Characterising Retroviral Restriction by TRIM Proteins

Sam Jack Fraser

University College London & the Francis Crick Institute

PhD Supervisor: Dr Jonathan Stoye

A thesis submitted for the degree of

Doctor of Philosophy

University College London

September 2016

I, Sam Jack Fraser, 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

Tripartite motif (TRIM) proteins are numerous in the human proteome, and a number of these molecules are known to restrict retroviral replication.

TRIM5α (T5α) is one such factor. It targets the viral capsid and imposes a block to infection between entry and reverse transcription. Capsid recognition is mediated by the C-terminal B30.2 domain, which contains surface-exposed loops of high amino acid variability. Restriction is then effected via proteasome recruitment and the induction of innate immune cascades. Although T5α is well-characterised in this respect, other factors – such as the highly divergent TRIM1 (T1) – remain poorly understood.

To further characterise the T1 restriction phenotype, chimeras of this protein and its non-restricting paralogue, T18, were generated by overlapping PCR. The restriction activities of the resulting molecules were then measured using an established flow cytometry assay. These experiments revealed that T1 also binds capsid via the B30.2 domain, although the majority of this region can be functionally replaced. Other aspects of T1 biology addressed in this work include the contribution of N-terminal components to restriction potency, and the relationship between protein expression level and restriction activity.

Following a number of attempts to generate a functional chimera of T1 and 5α, the latter half of this thesis explores how the spacing between capsid-binding and effector domains can influence restriction activity. To this end, a panel of mutations were made in the linker 2 (L2) region of T5α, and their effects on restriction measured. These experiments revealed that even small changes in interdomain spacing can have profound phenotypic consequences.

Collectively, this work reinforces the notion that TRIM family members share a common overall design, allowing individual components to be shuffled between them. At the same time, each molecule has been shaped by unique evolutionary pressures, which can render them sensitive even to relatively minor modifications.

Acknowledgements

I'd like to thank my supervisor, Jonathan Stoye, for giving me the opportunity to work in this area, and my thesis committee members – Kate Bishop, Ian Taylor and Peter Thorpe – for their ongoing guidance with the direction of the project.

I'm also indebted to all members of the Stoye & Bishop labs, past and present, for their support, advice, and the provision of countless protocols and reagents over the past four years: Melvyn Yap, Sada Ohkura, Wilson Li, Paula Ordoñez-Suárez, Marta Sanz-Ramos, Renata Varnaite, Bart Szafran, Paloma Fernandez, Ophélie Cosnefroy, Virginie Boucherit, Darren Wight, Madushi Wanaguru, Callum Donaldson, Harriet Groom and Seti Grambas. My thanks also to Neil Ball, Tom Flower and Dave Goldstone for their assistance with the structural aspects of this project, and to the former NIMR flow cytometry facility for their mentoring and technical support.

Finally, I'd like to thank my friends for making the past four years about more than just getting a PhD – and my mum for inspiring me to pursue one in the first place.

Table of Contents

List of Figures

List of Tables

Abbreviations

Δ(…)	Deletion of (…)
ψ	Packaging signal
aa	Amino acid
A3(A-H)	APOBEC subfamily 3 (members A-H)
AGM	African green monkey (*Chlorocebus aethiops*)
agmT5α	TRIM5α from the African green monkey (*Chlorocebus aethiops)*
AIDS	Acquired immunodeficiency syndrome
APOBEC	Apolipoprotein C mRNA-editing enzyme catalytic polypeptide
ATP	Adenosine triphosphate
B-MLV	B-tropic murine leukaemia virus
BCA	Bicinchoninic acid (assay)
bcT5α	TRIM5α from the brown capuchin (*Sapajus apella*)
BST-2	Bone marrow stromal antigen 2
CA	Capsid
CC	Coiled-coil (motif)
cDNA	Complementary deoxyribonucleic acid
CMV	Cytomegalovirus
COS	C-terminal subgroup one signature (domain)
Cryo-EM	Cryoelectron microscopy
CTD	C-terminal domain
cttT5α	TRIM5α from the cotton-top tamarin (*Saguinus oedipus*)
CypA	Cyclophilin A
dGTP	Deoxyguanosine triphosphate
DMEM	Dulbecco’s modified eagle medium
dNTP	Deoxynucleotide triphosphate
dNTPase	Deoxynucleoside triphosphate triphosphohydrolase
Dox	Doxycycline
dp	Decimal places
e7	Exon 7
eGFP	Enhanced green fluorescent protein
EIAV	Equine infectious anaemia virus
EM	Electron microscopy
Env	Envelope precursor protein
ESCRT	Endosomal sorting complexes required for transport
ERV	Endogenous retrovirus

eYFP	Enhanced yellow fluorescent protein
exp'ts	Experiments
FACS	Fluorescence-activated cell sorting
FFV	Feline foamy virus
FIV	Feline immunodeficiency virus
FN-III	Fibronectin type III (domain)
FV	Foamy virus
Fv1	Friend virus susceptibility 1
Gag	Group-specific antigen precursor protein
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
glycoGag	Glycosylated group-specific antigen precursor protein
GPI	Glycophosphatidylinositol
gRNA	Genomic ribonucleic acid
GTP	Guanosine triphosphate
HBV	Hepatitis B virus
HIV-1	Human immunodeficiency virus, type 1
HIV-2	Human immunodeficiency virus, type 2
HTLV	Human T-cell lymphotropic virus
huT5α	Human TRIM5α
hurhT5α	A chimera of human and rhesus TRIM5α
IFITM	Interferon-inducible transmembrane (protein)
IFN	Interferon
IFN-I	Type 1 interferon
IN	Integrase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRES	Internal ribosome entry site
kDa	Kilodalton
L2	Linker region 2
LB	Luria broth
LTR	Long terminal repeat
LxIY	Expression vector including: LTR, insert (x), IRES and eYFP
MA	Matrix
MDTF	*Mus dunni* tail fibroblasts
MERV	Murine endogenous retrovirus
MHR	Major homology region

MMTV	Mouse mammary tumour virus
MOI	Multiplicity of infection
MT	Microtubule
MW	Molecular weight
Mx2	Myxovirus resistance 2
N-MLV	N-tropic murine leukaemia virus
NB-MLV	NB-tropic murine leukaemia virus
NC	Nucleocapsid
NLS	Nuclear localisation signal
NMR	Nuclear magnetic resonance
nt	Nucleotide(s)
NTD	N-terminal domain
OM	Owl monkey (*Aotus trivirgatus*)
OMK	Owl monkey kidney (cells)
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
pbs	Primer-binding site
PBS-T	Phosphate-buffered saline supplemented with 0.1% Tween-20
PCR	Polymerase chain reaction
PFV	Prototypic foamy virus
PIC	Pre-integration complex
Pol	Polymerase precursor protein
PPT	Polypurine tract
PR	Protease
PVDF(-FL)	Polyvinylidene difluoride (for fluorescence)
RBCC	RING, B-box(es) and coiled-coil domains
REAF	RNA-associated early stage antiviral factor
rhT5α	TRIM5α from the rhesus macaque (*Macaca mulatta*)
RING	Really Interesting New Gene (domain)
RNase	Ribonuclease
RNP	Ribonucleoprotein
R	Repeat sequence in the 5’ and 3’ LTRs
RT	Reverse transcription; transcriptase
RTC	Reverse transcription complex

rtTA(3)	Reverse tetracycline-controlled trans-activator (third gen.)
SAMHD1	Sterile alpha motif and histidine/aspartic acid domain protein 1
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
SFV	Simian foamy virus
SIV	Simian immunodeficiency virus
SIVmac	Simian immunodeficiency virus from the rhesus macaque
SP1	Short spacer 1
SP2	Short spacer 2
ssRNA	Single-stranded ribonucleic acid
SU	Surface subunit of the envelope glycoprotein
T1	TRIM1 (MID2)
T1L	The long isoform of TRIM1
T1L-HA	The long isoform of TRIM1 with a C-terminal HA tag
T1S	The short isoform of TRIM1
T1S-HA	The short isoform of TRIM1 with a C-terminal HA tag
T5	TRIM5
T5α	TRIM5, alpha isoform
T5αCyp	An artificial chimera of rhT5α fused to owl monkey CypA
T18	TRIM18 (MID1)
TB	Terrific broth
TCyp	Fusion protein consisting of *TRIM5* and *CypA*
TM	Transmembrane subunit of the envelope glycoprotein
TRE(3G)	Tetracycline response element (third gen.)
TRIM	Tripartite motif
tRNA	Transfer ribonucleic acid
U3	Unique sequence in the 3’ LTR
U5	Unique sequence in the 5’ LTR
VR	Variable region
VSVg	Vesicular stomatitis virus glycoprotein

Chapter 1 Introduction

1.1 Retroviruses

1.1.1 Classification of retroviruses

The retroviruses (*Retroviridae*) are a large family of enveloped, positive-sense RNA viruses. They are characterised by a lifecycle involving reverse transcription of their genome into dsDNA, and the subsequent integration of this molecule into host chromatin. The *Retroviridae* were classically assigned to one of four groups – A, B, C or D – according to their particle morphology as observed by electron microscopy (Vogt, 1997). However, this system has since been displaced by a two-subfamily system, in which all retroviruses are divided between the *Orthoretrovirinae* and *Spumaretrovirinae* (Stoye et al., 2011) (Figure **1.1**). The former can be further divided into 6 genera – the α, β, γ, δ and ε-retroviruses and the lentiviruses – while the latter comprises a single grouping, owing to unique aspects in their morphology and replication (Lochelt and Flugel, 1996; Yu et al., 1996b).

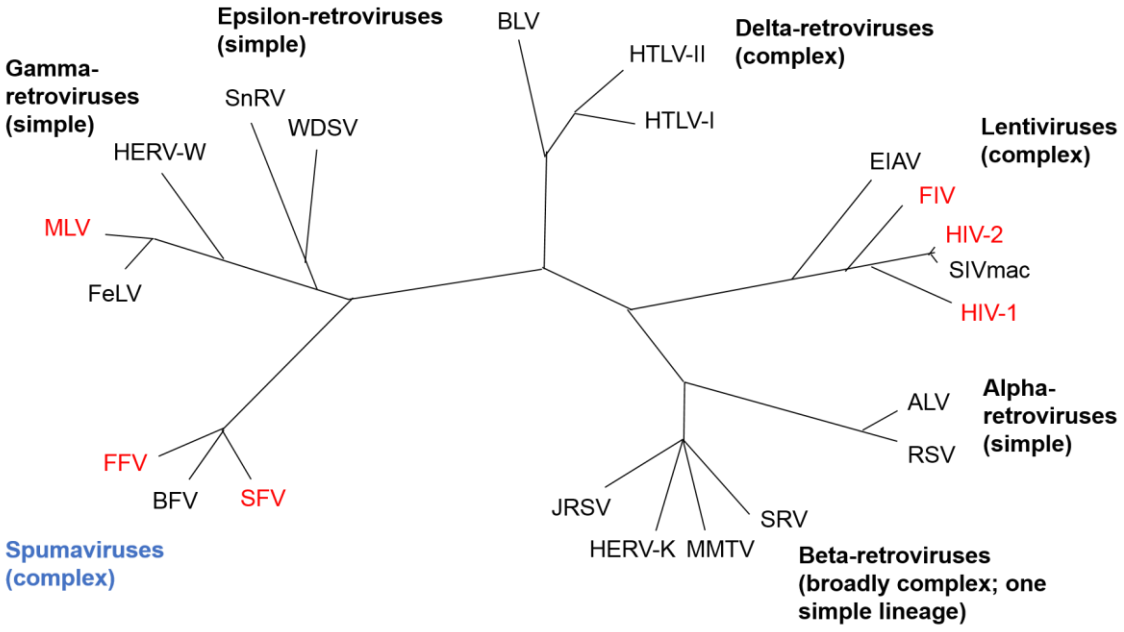


Figure 1.1: Phylogeny of the *Retroviridae*

The tree was derived by alignment of reverse transcriptase coding sequences. Simple retroviral genomes encode only the basic structural and catalytic proteins, while complex genomes contain additional, overlapping ORFs that encode various regulatory and accessory proteins (see Section **1.1.3**). Genera from the *Orthoretrovirinae* are written in black font, and those from the *Spumaretrovirinae* in blue. Viruses of relevance to this work are highlighted in red. Adapted from Weiss (2006).

1.1.2 Retroviral particles

Mature retroviral particles typically span 80–120 nm in diameter. They are diploid, possessing two copies of genomic RNA (gRNA) that dimerise via a 'kissing loop' structure and are packaged with nucleocapsid (NC) proteins (Skripkin et al., 1994; Clever et al., 1996). This complex, along with multiple viral enzymes, is contained within a hexameric lattice of capsid (CA) monomers to form the viral core. Different retroviruses exhibit distinct core morphologies: for example, while N-MLV and most other orthoretroviruses possess spherical cores, in the lentiviruses this structure is conical, and in epsilonretroviruses, cylindrical (Figure **1.2**).

In a mature virion, the viral core is encased in matrix (MA) proteins, and the entire structure surrounded by a lipid envelope. The envelope is studded with Env glycoproteins, which comprise a surface subunit (SU) that binds to the target cell, and a transmembrane subunit (TM) that mediates fusion between the viral and cellular membranes.

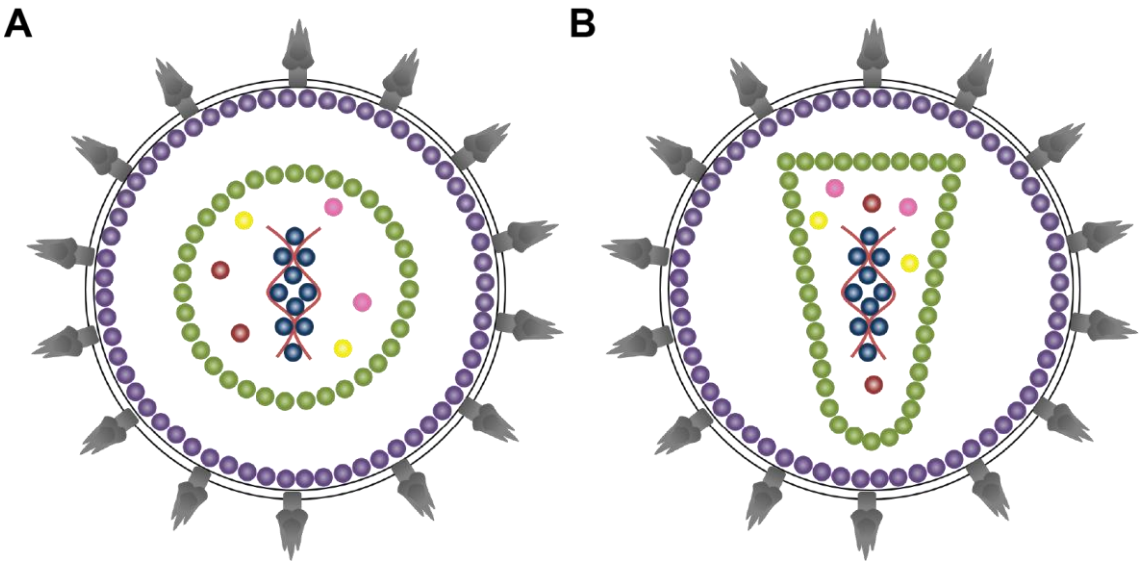


Figure 1.2: Cross-sections of a mature retrovirus with (A) spherical and (B) conical core morphology

gRNA is depicted in red; NC, blue; PR, pink; RT, red; IN, yellow; CA, green; MA, purple and Env, grey.

1.1.3 Retroviral genome organisation

The retroviral genome is a single-stranded, non-segmented, positive-sense RNA molecule, ranging from 7–12 kb in length. It exists within the virion as a homodimer, which is maintained through hydrogen bonding between 5' loop structures called dimer linkage sequences (Jones et al., 1993).

Like a eukaryotic transcript, the viral genomic RNA (gRNA) possesses a 5' methylguanosine cap and a 3' polyadenylated tail in order to support translation. Each end of the RNA has a repeat sequence (R) and a sequence unique to that terminus (U5 or U3). These sequences are duplicated during reverse transcription, resulting in a double-stranded cDNA that is flanked by *long terminal repeats* (LTRs) with U3-R-U5 architecture (Figure **1.3**). After integration, the 5' LTR directs transcription of the proviral DNA.

Immediately downstream of the U5 region in gRNA is a primer-binding site (pbs). This binds the tRNA responsible for priming minus-strand DNA synthesis during reverse transcription. Following the pbs, all retroviruses possess the *gag*, *pol* and *env* genes, in that order. *Gag* encodes the structural proteins of the virus; *pol*, the replicative enzymes, and *env*, the glycoproteins that stud the outer surface of the virion. Towards the 3' end of this molecule is a polypurine tract (PPT), which primes plus-strand DNA synthesis. Some retroviruses, including HIV-1, possess an additional PPT in the centre of the genome (cPPT) to permit dual initiation of plus-strand synthesis. All retroviral gRNAs also contain a packaging signal (ψ) to promote their specific encapsidation during virion assembly.

Complex retroviruses – such as HIV-1 – have a more elaborate genomic structure, containing additional genes encoded in overlapping reading frames (Figure **1.4**). The proteins that they encode perform diverse functions, including transcriptional trans-activation and immune evasion (Emerman and Malim, 1998).

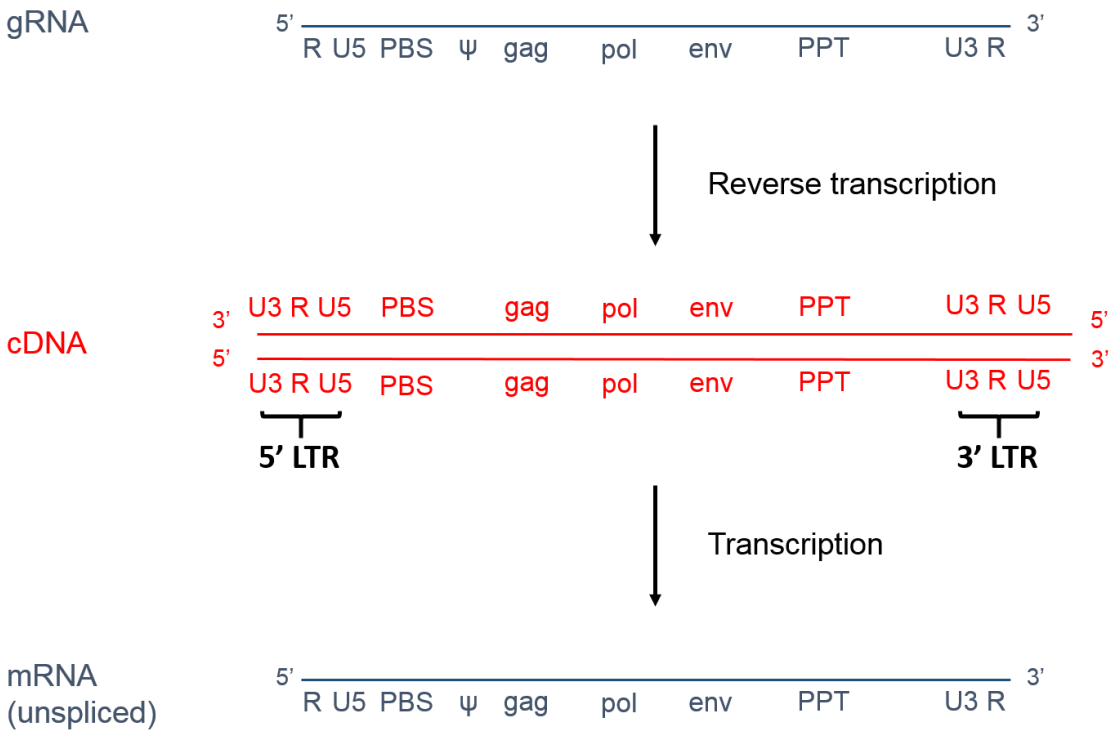


Figure 1.3: Retroviral nucleic acid metabolism

Ψ indicates the location of the packaging signal.

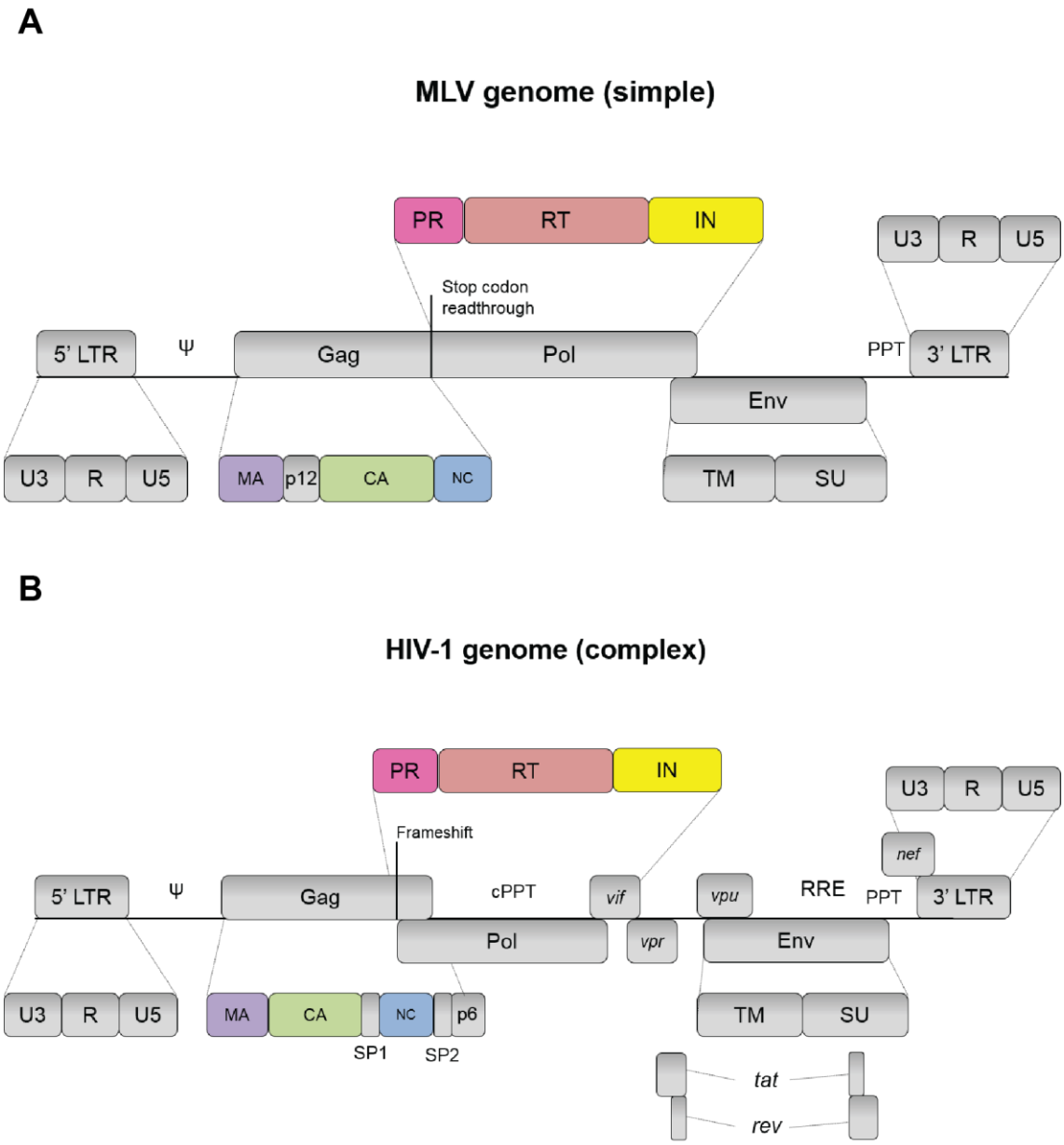


Figure 1.4: A comparison of the proviral genomes of (A) simple and (B) complex retroviruses

RRE: Rev-response element (described in Section **1.2.5**).

1.1.4 Retroviral proteins

Orthoretroviruses encode all of their structural and catalytic proteins within three open reading frames: *gag, pol* and *env*. Each of these is translated as a polyprotein (Gag, Gag-Pol or Env), which is cleaved into its constituent parts during virion maturation (see Section **1.2.8**). Spumaviruses also encode Gag, Pol and Env, although their post-translational cleavage is comparatively limited. This section will focus broadly on the structure and function of the orthoretroviral molecules; their spumaviral counterparts will be discussed in Section **1.1.6**.

Gag

The group-specific antigen (Gag) polyprotein contains most of the structural proteins of the virus. From the N-terminus, these are matrix (MA), capsid (CA) and nucleocapsid (NC). Each of these components is integral to viral function; however, Gag itself plays a number of roles in the viral lifecycle prior to cleavage, including the recruitment of gRNA, trafficking of viral components to the plasma membrane, and the assembly and budding of nascent virions. A detailed review of the role of Gag in these processes can be found in Freed (2015); the remainder of this section will deal with each of its cleavage products in turn.

MA targets Gag to the plasma membrane during the assembly of both HIV and MLV virions (Ono et al., 2000; Li et al., 2013a). To facilitate this process, the MA proteins of many retroviruses are myristoylated at their N-termini to permit hydrophobic interactions with the lipid bilayer (Rein et al., 1986; Bryant and Ratner, 1990; Liu et al., 2011b). It has been postulated that a patch of surface-exposed basic residues in the MA N-terminus might also contribute to this process, via electrostatic interactions with the acidic phospholipid head groups (Murray et al., 2005).

Although there is little sequence identity in the MA proteins of different retroviruses, there is strong conservation in their structural arrangement. This consists of four α-helices that come together to form a globular core, which is then capped by a three-stranded β-sheet. In the immature Gag polyprotein, a fifth helix links MA to the adjacent CA domain (Massiah et al., 1994; Conte and Matthews, 1998).

Crystal structures are available for MA from both HIV-1 and Moloney MLV (Momany et al., 1996; Riffel et al., 2002).

CA forms a hexameric lattice that both protects the viral genome and interacts with numerous cellular co-factors necessary for replication. For example, HIV-1 CA interacts with a host of factors involved in the nuclear import and trafficking of the viral pre-integration complex (PIC) (Price et al., 2012b; Chen et al., 2016). The structure and function of CA are particularly pertinent to this project and will therefore be treated separately in Section **1.1.5**.

NC is a highly basic protein that co-ordinates Zn^{2+} ions and engages in various critical protein-nucleic acid interactions. HIV-1 NC contains two zinc fingers with the canonical sequence $CX_4CX_4HX_4C$, separated by a short, basic linker (Darlix et al., 1995), while that of MLV contains only a single zinc finger. NC is required for a number of stages in the retroviral lifecycle, including genome dimerisation (Darlix et al., 1990); packaging of viral gRNA (Berkowitz et al., 1995); reverse transcription (Tsuchihashi and Brown, 1994; Cristofari and Darlix, 2002) and the integration of proviral DNA (Carteau et al., 1999). All of these functions hinge on the ability of NC to function as a nucleic acid chaperone, facilitating structural rearrangements within these molecules to maintain them in thermodynamically stable conformations (Rein et al., 1998; Levin et al., 2005).

Some retroviruses encode additional proteins within the *gag* reading frame. For example, HIV-1 encodes an unstructured peptide called p6 at the extreme C-terminus of Gag, as well as spacer peptides 1 and 2 (SP1/2) at the junctions between CA and NC, and NC and p6, respectively. SP1 and 2 contain cleavage sites that are important for proper maturation of the retroviral core (Accola et al., 1998; de Marco et al., 2012), while p6 recruits ESCRT proteins during budding (Pornillos et al., 2002; Sette et al., 2010) and interacts with the viral accessory protein, Vpr (Salgado et al., 2009). Meanwhile, MLV encodes p12 between MA and CA, which contributes to functions such as core stability and chromatin tethering (Wight et al., 2012).

Pol

The polymerase (Pol) polyprotein encodes the viral enzymes necessary for replication. From the N-terminus, these are protease (PR), reverse transcriptase (RT) and integrase (IN). In orthoretroviruses, Pol is invariably translated as part of a Gag-Pol polyprotein (160 kDa). Stepwise proteolytic cleavage then liberates each of the subunits during maturation of the virion (Section **1.2.8**).

PR is an aspartyl protease that functions as a homodimer. It first auto-catalytically cleaves itself from the Pol polyprotein, and then mediates a series of carefully-timed, high-fidelity proteolytic events necessary for the maturation of nascent virions (Wiegers et al., 1998; Goodenow et al., 2002). Each PR monomer consists of a β-hairpin followed by a wide loop, an α-helix and a second β-hairpin, all of which are present in duplicate; the active site of the enzyme lies in a cleft at the interface between monomers (Wlodawer and Erickson, 1993). PR structures are available for HIV-1, HIV-2, feline immunodeficiency virus (FIV) and simian immunodeficiency virus (SIV), among others (Wlodawer et al., 1989; Mulichak et al., 1993; Rose et al., 1993; Wlodawer et al., 1995). Furthermore, crystal structures snapshotting the mechanism of proteolysis by HIV-1 PR have been captured to about 1.3 Å resolution (Shen et al., 2012).

RT is the enzyme responsible for reverse-transcribing viral gRNA into an integration-competent cDNA. It harbours multiple functions to this end, including RNA- and DNA-dependent DNA polymerase activities, and an RNaseH activity for degradation of the template. HIV-1 RT forms an asymmetric heterodimer, with one subunit that comprises both polymerase and RNaseH domains (p66), and another that lacks the RNaseH component (p51). Interestingly, despite identical amino acid sequences, the polymerase domains of each subunit exhibit distinct tertiary structures. High-resolution crystal structures are available for the heterodimeric RT of HIV-1 and -2, and the monomeric RT of Moloney MLV (Jacobo-Molina et al., 1993; Das et al., 2008; Ren et al., 2002).

IN catalyses the steps necessary for the insertion of viral cDNA into host chromatin. It comprises three structural domains for this purpose: an N-terminal zinc-binding

domain, a catalytic core domain, and a C-terminal DNA-binding domain (Johnson et al., 1986; Engelman and Craigie, 1992). Following reverse transcription, IN forms a complex with the viral cDNA known as an *intasome*. Within this structure, the DNA undergoes a series of reactions in order to become integration-competent (Section **1.2.4**).

The structure of the PFV intasome was solved some time ago, revealing a homotetramer (a dimer of dimers) that tightly associates with two ends of the viral DNA (Hare et al., 2010; Maertens et al., 2010). Until earlier this year, it was believed that orthoretroviral IN also functions as a tetramer; however, recent characterisation of the intasome from mouse mammary tumour virus (MMTV), a β-retrovirus, has instead revealed an octameric architecture (Ballandras-Colas et al., 2016). This structure comprises a central tetramer with two flanking dimers, which associate with the core via their C-terminal domains (CTDs). These dimers are functionally important because they supply the target-DNA-binding activity of the intasome. This function cannot be provided in *cis* by the central tetramer due to a structurally restrictive linker region between the catalytic core and CTD of MMTV IN (compared to its longer, more flexible counterpart in PFV). The recent characterisation of the intasome of rous sarcoma virus (RSV), an α-retrovirus, revealed that this structure is also octameric, with one pair of integrase dimers engaging either end of the viral DNA for catalysis, while the other pair capture the target DNA ready for strand transfer. (Yin et al., 2016)

Env

Env encodes the envelope glycoproteins that stud the surface of the virion. These are required both for adsorption to cell surface receptors, and for the subsequent fusion of viral and host membranes.

The Env precursor of HIV-1 (gp160) is a heterodimer of surface (SU; gp120) and transmembrane (TM; gp41) subunits. This polyprotein is translated directly into the endoplasmic reticulum (ER), where it undergoes co-translational glycosylation of several key residues, before assembling into trimers in the ER lumen. The homotrimer then migrates to the Golgi apparatus, where each monomer is cleaved

into its constituent parts by subtilisin-like endoproteases (Hallenberger et al., 1997a). Crystal structures are available for HIV-1 Env in its native, trimeric form (Julien et al., 2013; Do Kwon et al., 2015).

HIV-1 regulatory and accessory proteins

Complex retroviruses such as HIV-1 possess a number of ORFs in addition to *gag, pol* and *env* (see Table **1.1**). These ORFs encode *regulatory proteins*, which are essential for replication, and *accessory proteins*, which are broadly dispensable *in vitro*, but may be required *in vivo*.

	Gene	Protein	Function(s)
Regulatory	*tat*	Tat	• Transcriptional trans-activation[1]
	rev	Rev	• Nuclear export of viral mRNA[2] • Temporal regulation of transcription[3]
Accessory	*nef*	Nef	• Downregulation of CD4 receptors[4] • Inhibition of T-cell activation[5] • Counteraction of SERINC3/5 (Section **1.3.2**)
	vpr	Vpr	• Nuclear localisation of the PIC[6] • Cell cycle modulation[7] • Counteraction of SAMHD1 (Section **1.3.4**)
	vif	Vif	• Counteraction of A3G (Section **1.3.3**)
	vpu	Vpu	• Downregulation of CD4 receptors[8] • Counteraction of tetherin (Section **1.3.8**)

Table 1.1 Functions of the regulatory and accessory proteins encoded by HIV-1

References: (1) Ruben et al., (1989); (2) Zapp and Green (1989); (3) Kim et al. (1989); (4) Garcia and Miller (1991); (5) Luria et al. (1991); (6) Heinzinger et al., 1994); (7) Jowett et al. (1995); (8) Chen et al. (1993).

1.1.5 Structure of the retroviral capsid

The retroviral capsid monomer (CA) is one of three structural proteins encoded within the *gag* reading frame. The structure of this protein is particularly pertinent to this thesis as it is the target of numerous retroviral restriction factors.

CA lies between MA and NC in the immature Gag polyprotein. During maturation, it is proteolytically liberated from this precursor and condenses into a core structure that is conical in HIV-1 and spherical in MLV (Ganser-Pornillos et al., 2004). Although retroviruses differ tremendously in the primary sequence of CA, the secondary and tertiary structures of this protein are remarkably well-conserved (de Marco et al., 2010a).

The HIV-1 CA monomer is split between a 150-residue N-terminal domain (NTD) and an 80-residue C-terminal domain (CTD) (Figure **1.5**). The former consists of a β-hairpin followed by seven α-helices, with a proline-rich loop between helices 4 and 5 that binds the peptidyl prolyl isomerase, cyclophilin A (CypA) (Gamble et al., 1996; Gitti et al., 1996; Du et al., 2011). Meanwhile, the latter comprises 4 α-helices, an unstructured region, and a core of highly conserved hydrophobic residues known as the *major homology region* (MHR), which coordinates conformational changes in CA during maturation (Gamble et al., 1997). Unsurprisingly, mutational inactivation of the MHR can have profound effects on virion morphology and infectivity (Purdy et al., 2008). The NTD and CTD are joined in the middle by a flexible linker, which is required for their correct orientation in higher-order capsid structures (Arvidson et al., 2003).

MLV CA is, in many ways, structurally reminiscent of its lentiviral counterpart. The NTD comprises a β-hairpin followed by six α-helices, the first three of which superimpose almost perfectly on the HIV structure (Mortuza et al., 2004); however, unlike HIV, MLV CA does not bind CypA. A crystal structure for the CTD of this protein is presently unavailable.

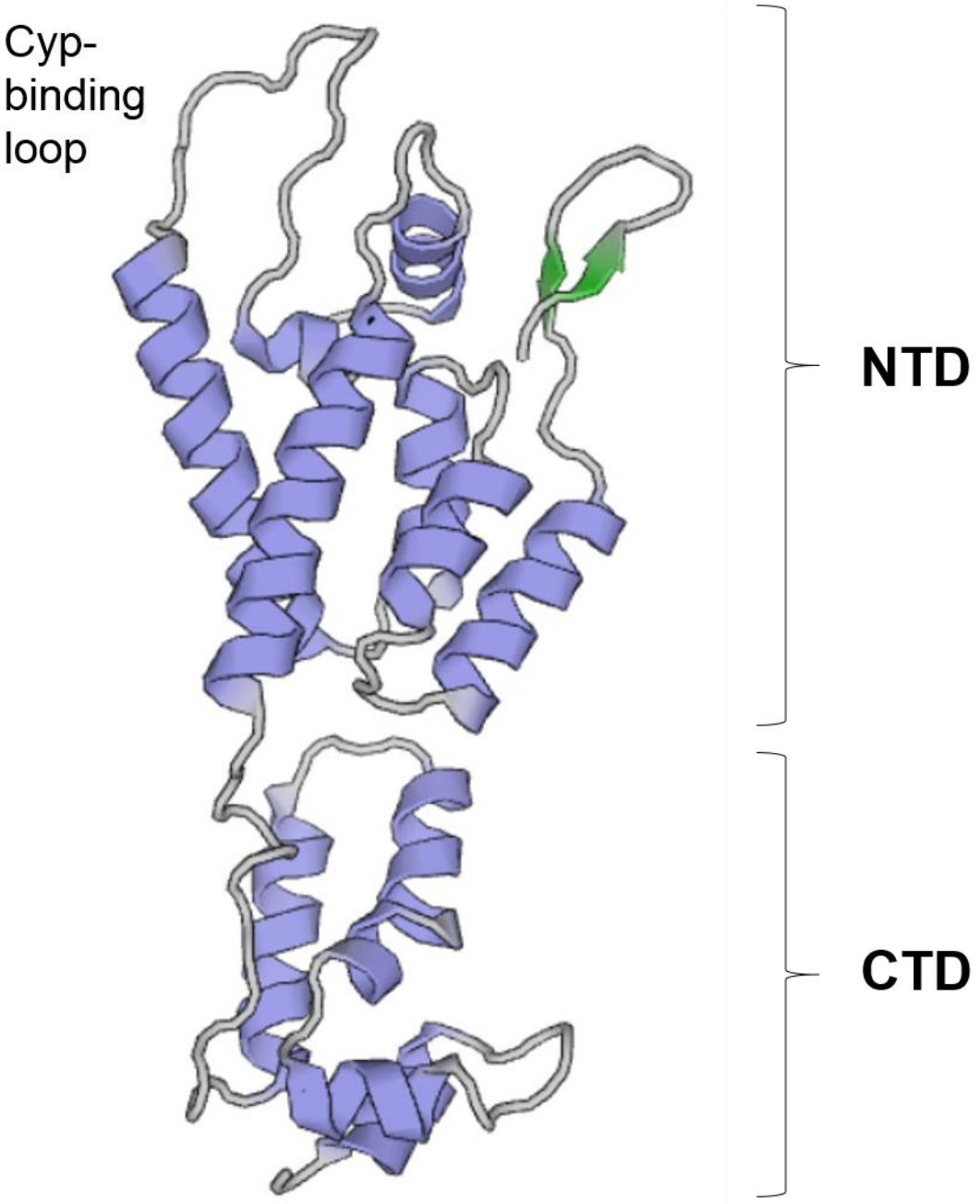


Figure 1.5: HIV-1 CA monomer

The monomeric structure of HIV-1 CA. The N-terminal β-hairpin is indicated in green, and α helices in blue. PDB accession code: 2M8N (Deshmukh et al., 2013)

The mature retroviral core is a lattice of CA hexamers, punctuated by occasional pentamers that offer curvature at the top and bottom of the structure. The mature core of HIV-1 consists of approximately 1500 CA monomers arranged in a hexameric lattice, with 12 pentamers that close the lattice into a fullerene cone: five at the apex and seven at the base (Briggs et al., 2004; Zhao et al., 2013). Recent work has highlighted the importance of water molecules in the capsid structure, both for stabilising inter-hexamer interactions and permitting conformational changes that are requisite at different stages of the lifecycle (Gres et al., 2015).

In the assembled lattice, each CA monomer is oriented such that the NTD is located on the outer surface and the CTD buried underneath. Thus, the majority of residues in CA that govern restriction factor sensitivity map to the NTD. Multiple interactions between CA monomers are responsible for preserving the integrity of this structure, including sixfold NTD-NTD interfaces within a hexamer (Ganser-Pornillos et al., 2007); two- and threefold CTD-CTD interfaces between hexamers (Ivanov et al., 2007; Byeon et al., 2009); and individual NTD-CTD interfaces within hexamers (Pornillos et al., 2009; Yufenyuy and Aiken, 2013). The latter interface also forms a binding pocket, which is necessary for interactions with numerous cellular cofactors, and is the target for various antiretroviral compounds (Bhattacharya et al., 2014; Price et al., 2014).

Recent data has revealed that HIV-1 CA is the most genetically fragile of any protein for which this property has been quantified (Rihn et al., 2013). In other words, it is highly intolerant of non-synonymous substitutions, with approximately 70% of amino acid changes yielding non-viable mutants. This low mutational robustness reflects the need for CA to form a diverse range of contacts in the mature core. For example, an identical monomer must engage in different sets of interactions depending on whether it contributes to a hexamer or a pentamer. In addition to these spatial restrictions, CA must also form contacts that are temporally conducive to the processes of uncoating and maturation, while interacting with a host of cellular co-factors that are critical for replication, including CPSF6, CypA, and components of the nuclear pore complex (Luban et al., 1993; Price et al., 2012a; De Iaco et al., 2013). The MLV capsid is also relatively intolerant of mutation (Auerbach et al., 2006; Auerbach et al., 2007).

In summary, CA is subject to strong purifying selection in order to preserve both structure and function. This renders it susceptible to immune recognition because it lacks the scope for diversification, perhaps explaining why nature has selected the capsid as an opportune antiretroviral target.

1.1.6 The *Spumaretrovirinae*

The *Spumaretrovirinae* (also known as foamy viruses; FVs herein) are a genus of retrovirus that were first described in the 1950s and isolated about twenty years later (Enders, 1954; Achong et al., 1971). While these viruses induce cytopathic effects *in vitro*, there is currently no evidence of pathogenicity in humans or other animals (Linial, 2000).

FVs are widespread among mammalian hosts, with isolates available from baboons, chimpanzees, gorillas, cats and cows, among others (Linial, 1999); (Meiering and Linial, 2001). The first to be discovered – *prototypic foamy virus* (PFV) – was isolated from a human nasopharyngeal cell line in the seventies (Achong et al., 1971) and was the first member of the genus to be cloned and sequenced (Flügel et al., 1987; Maurer et al., 1988). Eventual sequencing of other FV genomes revealed that these viruses are genetically distinct from all other retroviral genera – their closest known relatives being endogenous retroviruses of human and murine origin (Cordonnier et al., 1995). In fact, given certain aspects of their lifecycle and the fact that infectious particles carry DNA rather than RNA, FVs are often regarded as 'bridging the gap' between retroviruses and the *Hepadnaviridae*, the only other family of reverse-transcribing viruses.

Like all retroviruses, FV genomes possess the *gag, pol* and *env* ORFs. These genes encode proteins with the same basic functions as those already described. However, there are marked differences that distinguish the FV molecules from their counterparts in other retroviral genera.

Gag

In contrast to the orthoretroviral protein, FV Gag does not comprise individual MA, CA and NC subunits that are liberated during maturation. Instead, it undergoes a single, PR-dependent cleavage at the C-terminus to yield full-length (p71) and truncated (p68) species that appear in the viral capsid at a ratio between 1:1 and 1:4 (Enssle et al., 1997; Cartellieri et al., 2005). The presence of both molecules is critical for optimal infectivity, though not for particle release (Enssle et al., 1997). FV Gag is also distinctive in its low abundance of lysine residues, with the majority of the protein's basic content coming from arginines. Interestingly, however, R-K substitutions do not appear to have a deleterious effect on FV replication in culture (Matthes et al., 2011).

In place of the MA-CA-NC subdomain structure of orthoretroviral Gag, the FV protein possesses four coiled-coil (CC) domains that perform analogous functions (Matthes et al., 2011) (Figure **1.6**). At the N-terminus, CC1 appears to be involved in interactions between Gag and Env that are required for particle release, while CC2 mediates the homotypic Gag-Gag interactions that are necessary for capsid assembly (Tobaly-Tapiero et al., 2001). Definitive functions have not been assigned to CC3 and 4, although there are indications that the former mediates an interaction between FV Gag and the light chain of dynein motor protein complexes, which is necessary for the trafficking of incoming virions to the microtubule-organising centre (MTOC) (Petit et al., 2003). Like its orthoretroviral counterpart, FV Gag also harbours an L-domain for the recruitment of ESCRT proteins during egress.

FV Gag is distinct from that of orthoretroviruses in that it lacks zinc finger motifs for genome packaging. These are instead replaced with glycine-arginine-rich domains known as GR boxes, which reside at the C-terminus of Gag and bind DNA and RNA with equal affinity (Schliephake and Rethwilm, 1994; Yu et al., 1996c). The primate variants possess three GR boxes, which contribute to genome binding, reverse transcription and chromatin tethering (Tobaly-Tapiero et al., 2008; Müllers et al., 2011). A putative nuclear export signal (NES) has also been identified at the

N-terminus of FV Gag, indicating a role in the nuclear export of unspliced and singly spliced viral transcripts (Renault et al., 2011).

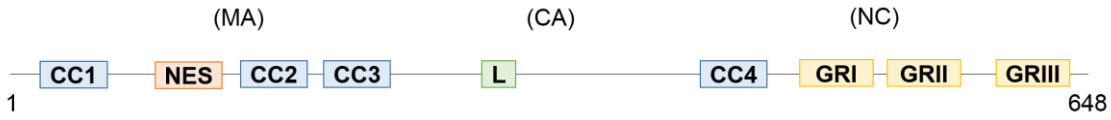


Figure 1.6: A typical FV Gag molecule

The Gag polyprotein of PFV compared to the MA-CA-NC subdomain architecture of orthoretroviral Gag. CC: coiled coil; L: late domain; GR: glycine-arginine box. Adapted from Müllers (2013).

A crystal structure of the N-terminal domain of PFV Gag (PFV-Gag-NTD) is available at 2.4 Å resolution (Goldstone et al., 2013). This revealed distinct structural divergence when compared to the orthoretroviral MA domain, despite conservation of function. While the MAs of other retroviruses possess a highly basic region – and often a myristate moiety – at their N-termini, neither of these components are present in PFV-Gag-NTD. Additionally, the tertiary structure of the latter comprises a mixed αβ topology with head and stalk domains. This stands in stark contrast to the predominantly α-helical, globular structure of orthoretroviral MA (Conte and Matthews, 1998). Despite these differences, however, the capsid-binding restriction factor, T5α (see Section **1.4.5**) appears to bind both PFV-Gag-NTD and the NTD of orthoretroviral CA (Yap et al., 2008; Goldstone et al., 2013).

Pol

In the orthoretroviruses, either frameshift or readthrough of the *gag* stop codon enables the neighbouring *pol* ORF to be translated as part of a Gag-Pol polyprotein; it is then liberated by PR during virion maturation. In FVs, however, no such fusion is detectable in infected cells, even when the active site of PR is mutated (Konvalinka et al., 1995b).

Instead, FV Pol is translated from a separate, singly spliced mRNA (Yu et al., 1996a), and its expression is regulated post-transcriptionally by way of a suboptimal splice acceptor site upstream of its start codon (Lee et al., 2008). While orthoretroviruses incorporate Pol into virions via covalent linkage to Gag, in FVs this is instead mediated by an interaction between Pol and the viral gRNA (Heinkelein et al., 2002). FV Pol is also peculiar in that undergoes only a single internal cleavage during maturation, yielding free IN and a PR-RT fusion (Pfrepper et al., 1998). Both molecules adopt a nuclear localisation (Imrich et al., 2000).

Env

Like the orthoretroviral protein, FV Env is translated from a spliced mRNA directly into the endoplasmic reticulum (ER), where it undergoes co-translational glycosylation of several key residues. However, FV Env is distinct from the orthoretroviral glycoprotein in that it retains the signal peptide that directs translation to the ER. This results in an Env precursor comprising an N-terminal signal peptide (termed the *leader peptide, LP,* from this point forward) in addition to the usual surface (SU) and transmembrane (TM) subunits. The LP is proteolytically liberated by furin or furin-like proteases as the protein translocates through the secretory pathway (Duda et al., 2004; Geiselhart et al., 2004).

FVs are largely intolerant of pseudotyping with heterologous Env glycoproteins, including those from MLV and VSVg (Lindemann et al., 1997; Pietschmann et al., 1999). This is attributable to an interaction between FV Gag and its cognate Env, which involves the N-termini of both partners and is essential for virion release (Wilk et al., 2001; Lindemann et al., 2001). A crystal structure is available for PFV Gag in complex with the Env leader peptide (Goldstone et al., 2013).

Accessory proteins

Like other complex retroviruses, FVs encode genes in addition to the *gag, pol* and *env* ORFs. FVs possess two such genes, both of which lie towards the 3' end of the genome.

Tas (formerly *bel-1*) encodes a 36 kDa transcriptional trans-activator with analogous function to the Tat protein of HIV-1. It harbours a C-terminal transcription activation domain and a centrally located DNA-binding domain (Blair et al., 1994; He et al., 1996), and binds to DNA sequences that contain conserved purine residues, but little other sequence identity (Kang et al., 1998). PFV Tas is indispensable for replication, and may also control the transcription of specific cellular genes (Baunach et al., 1993; Wagner et al., 2000).

Bet is an accessory protein found in all known foamy viruses. It has numerous functions, including conferring resistance to superinfection (Bock et al., 1998); negative regulation of proviral transcription (Meiering and Linial, 2002); and inhibition of restriction by the APOBEC (A3) family of enzymes (Löchelt et al., 2005). The latter function is analogous to that of HIV-1 Vif; however, unlike Vif, Bet acts by preventing the incorporation of A3 into virions (Lukic et al., 2013; Jaguva Vasudevan et al., 2013).

1.2 Retroviral replication

The retroviral lifecycle is complex and has been extensively described in a number of reviews (Perez and Nolan, 2001); (Amara and Littman, 2003); (Nisole and Saib, 2004). It can broadly be divided into an early phase, where virions enter the cell and travel to their site of genomic integration, and a late phase, where viral genes are expressed and progeny virions are synthesised, assembled and released. The entirety of this process is detailed in Figure **1.7** (below). Each stage is individually described in the sections that follow.

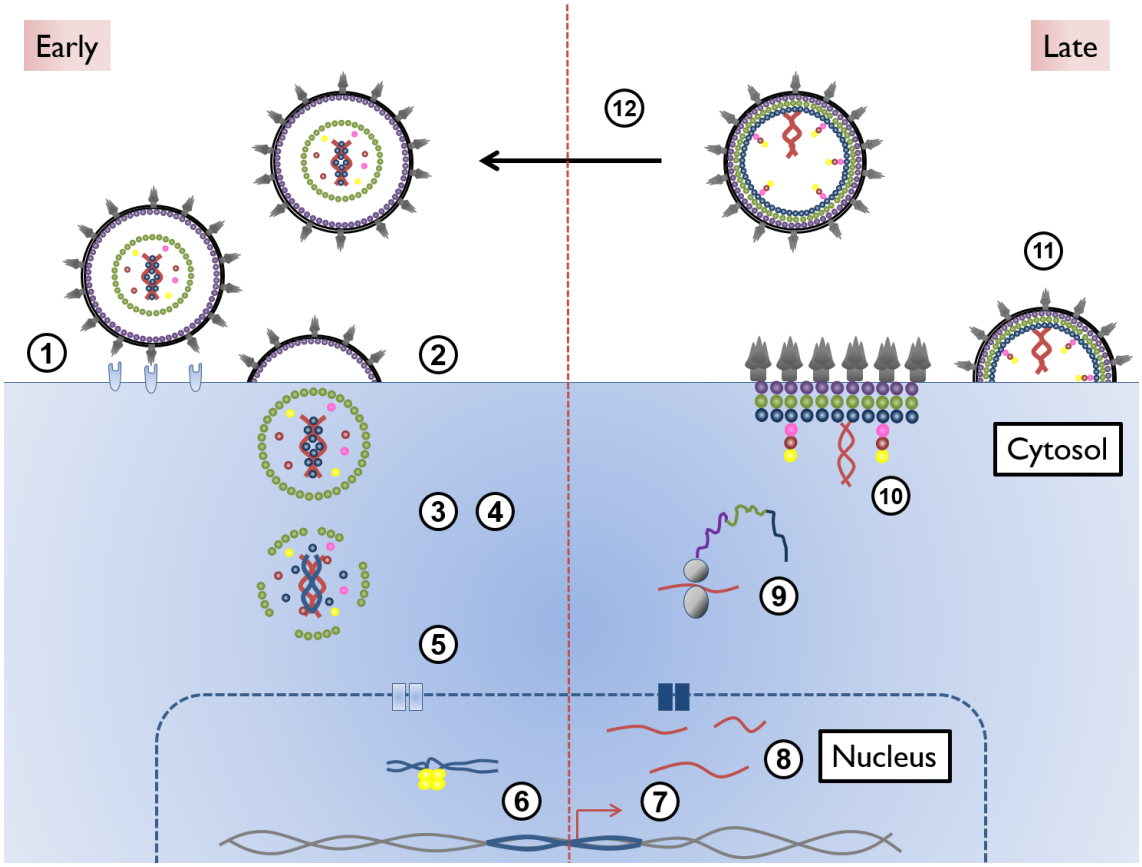


Figure 1.7: The retroviral lifecycle

(1) Adsorption; (2) cell entry; (3) uncoating; (4) reverse transcription; (5) nuclear trafficking and entry; (6) integration; (7) proviral transcription; (8) splicing and nuclear export; (9) translation; (10) assembly; (11) budding; (12) maturation.

1.2.1 Adsorption and entry

The initial encounter between a retrovirus and its target cell is mediated by weak interactions between the viral envelope glycoprotein, Env (or cellular proteins that have been incorporated into the virion membrane) and cell surface receptors: typically heparin for MLV (Walker et al., 2002), and heparin sulphate for HIV-1 (Saphire et al., 2001; Vivès et al., 2005). This provisional interaction is not essential for infection, although it does improve the efficiency of viral entry by bringing virions into close proximity with their primary receptor (Ugolini et al., 1999). Once a virion is adsorbed to the cell surface, a stronger interaction between Env and the viral receptor can proceed. This is sometimes supplemented by a secondary interaction with a co-receptor.

In the case of MLV, the receptor utilised determines the tropism of the strain. MLVs can be classified into ecotropic, xenotropic, polytropic and amphotropic subgroups depending on their host range. Ecotropic MLVs infect only mouse or rat cells using the mouse cationic amino acid transporter, mCAT-1 (Albritton et al., 1989). The remaining subgroups infect a broader range of mammalian hosts: amphotropic MLVs utilise the sodium-dependent phosphate transporter, Pit-2, to accomplish this (Kavanaugh et al., 1994), while poly- and xenotropic MLVs use different alleles of the Xpr1 cell-surface receptor (Kozak, 2010).

In the case of primate lentiviruses, entry involves engagement with the CD4 receptor, which is present on T-cells, macrophages, monocytes and dendritic cells – all of which are susceptible to infection by these viruses. Briefly, the gp120 subunit of Env binds to a membrane-distal region of CD4, thereby inducing a number of conformational changes in the former, first in the V1/V2 loops and subsequently in V3 (Kwong et al., 1998). CD4 binding also induces the formation of two double-stranded β-sheets (Chen et al., 2005) which, in combination with the reconfigured V3 loop, facilitate the engagement of a co-receptor.

Co-receptor binding by the V3 loop of gp120 is broadly considered the catalyst that triggers fusion of the viral and cellular membranes, specifically by externalising the hydrophobic gp41 fusion peptide of Env. The only co-receptors known to be

important for HIV-1 infection *in vivo* are the chemokine receptors CXCR4 (for X4-tropic viruses) or CCR5 (R5-tropic viruses); viruses that can engage either co-receptor are dubbed R5X4 viruses (Berger et al., 1998). With few exceptions, only R5 and R5X4 viruses are able to transmit between individuals (Keele et al., 2008). However, progression from R5 to X4 tropism *in vivo* is typically associated with rapid T-cell depletion and the onset of AIDS (Tersmette et al., 1989; Scarlatti et al., 1997).

Once receptor and co-receptor are engaged, HIV-1 co-opts underlying cytoskeletal components to 'surf' across the membrane to a site where membrane fusion can occur (Lehmann et al., 2005). Once a suitable region has been located, the exposed gp41 fusion peptide inserts into the host membrane, tethering the virion to the cell. This anchoring induces the three gp41 subunits of each Env trimer to fold at a hinge region, bringing their N- and C-termini together to form a six-helix bundle (6HB) (Chan et al., 1997). Because the N-termini are proximal to the cell membrane and the C-termini to the viral membrane, the formation of the 6HB brings the two partners together to create, and then stabilise, a fusion pore (Melikyan, 2008). This is the portal through which the viral core enters the cytoplasm.

1.2.2 Reverse transcription and uncoating

Following internalisation, the viral core undergoes numerous transformations in order to become integration-competent. These include the conversion of viral RNA into dsDNA (reverse transcription), the progressive displacement of capsid proteins (uncoating), and the trafficking of this reverse-transcribing structure, known as the reverse transcription complex (RTC), towards the nucleus. The precise order and interdependence of these events remains the subject of ongoing research (reviewed by (Campbell and Hope, 2015)); however, in the interests of clarity, reverse transcription and uncoating will be covered in this section, and nuclear trafficking in the subsequent one (Section **1.2.3**).

Reverse transcription

Reverse transcription (Figure **1.8**) is initiated by the annealing of a partially unwound host tRNA to an 18-nt primer binding site (pbs) at the 5' end of the viral genome. The tRNA species utilised for this purpose differs between retroviral genera. For example, while HIV-1 uses $tRNA^{lys3}$ (Wain-Hobson et al., 1985), MLV uses $tRNA^{pro}$ (Peters et al., 1977).

Minus-strand DNA synthesis is then initiated from the 3' end of the tRNA primer and progresses towards the 5' end of the genomic template, yielding a DNA-RNA hybrid. The RNaseH activity of RT degrades the RNA portion of this structure, leaving behind a single-stranded DNA species known as *minus-strand strong stop DNA*. This molecule possesses a repeat sequence (R) at its 3' end, which enables it to hybridise with the complementary sequence at the 3' end of the genomic template in a process called *first-strand transfer*. Following this jump, minus-strand DNA synthesis is completed up to the pbs, and RNaseH degrades the majority of the remaining template.

A short, purine-rich sequence towards the 3' end of the viral RNA, known as the *polypurine tract* (PPT), is able to resist degradation by RNaseH and can therefore serve as a primer for plus-strand DNA synthesis. RT proceeds from the 3' end of the PPT towards the 5' end of the minus-strand DNA, and then 18 nt into the unwound tRNA primer (up to and including the pbs) to yield a species called *plus-strand strong stop DNA*. Further progression along the tRNA template is prohibited by a 1-methyladenine (m^1A) residue.

Once the 3' tail of the tRNA has been copied, RT degrades the tRNA primer in its entirety, thereby liberating the 5' end of the minus-strand DNA and enabling the plus-strand DNA to detach and reanneal with the opposing pbs. Following this process – known as *second-strand transfer* – both minus- and plus-strand DNA synthesis are completed in full. The resulting molecule has duplications of the U3-R-U5 sequences at either end, known as *long terminal repeats*, or LTRs. The 5' LTR will ultimately serve as a promoter for transcription of the integrated provirus.

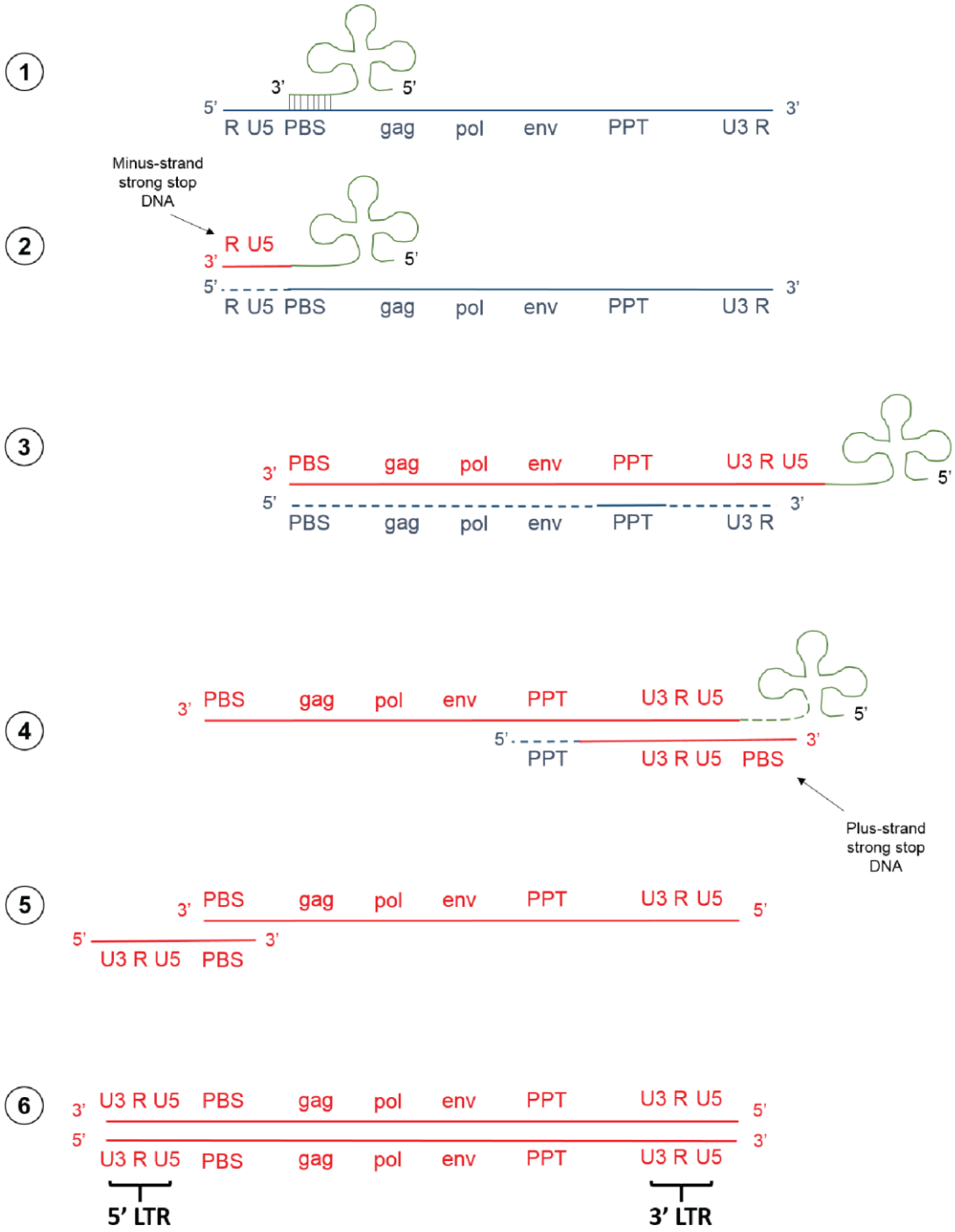


Figure 1.8: Reverse transcription

Viral genomic RNA (gRNA) is depicted in blue and nascent DNA in red. Dashed blue lines represent gRNA that has been degraded by RNase H. Steps depicted include (1) gRNA prior to reverse transcription; (2) synthesis of minus-strand strong stop DNA; (3) first-strand transfer; (4) synthesis of plus-strand strong stop DNA; (5) second-strand transfer; (6) completion of reverse transcription and the formation of LTRs.

Uncoating

Uncoating was classically defined as the complete dissociation of CA from the RTC shortly after viral entry (Aiken, 2006). However, the growing appreciation of the role that CA plays in later stages of the lifecycle – including reverse transcription (Forshey et al., 2002); shielding of viral nucleic acids from cytosolic DNA sensing (Gao et al., 2013; Lahaye et al., 2013); translocation across the nuclear membrane (Matreyek and Engelman, 2011; Matreyek et al., 2013) and targeting of proviral DNA to transcriptionally active sites in the genome (Koh et al., 2013) – has warranted a revision of this definition.

Although the exact timing of uncoating remains elusive, the fact that the HIV-1 core is about 20 nm too large to pass through a nuclear pore complex (Panté and Kann, 2002; Ganser-Pornillos et al., 2007), combined with the recent observation that some CA remains associated with the viral cDNA after nuclear entry (Peng et al., 2014; Hulme et al., 2015), indicates that it is likely to occur in two distinct phases: an initial loss of core integrity in the cytoplasm, followed by a complete dissociation of CA in the nucleus.

The precise mechanism through which this process occurs remains the subject of considerable ongoing research. Nevertheless, numerous hypotheses have been put forward in an attempt to reconcile the existing data, including mechanical stress imparted by the nascent viral cDNA during reverse transcription (Forshey et al., 2002), and a tug-of-war-like strain generated by the opposing microtubule motor proteins, dynein and kinesin-1 (Lukic et al., 2014; Pawlica and Berthoux, 2014). Uncoating within the nucleus is thought to be an active process mediated by transportin-3 (TNPO3) (Zhou et al., 2011). These models are supported by varying degrees of evidence and need not be mutually exclusive.

1.2.3 Nuclear trafficking and import

In parallel to reverse transcription and uncoating, the viral RTC must be trafficked towards the nucleus. To accomplish this, the virus exploits actin microfilaments for short-range movement near the cell periphery, and then microtubules (MTs) for the longer journey from the periphery to the nuclear membrane (Campbell and Hope, 2005); (Naghavi and Goff, 2007). Retroviruses specifically utilise stable MTs over their dynamic counterparts. These are typified by post-translational modifications such as detyrosination and acetylation, and are recognised by motor proteins as specialised tracks for long-range vesicle trafficking. This specific co-option of stable MTs explains the previously conflicting observation that HIV-1 is resistant to the pharmacological disruption of MT polymerisation (Sabo et al., 2013).

Upon completion of reverse transcription and partial uncoating, the resulting structure is known as the pre-integration complex (PIC). PICs are defined by their capacity for *in vitro* integration, and have offered a valuable system for the detailed characterisation of this process (Hansen et al., 1999). The HIV-1 PIC is trafficked to the nucleus by virtue of a number of viral karyophilic elements, including MA, IN and Vpr (Rivière et al., 2010). These components are central to the ability of HIV-1 and other lentiviruses to enter the nuclei of non-dividing cells. Other retroviruses, including MLV, depend on mitotic breakdown of the nuclear membrane for this process (Roe et al., 1993).

In order to enter the nucleus of a resting cell, the lentiviral PIC must harness components of the nuclear pore complex (NPC) (Fassati et al., 2003; Zaitseva et al., 2009; Yeung et al., 2009). In particular, Nup98, Nup153, Nup358 and TNPO3 are required for both nuclear import and proper trafficking within the nucleus: depletion of these factors shifts the distribution of integration sites from gene-dense to gene-poor regions (Schaller et al., 2011; Ocwieja et al., 2011; Di Nunzio et al., 2013). On the viral side of this interaction, CA is the major determinant for nuclear import and trafficking events within the nucleus. This is evidenced by the observation that the N57A and N74D mutations in HIV-1 CA cause a change in integration site preference from transcriptionally-active to -inactive regions,

phenocopying the knockdown of NPC components (Schaller et al., 2011; Ocwieja et al., 2011; Koh et al., 2013).

Nuclear import of the lentiviral PIC is a complex process involving a large complement of host proteins. During the early stages of infection, the viral core binds CPSF6 (Lee et al., 2012); this mediates interaction with the nuclear pore component, Nup358 (Bichel et al., 2013). Upon engagement of Nup358, the kinesin-1 motor protein KIF5B traffics the Nup358-bound core away from the nuclear pore (Dharan et al., 2016). Precisely how this facilitates nuclear import is currently unknown. Putative mechanisms include the reduction of PIC size by Nup358-mediated uncoating, as well as increased permeability of the nuclear membrane following the cytoplasmic relocalisation of this protein. Once the PIC is competent to access the nucleus, it is recognised as cargo by TNPO3, a member of the karyopherin-β family of nuclear transporters, and shuttled across the nuclear membrane (Chook and Süel, 2011).

On the nucleoplasmic side of the NPC, the PIC is recognised by another NPC component, Nup153 (Matreyek et al., 2013). Recent data from the Fassati lab suggests that this interaction maintains the integrity of the PIC (Chen et al., 2016), while CPSF6 targets it towards transcriptionally active regions of chromatin (Chin et al., 2015; Sowd et al., 2016). Once the PIC arrives at its genomic destination, TNPO3 triggers the displacement of any remaining CA (and any bound factors, including CPSF6), freeing the viral cDNA to interact with the necessary chromatin-tethering factors so that integration can ensue.

1.2.4 Integration

A defining step in the retroviral lifecycle is the integration of viral cDNA into the genome of an infected cell. However, nuclear entry does not guarantee successful integration. Viral DNA can undergo a number of circularisation reactions within the nucleoplasm that yield non-productive species and represent dead-ends for the virus (Farnet and Haseltine, 1991). These include 2-LTR circles, which occur when the cellular machinery ligates the viral DNA end-to-end (Li et al., 2001); 1-LTR

circles, resulting from homologous recombination between LTRs (Kilzer et al., 2003), and various autointegration products (Shoemaker et al., 1980).

However, the population of PICs that *do* remain integration-competent must be targeted and tethered to host chromatin. While the HIV-1 PIC associates with chromatin via the cellular co-factor LEDGF/p75 (Hombrouck et al., 2007), MLV relies on members of the chromatin-bound Bromodomain and Extra-Terminal (BET) family of proteins for this process (Gupta et al., 2013; Sharma et al., 2013; De Rijck et al., 2013). Mutagenesis and fluorescence microscopy have revealed that the virally-encoded p12 also contributes to the chromatin tethering of MLV PICs (Elis et al., 2012; Wight et al., 2012). These separate chromatin-targeting pathways result in distinct integration site preferences: while HIV-1 tends to integrate within the bodies of actively transcribed genes (Schroder et al., 2002), MLV integration is biased towards transcriptional start sites (Sharma et al., 2013).

Integration is catalysed by the virally-encoded integrase (IN) enzyme. IN functions as either an octamer (in α- and β-retroviruses) or a tetramer (in spumaviruses) and forms a complex with linear viral dsDNA known as the *intasome*. While chromatin-tethering factors are required to direct the intasome to the appropriate genomic location, the exact site of integration is governed by an interaction between residues in the CTD of IN and specific bases in the target DNA (Maertens et al., 2010; Serrao et al., 2014).

IN catalyses two sequential reactions that are necessary for integration (Figure **1.9**). First, it processes the 3' end of each strand of the viral DNA to reveal a conserved CA dinucleotide; this probably occurs within the cytoplasm (Fassati and Goff, 1999). Once the PIC has reached the site of integration, the second step – known as *strand transfer* – can proceed. In this step, the newly exposed hydroxyl groups on the viral DNA are used in the nucleophilic attack of a pair of phosphodiester bonds on the target DNA, and the 3' end of the viral DNA is simultaneously ligated to the 5' end of the host DNA (Vink et al., 1991). This results in a dinucleotide overhang and a single-stranded region of either 4 (MLV) or 5 (HIV-1) nucleotides on either side of the attack site, both of which are subsequently repaired by cellular enzymes.

Once integration is complete, the viral DNA is referred to as a *provirus* and the early phase of the lifecycle is complete.

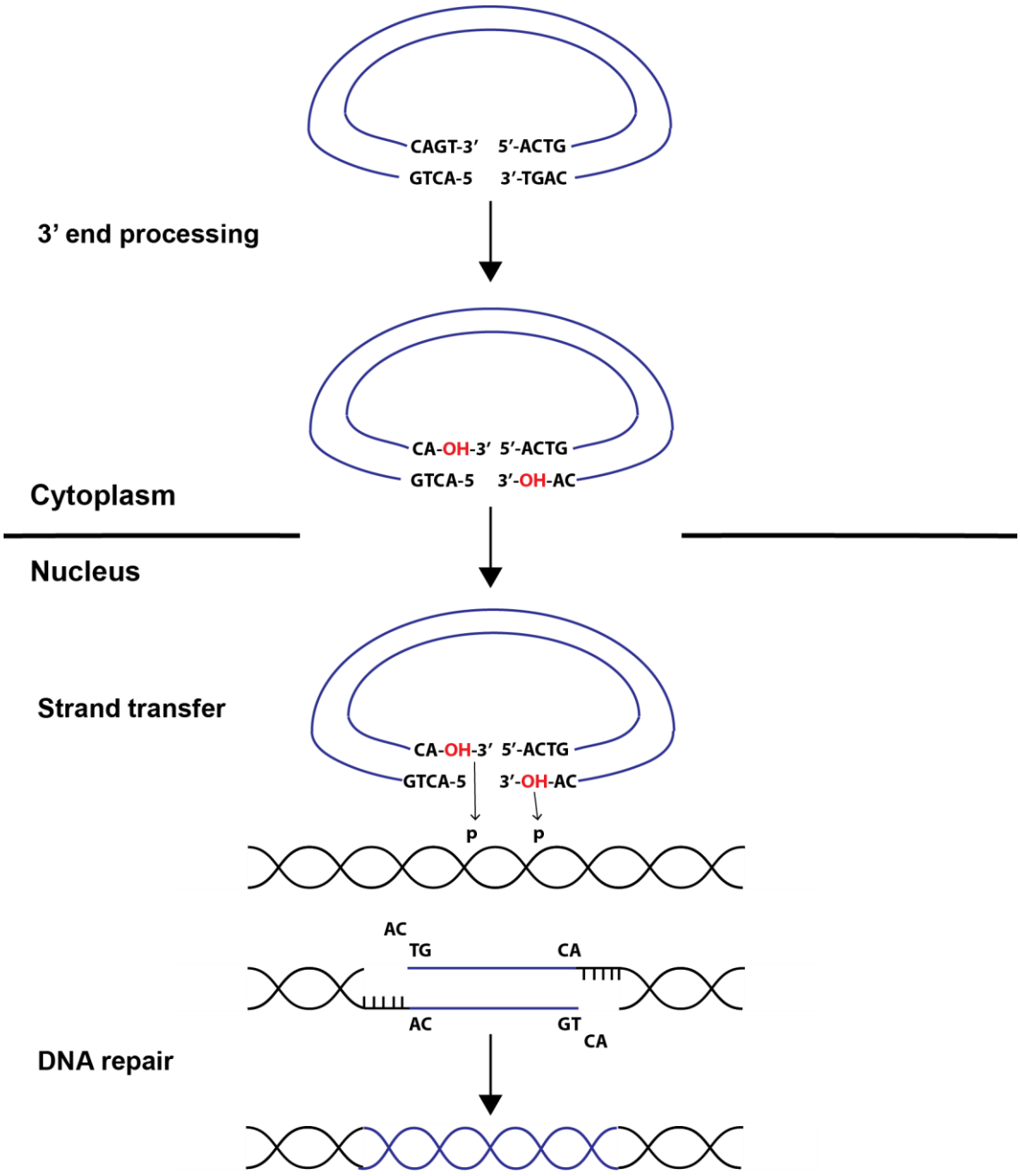


Figure 1.9: Processing and integration of viral cDNA

Adapted from Van Maele et al. (2006).

1.2.5 Transcription, splicing and nuclear export

Transcription of the provirus is typically initiated from a promoter sequence at the U3-R boundary of the 5' LTR. This element is usually sufficient to drive constitutive expression of the viral genome, although this is somewhat dependent on factors such as cell type and the exact site of integration (Feinstein et al., 1982). Rabson (1997) provides a detailed description of retroviral RNA synthesis; this section will offer only a brief overview of this process.

Transcription

Shortly following integration, a period of limited transcription from the HIV-1 provirus yields short, fully spliced mRNAs corresponding to the *tat*, *rev* and *nef* reading frames. Once sufficient Tat has been synthesised, it directs the transcription of longer, incompletely spliced mRNAs encoding *env, vif, vpr* and *vpu*, along with unspliced transcripts that serve both as a template for *gag-pol* translation, and as gRNA for progeny virions (Kim et al., 1989; Pomerantz et al., 1990).

Tat directs transcriptional transactivation by recruiting the transcription elongation factor, P-TEFb, which itself is a complex of cyclin T1 and Cdk9 (Wei et al., 1998). Tat binds the cyclin T1 component of this complex, as well as the trans-activation-responsive (TAR) region found at the 5' end of all viral transcripts. This brings Cdk9 in close proximity of the transcriptional machinery, enabling it to phosphorylate several residues within the C-terminal domain of the large subunit of RNA Polymerase II. These modifications substantially increase the processivity of the enzyme (Kim et al., 2002).

Transcription of the MLV provirus is less-well regulated, owing to the absence of a transcriptional trans-activator. Nevertheless, the U3 region contains a host of *cis*-regulatory sequences, including an E-box that binds basic helix-loop-helix (bHLH) transcription factors (Nielsen et al., 1992; Nielsen et al., 1994; Lawrenz-Smith and Thomas, 1995).

Splicing and nuclear export

Upon dissociation from the proviral template, all viral transcripts are modified with a 5' methylguanosine cap and a 3' polyA tail, and a subset are spliced to remove portions of the coding sequence. While MLV produces only unspliced and singly spliced mRNAs, the presence of twelve splice sites in the HIV-1 genome yields more than 40 different transcripts of varying abundance (Purcell and Martin, 1993).

As described on the previous page, HIV-1 transcription is temporally regulated. During the first wave of transcription, only completely spliced RNAs corresponding to *tat, rev* and *nef* are maintained, while any unspliced or singly spliced species are degraded upon synthesis.

While Tat is crucial for boosting subsequent levels of transcription, Rev serves to control the export of mRNAs once they are synthesised. Thus, as Rev protein levels accumulate with successive rounds of early transcription, longer mRNAs are eventually recognised and targeted for export. This occurs by virtue of a Rev-response element (RRE), which is only present in singly- and unspliced viral transcripts, and a nuclear export signal (NES) found roughly in the middle Rev (Fischer et al., 1995). The NES facilitates an interaction between the Rev-RNA complex and the karyopherin, Crm1, permitting the active transport of viral mRNA into the cytoplasm. Rev can then return to the nucleoplasm through an interaction between its N-terminal NLS and importin-β (Henderson and Percipalle, 1997).

Again, MLV lacks an equivalent *trans*-acting accessory protein. However, it does contain a *cis*-acting cytoplasmic accumulation element (CAE), which is found towards the 3' end of *pol* and mediates the export of transcripts through association with the nuclear export receptor, NXF1 (Sakuma et al., 2014). Consistent with this notion is the observation that inserting a CAE into the genome of HIV-1 facilitates the Rev-independent expression of Gag. Additionally, the 3' U3 region of the MLV genome appears to be required for the export of full-length transcripts (Volkova et al., 2014).

1.2.6 Translation

The synthesis of viral proteins is initiated upon recognition of the 5' methylguanosine cap – which is present on all viral transcripts – by the ribosome.

The *gag* reading frame encodes the structural proteins of the virus. In orthoretroviruses, its translation yields a polyprotein that is ultimately cleaved into matrix (MA), capsid (CA) and nucleocapsid (NC) subunits, along with a number of smaller peptides. As discussed in Section **1.1.6**, spumaviral Gag has a rather different structure and remains largely uncleaved throughout the viral lifecycle. In MLV, a heavily glycosylated form of this protein – glycoGag – is synthesised by translational initiation from an upstream CUG codon (Edwards and Fan, 1979); (Prats et al., 1989). This yields a protein with an 88 aa leader sequence appended to the N-terminus, which directs it to the Golgi apparatus for the addition of carbohydrate moieties. GlycoGag confers viral resistance to the SERINC3/5 and APOBEC restriction factors (see Sections **1.3.2** and **1.3.3**, respectively) (Stavrou et al., 2013; Rosa et al., 2015, Usami et al., 2015). Additionally, the Gag polyprotein of both MLV and HIV-1 is co-translationally myristoylated to facilitate association with the plasma membrane during assembly (Henderson et al., 1983; Bryant and Ratner, 1990).

In orthoretroviruses, the *pol* reading frame – including protease (PR), reverse transcriptase (RT) and integrase (IN) – is translated as part of a Gag-Pol polyprotein. This fusion is synthesised by translational readthrough of the *gag* stop codon, which is mediated by a pseudoknot structure in the transcripts of MLV (Wills et al., 1994) and by a -1 ribosomal frameshift in HIV-1 (Jacks et al., 1988). In both cases, readthrough occurs with approximately 5% efficiency. This phenomenon maintains a Gag to Gag-Pol ratio of approximately 20:1 in the producer cell, a ratio that is conserved among retroviruses and is important for late-phase events such as genome dimerisation and proteolytic processing (Shehu-Xhilaga et al., 2001). In spumaviruses, on the other hand, Pol is translated independently of Gag.

In all retroviruses, Env is synthesised from a spliced mRNA that lacks the former two reading frames. Translation of the HIV-1 Env precursor (gp160) occurs on the

rough endoplasmic reticulum, where it undergoes co-translational glycosylation of several key residues and post-translational cleavage of the N-terminal signal peptide that got it there. Upon entering the oxidising environment of the ER lumen, disulphide bonds form between the nascent gp160 monomers, facilitating their association into trimers (Earl et al., 1991). The trimeric precursor is then translocated to the Golgi apparatus, where it is cleaved into surface (SU; gp120) and transmembrane (TM; gp41) subunits by resident subtilisin-like endoproteases (Hallenberger et al., 1997b; Willey et al., 1988). This cleavage event weakens, but maintains, the Env trimer.

In addition to *gag, pol* and *env*, complex retroviruses possess additional ORFs encoding regulatory and accessory proteins. In HIV-1, these are *vif* and *vpr*, which are expressed from singly spliced transcripts; *tat, rev* and *nef*, from multiply spliced transcripts; and *vpu*, which is encoded on a bicistronic mRNA upstream of *env* (Schwartz et al., 1990).

1.2.7 Assembly

Once the full complement of structural and catalytic proteins have been synthesised, a complex orchestration of events leads to the assembly of immature virions.

Early interactions between the NC domain of Gag and the packaging signal (ψ) of unspliced viral transcripts facilitates the specific encapsidation of viral genomes, while excluding all cellular and spliced viral RNA species. MLV ψ consists of four stem-loops – two for packaging (SL-C/D) (D'Souza et al., 2001), and two for the initiation of genome dimerisation (DIS-1/2) (Ly and Parslow, 2002). Conversely, ψ is poorly defined in HIV-1, with a variety of mutations in the 5' UTR resulting in a reduction in packaging efficiency. Indeed, it has been postulated that a 'conformational switch' in this region shifts the RNA from a structure that is conducive to translation to one that promotes packaging (Lu et al., 2011b).

Following this initial encounter between Gag and gRNA, assembly of the viral core can proceed. For both MLV and HIV-1, this requires association with the plasma membrane via the myristoylated MA domain of Gag. Interestingly, this change in localisation reconfigures the Gag-RNA interaction from one involving primarily the 5' UTR of the latter, to one that spans the entire length of the genome (Kutluay et al., 2014).

The MA domain specifically targets the assembling viral complex to lipid rafts: domains within the membrane that are enriched for certain classes of lipid, including cholesterol, phosphatidyl serine and phosphatidyl inositol (4,5) bisphosphate (PIP_2) (Chan et al., 2008). While the incorporation of cholesterol is critical for virion infectivity (Liao et al., 2003), acidic phospholipids such as PIP_2 are required to form electrostatic interactions with basic residues in MA (Ono et al., 2004), to externalise the covalently attached myristate residue (Saad et al., 2006), and to competitively displace any RNA bound to this region (Chukkapalli et al., 2010). Collectively, these phenomena ensure tight anchoring of Gag to the plasma membrane. Once the viral complex is properly tethered, Gag polymerisation ensues. This yields an immature hexagonal lattice of 1100-1800 Gag molecules for MLV (Yeager et al., 1998) and 5000 molecules for HIV-1 (Briggs et al., 2004), collectively accounting for about 50% of total virion mass.

The curvature and flexibility of this lattice have posed substantial technical obstacles to defining its structure. However, a combination of electron microscopy and cryotomography recently accomplished this for the immature lattice of a truncated Gag from Mason-Pfizer monkey virus, a β-retrovirus (Bharat et al., 2012). Interestingly, this work revealed contacts between the unprocessed CA domains that are distinct from those found in the mature retroviral core. It is presently unclear whether the transition from an immature lattice to a condensed core involves an intermediate phase of Gag disassembly (Keller et al., 2013; Frank et al., 2015).

Env trimers are targeted to lipid rafts independently of Gag. Genetic evidence suggests that this is mediated by the cytoplasmic TM subunit of Env, which both directs it to the plasma membrane and ensures its incorporation into virions through

an interaction with MA (Yu et al., 1993; Cosson, 1996; Murakami and Freed, 2000). Interestingly, however, neither deletion of the TM subunit (Einfeld, 1996), nor the replacement of Env with a heterologous glycoprotein (Briggs et al., 2003), prevent the incorporation of Env in virions, indicating that a direct interaction between Env and MA is not essential for this process. Thus, a compensatory pathway must also exist, which may involve the recruitment of cellular proteins to act as a bridge between Env and MA (Checkley et al., 2011; Tedbury and Freed, 2014). In any case, the trimerisation of MA is probably important for Env incorporation (Tedbury et al., 2013).

The overall kinetics of virion assembly – from the initial detection of membrane-associated Gag to the formation of an immature lattice – have been estimated at 5-9 min on average, with an upper limit of around 20 min (Jouvenet et al., 2008; Ivanchenko et al., 2009).

1.2.8 Budding and maturation

Once all of the viral components have assembled at the plasma membrane, the immature virion must egress to the cell surface. This process is initiated by the polymerisation of Gag at the assembly site, which deforms the planar lipid bilayer into a spherical, virion-like configuration (Carlson et al., 2008). The resulting structure remains attached to the producer cell through a neck region. Complete membrane fission requires that this region be severed by components of the cellular ESCRT machinery.

The ESCRT pathway involves numerous protein complexes whose usual roles include cell abscission during cytokinesis and the biogenesis of multivesicular bodies (Hurley and Hanson, 2010). Retroviruses possess late assembly (L-) domains for the recruitment of these proteins. L-domains typically come in three flavours: P(T/S)AP, which is found in MLV MA and HIV-1 p6 (Garrus et al., 2001; Segura-Morales et al., 2005); LYPx_nL, which is found at the MLV MA-p12 junction and in HIV-1 p6 (Segura-Morales et al., 2005; Strack et al., 2003), and PPxY, which is present in MLV p12, but absent in HIV-1 (Yuan et al., 2000).

The viral L-domains recruit either the ESCRT-I heterotetramer or the ESCRT-associated protein ALIX, which themselves activate ESCRT-III protein complexes, including CHMP2 and 4 (Wollert and Hurley, 2010). These oligomers form long filamentous structures around the bud neck, which coil into a dome shape at the base of the virion (Shen et al., 2014). It is possible that this process is completed upon hydrolysis of ATP by the VPS4 complex (Babst et al., 1998; Scott et al., 2005), although this remains the subject of ongoing investigation. Alternatively, recent data from super-resolution microscopy suggests that ESCRT components are recruited within the head of the budding virion, where their selective remodelling might act as the driving force for membrane scission (Van Engelenburg et al., 2014).

Immediately following its release from the producer cell, the nascent virion undergoes a series of tightly regulated morphological changes that are necessary for subsequent infectivity. This is known as *maturation* (Kohl et al., 1988; Konvalinka et al., 1995a).

This process begins with the autocatalytic cleavage of PR from Pol shortly after the completion of budding. Immediately following the autoprocessing of PR, RT and IN are cleaved from one another, although it remains unclear whether this phenomenon occurs in *cis*, *trans*, or some combination thereof (Pettit et al., 2004). PR then catalyses sequential cleavage events within Gag, liberating each of its constituents in a stepwise manner. This temporal regulation is achieved by different rates of processing at each of the five Gag cleavage sites (Figure **1.10**). In HIV-1, these sites fall into three categories according to their rate of processing by PR: rapid (SP1/NC), intermediate (SP2/p6, MA/CA) and slow (NC/SP2, CA/SP1). By examining mutant virions that are defective for cleavage at each of these sites, it has been possible to deduce their respective contributions to maturation.

SP1/NC cleavage activates the fusogenic potential of Env (Wyma et al., 2004) and promotes the condensation of NC with the viral RNA (de Marco et al., 2010b). The resulting RNP particle is further processed by the SP2/p6 and NC/SP2 cleavages, which collectively free NC to act as a chaperone for the genome dimer (Kafaie et

al., 2008). Cleavage at the MA/CA junction liberates CA-SP1 from the membrane and causes the immature Gag lattice to disassemble; CA/SP1 cleavage then frees CA entirely, enabling it to form the conical structure that typifies the core of HIV-1 (de Marco et al., 2010b).

In addition to unshackling the individual domains of Gag, these events facilitate conformational changes within them that allow new interactions to be established. For example, the liberation of CA from MA enables the N-terminal 13 residues of the former to fold into a β-hairpin, thereby initiating the formation of mature CA hexamers.

The maturation process is less-well characterised for MLV virions, although it is known that cleavage between MA and p12 is not required for the formation of mature cores (Oshima et al., 2004). Cleavage at the p12/CA site, however, is crucial for the subsequent round of reverse transcription and integration (Rulli Jr et al., 2006).

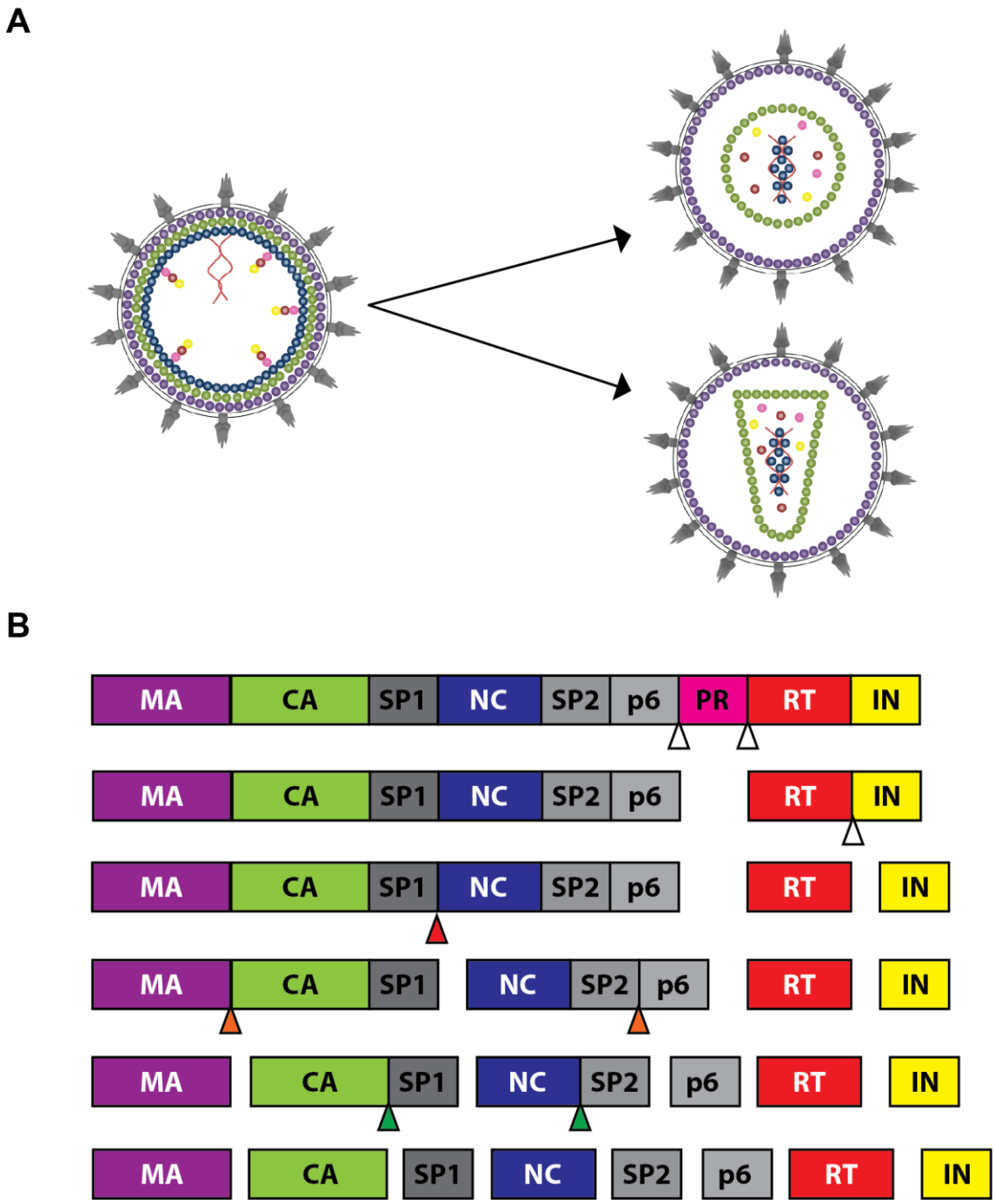


Figure 1.10: Retroviral maturation

(**A**) A schematic depicting the structural rearrangements that convert an immature virion (left) to a mature particle with either spherical (top right) or conical (bottom right) core morphology. Env trimers are shown in grey; MA, purple; CA, green; NC, blue; PR, pink; RT, red and IN, yellow. (**B**) Sequential, PR-mediated proteolysis of the Gag-Pol precursor. White triangles represent cleavage sites in Pol; red triangles, rapid cleavage sites in Gag; orange triangles, intermediate cleavage sites in Gag, and green triangles, slow cleavage sites in Gag.

1.2.9 Unique aspects of the *Spumaretrovirinae* lifecycle

The processes described in this chapter so far apply predominantly to viruses of the *Orthoretrovirinae* subfamily. The lifecycle of the spumaviruses is distinct in a number of ways, some of which unite them more closely with the *Hepadnaviridae*, a family of DNA viruses that includes hepatitis B virus (HBV). Some of these differences are highlighted in Table **1.2**.

	Orthoretrovirinae	*Spumaretrovirinae*	*Hepadnaviridae*
Reverse transcription	Occurs during early phase of the lifecycle; infectious particles contain RNA.	Occurs during late phase of the lifecycle; infectious particles contain DNA.[1]	Occurs during late phase of the lifecycle; infectious particles contain DNA.
Nuclear entry	Virions are not recycled intracellularly.	Intracellular 'recycling' of virions yields high copy-number in nucleus of the producer cell.[2]	Intracellular 'recycling' of virions yields high copy-number in nucleus of the producer cell.[3]
Integration	Essential for replication.	Essential for replication.[4]	Does not occur.
Transcription	Initiated from a single promoter.	Initiated from multiple promoters.[5]	Initiated from multiple promoters.[6]
Assembly	Gag binds RNA via NC domain.	Gag binds DNA and RNA with equal affinity.[7]	HBV core protein binds DNA and RNA with equal affinity.[8]
Budding & maturation	Budding occurs at the plasma membrane and maturation occurs extracellularly.	Both processes occur within intracellular compartments such as the ER.[9]	Both processes occur within intracellular compartments such as the ER.[10]

Table 1.2: A comparison of orthoretroviruses, FVs and hepadnaviruses

Orthoretroviral phenotypes are shaded in pink and hepadnaviral ones in blue. References: (1) Moebes et al. (1997); (2) Meiering et al. (2000); (3) Tuttleman et al. (1986); (4) Enssle et al. (1999); (5) Löchelt et al. (1993); (6) McLachlan (1991); (7) Yu et al. (1996c); (8) Nassal (1992); (9) Goepfert et al. (1999); (10) Ganem (1991).

1.3 Retroviral restriction factors

Restriction factors are host-expressed proteins which form the basis of a cell-intrinsic antiviral response (Figure **1.11**) (Yan and Chen, 2012). Intrinsic immunity is distinct from the innate and adaptive arms in that the proteins responsible are usually constitutively expressed, although they may be upregulated by interferon (Tanaka et al., 2006; Carthagena et al., 2009; Abdel-Mohsen et al., 2014). This chapter will describe each of the major retroviral restriction factors in turn, beginning with those that inhibit viral entry.

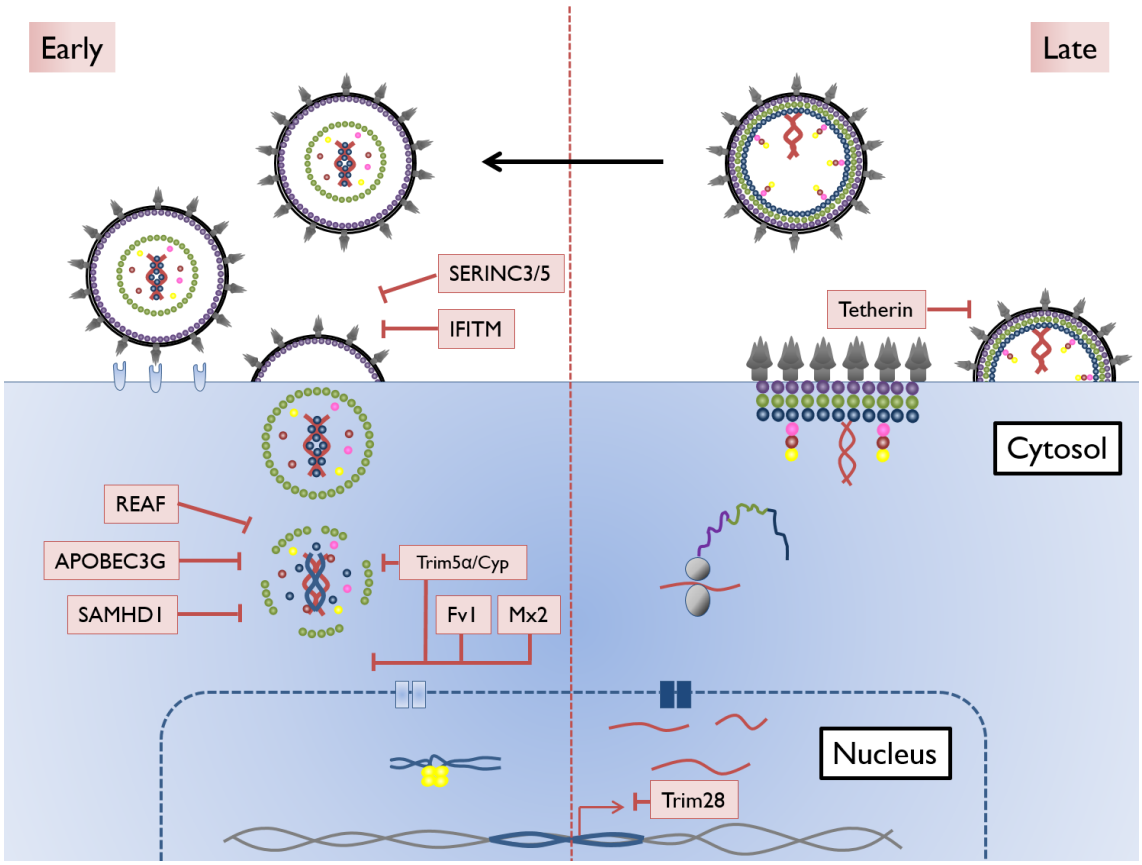


Figure 1.11: The retroviral lifecycle, illustrating the stage-specific blocks imposed by restriction factors

1.3.1 IFITMs

The interferon-inducible transmembrane (IFITM) family of proteins comprises five members, of which three – IFITM1, 2 and 3 – are induced by both type I and II interferon. These proteins are known to restrict the entry of a number of enveloped viruses, including influenza A virus, hepatitis C virus and HIV-1 (Brass et al., 2009; Lu et al., 2011a; Narayana et al., 2015).

The IFITMs consist of a cytosolic N-terminus, followed by two hydrophobic intramembrane domains separated by a conserved intracellular loop. IFITM1, 2 and 3 also undergo post-translational modifications of several residues, some of which are imperative for restriction (Perreira et al., 2013). The IFITMs block viral entry by preventing fusion between the host and viral membranes (Li et al., 2013b). One model for how they accomplish this is by reducing membrane fluidity, thereby preventing the changes in membrane curvature that are requisite for a hemifusion event.

1.3.2 SERINC3/5

Serine incorporator 3 and 5 (SERINC3 & 5) belong to a family of transmembrane carrier proteins that facilitate the incorporation of serine into lipids of the plasma membrane (Inuzuka et al., 2005). Their role in the Env-dependent restriction of HIV-1 was independently reported by two groups, both of whom were seeking to identify the mechanism through which HIV-1 Nef and the structurally distinct MLV glycoGag are able to enhance the infectivity of HIV-1 virions (Rosa et al., 2015; Usami et al., 2015).

The two groups discovered that Nef and glycoGag function by downregulating SERINC3 and 5 from the plasma membrane. This prevents their incorporation into budding virions, where they have an inhibitory effect on replication. A number of mechanisms for the restriction of HIV-1 by SERINCs have been postulated, including physical hindrance of Env trimer clustering and increasing the energy barrier for membrane fusion (Usami et al., 2015). It is also possible that these

proteins compromise infectivity by altering the lipid composition of the envelope (Waheed and Freed, 2010).

Although little is presently known about the mechanism through which Nef overcomes SERINC-mediated restriction, a preliminary observation suggested that it may target the latter to Rab7+ endosomal compartments for eventual degradation by the lysosome (Rosa et al., 2015). However, more recent findings have shown that the downregulation of SERINC by Nef is dispensable and insufficient for antagonism. This indicates that virion exclusion is not the only mechanism through which Nef counteracts SERINC-mediated restriction. (Trautz et al., 2016)

1.3.3 APOBEC family members

The APOBEC family of proteins comprises eleven members, all of which have cytosine deaminase activity. The substrate for these enzymes is typically single-stranded DNA, with a requirement for three bases preceding the target cytosine and a single one following it (Nabel et al., 2013). A conserved glutamic acid in the catalytic site promotes the hydrolytic deamination of cytosine by deprotonating water, thereby providing a hydroxide ion that can be used for nucleophilic attack of the pyrimidine ring. The net result is the replacement of the 4' amine with a carbonyl group – i.e. the conversion of cytosine to uracil.

The APOBEC family has a colourful evolutionary history, evidenced by the extensive variation in gene copy number among mammals (Conticello et al., 2005). Although all mammals possess activation-induced cytidine deaminase (AID) and APOBECs 1, 2 and 3, there is evidence of particular gene expansion among the primates, where tandem duplications in the APOBEC3 locus have yielded a total of seven protein-coding genes (A3A-H).

While it has been known for some time that the APOBEC family member *AID* contributes to adaptive immunity by promoting antibody diversity (Muramatsu et al., 1999), work over the past decade has uncovered numerous other members that contribute to retroviral restriction. APOBEC3G (A3G herein) was initially discovered as a dominant cellular factor that blocks the replication of Vif-deficient

HIV-1 in culture (Madani and Kabat, 1998; Sheehy et al., 2002). Later studies revealed that a number of other proteins in the A3 subfamily – including A3D, A3F and A3H – harbour similar activities (Hultquist et al., 2011; Ooms et al., 2013).

A3 enzymes are incorporated into budding virions in the producer cell via an RNA-bridged interaction with the NC domain of Gag (Apolonia et al., 2015). Shortly after budding, a substantial fraction of the incorporated restriction factor then becomes encapsidated in the viral core (Donahue et al., 2015).

Restriction itself is manifested in the target cell, specifically during reverse transcription. In the early stages of RT, degradation of the genomic RNA yields a single-stranded cDNA that is susceptible to modification by the encapsidated A3 enzyme. Deamination of this molecule provides an aberrant template for plus-strand synthesis, resulting in a high frequency of G-A mutations in the nascent proviral DNA (Harris et al., 2003; Mangeat et al., 2003).

A3 enzymes can also inhibit the elongation of reverse transcripts in a deamination-independent manner, by associating with viral ssRNA and imposing a physical barrier to the progression of reverse transcriptase along the template (Iwatani et al., 2007; Bishop et al., 2008). This steric hindrance requires a shift in the binding kinetics of the enzyme, from a fast off-rate (to permit rapid scanning for cytosine residues) to a much slower one (to pose a block to RT). This transition is mediated by a gradual oligomerisation of A3 (Chaurasiya et al., 2014), and is utilised to a greater extent by A3F than by A3G (Kobayashi et al., 2014). A3-mediated restriction of other retroviruses – including MLV, HTLV-1 and various foamy viruses – has also been reported (Delebecque et al., 2006; Rulli et al., 2008; Ooms et al., 2012).

Unsurprisingly, retroviruses have evolved various means of counteracting restriction by A3 proteins. The HIV-1 viral infectivity factor (Vif) is an accessory protein that renders the virus resistant to restriction by A3. Specifically, Vif recruits the Cullin5-ElonginB-ElonginC complex, which facilitates the polyubiquitination and proteasomal degradation of the restriction factor (Conticello et al., 2003). MLV antagonises A3 by way of a heavily glycosylated form of Gag (glycoGag), which

confers resistance by excluding the enzyme from incorporation into virions (Boi et al., 2014). Meanwhile, foamy viruses counter restriction using the accessory protein, Bet, which appears to employ a mechanism that is distinct from both Vif and glycoGag (Chareza et al., 2012).

The pressure to evade such viral countermeasures has resulted in numerous positively selected residues within genes of the A3 subfamily (Sawyer et al., 2004; Duggal et al., 2013). Interestingly, some of these predate the emergence of modern lentiviruses, implying that an ancient – and perhaps still extant – selective force has helped to shape the evolution of this protein.

1.3.4 SAMHD1

Sterile alpha motif (SAM) and histidine/aspartic acid (HD) domain-containing protein 1 (SAMHD1) is a restriction factor that inhibits HIV-1 and SIV replication in cells of the dendritic and myeloid lineages, and in resting $CD4^+$ T-cells (Laguette et al., 2011; Baldauf et al., 2012). The HD domain of the protein is a deoxynucleoside triphosphate triphosphohydrolase (dNTPase), which cleaves its substrate at the phosphoester bond between the α-phosphate and the 5' carbon of the ribose ring, thereby liberating the triphosphate moiety from its deoxynucleoside partner (Goldstone et al., 2011). This depletes the cellular pool of dNTPs, thus preventing the restricted virus from completing reverse transcription (Lahouassa et al., 2012).

The active form of SAMHD1 is a homotetramer with a total of 4 catalytic and 8 allosteric sites. Tetramerisation requires the C-terminal domain (CtD, residues 584-629) and is dependent on allosteric activation by (d)GTP. This induces conformational changes within the subunits that promote and stabilise their interaction with one another (Powell et al., 2011; Amie et al., 2013). A structure of the SAMHD1 tetramer in complex with dGTP has been solved to 1.8 Å resolution (Zhu et al., 2013).

While SAMHD1 limits the dNTP pool in differentiated and resting cells, it is inhibited in cycling cells to maintain sufficient dNTP concentrations for DNA synthesis. The inactivation of SAMHD1 occurs post-translationally through the phosphorylation of

T592 by the cyclin A2/CDK1 complex (Tang et al., 2015). This modification destabilises the tetramer by creating an electrostatic repulsion between the subunits, thereby inhibiting dNTPase activity. This has a knock-on effect on restriction, partly explaining why cycling $CD4^+$ T-cells are susceptible to HIV-1 infection despite sustained expression of SAMHD1 (Baldauf et al., 2012). Interestingly, T592 phosphorylation is reduced in response to IFN-I, suggesting that viral threat can directly result in the heightened activation of this restriction factor (Cribier et al., 2013).

In order to overcome SAMHD1 restriction in non-cycling cells, some retroviruses have evolved specialised countermeasures. The HIV-2/SIVmac accessory protein, *Vpx* is one such example. Vpx interacts with both the CtD of SAMHD1 and the Cullin4A complex (via the adaptor protein DCAF1) in order to promote the proteasomal degradation of the former (Schwefel et al., 2015). The SAMHD1-CtD thus exhibits strong signatures of positive selection (Laguette et al., 2012), the hallmark of an ongoing evolutionary arms race between a restriction factor and its viral antagonist.

1.3.5 REAF

RNA-associated early-stage antiviral factor (REAF) is a recently discovered restriction factor that acts in the early phase of the retroviral lifecycle, either during or immediately following the initiation of reverse transcription (Marno et al., 2014). The protein was first identified in a genome-wide siRNA screen (Liu et al., 2011a), where its knockdown in HeLa cells was found to result in >50-fold rescue of viral titres.

REAF has been shown to restrict HIV-1, HIV-2 and a number of SIVs, implying that its viral target is conserved at least among the primate lentiviruses. Although the mechanism of restriction is presently unclear, one hypothesis suggests that REAF associates with viral reverse transcription products and targets them for degradation (Marno et al., 2014). However, whether this interaction occurs directly or via an adaptor molecule remains to be seen.

Given the general observation that HIV-1 and other retroviruses typically evolve means of evading restriction factors, it is likely that such a countermeasure also exists for REAF. Consistent with this notion, REAF protein levels have been shown to decrease as soon as 1 h after viral challenge in HeLa cells (Marno et al., 2014), a phenotype that is lost with the addition of proteasome inhibitors. This implies that HIV-1 targets REAF for proteasomal degradation in a fashion that mirrors the retroviral antagonism of A3G and SAMHD1.

1.3.6 Capsid-targeting restriction factors: Fv1, T5α, TCyp and Mx2

The history of research on capsid-binding restriction factors can be traced to the discovery of the murine restriction factor, *Fv1*, a dominant gene that confers mice with resistance to MLV (Lilly, 1970; Rowe and Hartley, 1972; Odaka, 1975). More than twenty years following its discovery, the *Fv1* gene was cloned and sequenced by a positional cloning method (Best et al., 1996), and a sequence alignment revealed that it is derived from the *Gag* sequence of an endogenous retrovirus of the MERV-L family (Stoye, 1998).

The gene exists in two primary alleles, $Fv1^n$ and $Fv1^b$, with the former conferring resistance to B-tropic MLV and the latter usually to N-tropic, although it is modestly active against B-tropic virus when expressed above endogenous levels (Hartley et al., 1970; Bishop et al., 2001). The determinant responsible for the differential sensitivity of MLVs to alleles of *Fv1* was later mapped to a single amino acid (residue 110) in the MLV capsid (Kozak and Chakraborti, 1996), although additional determinants have since been identified (Stevens et al., 2004; Ohkura and Stoye, 2013).

Fv1 restricts MLV at a stage after reverse transcription but prior to genomic integration (Jolicoeur and Baltimore, 1976; Sveda and Soeiro, 1976). It does this, in part, by binding to the MLV capsid via its C-terminal recognition domain (Bishop et al., 2001). An exact mechanism of restriction is yet to be elucidated, but it is known that self-association of Fv1 via an N-terminal dimerisation domain is a prerequisite for restriction (Bishop et al., 2006; Yap et al., 2007). It is also evident

that the capsid target must exist in its mature and multimeric form for a successful Fv1 interaction to occur (Dodding et al., 2005).

A recent study characterising the restriction profiles of Fv1 orthologues from wild mice found that some are able to target non-MLV retroviruses, including lenti- and foamy viruses (Yap et al., 2014). This work revealed several novel determinants of restriction in the C-terminus of Fv1, and reinforced the notion that this restriction factor has been defending mice against episodes of retroviral challenge since its endogenisation, at least 4 million years ago (Yan et al., 2009).

Although Fv1 is a murine-specific factor, an Fv1-like activity became apparent in various primate cell lines in the late 1990s. These cells restricted N-, but not B-MLV, and, like Fv1, the factor responsible appeared to be both dominant and saturable, although its effects were manifested prior to reverse transcription. This activity was provisionally referred to as *Ref1* (Towers et al., 2000, 2002). Around the same time, several primate cell lines had been reported to block the replication of lentiviruses such as HIV-1 and SIVmac (Hofmann et al., 1999; Besnier et al., 2002; Cowan et al., 2002; Hatziioannou et al., 2003). Again, this phenotype was mediated by a factor that was dominant, saturable and acted pre-RT; it was denoted *Lv1*.

It was eventually discovered that a single protein was responsible for both of these phenomena. This was the alpha isoform of the *T5* gene (T5α herein). T5α was identified by screening a rhesus macaque cDNA library for genes that conferred human cells with resistance to HIV-1 infection (Stremlau et al., 2004). It has since been discovered that T5α is also capable of restricting N-MLV and a panel of lenti- and foamy viruses (Perron et al., 2004; Yap et al., 2004; Saenz et al., 2005; Yap et al., 2008).

A subsequent screen by the Luban group sought to identify the determinant for HIV-1 restriction in owl monkey kidney (OMK) cells (Sayah et al., 2004). Given former evidence that *Lv1* activity in OMK cells could be abolished using cyclosporine, a competitive inhibitor of CypA, the OMK cDNA library was screened for clones with homology to CypA. This screen yielded an unexpected result: a

fusion product combining the *T5* and *CypA* open reading frames, which is believed to have arisen via a LINE-I-mediated retrotransposition event that inserted a *CypA* cDNA into intron 7 of the *T5* gene. This fusion protein – TCyp – confers permissive human cells with anti-HIV-1 activity (Nisole et al., 2004). Such fusions have since been identified in several primate lineages (Brennan et al., 2008; Newman et al., 2008; Dietrich et al., 2010; Yu et al., 2013). A more thorough treatment of both T5α and TCyp can be found in Section **1.4.5.**

More recently, myxovirus resistance 2 (Mx2) was identified as an interferon-inducible factor with restriction activity against HIV-1 (Goujon et al., 2013; Kane et al., 2013; Liu et al., 2013). Mx2 interacts directly with the HIV-1 capsid via its N-terminal domain (Kong et al., 2014), and then effects restriction by stabilising the core, thus preventing uncoating (Fricke et al., 2014). The N-terminal domain is both necessary and sufficient for capsid recognition, as is evidenced by the observation that fusing the N-terminal 91 amino acids of Mx2 to $Fv1^b$ yields a chimera with potent anti-HIV-1 activity (Goujon et al., 2015). The C-terminal domain of Mx2 is responsible for dimerisation of the protein, a function that is also indispensable for its restriction activity (Fricke et al., 2014; Goujon et al., 2015).

Recent work has uncovered circulating variants of HIV-1 with capsid mutations that confer resistance to Mx2 (Wei et al., 2016). Correspondingly, sites of positive selection have been identified within Mx2 (Busnadiego et al., 2014; Sironi et al., 2014), implying an ongoing evolutionary arms race between the two.

1.3.7 TRIM28

TRIM28 (T28 herein) is a restriction factor that targets the integrated provirus of MLV and induces its transcriptional silencing (Wolf et al., 2008). T28 is recruited to a sequence element within the proviral DNA which closely overlaps with the conserved $tRNA^{pro}$ primer-binding site of the ecotropic MLV genome. This is mediated by members of the Krüppel-associated box (KRAB) family of zinc finger DNA-binding proteins, including ZPF809 (Wolf and Goff, 2009). Upon recruitment to the proviral DNA, T28 itself recruits a complex of chromatin modifiers that serve to repress local transcription, including a histone deacetylase complex, a histone

H3K9 methyltransferase, and the heterochromatin-associated protein, HP1 (Schultz et al., 2001, 2002).

T28 has also been shown to repress the transcription of endogenous retroviruses in neural progenitor cells (Fasching et al., 2015), a process that may be critical for proper brain development.

1.3.8 Tetherin

Tetherin (also known as bone marrow stromal antigen 2, BST-2) is an interferon-inducible restriction factor that causes fully formed virions to be retained on the surface of infected cells (Neil et al., 2008). It is a type II transmembrane protein with unusual topology, comprising an N-terminal transmembrane domain and a C-terminal glycophosphatidylinositol (GPI) anchor, linked by an extracellular coiled-coil motif (Hinz et al., 2010; Schubert et al., 2010). When a virion buds from a producer cell, one of the lipid anchors is incorporated into the viral membrane. This results in a physical link between cell and virus that prohibits the latter from being disseminated (Neil et al., 2008).

Unsurprisingly, lentiviruses have evolved numerous countermeasures to antagonise the activity of tetherin, including HIV-1 Vpu, HIV-2 Env and SIV Nef (Jia et al., 2009; Hauser et al., 2010). In the case of HIV-1 and -2, this involves sequestration of tetherin from the plasma membrane to a perinuclear compartment; less is known about the mechanism utilised by SIV Nef, although it is clear that the process is both clathrin-dependent (Serra-Moreno et al., 2013) and species-specific (Zhang et al., 2009).

1.4 The TRIM family

The tripartite motif (TRIM) family has at least 100 representatives in the human proteome, with functions as diverse as transcriptional regulation (Cammas et al., 2012); cell cycle control (Bell et al., 2013); embryonic development (Song et al., 2011) and apoptosis (Bernardi and Pandolfi, 2003). The family is defined by its tripartite N-terminus, which comprises a **R**ING finger, either one or two **B**-boxes, and a **c**oiled-**c**oil region (otherwise known as the *RBCC domain*). These modules appear in all family members with conserved order and spacing.

While TRIM proteins are united by their N-terminus, they diverge significantly at the C-terminus, where a number of domains and combinations thereof have been reported (Ozato et al., 2008). This has provided the basis for a classification system that divides human TRIMs into nine classes (Short and Cox, 2006) (Figure **1.12**). Interestingly, this system appears to have a phylogenetic basis, indicating that the RBCC domain was assembled and fixed prior to the acquisition of C-terminal partners.

1.4.1 The RING domain

The RING domain is a relatively short, cysteine-rich motif that coordinates two zinc ions and is found in a range of functionally diverse proteins (Borden and Freemont, 1996). It resides at the extreme N-terminus of all TRIM family members, typically within 10-20 amino acids of the initiator methionine (Torok and Etkin, 2001); (Reymond et al., 2001), and has the canonical sequence $CX_2CX_{9-39}CX_{1-3}HX_{2-3}CX_2CX_{4-48}CX_2C$, where C and H correspond to cysteine and histidine, respectively, and X denotes any amino acid. The RING finger has E3 ubiquitin ligase activity, mediating the transfer of ubiquitin from an E2 conjugating enzyme to a given target substrate and, in doing so, often targeting the substrate for destruction by the proteasome (Xu et al., 2003).

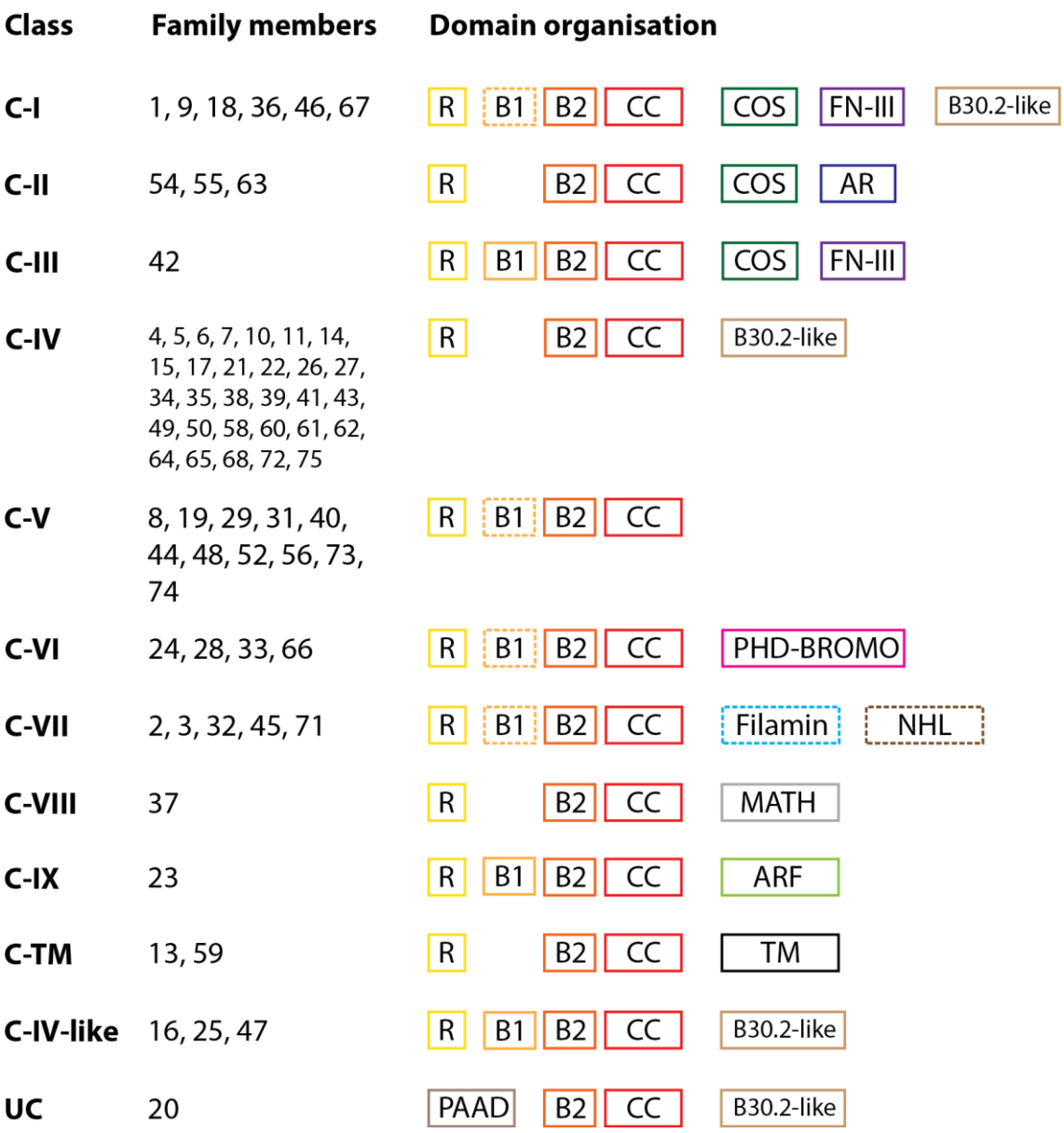


Figure 1.12: The classification of human TRIMs

The first nine subgroups were identified by (Short and Cox, 2006), and the remaining three described later. R: RING domain; B1: B-box1; B2: B-box2; CC: coiled-coil motif; COS: C-terminal subgroup one signature domain; FN-III: fibronectin type III domain; B30.2-like: either the full B30.2 domain (PRYSPRY), or the PRY or SPRY subdomain in isolation; AR: acid-rich region; PHD-BROMO: plant homeodomain and bromodomain; NHL: NHL repeat; MATH: meprin and TRAF homology domain; ARF: ADP ribosylation factor-like GTPase; TM: transmembrane domain. Dashed boxes represent domains that are absent in certain members of a given subgroup. Figure is adapted from Meroni (2012).

1.4.2 The B-boxes

Immediately downstream of the RING domain is the B-box, another zinc-binding motif that mediates protein-protein interactions. Given its structural and functional similarity to the RING finger, it has been proposed that the two domains diverged relatively recently from a common ancestor (Massiah et al., 2007).

B-boxes come in two flavours: type 2 B-boxes (B2), which are found in all TRIM proteins, and type 1 (B1), present in fewer than half. Despite their names, the two motifs share little primary sequence identity: the consensus for B1 is $CX_2CX_{6-17}CX_2CX_{4-8}CX_{2-3}C/HX_{3-4}HX_{5-10}H$[C5(C/H)H2]), while for B2 it is $CX_{2\_4}HX_{7-10}CX_{1-4}D/C_{4-7}CX_2CX_{3-6}HX_{2-5}H$[CHC(D/C)C2H2], where C, D and H correspond to cysteine, aspartic acid and histidine, respectively, and X denotes any amino acid.

The structure of B1 from T18 has been solved at high resolution, revealing an α-helix, two β-strands and three β-turns, which run from V117 to P164 and adopt a ββα topology in the tertiary structure (Figure **1.13A**) (Massiah et al., 2006). Like the RING domain, this structure coordinates two Zn^{2+} atoms: one by a cluster of cysteine residues (C119-C122-C142-C145) and the other by a tetrahedral arrangement of cysteine and histidine residues (C134-C137-H150-H155). The structure of T18 B2 has also been solved by multidimensional NMR spectroscopy; despite the divergence in primary sequence, it exhibited a similar tertiary fold to that of B1 (Figure **1.13B**) (Massiah et al., 2007). Crystal structures of B2 are also available for the human and rhesus orthologues of T5α (Diaz-Griffero et al., 2009; Goldstone et al., 2014b).

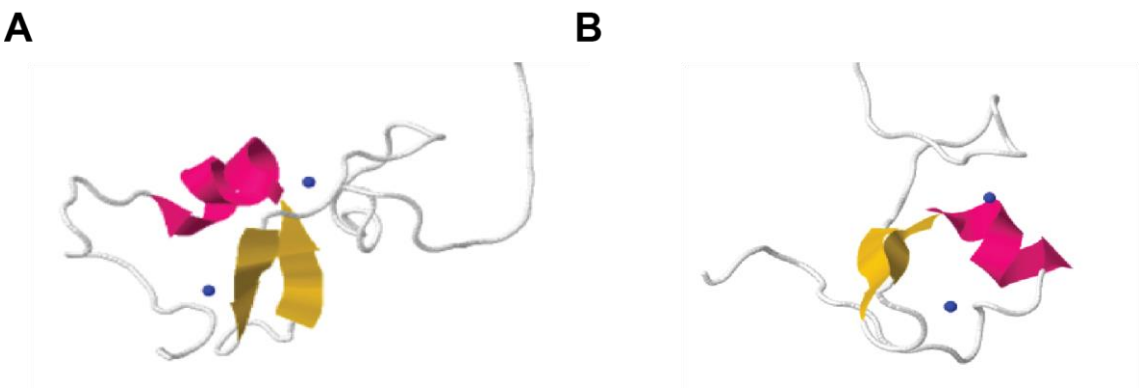


Figure 1.13: Solution structures of (A) B-box1 and (B) B-box2 from T18

Zinc atoms are shown in blue. PDB accession codes: 2FFW, 2DQ5, respectively. (Massiah et al., 2007, 2006)

B2 is the only B-box motif present in T5α, where it serves to mediate the higher-order oligomerisation of T5 dimers into a supramolecular assembly (Li and Sodroski, 2008). This is required to increase the avidity of the T5-capsid interaction, which itself is imperative for restriction (Perez-Caballero et al., 2005; Javanbakht et al., 2005). Recently, the B-box domain of rhT5α has been crystallised in both dimeric and trimeric form (Keown et al., 2016; Keown and Goldstone, 2016). This structure reveals two anti-parallel β-sheets per B-box monomer: the first sheet adopting a β1-β2 topology and the second sheet, β3-β4-α1-β5. Again, this motif coordinates two zinc ions in a tetrahedral fashion (Figure **1.14**).

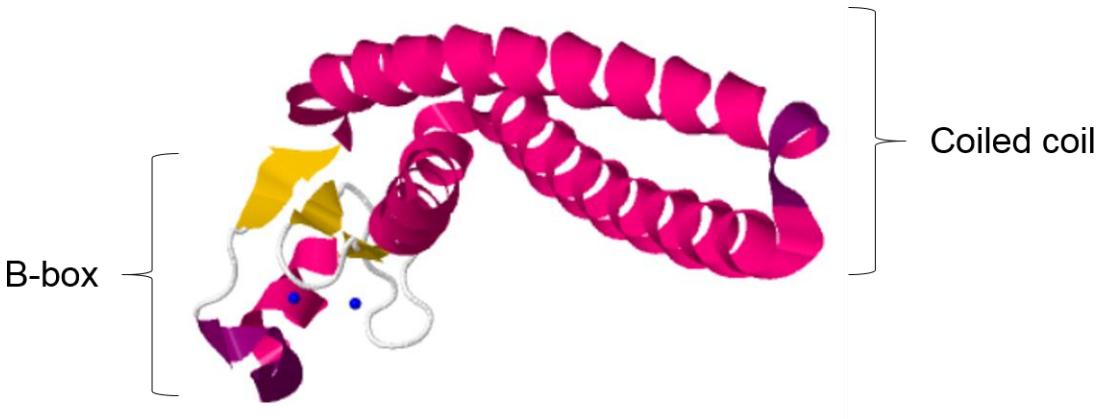


Figure 1.14: The B-box and coiled-coil of rhesus T5α

Zinc atoms are shown in blue. PDB accession code: 5F7T. (Wagner et al., 2016)

In the dimeric structure, the interface between monomers involves a salt bridge between E120 of one partner and R121 of the opposing one. This explains the former observation that charge-reversing mutations of surface-exposed residues (e.g. R121E) attenuate the restriction of HIV-1 by rhT5α, presumably by reducing its propensity for self-association (Diaz-Griffero et al., 2009). Incidentally, the fact that the B-box of rhT5α has been observed to form both dimers and trimers in solution suggests that this oligomerisation interface is relatively plastic in nature, a trait that may lend itself well to recognising the varying curvatures of divergent retroviruses (discussed further in Section **1.4.5**).

1.4.3 The coiled-coil motif

The coiled-coil (CC) motif is a highly α-helical region that comprises approximately 100 amino acids and follows the B-box(es) in all TRIM family members. Although this domain is invariably rich in hydrophobic residues (particularly leucine), the primary sequence conservation of CCs from different TRIMs is low.

The CC motif was once believed to consist of several smaller sub-domains punctuated by α-helices. However, more recent data from biochemical and crystallographic analyses of T25 have revealed that the domain actually constitutes a contiguous supercoil, with heptad and hendecad amino acid repeats that adopt a symmetrical 7-7-7-7-11-11-11-11-7-7-7-7 pattern (Sanchez et al., 2014). The supercoil is canonically left-handed at the ends but unwinds towards the middle, where it adopts a slightly right-handed topology. A multiple sequence alignment of 54 human TRIMs revealed that this arrangement is remarkably well conserved, despite limited sequence identity. Indeed, shortly after the characterisation of the T25 coiled-coil, a similar structure was observed in a CC-containing fragment of rhT5α (Goldstone et al., 2014b).

CC folding is dependent on short stretches of amino acids known as *trigger sites* (Steinmetz et al., 2007). These regions spontaneously form α-helices, thereby initiating a chain reaction that causes the remainder of the coiled-coil to ‘zip up’ into the appropriate conformation. This progressive supercoiling is often inextricably linked to the formation of oligomers (typically dimers or trimers) with other

CC-containing proteins. While all TRIM family members are capable of homodimerisation, the requirement for 'compatible' trigger sites limits the opportunity for heterodimerisation. Where this does occur, it is usually between closely related family members, such as T1 and 18 (Reymond et al., 2001) (see Section **1.4.6**).

In the case of T5α, the CC motif is both necessary and sufficient for dimerisation (Berthoux et al., 2005; Javanbakht et al., 2006). This is corroborated by the observation that restriction-defective T5α mutants act in a dominant-negative manner as long as they have an intact CC domain (Stremlau et al., 2004; Perez-Caballero et al., 2005). Given the low affinity of the T5-capsid interaction, dimerisation is indispensable for restriction (Langelier et al., 2008). Indeed, signatures of positive selection within the CC may reflect the conformational impact of dimerisation for restriction specificity (Sawyer et al., 2005).

1.4.4 The B30.2 domain

Members of the C-I and -IV subgroups of the TRIM family are defined by the presence of a C-terminal B30.2 (or PRYSPRY) domain. In T5α, this region is both necessary and sufficient for capsid recognition: substituting the rhT5α B30.2 domain into the human orthologue endows the latter with a rhesus-like ability to restrict HIV-1 (Perez-Caballero et al., 2005; Yap et al., 2005).

Crystal structures of the B30.2 domains from various TRIMs have been solved by a number of groups (Weinert et al., 2009; D'Cruz et al., 2013; Weinert et al., 2015). According to a model based on these structures, the domain consists of thirteen β-strands, with β1-3 residing in the recently acquired PRY domain and β4-13 in the ancestral SPRY domain. These strands form two opposing β-sheets, giving rise to a hydrophobic, sandwich-like core. The domain also contains four flexible loops, which protrude from one surface of the core and constitute the capsid-binding interface. These loops are regions of high amino acid variability and are correspondingly referred to as the variable regions (VRs). It is largely the species-specific polymorphisms within these regions that determine the panel of

retroviruses a given T5 orthologue is able to restrict (Stremlau et al., 2005; Ohkura et al., 2006).

A catalogue of data from crystallographic, NMR and *in silico* simulation studies has revealed that the B30.2 domain makes multiple contacts with its viral target. The prevailing model derived from these data positions VR2 and 3 within a single CA hexamer, while VR1 extends across the threefold axis to bridge adjacent hexamers (Biris et al., 2012, 2013; Kovalskyy and Ivanov, 2014). Consistent with this notion is the structural plasticity of VR1, which may adopt conformations that position the same residue up to 30 Å apart in space (Ohkura et al., 2006). It is likely that such flexibility has evolved to facilitate the recognition of divergent retroviral cores, which often differ in size, shape and curvature. Perhaps a further adaptation to this pleiomorphy is the fact that T5α need only engage with a minority of capsid subunits (~25%) for restriction to ensue (Shi et al., 2013).

Given their need to rapidly evolve with episodes of viral challenge, the VRs are hotspots of positive selection. Interestingly, there is a propensity for charged residues to be particularly subject to these selective forces, indicating that the interaction between a B30.2 domain and its CA partner is likely electrostatic in nature. For example, the removal of a single positive charge – R332P – in huT5α is sufficient to confer it with recognition of HIV-1 (Yap et al., 2005), while the analysis of N-MLV escape mutants reveals opposite charge preferences for the B30.2 domains of rhesus and human T5α (Ohkura et al., 2011; Ohkura and Stoye, 2013).

1.4.5 T5α and TCyp

T5α and -Cyp (collectively referred to as T5 herein) are related proteins derived from the *T5* gene. Both molecules restrict retroviral replication prior to reverse transcription in a capsid-dependent manner (Stremlau et al., 2004; Sayah et al., 2004). However, while T5α orthologues are abundant among mammals and other vertebrates (Song et al., 2005a; van der Aa et al., 2009; Pacheco et al., 2010); (Fletcher et al., 2010), TCyp is the product of retrotransposition events that substituted the B30.2 domain for cyclophilin A (CypA) in a handful of primate

lineages (Brennan et al., 2008; Virgen et al., 2008). This section will describe what is known about the nature of the T5-capsid interaction, as well as the mechanisms by which restriction is effected.

Capsid recognition

Retroviral cores are pleiomorphic structures that vary in hexamer number, pentamer distribution, and overall size and shape (Ganser-Pornillos et al., 2004; Benjamin et al., 2005; Heymann et al., 2008). Indeed, even virions from the same quasispecies can exhibit discordant geometries (Welker et al., 2000). This presents a challenge to capsid-binding restriction factors, which must recognise divergent retroviral targets with low primary sequence identity (Esteva et al., 2014).

Until recently, much of our understanding of the T5-capsid interaction was based on genetic data – particularly the isolation of viral escape mutants (Ohkura et al., 2011; Ohkura and Stoye, 2013). Biochemical analyses were hampered by technical difficulties in purifying native T5, as well as the need for both interacting partners to be present in multimeric form (Diaz-Griffero et al., 2006; Yap et al., 2007). Nevertheless, some insight was gleaned thanks to the development of various *in vitro* assays for capsid binding. These typically involved incubating the lysates from restriction factor-expressing cells with core-like assemblies of CA, and then detecting binding either through co-immunoprecipitation or electron microscopy (Stremlau et al., 2006; Hilditch et al., 2011). Encouragingly, these methods yielded binding data that correlated well with restriction phenotypes.

Since then, there have been numerous efforts to gain deeper structural insight into the T5-capsid interaction (Ganser-Pornillos et al., 2011; Goldstone et al., 2014b; Sanchez et al., 2014). Collectively, these studies have given rise to a model in which T5 forms elongated, antiparallel dimers that self-associate into a hexameric lattice (Figure **1.15A**); the resulting structure is geometrically complementary to the capsid target. Regions of electron density within this structure reveal that the coiled-coils form the longitudinal axes of each dimer, while the B-box2 domains sit at opposing ends, where they can mediate homotypic interactions with neighbouring dimers. The RING domains reside in close proximity to the B-box2

domains, but are thought to face outside the plane of assembly, where they can interact with components of the ubiquitination machinery.

In the T5α dimer, a hairpin turn immediately after the coiled-coil causes the adjacent linker 2 (L2) region to double back towards the twofold axis of the dimer, where it terminates in a four-helix bundle. This structure firmly anchors the B30.2 domains at the centre of the dimer (Figure **1.15**). Conversely, in TCyp, CypA is attached to the neighbouring L2 region via an unstructured linker. This confers the domain with a greater degree of conformational freedom (Goldstone et al., 2014b). It is also noteworthy that while the B-box2 domain is indispensable for T5α-mediated restriction (Javanbakht et al., 2005), this is not the case for TCyp, indicating that there are probably gross differences in the supramolecular structures of these two factors (Diaz-Griffero et al., 2007).

The above model was recently validated by the Sundquist group, who used electron microscopy to visualise native T5 proteins in complex with disulphide-crosslinked HIV-1 cores (Li et al., 2016c). Interestingly, this work revealed a distinct lack of uniformity within the T5 lattice, with hexamers spanning between 15 and 55 nm in diameter. This phenotype may have evolved to accommodate inherent irregularities on the capsid surface arising from different chemical microenvironments, in addition to the more overt deformations created by pentamers. A degree of flexibility in the T5 lattice may also be instrumental in permitting a single orthologue to adapt to the varying curvatures of distantly related viruses.

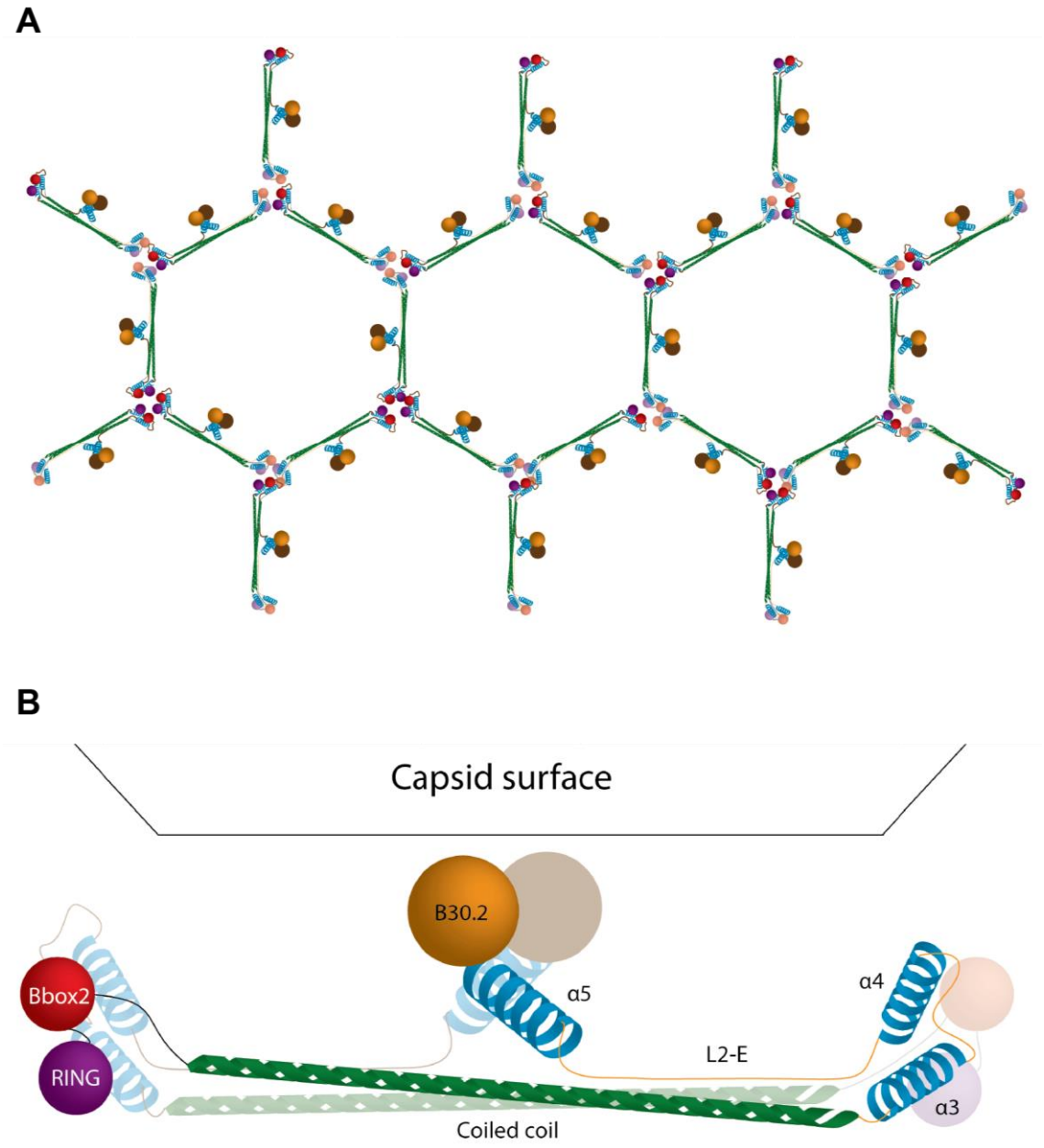


Figure 1.15: Prevailing models for the higher- and lower-order oligomerisation of T5α

(A) The positions of individual domains within the T5α lattice. The RING and B-box2 domains (purple and red) are positioned at the threefold vertices, while the coiled-coils (green) extend to form the hexamer edges and the B30.2 domains (red and brown) sit at the twofold axes of each edge. **(B)** Side-view of a single T5α dimer interacting with capsid. L2-E: an **e**xtended region within linker L2.

Mechanisms of restriction

Early models of T5-mediated restriction were largely centred on the RING domain and its associated E3 ubiquitin ligase activity. A number of groups proposed that T5 uses this activity to co-opt the proteasome and accelerate capsid uncoating, thereby imposing a block to the coupled process of reverse transcription (Perron et al., 2007; Kim et al., 2011). This was based on the observation that particulate capsids are converted to soluble capsid proteins more rapidly in cells expressing T5α (Diaz-Griffero et al., 2006; Stremlau et al., 2006), and that both mutational inactivation of the RING finger (Perez-Caballero et al., 2005) and the use of proteasome inhibitors (Butler et al., 2002) can increase viral titres under usually restrictive conditions.

While it was initially believed that CA served as the substrate for ubiquitination, later studies have shown that T5 undergoes autoubiquitination of several residues once it is bound to the viral core (Yamauchi et al., 2008). This gave rise to a model where translocation of Ub-T5 through the proteasome imparts mechanical stress on the attached virion, causing it to disassemble. Given the recent finding that RING domains function as dimers (Yudina et al., 2015), and the trimeric interface of RINGs at each vertex of the T5 lattice, it is feasible that this occurs via the mutual activation of two RING domains followed by ubiquitination of the third. This notion is supported by the observation that K45 and 50 are preferentially ubiquitinated in the rhT5α RING domain (Fletcher et al., 2015).

While proteasome inhibitors relieve the block to reverse transcription, the observation that virions accumulate in the cytoplasm of MG132-treated cells – coupled with an absence of 2-LTR circles under these conditions – point towards an additional pathway that is effected prior to nuclear entry (Campbell et al., 2008; Wu et al., 2006). This, combined with the recent observation that T5α can mediate the autophagic degradation of virions (Mandell et al., 2014a, b), suggests that T5 orchestrates a number of different restriction pathways, only some of which are proteasome-dependent.

In addition to the direct restriction of virions, it has been shown that T5 functions as an innate immune sensor, thereby contributing to the establishment of an antiviral state. Upon the detection of retroviral cores, T5α works with the E2 ubiquitin-conjugating enzyme complex, UBC13-UEV1A, to catalyse the synthesis of K63-linked polyubiquitin chains. These molecules activate the TAK1 complex, initiating a signalling cascade that culminates in the stimulation of NF-κB and AP1, and the resultant transcription of pro-inflammatory genes (Pertel et al., 2011; Lascano et al., 2016).

1.4.6 T1 and T18

T1 and 18 are paralogous members of the C-I subfamily of TRIM proteins (Meroni and Diez-Roux, 2005). Like T5α, both harbour a C-terminal B30.2 domain; however, the presence of two additional domains upstream of this region (COS and FN-III) makes them almost twice as long.

T1 and 18 map to mirrored loci on opposing arms on the X chromosome, suggesting that one arose from the other following an intrachromosomal duplication event (Perry et al., 1998; Buchner et al., 1999). The two proteins thus share 83% amino acid similarity (76% identity) and conserved domain boundaries, while the corresponding genes have 71% nucleotide identity and conserved splice sites (Perry et al., 1999).

T1 and 18 form both homo- and heterodimers via their CC motifs, a phenomenon that is required for their localisation to microtubules (Short et al., 2002) and is indispensable for function (Cainarca et al., 1999). Both proteins have been implicated in stabilising and organising microtubules, as well as anchoring other proteins to them (Short et al., 2002; Berti et al., 2004); mutations are thus linked to a host of disease phenotypes, including genetic disorders of embryonic midline structures (Cox et al., 2000), X-linked intellectual disabilities (Geetha et al., 2014) and end-stage breast cancer (Wang et al., 2016).

Although T1 and 18 are both integral to embryonic development, they have distinct spatiotemporal expression profiles (Buchner et al., 1999). T1 is expressed at a low level during early embryogenesis, where it is confined predominantly to the central

nervous system and the developing heart (particularly in the ventricle walls and septum). Expression of this protein increases in other organ systems – particularly the stomach, thyroid and kidney – later in embryogenesis. Conversely, T18 is ubiquitously expressed among early embryonic tissues, but is not detectable in the heart. Throughout embryogenesis, T18 is typically more abundant than T1.

Quite apart from these studies, the Stoye group have reported that T1 – but not T18 – is able to restrict both wild-type N-MLV and a capsid mutant (N82D) (Yap et al., 2004). This phenotype has since been reported by other groups (Uchil et al., 2008), but at the time of writing remains poorly characterised.

1.5 Aims of this project

Although the restriction phenotype of T1 was first described more than a decade ago, there have been no attempts to characterise this activity since. The main objectives of this thesis are thus to provide an insight into the basis of T1-mediated restriction (Chapter **3**) and to draw comparisons between T1 and its distantly related cousin, T5α (Chapter **4**). Furthermore, given that these proteins diverge significantly in terms of their size, Chapter **5** will explore the impact of intramolecular domain spacing on facilitating a productive interaction with capsid. Each of these chapters will be prefaced with a brief introduction of their own to provide context to the results described therein.

Chapter 2 Materials & Methods

2.1 Recombinant DNA

2.1.1 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) was used to amplify genes from plasmids, produce chimeric DNA and add tags to proteins.

A typical reaction consisted of 50 ng template DNA, 500 nM of each primer (Sigma), 200 µM dNTPs, and 2.5 U PfuUltra DNA polymerase (Agilent) in the supplied buffer, made up to a final volume of 50 µL with dH_2O. Cycling parameters included denaturation at 95°C for 2 min, followed by 25 cycles of denaturation (95°C for 1 min), primer annealing (57°C for 2 min), and extension (72°C for 3 min), before a final round of extension at 72°C for 10 min. The annealing temperature was raised to 60°C when using primers with a melting temperature (T_m) exceeding 65°C. Thermal cycling was performed using the PTC-100 thermal cycler (MJ Research).

All of the primers used in this project are listed in Appendix **7.1** of this thesis.

2.1.2 Overlapping PCR

Overlapping PCR is used to stitch together two fragments of DNA to produce a chimeric molecule. Briefly, each fragment is independently amplified from its parental construct using primers with cross-complementarity at the ends which are to be joined. The two fragments are then made contiguous in a second round of PCR (Figure **2.1**). Standard PCR conditions (Section **2.1.1**) were used for these reactions.

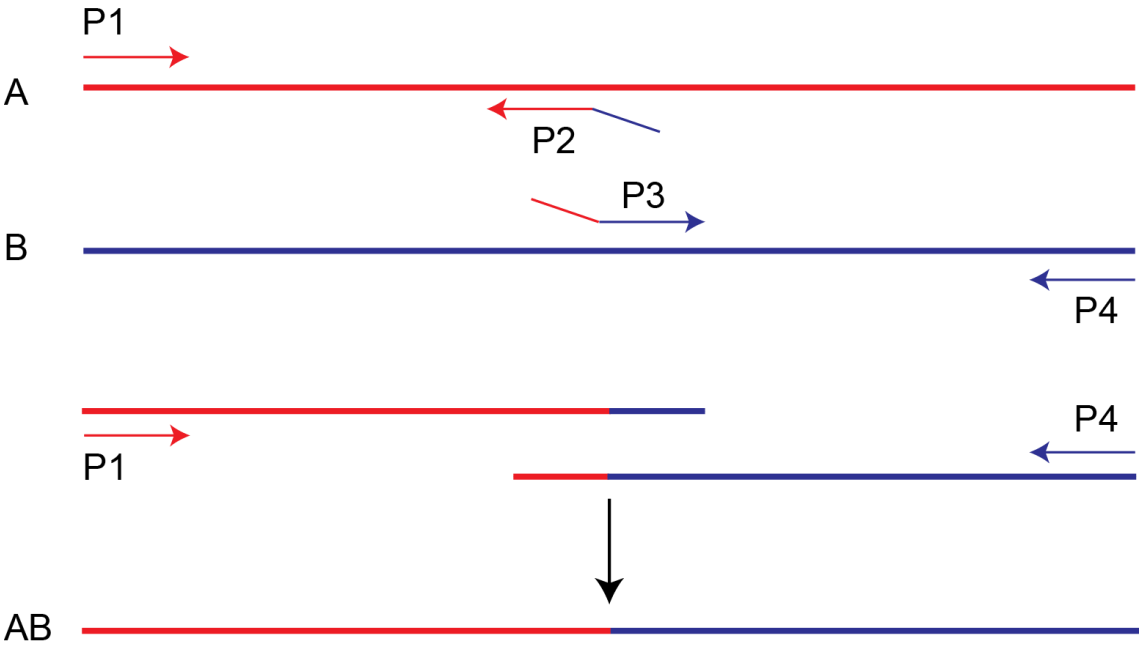


Figure 2.1: Overlapping PCR

Fragments from two constructs (A and B) are amplified using primer pairs P1-P2 and P3-P4, respectively. P2 and P3 have 5' ends that are complementary to the opposing template, enabling the fragments to anneal in the centre following the first round of PCR. Amplification of the resulting product with P1-P4 then yields a contiguous chimera.

2.1.3 Site-directed mutagenesis

Quikchange site-directed mutagenesis (Agilent) was used to introduce point mutations, insertions and deletions to a DNA template. A typical reaction consisted of 10 ng of template DNA, 125 ng of each primer (Sigma), 200 µM dNTPs, and 2.5 U PfuUltra DNA polymerase (Agilent) in the supplied buffer, made up to a final volume of 50 µL with dH_2O. Cycling parameters included an initial denaturation phase (95°C for 30 s), followed by 18 cycles of denaturation (95°C for 30 sec), primer annealing (55°C for 30 sec), and extension (68°C for 1 min per kb DNA). The reaction was then chilled at 4°C for 10 min. Thermal cycling was performed using the PTC-100 thermal cycler (MJ Research).

Following mutagenesis, the reaction was incubated with 20 U *DpnI* (NEB) at 37°C for 1.5 h. *DpnI* selectively digests methylated DNA, thereby removing any template material and leaving the mutant behind. After *DpnI* digestion, the remaining DNA was concentrated by ethanol precipitation (Section **2.1.10**) and transformed into

XL10-Gold ultracompetent *E. coli* (Agilent) that had been pre-treated with β-mercaptoethanol (4 µL per 50 µL cells).

2.1.4 Restriction digestion

DNA was digested with restriction enzymes for the purposes of screening PCR clones and providing sticky ends for insert-vector ligation (Section **2.1.12**).

A typical 20 µL reaction consisted of up to 1.5 µg DNA and 1 U of restriction enzyme in the supplied buffer (NEB; Roche). If performing a double-digest, then 0.5 U of each enzyme was used in a buffer compatible with both. Reactions were incubated at 37°C with gentle agitation for 1-3 h. The digested fragments were then resolved by electrophoresis in a 0.8% (w/v) agarose gel (Section **2.1.5**). If required for downstream PCR or ligation reactions, the desired bands were excised and purified from the gel (Section **2.1.6**).

2.1.5 Agarose gel electrophoresis

DNA fragments were separated by size using agarose gel electrophoresis. Agarose gels were made by dissolving 0.8-1.2% (w/v) agarose (Melford) in TBE Buffer (0.09 M Tris-HCl, 0.09 M borate, 2 mM EDTA) and adding SYBR Safe DNA stain (Invitrogen) at a final concentration of 1:2000.

DNA samples were mixed with 10x loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol) and loaded next to a ladder of DNA size markers (SmartLadder, Eurogentec). Gels were run at 120 V for 30 min and then visualised using the Doc-It transilluminator (UVP).

2.1.6 Extraction of DNA from agarose gels

When DNA fragments were required for overlapping PCR or DNA ligation, they were first resolved by agarose gel electrophoresis (Section **2.1.5**) and then extracted from the gel by placing it on a UV transilluminator and excising the

relevant band(s) with a scalpel. The excised bands were dissolved in Buffer QG (Qiagen) at 60°C for 10 min with vigorous shaking. Agarose and other contaminants were then removed using the MinElute Gel Extraction Kit (Qiagen), and the DNA was eluted in 10 µL of the supplied elution buffer.

2.1.7 Gateway cloning

The Gateway cloning system (Life Technologies) is a protocol that dispenses with the traditional cut-and-paste method for cloning DNA, instead relying on enzyme-directed recombination between proprietary recombination sequences. In this project, the Gateway system was used to clone restriction factor genes into various destination vectors.

The *LxIY* expression vector was the most frequently used (Figure **2.2**), where **L** represents the **L**TR promoter that drives constitutive expression of the insert, **x** represents the insert, **I** represents the internal ribosome entry site (**I**RES), and **Y** represents the e**Y**FP reporter gene. When inducible expression of the restriction factor was required, an alternative vector system was used (see Section **2.2.7**).

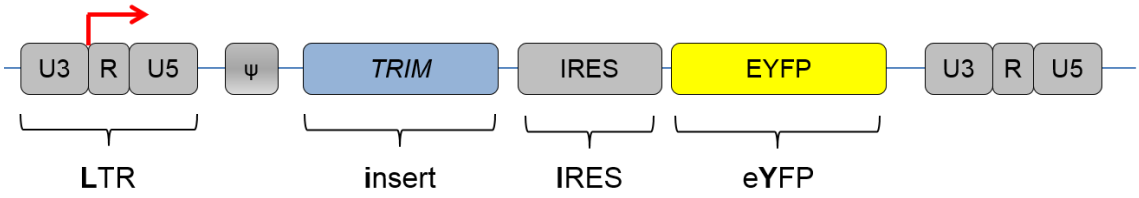


Figure 2.2: The LxIY vector used for restriction factor expression

Before putting a PCR product into the appropriate destination vector, it was first cloned into an entry vector by either BP recombination or directional TOPO cloning. In the former protocol, the PCR product is synthesised using modified primers that append attB sites to either end, facilitating enzyme-directed recombination between the attB-flanked PCR product and a supplied entry vector (pDONR221). In the latter, the tetranucleotide CACC is appended to the 5' end of a blunt-end PCR product to enable directional recombination into the pENTR-D-TOPO entry vector. This tetranucleotide has the additional benefit of embedding the start codon within

a Kozak sequence, thereby enhancing the translational efficiency of the gene (Kozak, 1987). In both cases, successful entry clones were identified by selection with kanamycin and sequencing with the M13 forward and reverse primers (GATC Biotech). These reactions are depicted graphically in Figure **2.3**.

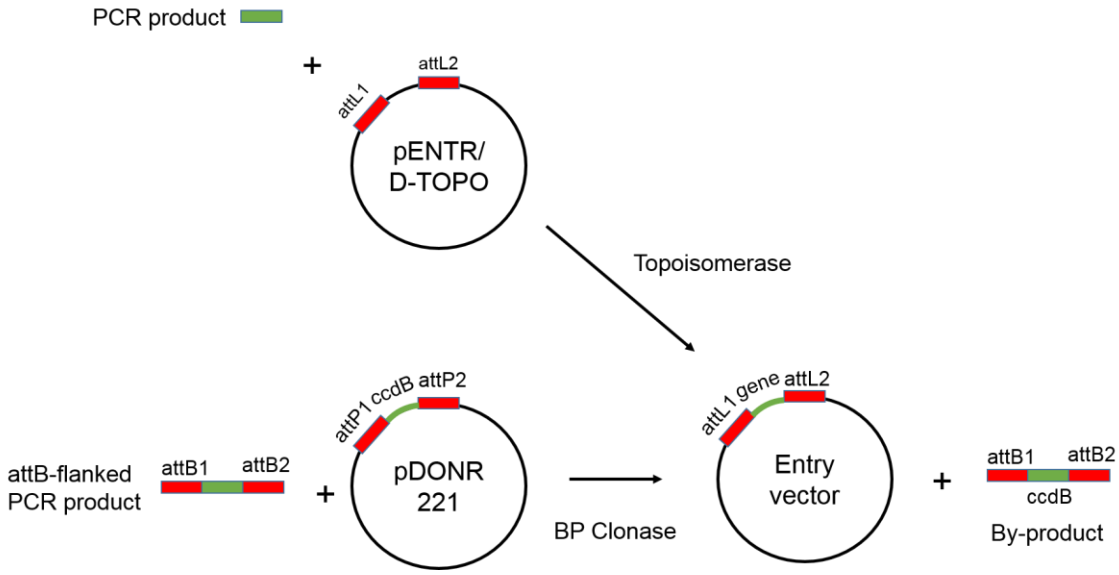


Figure 2.3: BP recombination and the TOPO reaction are used to clone PCR products into the entry vector

Both of the above reactions embed the PCR product between attL1 and attL2 sites. Because destination vectors carry complementary attR1 and attR2 sites, the PCR product can then be transferred from entry to destination vector via the LR recombination reaction (Figure **2.4**). Prior to LR recombination, the destination vector carries a suicide gene (*ccdB*) flanked by attR sites. The recombination reaction substitutes the desired insert into the destination vector, while transferring the suicide gene to the entry vector. The resulting expression clone can then be acquired by selection with ampicillin, while the suicide gene ensures removal of the entry clone.

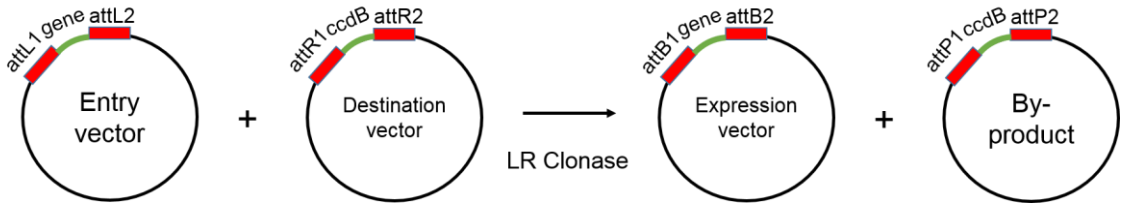


Figure 2.4: The LR reaction is used to transfer a PCR product from the entry vector to an appropriate destination vector

2.1.8 Transformation

Transformation was used to propagate plasmid DNA for various downstream applications, including screening PCR clones and transfecting mammalian cells. The bacteria used for most applications were One Shot TOP10 *E. coli* (Invitrogen), although XL-10 Gold ultracompetent cells (Agilent) were used to transform DNA from mutagenesis reactions (Section **2.1.3**), and Rosetta 2 DE3 competent cells (Novagen) were used to transform DNA for large-scale protein expression and purification (Section **2.4.1**).

Briefly, cells were defrosted in pre-chilled tubes and DNA was added at a volume not exceeding one-tenth that of the cells. This mixture was left on ice for 30 min, heat-shocked at 42°C for 30 s, and then immediately returned to ice for 2 min to enable cell recovery. 250 µL SOC media (2% w/v tryptone, 0.5% w/v yeast extract, 8.55 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose) was added to the cells, before incubation at 37°C for 1 h with agitation (225 rpm). 20-100 µL of the outgrowth was then plated on LB-agar supplemented with the appropriate antibiotic(s) (see Table **2.1**), before incubating at 37°C overnight.

Antibiotic(s)	Solvent	Final concentration	Source
Ampicillin-Nafcillin	Water	50 µg/ml each	Sigma
Chloramphenicol	Ethanol	25 µg/ml	Sigma
Kanamycin	Water	50 µg/ml	Sigma
Streptomycin	Water	50 µg/ml	Sigma

Table 2.1: List of antibiotics used for the selection of transformants

2.1.9 Propagation and purification of plasmid DNA

Following overnight incubation, a single colony was picked from a plate of transformants and added to 3 mL LB (1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7) supplemented with the appropriate antibiotic(s), before a second overnight incubation at 37°C with agitation (225 rpm).

If screening PCR clones, 2 mL of the overnight culture was pelleted at 13,000xg for 3 min, and plasmid DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. Clones were then screened by restriction digestion and agarose gel electrophoresis of the digested fragments.

For larger scale production of plasmid DNA, 1 mL of the initial outgrowth was added to 35 mL LB supplemented with the appropriate antibiotic(s) and grown overnight once more at 37°C with agitation (225 rpm). The cells were then pelleted at 4,000xg for 20 min and plasmid DNA was purified using the HiSpeed Plasmid Midi Kit (Qiagen).

2.1.10 Concentration of DNA by ethanol precipitation

DNA was concentrated by adding 0.1 volumes sodium acetate and 2.5 volumes 100% ethanol and chilling the mixture at -80°C for 30 min. The cooled mixture was then centrifuged at 18,000xg and 4°C for 15 min, and the supernatant replaced with 500 µL 70% ethanol. This mixture was centrifuged a second time under the same conditions, before removing the supernatant, air-drying the pellet for 5 min, and then resuspending it in 5 µL dH_2O. Concentrated DNA was either stored at -20°C or immediately transformed into bacteria.

2.1.11 Quantitation of DNA by spectrophotometry

DNA was quantified by absorbance at 260 nm using a NanoDrop 2000 UV spectrophotometer.

2.1.12 DNA ligation

Large-scale protein expression required the ligation of DNA into an appropriate expression vector, pET-47b (Novagen). To achieve this, both insert and vector were digested with *HindIII* and *XhoI* in the appropriate reaction buffer. The digested fragments were then resolved by electrophoresis on a 1% (w/v) agarose gel, and the relevant bands excised and purified (Section **2.1.6**). The volumes of insert and vector to use in the ligation reaction were calculated using the equation below, ensuring a minimum of 40 ng vector and an insert-to-vector ratio of ~8-10:

$$Insert\ (ng) = \frac{insert\ length\ (bp)}{vector\ length\ (bp)}\ x\ vector\ (ng)\ x\ ratio$$

The appropriate volumes of insert and vector were incubated with 1 µL T4 DNA ligase (NEB) in the supplied reaction buffer for 3 h at room temperature. The entire reaction was then transformed into One Shot TOP10 *E. coli* cells (Invitrogen), and the transformants plated on LB-agar supplemented with kanamycin. All ligation products were sequence-verified using the T7 forward and reverse sequencing primers (GATC Biotech).

2.1.13 DNA sequencing

DNA samples were shipped to GATC Biotech for overnight single-pass sequencing. All samples were diluted to 100 ng/µL in dH_2O before shipping, and sequencing primers supplied at 10 µM. The resulting chromatograms were analysed on SeqMan Pro (Lasergene) for sequence verification.

2.2 Cell culture & the restriction assay

2.2.1 Maintenance of cell lines

Human embryonic kidney (293T) cells and *Mus dunni* tail fibroblasts (MDTF) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Life Technologies) supplemented with 10% foetal calf serum (FCS), 100 U/mL penicillin and 100 µg/mL streptomycin. Cells were maintained at 37°C under conditions of 95% humidity and 5% CO_2, and passaged at a ratio of 1:10 every 3-4 days by adding 1 mL trypsin-versene (0.05% w/v trypsin and 0.53 mM EDTA in PBS) to the monolayer and resuspending the detached cells in 9 mL fresh DMEM.

2.2.2 Overview of the restriction assay

Restriction was measured using an established two-colour flow cytometry assay (Bock et al., 2000). Briefly, a vector carrying a restriction factor construct and an eYFP reporter was co-transfected into 293T cells with plasmids encoding the viral *Gag, Pol* and *Env* genes (Section **2.2.3**). This yielded 'delivery viruses' which could be used to transduce MDTF cells (Section **2.2.5**).

In parallel, 'tester viruses' – those viruses whose restriction was being measured – were generated by co-transfecting the appropriate *gag, pol* and *env* genes with an eGFP reporter construct. These viruses were used to infect MDTF cells 48 h after transduction with the restriction factor construct (Section **2.2.6**). Restriction of the tester virus by the transduced construct was then measured 48 h post-infection by flow cytometry (Section **2.3.3**). This entire process is depicted graphically in Figure **2.5**.

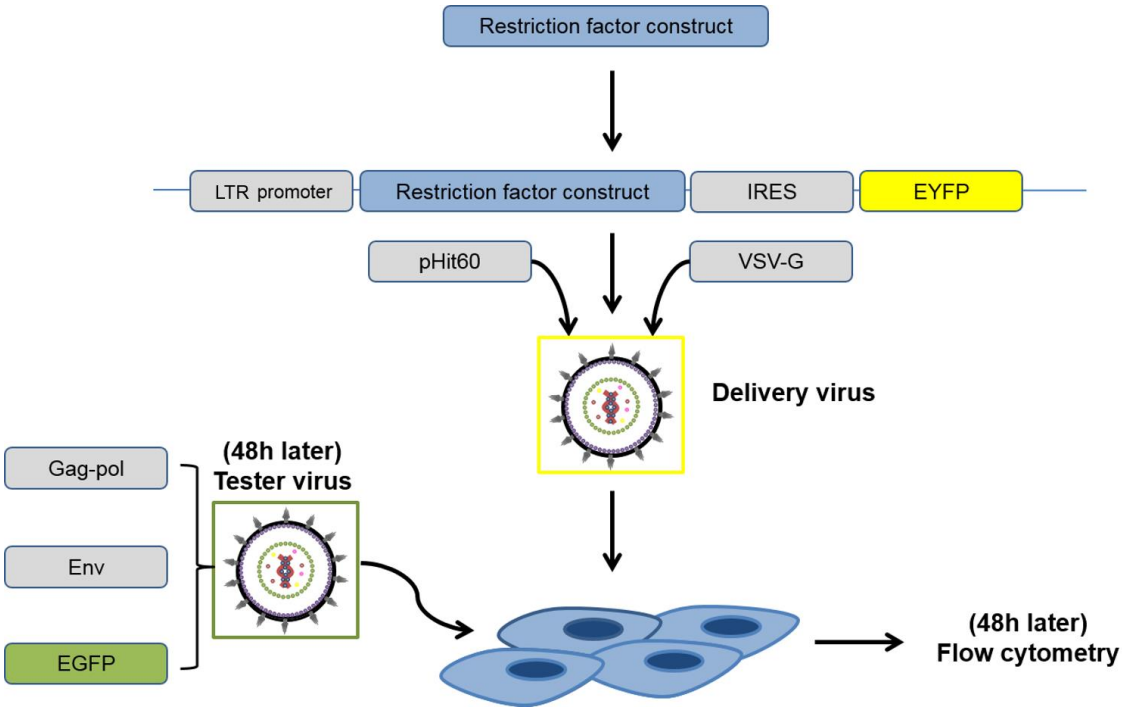


Figure 2.5: The two-colour restriction assay

2.2.3 Virus production by transient transfection

293T cells in the log phase of growth were seeded in 6 cm dishes (Nunc) at a density of 2 x 10^6 cells per dish and left overnight to adhere. The media was then replaced with 5.5 mL fresh DMEM, and the cells returned to incubation for at least 4 h prior to transfection. DNA was introduced to the cells by lipofection using TurboFect (Stratagene), according to the manufacturers' instructions. Briefly, equal masses of three plasmids (6 µg total DNA; see Section **2.2.4** for a description of the transfected constructs) were incubated with 600 µL serum-free DMEM and 12 µL TurboFect at room temperature for 20 min. The mixes were than added to the cells dropwise, and the cells returned to incubation overnight.

Approximately 15 h post-transfection, cells were stimulated with 5 mL fresh DMEM containing 10 mM sodium butyrate. Sodium butyrate enhances transcription by acting as a histone deacetylase inhibitor (Candido et al., 1978), thereby increasing viral titres. After 6 h, the media was replaced with 2.5 mL regular DMEM and the cells left to grow overnight. The next morning, viruses were harvested by drawing off the media with a 5 mL syringe (BD Plastipak) and passing it through a 0.45

micron filter (Sartorius) to remove cells and debris. The filtrate was divided into 250 µL aliquots, frozen at -80°C for at least 2 h, and then titrated on MDTF cells to determine the volume of virus required to give an MOI ≈ 0.7.

2.2.4 Plasmids used for virus production by transient transfection

As described in Section **2.2.3**, viruses were produced by transiently co-transfecting a total of 6 µg DNA. This was split equally between three plasmids when producing MLVs and lentiviruses, and between two plasmids when producing foamy viruses. Table **2.2** shows a list of the plasmids used to produce each type of virus. Note that all MLVs and lentiviruses were pseudotyped with VSVg Env. This both broadens the viral tropism, thereby permitting infection of any cell type, and standardises viral entry so that any restriction data obtained can be attributed purely to post-entry events. All foamy viruses were pseudotyped with SFV Env.

			Plasmid function		
			Gag-Pol	**Env**	**Reporter**
Tester viruses	**Gamma**	**B-MLV**	pCIG3 B	pVSVg	pIRES2-EGFP
		N-MLV	pCIG3 N	pVSVg	pIRES2-EGFP
		NB-MLV	pHit60	pVSVg	pIRES2-EGFP
	Lenti	**HIV-1**	p8.91	pVSVg	pCSGW
		HIV-2	pHIV2GagPol	pVSVg	pHIV2gfp
		FIV	pFP93	pVSVg	pGiNWF
	Foamy	**PFV**	POO1	pciSFV-1Envwt	*eGFP reporter is present in the Gag-Pol vector*
		SFV	SOO1	pciSFV-1Envwt	
		FFV	FOO3	pciSFV-1Envwt	
Delivery viruses			pHit60	pVSVg	*Expression vector with RF and eYFP reporter*

Table 2.2: Plasmids used in the transient transfection of 293T cells to produce retroviruses

2.2.5 Transduction of MDTF cells

MDTF cells were seeded in 12-well plates (Corning) at a density of 5 x 10^4 cells/well and left to adhere overnight. Approximately 24 h after seeding, the cells were transduced with delivery virus at an MOI ≈ 0.7 and returned to incubation for 48 h.

2.2.6 Infection of transduced MDTF cells

48 h post-transduction, MDTF cells were re-seeded in multiple wells at a density of 5 x 10^4 cells/well, and then infected in duplicate with the appropriate tester virus (MOI ≈ 0.7). After a further 48 h, the cells were harvested, pelleted and fixed for analysis by flow cytometry. A full description of the preparation of samples for flow cytometry, and the acquisition and analysis of flow cytometry data, can be found in Section **2.3**.

2.2.7 Regulation of restriction factor expression by doxycycline induction

To measure the relationship between the expression level of a restriction factor and the potency of restriction, C-terminally HA-tagged constructs were cloned into expression vectors with doxycycline-inducible promoters.

2.2.7.1 Development of a doxycycline-inducible vector system

Tet-On 3G is the third generation in a series of expression systems with doxycycline-inducible transcription. In contrast to the former Tet-On and Tet-On Advanced systems, Tet-On 3G uses the P_{TRE3G} promoter, which has a lower basal level of transcription (Loew, 2009), and the rtTA3 trans-activator, which has increased sensitivity to doxycycline (Zhou et al., 2006). These two features combine to minimise transcriptional leakiness while maximising the level of induction.

A former student in the Stoye lab generated an MDTF cell line constitutively expressing rtTA3 by stable transfection ('R18 cells' herein). Retroviral vectors were

also produced, carrying both the P_{TRE3G} promoter and attR sequences for Gateway cloning of restriction factor genes from their entry vectors (Section **2.1.7**).

Only two dox-inducible vectors were used in this project, denoted CS14 and CS15. In both vectors, the *eYFP* gene is placed furthest upstream in order to maximise reporter expression via cap-dependent translation. The restriction factor construct was then placed downstream, either following an IRES (CS14), or directly after *eYFP* without any intervening IRES (CS15). In the latter case, restriction factor translation can only occur via leaky ribosome scanning; thus CS15 provides a lower range of expression than does CS14. The structure of both inducible vectors, as well as the non-inducible vector used in all other restriction experiments (Bock et al., 2000), is shown in Figure **2.6**. A more thorough account of the design and generation of these vectors can be found in Li et al. (2016b). All of the protocols described herein are adapted from that work.

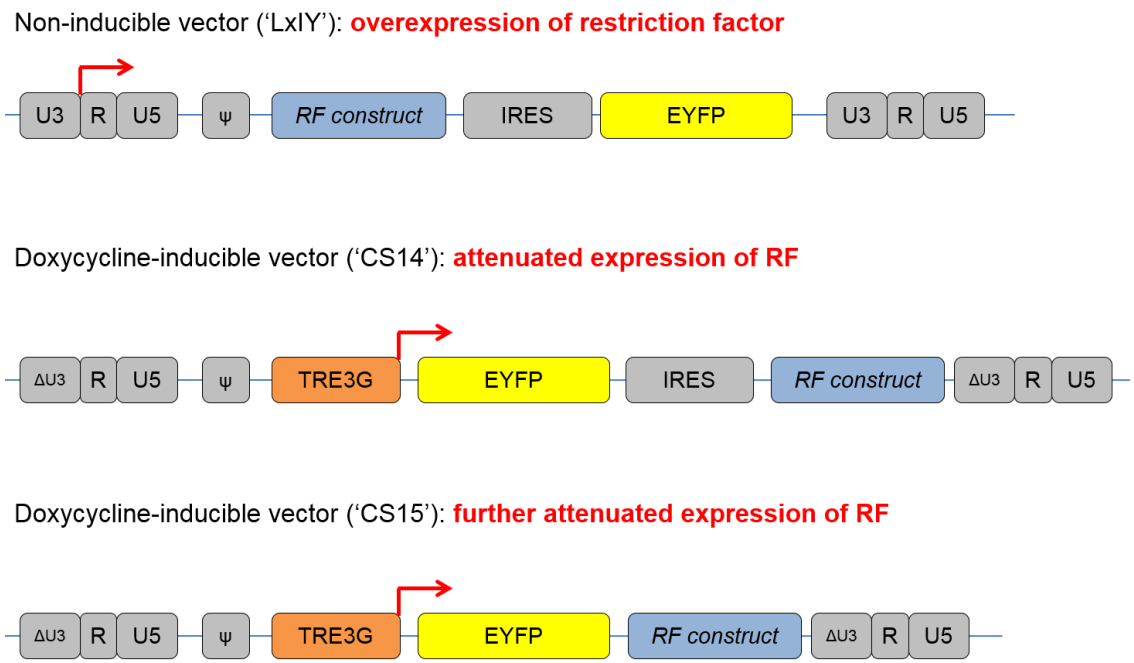


Figure 2.6: The non-inducible and doxycycline-inducible vectors used for restriction factor expression

The non-inducible vector *LxIY* was used to express restriction factor (RF) constructs in the vast majority of experiments described in this thesis. In this vector, transcription is initiated from the U3 promoter, and the upstream position of the RF construct relative to eYFP ensures its overexpression. However, when it was necessary to control the amount of restriction factor present in the cell, a *CS14* or *CS15* vector was used instead: in these vectors, the U3 promoter is deleted to permit transcription from a TRE3G element, and the RF construct is placed downstream of the reporter in order to limit its expression.

2.2.7.2 Measuring restriction by restriction factors under doxycycline-inducible expression

R18 cells were seeded in 12-well plates (Corning) at a density of 5 x 10^4 cells/well and left overnight to adhere. 24 h after seeding, the cells were transduced with delivery virus (MOI ≈ 0.7) carrying a restriction factor construct in either a non- or dox-inducible vector.

At least 48 h post-transduction, the cells were reseeded into 24-well plates (Corning) at a density of 2.5 x 10^4 cells/well. Those cells transduced with a non-inducible construct were reseeded in regular DMEM, while those transduced with a dox-inducible construct were re-seeded in DMEM containing the appropriate concentration of doxycycline. Approximately 24 h after re-seeding, the cells were infected with a tester virus (MOI ≈ 0.7) and, after a further 24 h, 5 µg doxycycline was added to boost reporter expression. Cells were harvested for flow cytometry between 24 and 48 h after the second exposure to doxycycline.

Restriction factor expression was quantified in parallel to the restriction assay by performing a western blot using a primary antibody against the C-terminal HA tag of the protein (Section **2.4.10**).

2.3 Flow cytometry

2.3.1 Preparation of samples for flow cytometry

Confluent MDTF cells were detached by adding 1 mL trypsin-versene and incubating at 37°C for 10 min. Once detached, the cells were fixed by resuspending in 3 mL 2% formaldehyde-PBS that had been pre-aliquoted into 5 mL round-bottom tubes (Corning). The cells were then pelleted at 400xg for 10 min and the pellet resuspended in 150 µL regular PBS, before filtering through a 35 µm nylon mesh (Corning) to remove clumped cells and debris.

2.3.2 Acquisition of data by flow cytometry

All flow cytometry was performed using the FACSVerse flow cytometer and associated software (BD Biosciences). Live cells were gated according to their forward-scatter versus side-scatter profile. Untransduced and single-colour controls were then used to gate GFP^-YFP^-, GFP^+YFP^-, GFP^-YFP^+ and GFP^+YFP^+ populations within the live cell population. Because GFP and YFP have overlapping emission spectra, a compensation matrix was employed to account for detection spillover. 10,000 cells were acquired for each experimental sample and the data outputted to FlowJo (Tree Star) for post-acquisition analysis.

2.3.3 Calculation of restriction from flow cytometry data

FlowJo was used to count the number of cells in each of the four live-cell populations: GFP^-YFP^-, GFP^+YFP^-, GFP^-YFP^+ and GFP^+YFP^+. These numbers were then exported to Microsoft Excel and plugged into the formula below to calculate the restriction ratio:

$$Restriction\ ratio = \frac{GFP^+YFP^+}{GFP^+YFP^+ + GFP^-YFP^+} \div \frac{GFP^+YFP^-}{GFP^+YFP^- + GFP^-YFP^-}$$

This formula represents a ratio of the percentage of infected (GFP^+) cells in populations expressing restriction factor (YFP^+) versus those not (YFP^-).

Values ≤0.3 were taken as restriction, ≥0.7 as absence of restriction, and intermediate values as partial restriction.

2.4 Protein expression, purification and analysis

2.4.1 Expression and harvesting of protein from *E. coli*

One colony was picked from the plate of transformants and grown in 100 mL terrific broth (12 g tryptone, 24 g yeast extract, 4 mL glycerol, 100 mL 0.17 M KH_2PO_4 plus 0.72 M K_2HPO_4) supplemented with kanamycin and chloramphenicol. The following day, large-scale cultures were made by transferring 7.5 mL of overnight culture to 750 mL fresh TB with kanamycin and chloramphenicol. These cultures were grown in 2 L flasks at 37°C with moderate agitation (180 rpm), until they reached an average optical density (OD) of 0.6. At this point, the temperature was reduced to 20°C, and 1 mM IPTG added to induce protein expression. The cultures were then left to grow overnight.

The following morning, the bacteria were pelleted at 4000xg (4°C for 30 min). The supernatant was decanted and the pellet weighed and resuspended in 7 mL lysis buffer (50 mM Tris pH8, 750 mM NaCl, 20 mM imidazole, 5 mM $MgCl_2$, 5 mM ATP, 0.5 mM TCEP, 0.2% v/v Triton) per gram of cells, by stirring at 4°C until homogenised. The lysate was sonicated for 10 min (power output 4, duty cycle 40%), and then pelleted at 56,000xg and 4°C for 1 hour. The supernatant was then filtered in preparation for affinity purification.

2.4.2 Protein purification by affinity to a nickel column

Following filtration, bacterial lysate was placed on a nickel (Ni^{2+}) column for affinity purification via the N-terminal His_6 tag. 100 column volumes were loaded to enable binding, and then the column was washed in an equivalent volume of Buffer A (50 mM Tris pH8.0, 500 mM NaCl, 20 mM imidazole, 0.5 mM TCEP) in order to remove unbound species. Upon completing the wash step, Buffer B (Buffer A + 300 mM imidazole) was used to elute the bound protein. The increased concentration of imidazole in Buffer B outcompetes His_6 for binding to the column, enabling any nickel-bound species to be displaced and eluted.

2.4.3 Protein purification by ion exchange chromatography

To remove any contaminating species after affinity purification, the eluate from the nickel column was subjected to a second round of purification by ion exchange chromatography.

Briefly, the concentration of salt in the eluate from the nickel column was reduced to 50 mM NaCl by diluting with Buffer C (50 mM Tris pH8.0, 0.5 mM TCEP), before loading onto a 55 mL Source Q anion exchange column which had been equilibrated in Buffer D (Buffer C + 50 mM NaCl). A linear gradient was used from 50 mM to 1 M NaCl over 10 column volumes, and the eluate collected in 10 mL fractions.

2.4.4 Protein purification by size exclusion chromatography

Following ion exchange chromatography, a final round of purification was performed by pooling and concentrating the relevant fractions, and then loading them onto an S200 16/60 size-exclusion column (GE Life Sciences) for gel filtration. The column was run with a flow rate of 0.3 mL min^{-1} and the protein collected in elution buffer (20 mM Tris pH8.0, 150 mM NaCl, 0.5 mM TCEP).

2.4.5 Expression and harvesting of protein from mammalian cells

To determine the protein expression levels for various restriction factor constructs, either MDTF cells (for non-inducible vectors) or R18 cells (for inducible vectors) were seeded at 5 x 10^4 cells/well in a 12-well plate. A delivery virus containing the relevant construct was then used to transduce the cells with an MOI ≈ 0.7. The transduced were left to grow for 48 h, before reseeding in T20 flasks (Thermo Scientific) at 1:5 dilution. After a further 48 h, the cells were transferred to a T80 flask and left to grow for a further 24 h.

Upon reaching approximately 70% confluence, the cells were trypsinised, resuspended in DMEM and centrifuged at 300xg for 15 min to obtain a cell pellet. The pellet was resuspended in 1 mL PBS and centrifuged again at 300xg for 5 min.

The supernatant was then aspirated, and the pellet resuspended in RIPA buffer (150 mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0; Sigma) with 1 protease inhibitor tablet added per 10 mL buffer.

The resuspension was incubated at 4°C for >30 min to allow complete lysis of the cells. The lysates were then centrifuged at 18,000xg for 10 min to separate the protein-containing supernatant from the cell debris. Finally, the supernatant was removed and diluted in RIPA buffer (1:5) ready for quantitation by the BCA assay (Section **2.4.7**).

2.4.6 Quantitation of total protein by spectrophotometry

Protein was quantified by absorbance at 280 nm using a NanoDrop 2000 UV spectrophotometer.

2.4.7 Quantitation of total protein using the BCA assay

Given the presence of detergent in RIPA buffer, total protein concentration in cell lysates was determined using the bicinchoninic acid (BCA) assay rather than the Bradford assay. Briefly, 5 µL of each lysate sample (diluted 1:5) was added to separate wells of a 96-well plate (BD Falcon) alongside standards of known concentration (2000, 1000, 500, 250, 125, 62.5, 31.25 and 0 µg/mL). 200 µL of prepared BCA reagent (Thermo Fisher) was then added to all of the wells, before shaking for 30 s and incubating at 37°C for 30 min. Following incubation, absorbance at 562 nm was measured using the Synergy 2 plate reader and associated software (Gen5). A curve of standards was produced, and total protein concentration in the lysates was interpolated from the curve.

2.4.8 Separation of proteins by SDS-PAGE

Proteins were separated by size using a 4-20% Mini-Protean TGX Precast Gel (Bio-Rad). Briefly, samples were mixed with dH_2O and 5x SDS loading buffer (250 mM Tris-HCl pH6.8, 10% SDS, 12.5% β-mercaptoethanol, 0.1% w/v

bromophenol blue, 50% glycerol) to achieve a final concentration of 1 mg/mL. The samples were then boiled at 95°C for 5 min to remove any higher-order structures that might interfere with migration through the gel. Approximately 10-30 µg of cell lysate or 10-100 ng of purified protein was loaded in each well, alongside a pre-stained protein ladder (PageRuler, Thermo Scientific). Gels were run at 100 V for 50-70 min, depending on the size of the target protein, and then either stained with Coomassie Brilliant Blue (Sigma) for immediate visualisation, or transferred to a PVDF membrane for western blotting (Section **2.4.9**).

2.4.9 Electro-transfer to a PVDF membrane

If using an SDS gel for western blotting, the proteins were first transferred to an Immobilon-FL PVDF membrane (Thermo Scientific) in SDS-free transfer buffer (5.8 g Tris, 2.9 g glycine, 200 mL methanol, 800 mL ddH_2O) at 20 V for 90 min. The membrane was then blocked in 1:1 PBS and Odyssey blocking buffer (LI-COR) for either 1 h at room temperature, or overnight at 4°C.

2.4.10 Western blotting by infrared detection

To enable the precise quantitation of a protein band, an infrared detection system was used in place of chemiluminescence.

Briefly, a blocked membrane was incubated with primary antibody diluted in Odyssey blocking buffer (LI-COR) supplemented with 0.2% Tween-20 for 1-3 h at room temperature, or overnight at 4°C. The membrane was then washed in PBS supplemented with 0.1% tween (PBS-T) for 4 rounds of 5 min with gentle agitation. Next, the membrane was incubated for 1 h at room temperature with the appropriate secondary antibody, diluted in the same buffer as the primary but with the addition of 0.01% SDS to reduce non-specific binding. The membrane was then washed 4 times in PBS-T as before. A list of primary and secondary antibodies used in this project can be found in Table **2.3**: List of antibodies used for western blots

Antibody	Source	Company	Dilution
Anti-HA	Rabbit polyclonal	Santa Cruz, Y-11, sc805	1:200
Anti-GAPDH	Rabbit monoclonal	Cell Signalling, 14C10, #2118	1:2000
Anti-rabbit-800	Goat polyclonal IRDye 800CW conjugated	LI-COR, #926-32211	1:10,000

Table 2.3: List of antibodies used for western blots

Because the secondary antibody used in this system is conjugated to an infrared dye, the membrane can be visualised using the 700 and 800 nm channels of a high-resolution infrared scanner (LI-COR). This method achieves more accurate quantitation than chemiluminescent detection because of its greater sensitivity and the improved linearity between signal intensity and amount of protein. Quantitation was performed by normalisation to a loading control using the Image Studio Lite software (LI-COR).

Chapter 3 Characterising retroviral restriction by T1

T5α has been extensively characterised since its first identification as a restriction factor, more than ten years ago (Stremlau et al., 2004). However, this molecule is far from the only TRIM family member to exhibit such a phenotype. In fact, at least 19 others have been shown to restrict divergent retroviruses at distinct stages in their lifecycles (Uchil et al., 2008).

T1 is one such factor. Its ability to inhibit N-MLV replication was first identified by the Stoye group (Yap et al., 2004, 2005), but this observation has not been characterised since. The aim of this chapter is thus to offer an in-depth analysis of the T1 restriction phenotype.

3.1 Murine T1 restricts N-MLV comparably to its primate orthologues

Previous work has shown that both human (hu) and African green monkey (agm) T1 are able to restrict N-MLV (Yap et al., 2004; 2005). However, prior to this work, it was not known whether this phenotype is shared by T1 orthologues of non-primate origin. To investigate this possibility, a cDNA expressing murine (m) T1 (Source Bioscience) was cloned into the LxIY expression vector and its restriction activity measured via the two-colour assay. When challenged with B- and N-MLV, mT1 restricted the latter but not the former (Figure **3.1**). This phenotype is reminiscent of that already described for the primate variants, confirming that the restriction activity of T1 indeed extends to other mammals.

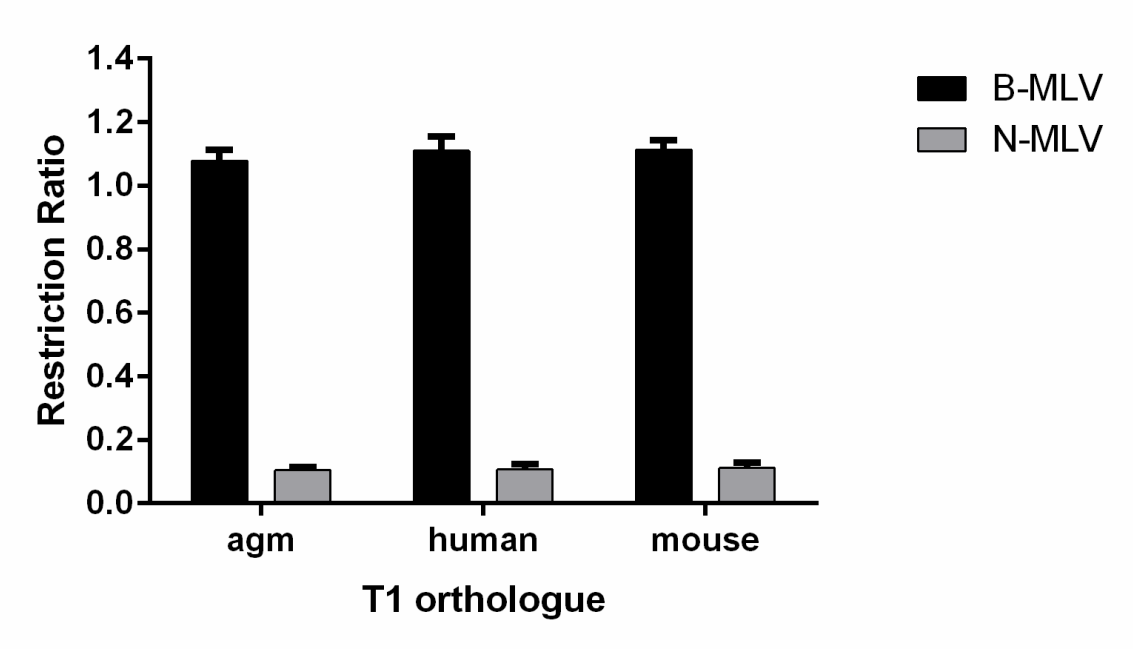


Figure 3.1: Restriction of N-MLV by the African green monkey (agm), human and murine orthologues of T1

Graph shows the mean and SEM from three independent experiments (*n* = 6).

3.2 T1 restricts a limited panel of retroviruses

It is well established that T5α contains signatures of positive selection, and is correspondingly able to restrict divergent retroviruses (Sawyer et al., 2005; Johnson and Sawyer, 2009). To investigate whether T1 has a similarly broad target range, the restriction of a panel of lentiviruses (Figure **3.2**) and foamy viruses (Figure **3.3**) by two isoforms of agmT1 (T1L and T1S) was examined via the two-colour assay. However, none of these viruses were restricted by either isoform, limiting the panel of known targets for T1 to N-MLV alone.

One explanation for this limited target range might be the integral role that T1 plays in the development of embryonic midline structures. Given the biologically fundamental nature of this process, it's highly likely that the *T1* gene has been subjected to negative selection over the course of its evolution. This would promote sequence conservation at the protein level, in stark contrast to the extensive diversification seen among T5α orthologues.

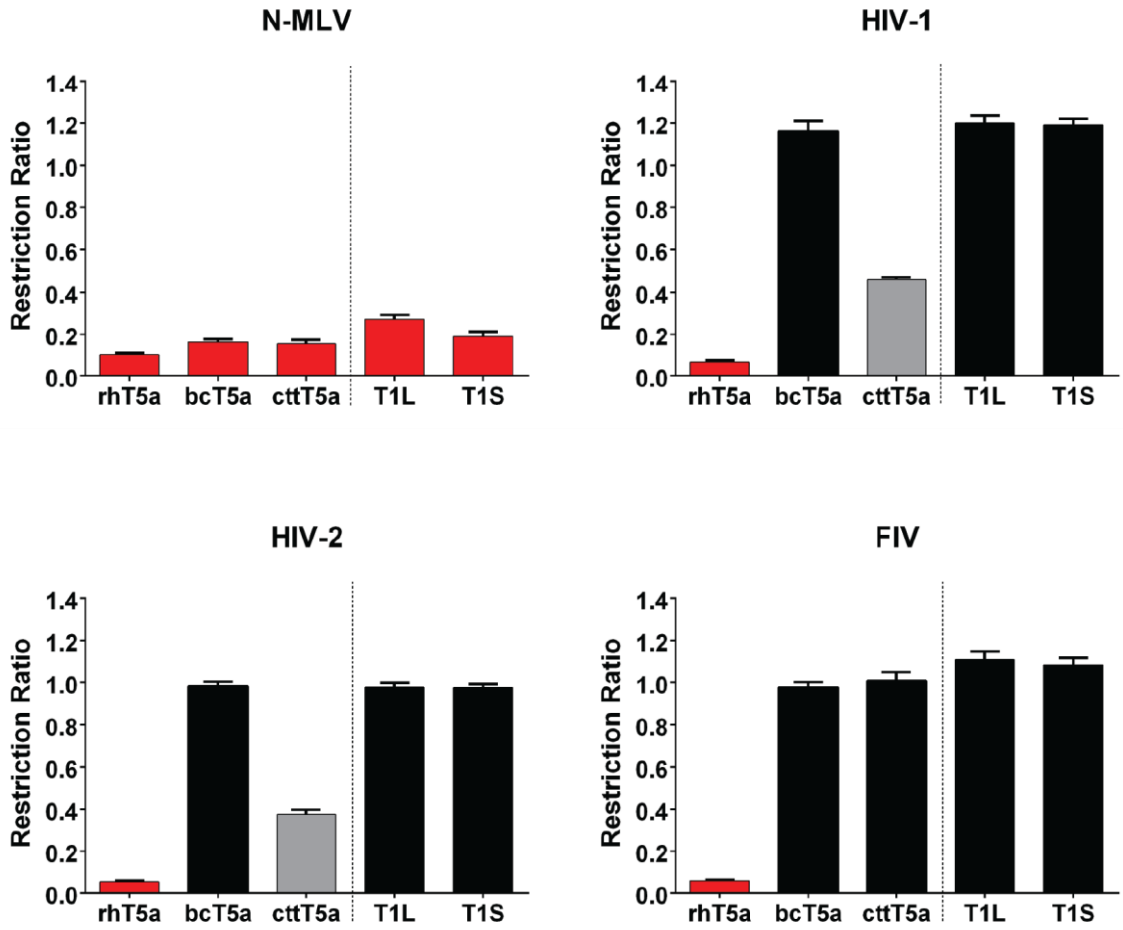


Figure 3.2: T1 is unable to restrict a panel of lentiviruses

Graphs show the restriction ratios obtained by challenging T1L and S from the African green monkey (*Chlorocebus aethiops*) with three lentiviruses: HIV-1, HIV-2 and FIV. N-MLV is included as a positive control (top left), and rhesus T5α (rhT5a), brown capuchin T5α (bcT5a) and cotton top tamarin T5α (cttT5a) are included as restriction factor controls. Red bars denote full restriction (ratio < 0.3), grey bars, partial restriction (ratio 0.3-0.7) and black bars, absence of restriction (ratio > 0.7). Graph for N-MLV shows the mean and SEM from four (rhT5a, T1L, T1S) or three (bcT5a, cttT5a) independent experiments (n = 8 and 6, respectively). Graphs for the lentiviruses show the mean and SEM from three (rhT5a, T1L, T1S) or two (bcT5a, cttT5a) independent experiments (*n* = 6 and 4, respectively).

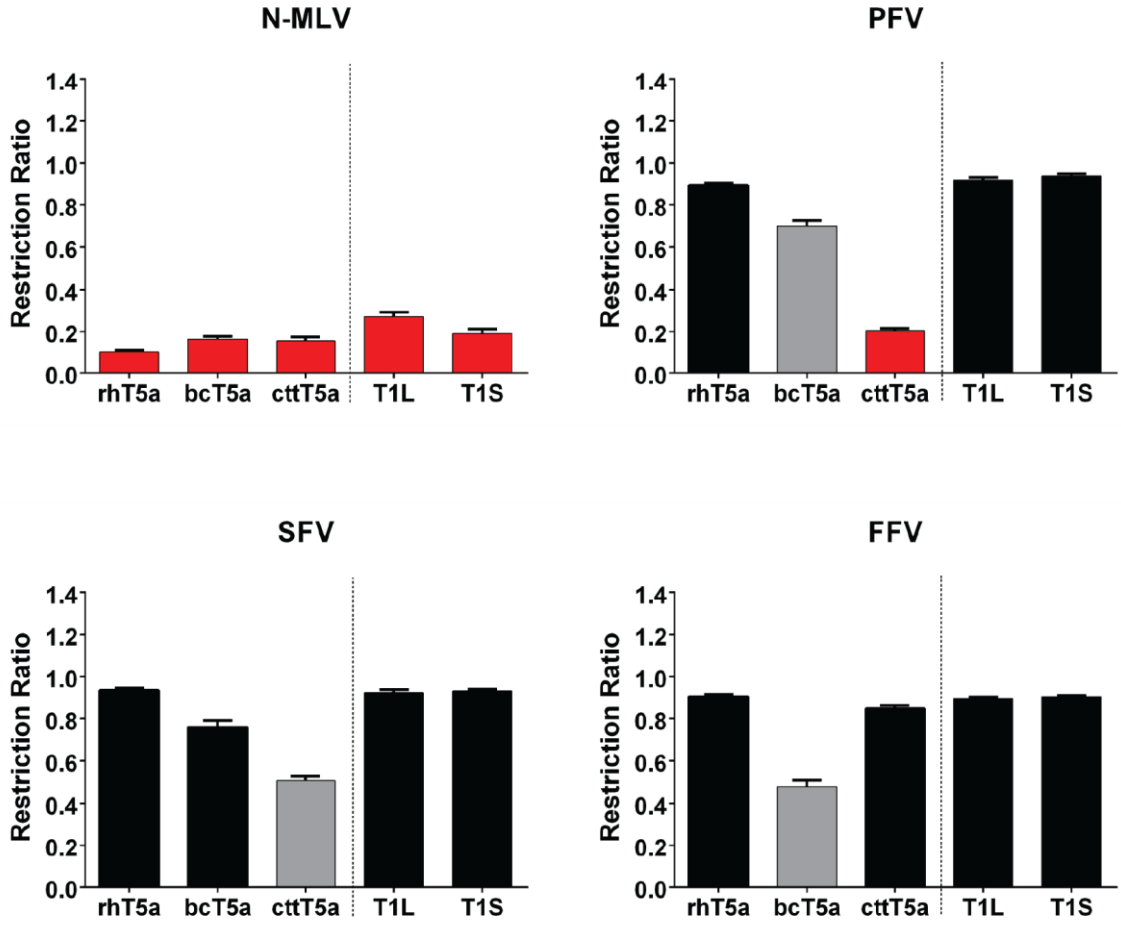


Figure 3.3: T1 is unable to restrict a panel of foamy viruses

Graphs show the restriction ratios obtained by challenging T1L and S from the African green monkey (*Chlorocebus aethiops*) with three foamy viruses: PFV, SFV and FFV. N-MLV is included as a positive control (top left), and rhesus T5α (rhT5a), brown capuchin T5α (bcT5a) and cotton top tamarin T5α (cttT5a) are included as restriction factor controls. Red bars denote full restriction (ratio < 0.3), grey bars, partial restriction (ratio 0.3-0.7) and black bars, absence of restriction (ratio > 0.7). Graph for N-MLV shows the mean and SEM from four (rhT5a, T1L, T1S) or three (bcT5a, cttT5a) independent experiments (n = 8 and 6, respectively). Graphs for the foamy viruses show the mean and SEM from four (PFV) or three (SFV, FFV) independent experiments (*n* = 8 and 7, respectively).

To determine whether *T1* is indeed constrained by negative selection, an alignment of T1 cDNAs from a host of mammals (Table **3.1**) was generated using PAML 4.8 (Yang, 2007) (**Error! Reference source not found.**). To avoid false positives, a 90 bp region present in certain splice variants was omitted from this analysis.

Primates	Human
	Chimpanzee
	Gorilla
	Orangutan
	Gibbon
	Rhesus macaque
	Baboon
	African green monkey
Rodents	Mouse
	Rat
	Kangaroo rat
	Guinea pig
	Squirrel

Table 3.1: T1 orthologues used to construct a phylogenetic tree

The selective pressures acting on a gene can be measured by calculating values for ω across all the sites in its DNA sequence, where ω equals the ratio of non-synonymous to synonymous substitutions at a particular site (dN/dS). $\omega \leq 1$ is indicative of negative selection, while $\omega > 1$ indicates positive selection. Codeml was used to conduct this analysis, with the help of Dr George Young (Figure **3.5**).

Briefly, ω values were calculated and then pairs of models – one allowing and the other disallowing positive selection – were applied to the values obtained. The models' relative fits, as likelihood ratio tests, were then compared using a χ^2 test (degrees of freedom = 2). The model pairs M2a vs M1a and M8 vs M7 were compared in this way, and each comparison performed using two methods of codon frequency calculation (F3x4 and F61). None of the alternative models provided a better fit to the data than the null hypothesis ($p \geq 0.584$ across all tests), confirming that *T1* is indeed highly conserved at the protein level. This might explain the limited ability of the molecule to recognise divergent retroviruses.

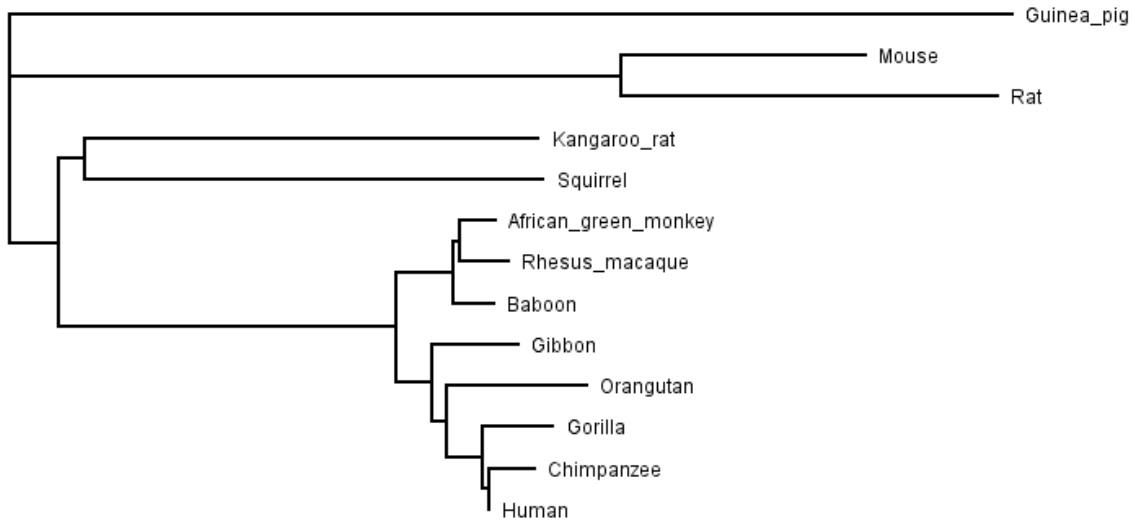
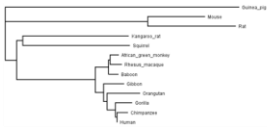


Figure 3.4: Phylogenetic tree of *T1* DNA sequences



1. Create a phylogenetic tree of DNA sequences.

CCGACTTCTCGT

2. Use this tree to deduce the ancestral state of each codon.

CCGACTTCTCGT
ω ω ω ω
CCGAGTTCTCGA
ω ω ω ω
CGGACTACTCGT

3. Align the modern sequences with the predicted ancestral one. Calculate ω values for each codon position using one or more methods of *codon frequency calculation*: F0 assumes that all codons are equally likely, F3X4 assigns ω values based on the average frequencies of nucleotides in a given codon, and F61 allows these parameters to be estimated freely. The latter two methods were used in our analysis of *T1*.

Neutral model (M1a, M7) [Null]
Allows two classes of sites:
Negatively selected ($0 < \omega < 1$), frequency p_0
Neutral sites ($\omega=1$), frequency p_1
$p_0 + p_1 = 1$

vs.

Selection model (M2a, M8) [Alternative]
Allows three classes of sites:
Negatively selected ($0 < \omega < 1$), frequency p_0
Neutral sites ($\omega=1$), frequency p_1
Positive sites ($\omega>1$), frequency p_2
$p_0 + p_1 + p_2 = 1$

4. Apply neutral and selection models to the ω values obtained. The M1a/M2a and M7/M8 pairs differ only in that the latter models assume a β probability distribution for ω values. This makes the M7/M8 pair less stringent, and thus more sensitive to positive selection.

$$D = -2\ln\left(\frac{\text{likelihood for null model}}{\text{likelihood for alternative model}}\right)$$

5. Use the 2x log-likelihood of each model to calculate a likelihood test ratio (*D*).

Degrees of Freedom	0.99	0.95	0.90
1	0.000	0.004	0.016
2	0.020	0.103	0.211
3	0.115	0.352	0.584
4	0.297	0.711	1.064
5	0.554	1.145	1.610
6	0.872	1.635	2.204
7	1.239	2.167	2.833
8	1.647	2.733	3.490

6. Compare *D* to a table of χ^2 values (at the appropriate degrees of freedom) to obtain a *p* value. Use the *p* value to either accept or reject the null hypothesis (i.e. the neutral model).

Figure 3.5: Workflow for measuring positive selection in *T1*

3.3 The short isoform of T1 restricts N-MLV more potently than the long

The *T1* gene undergoes alternative splicing to produce several transcripts, two of which encode proteins with a B30.2 domain. Due to an alternative splice donor site in the middle of exon 7, these isoforms differ exclusively in the length of their FN-III domains, with the longer one (T1L) possessing 30 amino acids in the centre of the domain (residues 430-459) that are lacking in the short (T1S) (Figure **3.6**). *T18* also undergoes alternative splicing, although its longest isoform has an FN-III domain of equivalent length to that of T1S.

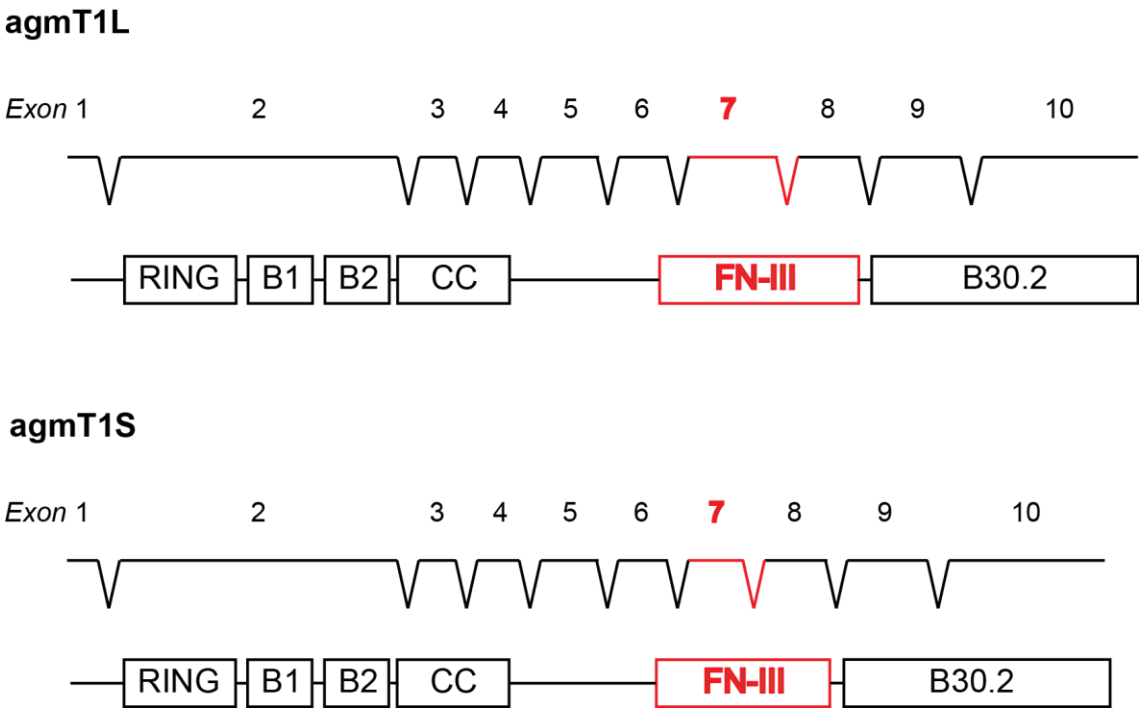


Figure 3.6: The intron-exon structures of agmT1L/S

At the start of this project, cDNAs for both agm isoforms were available in the lab, but only the short cDNAs were available for the human and mouse variants. Thus, long isoforms were produced for each of these orthologues by overlapping PCR, using the short cDNAs as templates, and a pair of internal primers with portions of the missing sequence appended to their 5' ends.

All six constructs were then cloned into the LxIY expression vector and their restriction activity measured via the two-colour assay. Figure **3.7** shows the FACS plots from a typical experiment. The plots clearly depict a more potent restriction of N-MLV by the short isoform than the long in all three cases. This observation was highly reproducible, reaching statistical significance for each orthologue tested (Figure **3.8A**). Again, B-MLV was not restricted in any case and served as a negative control in these experiments (Figure **3.8B**).

The FN-III domain is a protein-protein interaction module of about 100 amino acids (in T1S). It possesses a hydrophobic β-sandwich core, with flexible loops that protrude from either side to form the interaction surfaces. The domain is present in a range of animal proteins and shows remarkable evolutionary conservation at the structural level (Marino-Buslje et al., 1998). Since protein-protein interactions are pivotal for a number of biochemical phenomena, there are numerous mechanisms that might explain the relationship between the length of the FN-III domain and restriction activity.

To investigate whether this phenotype is the result of differences in protein expression, a C-terminal HA tag was appended to each of the agm isoforms to enable their detection and quantitation by western blotting. The tagged proteins restricted N-MLV comparably to their native counterparts (Figure **3.9A**).

The tagged constructs were transduced into MDTF cells to yield approximately 20-40% YFP+ cells (Figure **3.9B**), and the cells grown to about 70% confluency, before harvesting total protein for a western blot. As Figure **3.9C** shows, a slightly darker band was present for T1S-HA than T1L-HA; normalisation of these bands indicates that expression of the former is almost double that of the latter. Collectively, these data imply a correlation between protein concentration and restriction potency, which is underpinned by determinants in the FN-III domain.

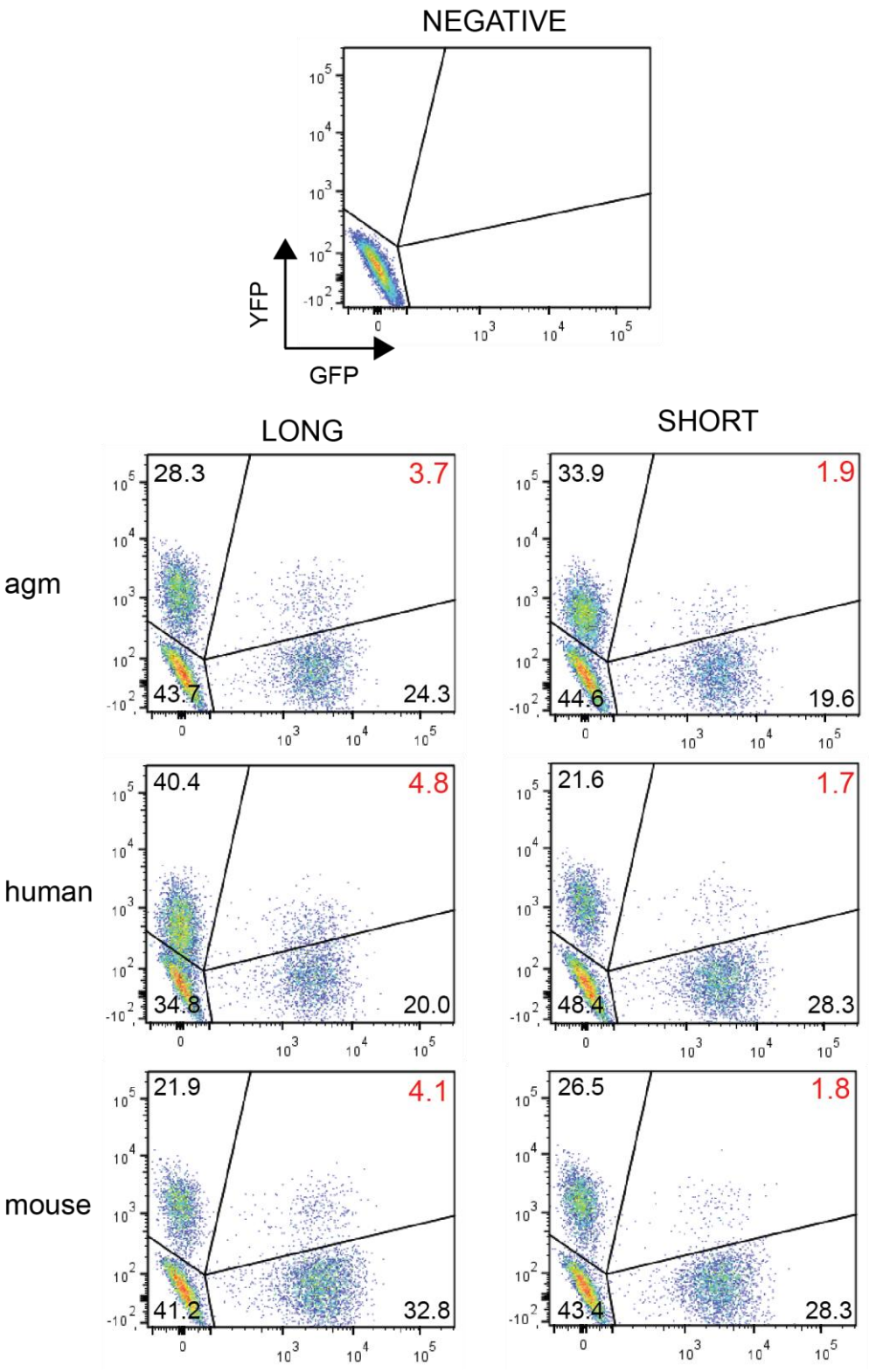


Figure 3.7: Typical FACS plots obtained when challenging the agm, human and murine orthologues of T1L/S with N-MLV

Values indicate the percentage of live cells that are present in each gate. The proportion of double-positive cells (i.e. $T1^+$ cells infected with N-MLV) is highlighted in red to emphasise the difference in potency between the two isoforms.

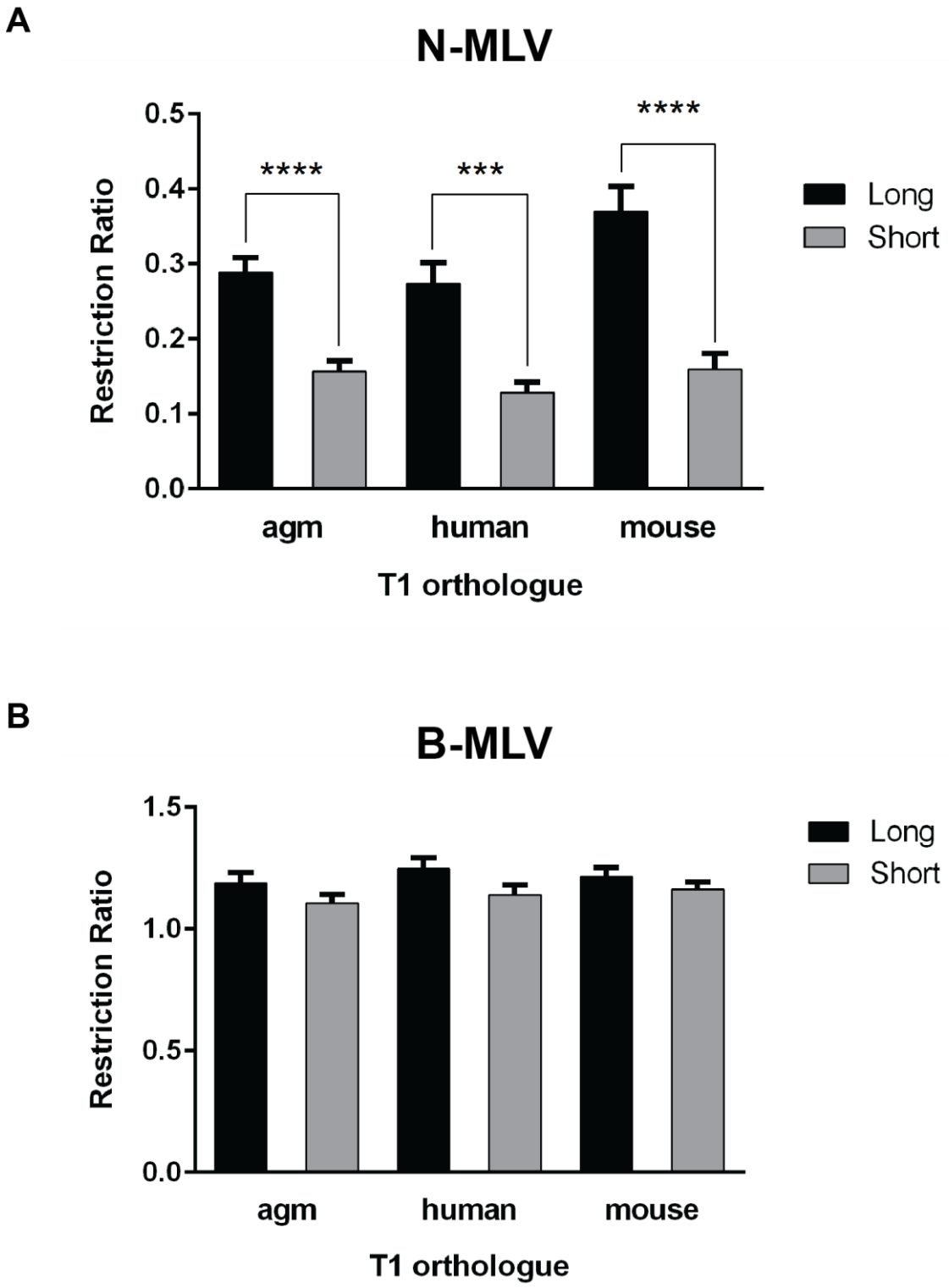


Figure 3.8: Restriction profiles of agm, human and murine T1L/S

Graph shows the mean and SEM from **(A)** at least five or **(B)** three independent experiments ($n \geq 10$ and $n = 6$, respectively). Statistical significance was established using an unpaired t-test: **** $P \leq 10^{-4}$ and *** $P \leq 10^{-3}$.

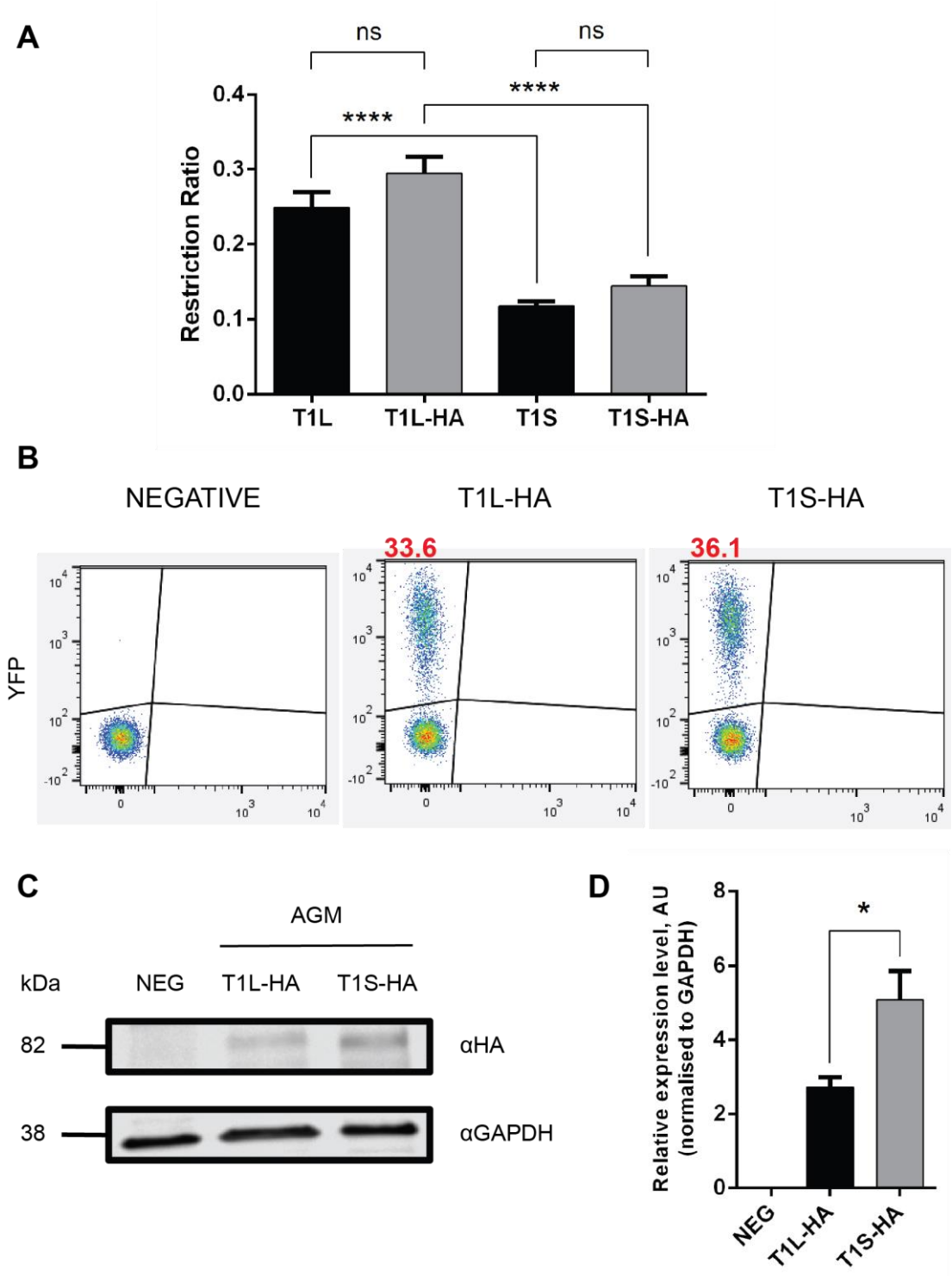


Figure 3.9: Quantitation of T1L/S protein expression

(A) HA-tagged T1L/S restrict N-MLV comparably to their native counterparts. Graph shows the mean and SEM from three independent exp'ts (*n* = 10). **(B)** FACS plots of the cell populations harvested for western blotting. **(C)** Representative western blot of T1L/S-HA, using GAPDH as a loading control. **(D)** Quantitation of bands obtained by western blotting. Graph shows the mean and SEM from three independent exp'ts (*n* = 3). In all cases, statistical significance was established using an unpaired t-test: **** $P \leq 10^{-4}$, * $P \leq 0.05$ and *ns* = not significant.

3.4 The majority of the T1 B30.2 domain can be functionally replaced with equivalent components from T18

It has long been known that T5α interacts with capsid via surface-exposed loops known as variable regions, which reside within the C-terminal B30.2 domain (Ohkura et al., 2006). Before this work, however, very little was known regarding the interaction between T1 and capsid. Previously, reciprocal chimeras of T1S and its non-restricting paralogue, T18, had been made at position 314 (Yap et al, 2004). Restriction data from these constructs revealed that the determinants for capsid recognition lie in the C-terminal half of the protein; however, given that the B30.2 domain is over 150 residues downstream of 314 (see Figure **3.10**), this work lacked sufficient resolution to identify the specific regions that confer this property.

Thus, an early aim of this project was to narrow down the regions within T1 that enable capsid recognition. To this end, reciprocal swaps were made between T1 and 18 again, but this time at the start of the B30.2 domain, and the restriction phenotypes of the resulting chimeras were measured. As expected, the T1 B30.2 domain was sufficient to confer T18 with activity, while the reciprocal construct was inactive (Figure **3.11**). This rules out any contribution of the COS and FN-III domains to capsid recognition.

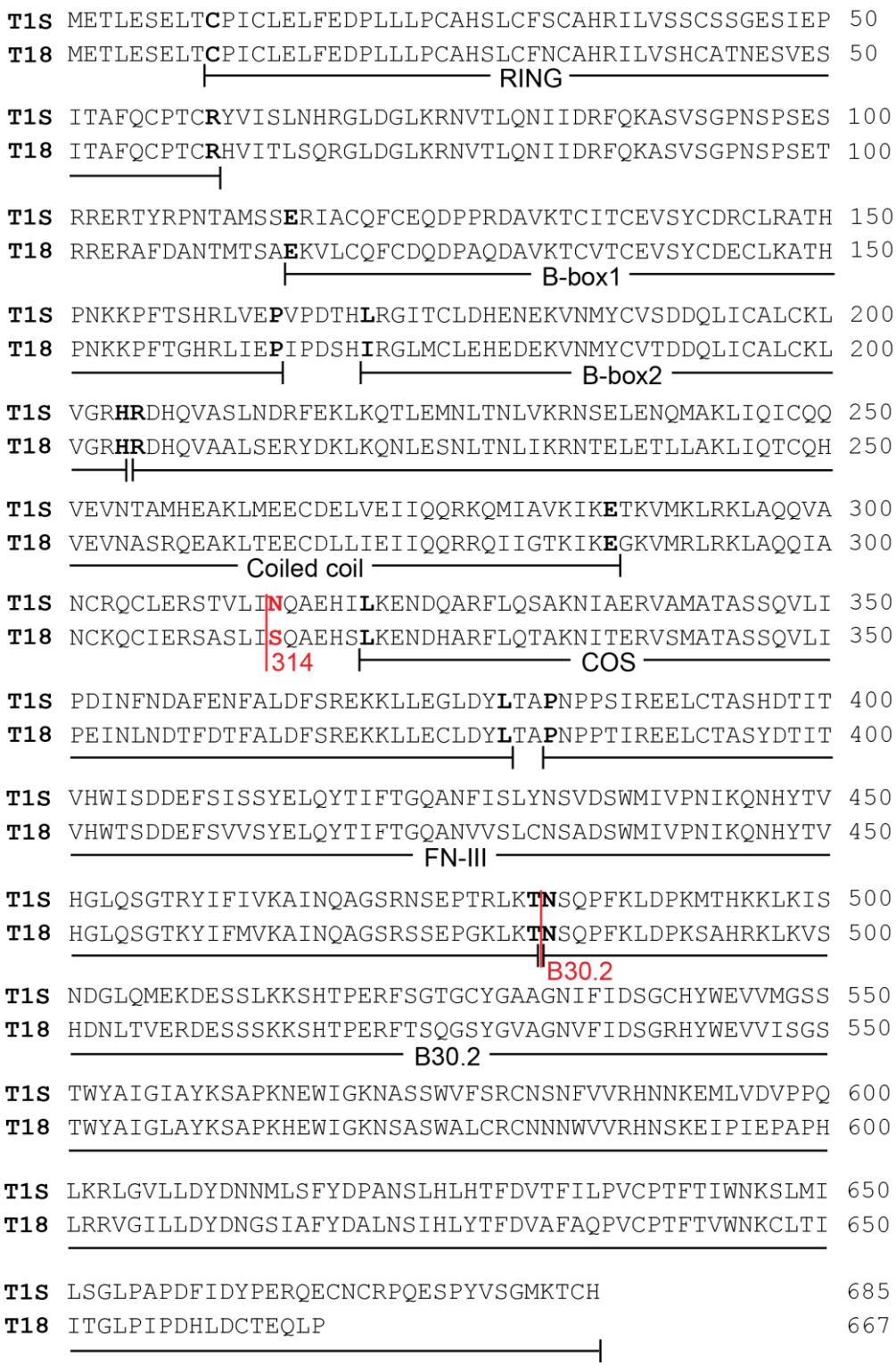


Figure 3.10: An alignment of the T1 and 18 protein sequences

The short isoform of T1 (T1S) was used for this alignment; both sequences are from the agm orthologues. Boundaries for the RBCC and B30.2 domains were established using the UniProt proteomics database, while those for the COS and FN-III domains were taken from Short and Cox (2006) and Li et al. (2016a), respectively. The positions at which the 314 and B30.2 chimeras were made are indicated in red.

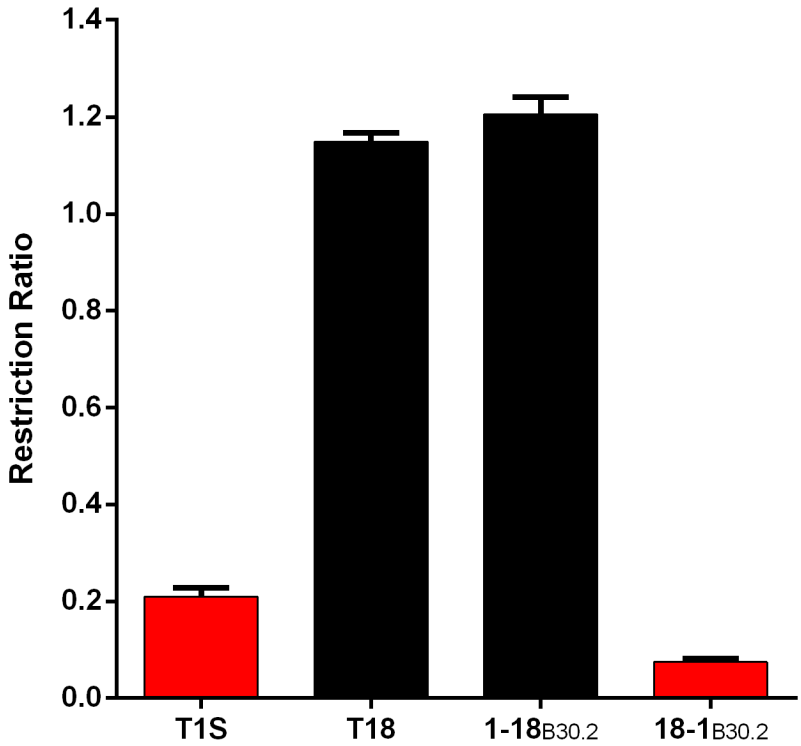


Figure 3.11: Restriction activities of B30.2-swapped chimeras of T1S and T18

Restriction of N-MLV by T1S, T18 and chimeras where the B30.2 domain is exchanged between these proteins. Graph shows the mean and SEM from five independent experiments (*n* = 10). Red bars denote full restriction (ratio < 0.3) and black bars, absence of restriction (ratio > 0.7).

Having confirmed that capsid recognition is indeed mediated by the B30.2 domain, the next goal was to assess the contribution of the variable regions ('VRs' herein) to this process. In order to do this, the boundaries of each VR were established using a published alignment of TRIM sequences (James et al, 2007) (Figure **3.12**). The T1 VRs were then individually substituted into 1-18$_{B30.2}$ in an attempt to rescue restriction (Figure **3.13**, Constructs **A-D**). Interestingly, however, none of these components were individually sufficient to confer capsid recognition. Following this observation, various combinations of T1 VRs were then substituted into 1-18$_{B30.2}$ in the following manner: VR1 and 2 (**E**); 2 and 3 (**F**); 1, 2 and 3 (**G**); and 1, 2, 3 and 4 (**H**). Surprisingly, however, these constructs were also inactive (Figure **3.13**).

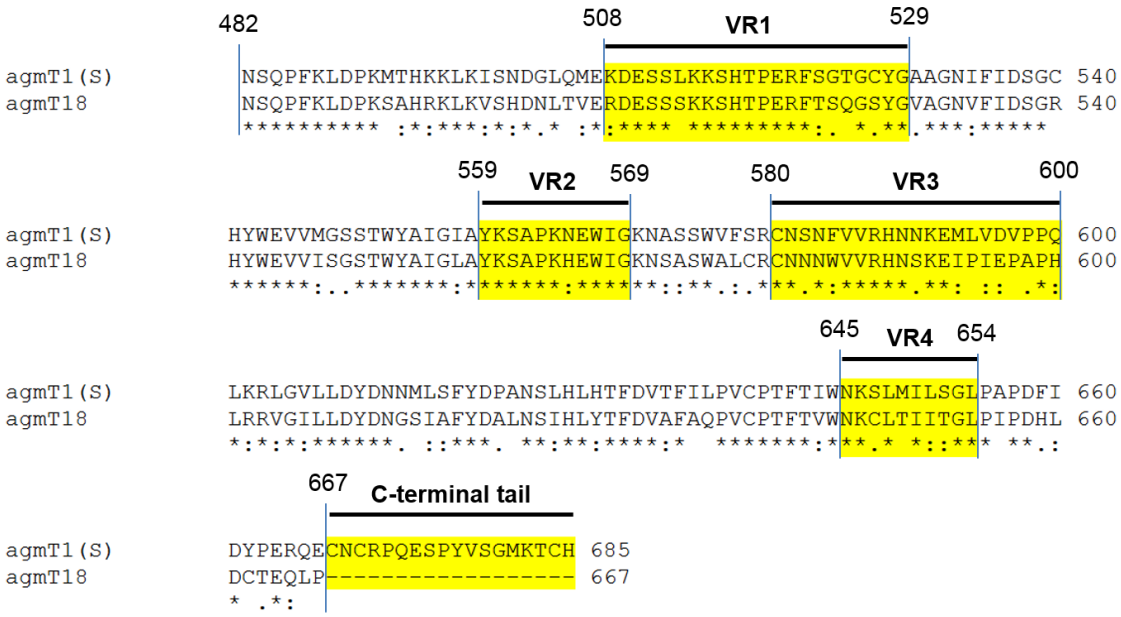


Figure 3.12: An alignment of the T1 and 18 B30.2 domains

VR boundaries were established by comparing the T1 sequence with a published alignment of TRIM proteins, including T5α and T18, in which the topology of the B30.2 domains had been established (James et al., 2007).

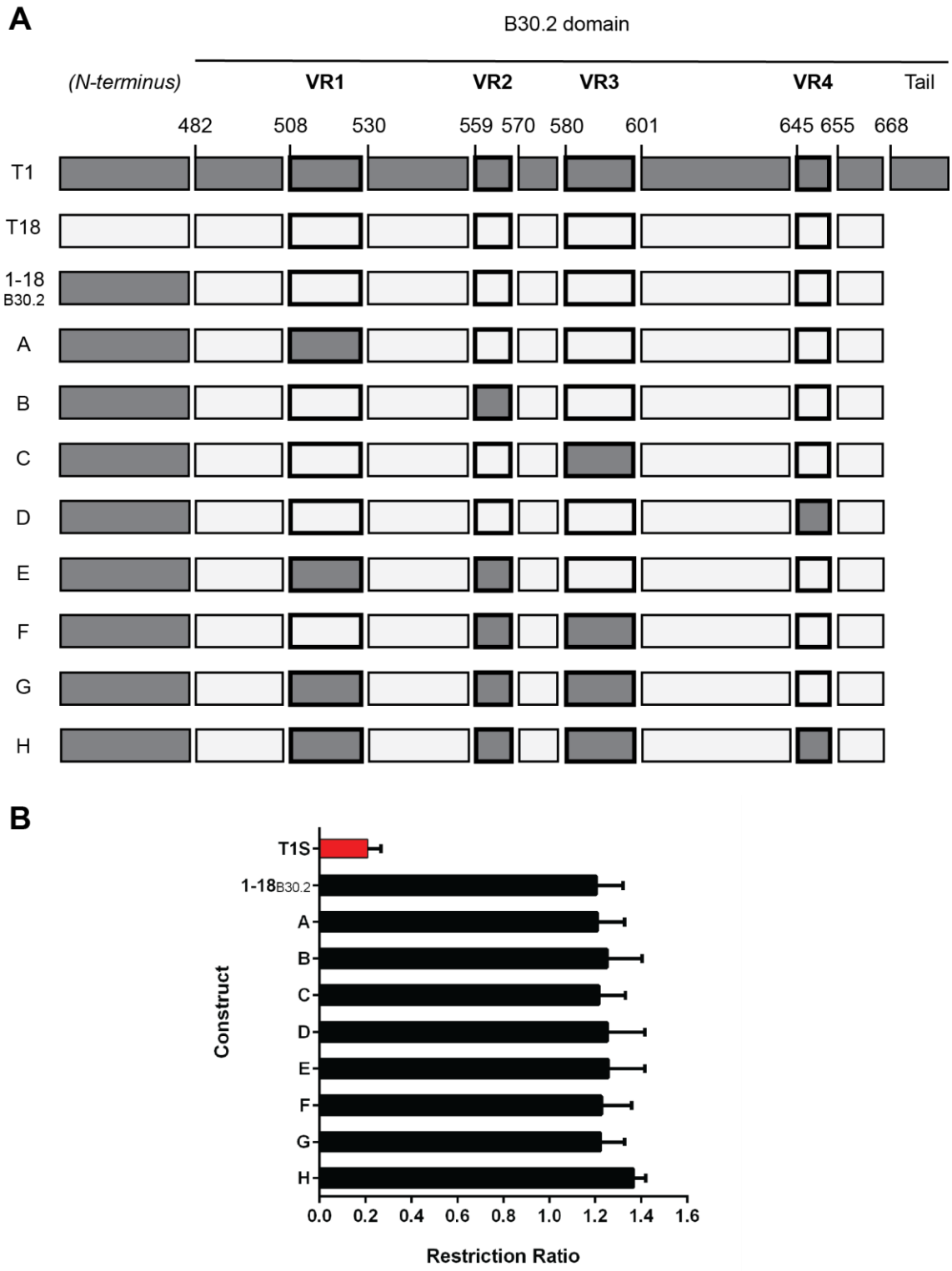


Figure 3.13: Single and combinatorial substitutions of the T1 VRs into 1-18$_{B30.2}$

(A) The structures of Constructs A-H (Section **3.4**). **(B)** Restriction of N-MLV by Constructs A-H. Graph shows the mean and SEM from five (T1S, 1-18$_{B30.2}$), three (A-G) or two (H) independent experiments (n = 10, 6 and 4, respectively). Red bars denote full restriction (ratio < 0.3) and black bars, absence of restriction (ratio > 0.7).

In addition to the VR sequences, the T1 and 18 B30.2 domains diverge at the C-terminus: T1 possesses an 18-residue C-terminal tail (CT) that is absent in T18. To investigate the possibility that the CT is required for restriction, the relevant sequence was removed from T1S and 18-$1_{B30.2}$, and appended to T18. However, the former two molecules remained restriction-competent without the tail, and the latter remained inactive despite it (Figure **3.14**). This indicates that the CT is neither necessary nor sufficient for capsid recognition.

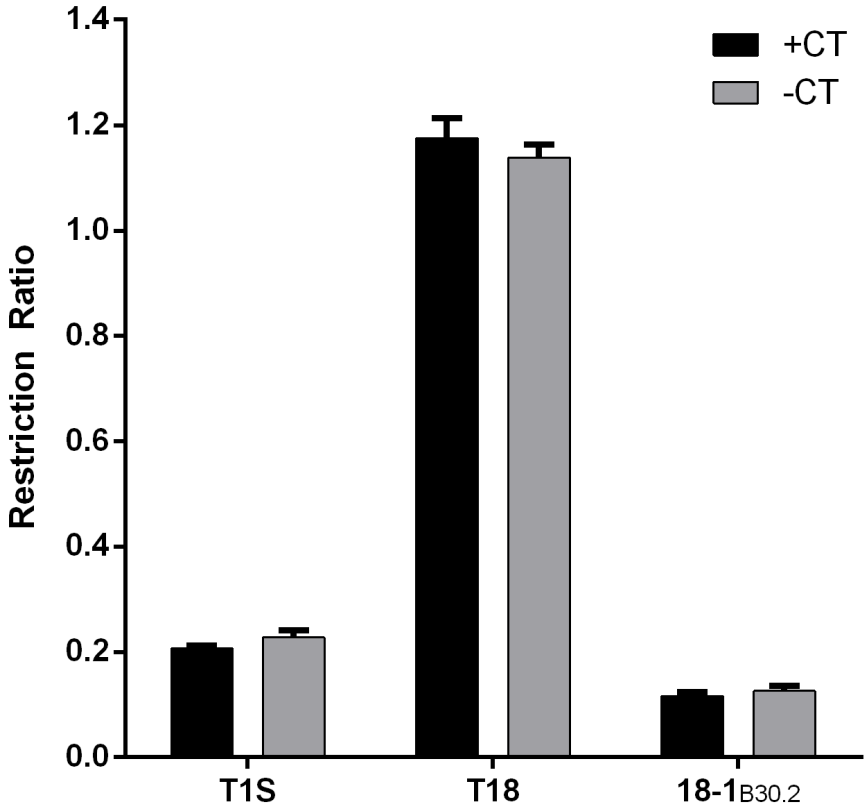


Figure 3.14: The C-terminal tail (CT) bears no impact on restriction
Graph shows the mean and SEM from three independent experiments (*n* = 6).

To account for the possibility that regions outside of the VRs and CT contribute to capsid recognition, a panel of chimeras were generated where bulk regions of T1 B30.2 were substituted into 1-$18_{B30.2}$ (Figure **3.15**, Constructs **I-Q**). Surprisingly, only three of these molecules – **N, P** and **Q** – acquired restriction activity.

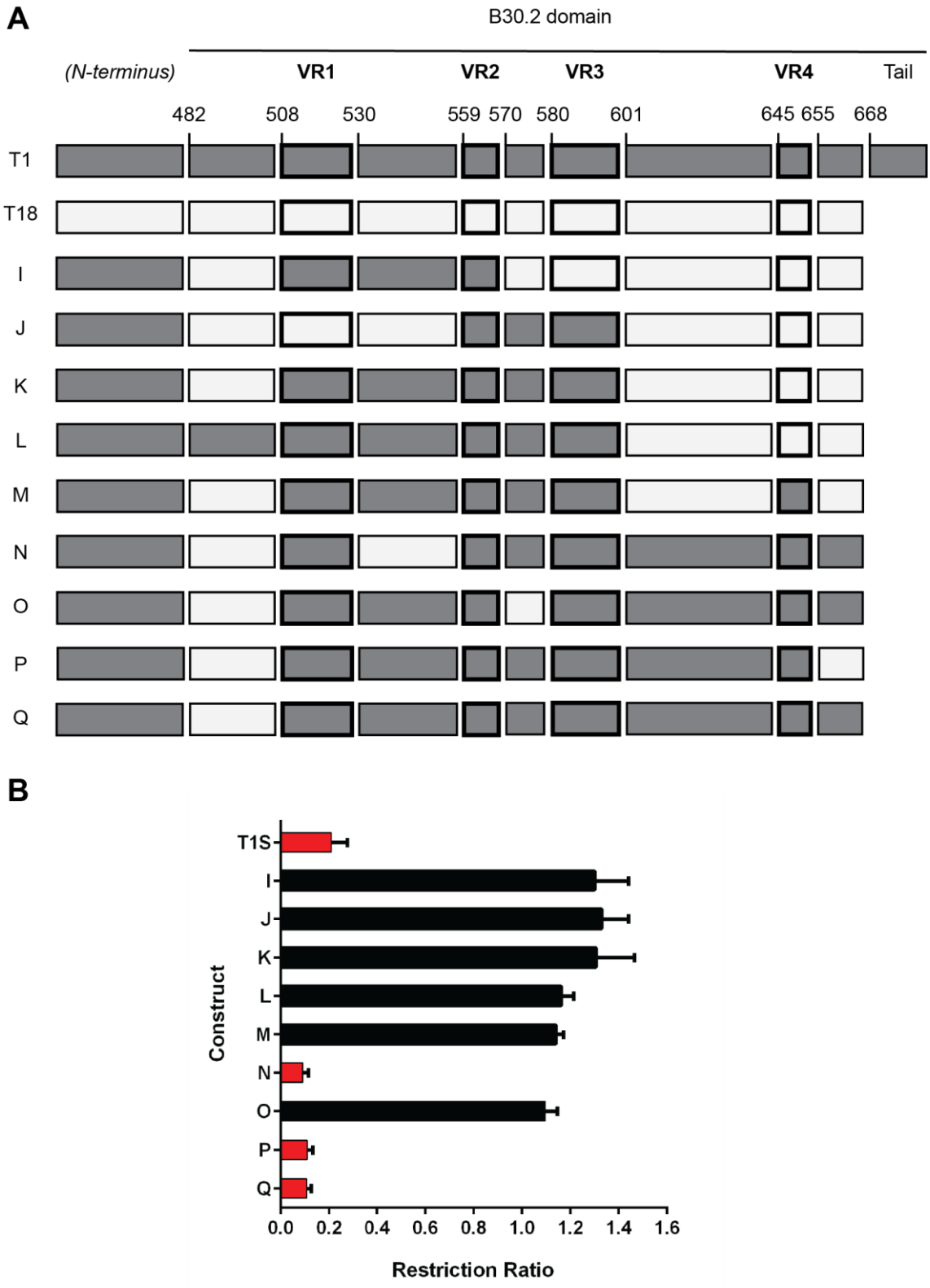


Figure 3.15: Bulk substitutions of the T1 B30.2 domain into 1-18$_{B30.2}$

(A) The structures of Constructs I-Q (Section **3.4**). **(B)** Restriction of N-MLV by Constructs I-Q. Graph shows the mean and SEM from four (T1S, Q), three (N-P) or two (I-M) independent experiments (*n* = 8, 6 and 4, respectively). Red bars denote full restriction (ratio < 0.3) and black bars, absence of restriction (ratio > 0.7).

In parallel to the above experiments, the T18 VRs were substituted into the 18-$1_{B30.2}$ chimera (Figure **3.16**, Constructs **R-U**). Interestingly, the presence of VR1, 3 or 4 from T18 had no effect on N-MLV restriction by this molecule, although this phenotype was attenuated by T18 VR2 (Construct **S**). This latter observation is given more attention in Section **3.5**.

Collectively, these data indicate that the majority of the T1 B30.2 domain can be functionally replaced with the equivalent regions from T18, with the exception of the region flanked by VRs 2 and 3 (Table **3.2**). The only other region that might contribute to this phenotype is the one between VRs 3 and 4. Unfortunately, it was not possible to examine this region within the timeframe of this project; however, a suitable construct has been synthesised at the time of writing and is currently being tested in the Stoye lab. However, since all of the constructs that lacked this portion of T1 were inactive, it is highly likely that this component is also necessary for capsid recognition.

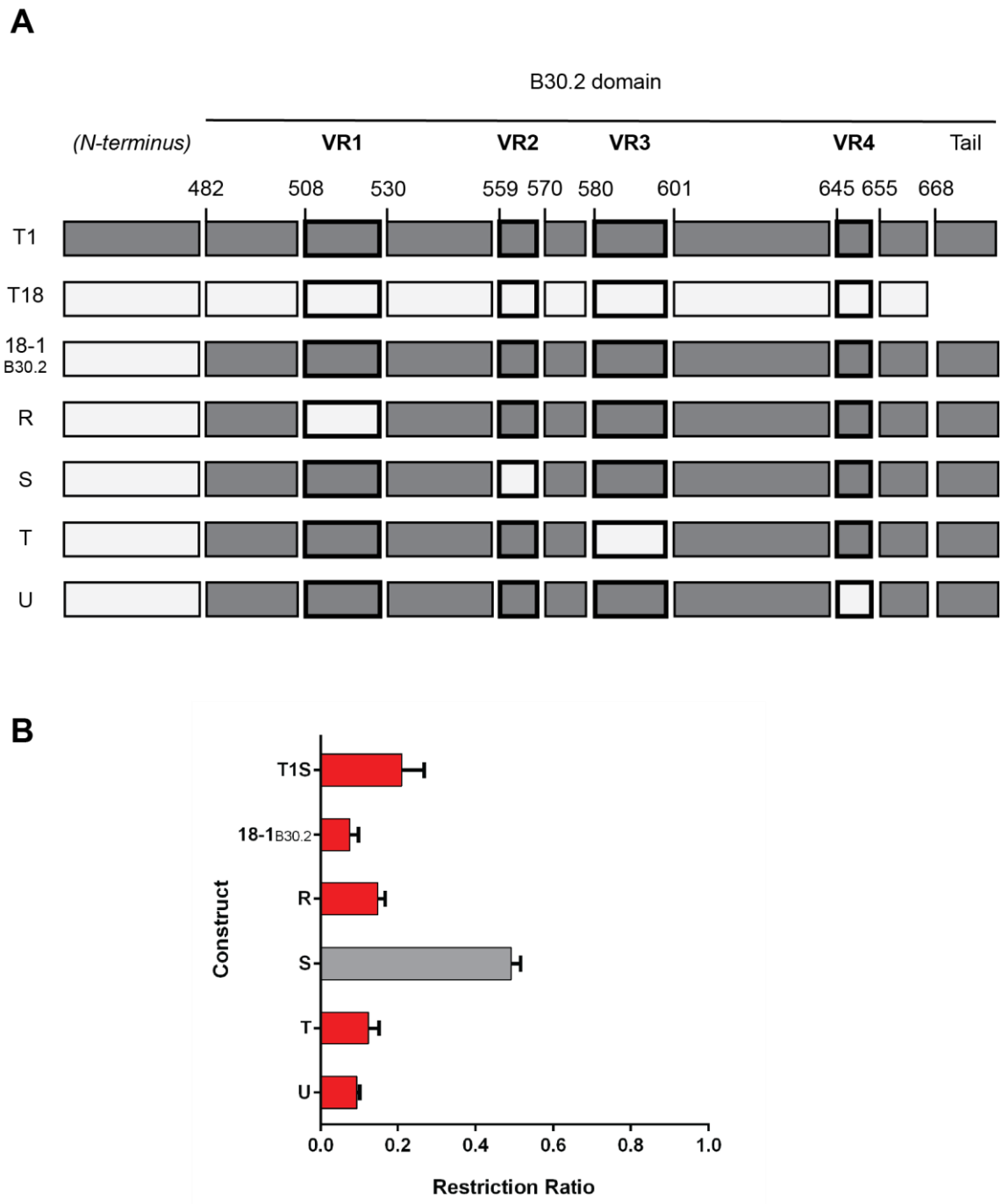


Figure 3.16: Single substitutions of the T18 VRs into 18-1$_{B30.2}$

(A) The structures of Constructs R-U (Section **3.4**). **(B)** Restriction of N-MLV by Constructs R-U. Graph shows the mean and SEM from five (T1S, 18-1$_{B30.2}$) or two (R-U) independent experiments (*n* = 10 and 4, respectively). Red bars denote full restriction (ratio < 0.3) and grey bars, partial restriction (ratio 0.3-0.7).

Region in T1 B30.2	Does equivalent region from T18 preserve function?	Construct(s) that demonstrate this	Cross-reference
Pre-VR1	Yes	**N, P** and **Q**	Figure **3.15**
VR1	Yes	**R**	Figure **3.16**
VR1-2	Yes	**N**	Figure **3.15**
VR2	Partially	**S**	Figure **3.16**
VR2-3	No	**O**	Figure **3.15**
VR3	Yes	**T**	Figure **3.16**
VR3-4	Unknown	**–**	**–**
VR4	Yes	**U**	Figure **3.16**
VR4-CT	Yes	**P**	Figure **3.15**
CT	Yes	T1S and 18-1$_{B30.2}$ ΔCT; **N, P** and **Q**	Figure **3.14** and Figure **3.15**

Table 3.2: The majority of the T1 B30.2 domain can be functionally replaced with equivalent regions from T18

CT = C-terminal tail.

3.5 T1 residue 595 is an important determinant of N-MLV capsid recognition

The data presented in Section **3.4** showed that T18 VR2 can attenuate restriction in the context of a T18-$1_{B30.2}$ chimera (Figure **3.16**, Construct **S**). Interestingly, the VR2 sequences of T1 and 18 differ by only a single amino acid: an asparagine in the former (T1L N595, T1S N565) is replaced by a histidine at the corresponding position in the latter (T18 H565).

To determine whether this phenotype is unique to the 18-$1_{B30.2}$ chimera, a single amino acid substitution was made in each of the wild-type T1 isoforms to produce T1L N595H and T1S N565H. Curiously, this modification yielded an even more dramatic phenotype: N-MLV restriction was lost in both cases, albeit to a greater extent for T1L (Figure **3.17**). As expected, the reciprocal construct (T18 H565N) did not acquire restriction activity.

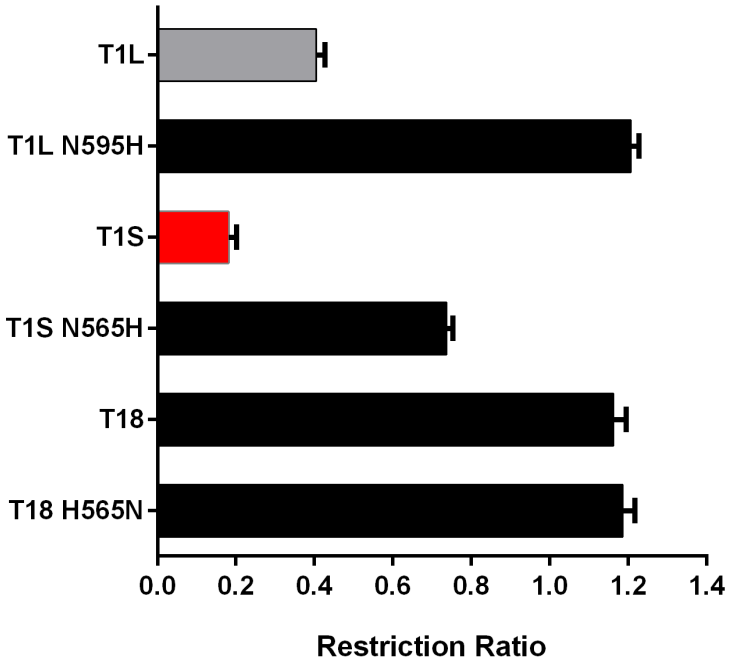


Figure 3.17: H595 (T1L), 565 (T1S) is sufficient to inhibit the restriction of N-MLV

Graph shows the mean and SEM from three (T1L/S constructs) or two (T18 constructs) independent experiments (*n* = 6 and 4, respectively). Red bars denote full restriction (ratio < 0.3); grey bars, partial restriction (ratio 0.3-0.7) and black bars, absence of restriction (ratio > 0.7).

Collectively, these data suggest that a histidine at this position weakens capsid recognition, and that this attenuation is partly relieved by the N-terminus of T18 (the latter phenomenon is discussed further in Section **3.6**).

To verify whether this effect is unique to histidine, a panel of mutants were generated where each of the remaining 18 standard amino acids were introduced at position 595 in T1L, and the effect on restriction measured. Interestingly, T1L was highly sensitive to these substitutions, with the vast majority completely inhibiting restriction activity (Figure **3.18**). In fact, only the acidic residues, D and E permitted restriction of N-MLV to a degree approaching that of wild-type.

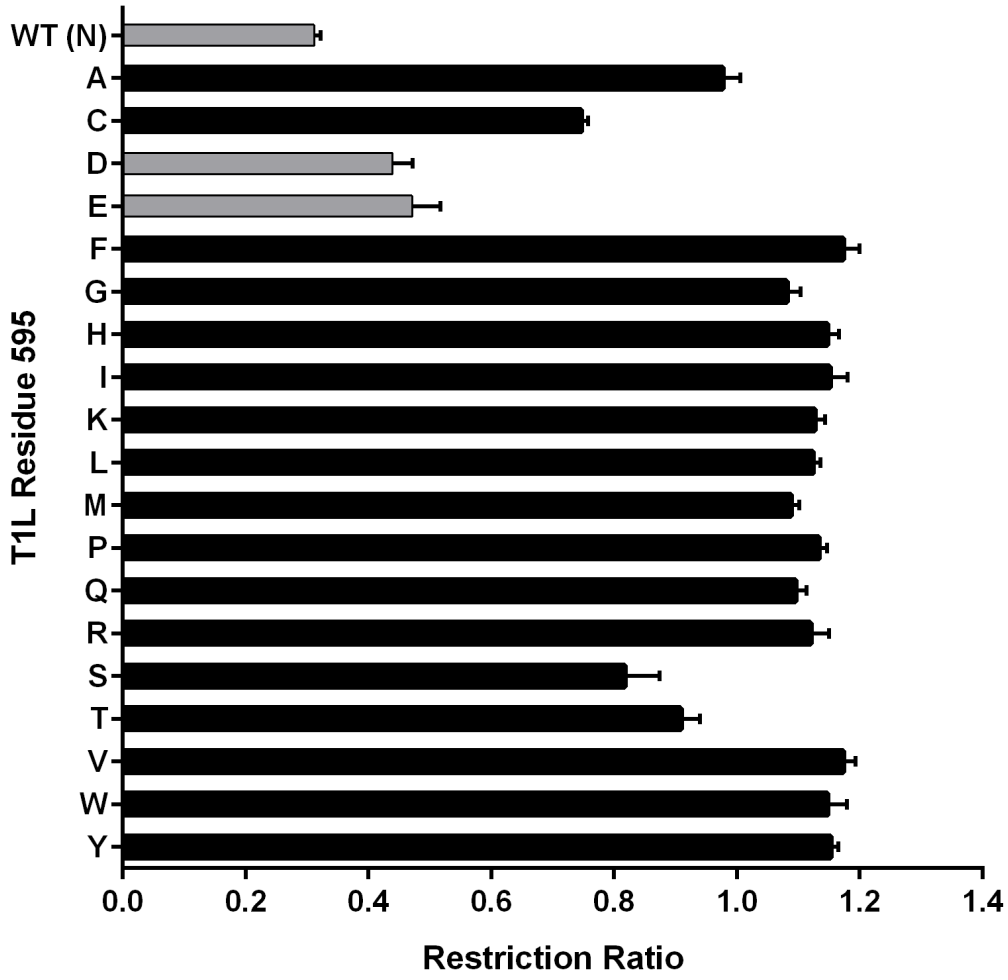


Figure 3.18: Restriction of N-MLV by T1L N595 mutants

Graph shows the mean and SEM from three independent experiments (*n* = 6). Grey bars denote partial restriction (ratio 0.3-0.7) and black bars, absence of restriction (ratio > 0.7).

3.6 The restriction of N-MLV by T1 is affected by N-terminal components

In Section **3.4**, the inhibitory effect of H595/565 in T1L/S was shown to be partially rescued by the T18 N-terminus. This phenomenon is corroborated by a previous observation by Yap et al. (2005), where a chimera of T18 and T1S made at position 314 (18-1_{314}) was shown to restrict N-MLV more potently than wild-type T1S. Given that this swap excludes the COS and FN-III domains, the only T18-derived regions that could possibly contribute to this phenotype are the RING, B-box and/or coiled-coil domains.

To identify which of these regions is responsible for the enhanced restriction potency of 18-1_{314}, the RING, B-box and coiled-coil domains of T1 were substituted into this molecule, first individually (Figure **3.19**, Constructs **A-C**), then in combination (**D-G**), and finally as a single unit (**H**). The majority of these molecules saw a return to the weaker phenotype of wild-type T1S – with the exception of those that retained the T18 B-boxes (**A, C** and **E**), indicating that this region is responsible for the enhanced potency of 18-1_{314}.

To verify this observation, the T18 B-boxes were introduced into full-length T1L and -S by overlapping PCR to produce T1L-B18 and T1S-B18. As expected, restriction was augmented in both cases, especially for T1L (Figure **3.19C**). This confirms that the T18 B-boxes, when paired with a T1 B30.2 domain, yield a molecule with a restrictive potential similar to that of T5α.

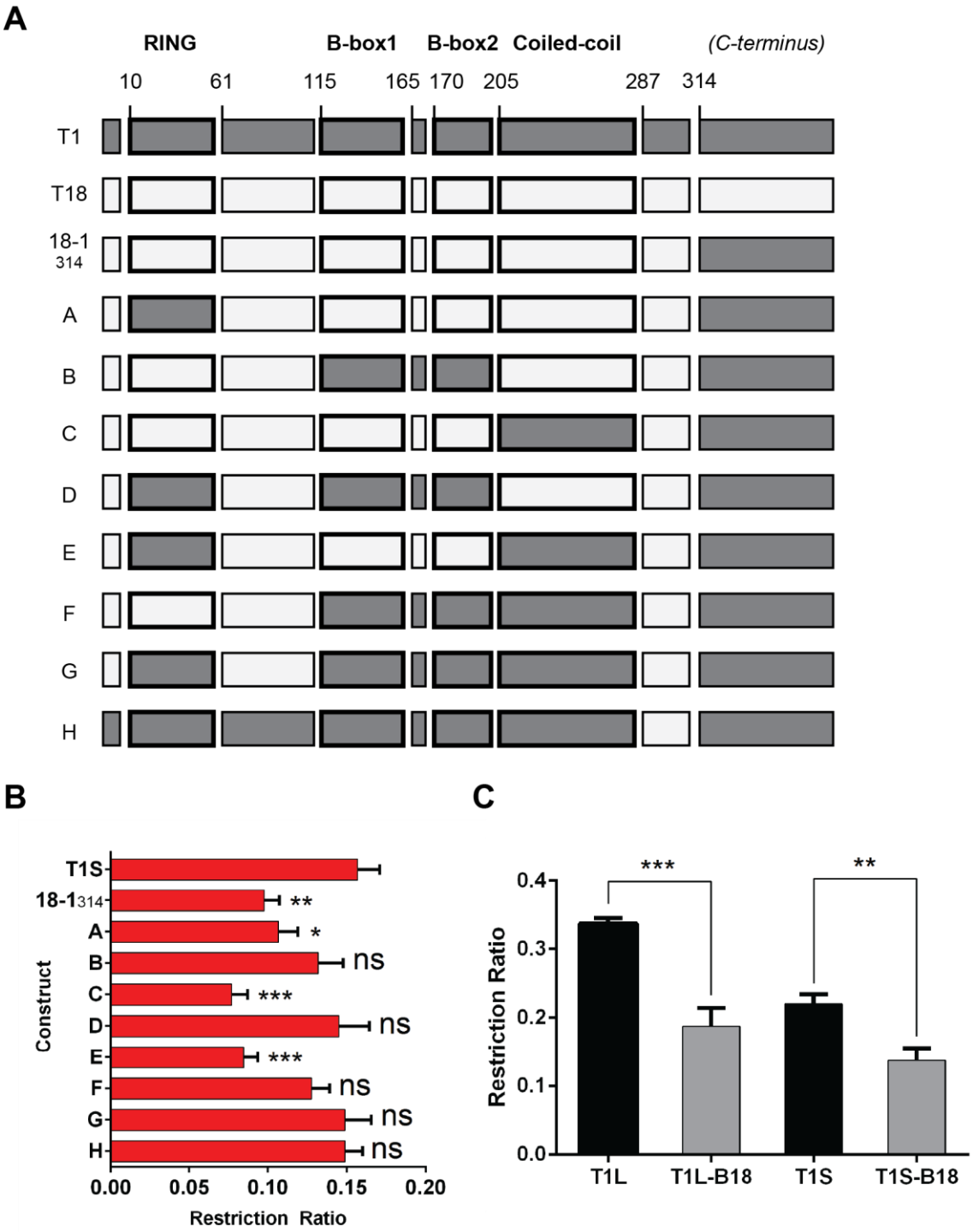


Figure 3.19: The T18 B-boxes augment N-MLV restriction by T1

(A) The structures of Constructs A-H (Section **3.6**). **(B)** Restriction of N-MLV by Constructs A-H. Graph shows the mean and SEM from five (T1S, 18-1_{314}, A-C) or four (D-H) independent experiments (*n* = 10 and 8, respectively). Where significance is shown, this is relative to T1S. **(C)** Restriction of N-MLV by T1L/S with and without the T18 B-boxes (B18). Graph shows the mean and SEM for three independent experiments (*n* = 6). In all cases, statistical significance was established using an unpaired t-test: *** $P \leq 10^{-3}$, ** $P \leq 10^{-2}$, * $P \leq 0.05$ and *ns* = not significant.

3.7 The restriction phenotypes of T1, 18-1_{314} and T5α are probably not artefacts of overexpression

All of the experiments described so far were performed using the LxIY vector, which is optimised for constitutive expression in mammalian cells. One limitation of this approach is the potential for artefactual restriction phenotypes resulting from protein overexpression.

To preclude the possibility that N-MLV restriction is a consequence of restriction factor overexpression, six constructs (huT5α, rhT5α, a chimera of the huT5α RBCC and rhT5α B30.2 domains [hurhT5α], agmT1L, agmT1S and 18-1_{314}) were fused to a C-terminal HA tag and cloned into the CS14 and 15 vectors previously described (Section **2.2.7**). These vectors differ from LxIY in two important respects: (1) the CMV promoter is replaced with P_{TRE3G}, permitting transcription only in the presence of doxycycline, and (2) the restriction factor gene is placed downstream of the *eYFP* reporter, thereby attenuating its recognition by the ribosome. CS15 provides the lower range of expression levels, while CS14 provides a level intermediate between that of CS15 and LxIY.

In a preliminary experiment, N-MLV restriction by each of the six constructs was tested in five conditions of increasing protein expression: (**1**) CS15 without dox; (**2**) CS14 without dox; (**3**) CS15 with dox; (**4**) CS14 with dox and (**5**) LxIY. Western blots were performed in parallel to quantify the relative amount of protein under each condition.

Conditions **1, 2** and **3** invariably yielded no detectable protein (Figure **3.20**) and a corresponding absence of N-MLV restriction (Figure **3.21**). However, the results of condition **4** varied with different constructs. RhT5α expression fell just within the limits of detection and restricted N-MLV only weakly at this concentration. In stark contrast, huT5α and hurhT5α were both more abundantly expressed and exhibited full restriction of N-MLV (albeit not quite as potently as when overexpressed in condition **5**). All three of the T1-derived constructs were detectable under condition **4** and exhibited partial restriction phenotypes, with 18-1_{314} approaching the threshold for full restriction.

To investigate whether these observations are unique to N-MLV restriction, the same experiment was performed using HIV-1. Given their inability to restrict the lentiviruses, all of the T1-derived constructs were excluded from this assay. Although huT5α is also inactive against HIV-1, this construct was included as a negative control here. As expected, the restriction data obtained were similar to those for N-MLV: conditions **1, 2** and **3** yielded no activity against HIV-1, while in condition **4** a modest restriction phenotype manifested, albeit not quite to the levels seen in **5** (Figure **3.21C**).

These initial experiments confirmed that inhibition of both N-MLV and HIV-1 can occur at modest levels of T1 and 5α expression, albeit not as potently as when these factors are overexpressed. However, this work used only a single concentration of doxycycline (1000 ng/mL), and therefore lacked sufficient resolution to determine a precise correlative relationship between protein expression level and restriction activity.

In order to fine-tune these data, CS14 vectors bearing either T1L or 18-1_{314} were transduced into R18 cells and then exposed to eight concentrations of doxycycline ranging from 10 to 2000 ng/mL. The lower limit for this experiment was chosen based on the previous observation that $Fv1^b$ exhibits a restriction phenotype that closely mirrors its endogenous counterpart when induced with 10 ng/mL doxycycline (Li et al., 2016b). Cells grown in dox-free media were included as a negative control, and cells transduced with the LxIY vector as a positive.

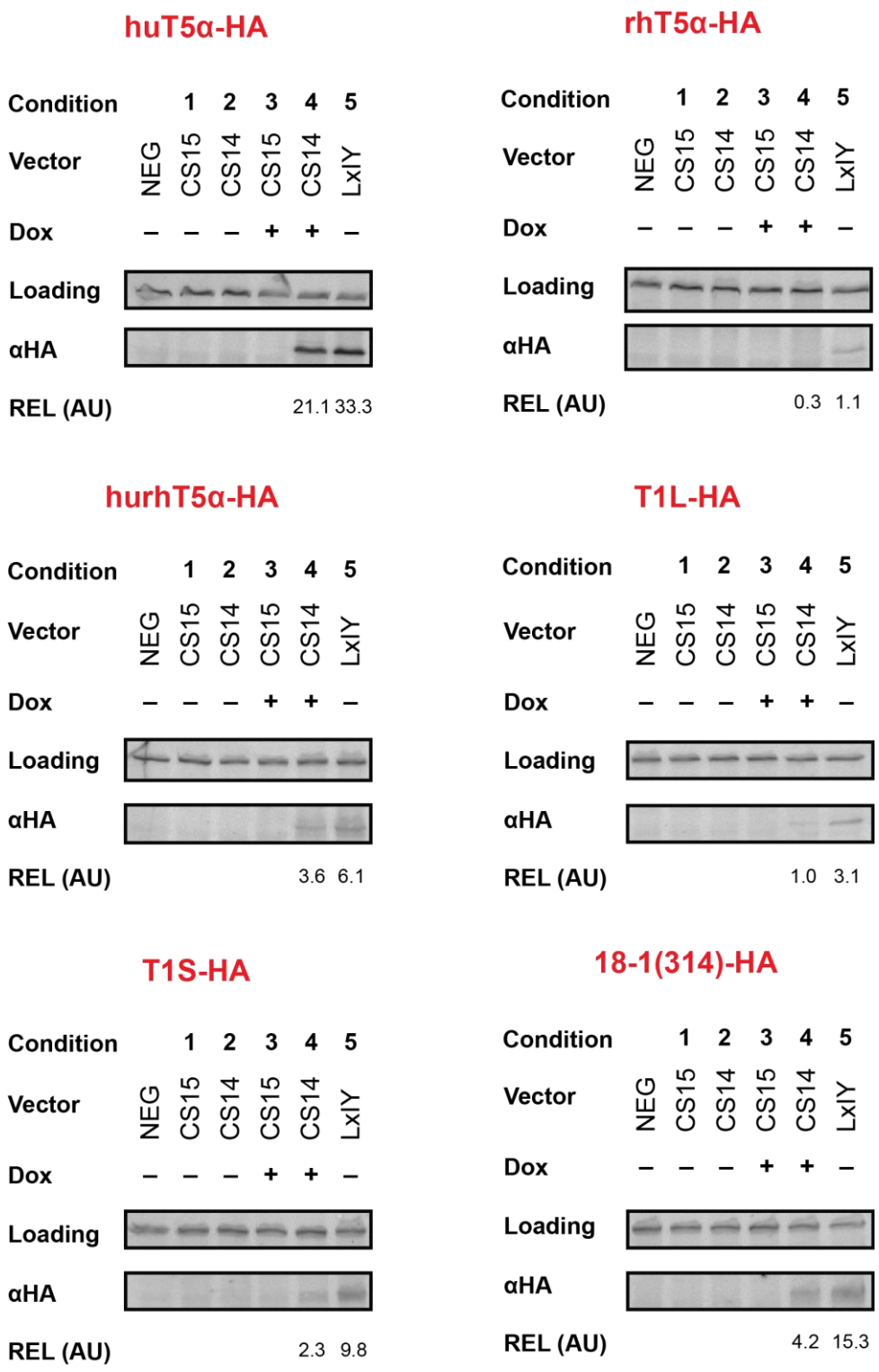


Figure 3.20: Expression of six restriction factor constructs under five conditions

The expression level of each construct was quantified by western blotting using an antibody against the C-terminal HA tag. Untransduced R18 cells were used as the negative control. The loading control is an extraneous band of unidentified origin that is detectable in all samples when αHA is used to probe R18 cell lysate. Where dox is present, it is added at a final concentration of 1000 ng/mL. REL: relative expression level (normalised to the loading control); AU: arbitrary units.

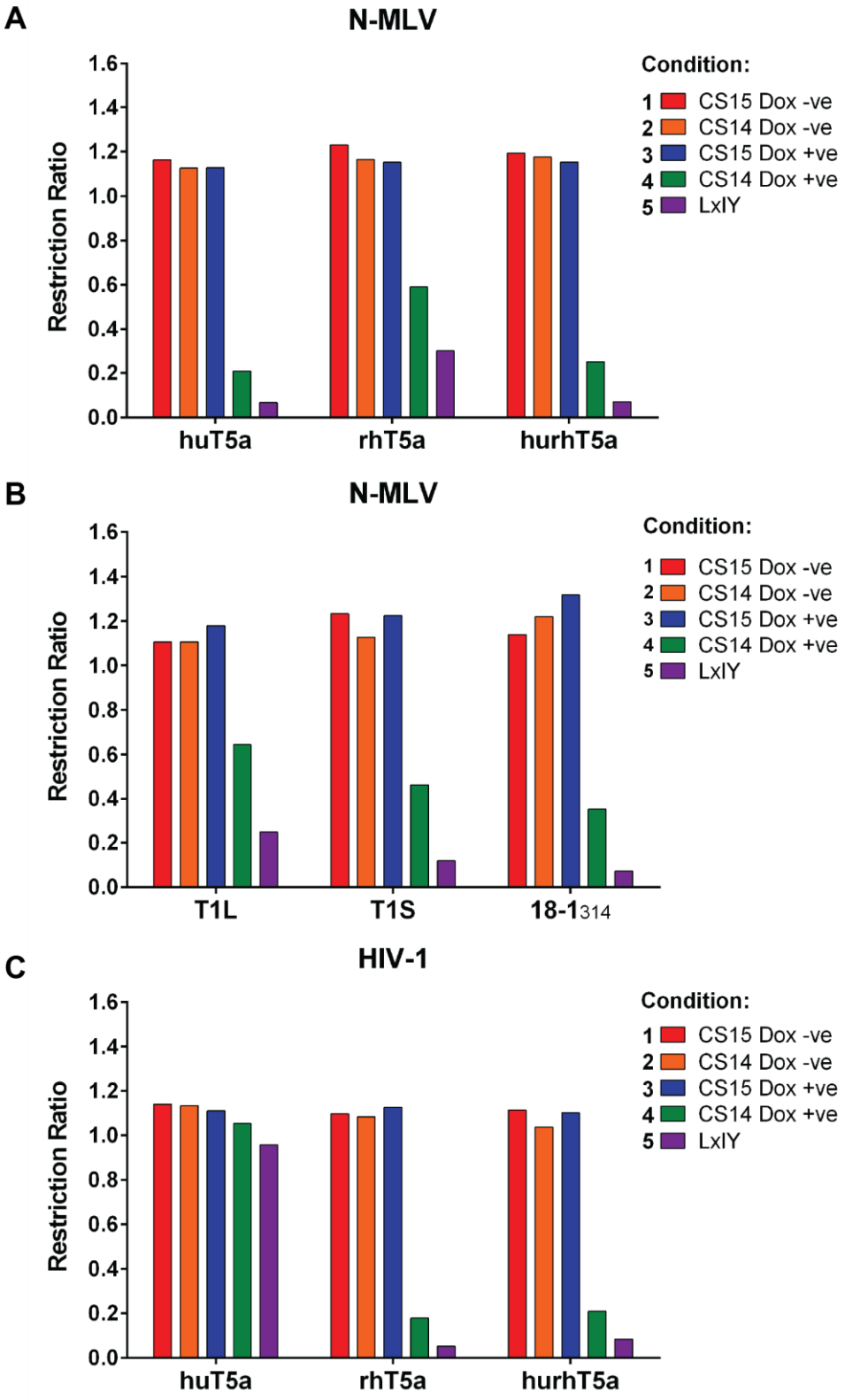


Figure 3.21: Restriction phenotypes of six constructs when expressed in inducible and non-inducible vector systems

Graph shows the mean values from a preliminary experiment (*n* = 2). (*n* = 1 for rhT5α in the CS15 Dox negative condition).

As Figure **3.22** shows, T1L became partially active at 50 ng/mL [Dox] and maintained this phenotype up to the maximum concentration used in this experiment (2000 ng/mL). The corresponding western blot shows that T1L protein became detectable at 50 ng/mL, and that this level of expression was roughly maintained at higher drug concentrations (with the exception of 1000 ng/mL); this aligns well with the restriction data. Protein levels were about sevenfold greater when T1L was overexpressed.

18-1_{314} was active at a slightly lower concentration of drug (20 ng/mL) and reached the threshold for full restriction at 100 ng/mL, before plateauing in a similar manner to T1L. However, the corresponding western blot did not yield values that marry well with these data. For example, the quantities of protein present following induction with 20 and 100 ng/mL [Dox] were roughly equivalent, despite the former yielding a partial restriction phenotype and the latter, full. Conversely, 100 and 2000 ng/mL [Dox] resulted in drastically different protein quantities, despite similar values for restriction. Moreover, there appeared to be greater expression in the presence of 2000 ng/mL [Dox] than from the LxIY vector, which is also discordant with the restriction data. It is possible that these discrepancies are due to technical issues with performing and/or quantifying the western blot; however, given that the blot for T1L was run in parallel, using an identical protocol, this seems unlikely. Unfortunately, time constraints prohibited further investigation into this issue.

Given the data presented in Section **3.6**, it is noteworthy that 18-1_{314} appeared to be more abundantly expressed than T1L in these experiments. To confirm that this is the case, both factors were probed on the same membrane to enable a direct comparison of protein quantity (Figure **3.23A**). Under the four conditions tested (20, 50 and 100 ng/mL doxycycline, plus overexpression), 18-1_{314} was visibly more abundantly expressed than T1L. Given that LxIY was used in the experiments described in Section **3.6**, the relative expression levels of each factor in this vector were quantified and compared (Figure **3.23B**). Indeed, 18-1_{314} was about twice as abundant as T1L when expressed from the LxIY vector. However, given that T1S expression is also about twice that of T1L under this condition (Section **3.3**), and that 18-1_{314} restricts N-MLV to a greater degree than T1S (Section **3.6**), this is unlikely to be the sole reason for the enhanced restriction potency of this chimera.

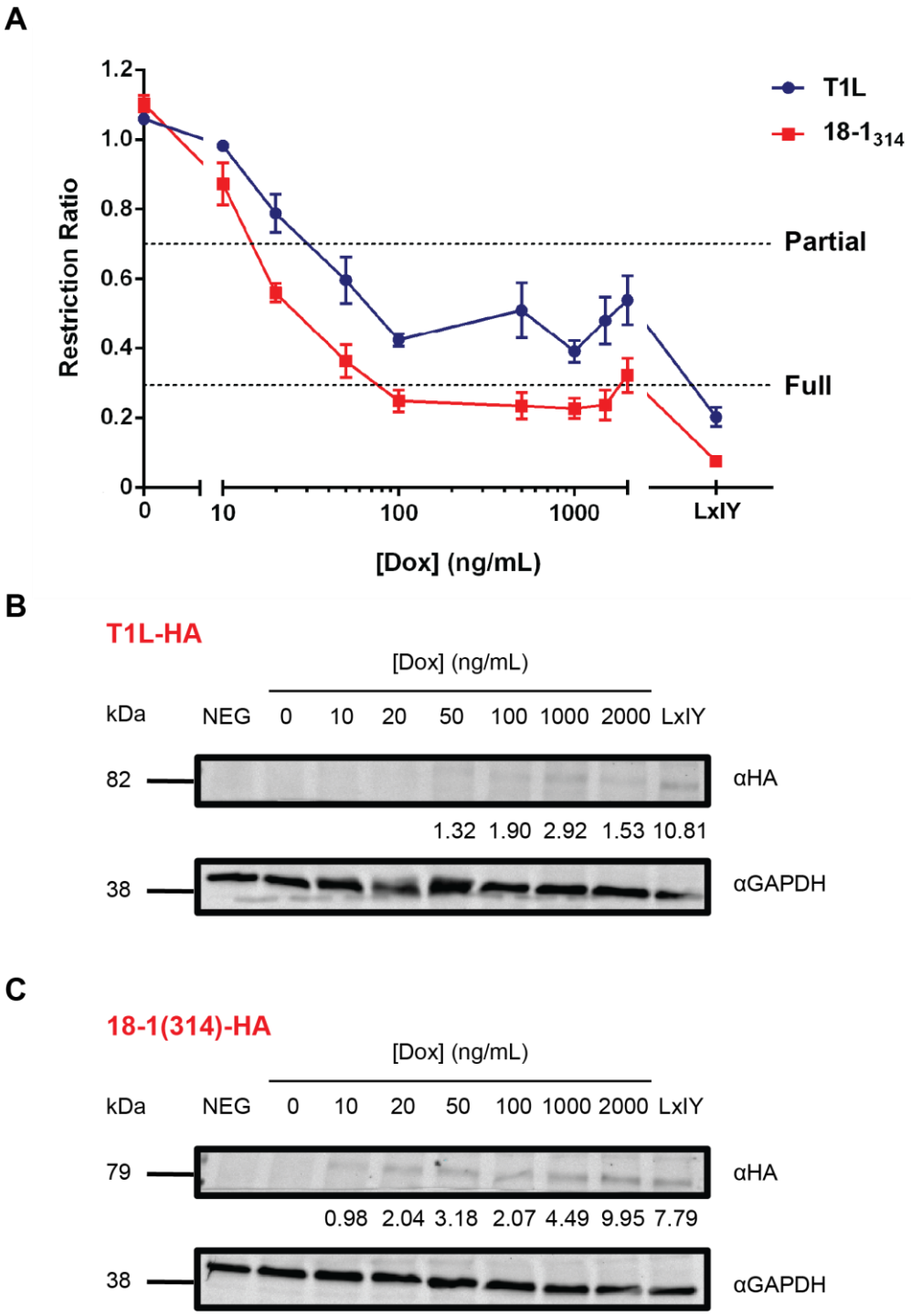


Figure 3.22: Restriction of N-MLV by T1L and 18-1_{314} under titrated doxycycline

(A) Restriction ratios are plotted for T1L and 18-1_{314} at nine concentrations of doxycycline (0, 10, 20, 50, 100, 500, 1000, 1500 and 2000 ng/mL). Restriction values obtained from overexpression are included as a positive control. Dotted lines mark the thresholds for partial (ratio 0.3-0.7) and full restriction (ratio < 0.3). Graph shows the mean and SEM from two independent experiments (n = 4). **(B, C)** Quantitative western blots of T1L (B) and 18-1_{314} (C) under eight conditions of expression. Relative amounts of protein (in arbitrary units) were calculated by normalisation to GAPDH.

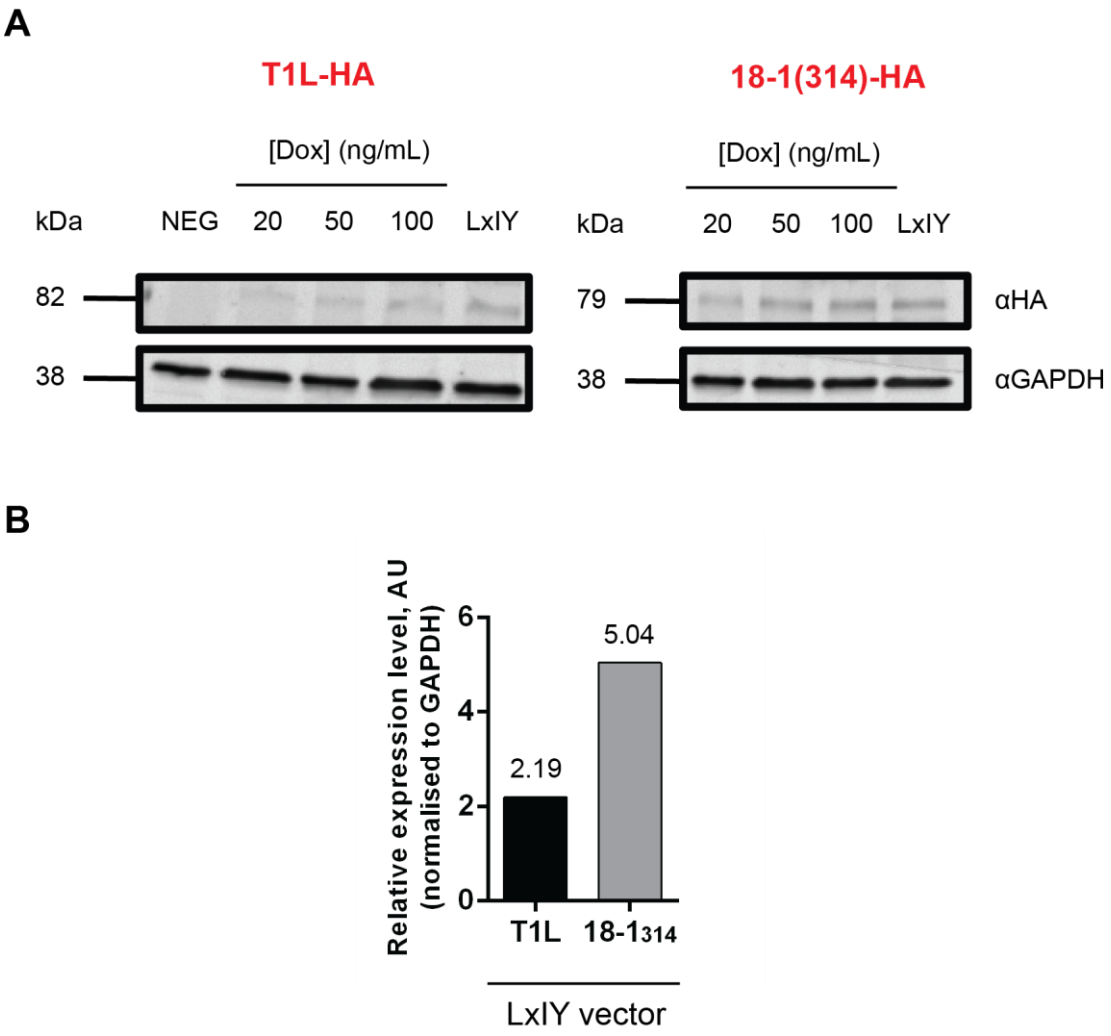


Figure 3.23: Comparative quantitation of T1L and 18-1$_{314}$

Each restriction factor was expressed under four conditions: three levels of Dox induction, plus overexpression. All eight samples were then run on the same gel and probed on the same membrane to enable a direct comparison of protein quantities. **(A)** Blots of T1L-HA and 18-1$_{314}$-HA, using GAPDH as a loading control. **(B)** Comparative quantitation of T1L-HA and 18-1$_{314}$-HA expressed from the LxIY vector. Values were obtained by normalisation to GAPDH.

3.8 Discussion

The ability of T1 to inhibit N-MLV replication was first described more than a decade ago (Yap et al., 2004); however, at the start of this project, there had been no further attempts to characterise this phenotype. The aim of this chapter was thus to offer a more detailed description of T1-mediated restriction. Each of the key findings arising from this work will now be addressed in turn.

T1 is constrained by negative selection

T5α contains numerous signatures of positive selection, and is correspondingly able to restrict divergent retroviruses (Song et al., 2005b; Yap et al., 2008; Fletcher et al., 2010). In order to verify whether T1 shares these properties, three orthologues of this protein were challenged with a panel of lenti- and foamy viruses. Unlike T5α, however, T1 was permissive to all of these viruses. In support of these data, our phylogenetic analysis of the *T1* gene reveals that it has been constrained by negative selective pressures over the course of its evolution.

These observations are perhaps unsurprising given that T1 participates in the development of embryonic midline structures (Cox et al., 2000), a biologically fundamental process that necessitates strict functional conservation. While T5α is free to diversify with successive rounds of viral challenge, the T1 protein sequence is constrained by its primary function. Indeed, it could be argued that the T1-mediated restriction of N-MLV is merely an incidental phenomenon: a by-product of the domain architecture and microtubular localisation of this protein.

The majority of the T1 B30.2 domain can be functionally replaced with the equivalent regions from T18

The species-specific restriction profiles of T5α orthologues have been mapped to surface-exposed loops in the C-terminal B30.2 domain, known as variable regions (VRs) (Ohkura et al., 2006; Yap et al., 2008; Kono et al., 2009). However, while the restriction activity of T1 had previously been mapped to the C-terminal half of the protein (Yap et al., 2005), the relative contribution of elements within this region

remained elusive. We therefore sought to gain further insight by making C-terminal chimeras of T1 and its non-restricting paralogue, T18.

Unsurprisingly, this work revealed that T1 also recognises capsid via the B30.2 domain. However, an extensive panel of chimeras generated within this region revealed that only a small fraction of T1-derived sequence is required for this phenotype, including VR2 and the sequences flanked by VRs 2-3 and 3-4. This stands in stark contrast to T5α, where VRs 1, 2 and 3 and numerous inter-loop regions have been reported to contribute to restriction specificity in different orthologues (Song et al., 2005b; Ohkura et al., 2006).

The requirement for T1 VR2 is particularly striking, given that this region differs by only a single amino acid between T1 and 18: an asparagine in the former (N595) is replaced by a histidine at the corresponding position in the latter (H565). To determine whether this effect is unique to histidine, the remaining 18 amino acids were substituted at position 595 of T1L. Interestingly, the results revealed a remarkably low mutational robustness of this region, with 17 out of 20 amino acids completely ablating restriction activity. In fact, the only residues that maintained restriction competence, besides the wild-type asparagine, were aspartic and glutamic acid, both of which are negatively charged at physiological pH. Surprisingly, even glutamine, which shares biochemical properties with both asparagine and glutamic acid, appeared to have an inhibitory effect on capsid recognition at this position. While there are no clear biochemical explanations for these patterns, it is hoped that the eventual structural characterisation of the T1 B30.2 domain (see Section **4.3**) will shed light on the molecular basis for these observations.

There are number of caveats to the above data. Firstly, in the absence of western blots for each mutant, we cannot be certain that the inactive constructs were adequately expressed. Moreover, even if they were, their inactivity could be attributed to a host of phenotypes that are unrelated to capsid binding, such as a change in localisation, or failure to initiate an effector pathway. While the genetic data in this chapter provide a good starting point for further investigation, they should be interpreted with caution until they are biochemically verified.

T1 protein concentration correlates with restriction potency

T1 is alternatively spliced to produce multiple RNA species, two of which encode proteins with a B30.2 domain. Interestingly, the shorter isoform (T1S) appears to restrict N-MLV with 2-3 fold greater potency than the long (T1L). While this is certainly a modest difference, it was highly reproducible; we therefore sought to identify the biological property that underpins it.

T1L and S differ in the length of their fibronectin type III (FN-III) domains by 30 residues. Since this domain is a protein-protein interaction module, and such interactions often regulate biochemical phenomena such as protein stability (Paci et al., 2012; Dar et al., 2014), the expression level of each isoform was quantified by western blotting. As expected, T1S was expressed at a level approximately double that of T1L, suggesting that a shorter FN-III domain favours increased protein concentration. This would endow the short isoform with a greater capacity to handle the incoming viral load. However, the possibility that other variables contribute to this phenomenon cannot be excluded. For example, determinants within the FN-III domain may influence the propensity for each isoform to associate with microtubules, which would in turn affect their abilities to intercept virions as they translocate towards the nucleus.

However, the FN-III domain is not the only determinant that appears to have dual effects on expression level and restriction potency. When generating the B30.2 domain chimeras, we found that the VR2 loop of T18 was more inhibitory in T1L/S than in an 18-$1_{B30.2}$ chimera, implying that determinant(s) in the T18 N-terminus can augment restriction. This observation aligns well with a previous finding by Yap et al. (2005), where a chimera made between T18 and T1S at position 314 restricted N-MLV more potently than wildtype T1S. By generating a panel of chimeras where N-terminal portions of T1 were substituted into 18-1_{314}, we identified the B-boxes as the determinant responsible for this enhanced potency. Subsequent western blotting revealed that, like T1S, 18-1_{314} is more abundantly expressed than T1L. Indeed, these data are in agreement with the observation that T18 expression typically exceeds that of T1 in the developing mouse embryo (Buchner et al., 1999).

Collectively, these data indicate that both the B-boxes and the FN-III domain can increase restriction potency by boosting protein expression levels. The fact that 18-1_{314} restricts N-MLV more potently than T1S suggests that these effects are additive, although it is difficult to conclusively establish this without a comparative blot of the two proteins. Furthermore, given that the relative difference between T1L and S is only slightly less than that between T1L and 18-1_{314}, it is likely that the T18 B-boxes also potentiate restriction in a concentration-independent manner. For example, it is conceivable that differences in the T1 and 18 B-box sequences (corresponding to 21% of residues) permit the latter to multimerise more efficiently than the former, thereby enhancing capsid binding through avidity effects.

The T1 restriction phenotype is probably not an artefact of overexpression

One limitation of the restriction assay developed by Bock et al. (2000), and used in the majority of experiments described in this chapter, is that it involves overexpression of the restriction factor in the transduced cell. This raises the possibility that the restriction profiles described thus far do not accurately reflect endogenous phenotypes. To assuage this concern, T1L, T1S and 18-1_{314} were each cloned into two separate, doxycycline-inducible vector systems: CS15 for low-level expression, and CS14 for intermediate. Both vectors have previously been used to correlate expression level and restriction specificity for the murine restriction factor, Fv1 (Li et al., 2016b).

Interestingly, while Fv1 exhibits modest restriction of MLV in the CS15 vector (Li et al., 2016b), the three constructs tested in this project were inactive under this condition. Nevertheless, all three acquired partial activity when expressed from the CS14 vector. While this provisional finding was interesting, it was derived using only a single concentration of doxycycline (1000 ng/mL). Previous work has shown that a small fraction of this dose (10 ng/mL) is required to yield an endogenous-like restriction phenotype for Fv1 (Li et al., 2016b). The restriction activities of T1L and 18-1_{314} were therefore tested at eight concentrations of doxycycline, taking 10 ng/mL as the lower limit.

T1L became partially active at 50 ng/mL doxycycline and maintained this level of activity up to 2000 ng/mL. In agreement with these data, T1L protein levels were roughly equivalent at these concentrations. 18-1_{314} showed a similar trend, although it became active at a slightly lower concentration of drug (20 ng/mL) and, as expected, restricted N-MLV more potently than T1L at all higher concentrations. Unfortunately, however, the corresponding western blot did not align well with these data, and time constraints prohibited further investigation of this issue.

It is important to be conscious of various limitations arising from the above experiments. While technical issues may be partly to blame for the poor quality of the quantitative data, it is also possible that western blotting is simply a less sensitive measure of protein concentration than the restriction assay. Furthermore, the apparent 'plateaus' in restriction activity observed upon induction with ≥50 ng/mL [Dox] probably reflect the upper limit for transcriptional induction in this system, since a more potent phenotype is achievable when the restriction factors are overexpressed from the LxIY vector.

Nevertheless, these experiments confirm the restriction-competence of T1 at lower concentrations of protein. However, assessing the biological relevance of these findings will necessitate a precise quantitation of the endogenous levels of T1.

Chapter 4 Searching for parallels between T1 and T5α

Given its limited target range and lack of positive selection, it seems unlikely that T1 plays a key role in the cellular defence against retroviruses. Instead, the utility of this protein lies in what it can teach us about the more prominent restriction factors – particularly T5α – by comparing what unites these proteins and what sets them apart.

The T1 and 5α protein sequences are highly divergent: the former is almost double the length of the latter due to two additional domains between the RBCC and B30.2 compartments (Figure **4.1A**). Furthermore, in Chapter **3**, it became apparent that while the variable regions are critical for capsid recognition in T5α (Ohkura et al., 2006; Anderson and Akkina, 2008), they are functionally replaceable in T1. Despite these differences, however, both factors are able to restrict N-MLV to comparable degrees. This suggests that either there are common features that unite the two proteins, or that evolution has devised two distinct mechanisms for recognising and/or restricting N-MLV. The focus of this chapter is thus to compare and contrast T1 and 5α in order to verify whether or not such parallels exist.

4.1 T1 and 5α can be fused to produce a molecule with restriction activity

Given that T1 and 5α each contain effector and capsid-binding domains, we reasoned that it might be possible to fuse them to generate a restriction-competent chimera. To explore this possibility, reciprocal chimeras were made between agmT1S and rhT5α at two positions – one immediately following the coiled-coil and the other at the start of the B30.2 domain (Figure **4.1B**). Their restriction activities were then measured using the two-colour assay (Figure **4.2**).

As expected, neither of the chimeras with a T1 B30.2 domain were able to restrict the lentiviruses. Surprisingly, however, their reciprocal counterparts were equally permissive to lentiviral replication. Of the four chimeras, one (5-$1_{B30.2}$) restricted N-MLV to a degree comparable to that of T1S. Collectively, these data indicate that it is possible to produce a restriction-competent chimera of T1 and 5α, although the position at which the swap is made has an impact on this phenotype.

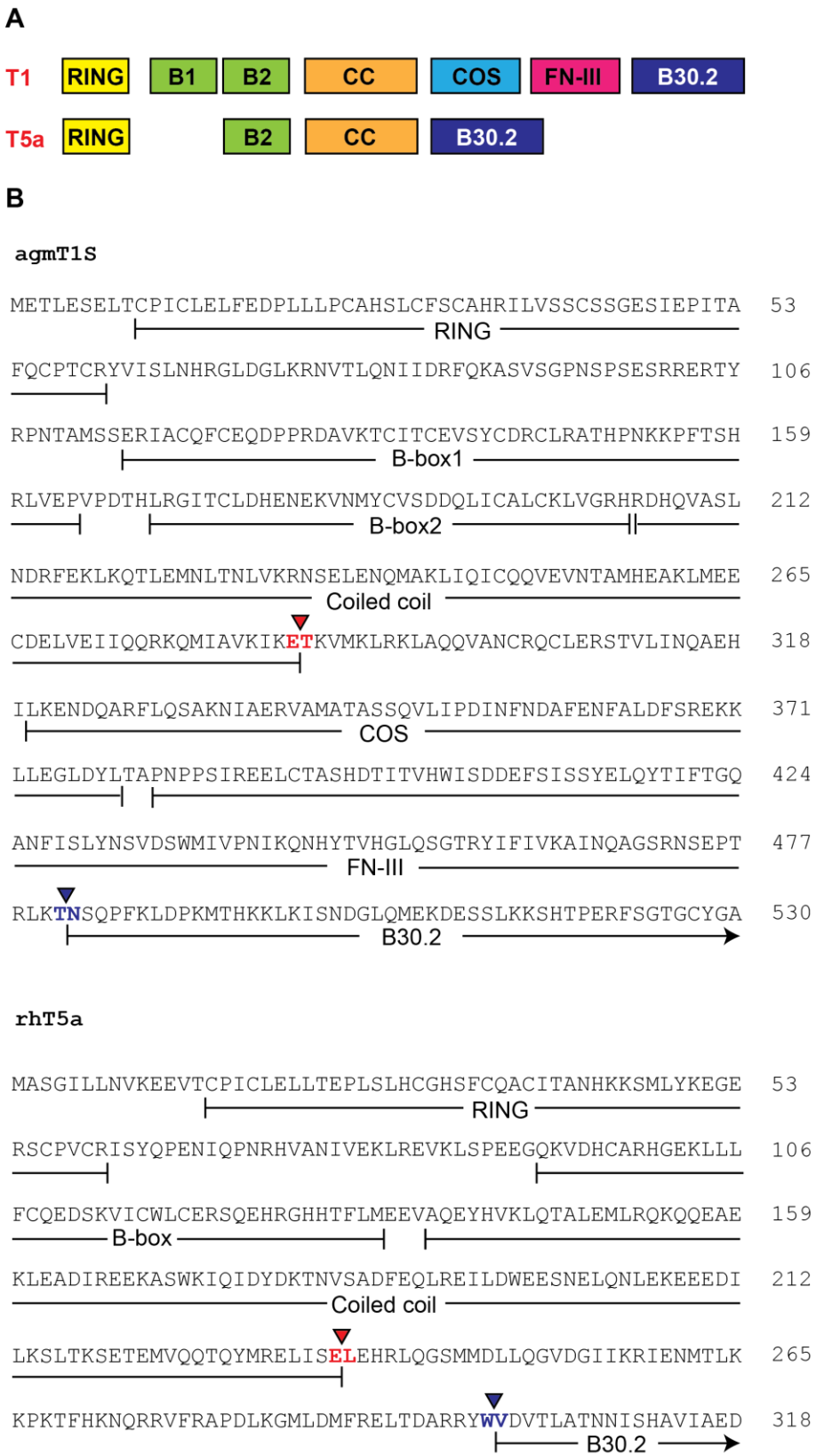


Figure 4.1: T1 compared with T5α

(A) Schematic comparison of T1 and T5α. **(B)** Positions at which the T1-5α reciprocal chimeras were made. Domain boundaries were established using the UniProt database (for T1) and Sastri et al. (2014) (for T5α).

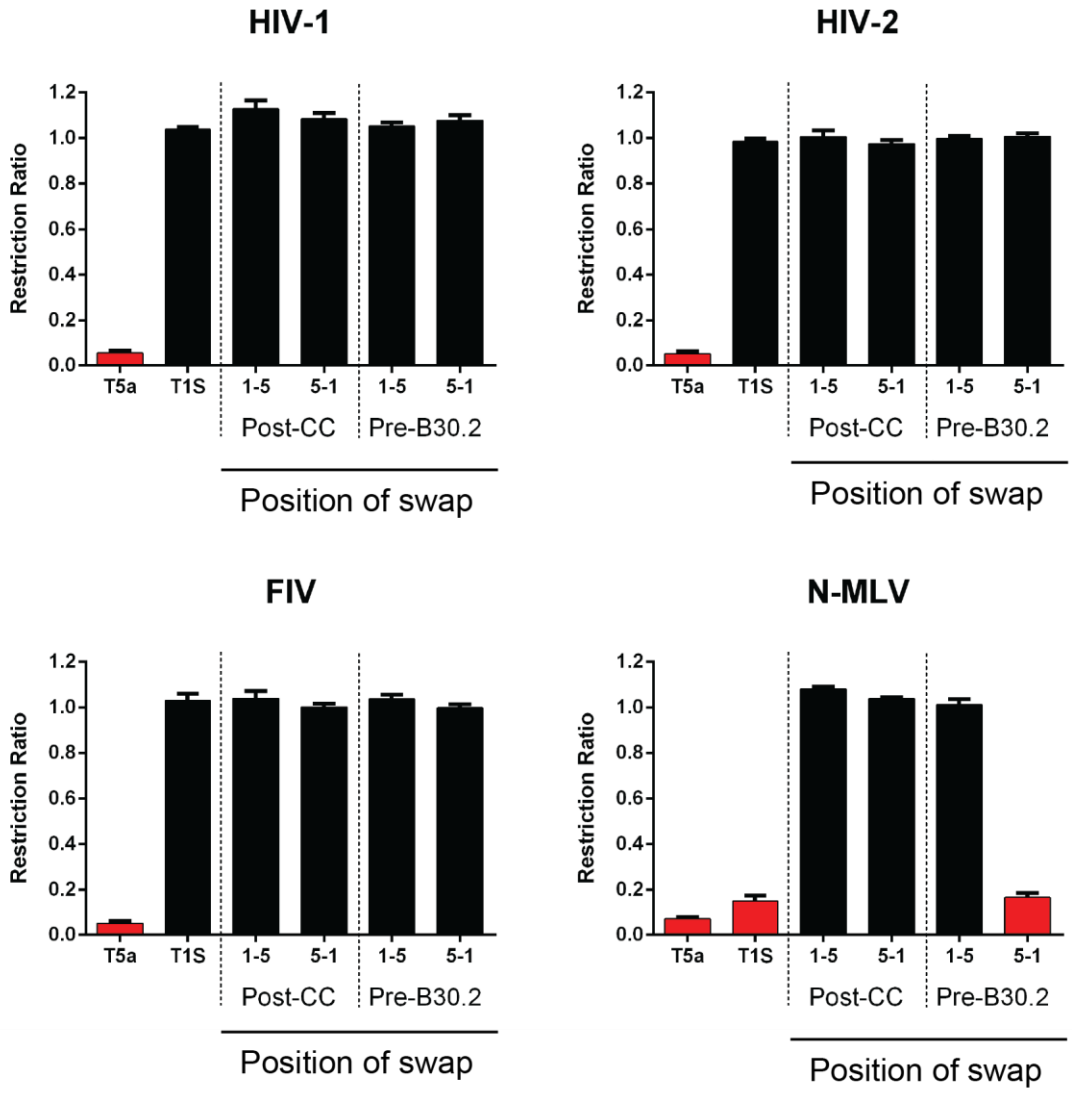


Figure 4.2: Restriction profiles for the T1-5α reciprocal chimeras

Graphs show the restriction activities of T5α (from the rhesus macaque), T1S (from the African green monkey), and reciprocal chimeras thereof. Red bars denote full restriction (ratio < 0.3) and black bars, absence of restriction (ratio > 0.7). Graphs show the mean and SEM from three independent experiments (n = 6).

4.2 A panel of N-MLV capsid mutants escape restriction by both T5α and T1

Having established that T1 and 5α are sufficiently similar in their overall design to produce a functional chimera (Section **4.1**), we next wanted to establish whether these proteins bind to similar epitopes on the N-MLV capsid.

Prior to the start of this project, the Stoye group identified a number of mutations in N-MLV CA that confer escape from restriction by rhT5α (Ohkura et al., 2011). This work revealed an extensive T5-binding interface, including residues in the formerly established Fv1-binding pocket (N82D, E92K, A95D), and in a cluster on the N-terminal side of this pocket (L4S, G8D, L10W). A subsequent study revealed that some of these mutations facilitated the acquisition of compensatory mutations that confer escape from huT5α (L10W/E100K and N82D/N113K) (Ohkura et al, 2013). Collectively, these findings indicated that the T5-binding interface spans a much broader surface on the N-MLV capsid than had previously been thought (Stevens et al., 2004).

To determine whether T1 recognises similar epitopes, we selected five mutants from the above experiments and measured their abilities to escape restriction by T1L, T1S and the 18-1_{314} chimera. In all cases, viral titres were adjusted to account for relative differences in infectivity. As Figure **4.3** shows, there was considerable variability in the degree of escape that each mutant accomplished. While N7K was able to escape restriction by T1L entirely, it retained some sensitivity to T1S and 18-1_{314}, albeit to a lesser extent than wild-type in both cases. A similar phenotype was observed for L10W. E92K was the most sensitive virus in the panel, retaining a degree of sensitivity to all three T1 constructs. Conversely, the double mutants were fully resistant to restriction. In many cases, these phenotypes were distinct from those previously described for T5α (discussed in Section **4.4**).

Collectively, these data indicate that T1 and 5α engage with distinct epitopes on the N-MLV capsid surface, although it is possible that certain residues and motifs contribute to both binding interfaces.

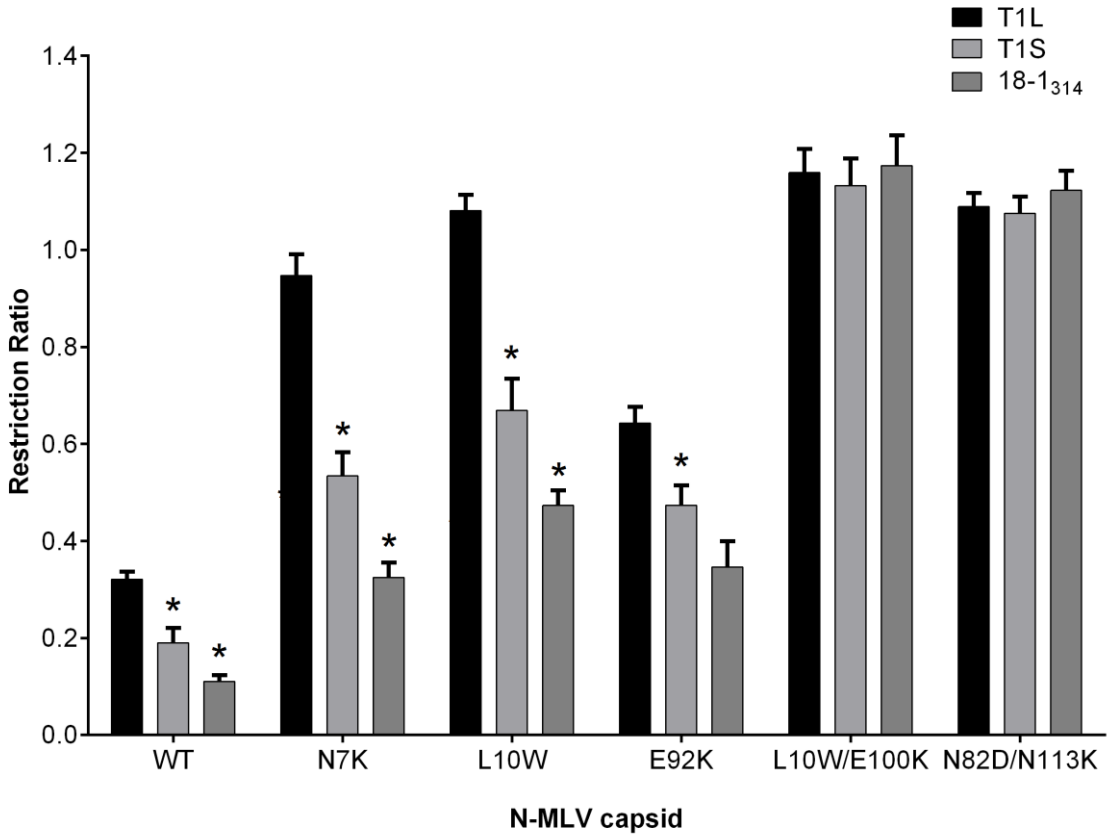


Figure 4.3: Sensitivity of N-MLV capsid mutants to restriction

Sensitivity of wild-type (WT) N-MLV and five capsid mutants to restriction by T1L, T1S and 18-1_{314}. Graph shows the mean and SEM from four (T1L) or three (T1S, 18-1_{314}) independent experiments (n = 8 and 6, respectively). Statistical significance was established using an unpaired t-test: * $P \leq 0.05$. Where significance is shown for T1S, this is relative to T1L, and for 18-1_{314}, relative to T1S.

4.3 Expression and purification of a recombinant T1 B30.2 domain

The fact that T1 and 5α bind to N-MLV through distinct sets of interactions (Section **4.2**) implies that the B30.2 domains of these proteins diverge at the structural level. Unfortunately, while crystal structures exist for several B30.2 domains, including that of T5α (Yang et al., 2012), T25 (D'Cruz et al., 2013), and T72 (Park et al., 2010), no such structure is currently available for T1. We therefore set out to express and purify a recombinant T1 B30.2 domain for crystallisation trials.

4.3.1 Expression of MBP-B30.2 in *E. coli*

To express sufficient amounts of the T1 B30.2 domain, the corresponding region was amplified from the *agmT1* gene using primers with inbuilt restriction sites (Table **4.1**). The PCR product was then digested and ligated into an expression vector (pET47) that includes N-terminal His_6 and MBP tags, to aid in protein purification and solubility, respectively.

Forward	5′- CGTAGA**AAGCTT**ACCCGACTAAAAACAAACAGCC
Reverse	5′- GTTCGA**CTCGAG**TTAATGACAGGTTTTCATCCC

Table 4.1: Primers used to amplify the B30.2-encoding region of *agmT1*

Emboldened sequences correspond to the restriction sites used to generate sticky ends for ligation: *HindIII* in the forward primer, and *XhoI* in the reverse. Bases upstream of these sites were selected at random.

The resulting construct (*MBP-B30.2* herein) was then sequence-verified and transformed into Rosetta 2 (DE3) cells (Novagen). This strain of *E. coli* is engineered to enhance the translation of eukaryotic proteins, thus circumventing the need for codon optimisation. A full description of the protocol used for protein expression and harvesting can be found in Section **2.4.1** of this thesis.

4.3.2 Purification of MBP-B30.2

Bacterial lysate was passed through a nickel column according to the protocol outlined in Section **2.4.2**. The trace obtained from the column is shown in Figure **4.4**: a clear peak of elution is visible soon after the imidazole concentration is increased. The load, flowthrough and eluate were then run in parallel on an SDS gel (Figure **4.7A**). A band of the expected size (67.1 kDa) was observed, consistent with the successful expression and purification of MBP-B30.2. However, the presence of multiple contaminating species warranted further purification by ion exchange chromatography (Figure **4.5**). Given the acidic isoelectric point of MBP-B30.2 (pI = 6.18), a Source Q anion exchange column was used for this purpose. A full description of this protocol can be found in Section **2.4.3**.

Given that a peak of elution from the anion exchange column was visible between fractions A8 and A12, an aliquot from each of these fractions was loaded on an SDS gel in order to verify protein yield and purity (Figure **4.7A**). The relevant fractions were then pooled and further purified by size exclusion chromatography. This final step was taken to minimise the presence of extraneous contaminants that might interfere with the crystallisation process. Figure **4.6** shows the trace obtained from the size exclusion column, with the peak of elution falling between fractions 23 and 29. These fractions were once again pooled, concentrated to 7.4 mg mL^{-1}, and then run on a gel to check protein yield and purity (Figure **4.7B**). Although lower molecular weight bands were visible upon staining, MBP-B30.2 was visibly the predominant species, with an estimated purity of >95%.

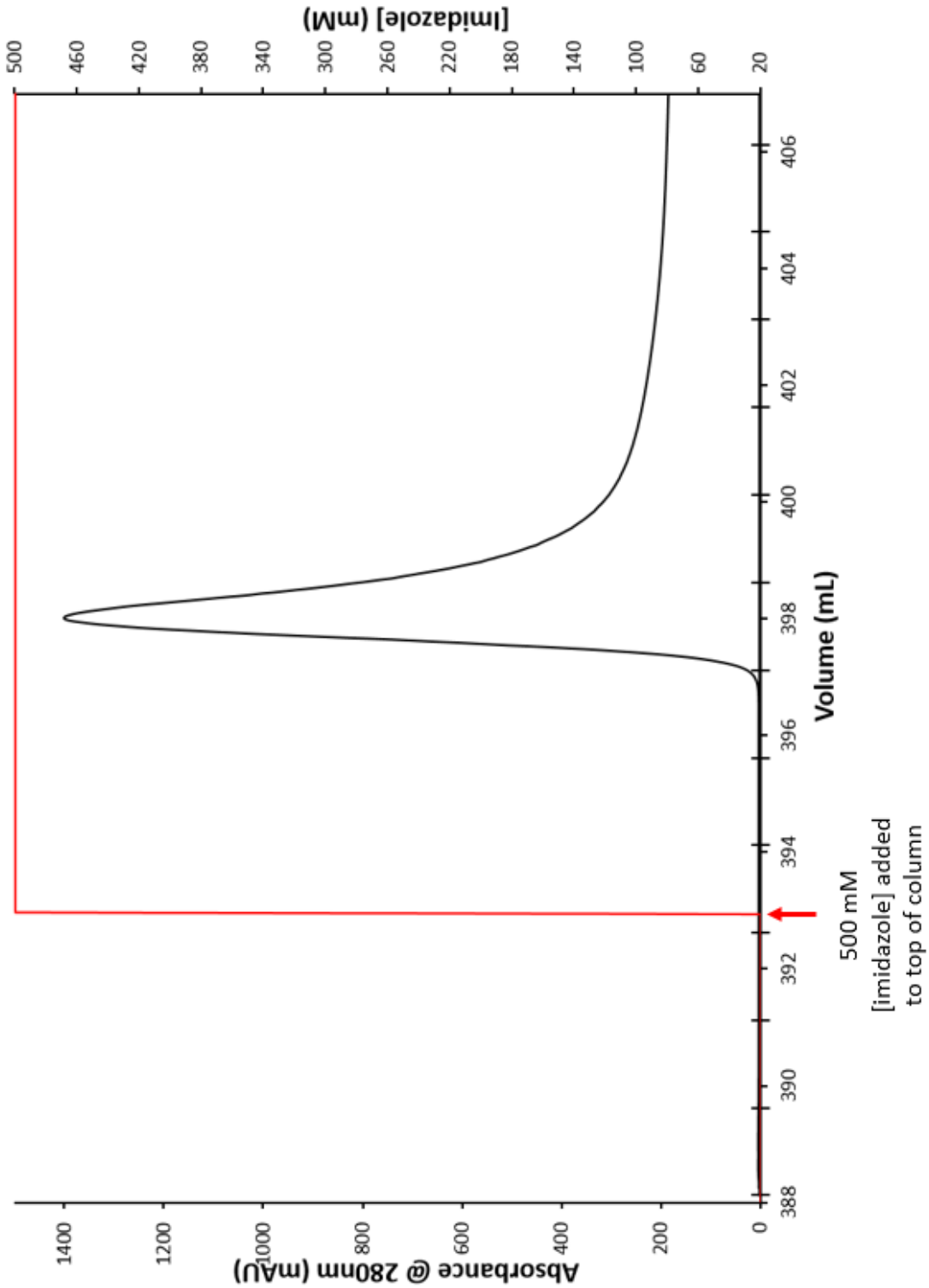


Figure 4.4: Trace from the affinity purification of MBP-B30.2

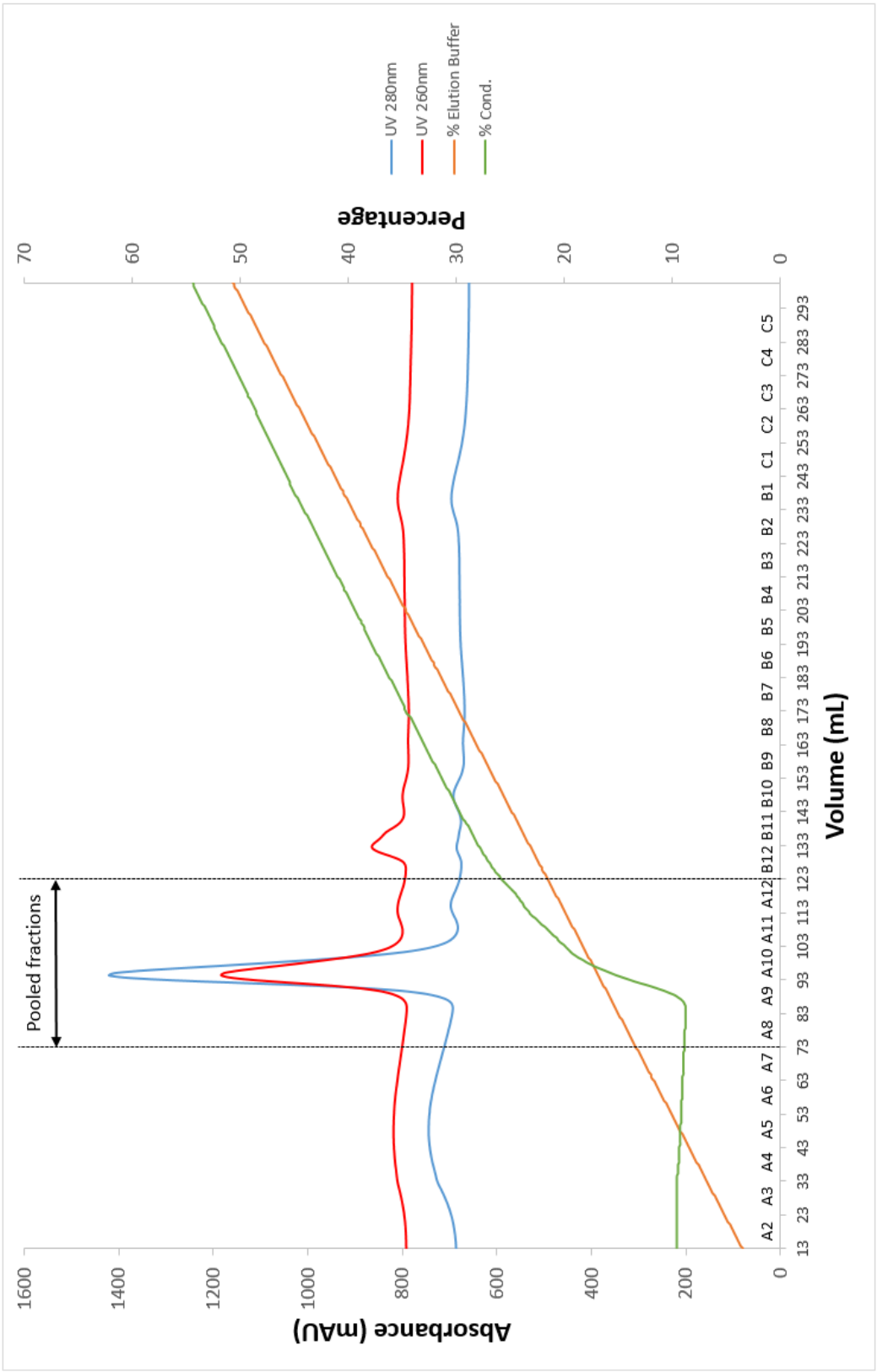


Figure 4.5: Trace from the ion exchange chromatography of MBP-B30.2

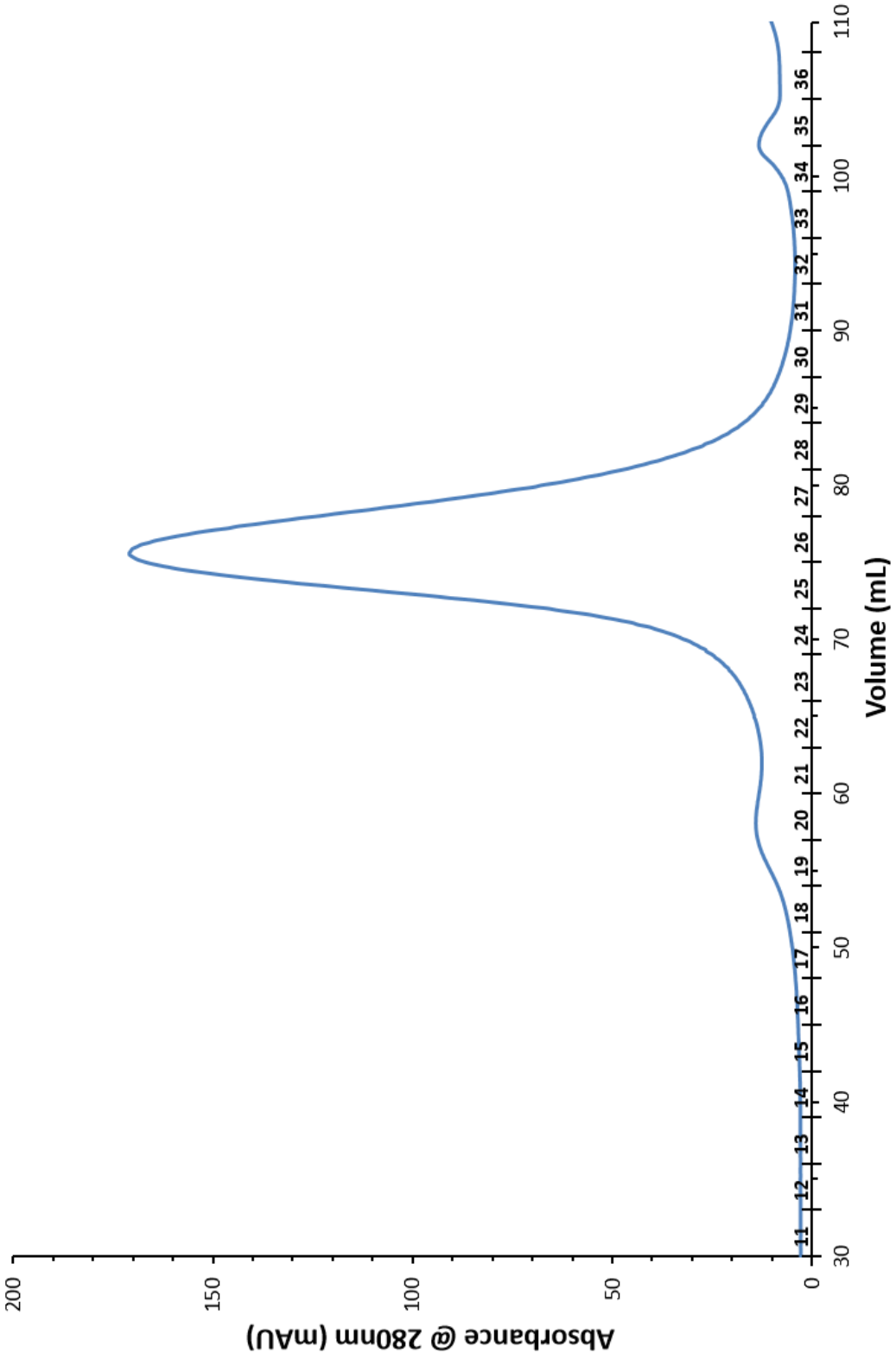
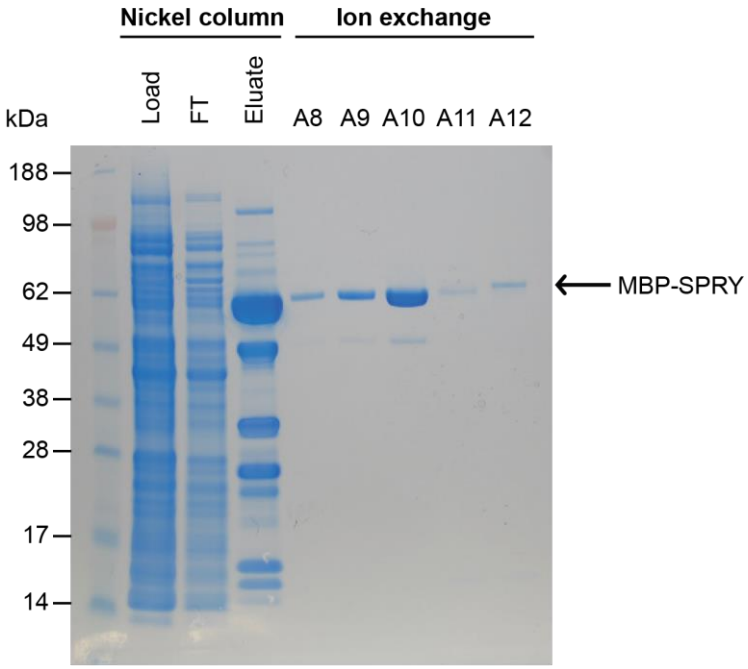


Figure 4.6: Trace from the size exclusion chromatography of MBP-B30.2

A



B

See Blue ladder, MES buffer

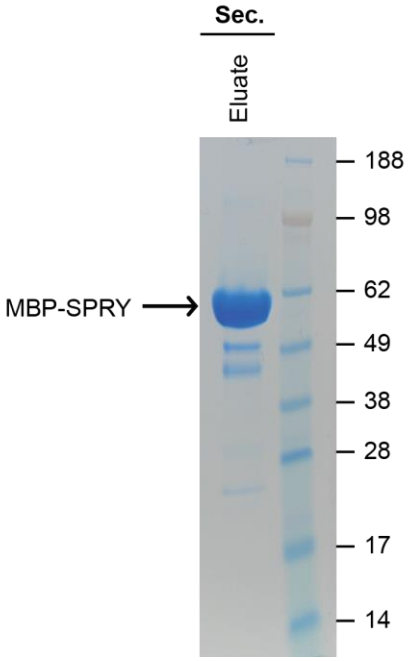


Figure 4.7: MBP-B30.2 after various stages of purification

MBP-B30.2 following **(A)** affinity purification (left) and ion exchange chromatography (right), and **(B)** size exclusion chromatography (sec)

4.3.3 Verification of MBP-B30.2 identity by mass spectrometry

To confirm the identity of the purified protein, an aliquot of eluate from the size exclusion column was sent for analysis by mass spectrometry. As Figure **4.8** shows, a large peak was observed at 67,140.9 Da. This value appears consistent with the theoretical mass of MBP-B30.2 following removal of the N-terminal methionine (67,131.5 Da), a common post-translational modification in bacteria (Giglione et al., 2004).

Possible explanations for the slight discrepancy between the observed and theoretical masses include oxidation of methionine residues within the protein, as well as the intrinsic error associated with mass spectrometry of high molecular weight proteins. Nevertheless, the presence of a large peak within 9 Da of the theoretical mass was taken as evidence that the correct protein had been purified.

A number of smaller peaks at both lower and higher molecular weights were also observed on the mass spectrum. Those at a lower MW probably correspond to species that have lost their N-terminal His_6 tag, but were able to co-purify with their full-length counterparts through homodimeric associations between MBP. Meanwhile, those at a higher MW likely represent proteins with gluconoylated or phosphogluconoylated His_6 tags. Such modifications are well-documented in *E. coli* expression systems (Geoghegan et al., 1999; Matthies et al., 2005).

4.3.4 Crystallisation trials of MBP-B30.2

To grow crystals of MBP-B30.2, neat (7.4 mg mL^{-1}) and diluted (1.85 mg mL^{-1}) samples were incubated in sparse matrix and systematic screens using the sitting drop method (Roksana Ogrodowicz, the Francis Crick Institute). The latter concentration was chosen given its reported effectiveness in growing crystals of the T5α B30.2 domain (Yang et al., 2012). The ratio of clear to precipitated drops obtained under these conditions indicated that both concentrations were favourable; however, no crystals grew within the timeframe of this project. A list of screens used in these trials can be found in Appendix **7.2**.

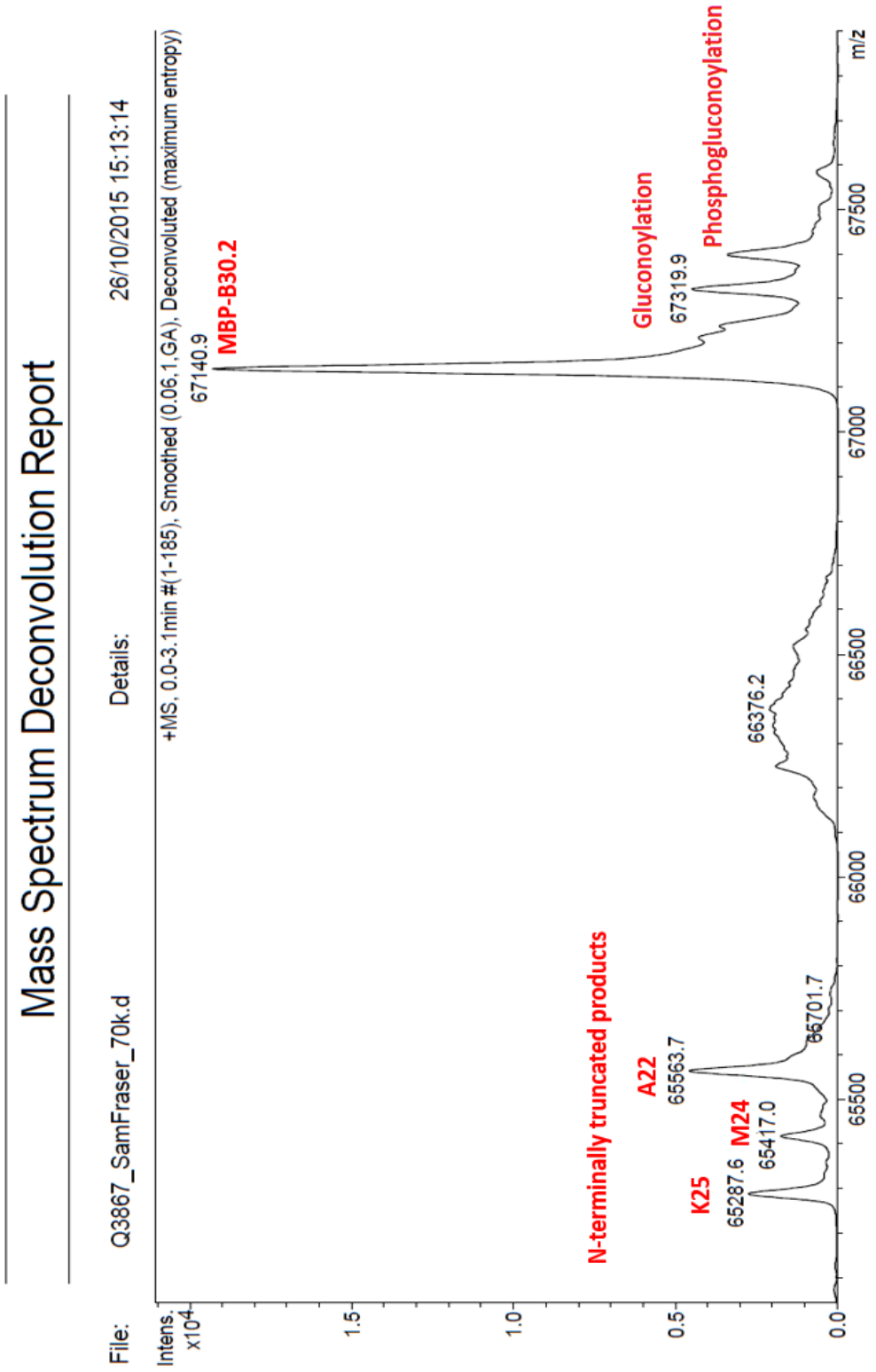


Figure 4.8: The mass spectrum deconvolution report for MBP-B30.2

4.4 Discussion

Although T1 and 5α are divergent proteins, they are united in their ability to restrict N-MLV. The aim of this chapter was to identify parallels between these factors that endow them with this common phenotype. Each of the key findings arising from this work will now be discussed in turn.

T1 and 5α can be fused to produce a restriction-competent chimera

Given that T1 and 5α have common capsid-binding and effector domains, we reasoned that it might be possible to fuse them to generate a restriction-competent chimera. To this end, reciprocal swaps were made between T1 and 5α, both immediately after the coiled-coil (1-5_{CC} and 5-1_{CC}) and at the start of the B30.2 domain (1-$5_{B30.2}$ and 5-$1_{B30.2}$). Although these chimeras were inactive against the lentiviruses, one of them (5-$1_{B30.2}$) did restrict N-MLV to a degree approaching that of wildtype T1S. This suggests that it is possible to generate a T5α-1 fusion with restriction activity, although the position at which the two molecules are joined has an impact on this phenotype.

One explanation for these data is the contribution of the region immediately preceding the B30.2 domain to capsid recognition. In T5α, this is known as Linker 2 (L2), a predominantly α-helical region that serves to align the neighbouring B30.2 domain with viral epitopes. The analogous region in T1 is much longer, owing to the presence of two intervening domains (COS and FN-III) between the coiled-coil and B30.2 domain. It is conceivable that capsid recognition by T5α B30.2 is contingent on the upstream L2 region, while T1 B30.2 can function adequately with either partner. In other words, while T5α is sensitive to the spacing between capsid-binding and effector domains, T1 is relatively flexible in this regard (this notion is covered in depth in Chapter **5**). This hypothesis would explain the ability of 5-$1_{B30.2}$ to restrict N-MLV, while the reciprocal construct is inactive.

However, this hypothesis does not provide a satisfactory explanation for the inactivity of 1-5_{CC} and 5-1_{CC}, both of which contain B30.2 domains that are paired with their native N-terminal neighbours. It has previously been shown through

molecular modelling that contacts between L2 and the preceding coiled-coil are required to maintain the integrity of the T5α dimer, thereby ensuring a productive interaction with capsid (Sastri et al., 2014). If this is the case, then one explanation for the loss-of-function for 1-5_{CC} and 5-1_{CC} might be the lack of complementarity between the L2 region (in the former), or the analogous region from T1 (in the latter), and the coiled-coil to which it is attached. This would prohibit the relevant region from docking onto the coiled-coil in the dimer, in turn preventing the B30.2 domains from efficiently engaging with their target. However, in the absence of structural data for these chimeras, it is difficult to conclusively establish the nature of the secondary and tertiary structures that they (fail to) form.

Nevertheless, the fact that one of these molecules was restriction-competent indicates that T1 and 5α share a common overall design. Both proteins are endowed with effector and capsid-binding domains, and these components appear to be at least somewhat cross-compatible, contingent on the position at which the two molecules are fused.

T1 and 5α recognise distinct epitopes on the N-MLV capsid surface

Having established that T1 and 5α can complement one another in a contiguous molecule, we next wanted to explore whether the two proteins recognise similar epitopes on the N-MLV capsid.

Work from the Stoye group has already begun to define regions on the N-MLV capsid that are recognised by T5α. This was accomplished by serial passage of virions in T5-expressing cells, and the subsequent isolation and characterisation of mutants that escape restriction (Ohkura et al., 2011; Ohkura and Stoye, 2013). These experiments revealed an extensive T5α-binding interface, spanning most of the capsid's outer surface and incorporating residues positioned up to 29 Å apart. In order to establish whether the T1 B30.2 domain binds a similar set of epitopes, five escape mutants were selected from the above experiments (N7K, L10W, E92K, L10W/E100K and N82D/N113K), and tested for their ability to escape restriction from T1L, T1S, and the 18-1_{314} chimera described previously. As shown in Table **4.2**, these data varied considerably between mutants.

	N-MLV capsid mutant				
	N7K	**L10W**	**E92K**	**L10W/E100K**	**N82D/N113K**
huT5α	Escape	Restriction	Escape	Escape	Escape
rhT5α	Escape	Escape	Escape	Escape	Escape
agmT1L	Escape	Escape	Partial escape	Escape	Escape
agmT1S	Partial escape	Partial escape	Partial escape	Escape	Escape
18-1$_{314}$	Partial escape	Partial escape	Partial escape	Escape	Escape

Table 4.2: Restriction phenotypes of five N-MLV capsid mutants
See Figure **4.3** for quantitative restriction data.

N7K falls outside of the formerly established Fv1-binding pocket (Mortuza et al., 2008); specifically, it resides in the apical loop of the conserved N-terminal β-hairpin (Mortuza et al., 2008). Given that this mutation has a destabilising effect on the viral core (and a corresponding impact on viral fitness), it has been postulated that escape from restriction is conferred by rapid transition of the virion through a TRIM-sensitive state (Ohkura and Stoye, 2013). While this model satisfactorily explains the loss of restriction by T5α and T1L, a question mark remains over the abilities of T1S and 18-1$_{314}$ to partially overcome this effect. In the latter case, it is conceivable that the T18 B-boxes endow a greater avidity effect than their counterparts in T1. This might enable restriction to be effected before N7K has sufficient time to disassemble. The case for T1S is less clear, however. While it's feasible that the increased expression of T1S over L (see Section **3.3**)

provides greater opportunity for restriction to occur within the timeframe of core integrity, this hypothesis hinges on T1S expression superseding that of T5α, which is also unable to restrict N7K prior to disassembly. Parallel quantitation of the two proteins will be necessary to confirm whether or not this is the case.

L10W maps to a shallow channel between the β-hairpin and helix α6 (the *β-α6 channel* herein) of MLV CA (Figure **4.9**). In the wildtype structure, the aliphatic sidechain of leucine forms the base of this channel; substitution for the bulkier tryptophan elevates the base by ~2.3 Å. Interestingly, this mutation inhibits restriction by rhesus, but not human, T5α (Ohkura et al., 2011), indicating that different T5 orthologues interact with distinct structures on the MLV capsid surface. Although T1L was unable to inhibit L10W, the fact that both T1S and 18-1_{314} remained restriction-competent confirms that this is not due to elements within the T1 B30.2 domain. Collectively, these data give rise to the rather bizarre notion that changes in the β-α6 channel can discriminate between the closely related rhesus and human T5αs, but not between the highly divergent B30.2 domains of huT5α and agmT1. This discrepancy is even more startling in light of the fact that L10W has been shown to evade restriction by a broad panel of primate T5αs, including variants from the New World monkeys (Ohkura et al., 2011). It is not immediately clear what sets the L10W-restricting factors apart from their inactive counterparts, especially given that members of the former group often have greater sequence identity with members of the latter than they do with each other.

As was the case with the N7K mutant, E92K has a substantial impact on viral fitness, although it does not have appear to have a corresponding effect on core stability (Ohkura and Stoye, 2013). This loss of fitness correlates with escape from restriction by both rhesus and human T5α. Interestingly, this effect appears to be charge-dependent: a basic residue at position 92 causes escape from T5α with a concomitant loss of fitness, while acidic residues confer the inverse phenotype (Ohkura and Stoye, 2013). Strangely, however, T1 appears to be relatively unhindered by this modification: in fact, E92K was the only mutant in the panel to retain sensitivity to all three T1 constructs. This observation implies that the T1 B30.2 domain is relatively insensitive to changes in charge, at least at this position in N-MLV CA.

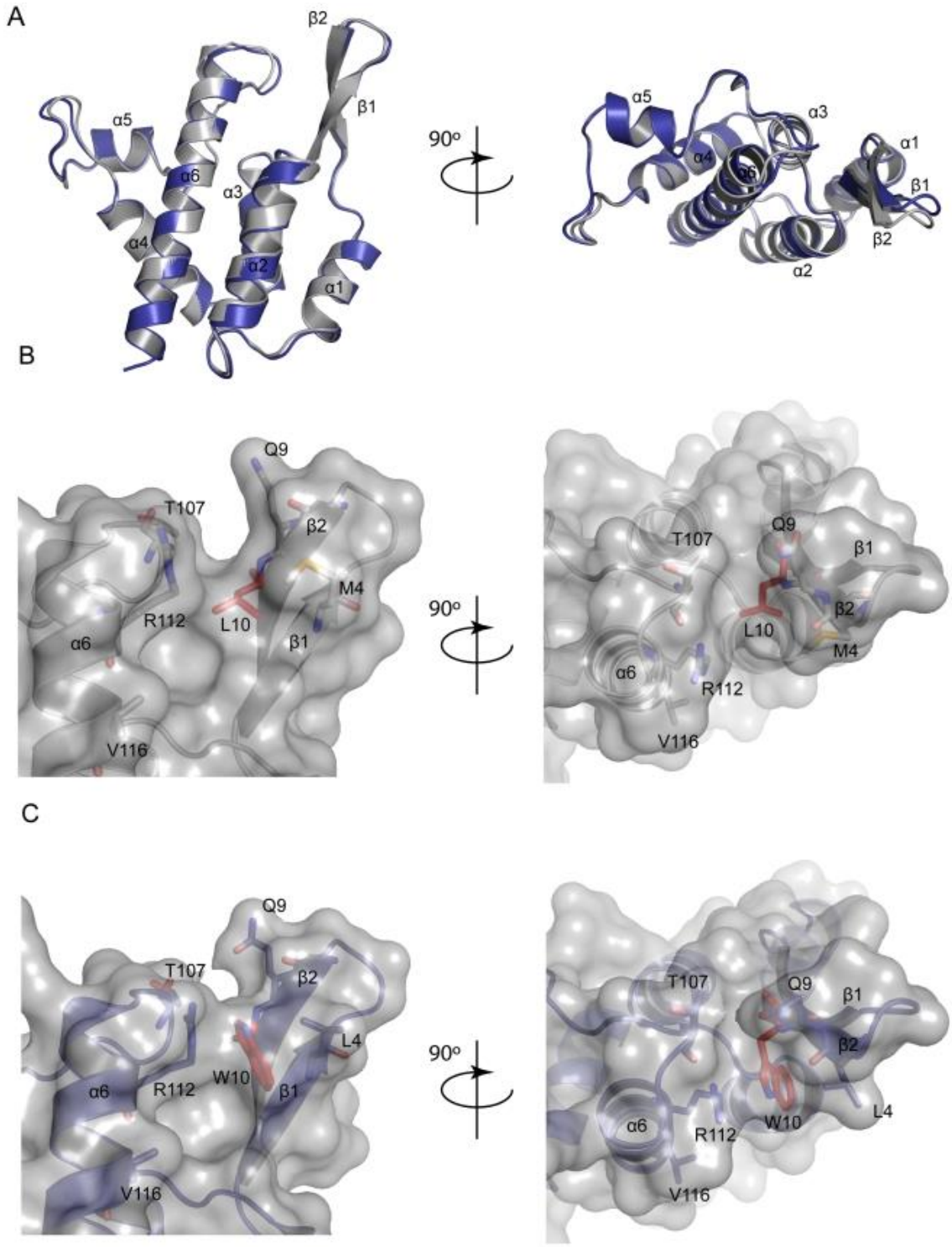


Figure 4.9: The impact of L10W on the structure of N-MLV CA

(A) Structural superimposition of wild-type N-MLV CA (grey) and the L10W mutant (blue). **(B)** The β-α6 channel in wild-type CA, with residue L10 indicated. **(C)** In the L10W mutant, the bulky side-chain of W10 obscures the base of this channel. Figure is taken from Ohkura et al. (2011), and reproduced here under the terms of the Creative Commons Attribution Licence.

The double mutants, L10W/E100K and N82D/N113K, were the only viruses in this panel to fully escape restriction by all five restriction factor constructs. While the molecular basis for these data is unclear, work from Ohkura & Stoye (2013) demonstrated the importance of charge for the escape of a similar mutant (N82D/E100K). Specifically, it was shown that removing arginine residues from the VR1 loop of huT5α (or introducing them at the corresponding positions in rhT5α) can restore restriction of this virus. While it is not clear whether charge-reversing mutations would have a similar impact on the restriction of L10W/E100K and N82D/N113K, it is noteworthy that the VR1 loop of T1 B30.2 contains 8 charged residues (3 acidic and 5 basic), providing ample opportunity to explore this possibility in the future.

Collectively, these data indicate that T1 and 5α engage with distinct epitopes on the N-MLV capsid surface, although it is likely that certain residues and motifs contribute to both binding interfaces. This probably reflects underlying structural differences in the B30.2 domains of the two factors. In order to confirm this hypothesis, we have successfully expressed and purified a recombinant form of the T1 B30.2 domain for crystallisation trials. Although no crystals were grown within the timeframe of this project, the high quality of protein obtained from the purification has provided ample raw material for further work in this area. Future attempts may benefit from using MBP as a seed to nucleate crystal formation, or perhaps introducing a cleavage site upstream of the B30.2 domain to enable removal of the tags after purification.

Collectively, the data presented in this chapter indicate that, despite their divergence, T1 and 5α share a common overall design that enables them to complement one another in a contiguous molecule. However, there are important differences that set these proteins apart, including their mechanisms of capsid recognition and their intramolecular requirements for effecting restriction. The latter point will be the central focus of the next chapter.

Chapter 5 Characterising the requirements for a productive TRIM-capsid interaction

Chapter **4** outlined efforts to produce a restriction-competent chimera of T1 and 5α. Given that all of the constructs described had intact capsid-binding and effector domains, it is curious that only one of them retained the ability to restrict N-MLV. This observation led us to consider the notion that it is not only the presence of certain domains that is important for restriction, but also the relative spacing between them.

Linker 2 (L2) is a 64 amino acid region that runs between the coiled-coil and B30.2 domain of T5α (Figure **5.1A**). L2 has been under extensive scrutiny in recent years, and a wealth of information is now available regarding its structure and contribution to the T5α dimer (Goldstone et al., 2014a; Sanchez et al., 2014). The N-terminal portion consists of two α-helices, α3 and α4, which are followed by an extended and largely unstructured region known as L2-E. L2-E runs for 24 residues before terminating in a final helix, α5, which forms the C-terminal connection with the B30.2 domain. In the T5α dimer, α3 and α4 interact with the B-boxes of the opposing monomer at either end of the dimer. L2-E then doubles back along the coiled-coil and runs towards the twofold axis of the dimer, thereby positioning α5 at the centre, where it serves to align the B30.2 domains with complementary epitopes on the capsid surface (Figure **5.1B**).

The equivalent region in T1 is about three times as long (195 residues), and includes a COS domain (60 aa) for microtubule binding and an FN-III domain (100 aa) for mediating protein-protein interactions. No structural information about the T1 dimer is currently available (nor, indeed, is it known whether T1 dimerisation is required for restriction at all), but it is likely that it differs substantially from T5α in this respect. This may go some way towards explaining the relative difficulty in producing a functional chimera of the two proteins.

The aim of the studies described in this chapter is to explore the impact of domain spacing on facilitating a productive interaction between TRIM and capsid. The main focus of these studies will be the L2 region of T5α, since this is shorter and better characterised than its counterpart in T1. For the purpose of this work, a *productive interaction* is defined as one that results in restriction of the virus concerned.

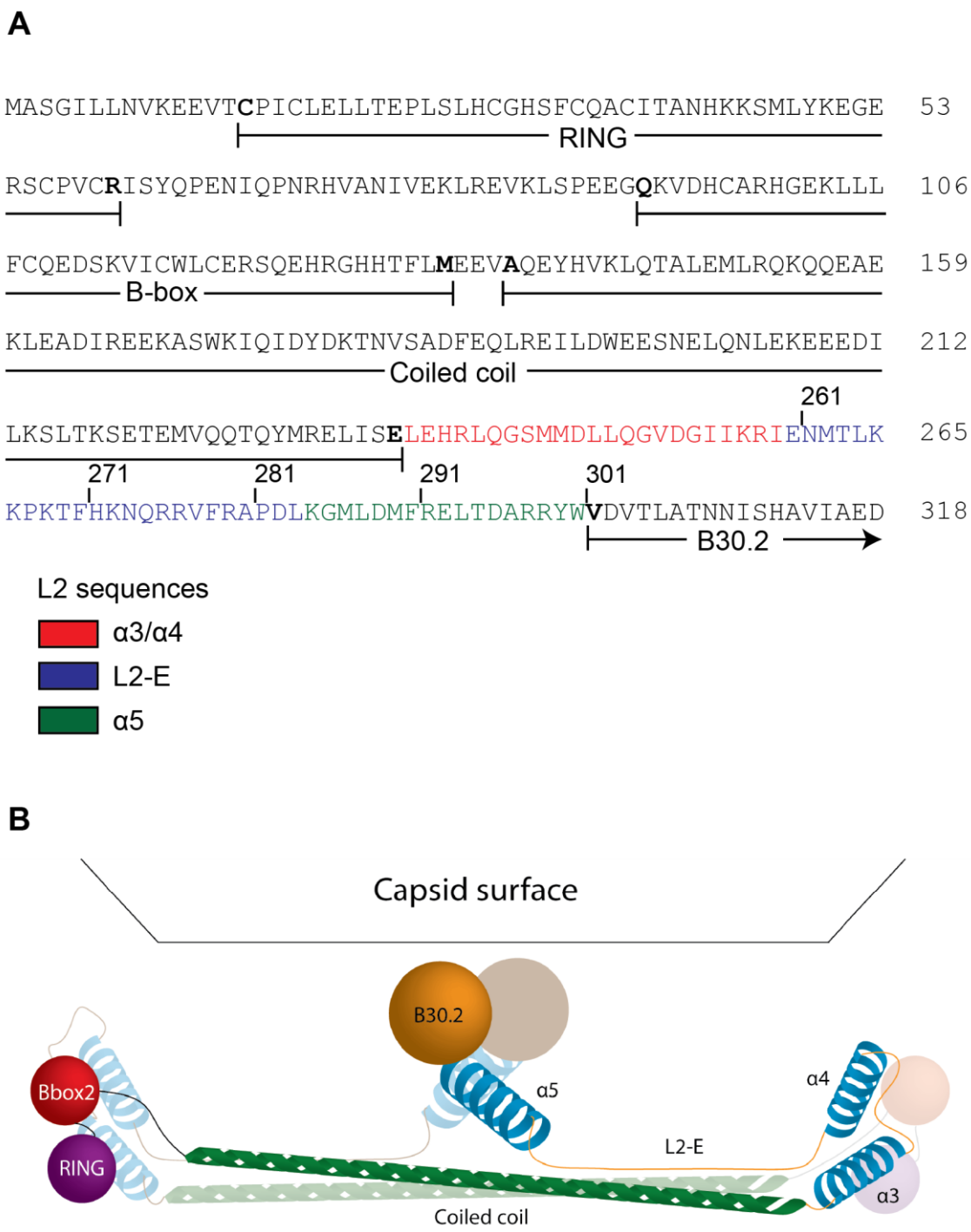


Figure 5.1: The T5α L2 region

(A) Position of L2 within the rhT5α protein sequence. **(B)** Side-view of a T5α dimer interacting with the capsid surface.

5.1 Rhesus T5α is largely intolerant of deletions in L2

To investigate the importance of L2 length for permitting a productive interaction between T5 and capsid, a panel of deletions were made in this region of rhT5α. The rhesus orthologue was chosen because it potently restricts a number of retroviruses, thereby providing a suitable baseline from which to compare any mutant phenotypes.

In an initial experiment, the entirety of L2-E and the N-terminus of α5 (L2-α5 herein) were deleted from rhT5α to yield Construct **A** (Table **5.1**; Figure **5.2**). Structural modelling of this mutation predicts repositioning of the B30.2 domains to either end of the dimer (Figure **5.3**). Unsurprisingly, this construct was inactive against all four of the viruses tested. Following this observation, three smaller deletions were made in L2-α5, each containing progressively larger amounts of L2-E, but all still lacking the N-terminus of α5 (**B-D**). However, these constructs were also largely inactive. In fact, only **D** – which retains all but two residues of L2-E – yielded a protein with any restriction activity at all: a partial phenotype against FIV.

To verify whether the inactivity of **B** and **C** was attributable to the missing N-terminus of α5, two more constructs were made with equivalent deletions in L2-E, but this time retaining helix α5 in its entirety (**H** and **I**). However, as Table **5.1** shows, the presence of α5 was insufficient to rescue restriction by these constructs, indicating that the portions of L2-E removed in **B** and **C** are necessary, either in size or sequence, for a productive capsid interaction.

To further investigate the partially restrictive phenotype of **D**, the region concerned (residues 282-290) was split into two halves, and deletions made of each, yielding **E** (Δ282-285) and **F** (Δ286-290). *In silico* modelling predicted similar structural manifestations for both mutations: the central position of the B30.2 domains is roughly maintained, but with a ~90° rotation relative to the axis of the coiled-coil (Figure **5.3**). Interestingly, **E** and **F** regained full activity against the lentiviruses, but were unable to restrict N-MLV.

To establish whether this phenotype is unique to the regions deleted in **E** and **F**, or simply a general consequence of shortening helix α5, a similarly sized deletion was made in the C-terminal half of α5 to produce **G**. However, **G** was entirely inactive, indicating that the ability of **E** and **F** to recognise the lentiviral capsids is attributable to dispensable residues in the N-terminus of α5, rather than a general tolerance of α5 shortening.

		HIV-1	HIV-2	FIV	N-MLV
	RhT5α (wild-type)	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.08 ± 0.02
A	RhT5α Δ(259-290)	1.13 ± 0.01	1.04 ± 0.02	1.21 ± 0.09	1.13 ± 0.01
B	RhT5α Δ(263-290)	1.12 ± 0.02	1.01 ± 0.01	1.19 ± 0.08	1.13 ± 0.02
C	RhT5α Δ(272-290)	1.09 ± 0.01	0.97 ± 0.03	1.32 ± 0.11	1.12 ± 0.03
D	RhT5α Δ(282-290)	0.81 ± 0.04	0.91 ± 0.03	0.44 ± 0.06	1.10 ± 0.01
E	RhT5α Δ(282-285)	0.05 ± 0.00	0.12 ± 0.03	0.07 ± 0.01	1.00 ± 0.03
F	RhT5α Δ(286-290)	0.05 ± 0.01	0.03 ± 0.00	0.04 ± 0.00	0.90 ± 0.07
G	RhT5α Δ(291-295)	1.08 ± 0.02	1.01 ± 0.01	1.03 ± 0.02	1.11 ± 0.01
H	RhT5α Δ(263-280)	1.12 ± 0.01	1.00 ± 0.02	1.30 ± 0.08	1.11 ± 0.01
I	RhT5α Δ(272-280)	1.13 ± 0.01	0.97 ± 0.03	1.28 ± 0.08	1.13 ± 0.01

Table 5.1: Restriction phenotypes of the rhT5α deletion constructs

Data show the mean and SEM (2 dp) from three independent experiments (*n* = 6). Black boxes represent full restriction; grey, partial, and white, absence of restriction

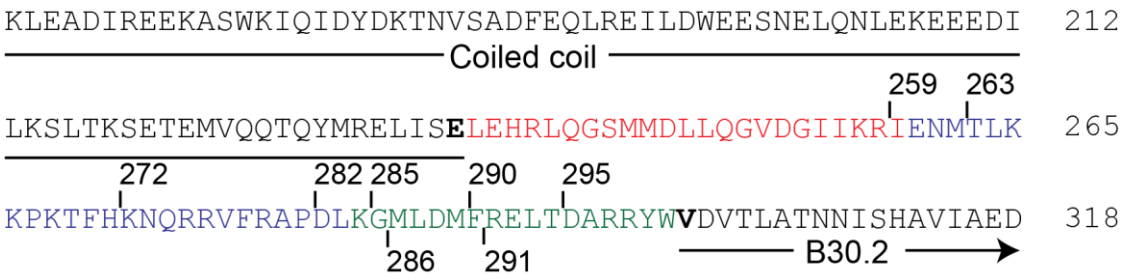


Figure 5.2: Positions of the removed portions in the rhT5α deletion constructs

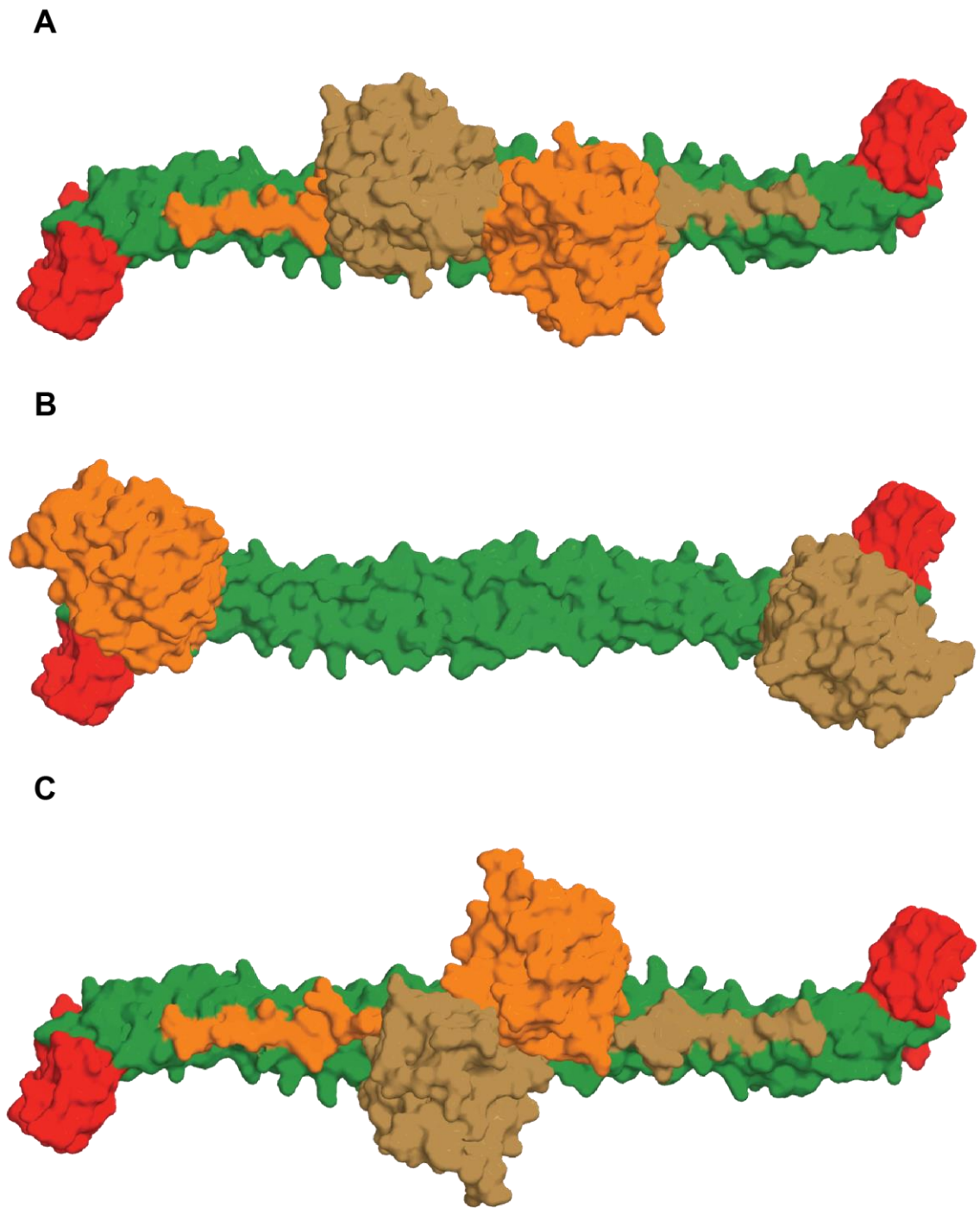


Figure 5.3: A structural model of wild-type and mutant T5α dimers

(A) Wild-type; **(B)** Δ259-290 (Construct **A**); (C) Δ282-285/286-290 (Constructs **E** and **F**). Models were constructed using the Pymol software. Credit to Dr Neil Ball at the Francis Crick Institute, Mill Hill Laboratory.

5.2 Rhesus T5α tolerates small extensions in α5

Having deleted portions of L2-α5, we next wanted to examine the consequences of extending it. To this end, three sequences were inserted roughly in the middle of the α5 helix, specifically between residues 289 and 290 (see Table **5.2**). These sequences were designed specifically to lengthen the helix while preserving its secondary structure.

	Conformational effects on B30.2 domains	
	Separation in dimer	Rotation relative to CC
289-***FRE***-290	8 Å	– 60°
289-***FREL***-290	12 Å	40°
289-***FRELFRL***-290	22 Å	– 20°

Table 5.2: Sequences inserted into the centre of helix α5 in rhT5α

The conformational effects of each insertion on the T5α dimer are shown in the right hand columns. Credit to David Goldstone (former Division of Molecular Structure, NIMR) for the design of these mutants.

Figure **5.4** shows the predicted effect of the largest insertion, *FRELFRL*, on the structure of the T5α dimer: the B30.2 domains are pushed apart in opposite directions and swivelled with respect to the axis of the dimer. Strikingly, the restriction data for these constructs were reminiscent of those seen with Constructs **E** and **F** in Section **5.1**: in all three cases, lentiviral recognition was unperturbed, while N-MLV recognition was lost (Table **5.3**).

	HIV-1	HIV-2	FIV	N-MLV
RhT5α (wild-type)	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.08 ± 0.02
RhT5α (289FRE290)	0.04 ± 0.00	0.05 ± 0.01	0.03 ± 0.01	0.85 ± 0.04
RhT5α (289FREL290)	0.03 ± 0.00	0.05 ± 0.01	0.03 ± 0.00	0.92 ± 0.03
RhT5α (289FRELFRL290)	0.04 ± 0.00	0.03 ± 0.00	0.03 ± 0.00	1.12 ± 0.01

Table 5.3: Restriction phenotypes of wild-type rhT5α and a panel of constructs with extended α5 helixes

Data show the mean and SEM (correct to 2 dp) from three independent experiments (*n* = 6). Black boxes represent full restriction; grey, partial restriction, and white, absence of restriction.

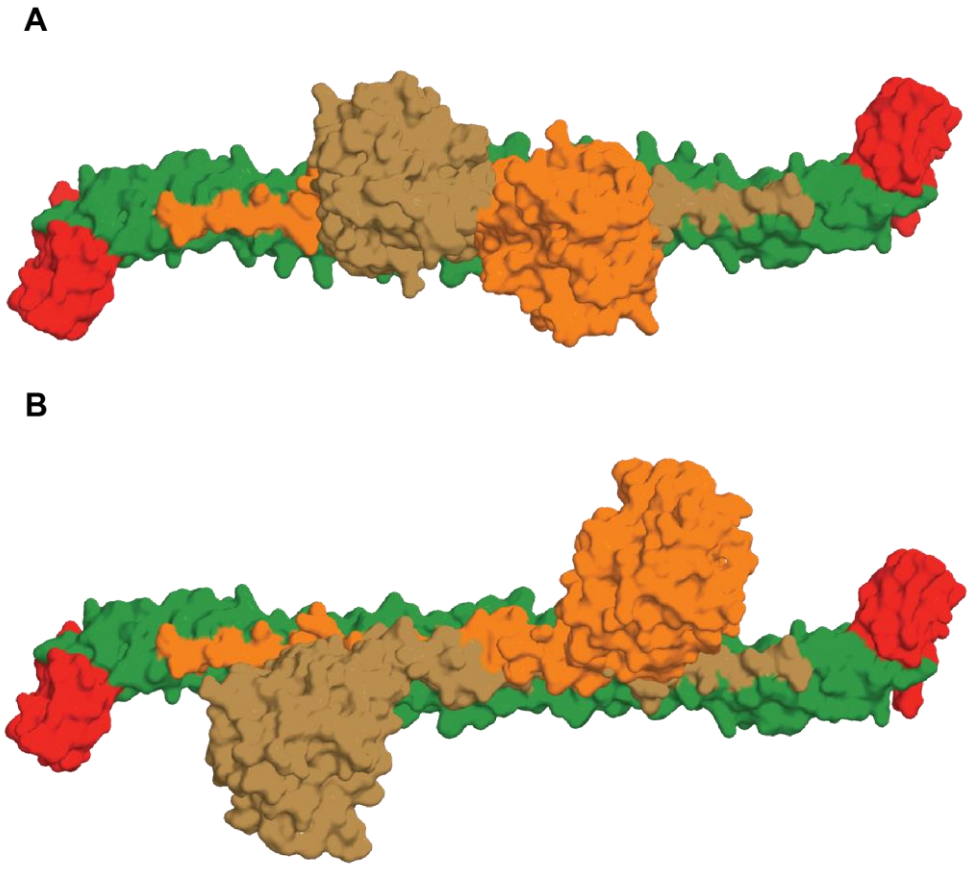


Figure 5.4: A structural model of wild-type and mutant T5α dimers

(A) Wild-type; **(B)** 289-FRELFRL-290. Models were constructed using the Pymol software. Credit to Dr Neil Ball at the Francis Crick Institute, Mill Hill Laboratory.

5.3 Disrupting the secondary structure of α5 has variable effects on restriction by rhesus and human T5α

Having altered the length of L2-α5, we next sought to examine the consequences of disrupting its secondary structure. To this end, an alignment of the α5 sequences from multiple primate T5αs was performed using an online protein alignment tool (Table **5.4**). This analysis revealed two highly conserved leucine residues that are likely to contribute to helix formation in this region. Thus, each of these residues was mutated to proline – which has a propensity to disrupt α-helices – and the effect on restriction measured.

		Helix α5
Hominidae	**Human** (*Homo sapiens*)	KGM**L**EVFRE**L**TDVRRYW
	Chimpanzee (*Pan troglodytes*)	KGM**L**EVFRE**L**TDVRRYW
	Gorilla (*Gorilla gorilla*)	KGM**L**EVFRE**L**TDVRRYW
	Orangutan (*Pongo pygmaeus*)	KGM**L**EVFRE**L**TDVRRYW
	White-cheeked gibbon (*Nomascus leucogenys*)	KVM**L**EELRE**L**RDVQHYW
Old world monkeys	**Rhesus macaque** (*Macaca mulatta*)	KGM**L**DMFRE**L**TDARRYW
	African green monkey (*Chlorocebus aethiops*)	KGM**L**DMFRE**L**TDVRRYW
	Sooty mangabey (*Cercocebus atys*)	KGM**L**DMFRE**L**TDVRRYW
New world monkeys	**Red-bellied tamarin** (*Saguinus labiatus*)	KAM**L**QAFKE**L**TEVQRYW
	Pygmy marmoset (*Callithrix pygmaea*)	KGM**L**QAFKE**L**TEVQRYW

Table 5.4: An alignment of α5 sequences from ten primate orthologues of T5α

Conserved leucine residues are highlighted in red. Alignment was performed using the Clustal Omega multiple sequence alignment tool (EMBL-EBI).

As Table **5.5**: shows, the effects of proline substitution on restriction were highly variable between constructs. In rhT5α, L287P resulted in a rather dramatic phenotype: HIV-2 restriction was lost entirely, while HIV-1 and FIV restriction were both attenuated. L293P was much better tolerated, with no discernible effect on lentiviral restriction. Recognition of N-MLV, however, was lost in both cases. In huT5α, the equivalent mutations had an even more profound effect, with substitution of either residue completely ablating restriction of this panel of viruses.

	HIV-1	HIV-2	FIV	N-MLV
RhT5α (wild-type)	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.08 ± 0.02
RhT5α (L287P)	0.33 ± 0.04	0.75 ± 0.04	0.48 ± 0.04	1.14 ± 0.01
RhT5α (L293P)	0.05 ± 0.00	0.05 ± 0.01	0.04 ± 0.00	1.13 ± 0.01

	HIV-1	HIV-2	FIV	N-MLV
HuT5α (wild-type)	1.17 ± 0.03	0.24 ± 0.01	0.52 ± 0.01	0.06 ± 0.01
HuT5α (L285P)	1.15 ± 0.01	1.03 ± 0.02	1.08 ± 0.03	1.11 ± 0.01
HuT5α (L291P)	1.18 ± 0.03	1.06 ± 0.02	1.01 ± 0.02	1.11 ± 0.01

Table 5.5: Restriction phenotypes of wild-type rhesus and human T5α and a panel of constructs with leucine-to-proline substitutions in helix α5

Data show the mean and SEM (correct to 2 dp) from two (N-MLV) or three (HIV-1, HIV-2, FIV) independent experiments (*n* = 4 and 6, respectively). Black boxes represent full restriction; grey, partial restriction, and white, absence of restriction.

5.4 CypA tolerates L2 deletions more readily than the B30.2 domain

Section **5.1** described the impact of L2 deletions on the restriction phenotype of rhT5α. To determine whether these observations are unique to restriction factors containing a B30.2 domain, we synthesised a chimera comprising the N-terminus of rhT5α (residues 1-300) fused to the CypA domain of owl monkey TCyp (Yap, unpublished work). This construct was designed specifically to incorporate the L2 region of rhT5α in its entirety, thereby enabling a comparison of the two capsid-binding domains in the same N-terminal background. The resulting molecule (T5α-Cyp herein) was restriction-competent (Table **5.6**, Construct **A**).

A number of deletions were then made in the L2 region of this construct by site-directed mutagenesis (Constructs **B-F**). As the data show, these modifications were far better tolerated than their equivalents in rhT5α. For all but the largest deletion (**F**), HIV-1 and FIV restriction were fully preserved. The pattern for HIV-2 restriction was a little less clear: interestingly, **B** restricted this virus more potently than the parental **A**, while **C-F** were permissive. This phenomenon is discussed in greater detail in Section **5.5**.

		HIV-1	HIV-2	FIV
A	T5α-Cyp	0.05 ± 0.01	0.68 ± 0.03	0.04 ± 0.01
B	T5α-Cyp Δ(292-299)	0.03 ± 0.00	0.05 ± 0.00	0.02 ± 0.00
C	T5α-Cyp Δ(282-299)	0.05 ± 0.01	1.01 ± 0.01	0.04 ± 0.00
D	T5α-Cyp Δ(272-299)	0.06 ± 0.00	1.02 ± 0.01	0.04 ± 0.01
E	T5α-Cyp Δ(263-299)	0.13 ± 0.01	1.05 ± 0.02	0.05 ± 0.00
F	T5α-Cyp Δ(259-299)	0.98 ± 0.01	1.02 ± 0.01	0.45 ± 0.03

Table 5.6: Restriction phenotypes of an artificial T5α-Cyp chimera and a panel of daughter constructs with L2 deletions

Data show the mean and SEM (correct to 2 dp) from two independent experiments (n = 4). Black boxes represent full restriction; grey, partial restriction, and white, absence of restriction.

5.5 The restriction specificity of TCyp is governed by multiple determinants

It is curious that removing just 9 amino acids from T5α-Cyp boosts HIV-2 restriction by more than tenfold, whilst having no concomitant effect on the inhibition of HIV-1 or FIV (Table **5.6**, Constructs **A** and **B**). This observation prompted us to conduct further enquiries into the determinants that govern restriction specificity in TCyp restriction factors. This section is thus divided into three parts, each one exploring a separate determinant.

5.5.1 Exon 7

The 9 amino acid difference between Constructs **A** and **B** falls within the C-terminal half of helix α5, and is encoded by exon 7 (e7) of the *T5* gene. TCyp fusions have independently arisen in several primate lineages, but due to differences in the site of CypA retrotransposition, the owl monkey variant (omTCyp) retains e7, while the rhesus macaque orthologue (rhTCyp) does not. These molecules also differ in their restriction profiles: the former potently restricts HIV-1, while the latter does not.

To establish whether this difference can be attributed to the presence or not of e7, this region was deleted from omTCyp and the restriction activity of the resulting molecule measured (Figure **5.5**). Curiously, we observed restriction of HIV-2 in these experiments, which is at odds with the restriction profile reported elsewhere (Wilson et al., 2008). Nevertheless, the inhibition of HIV-1 and FIV is consistent with the literature, and neither of these activities were affected by the removal of e7, indicating that this region is inconsequential for restriction by omTCyp. This stands in marked contrast to the situation with T5α-Cyp, where the presence of e7 appears to be inhibitory (Section **5.4**, Constructs **A** and **B**) and in rhT5α, where it is essential (Section **5.1**, Construct **G**).

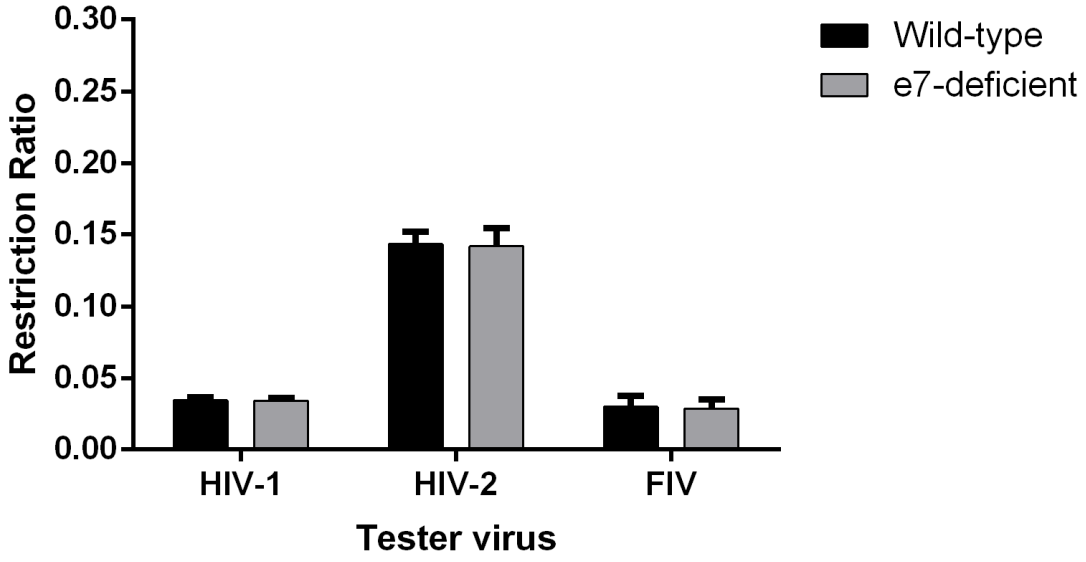


Figure 5.5: Restriction phenotypes of wild-type and e7-deficient omTCyp
Graph shows the mean and SEM from two independent experiments (*n* = 4).

Collectively, these data suggest that the impact of e7 on TRIM-mediated restriction is highly context-dependent, with influences from both the downstream capsid-binding domain and the target virus concerned. It is thus difficult to make any broad generalisations about its function on the basis of these data.

5.5.2 Residues in the active site of CypA

Modern rhTCyp has a strong affinity for the HIV-2 capsid, while the parental CypA molecule from which it is derived has a strong affinity for HIV-1 and only a weak one for HIV-2. One group has suggested that this affinity switch can be attributed to two mutations in the active site of the CypA domain (Price et al., 2009). The first of these (R69H) broadened specificity from HIV-1 to both viruses, while the second (D66N) enhanced HIV-2 restriction while ablating HIV-1 restriction entirely.

To test this model in T5α-Cyp and its derivatives, the D66N mutation was introduced to each of Constructs **A**-**F** (Table **5.6**), yielding **A1-F1** (Table **5.7**). The expectation was that this would reverse the phenotypes described in Section **5.4**: thus, in **C1-F1**, HIV-1 restriction would be lost and HIV-2 restriction gained.

Consistent with the findings of Price et al. (2009), this modification inhibited HIV-1 restriction; surprisingly, however, there was no concomitant gain-of-function against HIV-2. FIV restriction was much less affected by this change.

		HIV-1	HIV-2	FIV
A1	T5α-Cyp D66N	1.15 ± 0.02	0.99 ± 0.01	0.03 ± 0.00
B1	T5αCyp Δ(292-299) D66N	1.08 ± 0.01	0.76 ± 0.02	0.05 ± 0.00
C1	T5αCyp Δ(282-299) D66N	1.11 ± 0.03	0.97 ± 0.01	0.13 ± 0.01
D1	T5αCyp Δ(272-299) D66N	1.06 ± 0.03	1.01 ± 0.02	0.52 ± 0.03
E1	T5αCyp Δ(263-299) D66N	1.10 ± 0.01	1.00 ± 0.01	1.38 ± 0.01
F1	T5αCyp Δ(259-299) D66N	1.13 ± 0.02	1.02 ± 0.04	1.29 ± 0.00

Table 5.7: Restriction phenotypes of T5α-Cyp and its derivatives with the D66N active site mutation in CypA

Data show the mean and SEM (correct to 2 dp) from two independent experiments (*n* = 4). Black boxes represent full restriction; grey, partial restriction, and white, absence of restriction.

Next, the R69H mutation was introduced to **A** and **B**, yielding **A2** and **B2** (Figure **5.6**). This modification caused **A** to completely lose its limited ability to restrict HIV-2, although HIV-1 restriction was unaffected. **B** underwent a very slight attenuation of HIV-2 restriction – again, with no impact on the restriction of HIV-1. Finally, the two mutations were introduced to **A** and **B** in combination to produce **A3** and **B3**. Both resulting constructs lost activity against HIV-1. **A3** also lost the ability to restrict HIV-2, while for **B3** this function was merely attenuated.

Collectively, these data suggest that while R69H and D66N might have converted the restriction specificity of rhesus CypA from HIV-1 to -2, this is not the case for owl monkey CypA, at least in the context of this artificial chimera.

5.5.3 Leader sequence

In the splicing of *T5* and *CypA* to produce a TCyp fusion, the splice acceptor site lies slightly upstream of the *CypA* coding region. The resulting protein thus contains a short stretch of amino acids immediately N-terminal to the CypA domain known as the leader sequence. This sequence differs slightly between the owl monkey and rhesus orthologues (Table **5.8**).

	Leader sequence
Owl monkey (*Aotus trivirgatus*)	DAAAWD−LVASA
Rhesus macaque (*Macaca mulatta*)	DAAAEESPVLLA

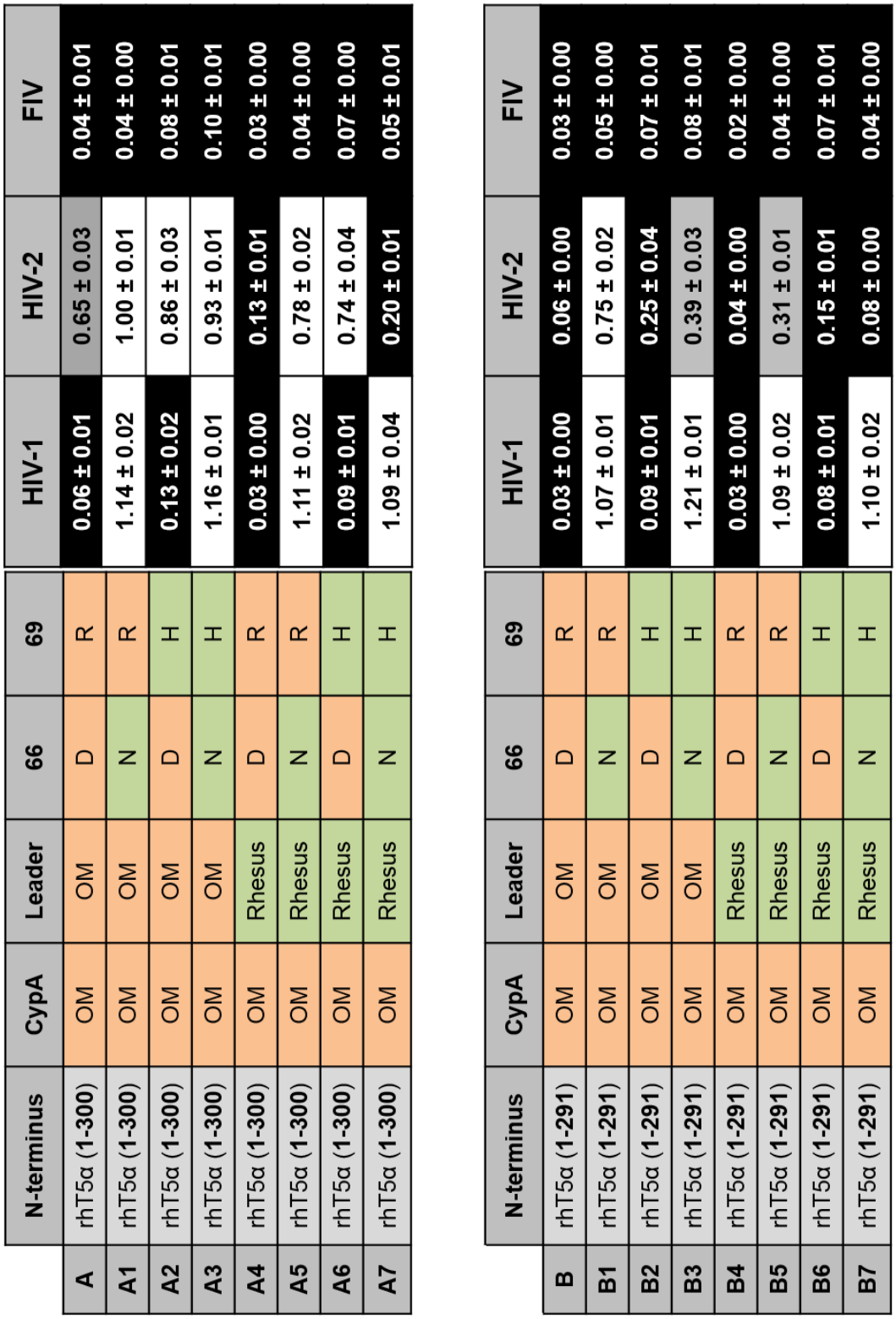
Table 5.8: The leader sequences of owl monkey and rhesus macaque TCyp

To investigate the impact of this region on restriction specificity, the owl monkey leader in Constructs **A** and **B** (Table **5.6**), was substituted for that of the rhesus macaque to produce **A4** and **B4** (Figure **5.6**). This modification had no effect on **B**; however, interestingly, it rescued the attenuated restriction of HIV-2 by **A**. This suggests that the inhibitory effect of e7 in this context can be reversed by simply substituting the downstream leader sequence.

Next, the impact of leader sequence substitution in the context of the mutations described in Section **5.5.2** was investigated. To this end, the rhesus leader was again introduced to **A** and **B**, but this time in combination with either D66N, R69H, or both (Figure **5.6**).

In the case of **A**, the rhesus leader lost the ability to rescue the inhibitory effect of e7 on HIV-2 restriction when either D66N or R69H was also present (compare **A4** with **A5** and **A6**, respectively). Curiously, though, when the rhesus leader was paired with the double mutant (**A7**), HIV-2 restriction was restored. Conversely, in the case of **B**, the rhesus leader rescued the attenuating effects of both D66N and R69H on HIV-2 restriction (compare **B5** with **B1**, and **B6** with **B2**, respectively). Interestingly, the double mutant (**B7**) restricted HIV-2 to a comparable degree as the one lacking either mutation (**B4**), despite the fact that each of these mutations attenuated HIV-2 restriction when present in isolation.

It is noteworthy that a strong inhibition of FIV was preserved for every construct described in this section. This served as a useful internal control, indicating that all of the constructs were expressed and active, albeit with modified restriction profiles.



	N-terminus	CypA	Leader	66	69	HIV-1	HIV-2	FIV
A	rhT5α (1-300)	OM	OM	D	R	0.06 ± 0.01	0.65 ± 0.03	0.04 ± 0.01
A1	rhT5α (1-300)	OM	OM	N	R	1.14 ± 0.02	1.00 ± 0.01	0.04 ± 0.00
A2	rhT5α (1-300)	OM	OM	D	H	0.13 ± 0.02	0.86 ± 0.03	0.08 ± 0.01
A3	rhT5α (1-300)	OM	OM	N	H	1.16 ± 0.01	0.93 ± 0.01	0.10 ± 0.01
A4	rhT5α (1-300)	OM	Rhesus	D	R	0.03 ± 0.00	0.13 ± 0.01	0.03 ± 0.00
A5	rhT5α (1-300)	OM	Rhesus	N	R	1.11 ± 0.02	0.78 ± 0.02	0.04 ± 0.00
A6	rhT5α (1-300)	OM	Rhesus	D	H	0.09 ± 0.01	0.74 ± 0.04	0.07 ± 0.00
A7	rhT5α (1-300)	OM	Rhesus	N	H	1.09 ± 0.04	0.20 ± 0.01	0.05 ± 0.01
B	rhT5α (1-291)	OM	OM	D	R	0.03 ± 0.00	0.06 ± 0.00	0.03 ± 0.00
B1	rhT5α (1-291)	OM	OM	N	R	1.07 ± 0.01	0.75 ± 0.02	0.05 ± 0.00
B2	rhT5α (1-291)	OM	OM	D	H	0.09 ± 0.01	0.25 ± 0.04	0.07 ± 0.01
B3	rhT5α (1-291)	OM	OM	N	H	1.21 ± 0.01	0.39 ± 0.03	0.08 ± 0.01
B4	rhT5α (1-291)	OM	Rhesus	D	R	0.03 ± 0.00	0.04 ± 0.00	0.02 ± 0.00
B5	rhT5α (1-291)	OM	Rhesus	N	R	1.09 ± 0.02	0.31 ± 0.01	0.04 ± 0.00
B6	rhT5α (1-291)	OM	Rhesus	D	H	0.08 ± 0.01	0.15 ± 0.01	0.07 ± 0.01
B7	rhT5α (1-291)	OM	Rhesus	N	H	1.10 ± 0.02	0.08 ± 0.00	0.04 ± 0.00

Figure 5.6: Restriction phenotypes of Constructs A (e7-proficient) and B (e7-deficient) and their derivatives.

Data show the mean and SEM (correct to 2 dp) from two independent experiments (n = 4). Black boxes represent full restriction; grey, partial restriction, and white, absence of restriction.

5.6 Discussion

While T1 and 5α share common capsid-binding and effector modules, they differ substantially in how these regions are linked to one another. In T5α, the coiled-coil and B30.2 domain are bridged by a 64-residue, highly α-helical Linker 2 (L2) region. However, the analogous region in T1 is approximately three times as long and comprises two additional domains.

Interestingly, while the T1 B30.2 domain is still able to interact with N-MLV when partnered with the L2 region of T5α, the reciprocal chimera (T5α B30.2 partnered with the equivalent portion of T1) is inactive (Section **4.1**). This implies that the relative position of the domains in T5α has been finely tuned over the course of evolution. The aim of this chapter was thus to characterise the impact of intramolecular domain spacing on the T5-capsid interaction. Each of the key findings arising from this work will now be discussed in turn.

Modifications to the L2 region have variable effects on restriction by rhT5α

To investigate the impact of L2 length on restriction by rhT5α, we deleted various portions within this region and measured the effect on restriction. These modifications were broadly deleterious, with the vast majority yielding proteins with no restriction activity. While the immediate implication of this observation is that the length of L2 influences capsid recognition, work published by the Campbell group identified two stretches of amino acids within this region (266-268 and 275-277) that are required for both cytoplasmic body localisation and retroviral restriction (Sastri et al., 2010). It is therefore possible that the loss-of-function exhibited by the mutants in this project is not related to changes in L2 length, but rather a direct consequence of removing these critical motifs.

While it is not possible to formally exclude this possibility, the fact that the removal of residues 282-290 (which excludes both of these motifs) also impedes restriction suggests that the modulation of L2 length alone has an impact on restriction activity. Indeed, this is unsurprising in the context of a T5α dimer, where substantial losses in L2 would cause a dramatic reorientation of the B30.2 domains relative to the

capsid surface, and the potential loss of important contacts between restriction factor and virus. Nevertheless, constructs with very small deletions (~ 4 aa) at the L2-E/α5 boundary, and in the N-terminus of α5 itself, retained restriction activity against the lentiviruses.

Having considered the impact of shortening the L2 region, we next turned our attention to lengthening it: specifically, by extending the α5 helix. To this end, three mutations (*FRE, FREL* and *FRELFRL*) were added to the centre of the helix in order to extend it by 1-2 additional turns. *In silico* modelling of the longest insertion in the context of the dimer revealed that this change not only separates the B30.2 domains by 22 Å, but also causes their rotation relative to the underlying coiled-coil. Despite these substantial conformational changes, however, all three mutants remained restriction-competent against the lentiviruses (although, again, N-MLV restriction was lost).

It has previously been reported that the α-helical content of L2 is critical for restriction by rhT5α, perhaps by enabling it to dock on the coiled-coil and thus orient the B30.2 domains appropriately (Sastri et al., 2010; 2014). However, this analysis focused on short stretches of amino acids within L2-E. We therefore sought to examine the effect of secondary structural modifications on the downstream α5 helix.

To do this, we identified two conserved leucine residues within α5 and substituted each for proline in both rhesus and human T5α. While leucine has a high propensity to form α-helices, the side-chain of proline imposes a steric barrier to helix formation. Surprisingly, the phenotypes arising from these mutations differed considerably depending on which position was mutated. While L287P completely inhibited restriction of HIV-2 by rhT5α (and attenuated that of HIV-1 and FIV), L293P bore no effect on lentiviral restriction. The impact of L287P is particularly surprising, given that deletion of the region including residue 287 (Δ286-290) bears no impact on lentiviral restriction. Yet again, anti-N-MLV activity was lost in both cases.

The equivalent mutations in huT5α had even more profound consequences: mutating either residue to proline ablated activity against both the lentiviruses *and* N-MLV. It is unclear whether the discrepant effects of proline mutagenesis on rhesus and human T5α are due to divergence in the surrounding sequences, or fundamental differences in the mechanism by which each orthologue binds to its target. In any case, further characterisation of these mutants in order to confirm the predicted changes in their secondary structures (perhaps by circular dichroism or NMR), alongside a western blot to confirm their expression, will be necessary before any meaningful conclusions can be drawn.

N-MLV restriction is highly sensitive to L2 modifications

An interesting phenotype exhibited by a number of the L2 mutants described in this chapter was a complete loss of N-MLV restriction, even when lentiviral restriction was preserved. This is particularly striking given the relatively conservative nature of these mutations: the deletion of just four amino acids at the L2-E/α5 boundary, for example.

This observation may reflect the fact that lentiviruses possess conical cores, while for N-MLV and other gammaretroviruses, this structure is spherical. In the absence of high-resolution structural data, it is hard to envisage precisely why a spherical core would place such tight restrictions on the T5α dimer, but it is possible that the higher degree of curvature on the capsid surface would position the viral epitopes in such a way as to constrain the conformational freedom of the B30.2 domains. Validating this hypothesis will require pairing the L2 mutants constructed in this project with other lenti- and gammaretroviruses, and perhaps members of other retroviral genera as well.

CypA tolerates L2 modifications more readily than the B30.2 domain

While the B30.2 domain consists of four variable loops – at least three of which make critical contacts with the capsid target (Ohkura et al., 2006) – the capsid-binding domain of TCyp need only interact with a single viral epitope: the Cyp-binding loop (Yoo et al., 1997). This endows the latter with the freedom to

sample multiple monomers within a CA hexamer while still maintaining a productive interaction. We therefore expected that modifications to the L2 region of TCyp might have a less severe impact on restriction than was the case for rhT5α.

To examine this possibility, we generated a chimera consisting of the N-terminal portion of rhT5α (up to and including helix α5) fused to the CypA domain of owl monkey TCyp (T5α-Cyp). In stark contrast to the data obtained from the rhT5α mutants, T5α-Cyp retained the ability to restrict HIV-1 and FIV even upon the removal of helix α5 and the majority of L2-E. However, HIV-2 restriction was lost in almost every case. Nevertheless, these data confirm that capsid recognition by CypA is relatively immune to modifications in the preceding L2 region, particularly when compared to the B30.2 domain.

The restriction specificity of TCyp is governed by multiple determinants

An interesting observation arising from the above experiments is that the removal of exon 7 (e7) from T5α-Cyp, which corresponds to just 9 amino acids, results in a >10-fold increase in the inhibition of HIV-2. This led us to consider the role of e7 in restriction.

To this end, we deleted the relevant region from wildtype omTCyp. Interestingly, we observed restriction of HIV-2 by both wildtype and e7-deficient omTCyp, which is discordant with the phenotype described by other groups (Wilson et al., 2008). The reasons for this discrepancy are unclear; however, the fact that HIV-1 and FIV restriction were also unaffected by the removal of e7 indicates that this region is dispensable for restriction by omTCyp. This is in marked contrast to the case for T5α-Cyp, where e7 has an inhibitory effect, and in rhT5α, where it appears to be at least partly required for restriction.

Collectively, these data paint a rather convoluted picture of the role of e7 in restriction. Recent work from the Zheng lab showed that e7 deficiency is associated with an inability to localise to cytoplasmic bodies and a reduced propensity for multimerisation (Liu et al., 2015). While this certainly explains the dependence of rhT5α on e7, it is not clear whether these functions are simply

dispensable in the case of omTCyp, or whether other regions compensate for the loss of e7 in this protein. It is also unclear why this region would inhibit HIV-2 restriction by the T5α-Cyp chimera. Further investigation into the role of e7 in a broader range of TRIM molecules is warranted before any meaningful conclusions can be drawn about its function.

It is interesting to note that while the majority of T5α-Cyp deletion constructs retained the ability to restrict HIV-1, most of them were inactive against HIV-2. The capsid structures of HIV-1 and -2 are broadly similar, although they differ in the conformation of the Cyp-binding loop due to the deletion of residue CA-88 in the latter. One group has proposed that rhTCyp acquired the ability to restrict HIV-2 at the expense of HIV-1 by acquiring two mutations in the active site of CypA (R69H/D66N) (Price et al., 2009). In particular, D66N is believed to have facilitated this interaction by inducing structural rearrangements within the active site that enabled it to accommodate the displaced Cyp-binding loop of HIV-2.

Although these analyses were based on the rhesus orthologue, we reasoned that the same mutations might have a similar effect in owl monkey CypA. We therefore introduced the D66N mutation to T5α-Cyp and each of its daughter constructs. Consistent with the findings of Price et al. (2009), D66N inhibited HIV-1 activity in all of these mutants; however, unlike the case for rhesus CypA, there was no concomitant gain-of-function against HIV-2.

Collectively, these observations indicate that the recognition of capsid by TCyp involves the cooperation of multiple components: both e7 and D66N yield different restriction phenotypes depending on their context. This prompted a more detailed enquiry into how the various determinants within TCyp interact to permit a productive capsid interaction. To this end, a panel of modifications were made to both T5α-Cyp and its e7-deficient counterpart. These included the D66N and R69H active site mutations, either alone or in combination; substitution of the owl monkey leader sequence for that of the rhesus macaque, and pairwise combinations thereof. This analysis yielded a number of apparently conflicting phenotypes that are difficult to reconcile.

In the case of full-length T5α-Cyp, HIV-1 restriction was ablated by D66N, but unaffected by R69H or substitution of the leader sequence. Strikingly, the attenuation of HIV-2 restriction imparted by e7 was rescued upon substitution of the owl monkey leader sequence for that of the rhesus macaque. The leader sequence is a short stretch of amino acids found just N-terminal to the CypA domain due to the presence of a splice acceptor site upstream of the *CypA* coding region. The owl monkey and rhesus leader sequences vary quite substantially; however, it is not immediately clear why the latter would rescue the attenuating effect of e7 on HIV-2 restriction while the former cannot.

When the same modifications were made to the e7-deficient form of T5α-Cyp, the resultant phenotypes were markedly different from the above. Again, HIV-1 restriction was intolerant of D66N, but resilient to both R69H and substitution of the upstream leader sequence. As expected, HIV-2 restriction was better preserved in the absence of e7; however, D66N ablated this phenotype, and both R69H and the D66N/R69H double mutant attenuated it. Nevertheless, pairing any of these mutants with a rhesus leader sequence restored HIV-2 restriction.

Given the lack of any obvious patterns in the above data, it is difficult to disentangle the contribution of each determinant to restriction. This issue is further compounded by the fact that FIV restriction was unaffected by the vast majority of modifications described in this chapter. This is particularly striking given that, like HIV-1, FIV interacts with CypA via a central proline residue (P90) in the Cyp-binding loop (Lin and Emerman, 2006). However, it is possible that divergence elsewhere in the loop contributes to the discrepant phenotypes observed for HIV and FIV in this chapter. Indeed, the Hatziioannou group have previously observed that FIV restriction is relatively unhindered by the addition of CsA, a competitive inhibitor of CypA, indicating that FIV might have an inherently higher affinity for CypA than its lentiviral cousins (Virgen et al., 2008). If this is the case, then it is perhaps unsurprising that FIV restriction is robust in the face of TCyp modifications.

To summarise, this chapter has provided insight into several intramolecular requirements for facilitating a productive interaction between TRIM and capsid. While it appears that these are most stringent for restriction factors with a B30.2 domain, there are also determinants immediately preceding the CypA domain that may affect restriction specificity. However, defining the precise contribution of each of these determinants is difficult in the absence of structural data.

Chapter 6 Conclusions

The focus of this thesis has been on retroviral restriction by TRIM proteins, with a particular emphasis on T1, T18, T5α and TCyp. Although these factors have many distinct characteristics, they are nevertheless united by a few common themes.

One idea that has cropped up repeatedly throughout this work is the notion that TRIM proteins share a common overall design. This is epitomised by the ability of the T1 B30.2 domain to function adequately with components from both T5α and T18 (Chapter **3** and **Chapter 4**), and the generation of a restriction-competent chimera between rhesus T5α and owl monkey TCyp (Chapter **5**).

Nevertheless, each of these factors has undoubtedly been shaped by unique evolutionary forces, and this has resulted in certain distinct requirements, particularly in terms of how the modules are positioned relative to one another. This likely reflects the need for TRIM molecules to complement regularly spaced epitopes on the capsid surface, while simultaneously engaging components of the ubiquitination and innate immune signalling pathways. Each factor has accomplished this in a slightly different way; as a result, introducing foreign components without consideration of the surrounding context may not always garner favourable results.

Another issue arising from this work surrounds the definition of the term *restriction factor* itself. While T5α is a prototypic member of this class of proteins, the antiviral activity of T1 is rarely considered. Indeed, most of the recent literature concerning T1 (or *MID2* as it is more commonly known) focuses on its role in non-infectious diseases, such as X-linked developmental disorders and advanced-stage breast cancer (Li et al., 2016a; Wang et al., 2016). Thus, the restriction activity of T1 may be a phenotype of circumstance: a mere by-product of the subcellular localisation and domain complement of this protein. Perhaps it would then be inaccurate to label it as a restriction factor *per se*, but rather a protein with an existing cellular function that happens to also have restriction activity. Indeed, it could be argued that the likes of T28, SAMHD1 and SERINC3/5 also fall into this camp.

While this might seem like a semantic issue, these definitions could inform future searches for proteins with restriction activity – some of which may have clinical potential. Currently, restriction factor candidates are identified using reverse genetic techniques, often in the form of RNAi screens. However, such techniques may result in a number of candidates being overlooked, either because their knockdown has unrelated effects on cell viability, or because they are compensated for by more dedicated restriction factors. Widening our definition will pave the way for novel approaches in the search for host-derived factors with antiviral activity.

In many ways, the findings discussed in this work can be related to what is already known for the murine restriction factor, Fv1. Much like the TRIM family members, Fv1 comprises both capsid-binding and multimerisation domains, although the precise mechanism by which restriction is effected remains elusive. The modularity of Fv1 has previously been demonstrated by generating a C-terminal fusion with the owl monkey CypA domain; the resulting molecule acquired restriction activity against both HIV and FIV (Yap et al., 2007; Schaller et al., 2007).

This restriction factor 'reprogramming' is mirrored in the case of T5α, where specificity can often be exchanged between primate orthologues (Stremlau et al., 2005; Yap et al., 2005; Anderson and Akkina, 2008), and replacement of the B30.2 domain with CypA has yielded novel factors with potent antiviral activity on several occasions (Sayah et al., 2004; Dietrich et al., 2010). Similarly, the work presented in this thesis demonstrates that the T1 B30.2 domain can function adequately when paired with the N-terminus of either T5α or T18, and in fact restriction is even augmented in the latter case. Moreover, previous work in the Stoye lab has shown that both T1 and T18 are restriction-competent when appended with a CypA domain (Yap et al., 2006).

Collectively, these data reinforce the notion that capsid-binding restriction factors are comprised of interchangeable units that promote capsid association and multimerisation (Yap et al., 2007). This opens the door to generating artificial restriction factors with tailored specificities and enhanced effector functions – some of which may have therapeutic potential.

Chapter 7 Appendix

7.1 Primer directory

7.1.1 Primers used in Chapter 3

To make the long isoform of murine T1 (Section **3.3**)

Name	Template	Sequence (5'-3')
mT1L P1	mT1S	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCATGGAAAC ACTGGAGTCAG
mT1L P2	mT1S	TACTGCTTCCTTACATTTCCTTATCTCTGGCCACAGGCCC CAACTACACCATGACTTACTGATGAAGTTAGCCTGGC
mT1L P3	mT1S	AAGGAAATGTAAGGAAGCAGTAAGCTGCTCAAGATTGGCC GGTGCGCCACGAGGCAAGTACAATTCAGTGGATAGCTGG
mT1L P4	mT1S	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATGGCAAG CTTTCATCC

To make the long isoform of human T1 (Section **3.3**)

Name	Template	Sequence (5'-3')
huT1L P1	huT1S	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCATGGGTGA AAGCCCAGCCTC
huT1L P2	huT1S	CTGCTTCCTTACATTTCCTTATCTCTGGCCACAGGCCCCA ACTACACCATGACTTACTGATGAAGTTAGCCTGGCC
huT1L P3	huT1S	AAGGAAATGTAAGGAAGCAGTAAGCTGCTCAAGATTGGCC GGGGCGCCACGAGGCCTGTATAATTCAGTAGACAG
huT1L P4	huT1S	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATGACAGG TTTTCATCC

To exchange the B30.2 domains of T1 and 18 (Section **3.4**)

Name	Template	Sequence (5'-3')
1-18$_{B30.2}$ P1	agmT1S	GGGGACAAGTTTGTACAAAAAAGCAGGCTCACCATGGAAA CACTGGAGTCTG
1-18$_{B30.2}$ P2	agmT1S	GTTGGCTGTTTGTTTTTAGTCGGGTAGGTTC
1-18$_{B30.2}$ P3	agmT18	ACTAAAAACAAACAGCCAACCATTTAAACTG
1-18$_{B30.2}$ P4	agmT18	GGGGACCACTTTGTACAAGAAAGCTGGGTTCACGGCAGCT GCTCTGTGCAGTCC
18-1$_{B30.2}$ P1	agmT18	GGGGACAAGTTTGTACAAAAAAGCAGGCTCACCATGGAAA CACTGGAGTCAG

18-1$_{B30.2}$ P2	agmT18	GTTGGCTGTTTGTCTTCAACTTCCCAGGCTC
18-1$_{B30.2}$ P3	agmT1S	GTTGAAGACAAACAGCCAACCCTTTAAATTG
18-1$_{B30.2}$ P4	agmT1S	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATGACAGG TTTTCATCCCAG

To exchange the variable regions of T1 and 18 (Section **3.4**)

Name	Template	Sequence (5'-3')
T1 VR1 FWD	1-18$_{B30.2}$*	CTTGACAGTAGAAAAGGATGAAAGCTCTCTCTAAAGAAGA GCCATACCCCAGAGAGGTTTAGTGGCACAGGGTGCTATGG GGTAGCTGGAAATG
T1 VR1 REV	1-18$_{B30.2}$*	CATTTCCAGCTACCCCATAGCACCCTGTGCCACTAAACCT CTCTGGGGTATGGCTCTTCTTTAGAGAGCTTTCATCCTTT TCTACTGTCAAG
T1 VR2 FWD	1-18$_{B30.2}$*	CTATTGGTCTTGCTTACAAATCAGCTCCAAAGAATGAATG GATTGGCAAGAACTCTGCTTC
T1 VR2 REV	1-18$_{B30.2}$*	GAAGCAGAGTTCTTGCCAATCCATTCATTCTTTGGAGCTG ATTTGTAAGCAAGACCAATAG
T1 VR3 FWD	1-18$_{B30.2}$*	GCGCTCTGCCGCTGCAATAGTAACTTCGTGGTGAGACATA ACAACAAGGAAATGCTGGTGGATGTGCCCCCACAGCTCCG GCGCGTG
T1 VR3 REV	1-18$_{B30.2}$*	CACGCGCCGGAGCTGTGGGGGCACATCCACCAGCATTTCC TTGTTGTTATGTCTCACCACGAAGTTACTATTGCAGCGGC AGAGCGC
T1 VR4 FWD	1-18$_{B30.2}$*	CCCACCTTCACGGTGTGGAACAAATCCCTAATGATCTTGT CTGGCTTGCCTATCCCAGACCATTTG
T1 VR4 REV	1-18$_{B30.2}$*	CAAATGGTCTGGGATAGGCAAGCCAGACAAGATCATTAGG GATTTGTTCCACACCGTGAAGGTGGG
T18 VR1 FWD	18-1$_{B30.2}$**	GATTGCAGATGGAGCGTGATGAATCATCATCTAAGAAGAG TCACACGCCTGAACGCTTCACCAGCCAGGGGAGCTATGGA GCAGCAGGAAATATATTC
T18 VR1 REV	18-1$_{B30.2}$**	GAATATATTTCCTGCTGCTCCATAGCTCCCCTGGCTGGTG AAGCGTTCAGGCGTGTGACTCTTCTTAGATGATGATTCAT CACGCTCCATCTGCAATC
T18 VR2 FWD	18-1$_{B30.2}$**	CAATTGGCATTGCCTATAAATCAGCCCCGAAGCATGAATG GATTGGGAAGAATGCCTCC
T18 VR2 REV	18-1$_{B30.2}$**	GGAGGCATTCTTCCCAATCCATTCATGCTTCGGGGCTGAT TTATAGGCAATGCCAATTG
T18 VR3 FWD	18-1$_{B30.2}$**	GTCTTCTCTCGCTGCAACAATAACTGGGTGGTGAGACACA ATAGCAAGGAAATCCCCATTGAGCCTGCTCCCCACCTGAA GCGTCTG
T18 VR3 REV	18-1$_{B30.2}$**	CAGACGCTTCAGGTGGGGAGCAGGCTCAATGGGGATTTCC TTGCTATTGTGTCTCACCACCCAGTTATTGTTGCAGCGAG AGAAGAC

T18 VR4 FWD	18-$1_{B30.2}$**	CCAACATTTACAATCTGGAACAAGTGTCTGACGATTATCACTGGGCTTCCTGCCCCAGATTTTATTG
T18 VR4 REV	18-$1_{B30.2}$**	CAATAAAATCTGGGGCAGGAAGCCCAGTGATAATCGTCAGACACTTGTTCCAGATTGTAAATGTTGG

* *Construct **A**, Section **3.4***

** *Construct **S**, Section **3.4***

To remove (or add) the C-terminal tail (Section **3.4**)

Name	Template	Sequence (5'-3')
+CT FWD	agmT18	CAGAGCAGCTGCCGTGCAACTGCAGGCCTCAAGAATCCCCTTATGTTTCTGGGATGAAAACCTGTCATTGAACCCAGCTTTC
+CT REV	agmT18	GAAAGCTGGGTTCAATGACAGGTTTTCATCCCAGAAACATAAGGGGATTCTTGAGGCCTGCAGTTGCACGGCAGCTGCTCTG
ΔCT FWD	agmT1S, 18-$1_{B30.2}$*	CCTGAGCGGCAGGAATAAACCCAGCTTTC
ΔCT REV	agmT1S, 18-$1_{B30.2}$*	GAAAGCTGGGTTTATTCCTGCCGCTCAGG

* *Construct **S**, Section **3.4***

To substitute position 595 in T1L (or 565 in T18) (Section **3.5**)

Name	Template	Sequence (5'-3')
T1L N595*n* FWD	agmT1L, agmT1S	CAAATCAGCTCCAAAG(NNN)GAATGGATTGGCAAG
T1L N595*n* REV	agmT1L, agmT1S	CTTGCCAATCCATTC(NNN)CTTTGGAGCTGATTTG
T18 H565N FWD	agmT18	CTTATAAATCAGCCCCGAAGAATGAATGGATTGGGAAGAAC
T18 H565N REV	agmT18	GTTCTTCCCAATCCATTCATTCTTCGGGGCTGATTTATAAG

To substitute the T18 RBCC domains for those of T1 (Section **3.6**)

Name	Template	Sequence (5'-3')
18-1$_{314}$ R1 P1	18-1$_{314}$*	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCATGGAAAC ACTGGAGTCAGAACTGACCTGTCCAATCTGCCTAGAG
18-1$_{314}$ R1 P2	agmT1S	TGATGACATGCCTGCAGGTAGGACACTG
18-1$_{314}$ R1 P3	agmT1S	TACCTGCAGGCATGTCATCACCCTCAGC
18-1$_{314}$ R1 P4	18-1$_{314}$*	TTAATGACAGGTTTTCATCCC
18-1$_{314}$ B1 P1	18-1$_{314}$*	CACCATGGAAACACTGGAGTCAG
18-1$_{314}$ B1 P2	18-1$_{314}$*	CAATTCGCTCGGCGGAGGTCATGGTGTTGG
18-1$_{314}$ B1 P3	agmT1S	GACCTCCGCCGAGCGAATTGCTTGCCAATTC
18-1$_{314}$ B1 P4	agmT1S	GATGATCGCGGTGACGACCCACCAGTTTGC
18-1$_{314}$ B1 P5	18-1$_{314}$*	GGGTCGTCACCGCGATCATCAGGTGGCAGC
18-1$_{314}$ B1 P6	18-1$_{314}$*	TTAATGACAGGTTTTCATCCC
18-1$_{314}$ CC1 P1	18-1$_{314}$*	CACCATGGAAACACTGGAGTCAG
18-1$_{314}$ CC1 P2	18-1$_{314}$*	GATGGTCTCGGTGCCGCCCAACCAGTTTAC
18-1$_{314}$ CC1 P3	agmT1S	GTTGGGCGGCACCGAGACCATCAGGTCGCATC
18-1$_{314}$ CC1 P4	agmT1S	CATCACCTTCCCCTCTTTGATTTTGACAGCG
18-1$_{314}$ CC1 P5	18-1$_{314}$*	CAAAATCAAAGAGGGGAAGGTGATGAGGCTTCGC
18-1$_{314}$ CC1 P6	18-1$_{314}$*	TTAATGACAGGTTTTCATCCC

** Construct **A**, Section **3.6***

To put the T18 B-boxes into T1 (Section **3.6**)

Name	Template	Sequence (5'-3')
T1-B18 P1	agmT1L, agmT1S	CACCATGGAAACACTGGAGTCTG
T1-B18 P2	agmT1L, agmT1S	GGACCTTCTCACTAGACATGGCGGTGTTGG
T1-B18 P3	agmT18	CATGTCTAGTGAGAAGGTCCTCTGCCAG
T1-B18 P4	agmT18	GATGGTCTCGGTGCCGCCCAACCAGTTTAC
T1-B18 P5	agmT1L, agmT1S	TGGGCGGCACCGAGACCATCAGGTCGCATC
T1-B18 P6	agmT1L, agmT1S	TTAATGACAGGTTTTCATCC

To add C-terminal HA tags to T1L, T1S and 18-1_{314} (Section **3.7**)

Name	Template	Sequence (5'-3')
T1-HA FWD	agmT1L, agmT1S	CACCATGGAAACACTGGAGTCTG
18-1_{314}-HA FWD	18-1_{314}*	CACCATGGAAACACTGGAGTCAG
HA REV	agmT1L, agmT1S, 18-1_{314}*	CTAAGCGTAATCTGGAACATCGTATGGGTAATGACAGGTTTTCATC

** Construct **A**, Section **3.6***

7.1.2 Primers used in Chapter 4

To make reciprocal chimeras of agmT1S and rhT5α (Section **4.1**)

Name	Template	Sequence (5'-3')
1-5$_{CC}$ P1	agmT1S	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGAAACACT GGAGTCTG
1-5$_{CC}$ P2	agmT1S	GATGCTCCAGCTCTTTGATTTTGACAGCG
1-5$_{CC}$ P3	rhT5α	AATCAAAGAGCTGGAGCATCGGTTGCAGGG
1-5$_{CC}$ P4	rhT5α	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAAGAGCTTG GTGAGCAC
5-1$_{CC}$ P1	rhT5α	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGCTTCTGG AATCCTGC
5-1$_{CC}$ P2	rhT5α	TAACCTTTGTTTCTGAGATGAGCTCTCTC
5-1$_{CC}$ P3	agmT1S	CATCTCAGAAACAAAGGTTATGAAACTGAG
5-1$_{CC}$ P4	agmT1S	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATGACAGG TTTTCATCC
1-5$_{B30.2}$ P1	agmT1S	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGAAACACT GGAGTCTG
1-5$_{B30.2}$ P2	agmT1S	TCACATCAACTGTTTTTAGTCGGGTAGGTTC
1-5$_{B30.2}$ P3	rhT5α	ACTAAAAACAGTTGATGTGACACTGGCTAC
1-5$_{B30.2}$ P4	rhT5α	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAAGAGCTTG GTGAGCAC
5-1$_{B30.2}$ P1	rhT5α	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGCTTCTGG AATCCTGC
5-1$_{B30.2}$ P2	rhT5α	GTTGGCTGTTCCAGTAGCGTCGGGCATCTG
5-1$_{B30.2}$ P3	agmT1S	ACGCTACTGGAACAGCCAACCCTTTAAATTGG
5-1$_{B30.2}$ P4	agmT1S	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATGACAGG TTTTCATCC

To amplify the B30.2 domain for expression in *E. coli* (Section **4.3**)

Name	Template	Sequence (5'-3')
MBP-B30.2 FWD	agmT1S	CGTAGAAAGCTTACCCGACTAAAAACAAACAGCC
MBP-B30.2 REV	agmT1S	GTTCGACTCGAGTTAATGACAGGTTTTCATCCC

7.1.3 Primers used in Chapter 5

To make deletions in rhT5α (Section **5.1**)

Name	Template	Sequence (5'-3')
Δ 259-290 FWD	rhT5α	GGCATCATTAAAAGGAGAGAGCTAACAG
Δ 259-290 REV	rhT5α	CTGTTAGCTCTCTCCTTTTAATGATGCC
Δ 263-290 FWD	rhT5α	GGATTGAGAACATGAGAGAGCTAACAG
Δ 263-290 REV	rhT5α	CTGTTAGCTCTCTCATGTTCTCAATCC
Δ 272-290 FWD	rhT5α	CCAAAAACTTTTCACAGAGAGCTAACAG
Δ 272-290 REV	rhT5α	CTGTTAGCTCTCTGTGAAAAGTTTTTGG
Δ 282-290 FWD	rhT5α	GTGTTTCGAGCTCCTAGAGAGCTAACAG
Δ 282-290 REV	rhT5α	CTGTTAGCTCTCTAGGAGCTCGAAACAC
Δ 282-285 FWD	rhT5α	GTGTTTCGAGCTCCTATGCTAGACATG
Δ 282-285 REV	rhT5α	CATGTCTAGCATAGGAGCTCGAAACAC
Δ 286-290 FWD	rhT5α	CCTGATCTGAAAGGAAGAGAGCTAACAG
Δ 286-290 REV	rhT5α	CTGTTAGCTCTCTTCCTTTCAGATCAGG
Δ 291-295 FWD	rhT5α	GCTAGACATGTTTGCCCGACGCTACTGG
Δ 291-295 REV	rhT5α	CCAGTAGCGTCGGGCAAACATGTCTAGC
Δ 263-280 FWD	rhT5α	GGATTGAGAACATGCCTGATCTGAAAGG
Δ 263-280 REV	rhT5α	CCTTTCAGATCAGGCATGTTCTCAATCC
Δ 272-280 FWD	rhT5α	CCAAAAACTTTTCACCCTGATCTGAAAGG
Δ 272-280 REV	rhT5α	CCTTTCAGATCAGGGTGAAAAGTTTTTGG

To make insertions in rhT5α (Section **5.2**)

Name	Template	Sequence (5'-3')
FRE FWD	rhT5α	GGAATGCTAGACATGTTTAGAGAGTTTAGAGAGCTAAC
FRE REV	rhT5α	GTTAGCTCTCTAAACTCTCTAAACATGTCTAGCATTCC
FREL FWD	rhT5α	GGAATGCTAGACATGTTTAGAGAGCTATTTAGAGAGCTAAC
FREL REV	rhT5α	GTTAGCTCTCTAAATAGCTCTCTAAACATGTCTAGCATTCC
FRELFRL FWD	rhT5α	GGAATGCTAGACATGTTTAGAGAGCTATTTAGACTATTTAGAGAGCTAAC
FRELFRL REV	rhT5α	GTTAGCTCTCTAAATAGTCTAAATAGCTCTCTAAACATGTCTAGCATTCC

To disrupt helix α5 in T5α (Section **5.3**)

Name	Template	Sequence (5'-3')
rhT5α L287P FWD	rhT5α	GATCTGAAAGGAATGCCAGACATGTTTAGAGAG
rhT5α L287P REV	rhT5α	CTCTCTAAACATGTCTGGCATTCCTTTCAGATC
rhT5α L293P FWD	rhT5α	GACATGTTTAGAGAGCCAACAGATGCCCGACGC
rhT5α L293P REV	rhT5α	GCGTCGGGCATCTGTTGGCTCTCTAAACATGTC
huT5α L285P FWD	huT5α	GATCTGAAAGGAATGCCAGAAGTGTTTAGAGAG
huT5α L285P REV	huT5α	CTCTCTAAACACTTCTGGCATTCCTTTCAGATC
huT5α L291P FWD	huT5α	GAAGTGTTTAGAGAGCCGACAGATGTCCGACGC
huT5α L291P REV	huT5α	GCGTCGGACATCTGTCGGCTCTCTAAACACTTC

To remove exon 7 (Section **5.5**)

Name	Template	Sequence (5'-3')
omTCyp Δe7 FWD	omTCyp	CTACAAGTGTTTAAAGACGCCGCCGCCTGG
omTCyp Δe7 REV	omTCyp	CCAGGCGGCGGCGTCTTTAAACACTTGTAG
rhT5α Δe7 FWD	rhT5α	GACATGTTTAGAGTTGATGTGACACTGG
rhT5α Δe7 REV	rhT5α	CCAGTGTCACATCAACTCTAAACATGTC

To mutate residues in the active site of CypA (Section **5.5**)

Name	Template	Sequence (5'-3')
CypA D66N FWD	T5αCyp*	GTGTCAGGGTGGTAACTTCACACGCCATAATG
CypA D66N REV	T5αCyp*	CATTATGGCGTGTGAAGTTACCACCCTGACAC
CypA R69H FWD	T5αCyp*	GGTGGTGACTTCACACACCATAATGGCACTGG
CypA R69H REV	T5αCyp*	CCAGTGCCATTATGGTGTGTGAAGTCACCACC

** An artificial chimera consisting of rhT5α fused to CypA from owl monkey TCyp (see Section **5.4**)*

To put the rhesus leader sequence into owl monkey CypA (Section **5.5**)

Name	Template	Sequence (5'-3')
Rhesus leader FWD	T5αCyp*	GACGCCGCCGCCGAAGAATCACCAGTACTTCTTGCCATGGTCAATCC
Rhesus leader REV	T5αCyp*	GGATTGACCATGGCAAGAAGTACTGGTGATTCTTCGGCGGCGGCGTC

** An artificial chimera consisting of rhT5α fused to CypA from owl monkey TCyp (see Section **5.4**)*

7.2 Screens used for crystallisation trials

Below is a list of screens used for crystallisation trials of the T1 B30.2 domain (Section **4.3.4**).

Screen	Company	External link
Anions	Qiagen	https://www.qiagen.com/products/protein/crystallization/compositiontables/pdf/1054288_ps_xtal_anions-suite.pdf
Classics	Qiagen	https://www.qiagen.com/products/protein/crystallization/compositiontables/pdf/1054292_ps_xtal_classics-suite.pdf
JCSG Core I	Qiagen	https://www.cicbiogune.es/services/userfiles/JCSG-Core-I-Suite.pdf
JCSG Core II	Qiagen	https://www.cicbiogune.es/services/userfiles/JCSG-Core-II-Suite.pdf
JCSG Core III	Qiagen	https://www.cicbiogune.es/services/userfiles/JCSG-Core-III-Suite.pdf
JCSG Core IV	Qiagen	http://services.mbi.ucla.edu/Crystallization/xtalscreens/docs/JCSGcoreIV_comp.pdf
PACT	Qiagen	https://www.qiagen.com/products/protein/crystallization/compositiontables/pdf/1057791_ps_xtal_pact-suite_update_160609_lowres.pdf
Proplex	Molecular dimensions	http://www.moleculardimensions.com/applications/upload/MD1-42%20ProPlex%20HT-96.pdf
Wizard I	Rigaku	http://www.rigakureagents.com/UploadDocuments/wiz1_rigaku_15.pdf
Wizard II	Rigaku	http://www.rigakureagents.com/UploadDocuments/wiz2_rigaku_15.pdf

References

ABDEL-MOHSEN, M., DENG, X., LIEGLER, T., GUATELLI, J. C., SALAMA, M. S., GHANEM, H. E.-D. A., RAUCH, A., LEDERGERBER, B., DEEKS, S. G., GÜNTHARD, H. F., WONG, J. K. & PILLAI, S. K. 2014. Effects of Alpha Interferon Treatment on Intrinsic Anti-HIV-1 Immunity In Vivo. *Journal of Virology,* 88**,** 763-767.

ACCOLA, M. A., HÖGLUND, S. & GÖTTLINGER, H. G. 1998. A Putative α-Helical Structure Which Overlaps the Capsid-p2 Boundary in the Human Immunodeficiency Virus Type 1 Gag Precursor Is Crucial for Viral Particle Assembly. *Journal of Virology,* 72**,** 2072-2078.

ACHONG, B. G., MANSELL, P. W. A., EPSTEIN, M. A. & CLIFFORD, P. 1971. An Unusual Virus in Cultures From a Human Nasopharyngeal Carcinoma. *Journal of the National Cancer Institute,* 46**,** 299-307.

AIKEN, C. 2006. Viral and cellular factors that regulate HIV-1 uncoating. *Current Opinion in HIV and AIDS,* 1**,** 194-199.

ALBRITTON, L. M., TSENG, L., SCADDEN, D. & CUNNINGHAM, J. M. 1989. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell,* 57**,** 659-666.

AMARA, A. & LITTMAN, D. R. 2003. After Hrs with HIV. *The Journal of Cell Biology,* 162**,** 371-375.

AMIE, S. M., BAMBARA, R. A. & KIM, B. 2013. GTP Is the Primary Activator of the Anti-HIV Restriction Factor SAMHD1. *Journal of Biological Chemistry,* 288**,** 25001-25006.

ANDERSON, J. & AKKINA, R. 2008. Human Immunodeficiency Virus Type 1 Restriction by Human–Rhesus Chimeric Tripartite Motif 5α (TRIM 5α) in CD34+ Cell-Derived Macrophages In Vitro and in T Cells In Vivo in Severe Combined Immunodeficient (SCID-hu) Mice Transplanted with Human Fetal Tissue. *Human Gene Therapy,* 19**,** 217-228.

APOLONIA, L., SCHULZ, R., CURK, T., ROCHA, P., SWANSON, C. M., SCHALLER, T., ULE, J. & MALIM, M. H. 2015. Promiscuous RNA Binding Ensures Effective Encapsidation of APOBEC3 Proteins by HIV-1. *PLoS Pathogens,* 11**,** e1004609.

ARVIDSON, B., SEEDS, J., WEBB, M., FINLAY, L. & BARKLIS, E. 2003. Analysis of the retrovirus capsid interdomain linker region. *Virology,* 308**,** 166-177.

AUERBACH, M. R., BROWN, K. R., KAPLAN, A., DE LAS NUECES, D. & SINGH, I. R. 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. & SINGH, I. R. 2007. Mutational Analysis of the N-Terminal Domain of Moloney Murine Leukemia Virus Capsid Protein. *Journal of Virology,* 81**,** 12337-12347.

BABST, M., WENDLAND, B., ESTEPA, E. J. & EMR, S. D. 1998. The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. *The EMBO Journal,* 17**,** 2982-2993.

BALDAUF, H.-M., PAN, X., ERIKSON, E., SCHMIDT, S., DADDACHA, W., BURGGRAF, M., SCHENKOVA, K., AMBIEL, I., WABNITZ, G., GRAMBERG, T., PANITZ, S., FLORY, E., LANDAU, N. R., SERTEL, S., RUTSCH, F., LASITSCHKA, F., KIM, B., KÖNIG, R., FACKLER, O. T. & KEPPLER, O. T. 2012. SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nature medicine,* 18**,** 10.1038/nm.2964.

BALLANDRAS-COLAS, A., BROWN, M., COOK, N. J., DEWDNEY, T. G., DEMELER, B., CHEREPANOV, P., LYUMKIS, D. & ENGELMAN, A. N. 2016. Cryo-EM

reveals a novel octameric integrase structure for betaretroviral intasome function. *Nature,* 530**,** 358-361.

BAUNACH, G., MAURER, B., HAHN, H., KRANZ, M. & RETHWILM, A. 1993. Functional analysis of human foamy virus accessory reading frames. *Journal of Virology,* 67**,** 5411-5418.

BELL, J. L., MALYUKOVA, A., KAVALLARIS, M., MARSHALL, G. M. & CHEUNG, B. B. 2013. TRIM16 inhibits neuroblastoma cell proliferation through cell cycle regulation and dynamic nuclear localization. *Cell Cycle,* 12**,** 889-98.

BENJAMIN, J., GANSER-PORNILLOS, B. K., TIVOL, W. F., SUNDQUIST, W. I. & JENSEN, G. J. 2005. Three-dimensional Structure of HIV-1 Virus-like Particles by Electron Cryotomography. *Journal of Molecular Biology,* 346**,** 577-588.

BERGER, E. A., DOMS, R. W., FENYO, E. M., KORBER, B. T. M., LITTMAN, D. R., MOORE, J. P., SATTENTAU, Q. J., SCHUITEMAKER, H., SODROSKI, J. & WEISS, R. A. 1998. A new classification for HIV-1. *Nature,* 391**,** 240-240.

BERKOWITZ, R. D., OHAGEN, A., HOGLUND, S. & 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**,** 6445-56.

BERNARDI, R. & PANDOLFI, P. P. 2003. Role of PML and the PML-nuclear body in the control of programmed cell death. *Oncogene,* 22**,** 9048-57.

BERTHOUX, L., SEBASTIAN, S., SAYAH, D. M. & LUBAN, J. 2005. Disruption of human TRIM5alpha antiviral activity by nonhuman primate orthologues. *J Virol,* 79**,** 7883-8.

BERTI, C., FONTANELLA, B., FERRENTINO, R. & MERONI, G. 2004. Mig12, a novel Opitz syndrome gene product partner, is expressed in the embryonic ventral midline and co-operates with Mid1 to bundle and stabilize microtubules. *BMC Cell Biology,* 5**,** 9-9.

BESNIER, C., TAKEUCHI, Y. & TOWERS, G. 2002. Restriction of lentivirus in monkeys. *Proc Natl Acad Sci U S A,* 99**,** 11920-5.

BEST, S., LE TISSIER, P., TOWERS, G. & STOYE, J. P. 1996. Positional cloning of the mouse retrovirus restriction gene Fv1. *Nature,* 382**,** 826-9.

BHARAT, T. A. M., DAVEY, N. E., ULBRICH, P., RICHES, J. D., DE MARCO, A., RUMLOVA, M., SACHSE, C., RUML, T. & BRIGGS, J. A. G. 2012. Structure of the immature retroviral capsid at 8[thinsp]A resolution by cryo-electron microscopy. *Nature,* 487**,** 385-389.

BHATTACHARYA, A., ALAM, S. L., FRICKE, T., ZADROZNY, K., SEDZICKI, J., TAYLOR, A. B., DEMELER, B., PORNILLOS, O., GANSER-PORNILLOS, B. K., DIAZ-GRIFFERO, F., IVANOV, D. N. & YEAGER, M. 2014. Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 18625-18630.

BICHEL, K., PRICE, A. J., SCHALLER, T., TOWERS, G. J., FREUND, S. M. V. & JAMES, L. C. 2013. HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. *Retrovirology,* 10**,** 81-81.

BIRIS, N., TOMASHEVSKI, A., BHATTACHARYA, A., DIAZ-GRIFFERO, F. & IVANOV, D. N. 2013. Rhesus monkey TRIM5alpha SPRY domain recognizes multiple epitopes that span several capsid monomers on the surface of the HIV-1 mature viral core. *J Mol Biol,* 425**,** 5032-44.

BIRIS, N., YANG, Y., TAYLOR, A. B., TOMASHEVSKI, A., GUO, M., HART, P. J., DIAZ-GRIFFERO, F. & IVANOV, D. N. 2012. Structure of the rhesus monkey TRIM5α PRYSPRY domain, the HIV capsid recognition module. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 13278-13283.

BISHOP, K. N., BOCK, M., TOWERS, G. & STOYE, J. P. 2001. Identification of the regions of Fv1 necessary for murine leukemia virus restriction. *J Virol,* 75**,** 5182-8.

BISHOP, K. N., MORTUZA, G. B., HOWELL, S., YAP, M. W., STOYE, J. P. & TAYLOR, I. A. 2006. Characterization of an amino-terminal dimerization domain from retroviral restriction factor Fv1. *J Virol,* 80**,** 8225-35.

BISHOP, K. N., VERMA, M., KIM, E.-Y., WOLINSKY, S. M. & MALIM, M. H. 2008. APOBEC3G Inhibits Elongation of HIV-1 Reverse Transcripts. *PLoS Pathogens,* 4**,** e1000231.

BLAIR, W. S., BOGERD, H. & CULLEN, B. R. 1994. Genetic analysis indicates that the human foamy virus Bel-1 protein contains a transcription activation domain of the acidic class. *Journal of Virology,* 68**,** 3803-3808.

BOCK, M., BISHOP, K. N., TOWERS, G. & STOYE, J. P. 2000. Use of a transient assay for studying the genetic determinants of Fv1 restriction. *J Virol,* 74**,** 7422-30.

BOCK, M., HEINKELEIN, M., LINDEMANN, D. & RETHWILM, A. 1998. Cells Expressing the Human Foamy Virus (HFV) Accessory Bet Protein Are Resistant to Productive HFV Superinfection. *Virology,* 250**,** 194-204.

BOI, S., KOLOKITHAS, A., SHEPARD, J., LINWOOD, R., ROSENKE, K., VAN DIS, E., MALIK, F. & EVANS, L. H. 2014. Incorporation of Mouse APOBEC3 into Murine Leukemia Virus Virions Decreases the Activity and Fidelity of Reverse Transcriptase. *Journal of Virology,* 88**,** 7659-7662.

BORDEN, K. L. & FREEMONT, P. S. 1996. The RING finger domain: a recent example of a sequence-structure family. *Curr Opin Struct Biol,* 6**,** 395-401.

BRASS, A. L., HUANG, I. C., BENITA, Y., JOHN, S. P., KRISHNAN, M. N., FEELEY, E. M., RYAN, B., WEYER, J. L., VAN DER WEYDEN, L., FIKRIG, E., ADAMS, D. J., XAVIER, R. J., FARZAN, M. & ELLEDGE, S. J. 2009. IFITM Proteins Mediate the Innate Immune Response to Influenza A H1N1 Virus, West Nile Virus and Dengue Virus. *Cell,* 139**,** 1243-1254.

BRENNAN, G., KOZYREV, Y. & HU, S. L. 2008. TRIMCyp expression in Old World primates Macaca nemestrina and Macaca fascicularis. *Proc Natl Acad Sci U S A,* 105**,** 3569-74.

BRIGGS, J. A. G., SIMON, M. N., GROSS, I., KRAUSSLICH, H.-G., FULLER, S. D., VOGT, V. M. & JOHNSON, M. C. 2004. The stoichiometry of Gag protein in HIV-1. *Nat Struct Mol Biol,* 11**,** 672-675.

BRIGGS, J. A. G., WILK, T. & FULLER, S. D. 2003. Do lipid rafts mediate virus assembly and pseudotyping? *Journal of General Virology,* 84**,** 757-768.

BRYANT, M. & RATNER, L. 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 523-527.

BUCHNER, G., MONTINI, E., ANDOLFI, G., QUADERI, N., CAINARCA, S., MESSALI, S., BASSI, M. T., BALLABIO, A., MERONI, G. & FRANCO, B. 1999. MID2, a Homologue of the Opitz Syndrome Gene MID1: Similarities in Subcellular Localization and Differences in Expression During Development. *Human Molecular Genetics,* 8**,** 1397-1407.

BUSNADIEGO, I., KANE, M., RIHN, S. J., PREUGSCHAS, H. F., HUGHES, J., BLANCO-MELO, D., STROUVELLE, V. P., ZANG, T. M., WILLETT, B. J., BOUTELL, C., BIENIASZ, P. D. & WILSON, S. J. 2014. Host and Viral Determinants of Mx2 Antiretroviral Activity. *Journal of Virology,* 88**,** 7738-7752.

BYEON, I. J., MENG, X., JUNG, J., ZHAO, G., YANG, R., AHN, J., SHI, J., CONCEL, J., AIKEN, C., ZHANG, P. & GRONENBORN, A. M. 2009. Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell,* 139**,** 780-90.

CAINARCA, S., MESSALI, S., BALLABIO, A. & MERONI, G. 1999. Functional Characterization of the Opitz Syndrome Gene Product (Midin): Evidence for Homodimerization and Association With Microtubules Throughout the Cell Cycle. *Human Molecular Genetics,* 8**,** 1387-1396.

CAMMAS, F., KHETCHOUMIAN, K., CHAMBON, P. & LOSSON, R. 2012. TRIM involvement in transcriptional regulation. *Adv Exp Med Biol,* 770**,** 59-76.

CAMPBELL, E. M. & HOPE, T. J. 2005. Gene Therapy Progress and Prospects: Viral trafficking during infection. *Gene Ther,* 12**,** 1353-1359.

CAMPBELL, E. M. & HOPE, T. J. 2015. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Micro,* 13**,** 471-483.

CAMPBELL, E. M., PEREZ, O., ANDERSON, J. L. & HOPE, T. J. 2008. Visualization of a proteasome-independent intermediate during restriction of HIV-1 by rhesus TRIM5α. *The Journal of Cell Biology,* 180**,** 549-561.

CANDIDO, E. P., REEVES, R. & DAVIE, J. R. 1978. Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell,* 14**,** 105-13.

CARLSON, L.-A., BRIGGS, J. A. G., GLASS, B., RICHES, J. D., SIMON, M. N., JOHNSON, M. C., MÜLLER, B., GRÜNEWALD, K. & KRÄUSSLICH, H.-G. 2008. Three-Dimensional Analysis of Budding Sites and Released Virus Suggests a Revised Model for HIV-1 Morphogenesis. *Cell Host & Microbe,* 4**,** 592-599.

CARTEAU, S., GORELICK, R. J. & BUSHMAN, F. D. 1999. Coupled Integration of Human Immunodeficiency Virus Type 1 cDNA Ends by Purified Integrase In Vitro: Stimulation by the Viral Nucleocapsid Protein. *Journal of Virology,* 73**,** 6670-6679.

CARTELLIERI, M., RUDOLPH, W., HERCHENRÖDER, O., LINDEMANN, D. & RETHWILM, A. 2005. Determination of the relative amounts of Gag and Pol proteins in foamy virus particles. *Retrovirology,* 2**,** 1-7.

CARTHAGENA, L., BERGAMASCHI, A., LUNA, J. M., DAVID, A., UCHIL, P. D., MARGOTTIN-GOGUET, F., MOTHES, W., HAZAN, U., TRANSY, C., PANCINO, G. & NISOLE, S. 2009. Human TRIM Gene Expression in Response to Interferons. *PLoS One,* 4**,** e4894.

CHAN, D. C., FASS, D., BERGER, J. M. & KIM, P. S. 1997. Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell,* 89**,** 263-273.

CHAN, R., UCHIL, P. D., JIN, J., SHUI, G., OTT, D. E., MOTHES, W. & WENK, M. R. 2008. Retroviruses Human Immunodeficiency Virus and Murine Leukemia Virus Are Enriched in Phosphoinositides. *Journal of Virology,* 82**,** 11228-11238.

CHAREZA, S., SLAVKOVIC LUKIC, D., LIU, Y., RÄTHE, A.-M., MÜNK, C., ZABOGLI, E., PISTELLO, M. & LÖCHELT, M. 2012. Molecular and functional interactions of cat APOBEC3 and feline foamy and immunodeficiency virus proteins: Different ways to counteract host-encoded restriction. *Virology,* 424**,** 138-146.

CHAURASIYA, K. R., MCCAULEY, M. J., WANG, W., QUALLEY, D. F., WU, T., KITAMURA, S., GEERTSEMA, H., CHANDENISE, S. B., HERTZ, A., IWATANI, Y., LEVIN, J. G., MUSIER-FORSYTH, K., ROUZINA, I. & WILLIAMS, M. C. 2014. Oligomerization transforms human APOBEC3G from an efficient enzyme to a slowly dissociating nucleic acid-binding protein. *Nat Chem,* 6**,** 28-33.

CHECKLEY, M. A., LUTTGE, B. G. & FREED, E. O. 2011. HIV-1 Envelope Glycoprotein Biosynthesis, Trafficking, and Incorporation. *Journal of Molecular Biology,* 410**,** 582-608.

CHEN, B., VOGAN, E. M., GONG, H., SKEHEL, J. J., WILEY, D. C. & HARRISON, S. C. 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature,* 433**,** 834-841.

CHEN, M. Y., MALDARELLI, F., KARCZEWSKI, M. K., WILLEY, R. L. & STREBEL, K. 1993. Human immunodeficiency virus type 1 Vpu protein induces degradation

of CD4 in vitro: the cytoplasmic domain of CD4 contributes to Vpu sensitivity. *Journal of Virology,* 67**,** 3877-3884.

CHEN, N.-Y., ZHOU, L., GANE, P. J., OPP, S., BALL, N. J., NICASTRO, G., ZUFFEREY, M., BUFFONE, C., LUBAN, J., SELWOOD, D., DIAZ-GRIFFERO, F., TAYLOR, I. & FASSATI, A. 2016. HIV-1 capsid is involved in post-nuclear entry steps. *Retrovirology,* 13**,** 1-16.

CHIN, C. R., PERREIRA, J. M., SAVIDIS, G., PORTMANN, J. M., AKER, A. M., FEELEY, E. M., SMITH, M. C. & BRASS, A. L. 2015. Direct visualization of HIV-1 replication intermediates shows that capsid and CPSF6 modulate HIV-1 intra-nuclear invasion and integration. *Cell Rep,* 13.

CHOOK, Y. M. & SÜEL, K. E. 2011. Nuclear import by karyopherin-βs: Recognition and inhibition. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1813**,** 1593-1606.

CHUKKAPALLI, V., OH, S. J. & ONO, A. 2010. Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 1600-1605.

CLEVER, J. L., WONG, M. L. & PARSLOW, T. G. 1996. Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. *Journal of Virology,* 70**,** 5902-5908.

CONTE, M. R. & MATTHEWS, S. 1998. Retroviral Matrix Proteins: A Structural Perspective. *Virology,* 246**,** 191-198.

CONTICELLO, S. G., HARRIS, R. S. & NEUBERGER, M. S. 2003. The Vif Protein of HIV Triggers Degradation of the Human Antiretroviral DNA Deaminase APOBEC3G. *Current Biology,* 13**,** 2009-2013.

CONTICELLO, S. G., THOMAS, C. J. F., PETERSEN-MAHRT, S. K. & NEUBERGER, M. S. 2005. Evolution of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases. *Molecular Biology and Evolution,* 22**,** 367-377.

CORDONNIER, A., CASELLA, J. F. & HEIDMANN, T. 1995. Isolation of novel human endogenous retrovirus-like elements with foamy virus-related pol sequence. *Journal of Virology,* 69**,** 5890-7.

COSSON, P. 1996. Direct interaction between the envelope and matrix proteins of HIV-1. *The EMBO Journal,* 15**,** 5783-5788.

COWAN, S., HATZIIOANNOU, T., CUNNINGHAM, T., MUESING, M. A., GOTTLINGER, H. G. & 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**,** 11914-9.

COX, T. C., ALLEN, L. R., COX, L. L., HOPWOOD, B., GOODWIN, B., HAAN, E. & SUTHERS, G. K. 2000. New mutations in MID1 provide support for loss of function as the cause of X-linked Opitz syndrome. *Human Molecular Genetics,* 9**,** 2553-2562.

CRIBIER, A., DESCOURS, B., VALADÃO, ANA LUIZA C., LAGUETTE, N. & BENKIRANE, M. 2013. Phosphorylation of SAMHD1 by Cyclin A2/CDK1 Regulates Its Restriction Activity toward HIV-1. *Cell Reports,* 3**,** 1036-1043.

CRISTOFARI, G. & DARLIX, J.-L. 2002. The ubiquitous nature of RNA chaperone proteins. *Progress in Nucleic Acid Research and Molecular Biology.* Academic Press.

D'CRUZ, A. A., KERSHAW, N. J., CHIANG, J. J., WANG, M. K., NICOLA, N. A., BABON, J. J., GACK, M. U. & NICHOLSON, S. E. 2013. Crystal structure of the TRIM25 B30.2 (PRYSPRY) domain: a key component of antiviral signalling. *Biochem J,* 456**,** 231-40.

D'SOUZA, V., MELAMED, J., HABIB, D., PULLEN, K., WALLACE, K. & SUMMERS, M. F. 2001. Identification of a high affinity nucleocapsid protein binding element

within the moloney murine leukemia virus Ψ-RNA packaging signal: implications for genome recognition1. *Journal of Molecular Biology,* 314**,** 217-232.

DAR, A., WU, D., LEE, N., SHIBATA, E. & DUTTA, A. 2014. 14-3-3 Proteins Play a Role in the Cell Cycle by Shielding Cdt2 from Ubiquitin-Mediated Degradation. *Molecular and Cellular Biology,* 34**,** 4049-4061.

DARLIX, J.-L., GABUS, C., NUGEYRE, M.-T., CLAVEL, F. & BARRÉ-SINOUSSI, F. 1990. Cis elements and Trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *Journal of Molecular Biology,* 216**,** 689-699.

DARLIX, J.-L., LAPADAT-TAPOLSKY, M., DE ROCQUIGNY, H. & ROQUES, B. P. 1995. First Glimpses at Structure-function Relationships of the Nucleocapsid Protein of Retroviruses. *Journal of Molecular Biology,* 254**,** 523-537.

DAS, K., BAUMAN, J. D., CLARK, A. D., JR., FRENKEL, Y. V., LEWI, P. J., SHATKIN, A. J., HUGHES, S. H. & ARNOLD, E. 2008. High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. *Proc Natl Acad Sci U S A,* 105**,** 1466-71.

DE IACO, A., SANTONI, F., VANNIER, A., GUIPPONI, M., ANTONARAKIS, S. & LUBAN, J. 2013. TNPO3 protects HIV-1 replication from CPSF6-mediated capsid stabilization in the host cell cytoplasm. *Retrovirology,* 10**,** 20-20.

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. & BRIGGS, J. A. 2010a. Conserved and variable features of Gag structure and arrangement in immature retrovirus particles. *J Virol,* 84**,** 11729-36.

DE MARCO, A., HEUSER, A.-M., GLASS, B., KRÄUSSLICH, H.-G., MÜLLER, B. & BRIGGS, J. A. G. 2012. Role of the SP2 Domain and Its Proteolytic Cleavage in HIV-1 Structural Maturation and Infectivity. *Journal of Virology,* 86**,** 13708-13716.

DE MARCO, A., MÜLLER, B., GLASS, B., RICHES, J. D., KRÄUSSLICH, H.-G. & BRIGGS, J. A. G. 2010b. Structural Analysis of HIV-1 Maturation Using Cryo-Electron Tomography. *PLoS Pathog,* 6**,** e1001215.

DE RIJCK, J., DE KOGEL, C., DEMEULEMEESTER, J., VETS, S., ASHKAR, S. E., MALANI, N., BUSHMAN, F. D., LANDUYT, B., HUSSON, S. J., BUSSCHOTS, K., GIJSBERS, R. & DEBYSER, Z. 2013. The BET family of proteins targets Moloney Murine Leukemia Virus integration near transcription start sites. *Cell Reports,* 5**,** 886-894.

DELEBECQUE, F., SUSPÈNE, R., CALATTINI, S., CASARTELLI, N., SAÏB, A., FROMENT, A., WAIN-HOBSON, S., GESSAIN, A., VARTANIAN, J.-P. & SCHWARTZ, O. 2006. Restriction of Foamy Viruses by APOBEC Cytidine Deaminases. *Journal of Virology,* 80**,** 605-614.

DESHMUKH, L., SCHWIETERS, C. D., GRISHAEV, A., GHIRLANDO, R., BABER, J. L. & CLORE, G. M. 2013. Structure and Dynamics of Full Length HIV-1 Capsid Protein in Solution. *Journal of the American Chemical Society,* 135**,** 16133-16147.

DHARAN, A., TALLEY, S., TRIPATHI, A., MAMEDE, J. I., MAJETSCHAK, M., HOPE, T. J. & CAMPBELL, E. M. 2016. KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. *PLoS Pathog,* 12**,** e1005700.

DI NUNZIO, F., FRICKE, T., MICCIO, A., VALLE-CASUSO, J. C., PEREZ, P., SOUQUE, P., RIZZI, E., SEVERGNINI, M., MAVILIO, F., CHARNEAU, P. & DIAZ-GRIFFERO, F. 2013. Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. *Virology,* 440**,** 10.1016/j.virol.2013.02.008.

DIAZ-GRIFFERO, F., KAR, A., LEE, M., STREMLAU, M., POESCHLA, E. & SODROSKI, J. 2007. Comparative requirements for the restriction of retrovirus infection by TRIM5alpha and TRIMCyp. *Virology,* 369.

DIAZ-GRIFFERO, F., LI, X., JAVANBAKHT, H., SONG, B., WELIKALA, S., STREMLAU, M. & SODROSKI, J. 2006. Rapid turnover and polyubiquitylation of the retroviral restriction factor TRIM5. *Virology,* 349**,** 300-15.

DIAZ-GRIFFERO, F., QIN, X.-R., HAYASHI, F., KIGAWA, T., FINZI, A., SARNAK, Z., LIENLAF, M., YOKOYAMA, S. & SODROSKI, J. 2009. A B-Box 2 Surface Patch Important for TRIM5α Self-Association, Capsid Binding Avidity, and Retrovirus Restriction. *Journal of Virology,* 83**,** 10737-10751.

DIETRICH, E. A., JONES-ENGEL, L. & HU, S. L. 2010. Evolution of the antiretroviral restriction factor TRIMCyp in Old World primates. *PLoS One,* 5**,** 0014019.

DO KWON, Y., PANCERA, M., ACHARYA, P., GEORGIEV, I. S., CROOKS, E. T., GORMAN, J., JOYCE, M. G., GUTTMAN, M., MA, X., NARPALA, S., SOTO, C., TERRY, D. S., YANG, Y., ZHOU, T., AHLSEN, G., BAILER, R. T., CHAMBERS, M., CHUANG, G.-Y., DORIA-ROSE, N. A., DRUZ, A., HALLEN, M. A., HARNED, A., KIRYS, T., LOUDER, M. K., O'DELL, S., OFEK, G., OSAWA, K., PRABHAKARAN, M., SASTRY, M., STEWART-JONES, G. B. E., STUCKEY, J., THOMAS, P. V., TITTLEY, T., WILLIAMS, C., ZHANG, B., ZHAO, H., ZHOU, Z., DONALD, B. R., LEE, L. K., ZOLLA-PAZNER, S., BAXA, U., SCHON, A., FREIRE, E., SHAPIRO, L., LEE, K. K., ARTHOS, J., MUNRO, J. B., BLANCHARD, S. C., MOTHES, W., BINLEY, J. M., MCDERMOTT, A. B., MASCOLA, J. R. & KWONG, P. D. 2015. Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol,* 22**,** 522-531.

DODDING, M. P., BOCK, M., YAP, M. W. & STOYE, J. P. 2005. Capsid processing requirements for abrogation of Fv1 and Ref1 restriction. *J Virol,* 79**,** 10571-7.

DONAHUE, J. P., LEVINSON, R. T., SHEEHAN, J. H., SUTTON, L., TAYLOR, H. E., MEILER, J., D'AQUILA, R. T. & SONG, C. 2015. Genetic Analysis of the Localization of APOBEC3F to Human Immunodeficiency Virus Type 1 Virion Cores. *Journal of Virology,* 89**,** 2415-2424.

DU, S., BETTS, L., YANG, R., SHI, H., CONCEL, J., AHN, J., AIKEN, C., ZHANG, P. & YEH, J. I. 2011. Structure of the HIV-1 Full-Length Capsid in a Conformationally-Trapped Unassembled State Induced by Small-Molecule Binding. *Journal of Molecular Biology,* 406**,** 371-386.

DUDA, A., STANGE, A., LÜFTENEGGER, D., STANKE, N., WESTPHAL, D., PIETSCHMANN, T., EASTMAN, S. W., LINIAL, M. L., RETHWILM, A. & LINDEMANN, D. 2004. Prototype Foamy Virus Envelope Glycoprotein Leader Peptide Processing Is Mediated by a Furin-Like Cellular Protease, but Cleavage Is Not Essential for Viral Infectivity. *Journal of Virology,* 78**,** 13865-13870.

DUGGAL, N. K., FU, W., AKEY, J. M. & EMERMAN, M. 2013. Identification and Antiviral Activity of Common Polymorphisms in the APOBEC3 locus in Human Populations. *Virology,* 443**,** 329-337.

EARL, P. L., MOSS, B. & DOMS, R. W. 1991. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *Journal of Virology,* 65**,** 2047-2055.

EDWARDS, S. A. & FAN, H. 1979. gag-Related polyproteins of Moloney murine leukemia virus: evidence for independent synthesis of glycosylated and unglycosylated forms. *Journal of Virology,* 30**,** 551-563.

EINFELD, D. 1996. Maturation and assembly of retroviral glycoproteins. *Curr Top Microbiol Immunol,* 214**,** 133-76.

ELIS, E., EHRLICH, M., PRIZAN-RAVID, A., LAHAM-KARAM, N. & BACHARACH, E. 2012. p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. *PLoS Pathog,* 8**,** 27.

EMERMAN, M. & MALIM, M. H. 1998. HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology. *Science,* 280**,** 1880-4.

ENDERS, J. P., TC. 1954. Propagation in Tissue Cultures of Cytopathogenic Agents from Patients with Measles. . *Proceedings of the Society for Experimental Biology and Medicine* 86**,** 277-86.

ENGELMAN, A. & CRAIGIE, R. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J Virol,* 66**,** 6361-9.

ENSSLE, J., FISCHER, N., MOEBES, A., MAUER, B., SMOLA, U. & RETHWILM, A. 1997. Carboxy-terminal cleavage of the human foamy virus Gag precursor molecule is an essential step in the viral life cycle. *Journal of Virology,* 71**,** 7312-7.

ENSSLE, J., MOEBES, A., HEINKELEIN, M., PANHUYSEN, M., MAUER, B., SCHWEIZER, M., NEUMANN-HAEFELIN, D. & RETHWILM, A. 1999. An active foamy virus integrase is required for virus replication. *Journal of General Virology,* 80**,** 1445-1452.

ESTEVA, M. J., AFFRANCHINO, J. L. & GONZÁLEZ, S. A. 2014. Lentiviral Gag Assembly Analyzed through the Functional Characterization of Chimeric Simian Immunodeficiency Viruses Expressing Different Domains of the Feline Immunodeficiency Virus Capsid Protein. *PLoS One,* 9**,** e114299.

FARNET, C. M. & HASELTINE, W. A. 1991. Circularization of human immunodeficiency virus type 1 DNA in vitro. *Journal of Virology,* 65**,** 6942-6952.

FASCHING, L., KAPOPOULOU, A., SACHDEVA, R., PETRI, R., JÖNSSON, M. E., MÄNNE, C., TURELLI, P., JERN, P., CAMMAS, F., TRONO, D. & JAKOBSSON, J. 2015. TRIM28 Represses Transcription of Endogenous Retroviruses in Neural Progenitor Cells. *Cell Reports,* 10**,** 20-28.

FASSATI, A. & GOFF, S. P. 1999. Characterization of Intracellular Reverse Transcription Complexes of Moloney Murine Leukemia Virus. *Journal of Virology,* 73**,** 8919-8925.

FASSATI, A., GORLICH, D., HARRISON, I., ZAYTSEVA, L. & MINGOT, J. M. 2003. Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. *Embo J,* 22**,** 3675-85.

FEINSTEIN, S. C., ROSS, S. R. & YAMAMOTO, K. R. 1982. Chromosomal position effects determine transcriptional potential of integrated mammary tumor virus DNA. *Journal of Molecular Biology,* 156**,** 549-565.

FISCHER, U., HUBER, J., BOELENS, W. C., MATTAJT, L. W. & LÜHRMANN, R. 1995. The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell,* 82**,** 475-483.

FLETCHER, A. J., CHRISTENSEN, D. E., NELSON, C., TAN, C. P., SCHALLER, T., LEHNER, P. J., SUNDQUIST, W. I. & TOWERS, G. J. 2015. TRIM5α requires Ube2W to anchor Lys63-linked ubiquitin chains and restrict reverse transcription. *The EMBO Journal,* 34**,** 2078-2095.

FLETCHER, A. J., HUÉ, S., SCHALLER, T., PILLAY, D. & TOWERS, G. J. 2010. Hare TRIM5α Restricts Divergent Retroviruses and Exhibits Significant Sequence Variation from Closely Related Lagomorpha TRIM5 Genes. *Journal of Virology,* 84**,** 12463-12468.

FLÜGEL, R. M., RETHWILM, A., MAURER, B. & DARAI, G. 1987. Nucleotide sequence analysis of the env gene and its flanking regions of the human spumaretrovirus reveals two novel genes. *The EMBO Journal,* 6**,** 2077-2084.

FORSHEY, B. M., VON SCHWEDLER, U., SUNDQUIST, W. I. & AIKEN, C. 2002. Formation of a Human Immunodeficiency Virus Type 1 Core of Optimal Stability Is Crucial for Viral Replication. *Journal of Virology,* 76**,** 5667-5677.

FRANK, G. A., NARAYAN, K., BESS JR, J. W., DEL PRETE, G. Q., WU, X., MORAN, A., HARTNELL, L. M., EARL, L. A., LIFSON, J. D. & SUBRAMANIAM, S. 2015. Maturation of the HIV-1 core by a non-diffusional phase transition. *Nat Commun,* 6.

FREED, E. O. 2015. HIV-1 assembly, release and maturation. *Nat Rev Micro,* 13**,** 484-496.

FRICKE, T., WHITE, T. E., SCHULTE, B., DE SOUZA ARANHA VIEIRA, D. A., DHARAN, A., CAMPBELL, E. M., BRANDARIZ-NUÑEZ, A. & DIAZ-GRIFFERO, F. 2014. MxB binds to the HIV-1 core and prevents the uncoating process of HIV-1. *Retrovirology,* 11**,** 1-14.

GAMBLE, T. R., VAJDOS, F. F., YOO, S., WORTHYLAKE, D. K., HOUSEWEART, M., SUNDQUIST, W. I. & HILL, C. P. 1996. Crystal Structure of Human Cyclophilin A Bound to the Amino-Terminal Domain of HIV-1 Capsid. *Cell,* 87**,** 1285-1294.

GAMBLE, T. R., YOO, S., VAJDOS, F. F., VON SCHWEDLER, U. K., WORTHYLAKE, D. K., WANG, H., MCCUTCHEON, J. P., SUNDQUIST, W. I. & HILL, C. P. 1997. Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein. *Science,* 278**,** 849-853.

GANEM, D. 1991. Assembly of hepadnaviral virions and subviral particles. *Curr Top Microbiol Immunol***,** 61-83.

GANSER-PORNILLOS, B. K., CHANDRASEKARAN, V., PORNILLOS, O., SODROSKI, J. G., SUNDQUIST, W. I. & YEAGER, M. 2011. Hexagonal assembly of a restricting TRIM5α protein. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 534-539.

GANSER-PORNILLOS, B. K., CHENG, A. & YEAGER, M. 2007. Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell,* 131**,** 70-9.

GANSER-PORNILLOS, B. K., VON SCHWEDLER, U. K., STRAY, K. M., AIKEN, C. & SUNDQUIST, W. I. 2004. Assembly properties of the human immunodeficiency virus type 1 CA protein. *J Virol,* 78**,** 2545-52.

GAO, D., WU, J., WU, Y.-T., DU, F., AROH, C., YAN, N., SUN, L. & CHEN, Z. J. 2013. Cyclic GMP-AMP Synthase is an Innate Immune Sensor of HIV and Other Retroviruses. *Science (New York, N.Y.),* 341**,** 10.1126/science.1240933.

GARCIA, J. V. & MILLER, A. D. 1991. Serine phosphorylation-independent downregulation of cell-surface CD4 by nef. *Nature,* 350**,** 508-511.

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., CÔTÉ, M., RICH, R. L., MYSZKA, D. G. & SUNDQUIST, W. I. 2001. Tsg101 and the Vacuolar Protein Sorting Pathway Are Essential for HIV-1 Budding. *Cell,* 107**,** 55-65.

GEETHA, T. S., MICHEALRAJ, K. A., KABRA, M., KAUR, G., JUYAL, R. C. & THELMA, B. K. 2014. Targeted Deep Resequencing Identifies MID2 Mutation for X-Linked Intellectual Disability with Varied Disease Severity in a Large Kindred from India. *Human Mutation,* 35**,** 41-44.

GEISELHART, V., BASTONE, P., KEMPF, T., SCHNÖLZER, M. & LÖCHELT, M. 2004. Furin-Mediated Cleavage of the Feline Foamy Virus Env Leader Protein. *Journal of Virology,* 78**,** 13573-13581.

GEOGHEGAN, K. F., DIXON, H. B. F., ROSNER, P. J., HOTH, L. R., LANZETTI, A. J., BORZILLERI, K. A., MARR, E. S., PEZZULLO, L. H., MARTIN, L. B., LEMOTTE, P. K., MCCOLL, A. S., KAMATH, A. V. & STROH, J. G. 1999. Spontaneous α-N-6-Phosphogluconoylation of a “His Tag” inEscherichia

coli:The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Analytical Biochemistry,* 267**,** 169-184.

GIGLIONE, C., BOULAROT, A. & MEINNEL, T. 2004. Protein N-terminal methionine excision. *Cellular and Molecular Life Sciences CMLS,* 61**,** 1455-1474.

GITTI, R. K., LEE, B. M., WALKER, J., SUMMERS, M. F., YOO, S. & SUNDQUIST, W. I. 1996. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science,* 273**,** 231-5.

GOEPFERT, P. A., SHAW, K., WANG, G., BANSAL, A., EDWARDS, B. H. & MULLIGAN, M. J. 1999. An Endoplasmic Reticulum Retrieval Signal Partitions Human Foamy Virus Maturation to Intracytoplasmic Membranes. *Journal of Virology,* 73**,** 7210-7217.

GOLDSTONE, D. C., ENNIS-ADENIRAN, V., HEDDEN, J. J., GROOM, H. C. T., RICE, G. I., CHRISTODOULOU, E., WALKER, P. A., KELLY, G., HAIRE, L. F., YAP, M. W., DE CARVALHO, L. P. S., STOYE, J. P., CROW, Y. J., TAYLOR, I. A. & WEBB, M. 2011. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature,* 480**,** 379-382.

GOLDSTONE, D. C., FLOWER, T. G., BALL, N. J., SANZ-RAMOS, M., YAP, M. W., OGRODOWICZ, R. W., STANKE, N., REH, J., LINDEMANN, D., STOYE, J. P. & TAYLOR, I. A. 2013. A Unique Spumavirus Gag N-terminal Domain with Functional Properties of Orthoretroviral Matrix and Capsid. *PLoS Pathog,* 9**,** e1003376.

GOLDSTONE, D. C., WALKER, P. A., CALDER, L. J., COOMBS, P. J., KIRKPATRICK, J. & BALL, N. J. 2014a. Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. *Proc Natl Acad Sci USA.*

GOLDSTONE, D. C., WALKER, P. A., CALDER, L. J., COOMBS, P. J., KIRKPATRICK, J., BALL, N. J., HILDITCH, L., YAP, M. W., ROSENTHAL, P. B., STOYE, J. P. & TAYLOR, I. A. 2014b. Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 9609-9614.

GOODENOW, M. M., BLOOM, G., ROSE, S. L., POMEROY, S. M., O'BRIEN, P. O., PEREZ, E. E., SLEASMAN, J. W. & DUNN, B. M. 2002. Naturally Occurring Amino Acid Polymorphisms in Human Immunodeficiency Virus Type 1 (HIV-1) Gag p7NC and the C-Cleavage Site Impact Gag-Pol Processing by HIV-1 Protease. *Virology,* 292**,** 137-149.

GOUJON, C., GREENBURY, R. A., PAPAIOANNOU, S., DOYLE, T. & MALIM, M. H. 2015. A Triple-Arginine Motif in the Amino-Terminal Domain and Oligomerization Are Required for HIV-1 Inhibition by Human MX2. *Journal of Virology,* 89**,** 4676-4680.

GOUJON, C., MONCORGÉ, O., BAUBY, H., DOYLE, T., WARD, C. C., SCHALLER, T., HUÉ, S., BARCLAY, W. S., SCHULZ, R. & MALIM, M. H. 2013. Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature,* 502**,** 10.1038/nature12542.

GRES, A. T., KIRBY, K. A., KEWALRAMANI, V. N., TANNER, J. J., PORNILLOS, O. & SARAFIANOS, S. G. 2015. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science,* 349**,** 99-103.

GUPTA, S. S., MAETZIG, T., MAERTENS, G. N., SHARIF, A., ROTHE, M., WEIDNER-GLUNDE, M., GALLA, M., SCHAMBACH, A., CHEREPANOV, P. & SCHULZ, T. F. 2013. Bromo- and Extraterminal Domain Chromatin Regulators Serve as Cofactors for Murine Leukemia Virus Integration. *Journal of Virology,* 87**,** 12721-12736.

HALLENBERGER, S., MOULARD, M., SORDEL, M., KLENK, H. D. & GARTEN, W. 1997a. The role of eukaryotic subtilisin-like endoproteases for the activation of human immunodeficiency virus glycoproteins in natural host cells. *J Virol,* 71**,** 1036-45.

HALLENBERGER, S., MOULARD, M., SORDEL, M., KLENK, H. D. & GARTEN, W. 1997b. The role of eukaryotic subtilisin-like endoproteases for the activation of human immunodeficiency virus glycoproteins in natural host cells. *Journal of Virology,* 71**,** 1036-1045.

HANSEN, M. S. T., SMITH, G. J., KAFRI, T., MOLTENI, V., SIEGEL, J. S. & BUSHMAN, F. D. 1999. Integration complexes derived from HIV vectors for rapid assays in vitro. *Nat Biotech,* 17**,** 578-582.

HARE, S., GUPTA, S. S., VALKOV, E., ENGELMAN, A. & CHEREPANOV, P. 2010. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature,* 464**,** 232-236.

HARRIS, R. S., BISHOP, K. N., SHEEHY, A. M., CRAIG, H. M., PETERSEN-MAHRT, S. K., WATT, I. N., NEUBERGER, M. S. & MALIM, M. H. 2003. DNA Deamination Mediates Innate Immunity to Retroviral Infection. *Cell,* 113**,** 803-809.

HARTLEY, J. W., ROWE, W. P. & HUEBNER, R. J. 1970. Host-range restrictions of murine leukemia viruses in mouse embryo cell cultures. *J Virol,* 5**,** 221-5.

HATZIIOANNOU, T., COWAN, S., GOFF, S. P., BIENIASZ, P. D. & TOWERS, G. J. 2003. Restriction of multiple divergent retroviruses by Lv1 and Ref1. *Embo J,* 22**,** 385-94.

HAUSER, H., LOPEZ, L. A., YANG, S. J., OLDENBURG, J. E., EXLINE, C. M., GUATELLI, J. C. & CANNON, P. M. 2010. HIV-1 Vpu and HIV-2 Env counteract BST-2/tetherin by sequestration in a perinuclear compartment. *Retrovirology,* 7**,** 1-16.

HE, F., BLAIR, W. S., FUKUSHIMA, J. & CULLEN, B. R. 1996. The human foamy virus Bel-1 transcription factor is a sequence-specific DNA binding protein. *Journal of Virology,* 70**,** 3902-8.

HEINKELEIN, M., LEURS, C., RAMMLING, M., PETERS, K., HANENBERG, H. & RETHWILM, A. 2002. Pregenomic RNA Is Required for Efficient Incorporation of Pol Polyprotein into Foamy Virus Capsids. *Journal of Virology,* 76**,** 10069-10073.

HEINZINGER, N. K., BUKINSKY, M. I., HAGGERTY, S. A., RAGLAND, A. M., KEWALRAMANI, V., LEE, M. A., GENDELMAN, H. E., RATNER, L., STEVENSON, M. & EMERMAN, M. 1994. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 7311-7315.

HENDERSON, B. R. & PERCIPALLE, P. 1997. Interactions between HIV rev and nuclear import and export factors: the rev nuclear localisation signal mediates specific binding to human importin-β1. *Journal of Molecular Biology,* 274**,** 693-707.

HENDERSON, L. E., KRUTZSCH, H. C. & OROSZLAN, S. 1983. Myristyl amino-terminal acylation of murine retrovirus proteins: An unusual post-translational protein modification. *Proceedings of the National Academy of Sciences of the United States of America,* 80**,** 339-343.

HEYMANN, J. B., BUTAN, C., WINKLER, D. C., CRAVEN, R. C. & STEVEN, A. C. 2008. Irregular and Semi-Regular Polyhedral Models for Rous Sarcoma Virus Cores. *Computational and mathematical methods in medicine,* 9**,** 197-210.

HILDITCH, L., MATADEEN, R., GOLDSTONE, D. C., ROSENTHAL, P. B., TAYLOR, I. A. & 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**,** 5771-6.

HINZ, A., MIGUET, N., NATRAJAN, G., USAMI, Y., YAMANAKA, H., RENESTO, P., HARTLIEB, B., MCCARTHY, A. A., SIMORRE, J.-P., GOTTLINGER, H. & WEISSENHORN, W. 2010. Structural basis of HIV-1 tethering to membranes by the Bst2/tetherin ectodomain. *Cell Host & Microbe,* 7**,** 314-323.

HOFMANN, W., SCHUBERT, D., LABONTE, J., MUNSON, L., GIBSON, S., SCAMMELL, J., FERRIGNO, P. & SODROSKI, J. 1999. Species-Specific, Postentry Barriers to Primate Immunodeficiency Virus Infection. *Journal of Virology,* 73**,** 10020-10028.

HOMBROUCK, A., DE RIJCK, J., HENDRIX, J., VANDEKERCKHOVE, L., VOET, A., DE MAEYER, M., WITVROUW, M., ENGELBORGHS, Y., CHRIST, F., GIJSBERS, R. & DEBYSER, Z. 2007. Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog,* 3.

HULME, A. E., KELLEY, Z., FOLEY, D. & HOPE, T. J. 2015. Complementary Assays Reveal a Low Level of CA Associated with Viral Complexes in the Nuclei of HIV-1-Infected Cells. *Journal of Virology,* 89**,** 5350-5361.

HULTQUIST, J. F., LENGYEL, J. A., REFSLAND, E. W., LARUE, R. S., LACKEY, L., BROWN, W. L. & HARRIS, R. S. 2011. Human and Rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H Demonstrate a Conserved Capacity To Restrict Vif-Deficient HIV-1. *Journal of Virology,* 85**,** 11220-11234.

HURLEY, J. H. & HANSON, P. I. 2010. Membrane budding and scission by the ESCRT machinery: it's all in the neck. *Nature reviews. Molecular cell biology,* 11**,** 556-566.

IMRICH, H., HEINKELEIN, M., HERCHENRÖDER, O. & RETHWILM, A. 2000. Primate foamy virus Pol proteins are imported into the nucleus. *Journal of General Virology,* 81**,** 2941-2947.

INUZUKA, M., HAYAKAWA, M. & INGI, T. 2005. Serinc, an Activity-regulated Protein Family, Incorporates Serine into Membrane Lipid Synthesis. *Journal of Biological Chemistry,* 280**,** 35776-35783.

IVANCHENKO, S., GODINEZ, W. J., LAMPE, M., KRÄUSSLICH, H.-G., EILS, R., ROHR, K., BRÄUCHLE, C., MÜLLER, B. & LAMB, D. C. 2009. Dynamics of HIV-1 Assembly and Release. *PLoS Pathogens,* 5**,** e1000652.

IVANOV, D., TSODIKOV, O. V., KASANOV, J., ELLENBERGER, T., WAGNER, G. & COLLINS, T. 2007. Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc Natl Acad Sci U S A,* 104**,** 4353-8.

IWATANI, Y., CHAN, D. S. B., WANG, F., MAYNARD, K. S., SUGIURA, W., GRONENBORN, A. M., ROUZINA, I., WILLIAMS, M. C., MUSIER-FORSYTH, K. & LEVIN, J. G. 2007. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Research,* 35**,** 7096-7108.

JACKS, T., POWER, M. D., MASIARZ, F. R., LUCIW, P. A., BARR, P. J. & VARMUS, H. E. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature,* 331**,** 280-283.

JACOBO-MOLINA, A., DING, J., NANNI, R. G., CLARK, A. D., JR., LU, X., TANTILLO, C., WILLIAMS, R. L., KAMER, G., FERRIS, A. L., CLARK, P. & ET AL. 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc Natl Acad Sci U S A,* 90**,** 6320-4.

JAGUVA VASUDEVAN, A. A., PERKOVIĆ, M., BULLIARD, Y., CICHUTEK, K., TRONO, D., HÄUSSINGER, D. & MÜNK, C. 2013. Prototype Foamy Virus Bet Impairs the Dimerization and Cytosolic Solubility of Human APOBEC3G. *Journal of Virology,* 87**,** 9030-9040.

JAMES, L. C., KEEBLE, A. H., KHAN, Z., RHODES, D. A. & TROWSDALE, J. 2007. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. *Proc Natl Acad Sci U S A,* 104**,** 6200-5.

JAVANBAKHT, H., DIAZ-GRIFFERO, F., STREMLAU, M., SI, Z. & SODROSKI, J. 2005. The contribution of RING and B-box 2 domains to retroviral restriction mediated by monkey TRIM5alpha. *J Biol Chem,* 280**,** 26933-40.

JAVANBAKHT, H., YUAN, W., YEUNG, D. F., SONG, B., DIAZ-GRIFFERO, F., LI, Y., LI, X., STREMLAU, M. & SODROSKI, J. 2006. Characterization of TRIM5alpha trimerization and its contribution to human immunodeficiency virus capsid binding. *Virology,* 353**,** 234-46.

JIA, B., SERRA-MORENO, R., NEIDERMYER, W., JR., RAHMBERG, A., MACKEY, J., FOFANA, I. B., JOHNSON, W. E., WESTMORELAND, S. & EVANS, D. T. 2009. Species-Specific Activity of SIV Nef and HIV-1 Vpu in Overcoming Restriction by Tetherin/BST2. *PLoS Pathog,* 5**,** e1000429.

JOHNSON, M. S., MCCLURE, M. A., FENG, D. F., GRAY, J. & DOOLITTLE, R. F. 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc Natl Acad Sci U S A,* 83**,** 7648-52.

JOHNSON, W. E. & SAWYER, S. L. 2009. Molecular evolution of the antiretroviral TRIM5 gene. *Immunogenetics,* 61**,** 163-76.

JOLICOEUR, P. & 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**,** 2236-40.

JONES, J. S., ALLAN, R. W. & TEMIN, H. M. 1993. Alteration of location of dimer linkage sequence in retroviral RNA: little effect on replication or homologous recombination. *J Virol,* 67**,** 3151-8.

JOUVENET, N., BIENIASZ, P. D. & SIMON, S. M. 2008. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature,* 454**,** 236-240.

JOWETT, J. B., PLANELLES, V., POON, B., SHAH, N. P., CHEN, M. L. & CHEN, I. S. 1995. The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *Journal of Virology,* 69**,** 6304-6313.

JULIEN, J.-P., CUPO, A., SOK, D., STANFIELD, R. L., LYUMKIS, D., DELLER, M. C., KLASSE, P.-J., BURTON, D. R., SANDERS, R. W., MOORE, J. P., WARD, A. B. & WILSON, I. A. 2013. Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science (New York, N.Y.),* 342**,** 10.1126/science.1245625.

KAFAIE, J., SONG, R., ABRAHAMYAN, L., MOULAND, A. J. & LAUGHREA, M. 2008. Mapping of nucleocapsid residues important for HIV-1 genomic RNA dimerization and packaging. *Virology,* 375**,** 592-610.

KANE, M., YADAV, S. S., BITZEGEIO, J., KUTLUAY, S. B., ZANG, T., WILSON, S. J., SCHOGGINS, J. W., RICE, C. M., YAMASHITA, M., HATZIIOANNOU, T. & BIENIASZ, P. D. 2013. Mx2 is an interferon induced inhibitor of HIV-1 infection. *Nature,* 502**,** 563-566.

KANG, Y., BLAIR, W. S. & CULLEN, B. R. 1998. Identification and Functional Characterization of a High-Affinity Bel-1 DNA Binding Site Located in the Human Foamy Virus Internal Promoter. *Journal of Virology,* 72**,** 504-511.

KAVANAUGH, M. P., MILLER, D. G., ZHANG, W., LAW, W., KOZAK, S. L., KABAT, D. & MILLER, A. D. 1994. Cell-surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium-dependent phosphate symporters. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 7071-7075.

KEELE, B. F., GIORGI, E. E., SALAZAR-GONZALEZ, J. F., DECKER, J. M., PHAM, K. T., SALAZAR, M. G., SUN, C., GRAYSON, T., WANG, S., LI, H., WEI, X., JIANG, C., KIRCHHERR, J. L., GAO, F., ANDERSON, J. A., PING, L.-H.,

SWANSTROM, R., TOMARAS, G. D., BLATTNER, W. A., GOEPFERT, P. A., KILBY, J. M., SAAG, M. S., DELWART, E. L., BUSCH, M. P., COHEN, M. S., MONTEFIORI, D. C., HAYNES, B. F., GASCHEN, B., ATHREYA, G. S., LEE, H. Y., WOOD, N., SEOIGHE, C., PERELSON, A. S., BHATTACHARYA, T., KORBER, B. T., HAHN, B. H. & SHAW, G. M. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 7552-7557.

KELLER, P. W., HUANG, R. K., ENGLAND, M. R., WAKI, K., CHENG, N., HEYMANN, J. B., CRAVEN, R. C., FREED, E. O. & STEVEN, A. C. 2013. A Two-Pronged Structural Analysis of Retroviral Maturation Indicates that Core Formation Proceeds by a Disassembly-Reassembly Pathway Rather than a Displacive Transition. *Journal of Virology,* 87**,** 13655-13664.

KEOWN, J. R. & GOLDSTONE, D. C. 2016. Crystal structure of the Trim5α Bbox2 domain from rhesus macaques describes a plastic oligomerisation interface. *Journal of Structural Biology*.

KEOWN, J. R., YANG, J. X., DOUGLAS, J. & GOLDSTONE, D. C. 2016. Characterisation of assembly and ubiquitylation by the RBCC motif of Trim5α. *Scientific reports,* 6**,** 26837.

KILZER, J. M., STRACKER, T., BEITZEL, B., MEEK, K., WEITZMAN, M. & BUSHMAN, F. D. 2003. Roles of host cell factors in circularization of retroviral dna. *Virology,* 314**,** 460-467.

KIM, J., TIPPER, C. & SODROSKI, J. 2011. Role of TRIM5alpha RING domain E3 ubiquitin ligase activity in capsid disassembly, reverse transcription blockade, and restriction of simian immunodeficiency virus. *J Virol,* 85**,** 8116-32.

KIM, S. Y., BYRN, R., GROOPMAN, J. & BALTIMORE, D. 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *Journal of Virology,* 63**,** 3708-3713.

KIM, Y. K., BOURGEOIS, C. F., ISEL, C., CHURCHER, M. J. & KARN, J. 2002. Phosphorylation of the RNA Polymerase II Carboxyl-Terminal Domain by CDK9 Is Directly Responsible for Human Immunodeficiency Virus Type 1 Tat-Activated Transcriptional Elongation. *Molecular and Cellular Biology,* 22**,** 4622-4637.

KOBAYASHI, T., KOIZUMI, Y., TAKEUCHI, J. S., MISAWA, N., KIMURA, Y., MORITA, S., AIHARA, K., KOYANAGI, Y., IWAMI, S. & SATO, K. 2014. Quantification of Deaminase Activity-Dependent and -Independent Restriction of HIV-1 Replication Mediated by APOBEC3F and APOBEC3G through Experimental-Mathematical Investigation. *Journal of Virology,* 88**,** 5881-5887.

KOH, Y., WU, X., FERRIS, A. L., MATREYEK, K. A., SMITH, S. J., LEE, K., KEWALRAMANI, V. N., HUGHES, S. H. & ENGELMAN, A. 2013. Differential Effects of Human Immunodeficiency Virus Type 1 Capsid and Cellular Factors Nucleoporin 153 and LEDGF/p75 on the Efficiency and Specificity of Viral DNA Integration. *Journal of Virology,* 87**,** 648-658.

KOHL, N. E., EMINI, E. A., SCHLEIF, W. A., DAVIS, L. J., HEIMBACH, J. C., DIXON, R. A., SCOLNICK, E. M. & SIGAL, I. S. 1988. Active human immunodeficiency virus protease is required for viral infectivity. *Proceedings of the National Academy of Sciences,* 85**,** 4686-4690.

KONG, J., XU, B., WEI, W., WANG, X., XIE, W. & YU, X.-F. 2014. Characterization of the amino-terminal domain of Mx2/MxB-dependent interaction with the HIV-1 capsid. *Protein & Cell,* 5**,** 954-957.

KONO, K., BOZEK, K., DOMINGUES, F. S., SHIODA, T. & NAKAYAMA, E. E. 2009. Impact of a single amino acid in the variable region 2 of the Old World monkey

TRIM5α SPRY (B30.2) domain on anti-human immunodeficiency virus type 2 activity. *Virology,* 388**,** 160-168.

KONVALINKA, J., LOCHELT, M., ZENTGRAF, H., FLUGEL, R. M. & KRAUSSLICH, H. G. 1995a. Active foamy virus proteinase is essential for virus infectivity but not for formation of a Pol polyprotein. *J Virol,* 69**,** 7264-8.

KONVALINKA, J., LÖCHELT, M., ZENTGRAF, H., FLÜGEL, R. M. & KRÄUSSLICH, H. G. 1995b. Active foamy virus proteinase is essential for virus infectivity but not for formation of a Pol polyprotein. *Journal of Virology,* 69**,** 7264-8.

KOVALSKYY, D. B. & IVANOV, D. N. 2014. Recognition of the HIV Capsid by the TRIM5α Restriction Factor Is Mediated by a Subset of Pre-Existing Conformations of the TRIM5α SPRY Domain. *Biochemistry,* 53**,** 1466-1476.

KOZAK, C. A. 2010. The mouse "xenotropic" gammaretroviruses and their XPR1 receptor. *Retrovirology,* 7**,** 101-101.

KOZAK, C. A. & CHAKRABORTI, A. 1996. Single amino acid changes in the murine leukemia virus capsid protein gene define the target of Fv1 resistance. *Virology,* 225**,** 300-5.

KOZAK, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research,* 15**,** 8125-8148.

KUTLUAY, S. B., ZANG, T., BLANCO-MELO, D., POWELL, C., JANNAIN, D., ERRANDO, M. & BIENIASZ, P. D. 2014. Global changes in the RNA binding specificity of HIV-1 Gag regulate virion genesis. *Cell,* 159**,** 1096-1109.

KWONG, P. D., WYATT, R., ROBINSON, J., SWEET, R. W., SODROSKI, J. & HENDRICKSON, W. A. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature,* 393**,** 648-659.

LAGUETTE, N., RAHM, N., SOBHIAN, B., CHABLE-BESSIA, C., MÜNCH, J., SNOECK, J., SAUTER, D., SWITZER, W. M., HENEINE, W., KIRCHHOFF, F., DELSUC, F., TELENTI, A. & BENKIRANE, M. 2012. Evolutionary and Functional Analyses of the Interaction between the Myeloid Restriction Factor SAMHD1 and the Lentiviral Vpx Protein. *Cell Host & Microbe,* 11**,** 205-217.

LAGUETTE, N., SOBHIAN, B., CASARTELLI, N., RINGEARD, M., CHABLE-BESSIA, C., SÉGÉRAL, E., YATIM, A., EMILIANI, S., SCHWARTZ, O. & BENKIRANE, M. 2011. SAMHD1 is the dendritic– and myeloid–cell–specific HIV–1 restriction factor counteracted by Vpx. *Nature,* 474**,** 654-657.

LAHAYE, X., SATOH, T., GENTILI, M., CERBONI, S., CONRAD, C., HURBAIN, I., EL MARJOU, A., LACABARATZ, C., LELIÈVRE, J.-D. & MANEL, N. 2013. The Capsids of HIV-1 and HIV-2 Determine Immune Detection of the Viral cDNA by the Innate Sensor cGAS in Dendritic Cells. *Immunity,* 39**,** 1132-1142.

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. & MARGOTTIN-GOGUET, F. 2012. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol,* 13**,** 223-228.

LANGELIER, C. R., SANDRIN, V., ECKERT, D. M., CHRISTENSEN, D. E., CHANDRASEKARAN, V., ALAM, S. L., AIKEN, C., OLSEN, J. C., KAR, A. K., SODROSKI, J. G. & SUNDQUIST, W. I. 2008. Biochemical Characterization of a Recombinant TRIM5α Protein That Restricts Human Immunodeficiency Virus Type 1 Replication. *Journal of Virology,* 82**,** 11682-11694.

LASCANO, J., UCHIL, P. D., MOTHES, W. & LUBAN, J. 2016. TRIM5 Retroviral Restriction Activity Correlates with the Ability To Induce Innate Immune Signaling. *Journal of Virology,* 90**,** 308-316.

LAWRENZ-SMITH, S. C. & THOMAS, C. Y. 1995. The E47 transcription factor binds to the enhancer sequences of recombinant murine leukemia viruses and influences enhancer function. *Journal of Virology,* 69**,** 4142-4148.

LEE, E.-G., KUPPERS, D., HORN, M., ROY, J., MAY, C. & LINIAL, M. L. 2008. A Premature Termination Codon Mutation at the C Terminus of Foamy Virus Gag Downregulates the Levels of Spliced pol mRNA. *Journal of Virology,* 82**,** 1656-1664.

LEE, K., MULKY, A., YUEN, W., MARTIN, T. D., MEYERSON, N. R., CHOI, L., YU, H., SAWYER, S. L. & KEWALRAMANI, V. N. 2012. HIV-1 Capsid-Targeting Domain of Cleavage and Polyadenylation Specificity Factor 6. *Journal of Virology,* 86**,** 3851-3860.

LEHMANN, M. J., SHERER, N. M., MARKS, C. B., PYPAERT, M. & MOTHES, W. 2005. Actin- and myosin-driven movement of viruses along filopodia precedes their entry into cells. *The Journal of Cell Biology,* 170**,** 317-325.

LEVIN, J. G., GUO, J., ROUZINA, I. & MUSIER‐FORSYTH, K. 2005. Nucleic Acid Chaperone Activity of HIV‐1 Nucleocapsid Protein: Critical Role in Reverse Transcription and Molecular Mechanism. *Progress in Nucleic Acid Research and Molecular Biology.* Academic Press.

LI, B., ZHOU, T. & ZOU, Y. 2016a. Mid1/Mid2 expression in craniofacial development and a literature review of X‐linked opitz syndrome. *Molecular Genetics & Genomic Medicine,* 4**,** 95-105.

LI, F., JIN, J., HERRMANN, C. & MOTHES, W. 2013a. Basic residues in the matrix domain and multimerization target murine leukemia virus Gag to the virological synapse. *J Virol,* 87**,** 7113-26.

LI, K., MARKOSYAN, R. M., ZHENG, Y.-M., GOLFETTO, O., BUNGART, B., LI, M., DING, S., HE, Y., LIANG, C., LEE, J. C., GRATTON, E., COHEN, F. S. & LIU, S.-L. 2013b. IFITM Proteins Restrict Viral Membrane Hemifusion. *PLoS Pathog,* 9**,** e1003124.

LI, L., OLVERA, J. M., YODER, K. E., MITCHELL, R. S., BUTLER, S. L., LIEBER, M., MARTIN, S. L. & BUSHMAN, F. D. 2001. Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *The EMBO Journal,* 20**,** 3272-3281.

LI, W., YAP, M. W., VOSS, V. & STOYE, J. P. 2016b. Expression levels of Fv1: effects on retroviral restriction specificities. *Retrovirology,* 13**,** 1-17.

LI, X. & 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**,** 11495-502.

LI, Y.-L., CHANDRASEKARAN, V., CARTER, S. D., WOODWARD, C. L., CHRISTENSEN, D. E., DRYDEN, K. A., PORNILLOS, O., YEAGER, M., GANSER-PORNILLOS, B. K., JENSEN, G. J. & SUNDQUIST, W. I. 2016c. Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. *eLife,* 5**,** e16269.

LIAO, Z., GRAHAM, D. R. & HILDRETH, J. E. K. 2003. Lipid Rafts and HIV Pathogenesis: Virion-Associated Cholesterol Is Required for Fusion and Infection of Susceptible Cells. *AIDS Research and Human Retroviruses,* 19**,** 675-687.

LILLY, F. 1970. Fv-2: Identification and Location of a Second Gene Governing the Spleen Focus Response to Friend Leukemia Virus in Mice. *Journal of the National Cancer Institute,* 45**,** 163-169.

LIN, T.-Y. & EMERMAN, M. 2006. Cyclophilin A interacts with diverse lentiviral capsids. *Retrovirology,* 3**,** 70-70.

LINDEMANN, D., BOCK, M., SCHWEIZER, M. & RETHWILM, A. 1997. Efficient pseudotyping of murine leukemia virus particles with chimeric human foamy virus envelope proteins. *Journal of Virology,* 71**,** 4815-20.

LINDEMANN, D., PIETSCHMANN, T., PICARD-MAUREAU, M., BERG, A., HEINKELEIN, M., THUROW, J., KNAUS, P., ZENTGRAF, H. & RETHWILM, A. 2001. A Particle-Associated Glycoprotein Signal Peptide Essential for Virus Maturation and Infectivity. *Journal of Virology,* 75**,** 5762-5771.

LINIAL, M. 2000. Why aren't foamy viruses pathogenic? *Trends in Microbiology,* 8**,** 284-289.

LINIAL, M. L. 1999. Foamy Viruses Are Unconventional Retroviruses. *Journal of Virology,* 73**,** 1747-1755.

LIU, F. L., KUANG, Y. Q., MU, D., ZHENG, H. Y., ZHU, J. W. & ZHENG, Y. T. 2015. The Effect of Exon 7 Deletion during the Evolution of TRIMCyp Fusion Proteins on Viral Restriction, Cytoplasmic Body Formation and Multimerization. *PLoS One,* 10**,** e0121666.

LIU, L., OLIVEIRA, N. M. M., CHENEY, K. M., PADE, C., DREJA, H., BERGIN, A.-M. H., BORGDORFF, V., BEACH, D. H., BISHOP, C. L., DITTMAR, M. T. & MCKNIGHT, Á. 2011a. A whole genome screen for HIV restriction factors. *Retrovirology,* 8**,** 94-94.

LIU, Y., KIM, Y.-B. & LÖCHELT, M. 2011b. N-Terminally Myristoylated Feline Foamy Virus Gag Allows Env-Independent Budding of Sub-Viral Particles. *Viruses,* 3**,** 2223-2237.

LIU, Z., PAN, Q., DING, S., QIAN, J., XU, F., ZHOU, J., CEN, S., GUO, F. & LIANG, C. 2013. The Interferon-Inducible MxB Protein Inhibits HIV-1 Infection. *Cell Host & Microbe,* 14**,** 398-410.

LOCHELT, M. & FLUGEL, R. M. 1996. The human foamy virus pol gene is expressed as a Pro-Pol polyprotein and not as a Gag-Pol fusion protein. *J Virol,* 70**,** 1033-40.

LÖCHELT, M., MURANYI, W. & FLÜGEL, R. M. 1993. Human foamy virus genome possesses an internal, Bel-1-dependent and functional promoter. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 7317-7321.

LÖCHELT, M., ROMEN, F., BASTONE, P., MUCKENFUSS, H., KIRCHNER, N., KIM, Y.-B., TRUYEN, U., RÖSLER, U., BATTENBERG, M., SAIB, A., FLORY, E., CICHUTEK, K. & MÜNK, C. 2005. The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 7982-7987.

LOEW, R. 2009. The use of retroviral vectors for tet-regulated gene expression in cell populations. *Methods Mol Biol,* 506.

LU, J., PAN, Q., RONG, L., LIU, S.-L. & LIANG, C. 2011a. The IFITM Proteins Inhibit HIV-1 Infection. *Journal of Virology,* 85**,** 2126-2137.

LU, K., HENG, X., GARYU, L., MONTI, S., GARCIA, E. L., KHARYTONCHYK, S., DORJSUREN, B., KULANDAIVEL, G., JONES, S., HIREMATH, A., DIVAKARUNI, S. S., LACOTTI, C., BARTON, S., TUMMILLO, D., HOSIC, A., EDME, K., ALBRECHT, S., TELESNITSKY, A. & SUMMERS, M. F. 2011b. NMR Detection of Structures in the HIV-1 5´-Leader RNA that Regulate Genome Packaging. *Science (New York, N.Y.),* 334**,** 242-245.

LUBAN, J., BOSSOLT, K. L., FRANKE, E. K., KALPANA, G. V. & GOFF, S. P. 1993. Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell,* 73**,** 1067-1078.

LUKIC, D. S., HOTZ-WAGENBLATT, A., LEI, J., RÄTHE, A.-M., MÜHLE, M., DENNER, J., MÜNK, C. & LÖCHELT, M. 2013. Identification of the feline foamy virus Bet domain essential for APOBEC3 counteraction. *Retrovirology,* 10**,** 76-76.

LUKIC, Z., DHARAN, A., FRICKE, T., DIAZ-GRIFFERO, F. & CAMPBELL, E. M. 2014. HIV-1 Uncoating Is Facilitated by Dynein and Kinesin 1. *Journal of Virology,* 88**,** 13613-13625.

LURIA, S., CHAMBERS, I. & BERG, P. 1991. Expression of the type 1 human immunodeficiency virus Nef protein in T cells prevents antigen receptor-mediated induction of interleukin 2 mRNA. *Proceedings of the National Academy of Sciences of the United States of America,* 88**,** 5326-5330.

LY, H. & PARSLOW, T. G. 2002. Bipartite Signal for Genomic RNA Dimerization in Moloney Murine Leukemia Virus. *Journal of Virology,* 76**,** 3135-3144.

MADANI, N. & KABAT, D. 1998. An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. *Journal of Virology,* 72**,** 10251-10255.

MAERTENS, G. N., HARE, S. & CHEREPANOV, P. 2010. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature,* 468**,** 326-329.

MANDELL, M. A., JAIN, A., ARKO-MENSAH, J., CHAUHAN, S., KIMURA, T., DINKINS, C., SILVESTRI, G., MÜNCH, J., KIRCHHOFF, F., SIMONSEN, A., WEI, Y., LEVINE, B., JOHANSEN, T. & DERETIC, V. 2014a. TRIM proteins regulate autophagy and can target autophagic substrates by direct recognition. *Developmental cell,* 30**,** 394-409.

MANDELL, M. A., KIMURA, T., JAIN, A., JOHANSEN, T. & DERETIC, V. 2014b. TRIM proteins regulate autophagy: TRIM5 is a selective autophagy receptor mediating HIV-1 restriction. *Autophagy,* 10**,** 2387-2388.

MANGEAT, B., TURELLI, P., CARON, G., FRIEDLI, M., PERRIN, L. & TRONO, D. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature,* 424**,** 99-103.

MARINO-BUSLJE, C., MIZUGUCHI, K., SIDDLE, K. & BLUNDELL, T. L. 1998. A third fibronectin type III domain in the extracellular region of the insulin receptor family. *FEBS Letters,* 441**,** 331-336.

MARNO, K. M., OGUNKOLADE, B. W., PADE, C., OLIVEIRA, N. M. M., O'SULLIVAN, E. & MCKNIGHT, Á. 2014. Novel restriction factor RNA-associated early-stage anti-viral factor (REAF) inhibits human and simian immunodeficiency viruses. *Retrovirology,* 11**,** 3-3.

MASSIAH, M. A., MATTS, J. A. B., SHORT, K. M., SIMMONS, B. N., SINGIREDDY, S., YI, Z. & COX, T. C. 2007. Solution Structure of the MID1 B-box2 CHC(D/C)C2H2 Zinc-binding Domain: Insights into an Evolutionarily Conserved RING Fold. *Journal of Molecular Biology,* 369**,** 1-10.

MASSIAH, M. A., SIMMONS, B. N., SHORT, K. M. & COX, T. C. 2006. Solution Structure of the RBCC/TRIM B-box1 Domain of Human MID1: B-box with a RING. *Journal of Molecular Biology,* 358**,** 532-545.

MASSIAH, M. A., STARICH, M. R., PASCHALL, C., SUMMERS, M. F., CHRISTENSEN, A. M. & SUNDQUIST, W. I. 1994. Three-dimensional Structure of the Human Immunodeficiency Virus Type 1 Matrix Protein. *Journal of Molecular Biology,* 244**,** 198-223.

MATREYEK, K. A. & ENGELMAN, A. 2011. The Requirement for Nucleoporin NUP153 during Human Immunodeficiency Virus Type 1 Infection Is Determined by the Viral Capsid. *Journal of Virology,* 85**,** 7818-7827.

MATREYEK, K. A., YÜCEL, S. S., LI, X. & ENGELMAN, A. 2013. Nucleoporin NUP153 Phenylalanine-Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to Mediate Lentiviral Infectivity. *PLoS Pathog,* 9**,** e1003693.

MATTHES, D., WIKTOROWICZ, T., ZAHN, J., BODEM, J., STANKE, N., LINDEMANN, D. & RETHWILM, A. 2011. Basic Residues in the Foamy Virus Gag Protein. *Journal of Virology,* 85**,** 3986-3995.

MATTHIES, A., NIMTZ, M. & LEIMKÜHLER, S. 2005. Molybdenum Cofactor Biosynthesis in Humans: Identification of a Persulfide Group in the Rhodanese-like Domain of MOCS3 by Mass Spectrometry. *Biochemistry,* 44**,** 7912-7920.

MAURER, B., BANNERT, H., DARAI, G. & FLÜGEL, R. M. 1988. Analysis of the primary structure of the long terminal repeat and the gag and pol genes of the human spumaretrovirus. *Journal of Virology,* 62**,** 1590-1597.

MCLACHLAN, A. 1991. Molecular Biology of the Hepatitis B Virus.

MEIERING, C. D., COMSTOCK, K. E. & LINIAL, M. L. 2000. Multiple Integrations of Human Foamy Virus in Persistently Infected Human Erythroleukemia Cells. *Journal of Virology,* 74**,** 1718-1726.

MEIERING, C. D. & LINIAL, M. L. 2001. Historical Perspective of Foamy Virus Epidemiology and Infection. *Clinical Microbiology Reviews,* 14**,** 165-176.

MEIERING, C. D. & LINIAL, M. L. 2002. Reactivation of a complex retrovirus is controlled by a molecular switch and is inhibited by a viral protein. *Proceedings of the National Academy of Sciences,* 99**,** 15130-15135.

MELIKYAN, G. B. 2008. Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology,* 5**,** 111-111.

MERONI, G. 2012. TRIM/RBCC proteins. *Springer*.

MERONI, G. & DIEZ-ROUX, G. 2005. TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. *BioEssays,* 27**,** 1147-1157.

MOEBES, A., ENSSLE, J., BIENIASZ, P. D., HEINKELEIN, M., LINDEMANN, D., BOCK, M., MCCLURE, M. O. & RETHWILM, A. 1997. Human foamy virus reverse transcription that occurs late in the viral replication cycle. *Journal of Virology,* 71**,** 7305-7311.

MOMANY, C., KOVARI, L. C., PRONGAY, A. J., KELLER, W., GITTI, R. K., LEE, B. M., GORBALENYA, A. E., TONG, L., MCCLURE, J., EHRLICH, L. S., SUMMERS, M. F., CARTER, C. & ROSSMANN, M. G. 1996. Crystal structure of dimeric HIV-1 capsid protein. *Nat Struct Mol Biol,* 3**,** 763-770.

MORTUZA, G. B., DODDING, M. P., GOLDSTONE, D. C., HAIRE, L. F., STOYE, J. P. & TAYLOR, I. A. 2008. Structure of B-MLV Capsid Amino-terminal Domain Reveals Key Features of Viral Tropism, Gag Assembly and Core Formation. *Journal of Molecular Biology,* 376**,** 1493-1508.

MORTUZA, G. B., HAIRE, L. F., STEVENS, A., SMERDON, S. J., STOYE, J. P. & TAYLOR, I. A. 2004. High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature,* 431**,** 481-485.

MULICHAK, A. M., HUI, J. O., TOMASSELLI, A. G., HEINRIKSON, R. L., CURRY, K. A., TOMICH, C. S., THAISRIVONGS, S., SAWYER, T. K. & WATENPAUGH, K. D. 1993. The crystallographic structure of the protease from human immunodeficiency virus type 2 with two synthetic peptidic transition state analog inhibitors. *Journal of Biological Chemistry,* 268**,** 13103-13109.

MÜLLERS, E. 2013. The Foamy Virus Gag Proteins: What Makes Them Different? *Viruses,* 5**,** 1023-1041.

MÜLLERS, E., UHLIG, T., STIRNNAGEL, K., FIEBIG, U., ZENTGRAF, H. & LINDEMANN, D. 2011. Novel Functions of Prototype Foamy Virus Gag Glycine-Arginine-Rich Boxes in Reverse Transcription and Particle Morphogenesis. *Journal of Virology,* 85**,** 1452-1463.

MURAKAMI, T. & FREED, E. O. 2000. The long cytoplasmic tail of gp41 is required in a cell type-dependent manner for HIV-1 envelope glycoprotein incorporation into virions. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 343-348.

MURAMATSU, M., SANKARANAND, V. S., ANANT, S., SUGAI, M., KINOSHITA, K., DAVIDSON, N. O. & HONJO, T. 1999. Specific Expression of Activation-induced Cytidine Deaminase (AID), a Novel Member of the RNA-editing

Deaminase Family in Germinal Center B Cells. *Journal of Biological Chemistry,* 274**,** 18470-18476.
MURRAY, P. S., LI, Z., WANG, J., TANG, C. L., HONIG, B. & MURRAY, D. 2005. Retroviral Matrix Domains Share Electrostatic Homology: Models for Membrane Binding Function throughout the Viral Life Cycle. *Structure,* 13**,** 1521-1531.
NABEL, C. S., LEE, J. W., WANG, L. C. & KOHLI, R. M. 2013. Nucleic acid determinants for selective deamination of DNA over RNA by activation-induced deaminase. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 14225-14230.
NAGHAVI, M. H. & GOFF, S. P. 2007. Retroviral proteins that interact with the host cell cytoskeleton. *Current opinion in immunology,* 19**,** 402-407.
NARAYANA, S. K., HELBIG, K. J., MCCARTNEY, E. M., EYRE, N. S., BULL, R. A., ELTAHLA, A., LLOYD, A. R. & BEARD, M. R. 2015. The Interferon-induced Transmembrane Proteins, IFITM1, IFITM2, and IFITM3 Inhibit Hepatitis C Virus Entry. *Journal of Biological Chemistry,* 290**,** 25946-25959.
NASSAL, M. 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *Journal of Virology,* 66**,** 4107-4116.
NEIL, S. J., ZANG, T. & BIENIASZ, P. D. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature,* 451**,** 425-30.
NEWMAN, R. M., HALL, L., KIRMAIER, A., POZZI, L. A., PERY, E., FARZAN, M., O'NEIL, S. P. & JOHNSON, W. 2008. Evolution of a TRIM5-CypA splice isoform in old world monkeys. *PLoS Pathog,* 4**,** 1000003.
NIELSEN, A. L., PALLISGAARD, N., PEDERSEN, F. S. & JØRGENSEN, P. 1992. Murine helix-loop-helix transcriptional activator proteins binding to the E-box motif of the Akv murine leukemia virus enhancer identified by cDNA cloning. *Molecular and Cellular Biology,* 12**,** 3449-3459.
NIELSEN, A. L., PALLISGAARD, N., PEDERSEN, F. S. & JØRGENSEN, P. 1994. Basic helix-loop-helix proteins in murine type C retrovirus transcriptional regulation. *Journal of Virology,* 68**,** 5638-5647.
NISOLE, S., LYNCH, C., STOYE, J. P. & YAP, M. W. 2004. A Trim5-cyclophilin A fusion protein found in owl monkey kidney cells can restrict HIV-1. *Proc Natl Acad Sci U S A,* 101**,** 13324-8.
NISOLE, S. & SAIB, A. 2004. Early steps of retrovirus replicative cycle. *Retrovirology,* 1**,** 9.
OCWIEJA, K. E., BRADY, T. L., RONEN, K., HUEGEL, A., ROTH, S. L., SCHALLER, T., JAMES, L. C., TOWERS, G. J., YOUNG, J. A. T., CHANDA, S. K., KÖNIG, R., MALANI, N., BERRY, C. C. & BUSHMAN, F. D. 2011. HIV Integration Targeting: A Pathway Involving Transportin-3 and the Nuclear Pore Protein RanBP2. *PLoS Pathogens,* 7**,** e1001313.
ODAKA, T. 1975. Genetic transmission of endogenous N- and B-tropic murine leukemia viruses in low-leukemic strain C57BL/6. *J Virol,* 15**,** 332-7.
OHKURA, S., GOLDSTONE, D. C., YAP, M. W., HOLDEN-DYE, K., TAYLOR, I. A. & 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**,** 31.
OHKURA, S. & STOYE, J. P. 2013. A comparison of murine leukemia viruses that escape from human and rhesus macaque TRIM5alphas. *J Virol,* 87**,** 6455-68.
OHKURA, S., YAP, M. W., SHELDON, T. & STOYE, J. P. 2006. All three variable regions of the TRIM5alpha B30.2 domain can contribute to the specificity of retrovirus restriction. *J Virol,* 80**,** 8554-65.
ONO, A., ABLAN, S. D., LOCKETT, S. J., NAGASHIMA, K. & FREED, E. O. 2004. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the

plasma membrane. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 14889-14894.

ONO, A., ORENSTEIN, J. M. & FREED, E. O. 2000. Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. *J Virol,* 74**,** 2855-66.

OOMS, M., BRAYTON, B., LETKO, M., MAIO, SUSAN M., PILCHER, CHRISTOPHER D., HECHT, FREDERICK M., BARBOUR, JASON D. & SIMON, V. 2013. HIV-1 Vif Adaptation to Human APOBEC3H Haplotypes. *Cell Host & Microbe,* 14**,** 411-421.

OOMS, M., KRIKONI, A., KRESS, A. K., SIMON, V. & MÜNK, C. 2012. APOBEC3A, APOBEC3B, and APOBEC3H Haplotype 2 Restrict Human T-Lymphotropic Virus Type 1. *Journal of Virology,* 86**,** 6097-6108.

OSHIMA, M., MURIAUX, D., MIRRO, J., NAGASHIMA, K., DRYDEN, K., YEAGER, M. & REIN, A. 2004. Effects of Blocking Individual Maturation Cleavages in Murine Leukemia Virus Gag. *Journal of Virology,* 78**,** 1411-1420.

OZATO, K., SHIN, D. M., CHANG, T. H. & MORSE, H. C., 3RD 2008. TRIM family proteins and their emerging roles in innate immunity. *Nat Rev Immunol,* 8**,** 849-60.

PACHECO, B., FINZI, A., MCGEE-ESTRADA, K. & SODROSKI, J. 2010. Species-Specific Inhibition of Foamy Viruses from South American Monkeys by New World Monkey TRIM5α Proteins. *Journal of Virology,* 84**,** 4095-4099.

PACI, A., LIU, X. H., HUANG, H., LIM, A., HOURY, W. A. & ZHAO, R. 2012. The Stability of the Small Nucleolar Ribonucleoprotein (snoRNP) Assembly Protein Pih1 in Saccharomyces cerevisiae Is Modulated by Its C Terminus. *The Journal of Biological Chemistry,* 287**,** 43205-43214.

PANTÉ, N. & KANN, M. 2002. Nuclear Pore Complex Is Able to Transport Macromolecules with Diameters of ~39 nm. *Molecular Biology of the Cell,* 13**,** 425-434.

PARK, E. Y., KWON, O.-B., JEONG, B.-C., YI, J.-S., LEE, C. S., KO, Y.-G. & SONG, H. K. 2010. Crystal structure of PRY-SPRY domain of human TRIM72. *Proteins: Structure, Function, and Bioinformatics,* 78**,** 790-795.

PAWLICA, P. & BERTHOUX, L. 2014. Cytoplasmic Dynein Promotes HIV-1 Uncoating. *Viruses,* 6**,** 4195-4211.

PENG, K., MURANYI, W., GLASS, B., LAKETA, V., YANT, S. R., TSAI, L., CIHLAR, T., MÜLLER, B. & KRÄUSSLICH, H.-G. 2014. Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *eLife,* 3**,** e04114.

PEREZ-CABALLERO, D., HATZIIOANNOU, T., YANG, A., COWAN, S. & BIENIASZ, P. D. 2005. Human tripartite motif 5alpha domains responsible for retrovirus restriction activity and specificity. *J Virol,* 79**,** 8969-78.

PEREZ, O. D. & NOLAN, G. P. 2001. Resistance Is Futile: Assimilation of Cellular Machinery by HIV-1. *Immunity,* 15**,** 687-690.

PERREIRA, J. M., CHIN, C. R., FEELEY, E. M. & BRASS, A. L. 2013. IFITMs restrict the replication of multiple pathogenic viruses. *Journal of Molecular Biology,* 425**,** 4937-4955.

PERRON, M. J., STREMLAU, M., LEE, M., JAVANBAKHT, H., SONG, B. & SODROSKI, J. 2007. The human TRIM5alpha restriction factor mediates accelerated uncoating of the N-tropic murine leukemia virus capsid. *J Virol,* 81**,** 2138-48.

PERRON, M. J., STREMLAU, M., SONG, B., ULM, W., MULLIGAN, R. C. & SODROSKI, J. 2004. TRIM5alpha mediates the postentry block to N-tropic murine leukemia viruses in human cells. *Proc Natl Acad Sci U S A,* 101**,** 11827-32.

PERRY, J., FEATHER, S., SMITH, A., PALMER, S. & ASHWORTH, A. 1998. The Human FXY Gene is Located Within Xp22.3: Implications for Evolution of the Mammalian X Chromosome. *Human Molecular Genetics,* 7**,** 299-305.

PERRY, J., SHORT, K. M., ROMER, J. T., SWIFT, S., COX, T. C. & ASHWORTH, A. 1999. FXY2/MID2, a gene related to the X-linked Opitz syndrome gene FXY/MID1, maps to Xq22 and encodes a FNIII domain-containing protein that associates with microtubules. *Genomics,* 62**,** 385-94.

PERTEL, T., HAUSMANN, S., MORGER, D., ZÜGER, S., GUERRA, J., LASCANO, J., REINHARD, C., SANTONI, F., UCHIL, P. D., CHATEL, L., BISIAUX, A., ALBERT, M., STRAMBIO-DE-CASTILLIA, C., MOTHES, W., PIZZATO, M., GRÜTTER, M. & LUBAN, J. 2011. TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature,* 472**,** 361-365.

PETERS, G., HARADA, F., DAHLBERG, J. E., PANET, A., HASELTINE, W. A. & BALTIMORE, D. 1977. Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. *Journal of Virology,* 2**1,** 1031-1041.

PETIT, C., GIRON, M.-L., TOBALY-TAPIERO, J., BITTOUN, P., REAL, E., JACOB, Y., TORDO, N., DE THÉ, H. & SAÏB, A. 2003. Targeting of incoming retroviral Gag to the centrosome involves a direct interaction with the dynein light chain 8. *Journal of Cell Science,* 116**,** 3433-3442.

PETTIT, S. C., EVERITT, L. E., CHOUDHURY, S., DUNN, B. M. & KAPLAN, A. H. 2004. Initial Cleavage of the Human Immunodeficiency Virus Type 1 GagPol Precursor by Its Activated Protease Occurs by an Intramolecular Mechanism. *Journal of Virology,* 78**,** 8477-8485.

PFREPPER, K.-I., RACKWITZ, H.-R., SCHNÖLZER, M., HEID, H., LÖCHELT, M. & FLÜGEL, R. M. 1998. Molecular Characterization of Proteolytic Processing of the Pol Proteins of Human Foamy Virus Reveals Novel Features of the Viral Protease. *Journal of Virology,* 72**,** 7648-7652.

PIETSCHMANN, T., HEINKELEIN, M., HELDMANN, M., ZENTGRAF, H., RETHWILM, A. & LINDEMANN, D. 1999. Foamy Virus Capsids Require the Cognate Envelope Protein for Particle Export. *Journal of Virology,* 73**,** 2613-2621.

POMERANTZ, R. J., TRONO, D., FEINBERG, M. B. & BALTIMORE, D. 1990. Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: A molecular model for latency. *Cell,* 61**,** 1271-1276.

PORNILLOS, O., ALAM, S. L., DAVIS, D. R. & SUNDQUIST, W. I. 2002. Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. *Nat Struct Mol Biol,* 9**,** 812-817.

PORNILLOS, O., GANSER-PORNILLOS, B. K., KELLY, B. N., HUA, Y., WHITBY, F. G., STOUT, C. D., SUNDQUIST, W. I., HILL, C. P. & YEAGER, M. 2009. X-ray structures of the hexameric building block of the HIV capsid. *Cell,* 137**,** 1282-92.

POWELL, R. D., HOLLAND, P. J., HOLLIS, T. & PERRINO, F. W. 2011. Aicardi-Goutières Syndrome Gene and HIV-1 Restriction Factor SAMHD1 Is a dGTP-regulated Deoxynucleotide Triphosphohydrolase. *The Journal of Biological Chemistry,* 286**,** 43596-43600.

PRATS, A.-C., DE BILLY, G., WANG, P. & DARLIX, J.-L. 1989. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *Journal of Molecular Biology,* 205**,** 363-372.

PRICE, A. J., FLETCHER, A. J., SCHALLER, T., ELLIOTT, T., LEE, K., KEWALRAMANI, V. N., CHIN, J. W., TOWERS, G. J. & JAMES, L. C. 2012a. CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication. *PLoS Pathogens,* 8**,** e1002896.

PRICE, A. J., FLETCHER, A. J., SCHALLER, T., ELLIOTT, T., LEE, K., KEWALRAMANI, V. N., CHIN, J. W., TOWERS, G. J. & JAMES, L. C. 2012b.

CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. *PLoS Pathog,* 8**,** 30.

PRICE, A. J., JACQUES, D. A., MCEWAN, W. A., FLETCHER, A. J., ESSIG, S., CHIN, J. W., HALAMBAGE, U. D., AIKEN, C. & JAMES, L. C. 2014. Host Cofactors and Pharmacologic Ligands Share an Essential Interface in HIV-1 Capsid That Is Lost upon Disassembly. *PLoS Pathog,* 10**,** e1004459.

PRICE, A. J., MARZETTA, F., LAMMERS, M., YLINEN, L. M. J., SCHALLER, T., WILSON, S. J., TOWERS, G. J. & JAMES, L. C. 2009. Active site remodeling switches HIV specificity of antiretroviral TRIMCyp. *Nature structural & molecular biology,* 16**,** 1036-1042.

PURCELL, D. F. & MARTIN, M. A. 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *Journal of Virology,* 67**,** 6365-6378.

PURDY, J. G., FLANAGAN, J. M., ROPSON, I. J., RENNOLL-BANKERT, K. E. & CRAVEN, R. C. 2008. Critical role of conserved hydrophobic residues within the major homology region in mature retroviral capsid assembly. *J Virol,* 82**,** 5951-61.

RABSON, A. A. G., BJ 1997. Synthesis and processing of viral RNA. *Retroviruses*.

REIN, A., HENDERSON, L. E. & LEVIN, J. G. 1998. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends in Biochemical Sciences,* 23**,** 297-301.

REIN, A., MCCLURE, M. R., RICE, N. R., LUFTIG, R. B. & SCHULTZ, A. M. 1986. Myristylation site in Pr65gag is essential for virus particle formation by Moloney murine leukemia virus. *Proceedings of the National Academy of Sciences of the United States of America,* 83**,** 7246-7250.

REN, J., BIRD, L. E., CHAMBERLAIN, P. P., STEWART-JONES, G. B., STUART, D. I. & STAMMERS, D. K. 2002. Structure of HIV-2 reverse transcriptase at 2.35-Å resolution and the mechanism of resistance to non-nucleoside inhibitors. *Proceedings of the National Academy of Sciences,* 99**,** 14410-14415.

RENAULT, N., TOBALY-TAPIERO, J., PARIS, J., GIRON, M.-L., COIFFIC, A., ROINGEARD, P. & SAÏB, A. 2011. A nuclear export signal within the structural Gag protein is required for prototype foamy virus replication. *Retrovirology,* 8**,** 6-6.

REYMOND, A., MERONI, G., FANTOZZI, A., MERLA, G., CAIRO, S., LUZI, L., RIGANELLI, D., ZANARIA, E., MESSALI, S., CAINARCA, S., GUFFANTI, A., MINUCCI, S., PELICCI, P. G. & BALLABIO, A. 2001. The tripartite motif family identifies cell compartments. *The EMBO Journal,* 20**,** 2140-2151.

RIFFEL, N., HARLOS, K., IOURIN, O., RAO, Z., KINGSMAN, A., STUART, D. & FRY, E. 2002. Atomic Resolution Structure of Moloney Murine Leukemia Virus Matrix Protein and Its Relationship to Other Retroviral Matrix Proteins. *Structure,* 10**,** 1627-1636.

RIHN, S. J., WILSON, S. J., LOMAN, N. J., ALIM, M., BAKKER, S. E., BHELLA, D., GIFFORD, R. J., RIXON, F. J. & BIENIASZ, P. D. 2013. Extreme genetic fragility of the HIV-1 capsid. *PLoS Pathog,* 9**,** 20.

RIVIÈRE, L., DARLIX, J.-L. & CIMARELLI, A. 2010. Analysis of the Viral Elements Required in the Nuclear Import of HIV-1 DNA. *Journal of Virology,* 84**,** 729-739.

ROE, T., REYNOLDS, T. C., YU, G. & BROWN, P. O. 1993. Integration of murine leukemia virus DNA depends on mitosis. *Embo J,* 12**,** 2099-108.

ROSA, A., CHANDE, A., ZIGLIO, S., DE SANCTIS, V., BERTORELLI, R., GOH, S. L., MCCAULEY, S. M., NOWOSIELSKA, A., ANTONARAKIS, S. E., LUBAN, J., SANTONI, F. A. & PIZZATO, M. 2015. HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. *Nature,* 526**,** 212-217.

ROSE, R. B., ROSE, J. R., SALTO, R., CRAIK, C. S. & STROUD, R. M. 1993. Structure of the protease from simian immunodeficiency virus: Complex with an irreversible nonpeptide inhibitor. *Biochemistry,* 32**,** 12498-12507.

ROWE, W. P. & HARTLEY, J. W. 1972. Studies of genetic transmission of murine leukemia virus by AKR mice. II. Crosses with Fv-1 b strains of mice. *J Exp Med,* 136**,** 1286-301.

RUBEN, S., PERKINS, A., PURCELL, R., JOUNG, K., SIA, R., BURGHOFF, R., HASELTINE, W. A. & ROSEN, C. A. 1989. Structural and functional characterization of human immunodeficiency virus tat protein. *Journal of Virology,* 63**,** 1-8.

RULLI JR, S. J., MURIAUX, D., NAGASHIMA, K., MIRRO, J., OSHIMA, M., BAUMANN, J. G. & REIN, A. 2006. Mutant murine leukemia virus Gag proteins lacking proline at the N-terminus of the capsid domain block infectivity in virions containing wild-type Gag. *Virology,* 347**,** 364-371.

RULLI, S. J., MIRRO, J., HILL, S. A., LLOYD, P., GORELICK, R. J., COFFIN, J. M., DERSE, D. & REIN, A. 2008. Interactions of Murine APOBEC3 and Human APOBEC3G with Murine Leukemia Viruses. *Journal of Virology,* 82**,** 6566-6575.

SAAD, J. S., MILLER, J., TAI, J., KIM, A., GHANAM, R. H. & SUMMERS, M. F. 2006. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proceedings of the National Academy of Sciences,* 103**,** 11364-11369.

SABO, Y., WALSH, D., BARRY, D. S., TINAZTEPE, S., DE LOS SANTOS, K., GOFF, S. P., GUNDERSEN, G. G. & NAGHAVI, M. H. 2013. HIV-1 induces the formation of stable microtubules to enhance early infection. *Cell Host & Microbe,* 14**,** 10.1016/j.chom.2013.10.012.

SAENZ, D. T., TEO, W., OLSEN, J. C. & POESCHLA, E. M. 2005. Restriction of feline immunodeficiency virus by Ref1, Lv1, and primate TRIM5alpha proteins. *J Virol,* 79**,** 15175-88.

SAKUMA, T., DAVILA, J. I., MALCOLM, J. A., KOCHER, J.-P. A., TONNE, J. M. & IKEDA, Y. 2014. Murine Leukemia Virus Uses NXF1 for Nuclear Export of Spliced and Unspliced Viral Transcripts. *Journal of Virology,* 88**,** 4069-4082.

SALGADO, G. F., MARQUANT, R., VOGEL, A., ALVES, I. D., FELLER, S. E., MORELLET, N. & BOUAZIZ, S. 2009. Structural Studies of HIV-1 Gag p6ct and Its Interaction with Vpr Determined by Solution Nuclear Magnetic Resonance. *Biochemistry,* 48**,** 2355-2367.

SANCHEZ, J. G., OKREGLICKA, K., CHANDRASEKARAN, V., WELKER, J. M., SUNDQUIST, W. I. & PORNILLOS, O. 2014. The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 2494-2499.

SAPHIRE, A. C. S., BOBARDT, M. D., ZHANG, Z., DAVID, G. & GALLAY, P. A. 2001. Syndecans Serve as Attachment Receptors for Human Immunodeficiency Virus Type 1 on Macrophages. *Journal of Virology,* 75**,** 9187-9200.

SASTRI, J., JOHNSEN, L., SMOLIN, N., IMAM, S., MUKHERJEE, S., LUKIC, Z., BRANDARIZ-NUÑEZ, A., ROBIA, S. L., DIAZ-GRIFFERO, F., WIETHOFF, C. & CAMPBELL, E. M. 2014. Restriction of HIV-1 by Rhesus TRIM5α Is Governed by Alpha Helices in the Linker2 Region. *Journal of Virology,* 88**,** 8911-8923.

SASTRI, J., O'CONNOR, C., DANIELSON, C. M., MCRAVEN, M., PEREZ, P., DIAZ-GRIFFERO, F. & CAMPBELL, E. M. 2010. Identification of residues within the L2 region of rhesus TRIM5α that are required for retroviral restriction and cytoplasmic body localization. *Virology,* 405**,** 259-266.

SAWYER, S. L., EMERMAN, M. & MALIK, H. S. 2004. Ancient Adaptive Evolution of the Primate Antiviral DNA-Editing Enzyme APOBEC3G. *PLoS Biol,* 2**,** e275.

SAWYER, S. L., WU, L. I., EMERMAN, M. & MALIK, H. S. 2005. Positive selection of primate TRIM5α identifies a critical species-specific retroviral restriction domain. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 2832-2837.

SAYAH, D. M., SOKOLSKAJA, E., BERTHOUX, L. & LUBAN, J. 2004. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature,* 430**,** 569-73.

SCARLATTI, G., TRESOLDI, E., BJORNDAL, A., FREDRIKSSON, R., COLOGNESI, C., KUI DENG, H., MALNATI, M. S., PLEBANI, A., SICCARDI, A. G., LITTMAN, D. R., MARIA FENYO, E. & LUSSO, P. 1997. In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. *Nat Med,* 3**,** 1259-1265.

SCHALLER, T., OCWIEJA, K. E., RASAIYAAH, J., PRICE, A. J., BRADY, T. L., ROTH, S. L., HUÉ, S., FLETCHER, A. J., LEE, K., KEWALRAMANI, V. N., NOURSADEGHI, M., JENNER, R. G., JAMES, L. C., BUSHMAN, F. D. & TOWERS, G. J. 2011. HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathogens,* 7**,** e1002439.

SCHALLER, T., YLINEN, L. M., WEBB, B. L., SINGH, S. & TOWERS, G. J. 2007. Fusion of cyclophilin A to Fv1 enables cyclosporine-sensitive restriction of human and feline immunodeficiency viruses. *J Virol,* 81.

SCHLIEPHAKE, A. W. & RETHWILM, A. 1994. Nuclear localization of foamy virus Gag precursor protein. *Journal of Virology,* 68**,** 4946-4954.

SCHRODER, A. R., SHINN, P., CHEN, H., BERRY, C., ECKER, J. R. & BUSHMAN, F. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell,* 110**,** 521-9.

SCHUBERT, H. L., ZHAI, Q., SANDRIN, V., ECKERT, D. M., GARCIA-MAYA, M., SAUL, L., SUNDQUIST, W. I., STEINER, R. A. & HILL, C. P. 2010. Structural and functional studies on the extracellular domain of BST2/tetherin in reduced and oxidized conformations. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 17951-17956.

SCHULTZ, D. C., AYYANATHAN, K., NEGOREV, D., MAUL, G. G. & RAUSCHER, F. J. 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & Development,* 16**,** 919-932.

SCHULTZ, D. C., FRIEDMAN, J. R. & RAUSCHER, F. J. 2001. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2α subunit of NuRD. *Genes & Development,* 15**,** 428-443.

SCHWARTZ, S., FELBER, B. K., FENYÖ, E. M. & PAVLAKIS, G. N. 1990. Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *Journal of Virology,* 64**,** 5448-5456.

SCHWEFEL, D., BOUCHERIT, VIRGINIE C., CHRISTODOULOU, E., WALKER, PHILIP A., STOYE, JONATHAN P., BISHOP, KATE N. & TAYLOR, IAN A. 2015. Molecular Determinants for Recognition of Divergent SAMHD1 Proteins by the Lentiviral Accessory Protein Vpx. *Cell Host & Microbe,* 17**,** 489-499.

SCOTT, A., CHUNG, H.-Y., GONCIARZ-SWIATEK, M., HILL, G. C., WHITBY, F. G., GASPAR, J., HOLTON, J. M., VISWANATHAN, R., GHAFFARIAN, S., HILL, C. P. & SUNDQUIST, W. I. 2005. Structural and mechanistic studies of VPS4 proteins. *The EMBO Journal,* 24**,** 3658-3669.

SEGURA-MORALES, C., PESCIA, C., CHATELLARD-CAUSSE, C., SADOUL, R., BERTRAND, E. & BASYUK, E. 2005. Tsg101 and Alix Interact with Murine

Leukemia Virus Gag and Cooperate with Nedd4 Ubiquitin Ligases during Budding. *Journal of Biological Chemistry,* 280**,** 27004-27012.

SERRA-MORENO, R., ZIMMERMANN, K., STERN, L. J. & EVANS, D. T. 2013. Tetherin/BST-2 Antagonism by Nef Depends on a Direct Physical Interaction between Nef and Tetherin, and on Clathrin-mediated Endocytosis. *PLoS Pathog,* 9**,** e1003487.

SERRAO, E., KRISHNAN, L., SHUN, M.-C., LI, X., CHEREPANOV, P., ENGELMAN, A. & MAERTENS, G. N. 2014. Integrase residues that determine nucleotide preferences at sites of HIV-1 integration: implications for the mechanism of target DNA binding. *Nucleic Acids Research,* 42**,** 5164-5176.

SETTE, P., JADWIN, J. A., DUSSUPT, V., BELLO, N. F. & BOUAMR, F. 2010. The ESCRT-Associated Protein Alix Recruits the Ubiquitin Ligase Nedd4-1 To Facilitate HIV-1 Release through the LYPX(n)L L Domain Motif. *Journal of Virology,* 84**,** 8181-8192.

SHARMA, A., LARUE, R. C., PLUMB, M. R., MALANI, N., MALE, F., SLAUGHTER, A., KESSL, J. J., SHKRIABAI, N., COWARD, E., AIYER, S. S., GREEN, P. L., WU, L., ROTH, M. J., BUSHMAN, F. D. & KVARATSKHELIA, M. 2013. BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proc Natl Acad Sci U S A,* 110**,** 12036-41.

SHEEHY, A. M., GADDIS, N. C., CHOI, J. D. & MALIM, M. H. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature,* 418**,** 646-650.

SHEHU-XHILAGA, M., CROWE, S. M. & MAK, J. 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J Virol,* 75**,** 1834-41.

SHEN, C. H., TIE, Y., YU, X., WANG, Y. F., KOVALEVSKY, A. Y., HARRISON, R. W. & WEBER, I. T. 2012. Capturing the reaction pathway in near-atomic-resolution crystal structures of HIV-1 protease. *Biochemistry,* 51**,** 7726-32.

SHEN, Q.-T., SCHUH, A. L., ZHENG, Y., QUINNEY, K., WANG, L., HANNA, M., MITCHELL, J. C., OTEGUI, M. S., AHLQUIST, P., CUI, Q. & AUDHYA, A. 2014. Structural analysis and modeling reveals new mechanisms governing ESCRT-III spiral filament assembly. *The Journal of Cell Biology,* 206**,** 763-777.

SHI, J., FRIEDMAN, D. B. & AIKEN, C. 2013. Retrovirus restriction by TRIM5 proteins requires recognition of only a small fraction of viral capsid subunits. *J Virol,* 87**,** 9271-8.

SHOEMAKER, C., GOFF, S., GILBOA, E., PASKIND, M., MITRA, S. W. & BALTIMORE, D. 1980. Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retrovirus integration. *Proceedings of the National Academy of Sciences of the United States of America,* 77**,** 3932-3936.

SHORT, K. M. & COX, T. C. 2006. Subclassification of the RBCC/TRIM superfamily reveals a novel motif necessary for microtubule binding. *J Biol Chem,* 281**,** 8970-80.

SHORT, K. M., HOPWOOD, B., YI, Z. & COX, T. C. 2002. MID1 and MID2 homo- and heterodimerise to tether the rapamycin-sensitive PP2A regulatory subunit, alpha 4, to microtubules: implications for the clinical variability of X-linked Opitz GBBB syndrome and other developmental disorders. *BMC Cell Biol,* 3**,** 4.

SIRONI, M., BIASIN, M., CAGLIANI, R., GNUDI, F., SAULLE, I., IBBA, S., FILIPPI, G., YAHYAEI, S., TRESOLDI, C., RIVA, S., TRABATTONI, D., DE GIOIA, L., LO CAPUTO, S., MAZZOTTA, F., FORNI, D., PONTREMOLI, C., PINEDA, J. A., POZZOLI, U., RIVERO-JUAREZ, A., CARUZ, A. & CLERICI, M. 2014. Evolutionary Analysis Identifies an MX2 Haplotype Associated with Natural Resistance to HIV-1 Infection. *Molecular Biology and Evolution,* 31**,** 2402-2414.

SKRIPKIN, E., PAILLART, J. C., MARQUET, R., EHRESMANN, B. & EHRESMANN, C. 1994. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 4945-4949.

SONG, B., GOLD, B., O'HUIGIN, C., JAVANBAKHT, H., LI, X., STREMLAU, M., WINKLER, C., DEAN, M. & SODROSKI, J. 2005a. The B30.2(SPRY) Domain of the Retroviral Restriction Factor TRIM5α Exhibits Lineage-Specific Length and Sequence Variation in Primates. *Journal of Virology,* 79**,** 6111-6121.

SONG, B., GOLD, B., O'HUIGIN, C., JAVANBAKHT, H., LI, X. & STREMLAU, M. 2005b. The B30.2(SPRY) domain of the retroviral restriction factor TRIM5alpha exhibits lineage-specific length and sequence variation in primates. *J Virol,* 79.

SONG, S., GE, Q., WANG, J., CHEN, H., TANG, S., BI, J., LI, X., XIE, Q. & HUANG, X. 2011. TRIM-9 functions in the UNC-6/UNC-40 pathway to regulate ventral guidance. *J Genet Genomics,* 38**,** 1-11.

SOWD, G. A., SERRAO, E., WANG, H., WANG, W., FADEL, H. J., POESCHLA, E. M. & ENGELMAN, A. N. 2016. A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. *Proc Natl Acad Sci USA,* 113.

STAVROU, S., NITTA, T., KOTLA, S., HA, D., NAGASHIMA, K., REIN, A. R., FAN, H. & ROSS, S. R. 2013. Murine leukemia virus glycosylated Gag blocks apolipoprotein B editing complex 3 and cytosolic sensor access to the reverse transcription complex. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 9078-9083.

STEINMETZ, M. O., JELESAROV, I., MATOUSEK, W. M., HONNAPPA, S., JAHNKE, W., MISSIMER, J. H., FRANK, S., ALEXANDRESCU, A. T. & KAMMERER, R. A. 2007. Molecular basis of coiled-coil formation. *Proceedings of the National Academy of Sciences,* 104**,** 7062-7067.

STEVENS, A., BOCK, M., ELLIS, S., LETISSIER, P., BISHOP, K. N., YAP, M. W., TAYLOR, W. & STOYE, J. P. 2004. Retroviral Capsid Determinants of Fv1 NB and NR Tropism. *Journal of Virology,* 78**,** 9592-9598.

STOYE, J. P. 1998. Fv1, the mouse retrovirus resistance gene. *Rev Sci Tech,* 17**,** 269-77.

STRACK, B., CALISTRI, A., CRAIG, S., POPOVA, E. & GÖTTLINGER, H. G. 2003. AIP1/ALIX Is a Binding Partner for HIV-1 p6 and EIAV p9 Functioning in Virus Budding. *Cell,* 114**,** 689-699.

STREMLAU, M., OWENS, C. M., PERRON, M. J., KIESSLING, M., AUTISSIER, P. & SODROSKI, J. 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature,* 427**,** 848-53.

STREMLAU, M., PERRON, M., LEE, M., LI, Y., SONG, B., JAVANBAKHT, H., DIAZ-GRIFFERO, F., ANDERSON, D. J., SUNDQUIST, W. I. & SODROSKI, J. 2006. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci U S A,* 103**,** 5514-9.

STREMLAU, M., PERRON, M., WELIKALA, S. & 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**,** 3139-45.

SVEDA, M. M. & SOEIRO, R. 1976. Host restriction of Friend leukemia virus: synthesis and integration of the provirus. *Proc Natl Acad Sci U S A,* 73**,** 2356-60.

TANAKA, Y., MARUSAWA, H., SENO, H., MATSUMOTO, Y., UEDA, Y., KODAMA, Y., ENDO, Y., YAMAUCHI, J., MATSUMOTO, T., TAKAORI-KONDO, A., IKAI, I. & CHIBA, T. 2006. Anti-viral protein APOBEC3G is induced by interferon-α stimulation in human hepatocytes. *Biochemical and Biophysical Research Communications,* 341**,** 314-319.

TANG, C., JI, X., WU, L. & XIONG, Y. 2015. Impaired dNTPase Activity of SAMHD1 by Phosphomimetic Mutation of Thr-592. *Journal of Biological Chemistry,* 290**,** 26352-26359.

TEDBURY, P. R., ABLAN, S. D. & FREED, E. O. 2013. Global Rescue of Defects in HIV-1 Envelope Glycoprotein Incorporation: Implications for Matrix Structure. *PLoS Pathogens,* 9**,** e1003739.

TEDBURY, P. R. & FREED, E. O. 2014. The role of matrix in HIV-1 envelope glycoprotein incorporation. *Trends in Microbiology,* 22**,** 372-378.

TERSMETTE, M., DE GOEDE, R. E. Y., LANGE, J. M. A., DE WOLF, F., EEFTINK-SCHATTENKERK, J. K. M., SCHELLEKENS, P. T. A., COUTINHO, R. A., GOUDSMIT, J., HUISMAN, J. G. & MIEDEMA, F. 1989. Originally published as Volume 1, Issue 8645ASSOCIATION BETWEEN BIOLOGICAL PROPERTIES OF HUMAN IMMUNODEFICIENCY VIRUS VARIANTS AND RISK FOR AIDS AND AIDS MORTALITY. *The Lancet,* 333**,** 983-985.

TOBALY-TAPIERO, J., BITTOUN, P., GIRON, M.-L., NEVES, M., KOKEN, M., SAïB, A. & DE THÉ, H. 2001. Human Foamy Virus Capsid Formation Requires an Interaction Domain in the N Terminus of Gag. *Journal of Virology,* 75**,** 4367-4375.

TOBALY-TAPIERO, J., BITTOUN, P., LEHMANN-CHE, J., DELELIS, O., GIRON, M.-L., DE THÉ, H. & SAÏB, A. 2008. Chromatin Tethering of Incoming Foamy Virus by the Structural Gag Protein. *Traffic,* 9**,** 1717-1727.

TOROK, M. & ETKIN, L. D. 2001. Two B or not two B? Overview of the rapidly expanding B-box family of proteins. *Differentiation,* 67**,** 63-71.

TOWERS, G., BOCK, M., MARTIN, S., TAKEUCHI, Y., STOYE, J. P. & DANOS, O. 2000. A conserved mechanism of retrovirus restriction in mammals. *Proc Natl Acad Sci U S A,* 97**,** 12295-9.

TOWERS, G., COLLINS, M. & TAKEUCHI, Y. 2002. Abrogation of Ref1 retrovirus restriction in human cells. *J Virol,* 76**,** 2548-50.

TRAUTZ, B., PIERINI, V., WOMBACHER, R., STOLP, B., CHASE, A. J., PIZZATO, M. & FACKLER, O. T. 2016. Antagonism of the SERINC5 Particle Infectivity Restriction by HIV-1 Nef Involves Counteraction of Virion-associated Pools of the Restriction Factor. *Journal of Virology*.

TSUCHIHASHI, Z. & BROWN, P. O. 1994. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *Journal of Virology,* 68**,** 5863-5870.

TUTTLEMAN, J. S., POURCEL, C. & SUMMERS, J. 1986. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. *Cell,* 47**,** 451-460.

UCHIL, P. D., QUINLAN, B. D., CHAN, W. T., LUNA, J. M. & MOTHES, W. 2008. TRIM E3 ligases interfere with early and late stages of the retroviral life cycle. *PLoS Pathog,* 4**,** 0040016.

UGOLINI, S., MONDOR, I. & SATTENTAU, Q. J. 1999. HIV-1 attachment: another look. *Trends in Microbiology,* 7**,** 144-149.

USAMI, Y., WU, Y. & GÖTTLINGER, H. G. 2015. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature,* 526**,** 218-223.

VAN DER AA, L. M., LEVRAUD, J.-P., YAHMI, M., LAURET, E., BRIOLAT, V., HERBOMEL, P., BENMANSOUR, A. & BOUDINOT, P. 2009. A large new subset of TRIM genes highly diversified by duplication and positive selection in teleost fish. *BMC Biology,* 7**,** 7-7.

VAN ENGELENBURG, S. B., SHTENGEL, G., SENGUPTA, P., WAKI, K., JARNIK, M., ABLAN, S. D., FREED, E. O., HESS, H. F. & LIPPINCOTT-SCHWARTZ, J. 2014. Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. *Science (New York, N.Y.),* 343**,** 653-656.

VAN MAELE, B., BUSSCHOTS, K., VANDEKERCKHOVE, L., CHRIST, F. & DEBYSER, Z. 2006. Cellular co-factors of HIV-1 integration. *Trends in Biochemical Sciences,* 31**,** 98-105.

VINK, C., YEHESKIELY, E., VAN DER MAREL, G. A., VAN BOOM, J. H. & PLASTERK, R. H. 1991. Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. *Nucleic Acids Res,* 19**,** 6691-8.

VIRGEN, C. A., KRATOVAC, Z., BIENIASZ, P. D. & HATZIIOANNOU, T. 2008. Independent genesis of chimeric TRIM5-cyclophilin proteins in two primate species. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 3563-3568.

VIVÈS, R. R., IMBERTY, A., SATTENTAU, Q. J. & LORTAT-JACOB, H. 2005. Heparan Sulfate Targets the HIV-1 Envelope Glycoprotein gp120 Coreceptor Binding Site. *Journal of Biological Chemistry,* 280**,** 21353-21357.

VOLKOVA, N. A., FOMINA, E. G., SMOLNIKOVA, V. V., ZINOVIEVA, N. A. & FOMIN, I. K. 2014. The U3 Region of Moloney Murine Leukemia Virus Contains Position-independent Cis-acting Sequences Involved in the Nuclear Export of Full-length Viral Transcripts. *Journal of Biological Chemistry,* 289**,** 20158-20169.

WAGNER, A., DOERKS, A., ABOUD, M., ALONSO, A., TOKINO, T., FLÜGEL, R. M. & LÖCHELT, M. 2000. Induction of Cellular Genes Is Mediated by the Bel1 Transactivator in Foamy Virus-Infected Human Cells. *Journal of Virology,* 74**,** 4441-4447.

WAGNER, J. M., ROGANOWICZ, M. D., SKORUPKA, K., ALAM, S. L., CHRISTENSEN, D., DOSS, G., WAN, Y., FRANK, G. A., GANSER-PORNILLOS, B. K., SUNDQUIST, W. I. & PORNILLOS, O. 2016. Mechanism of B-box 2 domain-mediated higher-order assembly of the retroviral restriction factor TRIM5α. *eLife,* 5**,** e16309.

WAHEED, A. A. & FREED, E. O. 2010. The Role of Lipids in Retrovirus Replication. *Viruses,* 2**,** 1146-1180.

WAIN-HOBSON, S., SONIGO, P., DANOS, O., COLE, S. & ALIZON, M. 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell,* 40**,** 9-17.

WALKER, S. J., PIZZATO, M., TAKEUCHI, Y. & DEVEREUX, S. 2002. Heparin Binds to Murine Leukemia Virus and Inhibits Env-Independent Attachment and Infection. *Journal of Virology,* 76**,** 6909-6918.

WANG, L., WU, J., YUAN, J., ZHU, X., WU, H. & LI, M. 2016. Midline2 is overexpressed and a prognostic indicator in human breast cancer and promotes breast cancer cell proliferation in vitro and in vivo. *Frontiers of Medicine,* 10**,** 41-51.

WEI, P., GARBER, M. E., FANG, S.-M., FISCHER, W. H. & JONES, K. A. 1998. A Novel CDK9-Associated C-Type Cyclin Interacts Directly with HIV-1 Tat and Mediates Its High-Affinity, Loop-Specific Binding to TAR RNA. *Cell,* 92**,** 451-462.

WEI, W., GUO, H., MA, M., MARKHAM, R. & YU, X.-F. 2016. Accumulation of MxB/Mx2-resistant HIV-1 Capsid Variants During Expansion of the HIV-1 Epidemic in Human Populations. *EBioMedicine*.

WEINERT, C., GRUTTER, C., ROSCHITZKI-VOSER, H., MITTL, P. R. & GRUTTER, M. G. 2009. The crystal structure of human pyrin b30.2 domain: implications for mutations associated with familial Mediterranean fever. *J Mol Biol,* 394**,** 226-36.

WEINERT, C., MORGER, D., DJEKIC, A., GRÜTTER, M. G. & MITTL, P. R. E. 2015. Crystal structure of TRIM20 C-terminal coiled-coil/B30.2 fragment: implications for the recognition of higher order oligomers. *Scientific reports* [Online], 5. Available: http://europepmc.org/abstract/MED/26043233
http://europepmc.org/articles/PMC4455283?pdf=render
http://europepmc.org/articles/PMC4455283

http://dx.doi.org/10.1038/srep10819 [Accessed 2015].
WEISS, R. A. 2006. The discovery of endogenous retroviruses. *Retrovirology,* 3**,** 67-67.
WELKER, R., HOHENBERG, H., TESSMER, U., HUCKHAGEL, C. & KRÄUSSLICH, H.-G. 2000. Biochemical and Structural Analysis of Isolated Mature Cores of Human Immunodeficiency Virus Type 1. *Journal of Virology,* 74**,** 1168-1177.
WIEGERS, K., RUTTER, G., KOTTLER, H., TESSMER, U., HOHENBERG, H. & KRÄUSSLICH, H.-G. 1998. Sequential Steps in Human Immunodeficiency Virus Particle Maturation Revealed by Alterations of Individual Gag Polyprotein Cleavage Sites. *Journal of Virology,* 72**,** 2846-2854.
WIGHT, D. J., BOUCHERIT, V. C., NADER, M., ALLEN, D. J., TAYLOR, I. A. & BISHOP, K. N. 2012. The Gammaretroviral p12 protein has multiple domains that function during the early stages of replication. *Retrovirology,* 9**,** 1-20.
WILK, T., GEISELHART, V., FRECH, M., FULLER, S. D., FLÜGEL, R. M. & LÖCHELT, M. 2001. Specific Interaction of a Novel Foamy Virus Env Leader Protein with the N-Terminal Gag Domain. *Journal of Virology,* 75**,** 7995-8007.
WILLEY, R. L., SMITH, D. H., LASKY, L. A., THEODORE, T. S., EARL, P. L., MOSS, B., CAPON, D. J. & MARTIN, M. A. 1988. In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity. *Journal of Virology,* 62**,** 139-147.
WILLS, N. M., GESTELAND, R. F. & ATKINS, J. F. 1994. Pseudoknot-dependent read-through of retroviral gag termination codons: importance of sequences in the spacer and loop 2. *The EMBO Journal,* 13**,** 4137-4144.
WILSON, S. J., WEBB, B. L. J., YLINEN, L. M. J., VERSCHOOR, E., HEENEY, J. L. & TOWERS, G. J. 2008. Independent evolution of an antiviral TRIMCyp in rhesus macaques. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 3557-3562.
WLODAWER, A. & ERICKSON, J. W. 1993. Structure-Based Inhibitors of HIV-1 Protease. *Annual Review of Biochemistry,* 62**,** 543-585.
WLODAWER, A., GUSTCHINA, A., RESHETNIKOVA, L., LUBKOWSKI, J., ZDANOV, A., HUI, K. Y., ANGLETON, E. L., FARMERIE, W. G., GOODENOW, M. M., BHATT, D., ZHANG, L. & DUNN, B. M. 1995. Structure of an inhibitor complex of the proteinase from feline immunodeficiency virus. *Nat Struct Mol Biol,* 2**,** 480-488.
WLODAWER, A., MILLER, M., JASKOLSKI, M., SATHYANARAYANA, B., BALDWIN, E., WEBER, I., SELK, L., CLAWSON, L., SCHNEIDER, J. & KENT, S. 1989. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science,* 245**,** 616-621.
WOLF, D. & GOFF, S. P. 2009. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature,* 458**,** 1201-1204.
WOLF, D., HUG, K. & GOFF, S. P. 2008. TRIM28 mediates primer binding site-targeted silencing of Lys1,2 tRNA-utilizing retroviruses in embryonic cells. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 12521-12526.
WOLLERT, T. & HURLEY, J. H. 2010. Molecular Mechanism of Multivesicular Body Biogenesis by ESCRT Complexes. *Nature,* 464**,** 864-869.
WU, X., ANDERSON, J. L., CAMPBELL, E. M., JOSEPH, A. M. & 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**,** 7465-70.
WYMA, D. J., JIANG, J., SHI, J., ZHOU, J., LINEBERGER, J. E., MILLER, M. D. & AIKEN, C. 2004. Coupling of Human Immunodeficiency Virus Type 1 Fusion to Virion Maturation: a Novel Role of the gp41 Cytoplasmic Tail. *Journal of Virology,* 78**,** 3429-3435.

XU, L., YANG, L., MOITRA, P. K., HASHIMOTO, K., RALLABHANDI, P., KAUL, S., MERONI, G., JENSEN, J. P., WEISSMAN, A. M. & D'ARPA, P. 2003. BTBD1 and BTBD2 colocalize to cytoplasmic bodies with the RBCC/tripartite motif protein, TRIM5delta. *Exp Cell Res,* 288**,** 84-93.

YAMAUCHI, K., WADA, K., TANJI, K., TANAKA, M. & KAMITANI, T. 2008. Ubiquitination of E3 ubiquitin ligase TRIM5 alpha and its potential role. *Febs J,* 275**,** 1540-55.

YAN, N. & CHEN, Z. J. 2012. Intrinsic Antiviral Immunity. *Nature immunology,* 13**,** 214-222.

YAN, Y., BUCKLER-WHITE, A., WOLLENBERG, K. & KOZAK, C. A. 2009. Origin, antiviral function and evidence for positive selection of the gammaretrovirus restriction gene Fv1 in the genus Mus. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 3259-3263.

YANG, H., JI, X., ZHAO, G., NING, J., ZHAO, Q., AIKEN, C., GRONENBORN, A. M., ZHANG, P. & XIONG, Y. 2012. Structural insight into HIV-1 capsid recognition by rhesus TRIM5α. *Proceedings of the National Academy of Sciences,* 109**,** 18372-18377.

YAP, M. W., COLBECK, E., ELLIS, S. A. & STOYE, J. P. 2014. Evolution of the Retroviral Restriction Gene Fv1: Inhibition of Non-MLV Retroviruses. *PLoS Pathogens,* 10**,** e1003968.

YAP, M. W., DODDING, M. P. & STOYE, J. P. 2006. Trim-cyclophilin A fusion proteins can restrict human immunodeficiency virus type 1 infection at two distinct phases in the viral life cycle. *J Virol,* 80**,** 4061-7.

YAP, M. W., LINDEMANN, D., STANKE, N., REH, J., WESTPHAL, D., HANENBERG, H., OHKURA, S. & STOYE, J. P. 2008. Restriction of Foamy Viruses by Primate Trim5α. *Journal of Virology,* 82**,** 5429-5439.

YAP, M. W., MORTUZA, G. B., TAYLOR, I. A. & STOYE, J. P. 2007. The design of artificial retroviral restriction factors. *Virology,* 365**,** 302-14.

YAP, M. W., NISOLE, S., LYNCH, C. & STOYE, J. P. 2004. Trim5alpha protein restricts both HIV-1 and murine leukemia virus. *Proc Natl Acad Sci U S A,* 101**,** 10786-91.

YAP, M. W., NISOLE, S. & STOYE, J. P. 2005. A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Curr Biol,* 15**,** 73-8.

YEAGER, M., WILSON-KUBALEK, E. M., WEINER, S. G., BROWN, P. O. & REIN, A. 1998. Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: Implications for retroviral assembly mechanisms. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 7299-7304.

YEUNG, M. L., HOUZET, L., YEDAVALLI, V. S. R. K. & JEANG, K.-T. 2009. A Genome-wide Short Hairpin RNA Screening of Jurkat T-cells for Human Proteins Contributing to Productive HIV-1 Replication. *Journal of Biological Chemistry,* 284**,** 19463-19473.

YIN, Z., SHI, K., BANERJEE, S., PANDEY, K. K., BERA, S., GRANDGENETT, D. P. & AIHARA, H. 2016. Crystal structure of the Rous sarcoma virus intasome. *Nature,* 530**,** 362-366.

YOO, S., MYSZKA, D. G., YEH, C.-Y., MCMURRAY, M., HILL, C. P. & SUNDQUIST, W. I. 1997. Molecular recognition in the HIV-1 capsid/cyclophilin A complex1. *Journal of Molecular Biology,* 269**,** 780-795.

YU, C. Q., NA, L., LV, X. L., LIU, J. D., LIU, X. M., JI, F., ZHENG, Y. H., DU, H. L., KONG, X. G. & ZHOU, J. H. 2013. The TRIMCyp genotype in four species of macaques in China. *Immunogenetics,* 65**,** 185-93.

YU, S. F., BALDWIN, D. N., GWYNN, S. R., YENDAPALLI, S. & LINIAL, M. L. 1996a. Human Foamy Virus Replication: A Pathway Distinct from That of Retroviruses and Hepadnaviruses. *Science,* 271**,** 1579-1582.

YU, S. F., BALDWIN, D. N., GWYNN, S. R., YENDAPALLI, S. & LINIAL, M. L. 1996b. Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. *Science,* 271**,** 1579-82.

YU, S. F., EDELMANN, K., STRONG, R. K., MOEBES, A., RETHWILM, A. & LINIAL, M. L. 1996c. The carboxyl terminus of the human foamy virus Gag protein contains separable nucleic acid binding and nuclear transport domains. *Journal of Virology,* 70**,** 8255-62.

YU, X., YUAN, X., MCLANE, M. F., LEE, T. H. & ESSEX, M. 1993. Mutations in the cytoplasmic domain of human immunodeficiency virus type 1 transmembrane protein impair the incorporation of Env proteins into mature virions. *Journal of Virology,* 67**,** 213-221.

YUAN, B., CAMPBELL, S., BACHARACH, E., REIN, A. & 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.

YUDINA, Z., ROA, A., JOHNSON, R., BIRIS, N., DE SOUZA ARANHA VIEIRA, D. A., TSIPERSON, V., RESZKA, N., TAYLOR, A. B., HART, P. J., DEMELER, B., DIAZ-GRIFFERO, F. & IVANOV, D. N. 2015. RING dimerization links higher-order assembly of TRIM5α to synthesis of K63-linked polyubiquitin. *Cell Reports,* 12**,** 788-797.

YUFENYUY, E. L. & AIKEN, C. 2013. The NTD-CTD intersubunit interface plays a critical role in assembly and stabilization of the HIV-1 capsid. *Retrovirology,* 10**,** 1742-4690.

ZAITSEVA, L., CHEREPANOV, P., LEYENS, L., WILSON, S. J., RASAIYAAH, J. & FASSATI, A. 2009. HIV-1 exploits importin 7 to maximize nuclear import of its DNA genome. *Retrovirology,* 6**,** 1742-4690.

ZAPP, M. L. & GREEN, M. R. 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature,* 342**,** 714-716.

ZHANG, F., WILSON, S. J., LANDFORD, W. C., VIRGEN, B., GREGORY, D., JOHNSON, M. C., MUNCH, J., KIRCHHOFF, F., BIENIASZ, P. D. & HATZIIOANNOU, T. 2009. Nef Proteins from Simian Immunodeficiency Viruses Are Tetherin Antagonists. *Cell Host & Microbe,* 6**,** 54-67.

ZHAO, G., PERILLA, J. R., YUFENYUY, E. L., MENG, X., CHEN, B., NING, J., AHN, J., GRONENBORN, A. M., SCHULTEN, K., AIKEN, C. & ZHANG, P. 2013. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature,* 497**,** 643-6.

ZHOU, L., SOKOLSKAJA, E., JOLLY, C., JAMES, W., COWLEY, S. A. & FASSATI, A. 2011. Transportin 3 Promotes a Nuclear Maturation Step Required for Efficient HIV-1 Integration. *PLoS Pathogens,* 7**,** e1002194.

ZHOU, X., VINK, M., KLAVER, B., BERKHOUT, B. & DAS, A. T. 2006. Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther,* 13**,** 1382-1390.

ZHU, C., GAO, W., ZHAO, K., QIN, X., ZHANG, Y., PENG, X., ZHANG, L., DONG, Y., ZHANG, W., LI, P., WEI, W., GONG, Y. & YU, X.-F. 2013. Structural insight into dGTP-dependent activation of tetrameric SAMHD1 deoxynucleoside triphosphate triphosphohydrolase. *Nat Commun,* 4.