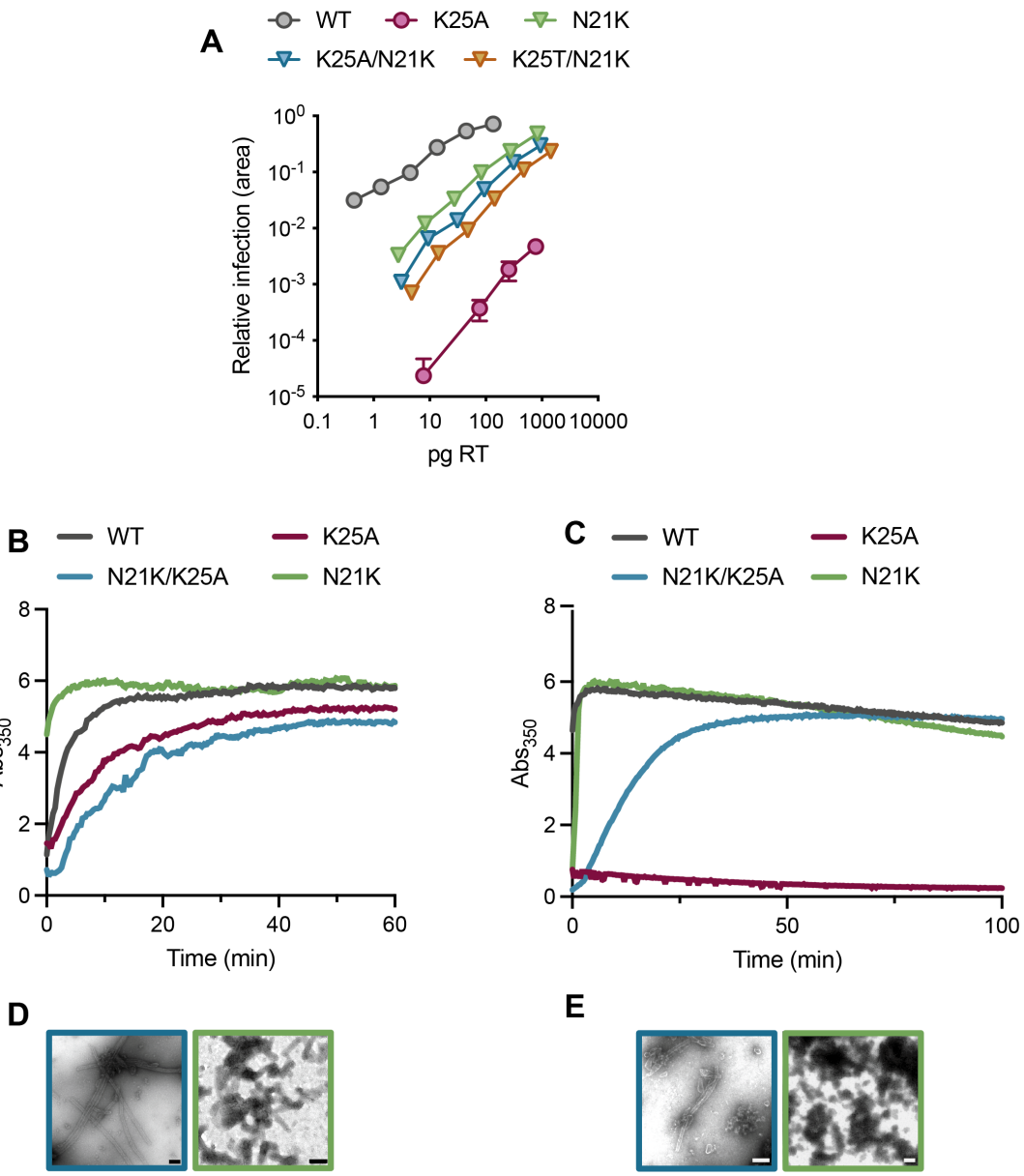
Appendix Figure 3: Representative tomogram slices of WT virion. Scale bar: 100 nm.



Appendix Figure 4: Representative tomogram slices of WT virion. Scale bars: 100 nm.



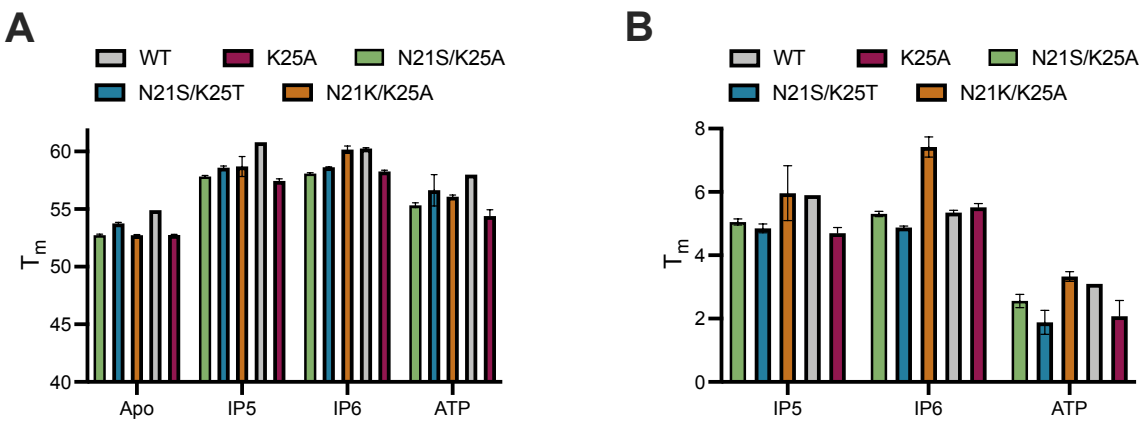
Appendix Figure 5: Cryo-ET Analysis of R18G and WT additional structures. Virions with mature lattices shown in **Figure 13** were further subdivided into: Multiple Cores (green), Single Cores (cyan), Cores with additional closed structure (orange), Cores with additional open structure (light orange), Multilayered Cores (blue). Scale bars: 100 nm.



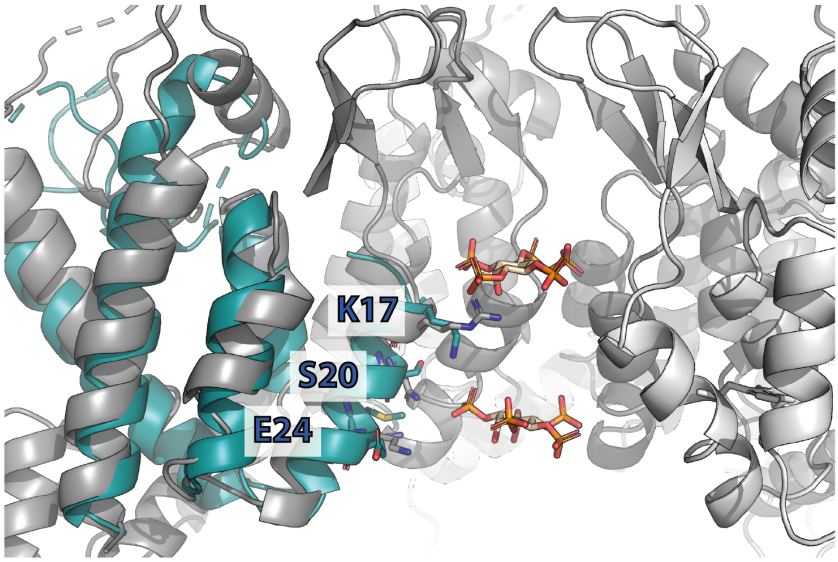
Appendix Figure 6: N21K mutants (A) Infectivity is measured using an Incucyte and determined as the proportion of cell area that is GFP +ve at a given viral dose (in ng RT).Error bars in infection experiments depict mean ± SD of three replicates from one experiment and representative of three independent experiments. (B-E) In vitro assembly reactions were monitored in real-time be measuring the absorbance at 350 nm. (B) Assembly reactions using 75 μM capsid protein and 2.5 M NaCl in 50 mM MES, pH 6.0 and the corresponding negative stain images (D). (C) Assembly reactions using 75 μM capsid protein and 1.25 mM IP6 in 50 mM MES, pH 6.0 and the corresponding negative stain images (E). Scale bars: 200 nm.

Table 3: Crystallographic data table for CA hexamer structures shown in Figure 21, Statistics in highest resolution shell are shown in parentheses.

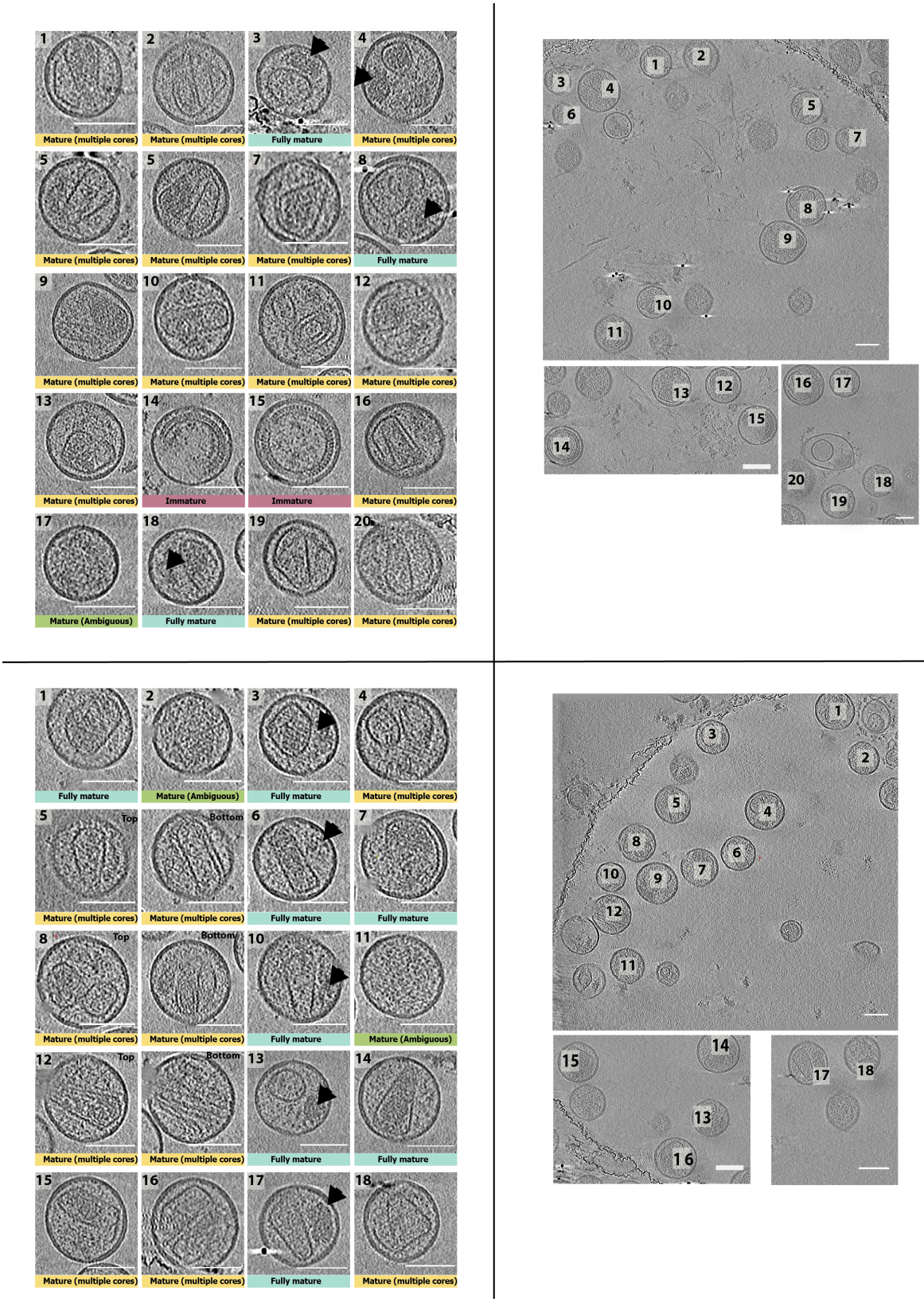
	N21S/K25T+ ATP	N21S/K25A+ IP6
Wavelength	0.9999	0.9999
Resolution range	45.79 - 2.05 (2.123 - 2.05)	46.14 - 1.94 (2.009 - 1.94)
Space group	P 6	P 6
Unit cell	89.895 89.895 56.619 90 90 120	90.8 90.8 56.973 90 90 120
Total reflections	32212 (3164)	36398 (3791)
Unique reflections	16478 (1632)	18461 (1935)
Multiplicity	2.0 (1.9)	2.0 (2.0)
Completeness (%)	99.79 (99.03)	92.50 (98.77)
Mean I/sigma(I)	6.84 (1.27)	10.44 (1.33)
Wilson B-factor	33.92	31.17
R-merge	0.06781 (0.6452)	0.04929 (0.7308)
R-meas	0.0959 (0.9125)	0.06971 (1.034)
R-pim	0.06781 (0.6452)	0.04929 (0.7308)
CC1/2	0.996 (0.42)	0.998 (0.362)
CC*	0.999 (0.769)	0.999 (0.729)
Reflections used in refinement	16478 (1634)	18461 (1935)
Reflections used for R-free	793 (86)	972 (117)
R-work	0.1967 (0.2903)	0.2090 (0.3147)
R-free	0.2424 (0.3316)	0.2465 (0.3555)
CC(work)	0.954 (0.749)	0.955 (0.661)
CC(free)	0.911 (0.575)	0.842 (0.610)
Number of non-hydrogen atoms	1787	1735
macromolecules	1599	1597
ligands	31	138
solvent	157	205
Protein residues	205	0.007
RMS(bonds)	0.009	0.96
RMS(angles)	1.31	98.99
Ramachandran favored (%)	97.99	1.01
Ramachandran allowed (%)	1.51	0
Ramachandran outliers (%)	0.5	0
Rotamer outliers (%)	0	4.71
Clashscore	5.26	35.52
Average B-factor	38.29	35.02
macromolecules	37.39	41.28
ligands	56.65	75.55
solvent	43.87	40.74



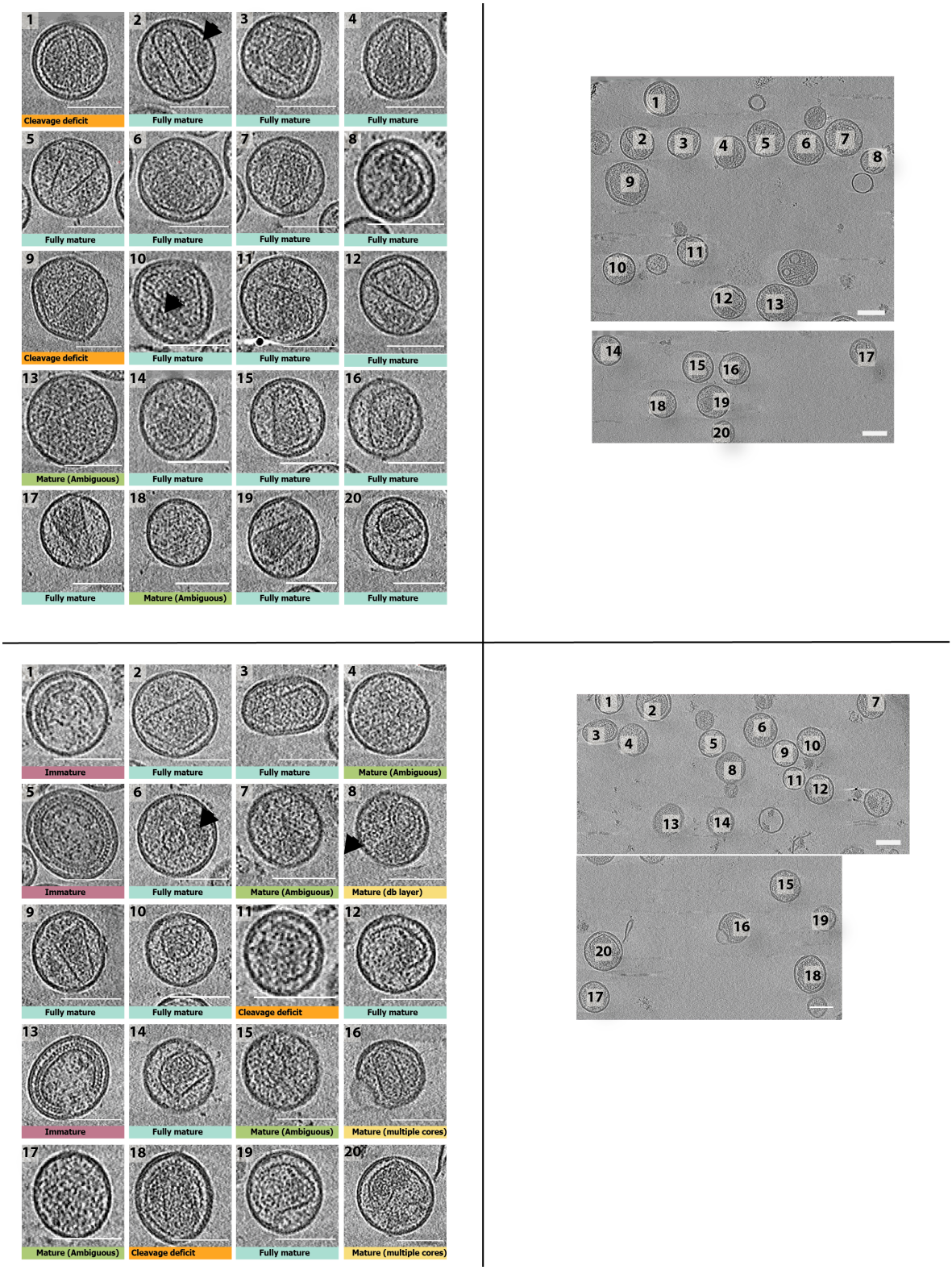
Appendix Figure 7: A) NanoDSF-derived melting temperature (ΔTM's) of WT and mutant hexamers B) Changes of NanoDSF-derived melting temperature (ΔTM's) of WT hexamers in presence of different IPs compared to the disulfide-stabilised hexamer in absence of ligands. Error bares depict mean SD of technical replicates.



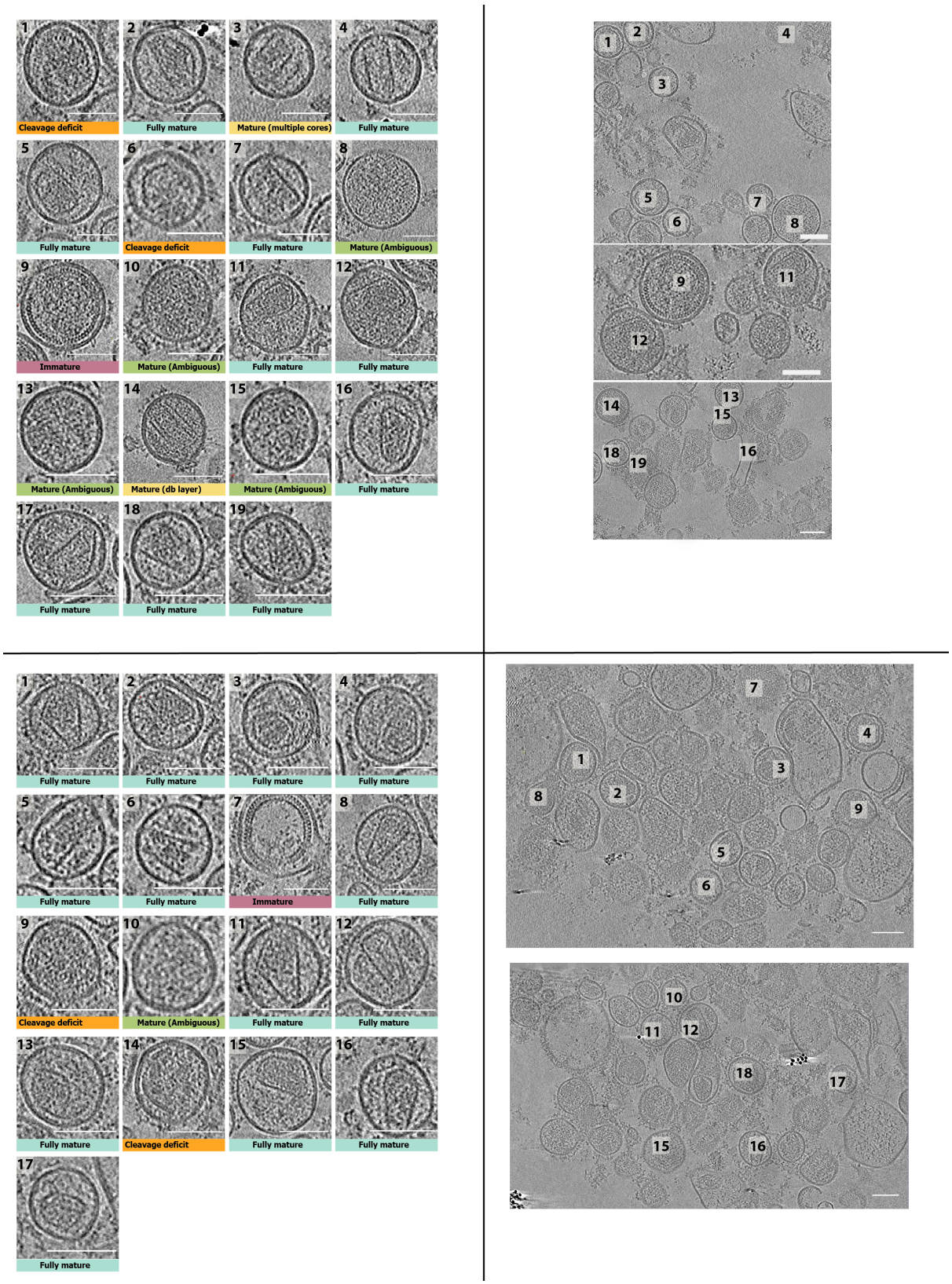
Appendix Figure 8: Comparison of the FIV and HIV mature hexamers. The FIV (5NA2) [337] mature hexamer is coloured in blue and residues corresponding to R18, N21, and K25 are labelled. The HIV structure from **Figure 10** A is coloured in grey.



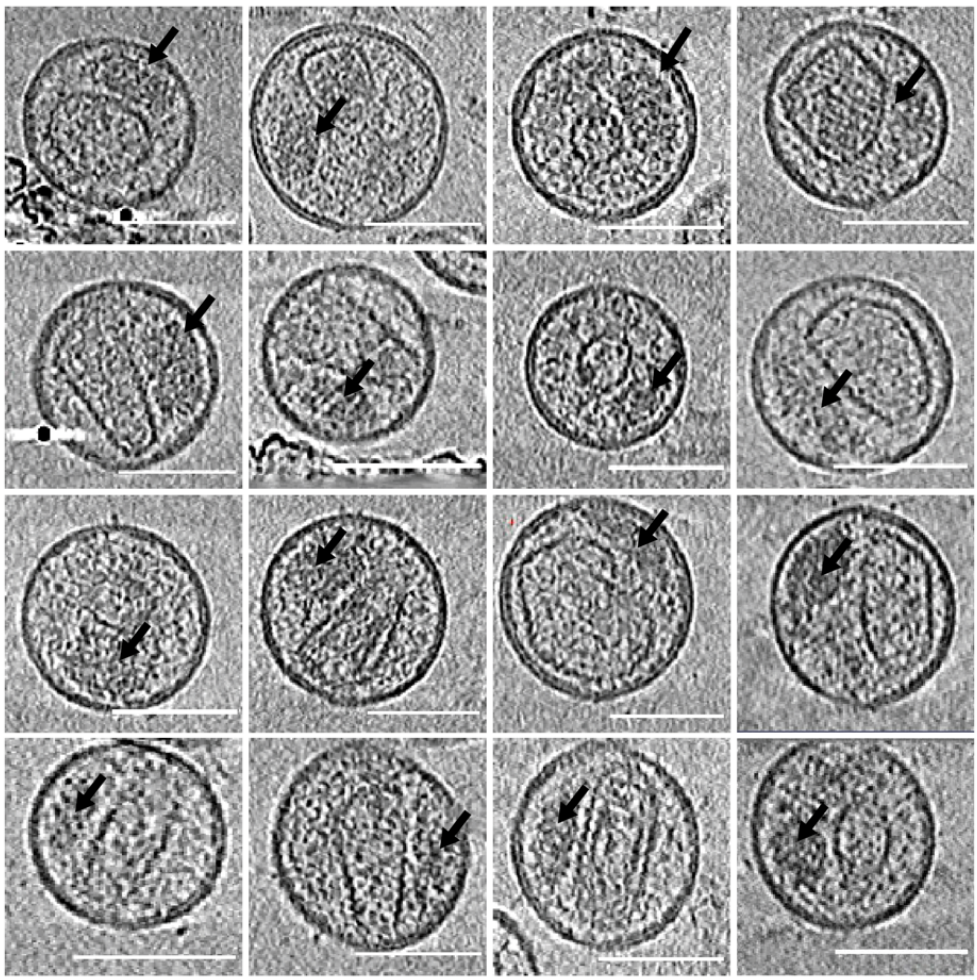
Appendix Figure 9: Representative tomogram slices of K158A virions. Scale bars: 100 nm.



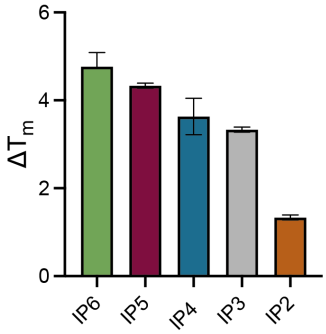
Appendix Figure 10: Representative tomogram slices of K158A/T8I virions. Scale bars: 100 nm.



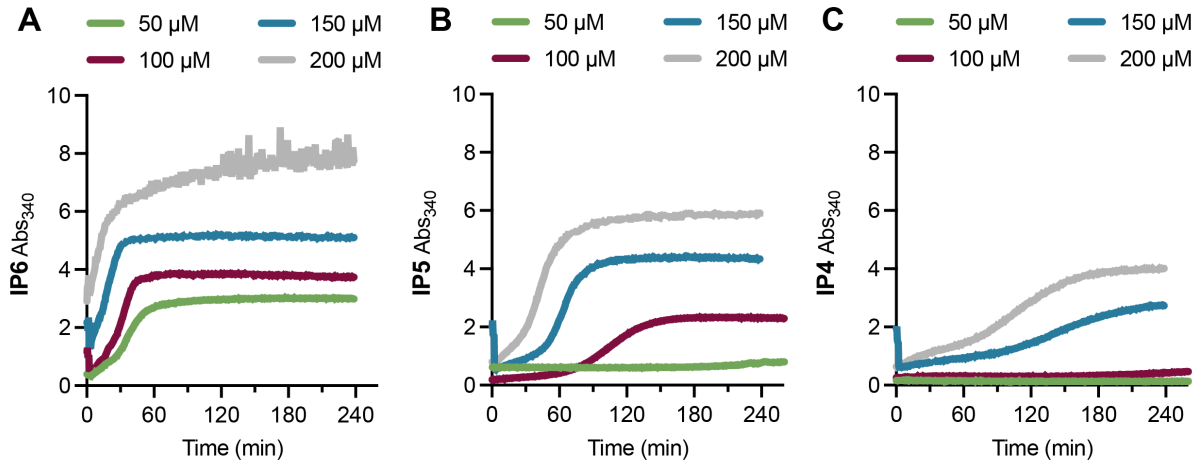
Appendix Figure 11: Representative tomogram slices of T8I virions. Scale bars: 100 nm.



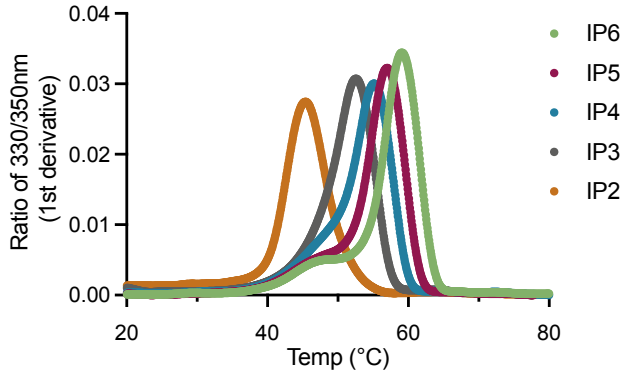
Appendix Figure 12: Example tomograms of K158A virions. Densities corresponding to possible RNPs are indicated by arrows. Scale bars: 200 nm.



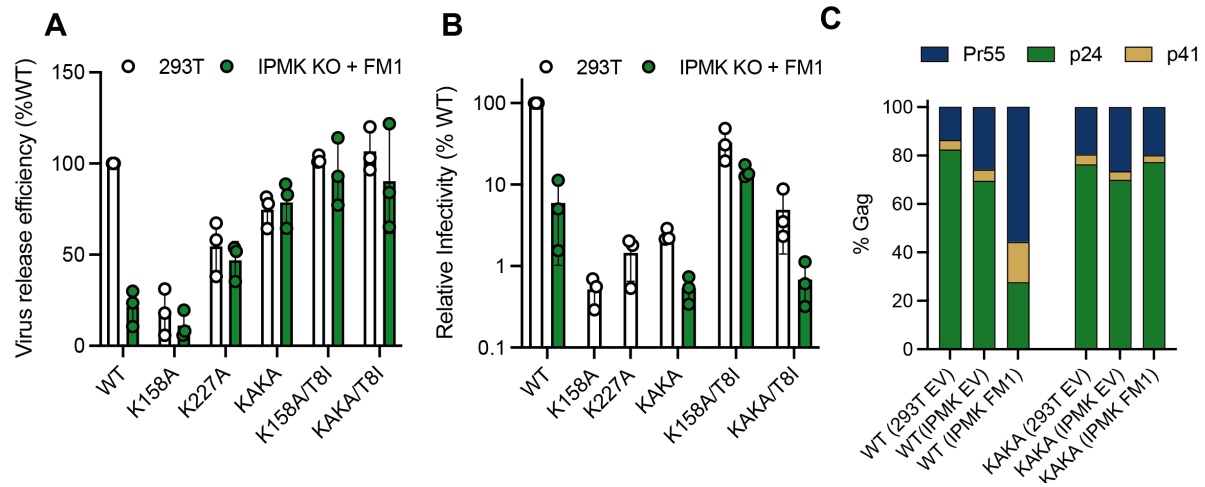
Appendix Figure 13: Changes of NanoDSF-derived melting temperature (ΔTM's) of WT hexamers in presence of different IPs compared to the disulfide-stabilised hexamer in absence of ligands. Error bares depict mean SD of technical replicates.



Appendix Figure 14: In vitro assembly of immature VLPs using recombinant ΔMA-CANC protein. IPs at the indicated concentrations and 7.5 μM RNA was added to 75 μM ΔMA-CANC and assembly monitored through absorbance changes at 350 nm at 25°C. IP6 (A), IP5 (B) and IP4 (C) were added.

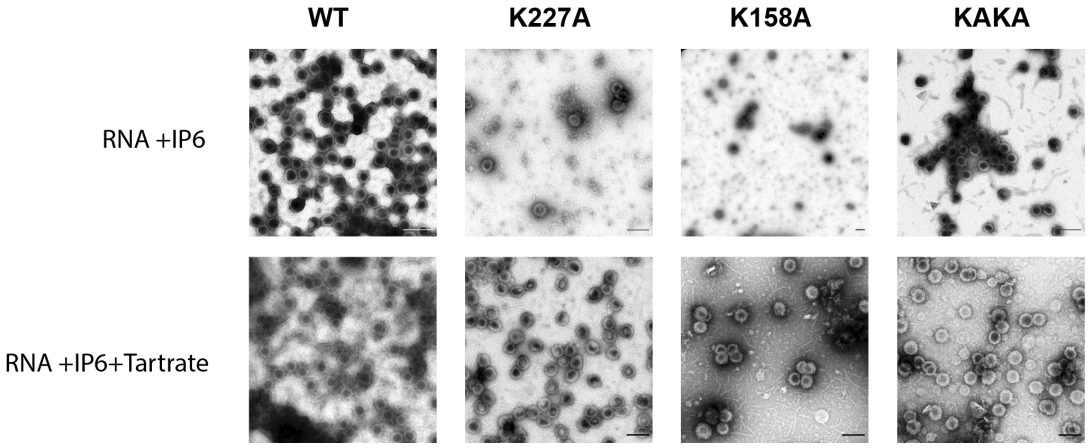


Appendix Figure 15: NanoDSF-derived melting curves of WT MA-CANC in presence of different IPs. Error bares depict mean SD of technical replicates.

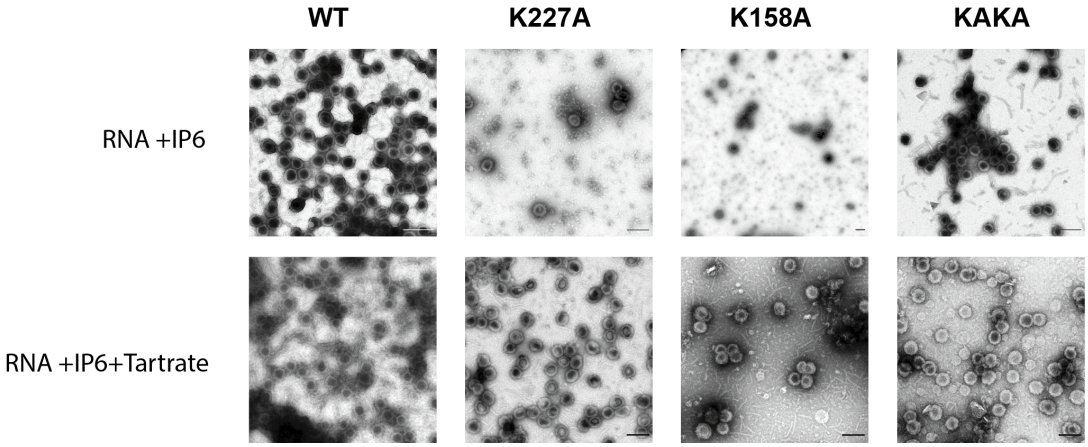


Appendix Figure 16: Simultaneously reducing cellular IP6 and the ability of HIV-1 to enrich it into virions amplifies infectivity defects. (A) Virus release efficiency of Gag mutants from either 293T cells or IPMK KOs overexpressing FM1, calculated as the percentage of particle-associated p24 (CA)

as a fraction of total (cell- + particle-associated Gag) normalised to WT virus. Error bars depict the SEM from at least three independent experiments. (B) Infectivity of Gag mutants normalised to the quantity of input virus as determined by 32P RT assay and relative to WT HIV-1. Error bars depict the SEM from three independent experiments. (C) Gag cleavage efficiency of WT and KAKA purified virions, calculated as the relative amount of p24 (CA), p41 and Pr55Gag as a percentage of total Gag in particle lysate. Experiments for panels (A) and (B) were performed by Alex Kleinpeter, Experiments for panel (C) were performed by Donna Mallery.



Appendix Figure 17: Negative stain EM of VLPs formed with ΔMA-CANC, IP6, RNA and tartrate. Scale bars: 200 nm.



Appendix Figure 18