Investigations into the function of the gammaretroviral protein p12 during the early stages of infection

Darren James Wight

June 2013

A Thesis Submitted to University College London for the Degree of Doctor of Philosophy

Division of Virology
National Institute for Medical Research
The Ridgeway
Mill Hill
Greater London
NW7 1AA

I, Darren James Wight 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

The retroviral structural proteins (encoded by *gag*) are not merely the architecture of the virus but carry out a diverse set of functions throughout the viral life cycle. Initially translated as a polyprotein, the Gag protein functions to encapsidate the genome and assemble retroviral particles at the budding sites. This polyprotein is then cleaved into the mature proteins: matrix (MA), capsid (CA) and nucleocapsid (NC) upon viral release. Most retroviruses also contain a small protein between MA and CA of unknown function. In Moloney murine leukaemia virus (Mo-MLV) this protein is p12. p12 contains the late-domain, required for recruiting ESCRT machinery to promote virus budding, but also contains regions that have been shown to be important during the early stages of infection

The aim of my PhD project was to establish the function of p12 during the early stages of infection and, in doing so, shed some light on the enigmatic processes occurring during these stages. Initially, the phenotypes of Mo-MLV virus-like particles (VLPs) containing previously identified p12 mutations were further characterised. Importantly, we found that these p12 mutations were clustered into two domains, both of which are required for Mo-MLV infectivity. Using genetic approaches, the order in which these domains act during infection was established. Furthermore, individual alanine mutants highlighted the residues within these domains that are essential for Mo-MLV infectivity.

The conservation of p12 function was analysed by creating analogous p12 mutations in other gammaretroviruses. While the whole C-terminal domain was essential, some viruses contained dispensable portions of the N-terminal domain. Chimeric viruses were constructed between Mo-MLV and Gibbon ape leukemia virus (GaLV) to further assess the conservation of p12 function.

Using biochemical, genetic and microscopic approaches, potential functions have been ascribed for the N- and C-terminal domains, allowing a model to be postulated for p12 function during the early stages of infection. This model proposes that the N-terminal domain of p12 binds to a virus specific factor which ties it to the pre-integration complex (PIC) and also stabilises the CA core. The C-terminal domain then comes into play during mitosis, tethering the PIC to the chromosomes and aiding integration.

Therefore, the last gammaretroviral Gag protein to be given a functional name has a large role to play in replication. Further understanding of p12, building on the work presented in this thesis, could aid in the design of gene transfer vectors with more precise integration targeting, and potentially will lead to improved understanding of early replication events that could enhance therapeutics against more complex retroviruses that cause human disease.

Acknowledgements

I would like to dedicate this thesis to my mum and nana. It's their love, guidance and support over the past 27 years that has gotten me here today.

Firstly, I would like to thank my supervisor, Kate Bishop, for her patience, guidance and inspiring me during my PhD. Current and past members of Kate's laboratory: Virginie, Harriet, Ophelie and Mirella, you guys have made the last three years in the lab really good fun (even when the experiments didn't work...!). Everyone in Jonathan Stoye's laboratory for their friendship, help and useful discussions. My fellow office workers on 'The Bridge' for the sauna-like temperatures (you know who you are!), fun conversations and lots of coffee!
In addition to the above, there are many other people around NIMR who have helped me during my PhD, and I would like to thank everyone who has pointed me in the right direction. Significantly, I would like to thank: Donald Bell and Kate Sullivan for their help with fluorescence microscopy, Liz Hirst for her time and expertise with the thin-section transmission electron microscopy and Ian Taylor for all his help and suggestions with the sucrose gradient work.

During my PhD I was lucky enough to work with our collaborators in Tel Aviv. I am hugely indebted to Eran Bacharach, Efrat Elis, Marcelo Erlich, Rachel, Oded and Irit, for the opportunity to work with them and their kind hospitality during my stay.

I must not forget to credit the wise people who helped set me on my current career path. Thank you to John Mason and Vassiliki Fotaki for mentoring me during my undergraduate degree and also giving me my first taste of real scientific research. Additionally, I would like to thank Mark Head and Mike Jones for the wonderful opportunities at the NCJDSU and their support of my decision to pursue a PhD.

Last, but by no means least, I would like to thank my friends and family. Especially my fellow PhD students/sandwich students/virologists/people at NIMR who have made working there great fun (Leonard, Lizzi, Dan, Sorrel, Christina, Gareth, Jo, Agathe aka 'righty', Mateo, Nick, Pat, Donald, Steve, Mike, Alex, Saira, Seb, Erhardt, Lesley, Rosemary, Seti, Rose, and anyone I may have forgotten to mention). Also my non-science friends who have helped restore normality to my life (Gouldy, Logan, Deeko, Joel, Chris, Charlie, Tino, Kev and Linzi). Finally a massive thanks to my family and my wonderful girlfriend Jara, I really couldn't have done it without you!

Contents

List of Figures

List of Tables

Abbreviations

AIDS	acquired immunodeficiency syndrome
APOBEC3G	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G
B-MLV	B-tropic murine leukaemia virus
BSA	bovine serum albumin
CA	capsid
CBS	chromatin binding sequence
CCD	catalytic core domain
cps	counts per second (luminescence)
CPSF6	cleavage and polyadenylation specific factor 6
cryo-EM	cryogenic electron microscopy
CsA	cyclosporin
CTD	C-terminal domain
CypA	cyclophillin A
DMEM	Dulbecco's modified eagle medium
EM	electron microscopy
Em	emission wavelength
EMCCD	electron multiplying charged coupled device
env	envelope gene
ESCRT	endosomal sorting complexes required for transport
Ex	excitation wavelength
FCS	foetal calf serum
FeLV	Feline leukaemia virus
FISH	fluorescent *in situ* hybridisation
FRAP	fluorescence recover after photobleaching
Fv1	Friend virus susceptibility 1
gag	group specific antigen gene
GaLV	Gibbon ape leukaemia virus
GFP	green fluorescent protein
gRNA	genomic RNA
HECT	homologous to the E6-AP carboxyl terminus
I-domain	interaction-domain
IN	integrase
kDa	kilo-Daltons
KoRV	Koala bear retrovirus
LacZ	β-galactosidase gene
LB	Lysogeny broth
L-domain	late-domain
LEDGF	lens epithelium-derived growth factor
LTR	long terminal repeat

MA	matrix
M-domain	membrane association domain
MMTV	mouse mammary tumour virus
Mo-MLV	Moloney murine leukaemia virus
MPMV	Mason Pfizer monkey virus
mU	milliunits
Mut	mutant
NC	nucleocapsid
NDS	normal donkey serum
Nef	negative regulatory factor
NLS	nuclear localisation sequence
N-MLV	N-tropic murine leukaemia virus
NoLS	nucleolar localisation sequence
NPC	nuclear pore complex
NTD	N-terminal domain
NUP	nucleoporin
PBS	phosphate buffered saline
PIC	pre-integration complex
PMT	photon multiplier tube
pol	polymerase gene
ppt	poly purine tract
PR	protease
PVDF	polyvinylidene fluoride
R	repeated region
Rev	regulator of expression of virion proteins
RFP	red fluorescent proteins
RfRV	Rhinolophus ferrumequinum retrovirus
RSV	Rous sarcoma virus
RT	reverse transcriptase
RTC	reverse transcription complex
SAMHD1	SAM domain and HD domain-containing protein 1
SDM	site directed mutagenesis
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sGFP	split GFP
SNV	spleen necrosis virus
SP1	spacer peptide 1
SP2	spacer peptide 2
Tat	trans-activator of transcription
TC	tetracystiene motif
TEM	Transmission electron microscopy
TNPO3	transportin 3

TRIM5alpha	Tripartite motif-containing protein 5
TSG101 (VPS23)	vacuolar protein sorting protein 23
U3	unique 3' region
U5	unique 5' region
UV	ultraviolet light
VSV-G	vesicular stomatitis virus glycoprotein
v/v	volume/volume
Vif	Viral infectivity factor
VLPs	virus like particles
Vpr	viral protein R
Vpu	viral protein unique
w/v	weight/volume
w/w	weight/weight
WDSV	Walleye dermal sarcoma virus
XMRV	Xenotropic murine leukaemia virus related virus
YFP	yellow fluorescent protein

Publications

The work presented in this thesis has led to the following publication:

Wight DJ*, Boucherit VC*, Nader M, Allen DJ, Taylor IA, Bishop KN. (2012). **The gammaretroviral p12 protein has multiple domains that function during the early stages of replication.** Retrovirology. Oct 4;9:83.

Chapter 1.

Introduction

Many viruses pose a significant health risk for both humans and animals, being the cause some of the largest epidemics in recent history. Viruses are completely reliant on a host for replication and have evolved to hijack various cellular processes to manipulate the host into producing progeny virus. Viral proteins mediate essential interactions between the virus and proteins of the host to carry out replication. A fuller understanding of the interactions between the host and virus can lead to the development of novel therapeutics against human disease. Furthermore, as many pathways exploited by viruses are essential for the cell, information on vital cellular processes can be derived from the study of virus replication.

1.1 An historical introduction to retrovirology

The discovery of retroviruses

The first isolation of a retrovirus came at the turn of the twentieth century, when Ellerman and Bang isolated a 'filterable agent' as the cause of a transmissible chicken leukosis (reviewed in (Coffin, Hughes, and Varmus, 1997b)). At the time, however, their work was largely ignored as the scientific community did not recognise leukaemia as a neoplastic disease until the 1930s (Van Epps, 2005). Similarly, the description by Peyton Rous, of a filterable agent capable of transmitting a neoplasmic disease between chickens was not widely regarded as having relevance to human disease (Rous, 1910; Van Epps, 2005). However, Rous was later vindicated as his discovery of 'tumour-inducing viruses' earned him a share of the Nobel Prize in Physiology or Medicine in 1966. As cell culture techniques improved (Temin and Rubin, 1958), additional oncogenic viruses were discovered, including those responsible for mammalian diseases (Bittner, 1936; Gross, 1957).

The proviral hypothesis and reverse transcriptase (RT)

Molecular characterisation of these filterable agents revealed them to be single stranded RNA viruses, although their replication was unlike other RNA viruses, known at the time. No double-stranded RNA was detected during infection and replication was found to be sensitive to inhibitors of DNA synthesis (Bader, 1965; Temin, 1963). Howard Temin proposed the proviral hypothesis, controversial at the time, stating that these

RNA viruses replicated through a DNA copy of their RNA genome (Temin, 1964). This hypothesis went against the popular central dogma of molecular biology, proposed by Francis Crick, that genetic information flows from DNA, to RNA, to protein. Temin's proviral hypothesis was corroborated and generally accepted upon the discovery of the virally encoded RNA-dependant DNA polymerase, reverse transcriptase (RT) (Baltimore, 1970; Temin and Mizutani, 1970).

Retroviruses and human disease

The first human oncogenic retrovirus discovered was human T-cell leukaemia virus (HTLV-1), found to be the cause of an abnormally high incidence of adult T-cell leukaemia on two islands in Japan (Poiesz et al., 1980). Not long after, the virus with causative links to the acquired immunodeficiency syndrome (AIDS) epidemic was isolated, originally known as LAV, HTLV-3 or ARV, but now known as human immunodeficiency virus 1 (HIV-1) (Barre-Sinoussi et al., 1983; Gallo et al., 1984; Levy et al., 1984). Since then much research has been devoted to HIV-1 replication and has led to the development of a therapy, called highly active anti-retroviral treatment (HAART) (Montagnier, 2010; Richman et al., 2009). Unfortunately, this treatment is life-long and cessation of HAART results in the rebound of viral titres (Chun et al., 2000). Access to HAART has improved since its introduction but the number of new HIV infections each year remains high. The search for an HIV-1 vaccine remains a high priority, however, despite much research no successful candidates have reached the clinic.

Retroviral vectors and gene therapy

Many congenital diseases are caused by the loss of a single gene product. The most straightforward approach to curing these types of genetic disease is by introduction of a functional copy of the defective gene into the individual. Retroviral based vectors hold great promise for gene therapy due to the fact that they stably integrate their genome into host chromosomes, although significant safety issues remain regarding their use in humans (Naldini et al., 1996). This issue was highlighted by the use of a gammaretroviral vector to supply a γC cytokine receptor subunit to patients with severe combined immunodeficiency X1, an inherited immune disease where T-cell and NK cells are absent in the affected individuals (Cavazzana-Calvo et al., 2000). While the treatment restored T-cells and NK cells in 9/10 patients, two patients developed leukaemia due to integration adjacent to the LM02 proto-oncogene promoter (Hacein-

Bey-Abina et al., 2003). Two further cases of leukaemia in patients from this trial have now been described, as well as cases from other similar gene therapy trials (Hacein-Bey-Abina et al., 2008; Howe et al., 2008). Interestingly, in all cases the mutagenesis causing the activated proto-oncogene expression is accompanied by an array of other genetic rearrangements (not directly caused by the gammaretroviral vector), which in combination have led to the development of leukaemia (Hacein-Bey-Abina et al., 2008; Howe et al., 2008).

Newer generations of gene transfer vectors (self-inactivating, SIN) containing deletions in the enhancer/promoter regions of the long terminal repeats (LTRs) have been created. In SIN vectors the therapeutic gene is driven from an internal cellular promoter, which reduces the chances of proto-oncogene activation (Montini et al., 2006). Insulator sequences can also be inserted before the cellular promoter to further reduce the chances of enhanced local proto-oncogene transcription (Ramezani, Hawley, and Hawley, 2008). While the new vector designs have increased the safety of retroviral-based vectors, further insight into integration targeting mechanisms of retroviruses will aid in the design of safer site-specific gene therapy vectors.

1.2 Retrovirus classification

The retrovirus family is divided into two subfamilies: the orthoretroviridae and the spumretroviridae. Orthoretroviridae are further divided into six genera: alpha-, beta-, gamma-, delta-, epsilon- retrovirus and lentivirus (Figure 1.1). The spumaretroviridae contain one genus named spumavirus, viruses from which are genetically and phenotypically very distinct from any genus of the orthoretroviridae, and will not form a focus of this study. Retroviruses are grouped into their genera based on a range of viral characteristics including: the physical morphology of the virus, genomic sequence and structure, tRNA species used for reverse transcription, the presence of accessory proteins and the site of virus assembly. The terms simple and complex retroviruses are often used to broadly subcategorise retroviruses, where the genomes of simple retroviruses lack accessory genes (see section 1.4.4).

The majority of the work described in this introduction derives from work on three prototypic retroviruses murine leukaemia virus (MLV, gammaretrovirus), Rous sarcoma virus (RSV, alpharetrovirus) and HIV (lentivirus). While these represent examples from

different genera and are thus somewhat divergent, there are similarities between many of their replication strategies.

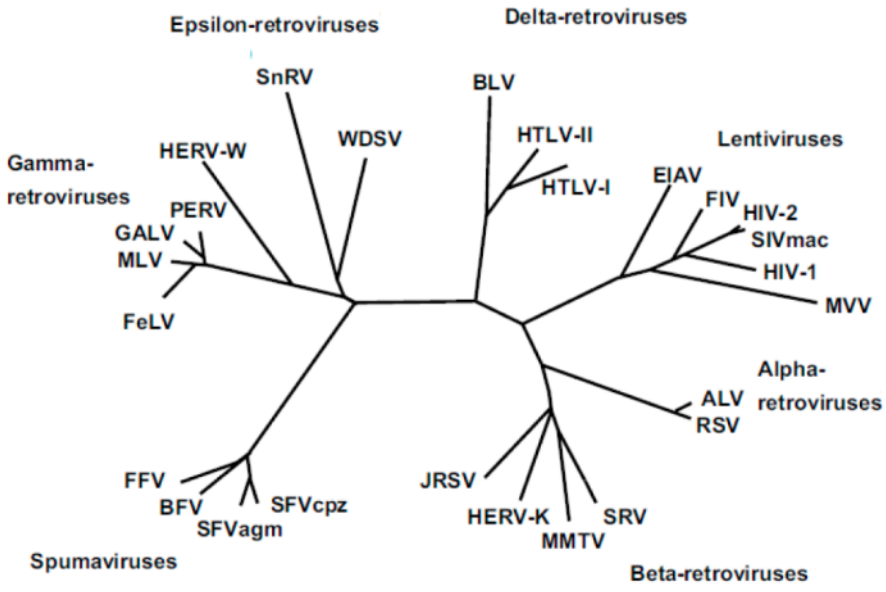


Figure 1.1: Phylogeny of retroviruses. Phylogenetic tree highlighting all seven retroviral genera from the Retroviridae family (adapted from (Weiss, 2006)).

1.3 Genome organisation

Retroviruses contain a dimer of viral RNA (gRNA) with a sedimentation coefficient of 70S (Fu, Gorelick, and Rein, 1994; Mangel, Delius, and Duesberg, 1974), which is specifically packaged into the virus during assembly. Similar to eukaryotic mRNA, the viral genome contains a 5' cap (Furuichi et al., 1975) and a 3' polyadenylyl tail (poly-A-tail) (Lai and Duesberg, 1972). During infection viral RNA is reverse transcribed into DNA, which is then integrated into the host DNA, to form the provirus. The genome is composed of non-coding regulatory sequences and sequences coding for the viral proteins (Figure 1.2).

1.3.1 Non-coding regions

The un-translated regions of the viral RNA include the unique 5' region (U5), unique 3' region (U3), repeated region (R), 5' *gag* leader (including the primer binding site, PBS) and the sequence 3' of *env* including the poly purine tract (ppt). These elements play roles in reverse transcription of the viral RNA into viral cDNA and transcription from the provirus (Gilboa et al., 1979). HIV-1 also contains an additional ppt called the cPPT (not shown in Figure 1.2), which is involved in synthesis of the plus-sense DNA during reverse transcription (Charneau and Clavel, 1991; Wain-Hobson et al., 1985). Reverse

transcription is initiated from a tRNA molecule bound to the PBS, which is $tRNA^{Pro}$ for MLV and $tRNA^{Lys}$ for HIV-1 (Harada, Peters, and Dahlberg, 1979; Peters et al., 1977; Ratner et al., 1985). After reverse transcription the viral DNA product contains identical U3-R-U5 regions at both ends (Shank et al., 1978), referred to as the LTRs. Post-integration the viral DNA remains flanked by the LTRs (Hughes et al., 1978), which contain promoter and enhancer sequences, important for transcription (Graves, Eisenman, and McKnight, 1985; Laimins et al., 1984; Ostrowski, Berard, and Hager, 1981).

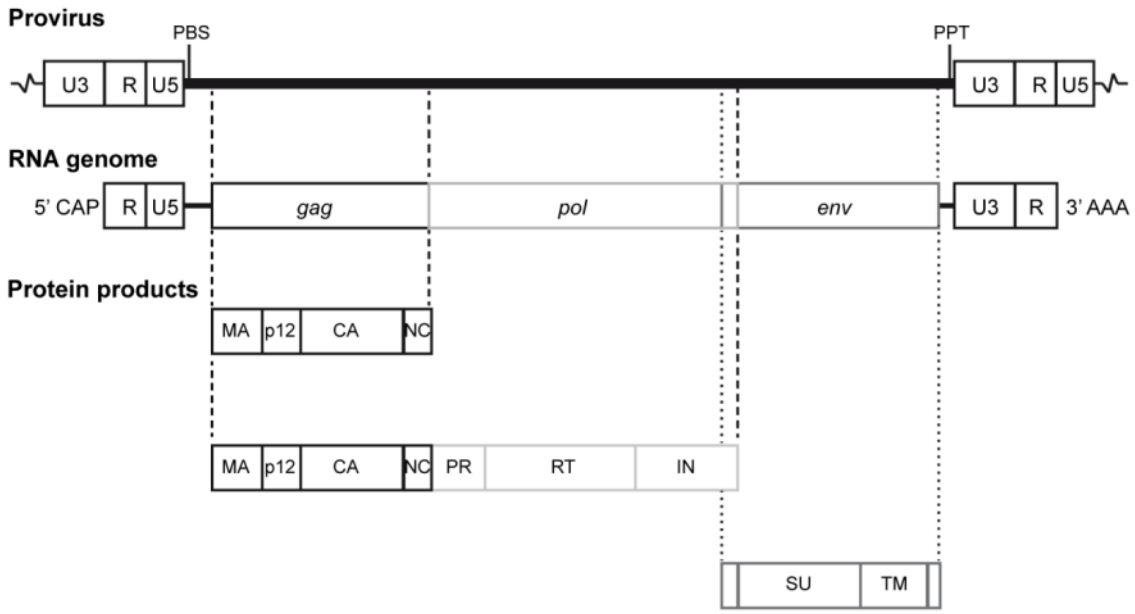


Figure 1.2: Structure of the gammaretroviral genome and provirus. The structure of the MLV provirus is shown above the genomic RNA. Below are the protein products from the two spliced RNAs. 'Glyco-Gag', initiated from a CUG start codon in the 5' *gag* leader sequence, not shown. (Adapted from (Petropoulos, 1997))**.**

1.3.2 Coding regions

All retroviruses contain the three main genes *gag, pol* and *env,* arranged in this order from 5' to 3' in the genome (Coffin and Billeter, 1976; Coffin, Hughes, and Varmus, 1997a). The *gag* gene encodes the structural proteins of the virus; *pol* encodes the viral enzymes and *env,* the components of the envelope glycoprotein complex. In gammaretroviruses translation of mRNA for *pol* is carried out by read-through of the stop codon at the end of *gag (Yoshinaka et al., 1985a; Yoshinaka et al., 1985b)*. Thus the number of Gag molecules considerably outnumber that of Pol (Jamjoom, Naso, and Arlinghaus, 1977). Other retroviruses, like HIV-1, generate Pol via a ribosomal frameshift (Jacks et al., 1988). In contrast to Gag and Pol, Env is translated from a

spliced mRNA. Genetically complex retroviruses also contain additional genes that code for regulatory and accessory proteins. These are generally situated either between *pol* and *env*, 3' to *env* (including the U3 region) or a combination joined by splicing of these regions (Coffin, Hughes, and Varmus, 1997a).

1.4 Viral proteins

This next section will briefly introduce the viral proteins. p12 is intentionally omitted from this section and will be described in greater detail in a later section.

1.4.1 Gag

The most abundant viral proteins in the particle derive from the *gag* gene. Retroviral Gag is synthesised on the ribosome as a polyprotein, which contains all the functions required to drive formation of virus-like particles (VLPs) (Gheysen et al., 1989). After viral release, protease (PR) cleaves Gag during maturation of the virus to produce the mature proteins: matrix (MA), capsid (CA) and nucleocapsid (NC) (Figure 1.2). Additionally, many retroviruses also contain other Gag cleavage products situated between the protein domains of Gag. Cleaved proteins between MA and CA are the largest of the additional Gag proteins and some retroviral genomes encode multiple proteins at this location (described in section 1.8).

In Gag at least three functional domains have been described as essential for production of VLPs: membrane binding-domain (M-domain), interaction domain (I-domain) and late-domain (L-domain). MA contains the M-domain and drives localisation of Gag to the plasma membrane. For many retroviruses, like MLV and HIV, this is done via a co-translational addition of a myristoyl group to the N-terminal glycine in MA (Bryant and Ratner, 1990; Gottlinger, Sodroski, and Haseltine, 1989; Henderson, Krutzsch, and Oroszlan, 1983; Rein et al., 1986). This modification is essential and mutation of the N-terminal glycine blocks MLV release and Gag accumulates in the cytoplasm (Rein et al., 1986). However, additional sequences in MA are also important for the stabilisation of the MA association with the plasma membrane (Ono and Freed, 1999; Spearman et al., 1994; Zhou et al., 1994). RSV is different from MLV and HIV in this regard, while MA is still the M-domain it does not require myristoylation (Schultz and Oroszlan, 1983; Verderame, Nelle, and Wills, 1996; Wills et al., 1991). After virus fusion and entry of the core to the cytoplasm, it is assumed that the majority of mature MA will remain with the virion membrane.

CA is the major structural protein for retroviruses and is composed of an N-terminal domain (NTD) and a C-terminal domain (CTD) connected by a linker. In Gag, CA contains part of the sequence making up the vital I-domain, important for Gag-Gag multimerisation. This sequence maps to the CTD, although some retroviruses do not require CA-CA contacts in Gag for assembly (Franke et al., 1994; Mammano et al., 1994; von Poblotzki et al., 1993; Weldon and Wills, 1993). Little sequence homology exists between CA proteins from different retroviruses and the most conserved region is the major homology region (MHR) in the CTD (Mammano et al., 1994). Despite the lack of sequence homology between retroviral CA, the NTDs all contain a remarkably similar tertiary structure (Gamble et al., 1996; Gitti et al., 1996; Mortuza et al., 2008; Mortuza et al., 2009). Structural conservation has even been observed in the CA NTD from resurrected ancient endogenous lentiviruses RELIK and PSIV, highlighting this structure has been highly conserved in the evolution of retroviruses (Goldstone et al., 2010). One major structural difference between gammaretroviral CA proteins and those from most lentiviruses is the absence of the so-called cyclophillin binding loop (Gamble et al., 1996). This loop is situated between helix 4 and 5 in the CA NTD and is extended in HIV-1 compared to MLV (Mortuza et al., 2004). Mature CA forms the core surrounding the ribonucleoprotein complex (RNP), which is released into the target cells.

The last core Gag protein is NC that also functions as the part of the I-domain during virus assembly. NC contains one (gammaretroviruses) or two (lentiviruses) conserved zinc-finger domains, which are required for packaging the viral RNA (Aldovini and Young, 1990; Gorelick et al., 1988; Henderson et al., 1981; Thomas and Gorelick, 2008). Deletion of NC causes a defect in virus release due to aberrant assembly of viral particles (Gheysen et al., 1989; Muriaux et al., 2004). Experiments have been performed on MLV that have shown that the NC and RNA interaction enables Gag multimerisation using the RNA as a scaffold (Muriaux et al., 2001). Moreover, replacement of NC function for virus release can be achieved, in an RNA-independent manner, by the addition of a multimerisation domain, like a leucine zipper (Accola, Strack, and Gottlinger, 2000; Johnson et al., 2002; Zhang et al., 1998). Therefore as a part of Gag, NC is a key driving force for Gag-Gag interactions essential for assembly. After maturation NC coats the viral RNA and is beneficial for reverse transcription through its nucleic acid chaperone activities (Thomas and Gorelick, 2008).

1.4.2 Pol

The Pol proteins are the virally encoded enzymes namely: protease (PR), reverse transcriptase (RT) and integrase (IN) (Figure 1.2). The translation of Pol is tightly regulated by retroviruses by mechanisms mentioned earlier. This results in a roughly 20:1 excess of translated Gag -to- Gag-Pol molecules (Jamjoom, Naso, and Arlinghaus, 1977). All the retroviral enzymes are therapeutic targets against HIV infection and inhibitors to each Pol protein have been approved for use in patients.

PR is an aspartic protease essential for triggering the maturation of the virus after budding. Mutation of the residues in the PR active site or deletion of PR results in viruses with an immature morphology, which as a result cannot establish an infection in target cells (Crawford and Goff, 1985; Katoh et al., 1985; Kohl et al., 1988). For cleavage, PR recognises diverse cleavage motifs between Gag and Pol proteins, and work from HIV-1 PR has indicated that this is through recognition of a conserved structure at the cleavage sites (Prabu-Jeyabalan, Nalivaika, and Schiffer, 2002).

RT was discovered independently by David Baltimore and Howard Temin in 1970 (Baltimore, 1970; Temin and Mizutani, 1970). It is an RNA-dependent DNA polymerase with RNase H activity, which can utilise an RNA or DNA template to synthesise DNA (Baltimore, 1970; Telesnitsky and Goff, 1993; Temin and Mizutani, 1970). For HIV-1 and HIV-2 the enzyme is active as a heterodimer composed of p66 and p51 subunits, with the RNase H domain present in the larger p66 subunit (Kohlstaedt et al., 1992; Rodgers et al., 1995). However, the gammaretroviral RT is active as a monomer (Das and Georgiadis, 2004; Nowak et al., 2013). RT is both necessary and sufficient for the synthesis of the retroviral DNA from the RNA genome (described later).

Integrase is essential for the establishment of the provirus in infected cells (reviewed in (Krishnan and Engelman, 2012)). This enzyme forms an intasome complex (dimer of IN dimers) with the viral LTR ends (Hare et al., 2010; Maertens, Hare, and Cherepanov, 2010) and provides two essential enzymatic activities: (i) 3' processing of the viral LTRs (Roth, Schwartzberg, and Goff, 1989) and (ii) the strand transfer catalysis (Brown et al., 1989; Fujiwara and Mizuuchi, 1988). Lentiviral IN complexes also form vital contacts with lens epithelium derived growth factor (LEDGF), important for integration site selection (detailed in a later section).

1.4.3 Env

The Env proteins are translated from a spliced full length mRNA and are trafficked to the endoplasmic reticulum (ER). The Env polyprotein contains two components: the surface (SU) and transmembrane (TM) domains. Within the ER the Env proteins form the Env glycoprotein complex, which is a trimer of Env for RSV, MLV and HIV (Einfeld and Hunter, 1988; Forster et al., 2005; Liu et al., 2008). Env is cleaved and then transported to the plasma membrane via the golgi apparatus. Functionally the SU component is required for Env-receptor interactions and the TM subunit is vital for membrane fusion.

1.4.4 Accessory proteins

Complex retroviruses, like those from the lentivirus genus, code for other proteins in addition to Gag, Pol and Env. Simple retroviruses like the gammaretroviruses do not code any accessory proteins (barring glyco-Gag) and thus the accessory proteins of HIV-1 will only be briefly mentioned.

Firstly, the gammaretroviral glyco-Gag protein is initiated from a CUG codon upstream of the Gag initiation codon (Prats et al., 1989). Not all gammaretroviruses encode glyco-Gag and it is not essential for replication in cell culture but appears essential for MLV infection *in vivo* (Low et al., 2007). In line with this, all the gammaretroviral vectors utilised in this thesis do not encode Glyco-Gag. Glyco-Gag has been implicated in directing viral assembly at lipid rafts through the cellular protein La and has been implicated in protecting MLV from APOBEC3 (Nitta et al., 2011; Stavrou et al., 2013). Additionally, it has been reported that glyco-Gag can substitute for Nef in HIV-1 suggesting it has a Nef like function for HIV-1 (Pizzato, 2010).

HIV-1 encodes the regulatory proteins Tat and Rev and the accessory proteins Vif, Vpr, Vpu and Nef. Tat increases proviral expression by binding to the transcribed RNA transactivating response (TAR) region from the viral promoter and enhances transcription (Berkhout and Jeang, 1989; Feng and Holland, 1988; Ott, Geyer, and Zhou, 2011). Rev binds to the Rev-response element (RRE) on the viral RNA and facilitates export of the unspliced viral RNA from the nucleus (Felber et al., 1989).

The accessory proteins are required *in vivo* but are generally dispensable for replication in many cell culture systems (Mustafa and Robinson, 1993). Vif has an important

function in counteracting the restriction factor apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) by binding and targeting it for proteasomal degradation (Conticello, Harris, and Neuberger, 2003; Sheehy et al., 2002; Sheehy, Gaddis, and Malim, 2003). Vpr is incorporated into virions and induces G_2/M cell cycle arrest in the target cells and has also has been linked to nuclear import of the pre-integration complex (PIC), although the latter role has been refuted (Heinzinger et al., 1994; Sharifi, Furuya, and de Noronha, 2012)). Vpu has been shown to counteract the late restriction factor tetherin, which prevents particle release, by down regulating tetherin cell surface expression (Neil et al., 2006; Neil, Zang, and Bieniasz, 2008; Van Damme et al., 2008). Vpu also degrades CD4 trapped in intracellular complexes with Env (Willey et al., 1992). Nef has multiple roles including down regulation of CD4 and MHC class I molecules. Additionally Nef is able to enhance viral DNA synthesis among other functions (reviewed in (Laguette et al., 2010)).

1.5 The retroviral life cycle

The retroviral life cycle is broadly separated into two parts, the early and late stages of infection. The early stages of infection comprise all the steps from virus binding to the target cell up to integration of the viral DNA. The late stages include transcription of the provirus through to budding and maturation (see Figure 1.3).

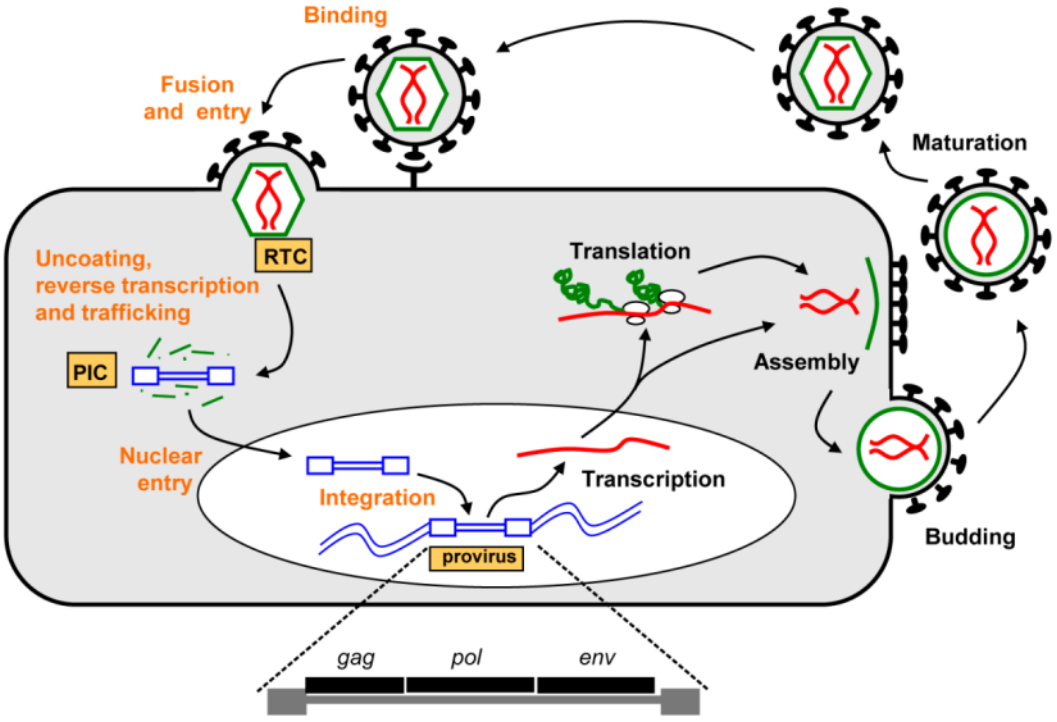


Figure 1.3: Schematic detailing the life cycle of a retrovirus. The main steps are written in bold with the early steps of the life cycle highlighted in orange. RTC- reverse transcription complex, PIC- pre-integration complex.

1.5.1 Retroviral entry

The major determinant of viral tropism is the receptor usage by the envelope glycoprotein complex on the surface of the virus. The tropism of the MLVs from laboratory mice are traditionally categorised in one of three classes: ecotropic, xenotropic or polytropic. Ecotropic MLVs have a host range that is restricted to host cells (murine and rat cells) and utilise mCAT1, a cationic amino acid transporter, for entry (Kim et al., 1991; Wang et al., 1991). Xenotropic MLVs are unable to infect the host cells of murine origin but are able to infect cells from other mammals (Levy, 1973). Xpr1 is the receptor molecule for MLV with the xenotropic envelope glycoprotein complex (Battini, Rasko, and Miller, 1999; Tailor et al., 1999). The last class, polytropic, have a wide host range and are able to infect both murine and non-murine cells also by exploiting the Xpr1 molecule for entry (Battini, Rasko, and Miller, 1999; Fischinger, Nomura, and Bolognesi, 1975; Tailor et al., 1999).

The initial binding of a virus to a target cell is mediated by non-specific interactions between the virion membrane and the membrane of the cell (Beer and Pedersen, 2007). Specific interactions between the Env glycoprotein complex and the receptor(s), triggers a cascade of steps leading to fusion and release of the core into the target cell. Engagement of the SU domain with the receptor isomerises the SU-TM disulphide bond resulting in the displacement of the SU subunit of the Env complex (Wallin, Ekstrom, and Garoff, 2004) This exposes the fusion peptide in the N-terminus of TM, which inserts into the host cell membrane. Insertion of TM into the host membrane triggers conformational changes in TM where the C-terminal heptad repeat folds against the N-terminal heptad repeat bringing the host and virion membrane into close proximity allowing fusion to occur (Fass, Harrison, and Kim, 1996).

For HIV-1, two cell surface molecules are utilised to gain entry to cells, CD4 and a co-receptor, either CXCR4 or CCR5, depending on the specific tropism of the virion Env complex (Dalgleish et al., 1984; Dragic et al., 1996; Feng et al., 1996). Interaction of the Env complex with CD4 induces conformational changes which expose the CCR5/CXCR4 binding pocket (Liu et al., 2008). Binding of the co-receptor leads to further changes and the gp41 (TM) inserts its fusogenic peptide into the cell membrane (Blumenthal, Durell, and Viard, 2012). A pore is formed which is widened to allow delivery of the viral core.

Retroviruses can also be pseudotyped with a variety of heterologous glycoproteins to broaden cell tropism. One that has been of particular use has been the vesicular stomatitis virus glycoprotein (VSV-G) (Naldini et al., 1996). VSV-G allows entry to a wide range of cells, as the receptor is thought to be a factor ubiquitously expressed in most cells (Cronin, Zhang, and Reiser, 2005).

1.5.2 Uncoating, reverse transcription, and the path to the nucleus

After release of the retroviral core into the cytoplasm a set of essential processes must occur before insertion of the viral genome into the host chromosomes. These processes are dynamic and evidence is beginning to suggest that they are all inter-connected. This next section will detail some of the key events occurring on transit to the nucleus.

Uncoating

Uncoating is a poorly understood step of the viral life cycle, where CA is lost from the retroviral reverse transcription complex (RTC)/PIC during trafficking to the chromosomes. When uncoating occurs, its kinetics and how much CA is actually lost during uncoating remain debated topics. Unfortunately, due to the difficulties in isolation and experimental manipulation of RTC/PICs these questions have not been very amenable to *in vitro* studies. Data have highlighted, however, that the rate of uncoating varies between cell types, indicating that uncoating is influenced by host factors (Arfi et al., 2009).

Different retroviruses appear to uncoat at varying rates. Early biochemical data suggested that the kinetics of HIV-1 uncoating was very fast and no CA could be detected with isolated RTCs very early after infection (Bukrinsky et al., 1993b; Fassati and Goff, 2001). On the other hand, in the case of MLV, CA was readily detected with the RTC/PIC during infection using the same biochemical techniques (Fassati and Goff, 1999). However, uncoating does eventually occur for MLV as little CA could be detected in nuclear fractions of infected cells or in the nucleus of mitotic cells by immunofluorescence (Elis et al., 2012; Fassati and Goff, 1999; Prizan-Ravid et al., 2010).

Despite evidence suggesting the rapid uncoating of the HIV-1 RTC, it seems likely that some CA does remain with the PIC until a later stage in the infection process. The evidence for this is: (i) immunofluorescence and EM showing CA associated with active RTCs and CA cores present at/around the nuclear pores during infection (Arhel et al., 2007; McDonald et al., 2002), (ii) CA appears to be the determinant of nuclear entry for HIV-1 and interacts with transportin 3 (TNP03) and a component of the nuclear pore complex, Nup358 (RANBP2) (Krishnan et al., 2010; Schaller et al., 2011; Yamashita and Emerman, 2004; Yamashita et al., 2007), (iii) CA can be detected in the nucleus of fractionated infected cells (Zhou et al., 2011) and (iv) the rapid uncoating induced by TRIM5alpha aborts infection before reverse transcription (Stremlau et al., 2006; Towers et al., 2000). The reason why CA is not biochemically isolated with HIV-1 RTCs is likely linked to RTC isolation protocols disrupting the CA core (McDonald et al., 2002).

The actual trigger for uncoating remains unknown. No CA-interacting molecules have been described for simple retroviruses that regulate/trigger uncoating. However, the host molecule cyclophilin A (CypA) could play a role in HIV-1 uncoating. CypA binding occurs on the CypA binding loop of the CA NTD for a range of lentiviral CA molecules (Gamble et al., 1996; Lin and Emerman, 2006). It was originally thought to be essential in HIV-1 producer cells, but now the general consensus is that CypA is required in the target cells (Hatziioannou et al., 2005; Luban et al., 1993; Thali et al., 1994). CypA is a peptidyl prolyl *cis*/*trans* isomerase that catalyses the *cis-trans* conformation of the G89-P90 bond in HIV-1 CA (Bosco et al., 2002). It has been speculated that CypA could cause uncoating by either (i) *cis/trans* isomerisation of the G89-P90 bond or (ii) destabilise the CA core by high levels of CypA binding (Gamble et al., 1996). Indeed, high levels of CypA can dismantle CA-tubes *in vitro* (Grattinger et al., 1999). However the precise role of CypA in HIV-1 infection and uncoating remains unknown.

Recent experiments have linked the process of uncoating to reverse transcription. Mutations or drugs blocking RT activity increase the stability of the HIV-1 CA core and the time required to uncoat (Hulme, Perez, and Hope, 2011; Yang, Fricke, and Diaz-Griffero, 2013) Likewise, mutations altering the stability of the CA core alter reverse

transcription, indicating they likely occur concomitantly and exert an influence on each other (Arfi et al., 2009; Forshey et al., 2002).

Reverse transcription

Formation of the viral cDNA is an essential step in the retroviral life cycle. RT catalyses the conversion of viral RNA to the linear double stranded DNA and exposure to higher levels of dNTPs is believed to be the trigger for reverse transcription to begin. Indeed, retroviruses can synthesise viral DNA if the virus is permeablised and exposed to dNTPs (Haseltine et al., 1976).

Synthesis of the negative strand begins from the incorporated tRNA bound to the PBS 3' to the R-U5 region of the viral RNA (Harada, Peters, and Dahlberg, 1979; Ratner et al., 1985). During synthesis of this so-called strong-stop DNA, the RNase H activity of RT digests the RNA template. Detection of the strong stop DNA species in infected cells is often used to quantify the amount of early reverse transcription products formed and has been used in this thesis. Next a translocation event occurs, referred to as minus-strand transfer, where the synthesised R-U5 region anneals to the 3' R region of the RNA. Minus strand DNA is synthesised up to the end of the RNA template (PBS region). The RNase H activity of RT digests the RNA template apart from a resistant stretch called the polypurine tract (ppt). Unlike MLV, HIV and some other lentiviruses have an additional ppt called the central ppt (cPPT). The cPPT seems to enhance the ability of HIV RT by acting as an additional priming site for plus strand DNA synthesis although it is not essential for replication (Charneau, Alizon, and Clavel, 1992; Charneau and Clavel, 1991).

The synthesis of the plus-strand DNA is initiated from the ppt and extends the U3-R-U5-tRNA region. RNase H activity degrades the tRNA molecule and the second translocation event occurs (plus-strand transfer), annealing the PBS of the plus-strand U3-R-U5-PBS to the minus-strand PBS. The plus and minus strands are then extended to completion so that both ends of the viral DNA are flanked by the identical LTR sequence. This forms the DNA 3' to the LTR which is used to detect late RT products as so-called second strand extension products. RT is both necessary and sufficient for *in vitro* reverse transcription, although other viral proteins and cellular factors have roles in the reverse transcription reaction *in vivo* (Hu and Hughes, 2012).

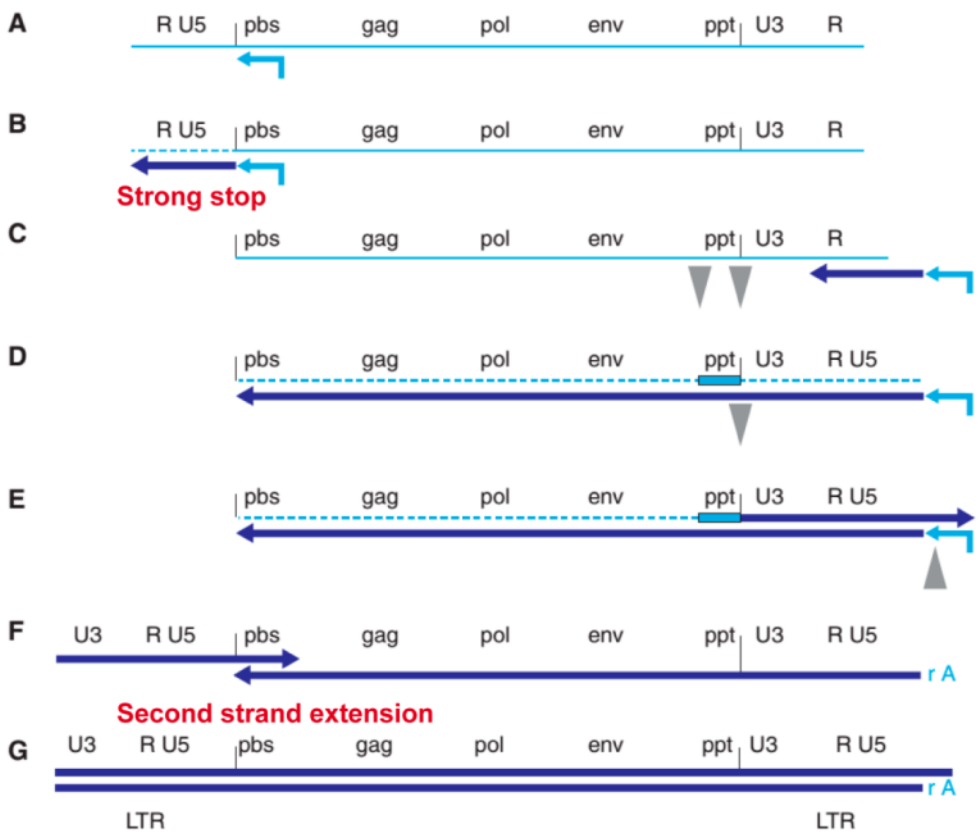


Figure 1.4: Reverse transcription of viral RNA to DNA. (A) Reverse transcription is initiated from the tRNA primer bound to the primer binding site (pbs). (B) RT catalyses the R-U5 of the minus strand (called strong stop DNA, written in red). The RNA template (baring the pbs) is degraded by the RNase H domain in RT. (C) Minus-strand transfer anneals the newly synthesised R region to the complementary R region at the 3' end of the RNA. (D) The minus strand is synthesised to the end of the pbs and the RNA template is degraded except for a resistant patch at the 3' end called the poly purine tract (ppt). (E) The ppt is then used to prime the synthesis of the plus strand LTR replicating the tRNA to produce the pbs. The tRNA molecule is then degraded. (F) The plus strand translocation event occurs annealing the plus strand and minues strand pbs. (G) RT completes synthesis of the plus and minus strands forming the double stranded DNA (forming a DNA detected by primers used in this thesis for second strand extension products, written in red). Figure adapted from (Hu and Hughes, 2012).

Trafficking

RTC/PICs must navigate through the relatively busy environment of the cell to reach the nucleus for viral DNA integration. These movements require active transport. Association of HIV-1 RTCs with the microtubule network has been shown using live immunofluorescence and correlative EM, revealing that RTCs use microtubule based transport for long distance movements (McDonald et al., 2002). Additionally, blocking dynein-mediated transport before core entry to the cell resulted in the peripheral localisation of the RTCs, further implicating microtubule based transport in movements

of the PIC (McDonald et al., 2002). One report suggested that MA is involved in the HIV-1 RTC transport by interaction with the actin network, which was blocked by inhibiting actin polymerisation. However, this effect was lost when virus was psuedotyped with VSV-G, suggesting actin may be involved in Env-specific entry events (Bukrinskaya et al., 1998). A more recent detailed study using FlAsH labelled HIV-1 IN described four types of movement, indicating HIV-1 was utilising both microtubule and actin based transport. This has implied that there may be some functional redundancy in the trafficking mechanisms utilised by retroviruses (Arhel et al., 2006).

Restriction factors targeting the early stages of infection

The early stages of infection are a particularly vulnerable time for retroviruses and many antiretroviral drugs target these early steps. Indeed, cells have evolved five main restriction factors targeting these stages, namely: tripartite-motif-containing 5 alpha (TRIM5alpha), Friend virus susceptibility gene 1 (Fv1), Friend virus susceptibility gene 4 (Fv4), APOBEC3G and SAM domain and HD domain-containing protein 1 (SAMHD1). Both Fv1 and TRIM5alpha were utilised as CA-core sensors in this thesis and their mechanism of action is described in more detail.

TRIM5alpha was discovered as a HIV-1 restriction factor from a screen of rhesus macaque genes, and blocks infection before reverse transcription (Stremlau et al., 2004; Stremlau et al., 2006). TRIM proteins come from a family with about 70 members, all which contain a similar domain architecture (Nisole, Stoye, and Saib, 2005). The N-terminus consists of a RING domain, either one of two B-boxes and a coiled-coil domain followed by a range of C-terminal domains depending on the TRIM protein (Figure 1.5A) (Nisole, Stoye, and Saib, 2005). TRIM5alpha has a PRYSPRY (or B30.2) in the C-terminus heavily involved in specifying the CA core recognised by the particular TRIM5alpha protein (Song et al., 2005; Stremlau et al., 2006; Yap, Nisole, and Stoye, 2005). The PRYSPRY contains three hypervariable loops (V1-V3) of different sizes and sequence depending on the TRIM5alpha species (Sawyer et al., 2005; Stremlau et al., 2005). These loops are important in the specificity of antiretroviral activity of TRIM5alpha and likely map key interfaces between the CA core and TRIM5alpha (Perez-Caballero et al., 2005; Sawyer et al., 2005; Stremlau et al., 2005; Yap, Nisole, and Stoye, 2005). The human TRIM5alpha is unable to restrict HIV-1 but is active against a range of retroviruses including N-tropic MLV. The viral

determinants of restriction map to the CA protein, although TRIM5alpha does not interact with monomeric CA proteins and requires a core-like structure (Cowan et al., 2002; Stremlau et al., 2006). Poor solubility has precluded structural analysis of theTRIM5alpha-CA interaction, however mapping of escape mutants has highlighted the interaction between CA and TRIM5alpha likely occurs over the entire outer face of the CA protein (Ohkura et al., 2011; Ohkura and Stoye, 2013).

A variant of TRIM5alpha, called TRIMcyp, has arisen from two independent retrotransposition events in owl monkeys and certain macaque species, where the PRYSPRY has been replaced by CypA (Brennan, Kozyrev, and Hu, 2008; Newman et al., 2008; Sayah et al., 2004; Virgen et al., 2008; Wilson et al., 2008). TRIMcyp therefore interacts with CA at the CypA binding loop and is a potent inhibitor of HIV-1 infection.

The exact mechanism behind TRIM5alpha's restriction activity remains poorly understood. Within the cell TRIM5alpha associates with the incoming core rapidly (<30 mins) and has been shown to prematurely uncoat the core, suggesting this causes the block to infection (Perez-Caballero et al., 2005; Stremlau et al., 2006). However, it has been shown in cells treated with proteasome inhibitors disassembly does not occur, reverse transcription proceeds but infection is aborted before nuclear entry (Diaz-Griffero et al., 2007; Wu et al., 2006). This has indicated that either TRIM5alpha has more than one mechanism of restriction or loss of the proteasome results in a kinetically slowed restriction, allowing reverse transcription to occur. Additionally, TRIM5alpha has been shown to act as a pattern recognition receptor for the retroviral CA lattice. Where association with the CA core increases the E3 ubiquitin ligase activity of the TRIM5alpha RING domain and with the UBC13–UEV1A complex forms free K63 ubiquitin chains, activating a cascade of downstream events and leading to AP-1 and NFκB activation (Pertel et al., 2011). However, it has been shown that the E3 ubiquitin ligase activity is not required for restriction (Perez-Caballero et al., 2005). Thus far the only aspects of TRIM5alpha action essential for restriction are: recognisation of the CA core through the PRYSPRY/CypA domain and multimerisation through the RING, B-box and coiled-coil domains. The B-box is important for higher order association of TRIM5alpha, which increases the efficiency of TRIM5alpha recruitment to the core (Diaz-Griffero et al., 2009; Li and Sodroski, 2008). Consistent with this notion, cryogenic electron microscopy (cryo-EM) images of a chimeric TRIM5alpha protein

(with the RING domain from TRIM21) formed on HIV-1 CA tubes have shown the formation of a superlattice over the CA hexagonal array, with the SPRY/B30.2 domain spread across the CA surface (Figure 1.5B) (Ganser-Pornillos et al., 2011).

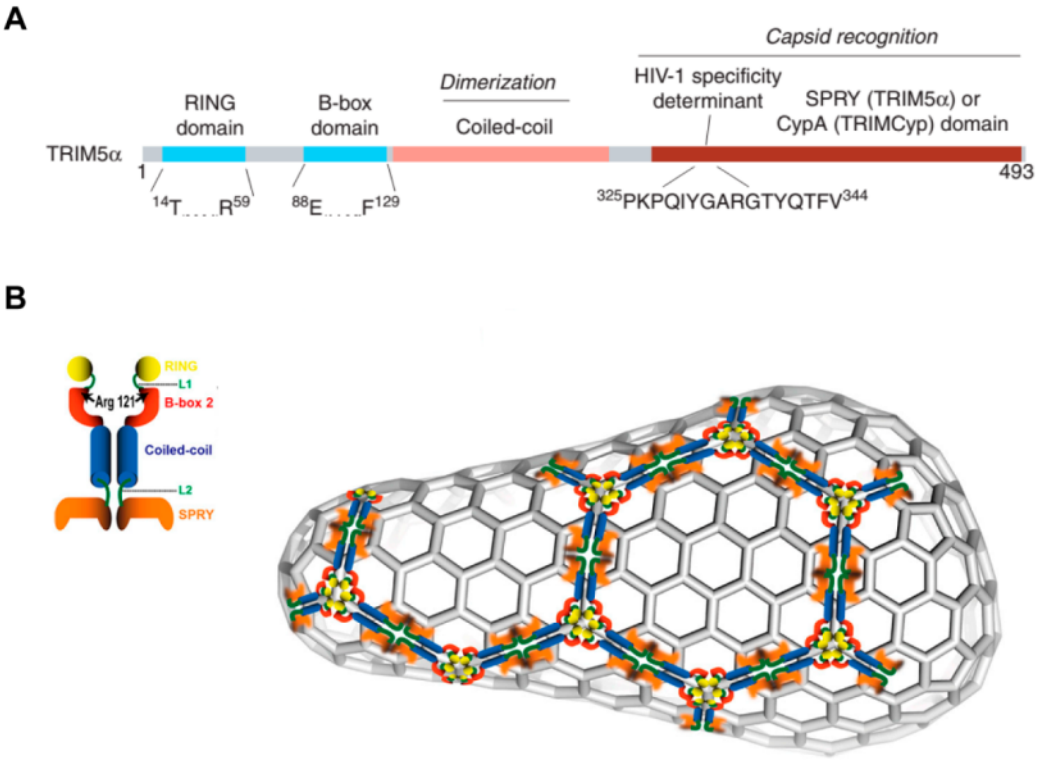


Figure 1.5: TRIM5alpha association on the CA core. (A) Schematic detailing the domains present in TRIM5alpha and TRIMcyp. Adapted from (Malim and Bieniasz, 2012). (B) Diagram displaying a proposed model of TRIM5alpha higher order association over the HIV-1 CA core. Adapted from (Ganser-Pornillos et al., 2011).

Fv1 was the first CA interacting restriction factor to be cloned but less is known about its mechanism of restriction than is known for TRIM5alpha. Fv1 was cloned in 1996 and is derived from the *gag* of an endogenous retrovirus (Best et al., 1996). It is a mouse gene and has two major alleles, $Fv1^{n}$ and $Fv1^{b}$. Cells expressing the former restrict B-tropic MLV and the expression of the latter restricts N-tropic MLV infection. Restricted viruses are blocked at a stage after reverse transcription but before integration (Jolicoeur and Baltimore, 1976). Viral determinants of restriction map to the CA protein (specifically amino acid 110) (Kozak and Chakraborti, 1996; Stevens et al., 2004) and the specificity of Fv1 restriction maps to the C-terminal domain (Bishop et al., 2001). Similar to TRIM5alpha, the N-terminal domain of the protein contains a predicted coiled-coil domain, has been shown to dimerise and this region is essential for

restriction activity (Bishop et al., 2006). Artificial restriction factors, where the N-terminal portion of Fv1 replaces the N-terminal part of TRIMcyp, are competent to restrict HIV-1 suggesting that Fv1 multimerisation in a fashion similar to TRIM5alpha is also essential for restriction (mentioned above) (Yap et al., 2007). The half-life of one of these restricting Fv1-Cyp proteins is much longer than TRIMcyp, indicating that the slower removal of Fv1-cyp may be what allows viruses restricted by Fv1 to reverse transcribe (Schaller et al., 2007). However, at present, poor solubility of the Fv1 protein means limited structural information exists and the exact mechanism of restriction remains ill-defined.

Both Fv1 and TRIM5alpha have been utilised as retroviral core sensors in cells due to the fact that restriction can be saturated by infecting with a high MOI of virus (Bassin et al., 1978; Besnier, Takeuchi, and Towers, 2002; Boone, Innes, and Heitman, 1990; Cowan et al., 2002; Duran-Troise et al., 1977; Towers, Collins, and Takeuchi, 2002), and have been utilised in this study.

Fv4 blocks infection with ecotropic MLV at the stage of viral entry to target cells (Suzuki, 1975). Fv4 is an endogenous gp70 molecule that is believed to compete with the viral Env glycoprotein for the ecotropic receptor mCAT1 (Ikeda et al., 1985; Ikeda and Odaka, 1984; Inaguma et al., 1991). However, the molecular mechanism behind the Fv4 and the mCAT1 interaction remains to be resolved.

APOBEC3G is specifically packaged into the HIV-1 virions (lacking Vif) during assembly through contacts between its N-terminal domain with RNA and NC (Malim and Bieniasz, 2012; Sheehy et al., 2002). Its action results in G-to-A hypermutation on the cDNA plus strand (Bishop et al., 2004) in addition to inhibition of cDNA synthesis by RT (Bishop et al., 2008).

Finally, the recently discovered restriction factor from monocytic cells, SAMHD1 (Laguette et al., 2011; St Gelais and Wu, 2011). The mechanism of action for SAMHD1 appears to be its enzymatic breakdown of dNTPs, depleting cellular levels of dNTPs and preventing reverse transcription, although the finer details of SAMHD1 restriction

remain under investigation (Goldstone et al., 2011; Kim et al., 2012; Lahouassa et al., 2012)

1.5.3 Nuclear entry and integration

Entry to the nucleus is a necessary step for retroviral replication and occurs by distinct methods for complex and simple retroviruses. The most commonly used nuclear entry marker is the presence of two-LTR circles in infected cells, which are believed to be formed by the non-homologous DNA end joining pathway (NEHJ) in the nucleus (Li et al., 2001; Shoemaker et al., 1981). Simple retroviruses, such as MLV, fail to infect non-cycling cells (Roe et al., 1993; Suzuki and Craigie, 2007). Mitosis and break down of the nuclear envelope is required for these viruses to access the host chromatin (Roe et al., 1993). Live microscopy of MLV PICs labelled with GFP-p12 have shown that a breach in the nuclear envelope is required for the MLV PIC to enter the nucleus (Elis et al., 2012). Complex retroviruses, however, are able to pass across an intact nuclear envelope, enabling infection of chemically arrested or terminally differentiated cells (Lewis, Hensel, and Emerman, 1992). Four HIV-1 viral factors were originally implicated as being the determinants allowing infection of non-dividing cells: MA, IN, Vpr and the central ppt (Arhel, Souquere-Besse, and Charneau, 2006; Bukrinsky et al., 1993a; Gallay et al., 1997; Heinzinger et al., 1994; Lu, Spearman, and Ratner, 1993; von Schwedler, Kornbluth, and Trono, 1994; Zennou et al., 2000). However, removal of all of these factors did not render HIV-1 unable to infect non-dividing cells (Yamashita and Emerman, 2005) and the current consensus is that HIV-1 CA is the determinant of nuclear entry (Yamashita and Emerman, 2004).

The HIV-1 PIC is believed to access the nucleus via the nuclear pore complex (NPC) which is composed of nuclear pore proteins (nucleoporins or NUPs). Two RNAi screens looking for host co-factors for HIV-1 infection identified a set of nucleoporins as essential for infection. Two nucleoporins were common to both studies: Nup153 and RanBP2 (Nup358) and additionally, a karyopherin TNP03 (Brass et al., 2008; Konig et al., 2008), but many others were also implicated between the two studies. It is becoming apparent that lentiviruses can use more than one configuration of NPC for nuclear entry and mutations to CA alter the NPC utilised (Lee et al., 2010; Schaller et al., 2011). This notion is consistent with CA being the main determinant for HIV-1 nuclear entry (Yamashita and Emerman, 2004; Yamashita et al., 2007), but suggests a substantial proportion of CA must remain with the PIC. One electron microscopy study and a

recent super resolution imaging study have shown what appear to be intact CA cores at/near the NPC and have suggested a major uncoating event occuring in/at the NPC (Arhel et al., 2007; Lelek et al., 2012). However, whether the core remains more or less intact until it reaches the NPC is widely debated.

TNPO3 was of particular interest from the RNAi screens as it is a karyopherin that transports proteins from the cytoplasm to the nucleus. Initially it was shown that HIV-1 IN was the interacting partner for TNPO3 but later the important viral interaction was shown to be with CA (Christ et al., 2008; Krishnan et al., 2010). The story of TNPO3 in HIV-1 infection has become more complicated. TNPO3 knock down does not significantly block nuclear entry but a stage after nuclear entry and before integration (Zhou et al., 2011). It has been hypothesised that TNPO3 may act as an export factor for residual CA that enters the nucleus with the PIC, and that not removing this residual CA interferes with integration (Zhou et al., 2011). The interference to integration when TNPO3 is depleted requires another CA interacting molecule cleavage and polyadenylation specific factor 6 (CPSF6), but the mechanism behind this is not fully understood (Fricke et al., 2013; Price et al., 2012).

Once inside the nucleus the HIV-1 PIC makes important contacts with the host factor LEDGF, which has been intentionally omitted from this section and will be described later. The actual catalysis of integration is performed by IN, which removes a dinucleotide from the 3' ends of the LTRs (Roth, Schwartzberg, and Goff, 1989) and mediates the strand transfer reaction, inserting the viral DNA ends into the host chromosomes (Brown et al., 1989; Fujiwara and Mizuuchi, 1988).

1.5.4 The late stages of the life cycle: from transcription to virus assembly

These steps of the life cycle are not particularly relevant to this thesis and will be briefly mentioned. The provirus is transcribed by the cellular RNA polymerase II through promoter and enhancer elements in the LTR (Laimins et al., 1984; Laimins, Tsichlis, and Khoury, 1984; Ostrowski, Berard, and Hager, 1981). In the nucleus the various spliced products of the viral RNA are generated via interactions with the splicing machinery. Export of the RNA from the nucleus occurs by different methods depending on the RNA species and on the virus. Fully spliced RNAs are transported constitutively as they resemble cellular RNAs, however, intron-containing viral RNAs must be actively exported (Cullen, 2003). Different retroviruses utilise different methods to

export the unspliced RNA from the nucleus. Simple retroviruses, like Mason-Pfizer monkey virus (MPMV), contain cis-acting elements in the 3' untranslated part of the viral RNA called constitutive transport elements (CTE) which drive the export of the unspliced RNA via the Tap/NXF1 export pathway (Bray et al., 1994; Ernst et al., 1997; Pasquinelli et al., 1997). Whereas complex retroviruses code for a regulatory protein (called Rev for HIV-1), which interacts with a cis-acting element in the viral RNA (rev response element (RRE) in HIV-1), and directs active transport via the CRM1 export pathway (Cullen, 2003; Felber et al., 1989). The mechanism behind the export of MLV viral RNA remains poorly defined but also likely requires a cis-acting CTE, potentially overlapping the packaging sequence (Basyuk et al., 2005; Smagulova et al., 2005).

Translation of the viral proteins occurs on the ribosomes and the gRNA is selected for encapsidation by the NC domain of Gag either during translation or soon after (Aldovini and Young, 1990; Berkowitz et al., 1995; Dorman and Lever, 2000; Jouvenet, Simon, and Bieniasz, 2009). From work on HIV-1 it appears as though assembly initiates in the cytoplasm where small Gag/gRNA complexes are formed (Jouvenet, Simon, and Bieniasz, 2009; Kutluay and Bieniasz, 2010). These Gag molecules and complexes traffic to portions of the plasma membrane and bind via interaction between MA and the phospholipid phosphatidylinositol-(4,5)-bisphosphate (PI(4,5)P2) (Ono et al., 2004). This interaction also changes the conformation of MA, exposing the myristate for insertion into the plasma membrane (Saad et al., 2006). Multimerisation of Gag through NC and CA domains of Gag drives assembly around the gRNA and nucleates the formation of a bud to protrude from the cell membrane (Jouvenet, Bieniasz, and Simon, 2008; Jouvenet, Simon, and Bieniasz, 2009).

1.5.5 Virus budding and release from the cell

The final intracellular stages of retroviral replication require the L-domain in Gag. Three types of retroviral L-domain have been described (i) PPxY (Wills et al., 1994), (ii) PTAP (Gottlinger et al., 1991) and (iii) $LYPx_nL$ (Puffer et al., 1997). MLV and RSV contain the PPxY type L-domains present in p12 and p2b, respectively, and the function of this L-domain will be described in detail later. The central responsibility of the L-domains is in recruiting members of the endosomal sorting complex for transport (ESCRT) pathway to excise the budding virus from the cell membrane (Figure 1.6).

It has been shown that the HIV-1 L-domain, which resides in the p6 portion of Gag, interacts with the ESCRT-I protein TSG101 (Garrus et al., 2001). This interaction is essential for HIV-1 budding and knock down of TSG101 significantly reduced virus release (Garrus et al., 2001). Interestingly, HIV-1 also contains another L-domain (also in p6) of the $LYPx_nL$ type that interacts with the ESCRT accessory protein ALIX (Strack et al., 2003). Interactions between Gag and the ESCRT-I factors serve to recruit other members of the ESCRT pathway, leading to downstream recruitment of ESCRT-III members to the budding site. Charged multivesicular body proteins (CHMPs) are the components of the ESCRT-III that are needed for the scission event, and co-depletion of CHMP4B and CHMP2A blocks HIV-1 release (Morita et al., 2011). The CHMP proteins are believed to form a protruding dome structure that pulls the membrane inwards, tightening the neck of the viral bud, promoting scission and release of the virus (Guizetti et al., 2011; Hanson et al., 2008).

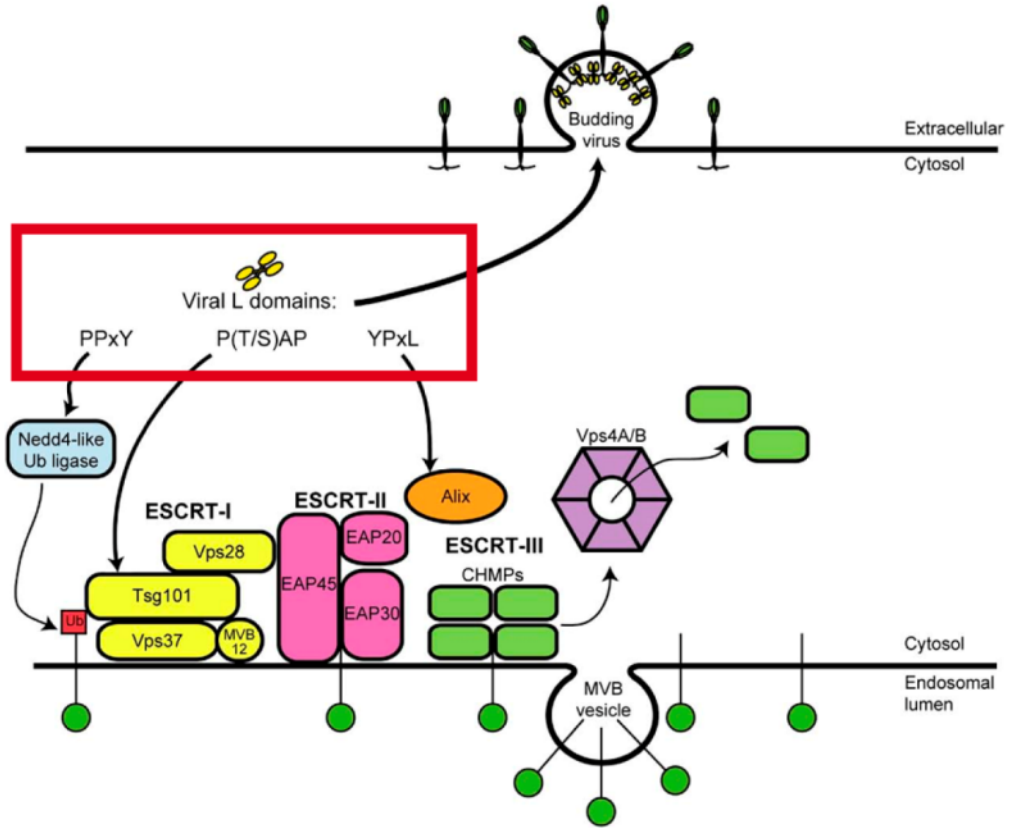


Figure 1.6: Retroviral L-domains and the ESCRT pathway. Diagram detailing how the three L-domains (in the red box) hijack the ESCRT pathway to direct the ESCRT machinery to the virion budding site. The Ub target of the Nedd4-like Ub ligases remains unknown. Adapted from (Chen and Lamb, 2008).

1.6 Maturation and formation of the retroviral CA core

Immediately after release from the cell, retrovirus particles undergo an essential maturation event. Triggered by PR, this transition can be observed by transmission EM (TEM), where an electron-lucent viral core surrounded by a 'train-track'-like organisation of Gag changes to an electron dense core of a defined shape (depending on the species of retrovirus) (Figure 1.7). Loss of PR activity blocks this conversion and renders the virus non-infectious (Crawford and Goff, 1985; Katoh et al., 1985; Kohl et al., 1988). Thus, maturation is an important therapeutic target and PR inhibitors (PI) have been in use for a number of years in the treatment of HIV-1 infection (Waheed and Freed, 2012). A new breed of antiretrovirals are currently under investigation, which instead of targeting PR activity, block maturation via alternative targets such as blocking the PR cleavage sites. One such drug is bevirimat (BMV), which blocks HIV-1 CA-SP1 cleavage by restricting PR access to the CA-SP1 junction (Waheed and Freed, 2012). Bevirimat blocks a late step in maturation and thin section EM of the virions display a phenotype similar to CA-SP1 cleavage mutants, where a dense RNP is seen, but the immature core does not dismantle for the subsequent core rearrangement (de Marco et al., 2010b; Li et al., 2003).

1.6.1 The immature Gag lattice

Original studies looking at the morphology of immature particles by EM showed a 'train-track'-like organisation of Gag under the virion membrane, projecting inwards from MA to the C-terminus of Gag (Figure 1.7) (Fuller et al., 1997; Wilk et al., 2001; Yeager et al., 1998). Estimates on the number of Gag molecules incorporated into virus particles vary considerably, with the most recent estimates being about 5000 molecules of Gag making up an immature HIV-1 virion (Briggs et al., 2004). However, these large values are likely to be overestimates as evidence is beginning to mount suggesting immature virus particles do not consist of a full sphere of Gag molecules (Briggs et al., 2009; Wright et al., 2007). The methods used to create immature Gag/CA shells range from *in vitro* forms of tubular or spherical CA to virions with inactive PR. These immature core structures have then been studied using cryo-EM. Due to the repeating nature of the immature core, averaging of the repeating units can form reasonably high resolution reconstructions of the immature lattice.

Hints that the immature CA core consisted of an ordered hexagonal lattice came from cryo-EM analysis of Moloney (Mo)-MLV PR negative particles (Yeager et al., 1998), as well as immature HIV-1 particles (Fuller et al., 1997). Immature lattices formed of CA-NC tubes have been created by preventing the formation of the β-hairpin in HIV-1, Mason-Pfizer monkey virus (MPMV) and RSV (de Marco et al., 2010a). The structures of these immature lattices have a similar hexameric form, which appear to arise from the ordered CA-NTD (de Marco et al., 2010a). However, differences do exist between different retroviral genera due the size of the spacer peptides in Gag and the spacing/distance to the nucleic acid (de Marco et al., 2010a). Recently, higher resolution cryo-EM reconstructions of MPMV immature CA-NC lattices have enabled fitting of known atomic structures of MPMV and HIV-1 CA domains into the electron densities, suggesting an immature hexamer of CA dimers (Bharat et al., 2012). This work has highlighted the potential interfaces within and between hexamers in the immature CA lattice, with the CA NTD-NTD interfaces important for inter-hexameric interactions and the CA CTD-CTD interfaces important for intra-hexameric interactions (Bharat et al., 2012).

Interestingly, recent *in vitro* studies on immature cores from a gammaretrovirus (XMRV) display a striking difference to other retroviruses (HIV-1 MPMV and RSV) in that the CA-NTD in the immature lattice appears to be more disordered (lacking hexameric order) and the hexameric order of the lattice comes from CA-CTD and NC (de Marco et al., 2010a; Hadravova et al., 2012). This implies that the interfaces supporting the immature Gag lattice in gammaretroviruses are likely to differ from those shown by Bharat *et al.* for MPMV immature CA-NC tubes.

1.6.2 The process of maturation and formation of the mature core

Upon Gag cleavage a huge rearrangement occurs inside the core of the virus (Figure 1.7) as highlighted in the models by Bharat *et al.*. Interestingly, only a fraction of the CA in the virus actually forms the mature core and for HIV-1 it has been estimated that the number is ~1500 CA molecules form the mature core (Briggs et al., 2003).

Mutations of the Gag PR cleavage sites in MLV and HIV-1 have demonstrated that the kinetics and efficiency of cleavage are important for the formation of the mature core (de Marco et al., 2012; Oshima et al., 2004). Studies into the cleavage rates of HIV-1

Gag *in vitro* have provided an estimation of the likely order of Gag cleavage (Pettit et al., 2002; Pettit et al., 2005; Pettit et al., 1994). Initially MA-CA-SP1 is cleaved from NC-SP2-p6, followed by NC-SP2 from p6, then MA from CA-SP1 and finally, CA from SP1 (Pettit et al., 2002; Pettit et al., 1994). Cleavage rates vary considerably for different sites where the cleavage of SP1 from NC-SP2-p6 is 400-fold faster than CA from SP1 (Pettit et al., 2002; Pettit et al., 1994). In MLV, the cleavage order appears to be different and p12 is separated from CA first (Naso et al., 1979). Release of the N-terminus of CA is an important step in maturation and results in the formation of the conserved β-hairpin by generating a salt bridge between the P1 and a conserved aspartate residue further downstream in CA (Gitti et al., 1996; Mortuza et al., 2004; Tang, Ndassa, and Summers, 2002; von Schwedler et al., 1998).

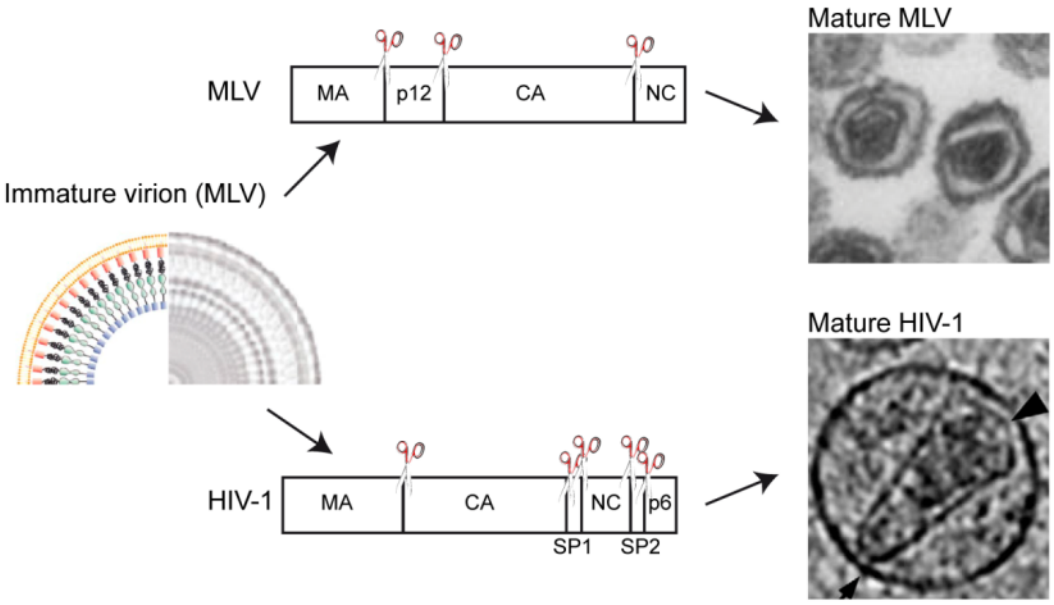


Figure 1.7: Maturation of retroviruses. Schematic showing the morphological transition between the immature and mature retroviral cores: The virion (immature virion (MLV)) on the left is an averaged cryo-EM image of an immature MLV virion displaying the domains in Gag (immature HIV-1 virions have a similar 'train track' like organisation of Gag). A schematic of MLV and HIV-1 Gag is shown with scissors marking the PR cleavage sites. The images on the right show EM micrographs of mature MLV and HIV-1 virions. (Images adapted from (Auerbach, Brown, and Singh, 2007; Briggs et al., 2006; Yeager et al., 1998)).

The mechanism behind formation of the mature core currently remains ill defined, but blocking the Gag cleavage sites in HIV-1 has enabled the analysis of potential maturation intermediates by cryo-EM. The processing of MA-CA-SP1 from NC-SP2-p6 allows condensation of the RNP (de Marco et al., 2010b) and based on the *in vitro* Gag processing kinetics and the *in vitro* condensation ability of the cleaved NC-containing

Gag products, suggests the condensation of the RNP is likely an initial stage in the maturation process (Mirambeau et al., 2007; Pettit et al., 2002). The cleavage of CA-SP1 from MA allows disassembly of the immature lattice but the mature CA lattice only forms when SP1 is cleaved from CA (de Marco et al., 2010b).

High resolution crystal structures of the hexameric CA-NTD of MLV and full-length HIV-1 CA (stabilised by thiol-linkages) have been reported (Mortuza et al., 2004; Pornillos et al., 2009). These high resolution structures have helped in delineating the intermolecular interactions between CA molecules which support the hexameric building blocks of the core. Interactions between residues in the β-hairpin and α helices 1-3 form a large intra-hexameric interface between neighbouring CA monomers, which stabilises the hexamers (Mortuza et al., 2004). Similarly, the cross-linked HIV-1 CA hexamer also contains this large intra-hexameric interface (Li et al., 2000; Pornillos et al., 2010; Pornillos et al., 2009). Adjacent hexamers were linked by CA CTD-CTD dimerisation interfaces, as have been seen in crystal structures of the CA CTD and cryo-EM reconstructions of retroviral CA lattices formed *in vitro* (Bharat et al., 2012; Gamble et al., 1997; Li et al., 2000; Pornillos et al., 2009).

One of the complexities of HIV-1 core formation, compared to formation of the MLV core, is the requirement to create a conical core. The formation of the CA cone is believed to be achieved by the insertion of 12 CA pentamers at the ends of the conical core. The structure of the HIV-1 CA pentamers have been solved, using a similar cross-linking strategy utilised for the HIV-1 CA hexamers, and a model of the HIV-1 conical core has been proposed (Pornillos, Ganser-Pornillos, and Yeager, 2011). In this model, it is proposed that the flexibility at the CA interfaces allows formation of the defined HIV-1 core shape. Recently it has been proposed that the pentamers of hexamers (a central CA pentamer surrounded by 5 hexamers) forms a more angular 'dome' surface which contributes to the formation of the conical ends of the HIV-1 CA core (Zhao et al., 2013). This study also highlighted the importance of the hydrophobic trimeric interface between the α helix 10 of the CA CTDs (Zhao et al., 2013), and this interface was necessary for the formation of the correct morphology of mature core. While our understanding of the interfaces important for the formation of the CA core have improved, the mechanism behind the formation of the mature core remains elusive.

1.6.3 Mutations and alterations affecting mature CA core formation

Many mutations have been described that have an effect on the formation of the mature retroviral CA core. Unsurprisingly, mutations in CA can affect the formation of retroviral cores. A whole range of mutations in Mo-MLV CA have been shown to be detrimental to the early stages of infection without significantly altering virus release, although the morphology of the CA core (by EM) has not always been assessed (Alin and Goff, 1996; Auerbach, Brown, and Kaplan, 2006; Auerbach, Brown, and Singh, 2007; Auerbach et al., 2003). Interestingly, one study showed that a range of insertions into the NTD of Mo-MLV CA had effects on the formation of a mature core with minimal affects on virus release (Auerbach, Brown, and Singh, 2007). Specifically, insertions in the β-hairpin caused an immature core morphology for many particles (many also displayed an aberrant core morphology), whereas mutations further downstream caused almost all virions to contain aberrant cores (Auerbach, Brown, and Singh, 2007). Similarly with HIV-1, when mutations were made in the N-terminal β-hairpin no conical cores were formed and many of these HIV-1 CA mutants displayed a condensed mass, likely the RNP, with no surrounding core (von Schwedler et al., 1998). Furthermore, similar aberrant core morphologies were also seen with deletion and insertion mutants in the NTD of HIV-1 CA (Dorfman et al., 1994; Reicin et al., 1996).

Some mutations in HIV-1 NC, in the zinc fingers, have also been reported to cause core morphology defects, although this is combined with significant defects in gRNA packaging, assembly and virus release (Grigorov et al., 2007). Moreover, mutation of the NC basic patch (Δ16-23 NC amino acids), N-terminal to the MLV zinc finger, also resulted in aberrant core morphologies where some particles even contained multiple cores (Muriaux et al., 2004). However, as with HIV-1 this mutant had a defect in gRNA packaging, assembly and release, indicating these mutations could potentially be affecting the formation of the immature Gag lattice during assembly.

Currently no host or viral factors have been described that actively participate in formation of the mature core. Very recently it has been shown that LEDGINs, which block the HIV-1 IN-LEDGF interaction, cause a late maturation defect where the majority of particles contain either: the RNP outside a formed core or no formed core (Desimmie et al., 2013). These drugs are able to induce multimerisation of Pol and the authors suggest that the drug influences the equilibrium of IN multimers within the virion, causing a defect in core formation (Desimmie et al., 2013). However, although unlikely, it was noted that the late effect of LEDGINs could be caused by inhibiting the

IN-LEDGF interaction at a late stage of infection (Desimmie et al., 2013). Interestingly, mutations in HIV-1 IN have also been described that cause a similar aberrant core morphology to the ones caused by LEDGINs (Engelman et al., 1995). Taken together this data suggests that alterations to IN can have detrimental effects on mature core formation, although whether the IN mutation causes a similar multimerisation of IN is currently unknown.

1.6.4 Stability of the CA core

The formation of a CA core of optimal stability has been shown to be essential for the early stages of infection. Mutations altering the stability of HIV-1 have significant effects on the ability to reverse transcribe in cells (Forshey et al., 2002). The CA mutations E45A and E128A/E132A both increase the stability of the HIV-1 CA core and result in a reduced ability to reverse transcribe in infected cells. Other mutations in CA were found to decrease the *in vitro* stability of the HIV-1 core and these also impacted on the ability to reverse transcribe or its kinetics *in vivo* (Forshey et al., 2002).
A similar effect has also been reported for RSV. Certain mutations in RSV CA MHR caused a reduction in virus release. However, more conservative amino acid changes, like the L117I mutant, caused a remarkably different phenotype, where virus was released but had problems with reverse transcription *in vitro* and *in vivo* (Cairns and Craven, 2001; Craven et al., 1995). Analysis of the resistance to detergent treatment of these RSV cores revealed that the mutations had made the core unstable (Craven et al., 1995). This again suggests that the formation of a core of optimal stability is essential for the early stages of infection.

1.7 Integration targeting

The design of effective and safe retroviral-based gene therapy vectors requires defined integration targeting. This was exemplified in human gene therapy trials utilising gammaretroviral-based vectors, which led to patients developing leukaemia (Cavazzana-Calvo et al., 2000; Hacein-Bey-Abina et al., 2003). Retroviral DNA integration into the host genome is not random and different retroviruses display preferences towards particular regions of the chromatin. MLV integrates close to transcriptional start sites around CpG islands (Mitchell et al., 2004; Wu et al., 2003), whereas HIV-1 has a preference for integration into actively transcribed genes (Schroder et al., 2002). Interestingly, the alpharetroviruses are the least selective and display only a weak preference for actively transcribed genes (Mitchell et al., 2004).

1.7.1 LEDGF and integration targeting

The targeting of the lentiviral intasome to chromosomal integration sites has been an area of intense investigation over the past decade. Biochemical pull down of a FLAG-tagged HIV-1 IN expressed in cells, isolated an important HIV-1 co-factor, LEDGF (Cherepanov et al., 2003). Inside the nucleus, HIV-1 IN binds to LEDGF via its integrase binding domain (IBD) (Cherepanov et al., 2004; Cherepanov et al., 2003; Maertens et al., 2003) and LEDGF then acts as a vital tether between IN and the host chromatin (Astiazaran et al., 2011; Busschots et al., 2005). The LEDGF interaction has been shown to occur with a variety of lentiviral IN proteins but not those from simple retroviruses like MLV (Busschots et al., 2005; Cherepanov, 2007).

LEDGF, through its interaction with IN and the host chromatin directly influences the integration pattern of lentiviruses, and lentiviruses infecting LEDGF-depleted cells display a different pattern of integration (Ciuffi et al., 2005; Llano et al., 2006; Marshall et al., 2007). One study has shown that the integration pattern of HIV-1 correlated with the majority of LEDGF chromatin binding sites, indicating that HIV-1 integration occurs near to the LEDGF chromatin docking sites (De Rijck et al., 2010). Interestingly, in cells depleted of LEDGF, integration targeting did not become random and another host factor with an IBD, called heptoma-derived growth factor-related protein 2 (HRP2), can drive the integration targeting in these cells when over-expressed (Cherepanov et al., 2004; Schrijvers et al., 2012; Vandegraaff et al., 2006). However, endogenous levels of HRP2 were unable to rescue the infectivity of HIV-1 in LEDGF knock down cells, suggesting its contribution in natural HIV-1 integration is minimal (Schrijvers et al., 2012).

No chromatin tethering factors for the MLV PIC had been described at the start of this project, although this has now changed and will be discussed in the light of my results in the following chapters. Previous work on HIV/MLV chimeras had, however, shown that inclusion of the MLV IN protein could change the HIV-1 integration pattern to one that is more similar to MLV (Lewinski et al., 2006). This suggested that the IN portion of MLV is a major driving force behind integration site selection. Interestingly, addition of the MLV Gag proteins into the HIV/MLV-IN chimera produced an integration pattern even closer to that of MLV, indicating that MLV Gag proteins are also involved in integration site selection (Lewinski et al., 2006). No further characterisation of the regions of Gag involved in MLV integration site selection has been performed.

A bioinformatics study had also revealed that MLV integration sites are areas of transcription factor binding (Felice et al., 2009). This suggested that transcription factors could be involved in MLV integration targeting. Moreover thirteen DNA-binding cellular factors have been shown to interact with MLV IN in yeast-two-hybrid screens including, amongst others, a panel of transcription factors TFIIE-β, B-ATF, Znfp15, Znfp38, Ankrd49 and AF9 (Studamire and Goff, 2008). However, little work has been performed on the effect of these transcription factors on MLV integration.

1.8 Gag proteins between MA and CA

Almost all genera of orthoretroviruses contain proteins between the MA and CA domains of Gag (Figure 1.8). These proteins often contain the L-domain and many have roles in viral assembly or influence the morphology of the CA core.

RSV is the prototypic alpharetrovirus and contains three proteins cleaved out from Gag between MA and CA: p2a, p2b and p10. p2b carries the RSV PPxY late domain and if removed results in a significant reduction in released virus (Xiang et al., 1996). The p10 domain of RSV Gag can convert *in vitro* assembled CA-NC tubes into spherical particles, indicating it is involved in particle morphology (Campbell and Vogt, 1997). *In vivo* p10 is involved in nuclear export of the RSV genome for packaging into assembling virions (Scheifele, Ryan, and Parent, 2005), but certain mutations in p10 also cause defects in core morphology (Scheifele et al., 2007).

The prototypic betaretrovirus mouse mammary tumour virus (MMTV) contains four proteins between MA and CA of Gag, namely: p21, p8, p3 and n (Hizi et al., 1989). Deletion of the MMTV p8 domain of Gag results in a lack of assembling virus in the cytoplasm (Zabransky et al., 2010). Deletion of the n domain of MMTV Gag results in the formation of tube-like structures in the cytoplasm and released from cells, suggesting it has a role in the morphology of particles during assembly (Zabransky et al., 2010). MPMV, another betaretrovirus, has three additional Gag proteins between MA and CA: p24/p16 (p24 is additionally cleaved to p16) and p12. Deletion of the p24/p16 region of MPMV Gag results in a budding defect as the two L-domains are positioned within p16 (PPPY and PSAP) (Gottwein et al., 2003; Yasuda and Hunter, 1998). The p12 protein is positionally analogous to p8/p3/n region of MMTV Gag (Sommerfelt, Rhee, and Hunter, 1992), and deletion of the p12 domain in MPMV Gag resulted in a defect in virus release (Sommerfelt, Rhee, and Hunter, 1992). Moreover,

betaretroviral particles initially assemble in the cytoplasm before trafficking to the membrane for release, and p12 mutant MPMV intracellular particles biochemically isolated from cells were unstable (Sommerfelt, Rhee, and Hunter, 1992). This indicated that the MPMV p12 protein is important for Gag-Gag multimerisation during assembly. This has also been shown with purified MPMV p12, which forms stable multimers *in vitro* (Knejzlik et al., 2007; Knejzlik et al., 2004) and is able to interact with itself in yeast-two-hybrid experiments, something that was not seen for the p8/p3/n protein for MMTV (Zabransky, Sakalian, and Pichova, 2005). All of which highlights a role for MPMV p12 in assembly.

The prototypic epsilon retrovirus walleye dermal sarcoma virus (WDSV) also contains a protein between MA and CA called p20 (which is also processed to p10), although currently little is known about the functions of Gag in this retrovirus (Holzschu et al., 1995).

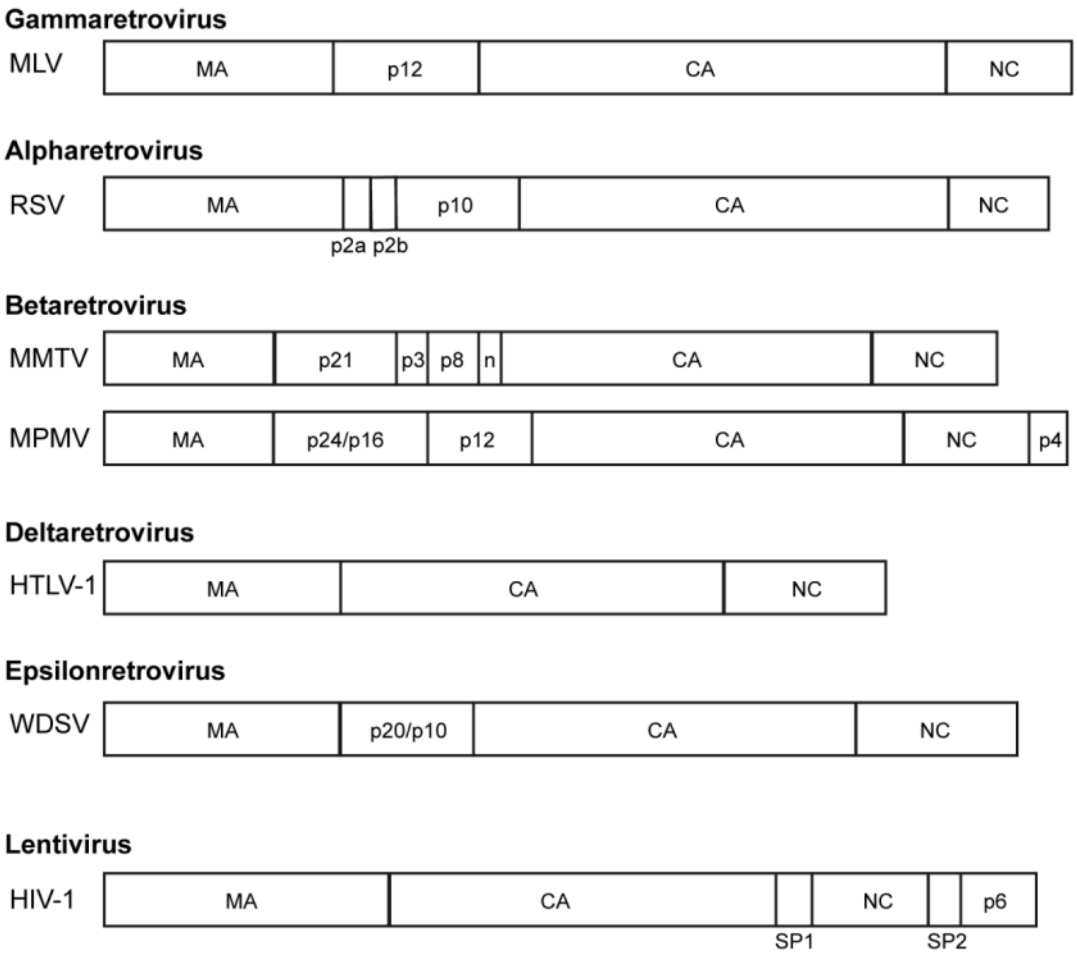


Figure 1.8: Gag protein composition. A diagram illustrating the protein domain structure for orthoretroviral Gag proteins, which are cleaved out during maturation.

1.8.1 p12

The protein situated between MA and CA in gammaretroviruses is the primary focus of this thesis, p12 (Figure 1.8). The mature p12 protein is a small polypeptide with high proline content (~20%). Nuclear magnetic resonance data suggests that p12 remains highly flexible when expressed as a fusion to the NTD of CA, suggesting it is unlikely to adopt a highly ordered structure in isolation (Kyere, Joseph, and Summers, 2008). Mutagenesis of p12 has revealed an essential requirement for this part of Gag during, both, the late and early stages of infection (Auerbach et al., 2003; Yuan et al., 2002; Yuan, Li, and Goff, 1999; Yueh and Goff, 2003).

1.8.2 p12 and the late stage of infection

L-domains are involved in recruiting members of the ESCRT pathway essential for virus release from the cell, as detailed earlier. MLV p12 contains the L-domain motif PPxY, which is utilised by a number of retroviruses (Bieniasz, 2006; Martin-Serrano et al., 2005; Yuan, Li, and Goff, 1999) and functions independent of its position in Gag (Yuan et al., 2000). Indeed, other retroviral L-domains can replace the MLV p12 L-domain function when inserted into p12 or other parts of Gag, suggesting it is a commonly utilised pathway by all retroviruses (Yuan et al., 2000). The L-domain in p12 has been shown to interact with the WW domain of cellular HECT ubiquitin ligases WWP1, WWP2, Itch and Nedd4 (Martin-Serrano et al., 2005), and MLV Gag has been shown to recruit WWP1 to the plasma membrane in an L-domain dependant fashion (Martin-Serrano et al., 2005). No interaction between WWP1 and WWP2 with members of the ESCRT pathway was seen in yeast-two-hybrid experiments, but indirect evidence for an interaction could be seen by analysing the localisation of WWP1, WWP2 and Itch in the presence of a dominant negative form of VPS4, a core component of the ESCRT pathway (Martin-Serrano et al., 2005). The budding of MLV via the PPxY L-domain interaction with WW domain proteins requires the HECT ubiquitin ligase activity, although the target of the ubiquitination is not a viral protein and currently remains unknown (Figure 1.6) (Martin-Serrano, 2007; Martin-Serrano et al., 2005; Zhadina et al., 2007).

1.8.3 The discovery of a role for MLV p12 during the early stages of infection

The first hint that p12 was involved in the early stages of MLV infection came from a study in 1984, when a deletion removing 48 amino acids in MLV Gag, 23 from the C-

terminus of MA and 25 from the N-terminus of p12, blocked infection at an early stage (Crawford and Goff, 1984). This mutant had reduced levels of linear viral DNA and no LTR circles *in vivo*, indicating a block affecting the efficiency of reverse transcription and entry of the viral DNA to the nucleus (Crawford and Goff, 1984). The *in vitro* endogenous RT activity assay indicated that the RNA genome was present and this mutant was competent for reverse transcription (Crawford and Goff, 1984). At the time, the effects on MA or p12 function could not be separated due to the nature of the deletion spanning MA and p12 sequence. However, subsequent mutagenesis experiments revealed that p12 does indeed have a function during the early stages of infection. Seven p12 mutants, four in the N-terminus (mutants 5-8) and three in the C-terminus (mutant 13-15), were blocked at an early stage of infection before entry to the nucleus, as determined by a lack (or severe reduction) of two-LTR circles in infected cells (Yuan, Li, and Goff, 1999).

Later experiments also found these regions of p12 were sensitive to insertions, with the N-terminal region slightly more sensitive than the C-terminus (Auerbach et al., 2003). Two p12 mutants in the N-terminus of p12 (mutant 6 and 8) also displayed a reduced ability to synthesise linear viral DNA *in vivo*, indicating that these mutants contain a slightly earlier block, reducing the efficiency of reverse transcription (Yuan, Li, and Goff, 1999). Mutant 6 packaged gRNA at levels equivalent to wild type, suggesting that RNA is available for reverse transcription *in vivo,* and this is not the defect with these mutants (Yuan, Li, and Goff, 1999).

1.8.4 p12 and nuclear entry of the MLV PIC

All the p12 mutants described by Yuan *et al.* were unable to access/remain in the nucleus based on a lack of two-LTR circles. Biochemical analysis of MLV p12 mutant 14 PICs isolated from cells has shown that this mutant is not significantly different from wild type PICs, indicating that mutation in the C-terminus of p12 does not affect the gross composition of the PIC (Yuan et al., 2002). Interestingly, PICs from the p12 mutant 14 are competent for integration *in vitro*, leading the authors to conclude that mutations in p12 are likely involved in localisation of the PIC to the nucleus or the specific sub-nuclear environment (Yuan et al., 2002). Furthermore, during the period of my PhD study, data has come to light stressing a role for MLV p12 in tethering the

MLV PIC to chromatin. This more recent data will be discussed in the context of my own work in the following chapters.

1.8.5 Interactions of p12 with host and viral factors

Nucleic acid

Purified MLV p12 has been shown to bind specifically to MLV RNA and not RNA from other retroviruses (Sen, Sherr, and Todaro, 1976; Sen and Todaro, 1977). Interestingly, the interaction of purified p12 did not alter the velocity sedimentation of the viral RNA, indicating the conformation of the RNA was not significantly altered by p12 binding (Sen, Sherr, and Todaro, 1976). The extent of p12-RNA binding was low with little p12 bound to the RNA, in the *in vitro* binding assays, with the authors calculated ratio of less than 15 p12 molecules per 70S dimeric genome (Sen, Sherr, and Todaro, 1976). Another study by the same group has also shown that the extent of phosphorylation of p12 dictates its affinity for RNA, where phosphorylated p12 had a higher affinity for RNA than unphosphorylated p12 or p12 saturated by phosphorylation (Sen, Sherr, and Todaro, 1977). Furthermore, it has been highlighted that the N-terminal sequence of MLV p12 has sequence homology with histone H5, which also binds to nucleic acid in a manner regulated by phosphorylation (Henderson, Gilden, and Oroszlan, 1979). However, the majority of p12 phosphorylation has been shown to be dispensable for the early function of p12 (Yueh and Goff, 2003).

Capsid

CA and p12 have been shown to colocalise in infected cells and immunoprecipitation of p12, from infected cells, can also co-precipitate CA, albeit at a low level compared to the total CA in the cells (Prizan-Ravid et al., 2010). However, work in our laboratory has so far not found a direct interaction between purified p12 and CA *in vitro* (Nader, M. unpublished data). Regardless, there is good genetic evidence to suggest that the function of p12 is linked to CA during the early stages of infection. Chimeric viruses constructed from spleen necrosis virus (SNV) and MLV could only infect cells if the CA and p12 (p18 in the case of SNV) were from the same species of virus (Lee and Nagashima, 2005). Therefore, while limited evidence for a direct interaction between p12 and CA exists, the chance that these proteins interact remains a possibility based on the genetic evidence from Lee *et al.*

Clathrin

Recently MLV has been shown to specifically incorporate clathrin via a DLL motif in p12 (Zhang et al., 2011). Knock down of clathrin using RNA interference reduced the infectivity of MLV by three-fold, whereas mutation of the DLL motif reduced infectivity by greater than 100-fold, indicating a large discrepancy between the effect of the mutation and of clathrin depletion (Zhang et al., 2011). However, the clathrin levels were only depleted to 20% of the wild type levels so this may explain the discrepancy. The authors observed a loss of mature p12 by immunoblotting when the DLL motif was mutated, and a reduced ability to saturate TRIM5alpha restriction (Zhang et al., 2011). This indicated that there was a morphological defect of the core combined with a loss of mature p12, when MLV was unable to incorporate clathrin (Zhang et al., 2011). A variety of retroviruses have now been shown to incorporate clathrin via motifs situated in varying positions in Gag and Pol (Popov et al., 2011; Zhang et al., 2011). However prevention of clathrin incorporation by removal of clathrin, or mutation of sequences driving incorporation, had varied effects on infectivity depending on the virus. Thus the exact function of clathrin incorporation is currently poorly understood.

1.8.6 Phosphorylation of p12

Some Gag proteins are phosphorylated by host proteins during infection. The function of these phosphorylation events remain poorly understood. p12 is the major phosphoprotein of MLV and phosphorylation is thought to mainly occur on serine residues (Ikuta and Luftig, 1988; Naso et al., 1979; Sen, Sherr, and Todaro, 1977). Mutation of S61 in p12 to alanine resulted in virus with significantly reduced levels of phosphorylation (~15% of wild type) and caused a block to infection at an early stage (Yueh and Goff, 2003). Mutation of other serines in p12 did not significantly reduce the levels of p12 phosphorylation, indicating that S61 is the major site of p12 phosphorylation (Yueh and Goff, 2003). Replacement of S61 with aspartic acid, to mimic the negative charge of phosphoserine, was unable to rescue the infectivity defect, and importantly revertant viruses derived from the S(61,65)A MLV mutant did not have an increase in phosphorylation (Yueh and Goff, 2003). This has been taken to indicate that the majority of p12 phosphorylation, per se, is not essential for p12 function during the early stages of infection.

A subset of the revertant viruses arising from the S(61,65)A p12 mutant contained additional positively charged residues in the C-terminus of p12 (mainly arginines) (Yueh and Goff, 2003). The authors decided to individually mutate the natural basic arginine patch in p12 (R66, R68, R70 and R71), which produced a mutants with a phenotype similar to the p12 mutants 13-15 (Yueh and Goff, 2003). Interestingly, the similar basic amino acid lysine could not replace the arginine, highlighting the importance of this C-terminal arginine patch for p12 function (Yueh and Goff, 2003).

1.9 Aims and Findings

The primary aim of this PhD project was to investigate the role of the gammaretroviral Gag protein p12 during the early stages of infection, and in doing so add to our understanding of the dynamic processes that occur during the early steps of the retroviral life cycle.

Mutations exhibiting a potent block to Mo-MLV infection during the early stages of infection have been previously described (Yuan, Li, and Goff, 1999). This project began by characterising these p12 mutants in an Mo-MLV vector system. Chapter 3 describes the phenotype of these Mo-MLV p12 mutant VLPs using a variety of assays. Importantly, Mo-MLV p12 was found to have two functional domains, which act in a non-independent manner during the early stages of infection. Additionally, the amino acids essential for function of these N- and C-terminal domains in Mo-MLV p12 were established.

Little is known about the function of p12 in other gammaretroviruses and chapter 4 details the analysis of the conservation of p12 function within the gammaretroviral genus. Similar assays to those utilised in chapter 3 have revealed that all gammaretroviruses require p12 during the early stages of infection. Closer analysis of gibbon ape leukaemia virus (GaLV) p12 revealed that it also contains two functional domains in the N- and C-terminus. Chimeras between Mo-MLV and GaLV were then generated to test if the N- and C-terminal domains serve common functions between gammaretroviruses.

Much work had been done by this point on p12 function and Chapter 5 details various experiments designed to establish the function of the N- and C-terminal domains of p12.

Importantly, functions were assigned to both domains based on the results of these experiments. To compliment the work carried out in the first three chapters, Chapter 6 details microscopy-based studies probing the localisation of p12 expressed in cells or as part of the retroviral PIC.

Chapter 2.

Materials and Methods

2.1 Recombinant DNA

2.1.1 Plasmid amplification in chemically competent cells

Plasmid DNA (<1μg) was transformed into chemically competent TOP10 *E.coli* cells and heat shocked for 30 s at 42 ^{o}C, then they were allowed to recover for 1 hr and plated on Lysogeny Broth (LB)-agar containing the selective antibiotic (either kanamycin 30 μg/ml [Sigma Aldrich] or ampicillin 50 μg/ml [Sigma Aldrich]). Plates were then incubated at 37 ^{o}C for >16hrs. A single clone was selected from the plate and grown in LB media containing the selective antibiotic at 37 ^{o}C for >16hrs.

2.1.2 Plasmid DNA purification

Purification of plasmid DNA from small bacterial cultures (2-3ml) was carried out by alkaline lysis followed by ethanol precipitation using the buffers (P1, P2 and P3) of the Plasmid Midi Kit (QIAGEN). Bacteria were pelleted and re-suspended in buffer P1 (100 μl). Cells were then lysed with buffer P2 (100 μl) and neutralised by incubation on ice for 10 min with buffer P3 (150 μl). Lysates were centrifuged at 13,000 xg for 5min, the supernatant transferred into 1 ml of pre-chilled 100 % ethanol and placed at -20 ^{o}C for 10 min. Samples were then centrifuged at 13,000 xg for 7 min and the supernatant aspirated. The pellet was washed with 70 % ethanol, centrifuged at 13,000 xg for 8 min and the supernatant aspirated. The pellet was air dried and re-suspended in dH_2O (50 μl).

Purification of plasmid DNA from larger bacterial cultures (50-100 ml) was carried out using the Midi Plasmid Purification kits (QIAGEN) according to the accompanying protocol. Plasmid DNA was resuspended in 100-200 μl of TE buffer (10 mM Tris-HCl [Fisher Scientific], 1 Ethylenediaminetetraacetic acid mM [EDTA, Sigma Aldrich] pH 8). DNA concentrations were determined by measuring the absorbance at 260 nm on a spectrophotometer (WPA). A list of all the plasmids utilised in this study can be seen in Appendix B.

2.1.3 Restriction enzyme digest

To verify the presence of the correct plasmid after purification, the DNA was digested with restriction endonucleases (New England Biolabs). Plasmid DNA (1μg) was digested with the selected restriction endonuclease in a 15 μl reaction containing the optimal buffer and at the optimal temperature for the enzyme for 1 hr. DNA loading dye (30 % glycerol, 0.25 % bromophenol blue 0.25% xylene cyanol [Sigma Aldrich], 3 μl) was added to the reaction and the DNA was separated by electrophoresis on a 1% agarose gel (Fisher Scientific) containing 0.5 μg/ml ethidium bromide (Bio-Rad). The gel was visualised using a UV transilluminator equipped with a digital camera (Bio-Rad).

2.1.4 Polymerase chain reaction (PCR)

Amplification of DNA plasmid fragments was carried out by PCR using Phusion DNA polymerase (Finnzymes). Typically, 5-10ng of plasmid DNA was subjected to PCR in a 50μl volume with the final concentrations of: 1X HF buffer (Finnzymes); 0.5 μM each primer (Sigma and Eurofins); 200 μM dNTPs (Roche); 1 U polymerase. The PCR was performed in a DNA engine thermal cycler (Bio-Rad) with the following cycling conditions: 98 ^{o}C for 30 s; 98 ^{o}C for 10 s; 65-72 ^{o}C for 20 s; 72 ^{o}C for 20s per kb and cycled 28 times with a final extension of 10 min. Amplified DNA fragments were purified using the QIAquick PCR purification kit (QIAGEN) according to the accompanying protocol and eluted in 30-50 μl dH_2O.

2.1.5 Cloning by restriction digest and DNA ligation

To clone DNA fragments into plasmids, the vector and DNA insert were digested with 0.8 μl of each restriction endonuclease (New England Biolabs) for 3 hr at the optimal temperature. The vector was additionally treated with 0.6 μl of alkaline phosphatase (Roche) for 30 min at 37 ^{o}C and another 0.6 μl for 30 min at 50 ^{o}C. Digested DNAs were separated on a 1 % agarose gel containing 0.5 μg/ml ethidium bromide and the required fragments of DNA excised on a UV transilluminator (UVP). Fragments were purified using the QIAquick Gel Extraction Kit (QIAGEN) according to the accompanying protocol. The digested insert was then ligated into the vector using the Rapid DNA ligation Kit (Roche) with the following protocol: 7 μl insert DNA; 1 μl vector DNA; 2 μl 5x DNA dilution buffer; 10 μl 2x T4 DNA ligation buffer and 1 μl T4

DNA ligase (5 U/μl). The reaction was then incubated at 20 ^{o}C for 10 min. Ligated plasmid DNA (2-4 μl) was then amplified in *E.coli* as in section 2.1.1.

2.1.6 Site-directed mutagenesis (SDM)

Substitutions and deletions were introduced into plasmids by extension of complementary oligonucleotides containing the desired changes using the Stratagene XL site directed mutagenesis kit (Aligent). Briefly, 30-50 ng of parental plasmid was amplified using 125 ng of each mutagenic primer (See Appendix A). The parental plasmid was removed by digestion with *Dpn* I (10 U, Roche) for 1 hr at 37 ^{o}C adding the same amount of enzyme after 1 hr. SDM product (2-8 μl) was amplified in *E.coli* as described in section 2.1.1. For reactions with a low yield of modified plasmid DNA, XL-10 gold cells (Aligent) were used due to their greater transformation capacity. Multiple clones from one transformation were grown in 2 ml of selective LB media for >16 hr at 37 ^{o}C and plasmid DNA was then prepared as in section 2.1.2. The resulting plasmid DNA was assessed by restriction enzyme digest and sequencing to confirm the presence of the desired change (for details of all the plasmids produced by SDM see Appendix B).

2.1.7 Production of chimeric Mo-MLV/GaLV plasmids by overlapping extension PCR

Chimeric p12 sequences were constructed by overlapping PCR from three DNA fragments amplified from either pKB4 (for Mo-MLV sequence) or pczGaLVgp (for GaLV sequence). All DNA fragments were amplified using Phusion DNA polymerase (Finnzymes) according to the manufacturer's instructions.

The amino acid sequences at the crossover junctions in the Mo-MLV based chimeras are as follows (i) *Mo-MLV/Ga-p12*: Mo-MLV MA(...SSLY/PALTDD…)GaLV p12 and GaLV p12(…DSTVIL/PLRA…)Mo-MLV CA; (ii) *Mo-MLV/Ga-Np12*: Mo-MLV MA(…SSLY/PALTDD…)GaLV p12 and GaLV p12(...LLSEPT/PPPY...)Mo-MLV p12; (iii) *Mo-MLV/Ga-Cp12*: Mo-MLV p12(...PPPY/PAALPP...)GaLV p12 and GaLV-p12(...DSTVIL/PLRA...)Mo-MLV CA. The final PCR fragments were amplified with primers for 5'-ccgcatggacacccagacca-3', and rev 5'-tggggcttctgcccgcgttt-3', and inserted into pKB4 between the *Bsr*GI and *Xho*I sites.

The amino acid sequence at the crossover junctions in the GaLV based chimeras are as follows: (i) *GaLV/Mo-p12:* GaLV MA(..PPIY/PALTPS..)Mo-MLV p12 and Mo-MLV p12-(..TTSQAF/PLRA..)GaLV CA; (ii) *GaLV/Mo-Np12:* GaLV MA(..PPIY/PALTPS..)Mo-MLV p12 and Mo-MLV p12(..EDPPPY/PAAL..)GaLV p12; (iii) *GaLV/Mo-Cp12:* GaLV p12(..SEPT/PPPYRD..)MoMLV p12 and Mo-MLV p12(..TTSQAF/PLRA..)GaLV CA. Due to the lack of suitable restriction sites in pczGaLVgp, a *Bsr*GI site was introduced into this plasmid 5' of *gag*, and an *Xho*I site 3' of *pol* was removed, both by site directed mutagenesis. The final chimeric fragments were amplified using primers, for 5'-gcagatatccagcacagtgg-3', and rev 5'-cagaagacgctccctacctg-3', and inserted into pczGaLVgp between the introduced *Bsr*GI site and a now unique *Xho*I site. The *Bsr*GI site was then removed by site directed mutagenesis.

2.1.8 DNA sequencing

All DNA sequencing was performed by Source Bioscience using ABI 3730 DNA sequencing technology. Sequencing primers used are listed in Table 2.1.

Table 2.1: Sequencing primers

Primer name	Plasmid	DNA target	Sequence (5'-3')
oKB128	pKB4 or derivatives of pKB4	p12	caagccctttgtacaccctaagcc
oKB129	pKB4 or derivatives of pKB4	p12	tctgggcgctcgaggggaaaagcg
oDW202	pKB4 or derivatives of pKB4	p12	atggacacccagaccaggtc
oVBrev1	pKB4 or derivatives of pKB4	CA	tctgctcatcctctgtcc
oDW203III	pcFeLVdw or pcFeLVgp	p12	accttatgtgaggccgaatg
oGaLVp12	pcGaLVgp	p12	caaggtcgctgttgtctctg
oXMRVp12	pHG1	p12	ctaaacccccctctcttaccg
oKB136	pCI3GN/B	p12	ggctattgcctatgaacc
oKB137	pCI3GN/B	p12	cagtttacctggatcctc

2.2 Cell culture and virus production

2.2.1 Cell culture

293T, D17, HeLa, *Mus dunni* (*M. dunni*), NIH3T3 (N-3T3), TE671, Balb/c 3T3 (B-3T3), U20S and U/R cells were maintained in Dulbecco's modified eagle medium (DMEM, Gibco) supplemented with 10 % fetal calf serum (FCS, BioSera) and 100

U/ml penicillin and 100 µg/ml streptomycin (Sigma) (from now on referred to as DMEM complete). U/R cells (a gift from Eran Bacharach) express mCAT-1, the ecotropic MLV receptor, under zeocin selection and were passaged with 0.1 mg/ml zeocin (Invitrogen). Cells were grown on 10 cm plates and were passaged every 48-72hr into fresh DMEM complete for a maximum of 28 passages. Cell stocks were stored in liquid nitrogen and were passaged 5-7 times after thawing before they were used in any assays.

2.2.2 Blocking the cell cycle in the Mitotic (M)-phase

To block at mitosis, cells were washed with DMEM complete and fresh DMEM complete with of nocodazole (0.4 µg/ml, Sigma) was added for 18 hr. To check the percentage of cells blocked in mitosis, cells were fixed with 4 % paraformaldehyde (Alfa Aesar) PBS for 20 min, washed three times with PBS and mounted on slides with vectashield (Vector labs) with 4',6-diamidino-2-phenylindole (DAPI, 1.5 µg/ml). Cells were imaged with a fluorescent microscope and the proportion of cells with condensed chromatin counted.

2.2.3 Transfection of cells with mCherry-p12 expression vectors

To express the mCherry-p12 fusion proteins, mammalian cells were transiently transfected using Turbofect *in Vitro* transfection reagent (Fermentas) according to the accompanying protocol. 293T, HeLa or *M. dunni* cells were seeded at $1x10^5$ cells per well, in a 12 well plate (Corning) containing a sterile 12 mm circular coverslip (VWR). The following day, 1-2 µg of mCherry-p12 expression plasmid (or control) was added to 180 µl of serum-free DMEM with 12 µl Turbofect (Fermentas), incubated for 20 min and added drop-wise onto the cells. After at least 16 hr cells were fixed and imaged using a fluorescent microscope (see section 2.8.7), or trypsinised, pelleted and prepared for western blotting as described in section 2.6.1.

2.2.4 Virus-like particle (VLP) production

Virus-like particles of Mo-MLV, N-MLV, B-MLV, XMRV, FeLV, GaLV and HIV-1 were prepared by co-transfection of 293T cells with a 1:1:1 ratio (2:1:1 for FeLV and 3:1:1 for GaLV) of three plasmids encoding the appropriate wild type or mutant Gag-Pol protein, VSV-G, and a reporter gene (β-galactosidase or GFP) respectively, using 7 µl (in 6-well plate) or 35 µl (for 10cm plate) polyethylenimine (PEI, PolySciences) as a

transfection reagent. To produce mixed viral particles, wild type and mutant Gag-Pol expression plasmids (or two different mutants) were added to the transfection mix at different ratios, keeping the total concentration of the Gag-Pol plasmid constant. After 24 hr, cells were washed and fresh DMEM complete was added for a further 16 hr. Virus-containing supernatants were harvested, passed through a 0.45 um filter (Sartorius), aliquoted and stored at -80 ^{o}C. In the abrogation assays, Turbofect (Fermentas) was used as transfection reagent. Approximately 18 hr post-transfection, cells were washed and sodium butyrate media (0.02 M sodium butyrate in DMEM complete) added for 6 hr before replacing it with fresh DMEM complete. VLPs were harvested after 16hr, as above.

2.2.5 Quantification of VLP reverse transcriptase activity

The reverse transcriptase (RT)-activity present in the harvested cell supernatant from VLP producer cells was quantified using the C-type RT activity Kit (CaVIDI) according to the accompanying protocol.

2.3 Infectivity assays

2.3.1 X-Gal staining of infected cells

To establish the infectious titre of *LacZ*-encoding VLPs, X-gal staining of infected cells was performed. Cells were seeded at $5x10^4$ cells per well in a 12-well plate (Corning). Cells were infected 24 hr later with 0.5 ml of a ten-fold dilution series of virus. The standard conditions for infection were 2 hr at 37 ^{o}C but for synchronised infections 2 hr at 4 ^{o}C or spinoculation (1000 xg) at 4 ^{o}C was used. After the infection, cells were washed three times with DMEM complete and incubated for 48-72 hr 37 ^{o}C. After the incubation, cells were washed once with PBS, fixed for 10 min in 2 % formaldehyde, 0.2% gluteraldehyde (Fisher Scientific and Sigma Aldrich, respectively) and stained overnight with the staining solution (0.4 mg/ml X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside, [Sigma Aldrich]), 4 mM $K_3Fe(CN)_6$ (Sigma Aldrich), 4mM $K_4Fe(CN)_6 \cdot 3H_2O$ (Sigma Aldrich), 2mM $MgCl_2$ (Fisher Scientific) in PBS). The following day the number of blue *LacZ*+ clones (groups of cells) were counted using a light microscope (Olympus). The titre in infectious units per ml (I.U./ml), was calculated by multiplying the number of clones by the dilution factor and then by 2 (to account for the fact that 0.5 ml virus was used in the infections).

2.3.2 Chemiluminescent detection of beta-galactosidase activity in infected cells

To determine the infectivity of *LacZ*-encoding VLPs a chemiluminescent assay for the presence of beta-galactosidase was used. D17 cells were seeded at $4x10^3$ cells per well in a 96-well plate (Corning). The cells were infected 24hr later with virus, normalised by their RT activity (infection was with 75 - 600 mU/ml virus). Infected cells were lysed 48-72 hr post-infection in 20 µl Tropix Lysis Buffer (Invitrogen) and stored at -20 ^oC. To assess the level of beta-galactosidase activity in the infected cell lysate 5 µl lysate was added to a white 96-well plate (Corning) and 50 µl of Tropix Galactostar reaction mixture (50:1 dilutent-to-substrate [Applied Biosciences]) added. The luminescence was measured every 10 min for 1 hr on a Tecan Safire^2 plate reader (Tecan Group Ltd) with an integration time of 1000 ms. For the luminescence readout a time point was selected from the kinetic run where the luminescence was still increasing exponentially to ensure that the chemiluminescent substrate was not limiting the enzymatic reaction. The mean luminescence from at least two uninfected controls was subtracted from the sample values to account for the background.

2.3.3 Quantification of GFP-encoding VLP infectivity by flow cytometry

VLPs encoding GFP were generated as described in section 2.2.4. *M.dunni* cells were plated at $5x10^4$ cells per well in a 24 well plate. The following day a 2-fold dilution series of the VLPs was added to the cells. Three days later the cells were trypsinised and transferred into flow cytometry tubes (BD Biosciences) with 3 % formaldehyde PBS (Fisher Scientific) and centrifuged at 530 xg for 10 min in a J6-MI centrifuge (Beckman Coulter). The supernatant was removed and the cells re-suspended by vortexing in 50µl of PBS 3mM Sodium azide (Sigma Aldrich). Infected GFP^+ cells were quantified on a FACS Calibur flow cytometer (BD Biosciences). Single cells were gated on in the forward scatter (FSC) versus side scatter (SSC) plot (G1). Fluorescence intensity detected in FL1 (excitation laser 488nm and emission filter 530/30nm) was plotted versus the cell count on a histogram and a second gate was set for GFP positive cells (G2). At least 10,000 events in G1 were recorded per sample. The percentage of infected cells for each sample was calculated in the Cell Quest software with the following formula: % of infected cells= $\left(\frac{G2}{G1}\right) \times 100$. The same gating and instrument settings (gains on detectors, thresholds, number of cells counted, etc) were used for all experiments.

2.4 Abrogation assay

The abrogation assay was used to test for the presence of the 'intact' CA core within cells infected with MLV VLPs. Tester VLPs were produced as described in section 2.2.4. The day before harvest of the VLPs, either TE671 or Balb/c3T3 cells, which express the restriction factors human TRIM5alpha and $Fv1^b$, respectively, were plated in 24 well plates (Corning) at $5x10^4$ cells per well. VLPs were harvested through sterile 0.45µm Minisart filters (Sartorius) and a 2-fold dilution series was made in DMEM. The dilutions of tester VLPs, from 1/2 to 1/128, were added to TE671 or B-3T3 cells and incubated for 4hr. Subsequently, cells were infected with a fixed volume of GFP-encoding indicator N-tropic MLV VLPs. The indicator GFP-encoding VLPs were titrated on *M.dunni* cells and the volume required to infect ~30% (~MOI 1) was used to infect the TE671 or B-3T3 cells, pre-infected with the tester VLPs. After 72 hr of infection the cells were prepared for flow cytometry as in section 2.3.3..

2.5 Isolation of MLV cores from virions (detergent layer gradient)

VLPs, from two 10 cm plates of 293T cells, were produced as described in section 2.2.4. The VLP containing cell supernatant was concentrated in a 14x89 mm polyallomer ultracentrifuge tube (Beckman Coulter), underlayed with a 2.5 ml 20 % (w/v) sucrose (VWR) cushion and by centrifuging in an SW41 rotor (Beckman Coulter) at 100,000 xg for 90 min using an Optima L-90K ultracentrifuge (Beckman Coulter) cooled to 4 ^oC. The supernatant and sucrose were aspirated and the viral pellet drained by placing the tube upside down on tissue paper for 10 min. The viral pellet was then re-suspended in 250 µl PBS on ice for 4-5 hr. Linear 10-42% (w/w) sucrose gradients were poured in 14x89mm polyallomer ultracentrifuge tubes using a Gradient Master 108 (BioComp Instruments, Inc). Layered on top of each gradient was 250 µl of 1 % Triton X-100 in 5 % (w/w) sucrose PBS or just 5 % (w/w) sucrose PBS (for VLP analysis) followed by 250 µl of 2.5 % (w/w) sucrose PBS. The re-suspended virus was gently layered on top (5 µl was kept for the input sample). The gradients were then centrifuged in an SW41 rotor at 100,000 xg for 16 hr using an Optima L-90K ultracentrifuge cooled to 4 ^oC. The next day the gradients were fractionated into 1 ml volumes using a BR-186 gradient fractionator with syringe pump (Brandel) connected to a RediFrac fraction collector (Pharmacia Biotech). For viral RNA analysis, fractions were collected in RNase-free tubes (Eppendorf) and 100 µl aliquots were treated and analysed as described in section 2.7.2 and 2.7.3, respectively. For protein analysis, 300

µl of each fraction was diluted in 100 µl 4X sodium dodecyl sulphate (SDS) loading dye (40 % (v/v) glycerol [Fisher Scientific], 12% SDS [Bio-Rad], 240 mM Tris pH 6.8 [Fisher Scientific], 0.4 M dithiothreitol [DTT, Sigma Aldrich], 0.0125 % bromophenol blue [Sigma Aldrich]). All samples were heated for 8 min at 100 ^{o}C and then analysed by SDS polyacrlyamide gel electrophoresis (PAGE) as described in section 2.6.3.

2.6 Protein analysis

2.6.1 Preparation of cell lysate for protein analysis

Cells were trypsinised from plates and transferred to 1.5 ml eppendorf tubes, centrifuged at 16,000 xg for 5 min. The supernatant was removed, the volume of the pellet was estimated and an equal volume of PBS added and vortexed. An equal volume of 0.3 M DTT (Sigma Aldrich) was added to the cell/PBS mixture and vortexed. Finally, one volume of 3x SDS loading dye was added (30 % (v/v) glycerol, 9 % SDS, 180 mM Tris pH 6.8, 0.0125 % bromophenol blue) and vortexed again. Samples were boiled at 100 ^{o}C for 10 min and analysed by SDS-PAGE as described in section 2.6.3.

2.6.2 Preparation of viral lysate for protein analysis

VLPs were harvested in the cell supernatant of transfected cells and passed through a 0.45 µm filter to remove cellular debris. For 6-well plate viruses: 1ml of VLP containing media was, underlayed with 200 µl 20 % (w/v) sucrose PBS and centrifuged at 16,000 xg for 2 hr at 4 ^{o}C. For larger volumes of virus (2-9 ml): VLP containing media was placed in a 14x89 mm polyallomer ultracentrifuge tube (Beckman Coulter), underlayed with 2.5 ml 20 % (w/v) sucrose PBS and centrifuged in an SW41 rotor (Beckman Coulter) at 28,500 rpm (~100,000 xg) using an Optima L-90K ultracentrifuge (Beckman Coulter) for 90 min at 4 ^{o}C. For the largest volumes of virus (10-30 ml): VLP containing media was placed in a 25x89 mm Ultra-Clear ultracentrifuge tube (Beckman Coulter), underlayed with 7.5 ml 20 % (w/v) sucrose PBS and centrifuged in an SW32 rotor at 28,600 rpm (100,000 xg) using an Optima L-90K ultracentrifuge (Beckman Coulter) for 90 min. After centrifugation the media and the sucrose cushion were aspirated and the viral pellet was left to dry up-side-down for 15 min. The virus pellet was re-suspended and lysed in 1X SDS loading dye (10 % (v/v) glycerol 3 % SDS, 60 mM Tris pH 6.8, 0.1 M DTT, 0.0125 % bromophenol blue). Virus lysate was boiled at 100 ^{o}C for 10 min and separated by SDS-PAGE as described in section 2.6.3.

2.6.3 Protein separation by SDS PAGE - Novex system

Samples were prepared as described in section 2.6.1 and 2.6.2 and then loaded on either a 10 % or 12 % acrylamide Novex NuPAGE gel (Invitrogen). Benchmark protein ladder (10µl, Invitrogen) was loaded as a molecular weight ladder. Gels were run at 180 V for 40 min in 1X MES running buffer (Invitrogen) in an XCell SureLock Mini-Cell (Invitrogen). To stain gels, 30 ml of SimplyBlue SafeStain (Invitrogen) was added, the gel was heated in a microwave for 30 s and was destained for 10-30 min with dH_2O. Gels were scanned with an Epson V700 scanner (Epson). For immunoblotting, proteins were transferred to either Immobilon-FL polyvinylidene difluoride (PVDF) or an Immobilon-P PVDF membrane (Millipore) for use with infrared labelled secondary antibodies or horseradish peroxidase (HRP) labelled secondary antibodies, respectively. Membranes were prepared by soaking briefly in methanol (Fisher Scientific) and rinsing in dH_2O. The wet blotting transfer was carried out in a Mini Trans-Blot Cell (Bio-Rad) in transfer buffer (25 mM Tris, 192 mM glycine [Sigma Aldrich], 30 % methanol) for either 60 min at 100 V or overnight (16 hr) at 16 V.

2.6.4 Protein separation by SDS PAGE - mini-PROTEAN 3 system

Tris-glycine gels containing 10 % or 12 % acrylamide resolving gel were poured into the mini-PROTEAN 3 gel casting system. Recipe of SDS-PAGE resolving gel: 12 % gel (12% acrylamide/bis 29:1 [Bio-Rad], 375 mM Tris pH 8.8, 0.1 % SDS, 0.1% ammonium persulfate [Bio-Rad], 0.04% Tetramethylethylenediamine [TEMED, Bio-Rad]) or 10 % gel (10% acrylamide/bis 29:1, 375 mM Tris pH 8.8, 0.1 % SDS, 0.1 % ammonium persulfate, 0.04 % TEMED). After the resolving gel had set a stacking gel was poured and a well comb inserted. Recipe of stacking gel: 5.1 % acrylamide/bis 29:1, 125 mM Tris pH 6.8, 0.1 % SDS, 0.1% ammonium persulfate, 0.1 % TEMED. Samples were loaded into wells on the gel with 10 µl of Benchmark (Invitrogen) molecular weight marker. SDS-PAGE was carried out for 60 min in running buffer (25 mM Tris, 192 mM glycine, 0.1 % SDS) buffer at 23 mA. Proteins were then transferred onto a PVDF membrane as in section 2.6.3.

2.6.5 Immunoblotting

After transfer onto a PVDF membrane (see section 2.6.3 and 2.6.4) the membrane was blocked in 1 % milk (Marvel) in PBS for 60 min. Primary antibodies were diluted in the antibody solution (1 % milk, 0.001 % Tween20 Sigma PBS) and membranes were

incubated in 4-5 ml of diluted antibody in a 50 ml falcon tube (Corning) on a roller mixer for 60 min (for antibody dilutions see Table 2.2). Membranes were then washed in PBS 0.001 % Tween20 four times (10 min each). Secondary antibodies were diluted in 5-10 ml of antibody solution (in a fresh 50 ml falcon tube) and the membrane was placed inside and incubated on a roller mixer for 60 min (for dilutions of secondary antibodies see Table 2.3). The membrane was washed four times with PBS 0.001 % tween20 (10 min each). For membranes probed with the secondary antibodies conjugated to infrared dyes, the bands were visualized using the Li-COR Odyssey imaging and quantitation system (LI-COR Bioscience). For membranes probed with secondary antibodies conjugated to horseradish peroxidase (HRP), the membranes were treated with Immobilon Western Chemiluminescent substrate (Millipore) for 1 min, then exposed on X-ray film (Scientific Laboratory Supplies). The X-ray film was processed with a Fujifilm FPM-3800A developer.

Table 2.2: Primary antibodies utilised in immunoblotting

Antibody	Source	Target	Host species	Dilution
anti-p12 monoclonal	Hybrodoma CRL-1890 (ATCC)	MLV p12	mouse	1:1,000
anti-p12 polyclonal	Gift from Jonathan Stoye	MLV p12	goat	1:2,000
anti-CA	Hybrodoma CRL-1912 (ATCC)	MLV CA	rat	1:5,000
anti-NC	Gift from Jonathan Stoye	MLV NC	goat	1:1,000
anti-MA	Gift from Alan Rein	MLV MA	rabbit	1:500,000
anti-FLAG	Sigma-Aldrich	FLAG epitope	rabbit	1:1,000
anti-dsRed	Clonetech	dsRed/mCherry/RFP	mouse	1:400

Table 2.3: Secondary antibodies utilised in immunoblotting

Antibody	Source	Host species	Dilution
anti-mouse IRDye800CW	LI-COR Bioscience	goat	1:5,000
anti-mouse HRP	Pierce	goat	1:20,000
anti-rat IRDye800CW	LI-COR Bioscience	goat	1:5,000
anti-rat HRP	Pierce	goat	1:20,000
anti-goat IRDye800CW	LI-COR Bioscience	donkey	1:5,000
anti-goat HRP	Vector Laboratories	Horse	1:20,000
anti-rabbit IRDye800CW	LI-COR Bioscience	goat	1:5,000
anti-rabbit HRP	Thermo Scientific	goat	1:10,000

2.7 Viral RNA and cDNA analysis

2.7.1 DNA isolation from infected cells

VLPs used for infections where the viral DNA was to be isolated were produced as described in section 2.2.4 with the following modification. The morning after transfection, the 293T cells were treated with 100 U of RQ1 RNase-free DNase (Promega) in the presence of 10 mM $MgCl_2$ (Fisher Scientific) for at least 6 hr. The DNase was washed away and fresh DMEM placed on the cells for the VLP harvest. For the infections, D17 cells were seeded at $3.5x10^5$ cells per well in a 6-well plate (Corning). The following day the virus was normalised by RT-activity, treated with 100 U of RQ1 RNase-free DNase in the presence of 10 mM $MgCl_2$ for 30 min at 37 ^oC, cooled to 4 ^oC for 10 min (as were the cells). The cells were then infected for 2 hr at 4 ^oC with gentle shaking allowing attachment of virions to the cells, but not entry. Cells were then washed with pre-warmed DMEM complete (apart from the 0 hr time point which washed with 4 ^oC DMEM complete) and replaced in the incubator. At various time points post-infection, cells were washed with PBS, flushed off the wells with PBS and pelleted at 10,000 xg for 5 min. The supernatant was aspirated and the pellet re-suspended in 200 µl PBS by vortexing and stored at -80 ^oC. The total DNA from infected cells was then isolated using the DNeasy Blood and tissue Kit (QIAGEN) according to the accompanying protocol.

2.7.2 Viral RNA isolation

To isolate viral RNA from sucrose gradients, a double phenol-chloroform extraction based protocol was adopted. To 100 µl of sample 10µl of RNase-free proteinase K (QIAGEN) was added and the sample was incubated at 56 ^oC for 15 min. Trizol reagent (1 ml, Invitrogen) was added and the sample vortexed. Chloroform (200 µl, Fisher Scientific) was then added and the tubes were mixed vigorously by hand for 15 s. The tubes were incubated at room temperature for 3 min and phase separation was performed by centrifugation at 12,000 xg at 4 ^oC. The aqueous phase, containing the RNA, was carefully removed and placed in a fresh 1.5 ml RNase/DNase-free eppendorf tube (Eppendorf). RNase-free glycogen (10µg, Invitrogen) was added as a carrier for the RNA and the sample vortexed. To precipitate the RNA, 500µl of propan-2-ol was added, the sample vortexed and centrifuged at 12,000 xg for 10 min at 4 ^oC. The supernatant was removed and the pellet was washed with 500 µl of 75 % ethanol (Fisher Scientific). The RNA was then pelleted at 7500 xg for 5 min at 4 ^oC. The ethanol was

aspirated and the RNA pellet air-dried for 10-15 min. The pellet was re-suspended in 50 µl of PCR grade H_2O by heating at 60 ^oC for 10 min. To remove the co-precipitated DNA, the sample was treated with 1 U of RQ1 RNase-free DNase (Promega) in RQ1 DNase 1X Reaction Buffer (40 mM Tris-HCl (pH 8.0), 10 mM $MgSO_4$ and 1 mM $CaCl_2$) at 37 ^oC for 30 min. Trizol reagent (500µl) was added and the phenol-chloroform extraction repeated, as above, with half the volumes of reaction components. The final air-dried RNA pellet was re-suspended in 25 µl of PCR-grade H_2O at 60 ^oC for 10 min. The RNA was then stored at -80 ^oC. To prepare the RNA for detection by quantitative PCR it was reverse transcribed using the High-capacity RNA-to-cDNA Kit (Applied Biosciences). Briefly, 9 µl of RNA was reverse transcribed in a 20 µl reaction at 37 ^oC for 60 min followed by 95 ^oC for 5 min and then 4 ^oC for 10 min using a DNA engine thermal cycler (Bio-Rad). The reverse transcribed viral cDNA was then stored at -20 ^oC.

2.7.3 Quantitative real time PCR for viral cDNA

To quantify the copy number of isolated viral cDNA products in infected cells or in the reverse transcribed viral RNA extracted from gradients, quantitative real time PCR was performed on samples generated from section 2.7.1 and 2.7.2. Minus strand strong stop reverse transcription products (Early RT product) were detected by amplification of the R-U3 region of Mo-MLV corresponding to nucleotides 997-1126 of the p*LacZ*-LTR plasmid using the primers oJWB45 (5'-gcgccagtcctccgatagactga-'3) and oJWB47 (5'-ctgacgggtagtcaatcactcag-'3) with probe oJWB38 (5'-FAM-atccgactcgtggtctcgctgttc-TAMRA-'3). This primer set was also used for detection of the reverse transcribed viral RNA, extracted from gradients. Second strand extension reverse transcription products (Late RT product) were detected by amplification of the R-*gag* region corresponding to the nucleotides 997-1240 of the p*LacZ*-LTR plasmid using the same forward primer, oJWB45, and probe oJWB38 but with the reverse primer oJWB48 (5'-acaatcggacagacacagataagttg-'3). Reactions were performed in triplicate on 2 µl of DNA in TaqMan Gene Expression Master Mix (Applied Biosystems) using 900 nM of each primer and 250 nM of probe. Quantitative real time PCR was performed on a Fast 7500 PCR system (Applied Biosystems) in 96-well plates (Applied Biosystems) using standard cycling conditions of 50 °C for 2 min, 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. To form a standard curve the p*LacZ*-LTR plasmid was diluted into purified gDNA from the target cell line (or PCR-grade H_2O for

the RNA isolated from gradients) to create a dilution series of samples that were used to calculate relative cDNA copy numbers and confirm the linearity of the assay.

2.7.4 MLV integration site sequencing

Integration sites, from cells infected with MLV based VLPs, were sequenced and annotated in Rick Bushman's laboratory. The infectious titre of the VLPs was established as in section 2.3.1 and HeLa cells were infected with an MOI of approximately 1.5 (or neat virus for p12 mutant 6 and mutant 14). Cells were harvested 48 hr later and the total DNA isolated as in section 2.7.1. DNA samples were then sequenced in the Bushman laboratory. Briefly, DNA samples were fragmented using Fragmentase (New England Biosciences) and DNA was purified using QiaQuick PCR purification kit (QIAGEN). The ends of the fragmented samples DNA were blunted and a polyA-tail added using Klenow 3'-5' exo- (New England Biolabs). Linkers were then ligated onto the fragmented sample DNA using T4 DNA ligase (New England Biosciences), purified using QiaQuick PCR purification kit and digested with *Dpn* I to remove vector DNA contamination. Nested PCR was performed to amplify LTR-genome DNA sequences, which were then gel purified and AMPure bead (New England Biolabs) -purified. The purified amplified DNA fragments were then analysed using 454 pyrosequencing. All sample handling, after the total DNA isolation from infected cells, and annotation to the human genome were performed by members of Rick Bushman's laboratory.

2.8 Microscopic analysis

2.8.1 Indirect immunofluorescence of VLPs fixed on coverslips

To detect viral proteins in VLPs bound to glass coverslips using indirect immunofluorescence, 300 µl of VLP containing media was placed on top of 12 mm #1.5 glass coverslips (VWR) in 24-well plates (Corning). The 24-well plate was spun at 1000 xg, 20 ^{o}C for 2 hr in a Legend RT centrifuge (Sorvall). Coverslips were gently washed with 500 µl PBS and then fixed with 500 µl of either (i) 3.7 % formaldehyde (Fisher Scientific) in piperazine-N,N′-bis(2-ethanesulfonic acid) buffer (PIPES,[Sigma Aldrich]) pH 6.7 for 10 min, (ii) 4 % paraformaldhyde (Alfa Aesar) in PBS for 20 min both at room temperature or (iii) 4 % paraformaldehyde for 5 min followed by -20 ^{o}C 100 % methanol (Fisher Scientific) for 5 min at -20 ^{o}C. For fixation (i) and (ii) the VLPs were permeabilised with 0.1 % Triton PBS for 10 min at room temperature.

Coverslips were washed twice with PBS 50 mM glycine (Fisher Scientific) and once with PBS followed by blocking with 5 % normal donkey serum (NDS) (Source Bioscience) in PBS for 30 min at room temperature. VLPs were then treated for 45 min with 400 µl primary antibody (Table 2.4) diluted in 1 % normal donkey serum (NDS, Source Bioscience) PBS. Coverslips were then washed three times with PBS and treated for 45 min with 400 µl of the secondary antibodies (Table 2.5) diluted in 1 % NDS PBS. The coverslips were kept in the dark as much as possible from here on. Coverslips were then washed three times and post-stain fixed with 400 µl 3.7 % formaldehyde in PIPES pH 6.7 for 10 min at room temperature. Coverslips were washed twice with PBS 50 mM glycine and once with PBS before being mounted in Prolong Gold mounting media (Invitrogen) on glass slides (Menzel-Gläser). The coverslips were finally sealed with clear nail varnish.

Table 2.4: Primary antibodies utilised for indirect Immunofluorescence.

Antibody	Source	Target	Host Species	Dilution
anti-p12 monoclonal	Hybridoma CRL-1890 (ATCC)	MLV p12	mouse	1:1,500
anti-p12 polyclonal	Gift from Jonathan Stoye	MLV p12	goat	1:2,000
anti-CA	Hybridoma CRL-1912 (ATCC)	MLV CA	rat	1:1,500
anti-NC	Gift from Jonathan Stoye	MLV NC	goat	1:2,000
anti-p24	ARP3243.3 (NIH AIDS Reagents program)	HIV CA (p24)	mouse	1:400
anti-myc 9E10	Gift from Eran Bacharach	c-Myc epitope	mouse	1:6 (hybridoma supernatant)
anti-CA polyclonal	Gift from Eran Bacharach	MLV CA	goat	1:1000

Table 2.5: Secondary antibodies utilised for indirect Immunofluorescence

Antibody	Source	Host Species	Dilution
Alexa Fluor 488 anti-mouse IgG (H+L)	Invitrogen	Donkey	1:2,000
Alexa Fluor 594 anti-mouse IgG (H+L)	Invitrogen	Donkey	1:2,000
Alexa Fluor 488 anti-rat IgG (H+L)	Invitrogen	Donkey	1:2,000
Alexa Fluor 594 anti-rat IgG (H+L)	Invitrogen	Donkey	1:2,000
Alexa Fluor 488 anti-goat IgG (H+L)	Invitrogen	Donkey	1:2,000
Alexa Fluor 594 anti-goat IgG (H+L)	Invitrogen	Donkey	1:2,000

2.8.2 Indirect Immunofluorescence in infected cells

Cells were seeded on sterile 13 mm coverslips (VWR) in a 12-well plate at $5x10^4$ cells per well. Cells were infected as specified in the text. The cells were then washed once

gently with PBS and fixed by either: (i) 4 % paraformaldehyde (AlfaAesar) for 20 min at room temperature or (ii) 4 % paraformaldehyde 5 min at room temperature followed by ice-cold methanol at -20 ^{o}C for 5 min. Cells were permeablised with 0.1 % Triton in PBS for 10 min and washed three times with PBS with 50 mM glycine. Cells were then stained as in section 2.8.1, except all antibodies were left on the cells for 1 hr and Prolong Gold (Invitrogen) with DAPI was used for mounting.

2.8.3 Labelling p12 with the biarsenical dye ReAsH

To follow p12 within infected cells, a tetracysteine tag (TC, sequence- CCPGCC) was cloned into the central portion of p12 to be stained with a biarsenical dye, FlAsH or ReAsH (Invitrogen). Two methods were used to label p12 containing the TC tag (referred to as p12-TC from now on): an *in vitro* cell-free labelling method and an *in vivo* labelling method. For the *in vitro* method, 1 ml of VLP containing media was treated with 50 µl of 1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES [Sigma Aldrich]) pH 7.5, 10 µl of 0.1 M β-mercaptethanol (Sigma Aldrich), 10 µl of 0.1 M *tris*(2-carboxyethyl)phosphine (TCEP [Sigma Aldrich]) and 10 µl of 1 mM 1,2-ethanedithiol (EDT [Sigma Aldrich]). After briefly mixing, 10 µl of 50 µM ReAsH-EDT_2 (Invitrogen) was added, mixed and incubated at room temperature for 2 hr. To clear the labelled VLPs from the free dye, VLPs were pelleted in an SW55 rotor (Beckman Coulter) at 40,000 rpm using an Optima L-90K ultracentrifuge (Beckman Coulter) for 30 min at 4 ^{o}C. The supernatant was then removed and the virus re-suspended in 50 µl (for staining on coverslips) or 500 µl for infections. For in vivo labelling VLPs were labelled in the producer cells. 16 hr post-transfection with the VLP plasmids the producer cells were washed twice with serum-free Opti-MEM I (Invitrogen) media. Serum-free Opti-MEM media containing 2.5 µM ReAsH-EDT_2 was added to the cells and labelled VLPs were harvested the next morning (>16 hr).

2.8.4 Detection of viral DNA by Fluorescent *in situ* hybridisation (FISH) combined with indirect immunofluorescence

D17 cells were seeded on sterile 13mm coverslips in the bottom of a 12-well plate at $5x10^4$ cells per well. The cells were then infected as specified in the text. Cells were fixed with 4 % paraformaldehyde for 10 min at room temperature, washed once with PBS and stored overnight in fresh PBS at 4 ^{o}C. The following day cells were permeabilised with 0.1 % Triton X100 in PBS for 10 min with gentle shaking. Then

they were washed in 0.1 M Tris buffer pH 7.4 and incubated in 20 % glycerol PBS for 20 min. Excess glycerol was removed by gently blotting the edge of the coverslip on tissue paper. Cells were dipped into liquid nitrogen three times for 10 s, transferred to a fresh 12 well plate and was washed with 0.1 % Triton X100 in PBS for 5 min. Cells were then blocked with 10% normal goat serum (NGS [Fisher Scientific]) in PBS for 30 min with gentle shaking. Primary antibodies were diluted in 25 mM Tris-buffered saline (TBS, Sigma Aldrich) and incubated on the cells for 1 hr with gentle shaking. Cells were washed with once with TBS and twice with PBS for 5 min each. Secondary antibodies were diluted in TBS and added to the cells for 40 min (all subsequent steps were carried out in the dark). Cells were washed again with TBS and PBS followed by a second fixation with 4 % paraformaldehyde for 1 min at room temperature. Cells were then washed twice with PBS and stored overnight in 70 % ethanol at 4 ^{o}C. The following day, the coverslips were removed, dried face up on tissue paper and placed in 6 well plates with 0.1 M HCl (Fisher Scientific) for 10 min. HCl was removed and replaced with 0.5 % Triton X100 in PBS for 10 min at 37 ^{o}C. The cells were then washed with ice-cold PBS three times for 5 min each and 1 ml fresh ice-cold PBS was added to the cells. Ice-cold fixer (1ml, 3:1 methanol:acetic acid [Fisher Scientific]) was dropped onto the cells and incubated for 20 min on ice. Fixer was removed and replaced with fresh ice-cold fixer for 15 min on ice. This replacement of fixer step was repeated four times in total. Coverslips were removed and left to air dry cell-side up on tissue paper. The coverslips were fixed to glass slides and left to air-dry overnight.

Biotin labelled probe was prepared using a nick translation kit (Roche) on a DNA fragment excised from p*LacZ*-LTR using the restriction endonucleases *Nhe*I and *Xmn*I. The probe was labelled for 90 min at 15 ^{o}C then placed on ice. Labelled probe was confirmed by denaturing a small fragment for 3 min at 95 ^{o}C and run on an agarose gel as in section 2.1.3. Probe (50ng per cell sample) was precipitated in a reaction containing: 20 µl dH_2O, 1 µl shredded (by pipetting up and down) salmon sperm (Roche), 1/10th of total volume with 3 M sodium acetate (Sigma Aldrich) and 2.5 volumes of ethanol. Precipitation was carried out at -70 ^{o}C for 1 hr. The probe was pelleted at 16,000 xg for 30 min at 4 ^{o}C, washed with 70 % ethanol and pelleted at 16,000 xg for 15 min at room temperature. The probe was allowed to air dry for 5 min and re-suspended in 4 µl (per 50ng probe) deionised formamide (Sigma Aldrich) for 10 min with constant agitation. Then 4 µl (per 50ng probe) of 2X hybridisation buffer (2%

dextral sulfate in 2X SSC buffer) was added and the reaction agitated for a further 10 min. The slides were then dehydrated using an series of ethanols (70 %, 90 % and 100 %) for 5 min each on a mixer. Slides were then dried at 37 ^{o}C and denatured in 70 % deionised formamide diluted in 2X SSC buffer (30 mM trisodium citrate, 0.3 M NaCl, pH 7) for 5 min at 75 ^{o}C. Slides were then dehydrated again (as above) and the probe denatured for 5 min at 75 ^{o}C followed by incubation on ice for 3 min. A Gene Frame (Thermo Scientific) was placed around the coverslips on the slides and 8 µl probe was spotted on to the cells. A clean coverslip was placed on top and the Gene Frame sealed. Hybridisation was carried out overnight at 37 ^{o}C in a moist chamber.

The following day, the Gene Frame was removed and the slides incubated in 50 % formamide 2X SSC to remove the top coverslips. Slides were washed three times with 50 % formamide 2X SSC (pre-warmed to 42 ^{o}C) for 5 min each. Then they were washed three times with 0.1 % SSC (pre-warmed to 60 ^{o}C) for 5 min each. Slides were blocked with 3 % BSA/4X SSC for 1 hr at 37 ^{o}C. To block non-specific avidin binding, cells were incubated with 1 mg/ml avidin (Sigma Aldrich) in 1X SSC buffer at 37 ^{o}C for 1 hr followed by three washes in 4X SSC 0.1 % Tween20 (pre-warmed to 42 ^{o}C). Circles were drawn around the coverslips using a hydrophobic barrier pen and a 1:1000 dilution of Streptavidin-AlexaFluor488 (Invitrogen) in 1 % BSA/4X SSC/0.1 % Tween20 was added for 30 min at 37 ^{o}C. The slides were then washed three times with 4X SSC/0.1% Tween20 (pre-warmed to 42 ^{o}C) for 5 min each. A drop of vectashield (Vector Laboratories) containing DAPI (1.5 µg/ml) was placed on top of the cells and a fresh coverslip was used to seal the cells.

2.8.5 Live imaging of cells infected with VLPs containing eGFP-p12

To visualise the movements of the Mo-MLV PIC, the PIC was labelled with eGFP-p12 or CA-YFP (Elis et al., 2012). Labelled Mo-MLV VLPs for these experiments were synthesised as described in section 2.2.4, but with a percentage of the Gag from pcDNA-GFP-p12 (30% of total Gag plasmids) or pB30Y (10 % of total Gag plasmids), which express cleavable GFP-p12 or CA-YFP fusions in Gag. These experiments were carried out in Eran Bacharach's laboratory at Tel Aviv University.

U/R H2A-RFP cells were seeded at 5x10^{4} cells per well on 4-chamber coverglass bottomed tissue culture dishes. Cells were either used 16 hr after seeding for infections

or 2 hr after seeding they were transfected with 1 µg S15mCherry using Polyjet DNA *In Vitro* Transfection Reagent (SignaGen Laboratories) according to the manufacturer's instructions. Fluorescent VLPs were diluted to the correct MOI (either 5 or 8, specified in the text) in DMEM complete containing 20 mM HEPES (Sigma) and warmed to 37 ^{o}C. The cells were then infected on the microscope and imaged through a x100 NA1.45 objective lens (Carl Zeiss Germany) using a Zeiss Axiovert 200 M microscope (Carl Zeiss Germany) with a Yokogawa CSU-22 spinning disk confocal head connected to an Evolve EMCCD camera (Photometrics). All images were recorded using the same exposure settings for GFP or YFP between the samples. After imaging, infected cells were washed with DMEM complete (2 hr post-infection) and returned to the incubator over night. Cells were then imaged again between 20-25 hr post-infection. Images were deconvolved using the No Neighbours algorithm in SlideBook 5 (3i) and exported as tiffs. GFP or YFP signal was scaled to emphasise the signal as appropriate for each experiment.

2.8.6 Immunofluorescence time-course of cells infected with N-terminal p12 mutants

Mo-MLV mutant VLPs containing a myc tag in p12 were synthesised as in section 2.2.4. U/R cells were seeded at 2x10^{4} cells per well on glass coverslips in 12 well plates. Cells were infected with Mo-MLV VLPs (MOI of 3, based on the RT-activity for p12 mutants) by spinoculation at 1000 xg for 2 hr at 4 ^{o}C. Cells were washed three times with cold DMEM complete and returned to the incubator at 37 ^{o}C. The cells were fixed at various times post-infection for 20 min using 4 % paraformaldehyde. Cells were then processed for indirect immunofluorescence as in section 2.8.2 with the following modifications: (i) samples were permeabilised for 2 min, (ii) samples were blocked in 3 % BSA in PBS, (iii) antibody solutions contained 1 % BSA instead of NDS, (iv) the myc antibody and CA antibody were applied sequentially rather than together and (v) no post-staining fixation was performed. Cells were imaged through a x100 NA1.45 objective lens (Carl Zeiss Germany) using a Zeiss Axiovert 200 M microscope (Carl Zeiss Germany) with a Yokogawa CSU-22 spinning disk confocal head connected to an Evolve EMCCD camera (Photometrics). Cells were chosen randomly based on the DAPI staining and a Z-stack was recorded through the imaged cells. All images were recorded using the same exposure settings for each fluorescent dye. The fluorescent signal was scaled by the same values for each fluorescent dye when constructing images from different samples.

2.8.7 Microscopy, image acquisition and analysis

Imaging of fixed mCherry-p12 transfected cells

Microscopy was performed using a DeltaVision Core widefield microscope with the x100 NA1.4 objective (Olympus). Images were recorded on an QuantEM EMCCD camera (Photometrics) using the same exposure settings for the mCherry signal. Excitation/emission filter wavelengths were: mCherry Ex555/25 nm Em605/52 nm and DAPI Ex360/40 nm Em455/50 nm. A Z-stack of images was taken at intervals of 0.2 μm and all images were analysed using Fiji (ImageJ). Images shown in this thesis show one Z-plane and fluorescent intensities are scaled by the same values for each sample.

Live imaging of mCherry-p12 transfected cells

Microscopy was performed using a DeltaVision Core widefield microscope with x63 NA1.4 objective (Olympus). Images were recorded every 5 min on an QuantEM EMCCD camera (Photometrics) using the same exposure settings for the mCherry signal. Excitation/emission filter wavelengths were: mCherry Ex555/25 nm Em605/52 nm and Hoechst 33342 Ex360/40 nm Em455/50 nm. All images were analysed using Fiji (ImageJ). Movies shown in this thesis are from one Z-plane and fluorescent intensities are scaled to emphasise the signal.

Imaging of VLPs on coverslips and infected cells

Microscopy was performed using a Deltavision core widefield micoscope with x100 NA1.4 objective (Olympus). Images were recorded on an QuantEM EMCCD camera (Photometrics) using the same exposure settings for each fluorescent dye between samples in an experiment. Excitation/emission filter wavelengths were: DAPI Ex Ex360/40 nm Em 455/50 nm; AlexaFluor488, GFP and YFP Ex490/20 nm Em525/36 nm and AlexaFluor594 Ex555/25 nm Em605/52 nm. For VLPs: eight simultaneous images were taken for each fluorescent channel. Images shown in the thesis are an average projection of the eight frames. For infected cells: A Z-stack of images were taken at intervals of 0.2 μm from the top to the bottom of the cell. Images shown in this thesis show one Z-plane and fluorescent intensities are scaled by the same values for each sample. All images were analysed by Fiji (ImageJ). Co-localisation analysis of VLPs on coverslips was carried out using the 'Co-localization finder' Plugin in Fiji, followed by manual counting of the co-localised puncta.

2.9 Electron Microscopy

2.9.1 VLP sample preparation for thin-section electron microscopy

Serum-free media containing VLPs (228ml) was produced in as in section 2.2.4. Six sets of 38 ml VLP containing media was pelleted in 14x89 mm polyallomer ultracentrifuge tubes (Beckman Coulter) at 100,000 xg for 60 min at 4 ^{o}C. Each pellet was re-suspended in 50 µl PBS for 30 min on ice. Re-suspended viral pellets (300µl total) were pooled and 300 µl of 5 % gluteraldehyde/0.2 M sodium cacodylate buffer (Fisher Scientific and Sigma Aldrich, respectively) was added, mixed and fixed for 60 min on ice. Virus was pelleted at 17,800 xg for 60 min at room temperature. The fixative was removed and 250 µl warm (37 ^{o}C) 2 % low melting point agarose (Fisher Scientific) placed on top and mixed. Virus was then centrifuged at 16,200 xg for 20 min in a centrifuge heated to 38 ^{o}C immediately followed by incubation on ice. After 30 min, ice-cold 2.5 % gluteraldehyde/0.1 M sodium cacodylate buffer was layered on top of the solid agarose (up to the top of the tube). This was left overnight on ice to completely set the agar.

2.9.2 Thin sectioning of VLPs and staining for transmission electron microscopy

The solid agarose containing the VLP pellet was removed from the eppendorf tube and washed with 0.1 M sodium cacodylate buffer for 10 min. The sample was post-fixed with 1 % osmium tetroxide for 90min, washed in 0.1 M sodium cacodylate buffer for 10 min and stained with 1 % aqueous uranyl acetate for 90 min. The samples were then dehydrated in an ethanol series (50 %, 75 %, 90 %, three times in 100 %, 5 min each) and twice in propylene oxide for 15 min each. Samples were embedded in medium Agar 100 resin and polymerised overnight at 70 ^{o}C. 50 nm-thick sections were cut from the resin block, placed on microscopy grids and stained with saturated ethanolic uranyl acetate and Reynold's lead citrate. After fixation, the sectioning and staining of the samples was performed by Liz Hirst in the Electron Microscopy department.

2.9.3 Transmission electron microscopy (TEM) and image acquisition

TEM micrographs of thin sectioned and stained VLPs were viewed at 20,000x magnification using a JOEL 1200EX transmission electron microscope (JOEL Ltd) operating at 82 kV. Images were captured on an Orius CCD camera (Gatan) using the auto-exposure mode. TEM micrographs were taken with the help and supervision of Liz Hirst.

Chapter 3.

Further investigations into the function of Mo-MLV p12 and the identification of early acting domains within p12

The importance of the MLV Gag protein, p12, during the early stages of retroviral infection has been known for over a decade (Yuan, Li, and Goff, 1999). The original study describing a role for p12 in the early stages of infection utilised a proviral Mo-MLV construct, expressing an infectious clone of Mo-MLV. Seven p12 mutants, four in the N-terminus and three in the C-terminus, were found to produce virus but were unable to cause a productive infection in susceptible cells (Yuan, Li, and Goff, 1999). Other studies have also found that these N- and C-terminal regions of p12 are intolerant to mutation (Auerbach et al., 2003; Yueh and Goff, 2003).

Mo-MLV single cycle vectors, encoding a reporter gene, offer a rapid and sensitive tool to assess the viability of mutant viruses during the early stages of infection (Soneoka et al., 1995). As Mo-MLV vectors can be easily pseudotyped using an envelope with a broad cell tropism, like the vesicular stomatitis virus glycoprotein (VSV-G), the phenotype of the p12 mutant virus like particles (VLPs) could be assessed in a range of different cellular contexts.

To validate the use of Mo-MLV VLPs, this chapter contains further analysis of the existing seven Mo-MLV p12 mutants, in a single cycle vector system. Importantly this analysis revealed the presence of two domains which function during the early stages of replication. Using genetic approaches, the relationship between these two domains was investigated and the residues essential for their function identified.

3.1 Investigating the effect of p12 mutations in an Mo-MLV vector system

All of the viruses used in this study are single cycle vectors pseudotyped with VSV-G and encoding the bacterial *LacZ* gene. A chemiluminescent substrate for the *LacZ* product, β-galactosidase, was used to assess the infectivity of Mo-MLV VLP mutants. Initially, the phenotype of the existing seven p12 mutants in Mo-MLV VLPs was investigated, to validate their use in place of viruses derived from the proviral vectors.

3.1.1 The phenotype of p12 mutants in an Mo-MLV vector system

The Mo-MLV Gag-Pol expression plasmid used here was produced by cloning the *gag-pol* region from pMD-MLV Gag-Pol into pcDNA3.1. The seven p12 mutant Gag-Pol expression plasmids were then generated by exchanging the *BsrG*I-*Xho*I fragment of *gag* from pNCS-PM5, -PM6, -PM7, -PM8, -PM13, -PM14, -PM15, -FLAG and -L-domain. The map and the position of these alterations in p12 are shown in Figure 3.1A.

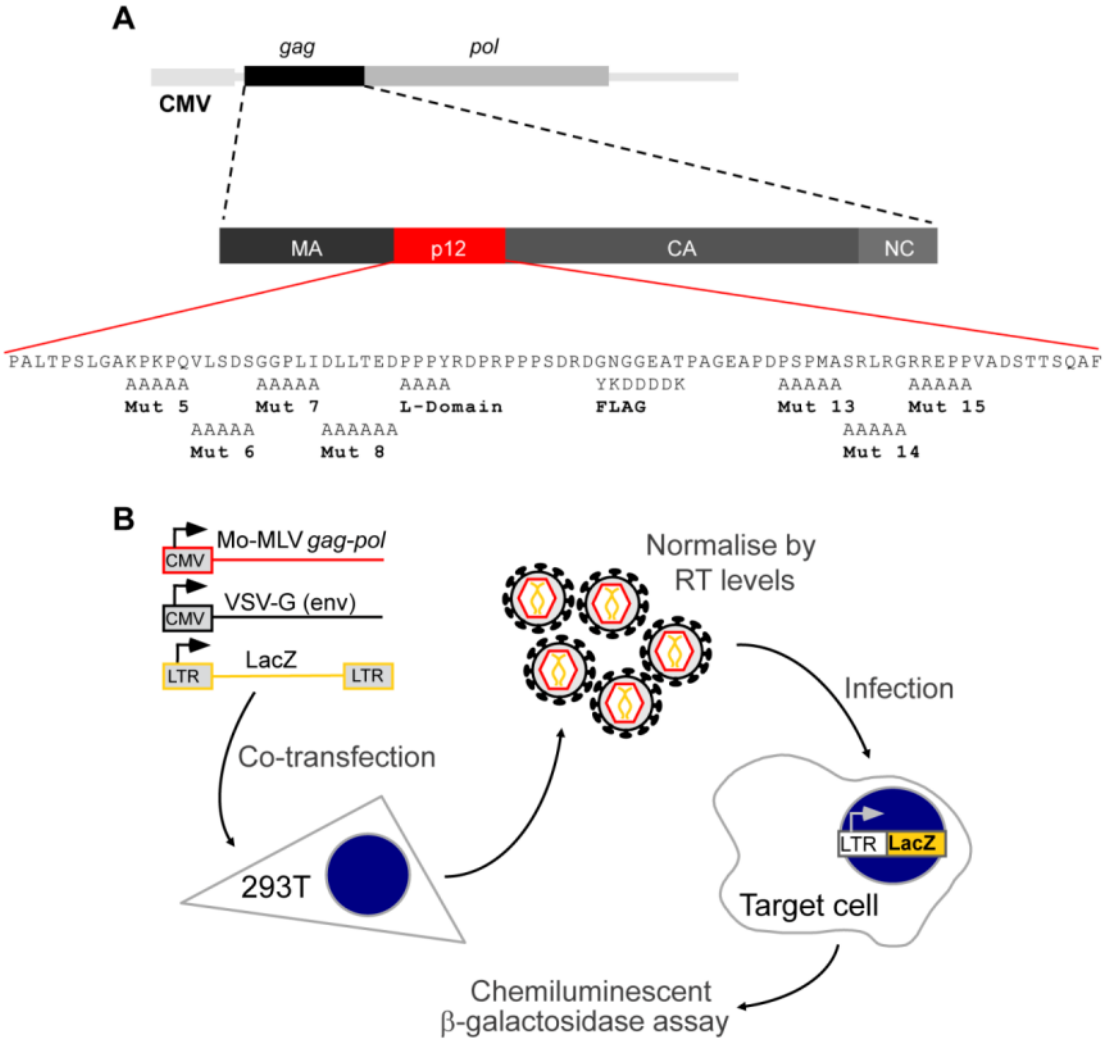


Figure 3.1: Production of Mo-MLV p12 mutant VLPs and assessment of their infectivity. (A) Position of p12 within the translated Gag polyprotein from the CMV-driven Mo-MLV Gag-Pol plasmid utilised in this study. Below the primary amino acid sequence of p12, the position of the existing p12 mutants and the FLAG tag in Mo-MLV p12 are shown (Yuan, Li, and Goff, 1999). (B) Schematic detailing how *LacZ*-encoding Mo-MLV VLPs were produced and their infectivity tested.

Mo-MLV VLPs were generated by transiently transfecting 293T cells with pMo-MLV Gag-Pol, pVSV-G and p*LacZ*-LTR (Figure 3.1B). After at least 36 hours, VLPs were harvested in the cell supernatant. To assess VLP production, a modified ELISA for reverse transcriptase activity (from now on called 'RT-ELISA') was performed on the harvested producer cell supernatant. The infectivity of the VLPs were determined by infecting target cells with equal RT-units of VLPs, incubating for 48-72 hours, followed by measuring the level of β-galactosidase activity in the cell lysate using a chemiluminescent substrate (Figure 3.1B). None of the existing p12 mutants, in this vector system, had a dramatic effect on VLP release (Figure 3.2A). This included the Mo-MLV VLPs with the FLAG-tagged wild type p12 (for the position of FLAG see Figure 3.1A). Consistent with the previous report using proviral constructs, mutant 8 had a minor effect (three-fold reduction compared to wild type) on VLP release (Figure 3.2A, dark green bar), likely due to its proximity to the Late(L)-domain (Yuan, Li, and Goff, 1999). Mutation of the L-domain itself affects the ability of Mo-MLV to recruit members of the ESCRT pathway required for efficient budding, and this mutant displayed a more significant (35-fold) reduction in VLP release (Figure 3.2A, grey bar) (Martin-Serrano et al., 2005; Martin-Serrano, Zang, and Bieniasz, 2003; Rauch and Martin-Serrano, 2011).

To gauge the effect of the p12 mutations on Gag polyprotein cleavage, immunoblotting was performed on VLPs using antibodies against CA and p12 (Figure 3.2B). All p12 mutants contained equal amounts of the cleaved CA protein (Figure 3.2B, top panel). Interestingly, the p12 monoclonal antibody recognised mutants 5 and 8 poorly, and could not recognise mutants 6 and 7 at all (Figure 3.2B, middle panel). This antibody was used in a previous publication to show that mature p12 was missing from viral particles containing mutations in the 'DLL' motif, which is mutated in mutant 8 (Zhang et al., 2011). Using both the monoclonal antibody and a polyclonal antibody against p12, it was shown that all the p12 mutants did, in fact, contain detectable mature p12 (Figure 3.2B, middle and bottom panel). Therefore the Mo-MLV VLPs containing alterations in p12 had no defect in Gag polyprotein processing by the viral PR. Interestingly the cleavage of MA-p12 appeared to be inefficient, even for wild type Mo-MLV, with a sizable proportion of the p12 present as the uncleaved MA-p12 fusion (Figure 3.2B, middle and bottom panel). This suggested that liberation of all p12 during maturation was not essential for infectivity.

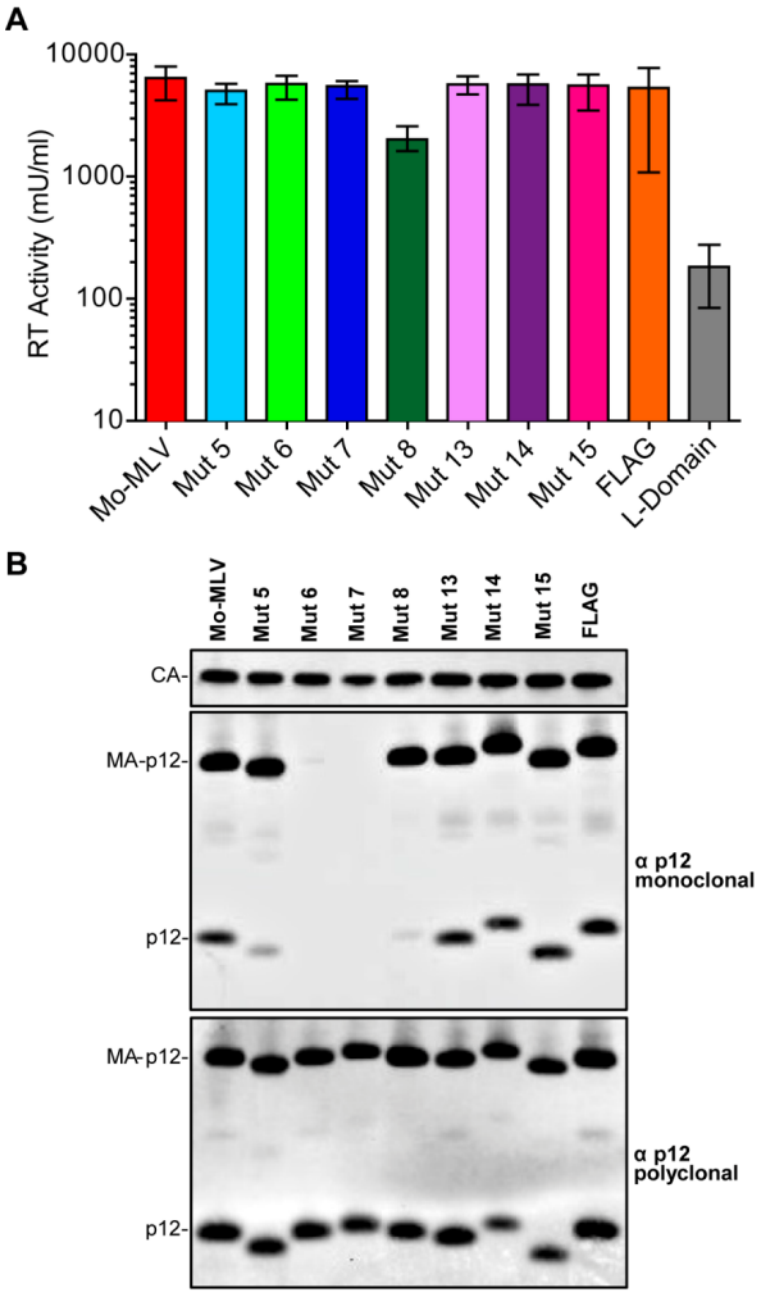


Figure 3.2: VLP release and Gag cleavage of Mo-MLV VLPs containing mutations in p12. (A) *LacZ*-encoding Mo-MLV VLPs were produced by transient transfection in 293T cells and VLP synthesis was measured by the viral reverse transcriptase activity quantified using an RT-ELISA. The histogram shows the mean RT activity of the harvested p12 mutant VLPs and the error bars indicate the range. (B) Equal RT-units of Mo-MLV VLPs were analysed by immunoblotting using anti-CA (top panel), anti-p12 monoclonal (middle panel) and anti-p12 polyclonal (bottom panel) antibodies.

To test the infectivity of the p12 mutant VLPs, D17 cells were infected with equal RT-units of mutant VLPs and the infectivity was determined using the chemiluminescent β-galactosidase assay. VLPs containing the FLAG-tagged wild type p12 were as infectious as wild type VLPs (Figure 3.3, orange bar), but the infectivity of the other

p12 mutants were reduced by 84-to-760-fold (Figure 3.3). This indicated that p12 mutants 5-8 and 13-15 in the Mo-MLV vector system recapitulated the potent infectivity defect seen for MLV p12 mutants derived from the proviral vectors.

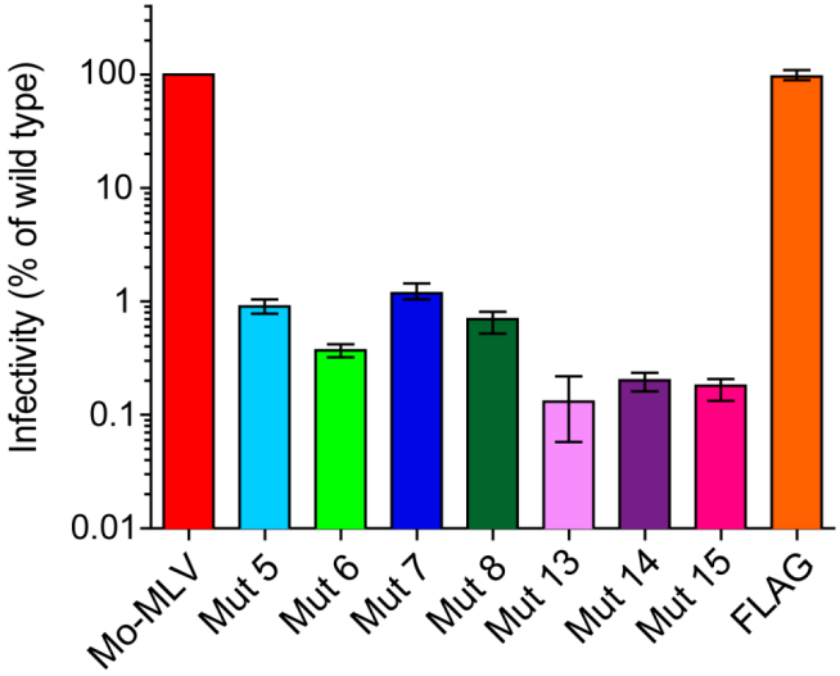


Figure 3.3: Infectivity of Mo-MLV VLPs bearing mutations in p12. D17 cells were infected with equal RT-units of p12 mutant VLPs and productive infection was quantified after at least 48 hours. The infectivity is displayed as a percentage of wild type Mo-MLV infectivity with the mean and range (highest and lowest values) of three independent experiments shown.

3.1.2 The cell species dependence of the p12 mutant infectivity phenotype

Many viral proteins have functions that are required in certain cellular environments, to counteract negative factors or provide functions that cannot be supplied by the host. As the Mo-MLV vector system is pseudotyped with VSV-G, this system allows observation of the p12 mutant phenotype in a range of cell species or types. A panel of cell lines, of different types and from different species, were infected with equal RT-units of Mo-MLV VLPs containing either wild type p12 or p12 mutant 5, 6, 8 or 13. The p12 mutants all showed a decrease in infectivity, when compared to the wild type, for all the cell lines tested (Figure 3.4). The smallest reduction in infectivity, between wild type and the most infectious mutant, was observed in CHO cells, where the reduction in infectivity was just six-fold. The largest difference was seen in the canine D17 cell line, where the infectivity was reduced by greater than 100-fold. The

preservation of the infectivity defect, over a wide range of cells, indicated that p12 did not function in a cell-specific manner.

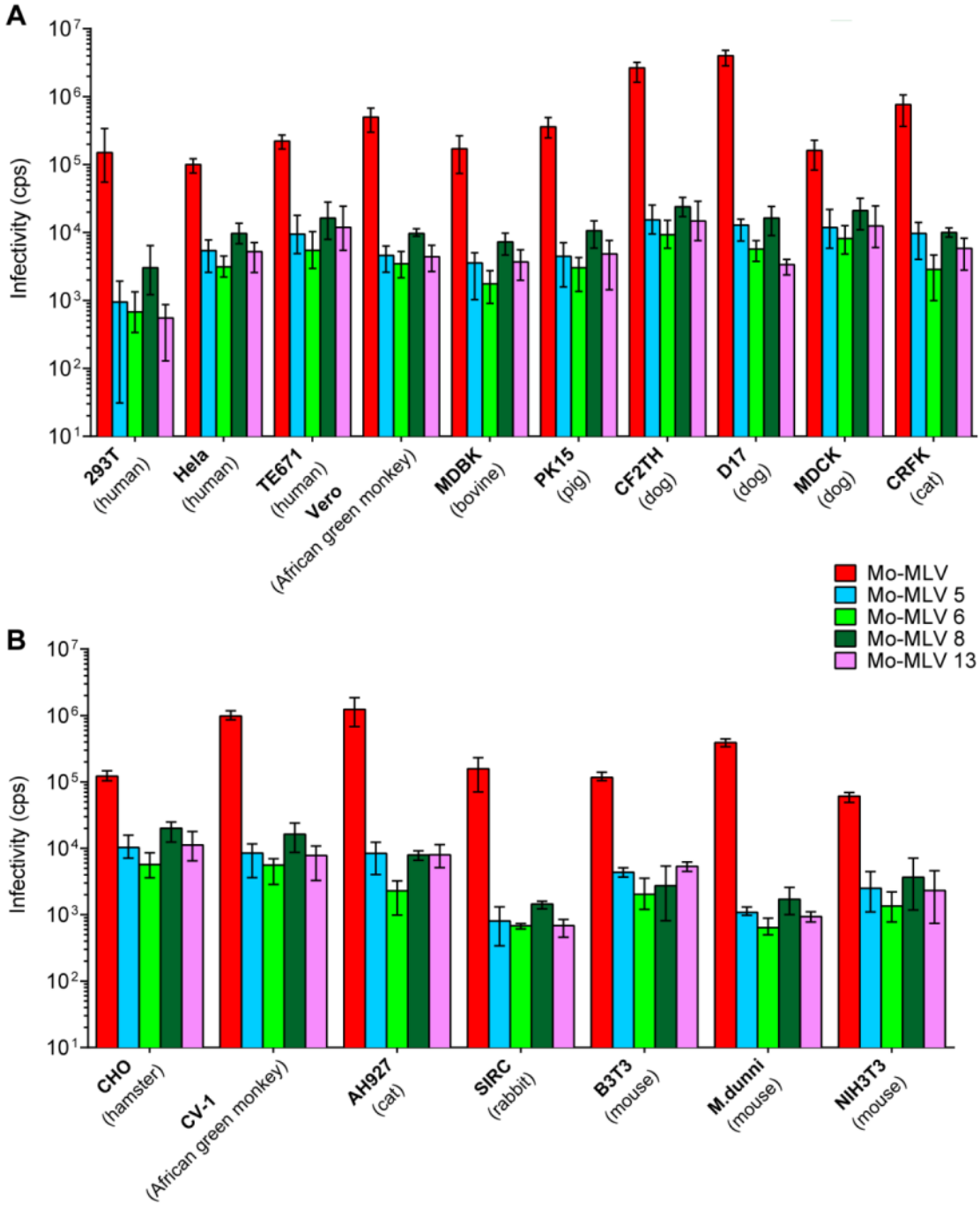


Figure 3.4: Activity of p12 mutants in a panel of cell lines. Equal RT-units of VLPs were used to challenge a panel of cell lines from different species of either (A) epithelial or (B) fibroblastic origin and productive infection was measured after 48 hours. Infectivity is displayed as luminescence counts per second (cps) with the mean and range of three independent experiments shown. (Data courtesy of Virginie Boucherit).

3.1.3 The ability of p12 mutants to reverse transcribe in infected cells

One of the hallmarks of retroviral infection is the reverse transcription of the viral RNA genome into double stranded viral DNA, which is then integrated into the host chromatin. To assess the ability of p12 mutants to reverse transcribe *in vivo,* D17 cells were infected with p12 mutant VLPs and total DNA was isolated from infected cells, at various time points post-infection. Quantitative PCR to detect both early (minus strand strong stop) and late (second strand extension) reverse transcription products was then performed.

For wild type VLPs, the presence of strong stop and second strand extension reverse transcription products appeared very early (~3 hours post infection) and peaked at around 6 hours post-infection (Figure 3.5, red line). The levels remained high throughout the time-course and were still present at two weeks post infection, indicating that the viral DNA had stably integrated into the host chromosomes.

All of the p12 mutants, barring mutant 6, synthesised both early and late viral DNA products with kinetics equivalent to that seen for wild type (Figure 3.5 A and B). However, two weeks after infection, viral DNA levels for all the p12 mutants were reduced to background levels, indicating they all failed to integrate (Figure 3.5 A and B, last points). As seen in NIH-3T3 cells (Yuan, Li, and Goff, 1999), mutant 6 failed to accumulate high levels of early, and subsequently late, reverse transcription products in D17 cells, implying this mutant has a defect in reverse transcription. Interestingly, mutant 8 was competent to reverse transcribe in D17 cells, in contrast to what has been described for mutant 8 in NIH-3T3 cells (Yuan, Li, and Goff, 1999). Unfortunately, a high background prevented the accurate measurement of viral DNA synthesis in NIH-3T3 cells infected with our mutant p12 VLPs (Virginie Boucherit, personal communication). The fact that mutant 8 could reverse transcribe in some cells and not in others suggested that the block with this mutant is probably concomitant with reverse transcription. Taken together, these results show that the Mo-MLV VLPs, generated from a Gag-Pol vector, have a phenotype that was comparable to that seen with mutant viruses derived from a proviral vector.

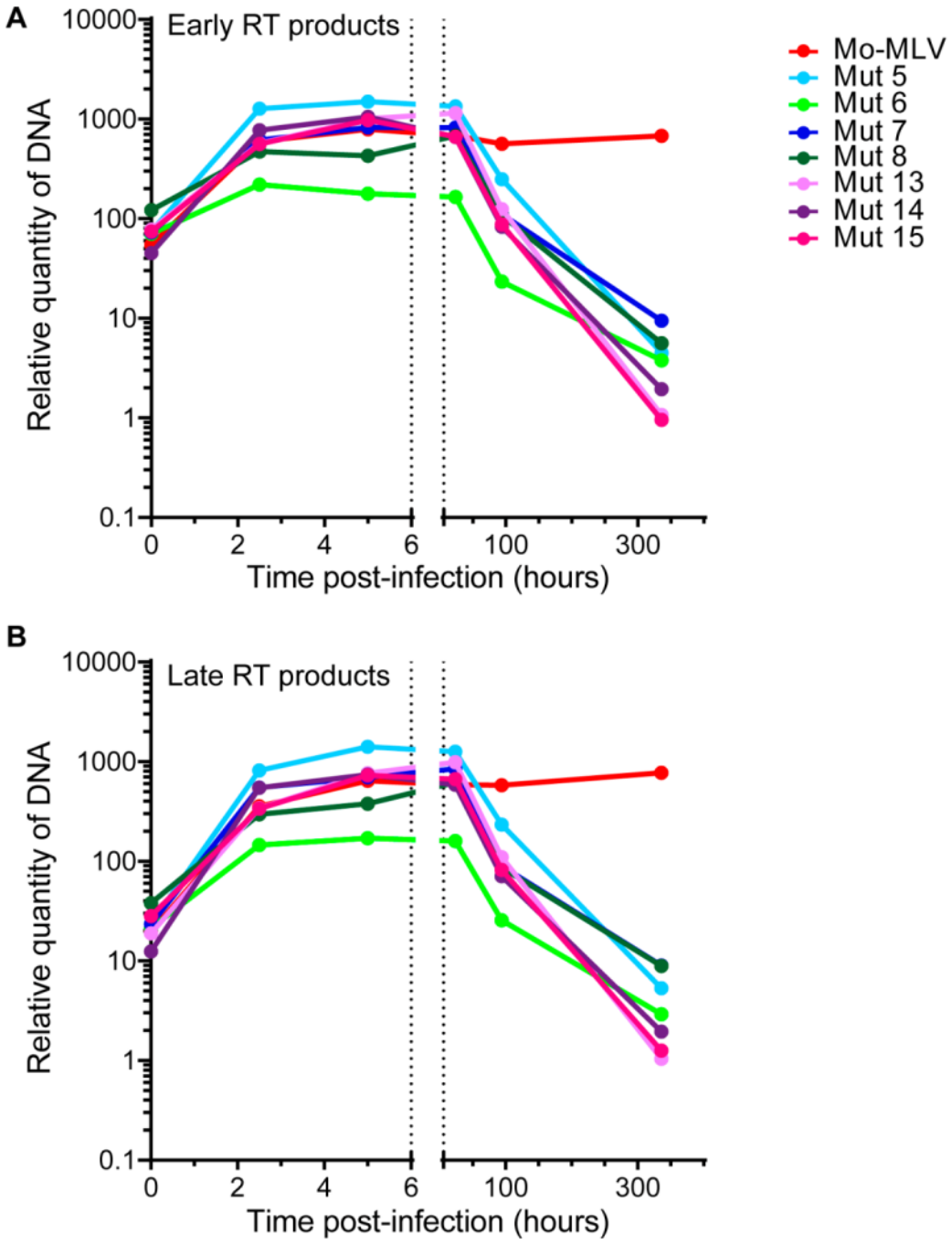


Figure 3.5: Ability of Mo-MLV p12 mutant VLPs to synthesise viral cDNA in infected cells. D17 cells were infected with equal RT-units of wild type or mutant VLPs and total DNA was isolated at various times post infection, as indicated. The relative amounts of (A) minus strand strong stop or (B) second strand extension reverse transcription products were measured by qPCR. Results are representative of at least three independent experiments.

3.2 Identification of two early acting domains in Mo-MLV p12

It had been assumed that the two regions, in the N- and C-terminus of p12, perform the same function and most of the published data concerns only alteration of the C-terminal

region of p12. No studies have addressed whether the N- and C-terminal regions of p12 carry out independent functions during the early stages of life cycle. Mutations to the N-terminal region of p12 have been shown to affect reverse transcription (Figure 3.5), but mutations to the C-terminus did not, therefore it is possible that these regions perform independent functions. This idea was investigated using genetic approaches.

3.2.1 Production of VLPs containing different p12 molecules from two separate plasmid sources

The p12 protein exists within the virion at an equimolar ratio with all of the other Gag proteins (MA, CA and NC), as Gag is incorporated into particles prior to cleavage and maturation. This does not necessarily mean that all of the packaged p12 molecules are needed by the virus for full infectivity. Therefore to establish how much wild type p12 was required in a virion for full infectivity, mixed VLPs were synthesised (Figure 3.6).

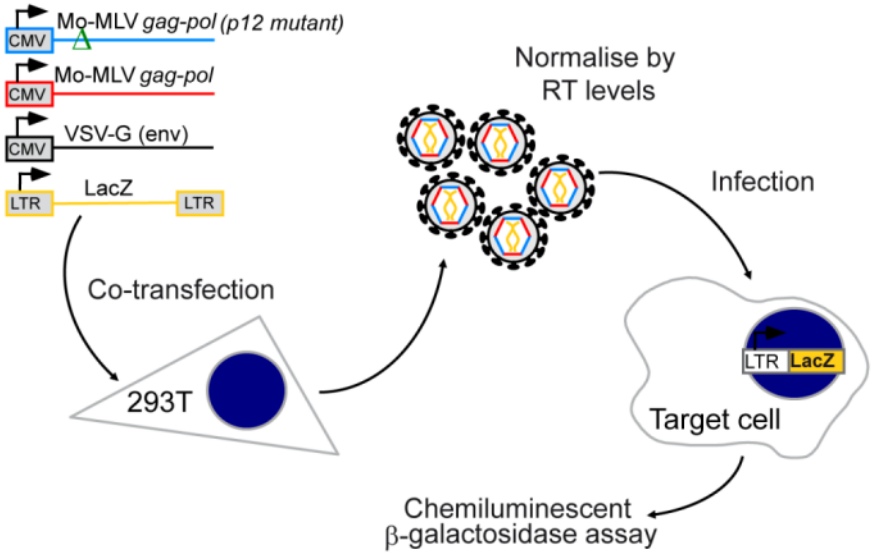


Figure 3.6: Producing mixed Mo-MLV VLPs with p12 from two Gag-Pol plasmid sources. To make Mo-MLV VLPs, which contain a mixture of wild type and mutant p12 (or two mutants), two Mo-MLV Gag-Pol expression plasmids (containing two different p12 species) were transfected into 293T cells along with the *LacZ*-LTR and VSV-G expression plasmids. VLPs were then harvested and the infectivity tested as detailed earlier.

In this case, instead of the Gag-Pol proteins coming from one plasmid source (as in Figure 3.1B), they come from two plasmid sources, one containing wild type p12 and

the other mutant p12 (or two p12 mutants). Theoretically, the ratio of different Gag proteins in the particles should reflect the ratio of the plasmids transfected. To test this, mixed Mo-MLV VLPs were synthesised containing wild type FLAG-tagged p12 and untagged mutant 6 p12 (Figure 3.7A). Equal RT-units of VLPs were analysed by immunoblotting using the anti-CA (Figure 3.7A, top panel) and anti-FLAG (Figure 3.7A, bottom panel) antibodies. The percentage of Mo-MLV Gag-Pol vector, containing the FLAG-tagged p12, in the transfection mixture is indicated below the immunoblot. For all of the mixed particles, the CA signal in the western blot remained constant (Figure 3.7A top panel), whereas the signal from the FLAG-tagged p12 decreased as the percentage of the Mo-MLV Gag-Pol, containing FLAG-tagged p12, decreased in the transfection mix (Figure 3.7A, bottom panel). This showed that the p12 present in the mixed VLPs came from the plasmids in a dose-dependent manner. Unfortunately this analysis cannot resolve between the following scenarios: (i) VLPs assembling with an actual mixture of Gag-Pol from the two plasmids transfected or (ii) two separate populations of non-mixed VLPs, each arising from one of the transfected Gag-Pol plasmids.

To distinguish between these two possibilities an experiment was devised to trans-complement the early defect induced by the p12 mutations using the p12 L-domain mutant. The L-domain function of p12 is completely independent of the early function of p12, as all of the p12 mutants exhibiting an early block in the life cycle can efficiently bud from the cell (Figure 3.2A). Therefore if true mixed particles were formed, the p12 mutant 6 should rescue budding and the L-domain mutant should rescue the early defect. This is indeed what was seen (Figure 3.7B). Mixed VLPs containing the p12 L-domain mutant and p12 mutant 6, were synthesised and D17 cells infected with equal RT-units of VLPs. The absence of the grey dashed line indicates that not enough VLPs were produced to test the infectivity, due to the L-domain defect. Therefore this synthesis protocol generated truly mixed particles containing Gag-Pol from the two transfected plasmids, in a dose-dependent manner.

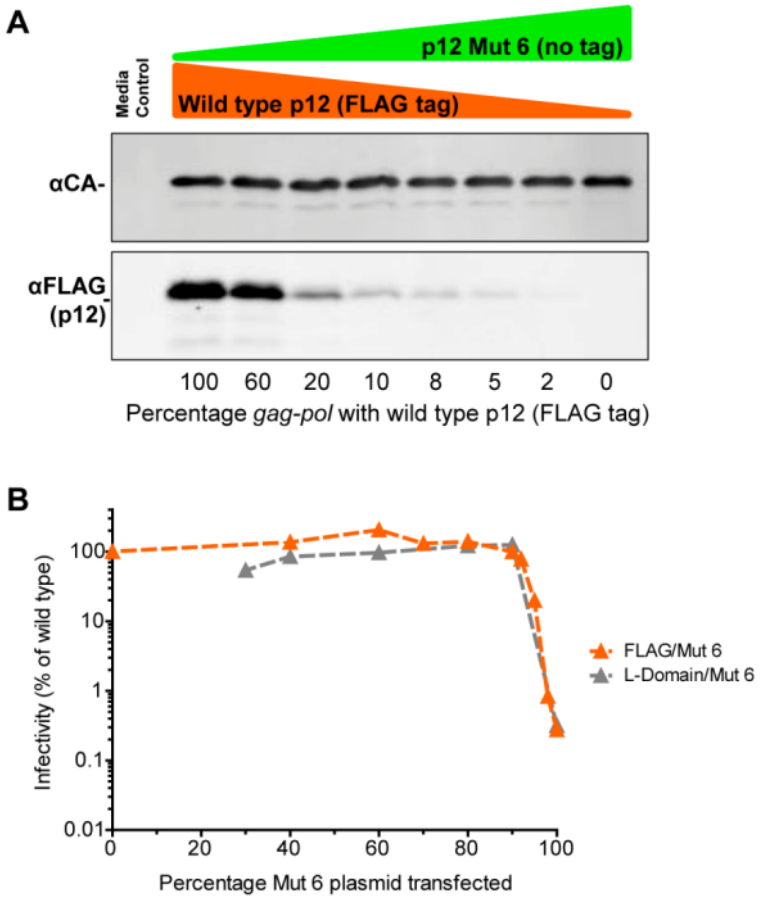


Figure 3.7: Gag protein contribution to mixed Mo-MLV VLPs made from two plasmid sources. (A) Mixed Mo-MLV VLPs were produced from two Gag-Pol expression plasmids, containing FLAG-tagged wild type p12 or untagged mutant 6 p12, and were analysed by immunoblotting with antibodies against CA (top panel) or FLAG (bottom panel). (B) Mixed Mo-MLV particles containing wild type FLAG-tagged p12 or L-domain mutant p12, mixed with mutant 6 p12 were produced. D17 cells were infected with equal RT-units of mixed Mo-MLV VLPs and productive infection was quantified after at least 48 hours. (Note: the absence of the grey dashed line indicate points where a lack of VLPs were produced for the infectivity to be tested). Infectivity is displayed as the percentage of wild type Mo-MLV infectivity and data is representative of at least three independent experiments.

3.2.2 The phenotype of mixed VLPs containing wild type and mutant p12

To establish the amount of wild type p12 required in MLV for full infectivity, mixed particles were synthesised with wild type and the p12 mutants. The infectivity of the resulting VLPs were assessed in D17 cells (Figure 3.8).

The infectivity of the p12 mutant mixed particles fell into two groups, with the exception of p12 mutant 5. The N-terminal mutants (except mutant 5) all maintained wild type levels of infectivity until over 90% of the p12 in the particle was mutant, after which the infectivity decreased sharply (Figure 3.8, light green, dark blue and dark green lines). In contrast to this, the C-terminal mutants exhibited a dramatic decline in infectivity, until they were completely non-infectious before 50% of the p12 in the particle was mutant (Figure 3.8, pink, purple and magenta lines).

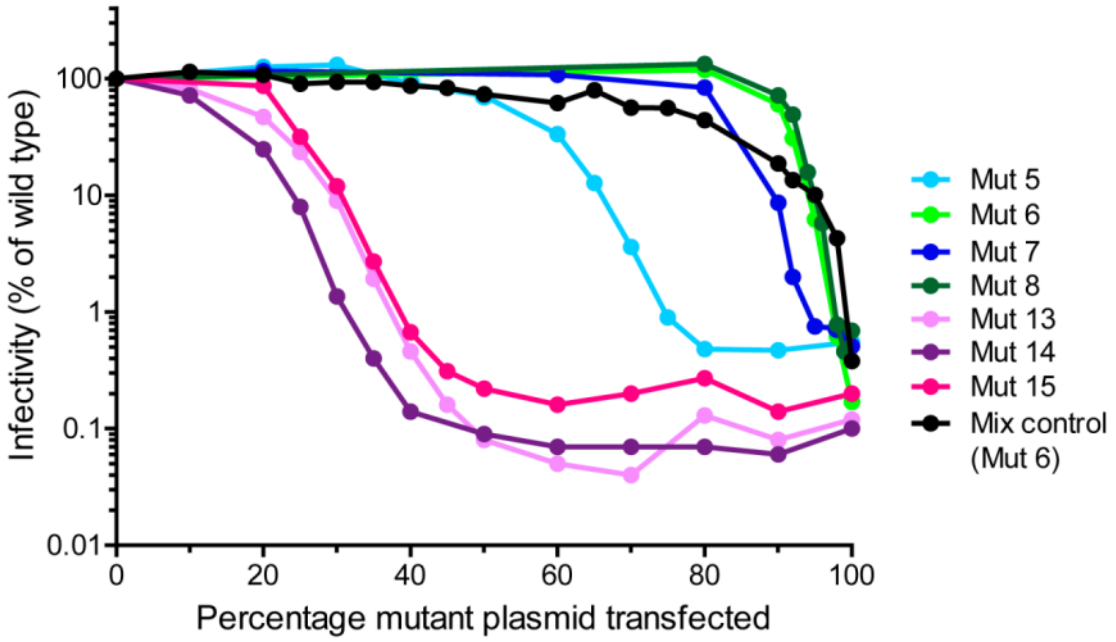


Figure 3.8: Infectivity of Mo-MLV mixed particles containing a mixture of wild type and mutant p12. Mo-MLV mixed particles, containing wild type p12 and a p12 mutant (see legend), were generated by transient transfection in 293T cells. (Note: the 'mix control' is not a mixed particle but a mixture of VLPs containing either all wild type p12 or mutant 6 p12, mixed at the percentages indicated by the x-axis.) D17 cells were infected with equal RT-units of VLPs and productive infection was quantified after at least 48 hours. Infectivity is displayed as the percentage of wild type Mo-MLV infectivity and data is representative of at least three independent experiments.

The p12 mutant 5 mixed particles remained as infectious as wild type until 60% of the p12 was mutant, then it fell until 80% of the p12 was mutant, and the particle was no longer infectious (Figure 3.8, light blue line). Therefore the N- and C-terminal p12 mutants have very differing effects on the infectivity of mixed particles. This has highlighted that only a small proportion of the p12 in the particle was required for full

infectivity and that the C-terminal p12 mutants were able to act as dominant negative mutants, interfering with wild type function.

3.2.3 The functional relationship between the N- and C-terminal regions of p12

The fact that the N- or C-terminal p12 mutants had such distinct phenotypes in mixed particles (Figure 3.8), and that the central portion of p12 was tolerant to mutation, suggested that p12 contains two functional domains (Auerbach et al., 2003; Yuan, Li, and Goff, 1999; Yueh and Goff, 2003).

To test if the N- and C-terminal regions of p12 acted independently, with regard to the infectivity of Mo-MLV, mixed particles were synthesised containing either p12 mutant 6 (N-terminal mutant) or 13 (C-terminal mutant) with all the other mutants. The infectivity of this panel of mixed VLPs was quantified and expressed as a percentage of wild type (Figure 3.9). When the p12 mutant 6 was mixed with FLAG-tagged wild type p12, the same phenotype was seen as when it was mixed with the untagged wild type p12 (compare: Figure 3.9A, orange line and Figure 3.8, light green line). Mixing any of the other p12 mutants with p12 mutant 6 in the mixed particles did not rescue infectivity (Figure 3.9A). Likewise, when p12 mutant 13 was mixed with FLAG-tagged wild type p12, the same dominant negative effect on infectivity was seen as when mixed with untagged wild type p12 (compare Figure 3.9B, orange line and Figure 3.8, pink line). Again, mixing any of the other p12 mutants with p12 mutant 13 failed to rescue the infectivity defect (Figure 3.9B). This suggested that the two domains did not function independently of each other and both were required on a single molecule of p12 for Mo-MLV infectivity.

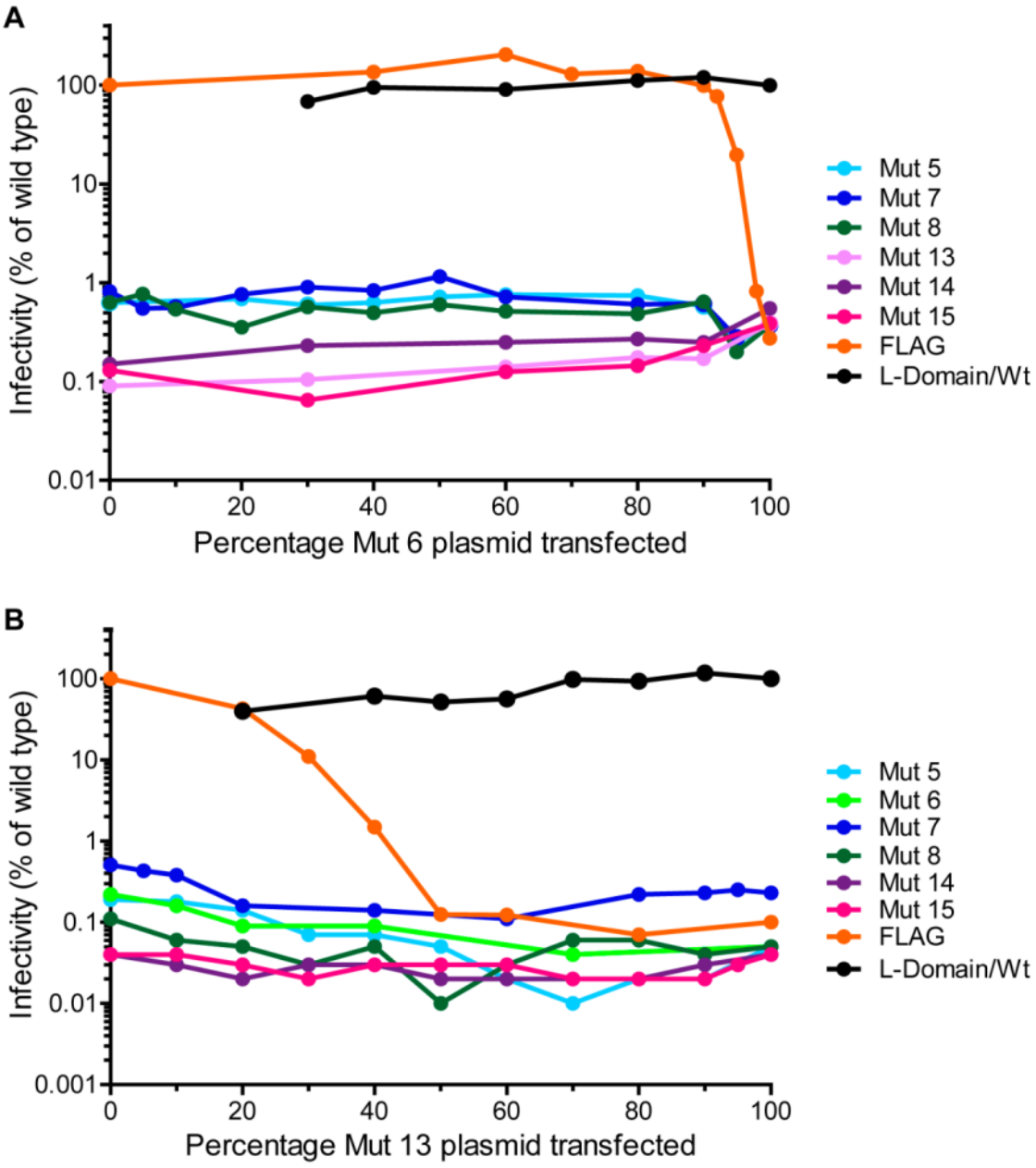


Figure 3.9: Infectivity of Mo-MLV mixed particles containing mixtures of two p12 mutants. Mixed Mo-MLV VLPs were produced by transient transfection with mixtures of (A) p12 mutant 6 or (B) p12 mutant 13 with all the other p12 mutants (see legend). (Note- The p12 L-domain mutant is mixed with wild type p12 as a control (A and B, black lines) and the percentage plasmid transfected, along the x-axis, in this case relates to wild type p12 plasmid transfected and not mutant 6 (A) or mutant 13 (B)). D17 cells were infected with equal RT-units of p12 mutant VLPs and productive infection was quantified after at least 48 hours. Infectivity is displayed as a percentage of wild type Mo-MLV and the data is representative of at least three independent experiments. Wt- wild type.

3.2.4 The effect of mutating both N- and C-terminal domains in p12

To assess which of the mixed particle phenotypes was dominant, two point mutations were introduced into p12 that knocked out the N- and C-terminal domain functions (as assessed by the infectivity of the point mutants, Figure 3.15 and Figure 3.18).

Mixed particles were synthesised and the infectivity quantified in D17 cells. Both the L16A and R66A point mutants infectivity phenotypes mirrored that of the larger substitution mutants where these residues reside, mutant 6 and 14 respectively (Figure 3.10, compare light green line and grey dashed line with green triangles, and purple line and grey dashed line with purple triangles). Interestingly, the phenotype of the p12 double mutant mixed particles, LR(16,66)AA, clearly resembled that of the N-terminal domain mutants (Figure 3.10, yellow line). Hence for the dominant negative effect to occur in mixed particles, p12 must contain an active N-terminal domain. This suggests that the N-terminus must interact with a limiting factor and/or its activity must precede that of the C-terminus.

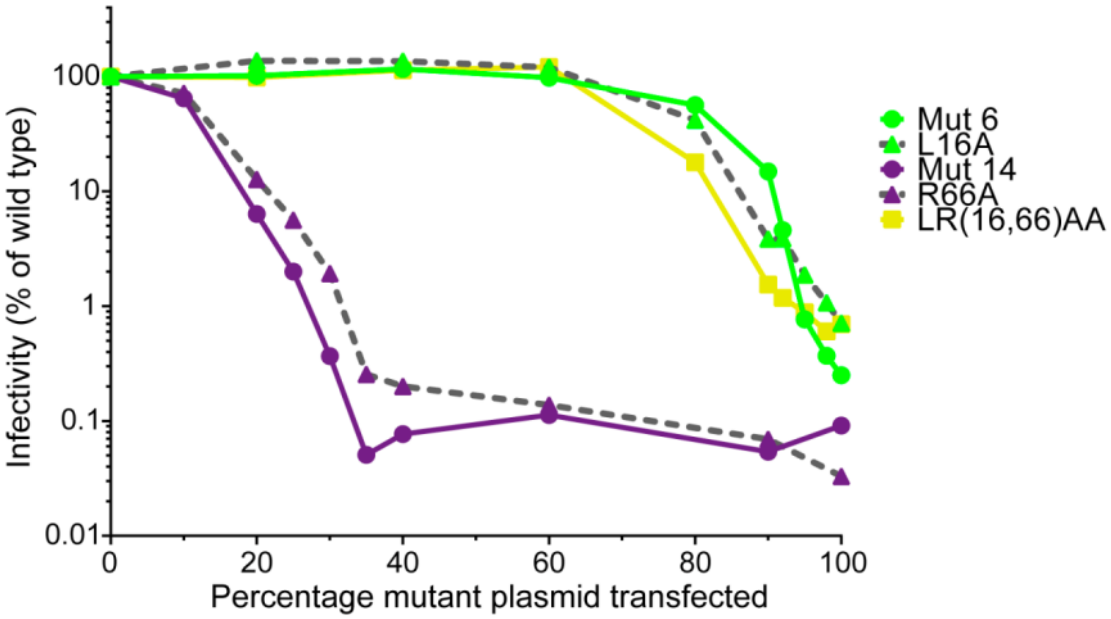


Figure 3.10: Infectivity of Mo-MLV mixed particles containing mutations in both the N- and C-terminal domains of p12. D17 cells were infected with equal RT-units of mixed VLPs and productive infection was quantified after at least 48 hours. Infectivity is displayed as a percentage of wild type Mo-MLV infectivity and the data is representative of at least three independent experiments.

3.2.5 The spatial and/or conformational requirements of the linker between the domains

The regions identified by mutation as the N- and C-terminal domains of Mo-MLV p12 are separated by 29 amino acids which, barring the L-domain (PPPY), have been shown to be tolerant to mutagenesis (Auerbach et al., 2003; Yuan, Li, and Goff, 1999; Yueh and Goff, 2003). This 'linker sequence' is a sizable proportion of the total protein length (>34% of the total amino acids), and it is possible that this stretch of amino acids is important for the optimal spacing and/or conformation of these two domains. To alter the size of the 'linker sequence' between these two domains two types of cloning were utilised (i) sequence excision by site-directed mutagenesis, and (ii) introduction of a unique *Asc*I endonuclease restriction site into p12 of the Gag-Pol vector, to introduce sequence into p12 (see Figure 3.11).

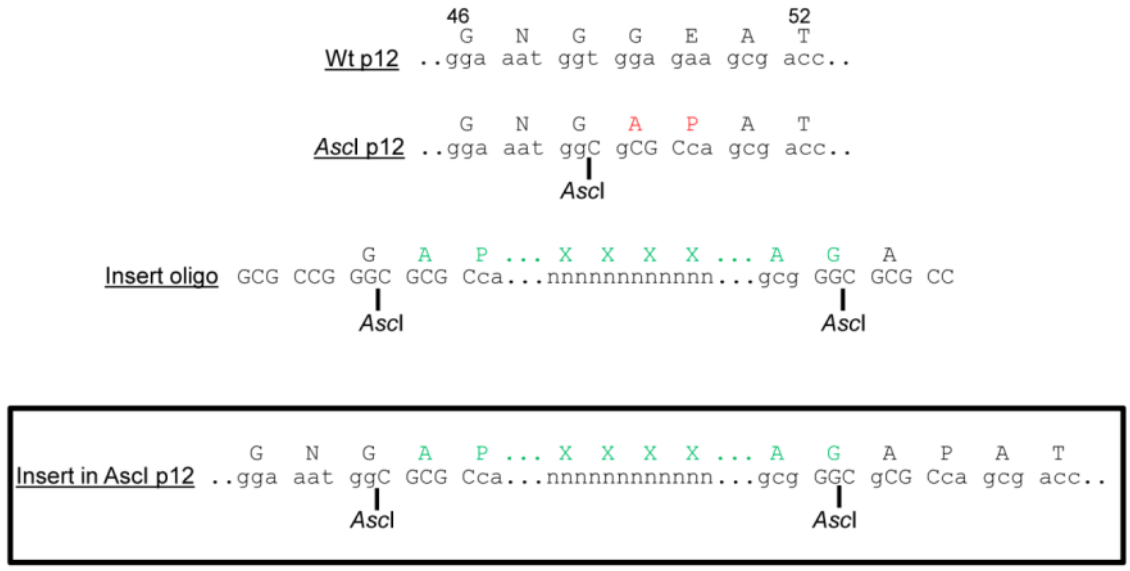


Figure 3.11: Generating large insertions in the central portion of Mo-MLV p12. An *Asc*I restriction endonuclease site was introduced into the central region of p12 by site directed mutagenesis (*Asc*I p12). Complementary oligonucleotides, including the desired sequence flanked by *Asc*I sites (Insert oligo), were annealed and digested with *Asc*I restriction endonuclease. The insert was then ligated into the *Asc*I digested Mo-MLV Gag-Pol plasmid containing the *Asc*I-p12 (Insert in *Asc*I-p12) using T4 DNA ligase.

Two deletions, of 11 and 21 amino acids, plus two insertions of 18 and 20 amino acids were generated using these methods. *LacZ*-encoding Mo-MLV VLPs containing the deletions or insertions in p12 were synthesised and the infectivity quantified. Neither insertion nor deletion of sequence between the N- and C-terminal domains of wild type p12 significantly reduced the infectivity of Mo-MLV (Figure 3.12). This indicated that the length of the 'linker sequence' after the L-domain in p12, but before the C-terminal domain, could be extended or shortened without affecting p12 function.

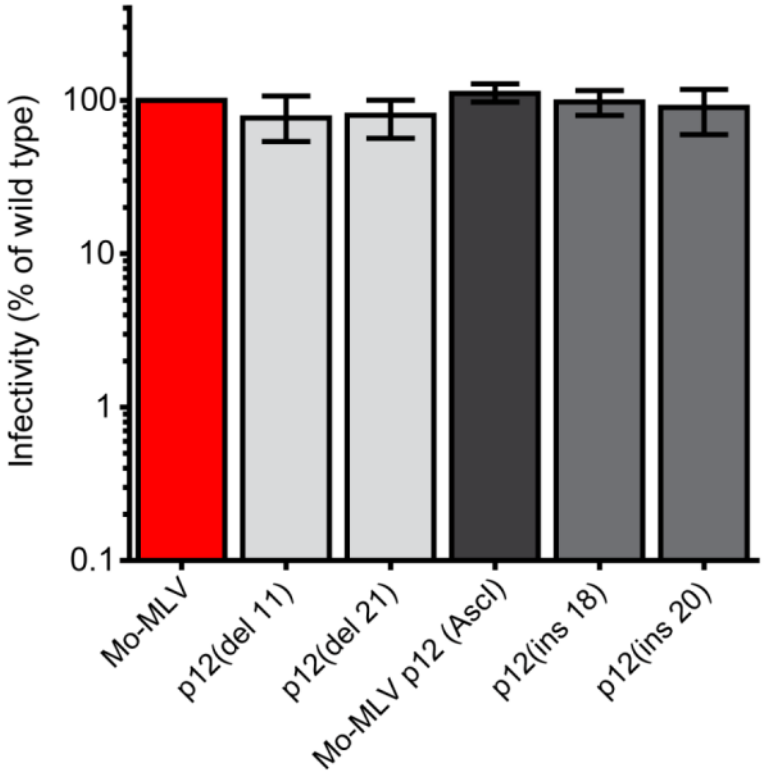


Figure 3.12: Infectivity of Mo-MLV containing insertions or deletions in the central portion of p12. Deletions or insertions were cloned into the central portion of p12, in the Mo-MLV Gag-Pol expression plasmid. D17 cells were infected with equal RT-units of VLPs and productive infection was quantified after at least 48 hours. The infectivity is displayed as a percentage of wild type Mo-MLV infectivity with the mean and range of three independent experiments shown.

A short peptide from GFP (sGFP11, RDHMVLHEYVNAAGIT) has been described as a genetic tag for expressed proteins, which when co-expressed with the non-fluorescent

remaining proportion of GFP (sGFP1-10) will self-complement to form the fluorescent GFP molecule (Cabantous, Terwilliger, and Waldo, 2005). In an attempt follow p12 movements in infected cells, the sGFP11 tag was cloned into p12 using the introduced *Asc*I restriction endonuclease site (Figure 3.11). Frustratingly, but interestingly, this sGFP11 tag in p12 decreased the infectivity of Mo-MLV by >1000-fold (Figure 3.13). This potent block to infectivity demonstrated that some sequences were not tolerated within this 'linker' region of p12. The disruption was likely due to the structure of this peptide compared to a string of alanines, as the sequence codes for a beta-strand from GFP (Kent, Childs, and Boxer, 2008).

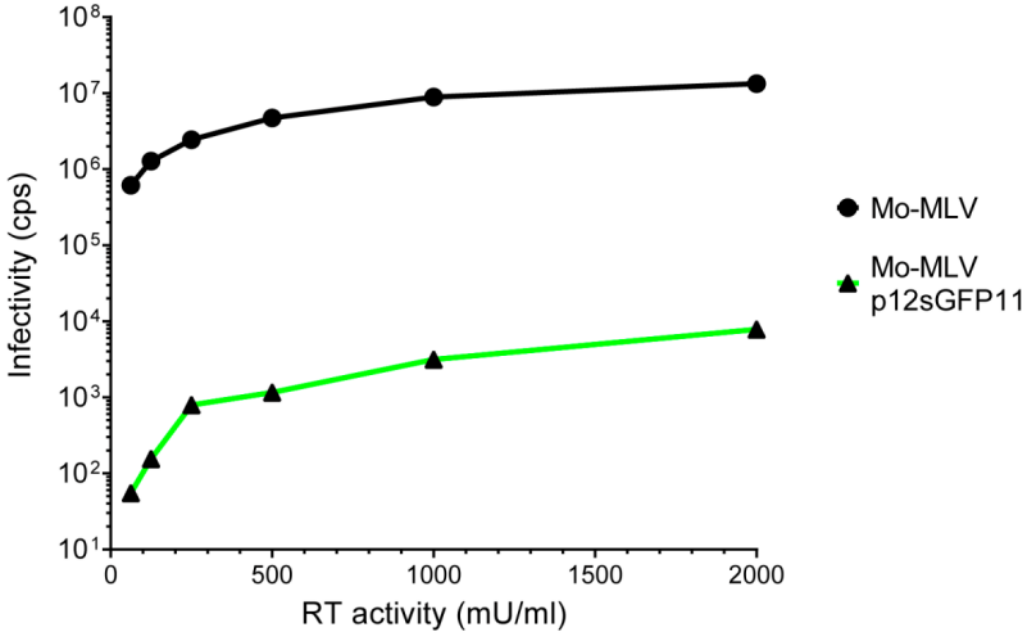


Figure 3.13: Infectivity of Mo-MLV containing the sGFP11 peptide in the central portion of p12. The 16 amino acid sGFP11 peptide (RDHMVLHEYVNAAGIT), which is part of the split GFP system (Cabantous, Terwilliger, and Waldo, 2005), was inserted into the central portion of p12 via the *Asc*I site. D17 cells were infected with a two-fold dilution series of the VLPs, based on RT activity, and productive infection was quantified after at least 48 hours. Infectivity is expressed as the luminescence counts per second (cps) for each dilution of VLPs. Data is representative of three independent experiments.

Therefore, although the 'linker region' of p12 was not sensitive to alanine mutagenesis and could be extended or cut in length, insertion of the sGFP11 has emphasised that some compositions of the linker were toxic to p12 function.

3.3 Identifying the essential residues within the early acting domains of p12

The two previous sections have built on previously published data on p12 function and have revealed that p12 contains two early acting domains that are functionally dependant. To date some isolated individual substitution mutants have been described, however not all of the amino acids in these domains have been assessed (Yueh and Goff, 2003).

3.3.1 The effect of individual alanine substitutions in p12 on Mo-MLV VLP release

To investigate which amino acids in the N- and C-terminal domains of p12 were essential for function, individual alanine mutations were generated in all the residues covered by the p12 mutations 5-8 and 13-15. All these p12 point mutant viruses produced similar levels of RT activity (data not shown), consistent with what was seen for the larger substitution mutants in p12. While p12 mutant 8 had a three-fold reduction in VLP production (Figure 3.2), this effect was not seen with any of the point mutants in this region of p12.

To assess the effect of the p12 mutations on Gag cleavage, a panel of p12 point mutant VLPs were concentrated and analysed by immunoblotting with a set of anti-MLV antibodies (Figure 3.14). None of the p12 point mutations in the N-terminal domain (Figure 3.14A) or the C-terminal domain of p12 (Figure 3.14B) affected the cleavage of Gag by the viral protease during maturation. The subtle changes in mobility of p12 (and MA-p12) reflected the changes in charge caused by the mutations. Thus, the point mutations in Mo-MLV p12 did not affect VLP release or Gag cleavage, indicating that these mutant viruses did not have major defects in the late stages of the viral life cycle.

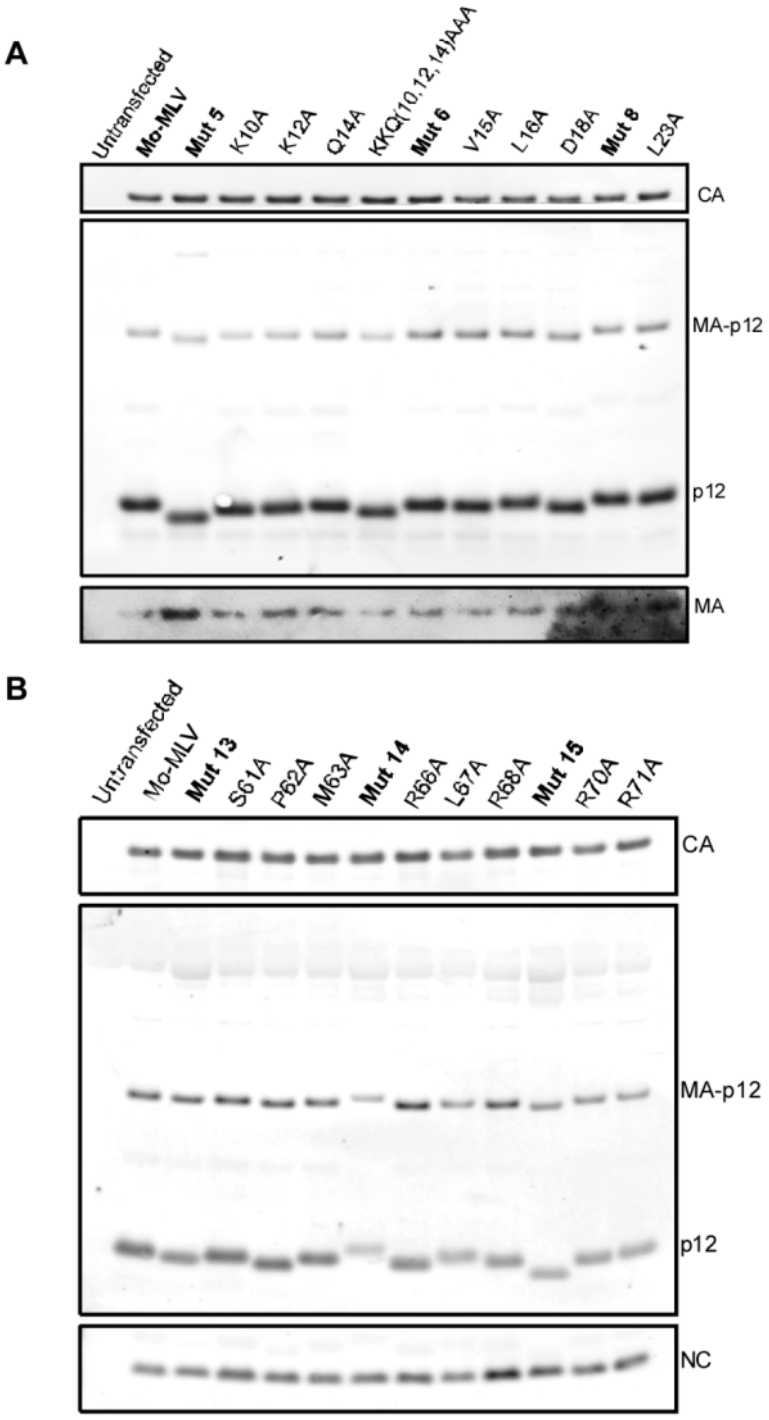


Figure 3.14: The effect of individual alanine mutations in p12 on Gag cleavage during virus maturation. A panel of Mo-MLV p12 mutant VLPs, with point mutations in the (A) N-terminus or (B) C-terminus of p12, were analysed by immunoblotting with anti-CA (A and B, top panel), anti-p12 polyclonal (A and B, middle panel), anti-MA (A, bottom panel) or anti-NC (B, bottom panel) antibodies.

3.3.2 Mapping the essential residues in the N-terminus of p12

The infectivity of Mo-MLV with point mutations in the N-terminal domain of p12 was assessed (Figure 3.15). Of the 21 amino acids mutated to alanines, only six reduced the infectivity by greater than 10-fold. These were V15 (19-fold), L16 (73-fold), L23 (15-fold), D25 (114-fold), L26 (14-fold) and L27 (36-fold). Nine mutations had negligible effect on the infectivity of Mo-MLV and the remaining seven mutations reduced the infectivity by 2-to-10-fold.

The fact that many of the mutations had minor effects on infectivity, suggested that the residues comprising the N-terminal domain of p12 contribute as a whole to the function of the domain. This was backed up by the mutations: K10A, K12A or Q14A which each reduced the infectivity by 2-to-10-fold, but when all three were mutated together the infectivity was reduced by greater than 100-fold compared to wild type (similar to the infectivity defect of p12 mutant 5) (Figure 3.15).

To investigate further the amino acid requirements in the N-terminal domain of p12, three amino acids in the mutant 6 region of p12 were chosen for further mutation to residues other than alanine. L16, S17 and D18 when mutated to alanine had a large, negligible and minor effect on infectivity, respectively. Therefore these three amino acids were mutated to a panel of residues with differing biochemical properties (F, K, L/V, Q and D/E) (Figure 3.16). None of the S17 changes had any effect on infectivity, indicating that this amino acid is not essential for Mo-MLV infectivity. When L16 was changed to valine or aspartic acid the infectivity defect was the same as seen for the alanine mutant (Figure 3.16). Interestingly, a change to phenylalanine, lysine or glutamine had less of an effect on infectivity (4-fold, 15-fold and 7-fold reduction compared to wild type, respectively).

The changes at position D18 fell into two groups: (i) mutations that had a greater effect on infectivity than the D18A change (F= 83-fold, K= 233-fold, and L= 286-fold reductions) and (ii) mutations that could function in the place of D18 (Q and E). Taken together this data suggests that some of the residues in the N-terminal domain (like S17), could accommodate a range of amino acids in their place with no detrimental effect on infectivity, whereas more critical residues (like L16 and D18) have some fundamental properties that could, to some degree, be mimicked by other amino acids.

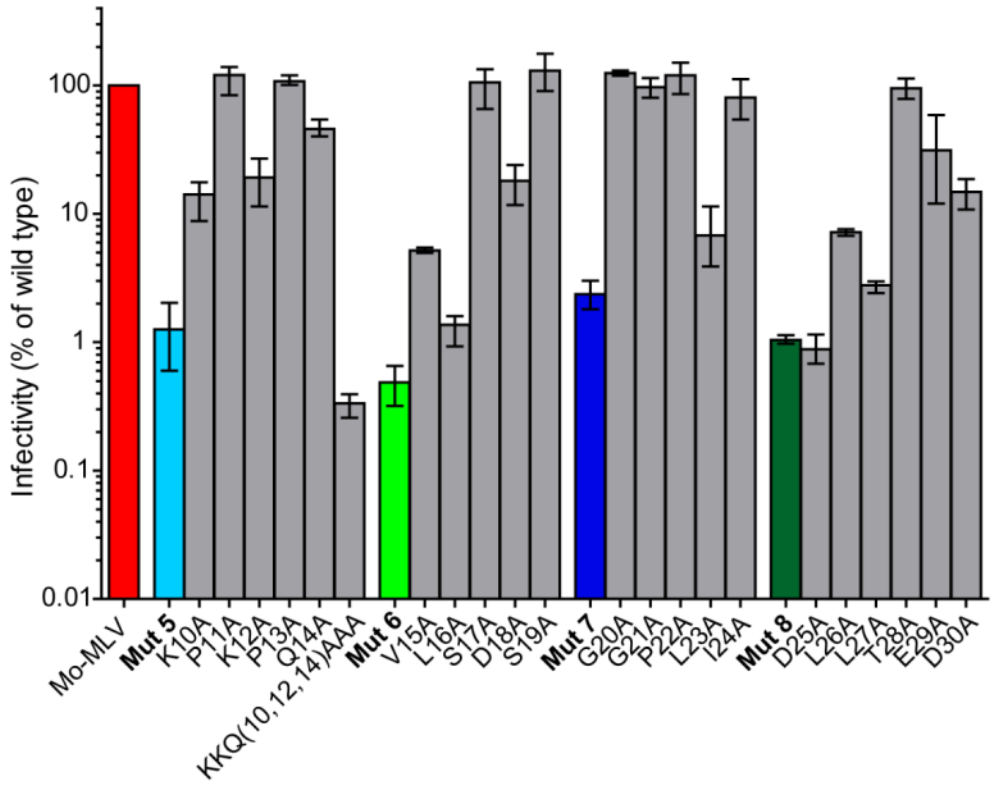


Figure 3.15: Infectivity of Mo-MLV containing individual alanine mutations in the N-terminal domain of p12. Individual alanine mutations in the N-terminal domain of p12 were generated and D17 cells were infected with equal RT-units of mutant VLPs. Productive infection was quantified after at least 48 hours. The infectivity is displayed as a percentage of wild type with the mean and range of three independent experiments shown.

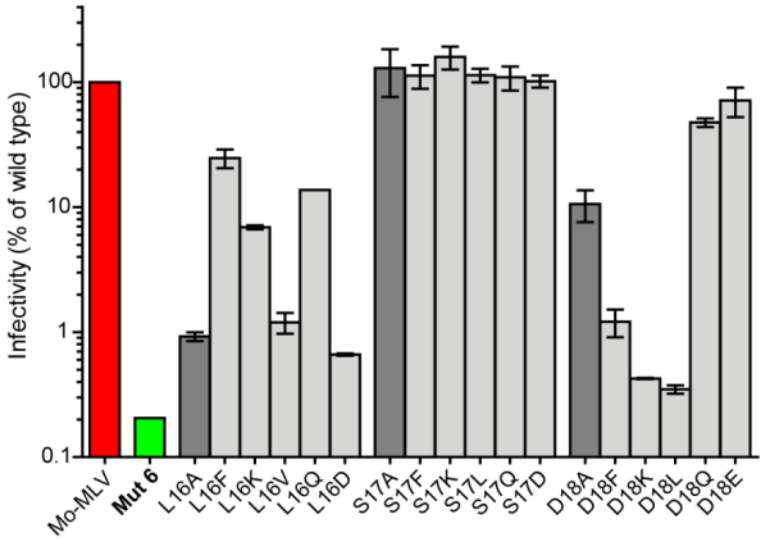


Figure 3.16: Infectivity of Mo-MLV containing expanded individual mutations in the mutant 6 region of p12. Individual mutations in the mutant 6 region of p12 were generated and D17 cells were infected with equal RT-units of mutant VLPs. Productive infection was quantified after at least 48 hours. The infectivity is displayed as a percentage of wild type with the mean and range of three independent experiments shown.

3.3.3 Investigating the role of p12 in Mo-MLV reverse transcription

Mo-MLV p12 mutant 6 was the only p12 mutant VLP that had a reduced ability to reverse transcribe in infected D17 cells (Figure 3.5). Three of the point mutants residing in the region covered by p12 mutant 6 had an effect on the infectivity of Mo-MLV. Therefore it seemed logical to investigate if these point mutations affected reverse transcription in infected cells. D17 cells were infected with equal RT-units of VLPs and the total DNA isolated at various times post-infection. Strong stop (Figure 3.17A) and second strand extension (Figure 3.17B) reverse transcription products were detected by qPCR.

Mutant 6 failed to accumulate significant levels of viral DNA products, as seen earlier (Figure 3.17). The three p12 point mutants V15A, L16A and D18A were all able to reverse transcribe in D17 cells with similar kinetics to wild type. The amounts of viral DNA synthesised by the V15A and L16A mutants were slightly reduced when compared to wild type across the time-course, but they were almost 10-fold higher than p12 mutant 6 levels (Figure 3.17). The level of viral DNA that was observed at the two week time point was significantly reduced for all the mutants (on a par with the infectivity defect of the mutants). In conclusion, the reduced ability to reverse transcribe only occurred when the larger mutant 6 region of p12 was altered, suggesting that the larger mutation increased the severity of the mutant phenotype.

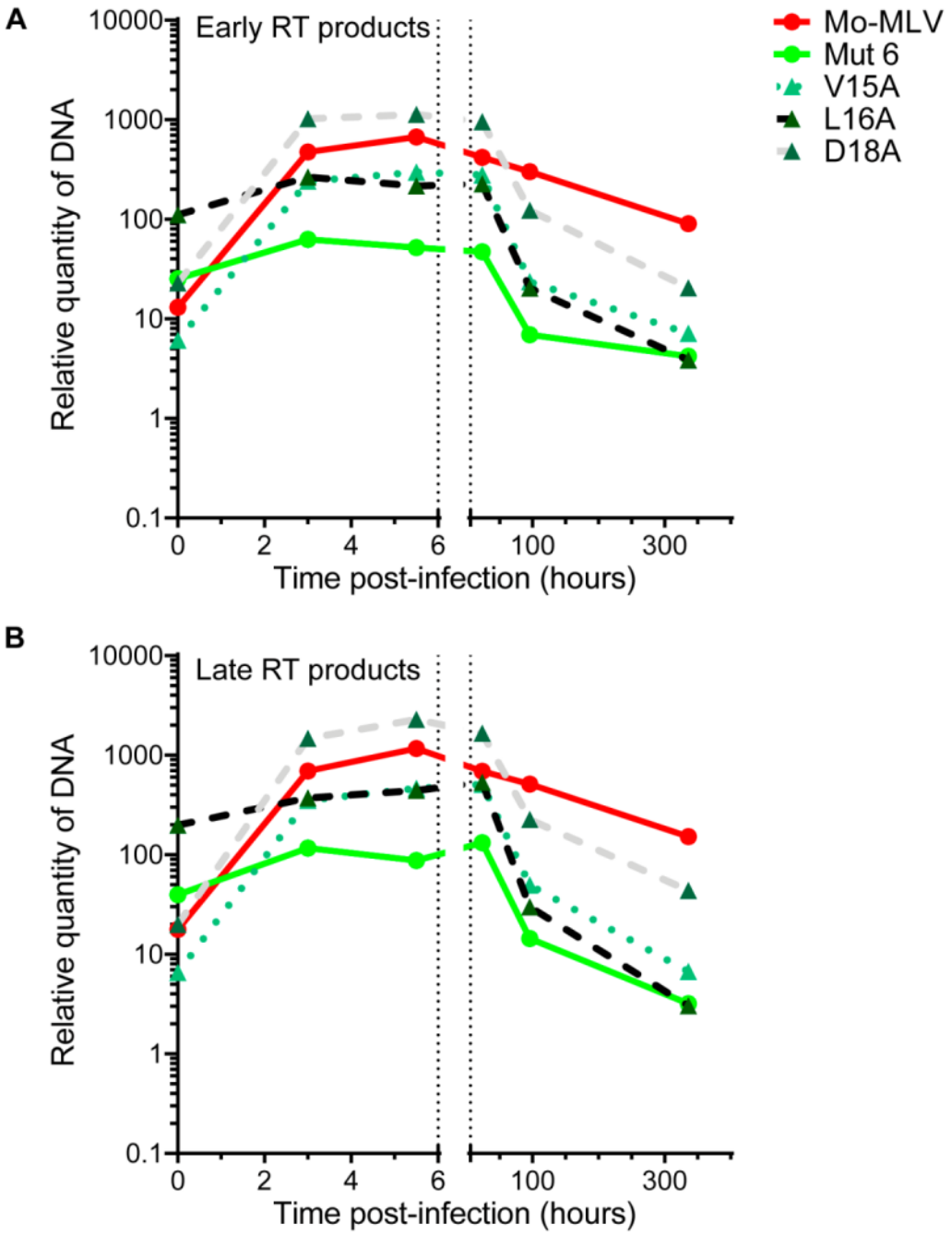


Figure 3.17: Ability of Mo-MLV VLPs containing individual alanine mutations in the mutant 6 region of p12 to synthesise viral cDNA in infected cells. D17 cells were infected with equal RT-units of wild type or mutant VLPs and total DNA was isolated at various times post infection, as indicated. The relative amounts of (A) minus strand strong stop or (B) second strand extension reverse transcription products were measured by qPCR. Results are representative of multiple independent experiments.

3.3.4 Mapping the essential residues in the C-terminus of p12

To complete the finer mutagenesis of Mo-MLV p12, all of the amino acids comprising the C-terminal domain of p12 were mutated individually to alanine. *LacZ*-encoding p12

point mutant Mo-MLV VLPs were synthesised and their infectivity quantified (Figure 3.18).

Of the 15 amino acids in the C-terminal domain (14 mutated as residue 64 is an alanine), six had no effect on the infectivity of Mo-MLV when mutated to alanine. Eight had an effect on the infectivity of Mo-MLV, and remarkably all inhibited the infectivity by greater than 100-fold, when compared to wild type infectivity (Figure 3.18). The least affected was the S61A mutant, although its infectivity was still reduced by 115-fold. In stark contrast to what was seen for the N-terminal domain, none of the mutations produced an in-between effect on infectivity. This indicated that a motif of eight essential amino acids were required for the function of the p12 C-terminal domain.

Of the amino acids mentioned above, mutation of the four arginines (R66, R68, R70 and R71) along with serine 61 have previously been shown to block replication of Mo-MLV at an early stage in the life cycle (Yueh and Goff, 2003). Attempts to rescue the block to replication caused by the S61A mutant included mutation to aspartic acid, to mimic the negative charge of phosphoserine (Yueh and Goff, 2003). This was unsuccessful, but could a similar phosphorylatable amino acid, threonine, function in the place of serine? A p12 S61T mutant was synthesised and the VLPs were also found to be non-infectious (Figure 3.19). This indicated that a phosphorylatable residue could not necessarily function in place of the serine at position 61, although the level of p12 phosphorylation was not quantified.

Revertant viruses from an Mo-MLV p12 S(61,65)A mutant were found to contain extra positive charges (mainly arginines), leading to the discovery that the four arginines in the C-terminus of p12 were essential for Mo-MLV infectivity (Yueh and Goff, 2003). These arginines in p12 could not be replaced by another positively charged amino acid, lysine (Yueh and Goff, 2003). To test if the other positively charged amino acid, histidine, could replace the arginines in the C-terminus of p12, R66 and R70 were mutated to histidine. In line with the lysine changes, histidine could not replace the function of the arginine residues (Figure 3.19). Therefore the charge alone was not the essential feature of these residues for the p12 C-terminal domain function. It is possible that these eight essential residues form a motif recognised by a cellular factor, and hence alterations to the motif affect this essential interaction.

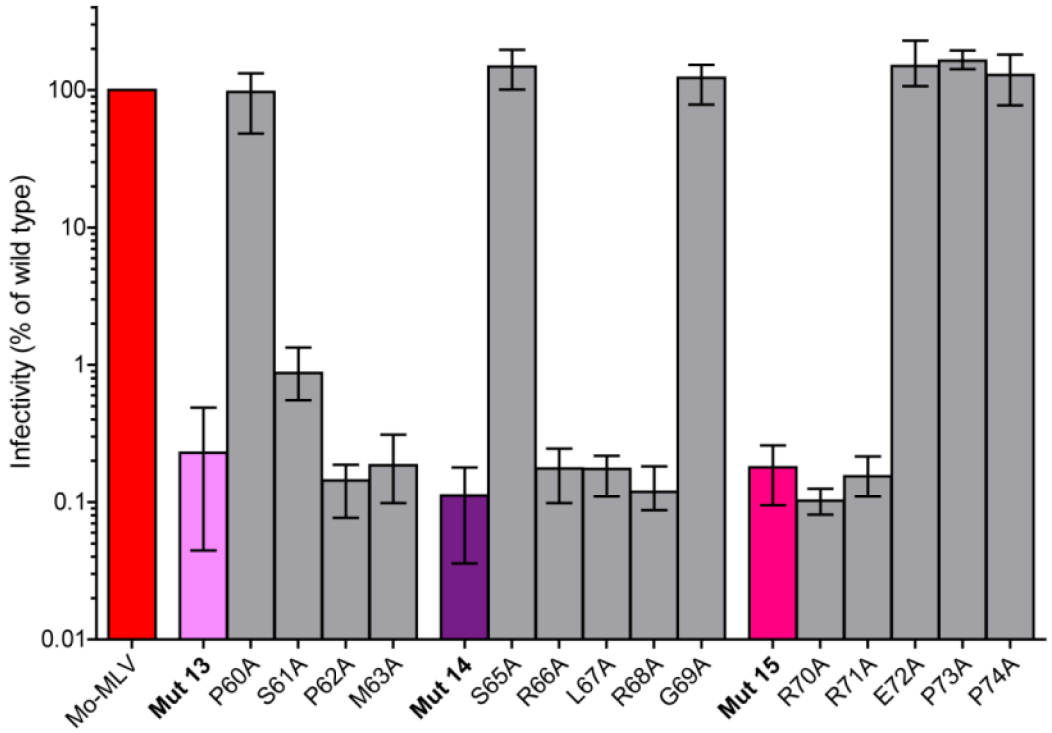


Figure 3.18: Infectivity of Mo-MLV containing individual alanine mutations in the C-terminal domain of p12. Individual alanine mutations in the C-terminal domain of p12 were generated and D17 cells were infected with equal RT-units of VLPs. Productive infection was quantified after at least 48 hours. The infectivity is displayed as a percentage of wild type with the mean and range of three independent experiments shown.

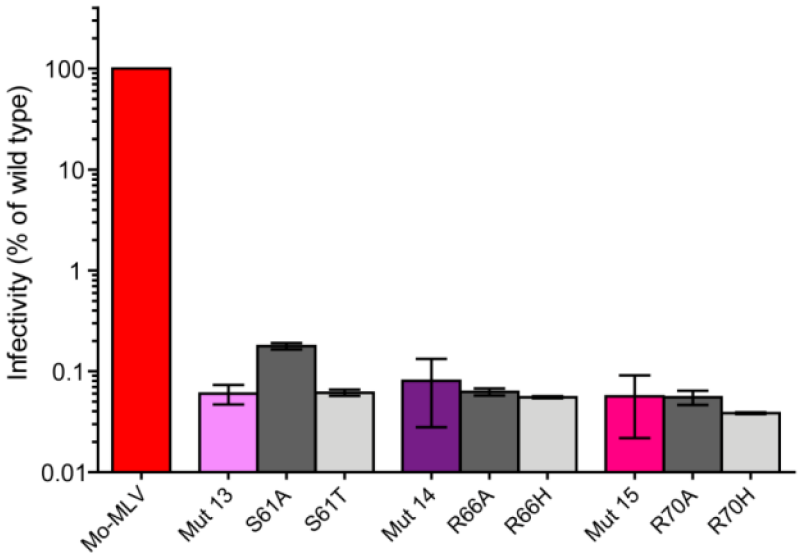


Figure 3.19: Infectivity of Mo-MLV containing extended individual mutations in the C-terminal domain of p12. Individual mutations in the C-terminal domain of p12 were generated and D17 cells were infected with equal RT-units of VLPs. Productive infection was quantified after at least 48 hours. The infectivity is displayed as a percentage of wild type with the mean and range of three independent experiments shown.

3.4 Summary

The experiments detailed here have shown that the p12 mutant Mo-MLV VLPs, synthesised from the Mo-MLV Gag-Pol vector system, have phenotypes reminiscent of those derived from the proviral plasmids (Yuan, Li, and Goff, 1999). The p12 mutants 5-8 and 13-15 were all blocked before integration of the viral DNA into the host chromosomes. Interestingly, the timing of the block, with regard to the synthesis of viral DNA in infected cells, varied between the experiments presented here and in the literature (Yuan, Li, and Goff, 1999). While the only p12 mutant VLP that failed to accumulate viral cDNA in infected cells was mutant 6, a previous study on the p12 mutants concluded that in addition to mutant 6, mutant 8 was unable to synthesise viral cDNA (Yuan, Li, and Goff, 1999). The viral cDNA in these former experiments was quantified from a different cell line, NIH-3T3, to the D17 cells used here. Unfortunately, a high background in the qPCR precluded the analysis of these cells infected with our p12 mutant VLPs (Virginie Boucherit, personal communication). The fact that p12 mutant 8 could reverse transcribe in D17 cells and not in N3T3 cells, indicates that reverse transcription occurs differently in different cells, which has also been reported for HIV-1 (Arfi et al., 2009).

Most importantly, the analysis of Mo-MLV mixed particles containing mixtures of wild type p12 and mutant p12, or two mutants, showed that the N- and C-terminal regions of p12 are actually two separate functional domains. Analysis of purified GST-tagged p12 by analytical ultracentrifugation (AUC) and multi-angle laser light scattering (MALLS) has shown that p12 is monomeric over a wide range of concentrations (Wight et al., 2012). Thus the dominant negative effect observed in mixed particles, containing wild type p12 and a C-terminal p12 mutant, was unlikely to be caused by inhibition of p12 multimerisation (with the caveat that the concentrations of p12 tested are probably below the levels present inside a virion). As Mo-MLV mixed particles containing a mixture of N- and C-terminal p12 mutants were non-infectious, the N- and C-terminal domains do not function independently of each other. For infectious particles, both domains must be unaltered on the same molecule of p12. Simultaneous mutation of both domains of p12 has shown that the N-terminal domain is required for the dominant negative effect of the C-terminal p12 mutants in mixed particles. Thus the N-terminus of p12 must function first and/or interact with a limiting factor.

The inhibition of MLV infectivity, in mixed particles composed of a C-terminal p12 mutant and wild type p12, occurred when very little C-terminal mutant p12 was in the

virus, indicating that the C-terminal mutant p12 molecule was more efficient than wild type p12 at carrying out the N-terminal domain interaction/activity. This suggests that the C-terminal domain of p12 has an inhibitory effect on the N-terminal domain of p12, although the significance of this inhibition for the virus is currently unknown.

Mutation of the 'linker' region, between the N- and C-terminal domains of p12, has revealed that it can be significantly shortened or increased in size without an impact on p12 function. Surprisingly, inclusion of the sGFP11 sequence was shown to be toxic to p12 function. The reason why this sequence is detrimental to p12 function is unclear, but potentially, due to the structure of the sGFP11 peptide, it may indicate that the central 'linker' region of p12 needs to remain flexible for p12 function. However other possibilities exist, for example the addition of extra charged residues into the 'linker' region may be detrimental p12 function.

A closer examination of the amino acids that were essential for the domain functions revealed a stark contrast between the N- and C-terminal domains. The C-terminal domain contained eight amino acids all of which are essential for infectivity, whereas the N-terminal domain contained a few essential amino acids (notably leucines and aspartic acids) surrounded by others that contributed to the optimal function of the domain (i.e. mutating them had a smaller but notable effect on infectivity). The C-terminal domain contains a motif of essential amino acids based around arginine residues, which could not be changed to residues with similar biochemical properties.

The N-terminal domain is more complicated: the larger substitution mutants, composed of five or six alanines, had a greater effect on the infectivity of Mo-MLV than many of the point mutants. It is entirely possible that the residues in the N-terminal domain may contribute to an optimal conformation and/or spacing within the domain. Therefore, the residues that reduce the infectivity of Mo-MLV by 2-to-10 fold when mutated to alanine may promote optimal conditions for domain function. Considering that alanine and valine, two small non-polar amino acids similar to leucine, could not function in place of the essential L16, showed that there is something special about this leucine residue. Interestingly, the larger non-polar amino acid phenylalanine could function (to some degree) in place of L16. One similarity between these two amino acids is the length of their side chains and the other two amino acids that could partly function in place of L16 also have longer side chains (although not the same as leucine) with one being polar (glutamine) and the other positively charged (lysine). Recently, the interaction of

TRIM5alpha with the CA core has been described as a poly-valent interaction, where multiple weak interactions combine to make a strong binding to the core surface (Ohkura et al., 2011). Could a similar principle be true for the N-terminal domain of p12? Identification of an interacting partner of the N-terminal domain of p12 would certainly help to answer this question.

Chapter 4.

The conservation of p12 function within the gammaretrovirus family

Almost all retroviruses contain some sequence between MA and CA in their *gag* gene, although it does not always code for a functional protein. In the case of HIV-1, a small (~16 amino acid) spacer peptide exists between MA and CA of no known function. Retroviruses from other genera, like the alpha-retrovirus RSV code for three peptides, p2a, p2b and p10, which serve functions in particle budding and morphology of the CA core (Scheifele et al., 2007; Spidel et al., 2004). All gammaretroviruses contain a 'p12-like' protein coded between MA and CA, which carries the 'PPPY' late-domain motif.

Experiments in the previous chapter have established that the p12 protein in Mo-MLV contains two functional domains that are essential for the early stages of infection. The requirement for p12 during this part of the viral life cycle in other gammaretroviruses has not been established. Therefore, this chapter describes investigations into the conservation of p12 function in gammaretroviruses. Importantly, the requirement for p12 function, during the early stages of infection, was shown to be conserved in the gammaretroviridae.

4.1 The effect of mutations in p12 on a panel of gammaretroviruses

A panel of gammaretroviral Gag-Pol vectors were available for this study: N-tropic MLV (N-MLV), B-tropic MLV (B-MLV), xenotropic MLV-related virus, feline leukaemia virus (FeLV) and gibbon ape leukaemia virus (GaLV). All the synthesised gammaretroviral VLPs contain the MLV LTR-driven *LacZ* reporter genome and were psuedotyped with VSV-G. Both, the FeLV and GaLV, have previously been assessed to ensure they can package the reporter genome from MLV (Melvyn Yap, personal communication). Therefore, the effect of p12 mutations in this panel of gammaretroviral VLPs could be investigated.

4.1.1 The conservation of p12 sequence within the gammaretroviral genera

Before generating mutations, the p12 sequences from the panel of gammaretroviruses were aligned using Clustal X2 (Larkin et al., 2007) (Figure 4.1). Manual fine-tuning of the outputs were made to ensure the most accurate alignments were constructed.

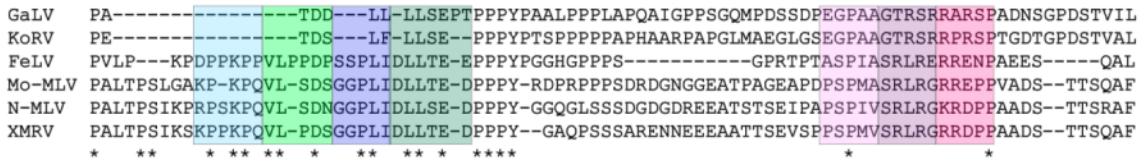


Figure 4.1: Alignment of Mo-MLV p12 with the p12 sequences from other gammaretroviruses. The Mo-MLV p12 amino acid sequence (GenBank: J02255.1) was aligned with other retroviruses using Clustal X2 (Larkin M. A. et al. 2007). (A) Alignment of Mo-MLV p12, with p12 from: GaLV (M26927.1), KoRV ([Koala retrovirus] AF151794.2), FeLV (M18247.1), N-MLV (K01203.1) and XMRV (EF185282.1). Positions of the p12 mutations 5-8 and 13-15 are shown by the coloured boxes.

The most similar p12 sequences to Mo-MLV p12 were the other MLV sequences. N-MLV (and B-MLV, the p12 sequence is identical) and XMRV p12 had 64.7% and 71.4% sequence identity when compared to Mo-MLV p12, respectively. FeLV p12 contained an 11 amino acid deletion in the centre of p12 and an overall 55.7% identity with Mo-MLV p12. The two most distantly related gammaretroviruses to Mo-MLV, GaLV and KoRV, had large deletions in the N-terminus of p12 and a low sequence identity with Mo-MLV p12 of 25.7% and 32.4%, respectively. The N-terminal domain of p12 had a high degree of similarity between the gammaretroviruses (barring GaLV and KoRV) and contained many leucines and aspartic acids (Figure 4.1A, green and blue boxes). The residues in the C-terminal domain of p12 appeared less well conserved (Figure 4.1A, pink and purple boxes). However, some arginines, which were shown to be essential for Mo-MLV infectivity, were preserved, although their position within the C-terminal domain varied.

4.1.2 VLP release from a panel of gammaretroviruses carrying mutations in p12

Similar p12 mutations were introduced into a panel of gammaretroviruses based on the sequence alignment in Figure 4.1A (see coloured boxes). As GaLV has a small N-terminal region, only two mutants were made in this region, mutant 8 and one called mutant N (TDDLL/AAAAA). *LacZ*-encoding gammaretroviral VLPs were synthesised

and VLP production was assessed by the level of RT activity. None of the gammaretroviral p12 mutants had a major defect in VLP production (Figure 4.2). Mo-MLV mutant 8 had a minor (three-fold) reduction in VLP release, which was also seen for N-MLV and B-MLV (~five-fold reduction, Figure 4.2 B and C, dark green bars), although not for FeLV and GaLV. Unfortunately, despite our best efforts, XMRV p12 mutant 8 could not be cloned (Virginie Boucherit, personal communication).

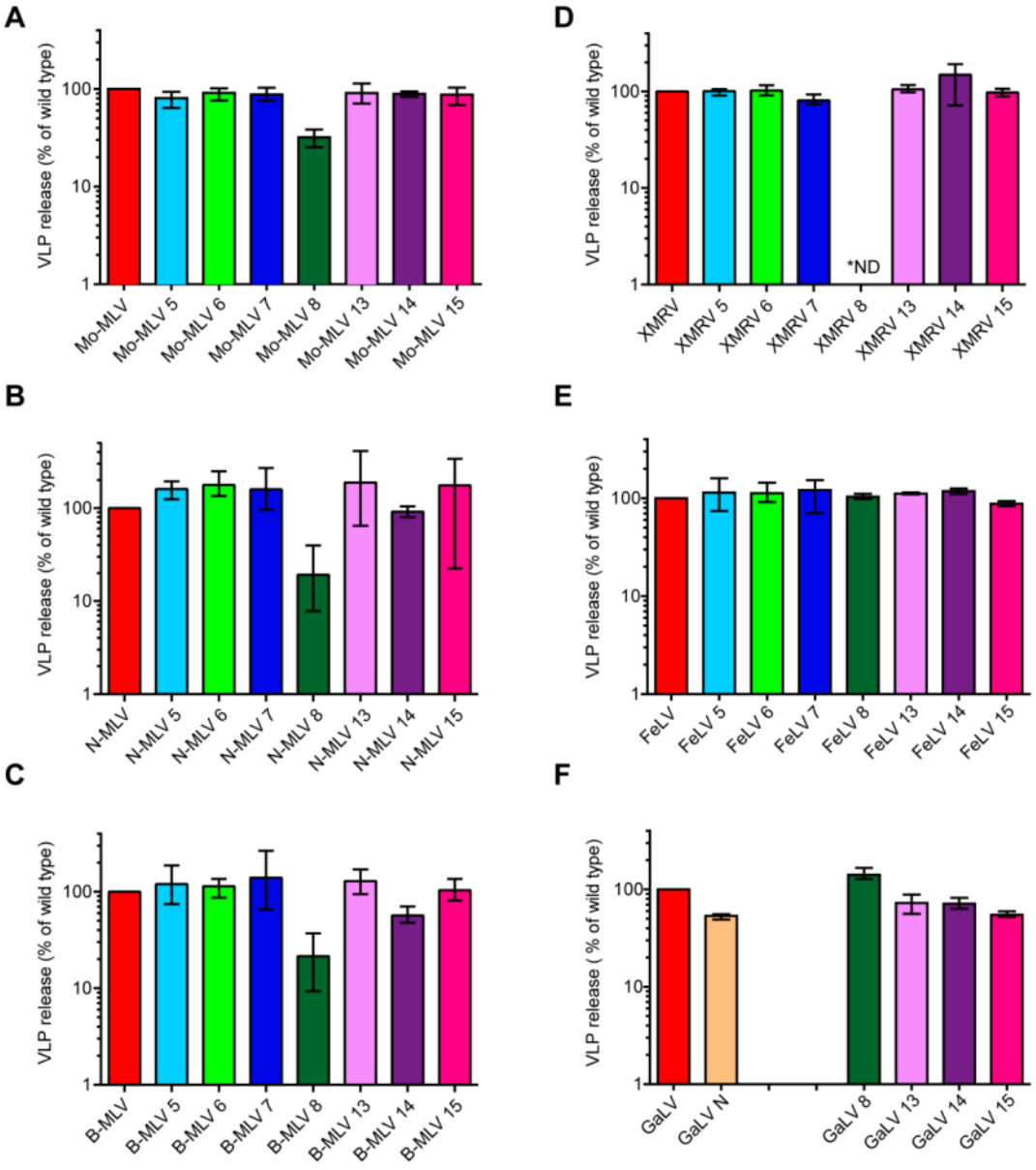


Figure 4.2: The effect of p12 mutations on VLP release for a panel of gammaretroviruses. The amount of (A) Mo-MLV, (B) N-MLV, (C) B-MLV, (D) XMRV, (E) FeLV and (F) GaLV VLPs released was quantified by measuring the RT activity using the RT-ELISA. The graphs show the percentage of p12 mutant VLPs released relative to wild type for each virus (based on the determined RT activity). The bars represent the mean and the range of three independent experiments. *ND- not determined.

4.1.3 The effect of p12 mutations on the infectivity of a panel of gammaretroviruses

To assess the effect of the p12 mutations on the ability of the gammaretroviruses to progress through the early stages of infection, the infectivity of our panel of p12 mutants were tested in D17 cells. All the gammaretroviruses tested were sensitive to alterations to the C-terminal region of p12 and these mutant viruses were non-infectious (Figure 4.3). The N-terminal domain was different in this regard: The p12 mutants 6 or 8 were non-infectious in all gammaretroviral backgrounds, but both N-MLV and B-MLV carrying p12 mutant 5 and 7 were much more infectious than Mo-MLV carrying these p12 mutations (Figure 4.3 A versus B and C, light and dark blue bars). The infectivity of these mutants was only reduced by 3-to-8.5-fold compared to wild type. This is interesting for p12 mutant 7 as the primary amino acid sequence in the p12 mutant 7 region is the same between these three MLVs (Figure 4.3A).

FeLV p12 mutant 5 was also infectious compared with Mo-MLV mutant 5, having only ~2-fold reduction in infectivity when compared to wild type FeLV (Figure 4.3E). Both the XMRV p12 mutants 5 and 7, while greater than 10-fold more infectious than XMRV p12 mutant 6 (Figure 4.3D blue bars versus light green bar), were still relatively non-infectious. Finally, mutation of the short N-terminal domain of GaLV p12 (mutant N) also rendered the virus non-infectious (Figure 4.3F, peach bar). Therefore, all of the gammaretroviruses analysed required p12 function during the early stages of infection, although some portions of the N-terminal region were less sensitive to mutation in certain gammaretroviruses.

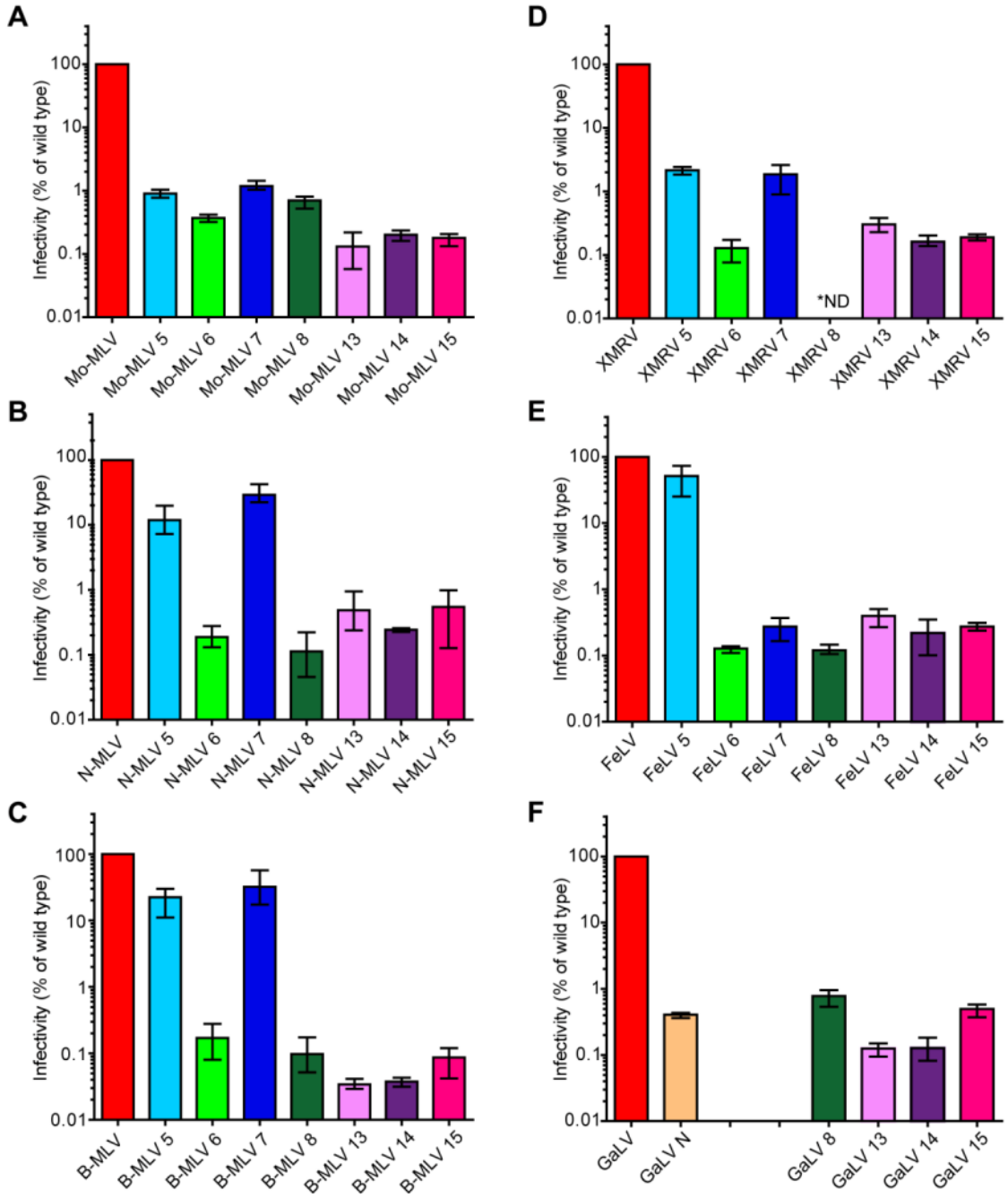


Figure 4.3: The infectivity of a panel of gammaretroviruses carrying mutations in p12. Equal RT-units of: (A) Mo-MLV, (B) N-MLV, (C) B-MLV, (D) XMRV, (E) FeLV and (F) GaLV VLPs were used to infect D17 cells. Productive infection was measured after at least 48 hours. The graphs show the mean infectivity, as a percentage of wild type, from three independent experiments with the error bars displaying the range. *ND- not determined.

4.2 Identifying the early acting domains in GaLV p12

From the panel of gammaretroviruses analysed, GaLV is the most distantly related to Mo-MLV. The p12 of GaLV has an interesting short N-terminal region, which, just like in Mo-MLV, was required during the early stages of replication (Figure 4.3F).

Therefore, to test the conservation of p12 function in the gammaretroviral genera further, a closer analysis of GaLV p12 function was undertaken.

4.2.1 The N- and C-terminal domains in GaLV p12

To assess if GaLV p12 also contained two functional domains, *LacZ*-encoding mixed GaLV particles containing wild type and mutant p12 were synthesised and the infectivity was quantified (Figure 4.4).

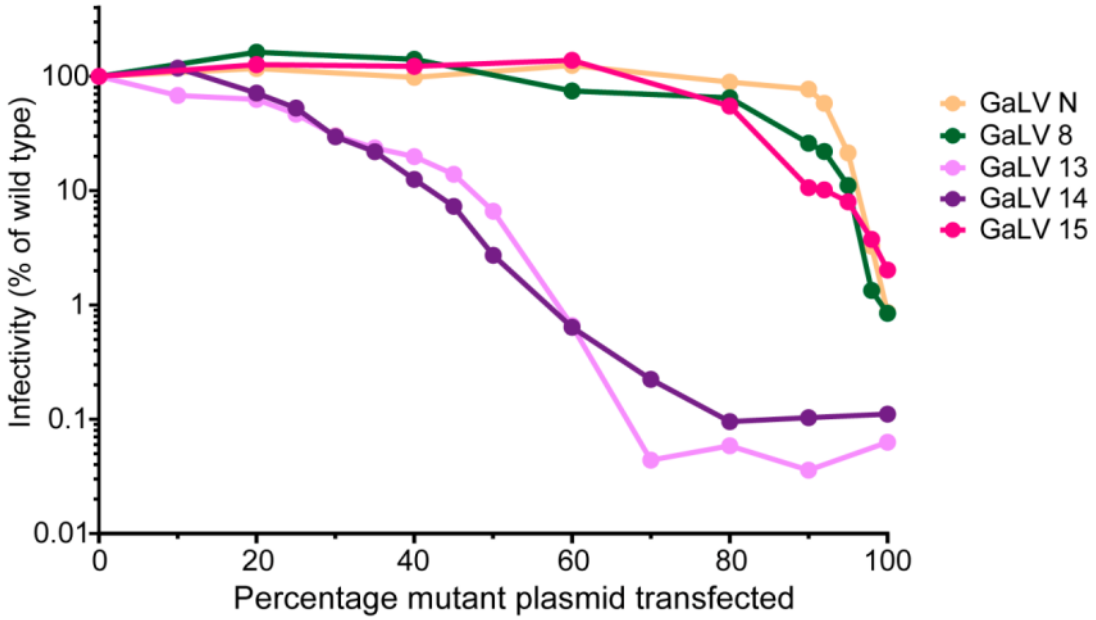


Figure 4.4: Infectivity of GaLV particles containing mixtures of wild type and mutant p12. D17 cells were infected with equal RT-units of mixed particles and productive infection was measured after at least 48 hours. Infectivity is displayed as a percentage of wild type GaLV infectivity and the data is representative of at least two independent experiments.

The GaLV N-terminal p12 mutant mixed particles displayed a phenotype similar to the N-terminal p12 mutants in Mo-MLV mixed particles (Figure 3.8). Little wild type p12 (~10-20%) was required in the mixed particles to preserve infectivity (Figure 4.4, dark green and peach lines). The C-terminal p12 mutants 13 and 14 (Figure 4.4, pink and purple lines) displayed the dominant negative effect seen with Mo-MLV C-terminal p12 mutants mixed particles, although the magnitude of the effect was reduced (the infectivity did not decrease sharply until 50% of p12 in the particle was mutant) (Figure 4.4, pink and purple lines). Furthermore, GaLV p12 mutant 15 showed no dominant negative effect and its phenotype resembled that of the N-terminal p12 mutant mixed particles (Figure 4.4, magenta line). Thus, the GaLV mixed particles showed that GaLV also likely contained two domains in p12, but there were differences. (i) The dominant

negative effect of the C-terminal mutants was restricted to mutant 13 and 14 and (ii) the strength of the dominant negative effect was less in GaLV compared to Mo-MLV.

To confirm the loss of the dominant negative effect, GaLV mixed particles were synthesised with wild type and a p12 point mutant in the mutant 15 region, R56A (that reduced the infectivity of GaLV by 71-fold, Figure 4.7B). As a control, mixed particles containing p12 mutant 14 and a point mutant in the mutant 14 region, R51A (Figure 4.7B), were synthesised (Figure 4.5). As expected, the phenotype of the point mutant, R51A, mirrored the dominant negative phenotype of the larger p12 mutant 14 (Figure 4.5, purple line and black dashed line with purple triangles), whereas the R56A point mutant mixed particles did not show a dominant negative effect, just like p12 mutant 15 (Figure 4.5, pink line and black dashed line with pink triangles). Therefore, mutation of this region of p12 disrupted p12 function, but additionally removed the dominant negative activity in GaLV mixed particles.

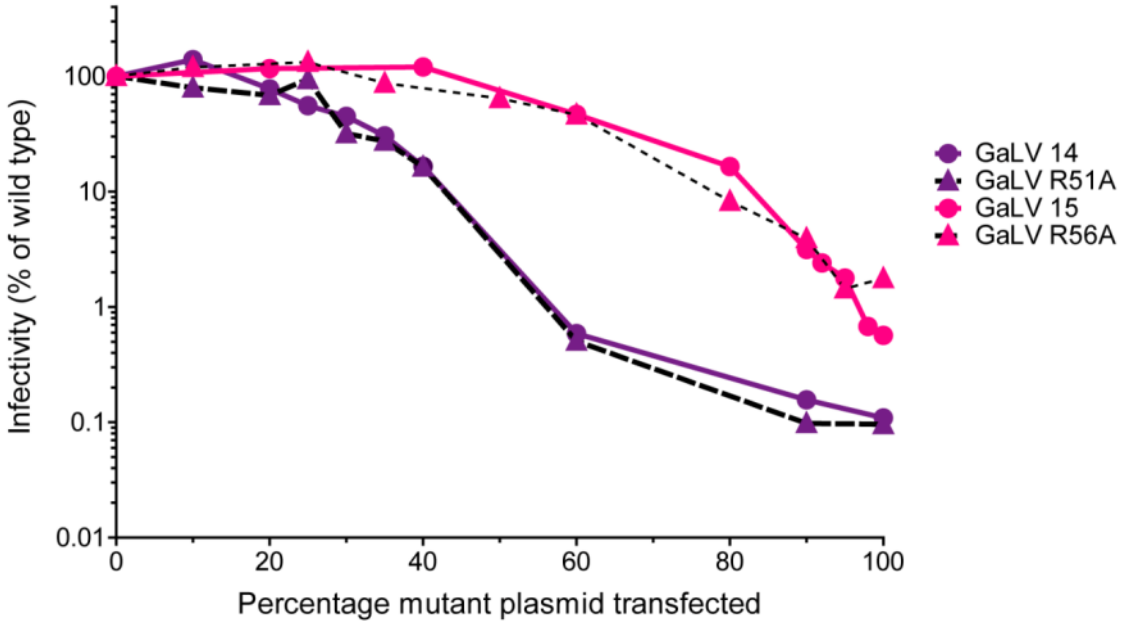


Figure 4.5: Infectivity of GaLV mixed particles containing mixtures of wild type and C-terminal p12 point mutants. D17 cells were infected with equal RT-units of mixed particles and productive infection was measured after at least 48. Infectivity is displayed as a percentage of wild type GaLV.

4.2.2 The essential residues in the N-terminal domain of GaLV p12

Individual point mutants in Mo-MLV p12 revealed the essential amino acids in the N- and C-terminal domains of p12. Thus, the same approach was taken with GaLV p12. Alanine substitutions were generated for the residues in the N-terminal domain of p12 (barring S8, E9, P10 and T11), *LacZ*-encoding VLPs were synthesised and their

infectivity was tested (Figure 4.6B). As a point of reference, the Mo-MLV N-terminal p12 sequence has been aligned with that of GaLV (Figure 4.6A).

Only one mutant displayed a major effect on infectivity and one had a minor effect (Figure 4.6B). Interestingly, these residues were leucine (600-fold reduction in infectivity) and aspartic acid (three-fold reduction). Therefore, as was seen with the N-terminal domain of Mo-MLV p12, leucine and aspartic acid residues were important for domain function. The larger p12 mutant N reduced the infectivity of GaLV by greater than 100-fold, but none of the point mutants in this region (mutant N) recapitulated this phenotype. This suggested that, as with Mo-MLV, the residues in the N-terminal domain of GaLV p12 were all contributing to an optimal conformation and/or spacing to the domain.

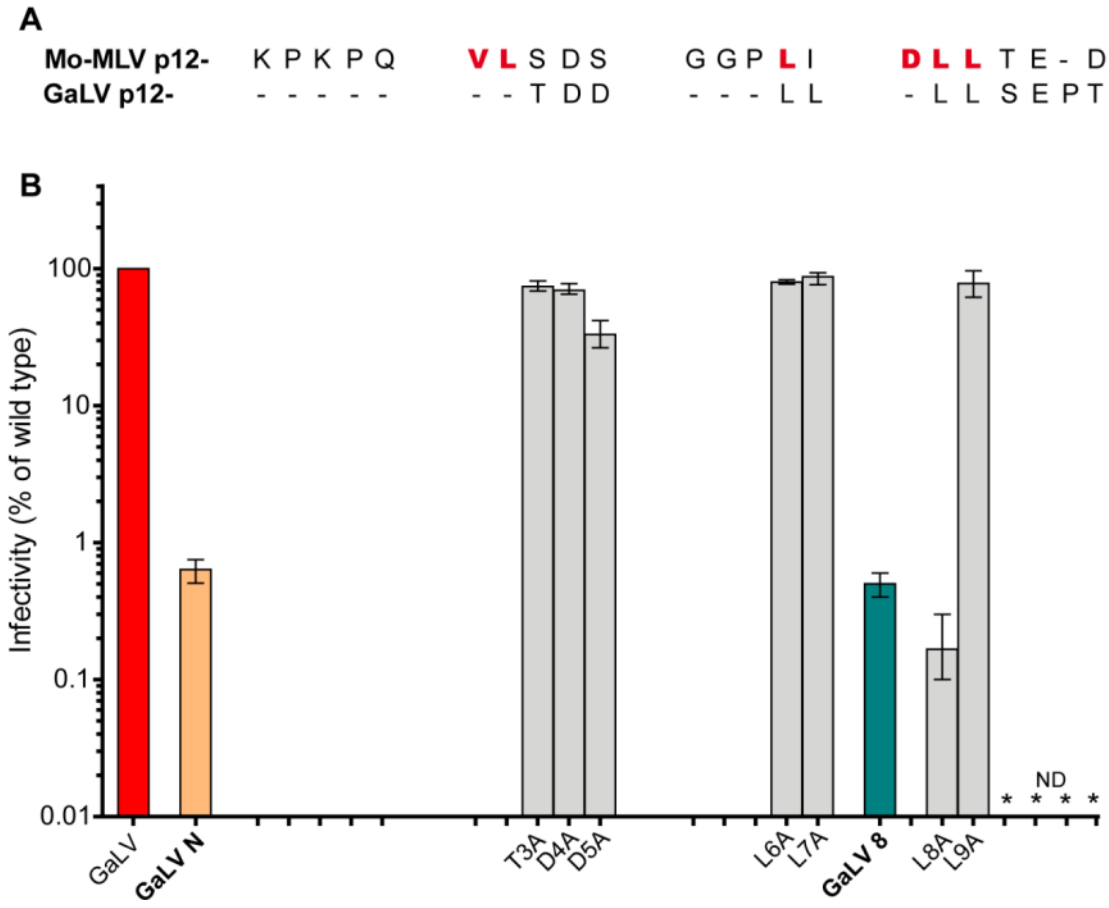


Figure 4.6: Infectivity of GaLV containing individual alanine mutations in the N-terminal domain of p12. (A) Primary amino acid sequence of the N-terminal domain of Mo-MLV p12 aligned with the corresponding amino acids of GaLV p12. Bold red amino acids indicate residues that when mutated to alanine, reduce the infectivity of Mo-MLV by greater than 10-fold. (B) D17 cells were infected with equal RT-units of VLPs and productive infection quantified after at least 48 hours. Infectivity is displayed as a percentage of wild type GaLV infectivity with the mean and range of three independent experiments shown. *ND- not determined.

4.2.3 The essential residues in the C-terminal domain of GaLV p12

To investigate which amino acids in the C-terminal domain of GaLV p12 were essential for function, individual alanine mutations were generated in GaLV p12 and the infectivity of the resulting VLPs tested (Figure 4.7). For comparison, the Mo-MLV p12 C-terminal domain sequence is aligned with that of GaLV p12 (Figure 4.7A).

Seven amino acid, when mutated to alanine, reduced the infectivity of GaLV by greater than 10-fold, namely: G45, P46, G49, T50, R51, R53 and R56 (Figure 4.7B). S52A had a minor effect on the infectivity of GaLV, reducing the infectivity by 7.3-fold (Figure 4.7B). Interestingly, unlike Mo-MLV, only three out of the four arginines in the C-terminus were essential for GaLV infectivity. Therefore, while the essential amino acids in the C-terminal domain of p12 were not conserved in position between Mo-MLV and GaLV, there were similarities.

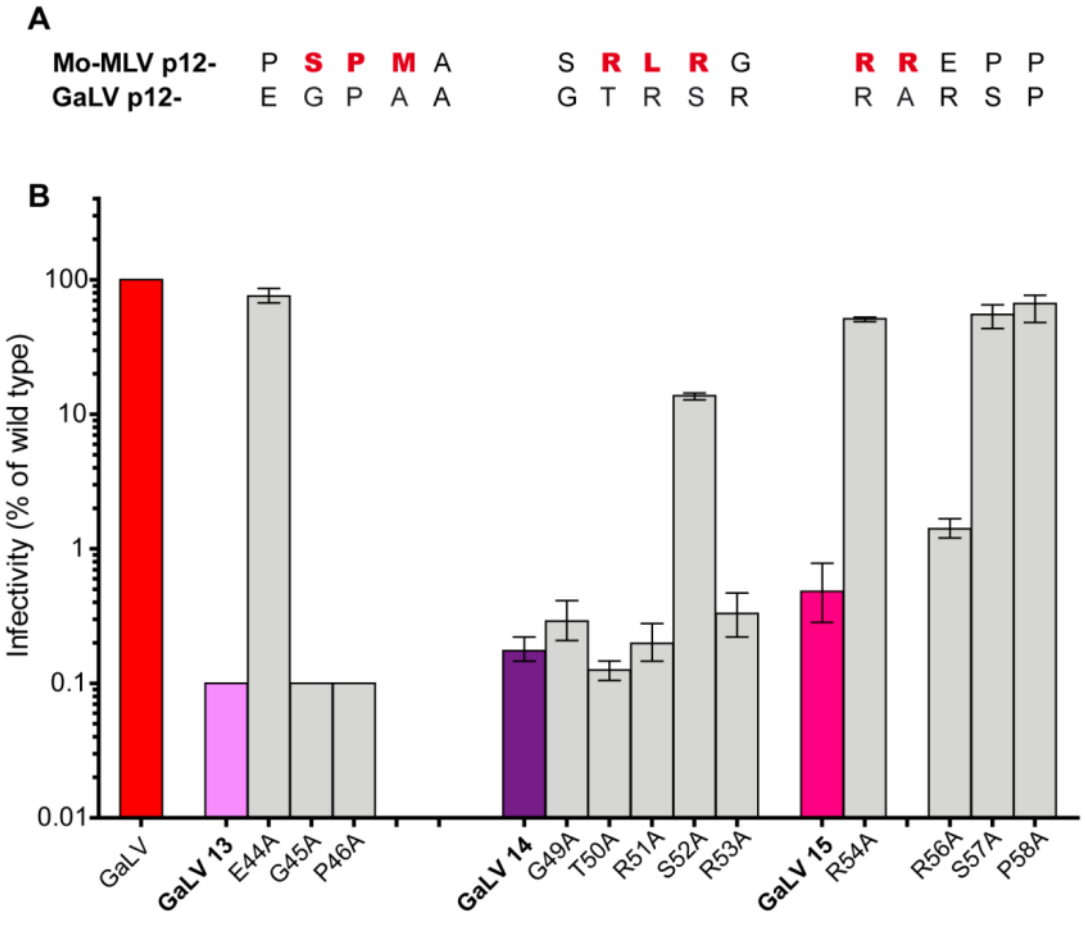


Figure 4.7: Infectivity of GaLV containing individual alanine mutations in the C-terminal domain of p12. (A) Primary amino acid sequence of the C-terminal domain of Mo-MLV p12 aligned with the corresponding amino acids of GaLV p12. Bold red amino acids indicate residues that when mutated to alanine reduce the infectivity of Mo-MLV by greater than 10-fold. (B) D17 cells were infected with equal RT-units of VLPs and the productive infection was quantified after at least 48 hours. Infectivity is displayed as a percentage of wild type GaLV infectivity with the mean and range of three independent experiments shown.

4.3 Investigating the conservation of the p12 domain function(s) between gammaretroviruses

The GaLV p12 protein is interesting in the fact that it contains a very short N-terminal domain when compared to MLVs or FeLV. The results from the p12 alanine mutagenesis highlighted similar essential amino acids between GaLV and Mo-MLV p12. Experiments were therefore designed to test if the p12 proteins could be swapped between the viruses.

4.3.1 The effect of GaLV p12 in Mo-MLV on VLP release

To test whether the GaLV p12 can function in the Mo-MLV background, a panel of chimeras were constructed, containing either the whole GaLV p12, or N- or C-terminal regions. (Figure 4.8A, Mo-MLV/Ga-p12, Mo-MLV/Ga-Np12 and Mo-MLV/Ga-Cp12, respectively). To restore the Mo-MLV MA-p12 protease cleavage site, a leucine was inserted between the A2 and T3 of the Mo-MLV/Ga-p12 and Mo-MLV/Ga-Np12 chimeras (Figure 4.8A, Mo-MLV/Ga-p12-CL and Mo-MLV/Ga-Np12-CL).

All the Mo-MLV chimeras released wild type amounts of VLPs (Figure 4.8B). Equal RT-units of chimeric VLPs were analysed by immunoblotting with antibodies against CA (Figure 4.8C, top left panel), MA (Figure 4.8C, top right panel) and two antibodies against p12 (monoclonal in Figure 1.8C bottom right panel and a polyclonal in Figure 4.8C, bottom left panel). All chimeric VLPs contained mature CA, but mature MA was not seen for the Mo-MLV/Ga-p12 and Mo-MLV/Ga-Np12 chimeras, indicating that the cleavage between MA-p12 had been disrupted (Figure 4.8C). When the Mo-MLV MA-p12 cleavage site was restored, mature MA was detected in the chimeras, indicating the processing defect had been reversed (Figure 4.8C). Using a combination of the p12 monoclonal and polyclonal antibodies (Figure 4.8C, lower panels), it could be seen that the Mo-MLV/Ga-Np12-CL and Mo-MLV/Ga-Cp12 contained mature p12. Unfortunately, the p12 antibodies did not cross react to GaLV p12, but as mature MA could be detected, it was assumed Mo-MLV/Ga-p12-CL must contain mature p12 (Figure 4.8C, top right panel). Therefore, the Gag polyprotein in the Mo-MLV chimeras was processed by the viral PR. However, cleavage between MA and p12 required the natural Mo-MLV p12 sequence (PALT).

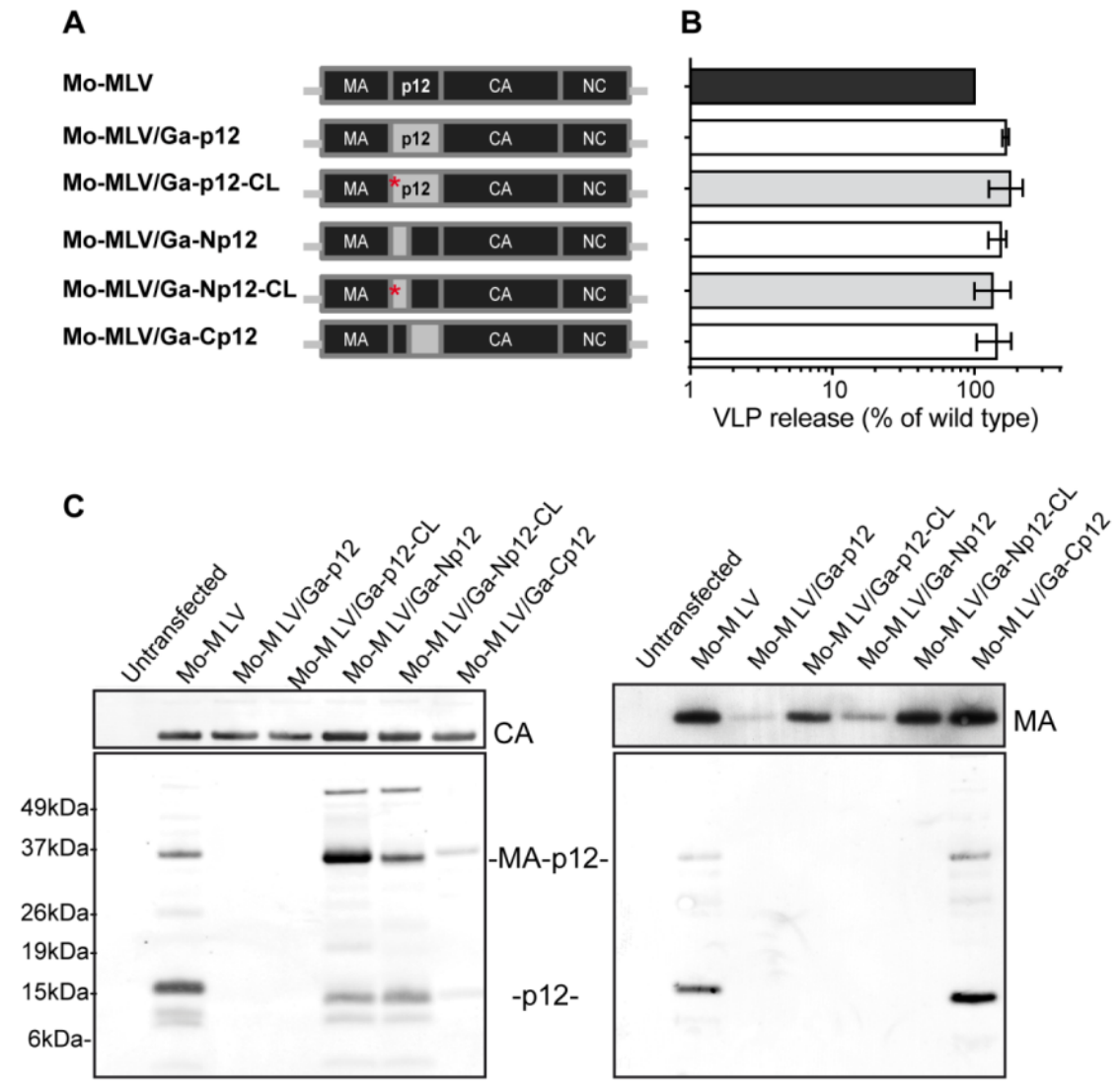


Figure 4.8: Mo-MLV chimeric VLPs containing GaLV p12. (A) Schematic detailing the regions of p12 swapped between Mo-MLV and GaLV (Asterisks- Mo-MLV/Ga-p12-CL and /Ga-Np12-CL contain a leucine inserted between the second alanine and third threonine of the GaLV p12, to restore the Mo-MLV protease cleavage site). (B) The level of VLPs released was measured by quantifying the RT activity by the RT-ELISA and is shown as the percentage of the wild type. The bars represent the mean and range of three independent experiments. (C) Equal RT-units of VLPs were analysed by immunoblotting with anti-CA (top left panel), anti-MA (top right panel), anti-p12 polyclonal (bottom left panel) and anti-p12 monoclonal (bottom right panel) antibodies.

4.3.2 The functionality of GaLV p12 domains in the Mo-MLV background

The ability of the Mo-MLV/Ga-p12 chimeras to pass through the early stages of the life cycle was assessed by quantifying their infectivity in D17 cells (Figure 4.9). Mo-MLV chimeras containing the N-terminal region of GaLV p12 were 833-fold less infectious

than wild type Mo-MLV (Figure 4.9, Mo-MLV/Ga-Np12). This reduction in infectivity remained even when the MA-p12 protease cleavage site in the chimera had been restored (Figure 4.9, Mo-MLV/Ga-Np12-CL). Unsurprisingly, the same reduction in infectivity was seen when the whole GaLV p12 was introduced into the Mo-MLV background (Figure 4.9, Mo-MLV/Ga-p12 and Mo-MLV/Ga-p12-CL). Interestingly, Mo-MLV containing the C-terminal region of GaLV p12 was infectious (Figure 4.9, Mo-MLV/Ga-Cp12). Therefore the C-terminal domain of GaLV p12, but not the small N-terminal domain of p12, was able to function in the Mo-MLV background.

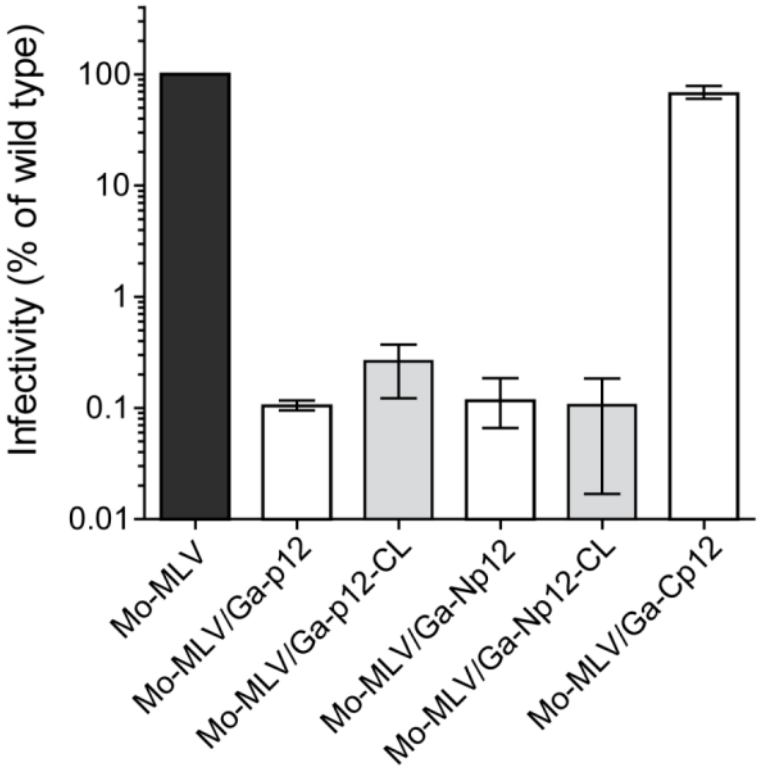


Figure 4.9: Infectivity of Mo-MLV chimeric VLPs with GaLV p12. D17 cells were infected with equal RT-units of chimeric VLPs and the infectivity was quantified. The Infectivity is displayed as a percentage of wild type infectivity with the mean and range of three independent experiments shown.

4.3.3 The effect of Mo-MLV p12 in GaLV on VLP release

As the N-terminus of GaLV p12 could not function in the Mo-MLV background, it was important to assess the inverted chimeras. To do this, GaLV chimeras were constructed with the full Mo-MLV p12, N-terminus of Mo-MLV p12 (N-terminal to PPPY) or C-terminus of Mo-MLV p12 (C-terminal to PPPY) (Figure 4.10A).

The GaLV chimera containing the C-terminus of Mo-MLV p12 did not have a defect in VLP release (Figure 4.10B, GaLV/Mo-Cp12). However, unlike the Mo-MLV chimeras, when GaLV contained the N-terminus of Mo-MLV p12 there was a dramatic reduction in the amount of chimeric VLPs released (Figure 4.10B, GaLV/Mo-Np12). Unsurprisingly, this was also seen with the chimera containing the entire Mo-MLV p12, although the effect was slightly less severe (Figure 4.10B, GaLV/Mo-p12). A *BsrG*I restriction endonuclease site was cloned into the GaLV Gag-Pol vector, 5' to the start of *gag,* to enable cloning of the chimeric fragments into the Gag-Pol vector. The reversion of this modification to the original sequence did not increase VLP production (data not shown).

To assess the Gag processing in the GaLV chimeras, VLPs were concentrated and analysed by immunoblotting with p12 monoclonal and polyclonal antibodies (Figure 4.10C, left and right panels, respectively). As mentioned above, neither antibody recognised the GaLV p12. Both the GaLV/Mo-p12 and GaLV/Mo-Np12 chimeras contained many bands, which likely represented Gag processing intermediates or processing errors (Figure 4.10C, left panel). The prominent band for 'mature' p12 in both these chimeras was slightly larger than that of Mo-MLV p12 in size. This was unexpected as the p12 amino acid sequence in Mo-MLV and GaLV/Mo-p12 is exactly the same. In addition, the p12 species in the GaLV/Mo-Np12 chimera has the same theoretical size as Mo-MLV p12 (8.7kDa). An explanation for this could be that the very faint bands, equal and smaller than Mo-MLV p12 (for GaLV/Mo-p12 and GaLV/Mo-Np12, respectively), represent the mature p12 species in these chimeras (Figure 4.10C, left and right panel, between 6kDa and 15kDa). In this scenario, the p12 species in the GaLV/Mo-Np12 must migrate slightly faster than the Mo-MLV p12 due to charge differences between the proteins.

The GaLV chimera with the C-terminus of Mo-MLV p12 had less processing intermediates than the other chimeras (Figure 4.10C, right panel). Again the size of the prominent band for 'mature' p12 in this GaLV/Mo-Cp12 was larger than would be expected based on a theoretical size of 7kDa, indicating that the mature protein may be the smaller faint p12 species on the immunoblot (Figure 4.10C, right panel, faint band between 6kDa and 15kDa). In summary, all of the chimeras did produce some form of

'mature' p12 separated from MA, but there were obvious processing defects, especially when the N-terminus of Mo-MLV p12 was in the GaLV background.

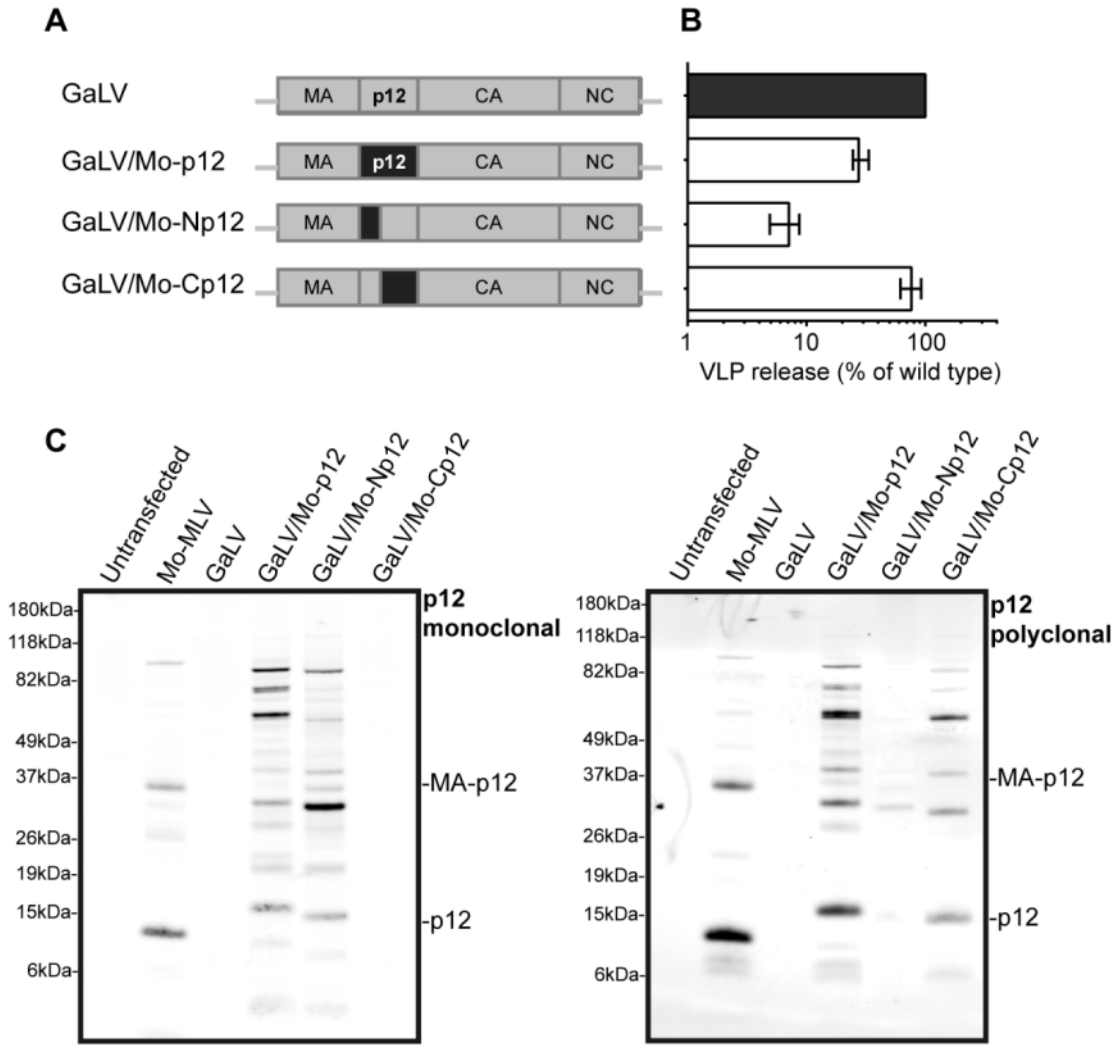


Figure 4.10: GaLV chimeric VLPs containing Mo-MLV p12. (A) Schematic detailing the regions of p12 swapped between GaLV and Mo-MLV. (B) The level of VLPs released was measured by quantifying the RT activity by the RT-ELISA and is shown as the percentage of the wild type. The bars represent the mean and range of three independent experiments. (C) Equal RT-units of VLPs were analysed by immunoblotting with anti-p12 polyclonal (right panel) and p12 monoclonal (left panel) antibodies.

4.3.4 The functionality of Mo-MLV p12 domains in the GaLV background

Despite the processing defects observed in this set of chimeras, all of the chimeras were relatively infectious compared to wild type GaLV (Figure 4.11). The GaLV chimera with the C-terminus of Mo-MLV p12 had an infectivity that was reduced by 3.3-fold compared to the wild type GaLV (Figure 4.11, GaLV/Mo-Cp12). Unsurprisingly, the GaLV chimera containing the whole Mo-MLV p12 displayed the same reduction in infectivity (Figure 4.11, GaLV/Mo-p12). Interestingly, the GaLV containing the N-

terminus of p12, which had an inhibitory effect on VLP release, was as infectious as wild type (Figure 4.11 GaLV/Mo-Np12). Thus, despite the chimeras having defects in VLP assembly/release and Gag cleavage, the released chimeric VLPs were infectious. Therefore, as was seen with the Mo-MLV chimeras, the C-terminus of p12 could function in the GaLV background, but, in contrast to the Mo-MLV chimeras, the Mo-MLV p12 N-terminus could also function in the GaLV background.

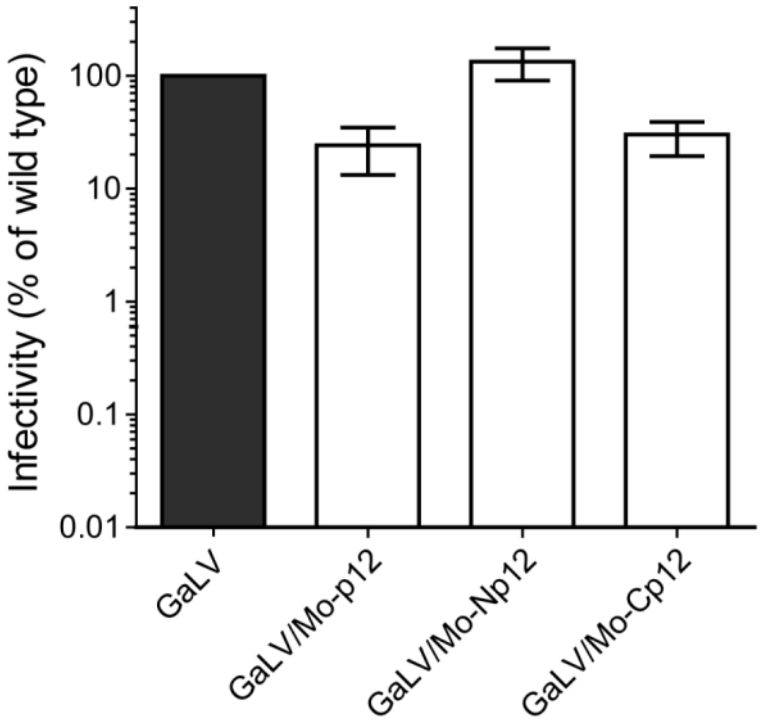


Figure 4.11: Infectivity of GaLV chimeric VLPs with Mo-MLV p12. D17 cells were infected with equal RT-units of chimeric VLPs and the infectivity was quantified. The infectivity is displayed as a percentage of wild type infectivity with the mean and range of three independent experiments shown.

4.4 Summary

This chapter looked at the function of p12 in other gammaretroviruses to establish if the function of p12 in early infection is conserved. By introducing p12 mutations, positionally analogous to those in Mo-MLV p12, into a panel of gammaretroviruses, the early function of p12 was shown to be conserved. Interestingly, while the whole C-terminal region was required for all of the gammaretroviruses analysed, some gammaretroviruses had portions of the N-terminal region which could be altered with minor effects on infectivity.

To test the conservation of p12 further, I looked closer at the GaLV p12. This p12 molecule is the most divergent from Mo-MLV p12 in the panel of viruses studied here, and it contains a very small N-terminal region when compared to the MLVs and FeLV. GaLV mixed particles showed that the N- and C-terminal p12 mutants acted in a similar way to the Mo-MLV mixed particles with the following differences: (i) the dominant negative effect of GaLV p12 mutants 13 and 14 was less severe when compared with Mo-MLV mutant 13 and 14, and (ii) GaLV p12 mutant 15 had no dominant negative activity in mixed particles. The lack of a dominant negative effect of p12 mutant 15 was confirmed using the point mutant R56A in mixed particles, which also lacked dominant negative activity.

The essential residues in the N- and C-terminal domains of GaLV p12 were similar to those in Mo-MLV p12. The N-terminal domain contained only one essential residue (a leucine) and an aspartic acid that was found to have a minor effect on infectivity when mutated to alanine. Interestingly, leucines and aspartic acids were the essential residues in the N-terminal domain of Mo-MLV. This suggests the possibility that the N-terminus of p12 is interacting with a similar factor between the two viruses. The C-terminal domain of GaLV p12, like with Mo-MLV, contained a run of essential amino acids. While the amino acid composition of these motifs varied between Mo-MLV and GaLV, some features were shared. (i) Both contained arginines that were essential (Mo-MLV four versus GaLV three), (ii) both had an essential proline at the N-terminus of the motif and (iii) a phosphorylatable residue was amongst the essential residues in the motif (Mo-MLV S61 and GaLV T50).

Using chimeric Mo-MLV or GaLV viruses it was shown that the C-terminus of p12 could be swapped between these viruses, whereas the N-terminal domain of GaLV was unable to function in the Mo-MLV background. This implied that the C-terminal domains of gammaretroviral p12 proteins serve a common function between viruses, and as speculation, the C-terminal domain may interact with some essential host factor. As the small N-terminus of GaLV p12 was unable to function in the Mo-MLV background, this pointed towards the N-terminus of p12 having a virus specific function, although other explanations exist. The fact that the Mo-MLV p12 N-terminus can function in the GaLV background would argue against this notion. However, due to the sheer size of the Mo-MLV p12 N-terminal domain, it could possibly adapt to provide the required function in GaLV.

It should be noted that there were assembly/release defects when GaLV contained the N-terminus of Mo-MLV p12. This potentially implicates the N-terminus of GaLV p12 as important during the late stages of the life cycle. However, it is entirely possible that this is merely a knock on effect of introducing extra sequence into the GaLV Gag. For example it was recently published that the runs of prolines in the C-terminus of MA and p12 produced an extended conformation of the MLV Gag poly-protein *in vitro* (Datta et al., 2011). The significance of this observation *in vivo* remains to be confirmed, but it is entirely possible that the added prolines (four extra) from the N-terminus of Mo-MLV p12 may affect the Gag polyprotein and thus assembly/release.

Chapter 5.

Investigations into the functions' of the early acting domains of the gammaretroviral protein p12

Indications towards the function of p12 have been around for some time. Mutation of the C-terminal domain of p12 results in PICs, which are biochemically identical to wild type, but are unable to physically integrate their viral genome during infection (Yuan et al., 2002; Yuan, Li, and Goff, 1999). In addition, mutation of the C-terminal domain of p12 does not affect the ability of the mutant PICs to integrate *in vitro* (Yuan et al., 2002). Therefore localisation to the chromosomal DNA appeared to be defective with these mutants. This idea was taken further using immunofluorescence. It was shown that, during mitosis, p12 was localised to the condensed chromatin in infected cells, something that was not seen for the C-terminal p12 mutant 14 (Prizan-Ravid et al., 2010).

It had been assumed, for a long time, that both the N- and C-terminal p12 mutants have a similar defect, and thus little work has been done on the N-terminal domain of p12. There does appear to be an alteration to core morphology (or stability) based on the inability of the 'DLL' mutant to be recognised by TRIM5alpha in infected cells (Zhang et al., 2011). Thus, the N-terminal domain of p12 may have a role in the morphology of the CA core. Additionally, the function of p12 has been genetically linked to CA by studying the infectivity of chimeras between MLV and spleen necrosis virus (SNV) (Lee and Nagashima, 2005).

This chapter describes various manipulations and assays to gain insight into the functions of the N- and C-terminal domains of p12. Importantly, a different function has now been assigned to each domain with the N-terminal domain of p12 participating in the formation of the mature retroviral core and the C-terminal domain playing a tethering function retaining the PIC on the condensed chromatin during infection.

5.1 Detecting the CA core within cells infected by p12 mutant VLPs

After fusion of the virion membrane and host cell membranes, the retroviral core, composed of a fullerene arrangement of CA hexamers, is released into the cytoplasm (Briggs et al., 2003; Mortuza et al., 2004). The CA core-recognising restriction factors, TRIM5alpha and Fv1, are believed to interact with the 'intact' CA core of incoming retroviruses, blocking replication before and after reverse transcription respectively (Cowan et al., 2002; Hofmann et al., 1999; Jolicoeur and Baltimore, 1976; Munk et al., 2002; Stremlau et al., 2006; Towers et al., 2000; Wu et al., 2006). Both these restriction factors interact with an 'intact' CA core very soon after its release into the cytoplasm and can be used as sensors for the correctly processed and formed mature retroviral core.

5.1.1 Abrogation of restriction factor activity by p12 mutant N-tropic MLV

Fv1 and TRIM5alpha can be saturated when a cell, expressing either factor, is infected with a high MOI of virus, blocking the restriction factor activity (Bassin et al., 1978; Boone, Innes, and Heitman, 1990; Duran-Troise et al., 1977; Towers, Collins, and Takeuchi, 2002). These cells can then be productively infected with a second (restriction factor-sensitive) virus, within a short time frame (4-6 hours) (Besnier, Takeuchi, and Towers, 2002; Boone, Innes, and Heitman, 1990; Hatziioannou et al., 2003). This phenomenon has been used to develop an assay to test for the correctly processed CA core early after infection, called the abrogation assay (Bassin et al., 1978; Besnier, Takeuchi, and Towers, 2002; Cowan et al., 2002; Duran-Troise et al., 1977), and was used to test the p12 mutants.

Mo-MLV is an NB-tropic virus and is not restricted by Fv1 or human TRIM5alpha thus the N-tropic virus N-MLV with the p12 mutants were tested (Hartley, Rowe, and Huebner, 1970; Towers et al., 2000). Balb/c 3T3 (B3T3, expressing $Fv1^b$) or TE671 (expressing human TRIM5alpha) cells were infected with a two-fold dilution series of the N-MLV VLPs. After 4 hours, the cells were infected with the second virus at a fixed dose (~MOI 1 in *M.dunni* cells) of GFP-encoding N-MLV. After 72 hours, the percentage of GFP^+ cells was quantified using flow cytometery. Infection of B3T3 cells with wild type N-MLV resulted in the abrogation of $Fv1^b$ activity and permitted

infection of the cells by the second indicator virus (Figure 5.1A, black line and triangles). In contrast, infection with the unrestricted B-MLV resulted in no abrogation and restriction of the indicator virus (Figure 5.1A, black line and squares). This represented a greater than 10-fold reduction in abrogation ability of B-MLV compared to N-MLV. The two C-terminal p12 mutants, mutant 13 and 15, displayed an abrogation profile similar to that of wild type N-MLV (Figure 5.1A, pink and magenta lines). Therefore the C-terminal p12 mutants in N-MLV contained CA cores that were recognised by Fv1. Two N-terminal p12 mutants, mutant 5 and 7, have a partial abrogation phenotype, reduced by 2.6-fold and 1.7-fold when compared to wild type N-MLV (Figure 5.1A, light blue and dark blue lines, respectively). This is not surprising as in N-MLV both these p12 mutants were relatively infectious compared to the other p12 mutants (Figure 4.3 B and C). Thus based on their infectivity, it would be expected that these mutant viruses would abrogate restriction, albeit at a slightly reduced level. In contrast, the two remaining N-terminal p12 mutants, mutant 6 and 8, are non-infectious and were unable to abrogate Fv1 restriction (Figure 5.1A, light and dark green lines). The same pattern was observed with the other core-targeting restriction factor, human TRIM5alpha (Figure 5.1B). This suggests that the CA cores from C-terminal p12 mutant N-MLV were processed correctly and available to restriction factors during infection, whereas the non-infectious N-terminal p12 mutant CA cores were not.

Mo-MLV p12 mutants 6 and 8 have been shown to have defects in reverse transcription (Chapter 3 and (Yuan, Li, and Goff, 1999)), and the inability to abrogate restriction factors could be linked to an inability to reverse transcribe in infected cells. Unfortunately, a high background prevented the measurement of viral DNA products in B3T3 or TE671 cells (Virginie Boucherit, personal communication). Therefore the ability of the N-MLV p12 mutants to reverse transcribe in D17 cells was assessed. Quantitative PCR on total DNA isolated from infected cells was performed as mentioned in chapter 3. All of the N-MLV p12 mutants were competent to reverse transcribe in D17 cells (Figure 5.2 A and B). (Note- a high background with mutant 8 was observed and thus a conclusion could not be made regarding this mutants ability to reverse transcribe). Hence it's less likely that the inability of the N-terminal p12 mutants 6 and 8 to abrogate restriction factors is linked to their ability to reverse transcribe *in vivo.*

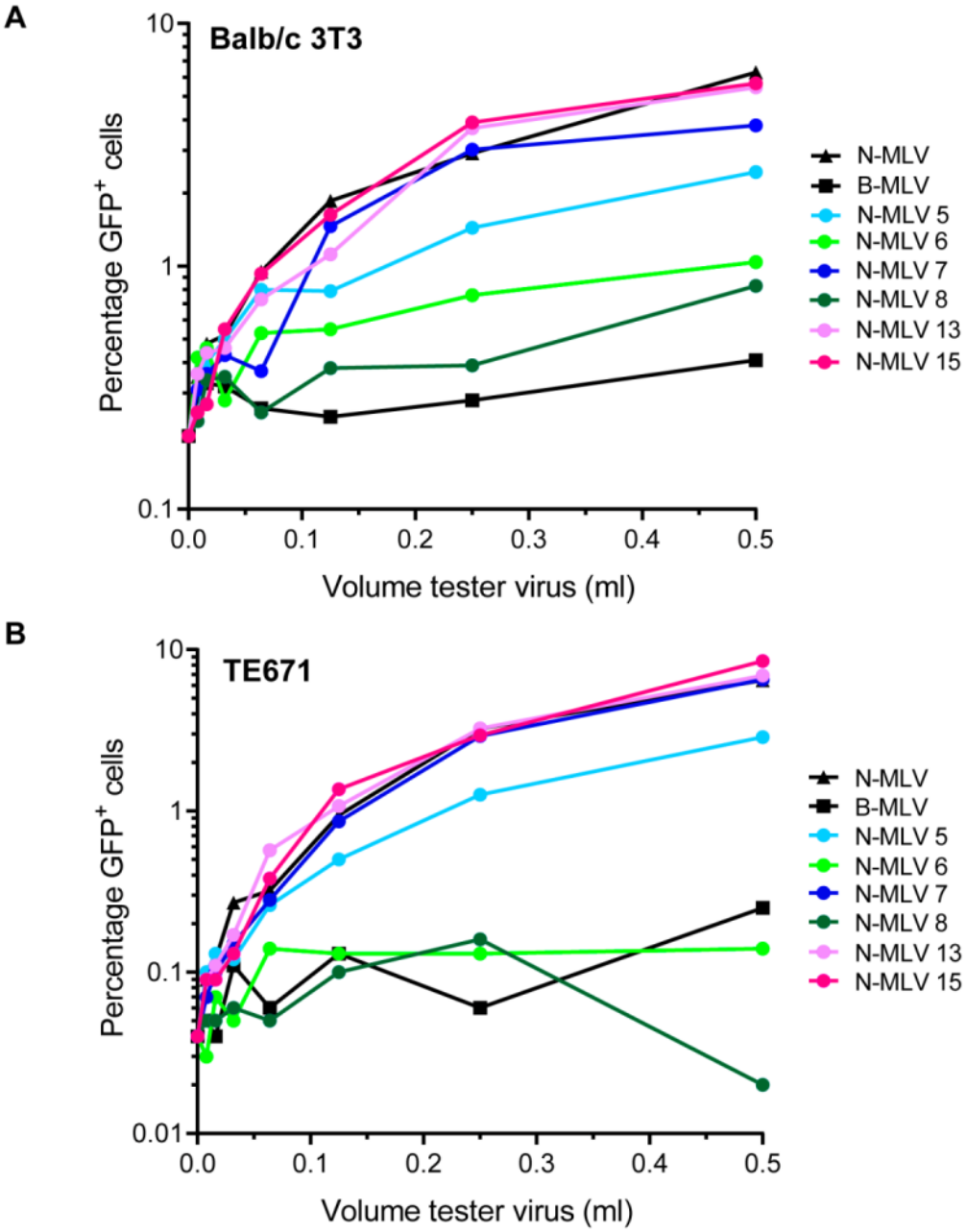


Figure 5.1: Abrogation of Fv1 and TRIM5alpha restriction factor activity by N-tropic MLV carrying mutations in p12. (A) Balb/c 3T3 cells, expressing $Fv1^b$, or (B) TE671 cells, expressing human TRIM5alpha, were infected with a two-fold dilution series of the tester VLPs (see key). Four hours later, cells received a second infection with a fixed volume of GFP-encoding N-MLV. The percentage of GFP^+ cells is plotted for each volume of tester virus. The graphs are representative of at least three independent experiments.

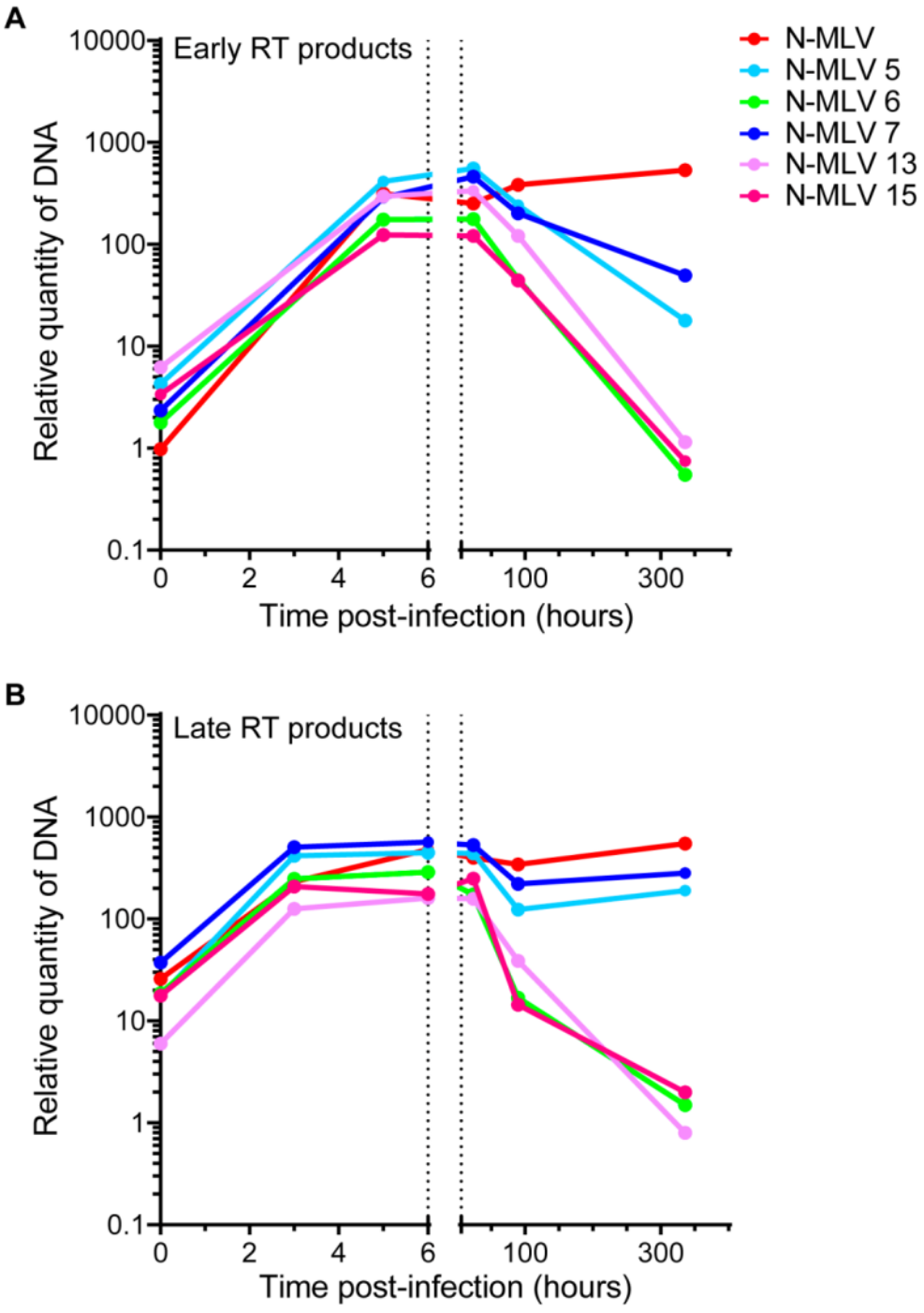


Figure 5.2: The ability of p12 mutant N-MLV VLPs to synthesise viral cDNA in infected cells. D17 cells were infected with equal RT-units of N-MLV VLPs containing mutations in p12 (see key). At various times post infection, DNA was isolated from the infected cells. The relative amounts of viral reverse transcription products (A) strong stop (Early product) and (B) second strand extension (Late product) were measured by qPCR. The graphs are representative of multiple independent experiments. Note- mutant 8 has been omitted due to a high background in the assay. (Data courtesy of Virginie Boucherit).

5.1.2 Abrogation of human TRIM5alpha activity by Mo-MLV with an altered restriction factor tropism

Due to the relatively high infectivity of the p12 mutants 5 and 7 in the N-MLV background, a general conclusion regarding the ability of the N-terminal p12 mutants to abrogate restriction factors could not be made. In an attempt to resolve this, mutations were made in the CA region (D82N, A110R and H117L) of the Mo-MLV Gag-Pol vectors to make the resulting VLPs (designated 'N/Mo') sensitive to restriction by $Fv1^b$ and human TRIM5alpha (Stevens et al., 2004).

To quantify the infectivity of the N/Mo VLPs, D17 or TE671 cells were infected with equal RT-units of VLPs (Figure 5.3). The N/Mo VLPs carrying mutations in p12 had a similar infectivity defect as seen for the Mo-MLV p12 mutants. Interestingly in D17 cells, N/Mo mutant 5 had an infectivity that was 40-fold reduced compared to wild type, which is considerably smaller than the infectivity defect of Mo-MLV p12 mutant 5 (Figure 5.3 and Figure 3.3). Moreover, the N/Mo p12 mutant 7 infectivity was 33-fold less than wild type, which is also considerably smaller than the infectivity defect of Mo-MLV p12 mutant 7 (Figure 5.3 and Figure 3.3). Therefore, despite the infectivity of N/Mo p12 mutant 5 and 7 not being elevated to the levels seen for these mutants in the N-MLV background, the changes in CA have increased their relative infectivity by a small but significant degree.

As would be expected, the CA changes have made the wild type N/Mo sensitive to human TRIM5alpha and the infectivity was reduced by 13.8-fold compared to Mo-MLV (Figure 5.3, TE671 cells, red bar with black stripes versus red bar). Thus, these N/Mo p12 mutants can be used in the abrogation assay with cells expressing human TRIM5alpha.

N/Mo VLPs with or without mutations in p12 were used as tester viruses in the TRIM5alpha abrogation assay. Wild type N/Mo was able to abrogate TRIM5alpha, albeit with a ~three-fold reduction in abrogation ability when compared to N-MLV (Figure 5.4, black line with triangles versus black line with rings). The N/Mo C-terminal p12 mutants, 13 and 15, were able to abrogate as well as wild type N/Mo (Figure 5.4, pink and magenta lines), consistent with what was seen for the C-terminal p12 mutants in the N-MLV background (Figure 5.1B). All of the N/Mo N-terminal p12 mutants were unable to abrogate TRIM5alpha restriction (Figure 5.4, light and dark green and blue lines). Therefore, alterations to the N-terminal domain of Mo-MLV p12

affect the ability of restriction factors to interact with the CA core, immediately after its release into the cytoplasm.

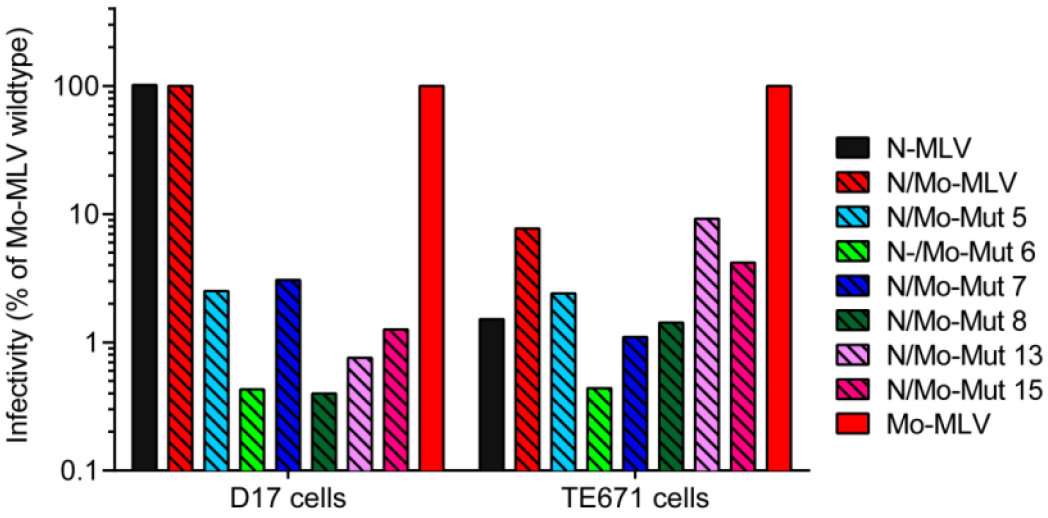


Figure 5.3: Infectivity of Mo-MLV with genetically altered restriction factor tropism. Three mutations were introduced into Mo-MLV CA (D82N, A110R and H117L) in the Mo-MLV Gag-Pol vector to make the resulting VLPs sensitive to Fv1 (called N/Mo). D17 cells (left side) and TE671 cells (right side) were infected with equal RT-units of VLPs and the infectivity quantified after at least 48 hours. The infectivity is displayed as the percentage of the wild type Mo-MLV infectivity for each cell type. The graphs are representative of multiple independent experiments.

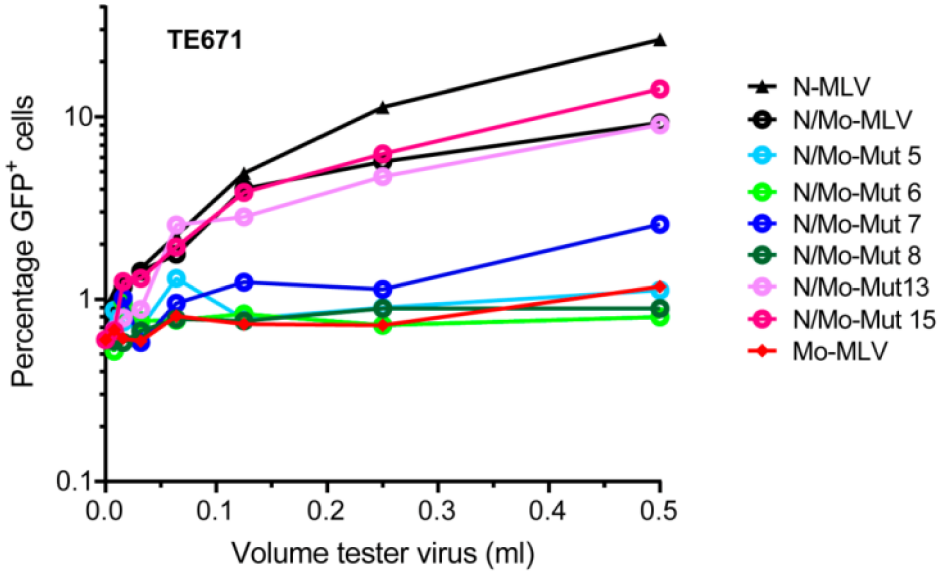


Figure 5.4: Abrogation of human TRIM5alpha restriction factor activity by N/Mo-MLV VLPs carrying mutations in p12. TE671 cells, expressing human TRIM5alpha, were infected with a two-fold dilution series of the tester VLPs (see key). Four hours later, cells received a second infection with a fixed volume of GFP-encoding N-MLV. The percentage of GFP^+ cells is plotted for each volume of tester virus. The graph is representative of at least three independent experiments.

One possible explanation for the failure of the N-terminal p12 mutants to abrogate restriction, is that mutation in p12 affects the processing and/or function of CA from the Gag polyprotein. To test if mutation in the N-terminal domain of p12 affects the processing of CA from Gag, mixed particles were synthesised from N-MLV Gag-Pol and B-MLV Gag-Pol, with or without the p12 mutant 6 (Figure 5.5A). All of these mixed particles contain 10% of wild type p12 and were as infectious as wild type N- or B-MLV (data not shown).

When only 10% of the CA in the mixed particle was N-tropic, TRIM5alpha could not recognise the core and its activity was not abrogated (Figure 5.5B, B6/N, golden brown line). When all of the CA in the mixed particle was N-tropic, and 90% of the Gag contained the upstream N-terminal p12 mutation, it was able to be recognised by TRIM5alpha and could abrogate restriction (Figure 5.5B, N6/N, maroon line). Finally, when all of the N-tropic CA in the mixed particle came from Gag that contained the upstream N-terminal p12 mutation, it was also able to abrogate restriction by TRIM5alpha (Figure 5.5B, N6/B, orange line). Therefore, when the *gag* sequence contained an N-terminal p12 mutation upstream to the CA sequence, it did not affect the processing or function of CA from the resulting Gag polyprotein with regards to forming a mature core recognised by TRIM5alpha.

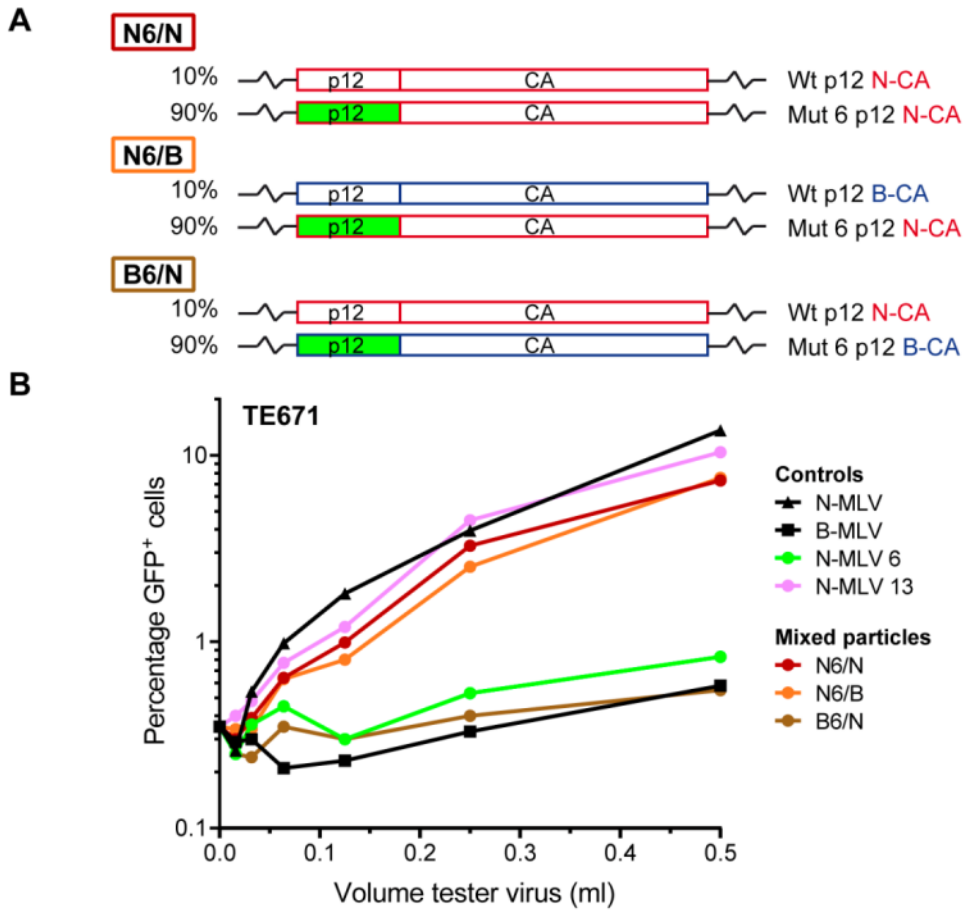


Figure 5.5: The effect of p12 mutation in *gag* on the ability of human TRIM5alpha to recognise CA in the abrogation assay. (A) Schematic detailing the Gag-Pol plasmids used to construct the mixed particles tested in this assay. The percentage of each plasmid transfected is shown on the left. (B) TE671 cells, expressing human TRIM5alpha, were infected with a two-fold dilution series of the tester VLPs (see key and A). Four hours later, cells received a second infection with a fixed volume of GFP-encoding N-MLV. The percentage of GFP^+ cells is plotted for each volume of tester virus. The graph is representative of at least three independent experiments.

5.1.3 Further investigations into the role of the N-terminal domain of p12 in restriction factor saturation

To assess if the affect of the p12 N-terminal domain on abrogation ability is independent of the C-terminal domain, the abrogation ability of mixed particles composed of an N- and C-terminal p12 mutant were tested. All the compositions of these mixed mutant particles are non-infectious, as was shown in chapter 3 (Figure 3.9).

All N-MLV mixed mutant particles were able to abrogate TRIM5alpha restriction when they contained as little as 10% of p12 with an active N-terminal domain (10% of a p12 C-terminal mutant) (Figure 5.6 A, B and C). Thus the affect of the N-terminal domain of p12 on abrogation ability was entirely independent of the C-terminal domain.

Some of the alanine point mutations in the N-terminal domain of Mo-MLV p12 only reduced the infectivity of the virus by 2-10-fold (Figure 3.15). The most significant reductions in infectivity were seen when leucines and aspartic acids were mutated to alanines (Figure 3.15). To see if the L16A point mutant, in the N-terminal domain of p12 was also unable to abrogate restriction, N/Mo VLPs containing this mutation were generated and used in the TRIM5alpha abrogation assay.

Both the L16A p12 mutant and the double domain mutant LR(16,66)AA, displayed the same phenotype. Both were able to partly abrogate human TRIM5alpha activity, which resulted in ~2-3-fold reduced abrogation ability compared to the controls (Figure 5.7A, grey dotted line with green rings and black dotted line with yellow rings).

N/Mo wild type appeared to have the same abrogation ability as these two mutants in this experiment, but when the percentage of GFP positive cells was plotted against the RT activity of each dilution of virus, it could be seen that this was due to less wild type N/Mo VLPs synthesised for use in this assay (Figure 5.7 A and B, compare black lines with black rings). As the RT activity is quantified retrospectively (the tester virus is used immediately after harvest), this check was used to ensure the phenotype observed in the abrogation assay was not due to reduced production of tester virus. Therefore, when the N-terminal domain of p12 was inactivated by a point mutation (L16A) the virus was able to partly abrogate restriction by TRIM5alpha.

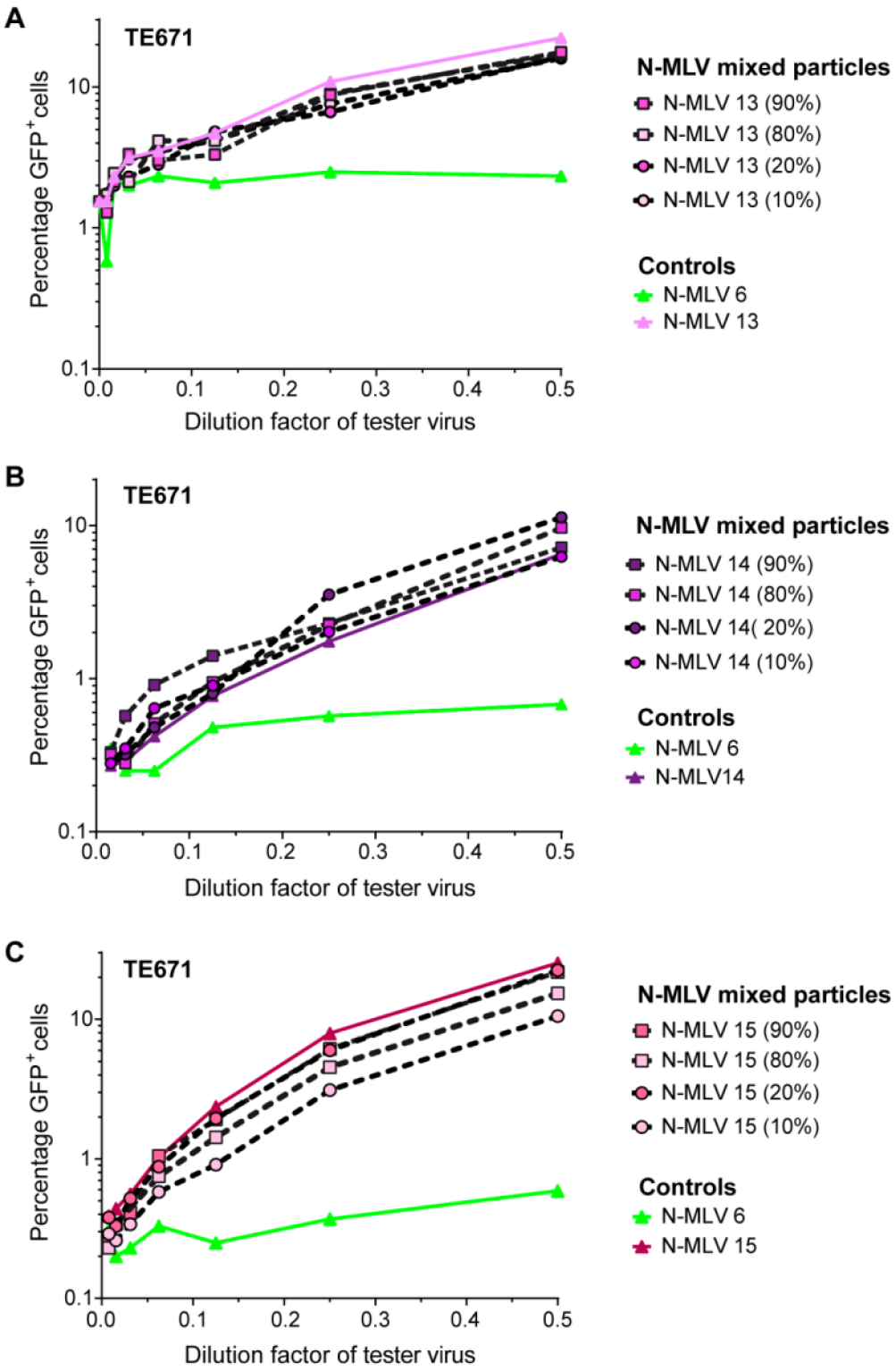


Figure 5.6: Abrogation of human TRIM5alpha restriction factor activity by VLPs containing a mixture of N- and C- terminal p12 mutants. TE671 cells, expressing human TRIM5alpha, were infected with a two-fold dilution series of (A) mutant 13, (B) mutant 14 or (C) mutant 15 tester mixed VLPs and controls (see key). (Note- The percentage of C-terminal p12 mutant in the N-MLV mixed particles is indicated in brackets). Four hours later, cells received a second infection with a fixed volume of GFP-encoding N-MLV. The percentage of GFP^+ cells is plotted for each dilution of tester virus. The graphs are representative of at least three independent experiments.

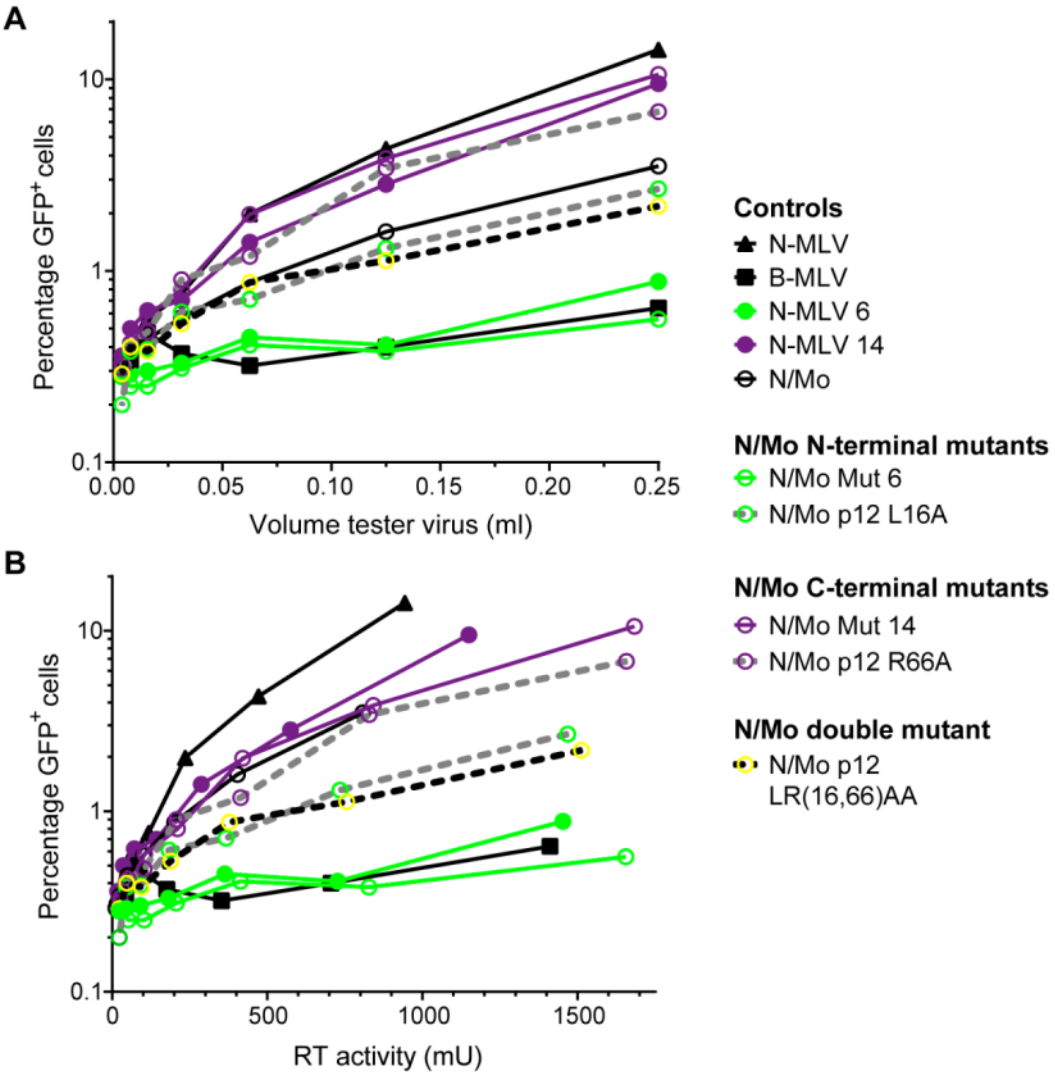


Figure 5.7: Abrogation of human TRIM5alpha restriction factor activity by N-tropic MLV with point mutations in the N- or C- terminal domains of p12. TE671 cells, expressing human TRIM5alpha, were infected with a two-fold dilution series of the tester VLPs (see legend) immediately after harvest. Four hours later, cells received a second infection with a fixed volume of GFP-encoding N-MLV. The percentage of GFP^+ cells are plotted against the volume of tester virus (A) or the RT activity of each virus dilution (B) from one experiment. The graphs are representative of two independent experiments.

Table 5.1: Summary of abrogation assay data.

Virus	Restriction factor	Abrogation ability	Figure
N-MLV	HuTRIM5alpha	+++	5.1
N-MLV 5	HuTRIM5alpha	+++	
N-MLV 6	HuTRIM5alpha	-	
N-MLV 7	HuTRIM5alpha	+++	
N-MLV 8	HuTRIM5alpha	-	
N-MLV 13	HuTRIM5alpha	+++	
N-MLV 15	HuTRIM5alpha	+++	
N-MLV	Fv1b	+++	5.1

N-MLV 5	Fv1b	+	
N-MLV 6	Fv1b	-	
N-MLV 7	Fv1b	+++	
N-MLV 8	Fv1b	-	
N-MLV 13	Fv1b	+++	
N-MLV 15	Fv1b	+++	
Mo-MLV	HuTRIM5alpha	-	5.4
N/Mo	HuTRIM5alpha	+++	
N/Mo 5	HuTRIM5alpha	-	
N/Mo 6	HuTRIM5alpha	-	
N/Mo 7	HuTRIM5alpha	-	
N/Mo 8	HuTRIM5alpha	-	
N/Mo 13	HuTRIM5alpha	+++	
N/Mo 14	HuTRIM5alpha	+++	
N/Mo 15	HuTRIM5alpha	+++	
N/Mo L16A	HuTRIM5alpha	+	5.7
N/Mo R66A	HuTRIM5alpha	+++	
N/Mo LR(16,66)AA	HuTRIM5alpha	+	
N6/N	HuTRIM5alpha	+++	5.5
N6/B	HuTRIM5alpha	+++	
B6/N	HuTRIM5alpha	-	
N13/N6 (0.9:0.1)	HuTRIM5alpha	+++	5.6
N13/N6 (0.8:0.2)	HuTRIM5alpha	+++	
N13/N6 (0.2:0.8)	HuTRIM5alpha	+++	
N13/N6 (0.1:0.9)	HuTRIM5alpha	+++	
N14/N6 (0.9:0.1)	HuTRIM5alpha	+++	
N14/N6 (0.8:0.2)	HuTRIM5alpha	+++	
N14/N6 (0.2:0.8)	HuTRIM5alpha	+++	
N14/N6 (0.1:0.9)	HuTRIM5alpha	+++	
N15/N6 (0.9:0.1)	HuTRIM5alpha	+++	
N15/N6 (0.8:0.2)	HuTRIM5alpha	+++	
N15/N6 (0.2:0.8)	HuTRIM5alpha	+++	
N15/N6 (0.1:0.9)	HuTRIM5alpha	+++	

+++ = 8-fold greater abrogation than negative control

+ = 3-8-fold greater abrogation than control

- = Less than 3-fold greater abrogation than control

5.2 Investigations into the state of the virion CA core before entry to the cytoplasm

All of the data compiled so far are consistent with the possibility that the mutations in the N-terminus of p12 have an effect on the very final stages of retroviral replication (maturation). Maturation is a complex process of rearrangement within the virion,

triggered by the cleavage of the Gag and Pol precursors into the mature proteins (Yoshinaka and Luftig, 1977). This rearrangement is vital for infectivity of the virus, and virions that cannot undergo this process are not infectious (Crawford and Goff, 1985; Katoh et al., 1985; Kohl et al., 1988). It has been shown so far that Mo-MLV N-terminal p12 mutants all bud from the cell and contain the processed Gag proteins, but they are unable to infect cells. As all of the N-terminal Mo-MLV p12 mutants (barring mutant 6) can reverse transcribe in infected cells, but cannot abrogate restriction factors, it is possible that the core is unstable or malformed. The experiments described in this next section will detail investigations aimed to assess the formation and stability of the pre-entry CA core from Mo-MLV p12 mutants.

5.2.1 Investigations into the effect of Gag processing on the early function of p12

In wild type Mo-MLV, a sizable proportion of p12 in the particle remains fused to MA (Figure 3.2). This indicated, that the processing between MA and p12 is inefficient, consistent with previous reports (Oshima et al., 2004). Consequently, it is possible that the active p12 molecule, during the early stages of infection, is the MA-p12 fusion. To assess this, a mutation was made at the end of MA (Y131D) in the Mo-MLV Gag-Pol expression vector to block the MA-p12 cleavage (Oshima et al., 2004).

To assess the liberation of free p12 from MA in VLPs containing the MA Y131D mutation, immunoblotting analysis was performed using a p12 polyclonal and MA antibodies (Figure 5.8A). This cleavage mutation blocked the majority of MA-p12 cleavage, as was seen by a loss of mature MA signal in the MA immunoblot (Figure 5.8A, bottom panel) plus an increase of MA-p12 and a significant reduction of free p12 in the p12 immunoblot (Figure 5.8A, middle panel).

The infectivity of p12 mutants, with and without the MA Y131D mutation, was quantified. Consistent with previous reports, the MA-p12 processing mutation had a minor (3.6-fold) effect on infectivity of wild type Mo-MLV (Figure 5.8B, red bar versus red/black checked bar) (Oshima et al., 2004). As expected, the addition of the MA Y131D mutation did not have an effect on the infectivity phenotype of the Mo-MLV p12 mutants (Figure 5.8B).

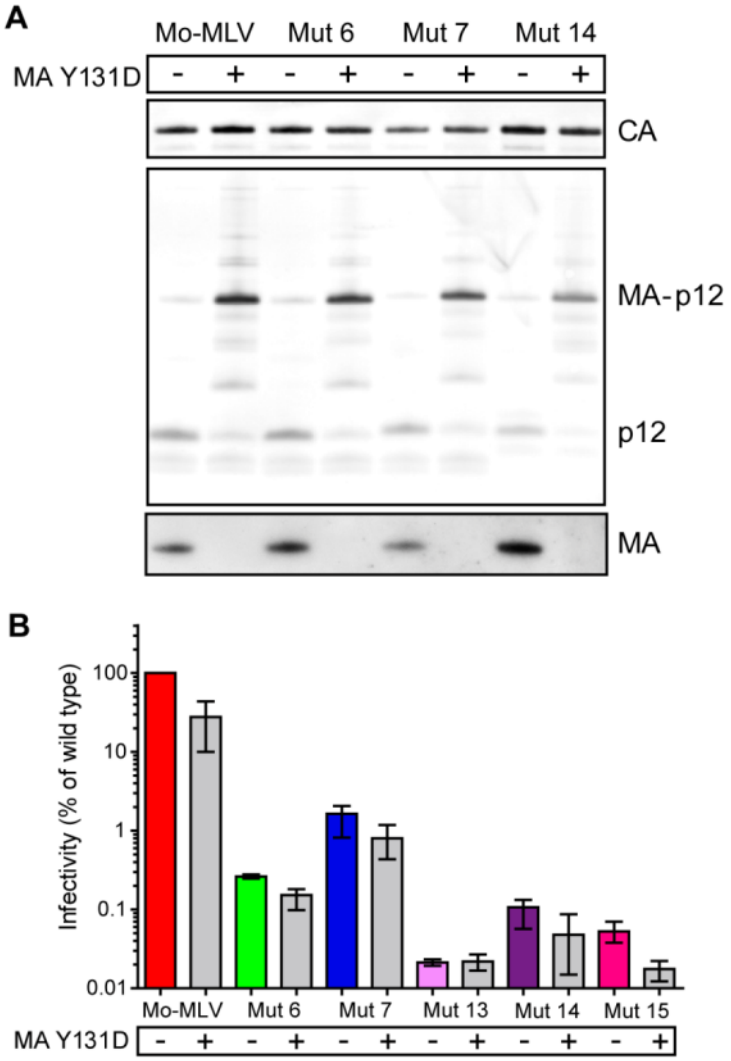


Figure 5.8: Gag processing and infectivity of Mo-MLV MA Y131D MA-p12 processing mutant. (A) Equal RT-units of VLPs containing Y or D at position 131 of MA, with or without p12 mutations, were analysed by immunoblotting with anti-CA (top panel), anti-p12 polyclonal (middle panel) or anti-MA (bottom panel) antibodies. (B) D17 cells were infected with equal RT-units of VLPs and infectivity was quantified. The graph indicates the mean and range infectivity of the VLPs, as a percentage of wild type, from three independent experiments.

To completely block MA-p12 cleavage, two additional mutations were introduced into p12, PA(1,2)DE. This changed the sequence at the MA-p12 border from YPA to DDE (called 'DDE' mutant). The processing of Gag in this mutant was assessed by immunoblotting with a panel of MLV antibodies (Figure 5.9 A and B). No mature MA protein was detected for the Mo-MLV DDE mutant, as was observed for the MA Y131D mutant (Figure 5.9A). A weak mature p12 band could be observed for the MA Y131D mutant but no band was detected for the Mo-MLV DDE mutant (Figure 5.9B), indicating the DDE mutant had completely blocked MA-p12 processing. The infectivity of the DDE mutant was, surprisingly, as infectious as the MA Y131D mutant (Figure 5.9C). Therefore, it is possible that p12 can function as the uncleaved MA-p12 fusion.

However, an increase in p12-containing Gag processing intermediates could be seen for the MA-p12 cleavage mutants (Figure 5.9B, marked by asterisks). These were smaller than MA-p12 and potentially could function relatively efficiently as the 'active' p12 species during the early stages of infection. As MA is anchored to the virion membrane, it would be assumed that the majority remains with the virion membrane after fusion and not travel with the PIC. Thus, as p12 is a component of the PIC, the latter scenario seems more plausible, where blocking MA-p12 gave rise to alternatively processed forms of p12 which could function in the place of mature p12.

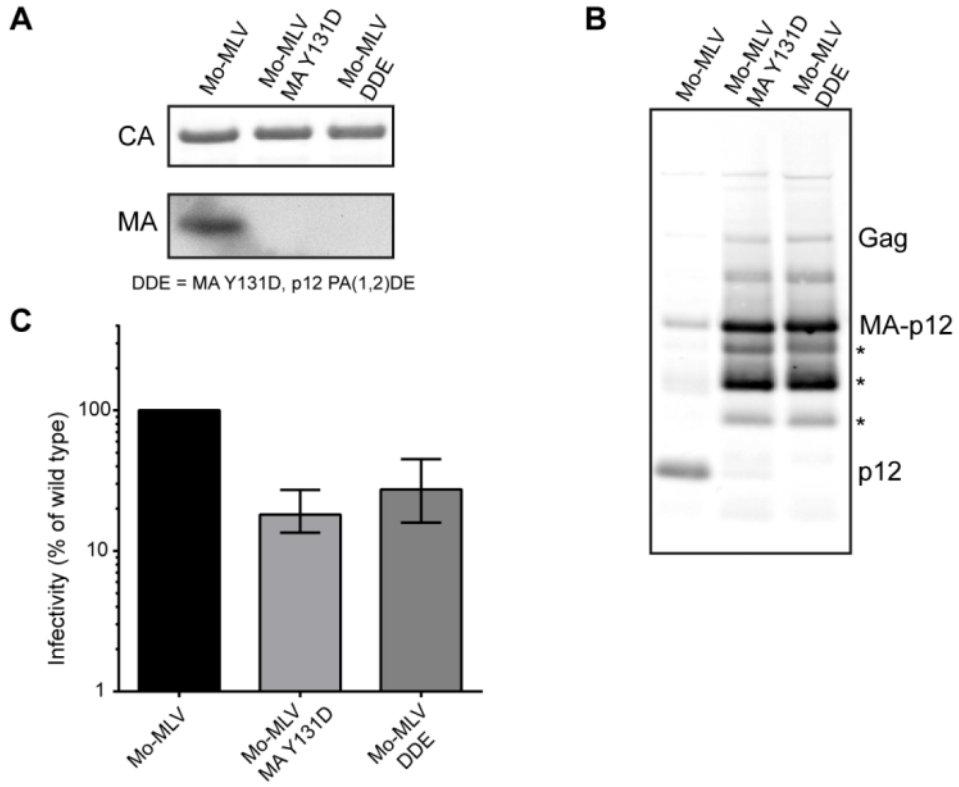


Figure 5.9: The ability of extra mutations at the MA-p12 border to further inhibit MA-p12 processing during maturation. (A) Equal volumes or (B) equal RT-units of VLPs were analysed by immunoblotting with anti-CA (A, top panel), anti-MA (A, bottom panel) or p12 monoclonal (B) antibodies. p12-containing Gag processing intermediates are represented by asterisks (*). (C) D17 cells were infected with equal RT-units of VLPs and infectivity was quantified after at least 48 hours. The graph indicates the mean and range infectivity of the VLPs, as a percentage of wild type, from three independent experiments.

Mo-MLV C-terminal p12 mutants in mixed particles display a dominant negative effect on infectivity (Figure 3.8). To test if fusing the C-terminal p12 mutants to MA abolished this activity, mixed particles were made where the Gag-Pol vector containing the p12 mutation also contained the MA Y131D MA-p12 processing mutation. Firstly, the N-terminal p12 mutants 6 and 7 were tested. Addition of the processing mutation had no effect on the infectivity phenotype of the mixed particles (Figure 5.10, green line and blue line versus green dashed line and blue dashed line, respectively). Interestingly,

when the Gag-Pol with the C-terminal p12 mutant contained the processing mutation, they lost their dominant negative effect on infectivity (Figure 5.11 A, B and C, compare solid coloured lines versus coloured dashed lines with black squares). Furthermore, when both the wild type and the C-terminal p12 mutant Gag-Pol vectors contained the processing mutation, the dominant negative effect was restored (Figure 5.11, A, B and C, coloured dot-dashed lines with black triangles).

This interesting result could highlight some important information regarding the N-terminal domain interaction. However due to the lack of some experimental evidence, the following assumptions will have to be made: (i) the functional p12 in the MA Y131D mutants is the alternatively processed forms of p12 and not MA-p12 and (ii) the alternatively processed forms of p12 are cleaved from Gag with slower kinetics than mature p12. Thus if these assumptions are correct and bearing in mind that the dominant negative effect requires an active N-terminal domain. Removal of the dominant negative effect of C-terminal p12 mutants by blocking the MA-p12 processing could indicate the following: (i) the N-terminal domain interaction occurs after cleavage of p12 from MA, (ii) either mature p12 is more adapted at making the N-terminal domain interaction, or has a kinetic advantage over the alternatively processed forms of p12 (mentioned above). Finally, the fact that the C-terminal mutants act as a dominant negative as long as the wild type p12 contains the same MA-p12 processing, suggests that the C-terminal p12 mutants are somehow more efficient at making the N-terminal interaction than wild type p12 molecules.

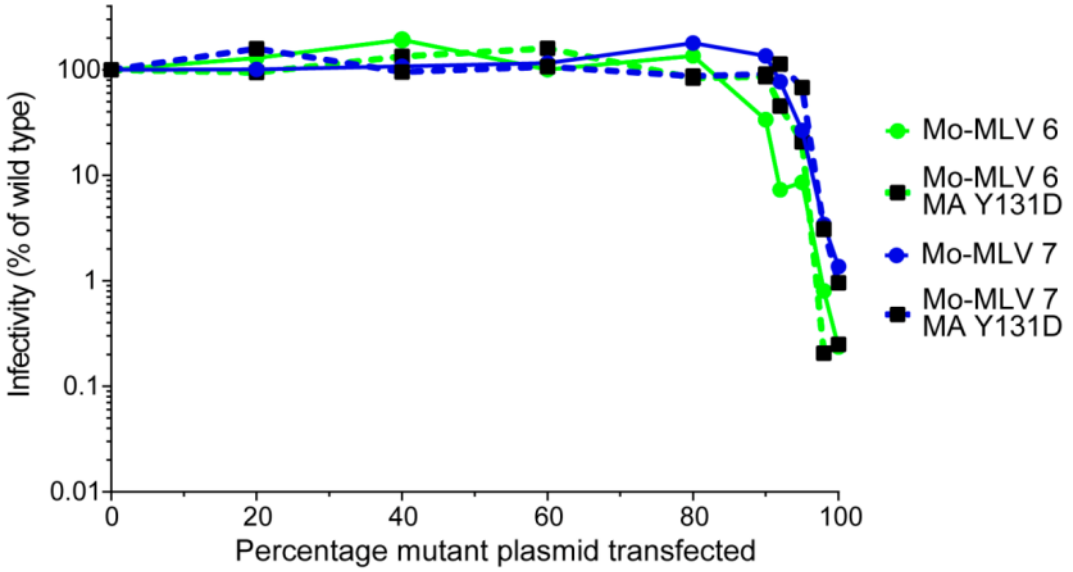


Figure 5.10: Infectivity of Mo-MLV mixed particles containing a mixture of wild type and N-terminal mutant p12 with a MA-p12 processing mutation. D17 cells were infected with equal RT-units of mixed VLPs (see key) and productive infection was quantified after at least 48 hours. Infectivity is plotted as a percentage of wild type Mo-MLV infectivity and the data is representative of at least two independent experiments.

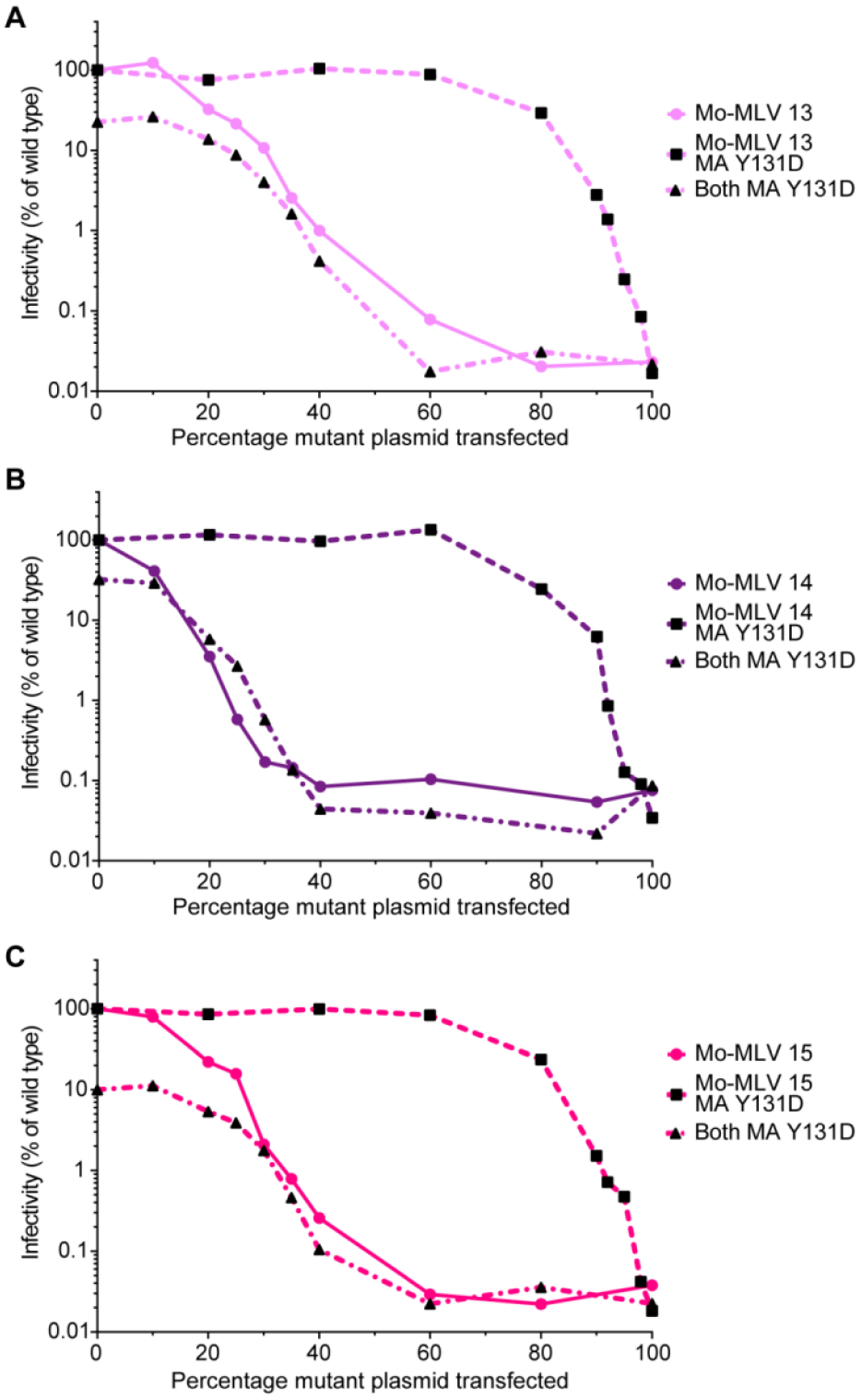


Figure 5.11: Infectivity of Mo-MLV mixed particles containing a mixture of wild type and C-terminal mutant p12 with a MA-p12 processing mutation. Mo-MLV mixed particles, containing wild type p12 and: (A) p12 mutant 13, (B) p12 mutant 14 or (C) p12 mutant 15, with or without the MA Y131D mutation (see key), were synthesised. An additional set of mixed particles were synthesised for each p12 mutant, where the wild type Gag-Pol also contained the MA Y131D mutation (A, B and C, 'both MA Y131D' triangle points with dot-dash lines). The infectivity was quantified and is plotted as a percentage of wild type Mo-MLV infectivity. The data is representative of at least three independent experiments.

5.2.2 Assessing the stability of the Mo-MLV CA core before entry to the cell

HIV-1 cores have been successful isolated from virions by centrifuging through a non-ionic detergent onto a sucrose gradient (Kotov et al., 1999). Early studies were able to isolate MLV cores at a sucrose density of about 1.28g/ml (Bolognesi, Luftig, and Shaper, 1973; Durbin and Manning, 1982). In contrast in a more recent study, mature MLV CA cores could never be isolated at the calculated density for the MLV core (1.25g/ml) (Fassati and Goff, 1999). Therefore a method to isolate Mo-MLV cores similar to that employed for MLV and HIV-1 was devised (shown in Figure 5.12A) (Bolognesi, Luftig, and Shaper, 1973; Durbin and Manning, 1982; Kotov et al., 1999). The gradient was fractioned from the top down using a syringe pump driven gradient fractionator (fraction 1= the top). The method used for forming the gradient and collecting the fractions displayed a relatively good linear gradient, which was consistent between experiments (Figure 5.12B).

Mo-MLV VLPs were run in gradients with and without Triton. The fractions were collected and analysed by immunoblotting with an anti-CA antibody. A peak of mature CA was detected in the gradient for both whole VLPs and the detergent stripped VLPs (Figure 5.12C). The peak of mature CA for whole VLPs migrated to a sucrose density of 1.148g/ml, whereas the detergent stripped VLPs migrated to a lower sucrose density of 1.113g/ml. It should be noted that the majority of the CA was seen in the pre-gradient and layers and the very top of the gradient (fraction 2) when VLPs were spun through detergent consistent with previous studies (Figure 5.12 C, fractions 1-3, note-these were diluted further to reduce the intensity of the bands) (Bolognesi, Luftig, and Shaper, 1973; Durbin and Manning, 1982; Fassati and Goff, 1999; Kotov et al., 1999). For detergent-stripped wild type Mo-MLV, the average sucrose density of the peak CA fraction from multiple experiments was 1.115g/ml (n=5) (Figure 5.12C, fraction 7). This sucrose density was significantly less than the sucrose density at which un-lysed MLV virions migrated to on the gradients (Figure 5.12C, 'No Triton' immunoblot). Thus, it was unlikely that this mature CA peak represented un-lysed virions. However, due to the low sucrose density of the fractions that contained the mature CA material compared with the published density of isolated MLV cores (Bolognesi, Luftig, and Shaper, 1973; Durbin and Manning, 1982), it was unlikely that the CA species isolated in these experiments represented whole cores and more likely represented partial cores.

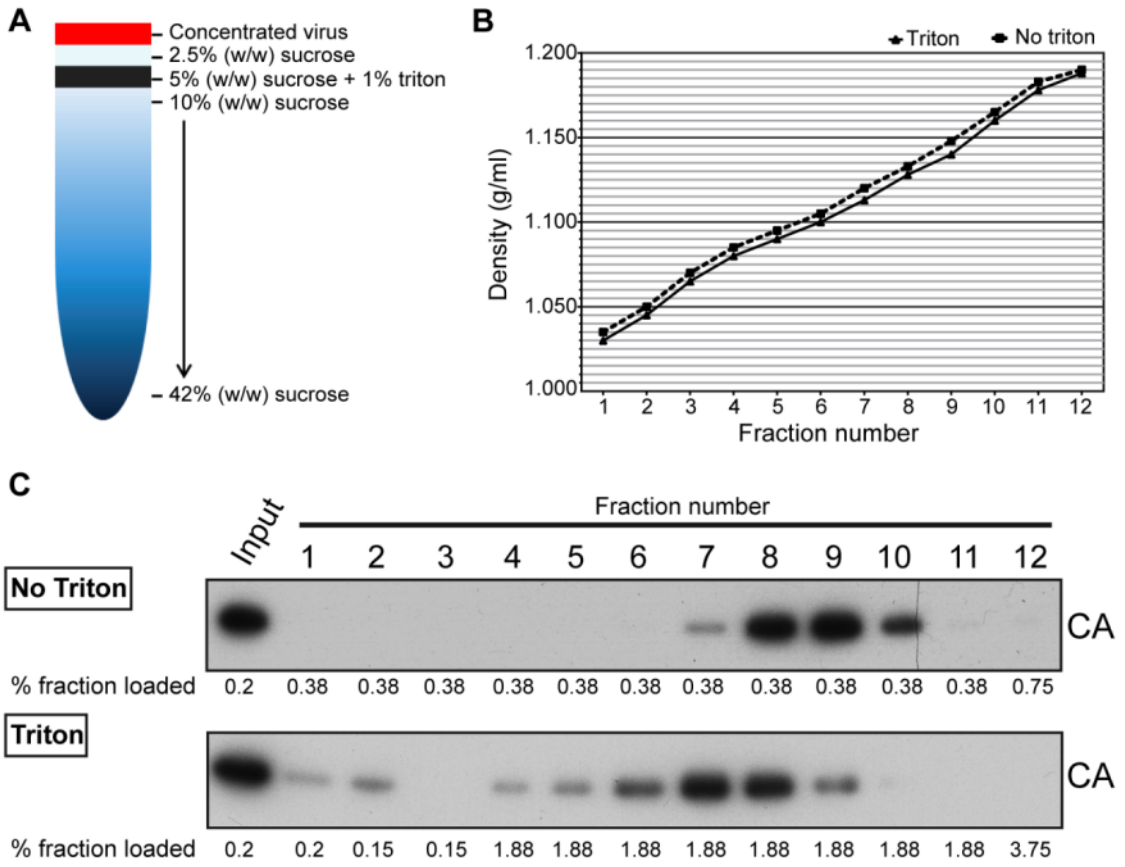


Figure 5.12: Attempted isolation of the MLV CA core from virions. To isolate MLV CA cores, virus was centrifuged through a layer of detergent onto a sucrose gradient. (A) Diagram showing preparation of gradients. For VLP analysis, triton was omitted. (B) Following centrifugation, the density of each fraction was measured using a refractometer and is plotted against the fraction number (fraction 1 = top of the gradient). (C) Wild type Mo-MLV VLPs were run on gradients with (bottom panel) or without Triton (top panel) and samples from each fraction were analysed by immunoblotting with an anti-CA antibody. The number below each fraction indicates the percentage of the total fraction loaded on the gel.

To test if the amount of virus loaded onto the detergent layer gradients affected the position of the CA peak, three gradients were run with increasing amounts of virus loaded (2 plates of virus, 4 plates of virus or 6 plates of virus). The fractions were analysed by immunoblotting with an anti-CA antibody (Figure 5.13). Loading twice the normal amount of virus (normal = two 10cm plates of virus) on the detergent layer gradient did not affect the position of the mature CA peak in the gradient (Figure 5.13A, middle panel fraction 7). Interestingly, when three times as much virus was loaded, the peak also remained at a similar sucrose density but an increase in CA material was seen further down the gradient (Figure 5.13A, bottom panel). Therefore if too much virus was loaded on the detergent layer gradient, further spreading of the CA material was observed, but importantly the peak CA fraction remained at the same sucrose density.

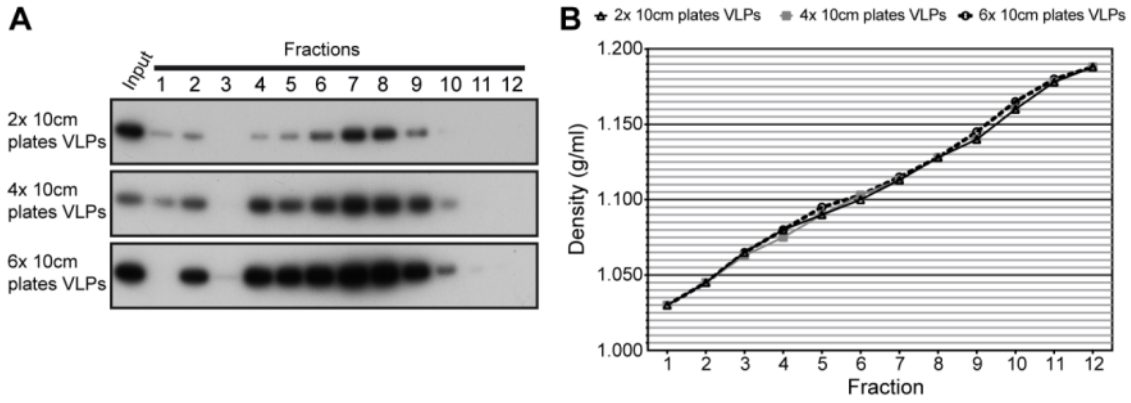


Figure 5.13: The effect of loading more virus onto the detergent layer gradient. Twelve 10cm plates of wild type Mo-MLV were produced by transient transfection of 293T cells, pooled together and concentrated over a 20% (w/v) sucrose cushion. Volumes equivalent to two 10cm plates (black circles), four 10cm plates (black squares) or six 10cm plates (black triangles) of virus were loaded onto separate detergent layer gradients. (A) The gradients were run and the fractions analysed by immunoblotting with an anti-CA antibody. Note- fractions 1-3 are diluted more to reduce the high signal from the 'free' CA monomers at the top of the gradient. (B) Graph displaying the density profile of the gradients from (A).

To establish if the viral RNA genome co-sediments with the CA peaks in this assay, RNA was isolated, reverse transcribed and the relative amount of viral genome was quantified using qPCR. Viral RNA could be detected in the wild type Mo-MLV detergent layer gradient in two peaks at fraction 4 (1.085g/ml) and fraction 11 (1.190g/ml) (Figure 5.14A). Fraction 4 and 11, contained little and no mature CA, respectively (Figure 5.14B). These peaks are relatively close to the where the 35S (~fraction 7) and 70S (~fraction 10) gRNA would be expected to migrate and possibly reflect partially NC-coated gRNA (based on the approximate calculated S-values for each fraction). Thus, the mature CA peak detected in the gradients did not co-migrate with the gRNA, suggesting they are not whole cores.

As it was possible to isolate some higher order form of mature CA from the core, using this assay, it seemed pertinent to assess if the p12 mutants also contained this CA core structure. The sucrose density of the fractions containing the peak of mature CA for the p12 N-terminal mutants was shifted to a lower sucrose density, when compared to wild type (Figure 5.15A). In contrast, the opposite was true for the C-terminal p12 mutants, where the peak was shifted to a higher sucrose density for mutants 14 and 15, although mutant 13 was similar to wild type (Figure 5.15A). The sucrose density profiles of the gradients, shown in Figure 5.15A, were plotted to emphasise their comparability (Figure 5.15B).

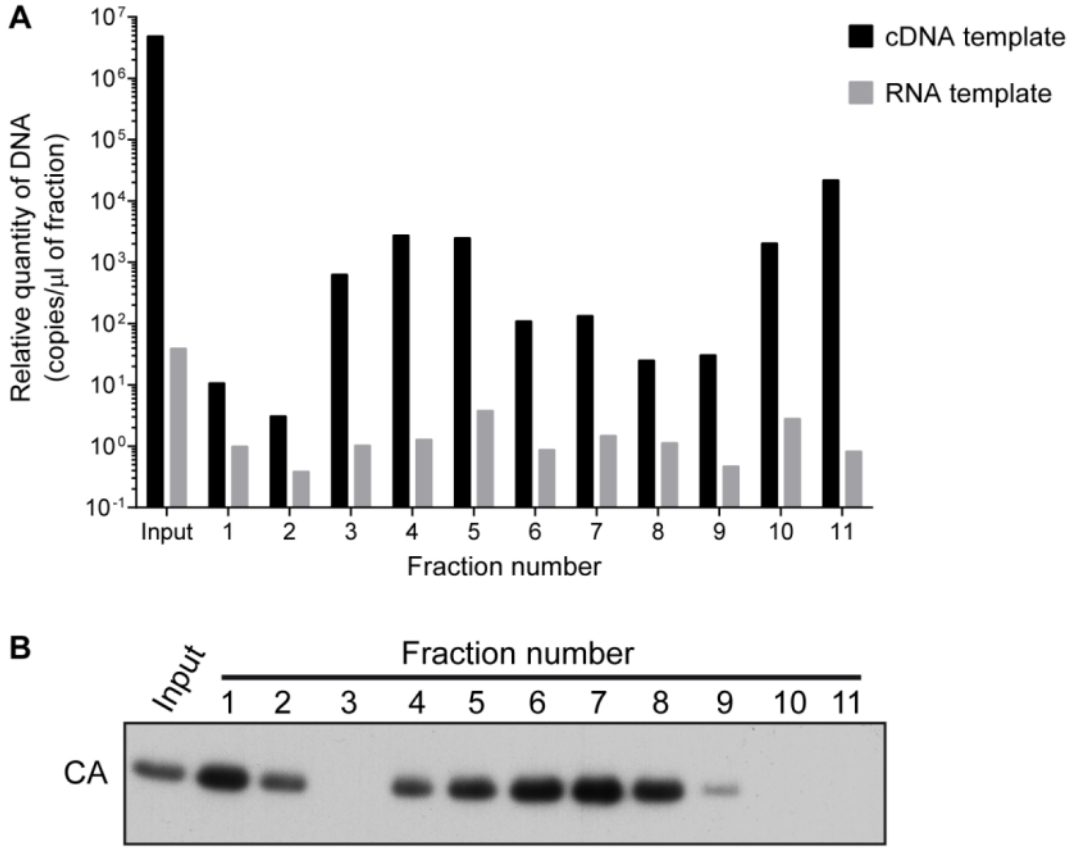


Figure 5.14: Position of the genomic viral RNA in detergent layer gradients. Mo-MLV VLPs were synthesised and spun through detergent into a gradient. (A) RNA isolated from each fraction by phenol-chloroform extraction was reverse transcribed and cDNA was quantified using qPCR. The RNA was used as a template control. (B) Protein from each fraction was analysed by immunoblotting using an anti-CA antibody.

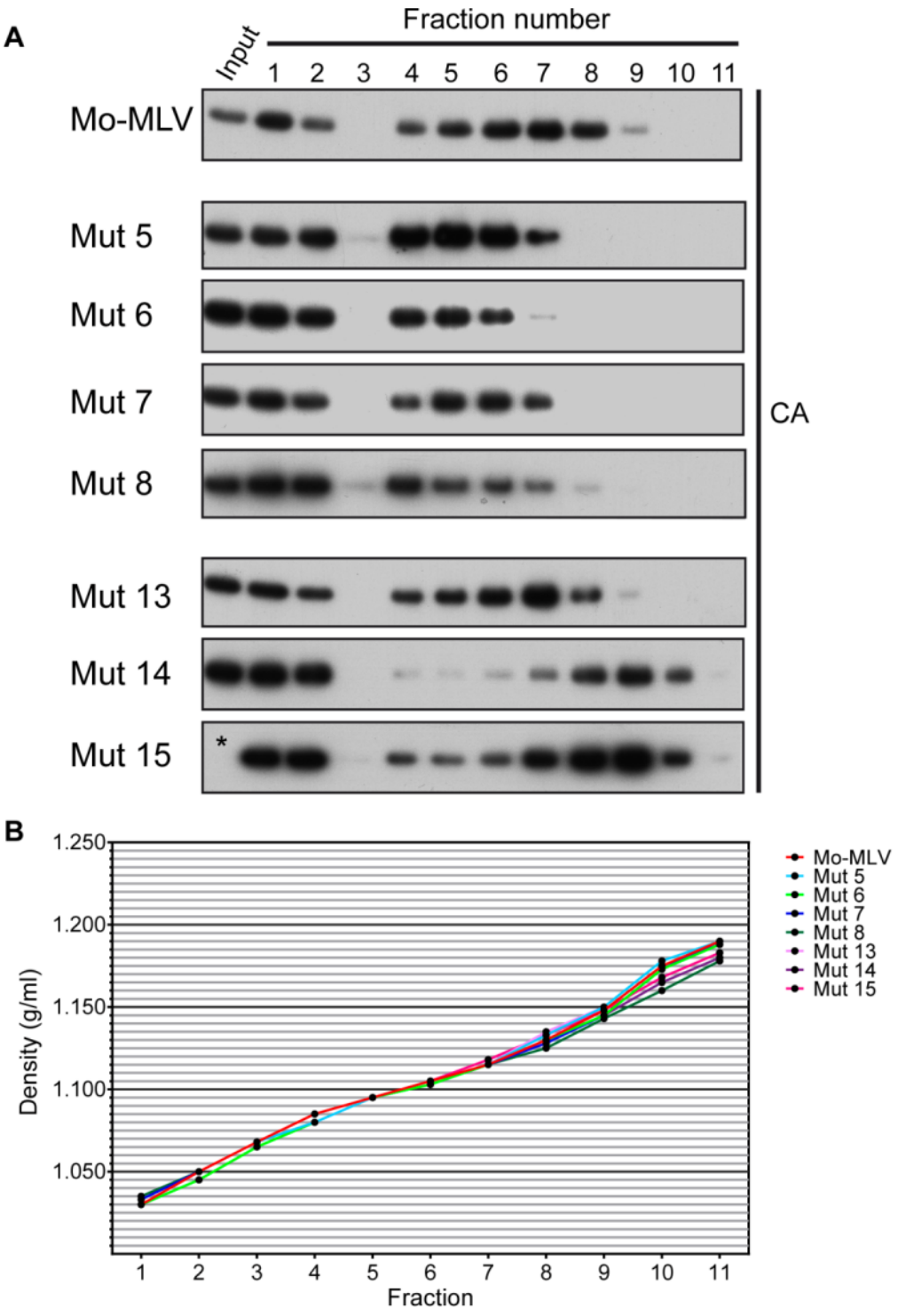


Figure 5.15: The effect of p12 mutations on the CA sedimentation profile in the triton spin through assay. Concentrated wild type or p12 mutant VLPs were layered onto triton spin through gradients. (A) Proteins were analysed by immunoblotting with an anti-CA antibody. Films were scanned with an Epson V700 scanner and representative western blots for wild type Mo-MLV and all the p12 mutants are shown. (B) A sample from each fraction was used to measure the refractive index to determine the sucrose densities. The graph shows the density profile of the all the gradients from (A). * input band missing due to over dilution of the sample.

To quantify the data from multiple experiments, the change in sucrose density of the mature CA peak fraction, between the p12 mutants and wild type, was calculated (Figure 5.16). This method standardises the difference observed for the p12 mutants in every experiment to the internal wild type Mo-MLV control. As could be seen in the representative blots (Figure 5.15A), the N-terminal p12 mutants have a sedimentation profile that was shifted to a lower sucrose density than wild type (Figure 5.16, light and dark, green and blue bars). Whereas, the C-terminal p12 mutants (barring mutant 13), all had a sedimentation profile that was shifted to a higher sucrose density (Figure 5.16, pink, purple and magenta bars).

Therefore, this assay showed that the N- and C-terminal p12 mutants had differing phenotypes. The lower sedimentation profile of the mature CA peak for the N-terminal p12 mutants suggested either: (i) they lacked a mature core within the virion or (ii) a less stable core was formed, which broke into smaller parts after removal of the virion membrane. Interestingly, two of the C-terminal mutants showed the opposite effect (mutant 14 and 15), indicating that the core could be more stable in these viruses.

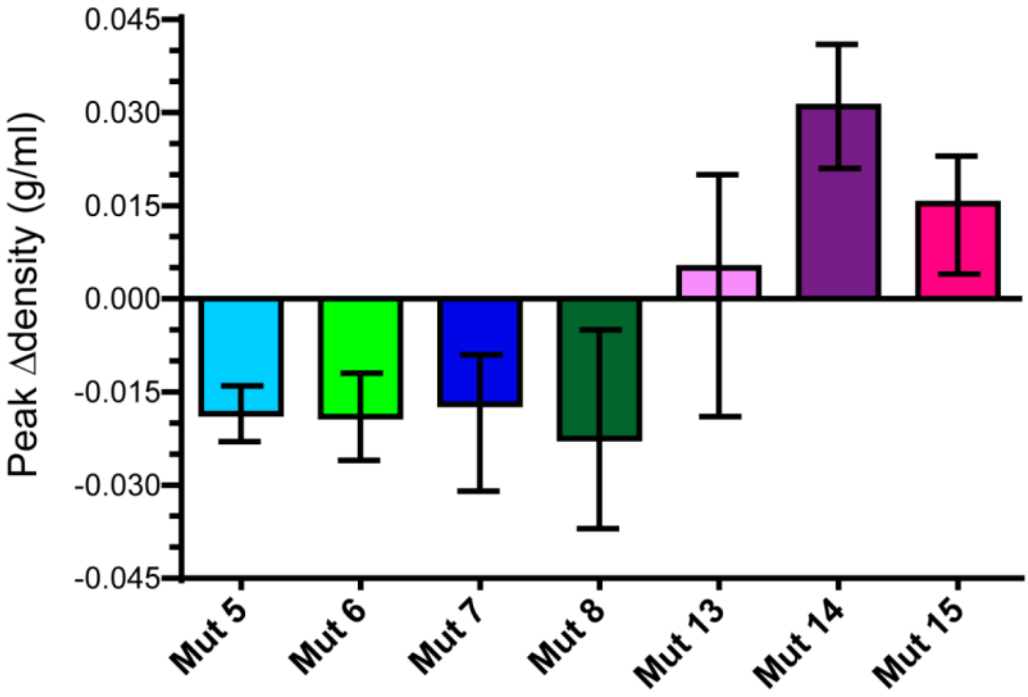


Figure 5.16: Quantification of the effect of p12 mutations on the CA sedimentation profile in the detergent layer gradient. The sucrose density of the CA peak, after fraction 4, was determined for each p12 mutant and compared to that of the wild type control run in the same experiment. The mean density difference between each p12 mutant and wild type are shown in the graph. The error bars indicate the range.

5.2.3 Ultrastructure of the Mo-MLV core within the virion

One reason why mutation of the N-terminal domain of p12 could result in a lack of abrogation of restriction factors, and cause the sedimentation profile of the partial cores to change, is if a mature CA core was not formed. To visualise if mutation of the N-terminus of p12 resulted in formation of an aberrant core, cell-free virus was prepared for thin-section transmission electron microscopy (TEM). The p12 mutant 6 was chosen as it displays the most severe form of the N-terminal domain defect (reduced reverse transcription *in vivo*). Wild type and p12 mutant 6 VLP preparations, used for TEM, displayed the expected phenotypes (Figure 5.17).

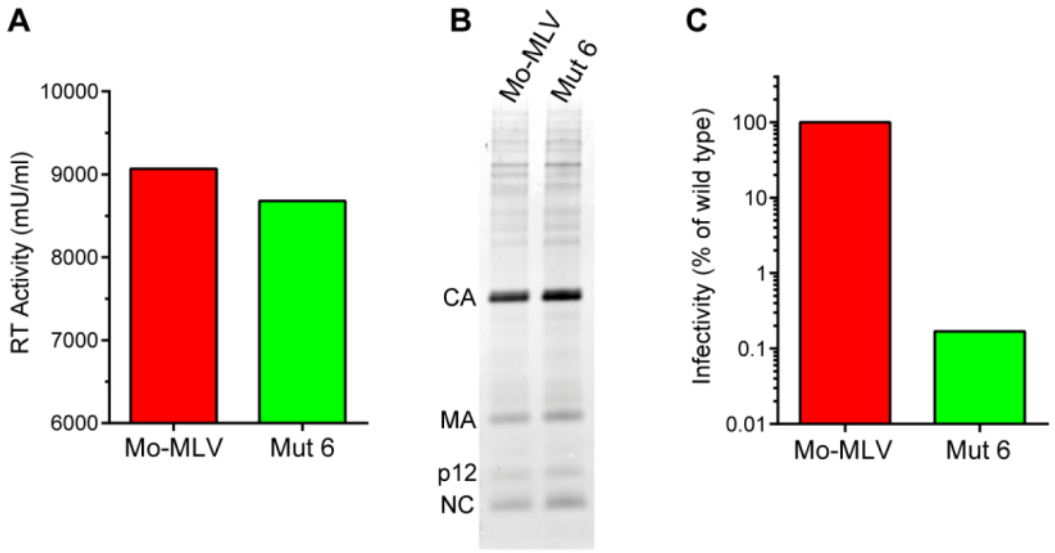


Figure 5.17: Characterisation of Mo-MLV VLPs for thin-section transmission electron microscopy. Sixty 10cm plates of 293T cells were transfected to produce either wild type or p12 mutant 6 Mo-MLV VLPs. The harvested VLPs were pooled and concentrated by pelleting in an ultracentrifuge. (A) Viral release was quantified using the RT-ELISA and is displayed as the RT activity mU/ml. (B) Viral proteins were separated by SDS-PAGE and the proteins stained with SimplyBlue stain. The gel was scanned using an Epson V700 scanner. (C) D17 cells were infected with equal RT-units of the VLPs and the infectivity was quantified. The infectivity is shown as the percentage of wild type.

VLP pellets were fixed, embedded in agar and thin 50nm sections were cut and stained on grids. Micrographs were taken at 20,000x magnification using a JOEL 1200EX TEM equipped with a CCD camera. Representative electron micrographs from wild type and p12 mutant 6 VLPs are shown adjacent to each other for comparison (Figure 5.18 A and

B). Looking at the wild type micrographs, an array of heterogeneity in VLP size and shapes were seen (Figure 5.18A). Importantly, the majority of the particles seen in these micrographs contained the characteristic electron dense core of a mature retroviral particle (Figure 5.18A) (Auerbach, Brown, and Kaplan, 2006; Auerbach, Brown, and Singh, 2007; Muriaux et al., 2004). In the micrographs of p12 mutant 6 VLPs, a few particles were seen containing a mature core (examples with green arrowheads), while the vast majority either lacked a core (examples with pink arrowheads) or contained some electron dense material that did not constitute an immature or mature core (examples with yellow arrowheads) (Figure 5.18B). This striking result highlighted p12 mutant 6 has a defect in the formation of a mature core.

To quantify this defect, the diameter of particles was measured and particles between 80-120nm in diameter were assigned a core morphology using a set of examples (Figure 5.19A). For wild type Mo-MLV particles, the majority (78.26%) displayed a mature core morphology, whereas only 12.50%, of the p12 mutant 6 virions contained a mature core phenotype (Figure 5.19B, green bars). The majority (56.85%) of the p12 mutant 6 virions displayed an aberrant core morphology (Figure 5.19B, red bar). Only a few wild type virions containing aberrant cores could be observed (9.13%) (Figure 5.19B, red bar). Similarly, for virions lacking a core, wild type had a minor proportion in this category (11.74%), whereas p12 mutant 6 had significantly more virions without a visible core (28.23%) (Figure 5.19B, orange bars). The remaining virions were scored as containing an immature core morphology (Figure 5.19B, blue bar). Therefore, the defect with p12 mutant 6 results in the inability to form a mature core, with 85.08% of the virions either lacking a core or displaying some type of aberrant core morphology. Thus, the N-terminal domain of p12 was involved in, but possibly not limited to, the efficient assembly of the mature CA core within the virion.

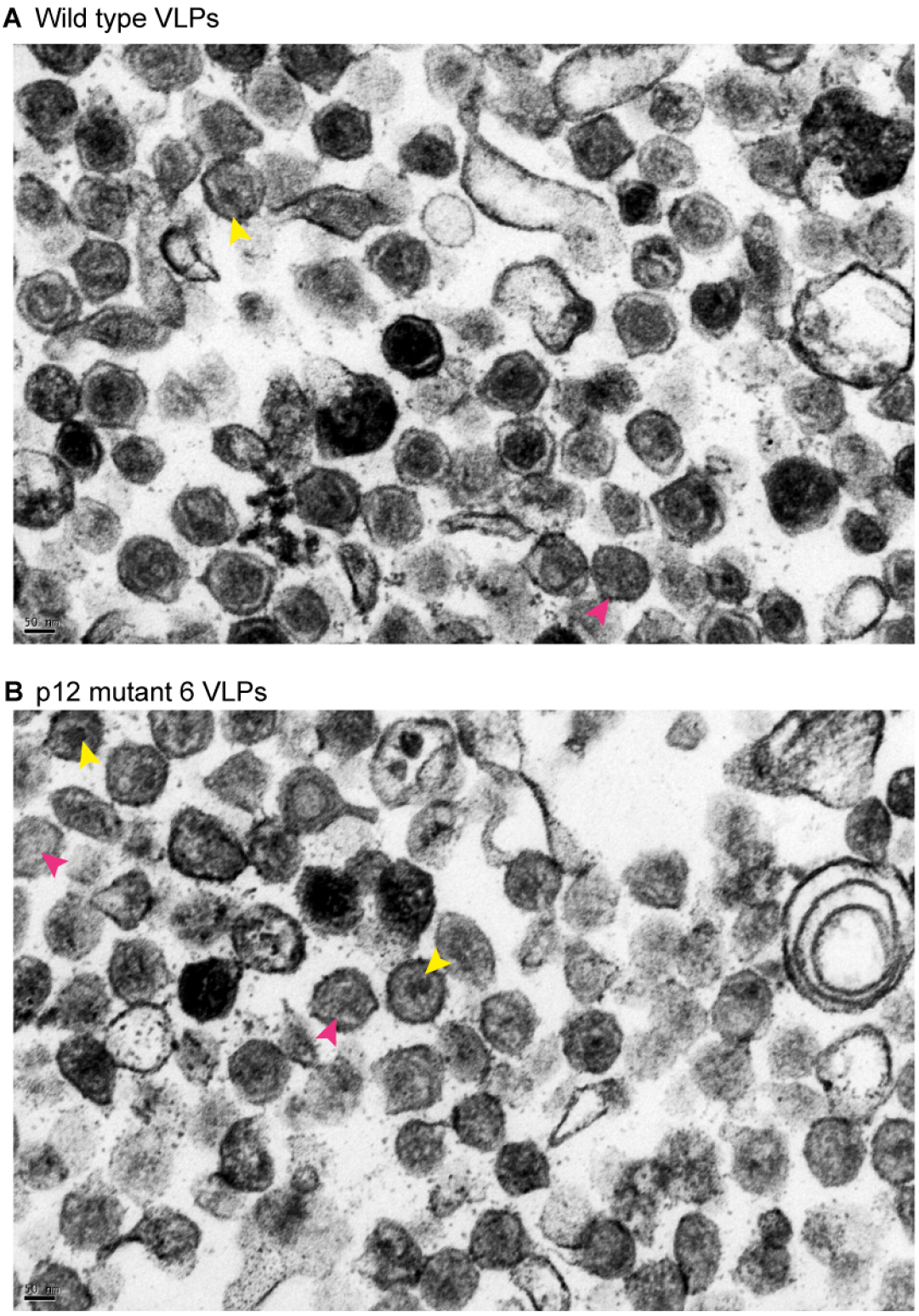
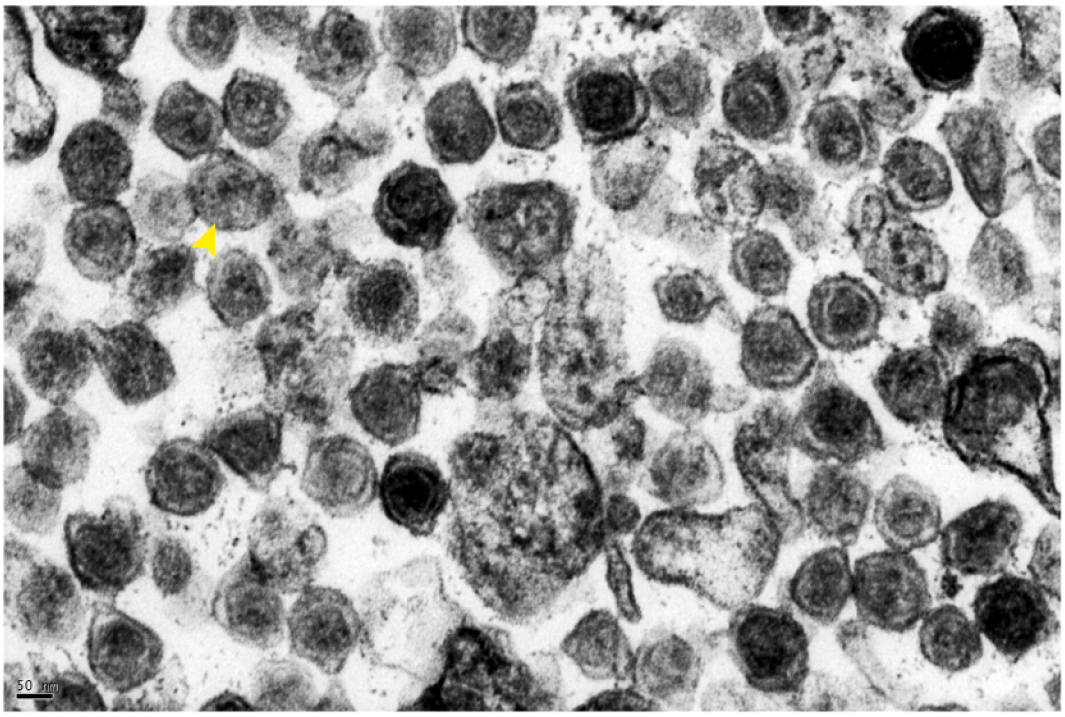
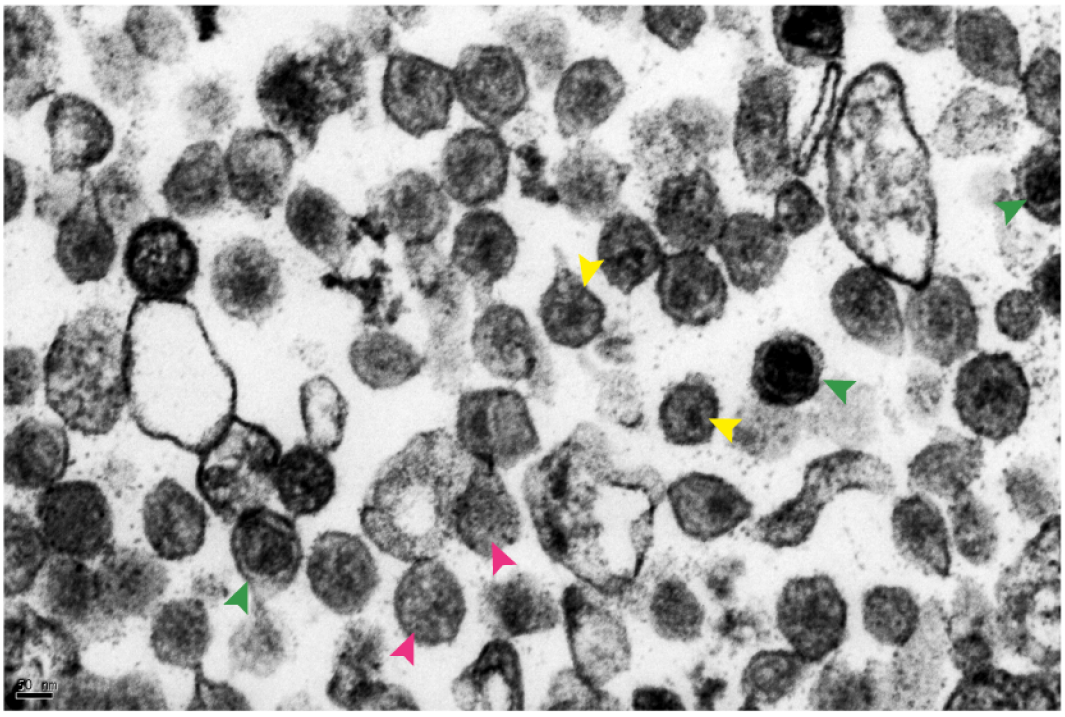


Figure 5.18: Ultrastructure of p12 mutant Mo-MLV VLPs by thin-section transmission electron microscopy. Wild type or p12 mutant 6 VLPs were concentrated and prepared for thin section electron microscopy. Micrographs were taken at 20,000x magnification using a JOEL 1200EX TEM equipped with a Gatan Orius CCD camera. Four representative micrographs of (A) wild type VLPs or (B) p12 mutant 6 VLPs are shown. Arrowheads indicate examples of: wild type-like VLPs in p12 mutant 6 micrographs (green arrowheads), VLPs lacking an identifiable core (pink arrowheads) or an aberrant core (yellow arrowheads). Scale bars are 50nm.

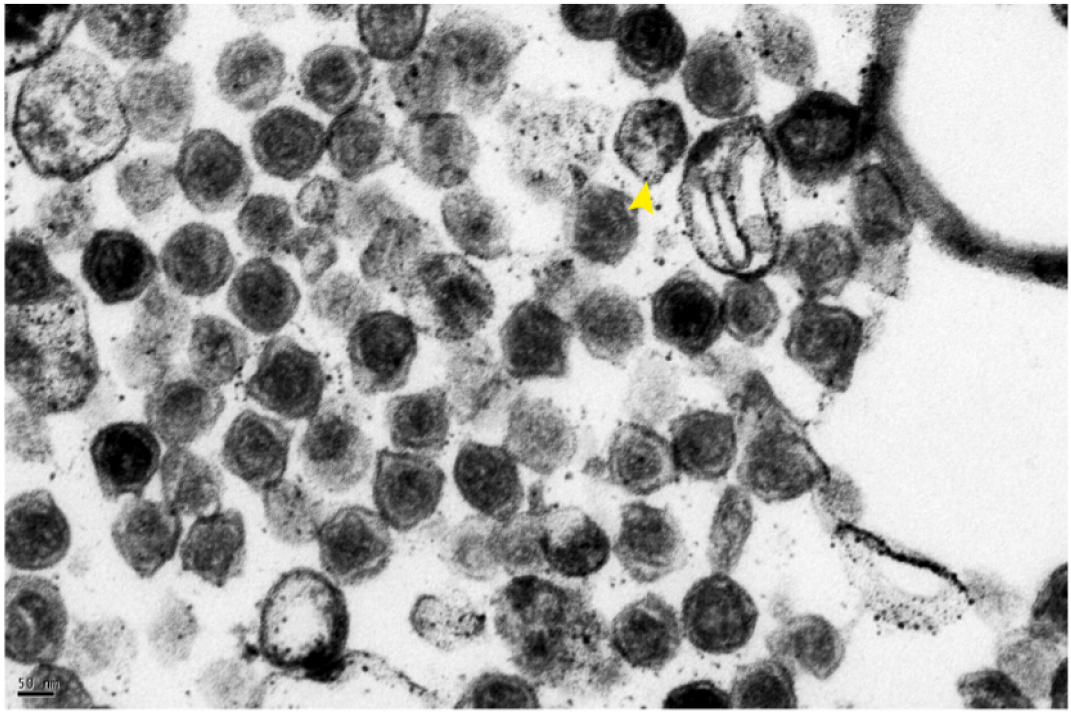
A Wild type VLPs



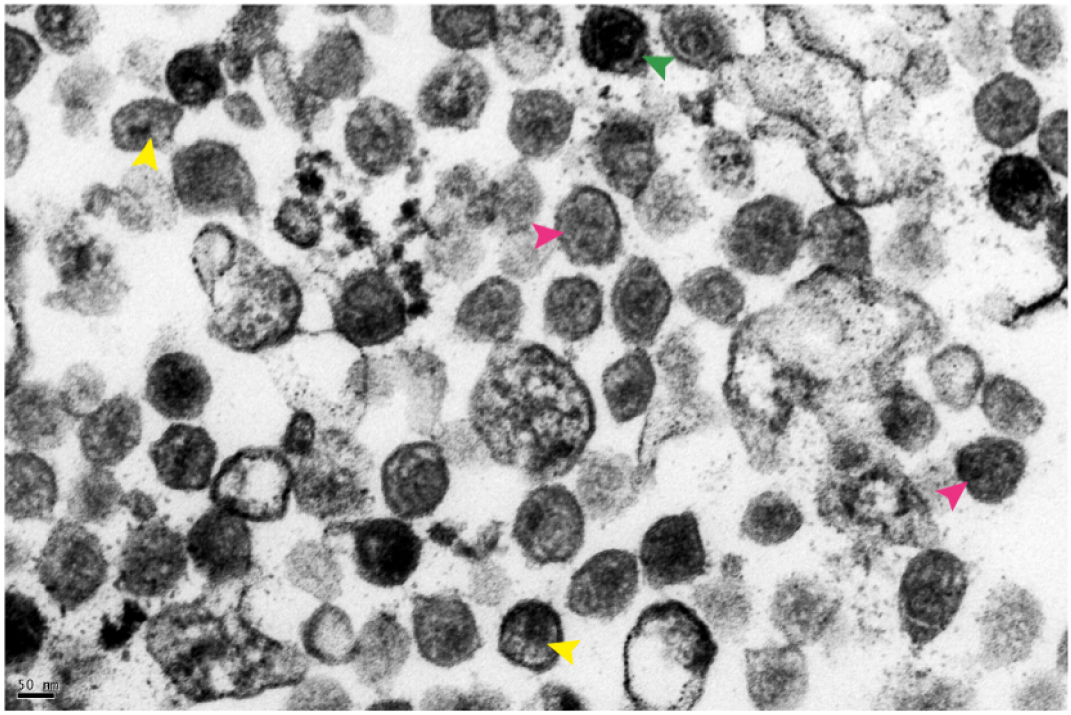
B p12 mutant 6 VLPs



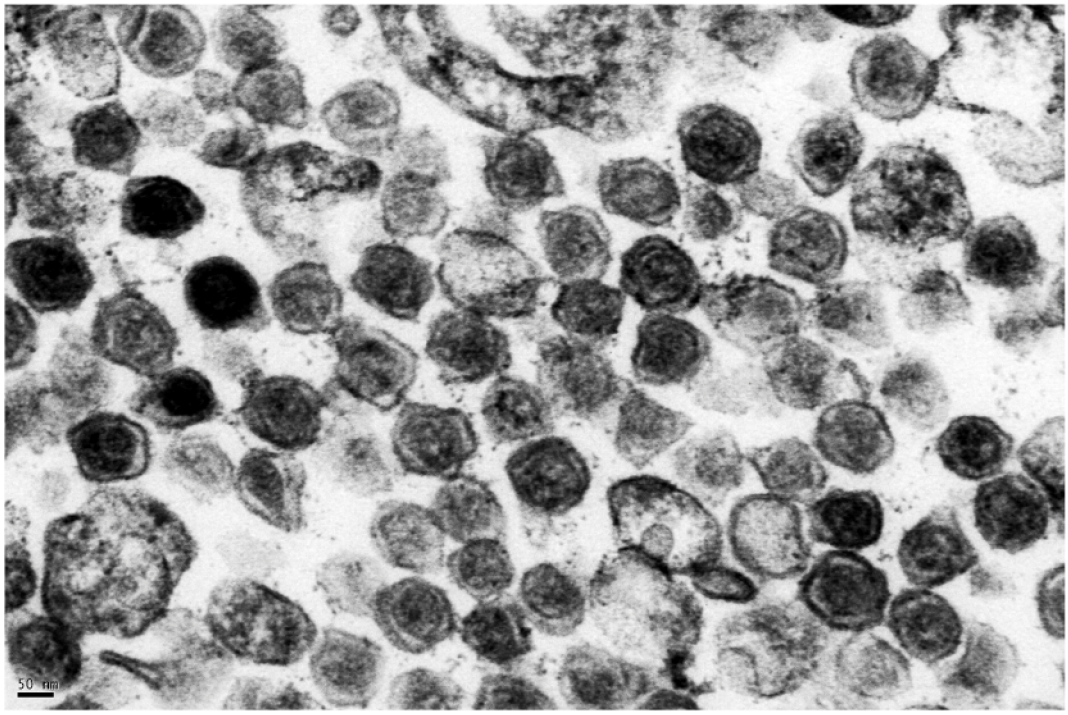
A Wild type VLPs



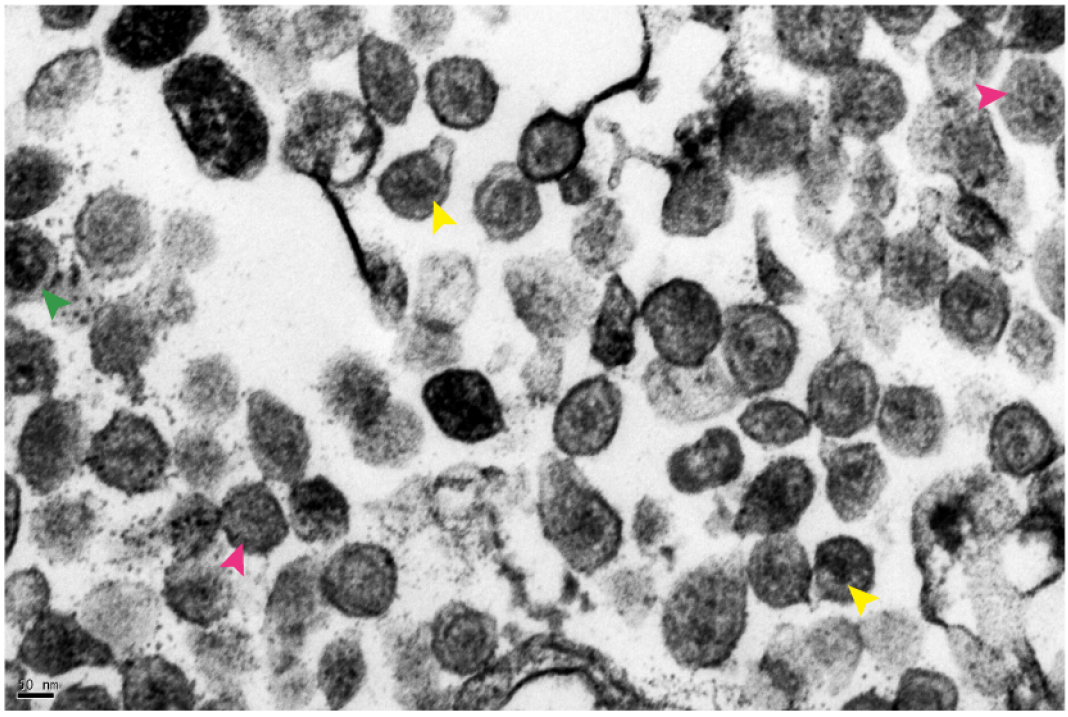
B p12 mutant 6 VLPs



A Wild type VLPs



B p12 mutant 6 VLPs



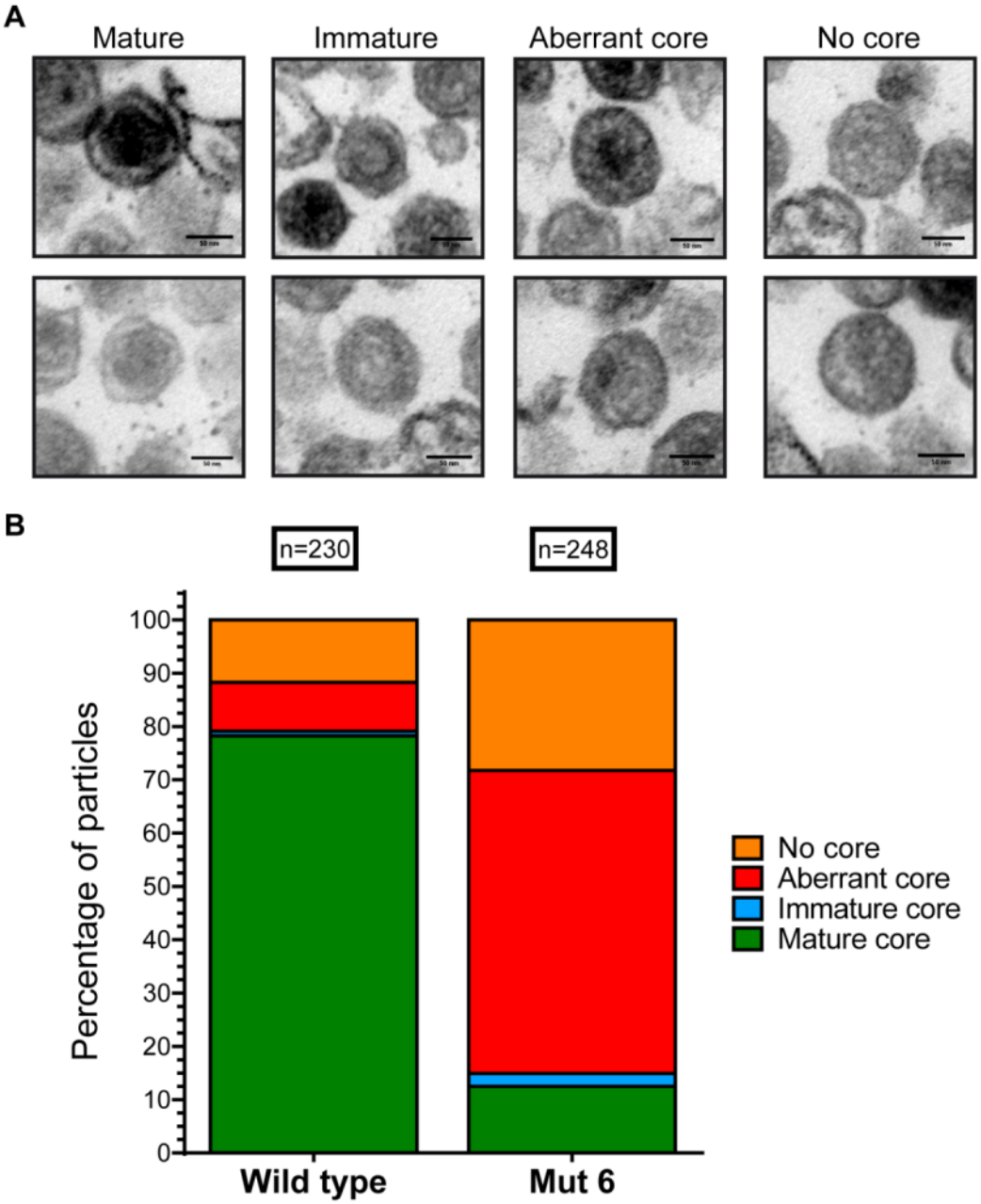


Figure 5.19: Quantification of p12 mutant Mo-MLV VLP core morphology by thin-section transmission electron microscopy. Micrographs of wild type or p12 mutant 6 VLPs were analysed and the core morphology of the virions (between 80-120nm in diameter) categorised. (A) Representative images showing the core morphology from each morphological category. (B) Quantification of the core morphology data from wild type and p12 mutant 6 virions.

5.3 Investigations into the function of the C-terminal domain of p12

As mutation of the C-terminal domain of p12 does not affect the ability of the virus to abrogate restriction, it can be assumed that the Mo-MLV core in these mutants is correctly formed. The intracellular PICs from cells infected with Mo-MLV containing p12 mutant 14 were biochemically similar to wild type, when analysed by equilibrium and velocity sedimentation (Yuan et al., 2002). Additionally, a later study using immunofluorescence, has highlighted that p12 could be seen on the condensed mitotic chromosomes during infection, a phenomenon that was not seen for p12 mutant 14 (Prizan-Ravid et al., 2010). Thus, genetic assays aimed at rescuing the C-terminal p12 mutant defect were performed.

5.3.1 The effect of nuclear localisation motifs in p12 on Mo-MLV viability

A few retroviral proteins have been described to have karyophilic motifs, essential for replication. One type of karyophilic motif is the nuclear (NLS) and nucleolar localisation motifs (NoLS), which are highly basic motifs directing nuclear trafficking of many proteins (Emmott and Hiscox, 2009; Lange et al., 2007). It has been shown that a motif based around essential arginines in the C-terminus of p12 was necessary for function. Could this motif in p12 have a role in nuclear localisation? However, barring a few examples, simple retroviruses like the gammaretroviruses only access the chromosomes during mitosis, thus if p12 does contain a karyophilic signal it is unlikely to direct active transport into the nucleus. (Jarrosson-Wuilleme et al., 2006; Liu et al., 2011; Roe et al., 1993).

To test if the C-terminal domain of p12 plays a karyophilic role during the early stages of infection, using the *Asc*I cloned into p12 (described in Chapter 3) an NLS and two NoLS were cloned into p12 of the Mo-MLV Gag-Pol vector. The release of wild type Mo-MLV VLPs, with or without the modifications, were quantified by the level of RT activity released (Figure 5.20A). Addition of the prototypic NLS from the simian vacuolating virus 40 (SV40) large T-antigen was significantly detrimental to wild type virus release, resulting in a 44-fold reduction in release (Figure 5.20A). The angiogenin NoLS and NF-kB inducing kinase (NIK) NoLS motifs in wild type Mo-MLV p12 also

had an effect on VLP release, but to a smaller degree than the SV40 T-antigen NLS (2.4-fold and 4.6-fold reductions, respectively) (Figure 5.20A). The infectivity of p12 mutant VLPs carrying these NoLS motifs in p12 was assessed. Insertion of the angiogenin NoLS into wild type p12 had only a minor, 1.8-fold, effect on infectivity (Figure 5.20B). Whereas, the NIK NoLS reduced the infectivity of wild type Mo-MLV by 145-fold (Figure 5.20B). The angiogenin NoLS had little effect on the infectivity of the p12 mutants, with no effect seen on the infectivity of the two N-terminal p12 mutants (Figure 5.20B). The C-terminal p12 mutant 14, had a further reduction in infectivity when carrying the angiogenin NoLS of 8.3-fold compared to p12 mutant 14 without this NoLS (Figure 5.20B). This infectivity was almost down to background levels, suggesting that this motif was actively taking p12 (and the PIC) away from where it needed to go during the early stages of infection. The fact that it did not affect wild type Mo-MLV infectivity, suggested that the motif in the C-terminal domain of p12 was dominant over this NoLS motif.

Interestingly, the NIK NoLS in p12 had the opposite effect on the infectivity of Mo-MLV. It reduced the infectivity of wild type Mo-MLV (145-fold), p12 mutant 6 (24-fold) and mutant 7 (22-fold), but increased the infectivity of p12 mutant 14 by 18-fold (Figure 5.20B). Therefore the effect of this motif was dominant over the C-terminal domain of p12, directing wild type Mo-MLV to a location not ideally suited for productive infection. However, this nuclear/nucleolar location could improve the chances of mutant 14 to integrate (up to the infectivity of wild type with the NIK NoLS). Thus, localisation of p12 to the nucleolus, while detrimental to wild type p12 function, could enhance the chances that p12 mutant 14 would integrate its viral DNA into the host cell chromosomes. However, it should be noted that the nuclear localisation of these viruses has not yet been tested.

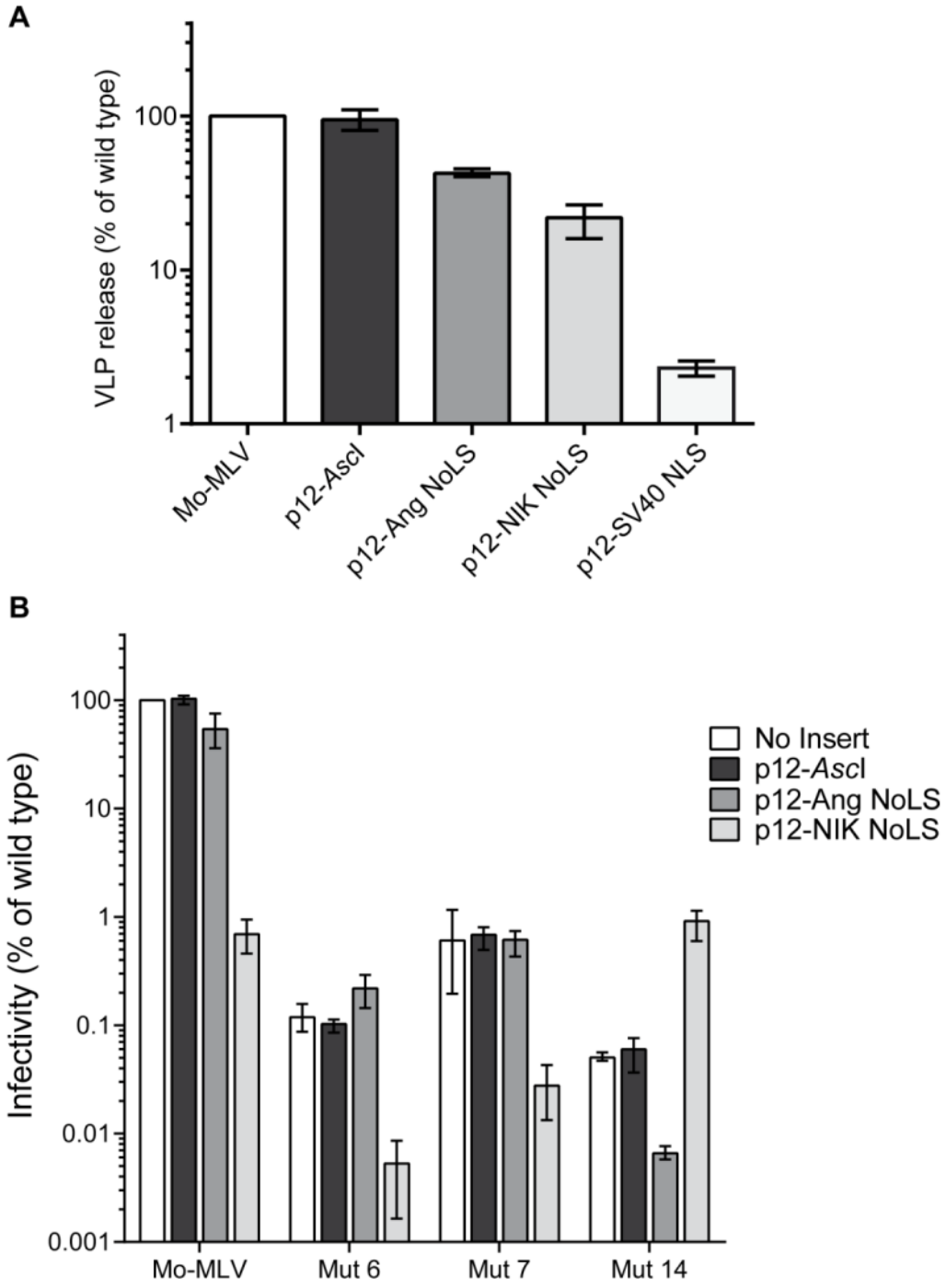


Figure 5.20: The effect of p12 containing different nuclear localisation signals on VLP release and infectivity. Using the *Asc*I restriction site introduced into p12 of the Mo-MLV Gag-Pol expression vector (p12-*Asc*I), different nuclear localisation motifs were cloned into p12. (A) RT activity was measured using the RT-ELISA and VLP release is shown as a percentage of the wild type VLP RT activity. The bars show the mean and range of three independent experiments. (B) The infectivity was quantified and is expressed as a percentage of wild type infectivity. The bars indicate the mean and range of three independent experiments. Ang NoLS = Angiogenin nucleolar localisation signal (IMRRRGL), NIK NoLS = NF-kB inducing kinase nucleolar localisation signal (RKKRKKK) and SV40 NLS = simian vacuolating virus 40 large T-antigen nuclear localisation signal (PKKKRRV).

5.3.2 The chromatin tethering function of the C-terminus of p12

The HIV-1 PIC is tethered to the host chromosomes by interactions with lens epithelium-derived growth factor (LEDGF) through IN (Cherepanov et al., 2003; Maertens et al., 2003). This interaction is essential for efficient HIV-1 integration. Based on the fact that p12 mutant 14 could not localise to condensed chromatin in infected cells (Prizan-Ravid et al., 2010), and a minor infectivity rescue could be achieved by addition of an NoLS, it was possible that p12 could be a LEDGF-like chromatin tether molecule for MLV.

A portion of the prototypic foamy virus (PFV) Gag, in the glycine-arginine-rich box II (GRII), is responsible for chromatin tethering of expressed PFV Gag, and of the incoming PFV PICs (Mullers et al., 2011; Tobaly-Tapiero et al., 2008). To see if the C-terminal domain of Mo-MLV p12 has a chromatin tethering function, the PFV chromatin binding sequence (CBS) was cloned into p12. Wild type and p12 mutant Mo-MLV VLPs, with or without the PFV CBS in p12, all produced similar levels of VLPs (none had greater than a two-fold defect in VLP production) (Figure 5.21A). Inclusion of the PFV CBS in wild type p12 did, however, reduce the infectivity of Mo-MLV by 16-fold (Figure 5.21B). Therefore, the PFV CBS targeted p12 (and the PIC) to the chromatin in a manner that was sub-optimal for integration. Inclusion of the PFV CBS in the N-terminal p12 mutants, mutant 6 or 7, did not have an effect on infectivity (Figure 5.21B). However, when the PFV CBS was inserted into any of the C-terminal p12 mutants, the infectivity was increased by greater than 100-fold (Figure 5.21B). In conclusion, this represented a significant rescue of the defect imposed by the mutation, indicating the C-terminal domain of p12 has a chromatin tethering function for the incoming PICs.

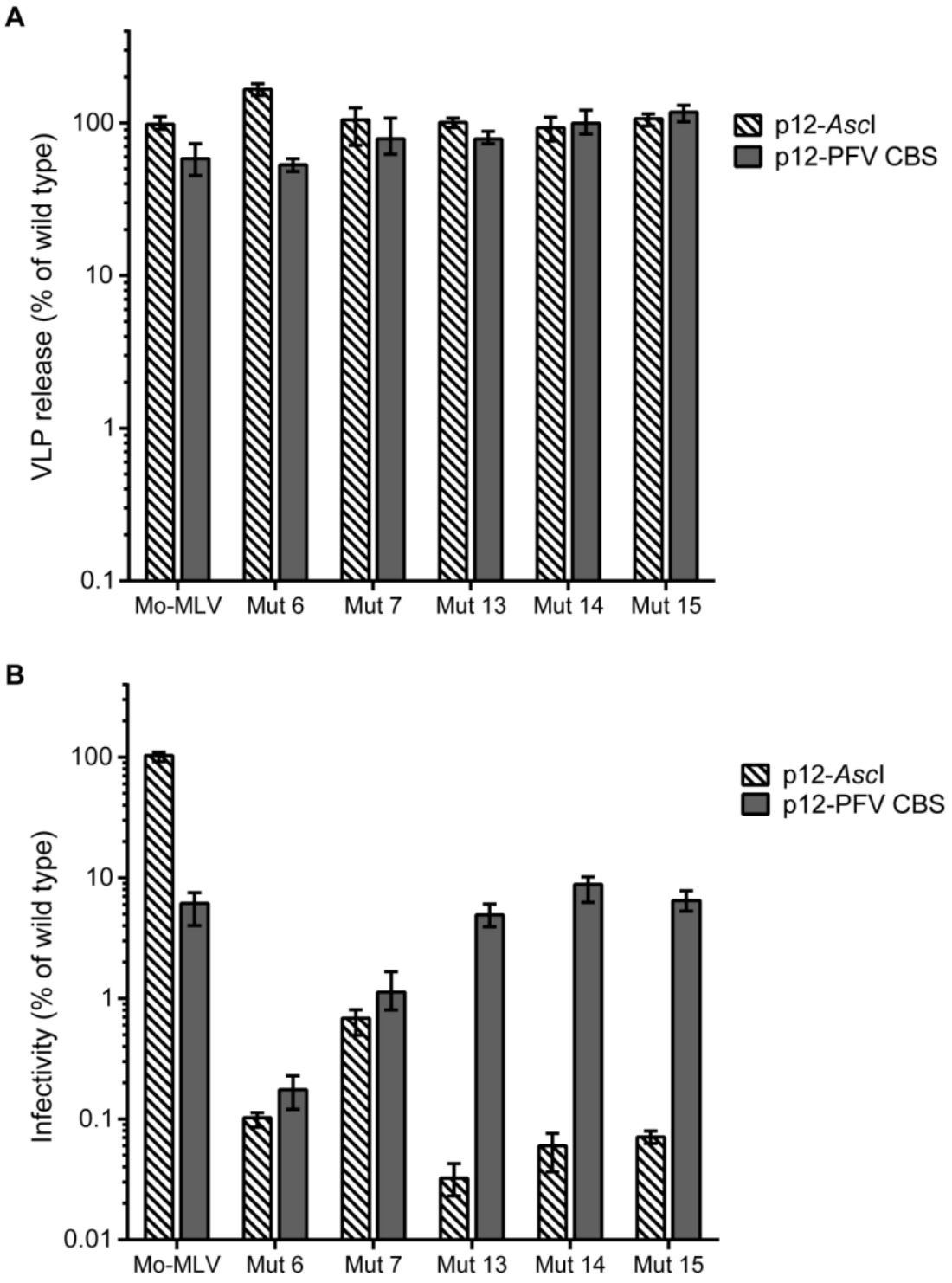


Figure 5.21: The effect on Mo-MLV of p12 containing the PFV CBS. Using the *Asc*I restriction site introduced into p12 of the Mo-MLV Gag-Pol expression vector (p12-*Asc*I), the PFV CBS (NQGGYNLRPRTYQP) was cloned into p12 (p12-PFV CBS). (A) RT activity was measured using the RT-ELISA and VLP release is shown as a percentage of the wild type VLP RT activity. The bars show the mean and range of three independent experiments. (B) The infectivity was quantified and is expressed as a percentage of wild type infectivity. The bars indicate the mean and range of three independent experiments.

5.3.3 The effect of p12 on viral DNA integration site selection

In addition to chromatin binding, LEDGF influences integration targeting for lentiviruses (Cherepanov, 2007; Ciuffi et al., 2005; Llano et al., 2006). MLV, similar to HIV-1, does not integrate randomly into the genome and MLV has a preference for transcriptional start sites (Mitchell et al., 2004). However, unlike HIV-1, no factors involved in integration targeting for simple retroviruses have been described.

To assess if p12 influences the integration targeting of MLV through p12's chromatin tethering function a panel of modified Mo-MLVs were synthesised, and titrated on HeLa cells (Figure 5.22A). HeLa cells were infected with a MOI~1.5 (less for p12 mutant 6 and 14) and the total DNA was isolated. The integration sites were then sequenced in the Bushman laboratory. After filtering out sequences that were non-unique and too short for annotation the unique integration sites (Figure 5.22B) were aligned to the human genome.

Using the genome annotation RefGene, the distance of the integration sites to the nearest gene were analysed (Figure 5.22C). For wild type Mo-MLV an enrichment of integration sites, located close to the start site of genes was found, as previously reported (Figure 5.22C) (Cattoglio et al., 2010; Mitchell et al., 2004). The same integration pattern was also seen for the Mo-MLV/Ga-Cp12 chimera and the Mo-MLV with the p12-PFV-CBS, with and without the p12 mutation 14 (Figure 5.22C). While p12 mutant 6 also followed the same pattern as wild type, p12 mutant 14 did not. Mutant 14 displayed a wider integration selection pattern, with some integration sites a considerable distance from a transcriptional start site (Figure 5.22C). Similarly, another genomic feature favoured for integration by Mo-MLV are CpG islands (Cattoglio et al., 2010). All of the viruses tested had a similar tendency to integrate near CpG islands, with the exception of p12 mutant 14, which displayed a slightly reduced tendency to integrate near CpG island midpoints (Figure 5.22D).

Therefore, targeting the Mo-MLV PIC to chromatin using MLV p12, the C-terminus of GaLV p12 or the PFV CBS did not influence the integration pattern of Mo-MLV. However interestingly, this data has suggested the loss of the C-terminus of p12 does have an influence on integration site selection, although the number of integration sites analysed for the p12 mutants is extremely low (due to their infectivity defect).

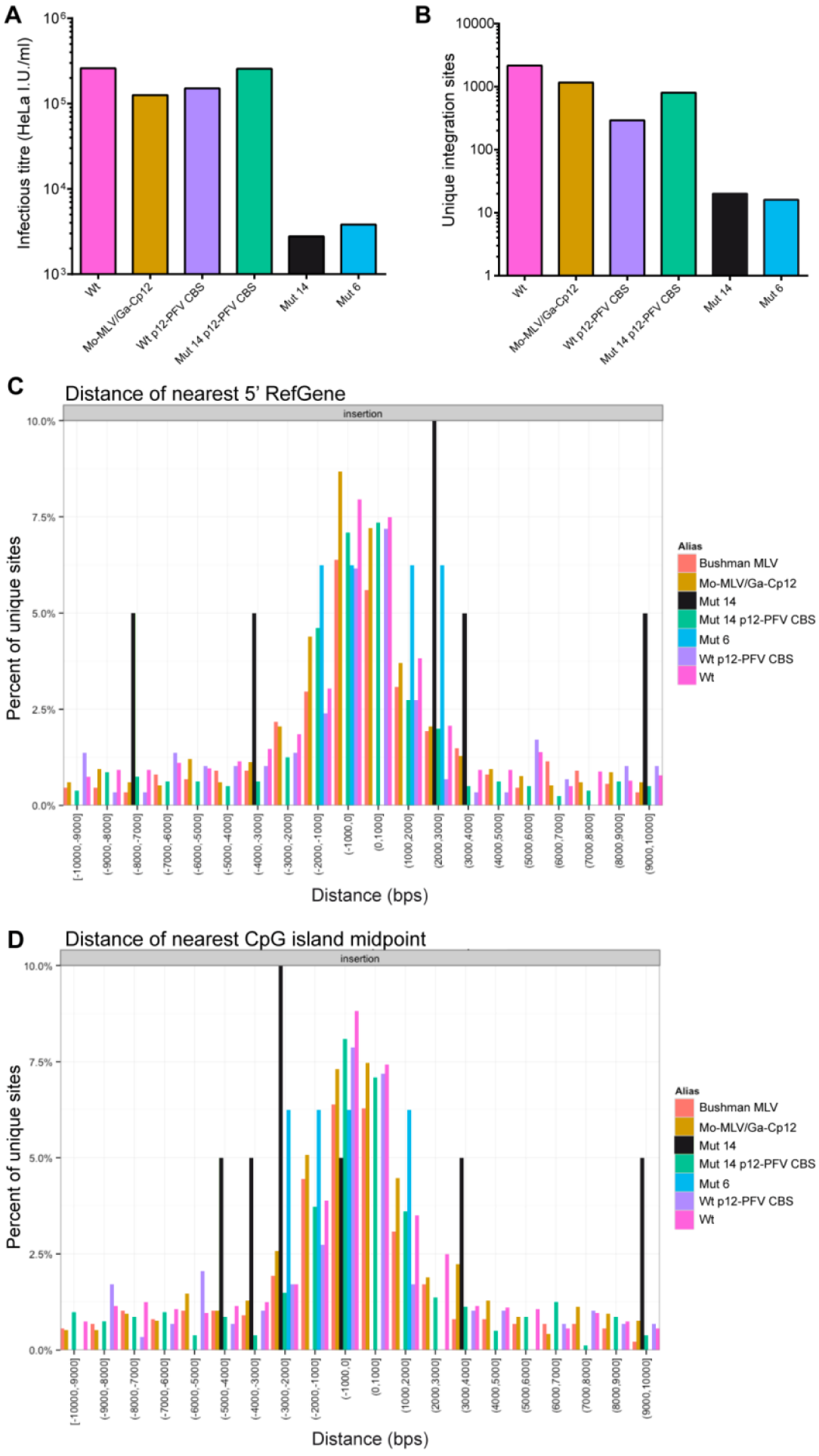

A
Infectious titre (HeLa I.U./ml)
Wt
Mo-MLV/Ga-Cp12
Wt p12-PFV CBS
Mut 14 p12-PFV CBS
Mut 14
Mut 6
B
Unique integration sites
C
Distance of nearest 5' RefGene
insertion
Percent of unique sites
Distance (bps)
Alias
Bushman MLV
Mo-MLV/Ga-Cp12
Mut 14
Mut 14 p12-PFV CBS
Mut 6
Wt p12-PFV CBS
Wt
D
Distance of nearest CpG island midpoint

Figure 5.22: The influence of p12 on integration site selection of Mo-MLV. (A) Infectious titre of viruses in HeLa cells (note- all VLPs except wild type were concentrated ten times). HeLa cells were infected at an approximate MOI~1.5, total DNA was isolated and sent to the Bushman laboratory for integration site sequencing. (B) Number of unique integration sites sequenced from the infected HeLa cells. (C) Distance of the integration sites for each virus to the 5' end of genes (determined using the RefGene annotation). (D) Distance of the integration sites for each virus to the nearest CpG island midpoint. Integration site sequencing and annotation was performed in the Bushman laboratory. Wt- wild type Mo-MLV, Mut 6- Mo-MLV p12 mutant 6 and Mut 14- Mo-MLV p12 mutant 14.

5.4 Summary

5.4.1 The function of the N-terminal domain of p12

Mutations in the N-terminal domain of p12 affected the ability of Fv1 and TRIM5alpha to interact with the CA core during infection. This implied that the CA core of the N-terminal p12 mutants was either (i) blocked from the restriction factor by the mutant p12 (ii) mislocalised within the infected cell, (iii) processed incorrectly and thus not recognisable by the restriction factors, (iv) unstable or (v) no recognisable core was present to be acted on by the restriction factors.

Firstly, it is unlikely that the N-terminal p12 mutants block restriction factor binding to the core, as it would be hard to explain why this would cause an infectivity defect. Secondly, it is also unlikely that the CA core of these p12 mutants are mislocalised within infected cells, as no specific sub-cellular localisation of retroviruses or restriction factors have been described.

Using mixed particles it was shown that the CA from *gag* that contained an N-terminal p12 mutation was correctly processed and could be identified by restriction factors. This indicated that p12 mutations *in cis* to CA did not alter CA processing or inclusion as part of the mature core. To assess the stability of the Mo-MLV p12 mutant core, isolation of MLV cores was attempted by centrifugation of VLPs through detergent onto a sucrose gradient. This method has allowed isolation of HIV-1 cores and the stability of CA mutant cores to be analysed (Forshey et al., 2002; Kotov et al., 1999). Unfortunately, whole cores were not isolated by this method, but a species of partial core, devoid of the genomic RNA, was isolated in the gradient. Importantly, the position of these partial cores within the gradient was reproducible between experiments, and were more buoyant than un-lysed VLPs. The N-terminal p12 mutants all migrated to a lower sucrose density than wild type, indicating that the core was either less stable or

poorly formed in these mutants. Interestingly, two of the C-terminal mutants (14 and 15) migrated to a sucrose density that was higher than wild type.

To assess the morphology of the mature core inside the N-terminal p12 mutants, thin section TEM was performed on purified mutant 6 virions. Perhaps surprisingly, only a minor proportion of p12 mutant 6 virions contained a morphologically mature core. The majority had aberrant core morphologies, and a sizeable proportion contained no visible cores. This hinted towards the involvement of p12 in the efficient formation of the mature CA core during maturation.

For p12 to have an effect on the CA core through its N-terminal domain, it would be expected to bind to some core-associated factor. No long-range interactions were reported between MLV p12 and the N-terminus of CA in an expressed fusion protein, but p12 and CA have been functionally linked using chimeric viruses (Kyere, Joseph, and Summers, 2008; Lee and Nagashima, 2005). Taken together with the chimeric data from chapter 4, it is possible that the N-terminal domain of p12 actually interacts, directly or indirectly, with the forming CA core within the virion.

Another point to consider is that the point mutant L16A, which significantly knocked down the infectivity of Mo-MLV, could partially abrogate restriction. This implied that the CA core of this p12 point mutant is actually present in the infected cells. The L16A mutant was 2.8-fold more infectious than p12 mutant 6 and the partial abrogation could reflect an increase in virions with mature cores formed by the L16A mutant. However, the abrogation and infectivity defects do not match, indicating that mutation of the N-terminus of p12 may cause another defect in addition to malformation of the mature core.

Finally, completely blocking the cleavage of p12 from MA had a very minor effect on infectivity of Mo-MLV, consistent with previous reports (Oshima et al., 2004). This suggested that either MA-p12 or more likely one of the Gag processing intermediates could carry out the early function of p12. In mixed particles, the dominant negative effect of the C-terminal p12 mutants was lost by blocking the processing of the dominant negative p12 molecule from MA. This showed that in the wild type virion the cleaved form of p12 was better suited to carry out the N-terminal domain interaction, and that this interaction occurred or was activated after MA-p12 cleavage.

5.4.2 The function of the C-terminal domain of p12

Localisation of p12 (and the PIC) to the nucleus via the NIK NoLS motif increased the infectivity of p12 mutant 14 by greater than 10-fold, with the caveat that this motif was toxic to wild type p12 function. Another NoLS, from angiogenin, was unable to partially rescue the p12 mutant 14 infectivity defect. The difference between the NIK NoLS and the angiogenin NoLS, is that the former is composed of two components, an NLS and an NoLS (Birbach et al., 2004). Therefore the NIK NoLS can shuttle between the nuclear and nucleolar compartments, potentially increasing the chances that the MLV PIC will interact with the chromosomes. The angiogenin NoLS is likely inactive in the context of the MLV PIC as it is unable to confer nucleolar localisation to β-galactosidase, but it can localise the smaller GFP molecule to the nucleolar compartment (Lixin et al., 2001).

Addition of a chromatin binding sequence, from PFV, into the C-terminal p12 mutants was able to significantly rescue the infectivity of Mo-MLV by at least 100-fold. This motif also had an effect on the infectivity of wild type Mo-MLV (>10-fold reduction), indicating that the location of the chromatin binding is sub-optimal for integration of Mo-MLV. Live imaging experiments have shown that the Kaposi's sarcoma-associated herpes virus (KSHV) LANA CBS rescues the chromatin binding of p12 mutant 14 (and partly rescues the infectivity defect) in infected mitotic cells, although it remains more tightly associated with the chromatin than wild type p12 after mitosis (Elis et al., 2012). The use of the PFV chromatin binding sequence or the C-terminal domain of GaLV p12 was unable to change the integration targeting of Mo-MLV, showing that p12-chromatin binding is not targeting Mo-MLV integration site selection. This has also recently been reported using a panel of heterologous CBS motifs in Mo-MLV p12 (Schneider et al., 2013). However, loss of the C-terminus of p12 reduced the integration tendency towards transcriptional start sites, something not seen for the N-terminal p12 mutant 6. However, it should be noted that the number of integration sites analysed were too small to make this observation statistically significant but it does warrant further investigation. If this trend remains after further experiments it means that p12 mediated docking does exert an effect on integration site selection, and that binding to the chromatin (by the motif in p12 or another CBS) allows the integration machinery to contact a co-factor that then determines site selection. This fits with the fact that wild

type Mo-MLV is 10-fold less infectious when p12 contains the PFV-CBS, indicating that the PFV-CBS could bind to the chromatin where less of the site selection co-factor exist.

The data I have shown here, together with data from Monica Roth's laboratory and the elegant experiments performed in Eran Bacharach's laboratory, indicate that p12 has a LEDGF-like chromatin tethering function for Mo-MLV. However, unlike the LEDGF-chromatin interaction for lentiviruses, the p12-chromatin interaction is not directly determining the site of Mo-MLV integration.

Chapter 6.

Investigations into the movements of the retroviral PIC during the early stages of infection

Work in the previous chapters has highlighted that p12 contains two domains, both which function during the early stages of infection. Genetic and biochemical assays have revealed that the N-terminal domain has a role in formation and/or stability of the retroviral core, while the C-terminus is required later in the life cycle, to bind the PIC to mitotic chromatin. To compliment the genetic and biochemical data compiled so far, the sub-cellular localisation of p12 and the CA core were investigated using fluorescence microscopy.

Studying the movements of the PIC in infected cells with fluorescence microscopy presents many problems. Of most significance is the non-productive uptake of virus particles into cells (Marechal et al., 1998; McDonald et al., 2002). Methods to exclude these defective particles from analysis were assessed. While the role of the C-terminal domain of p12 during the early stages of infection has been investigated thoroughly using microscopy, the phenotype of the N-terminal p12 mutants have not been assessed (Elis et al., 2012; Prizan-Ravid et al., 2010). During a collaborative stay at Tel Aviv University, the effect of the N-terminal p12 mutants on Mo-MLV PIC movements were studied using a combination of live cell imaging and fixed immunofluorescence.

6.1 Expression of p12 in mammalian cells

The only structural data on p12 comes from a fusion of p12 with the NTD of CA, and suggests that p12 displays a high degree of flexibility (Kyere, Joseph, and Summers, 2008). Thus p12 does not form a defined structure *in vitro*. Indeed, when untagged p12 was expressed in mammalian cells it was only weakly detectable by immunoblotting after transfection (Boucherit, V. unpublished data). This is likely due to its unstructured nature, when not part of the PIC, and subsequent degradation by the proteasome. Thus, analysis of the localisation of native p12 in mammalian cells by immunofluorescence was not an option. In order to stabilise p12 and to study its sub-cellular localisation, p12 was fused to a fluorescent protein.

6.1.1 The sub-cellular localisation of an mCherry-p12 fusion expressed in mammalian cells

The Mo-MLV p12 coding region of *gag* was amplified by PCR and cloned 3' to mCherry in a mammalian expression vector, resulting in a p12 fusion protein N-terminally tagged with mCherry, via a short linker (Figure 6.1A, mCherry-p12). A stop codon was introduced after the PPPY to produce an mCherry fusion to the N-terminus of p12 (Figure 6.1A, mCherry-Np12). To produce an mCherry fusion to the C-terminus of p12, an *Eco*RI restriction site was introduced into the central portion of p12 and the N-terminus of p12 was excised by digesting with this restriction enzyme (Figure 6.1A, mCherry-Cp12).

HeLa cells were transfected with the mCherry-p12 fusions or an mCherry molecule containing just the linker and were analysed by immunoblotting, 20 hours after transfection (using the p12 polyclonal antibody) (Figure 6.1B). As a control for p12 cleavage from mCherry, Mo-MLV VLPs were included. All three mCherry-p12 fusion proteins were expressed and displayed a molecular weight, by SDS-PAGE, similar to the calculated theoretical molecular weight. The mCherry-p12 transfected cells contained a prominent band at ~43kDa, slightly higher than the theoretical size of 36.5kDa. As p12 migrates as a 12kDa protein when separated by SDS-PAGE, despite its theoretical size of 8.7kDa, this slight discrepancy was to be expected (Figure 6.1B). Similarly, the prominent band in the cells transfected with the fusion proteins mCherry-Np12 or mCherry-Cp12 migrated at ~37kDa, slightly higher than the calculated molecular mass of 31.2kDa (Figure 6.1B). A non-specific band could be seen migrating at ~12kDa, the same size as mature p12, in all the HeLa cell samples (Figure 6.1B). Importantly, this band was not increased in the cells transfected with the mCherry-p12 fusion indicating that p12 was not specifically cleaved from the fusion protein.

Degradation products of the mCherry-p12 fusion proteins could be seen below the bands corresponding to the full-length fusion protein, and one was particularly prominent (~10kDa less that the fusion proteins) (Figure 6.1B, asterisks). This band likely represented a rate-limiting intermediate during the breakdown of the fusions proteins in mammalian cells. Unfortunately, despite our best efforts, the mCherry could

not be detected with an antibody that should cross-react to the mCherry protein (data not shown). Importantly, all p12 fusion proteins were expressed in mammalian cells, were approximately the expected sizes by SDS-PAGE, and p12 was not preferentially cleaved from mCherry.

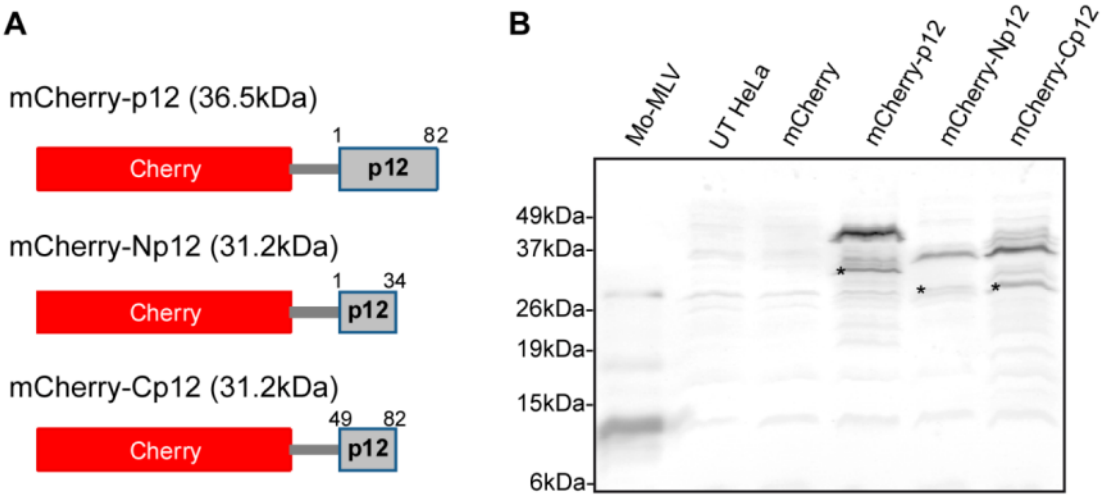


Figure 6.1: Expression of mCherry-p12 fusion proteins in mammalian cells. (A) Diagrams displaying the regions of p12 fused to mCherry. The molecular mass of the fusion proteins are displayed in brackets, adjacent to the name of the fusion protein. (B) HeLa cells were transfected with or without the fusion proteins, and the cell lysate analysed by immunoblotting with an anti-p12 polyclonal antibody. Viral lysate of wild type Mo-MLV was used as a control for cleavage of p12 from mCherry. UT- untransfected. *- specific degradation product

HeLa cells transfected with the mCherry-p12 fusion proteins were analysed by fluorescence microscopy to determine if p12 contains any karyophilic sequences determining sub-cellular localisation. Cells were fixed 20 hours post-transfection and treated with DAPI to stain the nuclear DNA (Figure 6.2).

Due to the relatively small size of mCherry it does not require active transport into the nucleus and could be seen throughout the cytoplasm and nuclear compartment (Figure 6.2B). Similarly mCherry-p12, mCherry-Np12 and mCherry-Cp12 were all distributed throughout the cytoplasm and nucleus (Figure 6.2 C, D and E, respectively). Interestingly, a few cells expressing the mCherry-Cp12 displayed some nucleolar accumulation of mCherry signal (Figure 6.2E, magenta box). However, this observation was relatively rare, indicating it may be due to overloading the cells with fusion protein. Therefore, p12 did not harbour significant karyophilic activity to direct the fusion proteins to specific sub-cellular compartments in these experiments.

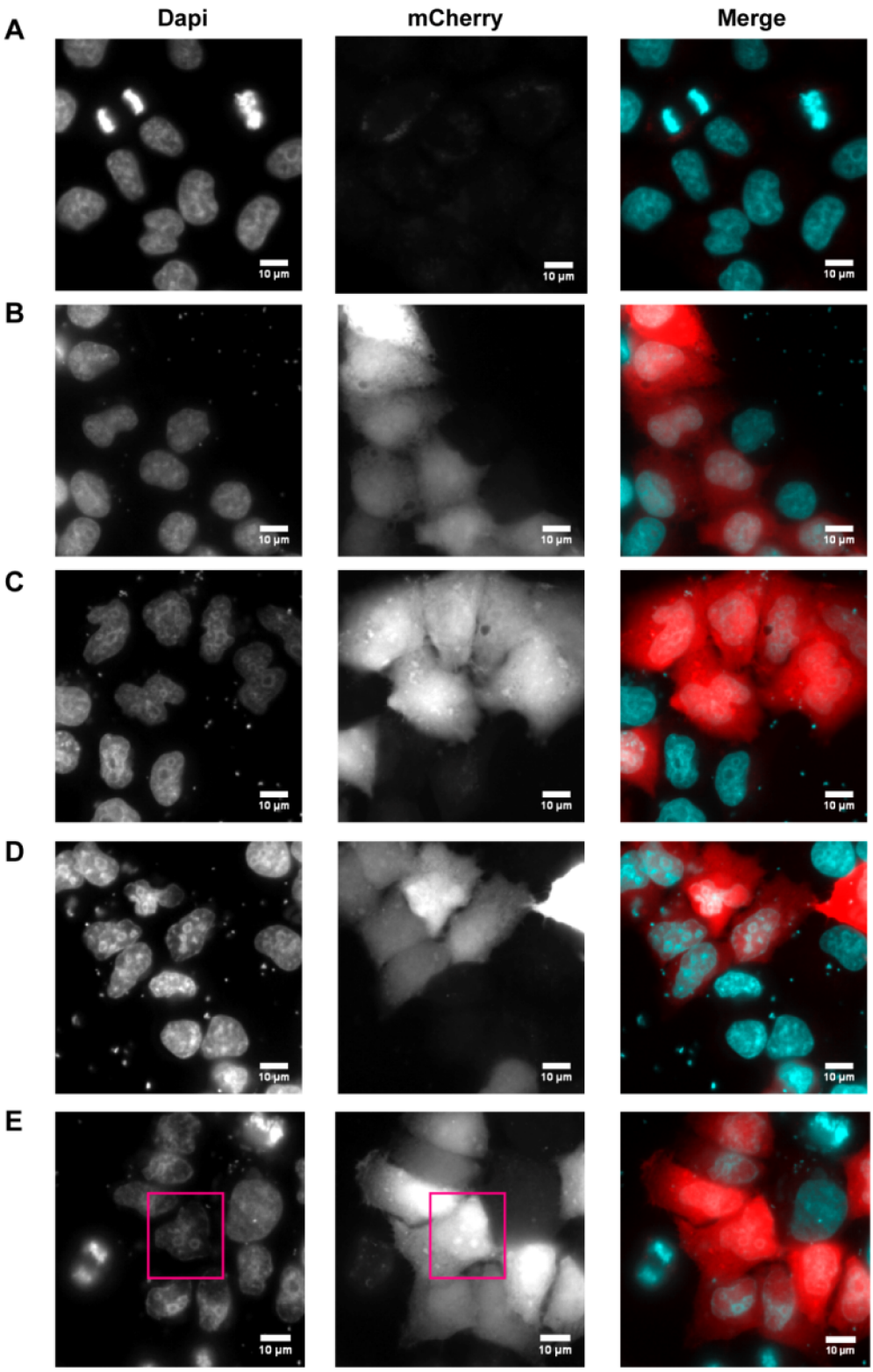


Figure 6.2: Sub-cellular localisation of mCherry-p12 fusions expressed in mammalian cells. HeLa cells were transfected with or without the mCherry-p12 fusion proteins. 20 hours after transfection the cells were fixed and analysed by fluorescent microscopy. Representative images of HeLa cells transfected with: (A) no plasmid, (B) mCherry, (C) mCherry-p12, (D) mCherry-Np12 and (E) mCherry-Cp12 are shown. Exposure settings were kept constant during acquisition and the fluorescence intensity of the mCherry signal has been scaled by the same values for each sample. The nucleolar accumulation of mCherry-Cp12 is highlighted in the magenta box.

6.1.2 The sub-cellular localisation of mCherry-p12 in mitotic cells

Unfortunately, no HeLa cells transfected with the fusion proteins were observed in mitosis when imaged using fluorescence microscopy. Therefore to observe transfected cells during mitosis, live imaging experiments were performed. *M.dunni* cells were transfected with the mCherry-p12 fusion constructs and the cells imaged over a ~24 hour time course. The *M.dunni* cells were chosen as they divide very rapidly and thus increase the number of mitotic events occurring. Cells were treated with Hoechst 33342, a membrane permeable DNA dye, so the chromosomes could be visualised. Representative movies can be seen in Appendix C. Unfortunately, no transfected mitotic cells could be seen during the time course. The temporal constraints of visiting multiple points (to record multiple cells) during the time-course precluded recording of enough cells, thus another approach was needed.

An array of drugs are available that bock the cell cycle at certain stages, nocodazole depolymerises microtubules and blocks cells in mitosis. 293T cells were transfected with the fusion proteins followed by incubation with nocodazole. This approach enabled the analysis of cells expressing the fusion proteins during mitosis (Figure 6.3). The control cells expressing mCherry did not show an accumulation of the cherry signal on the condensed chromosomes (Figure 6.3B). The histograms, on the right, display the fluorescent signal intensities along the line annotated in the images and have been scaled to emphasise the DAPI (DNA) and mCherry signal in the area around the condensed chromosomes (Figure 6.3). The chromosomes can be seen as the DAPI peaks in the histogram. The histogram of the mCherry control clearly showed that the mCherry signal did not peak in the proximity of the DAPI (chromosomal) peaks (Figure 6.3B). Similarly, for mCherry-p12 and mCherry-Cp12 no strong accumulation of mCherry signal could be detected on the chromosomes (Figure 6.3 C and D, respectively). However, avoidance of the DAPI peaks in the histograms displaying the mCherry-p12 and mCherry-Cp12 signal was not seen, in contrast to the mCherry control (Figure 6.3, histograms in C and D versus B). Therefore the mCherry-p12 was not excluded from the chromosomes, but due to the high signal in the space between the chromosomes, it is difficult to judge if p12 is driving a slight accumulation of the fusion proteins onto the chromosomes.

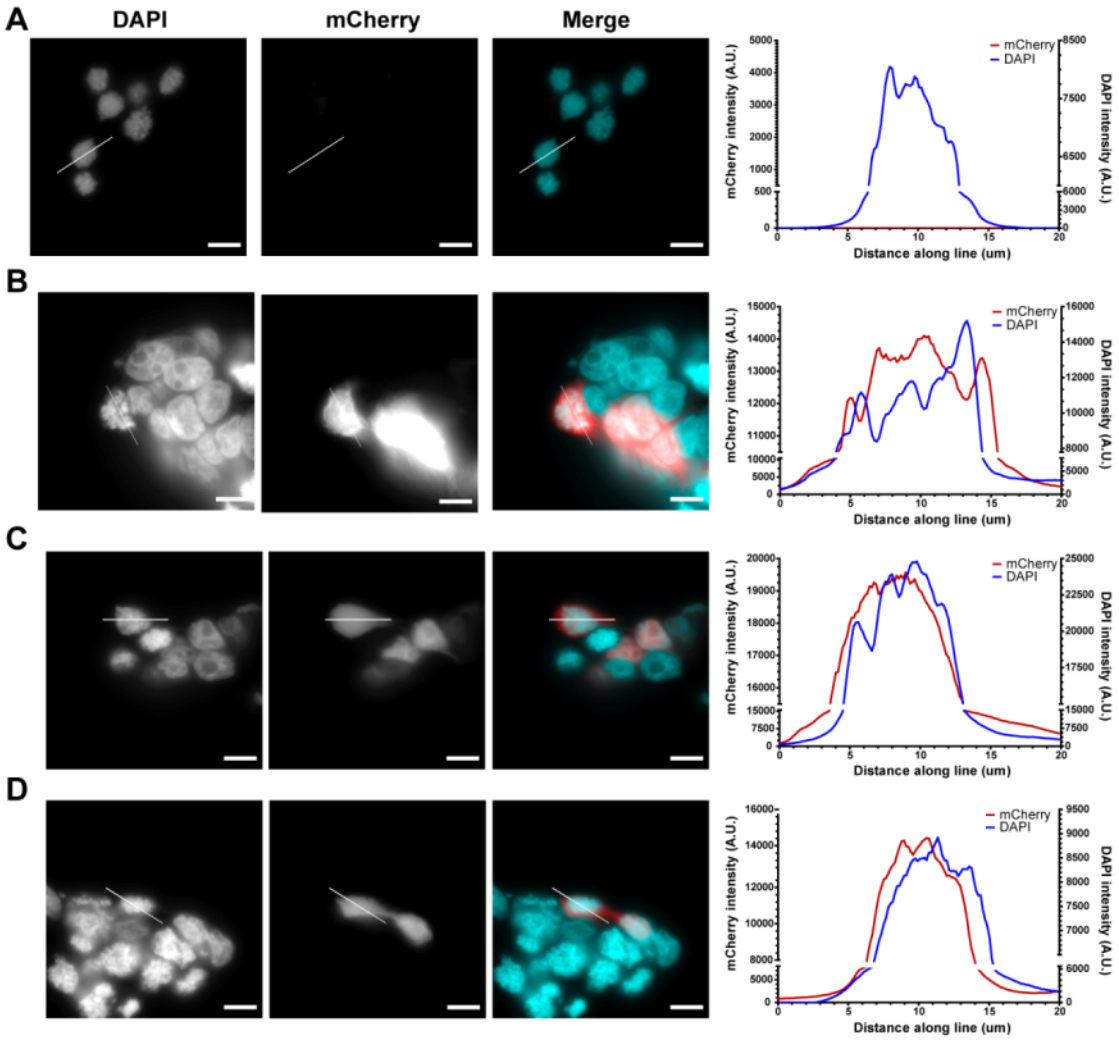


Figure 6.3: Sub-cellular localisation of mCherry-p12 fusions expressed in mammalian cells arrested in mitosis. 293T cells were transfected with the fusion proteins or controls and then arrested in mitosis using nocodazole. Representative images of 293T cells transfected with: (A) no plasmid, (B) mCherry, (C) mCherry-p12 and (D) mCherry-Cp12 are shown. The intensity profile of the mCherry and DAPI signal along the lines (left-to-right), drawn in the images, are displayed in the histograms on the right. Exposure settings were kept constant during acquisition and the fluorescence intensity of the mCherry signal has been scaled by the same values for each sample.

6.2 Visualising Mo-MLV p12 during infection

The use of fluorescence microscopy to follow viral proteins during different stages of infection has provided powerful insights into many aspects of the host-retrovirus interaction. Immunofluorescence has also been applied to study p12 and the Mo-MLV PIC during infection. The association of myc-tagged p12 with CA and the viral DNA during the early stages of infection has highlighted that p12 is part of the MLV PIC (Prizan-Ravid et al., 2010). As mentioned in Chapter 5, the interaction of the MLV PIC with condensed chromatin during infection in a p12-dependant manner has been elegantly shown using both fixed and live imaging fluorescence microscopy (Elis et al., 2012; Prizan-Ravid et al., 2010). Thus, application of microscopy techniques to study the movements of the Mo-MLV PIC bearing N-terminal p12 mutations may reveal more information regarding the defect in these mutants.

6.2.1 Non-specific uptake of Mo-MLV into cells via an Env-independent mechanism

It has been widely noted in the literature that in addition to viruses entering via a productive Env-mediated pathway, intact viral particles are taken up non-specifically into cells (Marechal et al., 1998; McDonald et al., 2002). These particles are not on an infectious pathway and can be considered as dead-end products, although their presence complicates the analysis of microscopy data.

To assess the magnitude of this problem with Mo-MLV VLPs, D17 cells were infected using equal RT-units of VLPs with or without VSV-G. After 12 hours, cells were fixed and immunostained with anti-p12 monoclonal and anti-CA antibodies, followed by secondary antibodies conjugated with alexafluor 594 and 488, respectively. No signal could be seen in the uninfected cells as expected (Figure 6.4A). However, cells infected with Mo-MLV VLPs with or without Env proteins looked remarkably similar (Figure 6.4 B and C). Large amounts of CA and p12 staining could be seen in the cytoplasm (in endosomes for VLPs lacking Env) for both, indicating Mo-MLV VLPs lacking Env proteins were efficiently taken up into cells in a non-specific manner.

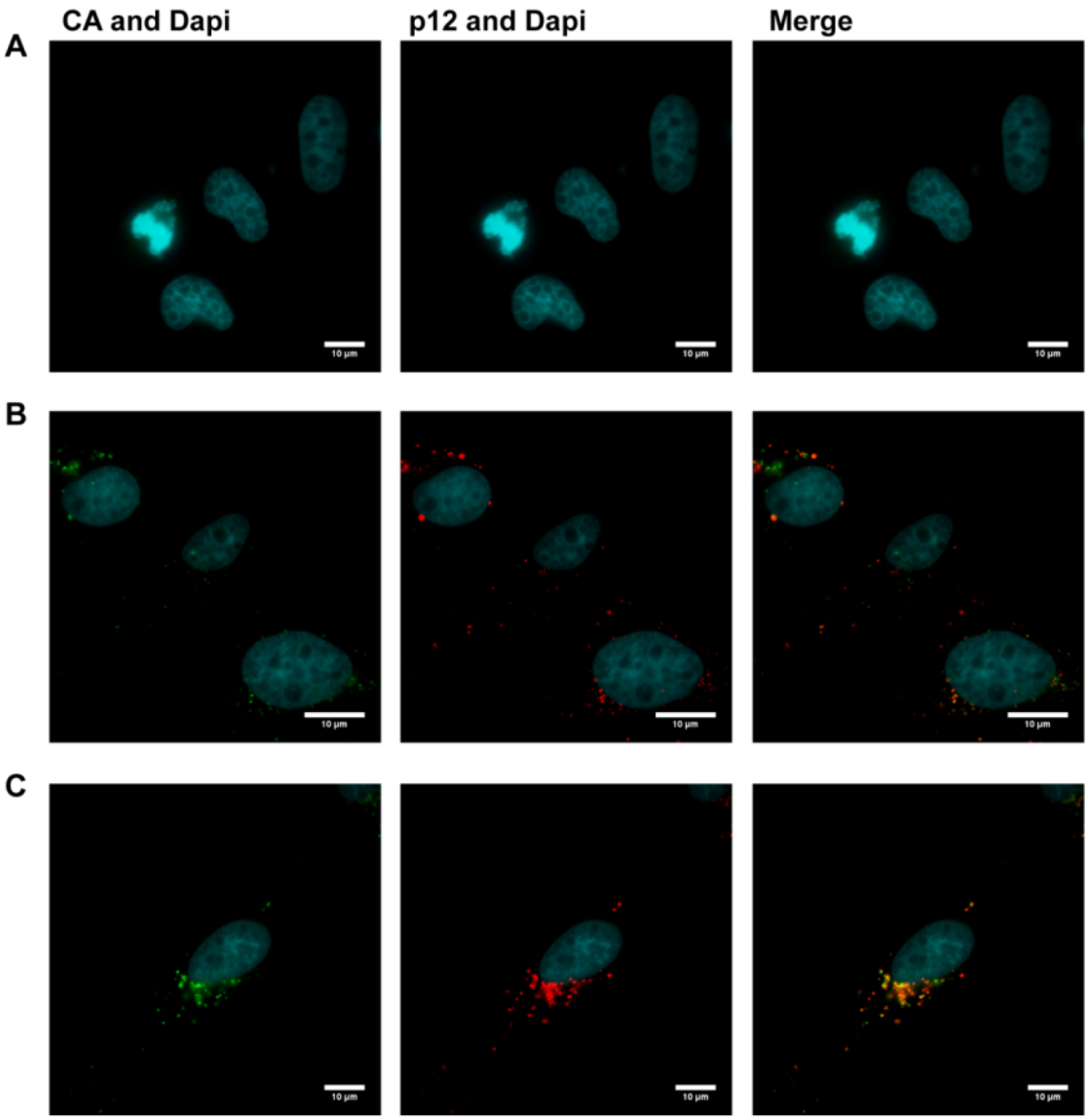


Figure 6.4: Non-specific uptake of Mo-MLV into D17 cells. 12 hours post-infection (A) uninfected D17 cells or D17 cells infected with equal RT-units of Mo-MLV with (B) the VSV-G envelope or (C) no envelope were analysed by indirect immunofluorescence using CA and p12 monoclonal antibodies. For each viral protein detected, the exposure settings were kept constant during acquisition and the fluorescence intensity has been scaled using the same values for each sample.

Prizan-Ravid *et al.* utilised cells (U/R) that stably express the ecotropic MLV receptor, mCAT-1, and viruses with the ecotropic envelope (Prizan-Ravid et al., 2010; Wang et al., 1991). Therefore, the non-specific uptake control in these experiments was infection of the parental U20S cells, which lack the mCAT-1 receptor. Remarkably, despite the large amount of internalised p12 puncta in the infected U/R cells, the non-specific uptake control had very few p12 puncta (Prizan-Ravid et al., 2010). Therefore it was possible that the Env protein somehow reduces non-specific uptake of virions into cells.

To test this, U20S cells were infected for two hours with equal RT-units of Mo-MLV containing either the ecotropic Env or no Env. Cells were washed thoroughly after infection, returned to the incubator for a further two hours, fixed and immunostained using the anti-CA antibody. Z-stacks of randomly chosen interphasic cells (picked based on the DAPI staining) were imaged and the Z-stacks projected into a 2D representation of the volume of the cell (Figure 6.5A, examples of the projected Z-stacks). The number of CA puncta per cell was much higher (four-fold increase) for Mo-MLV VLPs with no Env than VLPs that contained an Env complex that does not permit access to the cell (Figure 6.5B). Thus, it appears that in U20S cells, VLPs containing no Env were more readily taken up in an receptor-independent manner than VLPs bearing an Env complex that did not permit entry to U20S cells.

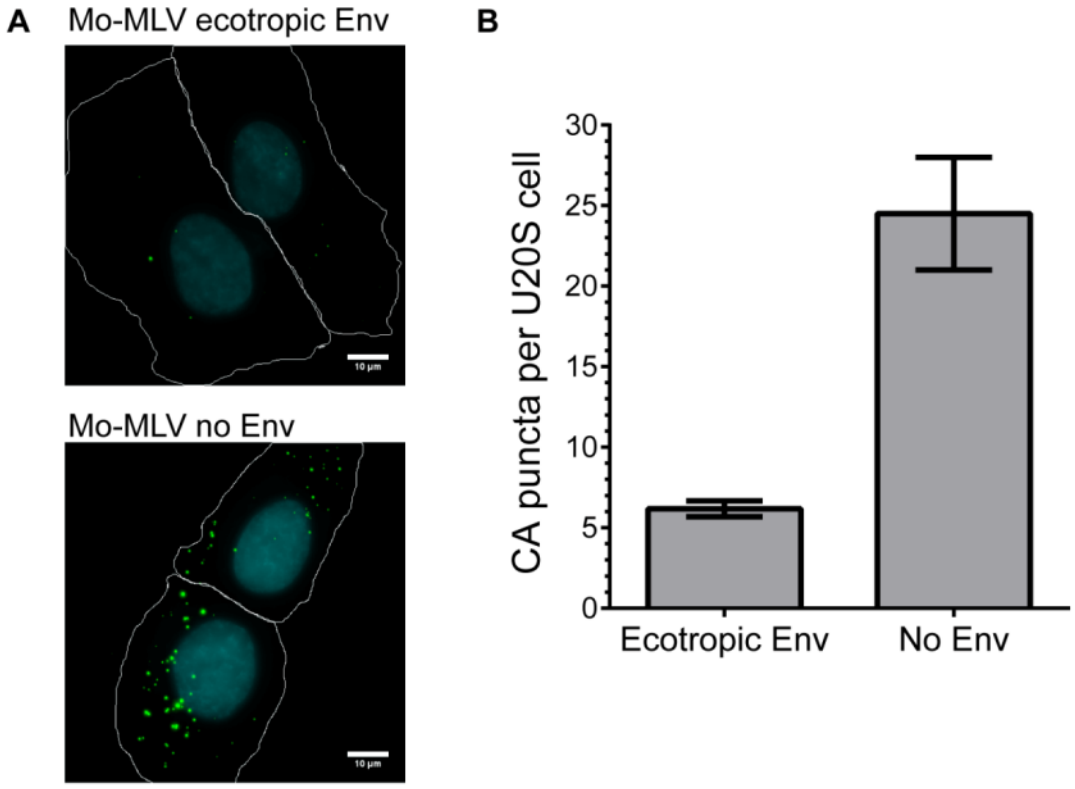


Figure 6.5: Non-specific uptake of ecotropic Mo-MLV into human U20S cells. Human U20S cells were infected with equal RT-units of Mo-MLV VLPs with an ecotropic envelope or Mo-MLV VLPs lacking an envelope. Cells were fixed two hours post-infection and analysed by indirect immunofluorescence using the anti-CA antibody. (A) Representative images of a compressed Z-stack of the infected U20S cells. (B) Two independent experiments were performed and a Z-stack of images through randomly chosen cells recorded. The total number of CA puncta per cell (over a fluorescence threshold) were counted. The mean of the two independent experiments are shown with the range. Note- the low green auto-fluorescence of cells was used to determine the cell borders.

6.2.2 Visualising PICs on a productive infection pathway in infected cells

As was shown in the previous section, Mo-MLV VLPs are non-specifically taken up into cells in an receptor-independent manner. While this problem is likely over-estimated by using VLPs containing no Env, it is evident that a significant number of VLPs enter the cell in a non-productive manner. Indeed, cells infected by a given infectious MOI of virus contain significantly higher numbers of discrete viral puncta than predicted from the MOI, suggesting that the replication of many particles must fail during the early steps of infection. Two separate methods were attempted to rule out the analysis of non-productive 'infections'.

Synthesis of viral DNA is an essential step before integration, thus, identification of PICs that contain viral DNA are more likely on a productive infection pathway. Viral DNA and proteins can be detected by using fluorescent *in situ* hybridisation (FISH) combined with immunofluorescence in infected cells (Prizan-Ravid et al., 2010; Roe et al., 1993). As all the N-terminal mutants, barring mutant 6, reverse transcribe in D17 cells a combined viral DNA FISH and immunofluorescence could be utilised during the early stages of infection. D17 cells were infected with a high MOI of Mo-MLV VLPs (MOI >10) and were stained with the anti-CA antibody followed by FISH against the viral DNA at 10 hours post-infection. A strong signal could be seen for the CA immunofluorescence in the cytoplasm but no viral DNA FISH signal could be seen through the high fluorescence background (Figure 6.6A). Indeed, uninfected D17 cells treated with the biotin labelled probe and the Streptavidin-FITC displayed a high level of background, which was most intense in the perinuclear region (Figure 6.6B).

As the background from the FISH staining appeared primarily in the cytoplasm and the perinuclear region, it was thought that the background may be due to binding of the Streptavidin-FITC to biotin-like molecules. To block this, cells were pre-treated with 1mg/ml avidin before FISH. Unfortunately, this did not reduce the background and no punctate signal for the viral DNA, co-localising with the p12 signal could be seen (Figure 6.6A). Hence, due to a high background, caused by non-specific binding of the Streptavidin-FITC, the viral DNA could not be detected in D17 cells infected with Mo-MLV VLPs.

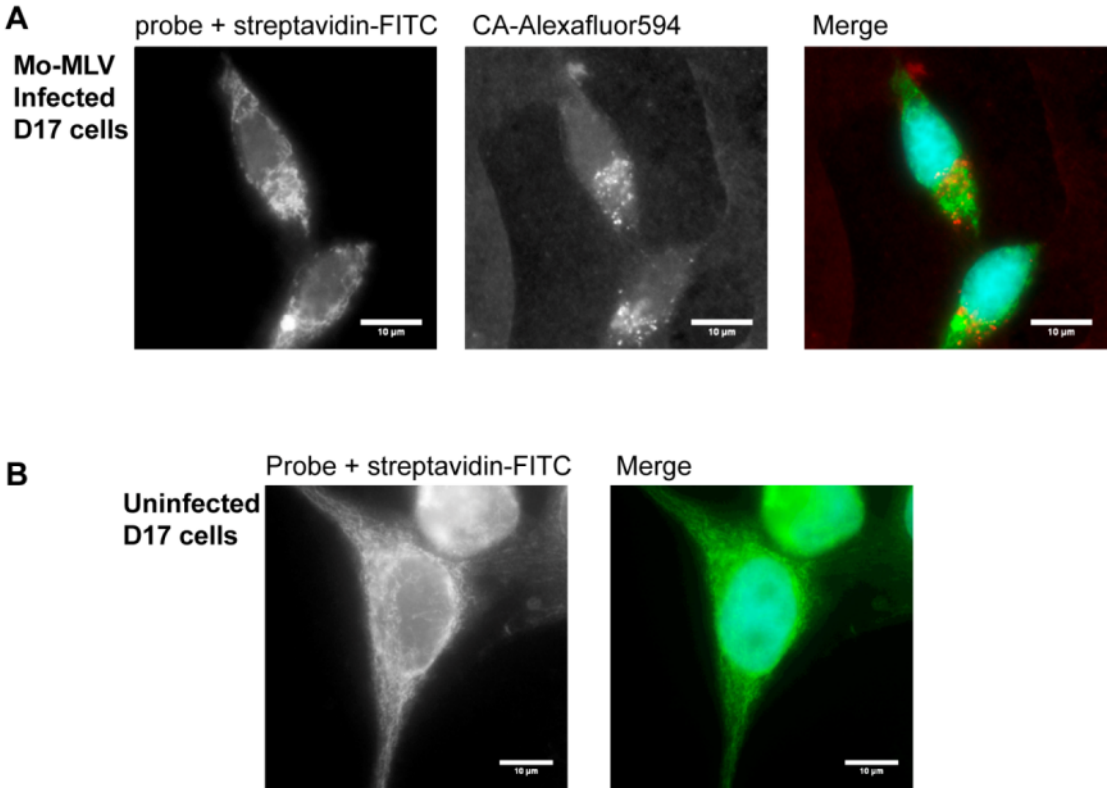


Figure 6.6: Attempted detection of the Mo-MLV PIC by fluorescent *in situ* hybridisation (FISH). (A) D17 cells were infected with a high MOI of Mo-MLV and were analysed by indirect immunofluorescence against CA, followed by FISH with a probe against the Mo-MLV LTR and *LacZ* genome of the VLPs. (B) Uninfected D17 cells were processed for FISH as in (A) without the immunofluorescence against CA. The exposure settings were kept constant for the FISH images.

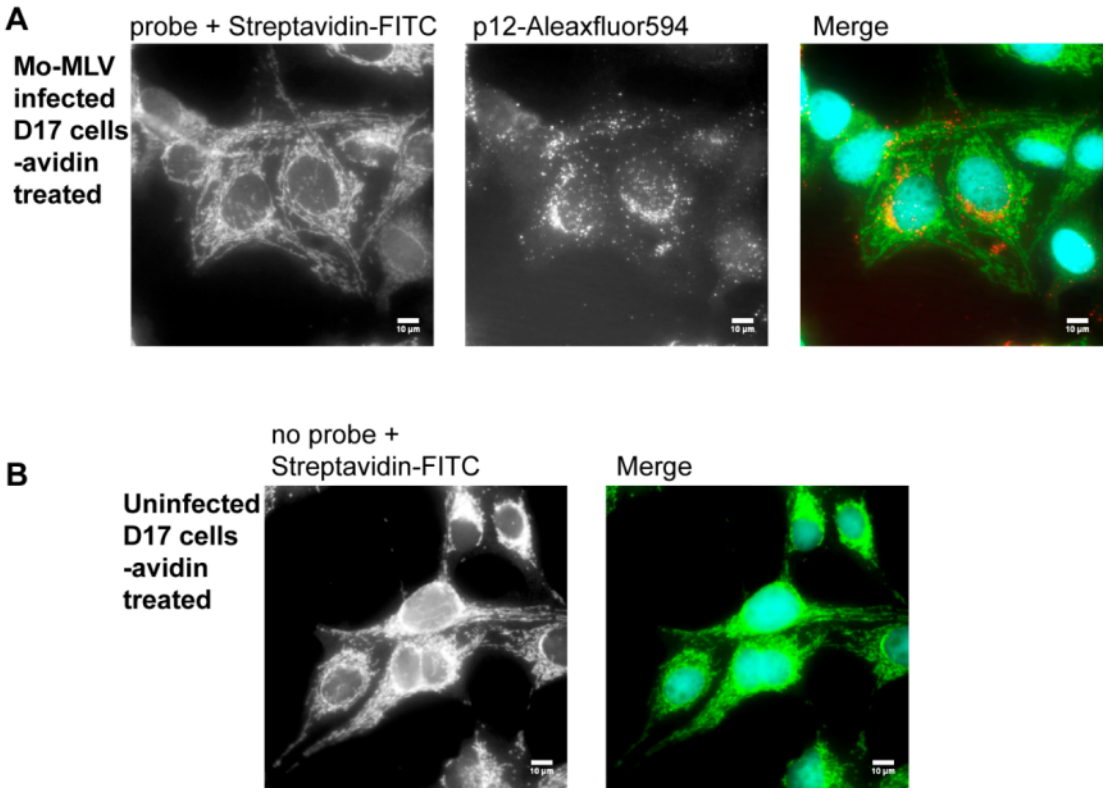


Figure 6.7: Reducing the background of fluorescent *in situ* hybridisation (FISH) of the Mo-MLV PIC in infected cells. (A) D17 cells were infected with a high MOI of Mo-MLV and were analysed by indirect immunofluorescence against p12, followed by FISH with a probe against the Mo-MLV LTR and *LacZ* genome of the VLPs. Before processing for FISH, cells were treated with 1mg/ml avidin. (B) Uninfected D17 cells were processed for FISH as in (A) without the immunofluorescence against p12. The exposure settings were kept constant for the FISH images.

A second approach to reduce the amount of non-specifically endocytosed virions analysed in microscopy has been done by labelling the HIV-1 virion membrane with GFP/mCherry fused to the 15 N-terminal membrane targeting amino acids of $p60^{c\text{-}Src}$ (Campbell et al., 2007). This method has allowed virions inside cells, containing the virion membrane, to be excluded from analysis. Therefore, this labelling method was applied to Mo-MLV. To assess the effect of co-transfection of the S15-mCherry plasmid on VLP assembly and release, VLP plasmids were transfected (1µg of each) with increasing amounts of the pS15-mCherry expression plasmid. More pS15-mCherry plasmid transfected dramatically reduced the amount of VLP released, as assessed by the RT activity in the supernatant (Figure 6.8A). Thus, excess S15-mCherry plasmid potently competed with the viral plasmids for transcription and translation. As lower doses of transfected S15-mCherry permitted synthesis of reasonable levels of virus

(Figure 6.8A), the infectivity of these 'labelled' viruses was tested. A 3.6-fold reduction in infectivity was seen for VLPs synthesised when co-transfected with 0.25µg of pS15-mCherry (Figure 6.8B). This effect was exacerbated when greater amounts of pS15-mCherry were transfected, which reduced the infectivity by 8-fold compared to wild type (Figure 6.8B). Therefore, transfection (and likely expression) of larger amounts of S15-mCherry during VLP synthesis was detrimental to Mo-MLV VLP release and infectivity.

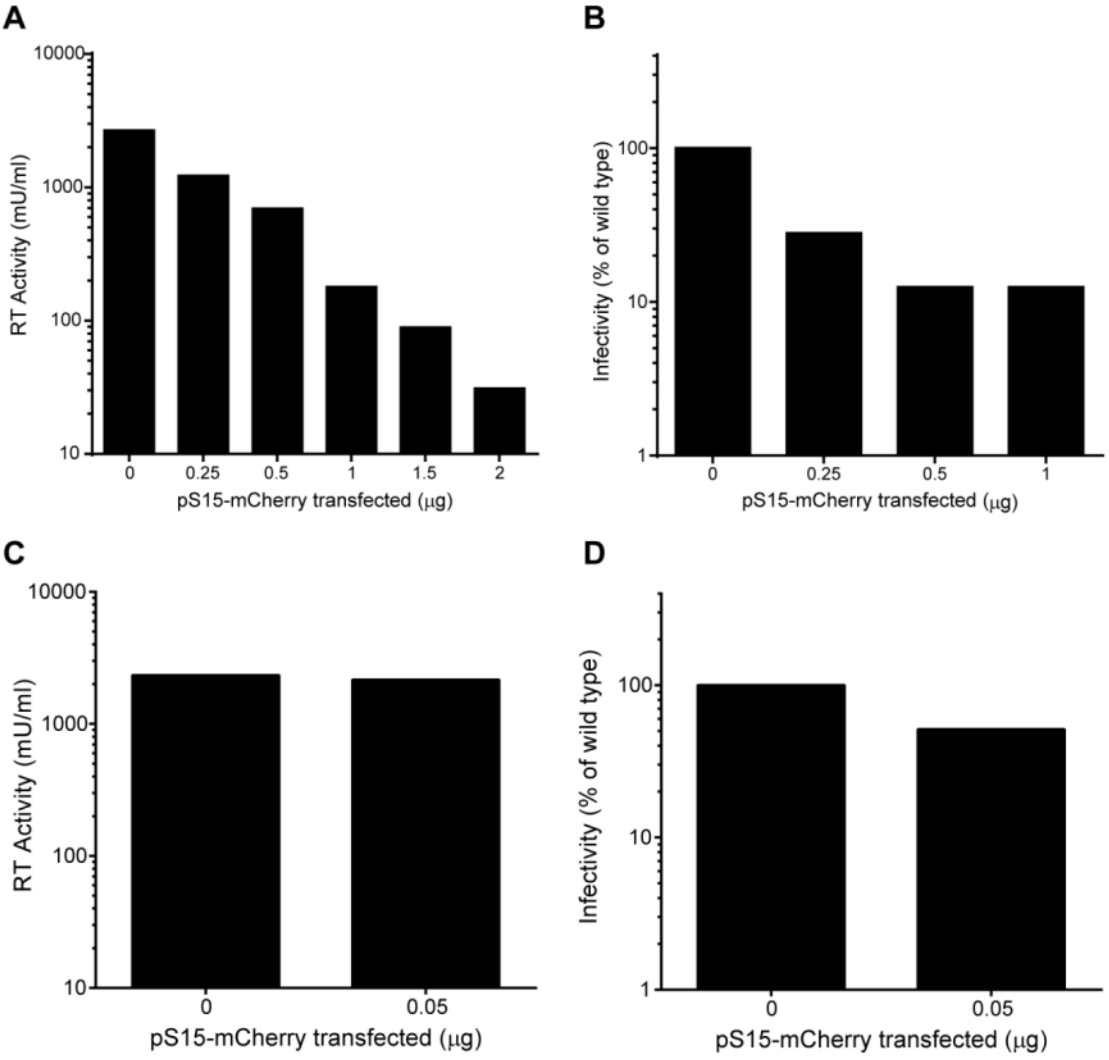


Figure 6.8: The effect of co-transfecting pS15-mCherry while synthesising Mo-MLV VLPs. Mo-MLV VLPs were synthesised with increasing amounts of co-transfected pS15-mCherry. (A) VLP release is shown as the level of RT-activity released from the transfected cells. (B) D17 cells were infected with equal RT-units of VLPs, infectivity was quantified and expressed as a percentage of Mo-MLV VLP infectivity without co-transfected pS15-mCherry. (C) VLP release of Mo-MLV co-transfected with lower levels of pS15-mCherry. (D) D17 cells infected with equal RT-units of VLPs, were quantified and expressed as a percentage of Mo-MLV VLP infectivity without co-transfected pS15-mCherry.

As Mo-MLV VLP production and infectivity was compromised using this labelling technique I was advised that transfecting very low levels of the S15mCherry plasmid was sufficient for HIV-1 labelling (Eleanor Gray, personal communication). Labelling Mo-MLV with 0.05μg of pS15-mCherry did not have an effect on VLP release (Figure 6.8C) and the infectivity was reduced by less than two-fold compared to un-labelled Mo-MLV (Figure 6.8D). Thus low levels of the pS15-mCherry plasmid were required to preserve Mo-MLV VLP viability.

HIV-1 VLPs released from the cells at early time points are the most efficiently labelled after transfection with the S15-mCherry plasmid (Ed Campbell, personal communication). Therefore to test how much Mo-MLV was labelled with the S15-mCherry, VLPs were harvested at different times post-transfection (media changed 24 hours post-transfection). As a control, HIV-1 VLPs were also labelled with S15-mCherry and harvested at the same time points.

VLP release was quantified by measuring the RT-activity (Figure 6.9 A and B). Labelling Mo-MLV by transfecting with 100ng or 50ng S15-mCherry had no discernible effect on the release of Mo-MLV when compared to the controls (Figure 6.9A, CON versus S15). When HIV-1 was labelled by transfecting 50ng of S15-mCherry a noticeable drop in VLP release was seen (Figure 6.9B, red and black solid lines with squares). Addition of 100ng of the control plasmid (pcDNA3.1) into the transfection also reduced the amount of HIV-1 VLP released (Figure 6.9B, black dashed line with triangles). When 100ng of S15-mCherry was transfected with the HIV-1 VLP plasmids, as time progressed less VLPs were produced. This suggested that larger quantities of S15mCherry plasmid were interfering with the synthesis/assembly/release of HIV-1 (Figure 6.9B, red dashed line with triangles).

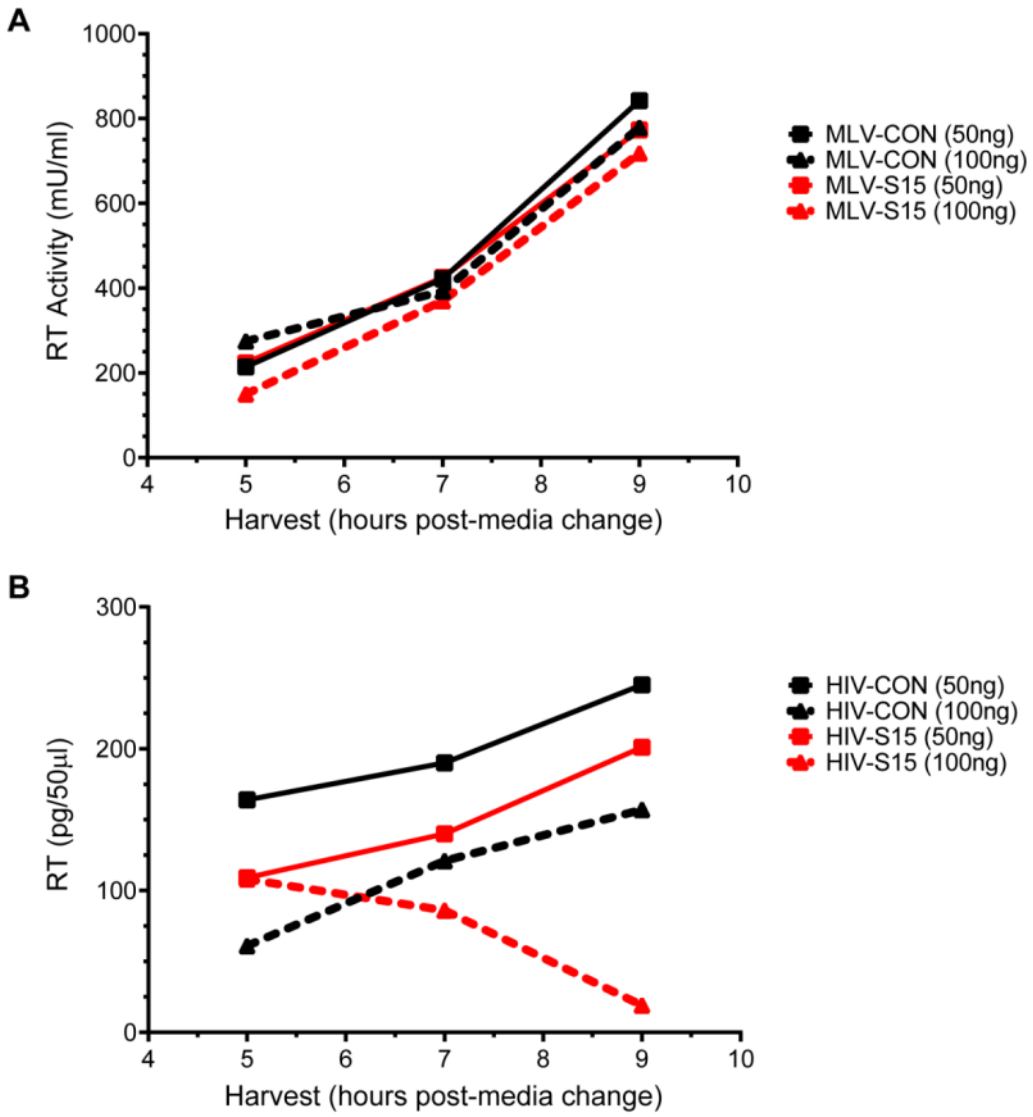


Figure 6.9: Mo-MLV and HIV-1 VLP release when co-transfected with pS15-mCherry. Mo-MLV and HIV-1 VLPs were synthesised co-transfected with pS15-mCherry (S15) or pcDNA3.1 (CON). (A) Time course of MLV VLP release (measured by RT-activity) when co-transfected with the pS15-mCherry or control plasmids (media change 24 hours post-transfection) (B) Time course of HIV-1 VLP release (measured by RT-activity, assay performed by Harriet Groom) when co-transfected with the pS15-mCherry or control plasmids (media change 24 hours post-transfection).

To test the efficiency of labelling, VLPs were spun onto coverslips, fixed and immunostained with an anti-p12 monoclonal antibody (Mo-MLV) or an anti-CA antibody (HIV-1). The control, HIV-1, was optimally labelled with the lower quantity of S15mCherry plasmid transfected when harvested 7 hours after the media change (Figure 6.10A, maroon line with squares). Using more of the S15mCherry plasmid resulted in two-fold decrease in labelling at this time point (Figure 6.10A, maroon dashed line and triangles). HIV-1 VLPs harvested very early (5 hours) and later (9

hours) had much less VLPs labelled with S15mCherry (Figure 6.10A, maroon solid and dashed lines). Mo-MLV was poorly labelled by S15mCherry using either quantity of plasmid DNA (Figure 6.10A, red solid and dashed lines). Unfortunately, these data preclude the use of this approach for identifying un-fused MLV virions.

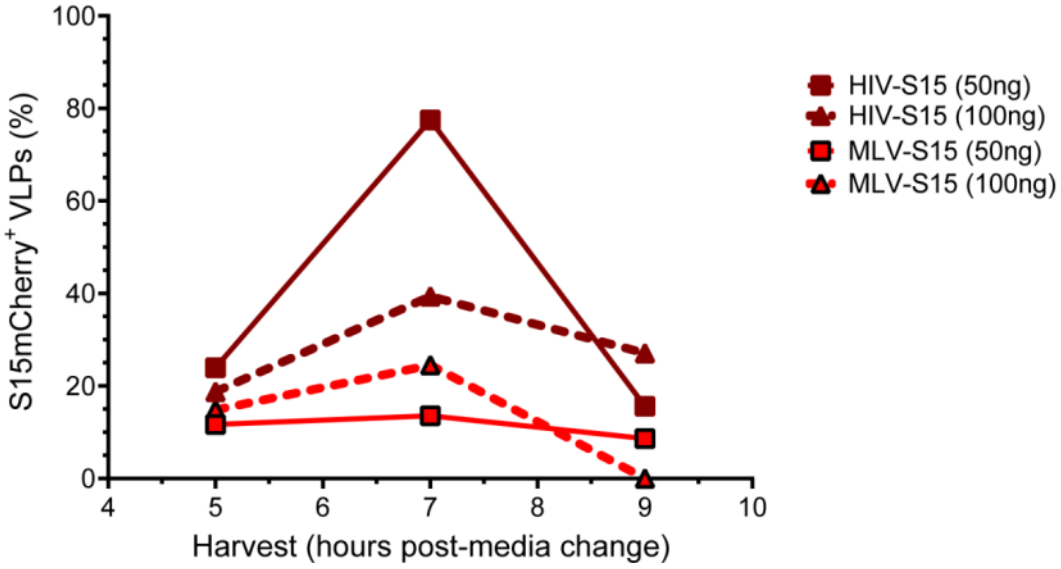


Figure 6.10: Labelling efficiency of Mo-MLV VLPs with S15-mCherry. Mo-MLV or HIV-1 VLPs co-transfected with 50ng or 100ng of pS15-mCherry during synthesis, were spun onto coverslips and analysed by indirect immunofluorescence with antibodies against p12 (for Mo-MLV) or CA (for HIV-1). (A) The percentage of $p12^+$ Mo-MLV VLPs and CA^+ HIV-1 VLPs labelled with S15-mCherry were calculated and plotted against the time of VLP harvest. At least 100 puncta were analysed for each virus/conditions and time point.

6.3 Construction of a system to follow the intracellular movements of the Mo-MLV PIC in real-time

Labelling of retroviral proteins, using a genetically encoded fluorescent protein, can provide a wealth of information about replication in real-time. HIV-1 Vpr is packaged into the virus through interactions with the p6 portion of the Gag precursor protein (Paxton, Connor, and Landau, 1993). Fusion of GFP to Vpr and expression of this fusion in HIV-1 producer cells labels a significant proportion of the viruses with GFP (McDonald et al., 2002). This labelling procedure has been widely used in the field to study the movements of HIV-1 during the early stages of infection. As MLV does not contain a trans-packaged accessory protein, like Vpr, fluorescent labelling methods for p12 and the MLV core were sought to study the early stages of MLV replication using real-time fluorescence microscopy.

6.3.1 Labelling p12 with the tetracysteine-specific dye ReAsH

Instead of tagging proteins with the relatively large eGFP (~27kDa), which can have detrimental effects to protein function, a smaller tag containing four cystienes can be cloned into the protein of interest. The biarsenical dyes, FlAsH or ReAsH, specifically bind to this tetracysteine motif (TC) and become fluorescent (Griffin and Adams, 1998). The ReAsH dye is particularly exciting as it is capable of photoconverting 3,3'-Diaminobenzidine (DAB) into a localised precipitate visible by electron microscopy (Lanman et al., 2008).

The smallest active TC-tag was inserted in the central portion of p12 ($_{45}$DGNGGE$_{50}$ > $_{45}$CCPGCC$_{50}$) of the Mo-MLV Gag-Pol vector using site directed mutagenesis. VLPs containing p12-TC, no p12-TC or a mixture of p12-TC and wild type p12 were generated and VLP release quantified (Figure 6.11A). VLPs containing only p12-TC could efficiently bud from the cell, indicating that the TC-tag in p12 did not affect VLP release (Figure 6.11A, black bars). VLP recovery, after the staining protocol (from Arhel *et al.* (Arhel and Charneau, 2009)), was good. This indicated that VLPs were not lost during the centrifugation and re-suspension step (Figure 6.11A, grey bars). Mo-MLV containing 100% p12-TC did, however, have a defect in infectivity, which resulted in a 3.8-fold reduction compared to wild type infectivity (Figure 6.11B, black bars). This infectivity defect was negligible when only 50% of p12 in the VLP contained the TC-tag (Figure 6.11B, black bars). Therefore the TC tag in the central portion of p12 had a small but significant impact on p12 function.

When the Mo-MLV VLPs were subject to the *in vitro* ReAsH labelling procedure, the infectivity of the VLPs was reduced. Even VLPs containing no p12-TC had a reduction in infectivity (1.7-fold reduction), indicating that the centrifugation step to remove unbound dye likely had a detrimental effect on viral infectivity (Figure 6.11B, grey bars). Interestingly, the VLPs that contained higher amounts of p12-TC had a greater loss of infectivity (Figure 6.11B, grey bars). Therefore, the *in vitro* staining protocol was detrimental to Mo-MLV infectivity, which was worse when more p12-TC was in the VLP.

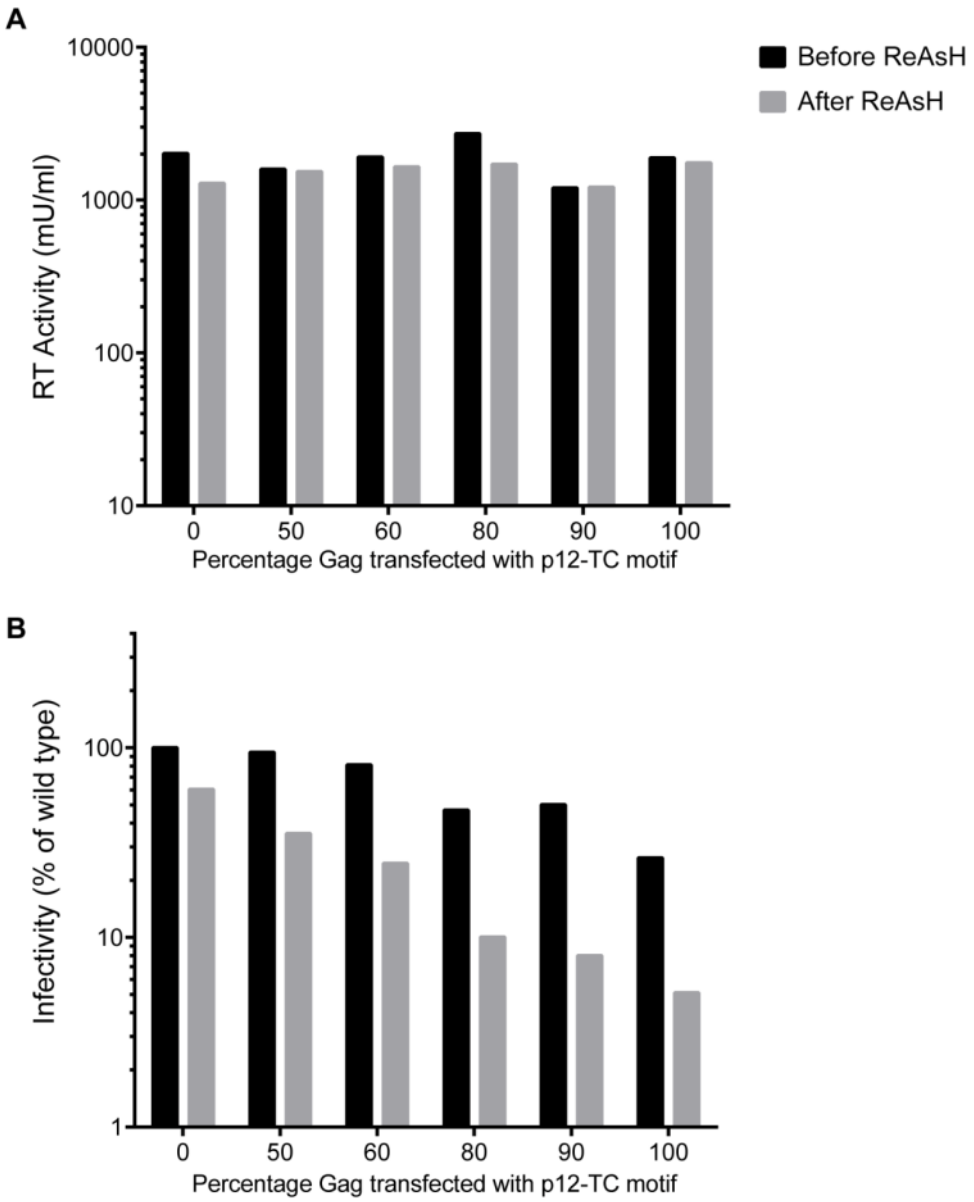


Figure 6.11: The effect of labelling p12-TC in Mo-MLV with ReAsH *in vitro*. Mo-MLV mixed particle VLPs, containing wild type p12 and p12 with a tetracysteine motif (p12-TC [CCPGCC]) inserted in the central portion of p12, were synthesised. (A) VLP release was measured by the level of RT activity produced (black bars, 'Before ReAsH'). After staining the amount of recovered VLP was quantified by the amount of RT activity (grey bars, 'After ReAsH'). (B) D17 cells were infected with equal RT-units of VLPs from before (black bars, 'Before ReAsH') and after (grey bars, 'After ReAsH') *in vitro* labelling with ReAsH. The infectivity is displayed as a percentage of unlabelled wild type Mo-MLV.

The efficiency of ReAsH labelling *in vitro* was assessed by immunostaining of VLPs bound to coverslips with the anti-p12 antibody. No specific (Figure 6.12, bottom panels), or non-specific ReAsH signal (Figure 6.12, top panels) could be seen for the VLPs stained with anti-p12. This suggested that the ReAsH dye was poorly bound to the TC motif in p12 using the *in vitro* labelling procedure. Large clumps of virus could

be seen, which was likely caused by the centrifugation step, and is probably part of the infectivity issue for viruses treated according to this protocol.

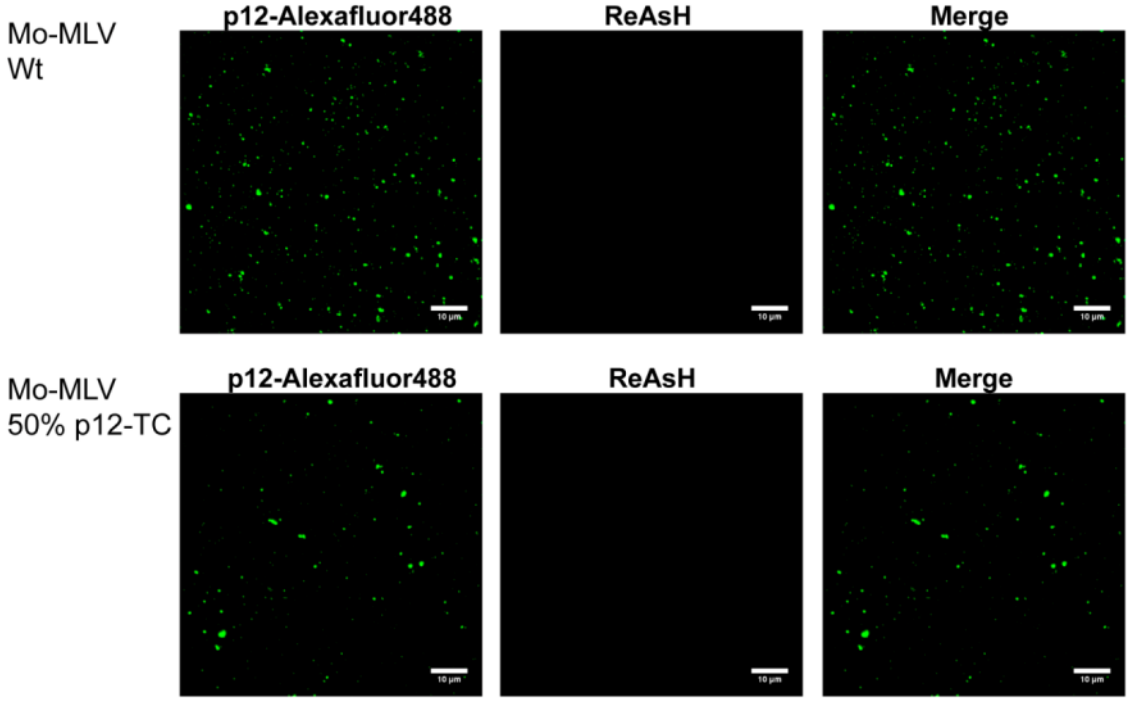


Figure 6.12: ReAsH labelling efficiency of p12-TC in Mo-MLV VLPs using an *in vitro* labelling protocol. Mixed Mo-MLV VLPs, with a 50/50 mixture of TC-tagged p12 to wild type p12, or wild type Mo-MLV VLP controls were stained with ReAsH *in vitro*. VLPs were analysed by indirect immunofluorescence with the p12 monoclonal antibody. A representative image of each VLP sample is shown. Exposure settings were kept constant for each fluorescent dye during acquisition and the fluorescence intensity of the signals have been scaled using the same intensity values.

Labelling the matrix protein (M-protein) from vesicular stomatitis virus with a biarsenical dye, for studying the early stages of infection, has been achieved by staining the virus in the producer cell, during assembly (Mire et al., 2009; Whitt and Mire, 2011). As the *in vitro* labelling technique was unsuccessful for ReAsH labelling of p12-TC in Mo-MLV, an *in vivo* labelling of the Mo-MLV p12-TC VLPs was attempted. Producer cells were treated with 2.5µM of ReAsH in serum-free media for 16 hours before VLP harvest. VLPs were concentrated, stuck to coverslips and analysed by immunofluorescence with a monoclonal antibody against p12.

Wild type Mo-MLV produced in the presence of ReAsH did not take up the dye non-specifically (Figure 6.13A, top panels). A few puncta of ReAsH material could be seen

in this control which likely represent aggregates of ReAsH, as they do not co-stain with the p12 antibody. Mo-MLV VLPs, containing 50% p12-TC, displayed a significant portion of VLPs labelled with ReAsH (Figure 6.13A, bottom panels), indicating that the *in vivo* technique labels p12-TC in Mo-MLV efficiently. Unfortunately, the infectivity of the labelled particles (Mo-MLV 50% p12-TC) was significantly reduced compared to Mo-MLV wild type (Figure 6.13B). This suggested that the binding of the ReAsH dye to the central region of p12 had a detrimental effect to p12 function.

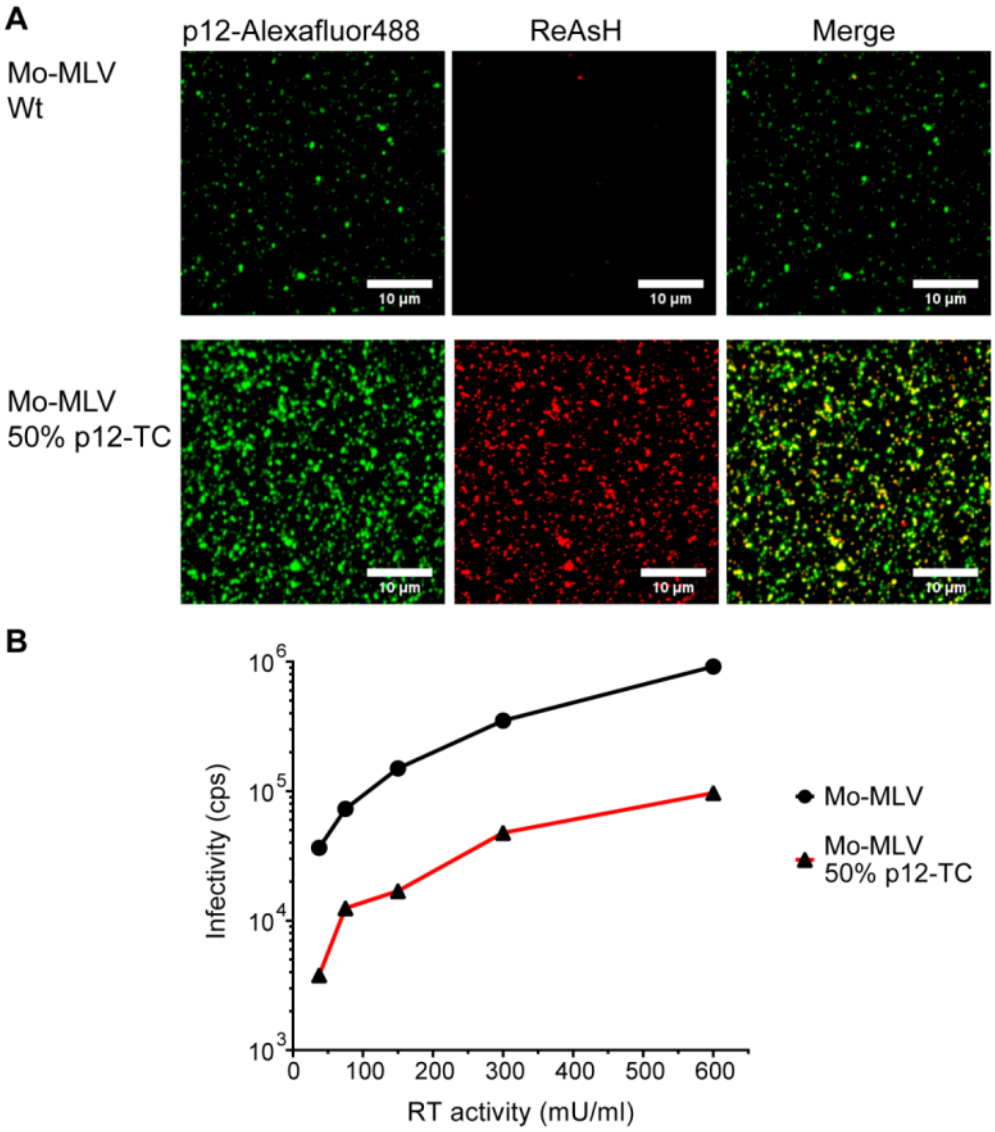


Figure 6.13: ReAsH labelling efficiency of p12-TC in Mo-MLV VLPs using an *in vivo* labelling protocol. Mo-MLV containing wild type p12 or a 50/50 mixture of wild type p12 and p12-TC were synthesised in the presence of 2.5µM ReAsH-EDT$_2$. (A) VLPs were harvested, concentrated, spun onto coverslips and analysed by indirect immunofluorescence with a p12 monoclonal antibody. A representative image of each sample is shown. (B) A two-fold dilution series (based on the RT activity) of the VLPs was used to infect D17 cells. The infectivity is displayed as the luminescence counts per second against the RT activity of the VLPs.

6.3.2 Labelling CA with a fluorescent protein to follow the PIC during infection

As labelling p12 using a TC tag based method did not work well, labelling of the core using a CA-YFP fusion protein was attempted. An expression vector coding for a modified B-tropic MLV Gag protein, where the last four amino acids of CA and the whole of NC have been replaced with YFP, was utilised. Mo-MLV particles were labelled with increasing amounts of this modified fluorescently labelled Gag. VLP release was quantified by measuring the RT activity of the labelled viruses (Figure 6.14A). Labelling with a small amount of the CA-YFP protein (5% transfected), did not grossly affect Mo-MLV VLP release (Figure 6.14A). When more of the Gag in the mixed VLP came from the modified CA-YFP plasmid, the release of VLPs was significantly reduced (Figure 6.14A, greater than 10%). This reduction in RT activity is probably in part due to less Gag-Pol being present, as the modified Gag plasmid does not encode *pol*. As CA is required for Gag-Gag association in MLV, which helps to drive assembly, the Gag-Gag interactions are likely disturbed to some degree when higher proportions of the CA molecules are fused to YFP. The infectivity of the labelled VLPs was assessed and remained high until 50% of the Gag came from the CA-YFP plasmid (Figure 6.14B). Thus, the released VLPs produced using this labelling method were viable.

To test the percentage of Mo-MLV VLPs labelled with CA-YFP, they were spun onto coverslips and analysed by immunofluorescence with the anti-p12 monoclonal antibody. The amount of p12 positive and YFP positive puncta were scored and displayed as the percentage of the total p12 positive puncta (Figure 6.14C). The optimal labelling of Mo-MLV VLPs occurred when 10% of the Gag transfected came from the CA-YFP modified Gag vector (Figure 6.14C). The percentage of puncta with co-localised p12 and YFP was high for all the mixed VLPs, when at least 10% of the transfected Gag was from the CA-YFP plasmid (Figure 6.14C). The slight caveat of this analysis is that all CA-YFP puncta should stain for p12, as the CA-YFP modified Gag contains p12. Thus some of the co-localised puncta will likely represent non-infectious particles. Importantly, VLPs containing no CA-YFP can be discerned by this method. Therefore the Mo-MLV VLPs were labelled efficiently with CA-YFP even when low levels of the labelling plasmid were co-transfected.

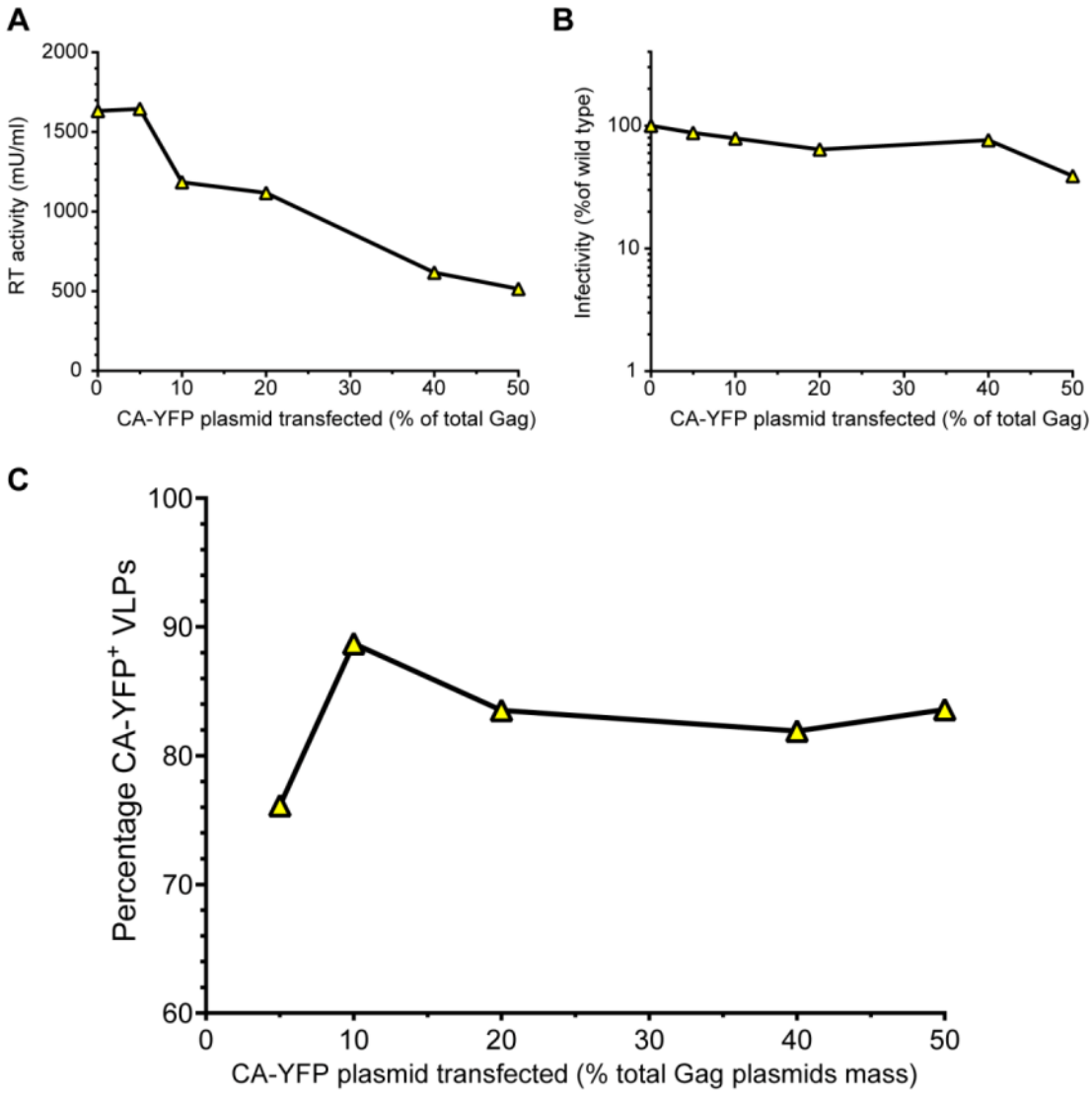


Figure 6.14: The effect of labelling Mo-MLV with a B-tropic CA-YFP fusion protein. Mo-MLV VLPs were labelled with CA-YFP by co-transfecting with the B-tropic CA-YFP modified Gag plasmid (pB30Y) during VLP synthesis. (A) VLP release is shown as the amount of RT activity released from producer cells for increasing amounts of CA-YFP *gag* transfected. (B) D17 cells were infected with equal RT-units of VLPs and the infectivity quantified. The infectivity is plotted as a percentage of Mo-MLV with no CA-YFP. (C) VLPs were spun onto coverslips and analysed by indirect immunofluorescence using a p12 monoclonal antibody. The percentage of p12 puncta positive for YFP was quantified and is displayed in the graph. The percentage of CA-YFP modified Gag transfected (along the x-axis), is expressed as a percentage of the total mass of the Gag containing plasmids transfected.

6.3.3 Labelling p12 with GFP to follow the PIC during infection

A previous attempt to label p12 by cloning the 16 amino acid fragment of the split GFP system (Cabantous, Terwilliger, and Waldo, 2005) into the central portion of p12 lead to significantly reduced infectivity (Chapter 3). However, the full length GFP molecule has been successfully attached to the N-terminus of p12, preserving the protease

cleavage site, in a modified Mo-MLV Gag expression vector (Elis et al., 2012). This GFP-p12 fusion has been used to label an infectious clone of Mo-MLV and visualise the incoming PIC in real-time during infection (Elis et al., 2012). Increasing amounts of this modified GFP-p12 Gag plasmid were used to label the Mo-MLV VLPs with GFP-p12.

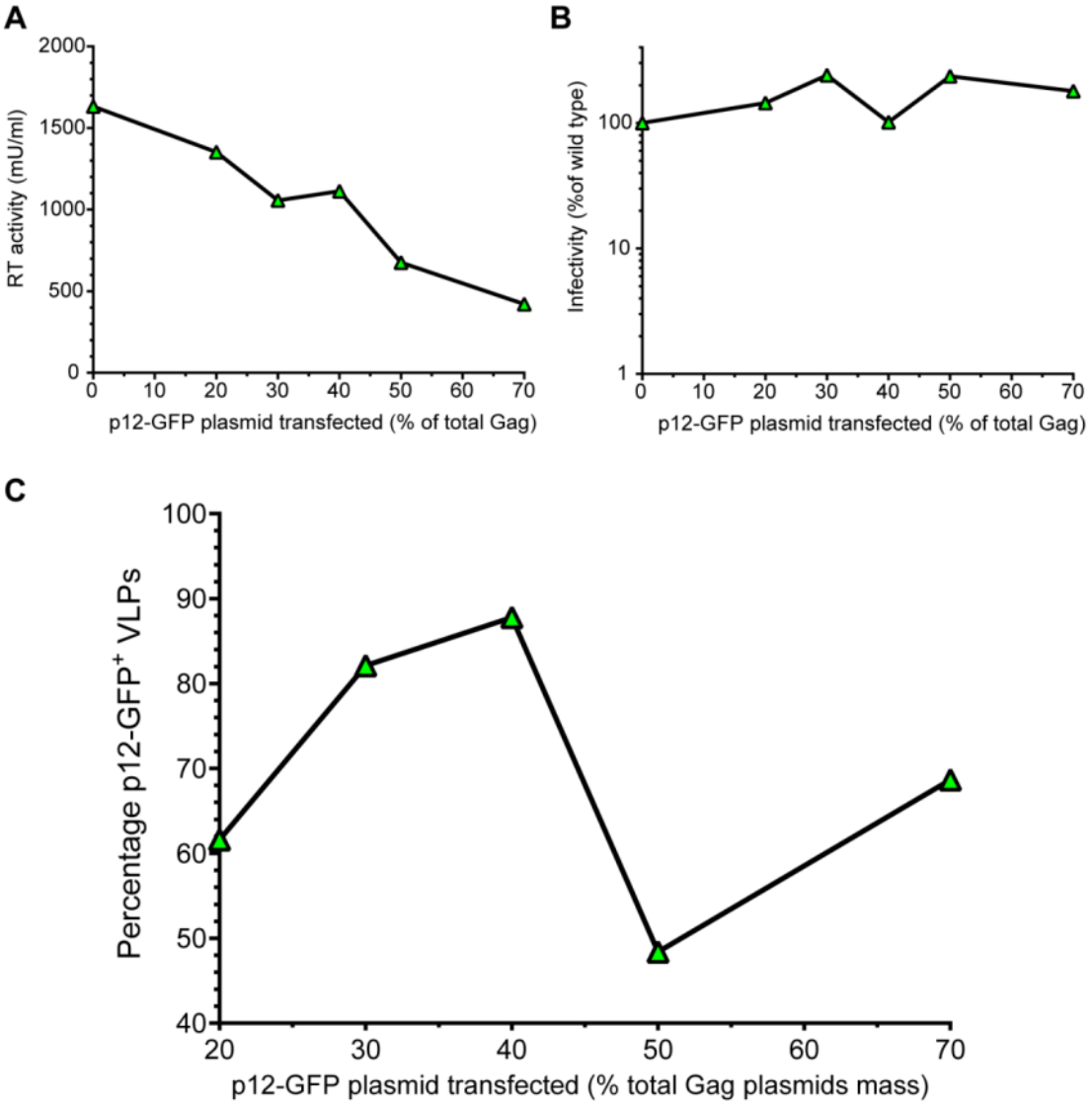


Figure 6.15: The effect of labelling Mo-MLV with eGFP-p12. Mo-MLV VLPs were labelled with eGFP-12 by co-transfecting with the MLV eGFP-p12 modified Gag plasmid during VLP synthesis. (A) VLP release is shown as the amount of RT activity released from producer cells for increasing amounts of GFP-p12 modified Gag transfected. (B) D17 cells were infected with equal RT-units of VLPs and the infectivity quantified. The infectivity is plotted as a percentage of Mo-MLV with no eGFP-p12. (C) VLPs were spun onto coverslips and analysed by indirect immunofluorescence using a p12 monoclonal antibody. The percentage of p12 puncta positive for GFP were quantified and displayed in the graph. The percentage of GFP-p12 modified Gag transfected (along the x-axis), is expressed as a percentage of the total mass of the Gag containing plasmids transfected.

The release of VLPs decreased in a linear fashion as the amount of the modified GFP-p12 Gag transfected was increased (Figure 6.15A). As with the CA-YFP modified Gag, this will be in part due to less Gag-Pol present in the VLPs, as the modified Gag does not encode *pol*. Labelled VLPs were as infectious as wild type even when 70% of the Gag came from the GFP-p12 plasmid (Figure 6.15B). Thus, while too much GFP-p12 Gag in the VLP appeared to have an effect on VLP synthesis/release, the produced VLPs were viable.

The efficiency of labelling VLPs, with the GFP-p12 modified Gag, was assessed by immunofluorescence with an anti-p12 monoclonal antibody on the VLPs. This analysis also suffers from the drawback that all GFP positive puncta will stain for p12, although some of these puncta may be solely composed of the modified Gag and are not viable VLPs. Regardless, when 30-40% of the Gag transfected came from the GFP-p12 modified Gag, the majority of VLPs contained measurable GFP-p12 signal (Figure 6.15C). Thus, this method could efficiently label Mo-MLV VLPs using GFP-p12, with minimal effect on virus viability.

6.4 The movements of Mo-MLV p12 and the PIC during the early stages of replication

To establish if any defect could be seen with the N-terminal p12 mutants, I set out to study the movements of the Mo-MLV PIC immediately after entry to the cell. I initiated this work during my visit to Eran Bacharach's laboratory in Israel. Using their expertise on following the movements of the Mo-MLV PIC in infected cells (Elis et al., 2012; Prizan-Ravid et al., 2010), two experimental approaches were attempted to analyse the N-terminal p12 mutants. One used synchronised infections of cells combined with immunofluorescence and the other using live imaging of GFP-p12 or CA-YFP labelled Mo-MLV in infected cells.

6.4.1 Following p12 labelled with GFP during the early stages of infection

GFP-labelled p12 in Mo-MLV has been successfully utilised to document the docking of p12 to chromatin in infected mitotic cells (Elis et al., 2012). Thus, attempts were

made to observe the movements of the N-terminal p12 mutants immediately after the release of the Mo-MLV cores into the cytoplasm, using live fluorescence microscopy. Wild type and N-terminal p12 mutant Mo-MLV VLPs labelled with GFP-p12 and containing an ecotropic Env were synthesised for use in these experiments. U/R-H2A-RFP cells, expressing histone 2A fused to RFP and the ecotropic MLV receptor (mCAT-1), were transfected with S15-mCherry to mark out the cell membrane. After 16 hours, cells were placed on the microscope and infected with GFP-p12 VLPs (MOI ~8, the wild type and p12 mutant 5 infections were MOI ~5). The imaging began immediately, looking for cells that contain GFP-p12 puncta in the cytoplasm.

Three types of intracellular movement were seen (i) direct movements, (ii) fast uncoordinated movements and (iii) still, shaky movements (Figure 6.16 and appendix D1). When the N-terminal p12 mutants were analysed, the same movements could be seen for all of them, although there appeared to be an increase in the amount of stationary puncta observed for p12 mutant 5 and 6 compared to wild type (appendix D2 versus D4, D6 and D8). Virtually all Mo-MLV GFP-p12 VLPs lacking an envelope protein (non-specifically taken up in vesicles) were stationary with occasional direct movements (appendix D10).

After imaging, cells were returned to the incubator and imaged the next day to see if a defect had become apparent for the p12 mutants. Interestingly, the N-terminal p12 mutants 5 and 6 displayed much less GFP-p12 puncta per cell when imaged between 20-25 hours post infection, and appeared to be less mobile when compared to wild type Mo-MLV (appendix D3 versus D5 and D7). However, p12 mutant 7 looked more like wild type than p12 mutant 5 or 6 (appendix D3 versus D9). VLPs non-specifically taken up into cells had formed large GFP aggregates in the perinuclear region (appendix D11), and were likely in the process of being degraded by the proteasome. In conclusion, this preliminary analysis showed that the p12 signal for the N-terminal mutants 5 and 6 appeared to diminish faster in infected cells than wild type. However, the phenotype appears quite subtle and would require analysis of a large number of cells to make the observation significant.

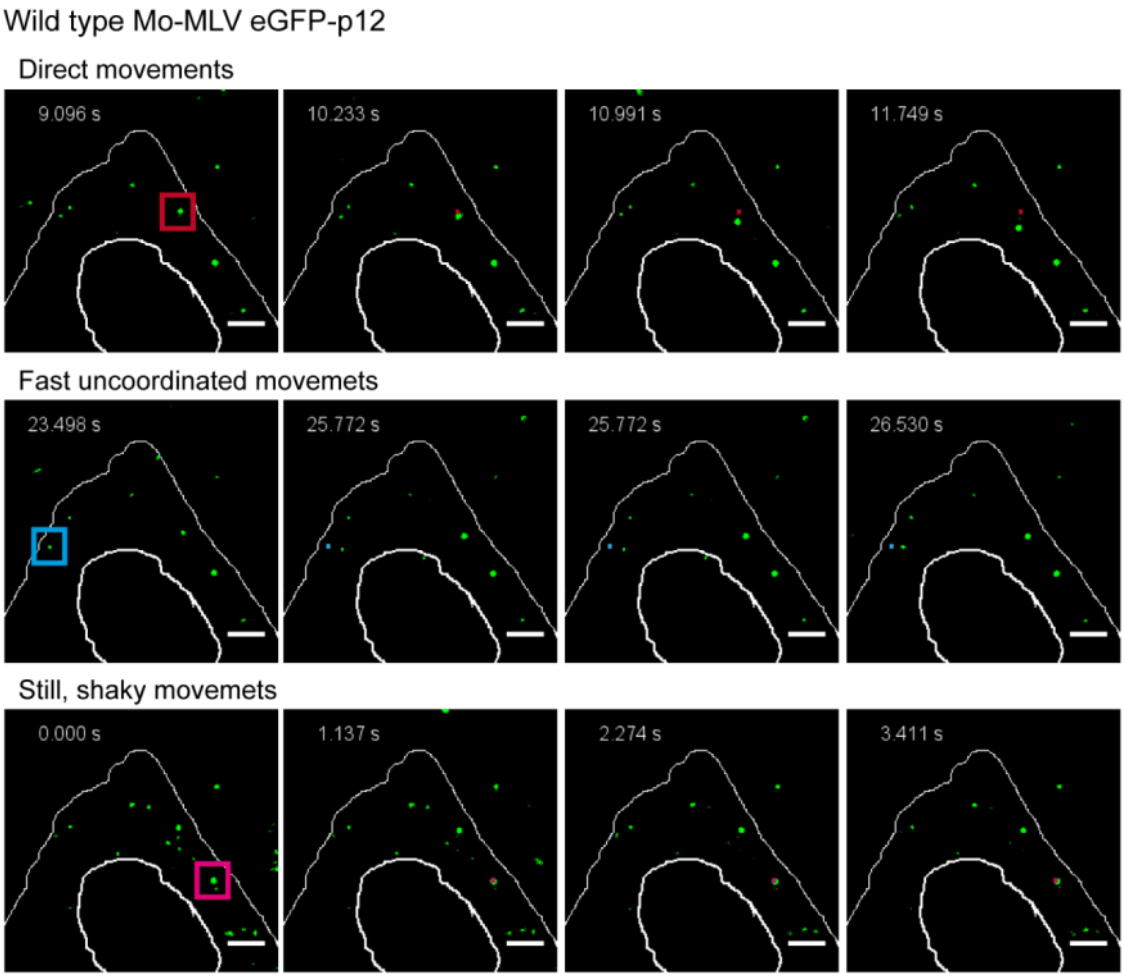


Figure 6.16: Intracellular Movements of Mo-MLV GFP-p12 puncta immediately after infection. U/R-H2A-RFP cells, transfected with S15mCherry, were infected with Mo-MLV GFP-p12 at an MOI ~5. Images were taken on a spinning-disk confocal microscope equipped with an EMCCD and scaled to highlight the signal and remove background noise. Displayed are movie frames showing the three typical intracellular movements of the GFP-p12 puncta (highlighted by a coloured box in the first frame): direct movements (top panels, red highlight), fast uncoordinated movements (middle panels, blue highlight) and still shaky movements (bottom panels, magenta highlight). The coloured dot in each frame indicates the starting position of the puncta in the first frame for comparison. The time (s) is displayed in the top left-hand corner. Scale bars- 5µm.

Interestingly, the infectivity of p12 mutant 5 and 7 was higher when the virus contained GFP-p12 (with the corresponding p12 mutation) compared to the untagged mutants (Figure 6.17), indicating that the fusion of GFP to p12 partly rescued the infectivity defect observed. This was not observed for p12 mutant 6, 8 or 14 (Figure 6.17).

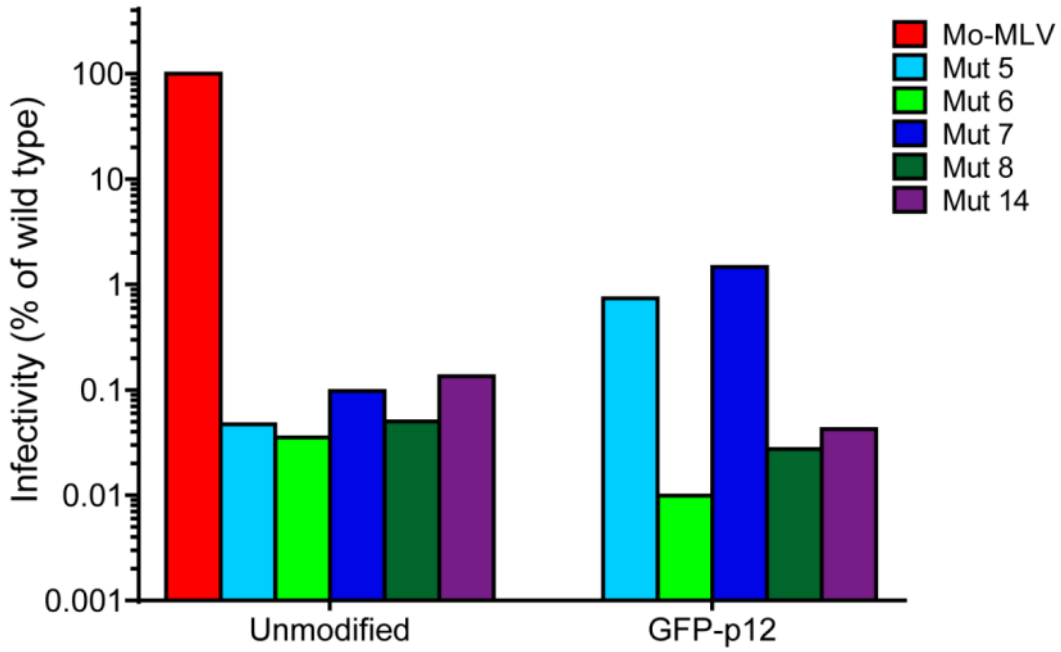


Figure 6.17: Infectivity of Mo-MLV p12 mutants containing GFP-p12. U/R cells were infected with p12 mutants with (GFP-p12) and without (unmodified) GFP-p12 and the infectivity was quantified. The infectivity of VLPs are shown as a percentage of the infectivity of wild type Mo-MLV without GFP-p12.

6.4.2 Following the Mo-MLV PIC labelled with CA-YFP during the early stages of infection

To validate the use of the CA-YFP for labelling the Mo-MLV PIC (from section 6.3.2), cells were infected on the microscope as above. Imaging of Mo-MLV labelled with CA-YFP revealed particles internalised in cells that displayed similar movements to the Mo-MLV particles labelled by GFP-p12 (appendix E1 versus D1 and D2). This indicated that the CA-YFP labelling technique did indeed label cores that enter the cell. Interestingly, and surprisingly, in mitotic cells, CA-YFP was seen docked on the chromosomes early after infection (Figure 6.18A and appendix E2). This was unexpected, as it had been assumed that little CA remains with the MLV PIC that associates with the chromatin. When the infection was observed again, after 18 hours, CA-YFP could again be seen attached to the condensed chromatin in mitotic cells (appendix E3).

GFP-p12 from Mo-MLV was shown to dock on condensed chromatin in mitotic cells after 12 hours, but has not been investigated at such an early time post-infection (Elis et al., 2012). Therefore, GFP-p12 labelled Mo-MLV VLPs were used to infect U/R-H2A-

RFP cells on the microscope and mitotic cells were imaged immediately after infection. As was seen for Mo-MLV labelled with CA-YFP, GFP-p12 was observed docked onto the condensed chromatin in mitotic cells at an early time point (Figure 6.18B and appendix E4). When p12 contained the mutation 14, neither the CA-YFP nor the GFP-p12 puncta associated with the condensed chromatin in infected cells, as previously reported (data not shown) (Elis et al., 2012). Thus, this early chromatin binding of the Mo-MLV PIC labelled with GFP-p12 or CA-YFP was mediated by p12 and some CA may remain with the nuclear associated PIC.

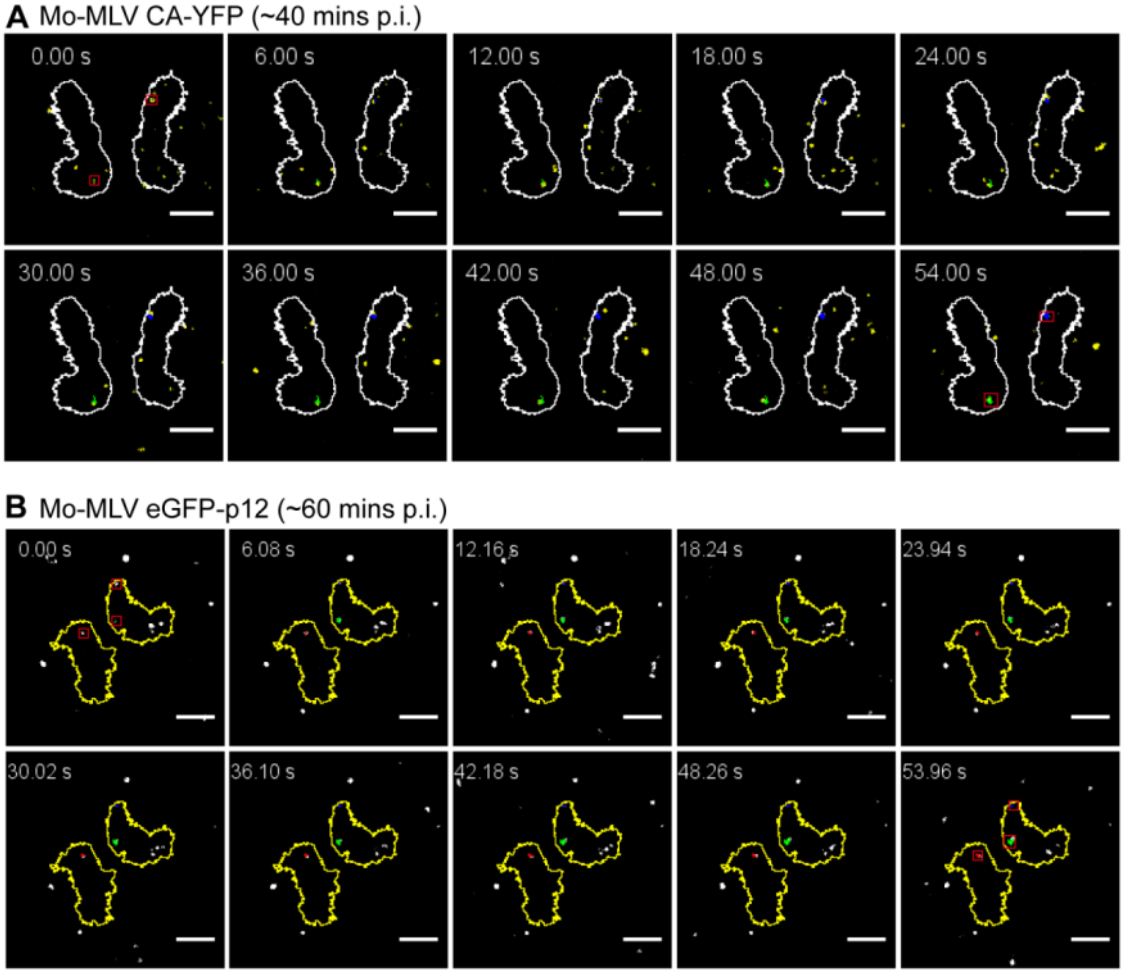


Figure 6.18: Docking of Mo-MLV GFP-p12 and CA-YFP on mitotic chromosomes early after infection. U/R-H2A-RFP cells were infected on the microscope with (A) Mo-MLV CA-YFP and (B) Mo-MLV GFP-p12 VLPs and mitotic cells were imaged. Displayed are time frames from representative infected cells. The docked puncta are marked out by red boxes and the tracks are progressively coloured with green, blue or red lines. The time (s) is displayed in the top left-hand corner. Scale bars- 5μm.

6.4.3 Early post-entry movements of the MLV PIC in infected cells analysed using immunofluorescence

Tagging proteins with fluorescent proteins can significantly alter their behaviour especially with regards to the their turnover in cells. Indeed, the infectivities of p12 mutants 5 and 7 tagged with GFP-p12 were higher than the untagged p12 mutants (Figure 6.17), indicating that the attachment of GFP to these p12 N-terminal mutants restored some p12 function. As the N-terminal p12 mutants had a defect in formation of the mature core (Chapter 5), indirect immunofluorescence was used to analyse the movements of p12 and CA immediately after core release into the cytoplasm. To normalise the detection of p12, a myc-tagged p12 (described previously, (Prizan-Ravid et al., 2010)) was cloned into the Mo-MLV Gag-Pol vectors and ecotropic Mo-MLV VLPs were synthesised. U/R cells were infected (MOI~3) at 4^{o}C using spinoculation (O'Doherty, Swiggard, and Malim, 2000) (to synchronise the infection), fixed and stained for p12 (myc) and CA using indirect immunofluorescence at various times post infection.

It should be noted that the immunostaining with the anti-CA antibody utilised in this set of experiments had a lot of non-specific background, and thus a false positive rate in uninfected cells of 1-2 CA puncta could be seen (data not shown). However this rate was not uniform and many uninfected cells contained no false positive CA signal. The highest background signal was seen around the nucleus (Figure 6.19A, top panels).

Cells infected with wild type Mo-MLV, at an early time point, contained mostly p12-myc puncta co-localised with the CA puncta (Figure 6.19A, top panels). As time progressed, less puncta were seen in the cells but most of the puncta remained both p12-myc and CA positive (Figure 6.19 B and C, top panels). In stark contrast, the N-terminal p12 mutants all contained very little p12-myc signal, despite the presence of CA signal even at the very early time points (Figure 6.19A). As time progressed this worsened until only a few CA puncta remained (Figure 6.19 B and C). Viruses attached to the coverslips could be found for the p12 mutants and these stained with both the p12-myc and CA antibodies, indicating the p12 mutations did not alter the antigenic properties of the myc tag (Figure 6.19B, Mut 6 yellow box). Thus, mutation of the N-terminus of p12 resulted in the loss of p12 from the internalised core (or aggregates of CA in the case of p12 mutant 6) very early after release into the cytoplasm.

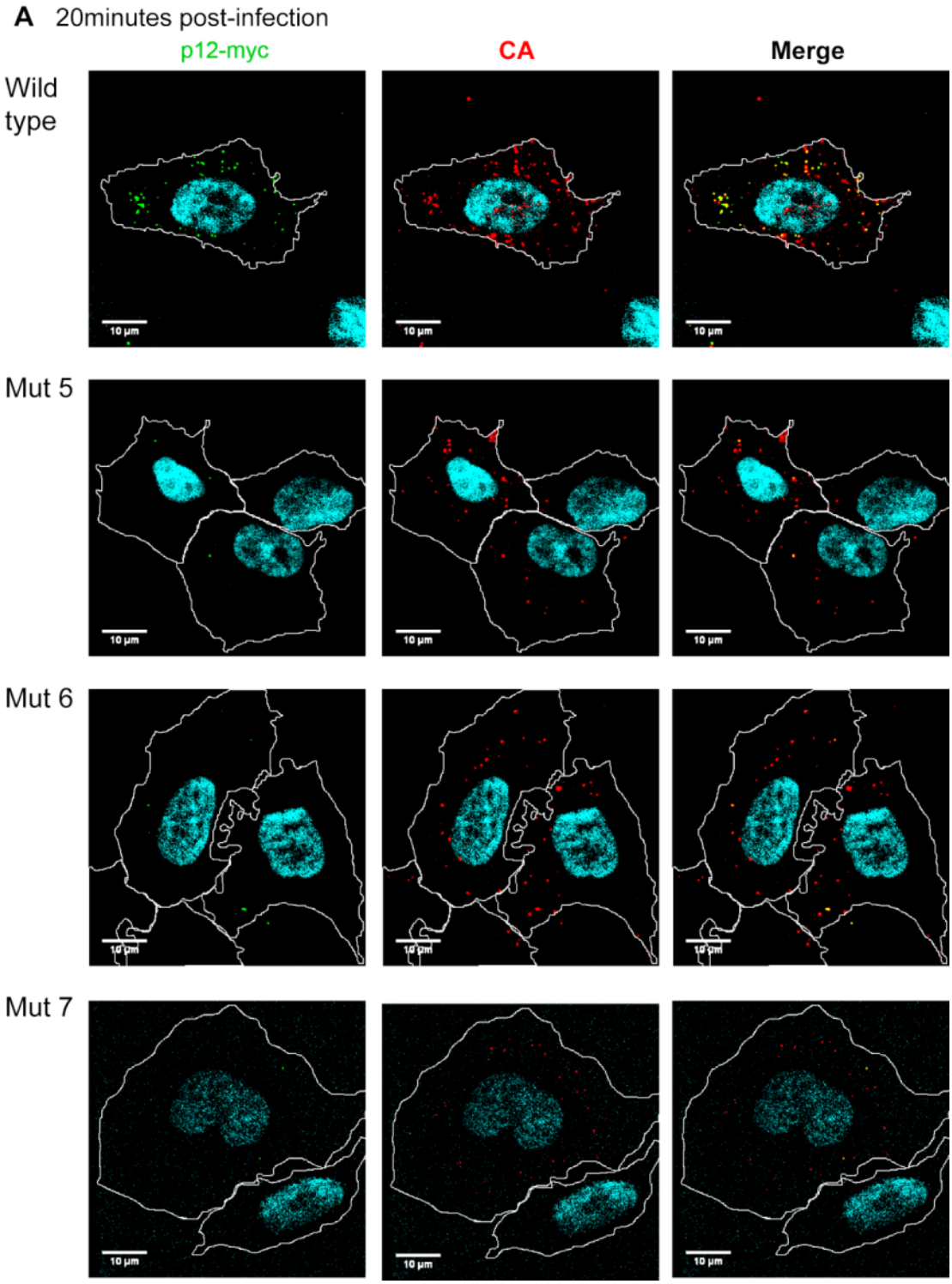
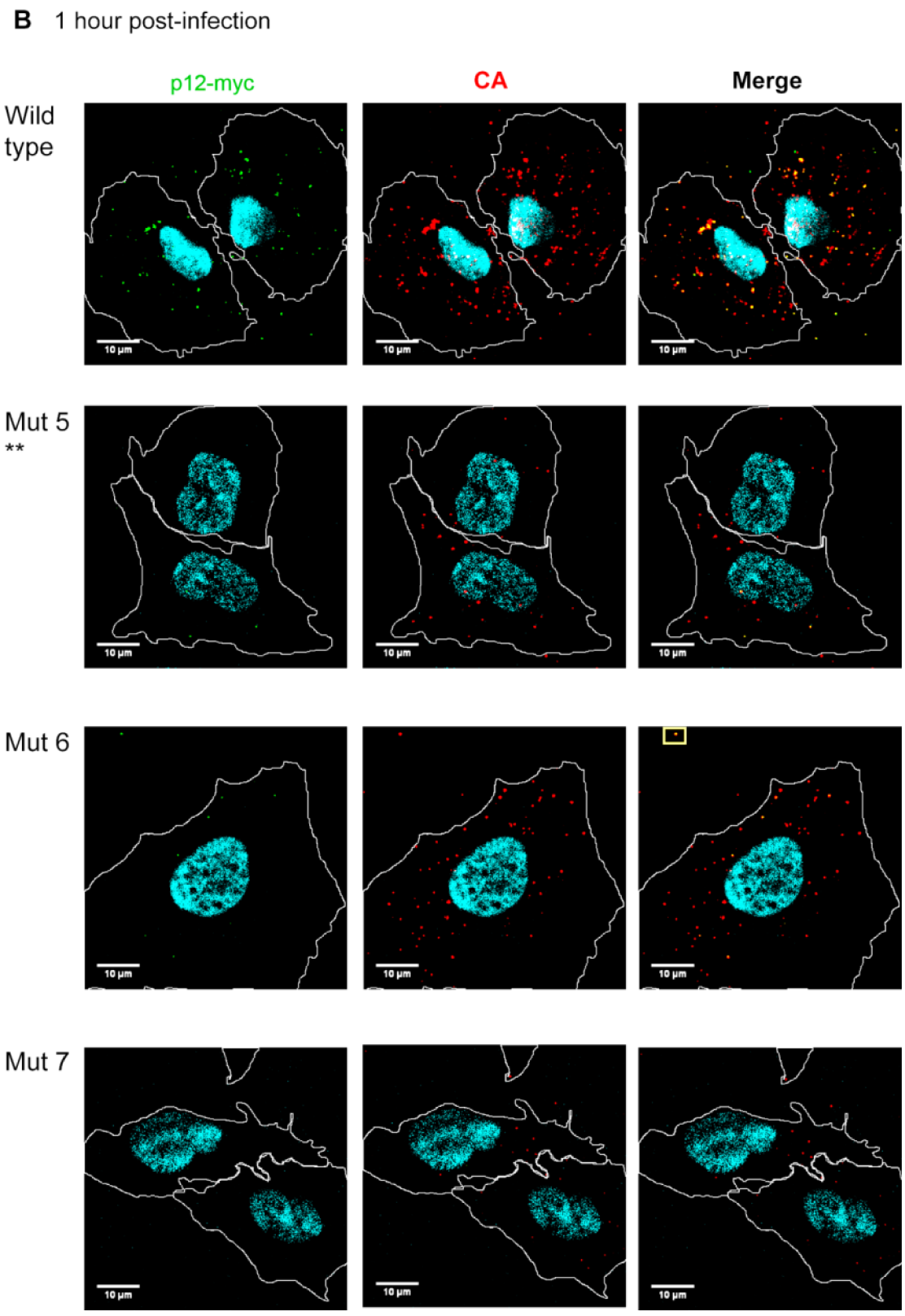
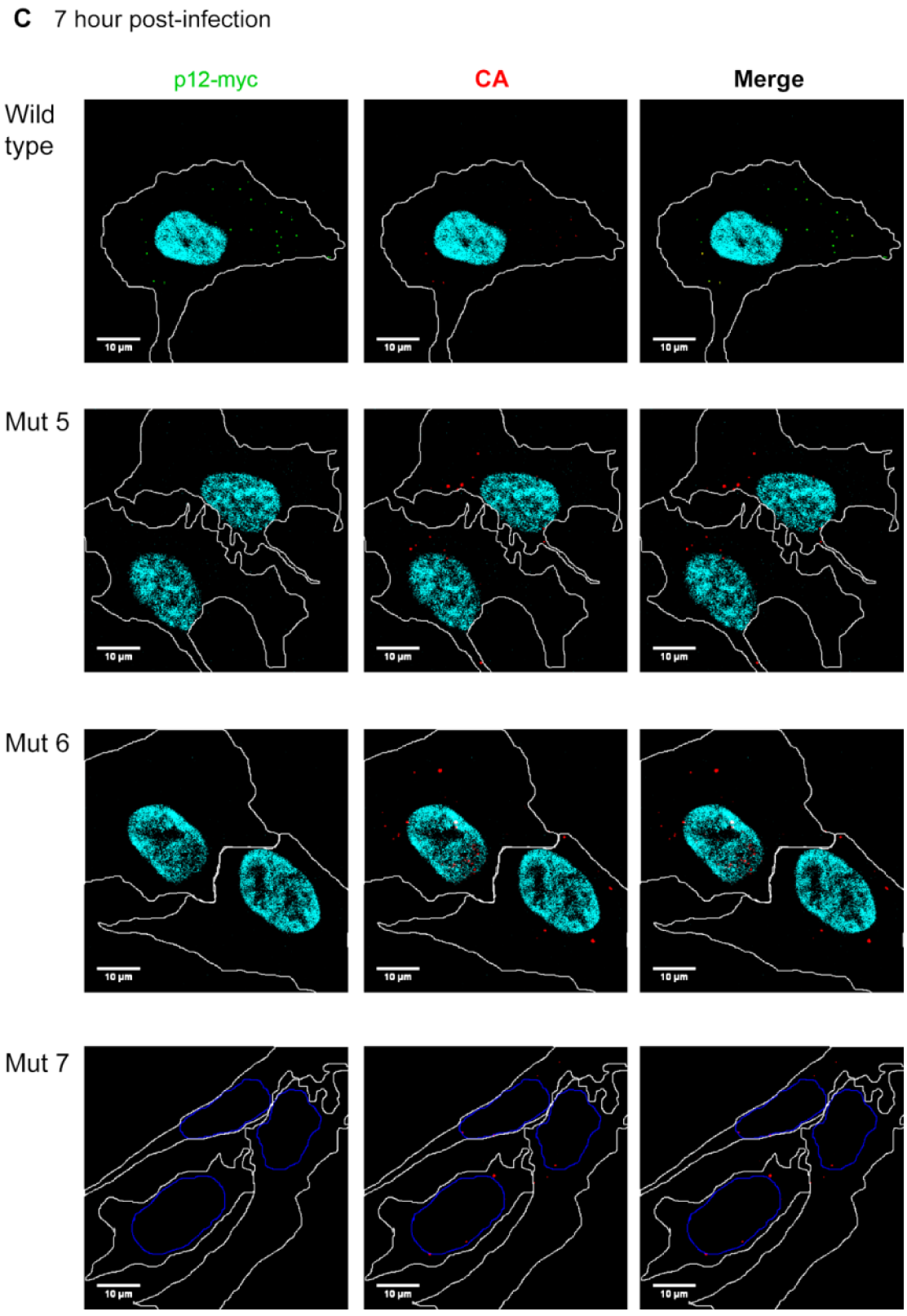


Figure 6.19: Movements N-terminal p12 mutants PICs in infected cells. Cells were infected with p12-myc Mo-MLV VLPs (MOI~3) at 4°C to synchronise the infection. Cells were fixed after (A) 20 minutes, (B) 1 hour and (C) 7 hours post-infection and immunostained with antibodies against the myc-tag (p12) and CA. Cells were imaged using a spinning disk confocal microscope with the same exposure settings for each staining. Representative images are shown with the fluorescence intensity scaled by the same values. The yellow box in (B) indicates a p12 mutant 6 VLP bound to the coverslip. ** Mutant 5 was fixed after 3 hours in this example. A trace of the nucleus is shown (blue lines) for mutant 7 in (C) as the dapi staining was weak causing a high background from the dapi channel, due to technical problems with the UV lamp.


B 1 hour post-infection
p12-myc
CA
Merge
Wild type
Mut 5 **
Mut 6
Mut 7
10 µm


C 7 hour post-infection
p12-myc
CA
Merge
Wild type
Mut 5
Mut 6
Mut 7
10 µm

6.5 Summary

In this chapter, the localisation of p12 expressed in cells and in cells infected with Mo-MLV was investigated. Expression of mCherry-p12, or the N- or C-terminal regions of p12 fused to mCherry, produced a stable fusion protein in mammalian cells but had no specific sub-cellular localisation, indicating that p12 does not confer a strong karyophilic influence on mCherry. As p12 has been shown to bind to condensed chromatin in mitotic cells, as part of the Mo-MLV PIC (Elis et al., 2012; Prizan-Ravid et al., 2010), the influence of the cell cycle on the localisation of mCherry-p12 was investigated. In contrast to a very recent report (Schneider et al., 2013), significant accumulation of mCherry-p12 was not seen on the mitotic chromatin. The discrepancy between the findings of Schneider *et al.* and those presented here is likely down to technical differences, which will be discussed in Chapter 7.

Unfortunately a high level of non-specific uptake (via Env-independent mechanisms) was seen in cells infected with Mo-MLV VLPs. This non-specific uptake has been widely documented and is a significant problem for the analysis of microscopy data (Marechal et al., 1998; McDonald et al., 2002). Interestingly, the non-specific uptake of Mo-MLV virus into cells lacking the appropriate receptor appeared low for Prizan-Ravid *et al* (Prizan-Ravid et al., 2010). Indeed, non-specific uptake of Mo-MLV VLPs with no envelope into cells was significantly higher than VLPs that contain an envelope protein, with no receptor present on the cells. Therefore the presence of an Env protein on the surface of the virion renders it less likely to be non-specifically taken up into cells. Whether this is a cell/Env type-specific phenomenon will require investigations with more combinations of cells and different Env proteins. Nevertheless, a large number of particles remain non-specifically taken up into cells, even if the use of an Env negative control over estimates the problem. Thus, methods to rule out Mo-MLV on a non-productive infection pathway were assessed.

Firstly, the detection of viral DNA and viral proteins in infected cells using combined immunofluorescence and FISH was attempted. Unfortunately, detection of the viral DNA using FISH was never successful due to a high background that resulted from non-specific binding of the avidin-fluorophore to some abundant biotin-like molecule in the cell. Blocking the cells with avidin did not reduce this background. Due to the unsuccessful nature of the preliminary experiments, and the time required to perform a single FISH experiment, further attempts were not pursued. A protocol using a

digoxigenin (DIG) labelled probe to detect the MLV viral DNA in infected cells has been more successful and could potentially be used in future studies (Sada Ohkura, personal communication).

Secondly, labelling of the Mo-MLV viral membrane with S15mCherry was attempted. This method has been used to distinguish viral particles that have undergone fusion and lost their viral membrane from particles non-specifically taken up into cells (Campbell et al., 2007). Unfortunately, co-expression of the S15mCherry plasmid with the Mo-MLV VLP plasmids caused a dramatic reduction in VLP release and affected the infectivity of any particles that were produced. Using very low levels of the S15mCherry plasmid resulted in wild type-like release and infectivity of Mo-MLV, but very few S15mCherry labelled particles could be detected by microscopy. Only HIV-1 control VLPs contained a significant number of labelled particles, although the labelling efficiency was lower than previously reported (Campbell et al., 2007). Consequently, Mo-MLV could not be successfully labelled with the S15mCherry molecule, indicating that Mo-MLV may bud from subtly different microdomains in the membrane compared to HIV-1, which have lower levels of the S15mCherry protein.

A small tetracysteine motif was cloned into p12 for labelling with a fluorescent dye (FlAsH or ReAsH) to follow the movements of the Mo-MLV PIC during infection. The *in vitro* labelling method described by Arhel *et al.* for labelling HIV-1 IN (Arhel and Charneau, 2009) was unsuccessful for labelling Mo-MLV containing the p12-TC, but a high level of labelling could be achieved when the VLPs were synthesised in the presence of ReAsH (*in vivo* labelling). Unfortunately, loading the dye onto p12 reduced infectivity by 10-fold compared to control Mo-MLV VLPs without p12-TC. Thus, similar to the result seen with the split GFP motif in p12 (Chapter 3) the ReAsH dye in the central 'linker' region of p12 significantly reduced infectivity. The binding of FlAsH or ReAsH to the TC motif induces the formation of a kink to allow disulphide bonds between the dye and the cystienes to form, hinting that structural alteration to the 'linker' region in p12 is not compatible with p12 function.

Labelling of Mo-MLV VLPs using either the GFP-p12, used by Elis *et al.* (Elis et al., 2012), or a CA-YFP molecule, was much more successful, with no significant effect on VLP release or infectivity at levels sufficient for labelling. These labelled VLPs were then used in live imaging experiments with our collaborators in Israel (Eran Bacharach's

laboratory). Live imaging of cells infected with GFP-p12 VLPs displayed three types of intracellular movements (i) direct movements, (ii) fast uncoordinated movements and (iii) still, shaky movements. Particles displaying movements of type (i) and (iii) could change between the two and frequently did, whereas GFP-p12 particles showing type (ii) movements always displayed this type of movement over the window of observation. Non-specifically endocytosed virus had a phenotype that was almost entirely of type (iii) based movements with occasional type (i) movements from some particles. I would suggest that the very fast uncoordinated movements (type ii) arise from free aggregates of protein including GFP-p12, which are small enough to pass about by diffusion within the cell, as their velocities were greater than fast microtubule based transport. In contrast, the slower and direct movements are likely microtubule based transport, which have been reported for HIV-1 (Arhel et al., 2006; McDonald et al., 2002). While lots of these type (ii) movements could be seen for wild type Mo-MLV GFP-p12 20 hours post-infection, little could be seen in the cells infected with p12 mutant 5 and 6 at the later time points. This suggested for these mutants that the majority of GFP-p12 that had entered the cells had been degraded by the later time point. Due to the small difference in phenotype observed with the Mo-MLV GFP-p12 mutants using live imaging, it is unlikely that this approach will shed new light on the N-terminal p12 mutant defect. Strangely, p12 mutants 5 and 7 had a higher infectivity when the VLPs contained the GFP-p12 and I will discuss this further in Chapter 7.

Unexpectedly, the live imaging of CA-YFP labelled Mo-MLV revealed it to be bound to condensed chromatin in a p12-dependant manner. This tethering could be visualised even at very early time points when Mo-MLV PICs were labelled using either GFP-p12 or CA-YFP. These early bound puncta, if they do represent PICs, will not likely have had time to reverse transcribe their genome, indicating that reverse transcription is not important for chromatin tethering. Indeed, p12-mediated chromatin tethering does not require packaging of the viral genome (Efrat Elis, personal communication). Thus, surprisingly, the data shown here suggests that a significant proportion of CA may remain with the MLV nuclear-associated PIC, although whether the CA-YFP fusion stabilises CA and causes an exaggerated phenotype will require further investigation.

Finally, indirect immunofluorescence was used to study the movements of p12 and CA within infected cells over a time course, and strikingly the signal for p12-myc was found to diminish rapidly after infection for the N-terminal p12 mutants (mutant 5, 6

and 7) compared to wild type Mo-MLV. However, the CA signal for the p12 mutants appeared to persist longer. This striking phenotype indicates that the observations with the GFP-p12 labelled PICs were indeed somewhat artifactual, in that the GFP-p12 signal could persist much longer inside infected cells than the 'native' p12 species (with a small tag). While this analysis should be regarded as preliminary, and quantification is required to make a solid conclusion, it looks as though mutation of the N-terminus of p12 does indeed result in the loss of p12 from the PIC early after infection. This suggests that the interaction at the N-terminus of p12 keeps p12 with CA and the PIC during the early steps of replication.

Chapter 7.

Discussion

Mo-MLV p12 has been shown to be an essential Gag protein. It is required at both early and late stages of the viral life cycle (Yuan, Li, and Goff, 1999), although the function of p12 during the early infection steps has remained enigmatic. The work presented in this thesis, and recently published data, have highlighted functions of p12 essential for progression through the early stage of infection. Two domains have been described in Mo-MLV p12, which are both required to be active on a single p12 molecule for infection. All gammaretroviruses analysed have a conserved requirement for p12 function during the early stages of infection, indicating that p12 function has been fixed in the gammaretroviral lineage. Importantly, using a combination of genetic and biochemical techniques, the N-terminal domain of p12 has been shown to function in formation of the mature CA core during viral maturation before entry to the target cell. Genetic approaches have shown that the C-terminus of p12 is vital for tethering the nuclear-associated MLV PIC to condensed chromatin. This chapter will discuss these findings in more detail, with an emphasis on their wider implications on the replication of retroviruses.

While the findings presented here have assigned important roles to p12 during infection the mechanism of p12 action remains ill defined. I will, therefore, finish by highlighting my model for p12 function and some key experimental approaches that will validate and expand on this model.

7.1 p12 N-terminal domain function during the early stages of infection

In Chapter 5, key experiments were described that triggered the investigation into the effect of p12 mutations on the structure/stability of the mature CA core. These experiments utilised CA-core sensing restriction factors (Fv1 and TRIM5alpha) to detect the incoming CA core from p12 mutants (Section 5.1). Surprisingly, this revealed that mutation of the N-terminus of p12 caused an inability to abrogate restriction factors. A few possible explanations exist as to why these p12 mutant cores failed to abrogate restriction: (i) inhibition of restriction factor binding to the CA core by mutant p12, (ii) incorrect localisation of the core within infected cells, (iii) the CA monomers in

the core were not processed correctly so that restriction factors could not recognise the core, (iv) the mature CA core was unstable, or (v) no mature CA core was formed inside the virion. These possibilities will be discussed in the following section.

7.1.1 Towards the cause of the p12 mutant abrogation defect

A mechanism where the N-terminal p12 mutants could inhibit restriction factor interaction with the CA core would be relatively complex. How this inhibition would also cause the infectivity defect for Mo-MLV is also unclear. It would, therefore, seem improbable that the N-terminal p12 mutants could block restriction factor interaction with the CA core and is unlikely the cause of the abrogation defect. Mislocalisation of the p12 mutant core within infected cells also seems less likely. Indeed, MLV that enters the cell using different entry pathways is restricted by Fv1 (Yap and Stoye, 2003). Thus, bypassing fusion at the plasma membrane (and fusing through endosomes in the cytosol) does not permit escape from restriction factors, suggesting a certain level of plasticity in their ability to 'locate' incoming CA cores. I have also shown that the CA monomers liberated from a Gag molecule containing a p12 mutation *in cis,* can be incorporated into a mature core recognised by restriction factors (Figure 5.5). Thus, p12 mutations upstream in Gag did not alter processing of CA from the Gag polyprotein, or the subsequent conformational changes required to form the mature core.

This leaves two possible causes for the abrogation defect: either the core is unstable or it is poorly formed/absent. *In vitro* disassembly assays have been used to assess the stability HIV cores with mutations in CA (Forshey et al., 2002). This assay requires isolation of cores from virions, by detergent removal of the membrane. The success of this purification for HIV-1 has a low yield (~16%) of cores, and cores from MLV appear to be less amenable to isolation by this method (Fassati and Goff, 1999). Cores isolated from Friend murine leukaemia virus and Rauscher murine leukaemia virus have been reported in the past, but more recent attempts to isolate cores from Mo-MLV using a range of detergents were unsuccessful (Bolognesi, Luftig, and Shaper, 1973; Durbin and Manning, 1982; Fassati and Goff, 1999) The work described in Chapter 5 also documented an inability to isolate whole cores from Mo-MLV virions (Section 5.2.2). However, partial cores could be isolated, migrating to a higher sucrose density than the free CA protein and a lower sucrose density than un-lysed virions (Figure 5.12). The

composition of these partial cores (i.e. how much CA they contain, etc) remains to be determined. This assay was used to test all the p12 mutants, and revealed that the N-terminal p12 mutants' partial cores had a lower apparent S-value than the wild type partial cores (Figure 5.15). Moreover, when Mo-MLV N-terminal p12 mutant RTCs were isolated from infected cells and analysed by velocity sedimentation they also had a lower apparent S-value than wild type RTCs (Boucherit, V. unpublished data). Interestingly, the C-terminal p12 mutants displayed the opposite phenotype, migrating to higher sucrose density fractions than wild type. Thus, either no mature core (but some aggregated form of CA) was present within the N-terminal p12 mutant virions or the core was less stable than wild type cores.

Strikingly, thin-section TEM of purified p12 mutant 6 virions found that relatively few mutant virions showed a recognisable mature core morphology (Section 5.2.3). Many of the particles displayed a smaller punctate mass in place of a mature core, indicating the N-terminal domain of p12 is required for the efficient formation of the mature CA core. This leads to the question: How does p12 aid in the formation of the mature core? Currently no co-factor molecules have been described required in the formation of a retroviral CA core. However, evidence does exist that alteration of CA molecules or addition of small molecules that interact with the CA core can affect core morphology. Auerbach *et al.* have described a similar aberrant core morphology, to the one observed for p12 mutant 6, for an insertion mutant in the second alpha helix of MLV CA-NTD (Auerbach, Brown, and Singh, 2007). From the hexameric structure of MLV CA-NTD, intermolecular interactions occur between alpha helix 2 and 3, and helix 2 and 1, indicating that the mutation described by Auerbach *et al.* may affect hexamer stability (Mortuza et al., 2004). It has also been shown that formation of the mature HIV-1 CA core is disrupted when virus is synthesised in the presence of a CA-targeting drug (PF-3450074) (Blair et al., 2010). The initial study found that the drug was able to promote CA multimerisation (Blair et al., 2010), however, a later report described that the drug destabilised the HIV-1 CA core (Shi et al., 2011). Thus currently, its precise mode of action remains controversial, but it has been shown to bind a pocket in the HIV-1 CA NTD (Blair et al., 2010).

Some mutations have also been described which do not affect CA directly but have an effect on the morphology of the CA core. One such mutation is blocking the PR

processing between NC and p6, which causes an aberrant core morphology despite no alteration to the liberation of mature CA (de Marco et al., 2012). Many explanations exist to the cause of this defect but it does highlight that altering the processing of Gag can have downstream effects on the formation of the mature core. It is quite unlikely that this is the cause of the p12 mutant 6 phenotype as no alteration to the cleaved Gag products is observed (Section 3.1.1). Indeed, completely blocking MA-p12 cleavage only had a minor effect on infectivity suggesting efficient cleavage of p12 from MA is not essential (Section 5.2.1).

Whether the p12 mutant 6 core phenotype is shared with the other N-terminal p12 mutants will required more TEM analysis. Mutant 6 is the only p12 mutant that could not efficiently reverse transcribe *in vivo* (Section 3.1.3), so the defect in core morphology may be more striking for this mutant. However, the biochemical data on the mutant partial cores from sucrose gradients would suggest that all the N-terminal mutants have altered cores. Whether this is due to an equivalent phenotype to p12 mutant 6 or an altered morphology making the cores less stable remains to be established.

7.1.2 The N-terminus of p12 retains p12 within the PIC

Both domains of p12 need to be active on a single p12 molecule for MLV to be infectious. It is known that the C-terminal domain is required for a late stage, tethering the PIC to the chromatin (discussed in Section 7.3), and that the N-terminal domain is required for the formation of the mature core. However, mixed particles containing a mixture of N- and C- terminal p12 mutants can abrogate restriction factors (indicating the CA core is correctly formed) but are non-infectious (Figure 3.9 and Figure 5.6). This implies that the N-terminal domain of p12 has a dual function, and it is this second function that needs to be *in cis* with an active C-terminal domain for MoMLV infectivity. I propose that this second function is in retaining p12 with the PIC and it is likely to be caused by the same interaction that the aids in the formation of the CA core. This notion is supported by fixed immunofluorescence studies of infected cells showing that p12 signal was lost very rapidly from the CA signal for p12 mutant 5, 6 and 7 when inside cells (Section 6.4.3), indicating these p12 mutants could not stay with the core (or malformed mutant core).

7.1.3 Sequence requirements in the N-terminal domain of p12

Individual alanine mutagenesis was performed on p12 from two gammaretroviruses, Mo-MLV and GaLV (Figure 3.15 and Figure 4.6). The N-terminal domains of both Mo-MLV and GaLV p12 contained essential leucine residues. Additionally, aspartic acid residues in the N-terminus of Mo-MLV p12 were also important for p12 function (one in the N-terminus of GaLV p12 also had a minor effect on infectivity when mutated). For both viruses, mutation of larger regions (5-6 alanine substitutions) had a more dramatic effect on infectivity than many point mutants, indicating the combined amino acid composition in the domain is likely important. This is exemplified by Mo-MLV p12 mutant 5, where three individual mutations (K10A, K12A and Q14A) have a minor effect on infectivity, but mutation of the whole region, or all three residues at once, cause a much bigger infectivity defect (Figure 3.15). Moreover, this was also seen with GaLV mutant N, where the D5A only caused a minor infectivity defect but the larger mutation N caused a much greater defect.

Certain regions of p12 were less sensitive to mutation in some gammaretroviruses (Figure 4.3), demonstrating the whole N-terminus is not essential for all gammaretroviruses. The regions that were less sensitive are covered by either p12 mutant 5 or 7 and, interestingly, Mo-MLV containing GFP-p12 with mutant 5 or 7 were more infectious than p12 mutant 5 or 7 without GFP (Figure 6.17). This was not seen for p12 mutant 6, 8 or 14. There are a couple of logical possibilities to the function of GFP in this minor rescue of infectivity: (i) GFP increased p12 (or the N-terminal domain) stability or (ii) the GFP molecule trapped the p12 mutant in the core. It seems unlikely that an added GFP molecule would trap p12 in the PIC as when a big enough gap in the core is formed, during infection, GFP-p12 would be expected to diffuse. Based on the fact that p12 mutant 5 and 7 are dispensable for some retroviruses and that the larger mutations have a bigger effect on infectivity than the point mutants, I would suggest that the residues in these region of p12 help to stabilise the N-terminal domain interaction. However, whether this stabilising action is through making weak contacts with the interacting factor or maintaining an optimal conformation of the domain will require more investigation.

7.1.4 Towards an interacting partner for the N-terminus of p12

It was shown that the N-terminal domain of p12 has an influence on the formation and, possibly, the stability of the CA core. The mechanism behind this effect on the CA core remains unknown, but it likely involves interaction of the N-terminus of p12 with another viral factor. No long range interactions between p12 and CA of MLV have been reported, when they were expressed as a fusion protein (Kyere, Joseph, and Summers, 2008). Indeed, work in our laboratory has found little evidence to suggest that p12 interacts with monomeric CA *in vitro* (Nader, M. unpublished data). However, studying the viability of SNV and MLV chimeras has provided genetic evidence linking p12 and CA function during the early stages of infection (Lee and Nagashima, 2005). A panel of chimeras between Mo-MLV and GaLV, analysed here, have also suggested a link between the N-terminal domain of p12 and another viral factor (Section 4.3). These chimeras, however, do not rule out the possibility of a link between p12 and a factor other than CA, but based on the data from Lee *et al* (Lee and Nagashima, 2005), it would seem reasonable that the functional link is with CA.

What does p12 interact with and could p12 interact with the CA lattice directly? Limited biochemical evidence exists suggesting an interaction between p12 and CA. Immunoprecipitation of p12 from infected cells co-precipitated CA, although only a very small proportion of the total CA in the cells (Prizan-Ravid et al., 2010). Viral DNA was also found to be immunoprecipitated with an antibody to p12, indicating p12 is in a complex with the viral DNA and CA during infection (Prizan-Ravid et al., 2010). Low levels of CA present in the pull down may indicate the interaction between p12 and the PIC is weak, being easily disrupted by the isolation protocols. Interestingly, when the restriction factor tropism of Mo-MLV was altered by making changes in CA (D82N, A110R and H117L) the infectivity of p12 mutant 5 and 7 increased (Section 5.1.2). This observation again suggests a link between p12 and CA. As p12 mutants 5 and 7 are infectious in the N- and B- MLV background, could further changes to Mo-MLV CA (making it even more similar to N- or B- MLV CA) increase the infectivity of p12 mutant 5 and 7 further? More experiments are required to determine if this is the case, but if further alterations to CA could rescue the infectivity defect of the p12 mutants this would imply an interaction between CA and p12.

A specific interaction between p12 and viral gRNA has been previously reported (Sen, Sherr, and Todaro, 1976; Sen, Sherr, and Todaro, 1977; Sen and Todaro, 1977). It was

shown that the affinity of p12 for RNA was determined by the level of phosphorylation, where too much or no phosphorylation reduced p12 affinity for gRNA (Sen, Sherr, and Todaro, 1977). If mutation of the N-terminus of p12 were to abolish p12 binding to RNA it is hard to imagine how this would cause the defect observed for the N-terminal p12 mutants, as in the wild type scenario RNA would be converted to DNA in the infected cell and likely result in the loss of p12 from the PIC. Unless, however, p12 were also to have an affinity for DNA, which has not been investigated. Either way, it seems unlikely that a lack of p12 binding to the gRNA in the virion would result in aberrant core morphologies.

Lastly, as p12 is required in the nucleus for chromatin tethering (Section 5.3.2), other viral factors that enter the nucleus could be prime binding targets for p12. One such factor is IN. IN is absolutely required for the formation of the provirus and as such is an essential component of the PIC. Recently, small molecule inhibitors of the HIV-1 IN-LEDGF interaction (LEDGINs) have been described (Christ et al., 2010). One of these compounds (CX05045) showed dramatic effects on the morphology of the HIV-1 core when viruses were synthesised in the presence of the drug (Desimmie et al., 2013). Many of the viral cores were 'empty' and the RNP was outside the core (Desimmie et al., 2013). The authors showed that this LEDGIN was increasing the multimerisation of IN both *in vitro* and in viral particles (Desimmie et al., 2013). Interestingly, mutations in IN causing these aberrant core morphologies have previously been described, although there is a significant reduction in the level of Pol incorporated into these IN mutant particles (Engelman et al., 1995). Instead of interacting with the CA core, p12 could potentially bind to MLV IN, and loss of this binding could result in premature IN multimerisation that is detrimental to the formation of the CA core. Identification of interacting partners for p12 is essential to establish a mechanism by which Mo-MLV p12 aids in the formation of the CA core.

7.2 Clathrin and the N-terminus of p12

One interaction between a host factor and p12 has been reported. Clathrin is incorporated into Mo-MLV via interactions with a 'DLL' motif in p12 (Zhang et al., 2011). These three amino acids are situated in the p12 mutant 8 region, and mutation of any of these amino acids reduces the infectivity by greater than 10-fold (Chapter 3). The biggest infectivity defect is seen with the D25A mutant (>100-fold decrease in

infectivity, Figure 3.15), which is far greater than the infectivity defect reported upon depletion of clathrin from the producer cell (although it should be noted that the depletion of clathrin was not complete) (Zhang et al., 2011). Furthermore, the reported loss of mature p12 from virions described by Zhang *et al.*, was artifactual due to the antibody utilised, as shown in Figure 3.2. However, as I have shown for mutant 8, the authors did find a morphological defect of the core, based on the reduced abrogation of TRIM5alpha with p12 D25A mutant (Zhang et al., 2011). Work in our laboratory has found that p12 mutant 6 can incorporate clathrin but has the same infectivity and abrogation defect as p12 mutant 8 (Nader, M. unpublished data). Due to the relatively small infectivity defect imposed by clathrin depletion and the fact that other mutations in the N-terminus of p12 do not affect clathrin incorporation yet still inhibit infectivity, it seems unlikely that reduced clathrin incorporation explains the defect in the N-terminal p12 mutants.

So what is the function of clathrin incorporation? It has now been shown that clathrin is incorporated into a wide range of retroviruses (HIV-1, HIV-2, SIVmac, MPMV and MLV) (Popov et al., 2011; Zhang et al., 2011). Motifs responsible for clathrin incorporation are situated in varying positions within Gag and Pol, in different viruses (Popov et al., 2011; Zhang et al., 2011), implying that the positional recruitment of clathrin is not necessary for clathrin function in the virion. One of the functions of clathrin in endocytosis is stabilisation of the curved structure of clathrin-coated pits, by forming a lattice-like structure (Ford et al., 2002; Kirchhausen, 2000). It is possible that clathrin could act as a scaffold during assembly aiding in the efficient formation and stabilisation of the virion bud. Indeed mutation of the SIVmac DLL motif, responsible for clathrin incorporation, and the late domain (PTAP), both in p6, resulted in wider non-spherical protrusions from producer cells (Zhang et al., 2011). This suggested that, in SIVmac, clathrin does indeed have a role in the stabilising the virion bud at the assembly sites (Zhang et al., 2011). A three-fold infectivity defect was also seen for SIVmac synthesised in cells depleted of clathrin, indicating a similar phenomenon could be occurring for MLV. While the incorporation of clathrin is potentially involved in the morphogenesis of the virion bud during assembly, the mechanism behind this effect remain to be established. It is also not clear why virus produced in cells with depleted clathrin have a reduced infectivity in target cells.

7.3 p12 C-terminal domain function during the early stages of infection

p12 has been shown to be part of the MLV PIC by co-localisation with the viral DNA and CA during the early stages of infection (Prizan-Ravid et al., 2010). Interestingly, mutation of the C-terminus of MoMLV p12 (mutant 14) resulted in PICs incapable of chromatin accumulation during mitosis (Elis et al., 2012; Prizan-Ravid et al., 2010). The PFV CBS, inserted in p12, was able to rescue the infectivity defect of C-terminal p12 mutants, and has shown that the C-terminus of Mo-MLV p12 has an essential chromatin tethering function (Section 5.3.2). Additionally, chromatin tethering of p12 mutant 14 could be rescued by addition of the LANA CBS from Karposis's sarcoma-associated herpes virus (KSHV) (Elis et al., 2012). Both these CBS motifs did not completely rescue the infectivity defect, indicating the targeting to the chromatin is sub optimal for MLV integration. Furthermore, Mo-MLV PICs were more rigidly bound to chromatin when p12 contained the LANA CBS, and the PICs remained associated with the chromatin after mitosis (not seen for wild type Mo-MLV) (Elis et al., 2012). Moreover, docked wild type GFP-p12 puncta on chromosomes appeared to wobble, indicating the p12-mediated tethering is dynamic and the PICs are not tightly bound (Section 6.4.2).

Very recently, a report has shown that the ability of a heterologous CBS to rescue the p12 mutant 14 infectivity defect correlates relatively well with the affinity for mitotic chromatin of GFP proteins containing the p12-CBS motifs (Schneider et al., 2013). In contrast to this recent report, no strong indication of mCherry-tagged p12 binding to mitotic chromatin was observed in this study (Section 6.1.2). Likely reasons for this discrepancy are as follows: (i) The chromatin accumulation via p12 binding is very weak, meaning that the amount of mCherry-p12 expressed in cells will likely have a great effect on the ratio of mitotic-to-cytoplasmic mCherry signal. (ii) In my experiments, widefield microscopy had to be used due to the higher sensitivity required to visualise the mCherry-tagged p12 molecules in cells. The widefield system utilises an electron multiplying charged coupled device (EMCCD) to record information whereas the available confocal systems use photon multiplier tubes (PMT), which have a very low quantum efficiency compared to EMCCDs (i.e. they cannot record low fluorescence signals). (iii) Widefield microscopy, in contrast to confocal, also records fluorescence from above and below the plane of focus, and thus subtle differences in signal between the cytoplasm and the mitotic chromatin would not be exposed. (iv) Schneider *et al.* used GFP to tag their proteins, which has a higher quantum efficiency than mCherry (Shaner, Steinbach, and Tsien, 2005), and the confocal microscope they

utilised potentially had an EMCCD detector or one of the new high sensitivity PMTs for data recording ((Schneider et al., 2013) not specified in the methods). Thus, it seems likely their experimental set up was primed for better detection of smaller differences in fluorescent signal between the cytoplasm and the chromatin.

Unfortunately, despite knowing the subcellular location of the p12 C-terminal interaction, to date, no interacting partners have been found (Schneider et al., 2013) (and Nader, M. and Efrat, E. unpublished data). Similar methodologies to those utilised so far for p12, have managed to show interactions between GFP tagged with either the PFV CBS or the KSHV LANA CBS and histones H2A and H2B (Tobaly-Tapiero et al., 2008). As the chromatin affinity of the PFV-CBS and LANA CBS is much stronger than that seen for p12, it is likely that either the interaction of p12 with the chromatin factor is weaker, or less of the p12-chromatin factor exists to interact with p12. Due to the transient affinity of p12 for mitotic chromatin, it is entirely possible that the C-terminus of p12 may interact with a modified histone protein only present during mitosis.

7.3.1 Integration site targeting

A better understanding of integration targeting is thought vital to the design of better and safer retroviral-based gene therapy vectors. HIV-1 is targeted to chromatin via interaction with LEDGF, which acts as an IN-to-chromatin link (Busschots et al., 2005; Cherepanov et al., 2003; Ciuffi et al., 2005; Maertens et al., 2003; Marshall et al., 2007). LEDGF is not merely a tether, but has a direct influence on the integration site selection by HIV-1 (Ciuffi et al., 2005; Llano et al., 2006). HIV-1 tends to integrate into the body of transcriptionally active genes (Berry et al., 2006; Lewinski et al., 2006; Mitchell et al., 2004), whereas MLV has a tendency to integrate close to the transcriptional start sites near CpG islands (Mitchell et al., 2004; Wu et al., 2003). Cells depleted of LEDGF, by RNAi knock down, change the integration pattern observed for HIV-1 (Ciuffi et al., 2005).

Whether p12 had a LEDGF-like function with regards to integration site selection was investigated (Section 5.3.3). In line with a very recent report (Schneider et al., 2013), targeting p12 to the chromatin using another chromatin binding motif did not alter the integration site selection for Mo-MLV. Interestingly, the residual integration events

seen for p12 mutant 14 displayed an intriguing integration pattern that appeared less favourable to transcriptional start sites, although it should be noted that the number of sites sequenced was very low. Encouragingly however, the N-terminal p12 mutant 6, which also had a very low number of integration events, displayed an integration pattern similar to wild type Mo-MLV. Thus, the reduced affinity for transcriptional start sites with p12 mutant 14 hints that the chromatin tethering of p12 (regardless of the specific CBS motif utilised) does influence integration site selection. I would hypothesise that targeting to the chromatin allows interaction of IN with a host co-factor that then influences the integration site selection. This hypothesis would fit with the data suggesting the factor targeting integration site selection for MLV is IN with a small but significant contribution from Gag (Lewinski et al., 2006).

7.3.2 Composition of the chromatin-associated PIC

The composition of the retroviral core undergoes many changes on its way to the host chromatin. On the path to the nucleus, the poorly understood process of uncoating occurs. The timing and extent of the CA loss from the cores remains ill defined, but appears to occur at different rates between retroviruses (Arfi et al., 2009; Fassati and Goff, 1999; Fassati and Goff, 2001; Hulme, Perez, and Hope, 2011). For MLV and simple retroviruses, nuclear envelope breakdown is required for access to the host chromosomes. The composition of the PIC in the nucleus is likely significantly different from the one present in the cytoplasm and, as a minimal requirement for tethering and integration, it must contain the following viral components: viral DNA, IN and p12. Little CA is seen in the nucleus of cells infected with MLV, using biochemical analysis (Fassati and Goff, 1999) and immunofluorescence (Elis et al., 2012). Surprisingly, when Mo-MLV was labelled with the CA-YFP fusion protein, CA-YFP could be seen docked on the mitotic chromatin as had been seen for GFP-p12 (Section 6.4.2). This implied that some fraction of CA remained with the nuclear-associated PIC. N-terminal p12 mutants labelled with GFP-p12 were more stable inside infected cells than native p12 (with a myc tag). This was also observed with mCherry-p12 expressed in cells, where expression of p12 with a small tag was unstable and was quickly degraded in stark contrast to mCherry-tagged p12. Therefore it is likely that, as with p12, the attachment of a YFP molecule to CA also has a stabilising effect on the protein and the presence of CA-YFP on the condensed chromatin could be an artefact of altered protein behaviour.

Curiously, attachment of the PIC to the chromatin could occur very rapidly after entry of the core into the infected cells. GFP-p12 was seen docked on the condensed chromatin less than one hour post infection (Section 6.4.2), meaning that reverse transcription is unlikely to have occurred before docking. Indeed, VLPs lacking the viral genome were also able to dock onto condensed chromatin (Efrat Elis, personal communication). Therefore, the presence of the viral DNA is not a requirement for p12 mediated docking on the chromatin. However, it is entirely possible that these early-docked PICs are not on a productive infection pathway and constitute dead end products.

7.4 Non-specific uptake of virus particles into cells

Viruses on non-productive infection pathways are a significant problem, especially in microscopy studies of infection (McDonald et al., 2002). Investigations into the amount non-specific uptake of virus into cells, in this thesis, have revealed that the amount of virus taken up into U20S cells was greater when virus contained no envelope protein compared to a redundant envelope protein (Section 6.2.1). This indicates that the use of VLPs without envelope proteins likely overestimates the number of non-specifically endocytosed particles. In agreement with this observation, a similar phenomenon has also been seen for HeLa cells incubated with NL43 HIV-1 and a ΔEnv version of this virus, where less NL43, containing a redundant Env, was non-specifically taken up into the HeLa cells than the ΔEnv virus (Marechal et al., 1998). Interestingly, when less of the ΔEnv HIV-1 was applied to the cells, the amount of internalised p24 (CA) dropped down to the level seen for viruses containing a redundant Env protein, indicating that it may have a dose-dependent effect (Marechal et al., 1998). However, more work is required with other cell/Env combinations to assess whether this is a common phenomenon or restricted to certain cell types. Additionally, investigations using a range of MOIs will help to determine if there is a dose-dependent effect to this non-specific virus uptake.

7.5 Conservation of p12 function in retroviruses

Many retroviruses contain Gag cleavage products between MA and CA. The early function of p12 was required for all the gammaretroviruses analysed, although not all gammaretroviruses required the whole of the N-terminus for function (Section 4.1.3). The alpharetrovirus, RSV, codes for three proteins between MA and CA: p10, p2a and

p2b. The p2b protein carries the 'PPPY' late-domain essential for the late stages of replication (Xiang et al., 1996). The p10 protein is interesting in that it determines the shape of *in vitro* formed particles (CA-NC) from tubes to spheres (Campbell and Vogt, 1997). This is also true *in vivo*, as RSV with mutations in the C-terminus of p10 (L219A) produce less virus (due to impaired nuclear export) but the viruses that are released contain aberrant core morphologies (Scheifele et al., 2007). Therefore, p10, while different from gammaretroviral p12 molecules, also has a role to play in formation of the CA core. MMTV is a betaretrovirus and contains four cleaved proteins between MA and CA, named: p21, p8, p3 and n (Hizi et al., 1989). Interestingly, deletion of the n domain of MMTV Gag results in the formation of tube-like structures assembling in the cytoplasm and released from cells (Zabransky et al., 2010). These two examples highlight other Gag cleavage products originating between MA and CA that have an effect on the formation/morphology of the CA core.

RfRV, an endogenous retrovirus from the greater horseshoe bat (*Rhinolophus ferrumequinum*), contains a p12-like sequence before the CA domain in Gag (Cui et al., 2012). The potential RfRV p12 sequence is considerably larger than Mo-MLV p12 (121 amino acids versus 84), although both sequences contained a similar proportion of prolines, 20.2% for Mo-MLV p12 and 26.4% for RfRV p12 (Figure 7.1). While the sequence identity was low (only 28.9%) this was similar to the identity of Mo-MLV p12 with GaLV p12. The 'PPPY' motif found in all gammaretroviral p12 proteins was also present in the potential RfRV p12 molecule. Interestingly, the N-terminus of the potential RfRV p12 contained conserved residues that were shown to be essential for Mo-MLV and GaLV p12 function (leucines and aspartic acids). After the 'PPPY' motif little sequence identity was shared between RfRV and Mo-MLV p12 until the start of the 'SPM' motif that was essential for Mo-MLV p12 function. This central 'linker' region in p12 displayed the largest sequence divergence between distantly related gammaretroviruses, indicating its composition is not under significant selective pressure. In additionally to the 'SPM' motif, many arginine residues are situated in the C-terminus of the RfRV p12 protein, which were essential for Mo-MLV and GaLV p12 function.

Sequence alignments of RfRV Gag and Pol with other gammaretroviruses showed RfRV to be ancestral to the exogenous MLVs (Cui et al., 2012). This suggests the possibility that the function of p12 has been preserved in gammaretroviruses before the

emergence of the exogenous MLVs. To test this experimentally, the Gag (or parts of Gag) of this ancient gammaretrovirus could be resurrected, as has been done for the endogenous lentiviral CAs from the rabbit endogenous lentivirus type K (RELIK) and prosimian immunodeficiency virus (PSIV) (Goldstone et al., 2010). This approach could test if the early function of p12 has actually been preserved in the gammaretroviral lineage.

```
      1                                                              48
Mo-MLV PALTPSLGAKP--KP---QVLSDSGG--P-LIDL-LTEDPPPYRDPRPPPSDRDGNG---
RfRV   PAV-PTAGPKPPEKPTQGPILQEGSDIYPSLIDLDLEETPPPYAPVAPLQRWRAPSPVTP
       **  *  * **  **     *       * **** * * ****    *    *

      49                                                             82
Mo-MLV GEATPAGEAP----DPSPMA--------------SRLRG--RRE--P---PVAD-STTSQAF
RfRV   LPPEAAAPSELPHSTASPMAPPPPGSPAPAQGPARGLRPRRRREETPEEEPSSSTSAGAPIL
            *          ****                 **   ***  *   *    *
```

Figure 7.1: Alignment of Mo-MLV p12 with RfRV p12 sequence. The Mo-MLV p12 amino acid sequence (GenBank: J02255.1) was aligned with that of RfRV (JQ303225.1) using Clustal X2 (Larkin M. A. *et al.* 2007). Numbers highlight the amino acid numbering from Mo-MLV p12.

7.6 A model for p12 function during the early stages of infection

Based on the findings presented in this thesis and the recent work published by others I would like to propose a model for the function of p12 during the early stages of infection (Figure 7.2). In a wild type context (Figure 7.2A), p12 is liberated from the Gag polyprotein by the viral protease and a proportion of the mature p12 molecules make essential interactions with a factor within the virion, through their N-terminal domain. The identity of this target and the nature of the interaction currently remain unknown. Interaction between p12 and its interacting partner aids in the formation of the mature core. After release of the viral core into the cytoplasm, p12 remains bound to the PIC (through the N-terminal domain interaction) and travels to the nucleus. At mitosis, the PIC reaches the proximity of the chromosomes and interaction of the C-terminus of p12 with host chromatin occurs. After mitosis, p12 releases from the chromatin and the integration complex is trafficked to the site of integration by host factors recruited during p12-mediated chromatin tethering.

In Mo-MLV with p12 mutant 6 (Figure 7.2B), p12 cannot make the N-terminal interaction at all, and thus the mature core is formed with a poor efficiency. This means

that upon infection, only a small fraction of viruses will contain correctly formed cores that can support reverse transcription (based on the reduced accumulation of viral DNA products with this mutant *in vivo*). Due to no interaction between p12 and PIC components, p12 will be lost from the PIC and no chromatin tethering can occur. For N-terminal p12 mutants 5, 7 and 8 (Figure 7.2C), the state of the CA core is currently unknown. Based on the ability of these p12 mutants to reverse transcribe, I would assume that they contain some aberrant core which can support reverse transcription but it is somewhat grossly malformed or unstable so that restriction factors cannot recognise/interact with it. Again when these mutant cores are released into the cytoplasm, a lack of interaction between p12 and the PIC results in the loss of p12 and no chromatin tethering.

Finally, when the C-terminal domain of p12 is mutated (Figure 7.2D), infection proceeds as wild type until nuclear envelope breakdown when p12 is unable to tether the nuclear-associated PIC to chromatin. Thus, the residual integration seen for this mutant likely represents chance interactions with an accessible region of the chromatin conducive to integration.

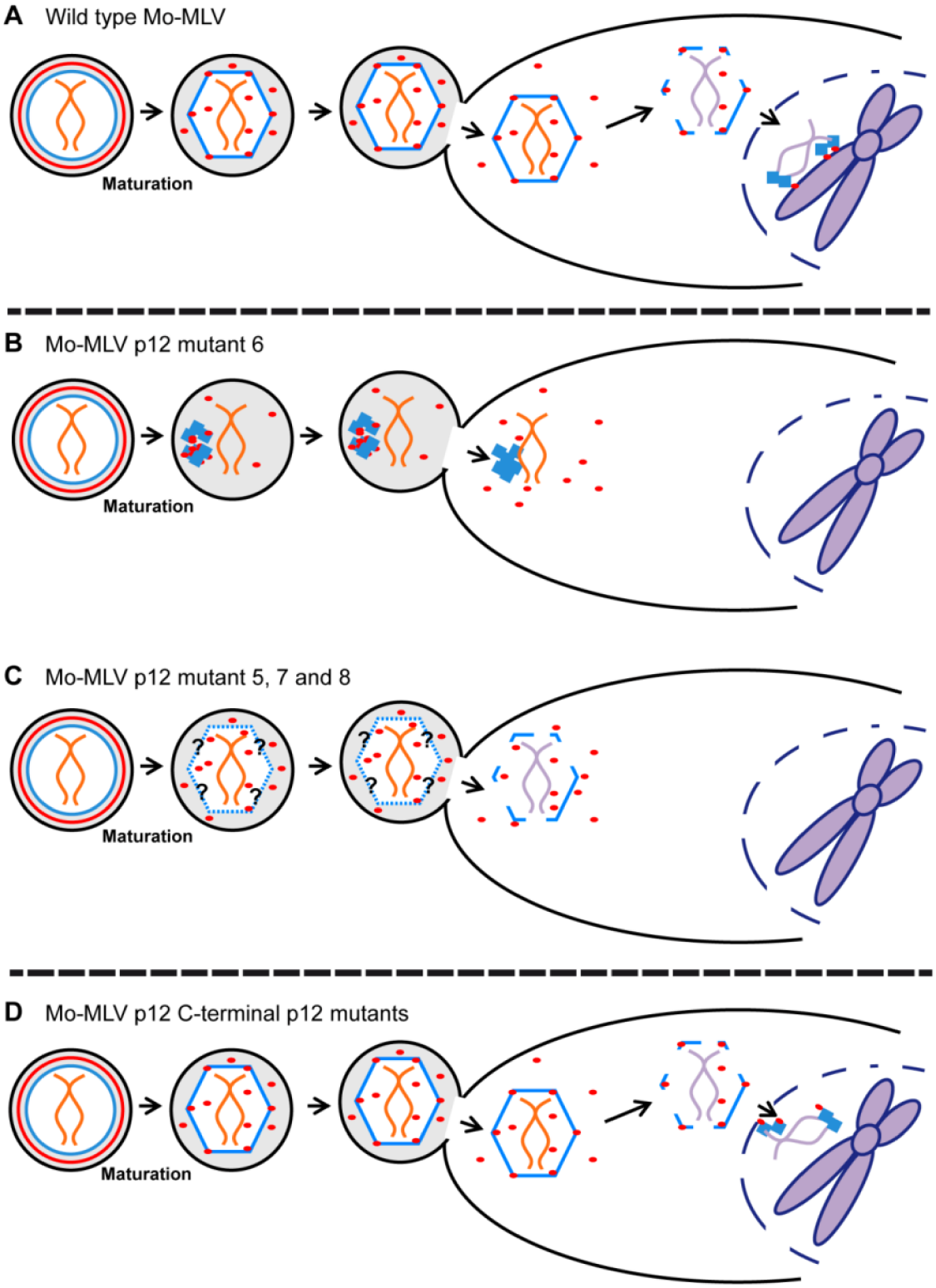


Figure 7.2: A model for the function of gammaretroviral p12 during maturation and the early stages of infection. Diagrams representing the stages of retroviral infection from cleavage of Gag and Pol to attachment of the PIC to the chromatin for Mo-MLV containing: (A) wild type p12, (B) mutant 6 p12, (C) mutant 5, 7 or 8 p12, or (D) C-terminal mutant p12. CA is coloured in blue, p12 in red, viral RNA in orange and the viral DNA in light purple.

7.7 Future work

One key question really needs to be answered to establish mechanisms for the action of p12:

- What are the viral and host factors that p12 interacts with?

A few experimental approaches are described below which may help to answer this question.

7.7.1 State of the CA core in the other p12 mutants

In order to determine the function of the N-terminus of p12, it is imperative that the state of the CA core is assessed in the other p12 mutants. Therefore, the same thin-section TEM analysis should be applied to N-terminal p12 mutants 5, 7 and 8 to analyse the morphology of the CA core in these mutants.

None of the work here has directly assessed the effect of the p12 mutations on the state of the immature Gag lattice. It is possible that the mutations in p12 cause a defect at a step before formation of the mature core. Mo-MLV p12 mutants, with an inactivate PR, could be used to analyse the immature cores by cryo-EM. Due to a greater degree of homogeneity in the morphology of immature viral particles, cryo-EM averaging techniques are more accurate, which will allow electron densities of the Gag molecules to be created for analysis. How much insight this approach would give is debatable and would really depend on the magnitude of any observed defect.

7.7.2 The composition of the isolated 'partial' cores

Partial Mo-MLV CA cores (devoid of the gRNA) were isolated in sucrose gradients by centrifugation of virions through a Triton layer. The sedimentation profile of these partial cores was different from free CA protein and un-lysed virions. However, the composition of these partial cores remain unknown. Detection of p12 and NC was unsuccessful due to low levels of protein present in the partial core fractions (data not shown). Concentration of protein from the fractions could be performed by trichloroacetic acid or methanol/chloroform precipitation. This would concentrate the proteins and potentially enable the detection of p12 and NC by immunoblotting. Determination of the amount of CA contained in the partial cores is a more complex problem. Samples could be run on non-denaturing gels to obtain some information about size of the CA complex, although this would work best if the partial cores are

smaller complexes (i.e. CA hexamers). If they are much larger than CA hexamers, then potentially dynamic light scattering (DLS) or cryo-EM of the partial cores may produce a better estimate on the size of these complexes. However these second methods may suffer due to the low concentration of partial cores isolated, in addition to the likely unstable nature of the cores. It is possible that the use of cross-linking reagents may help to stabilise the partial cores if they prove too unstable for analysis.

7.7.3 Identification of interacting partners for the N-terminus of p12

To establish a mechanism for the action of p12, identification of interacting partners for the N- and C-terminus of p12 is vital. Interactions of p12 with monomeric CA have been hard to detect *in vitro* (Nader, M. unpublished data). CA-interacting molecules can often have a low affinity for monomeric CA, especially if the interaction actually occurs with oligomeric forms of CA (eg. hexamers). Indeed, the CA-binding portion of CPSF6 binds monomeric CA with a very low affinity (Price et al., 2012). Thus, using oligomeric CA structures to probe for p12-CA interactions may be more fruitful. An ordered core-like structure of MLV CA on lipid-coated nanotubes has been reported (Hilditch et al., 2011). If p12 does indeed interact with CA, these assemblies may serve as better targets for assessment of p12 and CA-core interactions. CA-coated lipid nanotubes could be incubated with purified p12 and pelleted by centrifugation to see if p12 can bind. If the binding is disrupted by the centrifugal force required to pellet the CA-coated tubes, binding of GFP-p12 on the CA-coated tubes could be assessed by a fluorescence recovery after photobleaching (FRAP) technique, similar in principle to experiments performed by Schneider *et al.* to show immobilisation of GFP-p12 on condensed chromatin. Instead of transfecting cells, purified GFP-p12 could be applied to immobilised CA nanotubes and analysed by FRAP. If there is a delay in the fluorescent recovery, it suggests that the bleached GFP-p12 has remained bound to the CA-nanotube, indicating an interaction.

If these results indicate an interaction between p12 and CA core-like structures can occur, more advanced *in vitro* techniques could be used to establish the sites of binding. Hexamers of the N-terminal domain of N- and B-tropic MLV CA have been solved by X-ray crystallography (Mortuza et al., 2008; Mortuza et al., 2004). These N-MLV CA hexamer crystals could be used as a target for co-crystallisation of p12 with its binding surface. If crystallisation is not possible, p12 bound to the CA-coated nanotubes could

be investigated using cryo-electron microscopy (cryo-EM) in an attempt to localise the area of p12 core interaction. However, p12 may be too small for visualisation on the nanotubes by cryo-EM. Therefore, addition of a large protein to p12, like GFP, may make for easier identification of p12 on the nanotubes.

7.7.4 Identification of interacting partners for the C-terminus of p12

In an attempt to find an interacting partner for the C-terminal domain of p12, immunoprecipitation of cellular factors from mitotic cells should be attempted. Using cells arrested in mitosis may enrich lysates with the p12-chromatin interacting factor if the factor is only present during mitosis. It remains entirely possible that p12 tethers to the chromatin by actually interacting with the host DNA, thus chromatin immunoprecipitation (ChIP) may prove valuable in this investigation. Thus, if p12 binds to specific regions of DNA or DNA-associated host proteins this may be revealed.

7.7.5 The mechanism of integration site selection

Adding heterologous chromatin binding motifs in Mo-MLV p12 did not alter the targeting of Mo-MLV integration (Schneider et al., 2013). However, our preliminary data highlights an altered integration pattern for Mo-MLV C-terminal p12 mutant 14. The first thing that must be done is to confirm the phenotype by sequencing more integration sites. Due to the inefficiency of the p12-independant integration this will require a large volume of virus to be produced, to enhance the number of integration events. VLPs encoding GFP could be utilised which will allow fluorescent activated cell sorting to be used and enrich a population of infected cells for integration site analysis.

7.7.6 Position of p12 within the PIC and the composition of the nuclear-associated PIC

Fluorescence microscopy has undergone some significant advances over the past decade and a host of super resolution fluorescence microscopy techniques have been established. Theses technique use either physical or mathematical manipulation to surpass the diffraction limit of light, producing impressive spatial resolutions in the order of tens of nanometres. Using these super resolution techniques, I aim to test the localisation of p12 with respect to the PIC at various times leading up to docking on the chromatin. Potentially, whether p12 is located inside or outside of the PIC could be tested as well as any localised affinity of p12 with the CA core. From mixed particles,

little wild type p12 is needed in a particle for full infectivity, thus mixed VLPs with myc-tagged wild type p12 and an N-terminal p12 mutant could be studied, then only the 'functional' p12 will be analysed by microscopy using an anti-myc antibody. If a chromatin interacting partner for Mo-MLV p12 can be isolated, super resolution microscopy could be used to show an intimate association between the factors during infection, implicating a physiological significance to the interaction.

References

Accola, M. A., Strack, B., and Gottlinger, H. G. (2000). Efficient particle production by minimal Gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. *J Virol* **74**(12)**,** 5395-402.

Aldovini, A., and Young, R. A. (1990). Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J Virol* **64**(5)**,** 1920-6.

Alin, K., and Goff, S. P. (1996). Amino acid substitutions in the CA protein of Moloney murine leukemia virus that block early events in infection. *Virology* **222**(2)**,** 339-51.

Arfi, V., Lienard, J., Nguyen, X. N., Berger, G., Rigal, D., Darlix, J. L., and Cimarelli, A. (2009). Characterization of the behavior of functional viral genomes during the early steps of human immunodeficiency virus type 1 infection. *J Virol* **83**(15)**,** 7524-35.

Arhel, N., Genovesio, A., Kim, K. A., Miko, S., Perret, E., Olivo-Marin, J. C., Shorte, S., and Charneau, P. (2006). Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nature Methods* **3**(10)**,** 817-24.

Arhel, N. J., and Charneau, P. (2009). Bisarsenical labeling of HIV-1 for real-time fluorescence microscopy. *Methods in Molecular Biology* **485,** 151-9.

Arhel, N. J., Souquere-Besse, S., and Charneau, P. (2006). Wild-type and central DNA flap defective HIV-1 lentiviral vector genomes: intracellular visualization at ultrastructural resolution levels. *Retrovirology* **3,** 38.

Arhel, N. J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prevost, M. C., Allen, T. D., and Charneau, P. (2007). HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO Journal* **26**(12)**,** 3025-37.

Astiazaran, P., Bueno, M. T., Morales, E., Kugelman, J. R., Garcia-Rivera, J. A., and Llano, M. (2011). HIV-1 integrase modulates the interaction of the HIV-1 cellular cofactor LEDGF/p75 with chromatin. *Retrovirology* **8,** 27.

Auerbach, M., Brown, K., and Kaplan, A. (2006). A small loop in the capsid protein of Moloney murine leukemia virus controls assembly of spherical cores. *Journal of Virology* **80,** 2884-2893.

Auerbach, M. R., Brown, K. R., and Singh, I. R. (2007). Mutational analysis of the N-terminal domain of Moloney murine leukemia virus capsid protein. *J Virol* **81**(22)**,** 12337-47.

Auerbach, M. R., Shu, C., Kaplan, A., and Singh, I. R. (2003). Functional characterization of a portion of the Moloney murine leukemia virus gag gene by genetic

footprinting. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 11678-83.

Bader, J. P. (1965). The Requirement for DNA Synthesis in the Growth of Rous Sarcoma and Rous-Associated Viruses. *Virology* **26,** 253-61.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **226**(5252)**,** 1209-11.

Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**(4599)**,** 868-71.

Bassin, R. H., Duran-Troise, G., Gerwin, B. I., and Rein, A. (1978). Abrogation of Fv-1b restriction with murine leukemia viruses inactivated by heat or by gamma irradiation. *J Virol* **26**(2)**,** 306-15.

Basyuk, E., Boulon, S., Skou Pedersen, F., Bertrand, E., and Vestergaard Rasmussen, S. (2005). The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. *J Mol Biol* **354**(2)**,** 330-9.

Battini, J. L., Rasko, J. E., and Miller, A. D. (1999). A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction. *Proc Natl Acad Sci U S A* **96**(4)**,** 1385-90.

Beer, C., and Pedersen, L. (2007). Matrix fibronectin binds gammaretrovirus and assists in entry: new light on viral infections. *Journal of Virology* **81,** 8247-57.

Berkhout, B., and Jeang, K. T. (1989). trans activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the trans-acting-responsive hairpin: a quantitative analysis. *J Virol* **63**(12)**,** 5501-4.

Berkowitz, R. D., Ohagen, A., Hoglund, S., and Goff, S. P. (1995). Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric Gag polyproteins during RNA packaging in vivo. *J Virol* **69**(10)**,** 6445-56.

Berry, C., Hannenhalli, S., Leipzig, J., and Bushman, F. D. (2006). Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol* **2**(11)**,** e157.

Besnier, C., Takeuchi, Y., and Towers, G. (2002). Restriction of lentivirus in monkeys. *Proc Natl Acad Sci U S A* **99**(18)**,** 11920-5.

Best, S., Le Tissier, P., Towers, G., and Stoye, J. P. (1996). Positional cloning of the mouse retrovirus restriction gene Fv1. *Nature* **382**(6594)**,** 826-9.

Bharat, T. A., Davey, N. E., Ulbrich, P., Riches, J. D., de Marco, A., Rumlova, M., Sachse, C., Ruml, T., and Briggs, J. A. (2012). Structure of the immature retroviral capsid at 8 A resolution by cryo-electron microscopy. *Nature*.

Bieniasz, P. D. (2006). Late budding domains and host proteins in enveloped virus release. *Virology* **344**(1)**,** 55-63.

Birbach, A., Bailey, S. T., Ghosh, S., and Schmid, J. A. (2004). Cytosolic, nuclear and nucleolar localization signals determine subcellular distribution and activity of the NF-kappaB inducing kinase NIK. *Journal of Cell Science* **117**(Pt 16)**,** 3615-24.

Bishop, K. N., Bock, M., Towers, G., and Stoye, J. P. (2001). Identification of the regions of Fv1 necessary for murine leukemia virus restriction. *J Virol* **75**(11)**,** 5182-8.

Bishop, K. N., Holmes, R. K., Sheehy, A. M., and Malim, M. H. (2004). APOBEC-mediated editing of viral RNA. *Science* **305**(5684)**,** 645.

Bishop, K. N., Mortuza, G. B., Howell, S., Yap, M. W., Stoye, J. P., and Taylor, I. A. (2006). Characterization of an amino-terminal dimerization domain from retroviral restriction factor Fv1. *J Virol* **80**(16)**,** 8225-35.

Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M., and Malim, M. H. (2008). APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog* **4**(12)**,** e1000231.

Bittner, J. J. (1936). Some Possible Effects of Nursing on the Mammary Gland Tumor Incidence in Mice. *Science* **84**(2172)**,** 162.

Blair, W. S., Pickford, C., Irving, S. L., Brown, D. G., Anderson, M., Bazin, R., Cao, J., Ciaramella, G., Isaacson, J., Jackson, L., Hunt, R., Kjerrstrom, A., Nieman, J. A., Patick, A. K., Perros, M., Scott, A. D., Whitby, K., Wu, H., and Butler, S. L. (2010). HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog* **6**(12)**,** e1001220.

Blumenthal, R., Durell, S., and Viard, M. (2012). HIV entry and envelope glycoprotein-mediated fusion. *Journal of Biological Chemistry* **287**(49)**,** 40841-9.

Bolognesi, D. P., Luftig, R., and Shaper, J. H. (1973). Localization of RNA tumor virus polypeptides. I. Isolation of further virus substructures. *Virology* **56**(2)**,** 549-64.

Boone, L. R., Innes, C. L., and Heitman, C. K. (1990). Abrogation of Fv-1 restriction by genome-deficient virions produced by a retrovirus packaging cell line. *J Virol* **64**(7)**,** 3376-81.

Bosco, D. A., Eisenmesser, E. Z., Pochapsky, S., Sundquist, W. I., and Kern, D. (2002). Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A. *Proc Natl Acad Sci U S A* **99**(8)**,** 5247-52.

Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., Lieberman, J., and Elledge, S. J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. *Science* **319**(5865)**,** 921-6.

Bray, M., Prasad, S., Dubay, J. W., Hunter, E., Jeang, K. T., Rekosh, D., and Hammarskjold, M. L. (1994). A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc Natl Acad Sci U S A* **91**(4)**,** 1256-60.

Brennan, G., Kozyrev, Y., and Hu, S. L. (2008). TRIMCyp expression in Old World primates Macaca nemestrina and Macaca fascicularis. *Proc Natl Acad Sci U S A* **105**(9)**,** 3569-74.

Briggs, J. A., Grunewald, K., Glass, B., Forster, F., Krausslich, H. G., and Fuller, S. D. (2006). The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. *Structure* **14**(1)**,** 15-20.

Briggs, J. A., Riches, J. D., Glass, B., Bartonova, V., Zanetti, G., and Krausslich, H. G. (2009). Structure and assembly of immature HIV. *Proc Natl Acad Sci U S A* **106**(27)**,** 11090-5.

Briggs, J. A., Simon, M. N., Gross, I., Krausslich, H. G., Fuller, S. D., Vogt, V. M., and Johnson, M. C. (2004). The stoichiometry of Gag protein in HIV-1. *Nat Struct Mol Biol* **11**(7)**,** 672-5.

Briggs, J. A., Wilk, T., Welker, R., Krausslich, H. G., and Fuller, S. D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. *EMBO Journal* **22**(7)**,** 1707-15.

Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. (1989). Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc Natl Acad Sci U S A* **86**(8)**,** 2525-9.

Bryant, M., and Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc Natl Acad Sci U S A* **87**(2)**,** 523-7.

Bukrinskaya, A., Brichacek, B., Mann, A., and Stevenson, M. (1998). Establishment of a functional human immunodeficiency virus type 1 (HIV-1) reverse transcription complex involves the cytoskeleton. *Journal of Experimental Medicine* **188**(11)**,** 2113-25.

Bukrinsky, M. I., Haggerty, S., Dempsey, M. P., Sharova, N., Adzhubel, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., and Stevenson, M. (1993a). A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* **365**(6447)**,** 666-9.

Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G., and Stevenson, M. (1993b). Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A* **90**(13)**,** 6125-9.

Busschots, K., Vercammen, J., Emiliani, S., Benarous, R., Engelborghs, Y., Christ, F., and Debyser, Z. (2005). The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *Journal of Biological Chemistry* **280**(18)**,** 17841-7.

Cabantous, S., Terwilliger, T. C., and Waldo, G. S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology* **23**(1)**,** 102-7.

Cairns, T. M., and Craven, R. C. (2001). Viral DNA synthesis defects in assembly-competent Rous sarcoma virus CA mutants. *J Virol* **75**(1)**,** 242-50.

Campbell, E. M., Perez, O., Melar, M., and Hope, T. J. (2007). Labeling HIV-1 virions with two fluorescent proteins allows identification of virions that have productively entered the target cell. *Virology* **360**(2)**,** 286-93.

Campbell, S., and Vogt, V. M. (1997). In vitro assembly of virus-like particles with Rous sarcoma virus Gag deletion mutants: identification of the p10 domain as a morphological determinant in the formation of spherical particles. *J Virol* **71**(6)**,** 4425-35.

Cattoglio, C., Pellin, D., Rizzi, E., Maruggi, G., Corti, G., Miselli, F., Sartori, D., Guffanti, A., Di Serio, C., Ambrosi, A., De Bellis, G., and Mavilio, F. (2010). High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. *Blood* **116**(25)**,** 5507-17.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J. L., Bousso, P., Deist, F. L., and Fischer, A. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**(5466)**,** 669-72.

Charneau, P., Alizon, M., and Clavel, F. (1992). A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J Virol* **66**(5)**,** 2814-20.

Charneau, P., and Clavel, F. (1991). A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. *J Virol* **65**(5)**,** 2415-21.

Chen, B. J., and Lamb, R. A. (2008). Mechanisms for enveloped virus budding: can some viruses do without an ESCRT? *Virology* **372**(2)**,** 221-32.

Cherepanov, P. (2007). LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Research* **35**(1)**,** 113-24.

Cherepanov, P., Devroe, E., Silver, P. A., and Engelman, A. (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *Journal of Biological Chemistry* **279**(47)**,** 48883-92.

Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *Journal of Biological Chemistry* **278**(1)**,** 372-81.

Christ, F., Thys, W., De Rijck, J., Gijsbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J. C., Benarous, R., Cereseto, A., and Debyser, Z. (2008). Transportin-SR2 imports HIV into the nucleus. *Current Biology* **18**(16)**,** 1192-202.

Christ, F., Voet, A., Marchand, A., Nicolet, S., Desimmie, B. A., Marchand, D., Bardiot, D., Van der Veken, N. J., Van Remoortel, B., Strelkov, S. V., De Maeyer,

M., Chaltin, P., and Debyser, Z. (2010). Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat Chem Biol* **6**(6)**,** 442-8.

Chun, T. W., Davey, R. T., Jr., Ostrowski, M., Shawn Justement, J., Engel, D., Mullins, J. I., and Fauci, A. S. (2000). Relationship between pre-existing viral reservoirs and the re-emergence of plasma viremia after discontinuation of highly active anti-retroviral therapy. *Nature Medicine* **6**(7)**,** 757-61.

Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. *Nature Medicine* **11**(12)**,** 1287-9.

Coffin, J. M., and Billeter, M. A. (1976). A physical map of the Rous sarcoma virus genome. *J Mol Biol* **100**(3)**,** 293-318.

Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997a). Genetic Organization. *In* "Retroviruses" (J. M. Coffin, S. H. Hughes, and H. E. Varmus, Eds.). Cold Spring Harbor Laboratory Press, New York.

Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997b). "Retroviruses." Cold Spring Harbor Press, New York.

Conticello, S. G., Harris, R. S., and Neuberger, M. S. (2003). The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Current Biology* **13**(22)**,** 2009-13.

Cowan, S., Hatziioannou, T., Cunningham, T., Muesing, M. A., Gottlinger, H. G., and Bieniasz, P. D. (2002). Cellular inhibitors with Fv1-like activity restrict human and simian immunodeficiency virus tropism. *Proc Natl Acad Sci U S A* **99**(18)**,** 11914-9.

Craven, R. C., Leure-duPree, A. E., Weldon, R. A., Jr., and Wills, J. W. (1995). Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. *J Virol* **69**(7)**,** 4213-27.

Crawford, S., and Goff, S. P. (1984). Mutations in gag proteins P12 and P15 of Moloney murine leukemia virus block early stages of infection. *Journal of Virology* **49,** 909-17.

Crawford, S., and Goff, S. P. (1985). A deletion mutation in the 5' part of the pol gene of Moloney murine leukemia virus blocks proteolytic processing of the gag and pol polyproteins. *J Virol* **53**(3)**,** 899-907.

Cronin, J., Zhang, X. Y., and Reiser, J. (2005). Altering the tropism of lentiviral vectors through pseudotyping. *Curr Gene Ther* **5**(4)**,** 387-98.

Cui, J., Tachedjian, M., Wang, L., Tachedjian, G., Wang, L. F., and Zhang, S. (2012). Discovery of retroviral homologs in bats: implications for the origin of Mammalian gammaretroviruses. *J Virol* **86**(8)**,** 4288-93.

Cullen, B. R. (2003). Nuclear mRNA export: insights from virology. *Trends in Biochemical Sciences* **28**(8)**,** 419-24.

Dalgleish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**(5996)**,** 763-7.

Das, D., and Georgiadis, M. M. (2004). The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure* **12**(5)**,** 819-29.

Datta, S. A., Zuo, X., Clark, P. K., Campbell, S. J., Wang, Y. X., and Rein, A. (2011). Solution properties of murine leukemia virus gag protein: differences from HIV-1 gag. *J Virol* **85**(23)**,** 12733-41.

de Marco, A., Davey, N. E., Ulbrich, P., Phillips, J. M., Lux, V., Riches, J. D., Fuzik, T., Ruml, T., Krausslich, H. G., Vogt, V. M., and Briggs, J. A. (2010a). Conserved and variable features of Gag structure and arrangement in immature retrovirus particles. *J Virol* **84**(22)**,** 11729-36.

de Marco, A., Heuser, A. M., Glass, B., Krausslich, H. G., Muller, B., and Briggs, J. A. (2012). Role of the SP2 domain and its proteolytic cleavage in HIV-1 structural maturation and infectivity. *J Virol* **86**(24)**,** 13708-16.

de Marco, A., Muller, B., Glass, B., Riches, J. D., Krausslich, H. G., and Briggs, J. A. (2010b). Structural analysis of HIV-1 maturation using cryo-electron tomography. *PLoS Pathog* **6**(11)**,** e1001215.

De Rijck, J., Bartholomeeusen, K., Ceulemans, H., Debyser, Z., and Gijsbers, R. (2010). High-resolution profiling of the LEDGF/p75 chromatin interaction in the ENCODE region. *Nucleic Acids Research* **38**(18)**,** 6135-47.

Desimmie, B. A., Schrijvers, R., Demeulemeester, J., Borrenberghs, D., Weydert, C., Thys, W., Vets, S., Van Remoortel, B., Hofkens, J., De Rijck, J., Hendrix, J., Bannert, N., Gijsbers, R., Christ, F., and Debyser, Z. (2013). LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the virions. *Retrovirology* **10,** 57.

Diaz-Griffero, F., Kar, A., Perron, M., Xiang, S. H., Javanbakht, H., Li, X., and Sodroski, J. (2007). Modulation of retroviral restriction and proteasome inhibitor-resistant turnover by changes in the TRIM5alpha B-box 2 domain. *J Virol* **81**(19)**,** 10362-78.

Diaz-Griffero, F., Qin, X. R., Hayashi, F., Kigawa, T., Finzi, A., Sarnak, Z., Lienlaf, M., Yokoyama, S., and Sodroski, J. (2009). A B-box 2 surface patch important for TRIM5alpha self-association, capsid binding avidity, and retrovirus restriction. *J Virol* **83**(20)**,** 10737-51.

Dorfman, T., Bukovsky, A., Ohagen, A., Hoglund, S., and Gottlinger, H. G. (1994). Functional domains of the capsid protein of human immunodeficiency virus type 1. *J Virol* **68**(12)**,** 8180-7.

Dorman, N., and Lever, A. (2000). Comparison of viral genomic RNA sorting mechanisms in human immunodeficiency virus type 1 (HIV-1), HIV-2, and Moloney murine leukemia virus. *J Virol* **74**(23)**,** 11413-7.

Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayanan, C., Maddon, P. J., Koup, R. A., Moore, J. P., and Paxton, W. A. (1996). HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**(6584)**,** 667-73.

Duran-Troise, G., Bassin, R. H., Rein, A., and Gerwin, B. I. (1977). Loss of Fv-1 restriction in Balb/3T3 cells following infection with a single N tropic murine leukemia virus particle. *Cell* **10**(3)**,** 479-88.

Durbin, R. K., and Manning, J. S. (1982). The core of murine leukemia virus requires phosphate for structural stability. *Virology* **116**(1)**,** 31-9.

Einfeld, D., and Hunter, E. (1988). Oligomeric structure of a prototype retrovirus glycoprotein. *Proc Natl Acad Sci U S A* **85**(22)**,** 8688-92.

Elis, E., Ehrlich, M., Prizan-Ravid, A., Laham-Karam, N., and Bacharach, E. (2012). p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. *PLoS Pathog* **8**(12)**,** e1003103.

Emmott, E., and Hiscox, J. A. (2009). Nucleolar targeting: the hub of the matter. *EMBO Reports* **10**(3)**,** 231-8.

Engelman, A., Englund, G., Orenstein, J. M., Martin, M. A., and Craigie, R. (1995). Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J Virol* **69**(5)**,** 2729-36.

Ernst, R. K., Bray, M., Rekosh, D., and Hammarskjold, M. L. (1997). A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Molecular and Cellular Biology* **17**(1)**,** 135-44.

Fass, D., Harrison, S. C., and Kim, P. S. (1996). Retrovirus envelope domain at 1.7 angstrom resolution. *Nature Structural Biology* **3**(5)**,** 465-9.

Fassati, A., and Goff, S. P. (1999). Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. *J Virol* **73**(11)**,** 8919-25.

Fassati, A., and Goff, S. P. (2001). Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *Journal of Virology* **75**(8)**,** 3626-35.

Felber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T., and Pavlakis, G. N. (1989). rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci U S A* **86**(5)**,** 1495-9.

Felice, B., Cattoglio, C., Cittaro, D., Testa, A., Miccio, A., Ferrari, G., Luzi, L., Recchia, A., and Mavilio, F. (2009). Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. *PloS one* **4**(2)**,** e4571.

Feng, S., and Holland, E. C. (1988). HIV-1 tat trans-activation requires the loop sequence within tar. *Nature* **334**(6178)**,** 165-7.

Feng, Y., Broder, C. C., Kennedy, P. E., and Berger, E. A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **272**(5263)**,** 872-7.

Fischinger, P. J., Nomura, S., and Bolognesi, D. P. (1975). A novel murine oncornavirus with dual eco- and xenotropic properties. *Proc Natl Acad Sci U S A* **72**(12)**,** 5150-5.

Ford, M. G., Mills, I. G., Peter, B. J., Vallis, Y., Praefcke, G. J., Evans, P. R., and McMahon, H. T. (2002). Curvature of clathrin-coated pits driven by epsin. *Nature* **419**(6905)**,** 361-6.

Forshey, B. M., von Schwedler, U., Sundquist, W. I., and Aiken, C. (2002). Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J Virol* **76**(11)**,** 5667-77.

Forster, F., Medalia, O., Zauberman, N., Baumeister, W., and Fass, D. (2005). Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proc Natl Acad Sci U S A* **102**(13)**,** 4729-34.

Franke, E. K., Yuan, H. E., Bossolt, K. L., Goff, S. P., and Luban, J. (1994). Specificity and sequence requirements for interactions between various retroviral Gag proteins. *J Virol* **68**(8)**,** 5300-5.

Fricke, T., Valle-Casuso, J. C., White, T. E., Brandariz-Nunez, A., Bosche, W. J., Reszka, N., Gorelick, R., and Diaz-Griffero, F. (2013). The ability of TNPO3-depleted cells to inhibit HIV-1 infection requires CPSF6. *Retrovirology* **10**(1)**,** 46.

Fu, W., Gorelick, R. J., and Rein, A. (1994). Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. *J Virol* **68**(8)**,** 5013-8.

Fujiwara, T., and Mizuuchi, K. (1988). Retroviral DNA integration: structure of an integration intermediate. *Cell* **54**(4)**,** 497-504.

Fuller, S. D., Wilk, T., Gowen, B. E., Krausslich, H. G., and Vogt, V. M. (1997). Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Current Biology* **7**(10)**,** 729-38.

Furuichi, Y., Shatkin, A. J., Stavnezer, E., and Bishop, J. M. (1975). Blocked, methylated 5'-terminal sequence in avian sarcoma virus RNA. *Nature* **257**(5527)**,** 618-20.

Gallay, P., Hope, T., Chin, D., and Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc Natl Acad Sci U S A* **94**(18)**,** 9825-30.

Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., and et al. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**(4648)**,** 500-3.

Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **87**(7)**,** 1285-94.

Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **278**(5339)**,** 849-53.

Ganser-Pornillos, B. K., Chandrasekaran, V., Pornillos, O., Sodroski, J. G., Sundquist, W. I., and Yeager, M. (2011). Hexagonal assembly of a restricting TRIM5alpha protein. *Proc Natl Acad Sci U S A* **108**(2)**,** 534-9.

Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Cote, M., Rich, R. L., Myszka, D. G., and Sundquist, W. I. (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**(1)**,** 55-65.

Gheysen, D., Jacobs, E., de Foresta, F., Thiriart, C., Francotte, M., Thines, D., and De Wilde, M. (1989). Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* **59**(1)**,** 103-12.

Gilboa, E., Mitra, S. W., Goff, S., and Baltimore, D. (1979). A detailed model of reverse transcription and tests of crucial aspects. *Cell* **18**(1)**,** 93-100.

Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**(5272)**,** 231-5.

Goldstone, D. C., Ennis-Adeniran, V., Hedden, J. J., Groom, H. C., Rice, G. I., Christodoulou, E., Walker, P. A., Kelly, G., Haire, L. F., Yap, M. W., de Carvalho, L. P., Stoye, J. P., Crow, Y. J., Taylor, I. A., and Webb, M. (2011). HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* **480**(7377)**,** 379-82.

Goldstone, D. C., Yap, M. W., Robertson, L. E., Haire, L. F., Taylor, W. R., Katzourakis, A., Stoye, J. P., and Taylor, I. A. (2010). Structural and functional analysis of prehistoric lentiviruses uncovers an ancient molecular interface. *Cell Host Microbe* **8**(3)**,** 248-59.

Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a "zinc finger-like" protein sequence. *Proc Natl Acad Sci U S A* **85**(22)**,** 8420-4.

Gottlinger, H. G., Dorfman, T., Sodroski, J. G., and Haseltine, W. A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc Natl Acad Sci U S A* **88**(8)**,** 3195-9.

Gottlinger, H. G., Sodroski, J. G., and Haseltine, W. A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A* **86**(15)**,** 5781-5.

Gottwein, E., Bodem, J., Muller, B., Schmechel, A., Zentgraf, H., and Krausslich, H. G. (2003). The Mason-Pfizer monkey virus PPPY and PSAP motifs both contribute to virus release. *J Virol* **77**(17)**,** 9474-85.

Grattinger, M., Hohenberg, H., Thomas, D., Wilk, T., Muller, B., and Krausslich, H. G. (1999). In vitro assembly properties of wild-type and cyclophilin-binding defective human immunodeficiency virus capsid proteins in the presence and absence of cyclophilin A. *Virology* **257**(1)**,** 247-60.

Graves, B. J., Eisenman, R. N., and McKnight, S. L. (1985). Delineation of transcriptional control signals within the Moloney murine sarcoma virus long terminal repeat. *Molecular and Cellular Biology* **5**(8)**,** 1948-58.

Griffin, B., and Adams, S. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281,** 269-272.

Grigorov, B., Decimo, D., Smagulova, F., Pechoux, C., Mougel, M., Muriaux, D., and Darlix, J. L. (2007). Intracellular HIV-1 Gag localization is impaired by mutations in the nucleocapsid zinc fingers. *Retrovirology* **4,** 54.

Gross, L. (1957). Development and serial cellfree passage of a highly potent strain of mouse leukemia virus. *Proc Soc Exp Biol Med* **94**(4)**,** 767-71.

Guizetti, J., Schermelleh, L., Mantler, J., Maar, S., Poser, I., Leonhardt, H., Muller-Reichert, T., and Gerlich, D. W. (2011). Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* **331**(6024)**,** 1616-20.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B. H., Wintergerst, U., Velez, M. C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F. D., Fischer, A., and Cavazzana-Calvo, M. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *Journal of Clinical Investigation* **118**(9)**,** 3132-42.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J. I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L. E., Wissler, M., Prinz, C., Rabbitts, T. H., Le Deist, F., Fischer, A., and Cavazzana-Calvo, M. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**(5644)**,** 415-9.

Hadravova, R., de Marco, A., Ulbrich, P., Stokrova, J., Dolezal, M., Pichova, I., Ruml, T., Briggs, J. A., and Rumlova, M. (2012). In vitro assembly of virus-like particles of a gammaretrovirus, the murine leukemia virus XMRV. *J Virol* **86**(3)**,** 1297-306.

Hanson, P. I., Roth, R., Lin, Y., and Heuser, J. E. (2008). Plasma membrane deformation by circular arrays of ESCRT-III protein filaments. *Journal of Cell Biology* **180**(2)**,** 389-402.

Harada, F., Peters, G. G., and Dahlberg, J. E. (1979). The primer tRNA for Moloney murine leukemia virus DNA synthesis. Nucleotide sequence and aminoacylation of tRNAPro. *Journal of Biological Chemistry* **254**(21)**,** 10979-85.

Hare, S., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**(7286)**,** 232-6.

Hartley, J. W., Rowe, W. P., and Huebner, R. J. (1970). Host-range restrictions of murine leukemia viruses in mouse embryo cell cultures. *J Virol* **5**(2)**,** 221-5.

Haseltine, W. A., Kleid, D. G., Panet, A., Rothenberg, E., and Baltimore, D. (1976). Ordered transcription of RNA tumor virus genomes. *J Mol Biol* **106**(1)**,** 109-31.

Hatziioannou, T., Cowan, S., Goff, S. P., Bieniasz, P. D., and Towers, G. J. (2003). Restriction of multiple divergent retroviruses by Lv1 and Ref1. *EMBO Journal* **22**(3)**,** 385-94.

Hatziioannou, T., Perez-Caballero, D., Cowan, S., and Bieniasz, P. D. (2005). Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. *J Virol* **79**(1)**,** 176-83.

Heinzinger, N. K., Bukinsky, M. I., Haggerty, S. A., Ragland, A. M., Kewalramani, V., Lee, M. A., Gendelman, H. E., Ratner, L., Stevenson, M., and Emerman, M. (1994). 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**(15)**,** 7311-5.

Henderson, L. E., Copeland, T. D., Sowder, R. C., Smythers, G. W., and Oroszlan, S. (1981). Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses. *Journal of Biological Chemistry* **256**(16)**,** 8400-6.

Henderson, L. E., Gilden, R. V., and Oroszlan, S. (1979). Amino acid sequence homology between histone H5 and murine leukemia virus phosphoprotein p12. *Science* **203,** 1346.

Henderson, L. E., Krutzsch, H. C., and Oroszlan, S. (1983). Myristyl amino-terminal acylation of murine retrovirus proteins: an unusual post-translational proteins modification. *Proc Natl Acad Sci U S A* **80**(2)**,** 339-43.

Hilditch, L., Matadeen, R., Goldstone, D. C., Rosenthal, P. B., Taylor, I. A., and Stoye, J. P. (2011). Ordered assembly of murine leukemia virus capsid protein on lipid nanotubes directs specific binding by the restriction factor, Fv1. *Proc Natl Acad Sci U S A* **108**(14)**,** 5771-6.

Hizi, A., Henderson, L. E., Copeland, T. D., Sowder, R. C., Krutzsch, H. C., and Oroszlan, S. (1989). Analysis of gag proteins from mouse mammary tumor virus. *J Virol* **63**(6)**,** 2543-9.

Hofmann, W., Schubert, D., LaBonte, J., Munson, L., Gibson, S., Scammell, J., Ferrigno, P., and Sodroski, J. (1999). Species-specific, postentry barriers to primate immunodeficiency virus infection. *J Virol* **73**(12)**,** 10020-8.

Holzschu, D. L., Martineau, D., Fodor, S. K., Vogt, V. M., Bowser, P. R., and Casey, J. W. (1995). Nucleotide sequence and protein analysis of a complex piscine retrovirus, walleye dermal sarcoma virus. *J Virol* **69**(9)**,** 5320-31.

Howe, S. J., Mansour, M. R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M. H., Pike-Overzet, K., Chatters, S. J., de Ridder, D., Gilmour, K. C., Adams, S., Thornhill, S. I., Parsley, K. L., Staal, F. J., Gale, R. E., Linch, D. C., Bayford, J., Brown, L., Quaye, M., Kinnon, C., Ancliff, P., Webb, D. K., Schmidt, M., von Kalle, C., Gaspar, H. B., and Thrasher, A. J. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *Journal of Clinical Investigation* **118**(9)**,** 3143-50.

Hu, W. S., and Hughes, S. H. (2012). HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* **2**(10).

Hughes, S. H., Shank, P. R., Spector, D. H., Kung, H. J., Bishop, J. M., Varmus, H. E., Vogt, P. K., and Breitman, M. L. (1978). Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. *Cell* **15**(4)**,** 1397-410.

Hulme, A. E., Perez, O., and Hope, T. J. (2011). Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc Natl Acad Sci U S A* **108**(24)**,** 9975-80.

Ikeda, H., Laigret, F., Martin, M. A., and Repaske, R. (1985). Characterization of a molecularly cloned retroviral sequence associated with Fv-4 resistance. *J Virol* **55**(3)**,** 768-77.

Ikeda, H., and Odaka, T. (1984). A cell membrane "gp70" associated with Fv-4 gene: immunological characterization, and tissue and strain distribution. *Virology* **133**(1)**,** 65-76.

Ikuta, K., and Luftig, R. B. (1988). Detection of phosphorylated forms of Moloney murine leukemia virus major capsid protein p30 by immunoprecipitation and two-dimensional gel electrophoresis. *J Virol* **62**(1)**,** 40-6.

Inaguma, Y., Miyashita, N., Moriwaki, K., Huai, W. C., Jin, M. L., He, X. Q., and Ikeda, H. (1991). Acquisition of two endogenous ecotropic murine leukemia viruses in distinct Asian wild mouse populations. *J Virol* **65**(4)**,** 1796-802.

Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**(6153)**,** 280-3.

Jamjoom, G. A., Naso, R. B., and Arlinghaus, R. B. (1977). Further characterization of intracellular precursor polyproteins of Rauscher leukemia virus. *Virology* **78**(1)**,** 11-34.

Jarrosson-Wuilleme, L., Goujon, C., Bernaud, J., Rigal, D., Darlix, J. L., and Cimarelli, A. (2006). Transduction of nondividing human macrophages with gammaretrovirus-derived vectors. *Journal of Virology* **80**(3)**,** 1152-9.

Johnson, M. C., Scobie, H. M., Ma, Y. M., and Vogt, V. M. (2002). Nucleic acid-independent retrovirus assembly can be driven by dimerization. *J Virol* **76**(22)**,** 11177-85.

Jolicoeur, P., and Baltimore, D. (1976). Effect of Fv-1 gene product on proviral DNA formation and integration in cells infected with murine leukemia viruses. *Proc Natl Acad Sci U S A* **73**(7)**,** 2236-40.

Jouvenet, N., Bieniasz, P. D., and Simon, S. M. (2008). Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* **454**(7201)**,** 236-40.

Jouvenet, N., Simon, S. M., and Bieniasz, P. D. (2009). Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc Natl Acad Sci U S A* **106**(45)**,** 19114-9.

Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T., and Oroszlan, S. (1985). Murine leukemia virus maturation: protease region required for conversion from "immature" to "mature" core form and for virus infectivity. *Virology* **145**(2)**,** 280-92.

Kent, K. P., Childs, W., and Boxer, S. G. (2008). Deconstructing green fluorescent protein. *Journal of the American Chemical Society* **130**(30)**,** 9664-5.

Kim, B., Nguyen, L. A., Daddacha, W., and Hollenbaugh, J. A. (2012). Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. *Journal of Biological Chemistry* **287**(26)**,** 21570-4.

Kim, J. W., Closs, E. I., Albritton, L. M., and Cunningham, J. M. (1991). Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. *Nature* **352**(6337)**,** 725-8.

Kirchhausen, T. (2000). Clathrin. *Annual Review of Biochemistry* **69,** 699-727.

Knejzlik, Z., Smekalova, Z., Ruml, T., and Sakalian, M. (2007). Multimerization of the p12 domain is necessary for Mason-Pfizer monkey virus Gag assembly in vitro. *Virology* **365**(2)**,** 260-70.

Knejzlik, Z., Strohalm, M., Sedlackova, L., Kodicek, M., Sakalian, M., and Ruml, T. (2004). Isolation and characterization of the Mason-Pfizer monkey virus p12 protein. *Virology* **324**(1)**,** 204-12.

Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M., and Sigal, I. S. (1988). Active human immunodeficiency virus protease is required for viral infectivity. *Proc Natl Acad Sci U S A* **85**(13)**,** 4686-90.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992). Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**(5065)**,** 1783-90.

Konig, R., Zhou, Y., Elleder, D., Diamond, T. L., Bonamy, G. M., Irelan, J. T., Chiang, C. Y., Tu, B. P., De Jesus, P. D., Lilley, C. E., Seidel, S., Opaluch, A. M., Caldwell, J. S., Weitzman, M. D., Kuhen, K. L., Bandyopadhyay, S., Ideker, T., Orth, A. P., Miraglia, L. J., Bushman, F. D., Young, J. A., and Chanda, S. K. (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **135**(1)**,** 49-60.

Kotov, A., Zhou, J., Flicker, P., and Aiken, C. (1999). Association of Nef with the human immunodeficiency virus type 1 core. *J Virol* **73**(10)**,** 8824-30.

Kozak, C. A., and Chakraborti, A. (1996). Single amino acid changes in the murine leukemia virus capsid protein gene define the target of Fv1 resistance. *Virology* **225**(2)**,** 300-5.

Krishnan, L., and Engelman, A. (2012). Retroviral integrase proteins and HIV-1 DNA integration. *Journal of Biological Chemistry* **287**(49)**,** 40858-66.

Krishnan, L., Matreyek, K. A., Oztop, I., Lee, K., Tipper, C. H., Li, X., Dar, M. J., Kewalramani, V. N., and Engelman, A. (2010). The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase. *Journal of Virology* **84**(1)**,** 397-406.

Kutluay, S. B., and Bieniasz, P. D. (2010). Analysis of the initiating events in HIV-1 particle assembly and genome packaging. *PLoS Pathog* **6**(11)**,** e1001200.

Kyere, S. K., Joseph, P. R. B., and Summers, M. F. (2008). The p12 domain is unstructured in a murine leukemia virus p12-CA(N) Gag construct. *PloS One* **3**(4)**,** e1902.

Laguette, N., Bregnard, C., Benichou, S., and Basmaciogullari, S. (2010). Human immunodeficiency virus (HIV) type-1, HIV-2 and simian immunodeficiency virus Nef proteins. *Mol Aspects Med* **31**(5)**,** 418-33.

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Segeral, E., Yatim, A., Emiliani, S., Schwartz, O., and Benkirane, M. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **474**(7353)**,** 654-7.

Lahouassa, H., Daddacha, W., Hofmann, H., Ayinde, D., Logue, E. C., Dragin, L., Bloch, N., Maudet, C., Bertrand, M., Gramberg, T., Pancino, G., Priet, S., Canard, B., Laguette, N., Benkirane, M., Transy, C., Landau, N. R., Kim, B., and Margottin-Goguet, F. (2012). SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nature Immunology* **13**(3)**,** 223-8.

Lai, M. M., and Duesberg, P. H. (1972). Adenylic acid-rich sequence in RNAs of Rous sarcoma virus and Rauscher mouse leukaemia virus. *Nature* **235**(5338)**,** 383-6.

Laimins, L. A., Gruss, P., Pozzatti, R., and Khoury, G. (1984). Characterization of enhancer elements in the long terminal repeat of Moloney murine sarcoma virus. *J Virol* **49**(1)**,** 183-9.

Laimins, L. A., Tsichlis, P., and Khoury, G. (1984). Multiple enhancer domains in the 3' terminus of the Prague strain of Rous sarcoma virus. *Nucleic Acids Research* **12**(16)**,** 6427-42.

Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E., and Corbett, A. H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. *Journal of Biological Chemistry* **282**(8)**,** 5101-5.

Lanman, J., Crum, J., Deerinck, T. J., Gaietta, G. M., Schneemann, A., Sosinsky, G. E., Ellisman, M. H., and Johnson, J. E. (2008). Visualizing flock house virus infection in Drosophila cells with correlated fluorescence and electron microscopy. *Journal of Structural Biology* **161**(3)**,** 439-46.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**(21)**,** 2947-8.

Lee, K., Ambrose, Z., Martin, T. D., Oztop, I., Mulky, A., Julias, J. G., Vandegraaff, N., Baumann, J. G., Wang, R., Yuen, W., Takemura, T., Shelton, K., Taniuchi, I., Li, Y., Sodroski, J., Littman, D. R., Coffin, J. M., Hughes, S. H., Unutmaz, D., Engelman, A., and KewalRamani, V. N. (2010). Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7**(3)**,** 221-33.

Lee, S., and Nagashima, K. (2005). Cooperative effect of gag proteins p12 and capsid during early events of murine leukemia virus replication. *Journal of Virology* **79,** 4159-4169.

Lelek, M., Di Nunzio, F., Henriques, R., Charneau, P., Arhel, N., and Zimmer, C. (2012). Superresolution imaging of HIV in infected cells with FlAsH-PALM. *Proc Natl Acad Sci U S A* **109**(22)**,** 8564-9.

Levy, J. A. (1973). Xenotropic viruses: murine leukemia viruses associated with NIH Swiss, NZB, and other mouse strains. *Science* **182**(4117)**,** 1151-3.

Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. M., and Oshiro, L. S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225**(4664)**,** 840-2.

Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2006). Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog* **2**(6)**,** e60.

Lewis, P., Hensel, M., and Emerman, M. (1992). Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO Journal* **11**(8)**,** 3053-8.

Li, F., Goila-Gaur, R., Salzwedel, K., Kilgore, N. R., Reddick, M., Matallana, C., Castillo, A., Zoumplis, D., Martin, D. E., Orenstein, J. M., Allaway, G. P., Freed, E. O., and Wild, C. T. (2003). PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *Proc Natl Acad Sci U S A* **100**(23)**,** 13555-60.

Li, L., Olvera, J. M., Yoder, K. E., Mitchell, R. S., Butler, S. L., Lieber, M., Martin, S. L., and Bushman, F. D. (2001). Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *EMBO Journal* **20**(12)**,** 3272-81.

Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**(6802)**,** 409-13.

Li, X., and Sodroski, J. (2008). The TRIM5alpha B-box 2 domain promotes cooperative binding to the retroviral capsid by mediating higher-order self-association. *J Virol* **82**(23)**,** 11495-502.

Lin, T. Y., and Emerman, M. (2006). Cyclophilin A interacts with diverse lentiviral capsids. *Retrovirology* **3,** 70.

Liu, J., Bartesaghi, A., Borgnia, M. J., Sapiro, G., and Subramaniam, S. (2008). Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**(7209)**,** 109-13.

Liu, X. H., Xu, W., Russ, J., Eiden, L. E., and Eiden, M. V. (2011). The host range of gammaretroviruses and gammaretroviral vectors includes post-mitotic neural cells. *PloS one* **6**(3)**,** e18072.

Lixin, R., Efthymiadis, A., Henderson, B., and Jans, D. A. (2001). Novel Properties of the Nucleolar Targeting Signal of Human Angiogenin. *Biochemical and Biophysical Research Communications* **284**(1)**,** 185-193.

Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W., and Poeschla, E. M. (2006). An essential role for LEDGF/p75 in HIV integration. *Science* **314**(5798)**,** 461-4.

Low, A., Datta, S., Kuznetsov, Y., Jahid, S., Kothari, N., McPherson, A., and Fan, H. (2007). Mutation in the glycosylated gag protein of murine leukemia virus results in reduced in vivo infectivity and a novel defect in viral budding or release. *J Virol* **81**(8)**,** 3685-92.

Lu, Y. L., Spearman, P., and Ratner, L. (1993). Human immunodeficiency virus type 1 viral protein R localization in infected cells and virions. *J Virol* **67**(11)**,** 6542-50.

Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V., and Goff, S. P. (1993). Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell* **73**(6)**,** 1067-78.

Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *Journal of Biological Chemistry* **278**(35)**,** 33528-39.

Maertens, G. N., Hare, S., and Cherepanov, P. (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **468**(7321)**,** 326-9.

Malim, M. H., and Bieniasz, P. D. (2012). HIV Restriction Factors and Mechanisms of Evasion. *Cold Spring Harb Perspect Med* **2**(5)**,** a006940.

Mammano, F., Ohagen, A., Hoglund, S., and Gottlinger, H. G. (1994). Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J Virol* **68**(8)**,** 4927-36.

Mangel, W. F., Delius, H., and Duesberg, P. H. (1974). Structure and molecular weight of the 60-70S RNA and the 30-40S RNA of the Rous sarcoma virus. *Proc Natl Acad Sci U S A* **71**(11)**,** 4541-5.

Marechal, V., Clavel, F., Heard, J. M., and Schwartz, O. (1998). Cytosolic Gag p24 as an index of productive entry of human immunodeficiency virus type 1. *J Virol* **72**(3)**,** 2208-12.

Marshall, H. M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E., and Bushman, F. D. (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PloS one* **2**(12)**,** e1340.

Martin-Serrano, J. (2007). The role of ubiquitin in retroviral egress. *Traffic* **8,** 1297-303.

Martin-Serrano, J., Eastman, S. W., Chung, W., and Bieniasz, P. D. (2005). HECT ubiquitin ligases link viral and cellular PPXY motifs to the vacuolar protein-sorting pathway. *Journal of Cell Biology* **168**(1)**,** 89-101.

Martin-Serrano, J., Zang, T., and Bieniasz, P. D. (2003). Role of ESCRT-I in retroviral budding. *Journal of Virology* **77,** 4794-4804.

McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M., and Hope, T. J. (2002). Visualization of the intracellular behavior of HIV in living cells. *Journal of Cell Biology* **159**(3)**,** 441-52.

Mirambeau, G., Lyonnais, S., Coulaud, D., Hameau, L., Lafosse, S., Jeusset, J., Borde, I., Reboud-Ravaux, M., Restle, T., Gorelick, R. J., and Le Cam, E. (2007). HIV-1 protease and reverse transcriptase control the architecture of their nucleocapsid partner. *PloS one* **2**(7)**,** e669.

Mire, C. E., Dube, D., Delos, S. E., White, J. M., and Whitt, M. A. (2009). Glycoprotein-dependent acidification of vesicular stomatitis virus enhances release of matrix protein. *J Virol* **83**(23)**,** 12139-50.

Mitchell, R. S., Beitzel, B. F., Schroder, A. R., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* **2**(8)**,** E234.

Montagnier, L. (2010). 25 years after HIV discovery: prospects for cure and vaccine. *Virology* **397**(2)**,** 248-54.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., Sergi Sergi, L., Benedicenti, F., Ambrosi, A., Di Serio, C., Doglioni, C., von Kalle, C., and Naldini, L. (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nature Biotechnology* **24**(6)**,** 687-96.

Morita, E., Sandrin, V., McCullough, J., Katsuyama, A., Baci Hamilton, I., and Sundquist, W. I. (2011). ESCRT-III protein requirements for HIV-1 budding. *Cell Host Microbe* **9**(3)**,** 235-42.

Mortuza, G. B., Dodding, M. P., Goldstone, D. C., Haire, L. F., Stoye, J. P., and Taylor, I. A. (2008). Structure of B-MLV capsid amino-terminal domain reveals key features of viral tropism, gag assembly and core formation. *J Mol Biol* **376**(5)**,** 1493-508.

Mortuza, G. B., Goldstone, D. C., Pashley, C., Haire, L. F., Palmarini, M., Taylor, W. R., Stoye, J. P., and Taylor, I. A. (2009). Structure of the capsid amino-terminal domain from the betaretrovirus, Jaagsiekte sheep retrovirus. *J Mol Biol* **386**(4)**,** 1179-92.

Mortuza, G. B., Haire, L. F., Stevens, A., Smerdon, S. J., Stoye, J. P., and Taylor, I. A. (2004). High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature* **431**(7007)**,** 481-5.

Mullers, E., Stirnnagel, K., Kaulfuss, S., and Lindemann, D. (2011). Prototype foamy virus gag nuclear localization: a novel pathway among retroviruses. *J Virol* **85**(18)**,** 9276-85.

Munk, C., Brandt, S. M., Lucero, G., and Landau, N. R. (2002). A dominant block to HIV-1 replication at reverse transcription in simian cells. *Proc Natl Acad Sci U S A* **99**(21)**,** 13843-8.

Muriaux, D., Costes, S., Nagashima, K., Mirro, J., Cho, E., Lockett, S., and Rein, A. (2004). Role of murine leukemia virus nucleocapsid protein in virus assembly. *Journal of Virology* **78,** 12378-12385.

Muriaux, D., Mirro, J., Harvin, D., and Rein, A. (2001). RNA is a structural element in retrovirus particles. *Proc Natl Acad Sci U S A* **98**(9)**,** 5246-51.

Mustafa, F., and Robinson, H. L. (1993). Context-dependent role of human immunodeficiency virus type 1 auxiliary genes in the establishment of chronic virus producers. *J Virol* **67**(11)**,** 6909-15.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M., and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**(5259)**,** 263-7.

Naso, R. B., Karshin, W. L., Wu, Y. H., and Arlinghaus, R. B. (1979). Characterization of 40,000- and 25,000-dalton intermediate precursors to Rauscher murine leukemia virus gag gene products. *J Virol* **32**(1)**,** 187-98.

Neil, S. J., Eastman, S. W., Jouvenet, N., and Bieniasz, P. D. (2006). HIV-1 Vpu promotes release and prevents endocytosis of nascent retrovirus particles from the plasma membrane. *PLoS Pathog* **2**(5)**,** e39.

Neil, S. J., Zang, T., and Bieniasz, P. D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**(7177)**,** 425-30.

Newman, R. M., Hall, L., Kirmaier, A., Pozzi, L. A., Pery, E., Farzan, M., O'Neil, S. P., and Johnson, W. (2008). Evolution of a TRIM5-CypA splice isoform in old world monkeys. *PLoS Pathog* **4**(2)**,** e1000003.

Nisole, S., Stoye, J. P., and Saib, A. (2005). TRIM family proteins: retroviral restriction and antiviral defence. *Nat Rev Microbiol* **3**(10)**,** 799-808.

Nitta, T., Tam, R., Kim, J. W., and Fan, H. (2011). The cellular protein La functions in enhancement of virus release through lipid rafts facilitated by murine leukemia virus glycosylated Gag. *MBio* **2**(1)**,** e00341-10.

Nowak, E., Potrzebowski, W., Konarev, P. V., Rausch, J. W., Bona, M. K., Svergun, D. I., Bujnicki, J. M., Le Grice, S. F., and Nowotny, M. (2013). Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Research* **41**(6)**,** 3874-87.

O'Doherty, U., Swiggard, W. J., and Malim, M. H. (2000). Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. *Journal of Virology* **74,** 10074-10080.

Ohkura, S., Goldstone, D. C., Yap, M. W., Holden-Dye, K., Taylor, I. A., and Stoye, J. P. (2011). Novel escape mutants suggest an extensive TRIM5alpha binding site spanning the entire outer surface of the murine leukemia virus capsid protein. *PLoS Pathog* **7**(3)**,** e1002011.

Ohkura, S., and Stoye, J. P. (2013). A Comparison of Murine Leukemia Viruses That Escape from Human and Rhesus Macaque TRIM5alphas. *J Virol* **87**(11)**,** 6455-68.

Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K., and Freed, E. O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc Natl Acad Sci U S A* **101**(41)**,** 14889-94.

Ono, A., and Freed, E. O. (1999). Binding of human immunodeficiency virus type 1 Gag to membrane: role of the matrix amino terminus. *J Virol* **73**(5)**,** 4136-44.

Oshima, M., Muriaux, D., Mirro, J., Nagashima, K., Dryden, K., Yeager, M., and Rein, A. (2004). Effects of Blocking Individual Maturation Cleavages in Murine Leukemia Virus Gag. *Journal of Virology* **78,** 1411-1420.

Ostrowski, M. C., Berard, D., and Hager, G. L. (1981). Specific transcriptional initiation in vitro on murine type C retrovirus promoters. *Proc Natl Acad Sci U S A* **78**(7)**,** 4485-9.

Ott, M., Geyer, M., and Zhou, Q. (2011). The control of HIV transcription: keeping RNA polymerase II on track. *Cell Host Microbe* **10**(5)**,** 426-35.

Pasquinelli, A. E., Ernst, R. K., Lund, E., Grimm, C., Zapp, M. L., Rekosh, D., Hammarskjold, M. L., and Dahlberg, J. E. (1997). The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO Journal* **16**(24)**,** 7500-10.

Paxton, W., Connor, R. I., and Landau, N. R. (1993). Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. *J Virol* **67**(12)**,** 7229-37.

Perez-Caballero, D., Hatziioannou, T., Yang, A., Cowan, S., and Bieniasz, P. D. (2005). Human tripartite motif 5alpha domains responsible for retrovirus restriction activity and specificity. *J Virol* **79**(14)**,** 8969-78.

Pertel, T., Hausmann, S., Morger, D., Zuger, S., Guerra, J., Lascano, J., Reinhard, C., Santoni, F. A., Uchil, P. D., Chatel, L., Bisiaux, A., Albert, M. L., Strambio-De-Castillia, C., Mothes, W., Pizzato, M., Grutter, M. G., and Luban, J. (2011). TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* **472**(7343)**,** 361-5.

Peters, G., Harada, F., Dahlberg, J. E., Panet, A., Haseltine, W. A., and Baltimore, D. (1977). Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. *J Virol* **21**(3)**,** 1031-41.

Petropoulos, C. (1997). Retroviral Taxonomy, Protein Structures, Sequences, and Genetic Maps. *In* "Retroviruses" (J. M. Coffin, S. H. Hughes, and H. E. Varmus, Eds.). Cold Spring Harbor Laboratory Press, New York.

Pettit, S. C., Henderson, G. J., Schiffer, C. A., and Swanstrom, R. (2002). Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. *J Virol* **76**(20)**,** 10226-33.

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

Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A., and Swanstrom, R. (1994). The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J Virol* **68**(12)**,** 8017-27.

Pizzato, M. (2010). MLV glycosylated-Gag is an infectivity factor that rescues Nef-deficient HIV-1. *Proc Natl Acad Sci U S A* **107**(20)**,** 9364-9.

Poiesz, B. J., Ruscetti, F. W., Gazdar, A. F., Bunn, P. A., Minna, J. D., and Gallo, R. C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America* **77**(12)**,** 7415-9.

Popov, S., Strack, B., Sanchez-Merino, V., Popova, E., Rosin, H., and Gottlinger, H. G. (2011). Human immunodeficiency virus type 1 and related primate lentiviruses engage clathrin through Gag-Pol or Gag. *J Virol* **85**(8)**,** 3792-801.

Pornillos, O., Ganser-Pornillos, B. K., Banumathi, S., Hua, Y., and Yeager, M. (2010). Disulfide bond stabilization of the hexameric capsomer of human immunodeficiency virus. *J Mol Biol* **401**(5)**,** 985-95.

Pornillos, O., Ganser-Pornillos, B. K., Kelly, B. N., Hua, Y., Whitby, F. G., Stout, C. D., Sundquist, W. I., Hill, C. P., and Yeager, M. (2009). X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**(7)**,** 1282-92.

Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011). Atomic-level modelling of the HIV capsid. *Nature* **469**(7330)**,** 424-7.

Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002). Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure* **10**(3)**,** 369-81.

Prats, A. C., De Billy, G., Wang, P., and Darlix, J. L. (1989). CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J Mol Biol* **205**(2)**,** 363-72.

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

Prizan-Ravid, A., Elis, E., Laham-Karam, N., Selig, S., Ehrlich, M., and Bacharach, E. (2010). The Gag cleavage product, p12, is a functional constituent of the murine leukemia virus pre-integration complex. *PLoS Pathog* **6**(11)**,** e1001183.

Puffer, B. A., Parent, L. J., Wills, J. W., and Montelaro, R. C. (1997). Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein. *J Virol* **71**(9)**,** 6541-6.

Ramezani, A., Hawley, T. S., and Hawley, R. G. (2008). Combinatorial incorporation of enhancer-blocking components of the chicken beta-globin 5'HS4 and human T-cell receptor alpha/delta BEAD-1 insulators in self-inactivating retroviral vectors reduces their genotoxic potential. *Stem Cells* **26**(12)**,** 3257-66.

Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., and et al. (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**(6000)**,** 277-84.

Rauch, S., and Martin-Serrano, J. (2011). Multiple interactions between the ESCRT machinery and arrestin-related proteins: implications for PPXY-dependent budding. *J Virol* **85**(7)**,** 3546-56.

Reicin, A. S., Ohagen, A., Yin, L., Hoglund, S., and Goff, S. P. (1996). The role of Gag in human immunodeficiency virus type 1 virion morphogenesis and early steps of the viral life cycle. *J Virol* **70**(12)**,** 8645-52.

Rein, A., McClure, M. R., Rice, N. R., Luftig, R. B., and Schultz, A. M. (1986). Myristylation site in Pr65gag is essential for virus particle formation by Moloney murine leukemia virus. *Proc Natl Acad Sci U S A* **83**(19)**,** 7246-50.

Richman, D. D., Margolis, D. M., Delaney, M., Greene, W. C., Hazuda, D., and Pomerantz, R. J. (2009). The challenge of finding a cure for HIV infection. *Science* **323**(5919)**,** 1304-7.

Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C., and Harrison, S. C. (1995). The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A* **92**(4)**,** 1222-6.

Roe, T., Reynolds, T. C., Yu, G., and Brown, P. O. (1993). Integration of murine leukemia virus DNA depends on mitosis. *EMBO Journal* **12**(5)**,** 2099-108.

Roth, M. J., Schwartzberg, P. L., and Goff, S. P. (1989). Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence on IN function and terminal DNA sequence. *Cell* **58**(1)**,** 47-54.

Rous, P. (1910). A Transmissible Avian Neoplasm. (Sarcoma of the Common Fowl.). *Journal of Experimental Medicine* **12**(5)**,** 696-705.

Saad, J. S., Miller, J., Tai, J., Kim, A., Ghanam, R. H., and Summers, M. F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc Natl Acad Sci U S A* **103**(30)**,** 11364-9.

Sawyer, S. L., Wu, L. I., Emerman, M., and Malik, H. S. (2005). Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci U S A* **102**(8)**,** 2832-7.

Sayah, D. M., Sokolskaja, E., Berthoux, L., and Luban, J. (2004). Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* **430**(6999)**,** 569-73.

Schaller, T., Ocwieja, K. E., Rasaiyaah, J., Price, A. J., Brady, T. L., Roth, S. L., Hue, S., Fletcher, A. J., Lee, K., KewalRamani, V. N., Noursadeghi, M., Jenner, R. G., James, L. C., Bushman, F. D., and Towers, G. J. (2011). HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog* **7**(12)**,** e1002439.

Schaller, T., Ylinen, L. M., Webb, B. L., Singh, S., and Towers, G. J. (2007). Fusion of cyclophilin A to Fv1 enables cyclosporine-sensitive restriction of human and feline immunodeficiency viruses. *J Virol* **81**(18)**,** 10055-63.

Scheifele, L. Z., Kenney, S. P., Cairns, T. M., Craven, R. C., and Parent, L. J. (2007). Overlapping roles of the Rous sarcoma virus Gag p10 domain in nuclear export and virion core morphology. *J Virol* **81**(19)**,** 10718-28.

Scheifele, L. Z., Ryan, E. P., and Parent, L. J. (2005). Detailed mapping of the nuclear export signal in the Rous sarcoma virus Gag protein. *J Virol* **79**(14)**,** 8732-41.

Schneider, W. M., Brzezinski, J. D., Aiyer, S., Malani, N., Gyuricza, M., Bushman, F. D., and Roth, M. J. (2013). Viral DNA tethering domains complement replication-defective mutations in the p12 protein of MuLV Gag. *Proc Natl Acad Sci U S A*.

Schrijvers, R., Vets, S., De Rijck, J., Malani, N., Bushman, F. D., Debyser, Z., and Gijsbers, R. (2012). HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology* **9,** 84.

Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**(4)**,** 521-9.

Schultz, A. M., and Oroszlan, S. (1983). In vivo modification of retroviral gag gene-encoded polyproteins by myristic acid. *J Virol* **46**(2)**,** 355-61.

Sen, A., Sherr, C. J., and Todaro, G. J. (1976). Specific binding of the type C viral core protein p12 with purified viral RNA. *Cell* **7**(1)**,** 21-32.

Sen, A., Sherr, C. J., and Todaro, G. J. (1977). Phosphorylation of murine type C viral p12 proteins regulates their extent of binding to the homologous viral RNA. *Cell* **10**(3)**,** 489-96.

Sen, A., and Todaro, G. J. (1977). The genome-associated, specific RNA binding proteins of avian and mammalian type C viruses. *Cell* **10**(1)**,** 91-9.

Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat Methods* **2**(12)**,** 905-9.

Shank, P. R., Hughes, S. H., Kung, H. J., Majors, J. E., Quintrell, N., Guntaka, R. V., Bishop, J. M., and Varmus, H. E. (1978). Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. *Cell* **15**(4)**,** 1383-95.

Sharifi, H. J., Furuya, A. M., and de Noronha, C. M. (2012). The role of HIV-1 Vpr in promoting the infection of nondividing cells and in cell cycle arrest. *Curr Opin HIV AIDS* **7**(2)**,** 187-94.

Sheehy, A. M., Gaddis, N. C., Choi, J. D., and Malim, M. H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**(6898)**,** 646-50.

Sheehy, A. M., Gaddis, N. C., and Malim, M. H. (2003). The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature Medicine* **9**(11)**,** 1404-7.

Shi, J., Zhou, J., Shah, V. B., Aiken, C., and Whitby, K. (2011). Small-molecule inhibition of human immunodeficiency virus type 1 infection by virus capsid destabilization. *J Virol* **85**(1)**,** 542-9.

Shoemaker, C., Hoffman, J., Goff, S., and Baltimore, D. (1981). Intramolecular integration within Moloney murine leukemia virus DNA. *Journal of Virology* **40,** 164-172.

Smagulova, F., Maurel, S., Morichaud, Z., Devaux, C., Mougel, M., and Houzet, L. (2005). The highly structured encapsidation signal of MuLV RNA is involved in the nuclear export of its unspliced RNA. *Journal of Molecular Biology* **354,** 1118-1128.

Sommerfelt, M. A., Rhee, S. S., and Hunter, E. (1992). Importance of p12 protein in Mason-Pfizer monkey virus assembly and infectivity. *J Virol* **66**(12)**,** 7005-11.

Soneoka, Y., Cannon, P. M., Ramsdale, E. E., Griffiths, J. C., Romano, G., Kingsman, S. M., and Kingsman, A. J. (1995). A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Research* **23**(4)**,** 628-33.

Song, B., Gold, B., O'Huigin, C., Javanbakht, H., Li, X., Stremlau, M., Winkler, C., Dean, M., and Sodroski, J. (2005). The B30.2(SPRY) domain of the retroviral restriction factor TRIM5alpha exhibits lineage-specific length and sequence variation in primates. *J Virol* **79**(10)**,** 6111-21.

Spearman, P., Wang, J. J., Vander Heyden, N., and Ratner, L. (1994). Identification of human immunodeficiency virus type 1 Gag protein domains essential to membrane binding and particle assembly. *J Virol* **68**(5)**,** 3232-42.

Spidel, J. L., Craven, R. C., Wilson, C. B., Patnaik, A., Wang, H., Mansky, L. M., and Wills, J. W. (2004). Lysines close to the Rous sarcoma virus late domain critical for budding. *J Virol* **78**(19)**,** 10606-16.

St Gelais, C., and Wu, L. (2011). SAMHD1: a new insight into HIV-1 restriction in myeloid cells. *Retrovirology* **8,** 55.

Stavrou, S., Nitta, T., Kotla, S., Ha, D., Nagashima, K., Rein, A. R., Fan, H., and Ross, S. R. (2013). Murine leukemia virus glycosylated Gag blocks apolipoprotein B editing complex 3 and cytosolic sensor access to the reverse transcription complex. *Proc Natl Acad Sci U S A* **110**(22)**,** 9078-83.

Stevens, A., Bock, M., Ellis, S., LeTissier, P., Bishop, K. N., Yap, M. W., Taylor, W., and Stoye, J. P. (2004). Retroviral capsid determinants of Fv1 NB and NR tropism. *J Virol* **78**(18)**,** 9592-8.

Strack, B., Calistri, A., Craig, S., Popova, E., and Gottlinger, H. G. (2003). AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell* **114**(6)**,** 689-99.

Stremlau, M., Owens, C. M., Perron, M. J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature* **427**(6977)**,** 848-53.

Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., Diaz-Griffero, F., Anderson, D. J., Sundquist, W. I., and Sodroski, J. (2006). Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci U S A* **103**(14)**,** 5514-9.

Stremlau, M., Perron, M., Welikala, S., and Sodroski, J. (2005). Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. *J Virol* **79**(5)**,** 3139-45.

Studamire, B., and Goff, S. P. (2008). Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology* **5,** 48.

Suzuki, S. (1975). FV-4: a new gene affecting the splenomegaly induction by Friend leukemia virus. *Jpn J Exp Med* **45**(6)**,** 473-8.

Suzuki, Y., and Craigie, R. (2007). The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol* **5**(3)**,** 187-96.

Tailor, C. S., Nouri, A., Lee, C. G., Kozak, C., and Kabat, D. (1999). Cloning and characterization of a cell surface receptor for xenotropic and polytropic murine leukemia viruses. *Proc Natl Acad Sci U S A* **96**(3)**,** 927-32.

Tang, C., Ndassa, Y., and Summers, M. F. (2002). Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. *Nature Structural Biology* **9**(7)**,** 537-43.

Telesnitsky, A., and Goff, S. P. (1993). Two defective forms of reverse transcriptase can complement to restore retroviral infectivity. *EMBO Journal* **12**(11)**,** 4433-8.

Temin, H. M. (1963). The Effects of Actinomycin D on Growth of Rous Sarcoma Virus in Vitro. *Virology* **20,** 577-82.

Temin, H. M. (1964). Nature of the provirus of Rous sarcoma. *Journal of the National Cancer Institute Monographs* **17,** 557-570.

Temin, H. M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226**(5252)**,** 1211-3.

Temin, H. M., and Rubin, H. (1958). Characteristics of an assay for Rous sarcoma virus and Rous sarcoma cells in tissue culture. *Virology* **6**(3)**,** 669-88.

Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C. T., Sodroski, J., and Gottlinger, H. G. (1994). Functional association of cyclophilin A with HIV-1 virions. *Nature* **372**(6504)**,** 363-5.

Thomas, J. A., and Gorelick, R. J. (2008). Nucleocapsid protein function in early infection processes. *Virus Research* **134,** 39-63.

Tobaly-Tapiero, J., Bittoun, P., Lehmann-Che, J., Delelis, O., Giron, M. L., de The, H., and Saib, A. (2008). Chromatin tethering of incoming foamy virus by the structural Gag protein. *Traffic* **9**(10)**,** 1717-27.

Towers, G., Bock, M., Martin, S., Takeuchi, Y., Stoye, J. P., and Danos, O. (2000). A conserved mechanism of retrovirus restriction in mammals. *Proc Natl Acad Sci U S A* **97**(22)**,** 12295-9.

Towers, G., Collins, M., and Takeuchi, Y. (2002). Abrogation of Ref1 retrovirus restriction in human cells. *J Virol* **76**(5)**,** 2548-50.

Van Damme, N., Goff, D., Katsura, C., Jorgenson, R. L., Mitchell, R., Johnson, M. C., Stephens, E. B., and Guatelli, J. (2008). 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**(4)**,** 245-52.

Van Epps, H. L. (2005). Peyton Rous: father of the tumor virus. *Journal of Experimental Medicine* **201**(3)**,** 320.

Vandegraaff, N., Devroe, E., Turlure, F., Silver, P. A., and Engelman, A. (2006). Biochemical and genetic analyses of integrase-interacting proteins lens epithelium-derived growth factor (LEDGF)/p75 and hepatoma-derived growth factor related protein 2 (HRP2) in preintegration complex function and HIV-1 replication. *Virology* **346**(2)**,** 415-26.

Verderame, M. F., Nelle, T. D., and Wills, J. W. (1996). The membrane-binding domain of the Rous sarcoma virus Gag protein. *J Virol* **70**(4)**,** 2664-8.

Virgen, C. A., Kratovac, Z., Bieniasz, P. D., and Hatziioannou, T. (2008). Independent genesis of chimeric TRIM5-cyclophilin proteins in two primate species. *Proc Natl Acad Sci U S A* **105**(9)**,** 3563-8.

von Poblotzki, A., Wagner, R., Niedrig, M., Wanner, G., Wolf, H., and Modrow, S. (1993). Identification of a region in the Pr55gag-polyprotein essential for HIV-1 particle formation. *Virology* **193**(2)**,** 981-5.

von Schwedler, U., Kornbluth, R. S., and Trono, D. (1994). The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent T lymphocytes. *Proc Natl Acad Sci U S A* **91**(15)**,** 6992-6.

von Schwedler, U. K., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H., Davis, D. R., and Sundquist, W. I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO Journal* **17**(6)**,** 1555-68.

Waheed, A. A., and Freed, E. O. (2012). HIV type 1 Gag as a target for antiviral therapy. *Aids Research and Human Retroviruses* **28**(1)**,** 54-75.

Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., and Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**(1)**,** 9-17.

Wallin, M., Ekstrom, M., and Garoff, H. (2004). Isomerization of the intersubunit disulphide-bond in Env controls retrovirus fusion. *EMBO Journal* **23**(1)**,** 54-65.

Wang, H., Kavanaugh, M. P., North, R. A., and Kabat, D. (1991). Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. *Nature* **352**(6337)**,** 729-31.

Weiss, R. A. (2006). The discovery of endogenous retroviruses. *Retrovirology* **3,** 67.

Weldon, R. A., Jr., and Wills, J. W. (1993). Characterization of a small (25-kilodalton) derivative of the Rous sarcoma virus Gag protein competent for particle release. *J Virol* **67**(9)**,** 5550-61.

Whitt, M. A., and Mire, C. E. (2011). Utilization of fluorescently-labeled tetracysteine-tagged proteins to study virus entry by live cell microscopy. *Methods* **55**(2)**,** 127-36.

Wight, D. J., Boucherit, V. C., Nader, M., Allen, D. J., Taylor, I. A., and Bishop, K. N. (2012). The gammaretroviral p12 protein has multiple domains that function during the early stages of replication. *Retrovirology* **9,** 83.

Wilk, T., Gross, I., Gowen, B. E., Rutten, T., de Haas, F., Welker, R., Krausslich, H. G., Boulanger, P., and Fuller, S. D. (2001). Organization of immature human immunodeficiency virus type 1. *J Virol* **75**(2)**,** 759-71.

Willey, R. L., Maldarelli, F., Martin, M. A., and Strebel, K. (1992). Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J Virol* **66**(12)**,** 7193-200.

Wills, J. W., Cameron, C. E., Wilson, C. B., Xiang, Y., Bennett, R. P., and Leis, J. (1994). An assembly domain of the Rous sarcoma virus Gag protein required late in budding. *J Virol* **68**(10)**,** 6605-18.

Wills, J. W., Craven, R. C., Weldon, R. A., Jr., Nelle, T. D., and Erdie, C. R. (1991). Suppression of retroviral MA deletions by the amino-terminal membrane-binding domain of p60src. *J Virol* **65**(7)**,** 3804-12.

Wilson, S. J., Webb, B. L., Ylinen, L. M., Verschoor, E., Heeney, J. L., and Towers, G. J. (2008). Independent evolution of an antiviral TRIMCyp in rhesus macaques. *Proc Natl Acad Sci U S A* **105**(9)**,** 3557-62.

Wright, E. R., Schooler, J. B., Ding, H. J., Kieffer, C., Fillmore, C., Sundquist, W. I., and Jensen, G. J. (2007). Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO Journal* **26**(8)**,** 2218-26.

Wu, X., Anderson, J. L., Campbell, E. M., Joseph, A. M., and Hope, T. J. (2006). Proteasome inhibitors uncouple rhesus TRIM5alpha restriction of HIV-1 reverse transcription and infection. *Proc Natl Acad Sci U S A* **103**(19)**,** 7465-70.

Wu, X., Li, Y., Crise, B., and Burgess, S. M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**(5626)**,** 1749-51.

Xiang, Y., Cameron, C. E., Wills, J. W., and Leis, J. (1996). Fine mapping and characterization of the Rous sarcoma virus Pr76gag late assembly domain. *J Virol* **70**(8)**,** 5695-700.

Yamashita, M., and Emerman, M. (2004). Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. *Journal of Virology* **78,** 5670-5678.

Yamashita, M., and Emerman, M. (2005). The cell cycle independence of HIV infections is not determined by known karyophilic viral elements. *PLoS Pathogens* **1**(3)**,** e18.

Yamashita, M., Perez, O., Hope, T. J., and Emerman, M. (2007). Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells. *PLoS Pathog* **3**(10)**,** 1502-10.

Yang, Y., Fricke, T., and Diaz-Griffero, F. (2013). Inhibition of reverse transcriptase activity increases stability of the HIV-1 core. *J Virol* **87**(1)**,** 683-7.

Yap, M. W., Mortuza, G. B., Taylor, I. A., and Stoye, J. P. (2007). The design of artificial retroviral restriction factors. *Virology* **365**(2)**,** 302-14.

Yap, M. W., Nisole, S., and Stoye, J. P. (2005). A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Current Biology* **15**(1)**,** 73-8.

Yap, M. W., and Stoye, J. P. (2003). Intracellular localisation of Fv1. *Virology* **307**(1)**,** 76-89.

Yasuda, J., and Hunter, E. (1998). A proline-rich motif (PPPY) in the Gag polyprotein of Mason-Pfizer monkey virus plays a maturation-independent role in virion release. *J Virol* **72**(5)**,** 4095-103.

Yeager, M., Wilson-Kubalek, E. M., Weiner, S. G., Brown, P. O., and Rein, A. (1998). Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. *Proc Natl Acad Sci U S A* **95**(13)**,** 7299-304.

Yoshinaka, Y., Katoh, I., Copeland, T. D., and Oroszlan, S. (1985a). Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *Proc Natl Acad Sci U S A* **82**(6)**,** 1618-22.

Yoshinaka, Y., Katoh, I., Copeland, T. D., and Oroszlan, S. (1985b). Translational readthrough of an amber termination codon during synthesis of feline leukemia virus protease. *J Virol* **55**(3)**,** 870-3.

Yoshinaka, Y., and Luftig, R. B. (1977). Properties of a P70 proteolytic factor of murine leukemia viruses. *Cell* **12**(3)**,** 709-19.

Yuan, B., Campbell, S., Bacharach, E., Rein, A., and Goff, S. P. (2000). Infectivity of Moloney Murine Leukemia Virus Defective in Late Assembly Events Is Restored by Late Assembly Domains of Other Retroviruses. *Journal of Virology* **74,** 7250-7260.

Yuan, B., Fassati, A., Yueh, A., and Goff, S. P. (2002). Characterization of Moloney Murine Leukemia Virus p12 Mutants Blocked during Early Events of Infection. *Journal of Virology* **76,** 10801-10810.

Yuan, B., Li, X., and Goff, S. P. (1999). Mutations altering the moloney murine leukemia virus p12 Gag protein affect virion production and early events of the virus life cycle. *The EMBO Journal* **18,** 4700-10.

Yueh, A., and Goff, S. P. (2003). Phosphorylated Serine Residues and an Arginine-Rich Domain of the Moloney Murine Leukemia Virus p12 Protein Are Required for Early Events of Viral Infection. *Journal of Virology* **77,** 1820-1829.

Zabransky, A., Hoboth, P., Hadravova, R., Stokrova, J., Sakalian, M., and Pichova, I. (2010). The noncanonical Gag domains p8 and n are critical for assembly and release of mouse mammary tumor virus. *J Virol* **84**(21)**,** 11555-9.

Zabransky, A., Sakalian, M., and Pichova, I. (2005). Localization of self-interacting domains within betaretrovirus Gag polyproteins. *Virology* **332**(2)**,** 659-66.

Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., and Charneau, P. (2000). HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101**(2)**,** 173-85.

Zhadina, M., McClure, M. O., Johnson, M. C., and Bieniasz, P. D. (2007). Ubiquitin-dependent virus particle budding without viral protein ubiquitination. *Proc Natl Acad Sci U S A* **104**(50)**,** 20031-6.

Zhang, F., Zang, T., Wilson, S. J., Johnson, M. C., and Bieniasz, P. D. (2011). Clathrin facilitates the morphogenesis of retrovirus particles. *PLoS Pathog* **7**(6)**,** e1002119.

Zhang, Y., Qian, H., Love, Z., and Barklis, E. (1998). Analysis of the assembly function of the human immunodeficiency virus type 1 gag protein nucleocapsid domain. *J Virol* **72**(3)**,** 1782-9.

Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A. M., Schulten, K., Aiken, C., and Zhang, P. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**(7451)**,** 643-6.

Zhou, L., Sokolskaja, E., Jolly, C., James, W., Cowley, S. A., and Fassati, A. (2011). Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS Pathog* **7**(8)**,** e1002194.

Zhou, W., Parent, L. J., Wills, J. W., and Resh, M. D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J Virol* **68**(4)**,** 2556-69.

Appendix A: Mutagenesis Primers

Plasmid name	Plasmid details	Mutant name/ description	Mutation in p12 (unless otherwise stated)	Forward primer (5'-3')	Reverse primer (5'-3')
pCI G3-N and pCI G3-B	N-MLV and B-MLV Gag-Pol mammalian expression constructs	5	RPSKPQ(10-15)AAAAAA	gctcttaccccct ctataaaacccg cagctgctgcag ctgcggttctctcc gataatggcg	cgccattatcgg agagaaccgca gctgcagcagct gcgggttttatag aggggtaaga gc
		6	VLSDN(16-20)AAAAA	gaccttctaaacc tcaggctgccgc cgctgctggcgg acctctcattgac	gtcaatgagag gtccgccagca gcggcggcagc ctgaggtttaga aggtc
		7	GGPLI(21-25)AAAAA	acctcaggttctct ccgataatgccg cagctgccgctg accttctctcaga agaccct	agggtcttctgag agaaggtcagc ggcagctgcgg cattatcggaga gaacctgaggt
		8	DLLSED(26-31)AAAAAA	ggcggacctctc attgccgctgccg cagcagcccctc cgccgtacggag	ctccgtacggcg gaggggctgctg cggcagcggca atgagaggtccg cc
		13	PSPIV(61-65)AAAAA	gagattcctgccg ccgctgccgcag cgtctcgcctgcg gg	cccgcaggcga gacgctgcggc agcggcggcag gaatctc
		14	SRLRG(66-70)AAAAA	cctgccccctctc ccatagtggctgc cgcggcggcca aaagagacccg cccgcg	cgcgggcgggt ctcttttggccgc cgcggcagcca ctatgggagag ggggcagg
		15	KRDPP(71-75)AAAAA	cgcctgcggggc gcagcagccgc cgccgcggcag attcc	ggaatctgccgc ggcggcggctg ctgcgccccgca ggcg
pHG1	XMRV Gag-Pol mammalian expression construct	5	KPPKPQ(10-15)AAAAAA	cttacccccctctat aaagtccgcagc tgctgcggccgc ggttctccctgata gcggcgga	tccgccgctatc agggagaaccg cggccgcagca gctgcggacttta tagaggggtaag
		6	VLPDS(16-20)AAAAA	cctcctaagccc caggctgccgct gctgccggcgga cctctcat	atgagaggtccg ccggcagcagc ggcagcctggg gcttaggagg
		7	GGPLI(21-25)AAAAA	cccaggttctccc tgatagcgccgc agctgccgctga ccttctcacagag gatcc	ggatcctctgtga gaaggtcagcg gcagctgcggc gctatcagggag aacctggg
		13	PSPMV(60-64)AAAAA	tccgaggtttccc ccgctgctgccg cggcgtctcgact gcgggga	tccccgcagtcg agacgccgcgg cagcagcgggg gaaacctcgga
		14	SRLRG(65-69)AAAAA	tccccccttctcc catggtggctgc agcggcggcaa ggagagaccctc ccgcagcg	cgctgcgggag ggtctctccttgc cgccgctgcag ccaccatggga gaaggggggg a
		15	RRDPP(70-74)AAAAA	tctcgactgcggg gagcggcagcc gctgccgcagcg gactccacc	ggtggagtccgc tgcggcagcgg ctgccgctcccc gcagtcgaga
pFeLVdw	FeLV Gag-Pol mammalian expression construct	5	DPPKPP(7-12)AAAAAA	cgttctccccaag ccagccgccgc cgcagcggctgt gttaccgcctgat	atcaggcggtaa cacagccgctg cggcggcggct ggcttggggag

					aacg
		6	VLPPDP(13-18)AAAAAA	cagacccccccaaaccgcctgcggcagcggctgctgcttcctcccctttaattgatc	gatcaattaaaggggaggaagcagcagccgctgccgcaggcggttgggagggtctg
		7	SSPLI(19-23)AAAAA	aaccgcctgtgttaccgcctgatcctgctgccgctgcagctgatctcttaacagaagagccac	gtggctcttctgttaagagatcagctgcagcggcagcaggatcaggcggtaacacaggcggtt
		8	DLLTEE(24-29)AAAAAA	accgcctgatcctcttcccctttaattgctgccgcagcagcagcgccacctccctatccgggg	ccccggatagggaggtggcgctgctgctgcggcagcaattaaaggggaagaaggatcaggcggt
		13	SPI(51-53)AAA	aggaccccaaccgctgccgcggctgcaagccggcta	tagccggcttgcagccgcggcagcggttggggtcct
		14	SRLRE(55-59)AAAAA	cccaaccgcttccccgattgccgccgcggcagcggcacgacgagaaaaccctgctgaa	ttcagcagggttttctcgtcgtgccgctgccgcggcggcaatcgggagagcggttggg
		15	RRENP(60-64)AAAAA	ccccgattgccagccggctaaggaagcagcagcagccgctgctgaagaatctcaagcc	ggcttgagattcttcagcagcggctgctgctgcttcccttagccggctggcaatcgggg
pcGaLVgp	GaLV Gag-Pol mammalian expression construct	N	TDDLL(3-7)AAAAA	ccccccatctacccggcagcagccgccgcagccctcctctctgaacccacg	cgtgggttcagagaggagggctgcggcggctgctgccgggtagatgggggg
		8	LLSEPT(8-13)AAAAAA	gcaacagacgacttactcgccgccgctgcagccgcgcccccgcctatccggcggca	tgccgccggatagggcgggggcgcggctgcagcggcgagtaagtcgtctgttgc
		13	EGP(44-46)AAA	agcgatcctgcggcggcagccgctgggaccaggagt	actcctggtcccagcggctgccgccgcaggatcgct
		14	GTRSR(49-53)AAAAA	cctgaggggccagccgctgcggccgcggctgcccgtgcccgcagtccagca	tgctggactgcggggcacgggcagccgcggccgcagcggctggcccctcagg
		15	RARSP(54-58)AAAAA	gctgggaccaggagtcgcgctgccgccgctgcagcagacaactcgggt	acccgagttgtctgctgcagcggcggcagcgcgactcctggtcccagc
pKB4	Mo-MLV Gag-Pol mammalian expression construct	K10A	K10A	cactccttctctaggcgccgcaccaaacctcaagttctt	aagaacttgaggtttaggtgcggcgcctagagaaggagtg
		P11A	P11A	cttctctaggcgccaaagctaaacctcaagttcttt	aaagaacttgaggtttagctttggcgcctagagaag
		K12A	K12A	tccttctctaggcgccaaacctgcactcaagttctttct	agaaagaacttgaggtgcaggttggcgcctagagaagga
		P13A	P13A	tctaggcgccaaacctaaagctcaagttctttctg	cagaaagaacttgagctttaggtttggcgcctaga
		Q14A	Q14A	ctaggcgccaaacctaaacctgcagttctttctgacagtgg	ccactgtcagaaagaactgcaggttttaggtttggcgcctag
		KKQ(10,12,14)	KKQ(10,12,14)AAA	ctcactccttctct	cccccactgtca

		AAA		aggcgccgcac ctgcacctgcagt tctttctgacagtg gggg	gaaagaactgc aggtgcaggtgc ggcgcctagag aaggagtgag
		V15A	V15A	caaacctaaacc tcaagctctttctg acagtgggg	ccccactgtcag aaagagcttga ggtttaggtttg
		L16A	L16A	gccaaacctaaa cctcaagttgcttc tgacagtgggggc gc	gccccccactgt cagaagcaactt gaggtttaggttt ggc
		S17A	S17A	caaacctaaacc tcaagttcttgctg acagtgggg	ccccactgtcag caagaacttgag gtttaggtttg
		D18A	D18A	cctaaacctcaa gttctttctgctagt gggggccgctc atcgac	gtcgatgagcgg ccccccactagc agaaagaactt gaggtttagg
		S19A	S19A	ctcaagttctttctg acgctgggggg ccgctc	gagcggccccc cagcgtcagaa agaacttgag
		G20A	G20A	ctttctgacagtgc ggggccgctcat c	gatgagcggcc ccgcactgtcag aaag
		G21A	G21A	ctgacagtgggg cgccgctcatcg a	tcgatgagcggc gccccactgtca g
		P22A	P22A	gacagtggggg agcgctcatcga cc	ggtcgatgagcg ctcccccactgtc
		L23A	L23A	gacagtggggg gccggccatcga cctacttac	gtaagtaggtcg atggccggccc cccactgtc
		I24A	I24A	gtgggggccg ctcgccgaccta cttacaga	tctgtaagtaggt cggcgagcggc cccccac
		D25A	D25A	ggggccgctcat cgccctacttaca gaag	cttctgtaagtag ggcgatgagcg gcccc
		L26A	L26A	gggccgctcatc gacgcacttaca gaagaccc	gccgctcatcga cctagctacaga agacccccc
		L27A	L27A	gccgctcatcga cctagctacaga agacccccc	gggggtcttctg tagctaggtcgat gagcggc
		T28A	T28A	gccgctcatcga cctacttgcagaa gacccc	ggggtcttctgca agtaggtcgatg agcggc
		E29A	E29A	catcgacctactt acagcagaccc cccgccttatagg g	ccctataaggcg gggggtctgctgt aagtaggtcgat g
		D30A	D30A	atcgacctactta cagaagcccca ccgccttatagggg acc	ggtccctataag gcggtggggctt ctgtaagtaggtc gat
		P60A	P60A	gaggcaccgga cgcctccccaat gg	ccattggggagg cgtccggtgcct c
		S61A	S61A	gcaccggaccc cgccccaatggc at	atgccattgggg cggggtccggtg c
		P62A	P62A	ccggacccctcc gcaatggcatctc	gagatgccattg cggaggggtccgg
		M63A	M63A	ccggacccctcc ccagcggcatct cgcc	ggcgagatgcc gctggggaggg gtccgg
		S65A	S65A	ctccccaatggc agctcgcctacgt gg	ccacgtaggcg agctgccattgg ggag
		R66A	R66A	ccctccccaatg gcatctgccctac gtggg	cccacgtaggg cagatgccattg gggaggg
		L67A	L67A	caatggcatctcg cgcacgtggga gacggg	cccgtctcccac gtgcgcgagatg ccattg

		R68A	R68A	cccaatggcatct cgcctagctggg agacggg	cccgtctcccag ctaggcgagatg ccattggg
		G69A	G69A	ctcgcctacgtgc gagacgggagc c	ggctcccgtctc gcacgtaggcg ag
		R70A	R70A	cgcctacgtggg gcacgggagcc ccc	gggggctcccgt gccccacgtag gcg
		R71A	R71A	ctcgcctacgtgg gagagcggagc cccc	gggggctccgct ctcccacgtagg cgag
		E72A	E72A	tgggagacggg cgccccctgtgg	ccacaggggc gcccgtctccca
		P73A	P73A	ggagacgggag gcccctgtggcc	ggccacagggg cctcccgtctcc
		P74A	P74A	gacgggagccc gctgtggccgac	gtcggccacag cgggctcccgtc
		p12-TC	DGNGGE(45-50)CCPGCC	ccacccccttccg acaggtgctgcc ctggttgctgcgc gaccctgcggg agag	ctctcccgcagg ggtcgcgcagc aaccagggcag cacctgtcggaa gggggtgg
		p12-*AscI*	GGE(48-50)GAP	cgacagggacg gaaatggcgcg ccagcgacccct gcgggag	ctcccgcaggg gtcgctggcgcg ccatttccgtccct gtcg
		L16D	L16D	gccaaacctaaa cctcaagttgattc tgacagtgggggc	gccccccactgt cagaatcaactt gaggtttaggttt ggc
		L16F	L16F	aacctaaacctc aagttttttctgac agtgggggg	ccccccactgtc agaaaaaactt gaggtttaggtt
		L16K	L16K	gccaaacctaaa cctcaagttaagt ctgacagtgggg ggccg	cggcccccact gtcagacttaact tgaggtttaggttt ggc
		L16Q	L16Q	acctaaacctca agttcagtctgac agtgggggcc	ggcccccactg tcagactgaactt gaggtttaggt
		L16V	L16V	aacctaaacctc aagttgtttctgac agtgggggg	ccccccactgtc agaaacaacttg aggtttaggtt
		S17D	S17D	acctaaacctca agttcttgatgac agtgggggccg	cggcccccact gtcatcaagaac ttgaggtttaggt
		S17K	S17K	gcgccaaaccta aacctcaagttctt aaggacagtgg ggggcc	ggcccccactg tccttaagaactt gaggtttaggttt ggcgc
		S17Q	S17Q	gcgccaaaccta aacctcaagttctt caggacagtgg ggggcc	ggcccccactg tcctgaagaactt gaggtttaggttt ggcgc
		D18E	D18E	taaacctcaagtt ctttctgagagtg gggggcc	ggcccccactc tcagaaagaact tgaggttta
		S71T	S71T	gcaccggaccc caccccaatggc at	atgccattggggt ggggtccggtgc
		R66H	R66H	cccaatggcatct cacctacgtggg agac	gtctcccacgta ggtgagatgcca ttggg
		R70H	R70H	tctcgcctacgtg ggcatcgggag ccccctgtg	cacaggggct cccgatgcccac gtaggcgaga
		S17 F/V	S17 F/V	aaacctaaacct caagttcttnnkg acagtgggggg ccgctc	gagcggccccc cactgtcmnna agaacttgaggt ttaggttt
		D18 F/K/Q/L	D18 F/K/Q/L	gtcgatgagcgg ccccccactmn nagaaagaactt gaggtttagg	cctaaacctcaa gttctttctnnkag tgggggccgct catcgac

		Del Sm deletion p12	D43-P53 del	ccacccccttccg cgggagaggca	tgcctctcccgcg gaaggggtgg
		Del Lrg deletion p12	P37-A57 del	gccttataggga cccggacccctc cc	gggaggggtcc gggtccctataa ggc
		Changing tropism of Mo-MLV CA to N-tropic like	CA D82N	gcggtgcgggg caatgatgggcg cccc	ggggcgcccat cattgccccgca ccgc
			CA A110R	gggattacacca cccagcgaggta ggaaccacctag t	actaggtggttcc tacctcgctgggt ggtgtaatccc
			CA H117L	ggtaggaacca cctagtcttgtatc gccagttgctcct agc	gctaggagcaa ctggcgatacaa gactaggtggttc ctacc
		Insert L between p12 A2 and T3 of Mo-MLV/Ga-p12 or Ga-Np12 chimeras	PA>L<T	cctccctttatccg gcactcacagac gacttactc	gagtaagtcgtct gtgagtgccgga taaagggagg
		MA-p12 cleavage	MA Y131D	ccgcctcgatcct cccttgatccagc cct	agggctggatca agggaggatcg aggcgg
pKB4 p12AscI	Mo-MLV Gag-Pol mammalian expression construct with p12-AscI	AscI insert oligo	alanine insert (smaller)	gcgccgggcgc gccagcggcgg cggcggcagcg gcggcagctgcg gccgcagctgca gcgggcgcgcc gcgccg	cggcgcggcgc gcccgctgcag ctgcggccgca gctgccgccgct gccgccgccgc cgctggcgcgc ccggcgc
		AscI insert oligo	Angiogenin NoLS insert	gcgccgggcgc gccaatcatgag gaggaggggcc ttgcgggcgcgc cgcgccg	cggcgcggcgc gcccgcaaggc ccctcctcctcat gattggcgcgcc cggcgc
		AscI insert oligo	sGFP11 inset	gccgggcgcgc cacgtgaccaca tggtccttcatga gtacgtaaatgct gctgggattaca gcgggcgcgcc gcg	cgcggcgcgcc cgctgtaatccc agcagcatttac gtactcatgaag gaccatgtggtc acgtggcgcgc ccggc
		AscI insert oligo	NIK NoLS	gcgccgggcgc gccaaggaaga agaggaaaaag aaggcgggcgc gccgcgccg	cggcgcggcgc gcccgccttctttt tcctcttcttccttg gcgcgcccggc gc
		AscI insert oligo	SV40 T-antigen NLS	gcgccgggcgc gccaaagaaga aaaggagggtc gcgggcgcgcc gcgccg	cggcgcggcgc gcccgcgaccct ccttttcttctttgg cgcgcccggcg c
		AscI insert oligo	PFV CBS	gcgccgggcgc gccaaatcaag gaggatataatct tcgaccccgtact taccaacctgcg ggcgcgccgcg ccg	cggcgcggcgc gcccgcaggttg gtaagtacggg gtcgaagattat atcctccttgattt ggcgcgcccgg cgc
		AscI insert oligo	alanine insert (large)	gcgccgggcgc gccagcggcgg cggcggcagcg gcggcagctgcg gccgcagctgcg gcggcggcggc agcggcggcag ctgcggccgcag ctgcggccgcag cgggcgcgccg cgccg	cggcgcggcgc gcccgctgcgg ccgcagctgcg gccgcagctgc cgccgctgccgc cgccgccgcag ctgcggccgca gctgccgccgct gccgccgccgc cgctggcgcgc ccggcgc
pKB4 MA Y131D	Mo-MLV Gag-Pol mammalian expression construct	MA-p12 cleavage	MA Y131D p12 PA(1,2)DE	ccgcctcgatcct cccttgatgatga gctcactccttctc taggc	gcctagagaag gagtgagctcat catcaagggag gatcgaggcgg

pKB5 myc	pKB p12-myc plasmids	Mut 5 p12-myc	KPKPQ(10-14)AAAAA	ccctcactccttct ctaggcgccgca gctgcagctgca gttctttctgacag tgggggc	gccccccactgt cagaaagaact gcagctgcagct gcggcgcctag agaaggagtga ggg
pKB6 myc		Mut 6 p12- Myc and Mut 6 eGFP-p12	VLSDS(15-19)AAAA	gccaaacctaaa cctcaagctgctg ctgccgctgggg ggccgctcatcg a	tcgatgagcggc cccccagcggc agcagcagcttg aggtttaggtttg gc
pKB7 myc		Mut 7 p12- Myc and Mut 7 eGFP-p12	GGPLI(20-24)AAAAA	cctcaagttctttct gacagtgcggc ggcggccgccg acctacttacaga agacccc	ggggtcttctgta agtaggtcggcg gccgccgccgc actgtcagaaag aacttgagg
pKB8 myc		Mut 8 p12- Myc and Mut 8 eGFP-p12	DLLTED(25-30)AAAAAA	ggggggccgctc atcgccgcagct gcagcagcacc cccgccttatagg gacccaaga	tcttgggtccctat aaggcgggggt gctgctgcagct gcggcgatgag cggcccccc
pcDNACherry-p12	Cherry p12 expression plasmids	*EcoRI* removal in p12	A51F	agggacggaaa tggtggagaattc acccctgcggg	cccgcaggggt gaattctccacc atttccgtccct
pcDNACherry-p12		Stop codon in p12	R35stop	gaagacccccc gccttattgagac ccaagaccac	gtggtcttgggtct caataaggcgg ggggtcttc
pcGaLVgp	GaLV Gag-Pol mammalian expression construct	T3A	T3A	cccatctacccg gcagcagacga cttactcc	ggagtaagtcgt ctgctgccgggt agatggg
		D4A	D4A	ctacccggcaac agccgacttactc ctcc	ggaggagtaag tcggctgttgccg ggtag
		D5A	D5A	cccggcaacag acgccttactcct cctct	agaggaggagt aaggcgtctgttg ccggg
		L6A	L6A	cccggcaacag acgacgcactcc tcctctctgaa	ttcagagagga ggagtgcgtcgt ctgttgccggg
		L7A	L7A	acccggcaaca gacgacttagcc ctcctctctg	cagagaggag ggctaagtcgtct gttgccgggt
		L8A	L8A	gcaacagacga cttactcgccctct ctgaacccacgc cc	gggcgtgggttc agagagggcg agtaagtcgtctg ttgc
		L9A	L9A	gacgacttactcc tcgcctctgaacc cacgccc	gggcgtgggttc agaggcgagg agtaagtcgtc
		E44A	E44A	gatagtagcgat cctgcggggcca gccgctggg	cccagcggctg gccccgcagga tcgctactatc
		G45A	G45A	agtagcgatcct gaggcgccagc cgctgggacc	ggtcccagcgg ctggcgcctcag gatcgctact
		P46A	P46A	agcgatcctgag ggggcagccgct gggaccagg	cctggtcccagc ggctgccccctc aggatcgct
		G49A	G49A	gaggggccagc cgctgcgaccag gagtcgccgt	acggcgactcct ggtcgcagcgg ctggcccctc
		T50A	T50A	gggccagccgct ggggccaggag tcgccgtgcc	ggcacggcgac tcctggccccag cggctggccc
		R51A	R51A	ccagccgctggg accgcgagtcgc cgtgcccgc	gcgggcacggc gactcgcggtcc cagcggctgg
		S52A	S52A	gccgctgggacc agggctcgccgt gcccgcagt	actgcgggcac ggcgagccctg gtcccagcggc
		R53A	R53A	gctgggaccag gagtgcccgtgc ccgcagtcca	tggactgcgggc acgggcactcct ggtcccagc
		R54A	R54A	gggaccaggag tcgcgctgcccg cagtccagca	tgctggactgcg ggcagcgcgac tcctggtccc

		R56A	R56A	aggagtcgccgtgccgccagtccagcagacaac	gttgtctgctggactggcggcacggcgactcct
		S57A	S57A	agtcgccgtgccgcgctccagcagacaactcg	cgagttgtctgctggagcgcgggcacggcgact
		P58A	P58A	cgccgtgcccgcagtgcagcagacaactcgggt	acccgagttgtctgctgcactgcgggcacggcg
		Introduce *Bsr*GI site 5' to GaLV gag		ccccacccctcactgtacatctccccgcaggagaattc	gaattctcctgcggggagatgtacagtgaggggtgggg
		Remove *Xho*I site 3' of pol		gcaaaataagcggccgctaaagtctagagggcccgt	acgggccctctagactttagcggccgcttattttgc
pczGaLV/Mo-p12 chimeras	GaLV Gag-Pol chimera mammalian expression construct	Remove introduced *Bsr*GI site 5' of GaLV Gag		ccccacccctcactctgcttctccccgcaggagaattc	gaattctcctgcggggagaagcagagtgaggggtgggg

Appendix B: Plasmid List

Plasmid name	Description	Synthesis
pKB4	Mo-MLV Gag-Pol expression vector	Mo-MLV Gag-Pol from pMD-MLV cloned into pcDNA3.1
pKB5	Mo-MLV Gag-Pol expression vector with p12 mutant 5	*BsrGI/XhoI* fragment from pNCS PM5 cloned in pKB4
pKB6	Mo-MLV Gag-Pol expression vector with p12 mutant 6	*BsrGI/XhoI* fragment from pNCS PM6 cloned in pKB4
pKB7	Mo-MLV Gag-Pol expression vector with p12 mutant 7	*BsrGI/XhoI* fragment from pNCS PM7 cloned in pKB4
pKB8	Mo-MLV Gag-Pol expression vector with p12 mutant 8	*BsrGI/XhoI* fragment from pNCS PM8 cloned in pKB4
pKB13	Mo-MLV Gag-Pol expression vector with p12 mutant 13	*BsrGI/XhoI* fragment from pNCS PM13 cloned in pKB4
pKB14	Mo-MLV Gag-Pol expression vector with p12 mutant 14	*BsrGI/XhoI* fragment from pNCS PM14 cloned in pKB4
pKB15	Mo-MLV Gag-Pol expression vector with p12 mutant 15	*BsrGI/XhoI* fragment from pNCS PM15 cloned in pKB4
pKB16	Mo-MLV Gag-Pol expression vector with FLAG-tagged wild type p12	*BsrGI/XhoI* fragment from pNCS FLAG cloned in pKB4
pKB17	Mo-MLV Gag-Pol expression vector with p12 late domain mutant	*BsrGI/XhoI* fragment from pNCS L-domain cloned in pKB4
pKB4 K10A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P11A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 K12A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P13A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 Q14A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 KKQ(10,12,14)AAA	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 V15A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L16A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S17A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4

pKB4 D18A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S19A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 G20A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 G21A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P22A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L23A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 I24A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D25A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L26A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L27A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 T28A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 E29A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D30A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P60A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S61A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P62A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 M63A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S65A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 R66A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L67A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4

pKB4 R68A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 G69A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 R70A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 R71A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 E72A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P73A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 P74A	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 LR(16,66)AA)	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4 L16A
pKB4 L16D	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L16F	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L16K	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L16Q	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 L16V	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S17D	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S17K	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S17Q	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S17V	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 S61T	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 R66H	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 R70H	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4

pKB4 S17F	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D18F	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D18K	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D18Q	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D18L	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 D18E	Mo-MLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pKB4
pKB4 p12-TC	Mo-MLV Gag-Pol expression vector with tetracysteine motif in p12	Site directed mutagenesis of pKB4
pKB4 p12-AscI	Mo-MLV Gag-Pol expression vector with AscI site in p12	Site directed mutagenesis of pKB4
pKB4 p12 DelSm	Mo-MLV Gag-Pol expression vector with D43-P53 p12 deletion	Site directed mutagenesis of pKB4
pKB4 p12 DelLrg	Mo-MLV Gag-Pol expression vector with P37-A57 p12 deletion	Site directed mutagenesis of pKB4
pKB4 MA-D	Mo-MLV Gag-Pol expression vector with MA Y131D mutation	Site directed mutagenesis of pKB4
pKB4 DDE	Mo-MLV Gag-Pol expression vector with MA Y131D + p12 PA(1,2)DE mutations	Site directed mutagenesis of pKB4 MA-D
pKB4 sGFPp12	Mo-MLV Gag-Pol expression vector with sGFP11 motif in p12	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB4 p12-SV40	Mo-MLV Gag-Pol expression vector with SV40 T-antigen NLS in p12	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB4 p12Ang NoLS	Mo-MLV Gag-Pol expression vector with angiogenin NoLS motif in p12	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB6 p12Ang NoLS	Mo-MLV Gag-Pol expression vector with angiogenin NoLS motif in p12 mutant 6	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB7 p12Ang NoLS	Mo-MLV Gag-Pol expression vector with angiogenin NoLS motif in p12 mutant 7	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB14 p12Ang NoLS	Mo-MLV Gag-Pol expression vector with angiogenin NoLS motif in p12 mutant 14	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB4 p12NIK NoLS	Mo-MLV Gag-Pol expression vector with NIK NoLS motif in p12	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB6 p12NIK NoLS	Mo-MLV Gag-Pol expression vector with NIK NoLS motif in p12 mutant 6	AscI flanked oligos digested and cloned into pKB4 p12-AscI

pKB7 p12NIK NoLS	Mo-MLV Gag-Pol expression vector with NIK NoLS motif in p12 mutant 7	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB14 p12NIK NoLS	Mo-MLV Gag-Pol expression vector with NIK NoLS motif in p12 mutant 14	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB4 p12CBS	Mo-MLV Gag-Pol expression vector with PFV CBS motif in p12	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB6 p12CBS	Mo-MLV Gag-Pol expression vector with PFV CBS motif in p12 mutant 6	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB7 p12CBS	Mo-MLV Gag-Pol expression vector with PFV CBS motif in p12 mutant 7	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB13 p12CBS	Mo-MLV Gag-Pol expression vector with PFV CBS motif in p12 mutant 13	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB14 p12CBS	Mo-MLV Gag-Pol expression vector with PFV CBS motif in p12 mutant 14	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB15 p12CBS	Mo-MLV Gag-Pol expression vector with PFV CBS motif in p12 mutant 15	AscI flanked oligos digested and cloned into pKB4 p12-AscI
pKB4 p12mycE	Mo-MLV Gag-Pol expression vector with myc motif in p12	BsrGI/XhoI fragment cloned from pNCS p12-1xMycR (gift from Eran Bacharach)
pKB5 p12mycE	Mo-MLV Gag-Pol expression vector with sGFP11 motif in p12 mutant 5	BsrGI/XhoI fragment cloned from pNCS p12-1xMycR PM5 (gift from Eran Bacharach)
pKB6 p12mycE	Mo-MLV Gag-Pol expression vector with sGFP11 motif in p12 mutant 6	Site directed mutagenesis of pKB4 p12mycE
pKB7 p12mycE	Mo-MLV Gag-Pol expression vector with sGFP11 motif in p12 mutant 7	Site directed mutagenesis of pKB4 p12mycE
pKB8 p12mycE	Mo-MLV Gag-Pol expression vector with sGFP11 motif in p12 mutant 8	Site directed mutagenesis of pKB4 p12mycE
pKB14 p12mycE	Mo-MLV Gag-Pol expression vector with sGFP11 motif in p12 mutant 14	BsrGI/XhoI fragment cloned from pNCS p12-1xMycR PM14 (gift from Eran Bacharach)
pMo/N	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL	Site directed mutagenesis of pKB4

pMo/N Mut 5	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant 5	Site directed mutagenesis of pKB5
pMo/N Mut 6	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant 6	Site directed mutagenesis of pKB6
pMo/N Mut 7	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant 7	Site directed mutagenesis of pKB7
pMo/N Mut 8	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant 8	Site directed mutagenesis of pKB8
pMo/N Mut 13	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant 13	Site directed mutagenesis of pKB13
pMo/N Mut 15	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant 15	Site directed mutagenesis of pKB14
pMo/N L16A	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant L16A	Site directed mutagenesis of L16A
pMo/N R66A	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant R66A	Site directed mutagenesis of R66A
pMo/N LR(16,66)AA	Mo-MLV Gag-Pol expression vector with CA mutations DAH(82,110,117)NRL and p12 mutant LR(16,66)AA	Site directed mutagenesis of LR(16,66)AA
pcDNA3.1(A)p12eGFP	Mo-MLV modified Gag expression vector containing eGFP-p12	Gift from Eran Bacharach
pcDNA3.1(A)p12eGFP PM5	Mo-MLV modified Gag expression vector containing eGFP-p12 mutant 5	Gift from Eran Bacharach
pcDNA3.1(A)p12eGFP PM6	Mo-MLV modified Gag expression vector containing eGFP-p12 mutant 6	Site directed mutagenesis
pcDNA3.1(A)p12eGFP PM7	Mo-MLV modified Gag expression vector containing eGFP-p12 mutant 7	Site directed mutagenesis
pcDNA3.1(A)p12eGFP PM8	Mo-MLV modified Gag expression vector containing eGFP-p12 mutant 8	Site directed mutagenesis
pcDNA3.1(A)p12eGFP PM14	Mo-MLV modified Gag expression vector containing eGFP-p12 mutant 14	Gift from Eran Bacharach
pCIG3N	N-tropic MLV Gag-Pol expression vector	Gift from Jonathan Stoye
pCIG3N Mut 5	N-tropic MLV Gag-Pol expression vector with p12 mutant 5	Site directed mutagenesis of pCIG3N

pCIG3N Mut 6	N-tropic MLV Gag-Pol expression vector with p12 mutant 6	Site directed mutagenesis of pCIG3N
pCIG3N Mut 7	N-tropic MLV Gag-Pol expression vector with p12 mutant 7	Site directed mutagenesis of pCIG3N
pCIG3N Mut 8	N-tropic MLV Gag-Pol expression vector with p12 mutant 8	Site directed mutagenesis of pCIG3N
pCIG3N Mut 13	N-tropic MLV Gag-Pol expression vector with p12 mutant 13	Site directed mutagenesis of pCIG3N
pCIG3N Mut 14	N-tropic MLV Gag-Pol expression vector with p12 mutant 14	Site directed mutagenesis of pCIG3N
pCIG3N Mut 15	N-tropic MLV Gag-Pol expression vector with p12 mutant 15	Site directed mutagenesis of pCIG3N
pCIG3B	B-tropic MLV Gag-Pol expression vector	Gift from Jonathan Stoye
pCIG3B Mut 5	B-tropic MLV Gag-Pol expression vector with p12 mutant 5	Site directed mutagenesis of pCIG3B
pCIG3B Mut 6	B-tropic MLV Gag-Pol expression vector with p12 mutant 6	Site directed mutagenesis of pCIG3B
pCIG3B Mut 7	B-tropic MLV Gag-Pol expression vector with p12 mutant 7	Site directed mutagenesis of pCIG3B
pCIG3B Mut 8	B-tropic MLV Gag-Pol expression vector with p12 mutant 8	Site directed mutagenesis of pCIG3B
pCIG3B Mut 13	B-tropic MLV Gag-Pol expression vector with p12 mutant 13	Site directed mutagenesis of pCIG3B
pCIG3B Mut 14	B-tropic MLV Gag-Pol expression vector with p12 mutant 14	Site directed mutagenesis of pCIG3B
pCIG3B Mut 15	B-tropic MLV Gag-Pol expression vector with p12 mutant 15	Site directed mutagenesis of pCIG3B
pB30Y	B-tropic MLV modified Gag expression vector CA-YFP fusion (del of last 4 a.a. of CA and all of NC)	Gift from Melvyn Yap
pcFeLVgp	FeLV Gag-Pol expression vector	Gift from Jonathan Stoye
pFeLVdw	FeLV Gag-Pol expression vector	FeLV Gag-Pol cloned from pcFeLVgp into pcDNA3.1
pFeLV Mut 5	FeLV Gag-Pol expression vector with p12 mutant 5	Site directed mutagenesis of pFeLVdw
pFeLV Mut 6	FeLV Gag-Pol expression vector with p12 mutant 6	Site directed mutagenesis of pFeLVdw
pFeLV Mut 7	FeLV Gag-Pol expression vector with p12 mutant 7	Site directed mutagenesis of pFeLVdw

pFeLV Mut 8	FeLV Gag-Pol expression vector with p12 mutant 8	Site directed mutagenesis of pFeLVdw
pFeLV Mut 13	FeLV Gag-Pol expression vector with p12 mutant 13	Site directed mutagenesis of pFeLVdw
pFeLV Mut 14	FeLV Gag-Pol expression vector with p12 mutant 14	Site directed mutagenesis of pFeLVdw
pFeLV Mut 15	FeLV Gag-Pol expression vector with p12 mutant 15	Site directed mutagenesis of pFeLVdw
pcGaLVgp	GaLV Gag-Pol expression vector	Gift from Jonathan Stoye
pcGaLVgp Mut 5	GaLV Gag-Pol expression vector with p12 mutant 5	Site directed mutagenesis of pcGaLVgp
pcGaLVgp Mut 6	GaLV Gag-Pol expression vector with p12 mutant 6	Site directed mutagenesis of pcGaLVgp
pcGaLVgp Mut 7	GaLV Gag-Pol expression vector with p12 mutant 7	Site directed mutagenesis of pcGaLVgp
pcGaLVgp Mut 8	GaLV Gag-Pol expression vector with p12 mutant 8	Site directed mutagenesis of pcGaLVgp
pcGaLVgp Mut 13	GaLV Gag-Pol expression vector with p12 mutant 13	Site directed mutagenesis of pcGaLVgp
pcGaLVgp Mut 14	GaLV Gag-Pol expression vector with p12 mutant 14	Site directed mutagenesis of pcGaLVgp
pcGaLVgp Mut 15	GaLV Gag-Pol expression vector with p12 mutant 15	Site directed mutagenesis of pcGaLVgp
pcGaLVgp T3A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp D4A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp D5A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp L6A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp L7A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp L8A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp L9A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp E44A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp

pcGaLVgp G45A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp P46A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp G49A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp T50A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp R51A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp S52A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp R53A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp R54A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp R56A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp S57A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pcGaLVgp P58A	GaLV Gag-Pol expression vector with p12 mutant	Site directed mutagenesis of pcGaLVgp
pHG1-env	XMRV Gag-Pol expression vector	Gift from Harriet Groom
pHG1-env p12 mutant 5	XMRV Gag-Pol expression vector with p12 mutant 5	Site directed mutagenesis of pHG1-env
pHG1-env p12 mutant 6	XMRV Gag-Pol expression vector with p12 mutant 6	Site directed mutagenesis of pHG1-env
pHG1-env p12 mutant 7	XMRV Gag-Pol expression vector with p12 mutant 7	Site directed mutagenesis of pHG1-env
pHG1-env p12 mutant 13	XMRV Gag-Pol expression vector with p12 mutant 13	Site directed mutagenesis of pHG1-env
pHG1-env p12 mutant 14	XMRV Gag-Pol expression vector with p12 mutant 14	Site directed mutagenesis of pHG1-env
pHG1-env p12 mutant 15	XMRV Gag-Pol expression vector with p12 mutant 15	Site directed mutagenesis of pHG1-env
pczLacZ-LTR	MLV LTR-driven *LacZ* reporter expression vector	Gift from Kate Bishop
pLNCG	MLV LTR-driven GFP reporter expression vector	Gift from Jonathan Stoye

pczVSV-G	Vesicular stomatitis virus glycoprotein expression vector	Gift from Jonathan Stoye
pMoSAF	MLV ecotropic Env expression vector	Gift from Harriet Groom
pS15-mCherry	Mammalian expression vector for the first 15 amino acids of c-src fused to mCherry	Gift from Ed Campbell
pCMVTagcherryPaisy	mCherry-paisy fusion mammalian expression vector	Gift from Kate Bishop
pCMVTagcherry	mCherry-linker fusion mammalian expression vector	Gift from Kate Bishop
pCMVcherry-p12	mCherry-p12 fusion mammalian expression vector	p12 PCR amplified from pKB4 using primers oKB130-p12 for (5'-CTGAATTCccagccctcactccttctctag) and oKB131-p12 rev (5'-CTACTCGAGtcagaatgcctgcgaggtagtg). Digested with *EcoRI* and *XhoI* and ligated into pCMVcherry-paisy
pCMVcherry-Np12	mCherry-Np12 fusion mammalian expression vector	Site directed mutagenesis of pCMVcherry-p12
pCMVcherry-Cp12	mCherry-Cp12 fusion mammalian expression vector	Site directed mutagenesis of pCMVcherry-p12 to insert *EcoRI* site in p12. Digested out 5' p12 sequence with *EcoRI* digest
pKB4/Ga-p12	Mo-MLV Gag-Pol expression vector with GaLV p12 in place of MLV p12	Overlapping extension PCR
pKB4/Ga-p12-CL	Mo-MLV Gag-Pol expression vector with GaLV p12 in place of MLV p12 and insertion L between A2 and L3	Site directed mutagenesis of pKB4/Ga-p12
pKB4/Ga-Np12	Mo-MLV Gag-Pol expression vector with GaLV p12 N-terminus of in place of MLV p12 N-terminus	Overlapping extension PCR
pKB4/Ga-Np12-CL	Mo-MLV Gag-Pol expression vector with GaLV p12 N-terminus of in place of MLV p12 N-terminus and insertion L between A2 and L3	Site directed mutagenesis of pKB4/Ga-Np12
pKB4/Ga-Cp12	Mo-MLV Gag-Pol expression vector with GaLV p12 C-terminus of in place of MLV p12 N-terminus	Overlapping extension PCR
pczGaLVgp BsX	GaLV Gag-Pol expression vector with *Bsr*GI site 5' to *gag* introduced and *Xho*I 3' to *pol* removed	Site directed mutagenesis of pczGaLVgp
pczGaLV/Mo-p12	GaLV Gag-Pol expression vector with Mo-MLV p12 in place of MLV p12	Overlapping extension PCR

pczGaLV/Mo-Np12	GaLV Gag-Pol expression vector with Mo-MLV p12 N-terminus in place of GaLV p12 N-terminus	Overlapping extension PCR
pczGaLV/Mo-Cp12	GaLV Gag-Pol expression vector with Mo-MLV p12 C-terminus in place of GaLV p12 C-terminus	Overlapping extension PCR
pczGaLV/Mo-p12	GaLV Gag-Pol expression vector with Mo-MLV p12 in place of MLV p12 with 5' *Bsr*GI removed	Site directed mutagenesis of pczGaLV/Mo-p12
pczGaLV/Mo-Np12	GaLV Gag-Pol expression vector with Mo-MLV p12 N-terminus in place of GaLV p12 N-terminus with 5' BsrGI removed	Site directed mutagenesis of pczGaLV/Mo-Np12
pczGaLV/Mo-Cp12	GaLV Gag-Pol expression vector with Mo-MLV p12 C-terminus in place of GaLV p12 C-terminus with 5' BsrGI removed	Site directed mutagenesis of pczGaLV/Mo-Cp12

Appendix C-E: Legends

Appendix C: Live imaging of mCherry and mCherry-p12 expressed in mammalian cells. *M.dunni* cells were transfected with (C1) mock, (C2) mCherry or (C3) mCherry-p12 expression constructs and imaged from 12 hours post-transfection every 5 minutes for over 21 hours. Representative movies are shown. The fluorescent intensities of the mCherry signal are scaled by the same values between movies.

Appendix D: Live imaging of cells infected with GFP-p12 labelled Mo-MLV VLPs. U/R cells expressing H2A-RFP were transfected with S15mCherry (to mark the membrane) and infected with Mo-MLV VLPs labelled with GFP-p12. Movie (D1) shows the three types of GFP-p12 movements. In the red box- direct movements, the blue box- fast uncoordinated movements and the magenta box- still, shaky movements. Representative movies of early infections (20 minutes to 1 hour post-infection) and late infections (20-26 hours post-infection) with Mo-MLV containing: (D2 and D3) wild type p12; (D4 and D5) p12 mutant 5; (D6 and D7) p12 mutant 6; (D8 and D9) p12 mutant 7 and (D10 and D11) wild type p12 but no Env are shown. All fluorescent intensities were scaled by values to highlight signal over noise but are not equal between samples.

Appendix E: Live imaging of CA-YFP labelled Mo-MLV. U/R cells expressing H2A-RFP were infected with Mo-MLV labelled with CA-YFP. In (E1) movements of wild type Mo-MLV CA-YFP at 1 hour post-infection are shown. Note- the nucleus has been drawn in this movie from an image of H2A-RFP signal before the movie was recorded. (E2) shows a representative movie of wild type Mo-MLV CA-YFP docked on the mitotic chromatin early after infection. (E3) shows a representative movie of wild type Mo-MLV CA-YFP docked on the mitotic chromatin 18 hours post infection. (E4) shows a representative movie of wild type Mo-MLV GFP-p12 docked on the mitotic chromatin early after infection. All fluorescent intensities were scaled by values to highlight the signal over background noise but are not equal between samples.