Characterising lentiviral host interactions

Stephen Oliver Perry

University College London

Thesis submitted for the degree of Doctorate of Philosophy

September 2019

Declaration

I, Stephen Oliver Perry confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis

Abstract

Previously it was thought that the HIV-1M capsid undergoes rapid and spontaneous disassembly upon entering the cytoplasm. We now hypothesise that the HIV-1M capsid remains intact as it traverses the cytoplasm and plays a role in coordinating several stages in the life cycle via interactions with host cell proteins. These host proteins, which are essential for optimal infection, are termed cofactors. Here we investigate lentivirus-cofactor interactions, focusing on the cofactors IP6 and CypA.

Recently it has been found that the abundant inositol deviate IP6 is able to drive immature particle formation and stabilises the mature capsid core for HIV-1M. Taking a comparative virology approach, in combination with phylogenetic and mutagenesis studies, we investigated whether the use of IP6 is conserved throughout the lentivirus lineage or whether it is a specific cofactor for HIV-1M. We provide evidence that IP6 usage is conserved throughout the lentiviral lineage, via the mutagenesis of the IP6 binding site within the capsid of diverse lentiviruses, and is essential for particle production and reverse transcription.

Furthermore, we investigated CypA usage across the lentivirus lineage. We show that CypA usage is a feature which has evolved along the SIVCpz-HIV-1M branch, as replication of the parental SIVCpz virus to HIV-1M is independent of CypA. We also show that CypA usage in HIV-1M can be influenced by the position of the capsid NTD beta-hairpin. Finally, we investigate cellular factors which can influence HIV-1M CypA dependency in cell lines. Focusing on the cellular nuclease TREX-1, we manipulate its expression levels and measured CypA dependency of HIV-1M and the parental SIVCpz. We found that TREX-1 expression levels do no influence CypA dependency of either HIV-1M or SIVCpz.

Impact statement

In this study we have investigated the role the lentivirus capsid plays in viral replication. We have shown that the recently identified cofactor for HIV-1M, IP6, is a conserved cofactor for lentiviruses in general and is not specific to HIV-1M. Previously, it has been shown that IP6 is able to directly bind capsid residues and influence two distinct stages of the HIV-1M lifecycle. By binding to Lys 158 and Lys 227 in the immature Gag bundle, IP6 is able to promote particle assembly. Whereas by binding Arg 18 in the electrostatic channel of the mature capsid hexamer, IP6 is able to regulate the stability of the mature capsid core. We have shown that both IP6 binding sites, in the immature and mature capsid, are conserved throughout the lentivirus lineage and that disruption of either greatly inhibits lentiviral infection. The impact of this study is to broaden our understanding of the lifecycle of diverse lentiviruses.

These findings also have implication for novel drug design. It may be possible to design a drug which mimics IP6 and thus targets the electrostatic channel within the lentivirus capsid to alter uncoating dynamics to either promote premature uncoating in the cytoplasm or prevent uncoating completely, both of which we think would lead to non-productive infection. An IP6 analogue that binds at the electrostatic channel in the lentiviral capsid hexamer would not only be affective against HIV-1M but potentially against all lentiviruses which contain an electrostatic channel. As we have shown, the electrostatic channel is a highly conserved feature across the lentivirus lineage, an IP6 analogue would likely function as a broad spectrum antiviral.

Our study also provides evidence that escape from such as drug would be difficult. We have shown, by mutagenesis studies of the HIV-1M capsid, that the electrostatic channel is unable to tolerate the loss of Arg 18, even when substituted to the similar amino acid lysine. Furthermore, any mutation at position 18 within the HIV-1M capsid is highly

detrimental to infection. We have also shown that all examined lentiviruses, except for FIV, are unable to tolerate the loss of their electrostatic channel, again suggesting high fitness cost in viral escape of an IP6 analogue. In summary, my thesis has shown that IP6 is a conserved cofactor for viruses within the lentivirus lineage, making it an excellent therapeutic target against lentiviral infection.

Acknowledgments

First and foremost, I would like to thank Greg Towers for his supervision, training and support throughout my PhD. I would also like to thank Richard Goldstein, Clare Jolly and Joe Grove for their insightful engagement and comments during my thesis committee and lab meetings.

Secondly, I would like to thank Joe Grove, Isobel Honeyborne, Lucy Thorne and Lorena Zuliani-Alvarez. Without your continual guidance, support, willingness to answer countless questions and occasional tough love I don't think I would have ever finished my PhD.

To Lauren Harrison, David Sterling, Hataf Khan, Liane Fernandes, Tom peacock, Rob Lever and of course Chris Tie. Thank you all for the laughs, support and many memorable conversations. It just wouldn't have been the same without you guys. I would particularly like to thank Hale Tunbak for keeping me sane during the deepest darkest writer's block.

I also thank all the members of the Towers lab, past and present, who have supported me throughout my PhD.

Table of contents

List of figures

List of Tables

Abbreviations

A3F	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3F
A3G	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G
AIDS	Acquired immunodeficiency syndrome
AP2	Adaptor protein complex 2
APS	Ammonium persulfate
Arg	Arginine
BIV	Bovine Immunodeficiency virus
CAEV	Caprine arthritis encephalitis virus
CPSF6	Cleavage and polyadenylation specificity factor subunit 6
CRM1	Chromosome region maintenance 1
CsA	Cyclosporine A
CTD	Carboxy-terminal domain
CypA	Cyclophilin A
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DN25	Deoxyribonuclease
dNTP	Deoxynucleotide
DRC	Democratic Republic of Congo
EIAV	Equine infectious anemia virus
Env	Envelope
ESCRT	Endosomal sorting complexes required for transport
EtBR	Ethidium bromide
FCS	Fetal calf serum
FG	Phenylalanine-glycine

FISH	Fluorescence in situ hybridization
FIV	Feline immunodeficiency virus
FIV-Oma	FIV from Pallas cats
FIV-Pca	Feline immunodeficiency virus from domesticated cats
FIV-Pco	FIV from Pumas
FIV-Ple	FIV from lions
FSC-H	Front scatter height
Gag	Group specific antigen
GFP	Green fluorescent protein
Gly	Glycine
GTR	General time reversible
HFV	Human foamy virus
HIV-1	Human Immunodeficiency Virus 1
HIV-1M	Human Immunodeficiency Virus 1 Group M
HIV-1O	Human Immunodeficiency Virus 1 Group O
HIV-2	Human Immunodeficiency Virus 2
HR-C	Carboxy-terminal helical region of HIV-1 envelope
HR-N	Amino-terminal helical region of HIV-1 envelope
I.U./ngRT	Infectious Units per ng of reverse transcriptase
IFN	interferon
IP6	Inositol hexaphosphate
ISG	Interferon stimulated genes
ITC	Isothermal Titration Calorimetry
LB	Lysogeny broth

LINC	Linker of nucleoskeleton and cytoskeleton complex
LTR	Long terminal repeat
Lys	Lysine
MDM	Monocyte-derived macrophage
MHC	Major histocompatibility complex
MHR	Major-homology-region
MLV	Murine leukemia virus
MOI	Multiplicity of infection
MX2/MxB	Myxovirus resistance protein 2
Nef	Negative regulatory factor
NF-κB	Nuclear factor κ-light-chain-enhancer of activated B cells
NLEF	Negative elongation factor
NLS	Nuclear localisation signal
NPC	Nuclear pore complex
NTD	Amino-terminal domain
NUP	Nucleoporins
NUP153	Nucleoporin 153
NUP358	Nucleoporin 358
OD	Optical density
pbs	Primer binding site
PBS	Phosphate buffered saline
PBS-T	PBS-Tween
PFA	Paraformaldehyde
PI(4, 5)P2	Phosphatidyl inositol (4, 5) bisphosphate

PIC	Pre-integration complex
Pol	Polymerase
PPIase	Peptidyl-prolyl cis-trans isomerase
PTAP	Pro-Thr-Ala-Pro
qPCR	Quantitative PCR
R5	CCR5
RELIK	Rabbit Endogenous Lentivirus type K
REV	Regulator of expression of viral proteins
RNAPII	RNA polymerase II
RNP	Ribonucleoprotein
RRE	Rev response element
RSV	Rous Sarcoma Virus
RTC	Reverse transcription complexes
S.O.C	Super outgrowth media
SD	Standard deviation
SDM	Site directed mutagenesis
SDS	Sodium dodecyl-sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIV	Simian Immunodeficiency viruses
SIVCpz	Simian Immunodeficiency viruses of chimpanzees
SIVCPZBF	A SIVCpzPts (SIVCpzBF1167)
SIVCpzMB	A SIVCpzPtt (SIVCpzMB897)
SIVCpzMT	A SIVCpzPtt (SIVCpzMT145)
SIVCpzPts	SIVCpz from *Pan troglodytes schweinfurthii*

SIVCpzPtt	SIVCpz from *Pan troglodytes troglodytes*
SR proteins	Serine/arginine rich proteins
SSC-H	Side scatter height
TAE	Tris acetate EDTA
TAR	Transactivation response region
TAT	Viral transactivator protein
TCR	T-cell receptor
TEMED	Tetramethylethylenediamine
TNPO3	Transportin 3
Vif	Virion infectivity factor
Visna	Maedi visna virus
Vpr	Viral protein R
Vpu	Viral protein U
VSV-G	Vesicular stomatitis virus G envelope glycoprotein
WHO	World Health Organisation
X4	CXCR4
YPXL	Tyr-Pro-X-Leu,

1 Introduction

1.1 Retroviruses

1.1.1 Retroviridae

All viruses are obligate intercellular parasites, meaning they are reliant on host cell machinery for replication. Classified as group VI in the Baltimore system [1], the *Retroviridae* family is characterised by the use of virus encoded RNA dependent DNA polymerase to reverse transcribe their positive sense RNA genome into double strand DNA before its integration into host cell genomic DNA. Once the viral DNA genome is integrated into host genomic DNA, the provirus is formed. The provirus is the source of all new viral mRNAs and viral genomic RNA. The *Retroviridae* family can be divided into two main categories, simple and complex. Simple retroviruses including Alpha-, Beta- and Gammaretrovirus, only encode three open reading frames these being the group specific antigen (*gag*), polymerase (*pol*) and envelope (*env*) region [**Figure 1A**] which encode for the main viral structural proteins, the viral enzymes and the viral entry machinery respectively. Along with the three large open reading frames, complex retroviruses, Lenti- Spuma- and Deltaretrovirus encode several smaller accessory proteins [**Figure 1B** and **C**] which are used for a various function such regulating the cellular environment and antagonising restriction factors. Alongside the *gag*, *pol* and *env*, all retroviruses contain both a 5' and 3' long terminal repeats (LTR) [**Figure 1**] which can be further divided into U3, R and U5 regions. The LTRs are a length of repeating nucleotide sequence which plays a role in retroviral integration. Once integrated, the LTRs play a role in transcription control of the provirus. Originally the complex retroviruses were subdivided based on virion morphology [2] however modern sequencing methods have confirmed they are phylogenetically different. The most studied and best understood retrovirus is the lentivirus Human Immunodeficiency Virus 1 (HIV-1) causative agent of acquired immunodeficiency syndrome (AIDS).

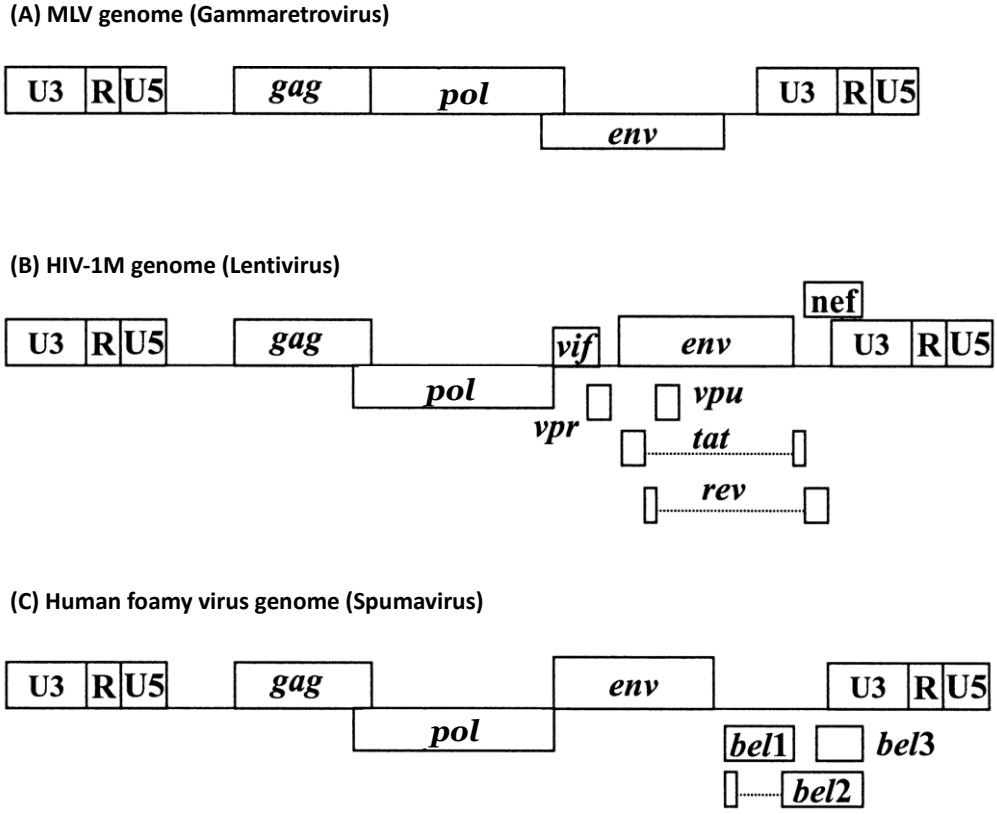


Figure 1: Genome comparison of simple and complex retroviruses
Schematic comparison of genome organisation between the simple gammretrovirus MLV (A) and complex lentivirus HIV-1M (B) and the spumavirus HFV (C). Adapted from Hu. W. S. et al. 2000

1.1.2 Lentiviruses

The study of retroviruses began at the turn of the twentieth century with the observation that cell free extract from chicken tumours could induce sarcoma formation within healthy fowl [3,4] and that equine infectious anemia was caused by a filterable agent [5], later the causative agents of these diseases would be named Rous sarcoma virus (RSV) and equine infectious anemia virus (EIAV) respectively, making EIAV the first identified lentivirus. Since then it has been found that lentiviruses are able to infect a wide range of mammalian lineages such as primates with HIV and simian immunodeficiency viruses (SIV), ovine-caprine lineage with caprine arthritis encephalitis virus (CAEV) and maedi visna virus (Visna), felid lineage with feline immunodeficiency virus (FIV), equid with

EIAV and bovine with bovine Immunodeficiency virus (BIV). Each lineage of lentivirus is associated with different diseases with varying severity. However, all lentiviruses show a characteristically long incubation times followed by chronic infection and disease. Due to this slow progression of disease, these viruses were given the prefix "*Lenti"*-meaning slow in latin. In the case of HIV-1M, the incubation time before disease is around 8 years [6] however there is a small portion of infected which have incubation periods of over 10 years [7]. However, it wasn't until the discovery of HIV-1 [8,9] that the field grew to prominence and real molecular details of infection started to be examined.

We now know that infection with lentiviruses is characterised by: (1) both latent and productive viral replication in macrophages and CD4 positive T cells [10–15], (2) immunological escape by various mechanisms, (3) variable long incubation times followed by chronic disease, (4) non-oncogenic infection and the use of several small virally encoded proteins to modulate the cellular environment. One defining feature of lentiviruses is that unlike many other members of the *Retroviridea* family including Gamma-, Alpha- and Spuma-, their replication is independent of cell proliferation as they can effectively infect dividing and non-dividing cells to the same extent [16–21]. To date, HIV-1M is the most studied human pathogen and still remains an area of high research activity.

1.1.3 The HIV/AIDS pandemic

It is currently estimated by the World Health Organisation that there are 36.9 million individuals worldwide infected with HIV-1 [22]. Originally identified in men who have sex with men in north America in the early 1980s [23], HIV/AIDS was quickly found to be in every region of the world [24]. Despite the HIV/AIDS pandemic being first reported in North America there was a period of silent spread from around 1920-1980 from now the

democratic republic of Congo (DRC). Although sexual transmission is by far the most prevalent form of transmission route, mother-child transmission during birth or breast-feeding can occur. Likewise, infection can also occur due to direct blood-blood contact as is common with injecting drug users.

Interestingly, HIV-1M is unique amongst mammalian lentiviruses as it's the only human lentivirus to become pandemic whereas the other human lentiviruses remain largely restricted to specific areas, e.g. Africa for HIV-2. A fundamental question being discussed in our lab is what adaptions occurred in HIV-1M to allow it to become pandemic.

1.1.4 Clinical progression of HIV-1M

After initial infection with HIV-1M the clinical course of disease can be measured in several ways, the appearance of opportunistic infections, wasting, decline of blood CD4+ T cells or by quantification of viral RNA via PCR. Progression of AIDS can be divided into different stages depending on symptoms and CD4+ T cells levels, a typical HIV-1 time course of infection can be seen in **Figure 2**.

One to two weeks after infection is known as the eclipse stage. During the eclipse stage neither viremia nor immune response can be detected and the infected individual is symptom free however viral replication is ongoing [25]. Two to four weeks post infection the patient enters the acute, or primary stage. This stage is characterised by high levels of viremia, a transient decrease in blood CD4+ T cells and often flu-like symptoms [25]. The high level of viremia is likely due to the failure of the early immune response to control the virus and the generation of large numbers of activated CD4+ T cells which act as target cells for HIV-1M infection. Towards the end of this stage there is a drastic decline in viremia to a set point, likely due to partial control by the adaptive immune system and

exhaustion of target cell population. During the clinical latency stage, one to ten years post infection, there is a slow but steady increase in viremia and an equally slow decline of CD4+ T cells. This stage is named latent but there is continual viral replication and infection of new target cells. However, this is largely asymptomatic of overt disease. Once the CD4+ T cell population is fully depleted, below 200 CD4+ T cells per ml of blood, opportunistic infections start to appear as the immune system cannot maintain control of them. This is accompanied by a rebound in viremia. This final stage is termed the AIDs stage and, if untreated, results in a mortality rate of ~100%.

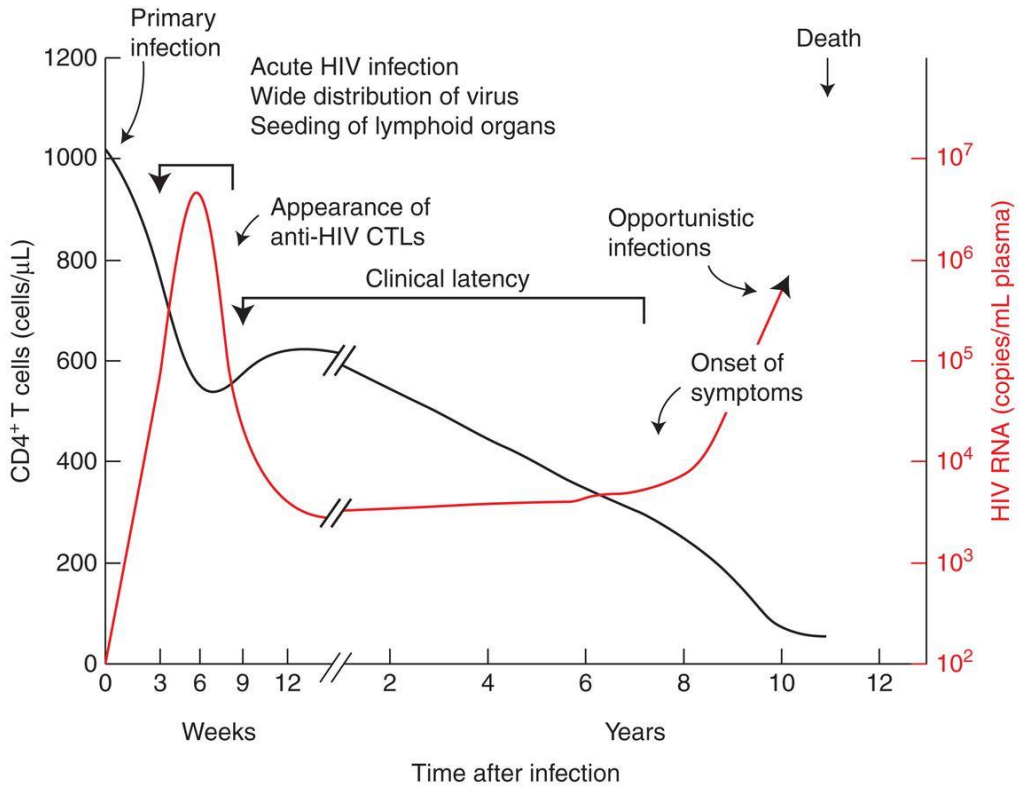


Figure 2: Typical time scale for disease progression in HIV-1
After a short acute increase in viral load and massive loss of CD4+ T cells which do not recover, there is an extended latency period where the virus replicates to high levels. This latency period can last up to 10 years before collapse of the immune system due to effect loss of CD4+ T cells. Taken from Fauci AS et al. 1997.

1.1.5 Lentiviral disease in other mammals

The diseases caused by lentiviral infection vary widely from virus to virus. However, the development of AIDS-like illness is not restricted to HIV-1. The parental virus to HIV-1M, SIVCpz (see section 1.1.6), was initially thought to be a non-pathogenic infection in its natural host, chimpanzees. However, careful tracking and observation of infected wild chimpanzees revealed an increased risk of mortality at any age and the development of an AIDS-like illness characterised by a decline in CD4+ T cell count [26,27]. The observation that SIVCpz is pathogenic directly challenged the view that all SIV infections are non-pathogenic in their natural host [28]. Development of AIDS isn't restricted to the HIV-1/SIVcpz lineage, FIV has been shown to cause AIDS within domesticated cats [29–32]. FIV infection of domesticated cats leads to disease progression with striking similarity to that of HIV-1M in humans and thus has been used as a natural model for AIDS development. Interestingly, non-domesticated felines, such as pumas and lions, which are infected with FIV don't seem to develop an AIDS-like disease [33,34].

Like HIV-1M, FIV has relatively recently entered the domesticated cat population via zoonosis as shown by the very low genetic diversity in the pol region of the FIV genome from domesticate cats [35]. One hypothesis is that lentiviruses only causes AIDS-like disease in newly acquired hosts, such as humans or domesticated felines. Lentiviruses, such as the SIV within monkey populations, which have been in a host for many millions of years are likely to become attenuated allowing the host to live longer and thus have a greater chance of transmitting the virus. It would be interesting to study the potential pathogenicity of SIVGor (another recently transmitted virus) in gorillas, but due to their critically endangered status study is difficult.

1.1.6 Zoonotic origins for human immunodeficiency viruses

The vast majority of emerging infectious disease are zoonoses and HIV-1 is no different. It is now thought that HIV-1M entered the human population around the 1920s, likely due to the consumption of bushmeat infected with SIVCpz [36]. This cross-species event gave rise to HIV-1M, the most prevalent HIV-1 which has nine subtypes A-D, F-H, J and K. Subtype C is most prevalent in Africa whereas subtype B is most prominent in Europe and North America [24,37]. Alongside the zoonosis which gave rise to HIV-1M there have been three additional cross species event giving rise to groups N, O and P. Groups N and P have had very limited spread compared to HIV-1M with only 20 [38,39] and 2 [40,41] recorded infections respectively. Whereas group O is the second most prominent, with 100000 estimated infections but remains largely restricted to West Africa [42].

Group M and N were transmitted to humans from chimpanzees. The chimpanzee lineage of lentiviruses, SIVCpz, was found to be formed by two highly divergent, but subspecies-specific, phylogenetic lineages [43]. The two phylogenetic clusters of SIVCpz have more than a 40% sequence difference and lineage specific protein signatures [44]. This level of divergences is at the same level of genetic difference seen between SIVs from related monkey species. These lineages of SIVCpz are found in eastern chimpanzees (*Pan troglodytes schweinfurthii*, SIVCpzPts) and western chimpanzees (*Pan troglodytes troglodytes*, SIVCpzPtt). Using phylogenetic analysis, it was found that HIV-M clusters closely with SIVCpzPtt and from that the evolutionary history of the virus was inferred [45].

Groups O and P were found to be transmitted to humans from gorillas [38]. However, the origins of SIVgor, again, was found to be from SIVCpzPtt [38]. There is only evidence for a single chimpanzee-gorilla transmission event but there is considerable genetic variability seen within SIVgor [38]. This would suggest that SIVgor has infected

gorillas far longer that HIV-1 has infected humans. Considering the similar levels of contact via bushmeat [36] It remains unclear as to why SIVCpzPtt has crossed the species barrier, twice directly into humans and twice indirectly via gorillas, but so far there has not been a reported zoonosis from SIVCpzPts.

The ancestry of HIV-1 has been traced back to SIVCpz but upon closer examination it was found that SIVCpz was itself the result of a zoonotic event. SIV have been found to infect more than 20 different species of African primate [46] but their evolutionary relationships are hard to determine due to high levels of recombination. Depending on what region of the genome is examined gives varying phylogenetic relationships. Using extensive phylogenetic analysis of various regions of the genome it was found that the Gag-Pol regions of SIVCpz clustered closely with that of SIVrcm, from red capped mangabeys, and Env with that of SIVgsn, from greater spot-nosed monkeys [47]. It is thought that the recombinant event, which gave raise to SIVCpz, occurred due to an ancient coinfection event within a chimpanzee. This zoonotic event must have occurred before the divergence of eastern and western chimp, which is thought to have occurred several thousand years ago [48], as both are infected with SIVCpz and now they have non-overlapping geographic ranges divided by rivers and other natural barriers [48].

HIV-1 is not the only human immunodeficiency virus which owes its origins to zoonotic event. It has been well described that HIV-2 is the result of a zoonotic event by SIVsm, from sooty mangabeys. Phylogenetic analysis has shown that the genome of HIV-2 is closely related to that of SIVsm [49]. Furthermore, it was found that the natural habitat of sooty mangabeys coincides with the epicentre of the HIV-2 epidemic [50,51] and that these monkey are often kept as pets in the area and hunted for food [52]. Since its discovery, there has been at least eight different lineages of HIV-2 identified, termed groups A-H.

Of these groups, only A and B have had any significant spread however both remain restricted to West Africa [53,54]. Groups C-H have only been reported in single individuals and are thought to not have any secondary transmission [50,51,55].

1.2 Lentiviral life cycle

Although lentiviruses are found to infect a diverse range of hosts their life cycles are highly similar [**Figure 3**]. Firstly, the viral envelope must fuse with the target cell membrane, mediated by the viral envelope proteins. Upon entry the viral capsid enters the cytoplasm and the +ssRNA genome is converted into dsDNA by the virally encoded RNA dependent DNA polymerase, reverse transcriptase. The precise location and environment in which reverse transcription takes place is still a highly controversial topic within the field. However, it is clear that the viral components must traverse the cytosol and enter the nucleus where the virally encoded integrase enzyme integrates the viral genome into host chromosomes. Once integration has occurred the cell is typically productively infected and the viral genome can be expressed leading to production of progeny virions which are released from the infected cell. Events prior to integration but after virus entry are often termed the early stages of infection. The capsid plays an essential role in the early life cycle by protecting the viral genome from cellular nucleases and pattern recognition receptors. This section will provide additional details about each stage in the life cycle of HIV-1M and a short overview of restriction factors the virus most overcome.

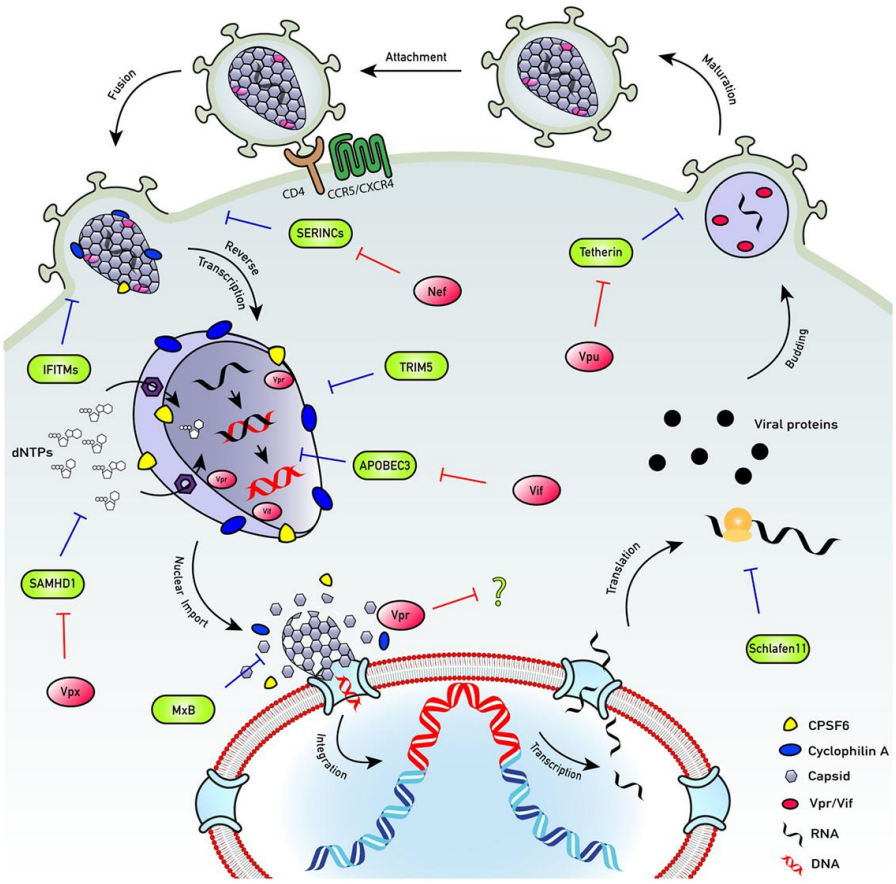


Figure 3: Lentivirus life cycle
The life cycle of lentiviruses is very similar with all the key steps being conserved throughout the lineage. In summary, the virus must enter the cells; reverse transcribe its RNA genome into DNA; enter the nucleus and integrate its genome into a host chromosome. During productive infection, HIV-1M must overcome many restriction factors which can inhibit infection if not avoided or antagonized. Adapted from R.P. Sumner et al. 2017.

1.2.1 Entry

The first step in the HIV-1M life cycle is the binding to and entry of target cells to initiate infection. This stage is dependent on the HIV-1M envelope protein, which is formed from a trimer of gp120 and gp41 heterodimers, to bind to the HIV-1M primary receptor CD4 [56]. First, however, HIV-1M attaches to the cell in a relatively non-specific manner, either driven by host proteins incorporated into the HIV-1M membrane or by gp120 interacting with the negatively charged heparan sulfate proteoglycans [57]. This non-specific attachment likely acts to bring HIV-1M into close proximity to CD4 and the co-receptors to increase entry efficiency. Once CD4 is bound, gp120 undergoes a structural rearrangement to reveal the co-receptor binding site [58,59]. HIV-1 can be classified

depending on which co-receptor the viruses uses. Either they will use CCR5 (R5) and be considered macrophage-trophic or CXCR4 (X4) and be T-Cell-tropic [60]. Interestingly, R5 predominate virus which is transmitted [61] and it isn't until late infection that X4 viruses become more prevalent in some, but not all, infected individuals [62,63]. Binding the co-receptor causes a structural change in the gp41 heterosdimer which reveals the fusion peptide which is then inserted into the host cell membrane. Once inserted, the fusion peptide folds at a hinge region which brings the amino-terminal helical region (HR-N) and a carboxy-terminal helical region (HR-C) together [64]. Due to this conformational change the viral membrane and host cell membrane are brought into close proximity which results in their fusion and leads to the capsid core being deposited into the cytoplasm.

1.2.2 Cytoplasmic trafficking

Upon entry into the target cytoplasm, HIV-1 must traverse to the nucleus to allow for integration and thus productive infection. However, HIV-1 must also avoid the cytoplasmic environment which contains innate immune sensors and nucleases which would both inhibit infection either by induction of innate immune response, (1) via DNA sensing mediated by cGAS-STING pathway [65] or, (2) by degradation of viral genomes via cellular nucleuses like TREX-1 [66], respectively. We now favour a hypothesis where the HIV-1M capsid remains intact during cellular transit and only uncoats at or after nuclear entry. This hypothesis is discussed in more detail in section 1.3.3. The movement of large macromolecules in the cytoplasm is highly restricted due to the steric hindrance caused by the highly crowed nature of the cytosol [67,68]. Therefore due to the large size of the HIV-1 capsid it is unlikely that it moves through the cell in a diffusion mediated style. As expected it was found, using GFP tagged viruses, that HIV-1M traffics in a directed

rapid manner towards the nucleus in a characteristically microtubule dependent fashion [69,70]. This is supported by the observation that disrupting microtubule mediate transport results in inhibition to HIV-1 infection [71,72]. Furthermore, using a genome wide screen, BICD2, a protein shown to play a role in the dynein-mediated transport of proteins [73], was found to be required for HIV-1 infection [74]. Once at the nuclear pore, particles accumulate and are ready for nuclear entry.

1.2.3 Reverse transcription

Although the exact location within the target cell where reverse transcription takes place is still yet to be determined conclusively, the mechanism has been well characterised. Reverse transcription begins before viral entry, however, it is only completed after viral entry into the target cell. Once entry is complete and the capsid core is deposited within the cell reverse transcriptase is able to complete the conversion of the RNA viral genome into double stranded DNA which will be used as the substrate for integration. The viral reverse transcriptase has two distinct enzymatic features which are both critical for reverse transcription. These are the RNA dependent DNA polymerase activity and the RNase H activity which is able to degrade the RNA component of a RNA-DNA duplex. The mature form of reverse transcriptase is a heterodimer [75,76] composed of the larger catalytic subunit p66 and the smaller structural subunit, p51 [77]. Furthermore, it has been shown that there is a relationship between reverse transcription and uncoating [78,79]. *In vitro* assays using atomic force microscopy have shown that as reverse transcription progresses the internal pressure within purified capsids increase, until the internal stress excessed the capsids structure strength leading to uncoating [78].

The DNA polymerase activity of reverse transcriptase is dependent on a primer to initiate DNA synthesis. Host cell tRNA3Lys has been found to act as the primer for

reverse transcription for HIV-1 by base pairing with a complementary sequence in the 5' end of the viral RNA genome, termed the primer binding site (pbs) [**Figure 4A]**. This tRNA is packaged within the mature virus [80] and is essential for revers transcription. Once reverse transcriptase is bound to the pbs via tRNA3Lys, DNA synthesis is initiated by the extension of tRNA3Lys in the 3' direction using the viral RNA genome as a template [**Figure 4B]**. This initial segment of reverse transcription is termed strong stop and can be isolated from mature virus particles before cellular entry [81–83] suggesting that reverse transcription actually begins within the viral particle. However, full length viral DNA construct can only be found in infected cells suggesting that reverse transcription can only progress so far within the virus particle. Once synthesis to strong stop is completed, the newly synthesised minus-strand DNA is transferred to the 3' end of the viral RNA genome [**Figure 4C]**. This is accomplished by the repeat sequences, present within the LTRs, annealing and allows for minus strand synthesis to continue along the whole length of the genome completing the minus strand synthesis [**Figure 4D]**. This stage is known a first strand transfer and is a major source of recombination [84–87]. After first strand transfer the minus-strand synthesis is completed and the positive-sense RNA template is degraded concurrently by the RNase H activity of reverse transcriptase [**Figure 4D]**. However, a poly-purine track, termed the ppt, is resistant to RNase H activity and isn't degraded. This ppt is essential for optimal reverse transcription [88] as it facilitates the synthesis of positive strand DNA by functioning as a primer for reverse transcriptase. DNA synthesis of the positive-strand DNA occurs towards the 5' end of the minus-strand [**Figure 4E]**. Once completed, the positive-strand is transferred to the minus-strand 3' end where the pbs anneals [**Figure 4F],** termed second strand transfer, and act as the primer allowing DNA synthesis along the remainder of the minus-strand **[Figure 4G]**.

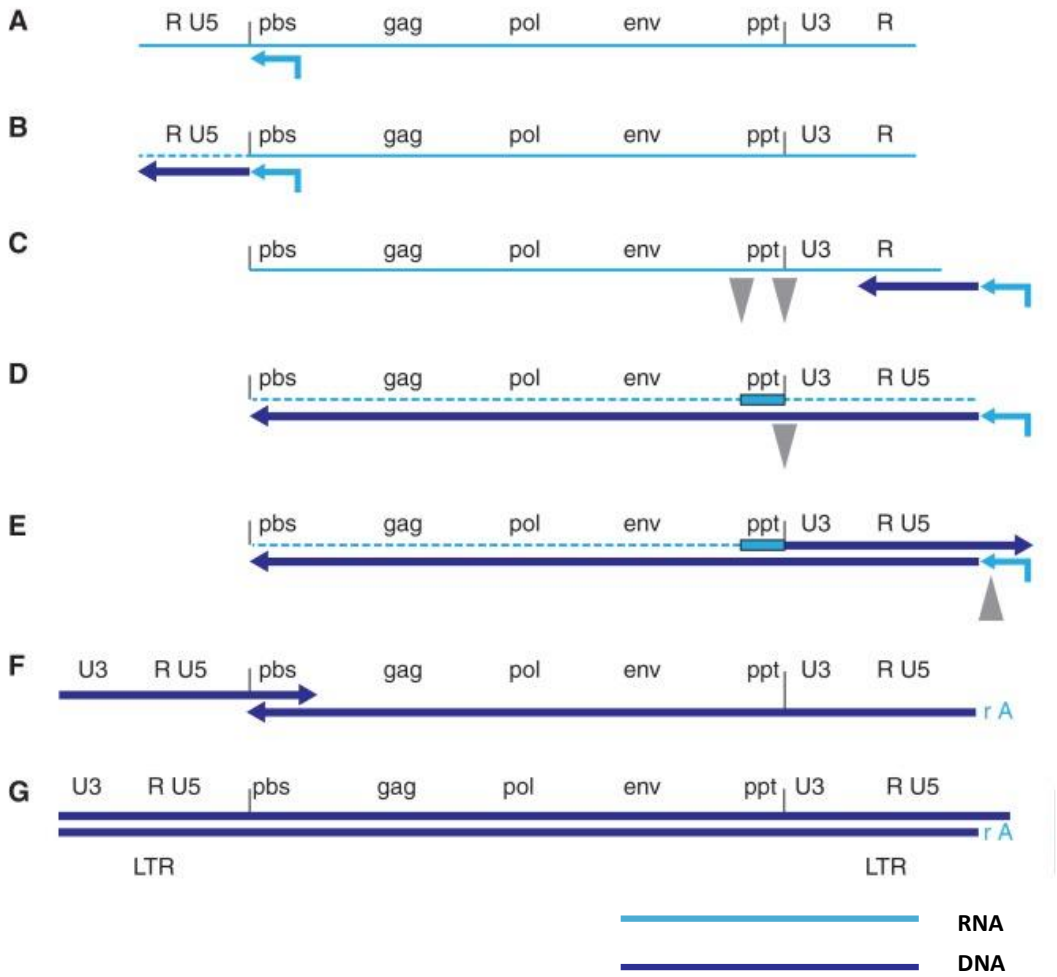


Figure 4: Reverse transcription of single stranded RNA lentivirus genome into double stranded DNA
Host cell tRNA binds to the primer binding site (A) allowing reverse transcriptase to initiate reverse transcription and generate minus strand DNA in the 5' direction (B). Once complete the minus strand DNA anneals to the 3' LTR (C) and reverse transcriptase generates the remaining DNA minus strands (D) which is accompanied by RNase H activity degrading the RNA genome. The Poly purine track is resistant to RNase H activity and acts as a primer for positive strand DNA synthesis in the 3' direction (E). The positive strand of DNA transfers and anneals to the 5' LTR of the minus strand (F) allowing reverse transcriptase to complete reverse transcription (G).

1.2.4 Nuclear entry

Nuclear entry is a complex and imperfectly understood process. In short, HIV-1M must cross the nuclear membrane before it can integrate. It is known that HIV-1M passes through nuclear pore complex (NPC) which functions to gate transport from the cytoplasm to the nuclear environment [89,90]. However, this process is limited by the internal

size of the NPC channel which is thought to only be able to passively transport proteins of less than 40-60kDa and about 40nm [91,92]. As current evidence suggests that the capsid core remains intact during cytoplasmic trafficking it remains unclear how the NPC can transport the large intact capsid core, which is about 50-60nm in width and 210nm in length [93], into the nucleus. It is known that many cofactors, which bind to capsid, influence nuclear entry (see section 1.4) but the exact order of events is still unclear. Currently, there are two major competing theories about how HIV-1M can overcome this limitation. (1) The HIV-M capsid core uncoats, fully or partially, at the NPC and only the viral genome and integrase enter the nucleus. (2) The entire capsid core enters the nucleus, promoted by interactions with nucleoporins and modulation of NPC size, before uncoating occurs. Consistent with this second hypothesis, capsid can readily be detected within the nucleus of infected cells and infection is lost if capsid doesn't enter the nucleus [94,95].

1.2.5 Integration

Integration of the DNA copy of the viral genome occurs after nuclear entry and is mediated by the virally-encoded integrase enzyme. The integrase enzyme is transported into the target cell nucleus by being packaged into the intact viral capsid. After reverse transcription has completed, it is known that the viral DNA becomes tightly associated with a octamer of integrase termed the intasome [96,97,98] which is known to be the functional unit of integrase, along with other viral protein, to form the pre-integration complex (PIC) [99]. Once in the nucleus, the PIC interacts with the host protein LEDGF/p75 [100] this is thought to aid in efficiency of integration by tethering the PIC to chromosomes and facilitating integration site selection [101,102]. In general, integration of lentiviruses is strongly biased towards transcription units of highly expressed genes [103–107]. This is

perhaps expected due to the conserved nature of the integrase-LEDGF/p75 interaction reported across the lentivirus genus [108,109]. Confirmation that LEDGF/p75 plays a role in selection of high expression gene integration sites came from both knockdown [110] and knockout [111] studies which both showed a statistically significant redistribution of integration sites to genes with lower expression. It is also known that the capsid binding cofactors CPSF6 and CypA also play a role in integration site selection [112], their role in infection is discussed in section 1.4.

Once an integration site is selected integrase catalyses two distinct reactions which result in the formation of a provirus. Firstly, integrase hydrolyses two nucleotides from each 3' end of the viral DNA [113] this creates a 3'-OH group. Next integrases uses the 3'-OH group to cleave the host DNA while simultaneously ligating the viral 3' ends to the newly formed host DNA 5' overhangs [114]. The resulting structure is a gapped recombinant intermediate with the 5' end of the viral genome unattached to the host DNA. Host cell DNA repair enzymes complete the integration event and form the provirus [115,116]. However, it is possible that viral DNA can be targeted by non-homologous end joining DNA repair machinery which results in the ligating of the LTRs at either end of the viral DNA. The resulting structure is termed a 2-LTR circle [117]. As 2-LTR circles are only formed within the nucleus of cells [118] and the unique LTR-LTR junction which can be assayed via qPCR [119], 2-LTR circles are often used as a marker for nuclear import of viral DNA. Once integrated into the genome the provirus is replicated at the same time as the cellular DNA during cell division. The provirus acts as a template for both viral transcripts for *de novo* viral protein synthesis, and for viral genome production.

1.2.6 Viral transcription and translation

Once integrated, transcription of the HIV-1 provirus is driven by promoters found in the 5' and 3' LTRs. Both LTRs are able to drive transcription, as Tat-inducible promoters, however the 5' LTR exerts a dominant control and likely acts as the primary promoter [120]. The both LTRs were found to contain a large number of transcription factor binding sites, which allow for the control of promoter activity in a context dependent manner [121] [122]. Essentially transcription begins when cellular transcription factors, such as SP1, NF-kB[123] or NFAT [124], are recruited to the U3-R regions of the 5' LTR. Once bound, these factors allow for RNA polymerase II (RNAPII) to bind to the TATA box within the LTR thus initiating transcription. This alone, however, leads to very low numbers of full length transcripts due to transcriptional pausing near the transcription start site. Once paused, RNAPII recruits negative elongation factor (NLEF) which results in premature transcription termination [125]. This is overcome by the viral transactivator protein (Tat) which binds at an RNA stem-loop structure within the 5' LTR called the transactivation response region (TAR). Once bound to the TAR, Tat recruits the P-TEFb complex which phosphorylates a serine in the CTD of RNAPII [126]. This results in the increased avidity for RNPII which allows for the production of full length transcripts [127,128].

Once transcription is complete these transcripts encounter a problem at nuclear export. It is known that un-spliced, intron containing pre-mRNAs, are retained in the nucleus, via interactions with splicing factors [129], until splicing is completed or they are degraded. As both the viral polyproteins, *env* and the viral genome are all incompletely spliced and therefore would be retained in the nucleus, HIV-1M needs to overcome this block to nuclear export. The regulator of expression of viral proteins (Rev) protein overcomes this block. It does this by binding to specific sequence termed the Rev response element (RRE), which is present in all incompletely spliced viral mRNAs, and promotes nuclear export of unspliced mRNAs [130–132]. Once bound to the RRE, Rev

multimerises [133] [134] allowing Rev to bind to chromosome region maintenance 1 (CRM1) thus promoting nuclear export [135]. HIV-1M does use fully spliced mRNAs, these encode the viral regulatory proteins Tat, Rev and Nef. Therefore, HIV-1M gene expression can be divided into two distinct stages, early and late. Early gene expression are those which are fully spliced and thus Rev independent whereas late genes are incompletely spliced and therefore dependent on Rev mediated nuclear export.

Once the mRNA is exported into the cytoplasm it is translated by host cell translation machinery. During translation of the Gag polyprotein, there is a 5-10% chance of a -1 frameshift event which results in the production of Gag-Pol instead [136]. This frameshift event has been shown to be caused by a poly-uridine track [137] in the Gag-Gag-Pol overlap region [136] due to the usage of a non-optimal Leu codon, which is recognised by a low concentration tRNA, and thus causes a ribosomal pause. During the ribosomal pause, there is a small chance of the ribosome shifting to bind to the UUU codon, at the -1 position, due to it being more thermodynamic stable [138]. It is thought that this frameshift mechanism is used by HIV-1 to ensure the correct amount of Gag and Gag-Pol is packaged made.

1.2.7 Viral assembly and budding

After translation the viral polyproteins, Gag and Gag-Pol, are trafficked to the plasma membrane of the infected cell. Viral proteins cluster at lipid rafts, which is the likely site of virion budding shown by the high level of cholesterol [139] [140] and specific incorporation of raft proteins into the viral membrane [141]. Once at the plasma membrane Gag becomes anchored there, a process dependent on myristoylation of the matrix domain of Gag and the matrix domain interacting with phosphatidyl inositol (4, 5) bisphosphate (PI(4, 5)P2) [142]. The matrix domain of Gag binds directly to PI(4, 5)P2 which causes a conformational

change to expose the myristoyl group, known as the myristoyl switch [143]. Once exposed the myristoyl group is able to interact, electrostatically, with PI(4, 5)P2 which strongly anchors Gag to the membrane [143]. Interestingly, the viral polyproteins do not polymerise until they reach the plasma membrane, thought to be induced by Gag-membrane interactions where anchored Gag-RNA complex act as nucleation sites [144], and instead are trafficked through the cell mainly as monomers [145]. The viral genome is selectively bound to Gag by the use of the packaging signal (Ψ) which is located near the 5'-UTR [146,147] and is essential for the formation of infectious particles [146,148,149]. It is thought that the packaging signal acts as a high affinity ligand which is able to outcompete other RNAs binding to Gag [149,150]. The ability of Gag to bind the viral genome is mediated by the nucleocapsid region of Gag. it has been shown that Nucleocapsid is able to bind HIV RNAs in isolated from other regions of Gag [151,152]. It is also known that Nucleocapsid contains two positively charged zinc finger domains, with the sequence Cys-X2-Cys-X4-His-X4-Cys where X can be any amino acid [153], which are able to recognise the viral RNA and are essential for packaging of the viral genome as when mutated no genome could be detected within produced particles [146,153]. Once Gag is bound to the viral genome, polymerisation initiates using the Gag-RNA complex as a nucleation site. Gag polymerisation leads to the formation of the immature Gag lattice where the Gag molecules are oriented radially with the matrix domains bound to the inner leaflet of the membrane and the carboxy-terminal of Gag towards the centre of the particle [154,155]. The viral envelope glycoprotein traffics to the plasma membrane via the secretory pathway and independently of Gag. The intracellular transmembrane domain of Env has been shown to play a role in lipid raft localisation [156,157] and, via interactions with the matrix domain of Gag, promote incorporation into the virion [158–160].

Once the immature virus particle is formed, the virion buds from the cell using host cell machinery. The endosomal sorting complexes required for transport (ESCRT)

factors facilitate viral budding. The HIV-1M P6 domain of Gag contains two essential late domains which play a role in recruiting ESCRT machinery and aid in budding. (1) The PTAP domain (Pro-Thr-Ala-Pro) in p6 mimics a cellular ESCRT-I recruiting motif [161]. This PTAP domain is able to directly bind and recruit the early ESCRT machinery, TSG101, to the site of budding [162,163]. (2) The YPXL (Tyr-Pro-X-Leu, where X can be various sequence and isn't a set length, also in p6. This YPXL domain is used to bind and recruit AIP1/ALIX, another ESCRT factor [164]. Once again, this YPXL domain is thought to mimic cellular AIP1/ALIX recruit signals [165]. Once both TSG101 and AIP1/ALIX are recruited they act to recruit downstream ESCRT complexes which in turn mediate membrane fission and particle release.

1.2.8 Particle maturation

The final stage of the HIV-1 life cycle is particle maturation. It is in this stage the virally encoded polyproteins which were packaged into the immature particle, Gag and Gag-pol, undergo proteolytic cleavage by the virally encoded protease to liberate their component proteins. Cleavage of Gag gives rise to the structurally proteins, matrix, capsid, nucleocapsid and p6 whereas Gag-Pol encodes the enzymatic proteins protease, reverse transcriptase and integrase. Once cleaved these proteins undergo a large rearrangement forming into the mature, fully infectious virus particle. The viral protease helps to regulate the site of maturation as its active form is a dimer [166–168] and can only weakly associate with itself when in the Gag-Pol polyprotein. However, these highly transient interactions are enough to allow auto-processing of Gag-Pol thus freeing protease to from stable dimers and cleave the remainder of the polyproteins [169]. Due to the rare nature of these highly transient interactions, it is likely that they only occurs within the tightly clustered environment found within the budded immature HIV-1M particle. This would prevent premature polyprotein cleavage, an event which was shown to renders the virus non-infectious by experiments where a linked dimer of protease was expressed [170].

The Gag polyprotein is cleaved at five sites and this occurs in a specific sequential fashion to aid correct particle formation [**Figure 5**]. It has been shown that cleavage is a highly ordered process, with different cleavage sites having different kinetics. There are three broad categories. Rapid, SP1-NC cleavage, intermediate, SP2-P6 and matrix-capsid and slow which is nucleocapsid-SP2 and capsid-SP1 [171]. Cleavage of the SP1-nucleocapsid site allows nucleocapsid to condense on the viral genome and forms the ribonucleoprotein (RNP) [172] followed by the cleavage the matrix-capsid sites which allows for the disassembly of the immature Gag lattice. Finally the nucleocapsid-SP2 and the capsid-SP1 sites, which once cleaved allows for the formation of the genomic RNA dimer [173] and the polymerisation of capsid and formation of the mature capsid core [172].

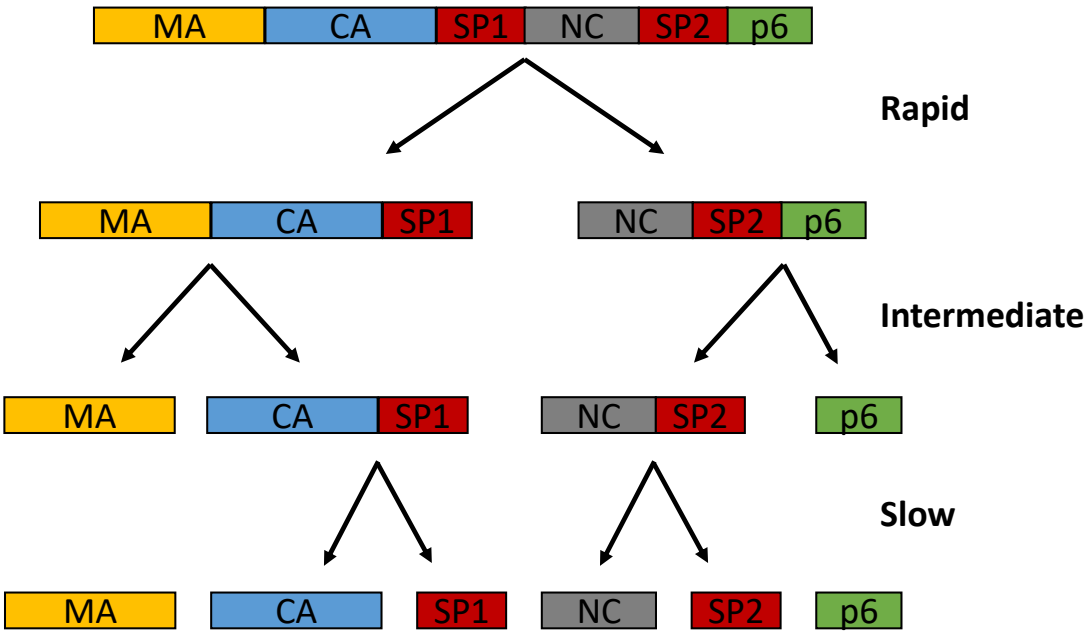


Figure 5: Sequential cleavage of HIV-1M Gag
Schematic overview of HIV-1M Gag cleavage during particle maturation. To aid in correct particle formation the cleavage sites in HIV-1M Gag and different kinetics of cleavage. Each cleavage site can be classed as either rapid, intermediate or slow.

1.2.9 Viral accessory proteins

HIV-1M encodes four viral accessory proteins, Vpu, Vpr, Nef and Vif. These accessory proteins function to modulate the cellular environment to promote conditions for optimal infection and aid HIV-1M in immune evasion. Here, what is known about each accessory protein and the role in plays in HIV-1M infection is described.

Viral protein U (Vpu) is a small, dimeric integral membrane protein which has been shown to be able to induce proteasomal degradation of CD4 via recruiting a ubiquitin ligase to the cytoplasmic tail of CD4 [174]. Unlike the negative regulatory factor (Nef) of HIV-1M, which also degrades CD4, Vpu targets the cytoplasmic tail of CD4 which is retained in the endoplasmic reticulum due to binding to newly synthesised viral envelope protein which is being transported to the plasma membrane. This degradation of CD4 from the membrane has been found to increase virion release and Env incorporation into particles [175,176] as cell surface CD4 can interfere with virus budding. Alongside this activity, Vpu has also been found to be able to counteract the interferon-stimulated gene tetherin and promote particle release [177–180]. Tetherin is a membrane protein associated with lipid rafts [181] which if not inhibited by Vpu leads to an accumulation of virus particles on the exterior of the cell [178]. The exact mechanism of Vpu's anti-tetherin activity is still be investigated. However it is known that the ability to degrade CD4 is dispensable for the anti-tetherin function [182] and that Vpu co-localises with tetherin in several cellular locations [180,179]. Taken together, this could suggest that Vpu disturbs tetherin protein sorting and thus prevents it from reaching the plasma membrane.

Nef is a myristoylated protein which has been found to be associated with the cellular plasma membrane. Like Vpu, Nef has been found to modulate cell surface CD4. Nef binds to the cytoplasmic tail of CD4 at the cellular membrane and recruits clathrin

adaptor protein complex 2 (AP2), this in turn leads to endocytosis of CD4 and ultimately lysosomal degradation [183]. Nef has also been found to modulate cell surface levels of both MHC class I receptor [184] and the T-cell receptor (TCR) [185]. Modulation of both MHC-1 and the TCR results in dysfunctional immune response towards HIV-1M infected cells. Loss of MHC-1 allows HIV-1 infected cells to escape recognition by cytotoxic T lymphocytes [186] whereas loss of TCR prevents immunological synapse formation [187] an essential part of T-cell activation. In both cases, Nef targets MHC-1 and the TCR to endosomes and thus prevents their proper protein soring and function [187–189]. Finally, the transmembrane protein SERINC has been found to potently inhibit HIV-1 infection by being incorporated into budding virions and preventing subsequent viral entry to target cell [190]. SERINC induced restriction of HIV-1M is thought to be overcome by Nef retargeting SERINC to endosomes [190,191].

Virion infectivity factor (Vif) is the HIV-1 accessory protein which inhibits the viral restriction factor apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (A3G) [192]. As a member of the APOBEC family of enzymes, A3G has the activity to deaminate cytidine to uridine [193]. It is this A3G function which acts to restrict HIV-1M infection. In the absence of Vif, A3G is able to interact with Gag and the viral mRNA genome during viral budding and package into progeny virions [194]. Once in the progeny virion, A3G functions to deaminate cytidines in the viral genome which, during reverse transcription, leads to guanosine to adenosine hypermutation [195,196] . Hypermutation of the viral genome leads loss of genetic integrity and although doesn't prevent infection of target cell it does potently prevent any further transmission. There is also growing evidence that A3G can interact with HIV-1M reverse transcriptase and directly inhibit reverse transcription [197–199]. Vif is able to counter this by recruiting a cullin-ring ubiquitin ligase to A3G and inducing its polyubiquitnation and ultimately its proteasomal degradation [200–202] before it can be packaged into the immature particle.

Viral protein R (Vpr) is packaged into virus particles during viral budding via specific interactions with the p6 regions of Gag [203] and is essential for infection of macrophages [204,205]. After target cell entry, Vpr shuttles to the nucleus and facilitates viral replication [206,207]. Perhaps the best reported function of Vpr is its ability to prevent cell cycle progression and arrest infected cells in G2 [208–210] a phase where the viral promoter, present in the 5'-LTR, is thought to be more active [211]. Vpr is thought to mediate this effect by activating the SLX4 complex, which plays a role in DNA damage repair but can also stall transcription forks [212]. Alongside this, Vpr has also been implemented in aiding in nuclear entry of HIV-1 in non-dividing cells such as macrophages via interactions with importin-alpha, which promote subsequent interactions with the nuclear pore [213–215].

1.3 Lentiviral capsid core

1.3.1 Retroviral capsid

Despite high level of divergence between capsid sequences, the retroviral capsid is a highly conserved feature of the *retroviridae* family. However, within the capsid there is a region with a strikingly high level of conservation, the major-homology-region (MHR). Except for a few minor polymorphisms the MHR is conserved within all mammalian and avian retroviruses [216]. Furthermore, the role of the retroviral capsid is also highly conserved. Each retrovirus encodes a capsid which forms protein shell that surrounds the viral genome.

Whilst all retroviruses encode a capsid shell, its morphology is variable depending on the virus. For example, the alpharetrovirus Rous sarcoma virus (RSV) was shown via electron microscopy to form essentially isometric cores with a polygonal appearance [217]. Even within the same retrovirus genus there is variation. EIAV and Visna virus seems to form a mixture of tubular or circular particles by electron microscopy [218,219] whereas HIV-

1M has been shown to form a fullerene core structure with a wide and narrow end [93,220–222]. Interestingly the gammaretrovirus murine leukemia virus (MLV) does not form a sealed structure. It has recently been shown, via a combination of X-ray crystallography and cryoelectron tomography, that the mature capsid core structure of MLV is composed of several multi-layered curved capsid sheets which wrap around the genome [223]. Of all these capsids, the structure of HIV-1M is most studied and perhaps best understood.

1.3.2 HIV-1M capsid structure

The HIV-1M capsid is the most studied lentivirus capsid, with less known about more distant lentivirus capsids. The HIV-1M immature particle is formed of ~5000 Gag molecules [224] which after proteolytic maturation, driven by the viral protease, form into the mature conical shell [**Figure 6A**] surrounding the RNA genome which can be observed as an electron-dense core, seen by electron microscopy. Interestingly, of the ~5000 capsid monomers packaged into the mature virion, only ~1200 are used in the formation of the capsid core [221]. It remains unclear as to the function of the abundant population of capsid monomers which aren't incorporated into the mature capsid core. The capsid protein contain two domains connected by a flexible linker region [**Figure 6C**]. The amino-terminal domain (NTD) is formed from seven alpha-helices and a Beta-hairpin [225]. Whereas the carboxy-terminal domain (CTD) contains only four alpha-helices [226]. The capsid monomers are able to assemble into higher order hexameric [**Figure 6B**] and pentameric assemblies.

Mature HIV-1M capsid cores contain ~250 hexameric subunits and exactly 12 pentamers [220,222]. These pentamers are always located at the vertices of the capsid with five at narrow end and seven at the broad end and allow the closure of the curved hexagonal lattice. Without pentamers open ended tubular assemblies are formed *in vitro*

[220]. The length of the capsid core is populated solely by hexamers. Using cryo-electron microscopy the dimensions of the mature capsid core have been determined [93]. They were 119 ± 11 nm along the length of the central axis and a diameter of 60 ± 8 nm at the thick end [93]. The mature capsid is thought to be stabilised by three main intramolecular interaction between capsid monomers. First is the 18 helix bundle that is formed in the centre of the hexamer by alpha-helices one-three in the NTD of each capsid monomer. Second is the CTD-CTD interaction at the dimerization interface which helps to link adjacent hexamers in the hexagonal lattice. Finally is the NTD-CTD interaction between neighbouring capsids monomers within a hexamer, helix 4 of the NTD interacts with the groove in the CTD.

Much of our understanding of higher order capsid structures comes from x-ray crystallography. Imaging of the capsid hexamer structure has been aided with the introduction of two disulphide bridges, by mutagenesis to cysteines of residue Ala 14 and Glu 45, that cross link the capsid protein, this has allowed for high resolution X-ray crystal structures of hexamers to be generated [227,228]. These cross linked structures were found to be accurate models as they were highly similar to the recent x-ray crystal structure of wildtype, non-cross linked, capsid hexamers [229]. Despite progress in the imagining of the capsid hexamers, imaging of the capsid pentamers remains poor. A disulphide cross linked pentamer structure was generated which has the expected 5-fold axis of symmetry [221]. However, unlike the cross-linked hexamer structure, the cross-linked pentamer structure did not agree with the recent cryo-electron tomography structures of both hexamers and pentamers in mature capsid core [222].

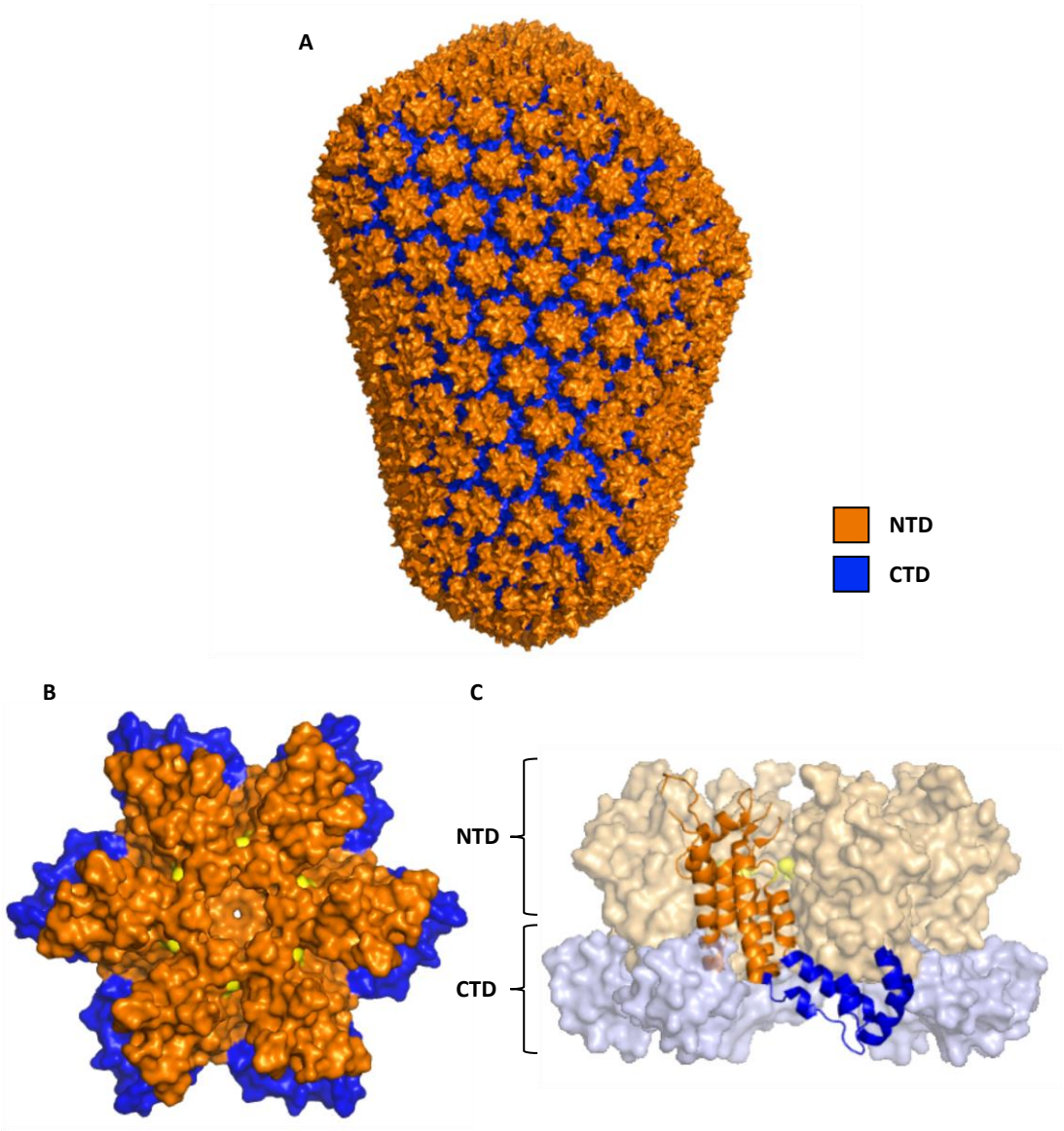


Figure 6: HIV-1M capsid structure
A) The mature capsid core contains ~250 hexameric units and 12 pentameric units. These pentamers are arranged so there are five at the narrow when and seven at the wide end. B) A capsid hexamer. C) A side-on look at a capsid monomer in the context of a hexamer. Through orange shows the NTD whereas blue shows the CTD. Taken from Zhao et al, 2013.

1.3.3 Role of capsid in HIV-1M lifecycle

Despite the HIV-1M capsid being one of the most intensively investigated proteins, its function in the HIV-1M life cycle still remains a point of major investigation within the field. The field is mainly focusing on the key question of when and how capsid uncoating occurs within the lifecycle. Currently, there are two popular hypothesises regarding the

capsids role in infection. (1) The early uncoating model, in which the capsid spontaneously disassembles post-entry and plays no further role in the life cycle with reverse transcription occurring within the cytoplasm after uncoating. (2) The encapsidated reverse transcription model, where the capsid play a role in the coordination of post-entry steps and provides a safe environment where reverse transcription can occur within the capsid. Currently, most of our understanding of the role of capsid in HIV-1M infection comes from work in cell lines and it is possible that the role of capsid within primary cells is different from that in cell lines, however this remains unclear.

Evidence for the encapsidated reverse transcription model is growing and can best be seen via capsid-cofactor interactions (more detailed review in section 1.4). In short, disruption of the CypA or CPSF6 capsid interaction, chemically, via mutagenesis of the capsid or by cofactor deletion, results in inhibition of infection in primary macrophages and T cells [230]. This can be countered by the observation that in cell lines, such as Ghost cells, these cofactor-capsid interactions are non-essential for infection. However, it has been shown that disruption of these interaction in cells line does cause a dramatic retargeting in integration site selection [112]. Another good example of the importance for capsid to remain intact is the stability mutants. Altering capsid stability to be either hyper stable, E45A, or unstable, P38A, leads to different reverse transcription kinetics and a replication defect in primary T cells [231]. If the capsid spontaneously uncoated upon cellular entry, mutations to make the capsid more unstable, i.e. uncoat faster, shouldn't alter infection dynamics.

Modern fluorescent microscopy techniques have improved our understanding of capsid uncoating further. One such technique uses fluorescently labelled oligomerised CypA which has the advantage to only be able to bind to the mature capsid structure and is lost upon uncoating [232]. Using this method it was shown that upon infection a large

proportion of viral cores uncoated, which is consistent with a spontaneous uncoating hypothesis. However, it was subsequently found that uncoating in the cytoplasm resulted in proteasomal degradation of the virus and that capsid was required for nuclear localisation of the particle [233]. Consistent with this observation, using YFP labelled APOBEC3F (A3F) has shown that the capsid stability mutants, P38A and E45A, both cause a decrease in the number of particles with nuclear envelope residence. Taken together, these observation support the hypothesis that particles that undergo premature uncoating are non-infectious [233,234].

Finally, it is known that the HIV-1M genome can be degraded by the host cellular exonuclease TREX-1 [235]. It is has been shown that in the absence of this degradation, the genome can be sensed by cGAS leading to production of both IFN-β and CXCL10 in a cGAS-STING dependent innate immune response [236]. In a model where the HIV-1M genome is free within the cytoplasm is it unclear how it is protected from degradation or being sensed as both events are restrictive to HIV-1M infection. The encapsidated reverse transcription model, provides a hypothesis on how the virus is able to avoid these events and be productively infectious. It is unclear how a model in which the capsid rapidly disassembles post entry explains the evidence above. One thing that is becoming clear is that the capsid of HIV-1M play an important role in infection, often by acting as a platform to form interactions with host cell factors.

1.3.4 HIV-1M electrostatic channel

Early models suggested that the HIV-1M capsid is impermeable and undergoes rapid and spontaneous disassembly upon entering the cytoplasm. However, we now hypothesise that the capsid is porous and remains intact throughout reverse transcription, and that uncoating is a highly coordinated event that occurs at or after nuclear import. Central to

this new hypothesis is the electrostatic channel in the mature capsid core. The electrostatic channel is formed at the six or five-fold axis of symmetry in each hexamer and pentamer of the capsid, respectively, and contains two structural features. A molecular iris formed by the N-terminal domain β-hairpin and an electropositive ring, formed by arginine 18 in HIV-1.

One major obstacle for the encapsidated reverse transcription model was the question of how dNTPs entered the capsid so that reverse transcription may take place. This question was answered by the study of the electrostatic channel, more importantly the electropositive Arg ring. Initially it was found that if purified capsid cores were supplied with dNTPs there was an accumulation of early reverse transcription products which were resistant to both DNase and RNases [237]. Investigating this observation, it was shown by X-ray crystallography that dNTP can directly interact with Arg 18 within the HIV-1M capsid [237]. Furthermore, fluorescence anisotropy experiments using fluorescently labelled dNTPs show very rapid on rates with HIV-1M capsid hexamers, as expected for electrostatic interactions. Mutations in the electrostatic ring results in a severe defect in infectivity which corresponds with a decrease in reverse transcriptase products [237,238] . This defect is thought to be caused by an inability to recruit dNTPs to the capsid core. The capsid mutant R18G has been shown to form normal mature conical capsid cores, using electronmicroscopy [238] and by being capable of abrogating TRIM5α -mediated restriction [237], but has a infectivity defect. The infectivity defect of R18G can be further dissected by the production of chimeric hexamers containing cores, which have a mixture of both wildtype and R18G capsid. As the ratio of glycine (Gly) increases at position 18 within mixed capsid hexamers the ability to bind dNTPs to the capsid is reduced in a stepwise fashion, as shown by fluorescence anisotropy [237]. Furthermore, as the ratio of Gly increases both the infectivity and reverse transcription simultaneously decrease.

Interestingly the electropositive ring has also been found to be able to bind to other charged molecules including ATP, IP6 and mellitic acid [237,239]. Of these, the most interesting is IP6 (more details in section 1.4.5). The IP6-capsid interaction has been shown to play a role in the regulation of capsid stability [239]. It is thought that the stabilising effect of IP6 in the mature capsid lattice is due to charge neutralisation. In the mature capsid lattice, IP6 binds the positively charged Arg 18 which are tightly clustered in the electrostatic channel at the six fold axis of symmetry and would exhibit strong charge repulsion. It was found that when the capsid core were bound to IP6 the thermal stability of the cores increase by 10 degrees and the half-life of the particles increased from minutes to hours [239]. More interesting still is the observation that IP6 bound purified capsid cores have a 100 fold increase in accumulation of reverse transcription products. Taken together, this suggest that IP6 is acting as a stability factor for the capsid core in the HIV-1M life cycle. Increased capsid core stability could be important during infection of immune sensing competent cells, such as macrophages, where premature uncoating leads to a restrictive immune response. The electrostatic channel, by binding IP6 at Arg 18, provides the mechanism in which HIV-1 regulates reverse transcription and stability simultaneously. It is clear that the electrostatic channel of HIV-1M is an important feature of the capsid core and plays a role in both control of reverse transcription and capsid stability, and thus indirectly to immune evasion.

1.3.4 Capsid targeted restriction factors

The HIV-1M capsid can be seen as a double edged sword, as it provides a safe environment for the genome but due to its large size and highly regular and ordered structure it can be targeted by innate immune sensors and restriction factors. Arguably, the study of lentiviral capsid targeting restriction factors began with the identification and

subsequent cloning of Fv1 [240,241], a capsid targeting [242–244] murine restriction factor able to potentially restrict infection with MLV [245]. Since then several such factors, including TRIM5, TRIMCypA and MX2, and been identified to restrict HIV-1M infection in a capsid dependent fashion.

TRIM5, originally termed Ref1 in humans [246] and Lv1 in non-human primates [247], was identified by overexpression of a macaque cDNA library in HeLa cells [248] and is perhaps the best studied HIV-1M capsid targeting restriction factor. Composed of four distinct regions, the N-terminal ring, B-box, a coiled-coil domain and a PRYSPRY domain, it was found that the PRYSPRY domain is responsible for capsid targeting [248,249]. Interestingly TRIM5 blocks infection at reverse transcription but when the proteasome inhibitor, MG132, is used, reverse transcription is not prevented but infection is [250]. Combined with the knowledge that TRIM5 restriction results in the loss of viral particles from the cell [251,252] and that TRIM5 is rapidly turned over by proteasomal degradation [253] it is now hypothesised that TRIM5 blocks infection in two stages. Firstly, TRIM5 binds the retroviral capsid and acts to stabilise the structure via the formation of a hexagonal lattice over the entire capsid core [254,255], this prevents infection but not reverse transcription. Secondly, TRIM5 is targeted for proteasomal degradation which results in the loss of both viral particles and reverse transcription products from the cell [251,252,256]. Although HIV-1M is able to avoid restriction by human TRIM5 [257,258] [see section 1.4.1] other non-primate lentiviruses such as EIAV and FIV [259,260] as well as the gammaretrovirus N-tropic MLV [261] are highly restricted by human TRIM5. Alongside this, TRIM 5 has been shown to be capable of inducing an innate immune response via the activation of TAK1 [262,263]. By interacting with the ubiquitin conjugating enzyme UBC13-UEV1A, TRIM 5 activates TAK1 via the synthesis of N-terminal anchored K63-linked ubiquitin chains [264] a process which is greatly enhanced by TRIM 5 interacting with the lentiviral capsid lattice [263]. Once activated, the kinase TAK1 complex stimulates

an AP1 and NF-kB driven immune response, which is characterised by production of CXCL10 [263]. Found in both old [265] and new [266] world primates TRIM-Cyp is a variant of TRIM5 where the PRYSPRY domain has been replaced by CypA. TRIM-Cyp is highly restrictive to HIV-1M infection and targets the CypA binding loop on the HIV-1M capsid exterior.

Myxovirus resistance protein 2 (MX2/MxB) is an interferon-inducible protein which has GTPase activity and is able to inhibit HIV-1M at nuclear entry [267–269] as observed by reduced formation of 2-LTR circles but not linear viral DNA. It has been shown that MX2 is able to bind directly to the capsid core [270] and that this binding to the capsid is essential for restriction of infection [270–272]. Consistent with this, is the observation that mutations in the HIV-1M capsid protein leads to resistance to Mx2.

SUN2, also known as UNC84B, is a membrane protein which spans the nuclear membrane [273,274] and was identified as a restriction factor against HIV-1M in an overexpression screen of interferon stimulated genes (ISGs) [275]. SUN2 is a component in the linker of nucleoskeleton and cytoskeleton complex (LINC) [276] which plays a role in controlling nucleus morphology, cell polarisation, organisation of chromosome tethering and DNA repair [277–279]. Overexpression of SUN2 inhibits HIV-1M and HIV-2 in both cell lines and in dendritic cells in a capsid specific fashion by preventing nuclear entry, as seen by the decrease in 2-LTR circle formation [280]. Mutations in the capsid of HIV-1M can render the virus resistant to SUN2 mediated restriction. Interestingly, transmitted/ founder viruses of HIV-1M were found to be resistant to SUN2 overexpression [280]. SUN2 has also be implicated in blocking HIV-1 transcription via chromatin modulation [281].

Recently a new capsid directed restriction factor has been reported that is active in T, B and myeloid cell lineages, termed Lv4 [282]. Although the factor is yet to be identified, the block imposed by Lv4 has been characterised. It has been found that Lv4

imposes a block to infection for the HIV-2/SIVsm lineage of lentiviruses after completion of reverse transcription and 2-LTR circle formation but before provirus formation [282]. This block was found to be capsid specific and independent of TRIM5, as depletion of TRIM5 did not rescue the viral infection.

1.4 Capsid Cofactors

It is thought that the HIV-1M capsid plays an important role in the regulation of several key stages of replication. This is thought, in part, to be due to the capsid core's ability to interact with several host factors, termed cofactors, to ensure optimal infection. So far it has been shown that the HIV-1M capsid core is able to bind to cyclophilin A (CypA) [283], cleavage and polyadenylation specificity factor subunit 6 (CPSF6) [284], nucleoporin 153[285,286] and 358 [112,287] (NUP153 and NUP358 respectively), transportin 3 [288,289] (TNPO3) and the inositol compound inositol hexaphosphate [239,290] (IP6). These cofactors have been shown to regulate immune evasion, nuclear entry, reverse transcription and capsid stability. Furthermore, there is growing evidence to suggest that cofactor dependency might be cell type dependent. In this section some of the key information regarding each cofactor is summarised.

1.4.1 CypA

CypA is a member of the peptidyl-prolyl cis-trans isomerase (PPIase) family, which catalyses the cis-trans isomerisation of proline peptide bonds. This process is thought to increase the speed of protein folding. However, the PPIase activity of CypA has not been implicated in its role in influencing HIV-1M infection. CypA is a highly abundant cellular protein, estimated to account for about 0.1-0.4% of total protein of the cell [291,292], and is localised to the cytoplasm [293]. Originally shown in 1993 to be able to bind to the HIV-1M

capsid in a yeast two hybrid screen [283], CypA has become one of the most intensely studied cofactors. It is known that CypA binds directly to the HIV-1M capsid NTD surface at the exposed CypA binding loop at Pro 90 [283,294,295]. In both the immature and mature structure, the CypA binding loop is exposed to the cellular environment, this leads to CypA being packaged into the progeny virions [296,297]. Along with HIV-1M, the closely related SIVCpz has also been shown to package CypA into progeny virions [298,283]. However the HIV-2/SIVsm lineage does not package CypA [296,298] and seems to have no dependence on it for replication. However the capsid-CypA interaction is not specific to HIV-1M lineage viruses. FIV and the SIVagm lineage have both been shown to be able to bind CypA in a TRIM-CypA restriction assay with cyclosporine A (CsA) able to rescue infection [299]. Furthermore, the recently discovered endogenous lentivirus, Rabbit Endogenous Lentivirus type K (RELIK) [300], which is estimated to be at least 12 million years old [301] was found to be able to bind CypA via its capsid protein [302]. The observation that RELIK is able to bind CypA suggests that the capsid-CypA interaction is ancestral characteristic and was lost by the HIV-2/SIVsm lineage rather than gained by the HIV-1/SIVCpz or FIV lineage.

The packaging of CypA into the progeny virion has been shown to have no effect on infection and rather it is target cell CypA which regulates infection [303]. Although each capsid monomer contains a CypA binding loop it has been found that, due to steric clashes, not all sites are simultaneously occupied within the mature capsid hexamer [304]. Since its discovery, CypA has been shown to play in various life cycle stages, including integration site selection, nuclear import pathway selection, immunevasion, capsid stability, control of reverse transcription and escape from restriction factors. Here we discuss some of the roles CypA is thought to play in the HIV-1M lifecycle.

Study of CypA-capsid interaction has greatly been aided by the use of the immunosuppressive drug cyclosporine (CsA) which chemically prevents the Cypa-capsid interaction [305] by acting as a competitive inhibitor. It has been shown that CsA outcompetes the HIV-1 capsid for the same binding site on CypA [295,306]. The CypA-capsid interaction can also be disrupted by mutagenesis of the capsid residue Pro 90 and depletion of CypA via transient or stable knockdown. In any case, disruption of the capsid-CypA interaction leads to an infection defect in various cell lines *in vitro* [298,307–310].

The role of CypA is still being determined, however it is known to influence several important stages in the HIV-1 lifecycle. In macrophages, disruption of the CypA-capsid interaction leads to the induction of a potent innate immune response, characterised by IFN-β and IP10, which is restrictive to infection [230]. This innate immune response was found to be dependent on viral DNA synthesis, via use of inactive reverse transcriptase mutants. However, disruption of the CypA-capsid interaction in cell lines doesn't cause such an immune response. Rather, it results in a shift in integration site selection for HIV-M. It was found that the capsid mutants P90A, which has a greatly reduced affinity for CypA [112], integrates into regions of higher gene density [112]. CypA has also been implicated in regulating capsid stability, although there have been conflicting reports. Using a combination of cryoEM and molecular dynamics, CypA has been shown to be able to bind to a non-canonical site and bridge two neighbouring capsid monomers in the hexamer lattice [304]. This interaction has been suggested to be able to increase the stability of the hexameric lattice [304]. Recent work using single-particle fluorescence microscopy to measure the half-life of individual HIV-1M particles found a small stabilising affect when CypA was added to their system [311]. Recently, there have been two papers which show that the role of CypA is to protect HIV-1M from human TRIM 5 alpha mediated restriction in primary cells such as macrophages and T cells [257,258]. The proposed model

is that CypA binds to the HIV-1M capsid and this interactions prevents TRIM5alpha binding and thus prevent restriction.

1.4.2 CPSF6

CPSF6 which was originally described as a component of mammalian cleavage factor 1 complex which plays a role in the cleavage and polyadenylation of mRNAs [312]. It is known that CPSF6 is able to bind directly to the HIV-1M capsid, in a conserved binding pocket in the NTD of capsid [284]. The CPSF6 binding pocket is exposed in both capsid monomer and hexamer. It was later determined that although CPSF6 can bind weakly to the monomeric form of capsid, the affinity is greatly increased for the hexameric form [313]. This observation suggests that it is the fully assembled hexamer which is the preferred binding target rather than the monomeric form. The processing of pre-mRNAs by CPSF6 is primarily nuclear in location and is not thought to influence HIV-1M replication. However, when CPSF6 is forced into being cytoplasmic then it potently restricts HIV-1 infection. It was this initial observation which lead to CPSF6 being considered a restriction factor rather than a cofactor when it was first described. This cytoplasmic redistribution of CPSF6 can be an achieved either by deletion of its CTD, which contains its nuclear localisation signal (NLS), or by depletion of the importin TNPO3 which facilitates CPSF6 nuclear import. More details about cytoplasmic redistribution of CPSF6 can be found in section 1.4.4. To further its classification as a restriction factor, depletion of CPSF6 in Hela cells did not result in an impairment for HIV-1M infection [314]. Furthermore, viral passage in cells expressing cytoplasmic CPSF6 results in the selection of the escape mutant N74D [314] which is unable to bind CSPF6 [284].

These studies, however, were misguided as they relied on the study of cytoplasmic CPSF6, a state not normally found within cell. After studying the CPSF6 interaction in

primary cells and under normal conditions the true role of CPSF6 in the HIV-1M lifecycle as been determined. Although there is no effect on infection after depletion of CPSF6 or disruption of the capsid-CPSF6 interaction, via N74D mutagenesis, in cell lines there is a drastic effect on integration site selection [112]. Under normal conditions, HIV-1M integrates into genomic regions which are gene dense and under active transcription. However the N74D capsid mutant, which is unable to bind CPSF6, was found to integrate in random genomic locations, often in transcriptionally quiet regions [112]. Furthermore, infection with N74D was found to be independent of the cofactors NUP358 and TNPO3 (section 1.4.3 and 1.4.4 respectively) which are known to influence nuclear entry. Taken with the data that during CPSF6 depletion there is no defect in infection of cell lines, such as HeLa cells, this would suggest that there is an alternative nuclear import pathway available to HIV-1M in which CPSF6 is the gate keeper. Further understanding came from study of the capsid-CPSF6 interaction in primary cells like human macrophages. It was found that either depletion of CPSF6 or distribution of the capsid-CPSF6 interaction, via capsid mutagenesis, leads to an innate immune response in macrophages, characterised by IFN-β and IP10, which is restrictive to infection [230]. Infection in macrophages with either depletion or disruption of the CPSF6-capsid interaction was rescued by an interferon receptor blocking antibody. Interestingly, if reverse transcription is prevented, via mutagenesis of reverse transcriptase, the innate immune response wasn't triggered suggesting DNA is the pathogen associated molecular pattern which is recognised by the cell. This data suggests that CPSF6 plays a role in HIV-1 immune evasion by being recruited to the capsid and preventing detection of viral DNA synthesis. Recently it has been shown, by fluorescence microscopy, that in macrophages nuclear entry is dependent on CPSF6 and without the capsid-CPSF6 interaction there is an accumulation of viral particles at the nuclear envelope which is associated with reduced infectivity [95].

1.4.3 NUP 153 and 358

The nuclear pore complex (NPC) is a large multi-protein structure which spans the nuclear membrane and is the site of nuclear import and export [89]. Nucleoporins (NUP) account for a large majority of proteins used to form the NPC. Many NUP contain phenylalanine-glycine (FG) repeats which extend into the cytoplasmic environment and form a mesh-like structure that acts as a binding site for many cargo proteins [315]. The FG repeats also allow for selectivity of cargo binding. Originally identified in several genome wide siRNA screens against HIV-1M infection [74,288,289], NUP153 is now an established cofactor for HIV-1M and component of the NPC. Several other NUP, including 358 and 98, have also been found to play a role in optimal infection for HIV-1M. Here we focus on NUP153 and 358.

Nup153, which is nucleoplasmic, was found to be able to bind, via its FG repeats, directly to the same site on the HIV-1M capsid as CSPF6 [286]. Confirmation that NUP153 is a cofactor needed for optimal infection of HIV-1M came with depletion studies. Depletion of NUP153 in cells lines causes a large defect in HIV-1M infection [285,316], which was thought to be due to a decrease in nuclear entry as seen by a decrease in 2-LTR circle formation and thus suggests a role for NUP153 in nuclear import. Direct observation and confirmation of this hypothesis came by the use of fluorescence microscopy in NUP153 depleted cells which showed that HIV-1M had a reduced ability to enter the nucleus [94]. Taken together this all supports the hypothesis that Nup153 play a role in optimal infection of HIV-1M by aiding in nuclear import.

NUP358, which extends into the cytoplasm, has also been found to bind directly to the HIV-1M capsid. NUP358 contains a Cyp-like domain which is able to bind directly to the capsid at the CypA binding loop [112]. Like Nup158, confirmation that NUP358 is

essential for optimal infection for HIV-1M came from transient and stable depletion of NUP358. It was found that in NUP358 depleted cells HIV-1M was inhibited for infection, again that block was found to be at the stage of nuclear entry, as shown by reduced 2-LTR circle formation [112,316,317]. Taking into account the orientations of both NUP158 and 358, nuclear and cytoplasmic respectively, one hypothesis is that Nup358 makes the initial interaction with the HIV-1M capsid which acts to tether the particle in close proximity to the NPC. Following this, NUP153 may then be able to bind the capsid and drive the nuclear import of the virus.

1.4.4 TNPO3

Originally the importin-β family member transportin-3 (TNPO3), a Ran-GTP-dependent nuclear import protein, was identified as a cellular factor essential for optimal infection for HIV-1M via genome wide siRNA screens [288,289]. Confirmation of the importance of TNPO3 for HIV-1M came with depletion experiments which showed a decrease in viral integration but no effect on reverse transcription [318,319]. There is also a small observed decrease in formation of 2-LTR circles in TNPO3 depleted cells suggesting a role in nuclear import for HIV-1M [112,320, 321]. Taken together, this would suggests that depletion of TNPO3 imposes two blocks to infection one at nuclear import and another, stronger, inhibition at integration. Initially it was thought that TNPO3 influenced the HIV-1M lifecycle by interactions with integrase [320,322], however it was shown that this ability did not influence infection [323,324]. Rather it was found that TNPO3's interaction with the capsid core is what influences infection [323]. Furthermore, it was found that point mutants in capsid can alter TNPO3 usage. The capsid mutant N57A and N74D were found to use a TNPO3 independent infection pathway [112]. Taken together, this all suggests that TNPO3 plays a role in nuclear translocation of HIV-1M. Modern fluorescence microscopy

techniques have also been used to show that when TNPO3 is depleted HIV-1M fails to enter the nucleus [94]. This is consistent with the hypothesis that the capsid core plays a role in nuclear import.

Although depletion of TNPO3 has been shown to be inhibitory to HIV-1M infection and to be able to interact with capsid, it is possible that TNPO3 action is indirect and due to changes in CPSF6 localisation. The endogenous function of TNPO3 is to mediate the nuclear import of serine/arginine rich proteins (SR proteins), via interactions with their RS domain [325,326], and which often play a role in the splicing of pre-mRNAs [327]. One such protein is CPSF6 which is a component of mammalian cleavage factor 1, whose role is cleavage/polyadenylation of mRNAs [312]. The CTD of CPSF6 was found to contain a RS domain and has been shown to be transported into the nucleus by TNPO3. Furthermore, it is known that depletion of TNPO3 leads to a redistribution of CPSF6 from the nucleus to the cytoplasm [328]. Once cytoplasmic, either by TNPO3 depletion or deletion of the CTD RS domain [314], CPSF6 is highly restrictive to HIV-1M infection [284,314] [see section 1.4.2]. Therefore is likely the inhibitory effect of depleting TNPO3 on HIV-1M infection is mediated by redistribution of CPSF6.

1.4.5 IP6

IP6 is widely found in cells at a high concentration between 10-100uM and is one of the most abundant inositol derivatives [329]. The reported physiological roles of IP6 are wide-ranging and seemingly unconnected. So far, it has been reported that IP6 is able to act as a neurotransmitter [330,331], activator of protein kinase C [332], inhibitor for protein phosphatase activity [333], iron transporter [334], cellular antioxidant [335], activator for enzymes conducting DNA repair [336], a regulator for mRNA nuclear export [337–339] and recently a cofactor for HIV-1M infection [239,290].

Along with the stabilising effect IP6 exerts on the mature capsid core (described in section 1.3.4), it has also been reported that IP6 plays a role in particle formation [290]. Interestingly, it was shown that IP5 and IP6 play a role in particle assembly of HIV-1M *in vitro* as early as 2001 [340] but it wasn't until recently, with the use of mutagenesis and high resolution crystallography, that the binding site for IP6 was determined. It is now known that IP6 directly binds to both the immature Gag [290], via the capsid region, and to the mature capsid hexamer [239]. Within the HIV-1M immature Gag, the capsid residues Lys 158 and Lys 227 were shown to form charge-charge interactions with IP6 [290]. Whereas in the mature capsid hexamer it was found that IP6 binds Arg 18 [239]. The interaction between IP6 and the immature Gag is thought to promote the formation of the immature Gag lattice neutralizing the repulsive positive charged Lys found at the centre of the six helix bundle. Consistent with this hypothesis is the observation that mutagenesis of the two Lys results in a large defect in particle formation [290]. During particle maturation, the viral protease cleaves the Gag polyprotein which firstly leads to the disruption of the six helix bundle and, secondly, the unmasking of the Arg 18 binding site in the mature capsid. Once unmasked, IP6 binds to Arg 18, again neutralising the charge repulsion created by the clustering of the six positively charged Arg at the six fold axis of symmetry. This charge neutralisation is thought to promote mature capsid lattice formation, improve mature capsid stability and allowing regulation of capsid stability.

One hypothesis that is gathering evidence is that the mature capsid core is an allosteric molecule and that the binding of cofactors results in conformational changes which help regulate capsid function. IP6 causes a strong stabilising effect on the capsid core [see section 1.3.4] however, capsid uncoating must occur for infection to take place. Therefore, IP6 is may be released at some point in the life cycle to promote uncoating. Other cofactors, such as CPSF6 or CypA, could function to alter IP6 binding, either

promoting or preventing, and thus ultimately regulate capsid uncoating. Currently little is known about cofactor-cofactor regulation.

1.5 Study Aims

It is known that HIV-1M infection is influenced by several host proteins which are able to recognise and bind directly to the capsid core. Those host proteins which have a beneficial effect on infection are termed cofactors whereas those with a negative effect are termed restriction factors (see sections 1.3.4 and 1.4). The vast majority of the work investigating the relationship between host proteins and lentivirus infection has been conducted using HIV-1M. Due to this, a lot is known about the interactions which influence HIV-1M infection but less is known about how and if these interactions influence the infection of other *lentiviruses*. It is this gap in the knowledge base that we aim to address.

This study aims to build on the hypothesis that the *lentiviral* capsid core remains intact after entry into the target cell and subsequently plays a role in influencing post-entry steps of the lifecycle via interactions with host proteins. We investigated whether host proteins and compounds known to influence HIV-1M infection are also able to influence infection with other *lentiviruses*. Primarily, this study will focus on the host factors IP6, CypA and TREX-1 and will attempt to answer the following questions. 1) Is the recently identified IP6 a general cofactor for the *lentivirus* lineage or a specific cofactor only for HIV-1M? As IP6 is required for optimal infection with HIV-1M in both the target cell and producer cells, we will investigate both stage of the lifecycle with other lentiviruses. 2) Is CypA an essential cofactor for other lentiviruses or is it specifically HIV-1M which has evolved to use CypA for optimal infection? Furthermore, we investigated whether the recently identified electrostatic channel, focusing on the NTD

beta-hairpin and the Arg18 ring, plays a role in regulating CypA usage in HIV-1M. 3) Does TREX-1 expression levels explain why HIV-1M has differential CypA dependencies in different cells? All three questions will be examined using a comparative virology methodology with viruses from across the *lentivirus* lineage.

2 Methods and Materials

2.1 Plasmid propagation

2.1.1 Agar plates

Agar was made by dissolving 14.8 g of agar powder (Millers) in 400 ml of dH_2O. Liquid agar was then sterilised by being autoclaved before being stored at 4 degrees. Antibiotic selection plates were made by melting sterilised agar and allowing to cool at room temperature before addition of appropriate antibiotic, either ampicillin at 100 ug/ml or kanamycin at 30 ug/ml. Once the antibiotic was added, agar was poured into 10 cm petri dishes and allowed to cool and solidify at room temperature. All agar plates were stored at 4 degrees.

2.1.2 Productions of competent HB101 cells for transformation

Glycerol stocks of commercially brought HB101 bacteria were used to inoculate 50 ml of liquid lysogeny broth (LB) which had been sterilised by autoclaving and that contains tryptone10 g/L, yeast extract 5 g/L, NaCl. The inoculated LB was then cultured overnight at 37 degrees. After 18 hours of culture, 5 ml was used to inoculate 200 ml of fresh LB, which was cultured at 30 degrees until an optical density (OD) of between 0.5-0.6 was reached. OD was examined using WPA UV1101 spectrophotometer (Biotech Jencons). Once the correct OD was reached the culture was incubated on ice for 10 minutes before being centrifuged at 3000 RPM for 10 minutes in a centrifuge cooled to 4 degrees. Supernatant was removed and bacterial pellet was resuspended in in 20 ml of TFB1 buffer (Table 1) before being centrifuged at 3000 RPM for 10 minutes at 4 degrees. Supernatant was removed and the pellet was resuspended in 2 ml of TBF2 buffer (Table 2) and incubated on ice for 10 minutes. Competent bacteria were aliquoted in 100 ul and snap

frozen with dry ice before storage at -80 degrees. Competency was checked by transforming 45 ul of bacteria with 10 pg of PUC19, 100+ colonies on selection plates indicates a high level of competency.

Table 1: TFB1 buffer components

KAC	30mM
RbCl	100mM
CaCl2	10mM
MnCl2	50mM
Glycerol	15% in dH_2O

All Chemicals were made up in dH_2O
TBF1 buffer is filter sterilised with a 0.22µM filter

Table 2: TFB2 buffer components

10mm PIPES	1ml of 0.1M stock
75mm CaCl2	7.5ml of 0. 1M stock
10mm RbCl	0.1ml of 1M stock
15% Glycerol	1.5ml of 100% stock

All Chemicals were made up in dH_2O
TBF2 buffer is filter sterilised with a 0.22µM filter

2.1.3 Bacterial transformation and plasmid miniprep

45 µl of competent bacteria were thawed on ice before 100 pg-1ug of plasmid DNA was added. Competent bacteria were then incubated on ice for 30 minutes before undergoing a 45 second heat shock at 42 degrees in a water bath. Following the heat shock, bacteria were incubated on ice for 5 minutes. For plasmids which encode ampicillin resistance, bacteria were plated on Agar selection plates containing 100 ug/ml ampicillin (Calbiochem). For plasmids which encoded kanamycin resistance bacteria were added to 250 µl of S.O.C (super outgrowth media, Invitrogen) and grown for 1 hour at 37 degrees on a shaking incubator before being plated on to agar plates containing 30 ug/ml kanamycin (Sigma). Once plated bacteria were grown for 18 hours at 37 degrees.

If transformed bacteria were being used in a miniprep, after 18 hours of incubation single colonies were picked and used to inoculate 5 ml of LB, containing the appropriate selection antibiotic, which was then cultured at 37 degrees for 18 hours on a shaking incubator. Once colonies were picked, the agar plates were stored at 4 degrees. Bacteria were pelleted via centrifugation at 3000 RPM for 8 minutes and the supernatant discarded. Plasmid DNA was then extracted using the QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions. Plasmid DNA concentration were measured using a Nanodrop NDW1000 (Thermo Fischer Scientific).

2.2 Western Blotting

2.2.1 Protein extraction

Adherent cells which were going to be used for protein extraction first had their growth media removed and discarded followed by a wash with 1 ml of PBS. Cells were detached from culture plate with 100 µl of Trypsin 0.25% +EDTA (Gibco Invitrogen) for a 12 well plate and 200 µl for a 6 well plate. Cells were counted using a haemocytometer and balanced for cell number before being transferred to an Eppendorf tube where they were pelleted at 13000 RPM for 3 minutes. Supernatant was removed and cells were resuspended in 100 µl PBS and 1x Laemmli Buffer (4% SDS, 20% glyercol, 0.004% bromophenol blue and 0.125M Tris). Samples were then boiled for 10 minutes and either stored at -20 degrees or incubated at 4 degrees for 10 minutes before being loaded onto a polyacrylamide gel for western blot.

2.2.2 Sodium dodecyl-sulphate polyacrylamide gels

Sodium dodecyl-sulphate (SDS) polyacrylamide gels were prepared the night before usage and stored at 4 degrees wrapped in wet tissue to prevent the gels from drying out.

Running gels were made up according to Table 3 whereas stacking gels were made up according to Table 4. Both running and stacking gels were thoroughly mixed after Tetramethylethylenediamine (TEMED) and ammonium persulfate (APS) were added and poured quickly after. Once the running gel was poured it was overlaid with 100% isopropanol to allow the gel to solidify without drying out the top. Once solidified the isopropanol was poured off and the top of the gel was washed with dH_2O before the stacking gel was poured and a 10- well comb inserted.

Table 3: SDS polyacrylamide running gels

Component	15% (10ml)	10% (20ml)
Water	2.3 ml	7.9 ml
30% acrylamide	5 ml	6.7ml
1.5M TRIS-HCl (pH 8.8)	2.5ml	5ml
10% SDS	100µl	200µl
10% APS	100µl	200µl
Temed	4µl	12µl

Table 4: SDS polyacrylamide stacking gels

Component	15% (10ml)
Water	3.4 ml
30% acrylamide	830µl
1M TRIS-HCl (pH 6.8)	630µl
10% SDS	50 µl
10% APS	50µl
Temed	5µl

2.2.3 Polyacrylamide gel electrophoresis and protein transfer

Once the polyacrylamide gel was set, the comb was removed and they were loaded into a Mini Protean (bioRad) cassette and submerged into 1x running buffer. Running buffer

was made up to 10x and contained 30 g of TRIS, 14 g of Glycine, 10 g SDS and made up to 1 litre with dH_2O before being diluted to 1x with dH_2O. 20 µl of sample was loaded into each well and ran alongside 5 µl of PageRuler plus Prestained Protein Ladder (Fermentas). Samples were chilled to 4 degrees before loading to avoid heat damage to the wells. Gels were first ran at 80 V for 15-20 minutes before the voltage was increased to 130 V and run for approximately 1.3 hours.

Once the gel was run it was removed from the glass plates and socked in 1x Trans-blot turbo buffer (BioRad) for 5 minutes alongside Pre-cut nitrocellulose membranes and buffer reservoirs (both from BioRad). The Trans-Blot Turbo buffer was made to 1 L so it contained 200 ml Trans-Blot Turbo buffer, 600 ml of dH_2O and 200 ml of 100% molecular grade ethanol. After 5 minutes in buffer, a transfer stack was assembled within a Trans-Blot Turbo transfer system (BioRad). Transfer was carried out using the pre-programmed mixed molecular weight settings, 7 minutes at 25 V.

After transfer completion, the transfer stack was disassembled and the nitrocellulose membrane transferred to a 2-5% milk solution made up in PBS-Tween (PBS-T) for 1 hour at room temperature. PBS-T contained 0.1% Tween in PBS. This is to block any non-specific antibody binding to the membrane during antibody probing. To prevent drying of the membrane, this blocking step was carried out in a 50 ml falcon tube on a rotating mixer.

2.2.4 Antibody probing

After the membrane was blocked with milk solution, it underwent 3 washes with 5 ml of PBS-T for 5 minutes before being transferred to a fresh 2-5% milk solution containing appropriate antibody at appropriate dilution. All antibody dilutions and species can be found in Table 5. Once in the antibody solution, the membrane was incubated at 4 degrees

overnight on a rolling platform. After incubation with primary antibody, the membrane underwent 3 washes with 5 ml of PBS-T for 5 minutes each. Membrane was then incubated with a complementary secondary antibody diluted in 2-5% milk solution for 1 hour at room temperature. Secondary antibody were conjugated to a fluorescent tag which allowed visualisation via LiCor machine (Li-Cor). Incubation with secondary antibody was carried out in the dark due to the light sensitivity of the fluorescent tag. Again, after incubation with secondary antibody, the membrane was washed 3 times with 5 ml of PBS-T for 5 minutes on a rocking platform. If the membrane was being used to blot for multiple proteins, then after development of 1st secondary, the membrane was washed and placed back into a fresh milk solution with the new primary and probed as described above. Quantification of western blot band intensity was done using ImageJ.

Table 5: Western blot antibody list

Antibody type	Target protein	Dilution	Species	Supplier
Primary	Capsid	1:2000	Mouse	CA183, AIDS reagents
	TREX-1	1:250	Rabbit	sc-271870, Santa Cruz
	CypA	1:5000	Rabbit	
	VCP	1:1000	Rabbit	sc-271870, Santa Cruz
	Tubulin	1:10000	Mouse	
Secondary (Licor)	Anti-mouse Anti-rabbit 680	1:15000	Various	Licor
	Anti-Mouse Anti-Rabbit 700	1:10000	Various	Licor

2.3 Lentiviral vector production and infection

2.3.1 Lentiviral vector preparation

HEK293T cells which were between 70-80% confluent were split 1:4 into 10cm^2 dishes and pen-strep free DMEM. For vector which used a three plasmid system e.g. 8.91 or FIV, a transfection mixture was made containing 1 µg packaging construct, 1 µg envelope construct and 1.5 µg of genome with 10.5 µl of Fugene 6 (Promega) made up to 250 µl with OptiMEM (Promega). The packaging construct contain the viral structural proteins such as Gag but lack packaging signals and so are only expressed in the producer cell. This is the same with the envelope construct which was always pMDG which encodes VSV-G. The genome construct contains the GFP or in the case of p-SIREN a ShRNA which is packaged into the vector, integrates into and is expressed in the target cell. All vectors used are single round meaning they can only perform 1 round of replication, with only the genes present in the genome construct being expressed in the target cells. Once the appropriate transfection solution is made up it was incubated at room temperature for 25 minutes and mixed vigorously before being added to the HEK293T culture media in a dropwise fashion. All transfection mixture and plasmid sets can be found in Table 6. The supernatants from the transfected cells was harvested at 48hr and 72hr post-transfection, filtered with a 0.48 µM filter and stored at -80 degrees.

For the production of mixed core hexamer vectors, HEK293T were transfected with a ratio of packaging constructs with the total amount of packaging construct equalling 1 µg of plasmid DNA. The ratio used were between 6:0 and 0:6 and we have assumed equal mixing and expression. All other plasmid quantities and techniques were kept the same as above.

Table 6: Vector plasmid composition

Vector	Packaging plasmid	Genome	Envelope
HIV-1M	P8.91	pCSGW	pMDG
HIV-1O (MVP)	P8.91-MVP Gag chimera (Not1-Bcl1)	pCSGW	pMDG
HIV-1O RBF)	P8.91-RBF206 Gag chimera (Not1-Bcl1)	pCSGW	pMDG
SIVCpzMT	P8.91-SIVCpzMT145 Gag chimera (Not1-Bcl1)	pCSGW	pMDG
SIVCpzBF	P8.91-SIVCpzBF1167 Gag chimera (Not1-Bcl1)	pCSGW	pMDG
SIVCpzMB	P8.91-SIVCpzMB897 Gag chimera (Not1-Bcl1)	pCSGW	pMDG
HIV-2	HIV-2 Pack	HIV-2 GFP	pMDG
FIV	FP93	GinSin	pMDG
EIAV	Pony3.2	Pony8GRST	pMDG
EIAV-RELIK	Pony3.2 with RELIK Capsid cloned in place of EIAV capsid	Pony8GRST	pMDG
shRNA	CMVi	pSIREN RetroQ	pMDG
Transgene	CMVi	pEXN	pMDG

2.3.2 Lentiviral vector titration

Appropriate adherent cells were detached from culture plate, 10 µl was then staining with trypan Blue (Sigma) and counted using a hemocytometry. For a 12 well plate $2.5x10^4$ cells ml^{-1} were seeded, for a 6 well plate $1x10^5$ cells ml^{-1} was used. Cell were seeded in full media. 24 hours after seeding the cells underwent a media change into full media containing polybrene at 1:1000 to improve infection rate by neutralising the charge on the cell membranes. Cells were then transduced with serially diluted lentiviral vector at 1:3 increments. Vector was diluted in DMEM containing polybrene. Unless stated otherwise, all titrations contained at least 6 virus doses. 48 hours post-transduction cells were trypsinised to detach them from the culture plate before being fixed in 4% (vol/vol) of paraformaldehyde in PBS and incubated for at least 10 minutes at room temperature before GFP expression was measured via flow cytometry on a BD Accuri C6 (BD Biosciences). A negative control, cells which are GFP negative, was always used in

parallel to aid in gating strategy for flow cytometry (section 2.3.3). All viruses are titrated onto appropriate cell lines to determine titre of the virus. Titre was determined by averaging infection units per ml at three doses ≤ MOI: 0.3. The calculation to determine infections units per ml can be seen below. MOI 0.3 was used as this represents the linear section of the titration curve where one infection event is likely caused by a single virus particle. Once a titre has been determined it is normalised by ng of RT allowing for titre to be expressed as infectious units per ng of RT (I.U./ngRT).

$$Infectious\ units\ per\ ml = \frac{(proportion\ of\ infected\ cells * cell\ number)}{Vol.of\ vector\ added}$$

2.3.3 Viral vector concentration by ultracentrifugation

After production of viral vectors via transfection of HEK293T cells with appropriate plasmids [**Table 6**] according to section 2.2.1, supernatants were aliquoted into ultracentrifuge tubes. Each ultra-centrifuge tube was first cleaned with 70% ethanol and two subsequent washes with sterile PBS before ~25 ml of viral vector supernatant was added which is under-layered by 5 ml of 20% sucrose. Ultra-centrifuge tubes were balanced to be within 0.01 g of each other before being placed into ultra-centrifuge buckets which have been cooled to 4 degrees. Before, loading the ultra-centrifuge, it was cooled, along with the rotor (SureSpin 630), to 4 degrees. Once the ultra-centrifuge was cooled, the buckers were placed into the rotor and spun at 23000 rpm for 2 hours at 4 degrees. The centrifuge deceleration was set to 5 to prevent vector pellet becoming dislodged at the end of the spin. Once completed the buckets were removed and their contents poured out in a biosafety hood. Excess media was allowed to drain out before vector pellet was resuspended in 100 µl of PBS before being frozen at -80 degrees.

For vector concentration with a bench top centrifuge, 1 ml of viral vector supernatant was added to an Eppendorf and spun at 4 degrees for 90 minutes at max speed

in a centrifuge pre-cooled to 4 degrees. Once the spin was complete, supernatant was aspirated and pellet was resuspended in 10 µl of PBS.

2.3.4 Quantification of Reverse Transcriptase Activity by qPCR

Quantification of Units reverse transcriptase was used to normalise viral titres. This was done using the protocol developed in Pizzato et al 2009 [341]. Viral vectors were lysed by mixing with 2x lysis buffer at a 1:1 ratio with a total volume of 10 µl. Lysis buffer components can be found in Table 7. Fresh Ribolock RNase inhibitor (Fermentas) was added before lysis of vector so a final concentration of 0.4 U/µl was achieved. Vector and lysis buffer were incubated for 10 minutes at room temperature before the lysis buffer was inactivated by the addition of 90 µl of water. All following steps were performed on ice to maintain reverse transcriptase stability. A master mix was made up according to Table 8 with enough made to run each sample in triplicate with 8 standards. From this master mix, 7.5 µl was aliquoted into each well on a 384 well plate before the addition of 5 µl of sample or standard. A negative control of qPCR grade water was included. The standards were made up using recombinant HIV reverse transcriptase (Millipore) and started at $1x10^8$ pU/µl and went to $1x10^2$ pU/µl. The recombinant HIV reverse transcriptase was resuspended in 10mM potassium phosphate, 1mM DTT and 20% glycerol at pH 7.4 before being stored at -80. When being used for standards, the HIV reverse transcriptase was diluted in qPCR grade water. The MS2 RNA was brought from Roche and stored at -80 degrees. Once all reaction mixtures were added, an optical cover was used to seal the plate before undergoing centrifugation for 2 minutes at 1200 RPM. The plate was then ran in a QuantStudio 5 - 384-Well Block (Thermofisher). The PCR programme used can be found in Table 9.

Table 7: Lysis buffer components

Chemical	Final concentration
TrisHCl pH 7.4	100mM
KCl	500mM
Triton X-100	0.25%
Glycerol	40%

Stored at -20 degrees

Table 8: SG-PERT reaction mixture

Component	Concentration of stock	Volume per reaction (µl)
Quantitect SYBR Green Master Mix	2X	6.25
Primer MS2 FW	100 µM	0.0625 (0.5 µM)
Primer MS2 RV	100 µM	0.0625 (0.5 µM)
MS2 RNA	700 pmol/ml	0.0625 (3.5 pmol/ml)
Ribolock RNAse inhibitor	4U/µl (Dilute the stock 10X in tissue culture PBS – fresh before adding to mastermix)	0.0625 (final 0.02U/µl) lower otherwise it would inhibit the reaction
Standards/Virus lysate (added in next step)	-	5
H20 (for SG-PERT)		1
Total Reaction Volume	-	12.5

Reaction volumes per well on a 384 well plate

Table 9: SG-PERT PCR programme

Step	Time	Temperature (°c)
Reverse transcription	20 min	42
Taq initial heat activation	15 min	95
3-step cycling:		
Denaturation	10 sec	95
Annealing	30 sec	60
Extension	15 sec	72
Number of cycles	40 cycles	

Florescence acquisition is done at the end of each elongation phase

2.3.5 Flow cytometry

Infection of GFP containing viruses was quantified on a BD Accuri C6 (BD Biosciences). First live cell were gated for using side scatter height (SSC-H) and front scatter height (FSC-H) which are a measure of internal complexity e.g. granularity and size respectively. The live cell gate were also used to gate out doublets e.g. event with a very high FSC-H and SSC-H. The alive population of cells was examined for GFP positivity using the FLH-1 laser which measure wavelengths of 515-545nm. A negative control sample, one which contains no GFP, was used to determine the threshold for GFP positivity. Once both gates were set, 10000 live cells were examined for GFP expression for each sample and expressed as a percentage of all cells examined.

2.3.6 Infection at set MOI

After the vector has been titrated on appropriate cells and a titre determined, it is possible to normalise experiments to infection. The volume required for a set MOI is calculated using the equation below.

$$Vol.(\mu l) for\ set\ MOI = \left(\frac{Desired\ MOI * cell\ number\ in\ well}{I.U.ml\ of\ vector}\right) * 1000$$

Once a volume has been determined, appropriate cells are seeded according to section 2.3.2. 24 hours after seeding, the cells are aspirated and washed with PBS before the addition of fresh media containing Polybrene at 1:1000. The volume of vector calculated for a set MOI is then added directly into the media before the plate is gently agitated by rocking to ensure mixing of vector and media. Flow cytometry is then used to measure levels of infection 72 hours after transduction with vector, as described in section 2.3.3.

2.3.7 Measuring late reverse transcription products

Twelve well plates were seeded with $2.5x10^4$ cells ml^{-1} U87 cells and transduced with DNAse (sigma DN25) treated lentiviral vector, in the presences of Polybrene, at a dose that will give an infection level of 10% or 30%. Six hours post transduction the cells are harvested and underwent DNA extraction using the Qiagen DNA blood mini kit. For reaction mixture see Table 10 and qPCR cycling parameters Table 11. A standard curve was generated using the same GFP primers and probes on CSGW between (1x10^9 and 1x10^1 copies/µl). Quantification of GFP via qPCR was using the following primers and probes GT139-F: 5'-CAACAGCCACAACGTCTATATCAT-3', GT140-R: 5'-ATGTTGTGGCGGATCTTGAAG-3', GFP-Probe: 5'-FAM-CCGACAAGCAGAAGAACGGCATCAA-TAMRA-3'. Parallel to this a vector titration was set up according to section 2.3 and ran alongside allowing the measurement of infection.

Table 10: qPCR reaction mixture for late RT products

Component/Stock Concentration	Volume (µl)
2x TaqMan Gene Expression Master Mix (ThermoFisher)	10
FAM/TAMRA Probe 7.5µM	0.5
Primer GT139 7.5µM	1
Primer GT140 7.5µM	1
dH_2O	2.5
Master Mix Total	15
DNA sample (20-150ng/µl) or plasmid standard (10^0-10^7 copies/µl)	5
Total	20

Table 11: qPCR programme for measuring RT products

Step	Temp (^oC)	Time	No. of Cycles
Polymerase Activation	95	10 mins	1
Denaturing	95	15 secs	40
Annealing/Extension	60	60 secs	

2.4 Cell culture

2.4.1 Adherent cell lines

HEK293T, U87 and CRFK cells were grown in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% foetal calf serum (FCS; Gibco) and penicillin-streptomycin at a final concentration of 50 µg/ml (Gibco). All cultured cells were incubated at 37 degrees and 5% CO2, except HEK293T cells, which were maintained at 10% CO2. All cells were routinely passaged to prevent over confluency. For adherent cell lines such as HEK293T, U87, CRFK and TE671 cells were maintained in 10 cm dishes. For passage of adherent cells, supernatant was removed and discarded before 5 ml of PBS was used to wash the cells followed by the addition of 2 ml of Trypsin 0.25% +EDTA (Gibco Invitrogen) and a 2-5 minute incubation at 37 degrees. Once cells were detached they were removed from culture plate and diluted into fresh a 10 cm plate containing fresh media.

2.4.2 Storage of cell lines

A confluent 10 cm plate was trypsinised with Trypsin 0.25% +EDTA (Gibco Invitrogen) and the cells pelleted via centrifugation at 1500 RPM for 5 minutes. The supernatant was discarded and the pellet resuspended in FCS containing 10% DMSO (SIGMA). 0.5 x 10^6 Cells were then aliquoted into cryovials in a total volume of ≥500 µl. Cryovials were then incubated at -80 degrees for 18-42 hours in a Mr Frosty (Thermo Scientific) which allowed for slow freezing. Once frozen, samples were transferred to liquid nitrogen for long term storage. When cells were rescued from liquid nitrogen they were thawed rapidly in a water bath at 37 degrees. Once thawed, cells were transferred to 10 ml of full media in a 10cm^2 dish. 24 hours later cells underwent a PBS wash and media change to remove the remaining DMSO and then cultured as normal.

2.5 Cloning and DNA Ligation

2.5.1 Site directed mutagenesis

For details of site directed mutagenesis (SDM) reaction mixture see Table 12. The reaction then underwent PCR using the following parameters initial heating at 92 degrees for 2 minutes before undergoing 12 cycles of 92 degrees of 1 minutes, 1 minute at 55 degrees and 2 minutes per kb of DNA at 68 degrees. Once the cycles were completed there was a final extension at 68 degrees for 35 minutes before being held at 4 degrees. All oligos and primers were purchased from Sigma whereas the dNTPs were from Promega. Sequence for SDM primers can be found in Table 13. After SDM the product underwent DPN1 (New England BioLabs) digestion for 2 hours at 37 degrees before gel extraction via QIAquick Gel Extraction Kit (Qiagen). The final product is then used to transform competent HB101 cells. Confirmation of successful mutagenesis was done via sequencing.

Table 12: SDM reaction mixture

Component	Volume (µl)
Molecular biology dd. Water	35
10x cloned PFU reaction buffer	5
dNTP mix (25mM each dNTP)	1
Template DNA (30ng)	3
Forward primer (10µM)	2
Reverse primer (10µM)	2
DMSO	2.5
PFU Turbo DNA polymerase 2.5U/µl	2

Table 13: SDM primers

SDM	Oligo sequences	Tm (degrees)
8.91 R18G	F:ATATCACCTGGAACTTTAAATGCATGGGTAAAAG R:ATTTAAAGTTCCAGGTGATATGGCCTGATGTAC	64
MVP R18G	F: CATATCCCCCGGGACTTTAAATG R: GCTTGATGTATCATTTGTCC	59
SIVCpzMT R18G	F:CATAATCCAATATCACCCGGAACCCTAAATG R:CATTTAGGGTTCCGGGTGATATTGGATTATG	63
SIVCpzBF R18G	F:GCATCAGCCAATTTCACCCGGAACTCTTAATGC R:GCATTAAGAGTTCCGGGTGAAATTGGCTGATGC	70
HIV-2 R17G	F: GCTGAGTCCCGGAACCCTAAATG R: GGTATATGGGTGTAGTTGC	62
EIAV R18G	F:GTTGGCCCGAGACTGATTGCGGCGTGGCTGAAACAG R:CGCAATCAGACCCGGGCCTATCGGTTCATATTCCTG	73
FIV K17G	F: ACTTGACCCAAGGATGGTGTCCATTTTTATG R: GCTACATATTGTGGTACTC	59
FIV K17R	F:CCACAATATGTAGCACTTGACCCAAGAATGGTGTCCATTTTTATG R:CTTTTCCATAAAAATGGACACCATTCTTGGGTCAAGTGC	69
EIAV RELIK G18R	F:GTTGGCCCGAGACTGATTGCGGCGTGGCTGAAACAG R:CAGCCACGCCGCAATCAGTCTCGGGCCAACCGGTTC	77
8.91 K227A	F:ACCCGGCCATgcaGCAAGAGTTTTG R:CCCCCCACTCCCTGACAT	68
MVP K227A	F:GCCAACTCACgcgGCAAAAATACTAC R:CCTCCTACTCCTTGACAG	59
SIVCpzMT K227A	F:ACCCTCACATgcaGCTAGAGTGCTG R:CCTCCTACCCCTTGACAA	61
SIVCpzBF K227A	F:ACCAGCACATgcaGCAAGGGTCCTAC R:CCCCCTACTCCCTGACAG	64
HIV-2 K227A	F:GCCAGGCCAGgcaGCTAGATTAATG R:CCACCTACCCCCTGACAG	64
EIAV K227A	F:TACAAAACAAgcaATGATGTTACCAATCATG R:GTTCCAATGTCTCTGCAAG	57
FIV K227A	F:ACCAGGATATgcaATGCAACTCC R:GAGCCTATTTCTTGACAAG	58

If primers were not listed for a mutant then it was previously made in the lab
F: Forward
R: Reverse

2.5.2 Restriction enzyme digestion assay

Digestion assays for both analytical purposes and cloning were carried out using the same protocol. Plasmids were digested at 2 µg in a total volume of 20 µl. Reaction mixture contained plasmid DNA, either NEB or Promega restriction enzymes at 5% (v/v), appropriate reaction buffers (provided at 10x by manufacturer) at 10% (v/v) and made up

to final volume with dH_2O. Reaction mixtures were then heated at 37 degrees for between 1 and 2 hours before being run on an agarose gel and used for gel extraction (section 2.5.3).

2.5.3 Agarose gel electrophoresis and gel extraction

1% agarose gels were made by heating 1 g of agarose powder (Sigma) in 100 ml of 1x TAE (Tris acetate EDTA). The TEA was made up in dH_2O. Once the agarose was fully dissolved in the TAE, the solution was allowed to cool before 1% (v/v) ethidium bromide (EtBr, provided by Sigma) was added. After addition of EtBr the gel was poured and allowed to set. Once set, DNA samples were mixed with loading dye (Thermo Scientific, at 6x) and run alongside 5 µl of GeneRuler kb DNA ladder (Thermo Scientific) at 100 volts for 1 hour. Bands were then visualised under UV light a BioDoc IT imaging system (UVP). If sampled needed to undergo gel extraction, digestions were ran in duplicate with one sample never being exposed to UV light to reduce the chances of UV induced DNA damage. DNA band of interest was excised from the gel then the DNA extracted using the QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions.

2.5.4 Sequencing

All mutagenesis products were confirms by sequencing the SDM product with specific primers. All sequencing was performed using the Mix2Seq overnight sequencing kit (Eurofins). Sequencing was checked against previously established plasmid maps and DNA sequence, using the DNADynamo software (Blue Tractor software), to confirm presents of intended mutations and no others.

2.6 Drug assays

2.6.1 Cyclosporine and other Cyclophilin targeting drugs

Transduction is set up according to section 2.3 with the addition of 5 µM cyclosporine (CsA), JW115, JW3-38, THW166, JW51 or equivalent volume of DMSO at the time of transduction. Infection and/or reverse transcription products are measured as above at either 48 hours or 6 hours post transduction respectively. JW115 and JW51 are both CsA analogues whereas JW3-38 and THW166 are a new class of cyclophilin A targeting drugs which have been termed Depsins. All of the drugs used, except CsA, are unpublished and have been designed and synthesized by Justin Warne and David Selwood (medicinal chemistry, UCL).

2.6.2 MTT assay

U87 cells were seeded at 2.5x10^4 cells per ml in a flat bottom 96 well plate. 24 hours after seeding, 5 µM of CsA, JW3-38, JW115, THW166 or JW51 or DMSO were added to media. Puromycin, G148 and 70% ethanol were used as positive control. 48 hours after addition of drug, cell media was aspirated and 50 µl of MTT solution in serum free media and incubated for 3 hours at 37 degrees. After incubation, 150 µl of MTT solvent was added to lysis the cells and incubated at room temperature for one hour before having absorbance read at a wavelength of 590 nm on a plate reader.

2.7 Generating Phylogenetic trees

Sequences were harvested from Genbank and aligned using MUltiple Sequence Comparison by Log- Expectation (Muscle) [342] within the Seaview software [343] to isolate the capsid coding regions. These were then used to generate phylogenetic trees. Trees were estimated by maximum likelihood using the general time reversible (GTR) substitution model and gamma-distributed substitution rate heterogeneity, as

implemented in RAxML 8 [344]. Confidence was assessed with 1000 non-parametric bootstrap replicates. The phylogenetic trees were then coloured using Chroma Clade [345].

2.8 Protein visualisation

All manipulations of protein structure were carried out in Pymol molecular visualiser. These include (1) colouring protein structures, (2) aligning protein structure from different *retroviruses* and (3) construction of capsid hexamer structure from published capsid monomers. Publicly available structures were fetched directly into Pymol software before construction of capsid hexamers. If the lentiviral capsid structure of interest was solved as a monomer, a hexamer was assembled using 5HGL as a scaffold structure. 5HGL was separated into individual monomers but not moved within solved structure. Solved monomer from another *retrovirus* was duplicated and aligned with separated 5HGL monomers so that each monomer of 5HGL was overlaid by that of another lentivirus. The PDB accession codes for all protein structures used in this work can be found in Table 14.

Table 14: Protein structural database codes

Virus	PDB code	Reference
HIV-1M	5HGL	[237]
HIV-1O	Unpublished (David A. Jacques)	N/A
SIVCpzPtt *	Unpublished (David A. Jacques)	N/A
HIV-2 *	2WLV	[346]
SIVMac *	4HTW	[347]
EIAV *	1EIA	[348]
FIV *	5NA2	[349]
RELIK *	2XGU	[302]
HTLV-1	1G03	[350]
BLV	4PH0	[351]
RSV	3TIR	[352]
MLV	3BP9	[353]
HIV-1M hexamer with IP6	6ES8	[239]
HIV-1M immature Gag with IP6	6BHR	[290]

N/A: Not applicable
(*) Construction of hexamer from solved monomers

2.9 Protein depletion and overexpression

2.9.1 Cloning into pSIREN

Where possible published shRNA sequences were used for cloning into pSIREN. If a published ShRNA sequence was not available a published SiRNA sequence was found and converted to ShRNA sequence using the Invitrogen Block-It RNAi server (https://rnaidesigner.thermofisher.com/rnaiexpress/construct.do?pid=3756125999796135117) which inserts CGAA hairpin loop. The top strand ShRNA sequence had the BAMH1 restriction site (GATC) added to the 5' end and the bottom strand had the ECOR1 site (AATT) added to the 5'end. These were then ordered from sigma fisher scientific and stored at -20 degrees until needed.

pSIREN was digested using BamH1 and EcoR1 (both New England Biolabs) for 3 hours at 37 degrees in NeBuffer 2.1 before being run on a agarose gel (Section 2.5.2 and 2.5.3). This digestion was ran in duplicate so during gel extraction on sample would not be exposed to UV and thus reduced the chance of UV induced dimerization of nucleotides. During gel extraction, one sample was removed and used as a guide for the second, non-UV exposed sample. The UV exposed sample was discarded. Digested pSIREN was extracted from agarose gel via the QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions. Dimerised linear oligos containing the shRNA were ligated into the linear pSIREN following the reaction mixture in Table 15. The ligation reaction was incubated at room temperature for 3 hours before being transformed into chemically competent bacteria (Section 2.1.3), confirmation of cloning was done via sequencing (Section 2.5.4). The oligos were linearized and annealed by combining each strand in a 1:1 ratio and then incubating for 30 second at 95 degrees before slowing cooling as follows 2 minutes at 72 degrees, 2 minutes at 37 degrees, 2 minutes at 25 degrees. Annealed oligos were stored at -20 degrees until used.

Table 15: T4 ligation reaction mixture

Linear pSIREN at 25ng/µl	2 µl
dsOligos	1 µl
10X T4 ligase buffer	1.5 µl
BSA	0.5 µl
Nuclease free water	9.5 µl
T4 DNA ligase (400U/µl)	0.5 µl

2.9.2 Transduction of ShRNA or transgenes and selection of transduced population

Both ShRNA and Transgene expression vectors were produced using a three plasmid system described in section 2.3 and the protocols used for transduction (section 2.3.2) and selection (section 2.8.2) were the same. For adherent cells a confluent 10 cm plate was split 1:4 and resuspended in 8 ml of media containing polybrene at 1:1000 before the addition of 1 ml of either pSIREN or pEXN 24 hours later. 24 hours post transduction the cells were washed in PBS and refeed 10 ml of media. After an additional 48 hours, 72 hours post transduction, appropriate selection antibiotic was added to the growth media. For pSIREN puromycin (Merck) was used at 2 µg/µl whereas G418 (Invitrogen) was used for pEXN, at 500 µg/ml. Alongside the transduced cells, a non-transduced control was used to ensure antibiotic selection was working. Selection would continue until the non-transduced cells were completely dead. Once dead, the selection antibiotic concentration was decreased to half and maintained for duration of selected cell usage.

Once selection was completed, cells underwent western blot (section 2.2) to confirm knockdown or overexpression. In all depletion or overexpression experiments, a control vector was used. For pSIREN, this contained a non-targeting shRNA. The pEXN control vector was pEXN with no insert cloned in and thus did not express any protein. As the pEXN construct contains a HA tag, it was possible to western blot for HA to confirm pEXN transduction. Additionally, to confirm the level overexpression or depletion was not lost or altered during passage of the cell, samples were collected for western blot before each experiment to assess expression for the protein of interest.

3 IP6 is a conserved structural cofactor throughout the lentivirus genus

3.1 Chapter 3 introduction

As detailed in section 1.3 there is growing evidence that the capsid of HIV-1M remains intact during the early stages of infection and that it influences post-entry life cycle steps through interactions with host cells proteins and compounds. Inositol hexaphosphate (IP6) has recently been described as a important interaction partner for HIV-1M [239,290]. It has been found that IP6 influences two critical stages in the HIV-1M life cycle, particle production and capsid stability. Using X-ray crystallography and mutagenesis the binding residues for IP6 have been identified. The capsid region within the immature Gag of HIV-1M binds IP6 using Lys 158 and Lys 227 and influence particle production [290], whereas the mature HIV-1M capsid binds IP6 using Arg 18 and influences capsid stability [239]. Of note is position Arg 18 in the mature capsid which has also been shown to be essential for infection via control of reverse transcription [237] and is present at the centre of the hexamer at the six fold axis of symmetry. Briefly, our current model for IP6 usage in the HIV-1M lifecycle is as follows; an infected cell produces immature Gag molecules which bind IP6 via their capsid section. This helps to drive assembly of the immature Gag lattice. Once the particle has budded, Gag undergoes proteolytic cleavage, via the viral protease, which reveals the IP6 binding site in the mature capsid, Arg 18. Once bound at Arg 18, IP6 functions as a stability factor, by neutralising the tightly clustered positively charged Arg 18 side chains, preventing premature uncoating of the capsid core and thus allows reverse transcription to take place within the protected environment of the capsid.

We took a comparative virology approach to investigate whether IP6 is a specific cofactor for HIV-1M or a general one for lentiviruses. We took a genetic approach

comparing several primate lentiviruses including HIV-1M, SIVCpzMT, SIVCpzBF, HIV-1O and HIV-2 alongside more distant non-primate lentiviruses, FIV and EIAV, to examine conservation of IP6 usage via mutagenesis and phylogenetic analysis of the lentiviral capsid.

Results

3.2 Conservation of the IP6 binding site in immature Gag

3.2.1 Lysine 158 and 227 are highly conserved throughout the lentivirus genus

It has been calculated that within HIV-1M the residues which bind IP6 in the immature Gag lattice, Lys 158 and Lys 227, are 98.8% and 99.4% conserved respectively [354]. We wanted to know whether this high level of conservation was unique to HIV-1M or whether it is a general feature of lentiviruses. Taking a phylogenetic approach we sought to determine whether the IP6 binding site within the immature Gag was conserved throughout the lentivirus genus. Capsid sequences which represent each major lineage of lentivirus, primate, feline, equine, bovine and ovine-caprine, were collected from Genbank and aligned. A maximum likelihood phylogenetic tree, using MLV as the root, was constructed before being coloured, using ChromaClade, for Lys 158 and Lys 227 in the HIV-1M capsid [**Figure 7**]. The vast majority of lentiviruses contained a Lys at both positions with the exceptions of the ovine-caprine lineage. The ovine-caprine lineage was found to lack the Lys at position 158 and instead contained a Thr or Asn, both polar but uncharged amino acids. This demonstrates conservation for key IP6 binding residues at the sequence level for the majority of lentiviruses.

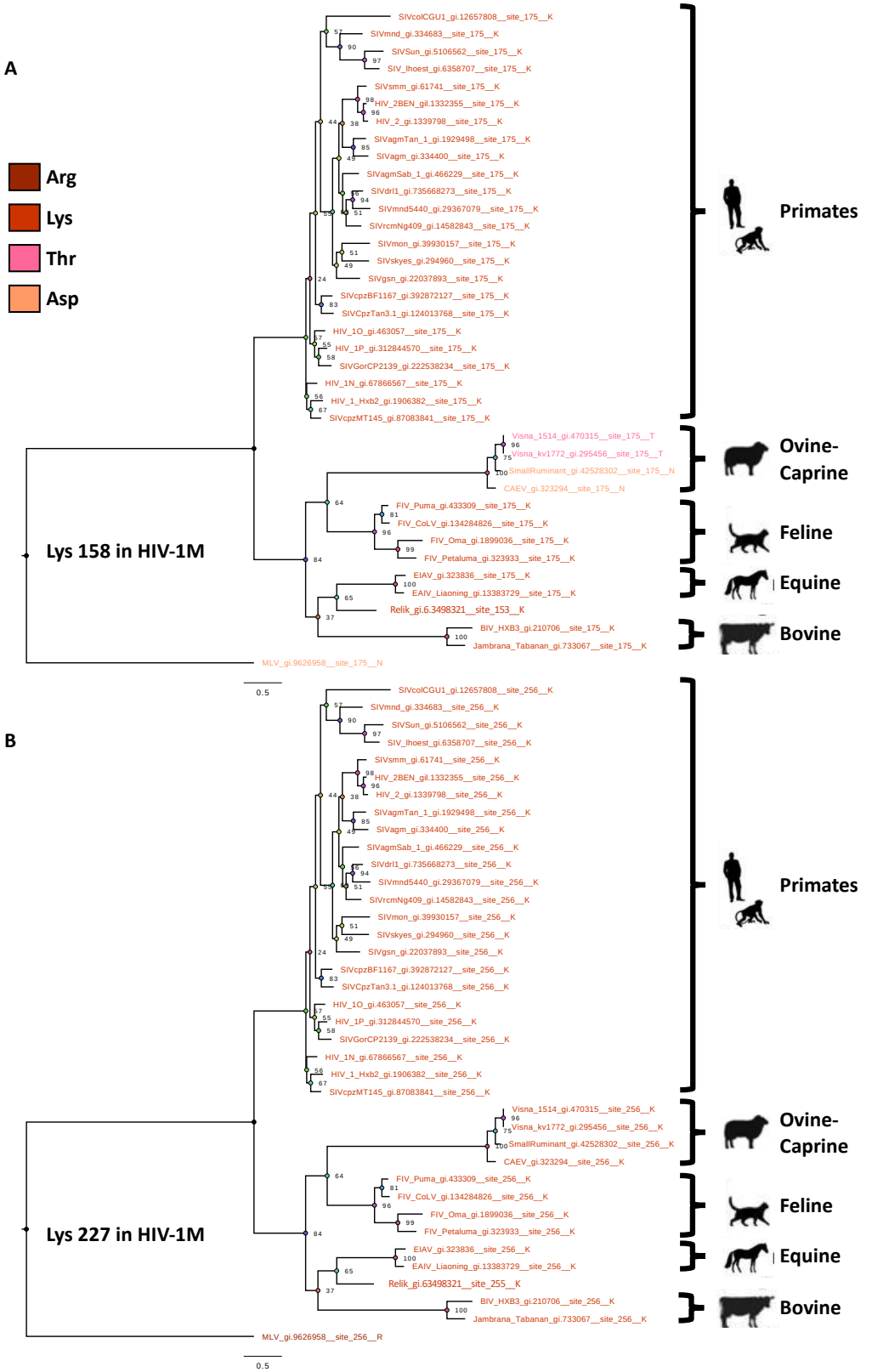


Figure 7: K157 and K227 are highly conserved throughout the lentivirus lineage A capsid phylogenetic tree was estimated by maximum likelihood using the GTR substitution model and gamma-distributed substitution rate heterogeneity. Confidence was assessed with 1000 non-parametric bootstrap replicates and shown as percentage on tree. The phylogenetic trees was then coloured using Chroma Clade according to residue 158 (A) or residue 227 (B) HIV-1M capsid.

3.2.2 Disruption of K227 via mutagenesis is detrimental to all lentiviruses

Having shown the conservation of Lys 227, we next wanted to determine whether it plays a role in the life cycle of other lentiviruses. It was previously demonstrated that mutagenesis of Lys 227 to a Ala in the HIV-1M capsid is highly detrimental to infection [290]. So we wanted to determine the effect of mutagenesis on the other lentiviruses. To investigate this, we chose members of the lentivirus genus allowing us to examine max diversity. Those selected include HIV-1M, both western (SIVCpzMT) and eastern (SIVCpzBF) lineage of SIVCpz, HIV-1O, HIV-2, FIV and EIAV. We mutagenized the homologous Lys residue which aligned with Lys 227 in HIV-1M to an Ala in all the panel members [**Figure 8**]. We found that all lentiviruses harbouring the K227A, or equivalent, mutation were highly defective in U87 cells with a large decrease in titre. The SIVCpz of eastern chimps showed the smallest defect with only a 10 fold decrease in titre [**Figure 8B**]. This decreased defect could be due to the presence of a three amino acid insertion which is harboured in all eastern chimp viruses. However, this is yet untested. All other panel members showed between a 100-1000 fold decrease in titre [**Figure 8B**].

Next we wanted to determine at what stage of the life cycle are the panel members defective. Previously it was shown that capsid mutants K158A and K227A in HIV-1M causes a decrease in particle formation via electron microscopy [290,355]. To determine if the other panel members also have a defect in particle production when harbouring the K227A (or equivalent) capsid mutation we examine reverse transcriptase activity in viral supernatants using the established SG-PERT assay [341]. It was found that all panel members had decreased reverse transcriptase activity in supernatant except for HIV-2, which was not significantly affected [**Figure 9A**]. It should be noted that the size of the defect observed in SG-PERT, which is about 10 fold for most panel members, does not fully explain the defect observed in infectivity [**Figure 8**], suggesting that fewer, less infectious particles are released. As it is known that unclevaged Gag-pol is less active in

reverse transcription we decided to examine Gag cleavage to see if there was a change in levels of uncleaved Gag which could contribute to a defect in SG-PERT. Viral supernatants from the primate lentiviruses were probed by western blot with the broad spectrum capsid antibody CA183, a monoclonal antibody which binds a highly conserved epitope and recognises most primate lentiviral capsids, [**Figure 9B**] each lane was normalised to volume of supernatant added. Quantification of total protein present in each lane, by ImageJ, showed that each K227A mutant had a reduced total protein [**Figure 9B**]. Although the reduction is small for HIV-1M and SIVCpzBF, about ~20%, it is consistent. Furthermore, previous studies have also shown a similar small effect size for HIV-1M K227A by western blot [356]. Both lineages of SIVCpz also show a change in cleavage pattern. Most notably, SIVCpzMT had an overall 50% decrease in Gag levels, mainly in the capsid band, however there was an increase in the matrix-capsid band. Whereas SIVCpzBF had a reduction in the capsid-matrix band. Interestingly the wild type SIVCpzBf showed a significant cleavage defect, with the presence of capsid-nucleoprotein band. This cleavage defect is likely due to the HIV-1M protease poorly recognising the cleavage sites within the HIV-1M-SIVCpzBF chimeric Gag constructs. Together this data suggests that there is a decrease in particle formation with the K227A mutants, as shown by decreased Gag and reverse transcriptase activity present in viral supernatants [**Figure 9**], and that this contributes to the defect observed in infectivity.

HIV-2 however seems to represent a different defect as reverse transcriptase activity was not reduced but there was a decrease in Gag in viral supernatants measured by western blot [**Figure 9**]. Taken with the observation that HIV-2 K226A has a large defect in viral titre relative to wild type HIV-2 [**Figure 8**], this suggests that for HIV-2 the defect is during infection rather than in particle production.

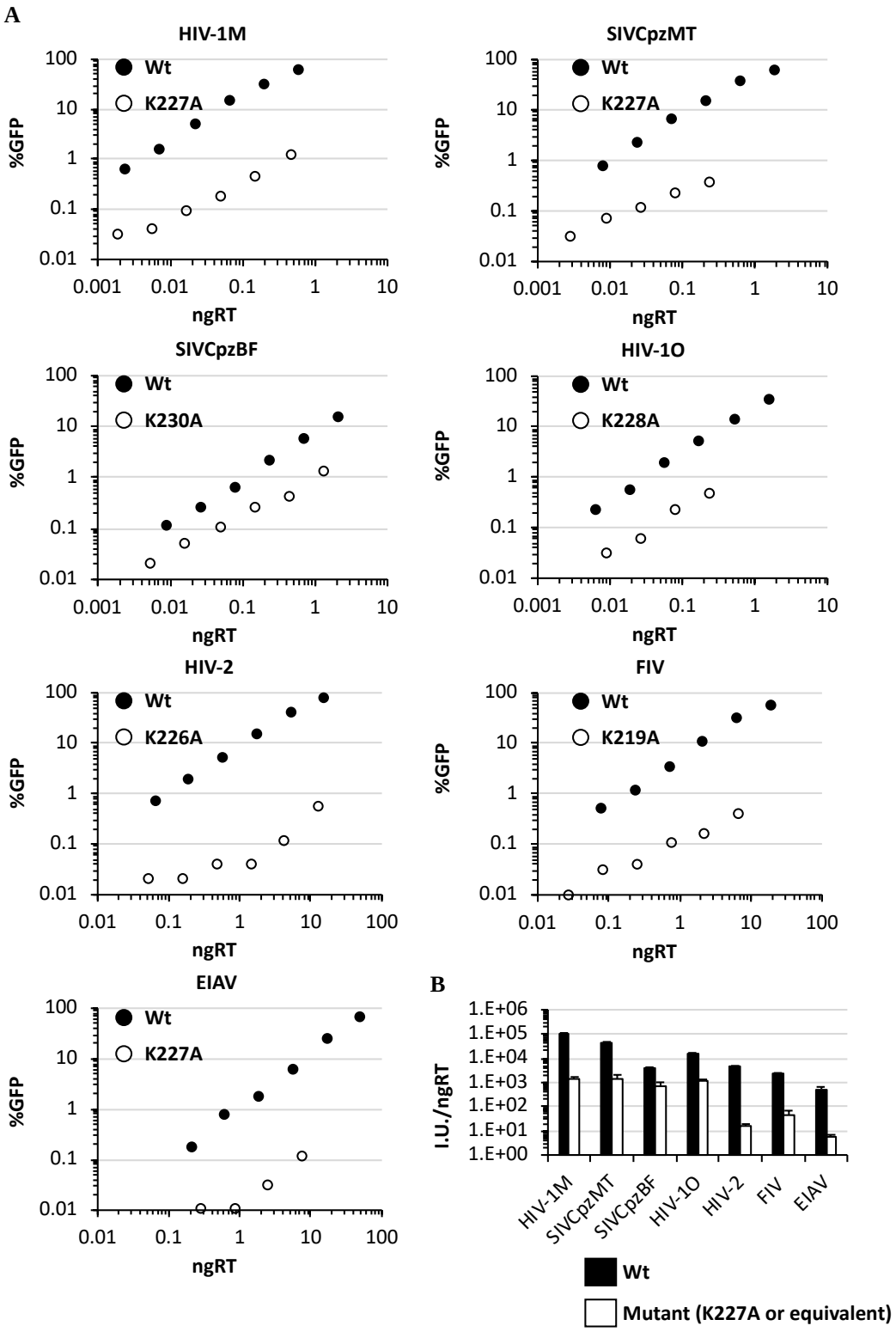


Figure 8: Mutagenesis of K227A is highly detrimental to infection to the primate, feline and equine lineage of lentiviruses

(A) Infectivity plots of VSV-G pseudotyped vectors derived from various lentiviruses either bearing wild type capsid (Wt) or the capsid mutant K227A, or equivalent, (mutant) as shown on U87 cells. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2.

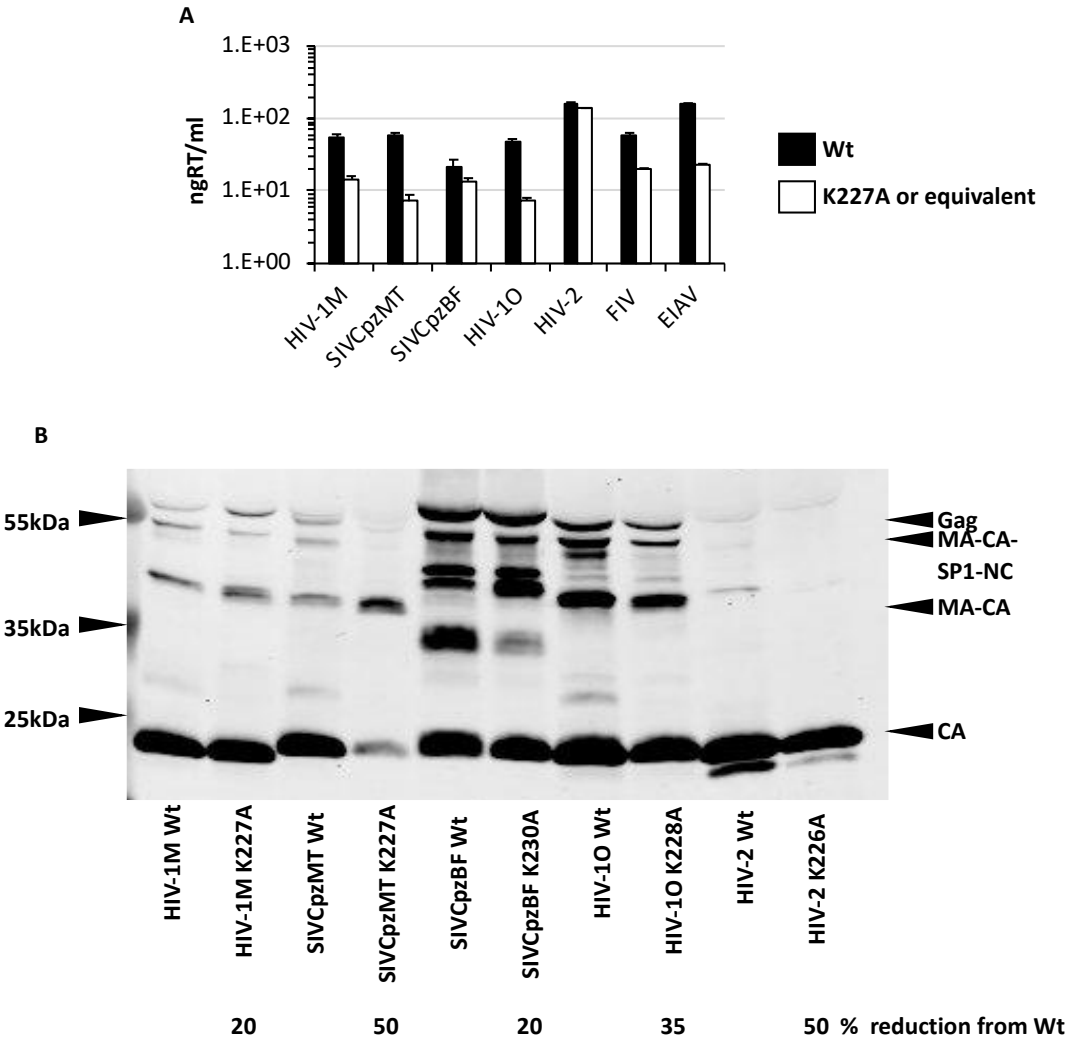


Figure 9: The K227A mutation causes a defect in particle production for some lentiviruses but not others

(A) Viral supernatant used in SG-PERT assay to determine reverse transcriptase activity of lentiviral vectors containing either wildtype (Wt) or K227A, or equivalent, capsid mutation (K227A or equivalent). N=2, +/- SD. (B) Viral supernatant of primate lentiviral vectors containing wildtype of K227A equivalent capsid mutation subjected to western blot detecting Gag.

3.3 Conservation of the IP6 binding site in the mature capsid hexamer

3.3.1 The electrostatic channel at the 6 fold axis of symmetry within the hexamer is highly conserved within the lentivirus lineage

IP6 has also been found to interact with the mature capsid of HIV-1M at Arg 18 [239]. Arg 18 is highly conserved within HIV-1M, calculated at 99.4% [354], and has been shown to play an important role in control of reverse transcription [237]. We sought to determine whether this high level of conservation was a general feature of lentiviruses or unique to

HIV-1M. We next used ChromaClade [345] to colour the phylogenetic tree from **Figure 7** according to which residue was present at position 18 in the HIV-1M capsid [**Figure 10A]**. With the exception of caprine-ovine lineage and Rabbit endogenous lentivirus type K (RELIK), which contained a Val and a Gly respectively at this position, all examined lentiviruses encode a positively charged amino acid at position 18, or equivalent [**Figure 10A]**. Arginine was most common at this position with both the primate and equine lineage exclusively encoding it. However, variation was observed in more distant lentiviruses such as those from the feline and bovine lineage. The bovine lineage exclusively contained Lys whereas the feline lineage contained a mix of Arg and Lys.

As sequence space does not equate to structural position we examined published structures of various lentiviral capsid hexamers to determine the location of the positively charged residue identified in **Figure 10A**. Visualisation of the hexamer structure either by solving hexamer structures by x-ray crystallography using cross linked capsid mutants to stabilise the recombinant hexamers or by solving the monomer structure and reconstructing a hexamer has revealed that each lentivirus examined has a distinctive hexamer morphology. Despite the difference in gross morphology between lentiviral capsid hexamers, the positively charged amino acid highlighted in the phylogenetic tree [**Figure 10A**] always occupied the same location within the hexamer, at the six fold axis of symmetry behind the N-terminal beta hairpin [**Figure 10B**]. The characteristic N-terminal beta hairpin and electropositive ring can be seen in all examined primate and feline lentiviruses hexamer structures showing structural conservation of this feature. The hexamer structure of EIAV shows the electrostatic channel but the NTD beta-hairpin is not resolved in the available structure due to the use of NTD purification tag during crystallisation.

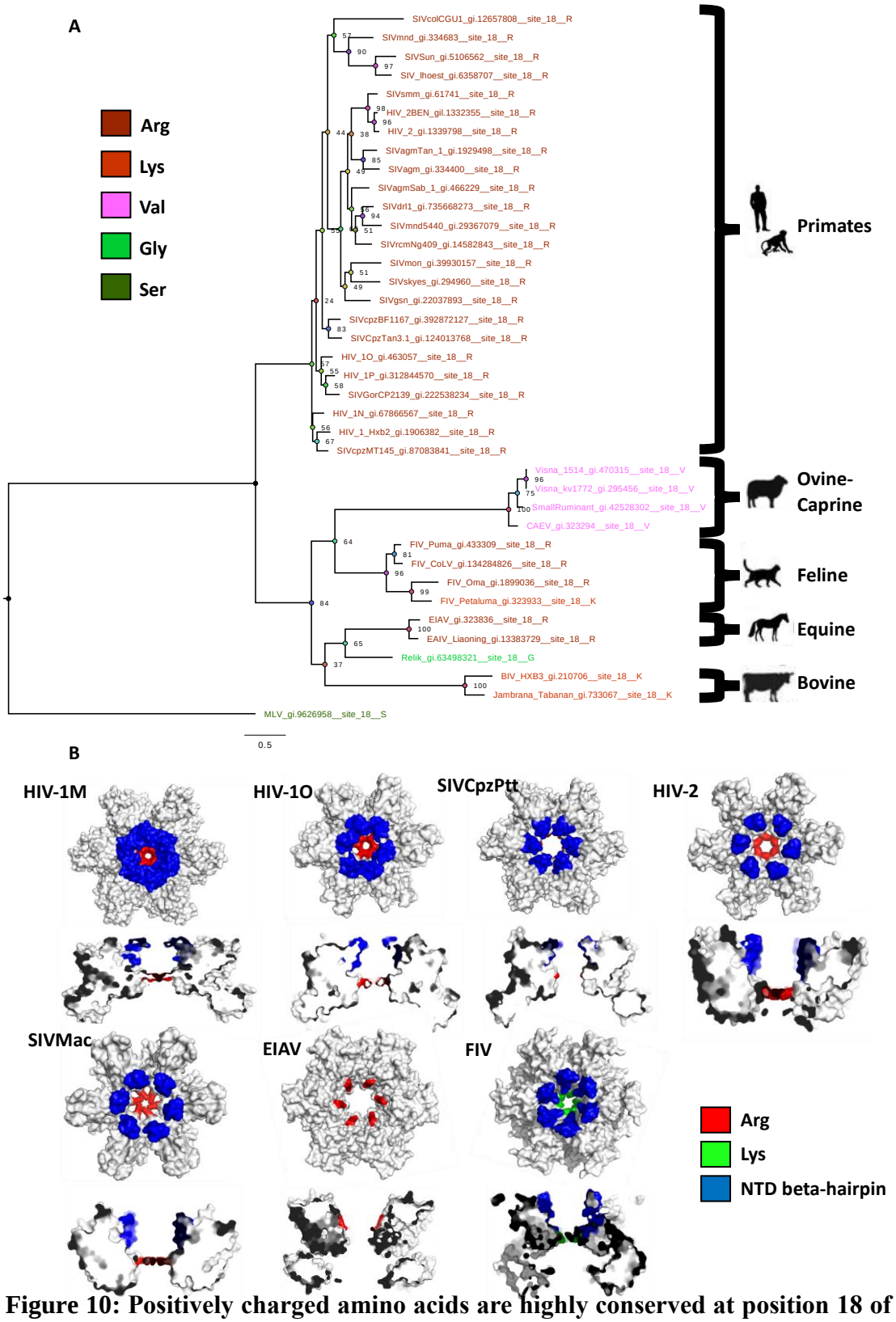


Figure 10: Positively charged amino acids are highly conserved at position 18 of the capsid throughout the lentivirus lineage

(A) A capsid phylogenetic tree was estimated by maximum likelihood using GTR substitution model and gamma-distributed substitution rate heterogeneity. Confidence was assessed with 1000 non-parametric bootstrap replicates and shown as percentage on tree. Tree was coloured, using chroma clade according to residue 18 in the HIV-1M capsid. (B) Published hexamer structures of different lentiviruses. In all examined structures the charged amino acid at position 18 or equivalent is at the six-fold axis of symmetry.

3.3.2 Disruption of the electrostatic channel via mutagenesis is detrimental to all lentiviruses

Next we wanted to address whether the conserved positively charged amino acid at position 18, or equivalent, plays a similar role in the life cycle for other lentiviruses. Taking advantage of previous work showing that as the ratio of Gly to Arg increases at position 18 at the six fold-axis of symmetry within the mixed capsid hexamers, infectivity drops in a step-wise fashion due to reduced reverse transcription [237]. We hypothesised that if we could observe the same relationship with other lentiviruses then we could infer a similar function of the electrostatic channel in their lifecycles. We made a series of mixed capsid hexamers [**Figure 11**] of wild type and R/K18G (or equivalent) for each of the panel members, which were R18G for SIVCpzMT, SIVCpzBF, HIV-1O and EIAV, R17G for HIV-2 and K17G for FIV, and measured infectivity and reverse transcription products [**Figure 12**]. When making the mixed capsid vectors it was assumed that there was equal mixing of wild type and mutant capsids within each hexamer. Consistent with D. Jacques et al [237], as the ratio of R18G increased within the HIV-1M prep the defect in infectivity and reverse transcription increased in a step-wise fashion [**Figure 12**]. We also observed that the closest relative of HIV-1M, the SIVCpz viruses, also demonstrated the step-wise pattern of reduction in titre and reverse transcription products. The non-primate lentiviruses, FIV and EIAV, both showed distinct defects. EIAV only begins to show a large defect at a ratio of 1:5. Whereas FIV shows the smallest defect even with the complete loss of the electrostatic channel, at a ratio of 0:6. EIAV was found to contain several other Arg residues in close proximity to Arg 18, which may play a role in moderating the effect of R18G. The differences in curve shape are likely due to differences to capsid morphology between lentiviruses. Taken together this data suggest a conservation of function of the electrostatic channel, in that it plays a critical role in the life cycle of several lentiviruses via its control on reverse transcription.

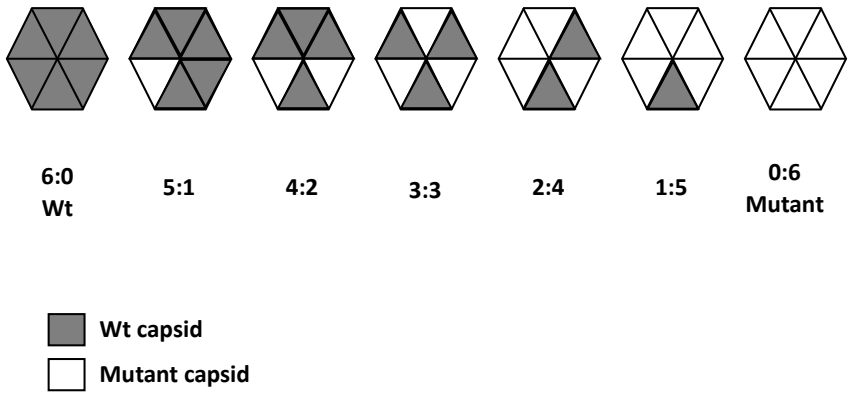


Figure 11: Schematic of mixed capsid hexamers
When producing mixed capsid hexamer particles, total amount of packaging plasmid transfected into 293T cells was kept constant but the ratio between wildtype and mutant containing plasmid was altered to give the shown ratios. It was assumed that there was equal mixing of mutant and wildtype construct.

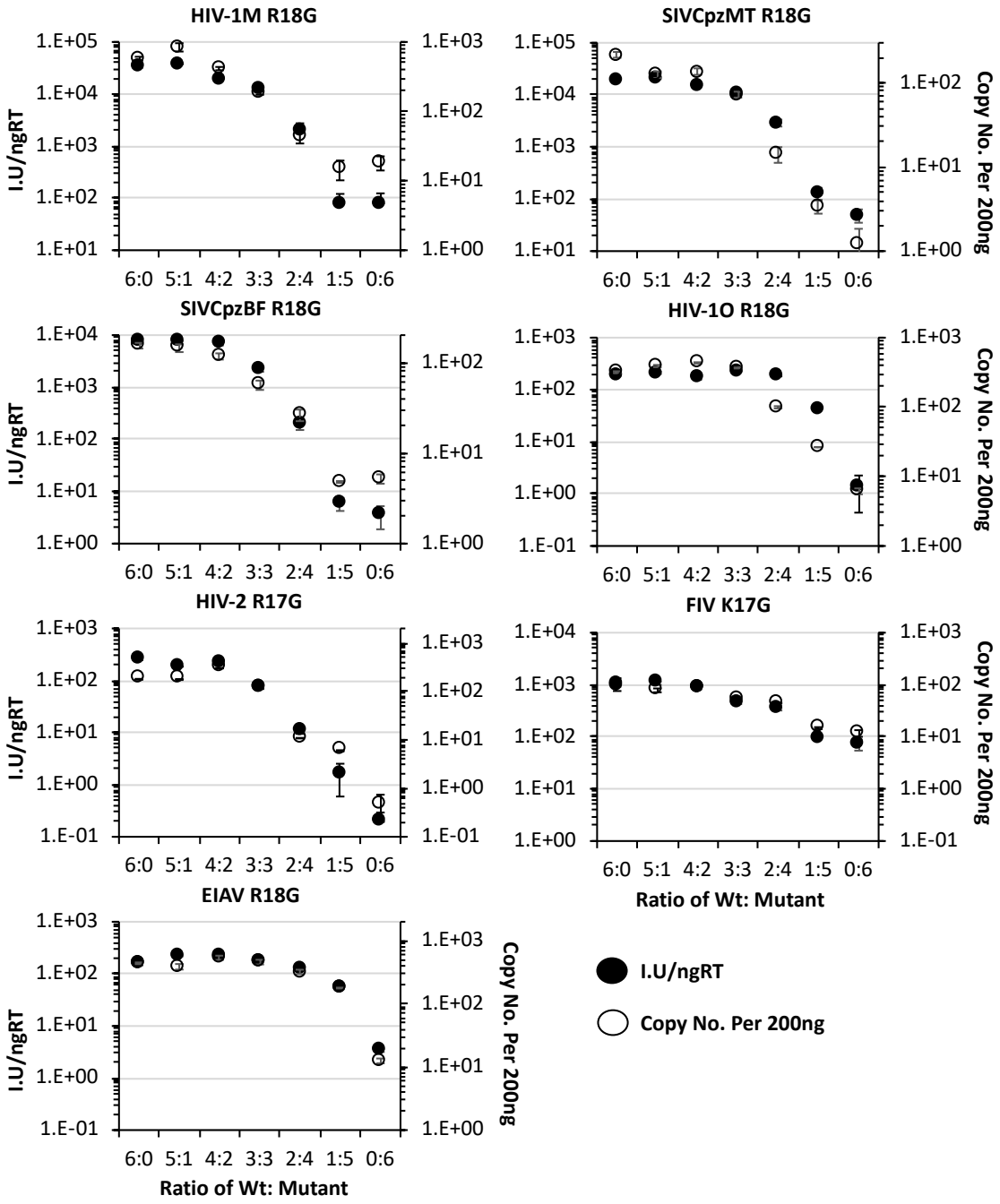


Figure 12: Disruption of the Arg electrostatic channel by R18G mutation causes a defect in reverse transcription which is mirrored in infectivity

Titres of mixed capsid VSV-G pseudotyped vectors derived from lentiviruses (I.U./ngRT) on U87 cells, +/- SEM N=2. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Titres determined from infections with equal dose of each ratio prep, dose was normalised by reverse transcriptase activity for MOI of 0.3 in fully wildtype capsid vectors. Reverse transcription products (Copy No. per 200ng) were measured at 6 hours post infection with mixed capsid vectors at an MOI of 0.3, +/- SEM N=2. Vectors contained a ratio of wildtype (Wt) and R18G, or equivalent, (mutant) capsid mutation. A ratio of 6:0 represents a fully wildtype capsid containing vector whereas a ratio of 0:6 only contains capsid bearing the R18G, or equivalent, mutation.

3.3.3 The defect caused by loss of the electrostatic channel is in reverse transcription and not particle production

After showing that all the viruses studied had a defect with the disruption of the electrostatic channel, we next sought to determine where the defect in the life cycle was. From data in **Figure 12** it remained unclear whether the reduction in reverse transcription, and thus infectivity, due to R18G was due to a defect in particle formation or maturation, as was the case for capsid mutant K227A. To test this we performed a western blot for capsid on HIV-1M for each ratio of R18G [**Figure 13**]. We found that there was no loss of capsid in viral supernatant as the ratio of Gly 18 increased. This demonstrates that the capsid mutant R18G is cleaved correctly in HIV-1M and that there isn't a decrease in particle maturation.

Next we measured the reverse transcription activity in viral supernatants for each of the mixed capsid vectors at each ratio for all panel members [**Figure 14**]. We found that there is a non-significant difference in reverse transcription activity in viral supernatant between any of the mixed capsid vectors for any panel members except for FIV. FIV showed a small reduction in reverse transcription activity as the ratio of Gly increased. Taken together, this data suggests that the defect for R18G (or equivalent) is post-entry in the life cycle at reverse transcription.

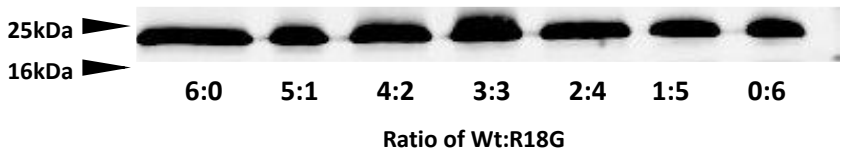


Figure 13: Disruption of the electrostatic channel in HIV-1M does not alter capsid release in viral supernatants
Western blot detecting capsid within mixed capsid cores of HIV-1M containing wildtype or capsid bearing the R18G mutation. Lanes were normalized by SG-PERT and contain the same reverse transcriptase value as was used in **Figure 12**. A ratio of 6:0 represents a fully wildtype capsid containing vector whereas a ratio of 0:6 only contains capsid bearing the R18G, or equivalent, mutation. Size marker positions are shown with molecular weights (KDa)

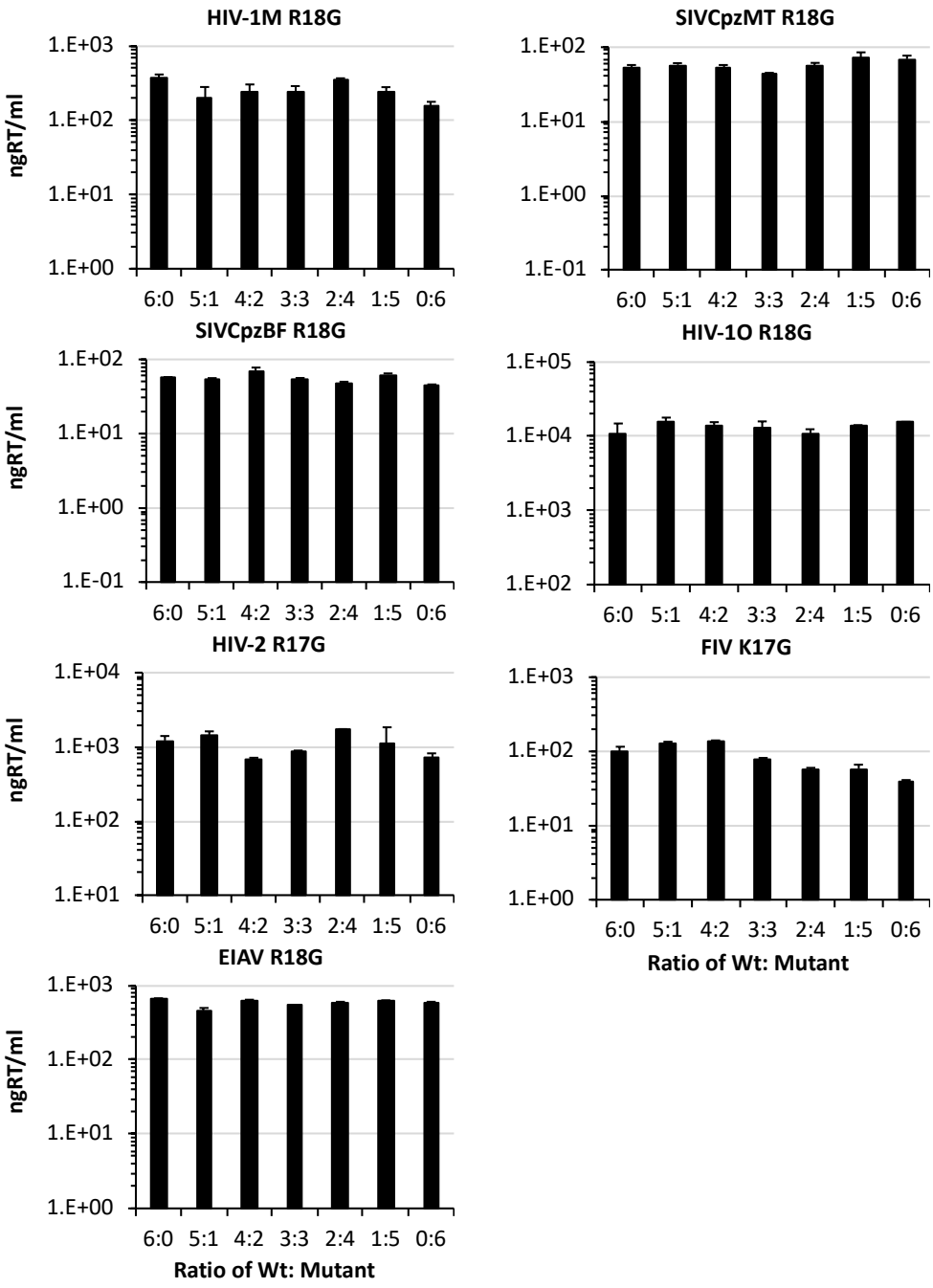


Figure 14: Disruption of the electrostatic channel did not alter particle release
Activity of lentiviral reverse transcriptase enzyme was measured from vector containing supernatant via SG-PERT assay. A ratio of 6:0 represents a fully wildtype capsid containing vector whereas a ratio of 0:6 only contains capsid bearing the R18G, or equivalent, mutation.

To confirm the defect was at the level of reverse transcription, was undertook linear regression analysis on the mixed capsid vectors at each ratio. Linear regression values give a mathematical measure of how well the data fits a linear relationship. An R^2 value of 1 representing a completely linear system whereas a low value shows little or no correlation between two factors. We used titre, reverse transcription products and reverse transcription activity in viral supernatants at each ratio for each of the mixed capsid vectors for all the panel members [**Figure 15**]. We found that for all of the panel members, there was no correlation, shown by the low R^2 value for the dotted line, between titre and reverse transcription activity in viral supernatant, with R^2 values of 0.02 for HIV-1M. Whereas there was a strong correlation between titre and reverse transcription products with R^2 values greater than 0.8. This was the case for all panel members except for FIV, which showed a high level correlation with both with an R^2 of 0.86 and 0.89 for titre to reverse transcription activity and titre to reverse transcription products respectively. From this we can conclude that the defect observed in the R18G mutant series is entirely due to a failure to reverse transcribe. FIV, however, represents a outlier, as it appeared to have a defect in both viral production and infectivity when the electrostatic channel is disrupted [**Figure 15**].

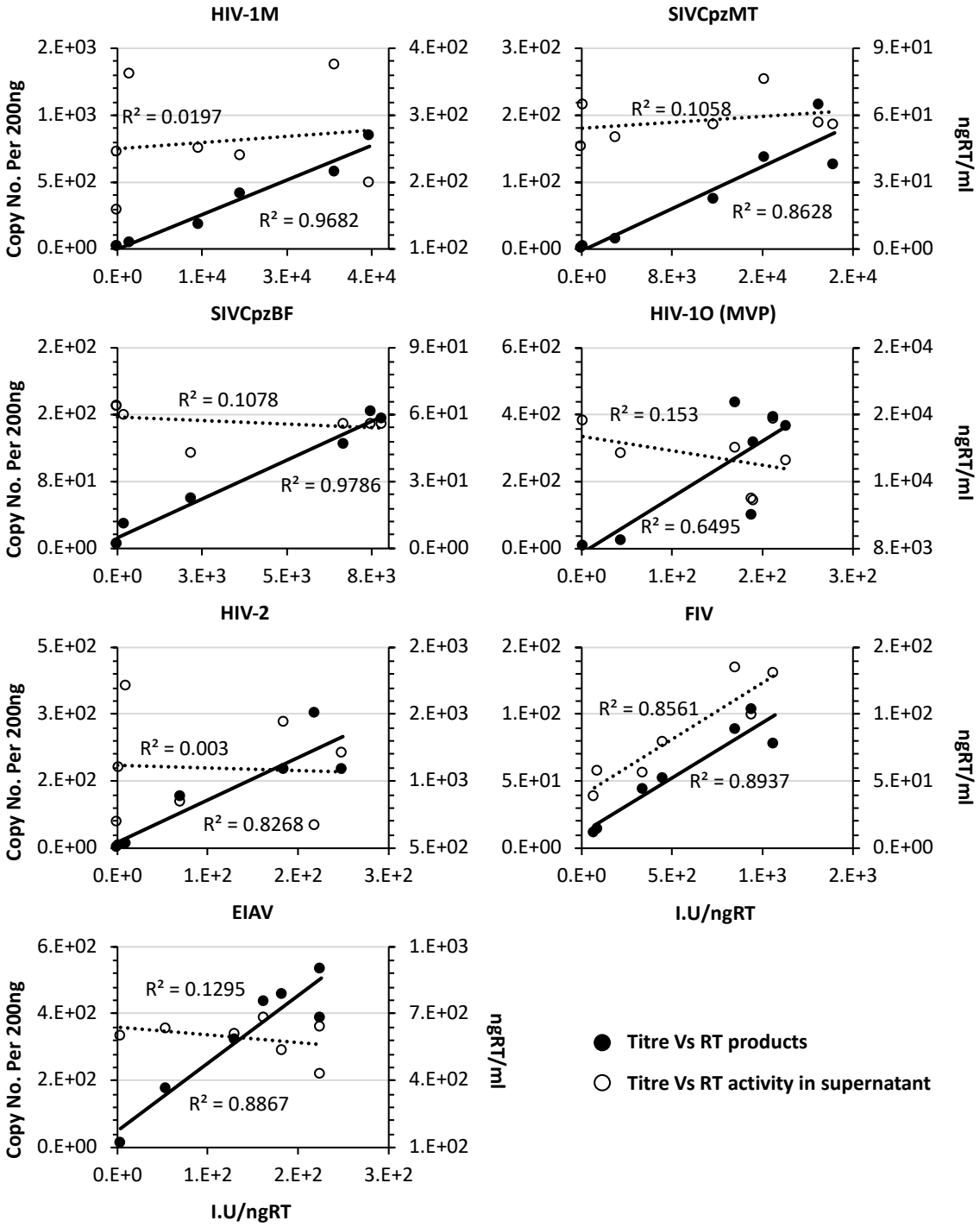


Figure 15: The defect in infectivity of R18G, or equivalent, mixed capsid viruses is due to a block at reverse transcription except in FIV

Linear regression analysis comparing relationship between lentiviral vector titre (I.U.ngRT) and reverse transcription products (Copy No. Per 200ng), shown in black and titre and reverse transcriptase activity in supernantant (ngRT/ml) shown in white, for each mixed core ratio. Titres and reverse transcription products taken from **Figure 12** and reverse transcriptase activity in supernatant taken from **Figure 14**.

3.3.4 A positively charged electrostatic channel is not unique to Lentiviruses

The retrovirus family contains seven different genera, *Alpharetrovirus*, *Betaretrovirus*, *Deltaretrovirus*, *Epsilonretrovirus*, *Gammaretrovirus*, *Lentivirus* and *Spumaretrovirus*. *Lentiviruses* are the best characterised and the focus of much research. We decided to investigate the other retrovirus genera for the existence of a charged electrostatic channel. To address this, we took a phylogenetic approach and constructed a maximum likelihood phylogenetic tree containing the capsid sequences from members of each genus of retrovirus [**Figure 16A**]. The capsid sequences were harvested from Genbank and aligned before being coloured, using Chroma-Clade, for Arg 18 in the HIV-1M capsid. For *Spumaretrovirus*, which contain an uncleaved Gag, Gag alignments were preformed and the region of the *Spumaretrovirus* Gag sequence which aligned with the capsid of HIV-1M was taken to be the *Spumaretrovirus* "capsid". As expected, there was a far greater level of diversity between viruses of different genera. However, we did observe that, along with *Lentiviruses Alpharetrovirus*, *Deltaretrovirus*, *Spumaretrovirus* and some *Betaretrovirus* did contain a positively charged amino acid at this position in capsid. All examined members of the *Spumaretrovirus* genus were found to contain Arg at this position whereas the *Alpharetrovirus*, *Betaretrovirus* and *Deltaretrovirus* genus contained mainly Lys [**Figure 16A**]. The *Gammaretrovirus* and *Epsilonretrovirus* genus, contained no positively charged residue at this position, suggestive that they don't contain an electrostatic channel.

To provide further evidence that the electrostatic channel is present in other *Retroviridae* members, we examined the published x-ray crystal structures of two *Deltaretrovirus*, human lymphotropic virus 1 and bovine leukaemia virus, one *Alpharetrovirus*, rous sarcoma virus and a *Gammaretrovirus*, murine leukaemia virus [**Figure 16B**]. Both *Deltaretrovirus* and the *Alpharetrovirus* contained the positively charged amino acid, shown in the tree **Figure 16A** to occupy the same structural location

as Arg 18 in HIV-1M capsid [**Figure 10B** and **Figure 16B**] likewise the Ser for MLV also occupied the same location in the structure [**Figure 16B**]. Due to lack of published structures of other retroviral capsid hexamers it remains to be determined whether the other retroviruses contain an electrostatic channel. It is also untested whether the electrostatic channel in other retroviruses plays a role in reverse transcription.

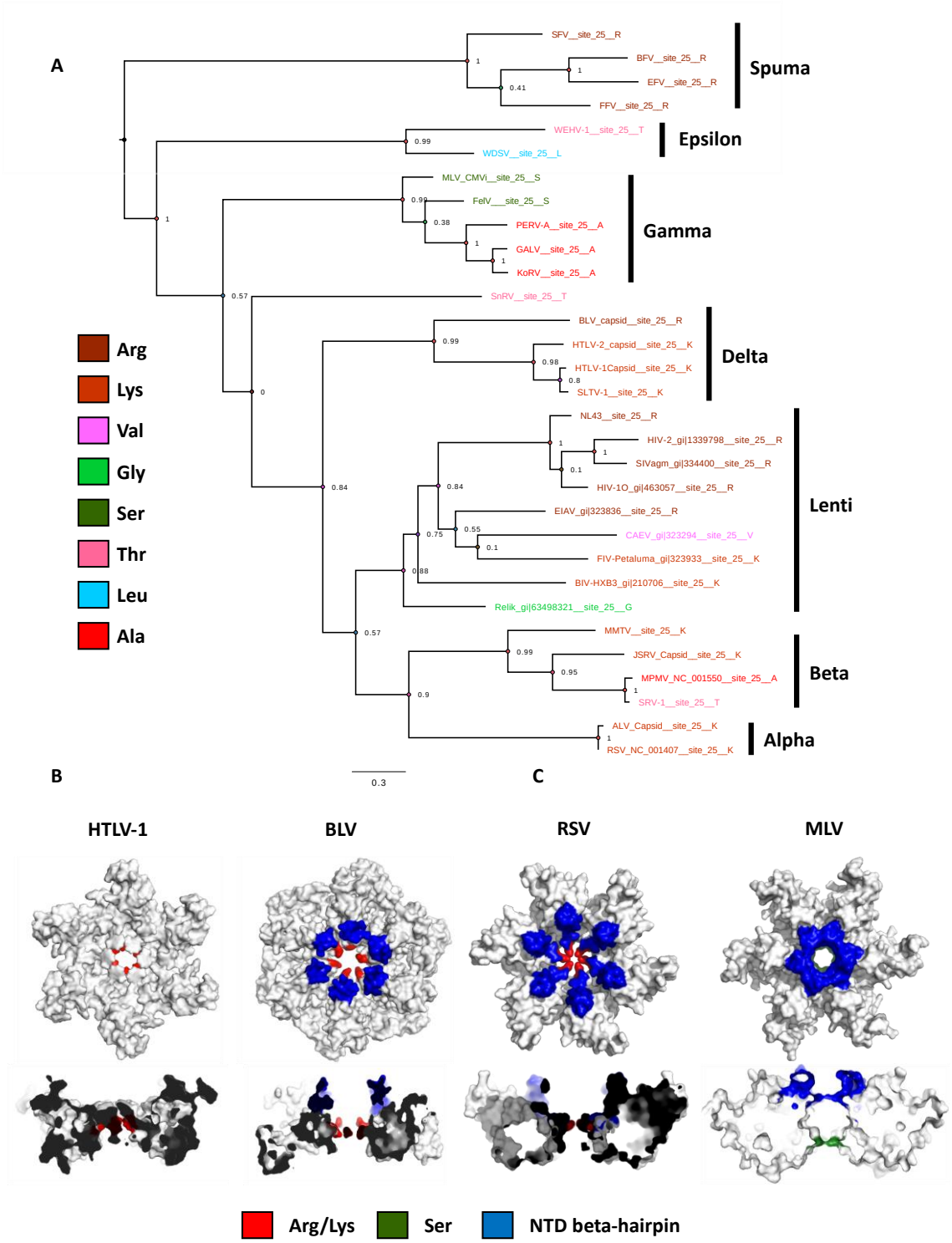


Figure 16: A positively charged amino acid at the six fold axis is not unique to the lentivirus linage

(A) A capsid phylogenetic tree was estimated by maximum likelihood using GTR substitution model and gamma-distributed substitution rate heterogeneity. Confidence was assessed with 1000 non-parametric bootstrap replicates and shown as a proportion on tree. Tree was coloured for residue 18 in the HIV-1M capsid. (B) Previously

3.4 Feline immunodeficiency virus

3.4.1 The feline lentiviral lineage can be divided by the positively charged residue in the electrostatic channel

We have shown that FIV behaviours differently to the other lentiviruses in the panel regarding its electrostatic channel. We found that FIV has the smallest defect when the electrostatic channel was mutated. Unlike the other panel members, FIV remained infectious once the electrostatic channel was lost through mutagenesis whereas the primate lentiviruses and EIAV became completely uninfectious [**Figure 12**]. Following from this we decided to take a more in-depth look at the FIV lineage. We harvested capsid sequences from each subtype of FIV from domesticated cats (FIV-Pca) and all unique capsid sequences from non-domesticated felines and constructed a phylogenetic tree. The tree contains FIV from various non-domesticated felines including pumas (FIV-Pco), lions (FIV-Ple) and pallas cats (FIV-Oma). The tree was then coloured for position 17 within the capsid of FIV-Pca [**Figure 17**]. Lys 17 in the FIV capsid is located at the 6 fold-axis of symmetry and structurally aligns with Arg 18 from the HIV-1M capsid [**Figure 18**].

The more comprehensive FIV phylogenetic tree allowed us to make two major observations. Firstly, position 17 within the FIV capsid completely divides the feline lineage with all FIV-Pca containing a Lys and all other FIVs containing an Arg [**Figure 17**]. Secondly, there seems to have been a switch from Arg to Lys along the branch leading to FIV-Pca which could suggest virus specific host adaption.

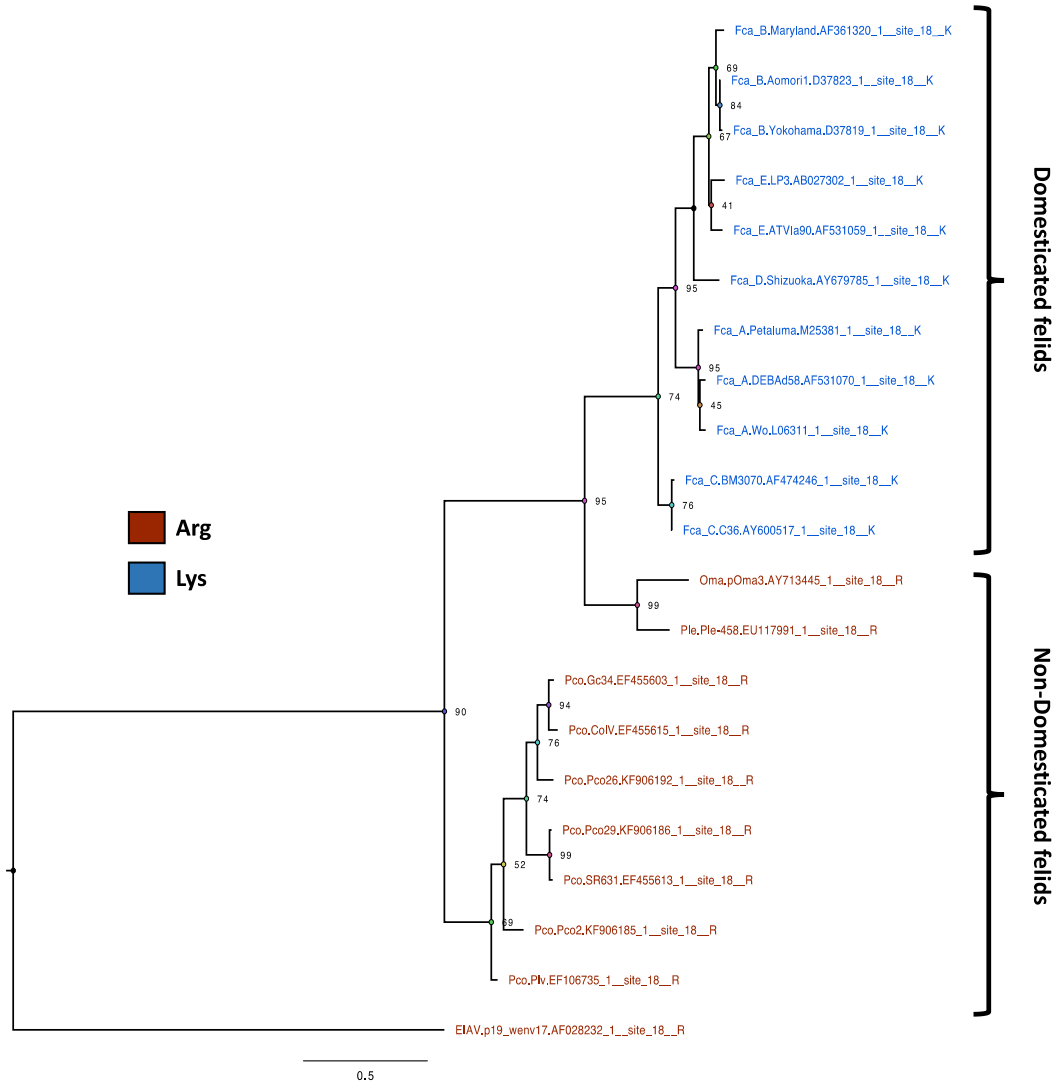


Figure 17: The Feline lentivirus lineage can be divided into FIVs of domesticated and non-domesticated felids by position 17 in the FIV capsid

A capsid phylogenetic tree was estimated by maximum likelihood using GTR substitution model and gamma-distributed substitution rate heterogeneity. Confidence in the tree model was assessed with 1000 non-parametric bootstrap replicates and is shown as percentage on each node. Tree was coloured for residue 17 in the FIV-Pca casid, using Chroma-Clade.

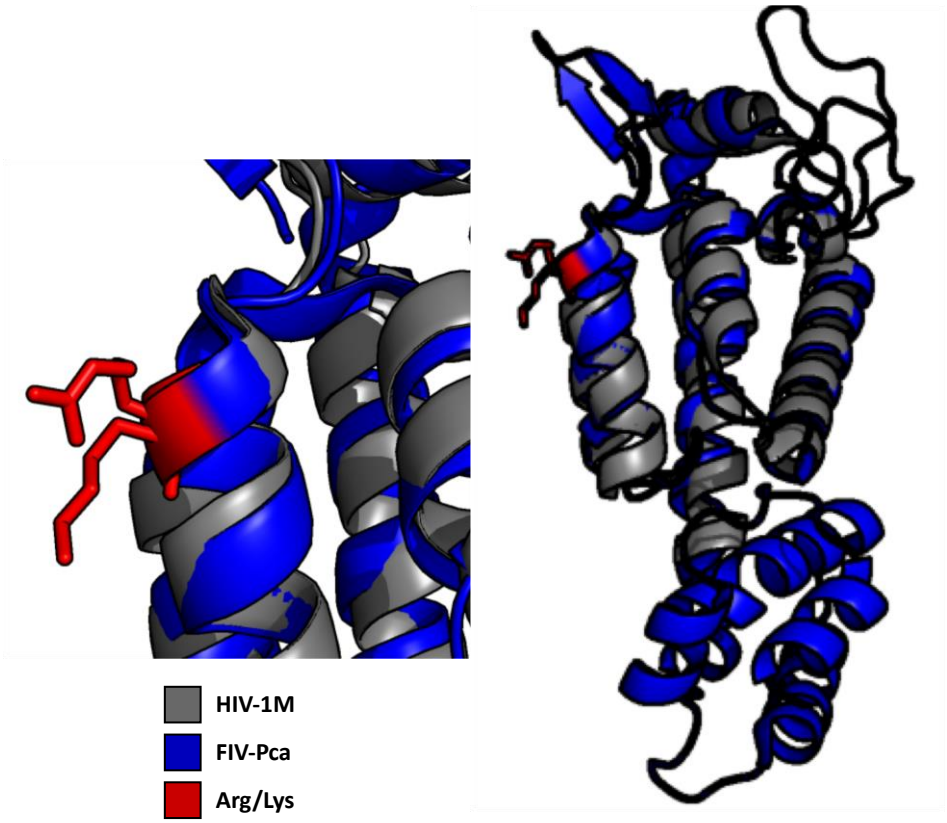


Figure 18: Residue Arg 18 in HIV-1M and Lys 17 in FIV occupy the same structural location in the mature capsid
Structural alignment of the previously published NTD of HIV-1M capsid (grey) and the FIV-Pca capsid (Blue).

3.4.2 FIV has a higher level of plasticity to its electrostatic channel

In HIV-1M there is a near complete conservation of Arg at position 18 within the capsid however other lentiviruses, such as FIV-Pca and BIV, form the electrostatic channel with a Lys, so we sought to determine the effect of mutagenizing HIV-1M's Arg to a Lys. Both amino acids are positively charged and Lys is naturally found in other lentiviruses including FIV-Pca and BIV so we tested the effect of making the capsid mutation R18K in the HIV-M capsid. Despite the high level of similarity between Arg and Lys, it was found that HIV-1M was unable to tolerate the capsid mutation R18K [**Figure 19A**]. Further mutagenesis at residue 18, to an Ala, revealed that HIV-1M seems to be highly intolerant to change at this position. Any change from Arg caused a severe defect to

infectivity due to a decrease in reverse transcription [**Figure 19B** and **Figure 12**]. This is perhaps not surprising considering the high level of conservation of Arg 18 with HIV-1M.

Whilst we were constructing the FIV phylogenetic tree [**Figure 17**] we noted a similar level of conservation in FIV-Pca for Lys 17 that HIV-1M has for Arg 18. So we decided to investigate whether FIV-Pca shared a similar level of intolerance for changes in its electrostatic channel. As FIV-Pca naturally harbours a Lys at position 17, we mutagenized it to an Arg [**Figure 19C and Figure 19D]** and tested the effect on infectivity. Where the HIV-M capsid mutant R18K had a 1000 fold defect in infectivity, FIV-Pca K17R only had a 5 fold defect [**Figure 19D].** Like HIV-1M, FIV-Pca has a very high level of conservation for residue 17 but unlike HIV-1M, it is able to tolerate changes at this position. Furthermore, we have already demonstrated that FIV is relatively resistant to the removal of the positive charge in its electrostatic channel [**Figure 12** and **Figure 19D]**. Taken together, these data suggest that the FIV-Pca capsid is more plastic as this position.

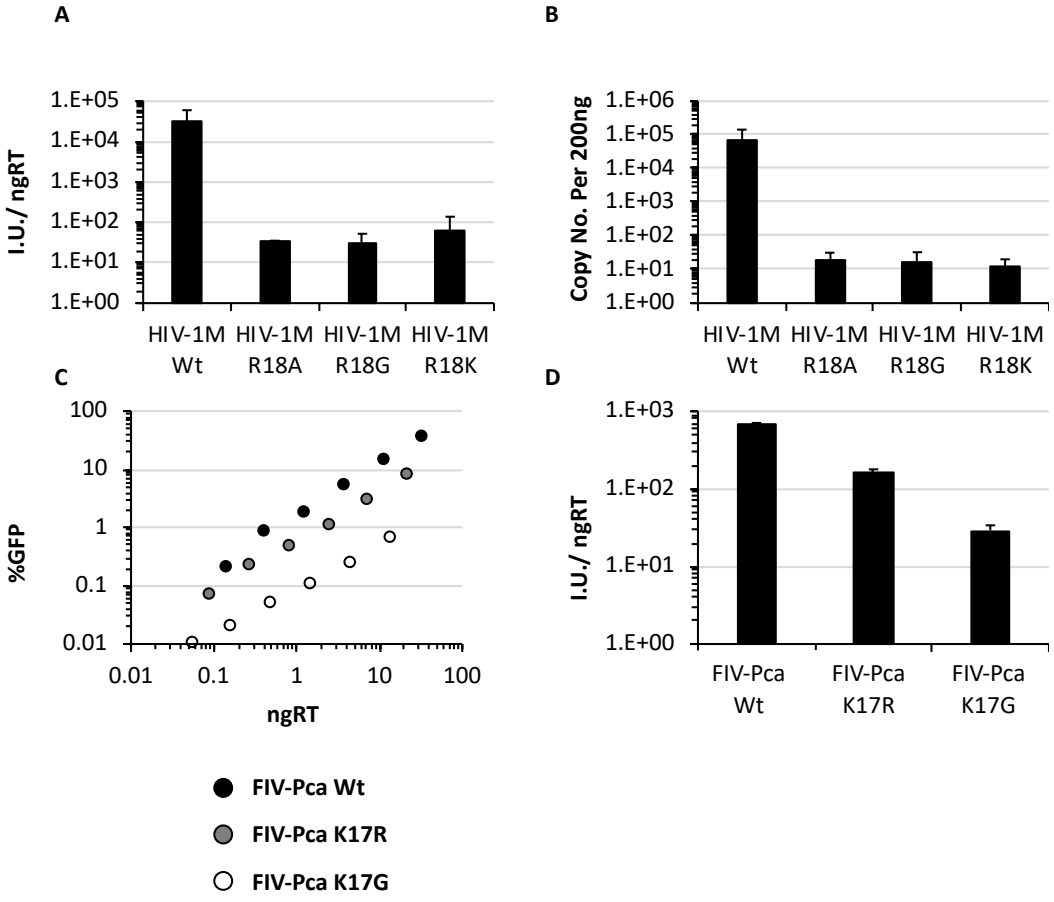


Figure 19: FIV-Pca can tolerate mutation of charged residues in electrostatic channel

(A) Titres determined from infections of VSV-G pseudotyped vectors derived from HIV-1M bearing either wildtype, R18A, R18G or R18K capsid mutations, using three doses of vector with MOI below 0.3, +/- SD N=2. (B) Reverse transcription products, measured 6 hours post infection with same vectors in (A) at MOI of 0.3. (C) Infectivity plots of VSV-G pseudotyped vectors derived from FIV-Pca either bearing wild type capsid (FIV-Pca Wt), K17R (FIV-Pca K17) containing or K17G (FIV-Pca K17G) containing capsid, on U87 cells. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (D) Titres determined from data in (C) using three doses of vector with MOI below 0.3, +/- SD N=2.

3.4.3 The IP6 binding sites in the immature Gag are conserved throughout the feline lineage of lentiviruses

Having shown that the IP6 binding site in the mature capsid of FIV is variable, perhaps linked to domestication, we examined the conservation of the IP6 binding site in the immature Gag of FIVs. Taking the more detailed phylogenetic tree of the FIV lineage [**Figure 17**] we coloured it for Lys 150 and Lys 219 in the capsid of FIV-Pca, the equivalent of Lys 158 and Lys 227 in HIV-1M identified in **Figure 7**, [**Figure 20**]. Despite seeing variation at the IP6 binding site in the mature capsid of FIVs, the IP6 binding site within the immature Gag was conserved throughout the feline lineage of lentivirus. In the immature Gag, both FIVs from domesticated and non-domesticated feline contained a Lys. The sequences of some of the FIV-Pca did not extend to Lys 227, however, enough sequence did for us to be confident that the Lys is fully conserved at this position.

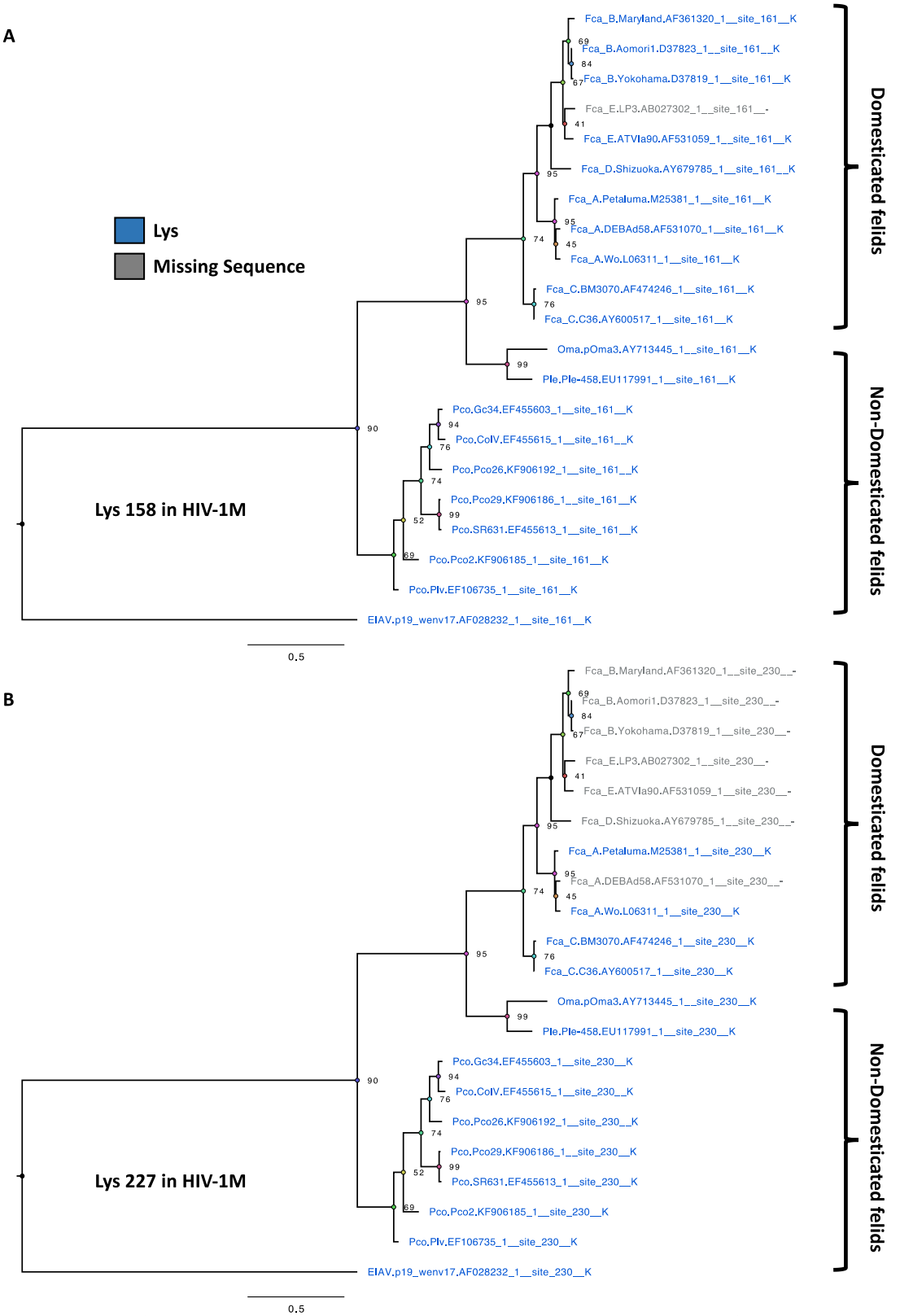


Figure 20: Conservation of the IP6 binding site in the immature Gag of FIVs
A capsid phylogenetic tree was estimated by maximum likelihood using GTR substitution model and gamma-distributed substitution rate heterogeneity. Confidence was assessed with 1000 non-parametric bootstrap replicates and shown as percentage on tree. Tree was coloured, using chroma clade, for residue 150 (A) and residue 219 (B) in the FIV-Pca capsid, equivalent to Lys 158 and Lys 227 in the HIV-1M capsid respectively.

3.5 RELIK's electrostatic channel

Rabbit endogenous lentivirus type K (RELIK) is an ancient lentivirus which is found in the genome of European rabbits and was thought to have invaded their genome over 12 million years ago [301]. The capsid of this ancient endogenous lentivirus has been found to have several attributes shared with extant lentiviruses. Most notably, the capsid of RELIK structurally aligns with that of HIV-1M and has been shown to be able to bind human CypA by Isothermal Titration Calorimetry (ITC) using recombinant NTD of the RELIK capsid [302]. Using RELIK we sought to time stamp the electrostatic channel and determine whether the electrostatic channel is a new or old adaption of the lentiviral capsid. In **Figure 10** we showed that RELIK contains a Gly at position 18 in its capsid and that this structurally aligns with Arg 18 in the HIV-1M capsid. Harvesting all available capsid sequences from RELIK we found that in the majority there was only a single nucleotide difference between Gly 18 in RELIK and an Arg codon. The majority of the RELIK sequences contained the codon GGA for residue 18 and with a single nucleotide change, to AGA, it would encode an Arg. Previously, the capsid for RELIK has been cloned into EIAV, the closest modern day lentivirus, however this causes a large defect in titre [357]. We hypothesise that this drop in titre is due to the loss of a positively charge residue at the six-fold axis of symmetry and thus an inability to use IP6. To test this, we restored the electrostatic channel in RELIK by mutagenizing Gly 18 to an Arg and measured infectivity in a highly permissive cell line, CRFK, [**Figure 21**]. We reasoned that a rescue in titre would be evidence that RELIK did contain an electrostatic channel and that it played a critical role in its life cycle. However, we observed that once we restored the electrostatic channel to the RELIK capsid, cloned into EIAV, there was no significant rescue to titre [**Figure 21A** and **Figure 21B].** To confirm that both the EIAV-RELIK and the EIAV-RELIK G17R construct were producing and releasing particles, we measured reverse transcriptase activity in viral supernatant [**Figure 21C**]. We observed no

significant difference in reverse transcriptase activity in viral supernatants for either EIAV-RELIK or EIAV-RELIK G17R. Taken together this suggest that restoring the electrostatic channel was not enough to rescue infectivity for the EIAV-RELIK chimera.

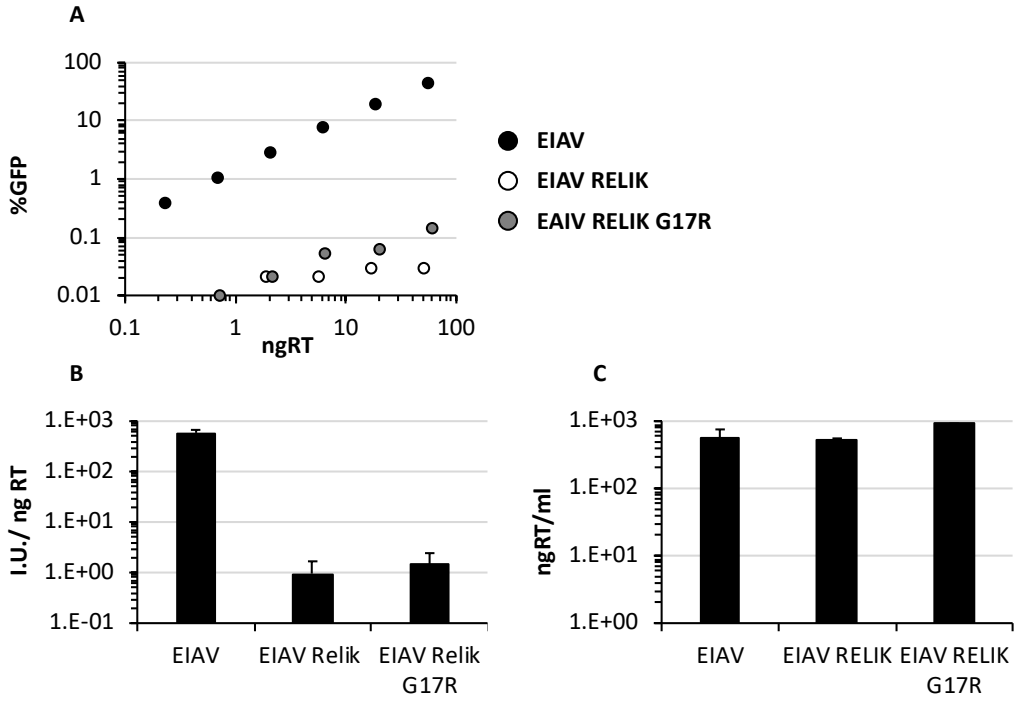


Figure 21: Mutagenesis to restore the electrostatic channel of RELIK does not rescue infectivity
A) Infectivity plots of VSV-G pseudotyped vectors derived from EIAV either bearing wildtype EIAV capsid (EIAV), RELIK capsid (EIAV-RELIK) or the RELIK capsid with the G17R mutation (EIAV-RELIK G17R) as shown on CrFK cells. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2. (C) Viral supernatants used in (A) were subjected to SG-PERT assay to measure reverse transcriptase activity in supernatant. Two technical repeats, +/- SD.

3.7 Chapter discussion

3.7.1 IP6 and the immature Gag lattice

We have shown that the IP6 binding site in the immature Gag is conserved throughout the lentiviral lineage. Consistent with previous work in the field [355, 290, 356], we have shown that disrupting the IP6 binding sites in the immature Gag, Lys 158 and Lys 227

equivalents, causes a defect in particle formation. Previous work has established this for HIV-1M [290] however, here we extend the observation to diverse primate and non-primate lentiviruses showing that in all viruses examined IP6 binding to the immature Gag is an essential step within their lifecycle as it functions to promote particle assembly [**Figure 8** and **Figure 9**]. From this, we conclude that IP6 is a cofactor for lentiviruses in general, and is not specific to HIV-1M, which acts to promote particle formation.

Lentiviruses have adapted to use IP6 to help assemble the immature Gag lattice and thus drive particle assembly. As all lentiviruses face the same challenge when it comes to assembling the Gag lattice, e.g. the tight clustering of Gag molecules, it may not be surprising to observe that this mechanism is highly conserved across the lineage. Instead, it is striking that some lentiviruses e.g. the Ovine-caprine lineage are lacking in one of the essential binding sites, Lys 158. As comparatively little is known about the ovine-caprine lineage of lentiviruses and as there isn't a readily available molecular clone for any of the lineage members, work with ovine-caprine lentiviruses becomes difficult. It should be noted that the ovine-caprine lineage does process Lys 227, it is yet to be determined what effect this has on particle assembly and whether or not it is IP6 dependent. It is possible, that a single IP6 binding site in the immature Gag of the Ovine-Caprine lineage is enough for efficient particle assembly. However, without a molecular clone it is yet unknown.

3.7.2 Function of the Arg 18 ring

As described in section 1.3 the electrostatic ring in the mature capsid has been shown to not only bind IP6 but also dNTPs. There are two non-exclusive hypothesis of the function of the Arg 18 ring. Firstly it binds IP6 to regulate the stability of the capsid core and secondly it binds to and imports dNTPs to fuel encapsidated reverse transcription. The

data presented in this chapter can be explained using either model. We have shown that by disrupting the Arg 18, or equivalent, in various lentiviruses we see a decrease in titre which is caused by a decrease in reverse transcription [**Figure 12** and **Figure 15**]. Either this is due to the mature capsid being unable to bind IP6, having a decreased stability and uncoating early thus preventing reverse transcription taking place. Likewise, the reverse transcription defect could be due to an inability of the capsids to bind and recruit dNTPs to the interior of the capsid. Despite not identifying the order of the defect e.g. inability to bind IP6 leading to uncoating or inability to fuel reverse transcription, we have shown that the electrostatic channel is essential for primate lentiviruses.

However, more distant lentiviruses to HIV-1M, such as EIAV and FIV, have a different dependence on the channel. We know that the life cycle of both EIAV and FIV vary considerably from that of HIV-1M, as seen in the difference in cellular cofactor usage. It has previously been shown that neither EIAV nor FIV require cofactors which are essential for optimal HIV-1M infection, these include NUP153 [285], NUP358, CypA [see chapter 4] and FIV also does not require TNPO3 however HIV-1M and EIAV do [314 323 358]. Taken together, it can be seen that the life cycle of FIV is strikingly different to that of HIV-1M. Therefore, it is not unexpected to see that FIV has a different dependence on the electrostatic channel. With the observation that FIV requires Lys 219, equivalent to Lys 227 in HIV-1M, for infectivity but not Lys 17 [**Figure 8** and **Figure 12**] it could be hypothesised that FIV undergoes a comparably early uncoating event to HIV-1M. Recent evidence is building to suggest that the capsid of HIV-1M remains intact during or even a short time after nuclear import [94,95], it has also been suggested that reverse transcription is completed within the nucleus. However, it might be possible that FIV undergoes uncoating at an earlier stage and thus is less dependent on the stabilising effect of IP6. Furthermore, the mature FIV capsid has been shown, by electron microscopy, to form spherical structures with a diameter of 40 – 80 nm [359] whereas the mature HIV-1M

capsid has a fullerene cone structure [93]. It is possible that due to the more uniform curvature of the FIV capsid there is less/ different need for an external stabilising factor like IP6. The FIV from non-domesticated felines, which unlike FIV-Pca, contains an Arg at position 17 and thus raises the question as to whether it requires IP6 as a stability factor for its mature capsid. An interesting prospect is during FIV-Pca adaption to domesticated cats, IP6 became a less essential cofactor as an alternative life cycle, possible involving an early uncoating event, evolved.

The role of Arg 18, or equivalent, within the lentiviral capsid is likely to be more complex than originally thought. The model where Arg 18 is used to import dNTPs into the capsid interior to fuel reverse transcription and the IP6 stabilisation model might not be mutually exclusive. Both models involves Arg 18 being important for electrostatic interactions which are characteristically very fast. Therefore it's possible that Arg 18 both imports dNTPs and stabilises the capsid via binding IP6. With ~250 hexameric electrostatic channels, it's possible that not all hexamers will be bound to IP6 simultaneously. Thus when some channels bind IP6 to stabilise the capsid, others might be being used to import dNTPs to fuel reverse transcription. It is possible that HIV-1M has evolved to use one feature, the Arg 18 ring, for multiple purposes. Also, all of the work presented above is thought of in the context of hexamers. Currently we have no data on the role of the pentameric channel and the role it may play in binding IP6 and dNTP import. It is possible that the pentamers might control dNTP influx when the hexamers control stability via IP6, or vice versa.

3.7.3 The electrostatic channel as a novel drug target

As mentioned previously the electrostatic channel within HIV-1M is highly conserved, ~99.4% and this work and others have shown that changes to position 18 are highly

detrimental for infectivity [**Figure 12**] [237] [238]. The electrostatic channel has also been shown, via X-ray crystallography, to be able to bind several compounds including IP6, dNTPs, NTP and Mellitic acid [237]. This suggest that the electrostatic channel may represent a novel drug target. This has the advantage of already having compounds which can be used in directed drug design and a high fitness costs to escape. As shown in this work, HIV-1M is unable to tolerate even a substitution to another positively charged amino acid at residue 18 [**Figure 19**] and removal of the positive charge is just as lethal for the virus. Thus to escape, compensatory mutations would need to be pre-existing in the virus before Arg 18 is changed as if that occur first the virus won't be able to replicate. This seems unlikely as any compensatory mutations may have a fitness cost themselves and be selected against. Supporting this is the observation that despite the huge amount of sequence data available for the HIV-1M capsid there is not a cluster of sequences which have lost the Arg and continue to spread.

Furthermore, as the electrostatic channel is widely conserved throughout the lentiviral lineage [**Figure 10**] it also represents a potential broad spectrum target. It may be possible to design a drug against HIV-1M but have it be effective against all the different clades of HIV-1M as well as HIV-1O and HIV-2 for example. It is known that not all subtypes in HIV-1M are equally susceptible to antiretroviral therapy as clade B, the most prominent in the western world. There is evidence that clade AG and C from HIV-1M are inherently less susceptible to protease inhibitors due to polymorphisms [private communication from Gupta lab]. However, the electrostatic channel is conserved in all clades and despite it not being characterised in other clades it can be assumed it's just as essential and intolerant of changes.

3.7.4 RELIK's electrostatic channel

RELIK is thought to have invaded the genome of the ancestor of European hares and rabbits about 12 million years ago [301]. Once a retroviruses becomes endogenous the virus will evolve under neutral selection, unless it is plays a role in the host organisms proliferation and survival, which will lead to the accumulation of inactivating mutations. As RELIK has not been reported to play a role in rabbit proliferation or survival, we assumed it to be under neutral selection and hence has lost its electrostatic channel due to somatic mutation. However, upon mutagenesis to restore the electrostatic channel, we saw no significant increase in viral titre [**Figure 21**]. Either this shows that RELIK does not have a charged electrostatic channel which would suggest the channel is a recent feature of lentiviruses or that there are other mutations within the RELIK capsid which render the EIAV-RELIK chimera poorly infectious. Interestingly, RELIK does contain the IP6 binding sites within the immature Gag, Lys 155 and Lys 222 [**Figure 7**] which are the equivalents to Lys 158 and Lys 227 in the HIV-1M capsid respectively, within its capsid. This could suggest that RELIK does at least use IP6 for particle assembly. It therefore seems unlikely that RELIK would not use IP6 in its mature capsid. Therefore, we conclude that RELIK must contain other mutations within its capsid that render the EIAV-RELIK chimera poorly infectious.

3.7.5 Murine leukaemia viruses and encapsidated reverse transcription

MLV is also thought to undergo encapsidated reverse transcription as shown by purified reverse transcription complexes (RTC), which were found to contain capsid protein, being able to initiate and complete reverse transcription [360]. However, unlike HIV-1M, when its capsids are purified its reverse transcription products are not protected from nucleases [237,360]. The MLV hexamer does contain a central channel but it lacking an

electropositive core, bearing Ser at position 17 [**Figure 10** and **Figure 16B**] and no other positively charged amino acids in the proximity. Notably MLV only infects dividing cells. One possibility is that expression of cellular nucleases, such as TREX-1, might be connected to cell cycle and regulated during cell division. Thus MLV may have linked uncoating to cell cycle to avoid cellular DNA sensor and nucleases. It would be interesting to determine how MLV is able to control its capsid stability without IP6.

3.8 Future work

We have provided evidence that IP6 is a conserved cofactor throughout the lentivirus lineage. We and other have proposed that IP6 is used to maintain capsid stability, prevent premature uncoating in an unfavourable environment such as the cytoplasm and promote encapsidated reverse transcription. Although we claim that other lentiviruses undergo encapsidated reverse transcription, no direct measurement of this has been made. To test the hypothesis, that encapsidated reverse transcription is a general feature of lentiviruses, it would be possible to perform the encapsidated reverse transcription assay in the presence or absence of nucleases with the panel of diverse lentiviruses harbouring the R18G, or equivalent, capsid mutation. Likewise, no direct measurements of capsid stability, or capsid core half-life were measured. Using modern single-particle fluorescence microscopy it would be possible to determine the half-life of individual capsid cores of various lentiviruses in the presence or absence of IP6. Furthermore, capsid stability could also be measured via thermal stability in the presence or absence of IP6 or the R18G capsid mutation.

We took a genetic approach to showing that IP6 is required for production of progeny virus, using the L227A mutant series. To strengthen our work, IP6 could be

depleted from producer cells, via Sh or Si RNA, and infectious titre, reverse transcriptase activity and Gag cleavage in viral supernatant measured.

We have shown that the electrostatic channel in the FIV of domesticated cats, FIV-Pca, is able to tolerate changes to the electrostatic ring. However, it would be interesting to see if FIV from non-domesticated cats, which naturally harbours an Arg at position 17 of the capsid, is also able to tolerate changes to the electrostatic channel. It would also be interesting to investigate whether FIV from non-domesticated cats is more FIV-Pca or more HIV-1M like in terms of cofactor usage.

4 CypA is important for human adapted lentiviruses

4.1 Chapter 4 introduction

As described in section 1.4 the HIV-1M capsid interacts with several host cofactors for optimal infection. Cyclophilin A (CypA) is one such cofactor. First identified during a yeast two hybrid screen using HIV-1 Gag as bait [283], CypA's role in the HIV-1M life cycle has been an area of intensive study. Since its discovery as a cofactor, CypA has been shown to bind to the exterior of the capsid core at a proline-rich loop region [294] and to play a role at various stages of the life cycle of HIV-1M, including integration site selection [112], immune evasion [230] and regulation of capsid stability [304]. It has been shown that CypA is important for infection in both primary macrophages [230] and T cells [293]. However its exact role in the HIV-1M lifecycle remains unclear. Furthermore, the role CypA plays in the life cycle of other lentiviruses is even less clear. Many investigations into CypA's role in HIV-1M life cycle have taken advantage of the immunosuppressive drug, cyclosporine (CsA), which has been shown to compete with the HIV-1M capsid for the same binding site on CypA [295,306] and therefore prevent the CypA-capsid interaction.

In brief, our model is that after cellular entry, HIV-1M binds CypA which it uses to help direct later stages in the life cycle, such as nuclear import pathway selection [112] and innate immune evasion [230]. Furthermore, we hypothesised that the binding of CypA may play an indirect role in regulating capsid stability via regulation of IP6 binding. Therefore, using a comparative virology approach with our lentivirus panel established in chapter 3, we sought to determine whether the capsid-CypA interaction is conserved and if this interaction is essential for optimal infection for other lentiviruses. Furthermore, we took a genetic approach to investigate whether there is any relation between CypA usage in HIV-1M and the recently described electrostatic channel within the mature capsid core with respect to IP6 binding.

Results

4.2 HIV-1M is uniquely adapted to use CypA for optimal infection

4.2.1 The ability to bind CypA is a general feature of lentiviruses

We wanted to determine the role CypA plays in the lifecycle of various lentiviruses. Within the capsid of HIV-1M Pro 90, which is the binding site for CypA [294], has been calculated to be conserved within >99.7% of sequences in the Los Alamos database [354]. This staggeringly high level of conservation prompted us to examine the conservation of Pro 90 in other lentiviruses and to investigate the role that the cofactor CypA plays in their lifecycles. First we approached the question of conservation from a phylogenetic standpoint. Using the phylogenetic tree generated in chapter 3, which contains representative members from each major lineage [**Figure 7**], we coloured the tree using Chroma-Clade for Pro 90 in the HIV-1M capsid [**Figure 22**]. We found that Pro 90 was highly conserved throughout the lentiviruses lineage with it being present in all examined lentivirus capsids.

Next we sought to determine if the panel of diverse lentiviruses, established in chapter 3, could functionally bind CypA. To examine this we cloned CypA, or the CypA like-domain of Nup358, into the human TRIM5 protein replacing its PRYSPRY domain. TRIM5 is known to be able to restrict lentiviral infection if it is able to bind to the lentiviral capsid, thus we reasoned if our viruses could bind CypA they would be restricted. To avoid any confounding factors, we overexpressed the Trim-fusion proteins in CrFK cells, a highly permissive feline cell line with a defective TRIM5 gene [361], as it's known that human TRIM5 can restrict non-primate lentiviruses such as FIV and EIAV [260]. After the transduction CrFK cells were selected before the panel of lentiviruses were titrated on to the cells seeking an infectivity defect [**Figure 23A**]. With the exception of HIV-2 for CypA, it was found that all of the panel members were inhibited suggesting that they bind both the CypA and the Nup358 TRIM fusion proteins but to varying

extents. HIV-2 showed no significant decrease in titre with the CypA TRIM fusion proteins, this is consistent with previous work showing HIV-2 is unable to bind CypA [296]. HIV-2 did, however, show a decrease in titre with the Nup358 TRIM fusion. Overall, the CypA-TRIM fusion protein was far more restrictive than the NUP358 fusion. The primate lentiviruses all had at least a 10 fold decrease in titre whereas the NUP358 fusion caused about a 5 fold decrease [**Figure 23B**]. EIAV showed the smallest defects with both fusion proteins. From the data, we conclude that not only is the binding site for CypA conserved amongst lentiviruses but also the ability to functionally bind it.

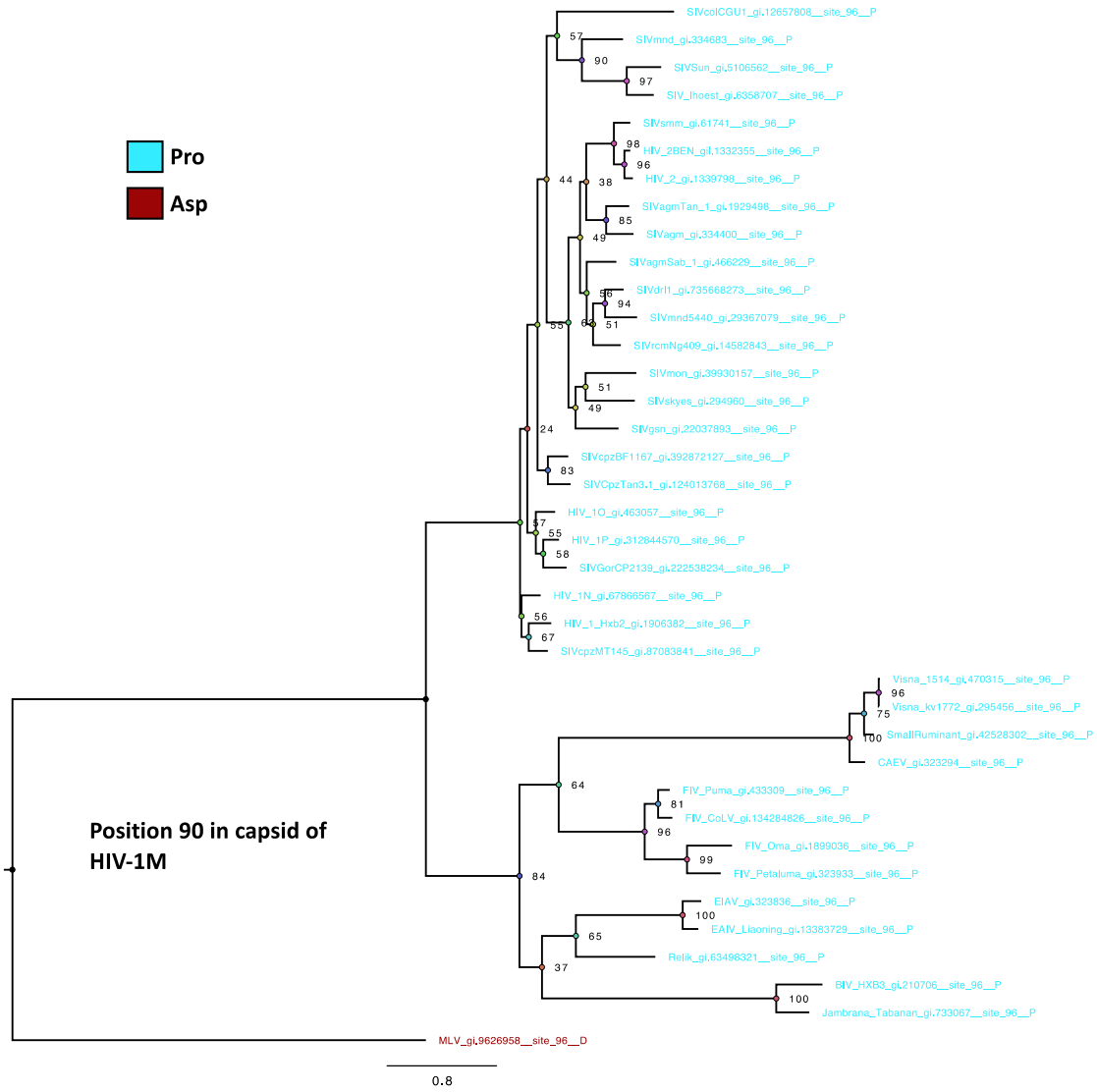


Figure 22: Proline 90 is highly conserved throughout the lentivirus lineage
A capsid phylogenetic tree was estimated by maximum likelihood using GTR substitution model and gamma-distributed substitution rate heterogeneity. Confidence was assessed with 1000 non-parametric bootstrap replicates and shown as percentages on the tree. The phylogenetic tree was then coloured using Chroma Clade according to residue 90 HIV-1M capsid.

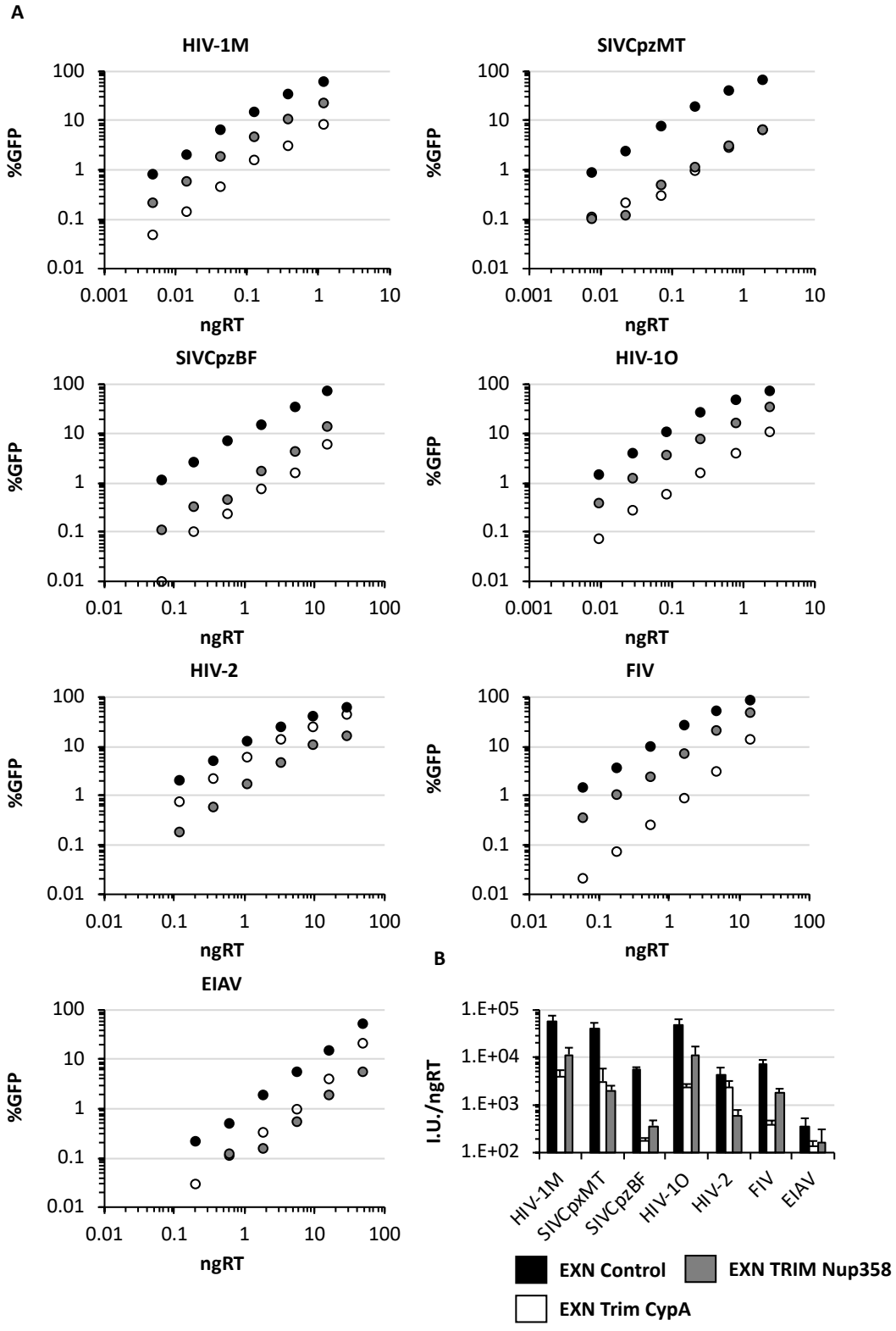


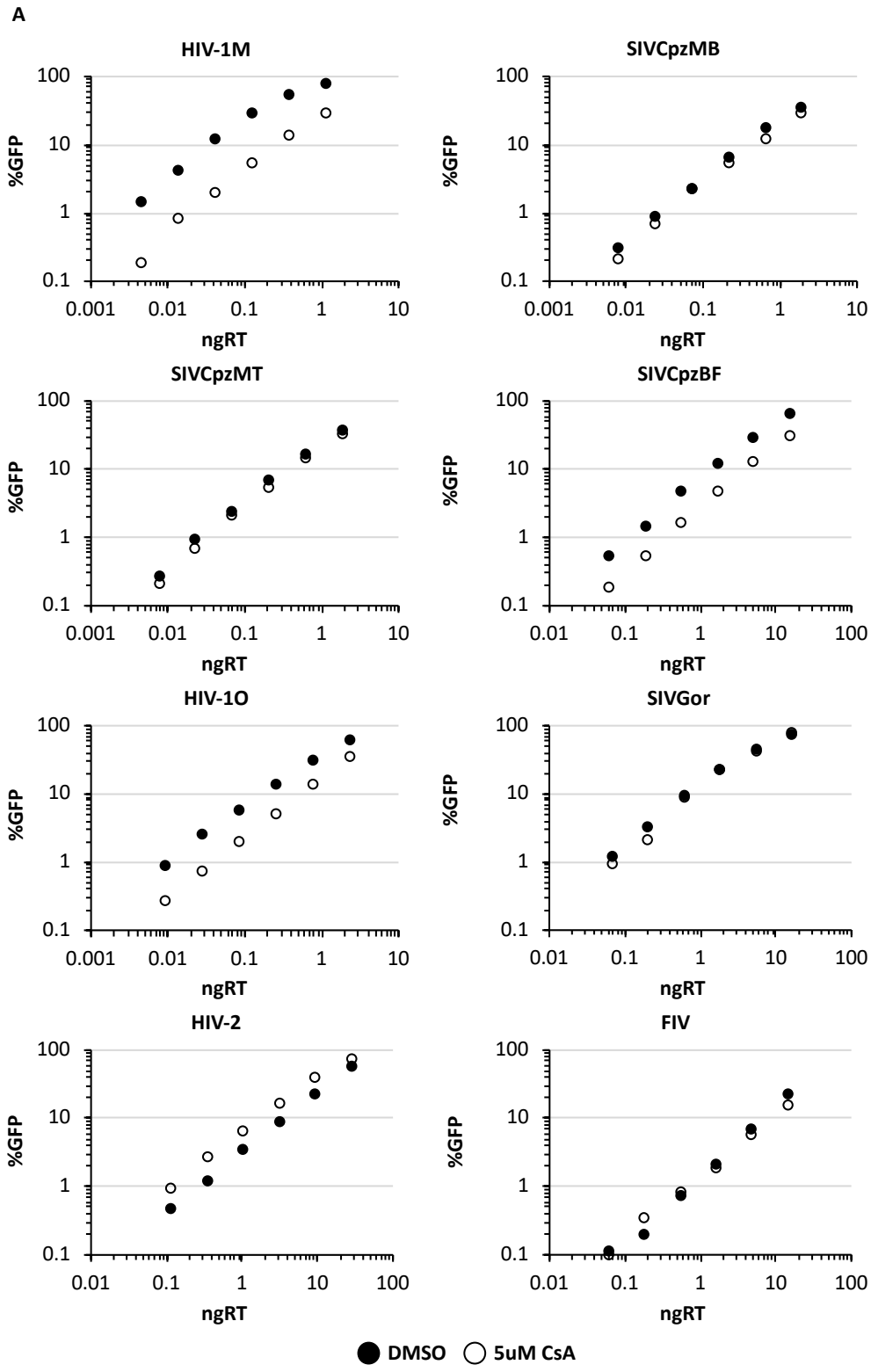
Figure 23: Binding CypA is not specific to HIV-1M
(A) Infectivity plots of VSV-G pseudotyped vectors derived from various lentiviruses as shown on CrFK cells expressing either empty vector (EXN Control) or TRIM5-CypA chimera (EXN TRIM-CypA) or TRIM5-Nup358 chimera (EXN TRIM Nup358). Infection determined by GFP expression (flow cytometry). Data representative of two biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2.

4.2.2 The CypA-capsid interaction is essential to HIV-1 viruses

Having shown that the panel of lentiviruses were able to bind CypA, we next sought to determine the effect of disrupting this interaction. CsA has been shown to chemically prevent CypA, but not NUP358, from interacting with the HIV-1M capsid [112]. Therefore 5µM CsA was added at time of infection of U87 cells with the lentiviral panel members and titre determined [**Figure 24**]. As expected HIV-1M had a 10-fold defect with the addition of CsA. However, contrary to previous work the western SIVCpz viruses, SIVCpzMB and SIVCpzMT, which are parental to HIV-1M [45], showed no defect in infetion in the presence of CsA [298] [**Figure 24B**]. Interestingly the eastern SIVCpz virus, SIVCpzBF, which has never been reported to transfer to humans, showed a consistent 3-fold defect on CsA addition. Furthermore, SIVGor, the precursor to HIV-1O [38], also showed no inhibition, whereas HIV-1O was inhibited by 2.5-fold on CsA addition. Neither of the non-primate lentiviruses, FIV and EIAV, showed any defect with CsA in U87 cells [**Figure 24B**].

Next, we wanted to determine at what stage in the life cycle CsA disrupts. It is already known that CsA acts post-entry, with HIV-1M production being unaffected when CsA is present [303]. Knowing this, we examined the likely post-entry step affected, reverse transcription. We infected U87 cells at an MOI of 0.1 and added 5µM CsA at the time of infection, and measured reverse transcription products and infection at 6 and 48 hours post transduction respectively [**Figure 25**]. Consistent with **Figure 24** HIV-1M was the most inhibited, with a 10-fold defect, on CsA addition, there was a small, but consistent 2-3 fold defect for SIVCpzBF and HIV-1O. Importantly, we found that for all the panel members, if there was a defect in infection it was fully explained by a defect of equal magnitude in reverse transcription [**Figure 25B**]. For example, HIV-1M had a 10 fold

defect in infection and a 10 fold defect in reverse transcription. Thus the complete defect is explained by the block in reverse transcription.



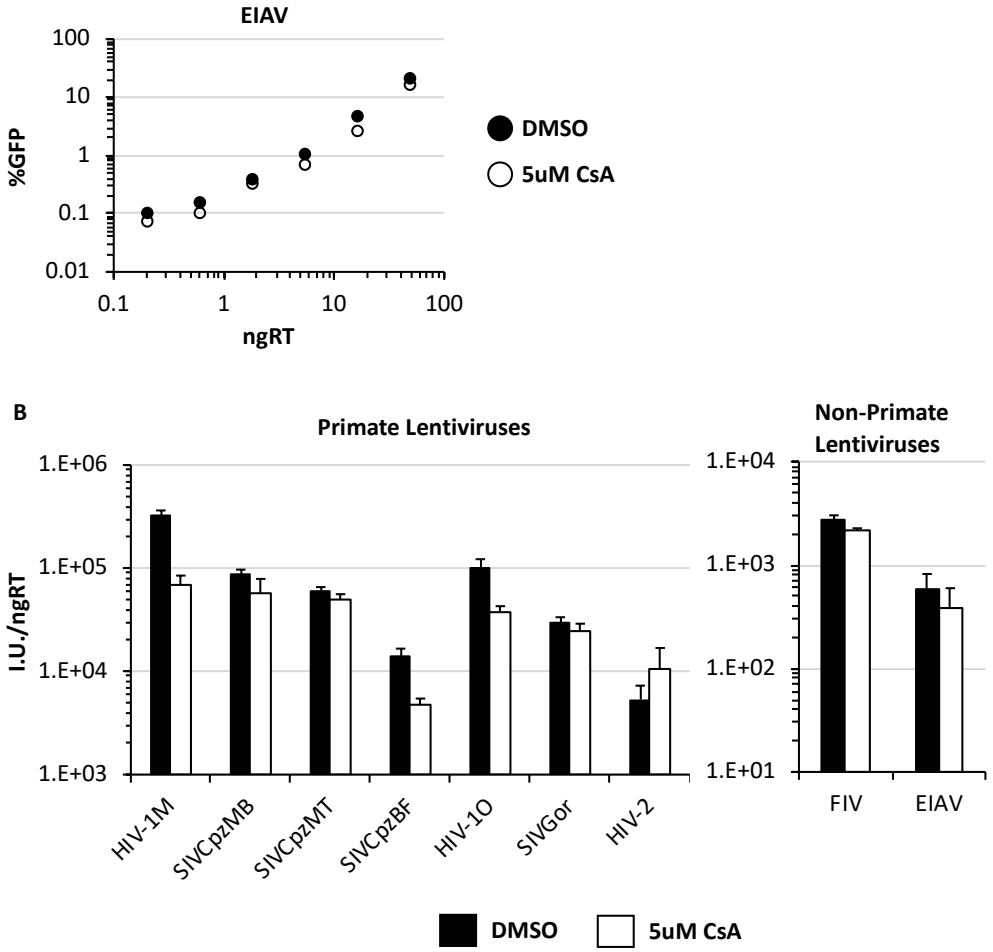


Figure 24: HIV-1M has the greatest dependency on CypA for infection of U87 cells

(A) Infectivity plots of VSV-G pseudotyped vectors, derived from lentiviruses, on U87 cells either in the presence or absence of 5µM CsA, added at the time of infection. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of three biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=3.

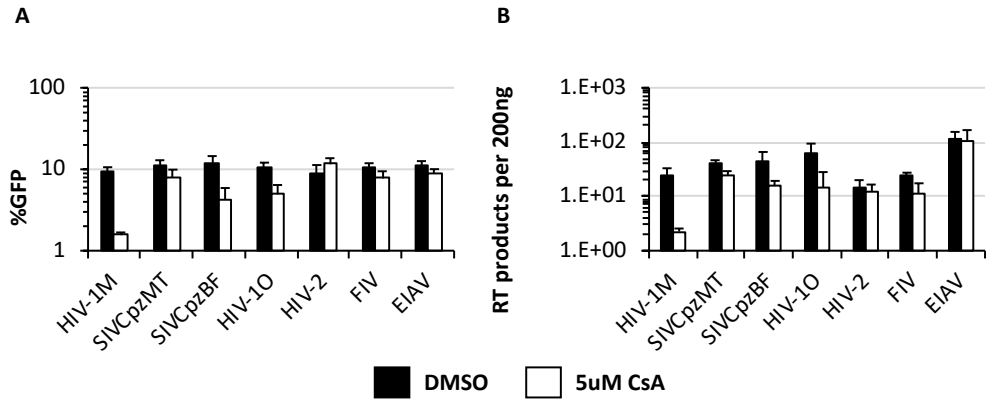


Figure 25: CsA causes a defect in infectivity by inhibiting reverse transcription

(A) GFP positive U87 cells after infection with single round VSV-G pseudotyped vectors at MOI 0.1 in the presence of 5µM CsA or DMSO. N=3, +/- SD. (B) Reverse transcription products 6 hours post infection with lentiviral vectors at MOI 0.1 from (A). N=3 with standard deviation shown with the error bars.

It has been reported that CsA prevents the HIV-1M capsid from interacting with CypB as well as CypA [283]. To determine the role that CypA plays in the other lentiviruses life cycles and to remove the confounding effect of other cyclophilins we produced a CypA CRISPR/Cas9 knockout U87 cell line. After single cell cloning, CypA knockout was confirmed by western blot [**Figure 26**], and the lentiviral panel members titrated comparing infection on the knockout and unmodified control U87 cells [**Figure 27**]. Interestingly, the results with the CypA knockout cells varied slightly from those using CsA. HIV-1M infection was as sensitive to CypA knockout, as it was to CsA treatment with a 10-fold defect in both. However, HIV-1O was insensitive to CypA Knockout despite being inhibited by CsA treatment. For SIVCpzBF, CypA knockout did not result in the level of inhibition observed with CsA treatment [**Figure 24** and **Figure 27B**]. These experiments raise the possibility that HIV-1O and SIVCpzBF could be dependent on different cyclophilins in U87 cells.

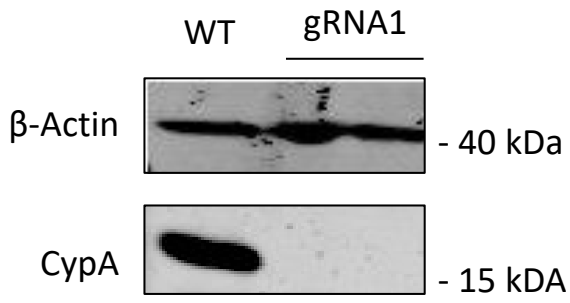


Figure 26: Knockout for CypA in U87 cells
Western blot detecting CypA in unmodified wild type (WT) U87 cells or in U87 after CRISPR/Cas9 knockout of CypA (by Lauren Harrison, Towers lab) and single cell cloning. Beta-Actin was detected as a loading control. Size marker positions are shown with molecular weights (KDa)

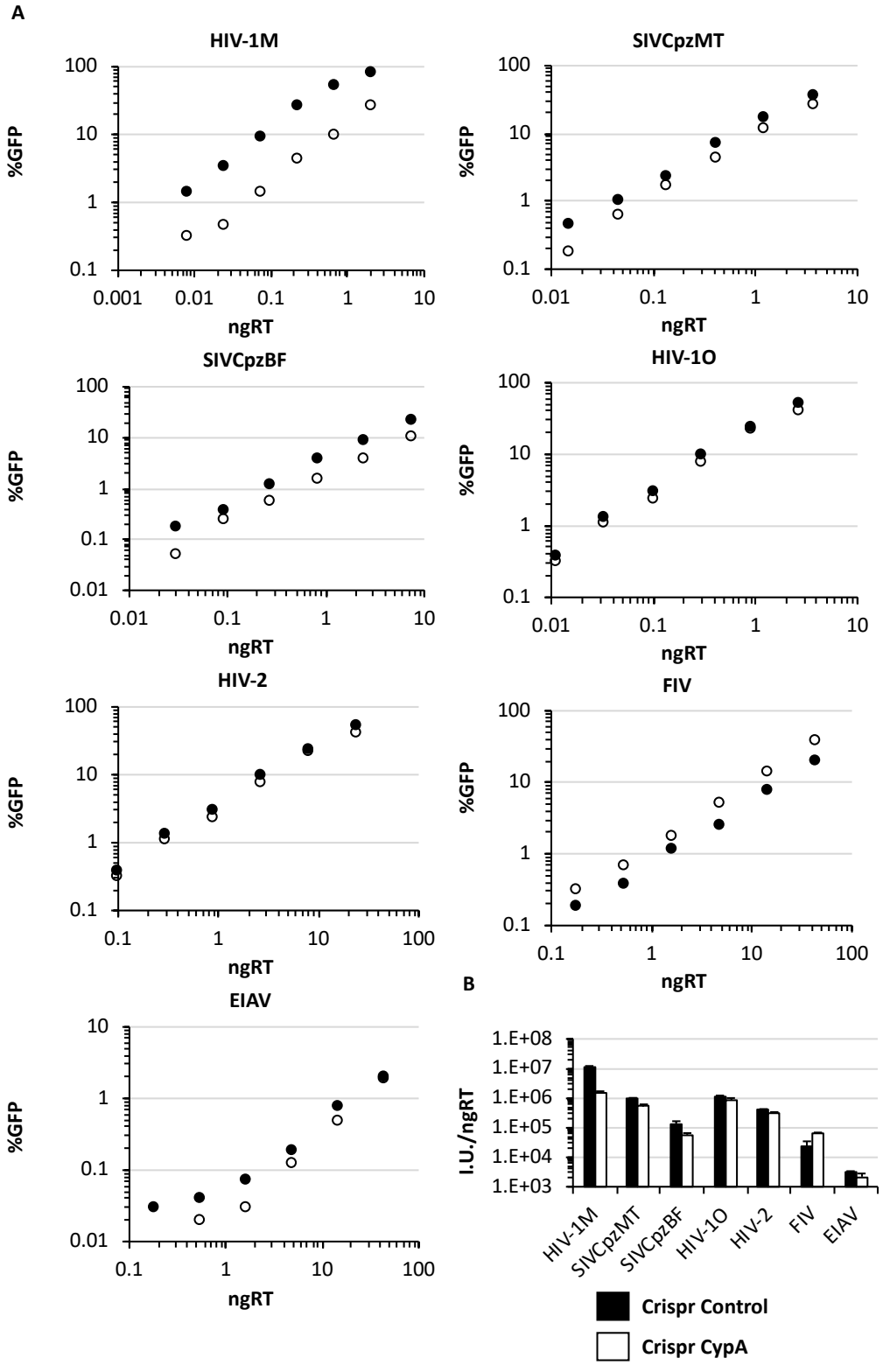


Figure 27: HIV-1M is most affected by Cripsr of CypA in U87 cells
(A) Infectivity plots of VSV-G pseudotyped vectors, derived from various lentiviruses, on either unmodified control U87 cells (Crispr Control) or on U87 cells with CypA knocked out (Crispr CypA). Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2.

Finally, we wanted to examine other drugs designed to prevent the Cyp-capsid interaction against the lentiviral panel. The Tower's lab has recently designed several analogues of CsA and a series of drugs of novel design, termed the depsins, to target the CypA-capsid interaction, some of which have altered CypA-binding specificities (unpublished data, Towers Lab). To determine if any of these new drugs had an increased viral range or potency compared to CsA, we tested a selection of the novel Cyp inhibitors against the virus panel in U87 cells [**Figure 28A**]. The U87 cells were selected because they give the biggest defect in infection for HIV-1M with CsA treatment compared to other cell lines (unpublished data, Towers lab). We infected U87 cells at an MOI of 0.1 and added the drugs at 5μM at the time of infection. Infection was measured at 48 hours post infection by flow cytometry for GFP expressing cells. All of the drugs tested were affective against HIV-1M, with 7-10 fold defect in infection [**Figure 28A],** however none had broader specificity than CsA. However, the despin JW3-38 was partly effective, 2 fold, at blocking infection of HIV-2. To ensure that the drugs were not toxic, we performed an MTT assay at 5μM with each of the drugs we used [**Figure 28B**]. At 5μM, none of the drugs tested showed any cytotoxic effect in U87 cells.

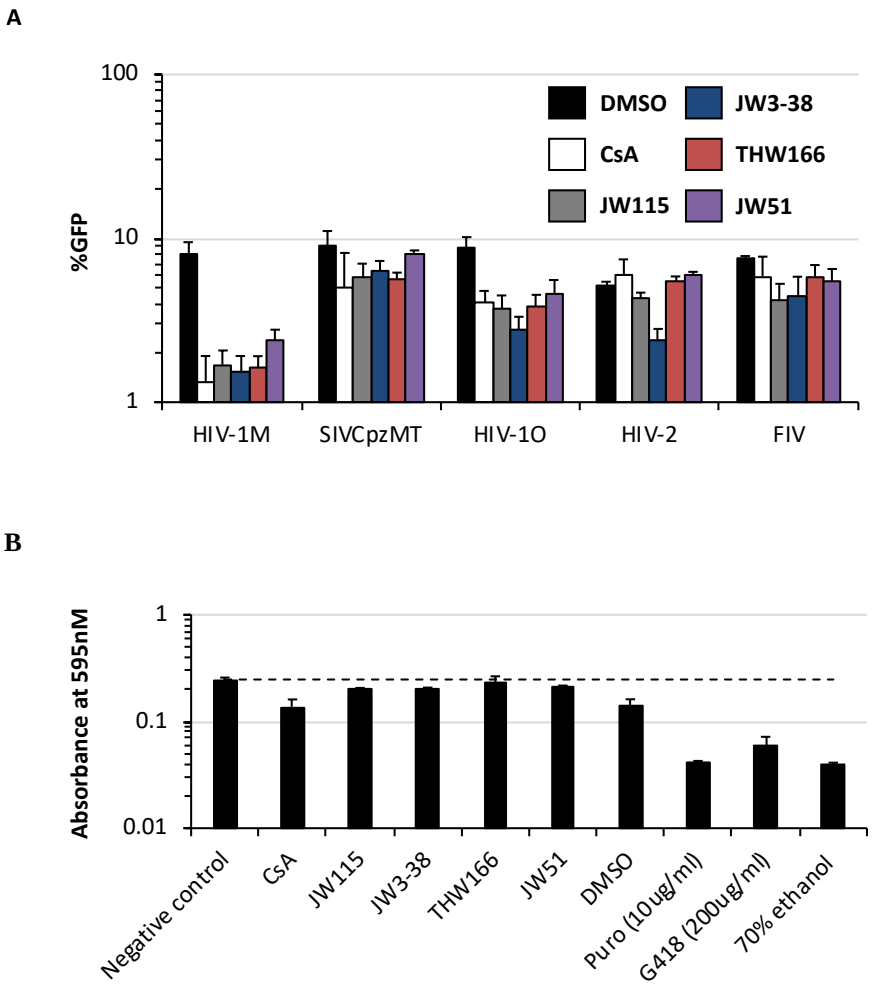


Figure 28: Novel drugs that target the CypA-capsid interaction do not affect CsA-resistant lentiviruses
(A) GFP positive U87 cells after infection with single round VSV-G psudeotyped lentiviral vectors at MOI of 0.1 in the presence of 5µM CsA, JW3-38, JW115, THW166 or JW51 or DMSO, added at time of infection. Error bars show +/- SD and N=2. (B) Shows a table of drug kinetics measured by Justin Warne using surface plasma resonance. The immunosuppressive nature of the drugs was determined by Lauren Harrison. (C) The cytotoxic effect of these drugs were examined at 5µM using an MTT assay. Absorbance was measured at 595nM 48 hours after addition of compound in wild type U87 cells.

4.3 SIVCpz is not dependent on CypA

4.3.1 SIVCpz has no defect in particle production in the presence of CsA

Interestingly we found that SIVCpz does not require CypA for single round infections. This is in contrast to Braaten et al [298], who showed that in a spreading infection within Jurkat cells a western SIVCpz, SIVCpzGab, was sensitive to CsA. The CsA induced

defect in replication was evident after 18-21 days post infection for SIVCpzGab, whereas CsA inhibited HIV-1M spread from day 10 onwards. This early study proposed that SIVCpz utilises CypA as a cofactor during infection. However, we have found that in single round infections of U87 cells, two western chimpanzee SIVCpz viruses show no defect to infection on the capsid-CypA is disrupted by CsA [**Figure 24** and **Figure 25**]. One possibility is that if the western SIVCpz had a defect in particle production caused by CsA it would explain the defect in spreading infection that isn't apparent in a single round experiment. For HIV-1M, it is known that target cells CypA is needed for optimal infection as the presence of CsA in infected producer cells does not alter HIV-1M infectivity [303]. However, no specific experiments have been done to test whether the CypA in producer cells is important for SIVCpz. To address whether CypA was required for particle production of SIVCpz, we produced SIVCpz in 293T cells in the presence of 5μM CsA, which was added at the time of transfection. Viral supernatants were harvested and virions were purified by ultra-centrifugation, to determine the level of packaged CypA by western blot [**Figure 29**]. This experiment aims to measure CypA incorporation into progeny virus particles. Consistent with previous work, CypA was found to be packaged into the viral particles of both western and eastern SIVCpz [296] [**Figure 29A**]. We found that with HIV-1M and both western and eastern SIVCpz viruses there was ~50-70% reduction in CypA present in viral particles produced in the presence of 5μM CsA, quantified by ImageJ, but no significant decrease in capsid present within viral supernatants [**Figure 29A**]. Having shown that CsA treatment of the 293T producer cells reduces incorporated CypA but not capsid release for both HIV-1M and SIVCpz, we confirmed there was no defect in particle release for HIV-1M or SIVCpz by measuring reverse transcriptase activity in viral supernatants. We found that there was no significant decrease in reverse transcriptase activity in viral supernatants for viruses produced in the presence of 5μM CsA or DMSO [**Figure 29D**]. We then measured the infectivity of the

viruses produced in the presence of CsA [**Figure 29B**]. As has previously been shown using HIV-1M, we found no significant difference in infectivity between SIVCpz produced in the presence of CsA or DMSO [**Figure 29C**]. Taken together, these data suggest that SIVCpz is similar to HIV-1M in that disrupting the CypA-capsid interaction in the producer cells does not impact particle production or infectivity in the target cells.

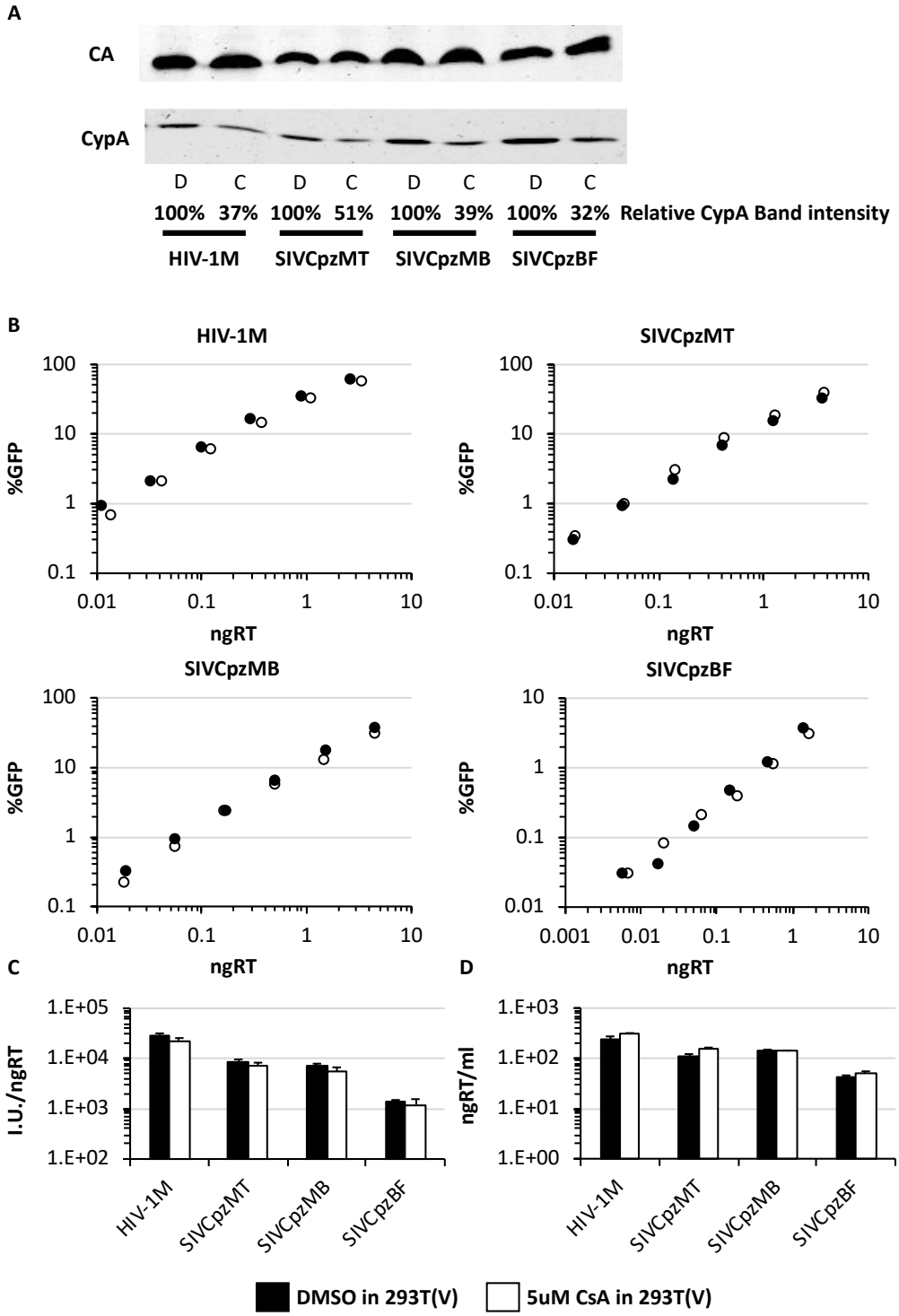


Figure 29: CsA does not affect particle production for SIVCpz
(A) Virions produced in the presence or absence of 5µM CsA were purified and subjected to western blot detecting capsid or CypA. (B) Infectivity plots of VSV-G pseudotyped vectors derived from various lentiviruses as shown on U87 cells, in the presence or absence or 5µM CsA. Representative of two biological repeats. (C) Titres determined from data in (B) from three doses of vector with MOI below 0.3. N=2. +/- SD (D) Reverse transcriptase activity was measured in viral supernatants, N=1. Error bars showing standard deviation.

4.3.3 SIVCpz use the same nuclear import pathway as HIV-1M

Having shown that in single round infection both the western SIVCpz were resistant to CsA and CypA knockout and eastern SIVCpz only had a small defect with both [**Figure 25** and **Figure 27**] we inferred that their replication is largely independent of CypA. We next wanted to determine whether SIVCpz infection was independent of other HIV-1M cofactors. It is known that if the CypA-capsid interaction of HIV-1M is disrupted, chemically with CsA or genetically with the capsid mutation P90A, infection becomes less dependent on the cofactors NUP358 and TNPO3 [112]. That is, CsA treatment rescues HIV-1M infection from the inhibition caused by depleting Nup358 or TNPO3 [112]. Alteration of the HIV-1M nuclear import pathway by CsA causes alteration in HIV-1M integration site selection [112]. Recruitment of CypA is therefore thought to direct HIV-1M into a particular nuclear import pathway and thus site of integration [112]. Hence, we sought to determine whether SIVCpz, a virus whose infection is naturally independent of CypA in single round infections but can still bind CypA, phenocopies HIV-1M P90A and is also independent of Nup358 and TNPO3. To address this we used previously described HeLa cells in which Nup358 or TNPO3 have been stably depleted [112]. We measured the permissivity of these depleted cells to single round infection with SIVCpz [**Figure 30**]. As has been previously reported, when NUP358 or TNPO3 are depleted, wild type HIV-1M infectivity is reduced by 10 fold [**Figure 30A**]. Unexpectedly, both SIVCpzMT and SIVCpzBF were sensitive to depletion of both NUP358 and TNPO3 [**Figure 30B** and **Figure 30C**] showing ~10 fold defect in infectivity. This is consistent with the observation that both eastern and western SIVCpz can functionally bind the NUP358 Cyp like domain in a TRIMCyp restriction assay [**Figure 23**]. Taken together, these data suggest that CypA does not play a role in nuclear import pathway selection in SIVCpz. These data also suggest that nuclear import pathway usage is conserved between SIVCpz and HIV-1M and is thus a more ancient feature of the lifecycle whereas usage of CypA

by HIV-1M may have more recently evolved and could represent an adaption to infect humans.

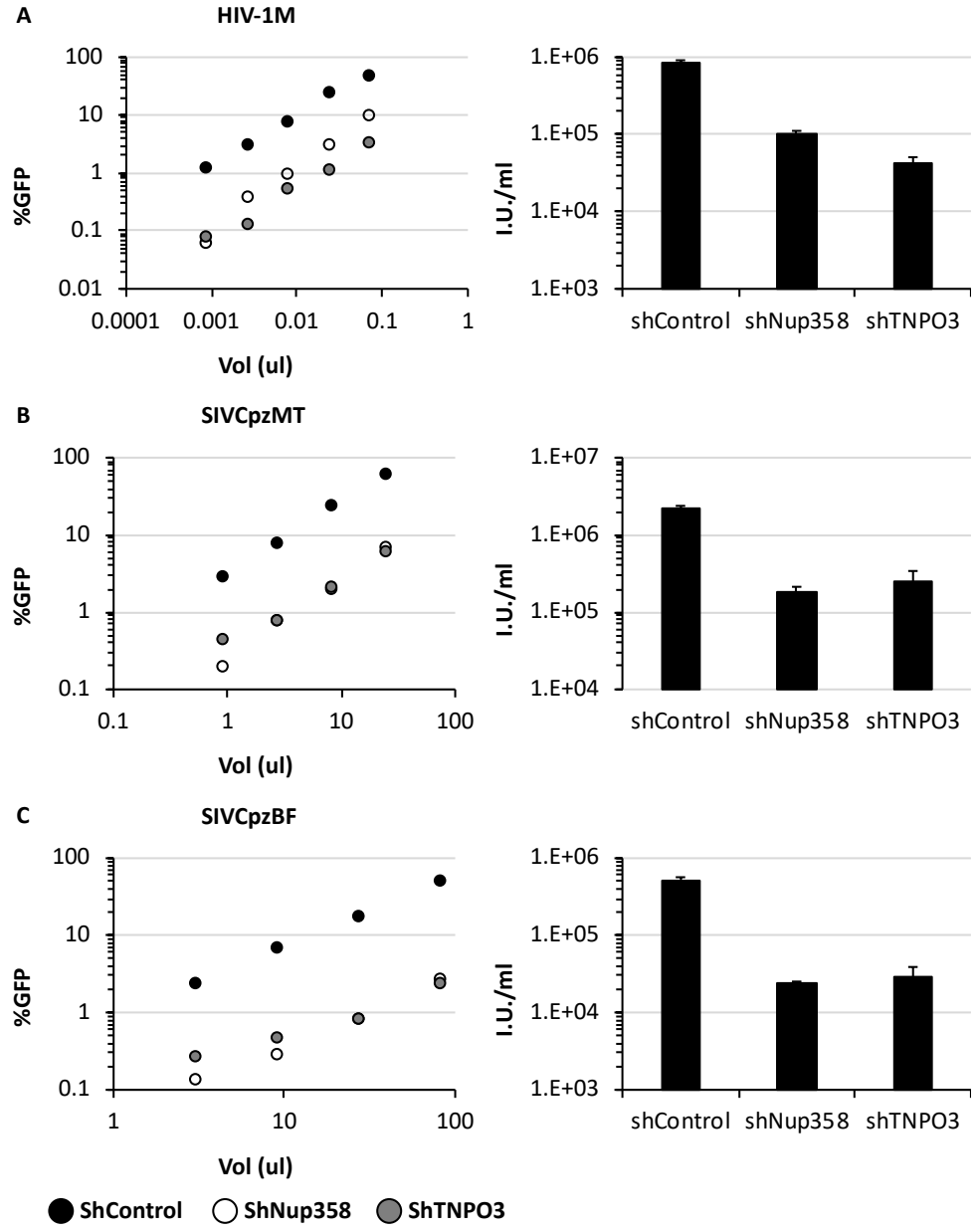


Figure 30: SIVCpz has similar dependency on Nup358 and TNPO3 as HIV-1M
(Left hand) Infectivity plots of VSV-G pseudotyped vectors derived from lentiviruses bearing Gag from HIV-1M (A), western SIVCpz (B) or eastern SIVCpz (C) as shown on wild type HeLa cells (ShControl), NUP358 depleted HeLa cells (ShNup358) or TNPO3 depleted HeLa cells (ShTNPO3). (Right hand) Vector titres determined from data in infectivity plots (left hand side) from three doses of vector with MOI below 0.3. N=1. +/-SD.

4.4 Mutations altering the position of the capsid N-terminal beta-hairpin affect HIV-1M CypA dependency

4.4.1 N-terminal beta-hairpin position regulates CypA usage

Recently it has been shown by x-ray crystallography that the NTD beta hairpin of the HIV-1M capsid hexamer is dynamic and can adopted at least two different conformations. These were termed open and closed because in the open conformation the electrostatic channel is open to the cytoplasmic environment, whereas in a closed conformation the NTD beta hairpin covers the channel opening with the arginine electrostatic core concealed beneath it [**Figure 31**] [237]. We hypothesise that in solution the NTD beta hairpin is dynamic and is regulated in an allosteric manner by cofactor interactions which shift the channel from open to closed, and vice versa. We propose that the beta-hairpin position may depend on which constellation of cofactors the capsid is interacting with at each stage within the life cycle. We sought evidence that CypA influences capsid conformation to a more closed or open state. To examine this we used capsid mutants, collectively termed the channel mutants, which have been established to shift the NTD beta hairpin into a more closed, H12Y [**Figure 31**], or open, Q50Y and Q50Y 120R [David Jacques, unpublished], conformation at least in X-ray crystal structures. We investigated the effect of disrupting the CypA-capsid interaction on the infectivity of these mutants in U87 cells using CsA. Cells were infected with the channel mutants alongside HIV-1O, which is also found to be in an open conformation by X-ray crystallography of capsid hexamers [David Jacques, unpublished], in the presence or absence of 5μM CsA [**Figure 32**]. We found that the position of the NTD beta hairpin determined whether HIV-1M was sensitive to CsA in U87 cells. The closed mutant, H12Y, responded in a similar way to wild type HIV-1M with a 10-fold decrease in titre on CsA addition. However, the two open channel mutants Q50Y and Q50Y 120R were insensitive to CsA, a phenotype similar to that of HIV-1O.

Next, we wanted to determine whether the defect observed for H12Y was the same as wild type HIV-1M which was found to be at the stage of reverse transcription [**Figure 25**]. To examine this, we infected U87 cells at an MOI of 0.1 and added 5μM of CsA at the time of infection. We quantified infection and reverse transcription products at 48 and 6 hours respectively [**Figure 33**]. As observed with the titrations, H12Y was the only channel mutant to be sensitive to CsA, this defect was found to be due to a decrease, of equal magnitude to infection, in reverse transcription products [**Figure 33B**]. The defect observed for reverse transcription in H12Y, like wild type HIV-1M, completely explained the loss of infectivity. Taken together, these data suggest that CypA may play a role in the allosteric regulation of the NTD beta hairpin and thereby reverse transcription.

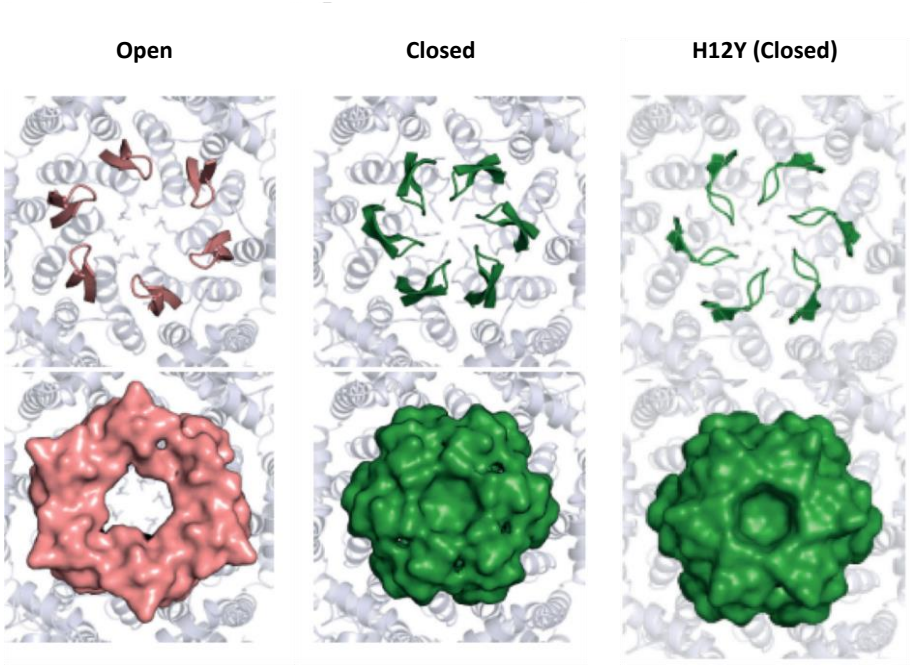


Figure 31: The N-terminal beta hairpin of the HIV-1M capsid hexamer can adopt at least two conformations
X-ray crystal structures of the capsid N-terminal beta hairpin of HIV-1M under different crystallisation conditions and X-ray crystal structure of the HIV-1M capsid bearing the H12Y mutation. Taken and adapted from D. Jacques. (2016) et al. [10].

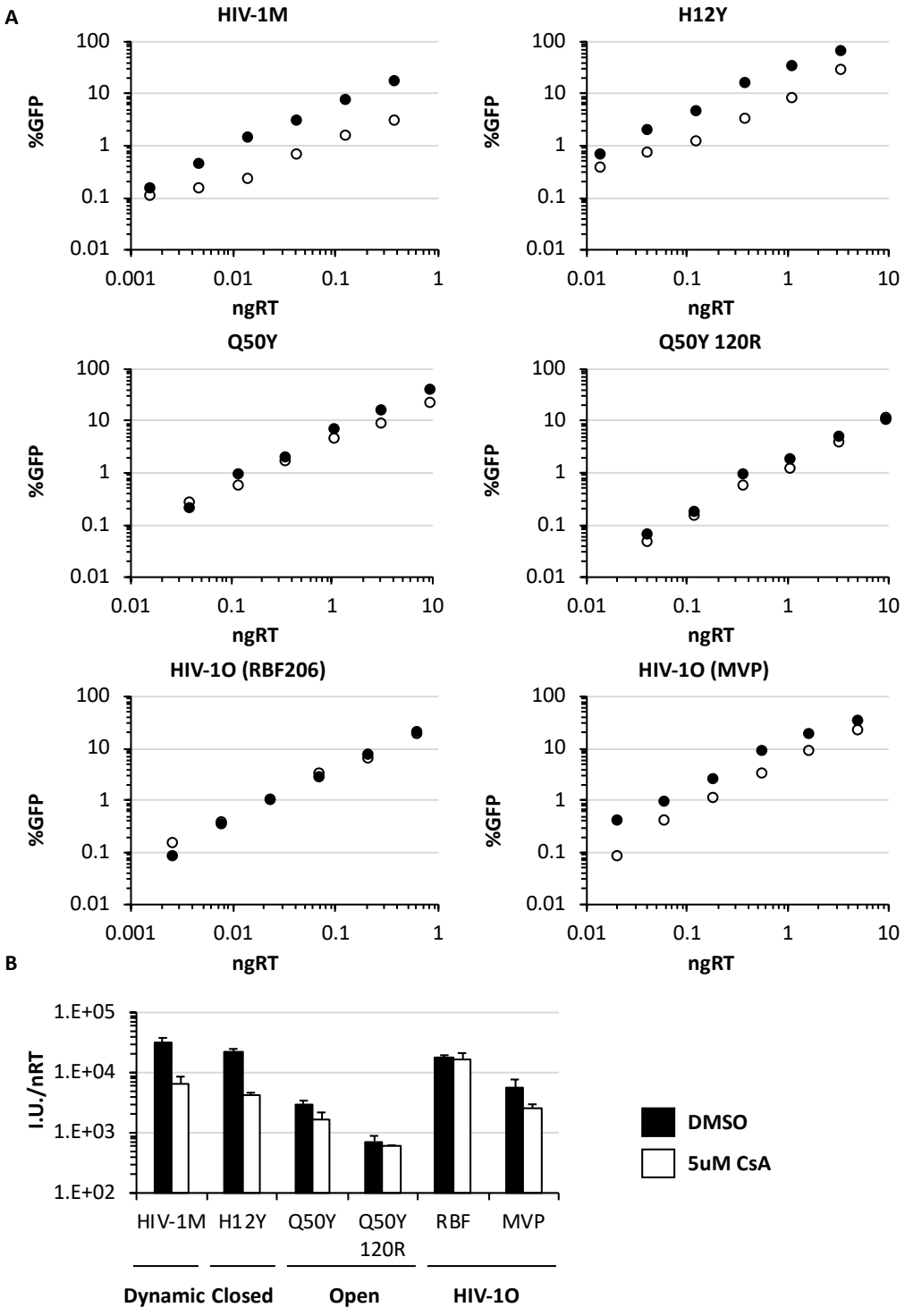


Figure 32: The position of the NTD beta-hairpin regulates CypA usage in HIV-1M

(A) Infectivity plots of VSV-G pseudotyped HIV-1M or HIV-1O derived vectors as shown on U87 cells in the presence or absence of 5μM CsA, added at time of infection. HIV-1M vectors either had wild type capsid (HIV-1M), a closed capsid (H12Y) or an open capsid channel (Q50Y and Q50Y 120R). Two different HIV-1O strains were used (RBF and MVP) both containing wild type capsid. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of four biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=4.

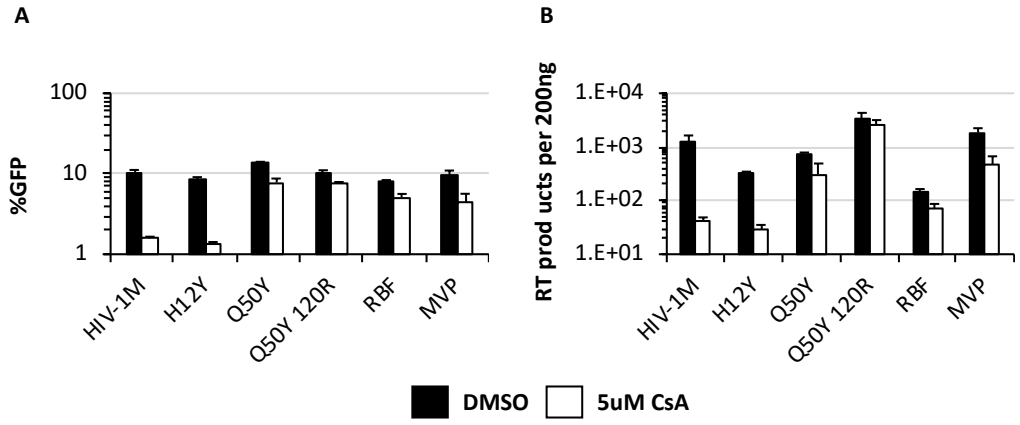


Figure 33: The CsA defect for the closed N-terminal beta-hairpin capsid mutant maps to reverse transcription
(A) GFP positive U87 cells after infection with VSV-G psudeotyped HIV-1M or HIV-1O derived vectors at MOI 0.1 in the presences or absence of 5μM CsA. HIV-1M vectors either had wild type capsid (HIV-1M), a closed capsid channel (H12Y) or an open capsid channel (Q50Y and Q50Y 120R). HIV-1O vectors contain wild type capsid. Infection determined by GFP expression (flow cytometry) 48 hours post infection. N=3, +/- SD. (B) Reverse transcription products measured 6 hours post infection, from infections shown in (A). N=3 and error bars showing standard deviation.

4.4.2 The ability to bind IP6 does not alter CypA dependency in U87 cells for HIV-1M

Based on the finding that there is a relationship between the position of the NTD beta hairpin of the HIV-1M capsid and CypA dependency, we hypothesized that CypA may play a role in the regulation of the electrostatic channel by inducing allosteric changes at the NTD beta hairpin. We wanted to determine if there was a relationship between the ability to bind IP6 and CypA dependency for HIV-1M. We hypothesised that CypA may open the electrostatic channel. We reasoned that with the open mutants, the channel is already open, so does not need to be opened leading to CypA independence. However, the closed mutant still needs CypA to open the channel and thus remained sensitive to CsA. We hypothesise that when the capsid is in an open state the Arg 18 ring is able to bind and transport nucleotides into the interior of the capsid to fuel reverse transcription and to regulate stability and uncoating through binding of IP6, with these processes being

regulated by capsid binding cofactors. However, closed capsid mutants may trap the capsid in a closed state with reduced dNTP import, hence poor DNA synthesis and reduced infectivity [237].

To examine this model further we used the capsid mutant R18G, which does not recruit IP6, with the mixed capsid viruses established in chapter 3 [**Figure 11**], and measured CsA sensitivity. As a control, we performed the same experiment with capsid mutant P90A, which has been shown to have a greatly reduced affinity for CypA and thus resistance to CsA [112]. We reasoned that if the IP6 free R18G mixed capsid viruses are resistant to CsA, this could be evidence for a relationship between ability to bind IP6 and CypA usage. U87 cells were infected at MOI 0.1 with either the R18G or the P90A mixed capsid viruses in the presence of 5μM CsA [**Figure 34**]. We observed that as the ratio of Gly increased at position 18, there was no significant change in CsA sensitivity [**Figure 34A**]. Whereas, as the ratio of Ala increased a position 90 there was a step-wise decrease in sensitivity to CsA [**Figure 34B**]. To confirm this observation, we repeated the experiment in the U87 CypA knockout cells [**Figure 35**]. Again, we observed that there was no significant change in the infectivity defect for the R18G mixed capsid viruses as the ratio of Gly increases at position 18 within the CypA knockout U87 cells [**Figure 35**A]. Whereas the P90A mixed capsid viruses showed a stepwise decrease in the infectivity defect as the ratio of Ala increased a position 90 [**Figure 35**B]. From this we can conclude that the ability to bind IP6 does not influence CypA usage in HIV-1M or vice versa.

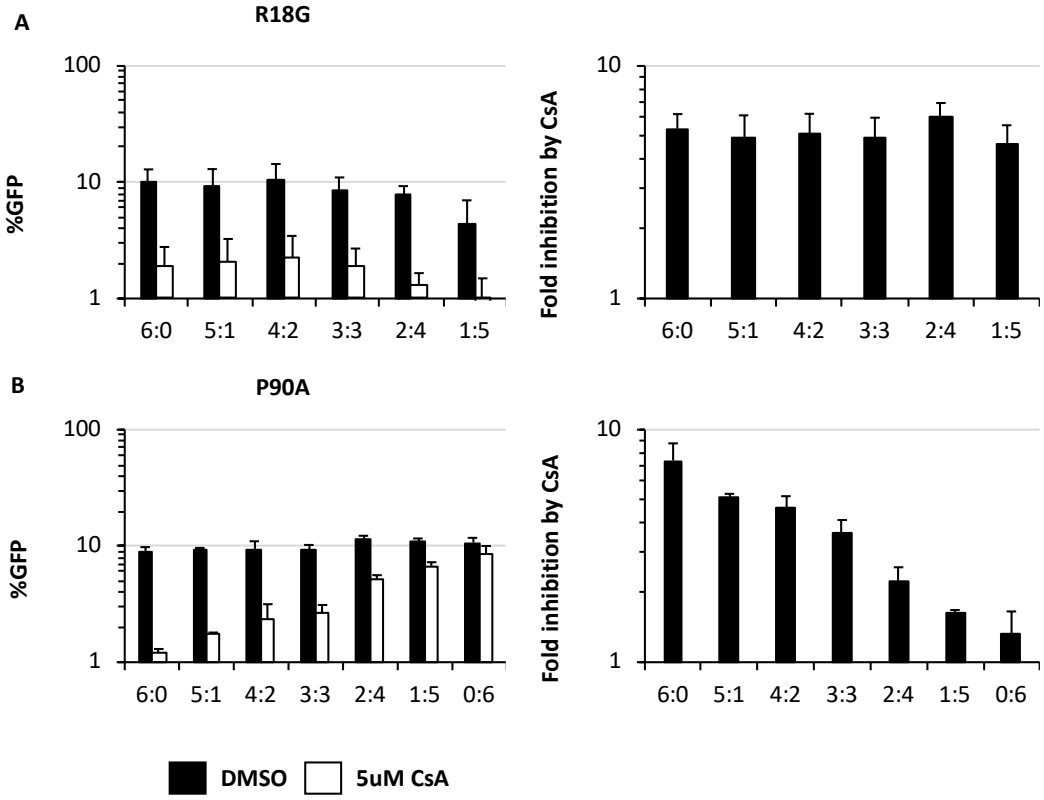


Figure 34: CsA sensitivity is independent of the ability to bind IP6 in HIV-1M
(Left hand) GFP positive U87 cells after infection with VSV-G psudeotyped mixed capsid vectors derived from HIV-1M at an MOI of 0.1 in the presence or absence of 5μM CsA. Mixed capsid vectors either contained wild type and R18G containing capsid (A) or wild type and P90A containing capsid (B). Infection determined by GFP expression (flow cytometry) 48 hours post infection. N=2, +/- SD. (Right hand) CsA sensitivity as expressed as fold inhibition to infection determined from (A). N=2 and error bars showing standard deviation.

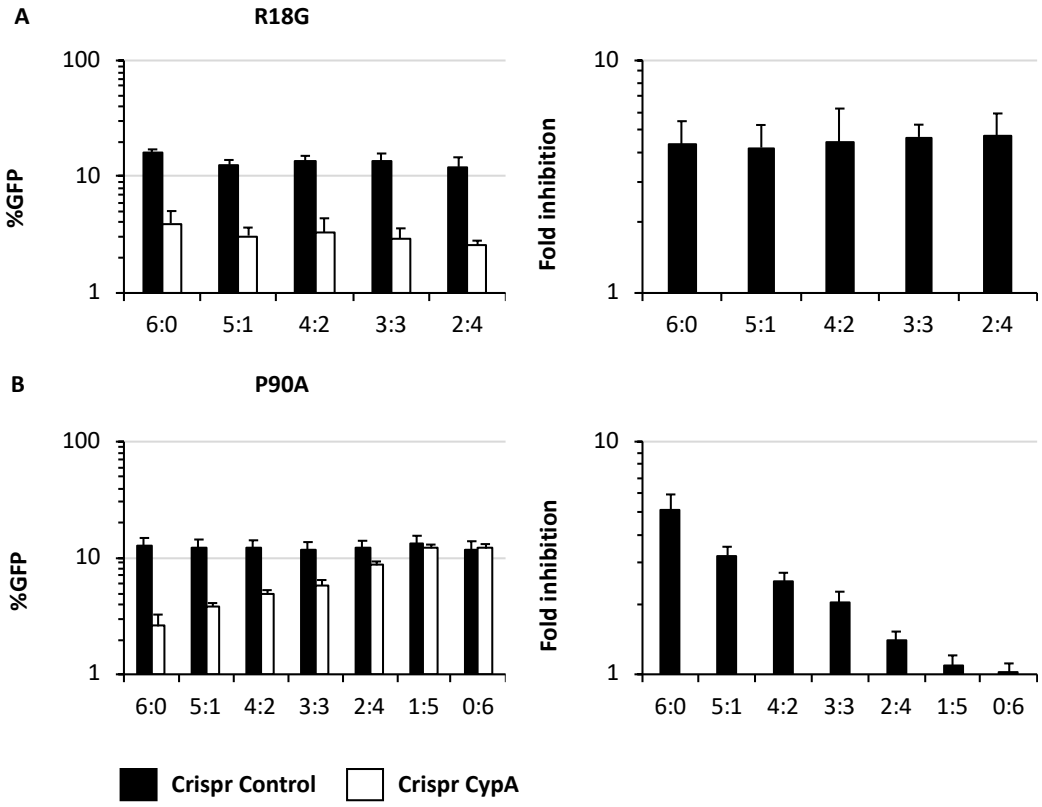


Figure 35: The ability to bind IP6 is independent of CypA usage in U87 cells
(Left hand) GFP positive wild type U87 (Crispr Control) or CypA knock out U87 (Crispr CypA) cells after infection with VSV-G psudeotyped mixed capsid vectors derived from HIV-1M at an MOI of 0.1. Mixed capsid vectors either contained wild type and R18G containing capsid (A) or wild type and P90A containing capsid (B). Infection determined by GFP expression (flow cytometry) 48 hours post infection. N=2, +/- SD. (Right hand) CypA dependency as expressed as fold inhibition to infection determined from (A). N=2 and error bars showing standard deviation.

4.5 Discussion

4.5.1 The role of CypA in the lentiviral HIV-1M life cycle

The exact role of CypA in the HIV-1M life cycle is still an area of active research. However, it is clear from previous work and this study that the CypA-capsid interaction is essential for optimal infection for HIV-1M. We have shown that if the CypA-capsid interaction is disrupted in the U87 cell line, chemically via CsA or via knocking out CypA, HIV-1M has a 10 fold defect in infection, which is explained by a matching defect in reverse transcription [**Figure 25** and **Figure 27**]. It is also known that in primary macrophages, disruption of the CypA-capsid interaction leads to a restrictive innate immune response [230]. Likewise, in primary T cells when the CypA-capsid interaction for

HIV-1M is disrupted the defect to infection maps to reverse transcription [Unpublished work, Towers lab]. However it is difficult to determine if the defect in primary T cells is accompanied by an innate immune response due to T cells needing to be activated before they can be infected. Here we propose a model in that the binding of CypA leads to allosteric changes within the HIV-1M capsid hexamer, which influences the N-terminal beta-hairpin. Mutagenesis of residue Asp 50 to a Tyr alters the capsid dynamics to favour a more open conformation, leading to loss of CsA sensitivity in U87 cells [**Figure 32** and **Figure 33**]. However, mutagenesis to force the capsid into a more closed conformation, H12Y, does not alter CsA sensitivity [**Figure 32** and **Figure 33**]. Taken together we can infer two properties of the HIV-1M capsid. Firstly, the hexamer must function as an allosteric complex as changes distal to the cyclophilin binding loop alter capsid behaviour, in this case loss of CsA sensitivity with the Q50Y and Q50Y 120R capsid mutants. Secondly, CypA functions to open the NTD beta-hairpin thus exposing the electrostatic channel to the cytoplasmic environment. We know that the N-terminal beta-hairpin of the HIV-M capsid is able to exist in an open or closed state by x-ray crystallography depending on crystallisation conditions [237] [**Figure 31**] and we hypothesise that in solution the NTD beta hairpin is dynamic able to shift between open and closed depending on what is bound to the capsid. In the capsid mutants Q50Y and Q50Y 120R the channel is already forced open, into a more HIV-1O like state, and so does not need to be opened, potentially by CypA, thus disruption of this interaction is not detrimental to infection. However, when the capsid channel is closed, with H12Y, CypA is required for capsid channel opening and thus does have a defect when this interaction is prevented chemically with CsA. It is possible that when the capsid is in more of a closed state the NTD beta-hairpin might trap IP6 in complex with the Arg 18 ring to maximise capsid stability and prevent premature uncoating and thus degradation of the viral genome by cytoplasmic nucleases and innate immune sensing. Closure of the electrostatic channel

may also prevent dNTP influx by blocking the electrostatic channel and delay reverse transcription, which again would aid in maintaining capsid stability as it is known that as reverse transcription progresses internal stresses and pressure increases within the capsid core of HIV-1M which promotes uncoating [78]. In support of this, H12Y is known to have delayed reverse transcription kinetics compared with wild type HIV-1M but is still able to produce full length reverse transcription products [unpublished work from the Towers Lab]. This reduced speed of reverse transcription could be due to decreased influx of dNTPs into the capsid core.

This proposed role of CypA in allosterically regulating capsid stability has implications for HIV-1M replication in innate immune competent cells, such as macrophages, where reverse transcription, and thus uncoating would need to be carefully controlled to avoid triggering and innate immune response. When the electrostatic channel is in an open conformation, reverse transcription could be able to progress at a faster rate and IP6 could freely dissociate from the capsid. Both loss of IP6 and more rapid reverse transcription may lead to faster or premature uncoating in the cytoplasm. HIV-1O whose electrostatic channel can only exist in an open state, as shown by x-ray crystallography in conditions where the HIV-1M beta hairpin is in a closed conformation [private communication from David Jacques and Leo James], provides additional evidence. It is known that HIV-1O and both the open channel mutants in the HIV-1M capsid induce a restrictive innate immune response in macrophages [unpublished data from the Towers lab]. We hypothesise that this innate immune response is due to unregulated reverse transcription and thus premature uncoating in the cytoplasm.

However, it is also possible that the loss of CsA sensitivity observed when the HIV-1M capsid is pushed to a more open confirmation [**Figure 32** and **Figure 33**] can be explained by TRIM5 mediated restriction. Human TRIM5 has been shown to potently

restrict non-human retroviruses such as FIV, EIAV and N-tropic MLV whereas other primate TRIM5's have been shown to restrict HIV-1M. previously, it was thought that human TRIM5 failed to bind HIV-1M and thus could not restrict infection. However, It has recently been shown that CypA binding to the HIV-1M capsid is able to protect the virus from TRIM5 binding and thus restriction and when the CypA-capsid interaction is lost HIV-1M is potentially restricted by human TRIM5 [257,258]. It is possible that the CypA does not induce allosteric changes within the capsid NTD beta hairpin rather the binding of CypA to the capsid is dependent on the conformation of the capsid. The open channel mutants Q50Y and Q50Y 120R are insensitive to CsA which suggests that their replication is independent of CypA. However, it remains unclear whether Q50Y and Q50Y 120R are still able to bind CypA, so it isn't known whether binding is prevented or the consequence of binding is. If the open channel capsid mutants are no longer able to bind CypA, this would support a TRIM5 explanation where the restriction in macrophage is due to a TRIM5 induced NFKB/AP1 [263] signalling and subsequent degradation of viral particles. If, however, the open channel mutants are still able to bind CypA then it would suggest a lack of allosteric control and premature uncoating hypothesis. If CypA was still able to bind the open channel mutants, then TRIM5 would be unable to bind and thus unable to restrict viral infection so the observed immune response within macrophages would likely be due to cytoplasmic uncoating and sensing of viral DNA.

In summary, HIV-1M has evolved to utilise CypA, whereas many other lentiviruses have little or no defect with CsA [**Figure 24**]. Therefore, we hypotheses that usage of CypA is to help control capsid stability, uncoating, reverse transcription and TRIM5 restriction likely to aid infection of immune competent cell and might represent a feature that is unique to HIV-1M which has important implications for tropism.

4.5.2 SIVCpz adaption to humans

The parental virus of HIV-1M is SIVCpz [45]. It is thought that during the early 1920s the chimpanzee virus crossed the species barrier and entered the human population in central Africa. Our data shows that the SIVCpz of western chimps is not dependent on CypA in single round infection in contrast to HIV-1M [**Figure 24** and **Figure 25**]. One hypothesis for this is that during the adaption to human infection, HIV-1M gained the dependence on CypA. This hypothesis starts to gain momentum when the transmission routes are considered for HIV-1M and SIVCpz. HIV-1M is predominantly a sexually transmitted virus however it is thought that SIVCpz is mainly transmitted during fighting. These two different routes of transmission mean the environment and cell type first encountered by the viruses are vastly different. During sexual transmission HIV-1M encounters a mucosal membrane which will contain resident macrophages, dendritic cells and other antigen presenting cells which are competent for sensing and triggering an immune response to HIV-1M infection. It is well known that HIV-1M is able to replicate in macrophages but if the CypA-capsid interaction is disrupted the virus is sensed and infection is blocked [230]. Once integrated into a macrophage or other cell present at the mucosal membrane, HIV-1M would be able to infect T-cells as the macrophage circulates. For SIVCpz however, infection during fighting is likely via blood-blood contact. This route of transmission allows the virus to directly enter the blood stream and avoid mucosal membranes and thus immune sensing cells such as macrophages, dendritic cells and other resident immune cells. It is likely, that infection via blood-blood contact leads to infection of a T-cell first whereas infection via mucosal membrane is likely a sensing competent cells. As T-cells are thought to not sense infection with HIV-1M it likely they don’t sense infection with SIVCpz.

Therefore, we propose a model where infection across a mucosal membrane needs additional layer of immunoevasion, which is provided by CypA, to establish infection in immune sensing cells. As we propose that HIV-1M has evolved to use CypA to avoid immune detection which allowed it to expand its tropism to myeloid cells and perhaps help to drive pandemicity. By avoiding triggering in macrophages HIV-1M may transmit more efficiently as an STI, efficiently infecting cells at mucosal membranes.

One question which remains is why the eastern SIVCpz, which has a small defect in infection with CsA treatment [**Figure 24** and **Figure 25**] and a similar level of contact with humans via bushmeat as western SIVCpz [36], has not transmitted to humans. One possibility is that the eastern SIVCpz lacks the ability to overcome other challenges when infecting macrophages that the western SIVCpz is able to. For example, eastern SIVCpz might be unable to overcome the low dNTP environment created by SAMHD1 [362]. It is likely that some yet undetermined feature of eastern SIVCpz lifecycle has prevented its jump to humans.

4.5.3 SIVCpz and nuclear import pathway selection

It is known that HIV-1M uses cellular cofactors, such as CypA, to help direct its nuclear import and subsequent integration site selection [112] .When the CypA-capsid interaction is prevented for HIV-1M, it is known that integration occurs in genomic regions of higher gene density [112] and infection becomes independent of the cofactors Nup358 and TNPO3 [112]. It is therefore interesting to observe that a virus whose infection is naturally independent of CypA, the western SIVCpz viruses, are still dependent on Nup358 and TNPO3 [**Figure 30**]. From this we know that HIV-1M is able to use alternative nuclear import pathway, which potentially SIVCpz is unable to use and is yet uncharacterised. Furthermore, we have shown that both eastern and western SIVCpz viruses are able to

bind CypA [**Figure 23** and **Figure 29**] however the consequence of binding is different to that of HIV-1M [**Figure 24**]. We hypothesis that along the evolutionary branch leading to HIV-1M from western SIVCpz, HIV-1M acquired new abilities to regulate its lifecycle via CypA interactions. For example, it's possible that SIVCpz is unable to allosterically regulate its electrostatic channel, like HIV-1M can but still uses CypA to help direct nuclear import pathway selection.

4.5.4 FIV and human TRIM5

FIV represents an interesting exception to the norm when it comes to CypA usage and CsA sensitivity in U87 cells. We have shown that FIV is able to bind CypA [**Figure 23**] but does not have a defect in infection, or reverse transcription, when the CypA-capsid interaction is prevent via CsA addition [**Figure 24** and **Figure 25**]. This however is not consistent with previous work showing that in jurkat T cells, FIV is sensitive to CsA to a similar extent as HIV-1M in single round infection [299]. Furthermore, it is known that FIV is restricted by human TRIM5 when it is expressed in CrFK cells [259,260]. As mentioned it has recently been shown that CypA binds to the HIV-1M capsid and prevents human TRIM5 mediated restriction, it is therefore possible the varying CsA sensitivity of FIV could be related to differential TRIM5 expression in cell lines. Assuming that CypA may also protect FIV from TRIM5 restriction then with the addition of CsA, FIV should be in a human TRIM5 susceptible state, however no defect is observed in U87 cells. Potentially, this could either be due to U87 cells expressing a low level of human TRIM5 or FIV already being restricted by human TRIM5 in U87 cells and thus CsA does not enhance restriction. It remains unclear as to why FIV does not seem to have a CsA phenotype in U87 cells but does in Jurkat cells, it is also unclear how a CypA binding lentivirus [**Figure 23**] is able to be restricted by human TRIM5. HIV-2 represents another

question in the field as it is unable to bind CypA [296,298] and only some strains are weakly, if at all, restricted by human TRIM5. It is possible that sensitively to human TRIM5 is not solely determined by ability to bind CypA, however this other factors remain unknown.

4.6 Future work

It would be interesting to determine whether the open channel mutants, Q50Y and Q50Y 120R are able to bind CypA or not. If they are able to bind, this would support a premature uncoating model where CypA helps to regulate capsid stability via the NTD capsid beta hairpin position. If the open channel mutants are unable to bind CypA then this would support a TRIM5 mediated restriction model. This could be achieved via western blotting for CypA incorporation in virus particles with the open channel capsid mutations. Functional binding of CypA for the channel mutants could also be measured using the TRIM-CypA restriction assay used in **Figure 23**. It would also be interesting to titrate the channel mutants onto the CypA knockout U87 cells to remove any confounding effects of CsA. Furthermore, an important question which underpins this, is to what level TRIM5 is expressed in U87 cells and whether TRIM5 effect on CypA usage is titratable. Production of a TRIM5 knockout cell line and subsequent reintroduction of TRIM5, via titration, would greatly aid in this further investigation. The use of TRIM5 mutants could also aid in the investigation, examining whether the signalling domain of TRIM5 is needed or whether it is just the ability to bind. N-tropic MLV, which is strongly restricted by human TRIM5 could also be used to determine TRIM5 expression in U87 cells in comparison to B-tropic MLV.

It remains unclear whether other capsid binding cofactors, such as CPSF6, NUP153 and NUP358 play any role in the regulation of the electrostatic channel. It would

be possible to investigate this by determining if the capsid channel mutants were sensitive to depletion of other capsid binding cofactors. Furthermore, X-ray crystallography could be used to determine the effect of cofactors on the electrostatic channel. By solving the crystal structure of the HIV-1M capsid hexamer in the presence other capsid binding cofactors, say CPSF6, under various conditions we could determine if they altered the position of the NTD beta hairpin and thus if they play a role in regulating the channel.

We have shown that infection with western SIVCpz is naturally independent of CypA, in single round infection as there is no infection defect when the CypA-capsid interaction is prevented. However, it is possible that as the SIVCpz viruses are able to bind CypA [**Figure 23**] that CypA still direct integration site selection and by only measuring infection and reverse transcription we missed the effect of CsA on the SIVCpz infection. It would be interesting to investigate whether SIVCpz exhibited the same pattern of integration site selection as HIV-1M or HIV-1M in the presence of CsA. This has previously been done for HIV-1M by measuring gene density at integration sites [112] but could also be examined by Fluorescence in situ hybridization (FISH). It would be possible to use FISH on the viral genome of SIVCpz and determine into what region it integrates. Again similar work has been done using HIV-1M and found that integration occurred in the outer shell of the nucleus in close proximity to a nuclear pore [363]. Using this technique with SIVCpz and HIV-1M with CsA we could determine whether the CypA-capsid interaction is important for integration site selection for SIVCpz.

Along the evolutionary branch leading to HIV-1M from SIVCpz, HIV-1M has gained the ability to use CypA to regulate its replication, as seen by CsA sensitivity and CypA knockout. It would be possible to determine the changes that have occurred along that branch to generate a list of statically significant amino acid substitutions by using the published SubRecon software [364]. Such as study might highlight changes in the capsid

that have allowed for altered consequence of CypA binding, as both SIVCpz and HIV-1M are able to bind CypA [**Figure 23**], but only HIV-1M is dependent on CypA for optimal infection. From this, a mutagenesis study could be undertaken to determine the key changes which have occurred that have increased HIV-1M dependency on CypA and thus might provide understanding on how HIV-1M expanded its tropism and become pandemic.

Furthermore, with the recent work showing that CypA is able to protect HIV-1M from human TRIM5 mediated restriction, it would be interesting to know what is happening in chimpanzees. Determining the relationship between chimpanzee TRIM5, SIVCpz and Cpz CypA would be interesting and might give insight to the chimpanzee-human zoonoses which gave rise to HIV-1M. Furthermore, we have investigated the effect of CsA on various lentiviruses in the context of human cells, it would therefore be interesting to examine the effect of CsA in cells from different species. For example, EIAV in an equine cell line.

5 Lentiviral CypA dependence in U87 cells is independent of TREX-1

5.1 Chapter 5 introduction

It remains unclear as to why some cells contain a high level of CypA dependency for HIV-1M and others do not. We know that CsA in THP-1 cells causes little to no defect in HIV-1M infection, HeLa and CrFK cells give an intermediate phenotype [unpublished data from the Towers lab] whilst in U87 cells CsA is highly inhibitory [see Chapter 4]. Recent data suggests that the HIV-1M capsid remains intact as it traverses the cytoplasm. This, combined with the knowledge that only target cell CypA influences infection, implies that differences in cytoplasmic environment could dictate CypA usage. One hypothesis is that CypA plays a role in regulating capsid stability and when the capsid-CypA interaction is disrupted capsid has reduced stability and uncoats prematurely uncoating releasing the genome into the cytoplasm. One hypothesis is that disrupting the CypA-capsid interaction always leads to premature uncoating, but the consequence of premature uncoating varies depending on cell type. In macrophages and U87 cells, premature uncoating is highly detrimental leading to failed infection, as well as a inhibitory innate immune response in macrophages. Whereas in THP-1 cells, and a lesser extent HeLa and CrFK, premature uncoating may not be as detrimental.

We hypothesise that in cell lines which are unable to sense and trigger an immune response to HIV-1M infection, such as U87, HeLa and CrFK cells, CypA usage and thus the result of premature uncoating is determined by the presence of cellular cytoplasmic nucleases. If the virus prematurely uncoats in the cytoplasm, these cellular nucleases will degrade the viral genome and thus prevent infection. In support of this hypothesis, it has been shown that DNA in the cytoplasm is rapidly degraded by host cellular nucleases [365] this was done by monitoring the degradation of cDNA plasmids inserted into the

cytoplasm of cells via microinjection [365]. Recently it has been reported that TREX-1, a host cellular nuclease, is able to degrade the HIV-1M genome when it is in the cytoplasm [235]. In the absence of this degradation, the HIV-1M genome can be sensed by cGAS, when present in the cytoplasm, which leads to production of IFN-b and CXCL10 in a cGAS-STING dependent innate immune response [236]. We hypothesise that HIV-1M has a high level of CypA dependency in cells which have highly active/abundant TREX-1, or other nucleuses, and therefore need the greatest protection for the genome when traversing the cytoplasm. Likewise, in cells in which HIV-1M infection is independent of CypA, we hypothesise that this is caused by reduced expression TREX-1 and other cellular nucleases leading to a less hostile cytoplasmic environment for viral DNA.

Here we focus on U87 cells and the role TREX-1 plays in determining CypA usage of HIV-1M. U87 cells were used as HIV-1M has a high dependency on CypA and they are thought to be unable to sense and trigger an immune response to HIV-1M infection. CypA dependency was determined via chemical inhibition of the capsid-CypA interaction, using CsA, and genetically, using the P90A capsid mutant.

Results

5.2 HIV-1M has a high level of CypA dependency in U87 cells

5.2.1 Infection in U87 cells with HIV-1M is dependent on CypA

It is well reported that HIV-1M has different dependencies on CypA in different cell lines. Measuring the effect of chemically inhibiting of the capsid-CypA interaction, via addition of CsA, is the most common measure of CypA dependency within cells. We sought to identify cellular factors which might influence CypA dependency of HIV-1M infection. First we identified two cells lines which had drastically different CypA dependencies, U87 and THP-1 cells. We decided to test U87 cells, an adherent cell line first obtained

from a human primary glioblastoma, against non-differentiated THP-1 cells, a human monocytic cell line derived from an acute monocytic leukemia grown in suspension. Using U87 and THP-1 cells we examined infection single round HIV-1M vector in the presence of 5μM CsA [**Figure 36**]. We found that in U87 cells there was a 10 fold defect in infectivity when CsA was added at the time of infection [**Figure 36**A], however no significant defect was observed after CsA treatment of THP-1 cells [**Figure 36**B]. To confirm that HIV-1M does not require CypA in THP-1 cells, we knocked out CypA via CRISPR/Cas9 and determined infectivity. Loss of CypA in THP-1 cells did not affect HIV1-M infectivity whereas in U87 cells loss of CypA caused a 10 fold defect in infection [**Figure 36**]. Taken together, this data shows that CypA is not essential for optimal infection in THP-1 cells whereas it is in U87 cells.

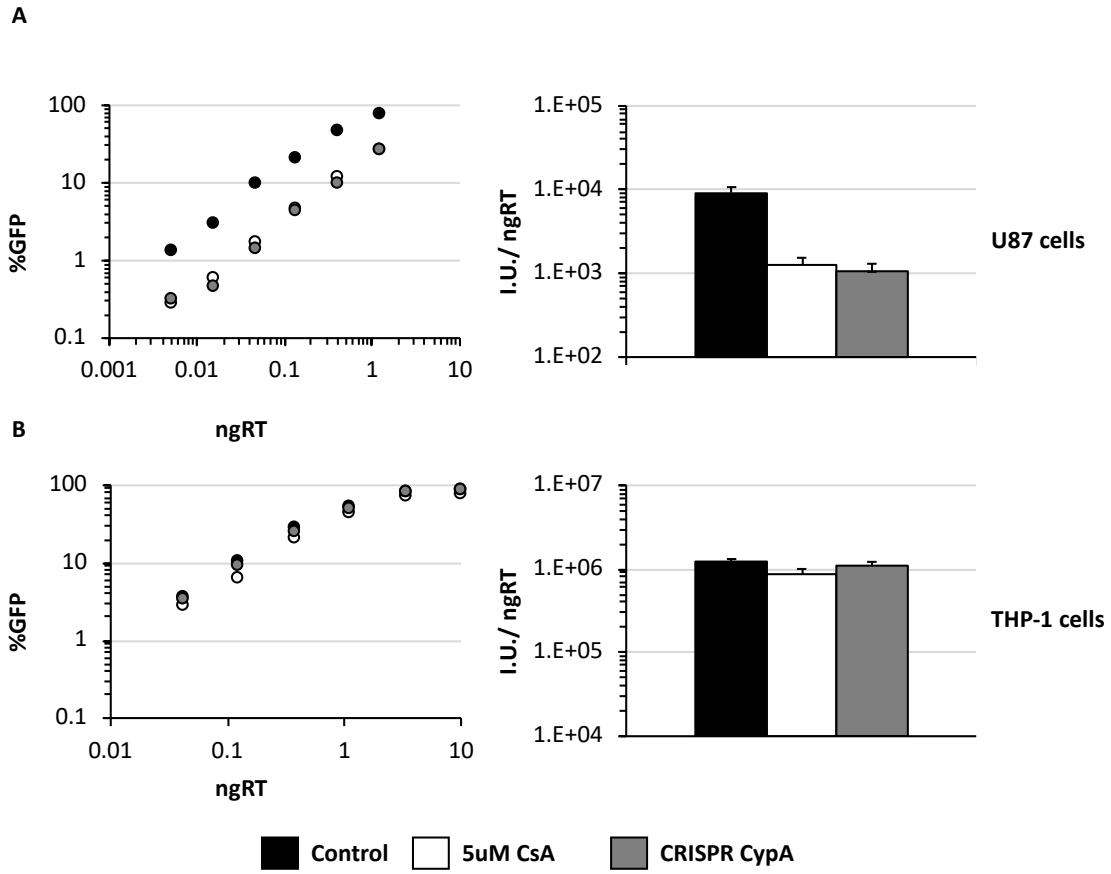


Figure 36: HIV-1M has a CypA dependency in U87 cells but not THP-1 cells
(Left hand) Infectivity plots of VSV-G pseudotyped HIV-M vector on either U87 cells (A) or THP-1 cells (B). For both U87 and THP-1 cells, infection was either performed in wild type cells (Control), in the presence of 5uM CsA (5uM CsA) which was added at the time of infection or in CypA knock out cells (CRISPR CypA). Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (Right hand) Titres determined from data in infectivity plots (left hand) using three doses of vector with MOI below 0.3, +/- SD N=2.

5.2.2 U87 cells express more TREX-1 than THP-1 cells

One possibility is that CypA helps to regulate capsid stability and so disruption of the CypA-capsid interaction is detrimental as it leads to premature uncoating. We hypothesise that premature uncoating would led to the HIV-1M genome being in an exposed state within the cytoplasm and therefore susceptible to degradation by cytoplasmic nucleases. Therefore, it is possible that cells in which HIV-1M has a high dependency on CypA are

those with a greater level of expression of TREX-1 thus the HIV-1M genome needs the greatest protection when traveling to the nucleus.

To address this we quantified TREX-1 expression levels, via western blot, between a cell line in which HIV-1M has a high CypA dependency, U87 cells, and a cell line with little or no CypA dependency, THP-1 cells [**Figure 37**]. We found that at the protein level U87 cells express slightly higher amounts of TREX-1, ~ 13% more TREX-1 relative to the tubulin control, as determined by ImageJ analysis. Although only a modest difference, we decided to examine the role TREX-1 plays in CypA dependency further.

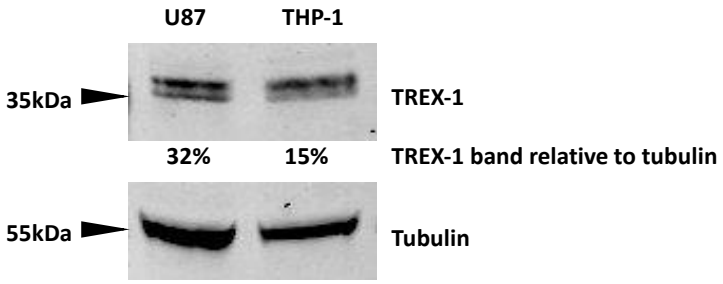


Figure 37: U87 cells express slightly higher amounts of TREX-1
Western blot detecting TREX-1 or tubulin as a loading control in wild type U87 or THP-1 cells. Lanes normalised by cell number. Size marker positions are shown with molecular weights (KDa)

5.3 Depletion of TREX-1 in U87 cells does not alter CypA dependency

5.3.1 TREX-1 depletion does not alleviate CsA induced restriction of HIV-1M

We sought to try and remove the CypA dependency in U87 cells by knocking down TREX-1. U87 cells express more TREX-1 compared to THP-1 cells **Figure 37**, relative to tubulin. Therefore, we hypothesised that by depleting TREX-1 in U87 cells we could make them more THP-1 like and thus decrease HIV-1M CypA dependency. U87 cells were transduced with ShRNAs targeting TREX-1 before antibiotic selection. After selection, TREX-1 protein expression levels were determined by western blot [**Figure 38A**]. Having depleted TREX-1, by ~70% quantified by ImageJ analysis, we next sought

to examine whether HIV-1M infection of the TREX-1 depleted cells was less sensitive to CsA. To determine this, wild type or TREX-1 depleted U87 cells were infected with HIV-1M vector in the presence or absence of 5μM CsA, added at the time of infection [**Figure 38B** and **C**]. We observed that CsA was similarly potent in wild type or TREX-1 depleted U87 cells. In both cases a 10 fold defect in HIV-1M infectivity was observed with no alleviation of CsA induced restriction upon TREX-1 depletion. To confirm the defect was not alleviated we measured reverse transcription products at 6 hours post infection [**Figure 38D** and **E].** As expected HIV-1M, CsA caused a 10 fold defect in infection and reverse transcription products compared to DMSO control which did not change upon deletion of TREX-1. These data show that depletion of TREX-1 in U87 cells does not rescue HIV-1M from CsA mediated restriction.

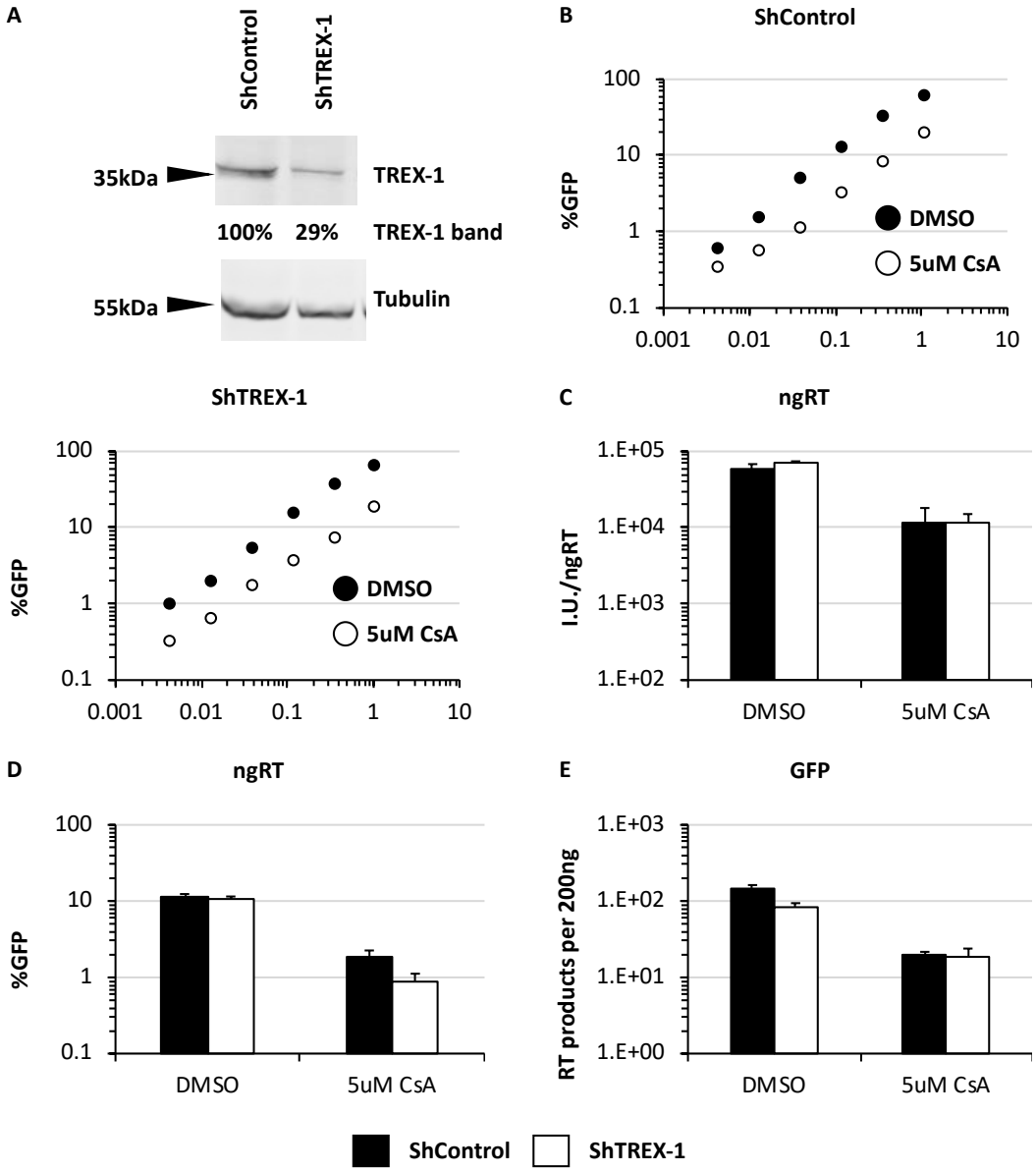


Figure 38: Depletion of TREX-1 in U87 cells does not rescue HIV-1M from CsA mediated restriction

(A) Western blot detecting TREX-1 or tubulin in control U87 (ShControl) cells or in U87 cells after ShRNA directed depletion of TREX-1 (ShTREX-1). Size marker positions are shown with molecular weights (KDa). (B) Infectivity plots of VSV-G pseudotyped vectors, derived from HIV-1M, on either empty vector expressing U87 cells (ShControl) or on TREX-1 depleted U87 cells (ShTREX-1) in the presence or absence of 5µM CsA, added at the time of infection. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (C) Titres determined from data in (B) using three doses of vector with MOI below 0.3, +/- SD N=3. (D) GFP positive wild type or TREX-1 depleted U87 cells 48 hours after infection with VSV-G pseudotyped vectors, derived from HIV-1M at an MOI of 0.1 in the presence or absence of 5µM CsA, added at the time of infection. N=2. +/- SD. (E) Reverse transcription products 6 hours from infections in (D). N=2, Error bars show standard deviation.

5.3.2 TREX-1 depletion does not alleviate the infectivity defect with P90A in U87 cells

Having shown that the CsA mediated restriction of HIV-1M infection is not alleviated by depletion of TREX-1 we next wanted to determine if the defect caused by the capsid mutant P90A was rescued. The capsid mutant P90A allows us to examine the effect of TREX-1 depletion without the confounder of any off target effects of CsA e.g. inhibition of calcineurin. As was done in section 5.3.1, TREX-1 depleted U87 cells were infected with wild type HIV-1M or HIV-1M bearing the P90A capsid mutation and titre determined at 48 hours post infection [**Figure 39A**]. As expected, in wild type U87 cells, P90A had a 10 fold defect in infection compared to wild type HIV-1M. However, this defect was not rescued upon depletion of TREX-1 [**Figure 39B**] which is consistent with TREX-1 depletion not rescuing HIV-1M from CsA mediated restriction [**Figure 38**]. It is also possible that by depleting TREX-1, any capsids which have prematurely uncoated, revealing the genome to the cytoplasmic environment, may not be degraded in TREX-1 depleted cells. The potential viral genome in the cytoplasm may undergo reverse transcription but fail to be productively infect the cell and so any additional reverse transcription would be missed by only measuring infection. To examine this, wild type or TREX-1 depleted U87 cells were infected at MOI of 0.1 with HIV-1M with either wild type or P90A containing capsid [**Figure 39C**] then reverse transcription products were measured at 6 hours post infection [**Figure 39D**]. We observed no difference between infection or reverse transcription products for either wild type HIV-1M or P90A. Supporting the data in the section above5.3.1 TREX-1 depletion does not alleviate CsA induced restriction we have shown that depletion of TREX-1 does not rescue the infection defect of the capsid mutant P90A, again suggesting that TREX-1 does not influence CypA dependence in U87 cells.

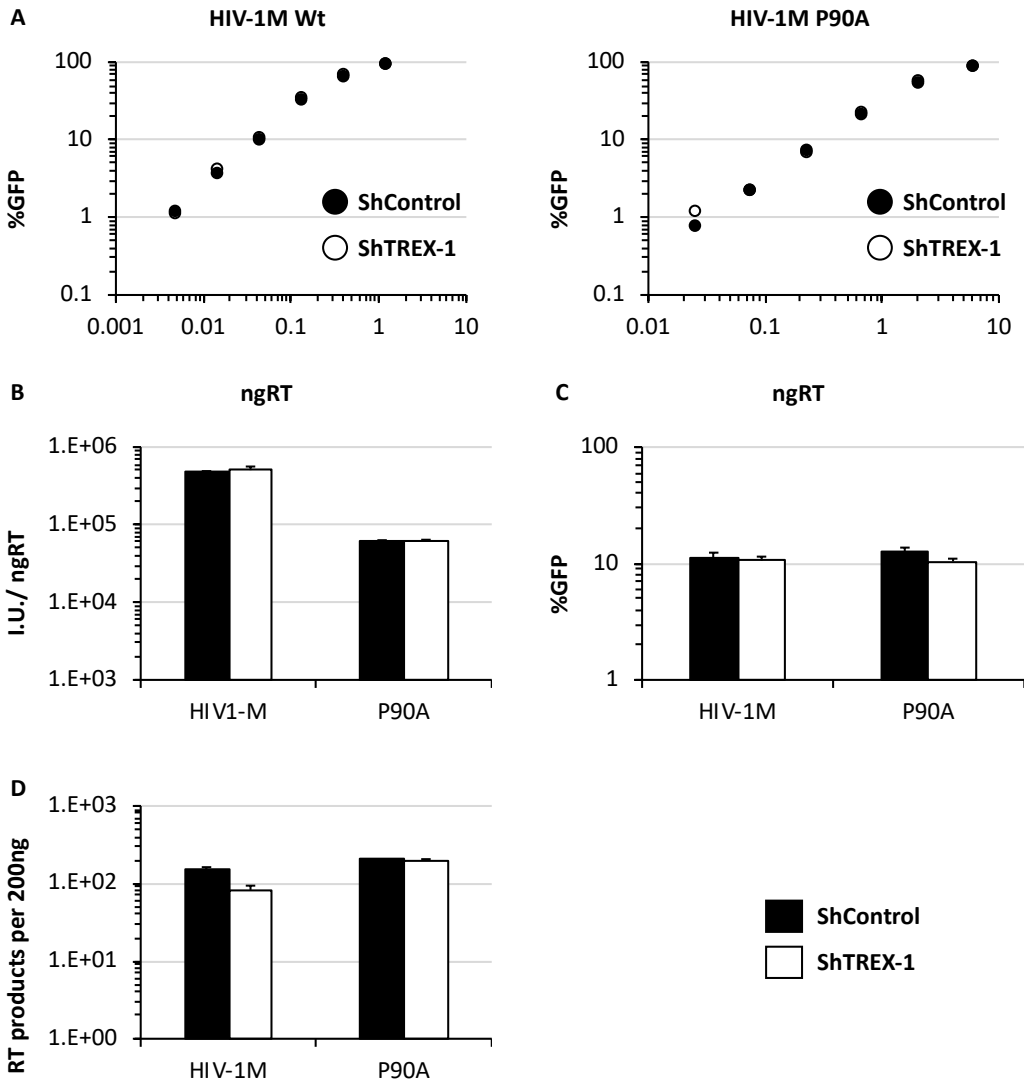


Figure 39: Depletion of TREX-1 in U87 cells does not rescue the HIV-1M capsid mutant P90A

(A) Infectivity plots of VSV-G pseudotyped vectors, derived from HIV-1M, bearing wither wild type capsid (HIV-1M) or the P90A capsid mutation (P90A) on either control U87 cells (ShControl) or on TREX-1 depleted U87 cells (ShTREX-1). Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2. (C) GFP positive wild type or TREX-1 depleted U87 cells 48 hours after infection with VSV-G pseudotyped vectors, derived from HIV-1M, either with wild type or P90A containing capsid, at an MOI of 0.1. N=2. +/- SD. (D) Reverse transcription products 6 hours from infections in (C). N=2, Error bars show standard deviation.

5.4 Overexpression of TREX-1 in U87 cells does not alter CypA dependency

5.4.1 TREX-1 overexpression does not enhance CsA induced restriction of HIV-1M

Having shown that TREX-1 depletion does not alter HIV-1M CypA dependency in U87 cells we next sought to determine if we could enhance CypA dependency. We hypothesised that CypA dependency of HIV-1M is determined by nuclease expression levels in the cytoplasmic environment. Therefore by increasing TREX-1 expression in U87 cells, we could test whether CypA dependency of HIV-1M increases. We produced stably transduced U87 cells overexpressing TREX-1 by transduction with gammaretrovirus containing the TREX-1 gene [**Figure 40A**]. The TREX-1 overexpression U87 cells were found, via western blot, to express 6.5 fold more TREX-1 compared to wild type U87 cells. HIV-1M was used to infect either wild type or TREX-1 overexpressing U87 cells in the presence or absence of 5μM CsA [**Figure 40B**]. We observed a 10 fold defect in infectivity with HIV-1M in wild type U87 cells with the addition of CsA, but no increase in inhibition in the TREX-1 overexpression U87 cells [**Figure 40C**]. The defect in infectivity seen in HIV-1M with the addition of CsA was of the same magnitude in both wild type and TREX-1 overexpressing U87 cells. Finally, we wanted to determine if reverse transcription was affected, so we infected either wild type or TREX-1 overexpressing cells with HIV-1M at an MOI of 0.1 [**Figure 40D**] before measuring reverse transcription products at 6 hours post infection [**Figure 40E**]. We found that CsA caused a defect of the same to magnitude to HIV-1M infection and reverse transcription. Overall, we observed no enhancement of CsA induced restriction on HIV-1M infection when TREX-1 was overexpressed in U87 cells.

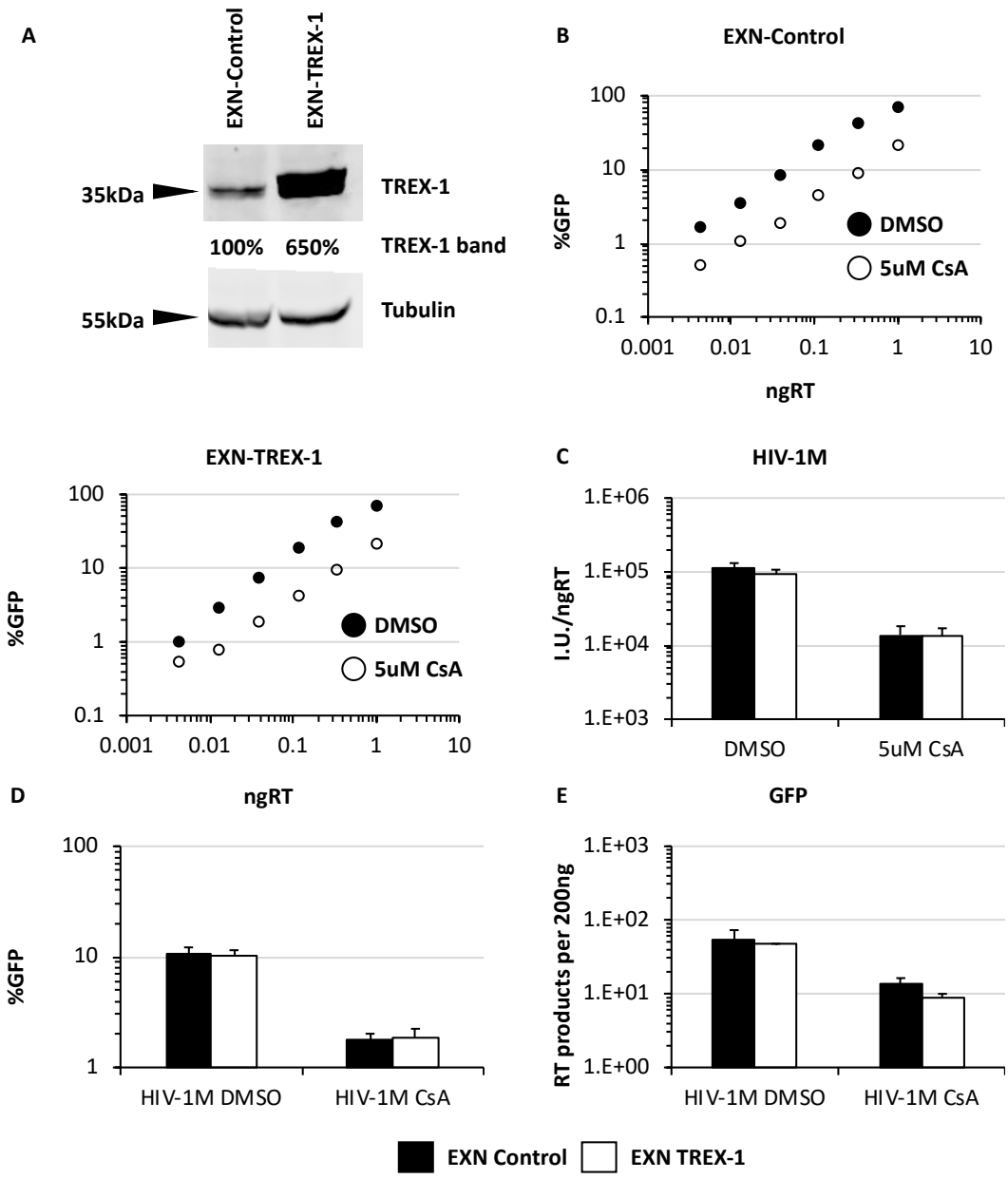


Figure 40: Overexpression of TREX-1 in U87 cells does not enhance CsA mediated restriction of HIV-1M infection

(A) Western blot detecting TREX-1 or tubulin as a loading control in empty vector expressing U87 (EXN Control) cells or in U87 cells after overexpression of TREX-1 (EXN TREX-1). Size marker positions are shown with molecular weights (KDa). (B) Infectivity plots of VSV-G pseudotyped vectors, derived from HIV-1M, on either empty vector expressing U87 cells (EXN Control) or on TREX-1 overexpressing U87 cells (EXN TREX-1) in the presence or absence of 5µM CsA, added at the time of infection. Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (C) Titres determined from data in (B) using three doses of vector with MOI below 0.3, +/- SD N=3. (D) GFP positive wild type or TREX-1 overexpressing U87 cells 48 hours after infection with VSV-G pseudotyped vectors, derived from HIV-1M at an MOI of 0.1 in the presence or absence of 5µM CsA, added at the time of infection. N=2. +/- SD. (E) Reverse transcription products 6 hours from infections in (D). N=2, Error bars show standard deviation.

5.4.2 TREX-1 overexpression does not enhance the infectivity defect of HIV-1M capsid P90A in U87 cells

Next we sought to determine whether TREX-1 overexpression was able to enhance the infectivity defect with the HIV-1M capsid mutant P90A. It is possible that both loss of Pro 90 and the CypA-Capsid interaction leads to a decrease in capsid stability which would result in exposing the genome to the cytoplasm via a premature uncoating event. We hypothesised that such an event would lead to the genome being targeting and degraded by TREX-1 and thus lead to a greater infectivity defect observed with P90A in U87 cells. To address this, we infected wild type or TREX-1 overexpression U87 cells with either HIV-1M bearing wild type or P90A containing capsid [**Figure 41A**] and determined the infectious titre [**Figure 41B**]. We found that in both wild type and TREX-1 overexpressing U87 cells, P90A showed the same behaviour, a 10 fold defect in infection. No enhancement of the P90A infection defect was observed on TREX-1 overexpression. To confirm this observation, we measured reverse transcription products at 6 hours post infection. U87 cells were infected at MOI 0.1 [**Figure 41C**] and reverse transcription products were measured at 6 hours post infection [**Figure 41D**]. We observed no significant difference in either infection or reverse transcription products of P90A compare to wild type HIV-1M during TREX-1 overexpression. These data shows that the infectivity defect observed with P90A is independent of TREX-1 expression levels.

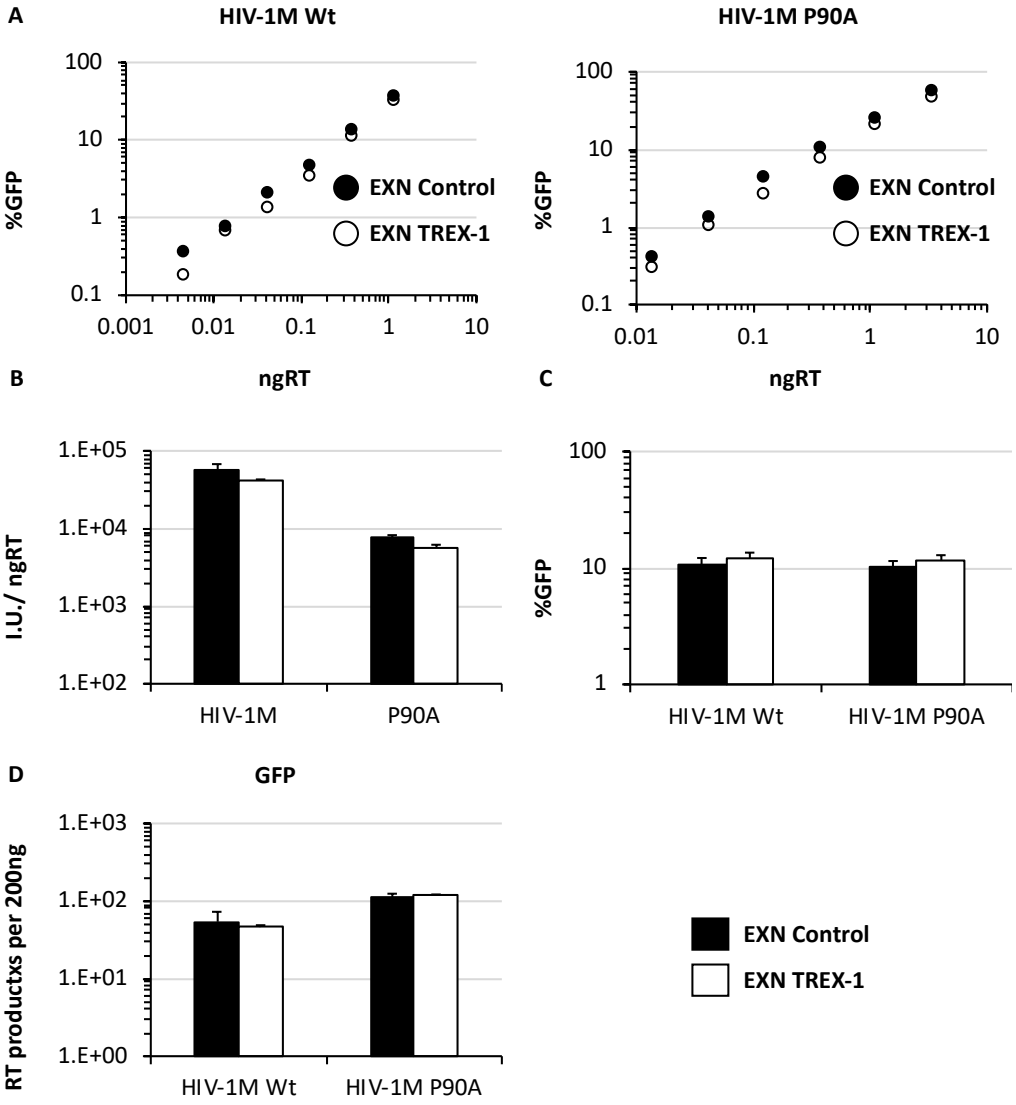


Figure 41: Overexpression of TREX-1 in U87 cells does not enhance the defect with of HIV-1M capsid mutant P90A

(A) Infectivity plots of VSV-G pseudotyped vectors, derived from HIV-1M, bearing either wild type capsid (HIV-1M) or the P90A capsid mutation (P90A) on either wild type U87 cells (EXN Control) or on TREX-1 overexpression U87 cells (EXN TREX-1). Infection determined by GFP expression (flow cytometry) 48 hours post infection. Data representative of two biological repeats. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2. (C) GFP positive wild type or TREX-1 overexpression U87 cells 48 hours after infection with VSV-G pseudotyped vectors, derived from HIV-1M, either with wild type or P90A containing capsid, at an MOI of 0.1. N=2. +/- SD. (D) Reverse transcription products 6 hours from infections in (C). N=2, Error bars show standard deviation.

5.5 Manipulation of TREX-1 expression levels does not sensitise SIVCpzMT to CsA

5.5.1 Overexpression of TREX-1 does not alter SIVCpzMT CsA insensitivity

We have previously shown that SIVCpzMT is insensitive to 5μM CsA in U87 cells [**Figure 24** and **Figure 25**] so we wanted to determine whether TREX-1 manipulation sensitised SIVCpzMT to CsA. By examining the effect of TREX-1 overexpression on SIVCpzMT, we can answer two question. Firstly, is the genome of SIVCpzMT contained in a TREX-1 protected environment and secondly, does TREX-1 play a role in regulating CypA usage in SIVCpzMT. It has previously been shown in the Towers lab that viruses thought to undergo early or premature uncoating, such as HIV-1O, have a ~2-3 fold defect in infection in U87 cells on TREX-1 overexpression [unpublished data from the Towers lab]. To answer both questions above, we infected either wild type or TREX-1 overexpressing U87 cells with SIVCpzMT in the presence or absence of 5μM CsA and measured infection [**Figure 42A**]. We found that neither TREX-1 overexpression alone nor TREX-1 overexpression combined with CsA treatment causes any significant infectivity defect to SIVCpzMT in U87 cells [**Figure 42B**]. This data suggests that SIVCpzMT does not undergo early or premature uncoating in the cytoplasm as TREX-1 overexpression in the absence of CsA caused no defect in titre. Potentially this suggests that SIVCpzMT can protect its genome from TREX-1 independently of CypA recruitment. We have also shown that TREX-1 overexpression does not alter the CsA phenotype of SIVCpzMT.

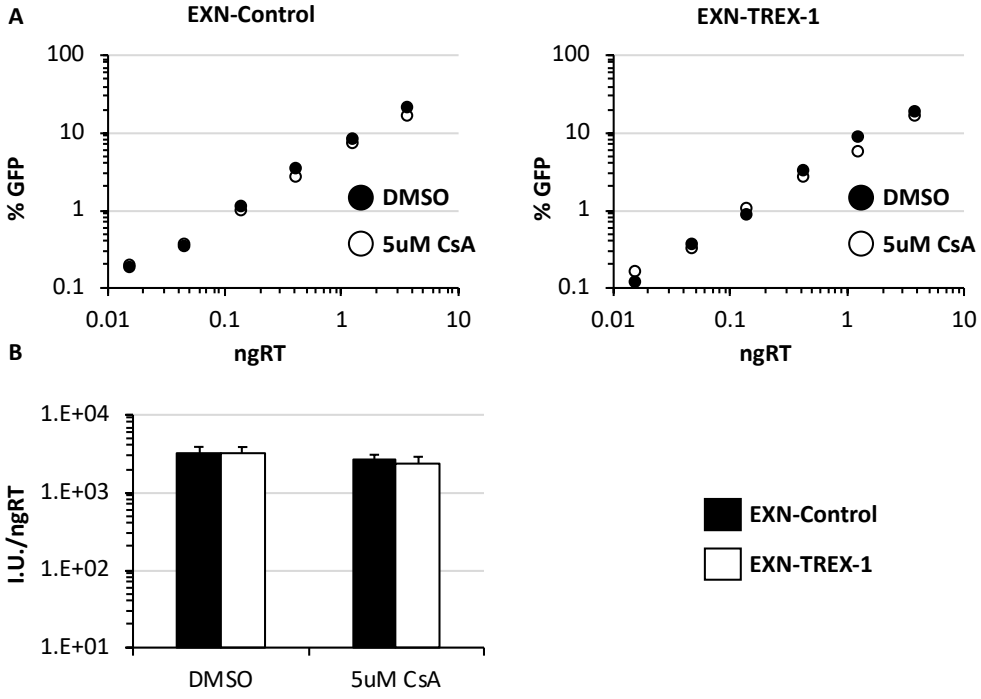


Figure 42: Overexpression of TREX-1 in U87 cells does not make SIVCpzMT sensitive to CsA mediated restriction

(A) Infectivity plots of VSV-G pseudotyped vectors, derived from SIVCpzMT on either control U87 cells (EXN Control) or on TREX-1 overexpression U87 cells (EXN TREX-1). Infection determined by GFP expression (flow cytometry) 48 hours post infection. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=1.

5.5.2 Depletion of TREX-1 in U87 cells does not alter the SIVCpzMT CsA phenotype

It is possible that U87 cells already contain a high enough level of TREX-1 to suppress any prematurely uncoated virus. So if SIVCpzMT was uncoating early, overexpressing TREX-1 would have no phenotype. Therefore we decided examine SIVCpzMT infection in the TREX-1 depleted U87 cells [**Figure 38A**]. Alongside this, we examined if SIVCpzMT became CypA dependent in a low TREX-1 environment. To address this, we infected U87 cells which are depleted for TREX-1, and examined the effect on SIVCpzMT infectivity [**Figure 43A**]. We found that the depletion of TREX-1 in U87 cells didn't cause any significant changes in SIVCpzMT infectivity compared to wild type U87 cells [**Figure 43B**]. We also examined the effect of 5μM CsA on SIVCpzMT in TREX-1 depleted U87 cells. Again, we found that there was no significate difference in CsA sensitivity in TREX-1 depleted U87 cells [**Figure 43B**]. Taken with section 5.5.1 we have shown that a virus which is not restricted by CsA, and therefore has no CypA dependence, cannot be sensitised to CsA via manipulation of TREX-1 expression.

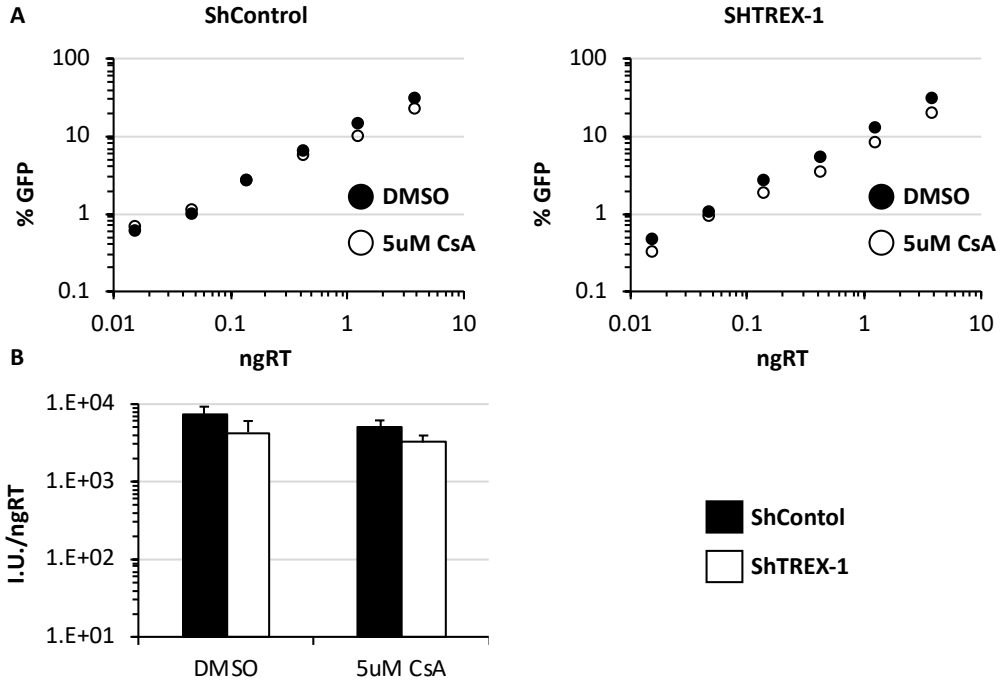


Figure 43: Depletion of TREX-1 in U87 cells does not make SIVCpzMT sensitive to CsA
(A) Infectivity plots of VSV-G pseudotyped vectors, derived from SIVCpzMT on either control (ShControl) or TREX-1 depleted U87 cells (ShTREX-1). Data representative of two biological repeats. Infection determined by GFP expression (flow cytometry) 48 hours post infection. (B) Titres determined from data in (A) using three doses of vector with MOI below 0.3, +/- SD N=2.

5.6 Discussion

5.6.1 TREX-1 does not play a role in determining CypA usage in U87 cells

Throughout this chapter we have shown that TREX-1 does not contribute in CypA usage of lentiviruses in U87 cells. We have shown that manipulations of TREX-1, with both overexpression and depletion, in U87 cells does not alter CsA mediated restriction of the CypA dependent virus HIV-1M [**Figure 38** and **Figure 40**]. We have also taken the opposite approach, taking a lentivirus, SIVCpzMT, which is known to be able to bind CypA [**Figure 23**] but not to have an infectivity defect when the CypA-capsid interaction is disturbed in U87 cells via addition of CsA [**Figure 24**, **Figure 25** and section 5.5] and

manipulated TREX-1 in an attempt to sensitise the virus to CsA. However, we were unable to reconstitute the CsA phenotype of HIV-1M in SIVCpzMT with manipulation of TREX-1 expression levels [**Figure 42** and **Figure 43**]. We have also shown that the HIV-1M capsid mutant P90A is not sensitive to overexpression of TREX-1 and nor does it have its titre rescued by depletion of the nuclease [**Figure 39** and **Figure 41**]. Likewise, P90A response to CsA is also unaltered by varying TREX-1 expression. We have therefore clearly shown that TREX-1 expression does not influence CypA dependency of lentiviruses within U87 cells.

It remains unclear as to why HIV-1M has a high level of CypA dependency for infection in some cells and not others. It is possible that TREX-1 is only one of several cellular cytoplasmic nucleuses which are able to degrade the HIV-1M genome. Thus other nucleases might be the cause of the hostile environment which lentiviruses aim to avoid. Recently it has been shown that CypA play a role in protecting HIV-1M from human TRIM 5 mediated restriction [See section 1.4.1]. Therefore, it is also possible that the varying CypA dependencies are due to different expression levels of TRIM 5 in different cells. However, the majority of the work describing the protective ability of CypA against TRIM 5 mediated restriction has been done in primary cells [257,258] such as macrophages and T cells. It is still unclear whether in cell lines, CypA protects from TRIM 5 mediated restriction to the same degree. THP-1 cells, in which HIV-1M has no CypA dependency [**Figure 36**] might express much lower levels of TRIM 5 compared to U87 cells. However, varying levels of TRIM 5 cannot fully explain the varying dependencies. For example, in the highly permissive CrKF cell line, which express a non-functional TRIM 5 [361] as seen by the permissively to TRIM 5 restricted viruses such as EIAV and N-tropic MLV, HIV-1M has an intermediate CsA phenotype [Unpublished data, Towers lab]. It is likely that several cellular factors contribute to CypA dependency during HIV-1M infection. We have clearly shown that TREX-1 does not contribute to CypA dependency

of HIV-1M. However, we cannot rule out the possibility that another cellular nuclease does contribute.

5.7 Future work

Although we have shown TREX-1 overexpression and depletion at the protein level, we did not confirm a functional effect of altering TREX-1 levels. It is possible that despite a ~7 fold increase in TREX-1 protein level, functionally there is no difference between wild type and TREX-1 overexpression U87 cells. Confirmation of altered TREX-1 functionally could be achieved by transfecting U87 cells with GFP encoding plasmid and measuring GFP positive cells. Preferably, this should be done as a titration to examine whether TREX-1 can be saturated by cytoplasmic DNA. It is known that the majority of vector particles uncoat upon entering the cell [232] and at a high MOI these non-infectious particles could deliver enough viral genome to saturate TREX-1 and thus mask an effect. Likewise, with TREX-1 depletion, it is unknown whether TREX-1 is present in excess within the cell. Potentially the TREX-1 depleted cells still contained enough TREX-1 for normal function and therefore we did not see any effect on lentiviral infection. If we were able to deplete TREX-1 further or produced a U87 TREX-1 knockout cell line we might have seen an effect on infection. Furthermore, currently it is unknown where overexpressed TREX-1 is localised with a cell. We assumed that overexpressed TREX-1 would remain in the cytoplasm but it is possible that it may enter the nucleus. By inserting a GFP tag into the expressed TREX-1 cDNA or by using a TREX-1 antibody with fluorescent tag it would be possible to determine cellular localisation via fluorescent microscopy.

Here we focused on primate lentiviruses from the SIVCpz-HIV-1M lineage. It would be interesting to determine how manipulation of TREX-1 expression levels affects

other lentiviruses in U87 cells. Of particular interest would be FIV, EIAV and HIV-2 as they are thought to undergo a different lifecycle to that of HIV-1M, as shown by difference in cofactor usage [see section 1.4].

Here we have investigated the effect of manipulating TREX-1 expression levels on the HIV-1M capsid mutant P90A. Investigating other capsid mutation, such as R18G, and the channel mutants [See chapter 4] might produce insight. Likewise, using capsid mutation known to alter uncoating kinetics, such as E45A and P28A which lead to a hyper stable or unstable capsid respectively [231], would also be of interest. Furthermore, we have only investigated the effect of manipulating TREX-1 expression levels on CsA ability to restrict HIV-1M infection. The use of PF74, a drug known to promote partial uncoating of the HIV-1M capsid core [311] might have altered potency when used in the context of TREX-1 depleted or overexpression cell lines.

Recent work has shown that CypA protects HIV-1M from restriction by TRIM5. This has been shown in primary cells including macrophages and T cells. It remains unclear whether TRIM 5 expression explains the differences seen between cell lines for CypA usage in HIV-1M. It would therefore be interesting to investigate the effect of manipulating TRIM 5 expression levels in U87 cells measuring CypA usage of various lentiviruses. This has been outlined in section 4.6.

6 Concluding remarks

6.1 Summary of key results

By taking a phylogenetic approach, we have shown that the recently identified IP6 binding sites within the HIV-1M capsid [239,290] are highly conserved throughout the lentivirus lineage. With the exception of the ovine-caprine lineage, all examined lentiviruses contain both IP6 binding sites within the immature Gag [**Figure 7**], capsid residues Lys 158 and Lys 227 for HIV-1M. Mutagenesis of these sites are highly detrimental to infection as they cause a large defect in particle production for all examined lentiviruses [**Figure 8** and **Figure 9**]. Furthermore, we have also shown that the IP6 binding site within the mature capsid, Arg 18 for HIV-1M, is also highly conserved throughout the lentivirus lineage [**Figure 10**]. Again, mutagenesis of the IP6 binding site within the mature capsid caused a large defect in infectivity, which was due to a block of equal magnitude in reverse transcription [**Figure 12**]. From these observations, we infer that IP6 is a conserved cofactor throughout the lentivirus lineage. A closer examination of the feline lineage of lentiviruses found that there was a directed switch by FIV-Pca to form its electrostatic channel with a Lys instead of an Arg, as all the FIVs from non-domesticated felines do. This is interesting as it was found that FIV-Pca was less dependent on its electrostatic channel and remained infectious even after its loss [**Figure 19**]. This is in contrast to HIV-1M, which is unable to tolerate any mutation at position 18 within its capsid.

We have shown that despite the ability to bind CypA is conserved in all examine lentiviruses, except EIAV and HIV-2, only HIV-1M has large defect in infectivity when the capsid-CypA interaction is disrupted by the use of CsA [**Figure 24**]. The infectivity defect observed in U87 cells with CsA treatment was found to be due to a block in reverse transcription [**Figure 25**]. Both the eastern SIVCpz, SIVCpzBF, and HIV-1O have a

small defect with CsA treatment in U87 cells [**Figure 24**] but these were found not to be due to CypA mediated affects as neither had any defect in U87 cells when CypA was knocked out [**Figure 27**]. Previous work has shown that a western SIVCpz virus is sensitive to CsA in spreading infection [298], however we found that in single round infection and particle production western SIVCpz was insensitive to CsA treatment [**Figure 24** and **Figure 29**]. Furthermore, we have shown that there is a relationship between the position of the capsid NTD beta-hairpin and sensitivity to CsA treatment. When HIV-1M had the capsid NTD beta-hairpin forced into a more open conformation, by Q50Y or Q50Y 120R mutagenesis, HIV-1M become insensitive to CsA treatment [**Figure 32** and **Figure 33**]. Whereas forcing the capsid NTD beta-hairpin into an closed conformation, H12Y, did not alter CsA sensitivity [**Figure 32** and **Figure 33**].

Additionally, we found that TREX-1 expression levels in U87 cells does not alter CsA sensitivity in either HIV-1M or the parental western SIVCpz virus. Neither overexpression nor depletion of TREX-1 in U87 cells was able to modulate CsA sensitivities of either HIV-1M or SIVCpzMT [**Figure 38**, **Figure 40**, **Figure 42** and **Figure 43**]. Furthermore, we also showed that manipulation of TREX-1 expression levels could not rescue the infectivity defect caused by the capsid mutant P90A in the HIV-1M capsid [**Figure 39** and **Figure 41**].

6.2 Model for IP6 usage within the lentivirus lineage

In this study, we have shown that IP6 is not just a cofactor for HIV-1M infection but rather a general cofactor for the lentivirus lineage. IP6 is known to bind to the HIV-1M capsid in both the immature Gag and in the mature capsid core [**Figure 44A and B**]. With both of these interactions, IP6 is critical for charge neutralisation and thought to regulate stability of the immature Gag bundle or the mature capsid core. We have shown that the

IP6 binding sites within both the immature Gag and mature capsids of various lentiviruses is highly conserved and disruption of these binding sites, via mutagenesis, is highly detrimental for viral infectivity [see Chapter 3]. From this, we infer that IP6 is a conserved cofactor throughout the lentivirus lineage. Here we outline our current working model for IP6 within lentivirus infection and note any significant variation observed.

Once a cell becomes productively infected with a lentivirus, it starts to produce Gag. In its immature form, Gag is unable to tightly cluster due to the positively charged residues located within helix six, Lys 158 and Lys 227 in HIV-1M. IP6 binding neutralises the charge of Lys 158 and Lys 227. This binding has two effects; 1), it neutralises the charge repulsion from the immature Gag molecules which aids in driving immature particle formation, and 2), the recruitment of IP6 into the immature lentivirus particle that later will be used to assemble the mature core [**Figure 44**]. All examined lentiviruses have a severe defect to infectivity when the IP6 binding site in the immature Gag is lost, by mutagenesis [**Figure 8**]. For all examined lentiviruses the defect caused by mutagenesis of the IP6 binding site in the immature Gag is due to decrease particle production [**Figure 9**.]. Interestingly, ovine-caprine lineage of lentiviruses naturally lack the Lys 158 IP6 binding site. It remains unclear what the role of IP6 is in the ovine-caprine lineage of lentiviruses.

Once the immature virus particle is formed, it undergoes particle maturation in which the viral protease cleavage the Gag polyprotein and leads to the formation of the mature particle. The maturation process leads to the unmasking of the IP6 binding site within the electrostatic channel of the capsid hexamer, Arg 18 for HIV-1M. Once unmasked, IP6 binds to Arg 18, for HIV-1M, and acts to neutralise the charge repulsion at the centre of the electrostatic channel caused by the tightly clustered Arg, or Lys for FIV-Pca, which results in increased capsid stability [see section 1.4.4]. The IP6 binding

site within the mature capsid hexamer is highly conserved throughout the lentivirus lineage [**Figure 10**]. Disruption of the IP6 binding site within the mature capsid in all examined lentiviruses, except for FIV, is highly detrimental to infection with a defect of equal magnitude in reverse transcription [**Figure 12**]. Interestingly, FIV is able to lose its electrostatic channel, seen by mutagenesis to Gly, and remain partly infectious [**Figure 19**.]. This is in contrast to HIV-1M, which is unable to tolerate any substitution for Arg 18. FIV could represent a virus that uses IP6 for particle assembly but is less dependent on the mature capsid stabilizing effect of IP6, potentially due naturally uncoating in the cytoplasm.

Currently, the field proposes a model in which the lentiviral capsid core remains intact as the virus traverse the cytoplasm. Central to this hypothesis is the electrostatic channel and IP6. HIV-1M has been shown to be able to reverse transcribe within its mature capsid core, it is likely that other lentiviruses are also able to do this. Encapsidated reverse transcription allows reverse transcription to take place within a protected environment and thus avoid a hostile cytoplasm [**Figure 44**C], which might contain immune sensors and DNases. However, it is known that as reverse transcription progresses, internal stresses within the capsid increase driving uncoating [see section 1.3]. Therefore maintenance of capsid stability is key to prevent premature uncoating and ensure productive infection [**Figure 44**D]. However, a hyper stable capsid would only function to trap the reverse transcription products within it and thus prevent infection. Therefore, lentiviruses need a mechanism to dynamically regulate capsid stability. IP6 is thought to provide the mechanism by which capsid stability and reverse transcription are both regulated.

Our lab proposes that IP6 usage is influenced by the constellation of cofactors which are bound to the capsid at any specific time. The lentivirus capsid is known to be

able to bind to several cofactors [see section 1.4], each of which have a specific cellular location e.g. CypA is cytoplasmic whereas NUP153 is in the NPC. By using cofactors to influence IP6 usage, lentiviruses would be able to "know" their cellular location and thus avoid premature uncoating whilst promoting uncoating at the correct location. Although the exact location where the lentiviral capsid core uncoats is still debatable, recent data suggests it is likely near or at the stage of nuclear import [see section 1.2.4 and section 1.3.] For example, it is possible that CypA, alongside its recently reported function to protect HIV-1M from human TRIM5 restriction, is able to regulate the position of the NTD beta-hairpin to maintain a more closed conformation retaining IP6 within the electrostatic channel of HIV-1M and thus act to prevent uncoating within the cytoplasm by maintaining optimal capsid stability. Once the capsid core reaches the NPC, CypA could be displaced by NUP358 which might results in the release of IP6 and thus a decrease in capsid core stability to promote uncoating. It is also possible that another cofactor functions to displace IP6 and decrease capsid stability. It would be interesting to determine the effect other lentivirus cofactors have on IP6 binding and thus capsid stability.

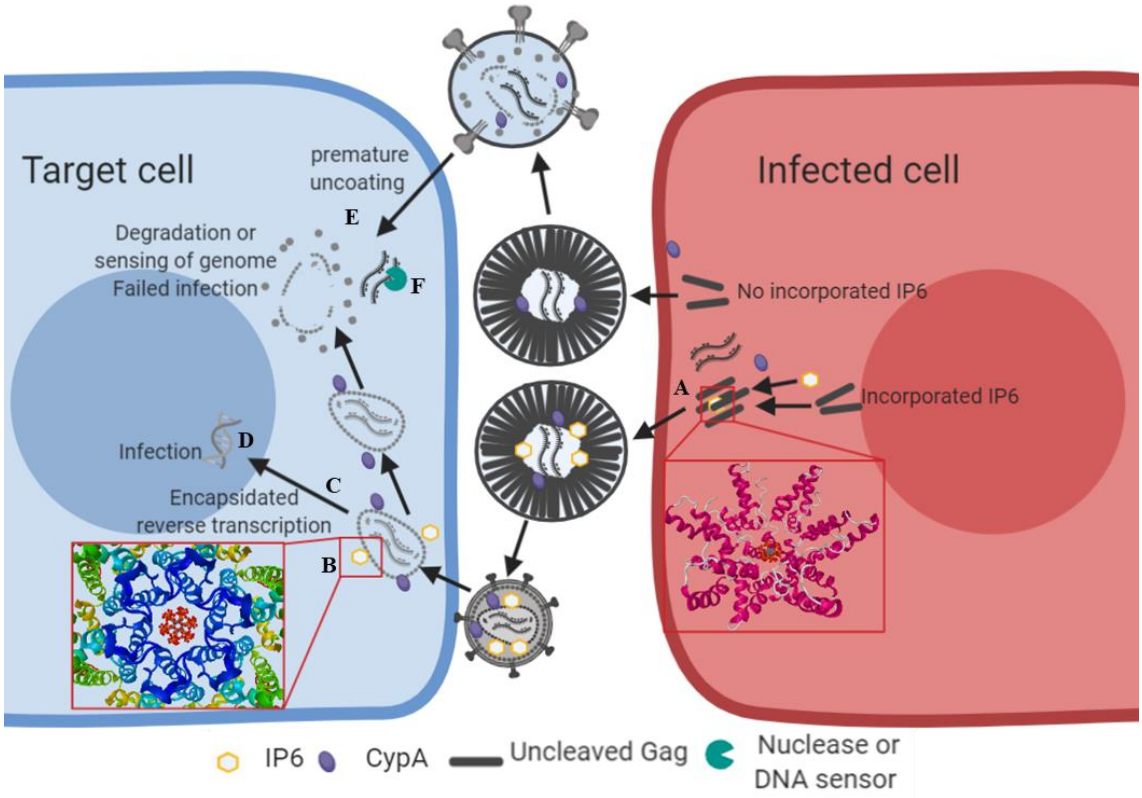


Figure 44: Schematic overview IP6 usage in the *lentivirus* lineage
Within an infected cell, IP6 binds to uncleavged Gag at residue Lys 157 and Lys 227 this neutralises the electrostatic repulsion force generated by the positively charged Lys side chains and promotes the formation of the immature Gag bundle (A), this aids in immature particle formation. Once Gag cleavage has occurred within the immature particle, the binding site for IP6 in the mature capsid, Arg 18, is unmasked. IP6 binds to Arg 18 and neutralises the electrostatic repulsion, generated by the positive charged Arg side chaians, and stabilises the mature capsid structure (B). After entry into the target cells cytoplasm, the capsid core remains intact where it influences later stages in the Lifecyle via interactions with host cell capsid binding cofactors. At this stage encapsidated reverse transcription starts (C). dNTPs enter the capsid core interior via the electrostatic channel. Encapsidated reverse transcription provides a mechanism in which the viral genome and reverse transcription products are protected from host cell nucleases and innate immune sensors. Once reverse transcription is complete, the viral DNA enters the nucleus and is integrated into the target cell genome resulting in productive infection (D). If IP6 is not present within the immature particle or is lost in the target cell, the mature capsid core will have a reduced stability and might resulting in premature capsid core uncoating (E). Premature uncoating is thought result in failed infection as the viral genome is released into the cytoplasm were it will either be degraded by cellular nuclease or sensed by innate immune sensors (F). We also propose that host capsid binding cofactors, like CypA, are able to influence IP6 binding to the electrostatic channel and thus play a role in regulating stability.

6.3 Future direction

Throughout this study, several questions have arisen but remain unanswered. Here I outline several of these questions which could be pursued in future work.

Perhaps the largest questions coming from the work presented in Chapter 3 is why FIV-Pca is able to tolerate the loss of its electrostatic channel and remain infections when all other examine lentiviruses have large infectivity defects when the electrostatic channel is lost [**Figure 12**]. Following from this, is the observation that FIV from non-domesticated felines, such as lions and pumas, form the electrostatic channel with an Arg instead of a Lys like FIV-Pca. This bring the question of whether FIV from non-domesticated felines is able to lose their electrostatic channel and remain infectious like FIV-Pca or if the loss will render the virus uninfectious like HIV-1M R18G. If FIV from non-domesticated felines is unable to tolerate the loss of the electrostatic channel then it would be interesting to determine why FIV-Pca has evolved to become less dependent on the electrostatic channel. This reduced dependency on the electrostatic channel could have implication for tropism and might suggest that FIV-Pca has evolved to replicate in cells that do not do DNA sensing that thus is able to uncoat in the cytoplasm and remain infectious. Furthermore, the bovine lineage is also of interest as the lineage solely uses Lys to form its electrostatic channel. It would be interesting to determine whether the dependence on the electrostatic channel of bovine lineage is more similar to HIV-1M or FIV-Pca.

The role of CypA, regarding the capsid hexamer NTD beta-hairpin, still remains unclear. In this study we propose a model in which CypA closes the electrostatic channel and thus aids in maintaining capsid stability. It would be interesting to investigate the allosteric role of CypA, and other capsid binding cofactors like CPSF6 for example, on the capsid and IP6 binding using crystallography or nuclear magnetic resonance. This

would allow the study of the effect of various cofactors on IP6 binding and the position of the NTD beta-hairpin. Ultimately, we would want to determine how capsid binding cofactors alter capsid stability and if they do, if it is mediated by altering IP6 binding to the electrostatic channel. Furthermore, the recently described function of CypA to protect HIV-1M from TRIM5 restriction [257,258] also raises interesting hypothesis [See section 1.3.4]. Primarily does TRIM5 expression levels explain why HIV-1M has different cofactor dependences in different cell lines?

7 References

1. Baltimore, D. Expression of animal virus genomes. *Bacteriol. Rev.* **35**, 235–41 (1971).

2. Weiss, R. A. Retrovirus classification and cell interactions. *J. Antimicrob. Chemother*. **37**, 1–11 (1996).

3. Rous, P. A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. *J. Exp. Med.* **13**, 397–411 (1911).

4. ROUS, P. A Transmissible Avian Neoplasm (Sarcoma of the Common Fowl). *J. Exp. Med.* **12**, 696–705 (1910).

5. Vallée, H., Carré, H. Sur la nature infectieuse de l'anémie du cheval. *C.R. Hebd. Seances Acad. Sci. Ser. D Sci. Nat.* **139**, 331–333 (1904).

6. Lui, K., Darrow, W. & Rutherford, G. A model-based estimate of the mean incubation period for AIDS in homosexual men. *Science.* **240**, 1333–1335 (2006).

7. States, U. & York, N. Long HIV-1 incubation periods and dynamics of transmission within a family. *Lancet* **1**, 134–136 (1990).

8. Gallo, R. *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).

9. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science.* **220**, 868–71 (1983).

10. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science.* **224**, 497–500 (1984).

11. Gendelman, H. E., Narayan, O., Molineaux, S., Clements, J. E. & Ghotbi, Z. Slow, persistent replication of lentiviruses: role of tissue macrophages and macrophage precursors in bone marrow. *Proc. Natl. Acad. Sci.* **82**, 7086–7090 (1985).

12. English, R. V, Johnson, C. M., Gebhard, D. H. & Tompkins, M. B. In vivo lymphocyte tropism of feline immunodeficiency virus. *J. Virol.* **67**, 5175–86 (1993).

13. Onuma, M. *et al.* Infection and dysfunction of monocytes induced by experimental inoculation of calves with bovine immunodeficiency-like virus. *J. Acquir. Immune Defic. Syndr.* **5**, 1009–1015 (1992).

14. Sellon, D. C., Perry, S. T., Coggins, L. & Fuller, F. J. Wild-type equine infectious anemia virus replicates in vivo predominantly in tissue macrophages, not in peripheral blood monocytes. *J. Virol.* **66**, 5906–5913 (1992).

15. Unger, R. E. *et al.* Detection of simian immunodeficiency virus DNA in macrophages from infected rhesus macaques. *J. Med. Primatol.* **21**, 74–81 (1992).

16. Yamashita, M. & Emerman, M. Retroviral infection of non-dividing cells: Old and new perspectives. *Virology* **344**, 88–93 (2006).

17. Lewis, P. F. & Emerman, M. Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.* **68**, 510–6 (1994).

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

19. Humphries, E. H. & Temin, H. M. Requirement for cell division for initiation of transcription of Rous sarcoma virus RNA. *J. Virol.* **14**, 531–46 (1974).

20. Bieniasz, P. D., Weiss, R. A. & McClure, M. O. Cell cycle dependence of foamy retrovirus infection. *J Virol* **69**, 7295–9 (1995).

21. Lewis, P., Hensel, M. & Emerman, M. Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J.* **11**, 3053–3058 (1992).

22. HIV Q/A - from https://www.who.int/features/qa/71/en/. *World Health Organisation* (2017).

23. Sones, M., Israel, H. L., Dratman, M. & Frank, J. H. Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men — Evidence of a New Acquired Cellular Immunodeficiency. *New Engl. Mournal Med.* **244**, 209–213 (1951).

24. Hemelaar, J. *et al.* Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect. Dis.* **19**, 143–155 (2019).

25. Fiebig, E. W. *et al.* Dynamics of HIV viremia and antibody seroconversion in plasma donors: Implications for diagnosis and staging of primary HIV infection. *AIDS* **17**, 1871–1879 (2003).

26. Etienne, L. *et al.* Characterization of a new simian immunodeficiency virus strain in a naturally infected Pan troglodytes troglodytes chimpanzee with AIDS related symptoms. *Retrovirology* **8**, 1–13 (2011).

27. Keele, B. F. *et al.* Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* **460**, 515–519 (2009).

28. Klatt, N. R., Silvestri, G. & Hirsch, V. Nonpathogenic simian immunodeficiency virus infections. *Cold Spring Harb. Perspect. Med.* **2**, (2012).

29. English, R. V *et al.* Development of clinical disease in cats experimentally infected with feline immunodeficiency virus. *J. Infect. Dis.* **170**, 543–552 (1994).

30. Pedersen, N. C., Ho, E. W., Brown, M. L. & Yamamoto, J. K. Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. *Science.* **235**, 790–793 (1987).

31. Yamamoto, J., Sparger, E., Ho, E. W. & Andersen, P. R. Pathogenesis of experimentally induced feline immunodeficiency virus infection in cats. *Am. J. Vet. Res.* **49**, 1246–1258 (1988).

32. Pedersen, N. C., Yamamoto, J. K., Ishida, T. & Hansen, H. Feline immunodeficiency virus infection. *Vet. Immunol. Immunopathol.* **21**, 111–129 (1989).

33. Packer, C. *et al.* Viruses of the Serengeti: Patterns of infection and mortality in African lions. *J. Anim. Ecol.* **68**, 1161–1178 (1999).

34. Lutz, H., Isenbügel, E., Lehmann, R., Sabapara, R. H. & Wolfensberger, C. Retrovirus infections in non-domestic felids: serological studies and attempts to isolate a lentivirus. *Vet. Immunol. Immunopathol.* **35**, 215–224 (1992).

35. Carpenter, M. A. & O'Brien, S. J. Coadaptation and immunodeficiency virus: lessons from the Felidae. *Curr. Opin. Genet. Dev.* **5**, 739–745 (1995).

36. Ziegler, S. *et al.* Mapping Bushmeat Hunting Pressure in Central Africa. *Biotropica* **48**, 405–412 (2016).

37. Hemelaar, J., Gouws, E., Ghys, P. D. & Osmanov, S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* **25**, 679–689 (2011).

38. D'arc, M. *et al.* Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1343-52 (2015).

39. Delaugerre, C., De Oliveira, F., Lascoux-Combe, C., Plantier, J. C. & Simon, F. HIV-1 group N: Travelling beyond Cameroon. *Lancet* **378**, 1894 (2011).

40. Vallari, A. *et al.* Confirmation of Putative HIV-1 Group P in Cameroon. *J. Virol.* **85**, 1403–1407 (2011).

41. Plantier, J. C. *et al.* A new human immunodeficiency virus derived from gorillas. *Nat. Med.* **15**, 871–872 (2009).

42. Peeters, M. *et al.* Geographical distribution of HIV-1 group O viruses in Africa. *AIDS* **11**, 493–498 (1997).

43. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS epidemic. *Cold Spring Harb. Lab. Press* (2011). doi:10.1101/cshperspect.a006841

44. Santiago, M. L. *et al.* Amplification of a Complete Simian Immunodeficiency Virus Genome from Fecal RNA of a Wild Chimpanzee. *J. Virol.* **77**, 2233–2242 (2003).

45. Gao, F. *et al.* Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **397**, 436–441 (1999).

46. Hahn, B. H., Shaw, G. M., De Cock, K. M. & Sharp, P. M. AIDS as a zoonosis: Scientific and public health implications. *Science* **287**, 607–614 (2000).

47. Bailes, E. *et al.* Hybrid Origin of SIV in Chimpanzees. *Science.* **300**, 1713–1713 (2003).

48. Morin, P. A. *et al.* Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science.* **265**, 1193–1201 (1994).

49. Hirsch, V. M., Olmsted, R. A., Murphey-Corb, M., Purcell, R. H. & Johnson, P. R. An African primate lentivirus (SIVsm closely related to HIV-2). *Nature* **339**, 389–392 (1989).

50. Gao, F. *et al.* Human infection by genetically diverse SIVSM-related HIV-2 in

West Africa. *Nature* **358**, 495–499 (1992).

51. Chen, Z. *et al.* Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J. Virol.* **71**, 3953–60 (1997).

52. Marx, P. A. A. *et al.* Isolation of a simian immunodeficiency virus related to human immunodeficiency virus type 2 from a West African pet sooty mangabey. *J. Virol.* **65**, 4480–4485 (1991).

53. Peeters, M., Toure-Kane, C. & Nkengasong, J. N. Genetic diversity of HIV in Africa: Impact on diagnosis, treatment, vaccine development and trials. *AIDS* **17**, 2547–2560 (2003).

54. Damond, F. *et al.* Quantification of proviral load of human immunodeficiency virus type 2 subtypes A and B using real-time PCR. *J. Clin. Microbiol.* (2001). doi:10.1128/JCM.39.12.4264-4268.2001

55. Chen, Z. *et al.* Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey t. *J. Virol.* **70**, 3617–3627 (1996).

56. Maddon, P. J. *et al.* The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47**, 333–348 (1986).

57. Saphire, A. C. S., Bobardt, M. D., Zhang, Z., David, G. & Gallay, P. A. Syndecans Serve as Attachment Receptors for Human Immunodeficiency Virus Type 1 on Macrophages. *J. Virol.* **75**, 9187–9200 (2001).

58. Kwong, P. D. *et al.* Structure of an HIV gp 120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).

59. Chen, B. *et al.* Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* **433**, 834–841 (2005).

60. Berger, E. A. *et al.* A new classification for HIV-1. *Nature* **391**, 240 (1998).

61. Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl. Acad. Sci.* **105**, 7552–7557 (2008).

62. Schuitemaker, H. *et al.* Biological Phenotype of Human Immunodeficiency Virus Type 1 Clones at Different Stages of Infection: Progression of Disease Is Associated with a Shift from Monocytotropic to T-Cell-Tropic Virus Populations. *J. Virol.* **66**, 1354–60 (1992).

63. Connor, R. I., Sheridan, K. E., Ceradini, D., Choe, S. & Landau, N. R. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. *J. Exp. Med.* **185**, 621–628 (2002).

64. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263–273 (1997).

65. Mankan, A. K. *et al.* Cytosolic RNA : DNA hybrids activate the cGAS – STING axis. *EMBO J.* **33**, 2937–2946 (2014).

66. Kumar, S., Morrison, J. H., Dingli, D. & Poeschla, E. HIV-1 activation of innate immunity depends strongly on the intracellular level of TREX1 and sensing of incomplete reverse transcription products. *J. Virol.* **92**, e00001-18 (2018).

67. Dix, J. A. & Verkman, A. S. Crowding Effects on Diffusion in Solutions and Cells. *Annu. Rev. Biophys.* **37**, 247–263 (2008).

68. Luby-Phelps, K. Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int. Rev. Cytol.* **192**, 189–221 (2000).

69. Arhel, N. *et al.* Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nat. Methods* **3**, 817–824 (2006).

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

71. Sabo, Y. *et al.* HIV-1 induces the formation of stable microtubules to enhance early infection. *Cell Host Microbe* **14**, 535–546 (2013).

72. Lukic, Z., Dharan, A., Fricke, T., Diaz-Griffero, F. & Campbell, E. M. HIV-1 Uncoating Is Facilitated by Dynein and Kinesin 1. *J. Virol.* **88**, 13613–25 (2014).

73. Hoogenraad, C. C. & Akhmanova, A. Bicaudal D Family of Motor Adaptors: Linking Dynein Motility to Cargo Binding. *Trends in Cell Biology* **26**, 327–340 (2016).

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

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

76. Lightfoote, M. M. *et al.* Structural characterization of reverse transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus. *J. Virol.* **60**, 771–775 (1986).

77. 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.* **90**, 6320–6324 (1993).

78. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C. & Rousso, I. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J. Virol.* **91**, e00289-17 (2017).

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

80. Huang, Y. *et al.* Primer tRNA3Lys on the Viral Genome Exists in Unextended and Two-Base Extended Forms within Mature Human Immunodeficiency Virus Type 1. *J. Virol.* **71**, 726–728 (1997).

81. Huang, Y. *et al.* Primer tRNA3Lys on the viral genome exists in unextended and two-base extended forms within mature human immunodeficiency virus type 1. *J Virol* **71**, 726–8 (1997).

82. Lori, F. *et al.* Viral DNA carried by human immunodeficiency virus type 1 virions. *J. Virol.* **66**, 5067–74 (1992).

83. Trono, D. Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses. *J. Virol.* **66**, 4893–900 (1992).

84. Yu, H., Jetzt, A. E., Ron, Y., Preston, B. D. & Dougherty, J. P. The nature of human

immunodeficiency virus type 1 strand transfers. *J. Biol. Chem.* **273**, 28384–28391 (1998).

85. Panganiban, A. T. & Fiore, D. Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science.* **241**, 1064–1069 (1988).

86. Hu, W. S. & Temin, H. M. Retroviral recombination and reverse transcription. *Science.* **250**, 1227–1233 (1990).

87. Van Wamel, J. L. B. & Berkhout, B. The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly. *Virology* **244**, 245–251 (1998).

88. Hungnesi, O., Tjøtta, E. & Grinde, B. Mutations in the central polypurine tract of HIV-1 result in delayed replication. *Virology* **190**, 440–442 (1992).

89. Fahrenkrog, B. & Aebi, U. The nuclear pore complex: Nucleocytoplasmic transport and beyond. *Nat. Rev. Mol. Cell Biol.* **4**, 757–66 (2003).

90. Fouchier, R. A. & Malim, M. H. Nuclear import of human immunodeficiency virus type-1 preintegration complexes. *Adv. Virus Res.* **52**, 275–299 (1999).

91. Pante, N. Nuclear Pore Complex Is Able to Transport Macromolecules with Diameters of 39 nm. *Mol. Biol. Cell* **13**, 425–434 (2002).

92. Mattaj, I. W. & Englmeier, L. Nucleocytoplasmic transport: the soluble phase. *Annu. Rev. Biochem.* **67**, 265–306 (1998).

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

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

95. Bejarano, D. A. *et al.* HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *Elife* **23**, e41899 (2019).

96. Engelman, A. N. & Cherepanov, P. Retroviral intasomes arising. *Curr. Opin. Struct. Biol.* **47**, 23–29 (2017).

97. Li, M., Mizuuchi, M., Burke, T. R. & Craigie, R. Retroviral DNA integration: Reaction pathway and critical intermediates. *EMBO J.* **25**, 1295–1304 (2006).

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

99. Brown, P. O., Bowerman, B., Varmus, H. E. & Bishop, J. M. Correct integration of retroviral DNA in vitro. *Cell* **8**, 347–56 (1987).

100. Cherepanov, P. *et al.* HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* **278**, 372–381 (2003).

101. Maertens, G. *et al.* LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.* **278**, 33528–33539 (2003).

102. Emiliani, S. *et al.* Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J. Biol. Chem.* **280**, 25517–25523 (2005).

103. Schröder, A. R. W. *et al.* HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–529 (2002).

104. MacNeil, A. *et al.* Genomic Sites of Human Immunodeficiency Virus Type 2 (HIV-2) Integration: Similarities to HIV-1 In Vitro and Possible Differences In Vivo. *J. Virol.* **80**, 7316–21 (2006).

105. Crise, B. *et al.* Simian Immunodeficiency Virus Integration Preference Is Similar to That of Human Immunodeficiency Virus Type 1. *J. Virol.* **79**, 12199–12204 (2005).

106. Kang, Y. *et al.* Integration Site Choice of a Feline Immunodeficiency Virus Vector. *J. Virol.* **80**, 8820–3 (2006).

107. Hacker, C. V. *et al.* The integration profile of EIAV-based vectors. *Mol. Ther.* **14**, 536–545 (2006).

108. Llano, M. *et al.* LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J. Virol.* **78**, 9524–9537 (2004).

109. Cherepanov, P. LEDGF/p75 interacts with divergent lentiviral integrases and

modulates their enzymatic activity in vitro. *Nucleic Acids Res.* **35**, 113–124 (2007).

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

111. Shun, M.-C. *et al.* LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **21**, 1767–1778 (2007).

112. Schaller, T. *et al.* HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog.* **7**, e1002439 (2011).

113. Pauza, C. D. Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. *Virology* **179**, 886–9 (1990).

114. Engelman, A., Mizuuchi, K. & Craigie, R. HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell* **67**, 1211–1221 (1991).

115. Yoder, K. E. & Bushman, F. D. Repair of Gaps in Retroviral DNA Integration Intermediates. *J. Virol.* **74**, 11191–11200 (2002).

116. Skalka, A. M. & Katz, R. A. Retroviral DNA integration and the DNA damage response. *Cell Death Differ.* **12**, 971–978 (2005).

117. Sloan, R. D. & Wainberg, M. A. The role of unintegrated DNA in HIV infection. *Retrovirology* **8**, 1–15 (2011).

118. Bukrinsky, M. I. *et al.* Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 6580–6584 (1992).

119. Butler, S. L., Hansen, M. S. T. & Bushman, F. D. A quantitative assay for HIV DNA integration in vivo. *Nat. Med.* **7**, 631–4 (2001).

120. Klaver, B. & Berkhout, B. Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. *J. Virol.* **68**, 3830–40 (1994).

121. Pereira, L. A., Bentley, K., Peeters, A., Churchill, M. J. & Deacon, N. J. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids Res.* **28**, 663–668 (2000).

122. Rafati, H. *et al.* Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLoS Biol.* (2011). doi:10.1371/journal.pbio.1001206

123. Nabel, G. & Baltimore, D. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* **326**, 711–713 (1987).

124. Kinoshita, S. *et al.* The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells. *Immunity* **6**, 235–244 (1997).

125. Natarajan, M. *et al.* Negative elongation factor (NELF) coordinates RNA polymerase II pausing, premature termination, and chromatin remodeling to regulate HIV transcription. *J. Biol. Chem.* **288**, 25995–26003 (2013).

126. Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H. & Jones, K. A. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92**, 451–462 (1998).

127. Kim, Y. K., Bourgeois, C. F., Isel, C., Churcher, M. J. & Karn, J. Phosphorylation of the RNA Polymerase II Carboxyl-Terminal Domain by CDK9 Is Directly Responsible for Human Immunodeficiency Virus Type 1 Tat-Activated Transcriptional Elongation. *Mol. Cell. Biol.* **22**, 4622–4637 (2002).

128. 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 (1996).

129. Legrain, P. & Rosbash, M. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell* **57**, 573–583 (1989).

130. Fischer, U. *et al.* Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. *EMBO J.* **13**, 4105–4112 (1994).

131. Felber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T. & Pavlakis, G. N. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc. Natl. Acad. Sci.* **86**, 1495–1499 (2006).

132. 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 (1989).

133. Zapp, M. L., Hope, T. J., Parslow, T. G. & Green, M. R. Oligomerization and RNA

binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. *Proc. Natl. Acad. Sci.* **88**, 7734–7738 (1991).

134. Malim, M. H. & Cullen, B. R. HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: Implications for HIV-1 latency. *Cell* **65**, 241–248 (1991).

135. Booth, D. S., Cheng, Y. & Frankel, A. D. The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. *Elife* **3**, e04121 (2014).

136. Jacks, T. *et al.* Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **21**, 280–3 (1988).

137. Wilson, W. *et al.* HIV expression strategies: Ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* **55**, 1159–1169 (1988).

138. Mouzakis, K. D., Lang, A. L., Vander Meulen, K. A., Easterday, P. D. & Butcher, S. E. HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res.* **41**, 1901–1913 (2013).

139. Aloia, R. C., Tian, H. & Jensen, F. C. Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes. *Proc. Natl. Acad. Sci. U. S. A.* **1**, 5181–5 (1993).

140. Chan, R. *et al.* Retroviruses Human Immunodeficiency Virus and Murine Leukemia Virus Are Enriched in Phosphoinositides. *J. Virol.* **82**, 11228–38 (2008).

141. Nguyen, D. H. & Hildreth, J. E. K. Evidence for Budding of Human Immunodeficiency Virus Type 1 Selectively from Glycolipid-Enriched Membrane Lipid Rafts. *J. Virol.* **74**, 3264–72 (2000).

142. 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.* **101**, 14889–94 (2004).

143. Saad, J. S. *et al.* Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl. Acad. Sci.* **103**, 11364–9 (2006).

144. Jouvenet, N., Simon, S. M. & Bieniasz, P. D. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc. Natl. Acad. Sci.* **106**, 19114–9 (2009).

145. Kutluay, S. B. & Bieniasz, P. D. Analysis of the initiating events in HIV-1 particle assembly and genome packaging. *PLoS Pathog.* **6**, e1001200 (2010).

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

147. Smyth, R. P. *et al.* Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat. Methods* **12**, 866–872 (2015).

148. Comas-Garcia, M., Davis, S. & Rein, A. On the Selective Packaging of Genomic RNA by HIV-1. *Viruses 2016,* **8**, 246 (2016).

149. El-Wahab, E. W. A. *et al.* Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat. Commun.* **2**, 4304 (2014).

150. Houzet, L. *et al.* HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. *Nucleic Acids Res.* **35**, 2695–704. (2007).

151. Webb, J. A., Jones, C. P., Parent, L. J., Rouzina, I. & Musier-Forsyth, K. Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: Implications for viral genomic RNA packaging. *RNA* **19**, 1078–88 (2013).

152. 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**, 7190–200 (1993).

153. Gorelick, R. J. *et al.* Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* **64**, 3207–11 (1990).

154. Carlson, L. A. *et al.* Cryo electron tomography of native HIV-1 budding sites. *PLoS Pathog.* **6**, e1001173 (2010).

155. Fuller, S. D., Wilk, T., Gowen, B. E., Kräusslich, H. G. & Vogt, V. M. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.* **7**, 729–738 (1997).

156. Yang, P. *et al.* The Cytoplasmic Domain of Human Immunodeficiency Virus Type 1 Transmembrane Protein gp41 Harbors Lipid Raft Association Determinants. *J. Virol.* **84**, 59–75 (2010).

157. Rousso, I., Mixon, M. B., Chen, B. K. & Kim, P. S. Palmitoylation of the HIV-1 envelope glycoprotein is critical for viral infectivity. *Proc. Natl. Acad. Sci.* **97**, 13523–13525 (2000).

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

159. Cosson, P. Direct interaction between the envelope and matrix proteins of HIV-1. *EMBO J.* **15**, 5783–8 (1996).

160. Murakami, T. & Freed, E. O. Genetic Evidence for an Interaction between Human Immunodeficiency Virus Type 1 Matrix and α-Helix 2 of the gp41 Cytoplasmic Tail. *J. Virol.* **74**, 3548–3554 (2000).

161. Pornillos, O. *et al.* HIV Gag mimics the Tsg101-recruiting activity of the human Hrs protein. *J. Cell Biol.* **162**, 425–434 (2003).

162. Pornillos, O., Alam, S. L., Davis, D. R. & Sundquist, W. I. Structure of the tsg101 uev domain in complex with the ptap motif of the hiv-1 p6 protein. *Nat. Struct. Biol.* **9**, 812–817 (2002).

163. Im, Y. J. *et al.* Crystallographic and functional analysis of the ESCRT-I /HIV-1 Gag PTAP interaction. *Structure* **18**, 1536–1547 (2010).

164. 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 (2003).

165. Vincent, O., Rainbow, L., Tilburn, J., Arst, H. N. J. & Penalva, M. A. YPXL/I is a protein interaction motif recognized by aspergillus PalA and its human homologue, AIP1/Alix. *Mol. Cell. Biol.* **23**, 1647–1655 (2003).

166. Weber, I. T. *et al.* Molecular modeling of the HIV-1 protease and its substrate binding site. *Science.* **243**, 928–31 (1989).

167. Lapatto, R. *et al.* X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature* **342**, 299–302 (1989).

168. Navia, M. A. *et al.* Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615–20 (1989).

169. 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 (2008).

170. Kräusslich, H. G. Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3213–3217 (1991).

171. 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**, 8917–27 (1994).

172. de Marco, A. *et al.* Structural analysis of HIV-1 maturation using cryo- electron tomography. *PLoS Pathog.* **6**, e1001215 (2010).

173. Ohishi, M. *et al.* The relationship between HIV-1 genome RNA dimerization, virion maturation and infectivity. *Nucleic Acids Res.* **39**, 3404–17 (2011).

174. Margottin, F. *et al.* A novel human WD protein, h-βTrCP, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol. Cell* **1**, 565–574 (1998).

175. Ross, T. M., Oran, A. E. & Cullen, B. R. Inhibition of HIV-1 progeny virion release by cell-surface CD4 is relieved by expression of the viral Nef protein. *Curr. Biol.* **9**, 613–621 (1999).

176. Lama, J., Mangasarian, A. & Trono, D. Cell-surface expression of CD4 reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner. *Curr. Biol.* **9**, 622–631 (1999).

177. Neil, S. J. D., Sandrin, V., Sundquist, W. I. & Bieniasz, P. D. An Interferon-α-Induced Tethering Mechanism Inhibits HIV-1 and Ebola Virus Particle Release but Is Counteracted by the HIV-1 Vpu Protein. *Cell Host Microbe* **2**, 193–203 (2007).

178. Neil, S. J. D., Eastman, S. W., Jouvenet, N. & Bieniasz, P. D. HIV-1 Vpu promotes release and prevents endocytosis of nascent retrovirus particles from the plasma membrane. *PLoS Pathog.* **2**, 354–367 (2006).

179. Neil, S. J. D., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425–430 (2008).

180. Van Damme, N. *et al.* The Interferon-Induced Protein BST-2 Restricts HIV-1 Release and Is Downregulated from the Cell Surface by the Viral Vpu Protein. *Cell Host Microbe* **3**, 245–252 (2008).

181. Kupzig, S. *et al.* Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. *Traffic* **4**, 694–709 (2003).

182. Schubert, U. *et al.* The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J. Virol.* **70**, 809–19 (1996).

183. Chaudhuri, R., Lindwasser, O. W., Smith, W. J., Hurley, J. H. & Bonifacino, J. S. Downregulation of CD4 by human immunodeficiency virus type 1 Nef is dependent on clathrin and involves direct interaction of Nef with the AP2 clathrin adaptor. *J. Virol.* **81**, 3877–3890 (2007).

184. Schwartz, O., Maréchal, V., Le Gall, S., Lemonnier, F. & Heard, J. M. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nat. Med.* **2**, 338–342 (1996).

185. Baur, A. S. *et al.* HIV-1 nef leads to inhibition or activation of T cells depending on its intracellular localization. *Immunity* **1**, 373–384 (1994).

186. Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D. & Baltimore, D. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* **391**, 397–401 (1998).

187. Thoulouze, M. I. *et al.* Human Immunodeficiency Virus Type-1 Infection Impairs the Formation of the Immunological Synapse. *Immunity* **24**, 547–561 (2006).

188. Roeth, J. F., Kasper, M. R., Filzen, T. M. & Collins, K. L. HIV-1 Nef disrupts MHC-I trafficking by recruiting AP-1 to the MHC-I cytoplasmic tail. *J. Cell Biol.* **167**, 903–913 (2004).

189. Lubben, N. B. *et al.* HIV-1 Nef-induced Down-Regulation of MHC Class I Requires AP-1 and Clathrin but Not PACS-1 and Is Impeded by AP-2. *Mol. Biol. Cell* **18**, 3351–3365 (2007).

190. Rosa, A. *et al.* HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. *Nature* **526**, 212–217 (2015).

191. Usami, Y., Wu, Y., Göttlinger, H. G. & Go, H. G. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature* **526**, 218–223 (2015).

192. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650 (2002).

193. Conticello, S. G., Thomas, C. J. F., Petersen-Mahrt, S. K. & Neuberger, M. S. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol. Biol. Evol.* **22**, 367–377 (2005).

194. Bogerd, H. P. & Cullen, B. R. Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation. *RNA* **14**, 1228–1236 (2008).

195. Harris, R. S. & Liddament, M. T. Retroviral restriction by APOBEC proteins. *Nature Reviews Immunology* **4**, 868–877 (2004).

196. Holmes, R. K., Malim, M. H. & Bishop, K. N. APOBEC-mediated viral restriction: not simply editing? *Trends in Biochemical Sciences* **32**, 118–128 (2007).

197. Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M. & Malim, M. H. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog.* **4**, e1000231 (2008).

198. Iwatani, Y. *et al.* Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res.* **35**, 7096–7108 (2007).

199. Wang, X. *et al.* The Cellular Antiviral Protein APOBEC3G Interacts with HIV-1 Reverse Transcriptase and Inhibits Its Function during Viral Replication. *J. Virol.* **86**, 3777–3786 (2012).

200. Mehle, A., Goncalves, J., Santa-Marta, M., McPike, M. & Gabuzda, D. Phosphorylation of a novel SOCS-box regulates assembly of the HIV-1 Vif-Cul5 complex that promotes APOBEC3G degradation. *Genes Dev.* **18**, 2861–2866

(2004).

201. Yu, X. F. X. *et al.* Induction of APOBEC3G Ubiquitination and Degradation by an HIV-1 Vif-Cul5-SCF Complex. *Science.* **302**, 1056–1060 (2003).

202. Kobayashi, M., Takaori-Kondo, A., Miyauchi, Y., Iwai, K. & Uchiyama, T. Ubiquitination of APOBEC3G by an HIV-1 Vif-Cullin5-Elongin B-Elongin C complex is essential for Vif function. *J. Biol. Chem.* **280**, 18573–18578 (2005).

203. Zhu, H., Jian, H. & Zhao, L. J. Identification of the 15FRFG domain in HIV-1 Gag p6 essential for Vpr packaging into the virion. *Retrovirology* (2004). doi:10.1186/1742-4690-1-26

204. Cohen, E. A. *et al.* Identification of HIV-1 vpr product and function. *J. Acquir. Immune Defic. Syndr.* **3**, 11–18 (1990).

205. Balotta, C., Lusso, P., Crowley, R., Gallo, R. C. & Franchini, G. Antisense phosphorothioate oligodeoxynucleotides targeted to the vpr gene inhibit human immunodeficiency virus type 1 replication in primary human macrophages. *J Virol* **67**, 4409–14 (1993).

206. Lu, Y.-L. L., Spearman, P. & Ratner, L. Human Immunodeficiency Virus Type 1 Viral Protein R Localization in Infected Cells and Virions. *J. Virol.* **67**, 6542–6550 (1993).

207. Desai, T. M. *et al.* Fluorescent protein-tagged Vpr dissociates from HIV-1 core after viral fusion and rapidly enters the cell nucleus. *Retrovirology* **12**, 1–20 (2015).

208. He, J. *et al.* Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. *J. Virol.* **69**, 6705–11 (1995).

209. Rogel, M. E., Wu, L. I. & Emerman, M. The human immunodeficiency virus type 1 vpr gene prevents cell proliferation during chronic infection. *J. Virol.* **69**, 882–8 (1995).

210. Jowett, J. B. *et al.* The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *J. Virol.* **69**, 6304–13 (1995).

211. Goh, W. C. *et al.* HIV-1 Vpr increases viral expression by manipulation of the cell

cycle: A mechanism for selection of Vpr in vivo. *Nat. Med.* **4**, 65–71 (1998).

212. Laguette, N. *et al.* Premature activation of the SLX4 complex by Vpr promotes G2/M arrest and escape from innate immune sensing. *Cell* **156**, 134–145 (2014).

213. Vodicka, M. A., Koepp, D. M., Silver, P. A. & Emerman, M. HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection. *Genes Dev.* **12**, 175–85 (1998).

214. Popov, S., Rexach, M., Ratner, L., Blobel, G. & Bukrinsky, M. Viral protein R regulates docking of the HIV-1 preintegration complex to the nuclear pore complex. *J. Biol. Chem.* **273**, 13347–13352 (1998).

215. Heinzinger, N. K. *et al.* The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 7311–7315 (1994).

216. Clish, C. B., Peyton, D. H. & Barklis, E. Solution structures of human immunodeficiency virus type 1 (HIV-1) and moloney murine leukemia virus (MoMLV) capsid protein major-homology-region peptide analogs by NMR spectroscopy. *Eur. J. Biochem.* **257**, 69–77 (1998).

217. Kingston, R. L., Olson, N. H. & Vogt, V. M. The organization of mature Rous sarcoma virus as studied by cryoelectron microscopy. *J. Struct. Biol.* **136**, 67–80 (2001).

218. Weiland, F., Matheka, H. D., Coggins, L. & Härtner, D. Electron microscopic studies on equine infectious anemia virus (EIAV). *Arch. Virol.* **55**, 335–340 (1977).

219. Boothe, A. D. & Van Der Maaten, M. J. Ultrastructural studies of a visna like syncytia producing virus from cattle with lymphocytosis. *J. Virol.* **13**, 197–204 (1974).

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

221. Pornillos, O., Ganser-Pornillos, B. K. & Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **469**, 424–7 (2011).

222. 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).

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

224. Briggs, J. A. G. *et al.* The stoichiometry of Gag protein in HIV-1. *Nat. Struct. Mol. Biol.* **11**, 672–5 (2004).

225. Gitti, R. K. *et al.* Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science.* **273**, 231–5 (1996).

226. Gamble, T. R. *et al.* Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science.* **278**, 849–53 (1997).

227. Pornillos, O. *et al.* X-Ray Structures of the Hexameric Building Block of the HIV Capsid. *Cell* **137**, 1282–1292 (2009).

228. Pornillos, O., Ganser-Pornillos, B. K., Banumathi, S., Hua, Y. & Yeager, M. Disulfide Bond Stabilization of the Hexameric Capsomer of Human Immunodeficiency Virus. *J. Mol. Biol.* **401**, 985–995 (2010).

229. Gres, A. T. *et al.* X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **349**, 99–104 (2015).

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

231. Forshey, B. B. M., Schwedler, U. Von, 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**, 5667–77 (2002).

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

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

234. Burdick, R. C. *et al.* Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog.* **13**, 1–38 (2017).

235. Yan, N., Regalado-Magdos, A. D., Stiggelbout, B., Lee-Kirsch, M. A. & Lieberman, J. The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat. Immunol.* **11**, 1005–1013 (2010).

236. Gao, D. *et al.* Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science.* **341**, 903–906 (2013).

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

238. Rihn, S. J. *et al.* Extreme Genetic Fragility of the HIV-1 Capsid. *PLoS Pathog.* **9**, e1003461 (2013).

239. Mallery, D. L. *et al.* IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *Elife* **7**, e35335 (2018).

240. Lilly, F. Susceptibility to two strains of Friend leukemia virus in mice. *Science.* **155**, 461–462 (1967).

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

242. Kozak, C. A. & Chakraborti, A. Single amino acid changes in the murine leukemia virus capsid protein gene define the target of Fv1 resistance. *Virology* **225**, 300–305 (1996).

243. Gautsch, J. W., Elder, J. H., Schindler, J., Jensen, F. C. & Lerner, R. A. Structural markers on core protein p30 of murine leukemia virus: functional correlation with Fv-1 tropism. *Proc. Natl. Acad. Sci.* **75**, 4170–4174 (1978).

244. Hopkins, N., Schindler, J. & Hynes, R. Six-NB-tropic murine leukemia viruses derived from a B-tropic virus of BALB/c have altered p30. *J Virol* **21**, 309–318 (1977).

245. Hartley, J. W., Rowe, W. P. & Huebner, R. J. Host-range restrictions of murine leukemia viruses in mouse embryo cell cultures. *J. Virol.* **5**, 221–5 (1970).

246. Towers, G. *et al.* A conserved mechanism of retrovirus restriction in mammals. *Proc. Natl. Acad. Sci.* **97**, 12295–12299 (2000).

247. Hatziioannou, T., Cowan, S., Goff, S. P., Bieniasz, P. D. & Towers, G. J.

Restriction of multiple divergent retroviruses by Lv1 and Ref1. *EMBO J.* **22**, 385–394 (2003).

248. Stremlau, M. *et al.* The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848–853 (2004).

249. Perez-Caballero, D., Hatziioannou, T., Yang, A., Cowan, S. & Bieniasz, P. D. Human Tripartite Motif 5 Domains Responsible for Retrovirus Restriction Activity and Specificity. *J. Virol.* **79**, 8969–8978 (2005).

250. Wu, X., Anderson, J. L., Campbell, E. M., Joseph, A. M. & Hope, T. J. Proteasome inhibitors uncouple rhesus TRIM5 restriction of HIV-1 reverse transcription and infection. *Proc. Natl. Acad. Sci.* **103**, 7465–7470 (2006).

251. Stremlau, M. *et al.* Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5 restriction factor. *Proc. Natl. Acad. Sci.* **103**, 5514–5519 (2006).

252. Kutluay, S. B., Perez-Caballero, D. & Bieniasz, P. D. Fates of Retroviral Core Components during Unrestricted and TRIM5-Restricted Infection. *PLoS Pathog.* **9**, e1003214 (2013).

253. Diaz-Griffero, F. *et al.* Rapid turnover and polyubiquitylation of the retroviral restriction factor TRIM5. *Virology* **349**, 300–315 (2006).

254. Wagner, J. M. *et al.* Mechanism of B-box 2 domain-mediated higher-order assembly of the retroviral restriction factor TRIM5α. *Elife* **5**, e16309 (2016).

255. Ganser-Pornillos, B. K. *et al.* Hexagonal assembly of a restricting TRIM5 protein. *Proc. Natl. Acad. Sci.* **108**, 534–9 (2011).

256. Rold, C. J. & Aiken, C. Proteasomal degradation of TRIM5α during retrovirus restriction. *PLoS Pathog.* **4**, e1000074 (2008).

257. Selyutina, A. *et al.* Cyclophilin A Prevents HIV-1 Restriction in Lymphocytes by Blocking Human TRIM5α Binding to the viral Core. *bioRxiv* 678037 (2019). doi:10.1101/678037

258. Kim, K. *et al.* Cyclophilin A protects HIV-1 from restriction by human TRIM5α. *bioRxiv* 587907 (2019). doi:10.1101/587907

259. Hatziioannou, T., Perez-Caballero, D., Yang, A., Cowan, S. & Bieniasz, P. D.

Retrovirus resistance factors Ref1 and Lv1 are species-specific variants of TRIM5 . *Proc. Natl. Acad. Sci.* **101**, 10774–10779 (2004).

260. Saenz, D. T., Teo, W., Olsen, J. C., Eric, M. & Poeschla, E. M. Restriction of Feline Immunodeficiency Virus by Ref1 , Lv1 , and Primate TRIM5 α Proteins. *Soc. v* **79**, 15175–15188 (2005).

261. Keckesova, Z., Ylinen, L. M. J. & Towers, G. J. The human and African green monkey TRIM5 genes encode Ref1 and Lv1 retroviral restriction factor activities. *Proc. Natl. Acad. Sci.* **101**, 10780–10785 (2004).

262. Uchil, P. D. *et al.* TRIM Protein-Mediated Regulation of Inflammatory and Innate Immune Signaling and Its Association with Antiretroviral Activity. *J. Virol.* **87**, 257–272 (2013).

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

264. Fletcher, A. J. *et al.* Trivalent RING Assembly on Retroviral Capsids Activates TRIM5 Ubiquitination and Innate Immune Signaling. *Cell Host Microbe* **24**, 761–775 (2018).

265. Wilson, S. J. *et al.* Independent evolution of an antiviral TRIMCyp in rhesus macaques. *Proc. Natl. Acad. Sci.* **105**, 3557–3562 (2008).

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

267. Liu, Z. *et al.* The Interferon-Inducible MxB Protein Inhibits HIV-1 Infection. *Cell Host Microbe* **14**, 398–410 (2013).

268. Kane, M. *et al.* MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature* **502**, 563–566 (2013).

269. Goujon, C. *et al.* Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **502**, 559–562 (2013).

270. Fricke, T. *et al.* MxB binds to the HIV-1 core and prevents the uncoating process of HIV-1. *Retrovirology* **11**, 1–14 (2014).

271. Schulte, B. *et al.* Restriction of HIV-1 Requires the N-Terminal Region of MxB as

a Capsid-Binding Motif but Not as a Nuclear Localization Signal. *J. Virol.* **89**, 8599–8610 (2015).

272. Buffone, C., Schulte, B., Opp, S. & Diaz-Griffero, F. Contribution of MxB Oligomerization to HIV-1 Capsid Binding and Restriction. *J. Virol.* **89**, 3285–3294 (2015).

273. Wang, Q., Du, X., Cai, Z. & Greene, M. I. Characterization of the Structures Involved in Localization of the SUN Proteins to the Nuclear Envelope and the Centrosome. *DNA Cell Biol.* **25**, 554–562 (2006).

274. Hodzic, D. M., Yeater, D. B., Bengtsson, L., Otto, H. & Stahl, P. D. Sun2 is a novel mammalian inner nuclear membrane protein. *J. Biol. Chem.* **279**, 25805–25812 (2004).

275. Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type 1 interferon antiviral response. *Nature* **472**, 481–485 (2011).

276. Sosa, B. A., Kutay, U. & Schwartz, T. U. Structural insights into LINC complexes. *Curr. Opin. Struct. Biol.* **23**, 285–291 (2013).

277. Starr, D. A. & Fridolfsson, H. N. Interactions Between Nuclei and the Cytoskeleton Are Mediated by SUN-KASH Nuclear-Envelope Bridges. *Annu. Rev. Cell Dev. Biol.* **26**, 421–444 (2010).

278. Chang, W., Worman, H. J. & Gundersen, G. G. Accessorizing and anchoring the LINC complex for multifunctionality. *Journal of Cell Biology* **208**, 11–22 (2015).

279. Rothballer, A. & Kutay, U. The diverse functional LINCs of the nuclear envelope to the cytoskeleton and chromatin. Chromosoma **122**, 415–429 (2013).

280. Donahue, D. A. *et al.* SUN2 Overexpression Deforms Nuclear Shape and Inhibits HIV. *J. Virol.* **90**, 4199–4214 (2016).

281. Sun, W.-W. *et al.* SUN2 Modulates HIV-1 Infection and Latency through Association with Lamin A/C To Maintain the Repressive Chromatin. *MBio* **9**, e02408-17 (2018).

282. Pizzato, M. *et al.* Lv4 Is a Capsid-Specific Antiviral Activity in Human Blood Cells That Restricts Viruses of the SIVMAC/SIVSm/HIV-2 Lineage Prior to Integration. *PLoS Pathog.* **11**, 1–29 (2015).

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

284. Price, A. J. *et al.* CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication. *PLoS Pathog.* **8**, e1002896 (2012).

285. Matreyek, K. A. & Engelman, A. The Requirement for Nucleoporin NUP153 during Human Immunodeficiency Virus Type 1 Infection Is Determined by the Viral Capsid. *J. Virol.* **85**, 7818–7827 (2011).

286. Matreyek, K. A., Yücel, S. S., Li, X. & Engelman, A. Nucleoporin NUP153 Phenylalanine-Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to Mediate Lentiviral Infectivity. *PLoS Pathog.* **9**, e1003693 (2013).

287. Bichel, K. *et al.* HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. *Retrovirology* **10**, 1–12 (2013).

288. König, R. *et al.* Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **135**, 49–60 (2008).

289. Brass, A. L. *et al.* Identification of host proteins required for HIV infection through a functional genomic screen. *Science.* **319**, 921–926 (2008).

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

291. Ryffel, B. *et al.* Distribution of the cyclosporine binding protein cyclophilin in human tissues. *Immunology* **72**, 399–404 (1991).

292. Marks, W. H., Harding, M. W., Handschumacher, R., Marks, C. & Lorber, M. I. The immunochemical distribution of cyclophilin in normal mammalian tissues. *Transplantation* **52**, 340–345 (1991).

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

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

295. Braaten, D., Ansari, H. & Luban, J. The hydrophobic pocket of cyclophilin is the

binding site for the human immunodeficiency virus type 1 Gag polyprotein. *J. Virol.* **71**, 2107–13 (1997).

296. Franke, E. K., Yuan, H. E. H. & Luban, J. Specific incorporation of cyclophilin a into HIV-1 virions. *Nature* **372**, 359–62 (1994).

297. Braaten, D., Franke, E. K. & Luban, J. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *J. Virol.* **70**, 3551–60 (1996).

298. 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 (1996).

299. Lin, T. Y. & Emerman, M. Cyclophilin A interacts with diverse lentiviral capsids. *Retrovirology* **3**, 1–12 (2006).

300. Katzourakis, A., Tristem, M., Pybus, O. G. & Gifford, R. J. Discovery and analysis of the first endogenous lentivirus. *Proc Natl Acad Sci U S A* **104**, 6261–6265 (2007).

301. Keckesova, Z., Ylinen, L. M. J., Towers, G. J., Gifford, R. J. & Katzourakis, A. Identification of a RELIK orthologue in the European hare (Lepus europaeus) reveals a minimum age of 12 million years for the lagomorph lentiviruses. *Virology* **384**, 7–11 (2009).

302. Goldstone, D. C. *et al.* Structural and Functional Analysis of Prehistoric Lentiviruses Uncovers an Ancient Molecular Interface. *Cell Host Microbe* **8**, 248–259 (2010).

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

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

305. Handschumacher, R. E., Harding, M. W., Rice, J., Drugge, R. J. & Speicher, D. W. Cyclophilin: A specific cytosolic binding protein for cyclosporin A. *Science.* **226**, 544–7 (1984).

306. Kallen, J. *et al.* Structure of human cyclophilin and its binding site for cyclosporin A determined by X-ray crystallography and NMR spectroscopy. *Nature* **353**, 276–9 (1991).

307. Karpas, A., Lowdell, M., Jacobson, S. K. & Hill, F. Inhibition of human immunodeficiency virus and growth of infected T cells by the immunosuppressive drugs cyclosporin A and FK 506. *Proc. Natl. Acad. Sci.* **89**, 8351–5 (2006).

308. Bartz, S. R., Hohenwalter, E., Hu, M. K., Rich, D. H. & Malkovsky, M. Inhibition of human immunodeficiency virus replication by nonimmunosuppressive analogs of cyclosporin A. *Proc. Natl. Acad. Sci.* **92**, 5381–5 (2006).

309. Billich, A. *et al.* Mode of action of SDZ NIM 811, a nonimmunosuppressive cyclosporin A analog with activity against human immunodeficiency virus (HIV) type 1: interference with HIV protein-cyclophilin A interactions. *J. Virol.* **69**, 2451–61 (1995).

310. Franke, E. K. & Luban, J. Inhibition of HIV-1 replication by cyclosporine A or related compounds correlates with the ability to disrupt the Gag-cyclophilin A interaction. *Virology* **222**, 279–282 (1996).

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

312. Rüegsegger, U., Beyer, K. & Keller, W. Purification and characterization of human cleavage factor Im involved in the 3′ end processing of messenger RNA precursors. *J. Biol. Chem.* **271**, 6107–13 (1996).

313. 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 (2014).

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

315. Frey, S., Richter, R. P. & Görlich, D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science.* **314**, 815–7 (2006).

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

317. Meehan, A. M. *et al.* A Cyclophilin Homology Domain-Independent Role for Nup358 in HIV-1 Infection. *PLoS Pathog.* **20**, e1003969 (2014).

318. 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. *J. Virol.* **86**, 5931–5936 (2012).

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

320. Christ, F. *et al.* Transportin-SR2 Imports HIV into the Nucleus. *Curr. Biol.* **18**, 1192–1202 (2008).

321. Logue, E. C., Taylor, K. T., Goff, P. H. & Landau, N. R. The Cargo-Binding Domain of Transportin 3 Is Required for Lentivirus Nuclear Import. *J. Virol.* **85**, 12950–12961 (2011).

322. De Houwer, S. *et al.* Identification of residues in the C-terminal domain of HIV-1 integrase that mediate binding to the transportin-SR2 protein. *J. Biol. Chem.* **287**, 34059–34068 (2012).

323. 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 (2010).

324. Cribier, A. *et al.* Mutations affecting interaction of integrase with TNPO3 do not prevent HIV-1 cDNA nuclear import. *Retrovirology* **8**, 104 (2011).

325. Lai, M. C., Lin, R. I., Huang, S. Y., Tsai, C. W. & Tarn, W. Y. A human importin-β family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *J. Biol. Chem.* **17**, 7950–7 (2000).

326. Lai, M.-C., Lin, R.-I. & Tarn, W.-Y. Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proc. Natl. Acad. Sci.* **98**, 10154–9 (2001).

327. Rüegsegger, U., Blank, D. & Keller, W. Human pre-mRNA cleavage factor Im Is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol. Cell* **1**, 243–53 (1998).

328. De Iaco, A. *et al.* TNPO3 protects HIV-1 replication from CPSF6-mediated capsid stabilization in the host cell cytoplasm. *Retrovirology* **10**, 1–18 (2013).

329. Sasakawa, N., Sharif, M. & Hanley, M. R. Metabolism and biological activities of inositol pentakisphosphate and inositol hexakisphosphate. *Biochem. Pharmacol.* **50**, 137–46 (1995).

330. Hanley, M. R. *et al.* Neural function: metabolism and actions of inositol metabolites in mammalian brain. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **320**, 381–98 (1988).

331. Vallejo, M., Jackson, T., Lightman, S. & Hanley, M. R. Occurrence and extracellular actions of inositol pentakis- and hexakisphosphate in mammalian brain. *Nature* **330**, 656–8 (1987).

332. Efanov, A. M., Zaitsev, S. V & Berggren, P. O. Inositol hexakisphosphate stimulates non-Ca2+-mediated and primes Ca2+-mediated exocytosis of insulin by activation of protein kinase C. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 4435–4439 (1997).

333. Larsson, O. *et al.* Inhibition of phosphatases and increased Ca 2+ channel activity by inositol hexakisphosphate. *Science.* **278**, 471–4 (1997).

334. Hawkins, P. T. *et al.* Inhibition of iron-catalysed hydroxyl radical formation by inositol polyphosphates: a possible physiological function for myo-inositol hexakisphosphate. *Biochem. J.* **294**, 929–924 (2015).

335. Graf, E., Empson, K. L. & Eaton, J. W. Phytic acid. A natural antioxidant. *J. Biol. Chem.* **262**, 11647–50 (1987).

336. 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**, 721–9 (2000).

337. Okada, M. & Ye, K. Nuclear phosphoinositide signaling regulates messenger RNA export. *RNA Biol.* **6**, 1–6 (2009).

338. York, J. D., Odom, A. R., Murphy, R., Ives, E. B. & Wente, S. R. A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science.* **285**, 96–100 (1999).

339. Odom, A. R., Stahlberg, A., Wente, S. R. & York, J. D. A role for nuclear inositol 1,4,5-trisphosphate kinase in transcriptional control. *Science.* **287**, 2026–9 (2000).

340. Campbell, S. *et al.* Modulation of HIV-like particle assembly in vitro by inositol phosphates. *Proc. Natl. Acad. Sci.* **98**, 10875–9 (2001).

341. Vermeire, J. *et al.* Quantification of Reverse Transcriptase Activity by Real- Time PCR as a Fast and Accurate Method for Titration of HIV , Lenti- and Retroviral Vectors. *PLoS One* **7**, e50859 (2012).

342. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. (2004).

343. Gouy, M., Guindon, S. & Gascuel, O. Sea view version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).

344. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

345. Monit, C., Goldstein, R. A. & Towers, G. J. ChromaClade: combined visualisation of phylogenetic and sequence data. *BMC Evol. Biol.* **19**, 186 (2019).

346. Price, A. J. *et al.* Active site remodeling switches HIV specificity of antiretroviral TRIMCyp. *Nat. Struct. Mol. Biol.* **10**, 1036–42 (2009).

347. McCarthy, K. R. *et al.* Gain-of-Sensitivity Mutations in a Trim5-Resistant Primary Isolate of Pathogenic SIV Identify Two Independent Conserved Determinants of Trim5α Specificity. *PLoS Pathog.* **9**, e1003352 (2013).

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

349. 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**, E335 (2017).

350. Cornilescu, C. C., Bouamr, F., Yao, X., Carter, C. & Tjandra, N. Structural analysis of the N-terminal domain of the human T-cell leukemia virus capsid protein. *J. Mol. Biol.* **4**, 783–97 (2001).

351. Obal, G. *et al.* Conformational plasticity of a native retroviral capsid revealed by x-ray crystallography. *Science.* **49**, 95–8 (2015).

352. Bailey, G. D., Hyun, J. K., Mitra, A. K. & Kingston, R. L. A structural model for

the generation of continuous curvature on the surface of a retroviral capsid. *J. Mol. Biol.* **417**, 212–23 (2012).

353. Mortuza, G. B. *et al.* Structure of B-MLV Capsid Amino-terminal Domain Reveals Key Features of Viral Tropism, Gag Assembly and Core Formation. *J. Mol. Biol.* **376**, 1493–1508 (2008).

354. Davey, N. E. *et al.* The HIV Mutation Browser: A Resource for Human Immunodeficiency Virus Mutagenesis and Polymorphism Data. *PLoS Comput. Biol.* **10**, (2014).

355. 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**, 5439–50 (2003).

356. Melamed, D. *et al.* The Conserved Carboxy Terminus of the Capsid Domain of Human Immunodeficiency Virus Type 1 Gag Protein Is Important for Virion Assembly and Release. *J. Virol.* **78**, 9675–9688 (2004).

357. Yap, M. W. & Stoye, J. P. Apparent effect of rabbit endogenous lentivirus type K acquisition on retrovirus restriction by lagomorph Trim5αs. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, (2013).

358. Wannes, T. *et al.* Interplay between HIV Entry and Transportin-SR2 Dependency. *Retrovirology* **8**, 1–17 (2011).

359. 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).

360. Fassati, A. & Goff, S. P. Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. *J. Virol.* **73**, 8919–8925 (1999).

361. McEwan, W. A. *et al.* Truncation of TRIM5 in the Feliformia Explains the Absence of Retroviral Restriction in Cells of the Domestic Cat. *J. Virol.* **83**, 8270–5 (2009).

362. Lahouassa, H. *et al.* SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat. Immunol.* **13**, 223–228 (2012).

363. Marini, B. *et al.* Nuclear architecture dictates HIV-1 integration site selection. *Nature* **521**, 227–31 (2015).

364. Monit, C. & Goldstein, R. A. SubRecon: Ancestral reconstruction of amino acid substitutions along a branch in a phylogeny. *Bioinformatics* **34**, 2297–2299 (2018).

365. Pollard, H. *et al.* Ca2+-sensitive cytosolic nucleases prevent efficient delivery to the nucleus of injected plasmids. *J. Gene Med.* **3**, 153–164 (2001).