



Vera Ortega, Walter (2018) *Approaching a Tat-Rev independent HIV-1 clone towards a model for research*. PhD thesis.

<https://theses.gla.ac.uk/30755/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>  
[research-enlighten@glasgow.ac.uk](mailto:research-enlighten@glasgow.ac.uk)

# Approaching a Tat-Rev Independent HIV-1 Clone Towards a Model for Research

Walter Vera Ortega



Submitted in fulfillment of the requirements for the Degree of

Doctor of Philosophy in Virology

Institute of Infection Immunity & Inflammation College of Medical, Veterinary and Life  
Sciences University of Glasgow

MRC - Centre for Virus Research

**August, 2018**

## Abstract

Human immunodeficiency virus type 1 (HIV-1) is responsible for the acquired immunodeficiency syndrome (AIDS), a leading cause of death worldwide infecting millions of people each year. Despite intensive research in vaccine development, therapies against HIV-1 infection are not curative and the huge genetic variability of HIV-1 challenges drug development. Current animal models for HIV-1 research present important limitations, impairing the progress of *in vivo* approaches. Macaques require CD8+ depletion or large portions of the genome to be replaced by sequences derived from simian immunodeficiency viruses to progress to AIDS, and the maintenance cost is high. Mice are a cheaper alternative, but need to be 'humanized' and breeding is not possible and knockout experiments are difficult.

The development of an HIV-1 clone able to replicate in mice is a challenging proposal. The lack of human co-factors in mice impedes function of the HIV-1 accessory proteins Tat and Rev, hampering HIV-1 replication. The Tat and Rev function can be replaced by constitutive/chimeric promoters, codon-optimized genes and the constitutive transport element (CTE), generating a novel HIV-1 clone able to replicate in mice without disrupting the amino acid sequence of the virus. By minimally manipulating the genomic 'identity' of the virus, we propose the generation of an HIV-1 clone able to replicate in mice to assist in antiviral drug development. My results have determined that murine NIH 3T3 cells are able to generate pseudotyped HIV-1 particles, but they are not infectious. Codon-optimized HIV-1 constructs are efficiently made in human HEK-293T cells in a Tat and Rev independent manner and capable of packaging a competent genome *in trans*. CSGW (an HIV-1 vector genome) efficiently generates infectious particles in the absence of Tat and Rev in human cells when 4 copies of the CTE are placed preceding the 3'LTR. HIV-1 replication competent genomes lacking *tat* expression and encoding different promoters are functional during the first cycle of replication when Tat is added *in trans*. Finally, we developed a replication competent HIV-1 clone lacking *tat* and *rev* genes and encoding 4CTEs that could be a future candidate for HIV research.

My results shown that the development of an HIV-1 Tat-Rev independent clone could become a candidate for HIV research in a near future, but further investigations are necessary before proposing our model as an alternative yet.

## Table of contents

<b>Abstract.....</b>	<b>I</b>
<b>Table of contents.....</b>	<b>II</b>
<b>List of tables.....</b>	<b>VII</b>
<b>List of figures.....</b>	<b>VIII</b>
<b>Acknowledgements.....</b>	<b>XII</b>
<b>Declaration.....</b>	<b>XIII</b>
<b>List of abbreviations.....</b>	<b>XIV</b>
<b>1. Chapter 1:Introduction.....</b>	<b>1</b>
1.1 Retroviruses.....	2
1.1.1 Classification.....	2
1.1.2 Genera.....	4
1.1.2.1 Lentiviruses.....	4
<b>1.2 Genome organization and virion characteristics.....</b>	<b>6</b>
<b>1.3 HIV-1 replication cycle.....</b>	<b>8</b>
1.3.1 Attachment, entry and uncoating.....	8
1.3.2 Reverse transcription.....	11
1.3.3 Nuclear import and integration.....	14
1.3.4 Transcription, translation and splicing.....	15
1.3.5 Packaging, assembly and release.....	16
<b>1.4 HIV-1 genome structure.....</b>	<b>19</b>
1.4.1 HIV-1 Long Terminal Repeat (LTR).....	20
<b>1.5 Origin and evolution of HIV-1.....</b>	<b>22</b>
<b>1.6 Regulatory proteins of HIV-1.....</b>	<b>25</b>
1.6.1 Trans-activator of transcription (Tat) and the Tat-TAR complex.....	25
1.6.2 Differences between human cyclinT1 (hCycT1) and murine cyclinT1 (mCycT1)...	26

1.6.3 Regulator of the expression of the viral proteins (Rev) and the Rev Response Element (RRE).....	28
1.6.4 Chromosomal maintenance 1 (CRM1) .....	29
<b>1.7 Rev-like proteins.....</b>	<b>32</b>
1.7.1 Constitutive transport element (CTE) .....	32
1.7.2 DRI and DR2 elements of avian retroviruses.....	33
1.7.3 Jaagsietke sheep retrovirus (JSRV).....	33
1.7.4 Regulator of export of mouse mammary tumor (MMTV).....	33
1.7.5 HERV-K.....	34
1.7.6 FIV Rev.....	34
1.7.7 HTLV.....	34
1.7.2 Regulator of export of mouse mammary tumor virus (MMTV) .....	35
<b>1.8 Antiretroviral drugs.....</b>	<b>37</b>
<b>1.9 HIV-1 animal models for research.....</b>	<b>37</b>
1.9.1 Humanized mice.....	37
1.9.2 Macaques and SHIVs.....	39
<b>1.10 Lentiviral vectors and self inactivating vectors (SIN) .....</b>	<b>42</b>
<b>2. Materials and methods.....</b>	<b>45</b>
<b>2.1 Cell culture.....</b>	<b>45</b>
2.1.1 Bacterial strains.....	45
2.1.2 Antibiotics.....	45
2.1.3 Eukaryotic cell lines.....	45
<b>2.2 Molecular biology.....</b>	<b>47</b>
2.2.1 Oligonucleotides.....	47
2.2.2 Enzymes.....	47
2.2.3 Cloning.....	48
2.2.4 Plasmids.....	51
2.2.5 Antibodies.....	51
<b>2.3 Chemicals and buffers.....</b>	<b>58</b>

2.3.1 Cell culture .....	59
2.2.2 Protein analysis.....	59
2.3.2 DNA analysis.....	59
2.3.3 RNA extraction.....	60
2.3.4 Western blot reagents.....	60
<b>2.4 Methods.....</b>	<b>60</b>
2.4.1 Cell culture.....	61
2.4.2. Transfection of human and murine cells .....	61
2.4.3 Assays for viral particle production and infectivity.....	61
2.4.4 Cloning.....	61
2.4.4.1 NHG clones.....	62
2.4.4.2 SYNGP clones.....	62
2.4.4.3 CSGW-flexi clones.....	63
2.4.4.4 CSGW_GP clones.....	63
2.4.4.5 DEST40 clones.....	64
2.4.4.6 MLV based retroviral vectors for <i>ex vivo</i> gene transfer.....	64
2.4.5 Ligation reactions.....	64
2.4.6 Bacteria transformation.....	65
2.4.6.1 Chemically competent <i>Escherichia coli</i> DH10B bacteria preparation.....	65
2.4.7 Miniprep cultures, DNA extraction and restriction digestion.....	65
2.4.8 Midiprep cultures.....	66
2.4.9 RNA extraction and purification .....	67
2.4.10 Reverse transcription and qPCR assays.....	67
2.4.11 PCR.....	67
2.4.12 Agarose gel electrophoresis.....	68
2.4.13 Western blot analysis.....	68
2.4.14 TZM-bl assays.....	69
2.4.15 Cell preparation for flow cytometric analysis analysis.....	70
<b>3. Chapter 3.....</b>	<b>71</b>

<b>3.1 Results.....</b>	<b>72</b>
3.1.1 CLTR and MLTR chimeric promoters support transcription and reverse transcription..	72
3.1.2 A 118 base pair fragment preceding the 3' end of CSGW is involved in viral infectivity.....	78
3.1.3 Viral particles derived from SYNGP clones are efficiently made in human and murine cells. Codon optimization and chimeric CLTR facilitate Tat and Rev independent VLP production from human and murine cells.....	83
3.1.4 Introducing the HIV-1 packaging signal ( $\Psi$ ) in SYNGP clones did not increase the efficiency of particle production.....	87
3.1.5 DEST40-tat drives Tat expression efficiently in TZM-bl cells.....	91
3.1.6 HIV-1 codon-optimized VLPs are infectious in a Rev independent manner when CSGW and DEST40_ <i>tat</i> are provided <i>in trans</i> .....	93
<b>3.2 Discussion chapter 3.....</b>	<b>97</b>
<b>4. Chapter 4.....</b>	<b>100</b>
<b>4.1 Results.....</b>	<b>101</b>
4.1.1 HIV-1 codon-optimized particles containing CSGW_Flexi clones were infectious in human cells when Tat, Rev or both were provided in trans.....	101
4.1.2 CLTR could be better at driving transcription whereas the MLTR more efficiently supports reverse transcription.....	104
4.1.3 The CLTR is Tat responsive.....	108
4.1.4 SYNGP clones encoding 4CTEs were not infectious in the absence of Tat and Rev, but infectivity was rescued when they were provided <i>in trans</i> .....	111
<b>4.2 Discussion chapter 4.....</b>	<b>115</b>
<b>5. Chapter 5.....</b>	<b>119</b>
<b>5.1 Results.....</b>	<b>120</b>
5.1.1 CSGW containing 4 copies of the Constitutive Transport Element (CTE) from Mason Pfizer Monkey Virus (MPMV) is able to generate infectious particles in a Tat-Rev independent manner in human cells.....	120
5.1.2 CSGW_4CTE clones encoding codon-optimized <i>gag-pol</i> HIV-1 genes in <i>cis</i> efficiently generated Gag, but VLPs produced were not infectious.....	124

5.1.3 Codon-optimized HIV-1 <i>gag-pol</i> with reduced number of CpGs was unable to generate HIV-1 particles.....	128
<b>5.2 Discussion chapter 5.....</b>	<b>131</b>
<b>6. Chapter 6.....</b>	<b>134</b>
<b>6.1 Results.....</b>	<b>135</b>
6.1.1 NHG $\Delta$ <i>tat</i> mutants do not express Tat in TZM-bl cells.....	135
6.1.2 Tat expressed <i>in trans</i> rescues NHG- <i>tat</i> mutants. The CLTR and MLTR support reverse transcription efficiently.....	140
6.1.3 NHG_ $\Delta$ <i>tat</i> clones were able to complete cycles of replication in a Tat-independent manner.....	145
6.1.4 NHG- <i>tat-rev-4CTE</i> clones were able to complete subsequent cycles of replication in a Tat and Rev independent manner, but infectivity was lost along with passage.....	150
<b>6.2 Discussion chapter 6.....</b>	<b>155</b>
<b>7. Chapter 7: Final reflections.....</b>	<b>159</b>
<b>7.1 Final reflections.....</b>	<b>160</b>
7.1.1 Summary.....	160
7.1.2 Potential for future research.....	162
<b>8. Chapter 8: References.....</b>	<b>160</b>
<b>8.1 References.....</b>	<b>168</b>

**List of tables**

Table 1.1: Classification of Retroviruses .....	3
Table 2.1: Oligonucleotides .....	48
Table 2.2: Restriction enzymes .....	51
Table 2.3: Plasmids .....	51
Table 2.4: Transfection volumes .....	61

## List of figures

Figure 1-1. HIV-1 particles budding and released at the host plasma membrane. ....	4
Figure 1-2. HIV-1 virion composition. ....	7
Figure 1-3. HIV-1 fusion process .....	9
Figure 1-4 . HIV-1 uncoating mechanism.....	10
Figure 1-5. HIV-1 reverse transcription mechanism .....	13
Figure 1-6. HIV-1 replication cycle.....	18
Figure 1-7. HIV-1 genome organization and open reading frames .....	19
Figure 1-8. HIV-1 LTR structure .....	22
Figure 1-9. Evolution of primate lentiviruses.....	24
Figure 1-10. TAR-Tat interaction complex with human co-factors Cyclin-T1 and Cdk-9..	26
Figure 1-11. Amino acid sequence comparison of hCycT1 and mCycT1.....	27
Figure 1-12. Predicted secondary structure for RRE from HXB3 strain with the major stem loops labeled .....	28
Figure 1-13. RNA processing along HIV-1 replication cycle.....	31
Figure 1-14. Human immune system mouse models.....	38
Figure 1-15. Schematic representation of HIV-1, SIV and SHIVs viral genomes.....	41
Figure 2-1. Plasmids maps employed in this thesis .....	53
Figure 2-2. Experiment protocol to generate replication competent HIV-1 particles or single cycle of replication viruses (VLPs) in HEK-293T cells.....	62
Figure 2-3. Diagram of a standard TZM-bl assay.....	70
Figure 3-1. Diagram of pCSGW, pCVR1-NLGP and pCMV_VSV-G.....	73
Figure 3-2. HIV-1 LTR, CLTR and MLTR composition .....	73
Figure 3-3. Location of the additional <i>SpeI</i> site at the beginning of 5'LTR in CSGW_Flexi_ΔU3.....	75
Figure 3-4. Mean titre of CSGW clones in PM1 cells 48 h.p.i. WB of cell lysates from transfected HEK-293T cells 48 hours post-transfection. WB for capsid in VLPs collected from the supernatant of transfected HEK-293T cells 48 h.p.t.....	77
Figure 3-5. Diagram of HIV-1 Gag polyprotein.....	79
Figure 3-6. Mean titre of CSGW clones in PM1 cells 48 h.p.i. WB of cell lysates from transfected HEK-293T cells 48 h.p.t. WB of viral particles 48 h.p.t. from supernatants in transfected HEK-293T cells.....	82
Figure 3-7. Diagram of SYNGP clones encoding a CLTR at the 5'end, 3'end or both. WB for p24 (CA) of transfected human HEK-293T and murine NIH 3T3 cells 48 hours post-transfection. WB of VLPs.....	86

Figure 3-8. Location of the HIV-1 packaging signal within the viral genome .....	88
Figure 3-9. Diagram of SYNGP clones encoding CLTRs, the HIV-1 packaging signal ( $\Psi$ ) and an IRES_GFP. WB for p24 (CA) of HEK-293T (human) and NIH 3T3 (murine) cells 48 hours post-transfection. WB of VLPs.....	90
Figure 3-10. TZM-bl assay 48 hours post-transfection of DEST40- <i>tat</i> . FACS of transfected TZM-bl with pDEST40- <i>tat</i> and pCRV1-EGFP. Blue cells foci per field.....	92
Figure 3-11. Mean titre of SYNGP clones in AA-2 cells.....	93
Figure 3-12. Diagram of the SYNGP clones. FACS of target cells (PM1) 48 h.p.i.....	96
Figure 4-1. Diagram of DEST40_ <i>tat</i> and DEST40_ <i>rev</i> expression cassettes .....	101
Figure 4-2. Diagram of CVR1_ NLGP and SYNGP vectors .....	101
Figure 4-3. Diagram of CSGW_ Flexi clones .....	102
Figure 4-4. Mean titre of CSGW_ Flexi clones in AA-2 cells 48 h.p.i. generated with Tat, Rev or both supplied in <i>trans</i> .....	103
Figure 4-5. Mean titre of CSGW_ Flexi clones in AA-2 cells 48 h.p.i. adding only DEST40-Tat <i>in trans</i> . Western blots for p55/p24 in transfected HEK-293T and VLPs collected from the supernatant. Mean titre of CSGW_ Flexi clones in AA-2 cells 48 h.p.i. adding only DEST40-Rev <i>in trans</i> . Western blots for p55/p24 in transfected HEK-293T and VLPs collected from the supernatant. Mean titre of CSGW_ Flexi clones in AA-2 cells 48 h.p.i. adding DEST40-Tat and DEST40-Rev <i>in trans</i> . Western blots for p55/p24 in transfected HEK-293T and VLPs collected from the supernatant. Mean titre of CSGW_ Flexi clones in AA-2 cells 48 h.p.i. without providing Tat and Rev <i>in trans</i> . Western blots for p24 (CA) of VLPs.....	107
Figure 4-6. Diagram of SYNGP_ 4CTE clones.....	108
Figure 4-7. SYNGP clones encoding CLTRs, the HIV-1 packaging signal, codon-optimized HIV-1 <i>gag-pol</i> genes, an IRES-GFP cassette and 4 copies of the CTE. The envelope was provided <i>in trans</i> with pCMV-VSV-G to generate HIV-1 pseudotyped particles. Mean Fluorescence Intensity (MFI) in HEK-293T cells. Diagram of Tat responsiveness in the CLTR.....	110
Figure 4-8. Diagram of SYNGP_ 4CTE clones and DEST40_ <i>tat</i> and DEST40_ <i>rev</i> plasmids .....	111
Figure 4-9. Diagram of SYNGP_ 4CTE clones. Mean titre of SYNGP_ 4CTE clones in AA-2 cells 48 h.p.i. ....	113
Figure 4-10. Mean titre of SYNGP_ 4CTE clones in AA-2 cells 48 h.p.i. WB for p55/p24 from transfected HEK 293T and VLPs collected from the supernatant 48 h.p.i. ....	114
Figure 5-1. Diagram of CSGW and CSGW_ 4CTE .....	120

Figure 5-2. Diagram of CSGW_4CTE encoding a competent HIV-1 LTR (CSGW_4CTE_3LTR) .....	121
Figure 5-3. Murine RAW 264.7 cells infected with CSGW_4CTE viruses produced in HEK-293Ts. ....	122
Figure 5-4. CSGW clones including 4 copies of the CTE. Codon-optimized <i>gag-pol</i> genes are provided by pSYNGP and the envelope by pCMV-VSV-G. Mean titre of CSGW_4CTE clones in AA-2 cells. DEST40_ <i>tat</i> and DEST40_ <i>rev</i> were added <i>in trans</i> . WB from transfected HEK-293T and viral particle production 48 hours post transfection. Mean titre of CSGW_4CTE construct encoding a competent LTR at the 3' end (CSGW_4CTE_3LTR) .....	123
Figure 5-5. Diagram of CSGW_4CTE_3LTR vector encoding HIV-1 codon-optimized <i>gag-pol</i> genes (Flexi_GP_4CTE_3LTR) .....	124
Figure 5-6. Diagram of the CSGW_Flexi_GP clones encoding the HIV-1 <i>gag-pol</i> codon-optimized genes from pSYNGP, the CLTR/MLTR and 4 copies of the CTE. Mean titre of Flexi_GP clones in AA-2 cells. Western Blot for p55/p24 in transfected HEK-293T and VLPs collected from the supernatant.....	127
Figure 5-7. Diagram of CSGW_4CTE encoding HIV-1 codon-optimized <i>gag-pol</i> genes with reduced CpG number (Flexi_WV-GP_4CTE).....	128
Figure 5-8. Western Blot for p24 (CA) of transfected HEK-293Ts using an HIV-1 codon-optimized <i>gag-pol</i> with reduced CpGs and VLPs production. Western blot for p55 and p24 in HEK 293T. Mean titre of viruses encoding HIV-1 codon-optimized <i>gag-pol</i> with reduced CpG number .....	130
Figure 6-1. Diagram of NHG vector, encoding an HIV-1 replication competent genome. ....	135
Figure 6-2. Tat exon 1 location in pNHG and mutations performed in <i>tat</i> exon 1 .....	136
Figure 6-3. Diagram of NHG_ <i>Δtat-F</i> and NHG_ <i>Δtat-D</i> clones encoding a 3'LTR, CLTR and MLTR.....	137
Figure 6-4. Diagram of a standard TZM-bl assay. Transfection into TMZ-bl cells to detect Tat expression from NHG- <i>tat</i> mutants.....	139
Figure 6-5. Standard experiment to generate replication competent HIV-1 particles in human HEK-293T cells .....	140
Figure 6-6. Mean titre of NHG clones encoding <i>tat</i> in <i>cis</i> (left), NHG_ <i>Δtat-F</i> and NHG_ <i>Δtat-D</i> clones in PM1 cells. Western blot for p24 (CA) of transfected HEK-293T cells and viral particle production .....	143

Figure 6-7. TZM assay to detect Tat expression in our NHG- <i>Atat</i> mutants. GFP expression from infected TZM cells (48 h.p.i.) with NHG clones encoding <i>tat</i> in <i>cis</i> NHG- <i>Atat-F</i> and NHG- <i>Atat-D</i> clones.....	144
Figure 6-8. Growth curves of NHG_ <i>Atat-F</i> _3CLTR and NHG_ <i>Atat-D</i> _3CLTR viruses in AA-2. Red arrows indicate shutdown period in the CL3 facilities. The right curve represents the growth of a fresh virus after the shutdown. ....	146
Figure 6-9. HIV-1 adaptation in human AA-2 cells. ....	148
Figure 6-10. TZM cells inoculated with NHG_3LTR, NHG_3CLTR, NHG_3MLTR and NHG_ <i>Atat</i> viruses after 11 days of passage.....	149
Figure 6-11. NHG_ <i>Atat-D-Δrev</i> viruses encoding 4CTEs at the 3' end .....	151
Figure 6-12. Maps of the replication competent The NHG_ <i>Atat-D_Δrev_4CTE</i> viruses encoding a 43 nt deletion in <i>tat</i> exon 1 ( <i>Atat-D</i> ) and a deletion of the first exon of <i>rev</i> ( <i>Δrev</i> ), 4CTEs and chimeric LTRs. B) GFP expression of NHG- <i>Atat-D-Δrev_4CTE</i> viruses in AA-2 cells 4 d.p.i.....	153
Figure 6-13. GFP expression in NHG_ <i>Atat-D_Δrev_4CTE_3CLTR</i> virus 12 d.p.i.....	154

## Acknowledgements

I'd like to acknowledge my supervisor Sam Wilson for his advices and contribution for the development of this thesis. Thus, I'll be always grateful be have done a PhD in Glasgow University within the CVR, thanks for this amazing opportunity Glasgow Uni!. Firstly, I want to really thank Idoia, who helped me with no hesitation and made me feel as part of her family. She was always caring about me since the first day I arrived to Scotland, and also outside the lab during these three years. I also would like to thank Idoia for teaching me how science is done and I wish you the success in science you deserve. Moreover, I am also indebted with my 'canarión' Nico, my favourite greek character Kostas aka 'the beast' and my best friend from Talavera de la Reina Virginia, who helped me as well and I really had fun with all of you during my PhD. They were always there for helping me and I really appreciate their support and patience. I'll miss the coffee breaks, the laughs we shared together and the whisky tasters sessions!. I'll also miss Junjie and Tristan, really nice guys and friends. I remember when I firstly arrived to Glasgow and did not understand what you guys were saying to me, but we finally become very good friends and I hope our friendship remains forever despite the distance.

Special thanks must go to my parents Arturo and Ana, my grandparents Elda, Emiliano and Arturo, my brother Héctor, my lovely Sara and her parents also. They were always supporting me throughout my PhD despite the distance sparing my country from the UK. I'd like to acknowledge Sara for coming during one year to Scotland to support my project and make me feel much better about living in a foreign country alone and dealing with the pressure from my project. I'll never forget the huge effort you all made to deal the distance and how difficult it was since the beginning. Special thanks to my assessors Brian and Jim for his contribution and ideas along my PhD, the discussions about my work were really helpful. Thanks to my lab teammates for their help within the lab. Thanks to my ex-teammates in the Outlands basketball team, Ray, PJ, Yoann, Pablo and the rest of the guys who really helped me to disconnect from the lab and keep enjoying basketball. I'll really miss PJ's complaints for everything, and Ray's jokes. I'll never forget the people who recommended and helped me to perform a PhD. Thanks to Agustín Valenzuela and David Gillespie about giving me the opportunity to work in science and become a scientist. Thanks to all my friends from Tenerife who cheered me up everyday: Cristian, Sergio, Javi, Alexander, Bea and Lucas. Lastly, I just want to say thanks again and hope I can become a doctor and feel proud of my work and myself.

**Declaration**

Walter Vera Ortega, declare that, except where explicit reference is made to the contribution of others, that this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature

A black rectangular box redacting the signature of the author.

Printed name

Walter Vera Ortega

**List of abbreviations**

AIDS	acquired immunodeficiency syndrome
AAV	adeno-associated virus
APOBEC	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
BLT	humanized bone marrow liver thymus mice
BLV	bovine leukemia virus
CA	capsid
CCR5	CC chemokine receptor type 5
CD4	cluster of differentiation 4
CDK9	cyclin-dependent kinase 9
cDNA	complementary DNA
CDS	coding sequence
CIAP	calf intestinal alkaline phosphatase,
CMV	cytomegalovirus
CLTR	CMV-R-U5 chimeric promoter
CPE	cytopathic effect
CRM1	chromosomal maintenance 1
CTD	C-terminal domain
CTE	constitutive transport element
CXCR4	CXC chemokine receptor type 4
CycT1	cyclin T1
DCAF-1	DNA binding protein 1
DIS	dimerization site
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
dsRNA	double stranded RNA
EDTA	ethylenediaminetetraacetic acid
env	envelope gene
ERVs	endogenous retroviruses
ESCRT	endosomal sorting complexes required for transport

FCS	fetal calf serum
FeLV	feline leukemia virus
FIV	feline immunodeficiency virus
Gag	group specific antigen
GFP	green fluorescent protein
gp120	HIV-1 envelope glycoprotein
gp41	HIV-1 envelope glycoprotein
h.p.i.	hours post-infection
h.p.t.	hours post-transfection
HAART	highly active antiretroviral therapy
HERVs	human endogenous retroviruses
HIV-1	immunodeficiency virus type 1
HTLV	human T-cell leukemia virus
Imp	importin
IN	integrase
INBI	integrase binding inhibitor
INSTI	integrase strand transfer inhibitor
IFN	interferon
INS	inhibitory/instability RNA sequences
IRES	internal ribosomal entry site
LEDGF	lens epithelium-derived growth factor
LTR	long terminal repeat
MA	matrix
MIF	mean fluorescence intensity
MLV	murine leukemia virus
MLTR	MLV-R-U5 chimeric promoter
MMTV	mouse mammary tumor virus
MOPS	3-(N-morpholino)propanesulfonic acid
MPMV	Mason-Pfizer monkey virus
mRNA	messenger RNA
NC	nucleocapsid
ncRNA	non-coding RNA
Nef	negative factor
NES	nuclear export signal

NFAT	nuclear factor of activated T-cells
NIS	nuclear inhibitory signal
NLS	nuclear localization signal
NNRTI	non-nucleoside transcriptase inhibitor
NRTI	nucleoside analog reverse-transcriptase inhibitor
NtART	nucleotide analog reverse-transcriptase inhibitor
NPC	nuclear pore complex
Nt	nucleotide
Nup	nucleophorin
ORF	open reading frame
P-TEFb	human transcription elongation factor
PBS	primer binding site/ phosphate buffered saline
PCR	polymerase chain reaction
PEI	polyethylenimine
PIC	pre-integration complex
Pol	polymerase
PPT	polypurine track
PR	protease
qPCR	quantitative PCR
R	R region
RCF	relative centrifuge force (g)
RCL	replication competent lentivirus
Rev	regulator of the expression of the viral proteins
RFP	red fluorescent protein
RH	ribonuclease H
RNA	ribonucleic acid
Rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute
RSV	Rous sarcoma virus
RT	reverse transcriptase
RTI	reverse transcriptase inhibitor
SCID	severe combined immune deficiency
SD	splice donor
SDS	sodium dodecyl sulphate

SFFV	spleen focus-forming virus
SFV	simian foamy virus
SHIV	SIV/HIV chimera
SIN	self-inactivating vector
SIV	simian immunodeficiency virus
ssRNA	single stranded RNA
SU	surface protein
TAE	tris-acetic acid-EDTA
TAR	trans-activator response element
Tat	trans-activator protein
TCR	T-cell receptor
TM	trans membrane protein
TRIM5 $\alpha$	tripartite motif-containing protein 5
tRNA	transfer RNA
U3	unique to the 3' region
U5	unique to the 5' region
vif	viral infectivity factor
VLP	viral-like particle
Vpr	viral protein R
Vpu	viral protein U
Vpx	viral protein x
VSV	vesicular stomatitis virus
VSV-G	vesicular stomatitis virus G protein
WDSV	Walleye dermal sarcoma virus
WPRE	post-transcriptional regulatory element of Woodchuck Hepatitis virus
Wt	wild type
$\psi$	packaging signal

Science is a way of thinking much more than  
it is a body of knowledge

# **Chapter 1**

## **Introduction**

## 1. Introduction

### 1.1 Retroviruses

#### 1.1.1 Classification

The family *Retroviridae* comprises a diverse family of enveloped RNA viruses divided into seven genera: *Alpharetrovirus*, *Betaretrovirus*, *Gammaretrovirus*, *Deltaretrovirus*, *Epsilonretrovirus*, *Lentivirus* and *Spumavirus* (Table 1.1), which possess the ability to reverse transcribe a single stranded RNA (ssRNA) genome through a reverse transcription mechanism to generate double stranded DNA (dsDNA), which subsequently integrates into the host cell genome becoming a proviral DNA (Dahlberg 1988). The enzyme capable of reverse transcription is called reverse transcriptase (RT), and was discovered by David Baltimore (Baltimore 1970) and Howard M. Temin, awarded with the Nobel prize jointly with and Renato Dulbecco (which demonstrated that the infection of normal cells with some types of viruses (oncoviruses) were able to incorporate virus-derived genes into the host-cell genome, generating tumoral phenotypes in infected cells) in 1975 for their discoveries concerning the interaction between tumour viruses and the genetic material of the cell. The discovery of reverse transcriptase is one of the most important of the medicine discipline, which contributed to the generation of novel techniques in molecular biology.

Retroviruses can be divided into two different groups:

(i) Endogenous retroviruses (ERVs) are transmitted vertically through the germ line of the host, and comprise about 5%-8% of the human genome (Belshaw et al. 2004). A possible explanation for the origin of ERVs could be a result of exogenous retroviruses infecting the germ line cells at some point in the past, where ERVs can later become fixed in the gene pool of the host population and inherited by its offspring as a novel allele (vertical transmission). Analysis of sequenced genomes has revealed that between four and ten per cent of vertebrate DNA is made up of retroviral remnants (Stoye et al. 2012). The human ERVs (HERV) integration could severely affect the function of nearby genes, where the transcriptional orientation in relation to the host gene can act as a promoter or repressor of neighbouring genes (Jern et al. 2008), and HERVs that have been integrated within introns provide an alternative transcription initiation or termination sites, generating truncated host mRNA transcripts (Cohen et al. 2009). Several reports have described that CpG DNA

methylation participates in the transcriptional silencing of HERVs, and can also reverse the silencing of ERVs long terminal repeats (LTRs) (Isbel et al. 2012).

(ii) Exogenous retroviruses whose genetic information is sufficient for the generation of replication competent viruses capable of transmitting from one individual to another from the same specie (horizontal transmission). These retroviruses need to propagate through several individuals in order to survive and evolve along episodes of infection in novel infected hosts. The success of the infection of exogenous retroviruses is closely related to the virulence and pathogenicity within the host to complete the cycle of replication and spread in the infected host. One example of an exogenous retrovirus that can cause severe disease is the Human Immunodeficiency Virus 1 (HIV-1), the causative agent of Acquired Immunodeficiency Syndrome (AIDS).

Therefore, retrovirus research has led to the discovery of oncogenes and, among other things, has also enhanced our understanding of eukaryotic gene expression and cancer genetics (Vogt 2012). For example, the mouse mammary tumor virus (MMTV) was identified as the causative agent of mammary carcinoma in mice following milk transmission from mother to suckling offspring (Spiegelman et al. 1972). Retroviruses are valuable research tools in molecular biology, and they have been used successfully in gene delivery systems.

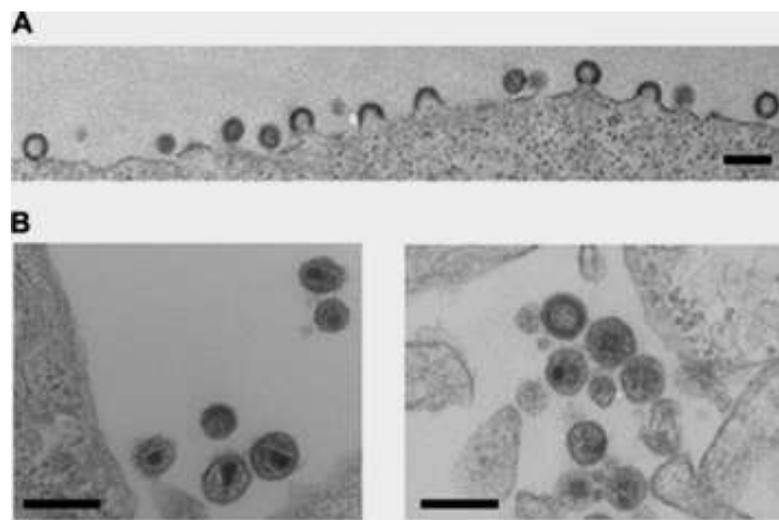
**Table 1.1** Classification of Retroviruses

Genus	Examples	Host
<i>Alpharetrovirus</i>	Rous sarcoma virus (RSV)	<i>Gallus gallus</i>
<i>Betaretrovirus</i>	Mouse mammary tumour virus (MMTV) Jaagsiekte sheep retrovirus (JSRV) Mason-Pfizer monkey virus (MPMV)	<i>Mus musculus</i> <i>Ovis aries</i> <i>Macaca mulatta</i>
<i>Gammaretrovirus</i>	Murine leukemia virus (MLV) Feline leukemia virus (FeLV) Python endogenous retrovirus (PyERV)	<i>Mus musculus</i> <i>Felis catus</i> <i>Python molurus</i>
<i>Deltaretrovirus</i>	Human T-cell leukemia virus type 1 (HTLV-1) Bovine leukemia virus (BLV)	<i>Homo sapiens</i> <i>Bos taurus</i>
<i>Epsilonretrovirus</i>	Walleye dermal sarcoma virus (WDSV)	<i>Stizostedion vitreum</i>
<i>Lentivirus</i>	Human immunodeficiency virus type 1 (HIV-1) Simian immunodeficiency virus (SIV) Feline immunodeficiency virus (FIV)	<i>Homo sapiens</i> <i>Macaca mulatta</i> <i>Macaca nemestrina</i> <i>Felis catus</i>
<i>Spumavirus</i>	Simian foamy virus (SFV)	<i>Macaca mulatta</i> <i>Macaca nemestrina</i>

## 1.1.2 Genera

### 1.1.2.1 Lentiviruses

*Lentiviruses* are classified into five groups according to the vertebrate hosts they infect. Lentiviruses are retroviruses with additional accessory and regulatory proteins, that differ in morphology and genetic composition: feline immunodeficiency virus (FIV) encodes 6 genes, and the human immunodeficiency virus 1 (HIV-1) genome encodes 9 genes in total. Several lentiviruses such as HIV-1 can deliver large amounts of genetic information into the DNA of host cells and replicate in non-dividing cells, causing slowly progressing diseases (Sakuma et al. 2012). This feature (infecting nondividing cells) converts the lentiviruses into a useful tool for gene therapy, but integrating the lentiviral DNA into the host cell genome can disturb the function of host cell genes by leading to the repression or overexpression of such genes and insertional mutagenesis (Bokhoven et al. 2009). In addition, some patients treated with retroviral vectors to cure their diseases subsequently developed vector-related leukemia diseases in a significant number of patients, demonstrating that the future success of these vectors requires more research to understand their biology in depth (Escors et al. 2011). Thus, MLV-based vectors employed in cancer therapies have been related to DNA replication instability and DNA damage in severe combined immune deficiency (SCID) patients (Bester et al. 2013)



**Figure 1-1.** HIV-1 particles budding and released at the host plasma membrane. B) EM representative sections depicting normal mature virions released from wild-type HIV-1. Some aberrant HIV-1 particles are observed in image B (right). Photos taken from Bartonova, V. y col. (2008) *J. Biol. Chem.* 283:32024-33033

Other approaches in gene therapy are related to adeno-associated virus 2 (AAV-2), developed to establish long-term gene expression. The replication cycle of AAV-2 involves a transformation of single stranded DNA genome into a double stranded DNA suitable for gene expression (Fisher et al. 1996). Thus, the serotype tropism comprises a clinical problem that impedes AAV-2 vectors to infect target cells, which are not susceptible to infection. There are several limitations in AAV-2 gene therapy such as the low expression of cellular receptors and co-receptors in non-permissive cells and the intracellular trafficking process after infection, but these inconveniences have been approached with current recombinant AAV vectors (rAAV) (based on AAV-2 vectors) and increasing genome capacity and gene expression (Fisher et al. 1997). Moreover, the AAV-2 vectors integration occurs at low frequency meanwhile lentiviral vectors integration can integrate at many sites along the human genome contributing to the generation of mutagenesis in patients, but the limitations of AAV-2 serotypes and genomic capacity represent relevant disadvantages that lentiviral vectors can overcome.

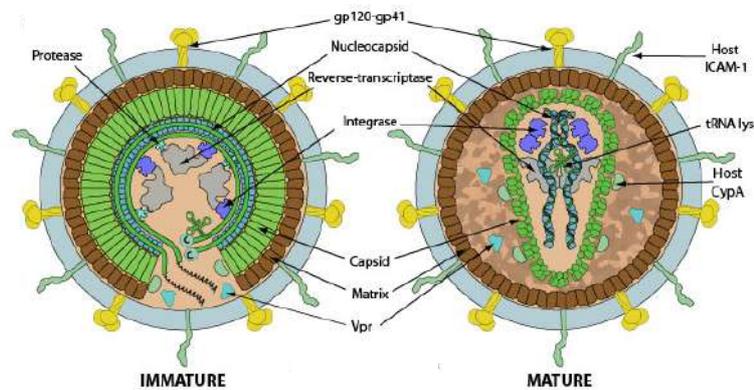
The AAV vectors have demonstrated the greatest clinical success *in vivo* given the existence of many serotypes and capsid variants developed that target a wide variety of cells, but the limitations to carry large amounts of genetic material need to be developed to increase the potential in gene therapy using AAV-2 vectors.

## 1.2 Genome organization and virion characteristics

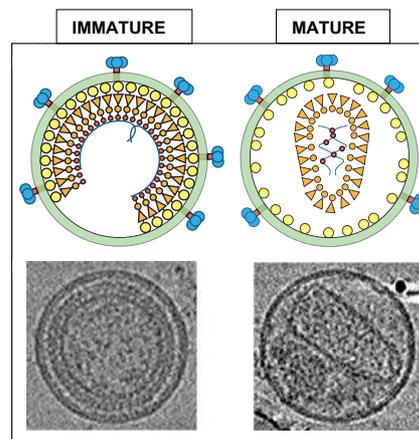
Retroviruses are a large and diverse family of infectious agents classified by virion morphogenesis and replication cycle. However, all intact retroviruses share three genes: *gag*, *pol* and *env* (Coffin 1979), which are precursors of different viral proteins. In the HIV-1 genome, the capsid (CA/p24), matrix (MA/p17), nucleocapsid proteins (NC/p7), p2, p1 and p6 are encoded in *gag*; protease (PR), reverse transcriptase (RT), ribonuclease H (RH) and integrase (IN) are encoded in *pol*; and transmembrane (gp41) and surface proteins (gp120) are encoded in *env*. The CA and NC proteins compose the viral core, and the MA lines the inner surface of viral envelope; the PR, RT, RH and IN proteins are involved in viral gene expression and integration, and the envelope proteins determine the tropism of the virus. The PR is essential for the cleavage of Gag polyprotein to separate matrix, capsid and nucleocapsid proteins, and the RT enzyme synthesizes a dsDNA from the viral RNA genome (through a reverse transcription mechanism), which will be subsequently integrated into the host genome.

The HIV-1 virion possess a similar structure: enveloped particles about 100nm in diameter, containing two molecules of single-stranded RNA about 7 to 10 kilobases in length, linked by their 5' end to form a dimeric RNA structure (Greatorex 2004). The viral envelope requires conformational changes required for the process of attachment and fusion with the host membrane (Riedel et al. 2017). Thus, HIV-1 matrix is composed of amino terminal myristyl acids creating the isometric structure of the particles, which are also associated with the viral envelope (Conte et al. 1998; Sun et al. 2014). The viral capsid is composed of Gag proteins, which associate with the viral membrane. There are several elements that are located inside the capsid in HIV-1 virion: two RNA molecules associated with the nucleocapsid proteins (Zuber et al. 2000). (Figure 1-2).

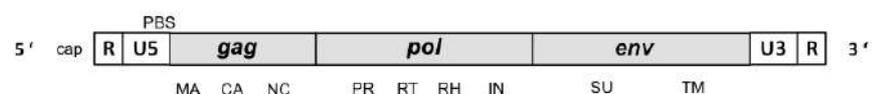
A



B



C



#### Non-coding sequences

<b>R</b>	repeat sequence
<b>PBS</b>	primer binding site
<b>U5</b>	unique sequence at 5' end
<b>U3</b>	unique sequence at 3' end

#### Genes

<b>gag</b>	group-specific antigen
<b>pol</b>	polymerase
<b>env</b>	envelope

**Figure 1-2.** A) HIV-1 virion composition. B) Schematic representation of the structure of mature and immature HIV-1 virions (upper panel) and their corresponding tomographic reconstructions (bottom panel). C) Genome organization of a standard retrovirus. Diagrams taken from Swiss Institute of Bioinformatics (2015); [https://viralzone.expasy.org/?outline=all\\_by\\_species](https://viralzone.expasy.org/?outline=all_by_species) and Karlsson Hedestam, G. B., Fouchier, R. A. M., Phogat, S., Burton, D. R., Sodroski, J., & Wyatt, R. T. (2008). The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nature Reviews Microbiology*, 6, 143. Taken from <http://dx.doi.org/10.1038/nrmicro1819>.

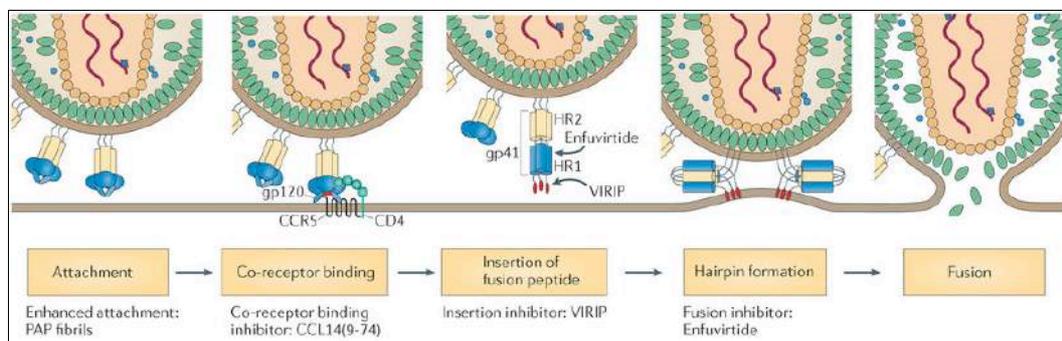
### 1.3 HIV-1 replication cycle

#### 1.3.1 Attachment and uncoating

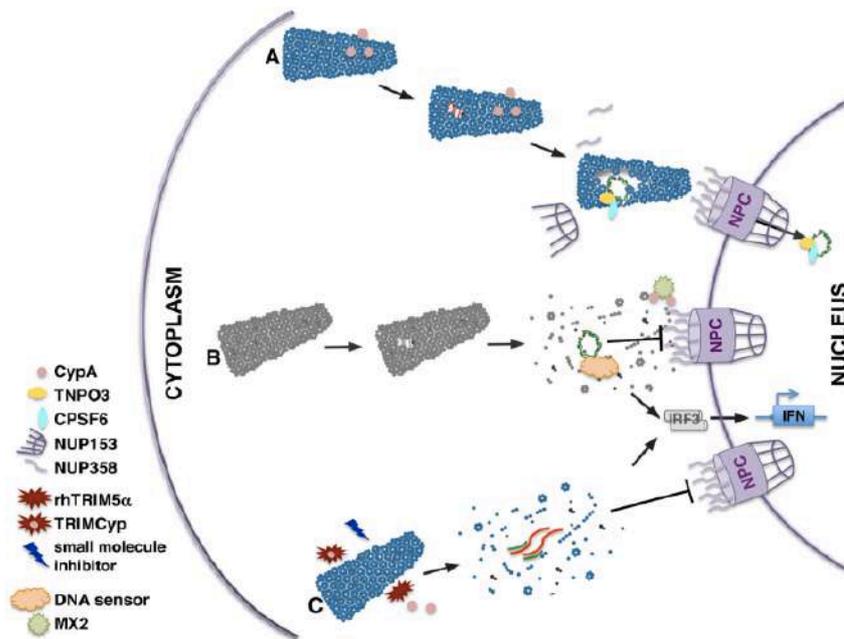
The HIV-1 replication cycle starts with the attachment of HIV-1 via the gp120 envelope protein to the CD4 receptor and co-receptors CCR5 or CXCR4, depending of the tropism of the virus (Feng et al. 1996; Clapham et al. 2001). Despite all HIV-1 strains recognizing the CD4 receptor, viral entry requires coreceptors (usually CXCR4 and/or CCR5) in order to undergo fusion and entry into the host cell (Clapham et al. 1991; Dittmar et al. 1997). Both CXCR4 and CCR5 are chemokine receptors involved in the migration of several leukocytes to inflammation sites and also specific niches in lymphoid organs (Mackay, 1996). Nowadays, CCR5 is recognized as the major co-receptor for the virus and R5 tropism appears to be more prevalent than CXCR4. Approximately 50% of infected patients undergo the switch from R5 virus (the majority in primary infections) to R5X4 tropism, which is associated with a progressive decline in the CD4<sup>+</sup> cell count and rapid disease progression. The HIV-1 tropism switch is determined by the highly variable and flexible V3 loop region of the viral gp120 Env protein (Briggs et al. 2000), specifically the amino acids between Cys263 and Cys296. The interaction between gp120 and CCR5/CXCR4 suggests that the accumulation of amino acid replacements within the V3 loop is responsible for the eventual switch in co-receptor affinity (Masso et al. 2010). However, the ‘gatekeeper’ mechanism that restrict HIV-1 transmission could be related with superimposition of multiple imperfect gatekeepers (Margolis et al. 2006). Thus, the CCR5 inhibitors employed in clinical trials represent a powerful tool to provide new insights about HIV-1 envelope tropism (Hemelaar et al. 2006).

The envelope HIV-1 glycoprotein (Env) is synthesized as a gp160 precursor of 850 amino acids. After the cleavage of gp120 precursor, the gp120 and gp41 proteins are generated, following folding and glycosylation within the endoplasmic reticulum (Wyatt et al. 1998). The gp160 precursor is cleaved by host furin-like proteases to produce the gp120 and gp41 subunits (Mao et al. 2012). This cleavage requires several arginine and lysine residues in the consensus amino-acid sequence Arg-X-Lys/Arg-Arg located at the cleavage site (McCune et al. 1988; Hallenberger et al. 1992). The Env glycoprotein consists of a trimer of heterodimers containing three gp41 and three gp120 molecules (each trimer), where the gp41 subunit remains hidden within the viral envelope by non-covalent interactions between gp120 and gp41 subunits (Mao et al. 2012).

When the viral envelope binds the cellular CD4 receptor and co-receptor, CCR5/CXCR4 conformational changes are triggered in envelope the proteins (gp120 and gp41), exposing the fusion peptide within the gp41 subunit that drives the fusion of the viral and cell membranes (Kong et al. 2016). Following penetration of the host-cell, the process termed ‘uncoating’ occurs. The HIV-1 uncoating mechanism is a highly controversial area and still uncertain. Several reports suggest that uncoating of the viral capsid occurs within the cell cytoplasm, releasing the viral genome, which will be reverse transcribed within the reverse transcription complex to generate dsDNA (Iordanskyi et al. 2007). However, there are multiple lines of evidence suggesting that the HIV-1 capsid remains intact (or partially intact) as it traverses the cytoplasm during the early stages of infection. The intact viral capsid is required to engage the nuclear envelope, and assembled capsid proteins interact with the nuclear pore complex (Jacques et al. 2016; Ambrose et al. 2014). The uncoating may occur within the nucleus, and reverse transcription could take place within the HIV-1 virion during cytoplasmic transit before it reaches the host nuclear membrane. Moreover, mutations that alter the CA stability severely reduce reverse transcription (Rihn et al. 2013).



**Figure 1-3.** HIV-1 fusion process. Both gp120 and gp41 are required during this process. The gp41 subunit is shown in green and gp120 in blue. Diagram taken from Jan Münch, Ludger Ständker, Wolf-Georg Forssmann & Frank Kirchhoff (2014). Discovery of modulators of HIV-1 infection from the human peptidome. *Nature Reviews Microbiology* 12, 715–722 (2014) doi:10.1038/nrmicro3312



**Figure 1-4.** HIV-1 uncoating mechanism linked to reverse transcription requires nuclear import and host cell factors. A) Capsid core opening. B) Premature uncoating. C) Uncoating acceleration by host viral inhibitors such as TRIM5 $\alpha$ , that causes a reduction in reverse transcription levels and nuclear import of viral DNA. Diagram taken from Zandrea Ambrose and Christopher Aiken (2014). HIV-1 Uncoating: Connection to Nuclear Entry and Regulation by Host Proteins. *Virology*. 2014 April ; 0: 371–379. doi:10.1016/j.virol.2014.02.004.

### 1.3.2 Reverse transcription

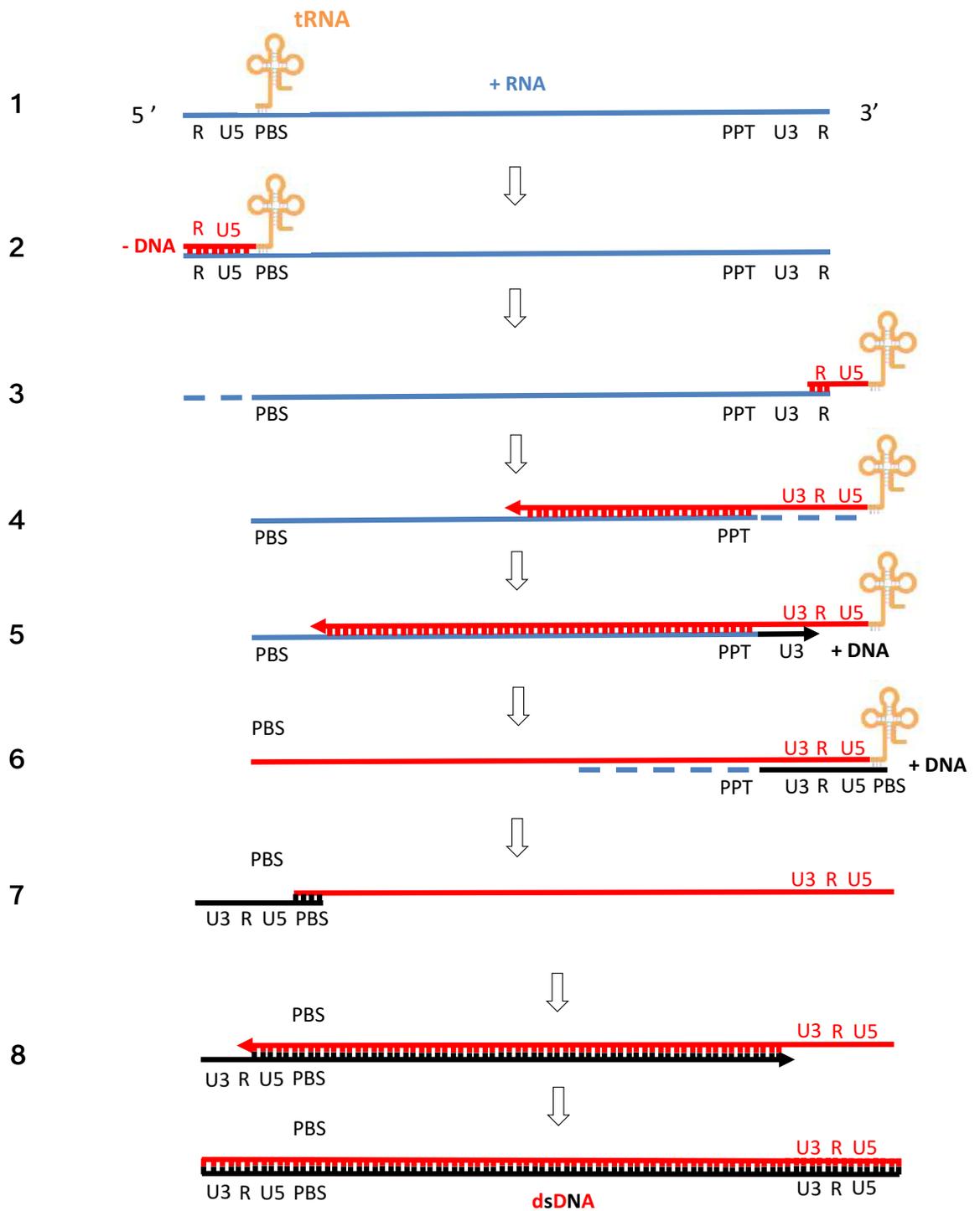
The place where reverse transcription takes place in the cell is still controversial. There are reports suggesting that reverse transcription occurs immediately after the viral genome is released from the viral capsid into the target cell cytoplasm (Mamede et al. 2017). However, more recent studies suggest that reverse transcription may occur within the viral capsid during cytoplasmic transit and inhibiting reverse transcription may delay uncoating (Figure 1-4) (Hulme et al. 2011).

The HIV-1 genome is composed of two identical molecules of ssRNA (positive sense) that act as a template for reverse transcription (Canard et al. 1997). Reverse transcription is a complicated process involving several strand transfer events required to generate the complete LTRs in the DNA provirus, which are not present in the viral RNA (O'Neil et al. 2002). The viral enzymes RNase H and reverse transcriptase (RT) are essential for reverse transcription, which generates dsDNA in the host cell cytoplasm that will be trafficked to the nucleus and will subsequently become integrated into the host genome (Figure 1-5). The steps involved in reverse transcription are the following:

1. Minus DNA strand synthesis starts at the 3'-OH of a cellular tRNA acting as a primer that binds the PBS in the virus RNA.
2. The first strand DNA is synthesized until the RT enzyme reaches the terminus of the 5' end, generating a small ssDNA fragment around 100-150 bases (this termination is termed 'strong-stop DNA').
3. Viral RNase H degrades the RNA fragment used as template and the ssDNA (which includes the reverse transcribed R region) 'jumps' and attaches to the R region in the 3' end of the viral RNA.
4. Next, this intermediate DNA is extended towards the 5' end of the viral RNA genome (terminating at the PBS), and the RNA strand used as template is now degraded by the RNase H except the polypurine tract (PPT), which is resistant to RNase H activity.
5. This RNA fragment containing the PPT acts as a primer for the generation of the +DNA strand synthesis, which is generated until it reaches the 3'-end of the -DNA strand.
6. Now RNaseH degrades all viral RNA remaining, exposing complementary sequences near the 3'-end of the +DNA strand.

7. The +DNA strand is transferred to the 5'-end of the -DNA strand, annealing to the complementary PBS at the 5' end.
8. Finally, the synthesis of both DNA strands is completed, with each strand serving as template for the other strand. Once the dsDNA is synthesized, the provirus will be transferred into the nucleus to be integrated into the host genomic DNA.

**Figure 1-5.** Reverse transcription mechanism. Diagram adapted from Virology: Principles and Applications, John Carter, Venetia Saunders.



### 1.3.3 Nuclear import and integration

Following reverse transcription, the pre-integration complex traverses the nuclear envelope and subsequently the viral dsDNA is integrated into the host cell genome (Endsley et al. 2014). The dsDNA remains associated with viral (CA, MA and Vpr) and cellular proteins at the pre-integration complex (PIC) within the cytoplasm (Figure 1-6). The PIC is translocated to the host nucleus through nuclear pore complexes (NPCs) prior to integration into the host genome (Bukrinsky et al. 1998). There are several host proteins involved in the translocation of the PIC into the host nucleus, such as importins (Imp) and nucleoporins (NUPs). The viral integrase recruits importin  $\alpha$  and  $\beta$  to import the viral dsDNA into the nucleus, and Nup153 is also involved in nuclear import (Woodward et al. 2009). Thus, the viral CA from the PIC could be an essential factor for HIV-1 nuclear import because a reduction in CA levels within the PIC results in a reduction of viral replication in cell cycle arrested cells (Yamashita et al. 2004).

Interestingly, another study showed that Imp $\alpha$  dependent nuclear localization of Vpr does not require Imp $\beta$ , and Vpr also interacts with several Nups and localizes to the nuclear envelope (Nitahara-Kasahara et al. 2007). However, the MA protein located in the PIC is also required for nuclear import via an IN interaction. Therefore, the viral DNA is also involved in nuclear integration of the PIC. The polypurine tract from HIV-1 cDNA (cPPT) forms a triple stranded DNA structure called the DNA flap that seems to be involved in nuclear import, and mutations in the DNA flap may contribute to retain the cDNA within the viral coat (Arhel et al. 2007).

Following the PIC formation, the viral dsDNA integrates in the chromatin of mitotic and non-mitotic target cells overcoming the nuclear membrane barrier, and allowing HIV-1 to successfully replicate in dividing and non-dividing cells (Gallay et al. 1997). HIV integration location is influenced by DNA structure and host factors such as the lens epithelium-derived growth factor (LEDGF/p75), a chromatin binding protein that interacts with HIV-1 integrase through specific binding domains and targets HIV-1 to transcriptionally active regions of the host chromosome (Marshall et al. 2007). Recent studies shown that knockout experiments for LEDGF severely reduce HIV infectivity and replication, suggesting that the HIV IN-LEDGF interaction could be a potential candidate for antiviral therapy (Llano et al. 2006).

During integration, IN cleaves specific phosphodiester bonds near the viral DNA end, and then uses the resulting DNA 3'-OH groups during strand transfer to cut chromosomal target DNA (Cherepanov et al. 2014). However, the integration site location of HIV-1 DNA is still controversial (Woodward et al. 2009; Gonçalves et al. 2016), but recent evidence suggests that specific integration sites are related to cell cycle control and proliferation of HIV-1 infected cells (Sunshine et al. 2016; Maldarelli et al. 2014).

The viral integration into host DNA occurs at preferred sites that are associated with active transcription (Lewinski et al. 2006), where the HIV integrase associates with the cellular chromatin factor LEDGF/p75. This interaction represents an antiviral target to develop drugs that counteract this interaction, leading to the development of an integrase inhibitor called raltegravir that was approved for clinical use in 2007 for the treatment of HIV-AIDS by allosteric inhibition of the strand transfer reaction of the viral integrase (Summa et al. 2008). However, the raltegravir resistance evolves rapidly in clinic even in the presence of highly active antiretroviral therapy (HAART). Thus, the drugs targeting the LEDGF/p75-IN interaction are small molecule inhibitors (LEDGINs) binding to this dimer to inhibit viral replication with a dual mechanism of action: the inhibition of LEDGF/p75-IN interaction and modifying the multimerization state of IN by allosteric inhibition within the catalytic activity. For example, the LEDGF/p75 binding to the pocket in the dimer interface of HIV-IN catalytic core domain can be inhibited by D77, a potential drug that interacts with LEDGF/p75 binding pocket (Du et al. 2008).

#### 1.3.4 Transcription, translation and splicing

There are two gene expression phases in the HIV-1 life cycle: early and late expression. HIV-1 gene expression is highly regulated by splicing to generate the different viral proteins encoded by the genome. At early expression phases, HIV-1 uses three different reading frames, which overlap each other, allowing the control of gene expression in multiple pathways (Locker et al. 2011). Transcription from the viral genome is initiated when cellular transcription factors bind to the proviral LTR promoter located at the 5' region of the viral genome (Kutky et al. 2017). Then, viral transcripts are polyadenylated at the R-U5 junction within the 3' LTR. Most of these nascent RNAs are spliced in early infection, transported and translated in the cytoplasm generating Tat, Rev and Nef proteins (Strebel 2003). The regulatory proteins Tat and Rev are essential for the viability of HIV-1

and are translated in the cytoplasm following integration and these proteins traffic to the nucleus, where they exert their functions (Sawaya et al. 2000). However, Nef remains in the cytoplasm and is myristoylated close to the N-terminus (Mishra et al. 2011; Love et al. 1998; Bentham et al. 2006) causing it to associate with cellular membranes.

Most of the regulatory mechanisms of HIV gene expression are controlled by splicing, which seems to be independent of Tat (Leblanc et al. 2013; Debaisieux et al. 2012). In addition, spliced viral RNAs are transported to the cytoplasm during early stages of infection in a Rev independent manner (Dayton 2004). However, the export of full length viral RNAs are Rev dependent. Rev protein interacts with the Rev Response Element (RRE) located in the *env* gene RNA (Sherpa et al. 2015). The formation of the Rev-RRE complex is essential for unspliced viral RNA export to the cytoplasm to permit translation of viral proteins (Fernandes et al. 2012) and provide full length RNA-genomes to be packaged into nascent virions. Cellular co-factors are also required in this step such as the chromosomal maintenance 1 (CRM1) (Sherer et al. 2011).

### 1.3.5 Packaging, assembly and release

The assembly for viral particles production requires two copies of the viral RNA genome, a cellular tRNA<sup>Lys,3</sup>, Gag and Env proteins and the viral enzymes RT, PR and IN. The viral packaging and assembly occurs at the plasma membrane, where Gag proteins form multimeric complexes and bind to multiple molecules including the viral genome (Nikolaichik et al. 2017). The polyprotein precursors Gag and Gag-Pol regulate the events for viral assembly: Gag assembly leads to the generation of immature particles, and Gag molecules are extended radially in the immature virions, with their N-terminal MA domains bound to the inner membrane. HIV-1 assembly and budding appears to occur predominantly at the plasma membrane. However, several parameters of HIV-1 assembly are challenging to determine, and some have been controversial (Jouvenet et al. 2011)

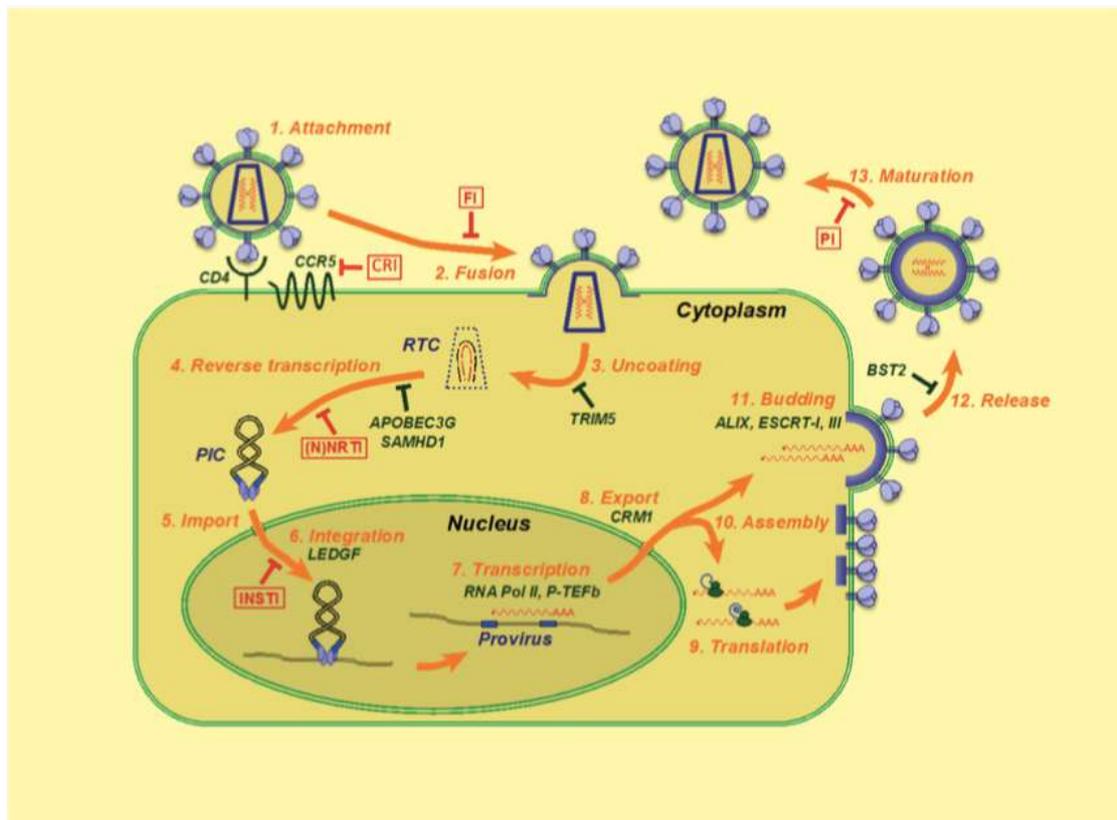
The NC and MA proteins are essential for the generation of immature virions, and the maturation of HIV virions occurs when the PR cleaves Gag and Gag-pol polyproteins generating MA, CA, NC, p6, PR, RT, and IN proteins during budding process (Scarlata et al. 2003). The viral protease cleaves Gag and Gag-pol polyproteins generating MA, CA, NC, p6, PR, RT, and IN proteins during the budding process to generate immature viral particles that will be released and subsequently rearranged into the mature infectious form.

This cleavage triggers the dimeric formation of the viral RNA genome and the NC protein, which contains basic residues with zinc fingers that bind a dimeric, unspliced HIV-1 RNA called the packaging signal ( $\psi$ ) (Dannull et al. 1994). Approximately 95% of the translating ribosomes produce Gag, while 5% produce a Gag-Pol fusion protein originating from a ribosomal frameshift sequence located at the end of *gag*, which allows the expression of the Pol viral polyprotein (Hung et al. 1998; Biswas et al. 2004). In the case of *gagpol*, the GagPol polyprotein is translated when a ribosome frameshift occurs due to the slippery sequence U UUU UUA located at the end of *gag* (Mathew et al. 2015). Finally, the envelope gp160 proteins are synthesized, glycosylated and cleaved by host furin-like proteases to produce the gp120 and gp41 subunits (Sharma et al. 2015). The Rev-dependent full-length HIV-1 genome dimerizes via base pairing between complementary sequences at the 5' end of each RNA strand, stabilized by the kissing-loop complex (Checkley et al. 2011).

The rest of viral proteins translocate to the cell surface to assemble into new immature virus particles that are released by the ESCRT (endosomal sorting complexes required for transport) machinery (Morita et al. 2007). Capsid proteins bind the host protein cyclophilin A and Gag C-terminal p6 domain bind to Vpr (Zhu et al. 2004). Vif could be packaged into virions as a component of the viral nucleoprotein complex (Khan et al. 2001), but there are other data suggesting that Vif packaging is irrelevant and is required for sustained HIV-1 infection to counteract antiviral factors within infected cells (Salter et al. 2014).

The cellular intrinsic immunity family of cytidine deaminases APOBEC3G (A3G) are involved in antiviral mechanisms of mRNA modification by deamination of the cytidines (cytidine to uridine) of the target viral RNA to impair viral replication within the infected cell (Sheehy et al. 2002). Once the viral RNA is released into the cytoplasm the A3G proteins edit the viral genome sequence to inhibit viral infection, reducing the accumulation of nascent HIV-1 reverse transcripts in newly challenged cells. These effects have been described for early and late reverse transcription intermediates, and it appears that the magnitude of the defects increases as later cDNA products are measured (Anderson et al. 2008). The A3G proteins are encapsidated by viral budding particles, and their function is exerted in the next cycle of replication. However, the HIV-1 protein Vif counteracts APOBEC3G function binding and subsequently inducing polyubiquitylation, leading to proteosomal degradation of A3G to diminish the pool of A3G available that

could be incorporated into the nascent particles. Moreover, Vif protein recruits cullin5-elongin B/C-Rbx ubiquitin ligase, serving as an adaptor to recruit the ligase complex to its substrate (Yu et al. 2003). Moreover, the Asp-Pro-Asp motif at positions 128-130 in A3G are critical for Vif interaction (Huthoff et al. 2007).

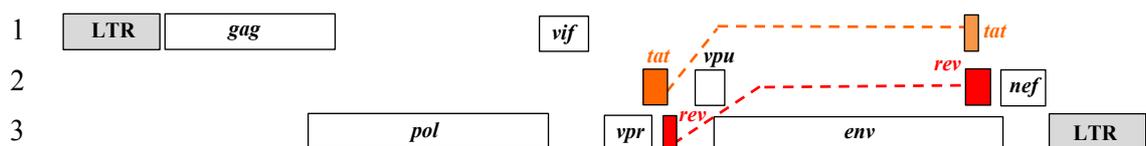


**Figure 1-6.** HIV-1 replication cycle. Taken from A. Engelman and P. Cherepanov (2012). The structural biology of HIV-1: mechanistic and therapeutic insights. *Nature Reviews Microbiology* 10, 279–290(2012) doi:10.1038/nrmicro2747

## 1.4 HIV-1 genome structure

HIV-1 possesses the genetic structure shared by all retroviruses: *gag*, *pol* and *env*, but also encodes 6 additional genes: *vpu*, *vpr*, *vif*, *tat*, *rev* and *nef* (Figure 1-7). These genes are involved in many roles, such as controlling virus gene expression, transporting virus components within the cell or modifying the host immune response (Bohne et al. 2007). There are three open reading frames along the HIV-1 genome that overlap each other, allowing the control of gene expression in multiple pathways. At first, Vif and Nef proteins are generated upon early infection to counteract antiviral activity from the host cell and downregulate multiple cell surface markers including CD4 (Mlcochova et al. 2015). The Vpu protein counteracts the host protein tetherin reducing the levels at the infected cell surface (Neil et al. 2008; Van Damme et al. 2008). Thus, Vpu also reduces the levels of tetherin within infected cells when is overexpressed (Bartee et al. 2006). However, the role of Vpr protein still controversial, but it mediates many processes related to evade immune response and the persistence of HIV-1 infection (Conti et al. 1998), and also induces G2 cell cycle arrest interacting with the DNA binding protein 1 (DCAF-1) factor, leading to immune cell dysfunction (Belzile et al. 2007). Moreover, Vpr could be also involved in pathogenesis (Gibbs et al. 1994)

### ORFs



**Figure 1-7.** HIV-1 genome organization and open reading frames. Adapted from Virology: Principles and Applications, John Carter, Venetia Saunders

The HIV regulatory proteins Tat (trans-activator of transcription) and Rev (regulator of expression of viral proteins) are produced in early stages after infection and stimulate viral transcription and viral RNA transport respectively.

### 1.4.1 HIV-1 Long Terminal Repeat (LTR)

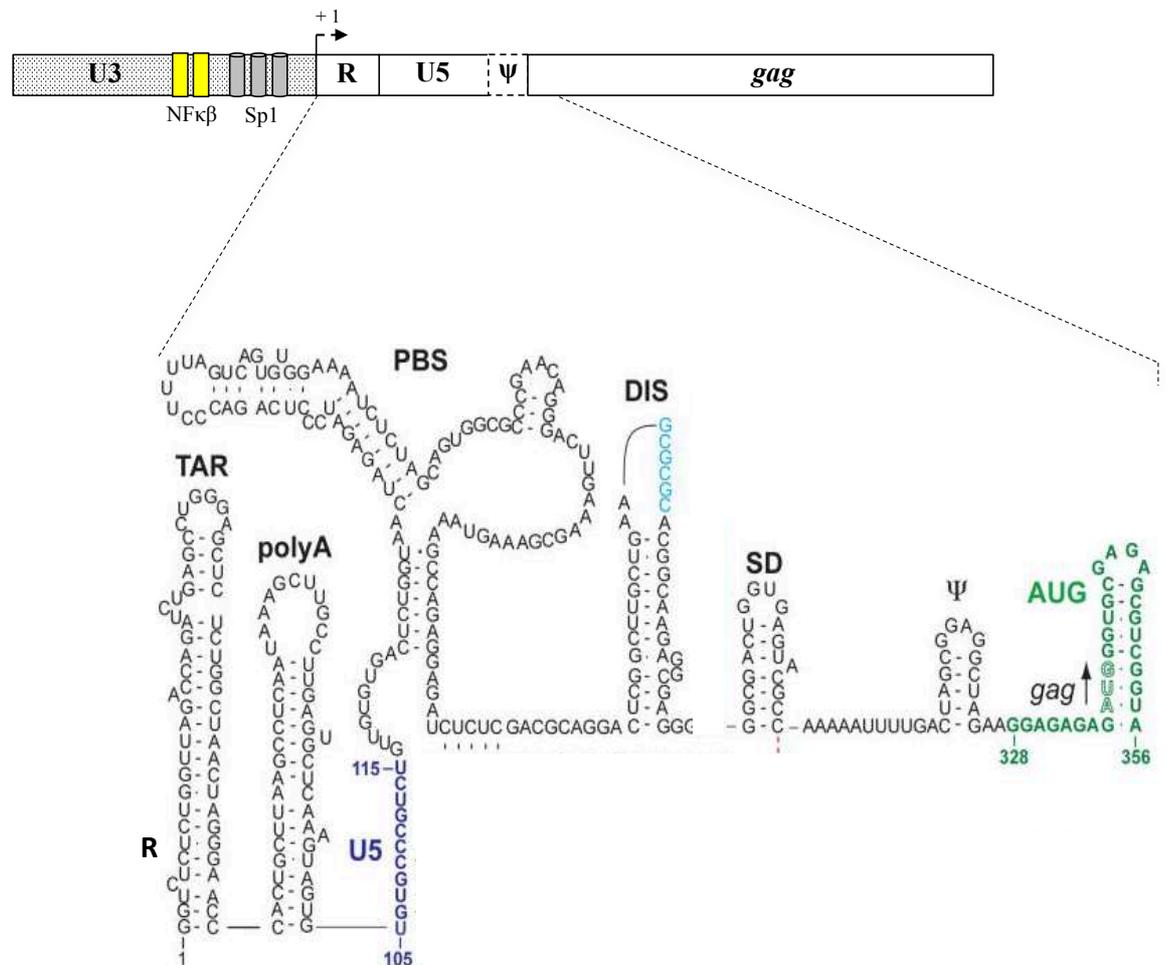
HIV-1 encodes 9 genes in total that are flanked by two Long Terminal Repeats (LTRs) at the 5' and 3' ends, where the 5' LTR controls viral gene expression by acting as a promoter (Klaver et al. 1994). These two LTRs are essential to perform the reverse transcription mechanism (section 1.3.2). Each HIV-1 LTR is ~634bp in length, composed of three regions (as a provirus): unique to the 3' end (U3), R region (R) and unique to the 5' end (U5). Before reverse transcription takes place, the 5' end of the RNA genome only encodes the R-U5 region, and the 3' end encodes the U3-R (Figure 1-7 and Figure 1-8). However, the U5 and U3 regions are duplicated at both ends of the genome during reverse transcription, where the resultant dsDNA provirus contains two identical copies of the tandem U3-R-U5 (Figure 1-8).

The U3 region contains *cis*-acting binding sites for cellular transcription factors, but viral transcription starts (by definition) at the start of the R region, which is also essential for viral reverse transcription (Ohi et al. 2000; Berkhout et al. 2001). There are two secondary RNA structures in the R region, such as the TAR loop and the polyadenylation signal that are required for efficient initiation and termination of viral transcription. The U5 region possesses several elements involved in many relevant functions: the primer binding site (PBS), the dimerization site (DIS), the splice donor (SD) and the packaging signal ( $\psi$ ) (Figure 1-8). A cellular tRNA (Lys3) binds to the PBS of the viral genome, acting as an RNA primer for reverse transcription. The DIS joins the two molecules of single stranded viral RNA, generating a 'kissing-loop' stabilizing the joint. The SD region is essential for viral RNA splicing, and the packaging signal interacts with Gag proteins to incorporate the full-length viral genomes into the nascent virions (Dahiya et al. 2014).

The HIV-1 LTR also presents DNA binding elements for cellular transcription factors that activate or repress the viral gene expression (Montano et al. 1996). There are two adjacent NF-kappa  $\beta$  sites located in the U3 region: the NF-kappa  $\beta$  protein allows the virus to be responsive to the activation state of the infected T cell. The mechanism of activation of NF-kappa  $\beta$  involves the formation of reactive oxygen intermediates that trigger the phosphorylation and degradation of I $\kappa$ B (Deptala et al. 1998), which triggers the translocation of NF-kappa  $\beta$  to the nucleus where it induces the expression of T-cell specific genes. (Stroud et al. 2009; Valentin et al. 1989). Mutations or deletion of the NF-

kappa  $\beta$  sites within HIV-1 LTR abolish reporter gene activation and host transcription factors such as nuclear factor of activated T-cells NFAT1 and NFAT5 are unable to activate HIV-1 transcription through these sites (Kinoshita et al. 1997; Cron et al. 2000).

Additional binding sites for constitutive transcription factors such as SP-1, Lef, Ets, and for inducible transcription factors NF-AT and AP-1 are also encoded within the LTR. The three SP-1 binding sites are essential for the function of the LTR, and removal of these sites results in a huge decrease in transcription levels (Harrich et al. 1989). The HIV-1 LTR is unable to reach sufficient levels of viral transcription independently, but when HIV-1 Tat protein binds to the LTR it enhances viral transcription ~200-300 fold (Southgate et al. 1991). In the absence of Tat, the stimulation of viral transcription (carried out by cellular transcription factors) leads to the generation of short transcripts. Host proteins such as cyclin T1 (cycT1) and cyclin-dependent kinase 9 (CDK9) are recruited to form a complex with Tat that stimulates viral transcription. The secondary RNA structures of the LTR are essential for HIV-1 replication, and conformational changes severely affect the viability of the virus (Mueller et al. 2014).



**Figure 1-8.** HIV-1 LTR structure. There are several secondary RNA structures along the LTR. The R region contains the TAR loop and the polyadenylation signal, meanwhile the U5 region encodes the primer binding site (PBS), the dimerization site (DIS), the splice donor (SD) and the packaging signal ( $\psi$ ). Adapted from Kun Lu, Xiao Heng and Michael F. Summers. *J Mol Biol.* 2011 Jul 22; 410 (4):609-633, L J Chang, E McNulty.

### 1.5 Origin and evolution of HIV-1

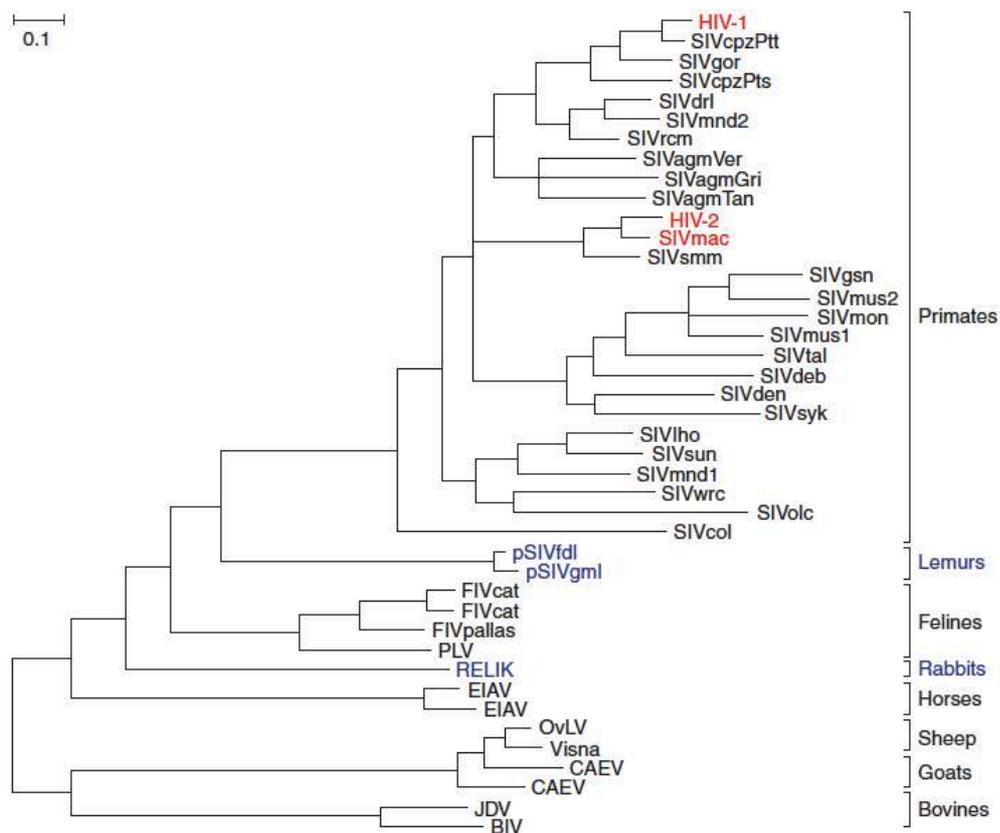
Human immunodeficiency virus (HIV), the causative agent of acquired immunodeficiency syndrome (AIDS) was identified in 1983 by Luc Montagnier and Françoise Barré-Sinoussi, revealing a novel lymphadenopathy-associated virus (LAV) from a patient suffering from AIDS (Barré-Sinoussi et al. 1983). There are two main types of human immunodeficiency viruses, which are designated HIV-1 and HIV-2. The closest relatives of HIV-1 are simian immunodeficiency viruses (SIVs) infecting wild-living chimpanzees (*Pan troglodytes troglodytes*) and gorillas (*Gorilla gorilla gorilla*) in west central Africa (Letvin et al. 1985). Recent phylogenetic studies have revealed that

chimpanzees were the original host, and that multiple HIV-1 have resulted from independent cross species transmission events (Sharp et al. 2010, Rambaut et al. 2004) (Figure 1.9). HIV-1 strains are classified into 4 groups with different prevalence: group M, N, P and O. The most prevalent infections in the world are related to the group M (>90%), meanwhile groups N and O are restricted to central Africa countries such as Cameroon (Triques et al. 1999).

HIV-1 is closely related to chimpanzee SIV (SIVcpz), originally from Western and Central African species (Gao et al. 1999). These regions contain two different chimpanzee sub-species respectively: *Pan troglodytes troglodytes* and *Pan troglodytes schweinfurthii*. Interestingly SIVcpz from *P.troglodytes* shares many similarities with HIV-1 M group, and HIV-2 is more similar to SIV infecting *Sooty mangabeys* (SIVsm), from West Africa. (Sharp et al. 2011). For a long time, it was suspected that SIVs caused little disease in their natural hosts (in contrast to HIV-1 and AIDS in humans). However, recent studies have revealed new insights about SIV pathogenicity: some chimpanzees infected with SIVcpz (as humans infected with HIV-1) present a progressive CD4<sup>+</sup> T-cell loss, lymphatic tissue destruction and premature death (Sharp et al. 2010). This evidence suggests that SIV infections could affect lifespan of chimpanzees and severely affect their health in the wild in a similar way as AIDS does in humans. HIV-1 is a diverse and rapidly evolving pathogen, and the extensive genetic variation of HIV-1 lineages differs between individual hosts, representing a major roadblock to vaccine development (Maldarelli et al. 2013). This variability is explained by the high frequency of errors introduced during reverse transcription in each replication cycle (average error rate per detectable nucleotide incorporated of ~1/1700) and also recombination (Roberts et al. 1988), increasing the genetic variability of the population, likely contributing to the evasion of specific immune responses (Mostowy et al. 2011).

A great deal of recombination has occurred among diverse SIV lineages during the evolution of primate lentiviruses, and individuals infected with more than one genetically distinct HIV-1 tend to harbour hybrid viruses (Courgnaud et al. 2002; Chen et al. 2006). However, there is little evidence supporting a relationship between genetic diversity and pathogenesis. The viral population present within a infected person consists of a complex mixture of heterogeneous strains, with different antigenic and phenotypic features (Meyerhans et al. 1989). Thus, the high frequency of recombination in HIV assists the

virus in adapting to and overcoming selective pressures, hampering the development of antiretroviral therapies (Domingo et al. 2000). However, HIV-1 exhibits reduced viral diversity following transmission to a new host and only one (or a few viruses) establish the initial infection within the new individual (transmission bottleneck) (Joseph et al. 2015).



**Figure 1-9.** Evolution of primate lentiviruses. This tree shows the evolutionary relationship among Pol sequences derived from mammalian lentiviruses. HIV-1 and HIV-2 appeared after independent cross-transmission events from SIVs. The exogenous retroviruses are depicted in black; the endogenous retroviruses are shown in blue and HIV-1, HIV-2 and SIVmac are highlighted in red. Host species are shown at the right. Image taken from Sharp, P. M., & Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harbor Perspectives in Medicine*, 1(1), 1–22. <https://doi.org/10.1101/cshperspect.a006841>

## 1.6 Regulatory proteins of HIV-1

The human immunodeficiency virus type 1 encodes regulatory genes that are not present in other retroviruses such as murine leukemia virus (MLV). However, the origins of HIV-1 regulatory genes *tat* and *rev* remain unknown (Magin et al. 1999). The *tat* and *rev* genes are mainly involved in viral transcription and transport/stability of the viral RNA respectively. Both genes are essential to complete the viral life cycle, and their absence negatively affects viral transcription and the export/stability of full-length (unspliced) viral genomes from the host-cell nucleus.

Multiple obstacles preventing HIV- 1 replication in non-human cells are related to Rev and Tat proteins; Rev is mainly involved in RNA transport within the host cell (after the HIV genome has been transcribed), but Rev also seems to have additional functions such as stabilizing viral transcripts (Kashanchi et al. 1994).

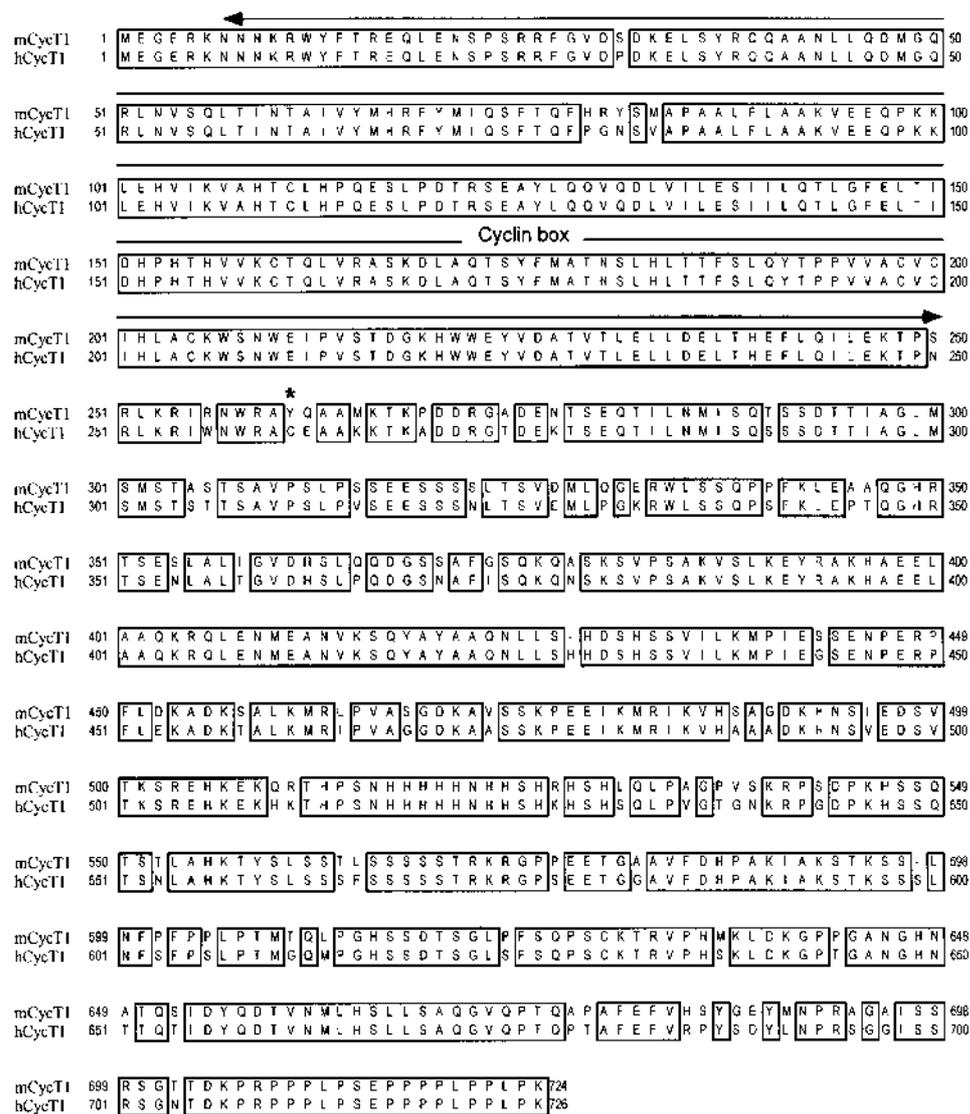
### 1.6.1 Trans-activator of transcription (Tat) and the Tat-TAR complex

The human immunodeficiency virus Tat protein ensures high levels of HIV transcription mediated by human transcription elongation factor (P-TEFb), which interacts with Tat and phosphorylates the C-terminal domain (CTD) of RNA polymerase II facilitating higher processivity of RNA polymerase II during the elongation phase. (Zhou et al. 2002). The formation of a ternary complex between TAR RNA, Tat, CDK9 and cyclinT1 is essential to enhance viral transcription. In the absence of Tat, viral transcription levels are considerably reduced because the polymerase engages poorly with the LTR, causing the complex to frequently 'stall'. However, the viral genome can be completely transcribed in its absence, albeit extremely inefficiently (Shen et al. 2000).

When viral transcription is initiated after the integration of the dsDNA provirus into the host genome, a short RNA (~65 bp in length) is synthesized. This 65bp RNA contains an RNA regulatory signal loop called transactivator response element (TAR), located within the 5'LTR of the viral genome (Jeang et al. 1990). The TAR loop acts as a binding site for essential cellular co-factors, such as the P-TEFb, a kinase complex including a regulatory cyclin T1 subunit (CycT1) and a catalytic kinase subunit (CDK-9) (Figure 1-10) (Price et al. 2006). Both subunits are recruited to the TAR-Tat complex through direct interactions with Tat (Rana et al. 2002). Mutations altering the upper stem-loop of TAR severely affect



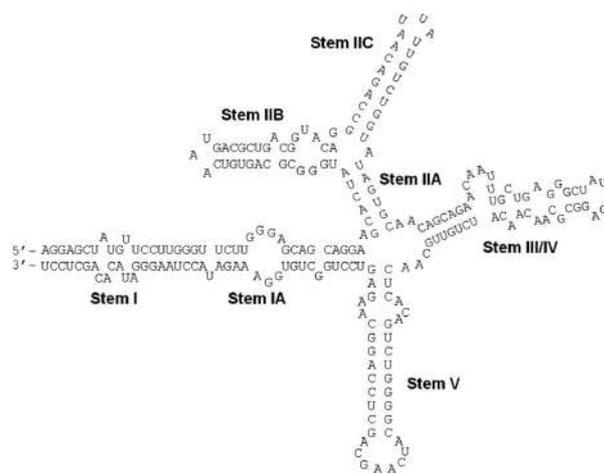
The Tat and hCyclinT1 interaction is important, but there are additional residues that are critical to facilitate binding of the complex to TAR RNA (trans-activator response element). There are zinc interactions involved in Tat-cyclinT1 interaction, where zinc-independent complexes dramatically reduce the ability of this complex to bind TAR RNA (Garber et al. 1998). However, there are also additional cellular proteins involved in HIV-1 transcription, such as the chaperone Hsp7, which generate complexes that stabilize the cyclinT1-P-TEFb complex and also contribute to the folding of hCyclinT1 (O’Keefe et al. 2000).



**Figure 1-11.** Amino acid sequence comparison of hCycT1 and mCycT1. The identical amino acids are boxed and the conserved cyclin box indicated. The crucial single residue at position 261 is indicated by an asterisk. The sequence of murine CycT1 has been deposited in the DDBJ/EMBL/GenBank database (accession No. AF095640). Figure taken from Bieniasz, P. D., Grdina, T. A., Bogerd, H. P., & Cullen, B. R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. *The EMBO Journal*, 17(23), 7056–65. <https://doi.org/10.1093/emboj/17.23.7056>

### 1.6.3 Regulator of the expression of the viral proteins (Rev) and the Rev Response Element (RRE)

The transition from early to late transcripts is mediated by the HIV-1 Rev protein, which is mainly involved in the nuclear export of incompletely spliced/unspliced viral RNA. Rev is a 19-kDa phosphoprotein located predominantly in the nucleus (Cochrane et al. 1990) and is composed of several domains. The N-terminal region encodes the nuclear localization signal (NLS), responsible for Rev nuclear location. The N-terminal domain also has a nuclear inhibitory signal (NIS), which controls the distribution and activity of Rev (Kubota et al. 1998). The C-terminal domain mediates the Rev effector functions *in vivo* encoding the nuclear export signal (NES) that interacts with cellular proteins involved in viral mRNA export (Malim et al. 1989). The RRE, located downstream the *env* gene, is essential to export the full-length viral RNAs towards the cytoplasm (Kjems et al. 1992).



**Figure 1-12.** Predicted secondary structure for RRE from HXB3 strain with the major stem loops labeled. This structure is predicted from chemical and enzymatic probing studies in combination with mobility shift assays to determine the binding region for HIV-1 Rev protein. Taken from Fernandes, J. D., Jayaraman, B., Frankel, A. D., & Location, G. (2012). © 2012 Landes Bioscience.

There are some instability regions within the HIV genome that are crucial in gene expression. These regions called (INS) are AU rich located in the main viral gene sequences: *gag*, *pol*, and *env*, requiring the presence of Rev for efficient RNA expression (Keating et al. 2009). The RRE is present in the unspliced and singly spliced viral transcripts, but is absent from the multiply spliced transcripts. In addition, it is necessary for the formation of a RRE-Rev complex to export the unspliced and singly spliced viral mRNAs (Suhasini et al. 2009).

The RRE is extremely well conserved across subtypes of the major group M, with conservation pressure from both the RNA structure and the overlapping Env protein reading frame (Fernandes et al. 2012), and chemical and enzymatic probing studies confirmed the secondary structure and in combination with mobility shift assays, established its importance in sequence-specific Rev-binding (Kjems et al. 1991). Moreover, the accuracy of secondary structure prediction structures can be improved using experimental mapping data that provide information about the pairing status of single nucleotides using high-throughput sequencing from transcriptomes (Spasic et al. 2018).

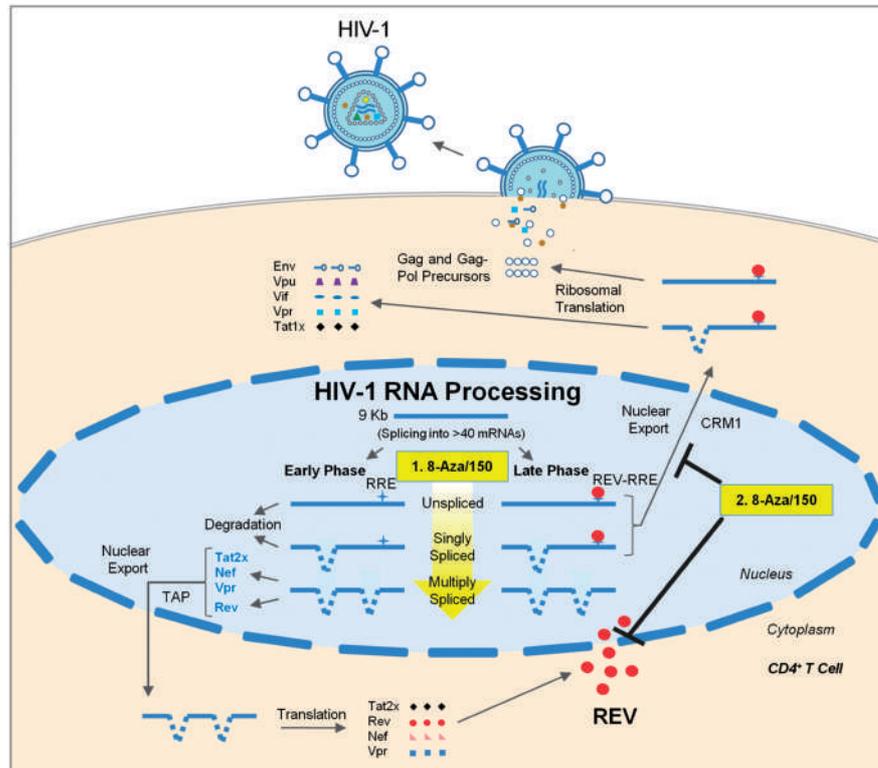
#### 1.6.4 Chromosomal maintenance 1 (CRM1)

There are cellular cofactors, such as the Chromosomal Maintenance 1 (CRM1) and importin- $\beta$  that are involved in the formation of the RRE-Rev complex (Figure 1-12 and Figure 1-13). CRM1 is a nuclear transport receptor regulated by Ran GTPase that engages with the NES of Rev allowing the export of viral mRNA, meanwhile the importin- $\beta$  imports Rev protein to the nucleus (Fukuda et al. 1997; Truant et al. 1999). The formation of the multimeric complex (Rev/RRE and CRM1) enables the export of full-length and singly spliced viral mRNAs to the cytoplasm (Figure 1-13).

However, substitutions between nearly identical orthologs of CRM1 restrict HIV replication in different species that may impair key post-transcriptional properties in murine cells (Booth et al. 2014; Nagai-Fukataki et al. 2011). The ability of human CRM1 (hCRM1) to stimulate HIV-1 production requires a species-specific configuration of amino acids on the convex surface of CRM1, and recent phylogenetic analysis show that this feature could be conferred only in higher primates (Sherer et al. 2011). Moreover, this configuration of amino acids within hCRM1 can be transferred into murine CRM1 (mCRM1), overcoming the deficiencies of mCRM1 to generate virus particle production in murine cells. However, the mCRM1 fails to engage Rev multimers to form Rev-RNA multimeric complexes that are required for nuclear export (Aligeti et al. 2014). The major block to HIV-1 mRNA nuclear export in rodent cells is weak interactions between mCRM1 and Rev/RRE complexes occurring within the nucleus. Thus, the hCRM1 could also be playing additional Rev-independent roles in the HIV-1 replication cycle.

The HIV export mechanism relies on the host protein CRM1 to export its unspliced and partially spliced RNA transcripts. Recent findings have implicated a DEAD-box RNA

helicase termed DDX3 as a member of the export complex that can bind to CRM1 (Yedavalli et al. 2004). DDX3 promotes nuclear export and translation of the HIV-1 unspliced mRNA, and knockdown of DDX3 was shown to strongly inhibit HIV-1 replication. An important issue regarding binding of DDX3 to CRM1 is the uncertainty in the sequence of events occurring throughout the assembly and disassembly of HIV-1 mRNA export complex, and DDX3 may associate with the HIV-1 Rev-CRM1 export complex at later step of transport (Mahboobi et al. 2015). Thus, DDX3 promotes translation of viral mRNAs containing the HIV-1 5' LTR, especially with the transactivation response element (TAR) and also interacts with HIV Tat protein. Moreover, the DDX3 is recruited to the TAR hairpin by interaction with viral Tat to facilitate HIV-1 mRNA translation (Lai et al. 2013). However, recent studies have shown DDX3 as an HIV-1 sensor that bound abortive HIV-1 RNA after HIV-1 infection and induced dendritic cells maturation and type I interferon responses via the signaling adaptor MAVS and DDX3 is crucial in translation initiation of HIV-1 mRNA in dendritic cells (Gringhuis et al. 2017).



**Figure 1-13.** RNA processing along HIV-1 life cycle. Rev attaches the RRE located in the viral RNA and the host factor CRM1 for nuclear import and export. This complex is essential for full-viral RNA export, but small RNA transcripts are exported in a Rev independent manner. Diagram adapted from *Nucleic Acids Research*, 2013, Vol. 41, No. 20 9471–9483 doi:10.1093/nar/gkt727, Characterization of novel inhibitors of HIV-1 replication that function via alteration of viral RNA processing and rev function

## 1.7 Rev-like proteins founded in retroviruses

A key issue for retroviruses is how they transport unspliced genome-length RNA to the cytoplasm within infected cells. The post-transcriptional events of RNA include the presence of the regulated assembly of the ribonucleoprotein complex (RNP) involved in post-transcriptional, processing, transport and translation events. The expression of viral proteins from unspliced, incompletely spliced and fully spliced transcripts result essential for retroviruses to develop strategies of control the extent of RNA splicing. However, the wide range of retroviruses comprises different mechanisms of RNA export and transport that involve viral *cis* and *trans*-acting elements. The simple retroviruses present *cis*-acting RNA sequences that recruit cellular *trans*-acting factors, and also *trans*-acting viral factors. However, complex retroviruses have evolved unique mechanisms to export unspliced mRNAs using viral proteins that interact with cellular RNA transport machinery.

The majority of cellular mRNA export is facilitated by the recruitment of the export factor NXF1 (Nxf1-Nxt1/Tap-p15) via TREX-1, recruited in a splicing dependent manner to allow the mRNA to traverse the nuclear pore complex (NPC) (Viphakone et al. 2012). There are adapter proteins such as CRM1, which is involved in RNA export through the NPC to the cytoplasm and SR proteins that recruit Nxf1 to be spliced mRNA in human cells (Huang et al. 2004). Some retroviruses like HIV use the CRM1 pathway, forming a Rev-RRE-CRM1 complex that permits full length viral RNAs to be exported towards the cytoplasm.

### 1.7.1 Constitutive transport element (CTE)

The CTEs are highly conserved *cis*-acting elements encoded in simian *Betaretroviruses* that stimulate nuclear export of spliced mRNAs (Ernst et al. 1997). The major representative of this group is the Mason-Pfizer monkey virus (MPMV), which uses the CTEs (located between the *env* gene and the 3' LTR) to export efficiently its viral mRNA within the host cell. The CTEs represent inverted direct repeat sequences that adopt stable stem-loop conformations, where the nucleotide sequence is crucial (Rizvi et al. 1996). The absence of *trans*-acting RNA regulatory proteins in MPMV suggests the likely existence of different host requirements for the HIV-1 Rev system and the CTE. Moreover, the CTE can promote the export of HIV-1 derived RNAs dependent on Rev-RRE (Ernst et al. 1997;

Wodrich et al. 2001). Thus, the CTE binds to the cellular *trans*-acting factor NXF1/Tap, which directs nuclear export of the RNA- protein complex to the cytoplasm.

### 1.7.2. DR1 and DR2 elements of avian retroviruses

The presence of *cis*-acting elements are also present in other retroviruses such as the Rous sarcoma virus (RSV), which encodes two repeated regions flanking the *src* gene: the downstream region called DR2 and a second repeat located upstream (DR1). Both repeated regions are involved in viral RNA transport and are conserved in other avian retroviruses (Czernilofsky et al. 1983). The *cis*-acting DR elements manage retroviral RNAs to one export pathway or another, and the absence of DR elements severely impairs RSV replication in avian cells, but some unspliced RNAs are exported and translated in their absence (Simpson et al. 1997). Moreover, the lack of DR elements negatively affects the levels of packaged virion RNA, impairing cytoplasmic accumulation of unspliced viral RNA that is necessary for synthesis of virion proteins and packaging of viral RNA (Ogert et al. 1996). Thus, avian leukosis virus (ALV) particles can be formed in mammalian cells when the HIV-1 RRE is provided (in *cis*) and Rev (in *trans*), allowing the RNA export via the CRM1 pathway (Nasioulas et al. 1994).

### 1.7.3 Jaagsiekte sheep virus (JSRV)

The JSRV is a simple betaretrovirus encoding a regulatory protein called Rej, which is responsible for the correct synthesis of Gag polyprotein from unspliced RNA and required for translation of unspliced viral RNA (Nitta et al. 2009). The Rej protein interacts with a Rej-responsive element (RejRE) located at the end of the 3'end of *env* gene, which is involved in viral RNA transport and cytoplasmic accumulation within infected cells. This mechanism of RNA transport reminds the HIV-1 RNA transport pathway via Rev-RRE interaction.

### 1.7.4 Regulator of export of mouse mammary tumor virus (MMTV)

The mouse mammary tumor virus (MMTV) encodes a functional homolog of the HIV-1 nuclear export trans-activator Rev that is also involved in nuclear transport of viral RNA. Functional homologs have been founded in other betaretrovirus such as the human endogenous retrovirus K (HERV-K). The regulatory protein from MMTV was termed Rem (Indik et al 2008; Dudley et al. 2005). Moreover, the location of Rem within the

MMTV genome correlates to other Rev-like proteins and can also replace the Rev/RRE complex. However, Rem requires the presence of MMTV env-U3 junction and the CRM1 pathway to be functional (Cullen 2003).

#### 1.7.5 HERV-K

The HERV-K encodes a nuclear RNA export factor termed K-Rev that is the functional equivalent of HIV Rev protein. Thus, the K-Rev protein interacts with a *cis*-acting element called K-RRE that allows the nuclear export of viral RNA, and its sequence is conserved across known HERV-K genomes (Yang et al. 2000). Moreover, the HERV family also encodes an accessory protein termed Rec (Lower et al. 1995) whose functions are similar to those of Rem in mouse mammary tumor virus (MMTV) or Rev in human immunodeficiency virus (HIV).

#### 1.7.6 FIV Rev

There are sequence similarities between the C-terminal domain of HIV-1 RRE and FIV RRE according to the nuclear export signal. In primate lentiviruses this is a conserved 11 amino acid, leucine rich region that mediates binding to CRM1, but in FIV the region is longer, and the leucines are spaced further apart and are interspersed with other hydrophobic residues (Dong et al. 2009). However, the mechanism of RNA export in FIV is closely related to HIV-1 pathway where the generation of the Rev-RRE-CRM1 complex is essential to transport unspliced viral RNAs outside the nucleus.

#### 1.7.7 HTLV

The human T-cell leukemia virus (HTLV) encodes a viral protein Rex that allows the transport of genome-length unspliced mRNA and partially spliced mRNAs to be exported from the host nucleus towards the cytoplasm by an interaction with a *cis*-acting element (RxRE) within the viral genome (Younis et al. 2005). Thus, Rex is regulated by phosphorylation to ensure efficient viral gene expression and provide the best conditions for virion production. Some RNA species can be exported to the cytoplasm in a Rex independent manner, but this fact severely affect viral mRNA trans-activation.

## 1.8 Antiretroviral drugs

The current antiretroviral treatments fail to eradicate HIV infection in patients despite the combination of different drugs targeting several proteins of the virus. However, the HAARTs increase lifespan from infected patients and also prevent HIV transmission by controlling the viral load (reducing the plasma viremia) to preserve the immunologic function. Generally the best therapy for HIV infection comprises a combination of several classes of drugs such as integrase, protease and reverse transcriptase inhibitors and fusion inhibitors to counteract viral infection. Current HIV therapies are designed to target different viral proteins to counteract HIV infection in different steps of the replication cycle simultaneously. The reverse transcriptase inhibitors (RTIs) target the viral RT, impairing the transformation of the viral RNA genome into dsDNA within infected cells.

The RTIs comprise two groups of analogues of deoxynucleotides that interfere with the viral dsDNA synthesis during reverse transcription: nucleoside analog reverse-transcriptase inhibitors (NARTIs), nucleotide analog reverse-transcriptase inhibitors (NtARTs) and non-nucleoside analog reverse-transcriptase inhibitors (NNRTIs). The NARTIs and NNRTIs compete with the natural deoxynucleotides to be incorporated in the growing viral DNA chain. Thus, the NARTIs lack a 3'-hydroxyl group on the deoxyribose moiety and after their incorporation into the nascent DNA strand impede the next incoming deoxynucleotide, which cannot form the next 5'-3' phosphodiester bond needed to continue the extension of the DNA chain. The NARTIs also impair cellular DNA synthesis acting as chain terminator for viral and cellular DNAs, resulting in the appearance of toxicity and side effects. However, the NNRTIs have a different mode of action, blocking reverse transcriptase binding directly to the viral enzyme and blocking its capacity to perform reverse transcription.

Moreover, the HIV protease inhibitors represent one of the most important antiretroviral groups due to lower level of resistance compared to NNRTI therapy (Riddler et al. 2008). The protease inhibitors interact with the HIV protease to impede its function, which is essential in viral maturation. However, the HIV protease inhibitors induce side effects at long-term such as dyslipidemia or cardiovascular diseases (Bozzette et al. 2003). These side effects have been reduced in new protease inhibitors such as RO1, which also has about 5 to 10 higher fold efficiency compared to other protease inhibitors.

Other drugs employed in HIV therapies are the fusion inhibitors that block the HIV envelope from merging with the host CD4 cell membrane, preventing HIV from entering the CD4 cells. The fusion inhibitors are used in combination therapy when other treatments previously failed. Enfuvirtide is a peptide based on the HIV gp41 sequence that prevents the creation of an entry pore for the capsid of the virus by gp41 direct interaction. Moreover, Enfuvirtide is the only approved fusion inhibitor for clinical usage, and prevents membrane fusion by competitively binding to gp41 and blocking the formation of the post-fusion structure (Lalezari et al. 2003). There are also CCR5 antagonist such as Maraviroc to impede HIV entry. Finally, integrase inhibitors target the viral integrase and are divided in two groups: integrase strand transfer inhibitors (INSTIs) and integrase binding inhibitors (INBIs). The INSTIs are currently used in clinical therapies, blocking the insertion of viral dsDNA into the host cell DNA, meanwhile the INBIs interact directly with the viral DNA-IN formation binding directly to the complex, but they still experimental.

## 1.9 HIV-1 animal models for research

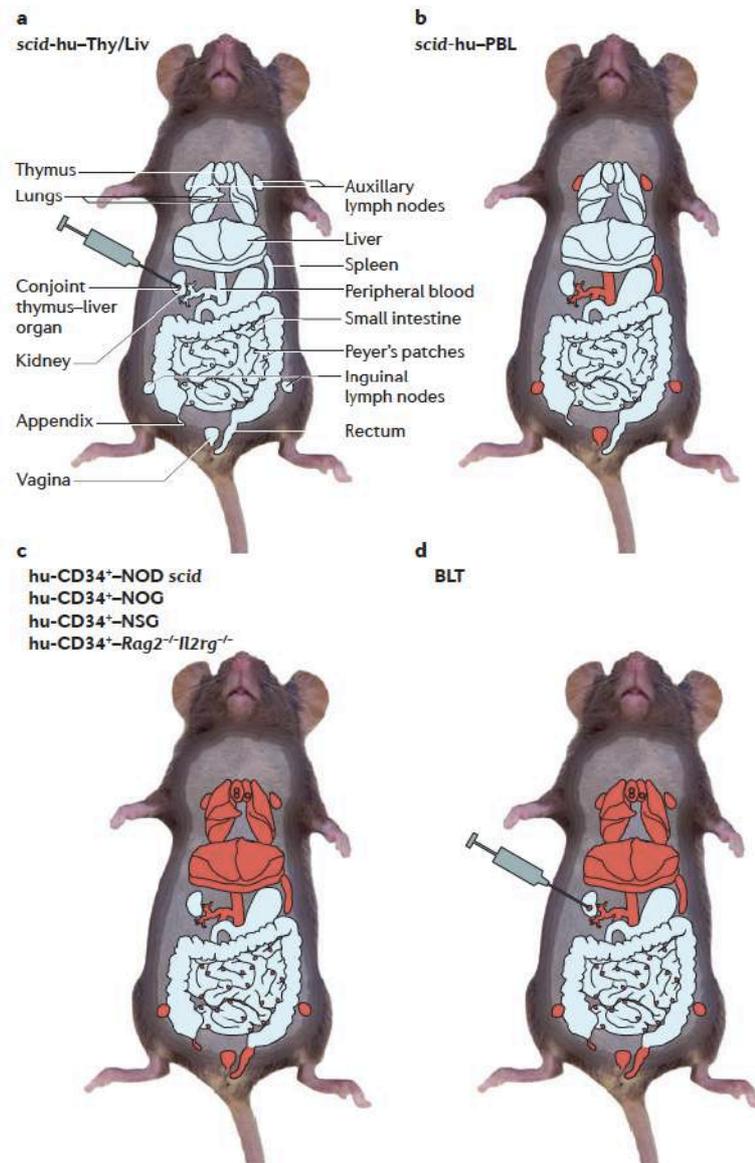
The development of an effective vaccine against HIV-1 represents a difficult and ambitious aim, with limited reported success despite 20 years of intensive research (Finzi et al. 1999; Whitney et al. 2004; Burton et al. 2017). Currently, there is not an appropriate animal model to study HIV-1 *in vivo*, although humanized mice are able to allow HIV-1 replication and other macaque models exist (Borkow 2005). Simulating the conditions of HIV infection in animal models is a hard process because current animal models are unable to reproduce HIV-1 infection *in vivo*.

### 1.9.1 Humanized mice

Humanized mouse models have been used for HIV-1 research, but they require surgical procedures to implant human cells to make them susceptible to HIV infection. Thus, they cannot be bred, and murine cells are unable to sustain viral replication under normal conditions (Feinberg et al. 2002). Moreover, the RNA instability of HIV-1 severely affects the synthesis of Gag precursors in murine cells, and Tat is inactive because Tat requires the presence of human factors not present in mice to transactivate the LTR. Alternatively, there may be additional, as yet unidentified, factors that restrict HIV-1 replication in murine cells. Transgenic mice are susceptible to infection, but HIV-1 cannot replicate in murine CD4 T cells and the viral loads are almost undetectable in these animals (Hatzijioannou et al. 2012). The knockout mice lacking HIV-1 restriction factors have revealed new insights about mechanisms to counteract the infection and the discovery of novel restriction factors such as tetherin (Neil et al. 2008; Liberatore et al. 2011). However, the current humanized mice models still have many limitations to approach, such as the limitation to determine the potential of drugs that could prevent vaginal HIV-1 transmission, or the persistence of murine innate immunity (Vaselinovic et al. 2014).

No humanized mice models have demonstrated humanization of the gut mucosa, and rectal transmission effects of HIV-1 infection on human cells in the GALT can not be evaluated using humanized bone marrow liver thymus (BLT) mice models (Denton et al. 2009). Despite HIV research in humanized mice has entered new insights and manners of investigation they present limitations such as the lack of a full repertoire of human

immune cells and a functioning immune system. Thus, humanized mice need surgical procedures to implant human cells, but the innate immunity from mice persist and humanized mice are unable to generate humoral immune responses that lead to class switching from immunoglobulin M (IgM) to immunoglobulin G (IgG) antibody production after infection (Akkina 2013). In addition, engrafted humanized mice poorly develop lymph nodes and germinal centers (Yu et al. 2017; Brehm et al. 2013).



**Figure 1-14.** Human immune system mouse models. These animal models are currently used in cancer research, but they are also useful to investigate how the immune system reacts against viral infections *in vivo*. The conjoint human thymus–liver organ formed following injection of human fetal liver and thymus cells into the mouse renal capsule. The engraftment is achieved by injecting human hematopoietic (CD34+) stem cells derived from cord blood, fetal liver or adult blood. These mice produce human immune cells that disseminate to numerous sites, among which are the peripheral blood, liver, lung, vagina and rectum. Diagram Taken from Theodora Hatzijioannou and David T. Evans (2012). Animal models for HIV/AIDS research. *Nature Reviews Microbiology*. 2012 Dec; 10(12): 852-867.doi: 10.1038/nrmicro2911

### 1.9.2 Macaques and SHIVs

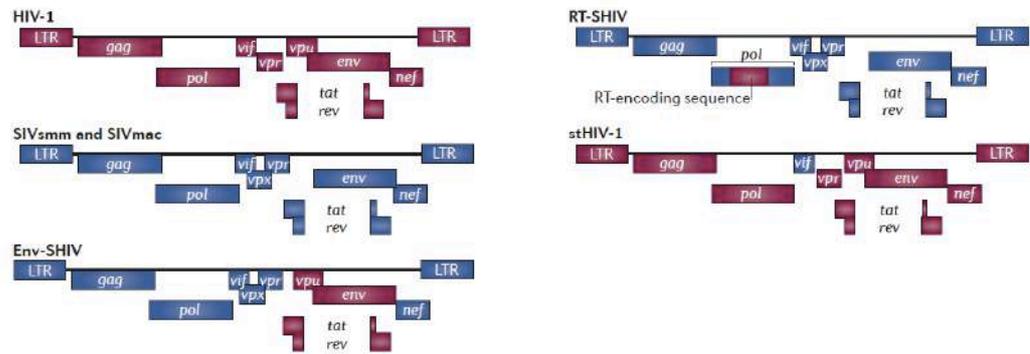
There are other alternatives to humanized mice such as macaques, which currently represent the ‘best’ models to investigate HIV-1 *in vivo*. The stHIV strains (substitution of SIV genes with those from HIV-1) are currently used in macaques, reflecting more accurately the features of the virus responsible for the AIDS. Rhesus macaques are normally non-permissive due to TRIM5 (Stremlau et al. 2004). TRIM5 $\alpha$  is a post-entry restriction factor that is constitutively expressed in primate cells, and upregulated by type I interferon (IFN) (Nisole et al. 2008). Thus, old world monkey cells express a fusion protein TRIM5 generated by transposition of a cyclophilin A mRNA within a TRIM5 locus named TRIMCyp, which specifically interferes with HIV-1 uncoating process. However, current investigations have not determined the induction of TRIMCyp by IFN. Therefore, pig-tailed macaques are commonly used in HIV-1 research, but require CD8 depletion to progress to simian AIDS (Joag et al. 1996).

However, rhesus and pig-tailed macaques infected with simian immunodeficiency virus (SIV) or SIV/HIV chimeras encoding (SHIVs) are the most representative animal models (Hatzijioannou et al. 2009). There are several types of SHIVs, classified by the HIV-1 gene substitution: SHIVs expressing HIV-1 *pol*, SHIVs expressing HIV-1 *env* and simian tropic HIV-1 (stSHIVs, HIV-1 engineered to replicate in macaques). The stSHIV represent a useful tool to overcome intrinsic barriers for viral replication, but more intensive research is essential to approach the entire potential of stHIV-1 strains to transform non-human primates for HIV-1 vaccine development. The SHIV models have been useful to determine pathogenesis and transmission of HIV-1, and also to test vaccines in infected animals. However, there are limitations to SHIV infection in macaques, such as the low genetic diversity that HIV-1 subtype B envelope displays in these animals and the presence of the HIV-1 regulatory genes *tat* and *rev* (Ndungu et al. 2001), which are encompassed by the *env* gene. Moreover, the SHIVs lack the HIV-1 gene *vpu*, having *vpx* instead, which counteracts the restriction factor SAMHD1 present in macrophages and dendritic cells (Hrecka et al. 2011) altering the cell tropism relative to HIV-1. SAMHD1 is a deoxynucleoside triphosphohydrolase expressed in myeloid cells that hydrolyzes the cellular deoxynucleotide triphosphates to a level below that which is required for reverse transcription. The Vpx protein recruits SAMHD1 to a cullin4A-RING E3 ubiquitin ligase (CRL4), which targets the enzyme for proteasomal degradation (Hofmann et al. 2012).

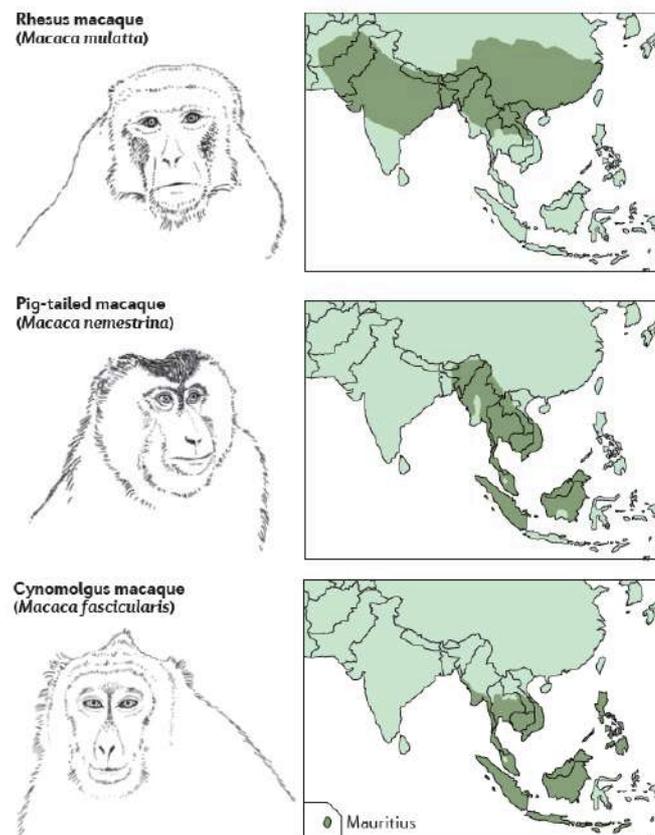
The huge genetic variability of HIV-1 has hampered approaches to develop a vaccine, despite the advances made in primate models (Kimata 2014). A major problem remains the absence of a cheap and simple animal model capable of determining whether a vaccine could be curative against HIV-1 infection in humans. Although therapeutic control of HIV replication allows the immune system to partially restore and delays disease progression, the cure of HIV infection remains still unachievable. However, there is an HIV-infected patient with acute myeloid leukemia that received two stem cell transplants in 2007 and 2008 from a donor naturally immune to HIV, and has remained off antiretroviral therapy since the transplant. This patient is considered cured of HIV-1 infection, but still raising controversial opinions whether there is no trace of the virus in his body or whether he simply no longer needs treatment (Allers et al. 2018). The HIV-infected patients homozygous for CCR5 gene variant (CCR5 $\Delta$ 32/ $\Delta$ 32) are naturally resistant to CCR5-tropic strains, but the cured patient presented CXCR4-tropic variants before the transplant, so it is reasonable to hypothesize that HIV from the viral reservoir may reseed the body once the immune system has efficiently been restored with X4 HIV-susceptible target cells (Hütter et al. 2009).

However, the success of HIV-1 curative therapies requires the usage of animal models, but their current limitations severely impede the potential of vaccine development to cure HIV infection. Perhaps we could transfer the discoveries from the ‘cured’ patient into animal models to investigate and reveal new insights about a cure that we could export to humans.

A



B



**Figure 1-15.** A) Schematic representation of HIV-1, SIV and SHIVs viral genomes. The HIV-1 genes are coloured in red, and the SIV genes in blue. B) Macaque species currently used in AIDS research and their geographical location. The geographical range of the rhesus macaque (*Macaca mulatta*) exceeds that of all other primate species except humans, extending from western India and Pakistan across China. Distinct populations of rhesus macaques can be differentiated on the basis of mitochondrial DNA sequences or SNP. Figures taken from Hatzioannou, T., & Evans, D. T. (2012). Animal models for HIV/AIDS research. *Nature Reviews Microbiology*, 10(12), 852–867. <https://doi.org/10.1038/nrmicro2911>

### 1.10 Lentiviral vectors and self inactivating vectors (SIN)

A self-inactivating (SIN) vector is a retroviral derived vector containing a non-functional or modified 3' Long Terminal Repeat (LTR) sequence. This sequence is copied to the 5' end of the vector genome during reverse transcription, resulting in a transcriptional inactivation of the provirus after the first cycle of replication (Yu et al. 1986). The SIN vectors were generated to develop gene transfer experiments in order to investigate the regulation of gene expression, also providing a safe alternative to be employed for gene therapy. The recombination events required for these vectors to develop a replication competent genome during the first cycle of replication (Gage et al. 1998) are unlikely, and SIN lentiviral vectors can transduce non-dividing cells, integrating into the host chromosome.

There were other lentiviral vectors used before SIN vectors were created: first and second generation lentiviral vectors. This classification is based on the safety features that each type of vector possesses, where the essential genes required for virus production are split across several plasmids (3 for 2<sup>nd</sup> generation and 4 for 3<sup>rd</sup> generation vectors). Each lentiviral vector encodes the following components: a lentiviral transfer plasmid (the vector or vector-genome) encoding the insert of interest, a packaging plasmid and an envelope. The generation of lentiviral vectors could be potentially used to approach novel human gene therapies and vaccines. Moreover, the use of lentiviral and retroviral vectors for human gene therapy has been supported by preclinical and clinical trials (Bif et al. 2013). However, lentiviral vector safety is still uncertain and they could upregulate the expression of genes flanking the integration site, especially when carrying strong enhancer/promoter sequences in internal positions (Ramezani et al. 2008) or cause insertional mutagenesis. Three generations of HIV-based lentiviral vector packaging systems have been developed for production of lentivectors by transient transfections, where the structural and enzymatic components of the virion come from HIV-1. However, the envelope glycoprotein is derived from vesicular stomatitis virus (VSV) due to the high stability and pantropic features of its G protein.

The 1<sup>st</sup> generation of lentiviral vectors encompasses all HIV-1 genes except *env* gene, and the 2<sup>nd</sup> generation of lentiviral vectors lack HIV-1 accessory genes from the packaging plasmid (*nef*, *vpr*, *vpu* and *vif*), increasing the safety threshold compared to the first

generation of lentiviral vectors. The regulatory genes *tat* and *rev* are also encoded, fulfilling transcriptional and post-transcriptional functions.

The third generation of lentiviral vectors provide an improved biosafety compared to second and first lentiviral vectors, only encoding HIV-1 *gag*, *pol* and *rev* genes, and the 5'LTR HIV promoter has been substituted by a chimeric promoter to ensure transcription in the absence of Tat. However, it is easier to work with one packaging plasmid such as the pCVR1\_NLGP, which encodes the HIV-1 *gag*, *pol*, *tat*, and *rev* genes. Thus, there are self inactivated vectors (SIN) that present a deletion in the enhancer region of the 3'U3 of the LTR, generating an aberrant LTR at the 5' U3 end after reverse transcription during the first cycle of replication. This feature reduces the likelihood of SIN vectors recombining (compared to first and second generation of lentiviral vectors) and also makes them much safer. Crucially, the resultant provirus after RT is transcriptionally inactive and unable to be converted into a full length RNA. In addition, the 3<sup>rd</sup> generation of lentiviral vectors reduces the risk of tumorigenesis via promoter insertion (Cesana et al. 2014) and reduces the probability of regenerating a replication competent lentivirus (RCL). Thus, this system efficiently generates stable packaging cell lines with selected characteristics.

In summary, animal models are a powerful tool to investigate and develop novel approaches to eradicate the virus from infected patients despite their limitations. The aim of this project was to develop an HIV-1 replication competent clone able to replicate in murine cells to overcome the limitations in current animal models preserving the 'identity' of HIV. The development of an HIV-1 clone able to replicate in mice is a challenging proposal. The lack of human co-factors in mice impedes function of the HIV-1 accessory proteins Tat and Rev, hampering HIV-1 replication. However, Tat and Rev function can be replaced by constitutive/chimeric promoters, codon-optimized proteins and the constitutive transport element (CTE), without disrupting the amino acid sequence of the virus. We propose the generation of an HIV-1 clone able to replicate in mice to assist in antiviral drug development by minimally manipulating the genomic 'identity' of the virus. Moreover, changing the current perspective of adapting the host to the virus (current animal models) instead adapting the virus to the host (our aim) could solve most of the issues related to mice and macaques models, contributing to the generation of an HIV-1 clone that could be used to develop curative vaccines and therapies in animal models.

## **Chapter 2**

### **Materials and methods**

## 2. Materials

### 2.1 Cell culture

#### 2.1.1. Bacterial strains

The bacterial strain employed in this project for maintenance and growth of plasmid stocks was *Escherichia coli* DH10B. Competent DH10B bacteria were generated in the lab using a protocol described in section 2.4.6.1.

#### 2.1.2 Antibiotics

All antibiotics employed in this thesis were provided by Melford Laboratories. Ampicillin (200mg/ml stock concentration) and Kanamycin (100mg/ml stock concentration) were used for bacterial selection. Puromycin and Neomycin (G418) were used to select transduced cell lines. Gentamicin (20µg/ml final concentration) was added into all cell culture media to prevent contamination in the cell lines employed in this project.

#### 2.1.3 Eukaryotic cell lines

HEK-293T: cell line originally derived from human embryonic kidney cells. They are highly transfectable and useful to generate a variety of viral particles. This “packaging cell line” is cultured in DMEM (Dulbecco’s modified Eagle’s medium) with 10% fetal calf serum (FCS) and gentamicin [20µg/ml final concentration]. Reference: ATCC cell data sheet HEK-293T.

TZM-bl: HeLa cells expressing HIV-1 receptor (CD4) and co-receptors (CCR5/CXCR4). These cells are highly permissive to infection by HIV, SIV and SHIV, and contain an integrated reporter gene for firefly luciferase and *E. coli* β-galactosidase under control of an HIV-1 LTR, which enables simple and quantitative analysis of Tat expression using either β-gal or luciferase as a reporter. They are maximally sensitive to HIV infection by including DEAE-dextran (10 and 100 µg/ml final concentration) in the infection medium (adapted from NIH AIDS Reagents data sheet). TZM-bl are cultured in DMEM with 10% fetal calf serum (FCS) and gentamicin [20µg /ml final concentration]. Reference: *Montefiori - Protocol for Measuring Neutralizing Antibodies Against HIV-1, SIV and SHIV Using a Luciferase Reporter Gene Assay in TZM-BL Cells (2006)*.

NIH-3T3: cell line established from primary mouse embryonic fibroblast cells which are sensitive to murine leukaemia virus. NIH-3T3s a suitable transfection host and highly

sensitive to sarcoma virus focus formation and leukaemia virus propagation and has proven to be very useful in DNA transfection studies. Cultured in DMEM with 10% fetal calf serum (FCS) and gentamicin [20 µg/ml final concentration]. Reference: ATCC cell data sheet NIH-3T3.

p/m40 AK: murine cells derived from a thymic lymphoma that grew out of a p53 null mouse. AK are the initials of Anna Kilbey PhD, who made the cell line. They are cultured in RPMI (Roswell Park Memorial Institute medium) with 50µM of β-mercaptoethanol (final concentration), 10% fetal calf serum (FCS) and gentamicin [20µg/ml final concentration]. Reference: *Blyth et al., 1995 Oncogene 10: 1717-1723*.

PM1: clonal derivative of HUT 78. These cells express HIV-1 CD4 receptor, CXCR4 co-receptor and also CCR5 receptor at lower levels. PM1 are permissive for growth of macrophage and T-cell trophic viruses. Cultured in RPMI with 10% fetal calf serum (FCS) and gentamicin [20µg /ml final concentration]. Taken from NIH AIDS Reagents data sheet.

AA-2: AK-dCK subclone of AA cells derived from the WIL-2 human splenic EBV +B lymphoblastoid cell line. AA-2 are an isolated from the spleen of a Caucasian male with hereditary spherocytic anemia, expressing high levels of HIV CD4 receptor. This cell line is remarkably permissive for HIV-1 infection, extremely sensitive to virus cytopathic effects and highly permissive for HIV infection. Cultured in RPMI with 10% fetal calf serum (FCS) and gentamicin [20µg /ml final concentration]. Reference: NIH AIDS Reagents data sheet.

RAW 264.7: murine macrophages. This line was established from a tumour induced by Abelson murine leukaemia virus. High efficiency for DNA transfection, sensitivity to RNA interference, and supports replication of murine noroviruses. When this line was established, it was described as not secreting detectable virus particles and negative using the XC plaque formation assay. Based on a published study by Dr. Janet W. Hartley in 2008, this line was demonstrated to express ecotropic and polytropic MuLV, and is positive using the XC plaque assay for virus replication. The base medium for this cell line is DMEM plus fetal calf serum (10%) and gentamicin [20µg /ml final concentration]. Reference: ATCC cell data sheet RAW 264.7.

## **2.2 Molecular biology**

### 2.2.1 Oligonucleotides

The synthetic oligonucleotides were provided by Integrated DNA Technologies at 25nM scale with standard desalt purification. Oligonucleotides were designed using DNA Dynamo Sequence Analysis Software. Sequences are shown in Table 2.1.

### 2.2.2 Enzymes

Standard PCRs were carried out using Pfu Turbo DNA Polymerase® (Agilent Technologies, 600250), and Brilliant III Ultra Fast QPCR Master Mix® (Agilent 600880) was used for qPCR assays. Reverse transcription assays were developed using Superscript III® (Invitrogen 18080044) and also the RNase Out Ribonuclease Inhibitor (Invitrogen 10777.019).

**Table 2.1.** Oligonucleotides

NAME	SEQUENCE
WV-001-3LTR-NHG-XhoI-F	CTCTCTCTCGAGTGGAAAGGGCTAATTCCTACTCCAACG
WV-002-3LTR-U5-NotI-RC	AGAGAGGGCGCCGCTGCTAGAGATTTTCCACACTGACTAAAAGGG
WV-003-3LTR-PGK-XhoI-F	CTCTCTCTCGAATAATTCTAGGGTAGGGGAGGGCGCTTTTCCAAGGG
WV-004-3LTR-CMV-XhoI-F	CTCTCTCTCGAGGTTGACATTGATTATTGACTAGTTATTAATAG
WV-005-3LTR-CMV-R-NotI-RC	AGAGAGGGCGCCGCTGAAGCACTCAAGGCAAGCTTTATTGAGGC
WV-006-deltat-1	AAGCAAAGATCATCAGGGATTATGG
WV-007-deltat-2	TTTGTATGAAACAACTTAAGCTTAGGCTCATTACTGGATGCTACCAGGGCTTAGT
WV-007-deltat-2	TTTGTATGAAACAACTTAAGCTTAGGCTCATTACTGGATGCTACCAGGGCTTAGT
WV-008-deltat-3	ATCCAGTAAGTGAGCCTAAGCTTAAGTTTGTTCATAACAAAAGCCTTAGGCATCTCT
WV-009-deltat-4	ACACATGGCTTTAGGCTTTGATCCC
WV-010-tat-F	CTCTCTGGCCGAGAGGGCCATGGAGCCAGTAGATCTAGACTAG
WV-011-Tat RC	TCTCGGCCAGAGAGGCCCTATTCCTTCGGGCTGTCCGGTCCCCTCGGGATTGGGAGGTGGGT GCTTTGATAGAGAACTTGATGAG
WV-012-deltat 2-II	TTCCAGGGCTTTATCTAGGATCTACTGGCTCCGTTTCTGTCTCTCTGTGCGAGTAAC
WV-013-deltat 3-II	GCAAGAAACGGAGCCAGTAGATCTAGATAAAGCCCTGGAAGCATCCAGGAAGTCAGCC
WV-014-seq LTR F	GACCGCCGCGGGATCACTCTCGGC
WV-015-3LTR-F	GCGCGATCACATGGTCTGCTGGAGTTCG
WV-015-seq LTR F2	GCGCGATCACATGGTCTGCTGGAGTTCG
WV-016-SYNGP_seq-RC	GCCACGGTGTGTACAGGCTGCGCAGCTCCTCGC
WV-017-SYNGP_seq-F	GCGGATCGTGGACATCATCGCGACCGACATCCAGAC
WV-018-CLTR-Eco-RC	AGAGAGGAATTCTGTAGAGATTTTCCACACTGACTAAAAG
WV-019-CMV-NotI-F	CTCTCTGCGCCGCGTTGACATTGATTATTGACTAGTTATTAATAG
WV-020-CLTR-HpaI-RC	AGAGAGGTTAACCTGTAGAGATTTTCCACACTGACTAAAAG
WV-021-5CLTR seq	GGCCTTTGTCTCACATGGCTCGACAG
WV-022-3CLTR seq RC	GCCAGCCCGGATCGATCCTTATCCG
WV-023-IRES_GFP-F	CTCTGCGGCCGCTCGAGCCGCCCTCTCCCTCCCCCCCCCTAAC
WV-024-IRES_GFP-RC	CTCTCTGCGGCCGCGAATTCCTGATTATGATCTAGAGTCGCGG
WV-025-gag.1-NdeI-F	GGCAGTACATCAAGTGTATCATATGCCAAGTACG
WV-026-gag 2"-RC	GCCGCCGACAGTACCAGCTCTCGCACCCATCTCTCTCTTCTAGCCTCCGCTAGTC
WV-027-gag 3"-F	TAGAAGGAGAGAGATGGGTGCGAGAGCGTGGTACTGTCTGGGCGGCGAGCTGGACCGCT
WV-028-gag 4-MluI-RC	TTTCTAGGGGCCCTGCAGTTGCGG
WV-029-env seq 5' F	GGCGTAATCTGCTGCTTGCACAAACAAAAAAC
WV-030-env seq 5' RC	CTCAAGGATCTTACCCTGTTGAGATCCAG
WV-031-signal weave sq-F	CGCGTTACATAACTTACGGTAAATGGCCCGC
WV-032-signal weave sq-RC	GAGGGGTGCTGCGCAAGAGTGTACTGAG
WV-033-signal weave sq-F"	GCAGTGGCGCCGAACAGGGACTTGAAAGC
WV-034-signal weave sq-RC"	CGTTTTGTAGAACCGGTCCACGTAGTCGCG
WV-035-psi seq	GCCTTCCCTTGTGGGAAGGCCAGATCTTCC
WV-036-psi seq RC	GTCGAGAGAGCTCCTCTGGTTTCCC
WV-038	CTCTAGCAGAAATTCGCCACCATTGGGCGCCGTTAAC
WV-039-IRES-GFP.NotI F	GCGGCCGCTTCAAACTTTCAGTTCGACGGTACCAGC
WV-040-3CLTR XhoI.NotI RC	CTCTCTCTGCGGCCGCTCGAGTTTACTTGTACAGCTCGTCCATGCCGAGAG
WV-041-IRES-EGFP NotI F	CTCTCTGCGGCCGCGCCGCGGCTCCGCCCTCTCCCTC
WV-042-IRES-EGFP XhoI NotI RC	CTCTCTGCGGCCGCTCGAGTTACTTGTACAGCTCGTCCATGCCGAG
WV-043-3CLTR XhoI F	CTCTCTCTCGAGTCAATATTGGCCATTAGCCATATTATTC
WV-044-IRES-EGP seq F	CAGAATTTCCGGTCTACTACAGGGACAGCAG
WV-045-3CLTR RC seq	GGCGATCGGTGCGGGCTCTTCGCTATTAC
WV-046-co env seq 1	CTGAGTGGGTGGAGACTGAAGTTAGGCCAGC
WV-047-co env seq 2	CATCACACCAGCATCAGGGATGAGGTGCAG
WV-048-co env seq 3	CATCAACATGTGCGAGGAGGTGGCAAGGC
WV-049-co env seq 4	CACGTCCACCTCCATCTCTTCTCAGCACC
WV-050-MLV/HIV LTR 1 XhoI	CTCTCTCTCGAGTGGAAAGGGCTAATTTGGTCCAAAAAGACAAG
WV-051-MLV/HIV LTR 2	CCTGACCTTGATCTGAACCTTCTCTTCTTTGTAGTACTCCGGATGCAGCTC
WV-052-MLV/HIV LTR 3	AGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCC
WV-053-MLV/HIV LTR 4	CACTCCCCAGTCCCGCCAGGCCACAAGCAGAAGCGCGCAACAGAAGCGAG
WV-054-MLV/HIV LTR 5	TGTGGCTTGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGTACATATAAGC
WV-055-MLV/HIV LTR 6 XbaI	GAGAGAGTCTAGAGTCCCTGTTCCGGGCGCCAC
WV-056-restriction sites psi-IG-3CLTR-F	TCGAGTTAATTAACCTGCAGGGCGCGCCGCGATCGCT
WV-057-restriction sites psi-IG-3CLTR-RC	TCGAAGCGATCGGGCGGCCCTGCAGGTTAATTAAC
WV-058-CTE-F	CTCTCTGTGCACTTAATTAAGGATCCACCTCCCCTGTGAGCTAGACTGGAC
WV-059 CTE-RC	AGAGAGGTCGACCCTGCAGGGCTAGCAGATCTACACATCCCCTCGGAGGCTGCGCCTG
WV-059 CTE-RC	AGAGAGGTCGACCCTGCAGGGCTAGCAGATCTACACATCCCCTCGGAGGCTGCGCCTG
WV-060-MLV/HIV LTR 6 NotI	AGAGAGGGCGCCGCTCCCTGTTCCGGGCGCCAC
WV-061 seq SYNGP 01	CTGTGCGGTTTCGCCACCTCTGACTTTGAG
WV-062 seq SYNGP 02	CGAGTCTCGCTCGAGAGAGCTCCTCT
WV-063 seq SYNGP 03	GTCAGTGTGAAAAATCTCTAGCAGTG
WV-064 seq SYNGP 04	GGATCGTGTGCATCCGGTGCACGCAG
WV-065 seq SYNGP 05	GGCTGTTGAAATGTGAAAGGAAGG
WV-066 seq SYNGP 06	GATTGGGCTGAGAACCCGTACAACAC
WV-067 seq SYNGP 07	GTTGAACCTGGGCGAGTCAGATTTACCCAG
WV-068 seq SYNGP 08	CACCACCAACCAGAAGACTGAGCTGCAGG
WV-069 seq SYNGP 09	CAAGGTTATCTGGTAGCCCTCCATG
WV-070 seq SYNGP 010	GCTCTCTCAAGCGTATTCAACAAGGG
WV-071 seq SYNGP 011	CAGCCACAACGTCTATATCATGGCCGAC
WV-072 seq SYNGP 013	GCAATTGTTGTTGTTAACTGTAGAG
WV-073 seq SYNGP 014	GACCGAGATAGGGTTGAGTGTGTTTC
WV-074 seq SYNGP 015	CTGGAATAGCTCAGAGGCCGAGGCG
WV-075 seq SYNGP 016	CTCTTCTCTCAGATCATCTCTGATCCAG

WV-076 seq SYNGP 017	GAGACGGTCACAGCTTGTCTGTAAGC
WV-077 seq SYNGP 018	CTGTGACTGGTGAGTACTCAACCAAGTC
WV-078 seq SYNGP 019	CAGTGAGGCACCTATCTCAGCGATCTGTGTC
WV-079 seq SYNGP 020	CTCATAGCTCACGCTGTAGGTATCTCAG
WV-080 seq SYNGP 021	CGATAGTTACCGGATAAAGGGCAGCG
WV-081 seq SYNGP 022 CTE-POLY	TCAATAATCAATGTCAACAGCGATC
WV-082-CSGW delta NheI 01	CTCTCTGAATTCCTGCAGCCCCGATAAAAATAAAAG
WV-083-CSGW delta NheI 02	AAGACCCACCTGTAGGTTATGGCAAGCTGGCTGCAGTAACGCCATTATTGCAAGGCAT
WV-084-CSGW delta NheI 03	ATGCCTTGCATAAATGGCGTTACTGCAGCCAGCTTGCCATAACCTACAGTGGGGTCTT
WV-085 CSGW delta NheI 04	AGAGAGGGATCCCCCGGGCGACTCAGTCTGTGC
WV-086 CSGW HpaI	CTCCCCCTGAACCTGAAACATAAAAATGAATGC
WV-087 CSGW SpeI RC	TCTTTGGGAGTGAATTAGCCCTTCCAACTAGTAAGTTGTAGTACATATTTAACAAATA
WV-088 CSGW SpeI F	TATTTGTTAAATATGTACTACAAACTACTAGTTGGAAGGGCTAATCACTCCCAAAGA
WV-089 CSGW NruI RC	AGAGAGTCGCGATCTAATCTCCCCCGCTTAATAC
WV-090 CSGW psi BclI F	GGCGACTGGTGTAGTACGCCTGATCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAG
WV-091 CSGW psi BclI RC	CCGCTAGTCAAAAATTTTGTATCAGGCGTACTACCAGTCGCGCCCTCGCCTCTTGCC
WV-092-MLV repeated seq F	GAAGAGAAGTTCAGATCAAGGTGAG
WV-093-MLV repeated seq RC	GGACCCGATCTGGGGACCATCTGTTC
WV-094 KpnI LTR/MLTR F	CTCTCTGGTACCTGGAAGGGCTAATTTGGTCCC
WV-095 NheI LTR/CLTR RC	AGAGAGGCTAGCTGCTAGAGATTTTCCACTGAC
WV-096 NheI MLTR RC	AGAGAGGCTAGCGTCCCTGTTCCGGCGCCACTG
WV-097 KpnI CLTR F	CTCTCTGGTACCTTGACATTGATTATTGAC
WV-098 5' SpeI MLTR F	CTCTCTACTAGTTGGAAGGGCTAATTTGGTCC
WV-099 5' NruI MLTR RC	AGAGAGTCGCGAGTCCCTGTTCCGGCGCCAC
WV-100 5' SpeI CLTR F	CTCTCTACTAGTTGACATTGATTATTGACTAG
WV-101 5' NruI CLTR RC	AGAGAGTCGCGATGCTAGAGATTTTCCACTGAC
WV-102 5 LTR seq	CGTCCAGGTCTGAAGATCAGCGGCCG
WV-103-CSGWseq	TGGAGGTGCATCCATGACACAGGC
WV-104 CSGWdelNheI ApaI F	GGTGATTGGCTACCGTGGCGCAACTG
WV-105 5 CLTR psi NruI RC	AGAGAGTCGCGATACCGACGCTCTCGACCCATCTC
WV-106 5 CLTR SpeI F	CTCTCTACTAGTGTGGCATTGATTATTGACTAG
WV-107 5MLTR SpeI F	CTCTCTACTAGTTGGAAGGGCTAATTTGGTCCC
WV-108	GAGGAGACCAAAGGAAAGCGAAAGTTGTCCCTGTTCCGGCGCCACTGCTAGAG
WV-109	TTGAAAGCGAAAGGAAACCAGAGGAGCTCTCTCGACGAGGACTCGGCTTGC
WV-110 3LTR seq RC	GCCATCTCCTTGCATGACCATCTCTTG
WV-111 3MLTR seq F	GACGCTCTTCTGCTACGTCCTCTCGG
WV-112 Rev exon 1 SfiI F	CTCTGGCCACTAGGGCCATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAGAACAC
WV-113 Rev exon 2 SfiI RC	TCAGACTCATCAAGCTTCTCTATCAAAGCAACCCACCTCCCAATCCCGAGGGGACCCG
WV-114 Rev exon 1 SfiI F	AGAGGGCCCGGGCCCTATCTTTAGTTCTGACTCCAATAC
WV-115 Rev exon 2 SfiI RC	CTCTGGCCGAGAGGGCCATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAGAACAC
WV-116 4CTE F PstI	AGAGGGCCTCTCTGGCCCTATTCTTTAGTTCTGACTCCAATAC
WV-117 4CTE RC SbfI	CTGTACAAGTAACCTGACTTAATTAAGTGCAGACCTCCCTG
WV-118 RRE flexi F PstI	AATTAACCTCGACCCCTGGCAGGGCTACCTGCAGGACACATCCCT
WV-119 RRE flexi RC SbfI	CTCTCTGTCAGCAGGAGCTTTGTTCTTGGGTTT
WV-120 CTE F	AGAGAGCTGCAGGAGGAGCTGTTGATCCTTTAG
WV-121 CTE MluI RC	CTCTCTGTCAGGTAACCTGACTTAATTAAGGATCCACC
WV-122 MLV-R-U5 SpeI-KpnI F	CTCTCTGTCAGACGCGCTTAATTAACCTGACCCCTGGCAGGGCTAGCAGATC
WV-123 MLV-R-U5 2	CTCTACTAGTGGTACCGAATGAAAGACCCACCTGTAGGTTTG
WV-124 MLV-R-U5 3	CTCAGATCTGGTCTAACCAGAGAGACCAATGAAAGACCCCGCTGACGGGTAGTC
WV-125 MLV-R-U5 NruI-NheI RC	GGGTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGG
WV-126 MLV-R-U5 XhoI F	AGAGAGTCGCGAGCTAGCTGCTAGAGATTTTCCACTGACTAAAAG
WV-127 MLV-R-U5 NotI RC	CTCTCTCGAGGAATGAAAGACCCACCTGTAGGTTTG
WV-128 SFFV EcoRI F	AGAGGGCGCCGCTGCTAGAGATTTTCCACTGACTAAAAG
WV-129 SFFV NdeI-MfeI RC	CTCTCTGAATTCGCGATAAAAATAAAAGATTTTATTAG
WV-130 lack U3 KpnI RC	AGAGAGCAATTGCATATGCCCGGGCGACTCAGTCTGTCCG
WV-131 lack U3 KpnI F	AGAGAGGGTACCAGTACAAGCAAAAAGCAGATC
WV-132 U3 lack NheI F	CTCTCGGTACCTTTAAGACCAATGACTTACAAGGC
WV-133 U3 lack NheI RC	CTCTCTGCTAGCGCATCTAGAATTAATCCGTG
WV-134 NHG-3MLTR F	AGAGAGGCTAGCGATGACCCCTGCTGATTGGTTCCG
WV-135 NHG-3MLTR RC	CATGGTCTGCTGAGTTCGTGACCCG
WV-136 NHG-5MLTR RC	GTCCCTGTTCCGGCGCCACTGTAGAG
WV-137 NHG-5MLTR F	CATCTCTCCTTCTAGCCTCCGCTAGTC
WV-138 tat exon 1 F	TGGAAGGGCTAATTTGGTCCCAAAAAGACAAG
WV-139 tat exon 1 RC	CTTACGGGGTACTTGGGCAGGAGTG
WV-140 tat exon 2 F	CTTTCTGCACCTTACCTCTTATGCTTG
WV-141 tat exon 2 RC	GATGTTTCAATAAATATTACAGGGCTGC
WV-142 CCR5 F SfiI	GAGTAAGTCTCTCAAGCGGTGGTAGC
WV-143 CCR5 RC SfiI	CTCTCTGGCCATTAAGGCCATGGATTATCAAGTGCAAGTCCAATC
WV-144 gag SpeI F	AGAGAGGGCCGAGGCGCCTACAAGCCACAGATATTTCTGTG
WV-145 NdeI pol RC	CTCTCTACTAGTATGGGCGCCGCGCCAGCGTGTG
WV-146 NdeI F	AGAGAGCATATGCTAATCCTCATCTGTCTGCTGCC
WV-147 NdeI RC	GAGAGATGGGTGCGAGAGCGTCAATATATGGGGGAGAATTAGATCGCGATG
WV-148 gag NdeI F	CATCGCATCTAATCTCCCCCATATGTAAGTCTGACGCTCTCGACCCATCTCTC
WV-149 XhoI F	CTCTCTCATATGATGGGCGCCCGCCAGCGTGC
WV-150 NheI RC	CTCTCTCGAGGGAATTAATTCGAGCTCGG
	AGAGGGTACCCAGATCCTCTACGCGGACCTCT

WV-151 4CTE BamHI F	GCAGGTAACCTCGACTTAATTAAGGATCCACCTCCC
WV-152 4CTE XhoI RC	AGAGAGCTCGAGTCCGACCCTGGCAGGGCTAGCAGATCTACACATCCCTCGGAG
WV-153 XhoI F	CTCTCTCGAGACTCGACTTAATTAAGGATCCACCTCCCCTGTGAG
WV-154	TAATGAGTCTGACTGCCTTAGGAGGTCTATCGTCGCTGTCTCCGCTATCTTCTGCCAT
WV-154 RC REV	TAATGAGTCTGACTGCCTTAGGAGGTCTATCGTCGCTGTCTCCGCTATCTTCTGCCAT
WV-155 XhoI CTE F	CCCCGAAAAGTGCCACCTCGAGTCCGACTTAATTAAGGATC
WV-156 XhoI CTE RC	GCTATGACCATGATTACGCCAAGCTCTAGCTCGAGGTCGACCCTGGCAGGGCTAGCAG
WV-157 REV F	GTAAGTAATACATGTAATGCAACCTATACAAATAGCAATAGTAGCATTAGTAGTAGC
WV-158 REV RC	GGTTGCATTACATGTATTACTTACAGGAGATGCCTAAGGCTTTTGTATGAAACAAAC
WV-159 PacI-MluI F	GGCCAACCTCATGACCGGTAGATCTTAGCCACTTTTTTAAATTAAGGGGGGACTCGAG
WV-160 PacI-MluI RC	CTCGAGTCCCCCTTAAATTAAGGATGGCTAAGATCTACGCGTATGAGTTGGCC
WV-161 CD4 SfiI F	CTCTGGATCCGGCCAGAGAGGCCATGAACCGGGGAGTCCC
WV-162 CD4 SfiI RC	AGAGGGCCGCGGCCCTCTCTGGCTCAAAATGGGGCTACATGTC
WV-163 3LTR seq	GTAGCATCTGGCTTACTGAAGCAGAC
WV-164 EGFP seq F	CCGACAAGCAGAAGAACGGCATCAAG
WV-165	CTCTCTCGAGACTAGTTGGAAGGGCTAATTCAC
WV-166	TGCAACTGCAAGAGGGTTTATTGGATACACGGGTACCCGGGCGACTCAGTCAATCGGAGGACTC
WV-167	GCGCAGTACAGGCAAAAAGCAGCTGCTTATATG
WV-168	GTATCCAATAAACCTCTTGCAGTTGCAAGTAGTGTGTGCCCGTCTGTTGTGTG
WV-169 EGFP qPCR F	GAGAGGCTAGCTCGGATCTAATCTCCCCCGC
WV-170 EGFP qPCR RC	GGAGCGCACCATCTTCTTCA
WV-171 LTR-R XhoI F	AGGGTGTGCGCCTCGAA
WV-172 LTR-R NotI RC	CTCTCTCGAGTGGAAAGGGCTAATTCACTCC
WV-173 SFFV-GFP F	AGAGGGCCGCGCTGCTAGAGATTTTCCACAC
WV-174 SFFV-GFP RC	CTCTAGATCTGAATTCCTGCAGCCCCG
WV-175 CXCR4 BamHI 1	AGAGAAGCTTGGCGGCGCTTACTTGTAC
WV-176 CXCR4 2	CTCTGGATCCATGGAGGGGATCAGTATATAC
WV-177 CXCR4 3	GCATAGAGGATGGGGTTCAGACAACAGTGG
WV-178 CXCR4 NotI 4	CCACTGTTGTCTGAACCCATCCTCTATGC
WV-179 SFFV seq F	AGAGGGCCGCGCTTAGCTGGAGTGAAGAACTTG
WV-180 CD4 seq	CCTGAAATGACCCTGCGCCTTATTG
WV-182 CXCR4 BglII F	CTCCTTCCCACTCGCCTTTACAG
WV-183 CXCR4 BglII RC	CTCTAGATCTTGGAGGGGATCAGTATATAC
WV-184 LNCX SphI F	AGAGAGATCTTTAGCTGGAGTGAAAACCTG
WV-185 LNCX 2 RC	CTCTGCATGCCGACGCGGAGGATCTCG
WV-186 LNCX 3 F	CGTAATTGATTACTATTACGAGATCTAGAGTCCCCTCAGAAGAAC
WV-187 LNCX 4 BglII RC	GTCTTCTGAGCGGACTCTAGATCTCGTAATAGTAATCAATTACG
WV-188 LNCX XhoI RC	AGAGAGATCTGAGTCCGGTAGCGTAGCGGATC
WV-189 NHG 3LTR seq	AGAGCTCGAGATCTGAGTCCGGTAGCGC
WV-190 5' LTR seq (NHG) F	CTCTGACGGCAGTTTACGAGAGAGATG
WV-191 CXCR4 WV SfiI F	TATTACCAC TGCCAATTACCTGTGG
WV-192 CXCR4 WV SfiI RC	CTCTGGCCGAGGGCCATGGAGGGGATCAGTATATACAC
WV-193 EGFP SfiI F	AGAGGGCCATTAAGGCCTTAGCTGGAGTGAAAAC
WV-194 EGFP SfiI RC	CTCTGGCCGAGGGCCATGGTGAGCAAGGGCGAGGAGC
WV-195 PCR 4 RC	AGAGGGCCATTAAGGCCTTAGCTGTACAGCTCGTCCATGC
WV-196 PCR 3 F	AGAGGAATTCTTATTATGGCTTCCACTCC
WV-197 PCR 2 RC	GGCGAGCAGACAGGATGAGGATTAGCACATGAAAAGATTAGTAAAACAC
WV-198 PCR 1 F	GGTGTTTTACTAATCTTTTCCATGTGCTAATCTCATCTGCTGCTCGCC
WV-199 EcoRI RC	CTCTGAATTCGTTGGCATTGATTATTGACTAG
WV-200 EcoRI F	GAATTCACTACTAAGTTTGTAGTACATATTTAAC
WV-201 RC	GTAAAATATGTACTACAACTTAGTAGTGAATTC
WV-202 NdeI F	CTGTTTTAACCTGCAGGATGTGGTATTC
WV-203 NotI RC	CTCTCATATGATGGGTGCGAGAGCGTCCGTTAATG
WV-204 CXCR4 seq	AGAGGGCCGCGCTTAATCCTCATCTGCTACTTGCC
WV-205 psi gag NdeI F	CTGCCTGGTATTGTCATCCTGTCTC
WV-206 CpG BssHII 1 F	CTCTCATATGAAAAATTTGACTAGCGGAGGCTAG
WV-207 CpG RC 2	CTCTGCGCGCACGGCAAGAGGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTGACTAG
WV-208 CpG RC 3	CGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCCGTTAATG
WV-209 CpG EcoRI 4 RC	GGTGTTTTACTAATCTTTTCCATGTGTTAATCTCATCTGCTACTTGCC
WV-210 seqGP	GGCAAGTAGACAGGATGAGGATTAACACATGAAAAGATTAGTAAAACAC
WV-211 seqGP	AGAGGAATTCTTATTATGGCTTCCACTCCTG
WV-212 seqGP	CAGTTTATGACAGGCCAATAGCAC
WV-213 seqGP	CTCAGATCACTCTTTGGCAGCGACC
WV-214 seqGP	GACTTGTATGTGGGAGTACCTG
	GACAGACAGGGGAAGACAGAAGGTAG
	CAGCAGTCCAGATGGCTGTCTTCAATC
WL-191 NHG BamHI F	TATTCATAATGATAGTAGGAGGCTTGG

### 2.2.3 Cloning

2x Ligation buffer (Promega, C671) and T4 Ligase (Promega, C1263) were used for restriction digest cloning reactions. Restriction enzymes were purchased from Promega and New England Bio Labs, shown in Table 2.2.

**Table 2.2.** Restriction enzymes

	PROMEGA (UK)	New England Biolabs (UK)
	Catalog Number	Catalog Number
Alkaline Phosphatase, Calf Intestinal (CIAP)	M1821	
Apal	R6361	
BamHI	R6021	
BglIII	R6081	
EcoRI	R6011	
EcoRV	R6351	
HindIII		R0104
HpaI	R6301	
KpnI	R6341	
MluI	R6381	
NdeI	R6801	
NheI	R6501	
NotI	R6431	
PacI		R0547
PstI	R6111	
SalI	R6051	
SbfI	R0642	
SfiI		R0123
SnaBI		R0130
SpeI	R6591	
XbaI	R6181	
XhoI	R6161	

### 2.2.4 Plasmids

All plasmids used in this thesis are shown in Table 2.3

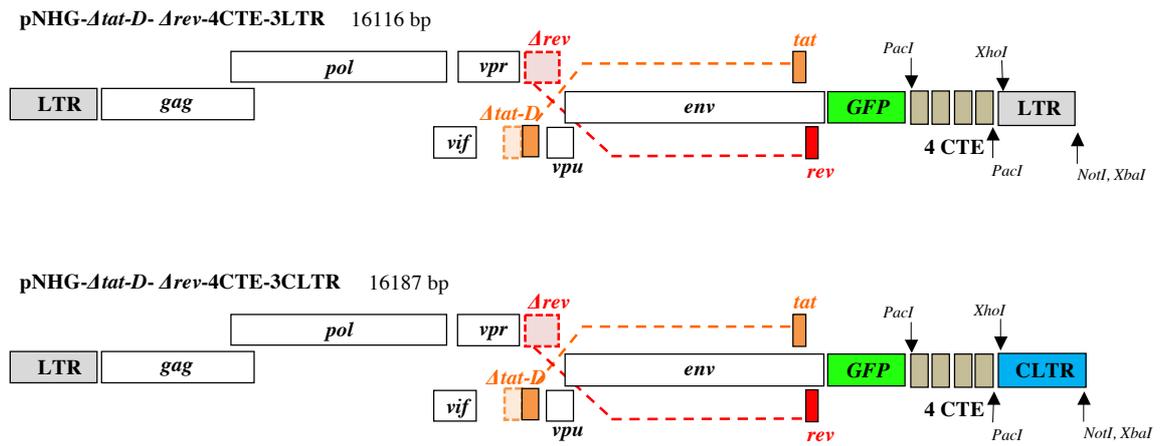
**Table 2.3.** Plasmids

Plasmids	Description	Source
pCNCG	MLV-derived vector expressing EGFP from the CMV promoter	Generous gift from G. Towers
pCSGW	HIV-1-derived vector that provides the cis-acting information required for RNA packaging into virions and for subsequent reverse transcription and for subsequent reverse transcription and integration. It expresses GFP from an internal CMV promoter and the U3 region has been removed from the 3' LTR	A. Thrasher, Institute of Child Health, University College London, UK
pCSGW_4CTE	pCSGW encoding 4 copies of the Constitutive Transport Element preceding the 3' ΔU3 region from the 3' LTR	Made during PhD
pCSGW_Flexi_GP_4CTE_3LTR	pCSGW encoding the codon-optimized HIV-1 gag-pol genes, an IRES_GFP and 4 copies of the Constitutive Transport Element preceding the 3' LTR	Made during PhD
pCSGW_Flexi_GP_4CTE_3CLTR	pCSGW encoding codon-optimized HIV-1 gag-pol genes, an IRES-GFP and 4 copies of the Constitutive Transport Element preceding the 3' CLTR	Made during PhD
pCSGW_Flexi_GP_4CTE_3MLTR	pCSGW encoding codon-optimized HIV-1 gag-pol genes, an IRES-GFP and 4 copies of the Constitutive Transport Element preceding the 3' MLTR	Made during PhD
pCSGW_Flexi_ΔU3	pCSGW derivative including a unique <i>SpeI</i> site in the 5528 position; the <i>NheI</i> site located at position 7842 has been deleted. These modifications facilitate the insertion of different LTRs at the 5' and 3' ends	Made during PhD
p5CLTR_Flexi_ΔU3	pCSGW-Flexi encoding a 5'CLTR	Made during PhD
p5MLTR_Flexi_ΔU3	pCSGW-Flexi encoding a 5'MLTR	Made during PhD

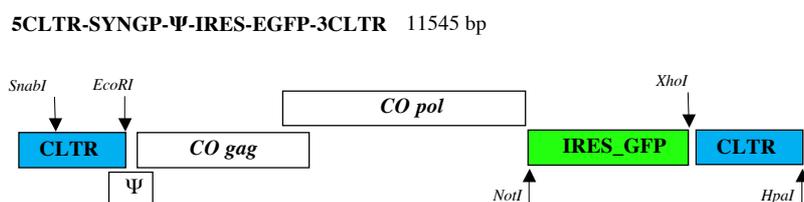
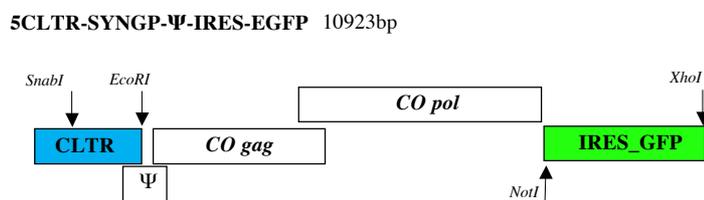
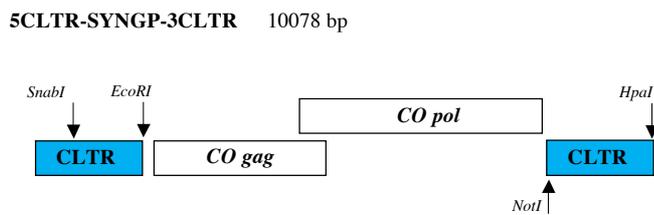
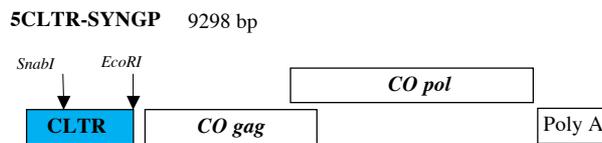
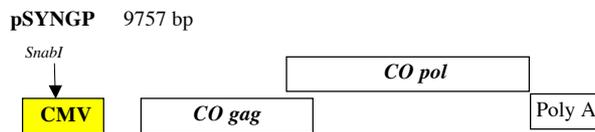
pCRV1_NL_gag-pol	Expresses the HIV-1 structural proteins under the control of the CMV immediate-early promoter. The gag region encodes genes which comprise the capsid proteins; the pol region encodes the reverse transcriptase and integrase proteins	Cell biolabs
pDEST40	Expression vector designed to allow high-level, constitutive expression of the gene of interest in a variety of mammalian hosts	Invitrogen
pDEST40_tat	pDEST40 encoding tat gene	Made during PhD
pDEST40_rev	pDEST40 encoding rev gene	Made during PhD
pDEST40_4CTE	pDEST40 encoding 4 copies of the CTE from MPMV	Made during PhD
pHBXB2_env	Construct expressing HXB2 gp160	NIH AIDS Reagent Program Catalog number 12363
pIRES2_EGFP	Contains the internal ribosome entry site (IRES) of the encephalomyocarditis virus (ECMV) between the MCS and the enhanced green fluorescent protein (EGFP) coding region. This permits both the gene of interest (cloned into the MCS) and the EGFP gene to be translated from a single bicistronic mRNA	ClonTech
pLNCX2	Mouse Moloney murine sarcoma virus vector used for transduction of several cell lines. Contains a Neomycin resistance gene	AddGene
pLNCX_CXCR4	pLNCX2 derivative encoding a HIV-1 CXCR4 co-receptor	AddGene
pLPSG	Mouse Moloney murine sarcoma virus vector used for transduction of several cell lines. Contains a Neomycin resistance gene	AddGene
pLPSG-CD4	pLPSG derivative encoding a HIV-1 CD4 receptor	AddGene
pNHG	Replication-competent HIV-1 clone in which Nef was replaced with EGFP	NIH AIDS Reagent Program Accession number JQ585717
pNHG_3LTR	pNHG derivative encoding a polylinker preceding the 3'LTR	Made during PhD
pNHG_3CLTR	pNHG derivative encoding a 3' CLTR instead the 3'LTR	Made during PhD
pNHG_3MLTR	pNHG derivative encoding a 3' MLTR instead the 3'LTR	Made during PhD
pNHG_Atata-F_3LTR	pNHG encoding a frameshift mutation in the first exon of <i>tat</i>	Made during PhD
pNHG_Atata-F_3CLTR	pNHG encoding a frameshift mutation in the first exon of <i>tat</i> and a 3'CLTR	Made during PhD
pNHG_Atata-F_3MLTR	pNHG encoding a frameshift mutation in the first exon of <i>tat</i> and a 3'MLTR	Made during PhD
pNHG_Atata-D_3LTR	pNHG encoding a 43nt deletion in the first exon of <i>tat</i> in the first exon of <i>tat</i>	Made during PhD
pNHG_Atata-D_3CLTR	pNHG encoding a 43nt deletion in the first exon of <i>tat</i> and a 3'CLTR	Made during PhD
pNHG_Atata-D_3MLTR	pNHG encoding a 43nt deletion in the first exon of <i>tat</i> and a 3'MLTR	Made during PhD
pNHG_Atata-D_Δrev_4CTE_3LTR	pNHG encoding a 43nt deletion in the first exon of <i>tat</i> . The <i>rev</i> exon 1 has been removed and 4CTEs are preceding the 3' LTR	Made during PhD
pNHG_Atata-D_Δrev_4CTE_3CLTR	pNHG encoding a 43nt deletion in the first exon of <i>tat</i> . The <i>rev</i> exon 1 has been removed and 4CTEs are preceding the 3' CLTR	Made during PhD
pNHG_Atata-D_Δrev_4CTE_3MLTR	pNHG encoding a 43nt deletion in the first exon of <i>tat</i> . The <i>rev</i> exon 1 has been removed and 4CTEs are preceding the 3' MLTR	Made during PhD
pNL4-3	Full-length replication-competent HIV-1 clone	NIH AIDS Reagent Program Catalog number 12363

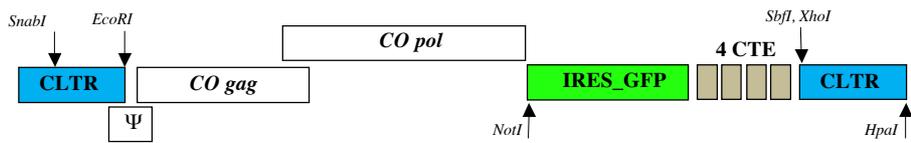
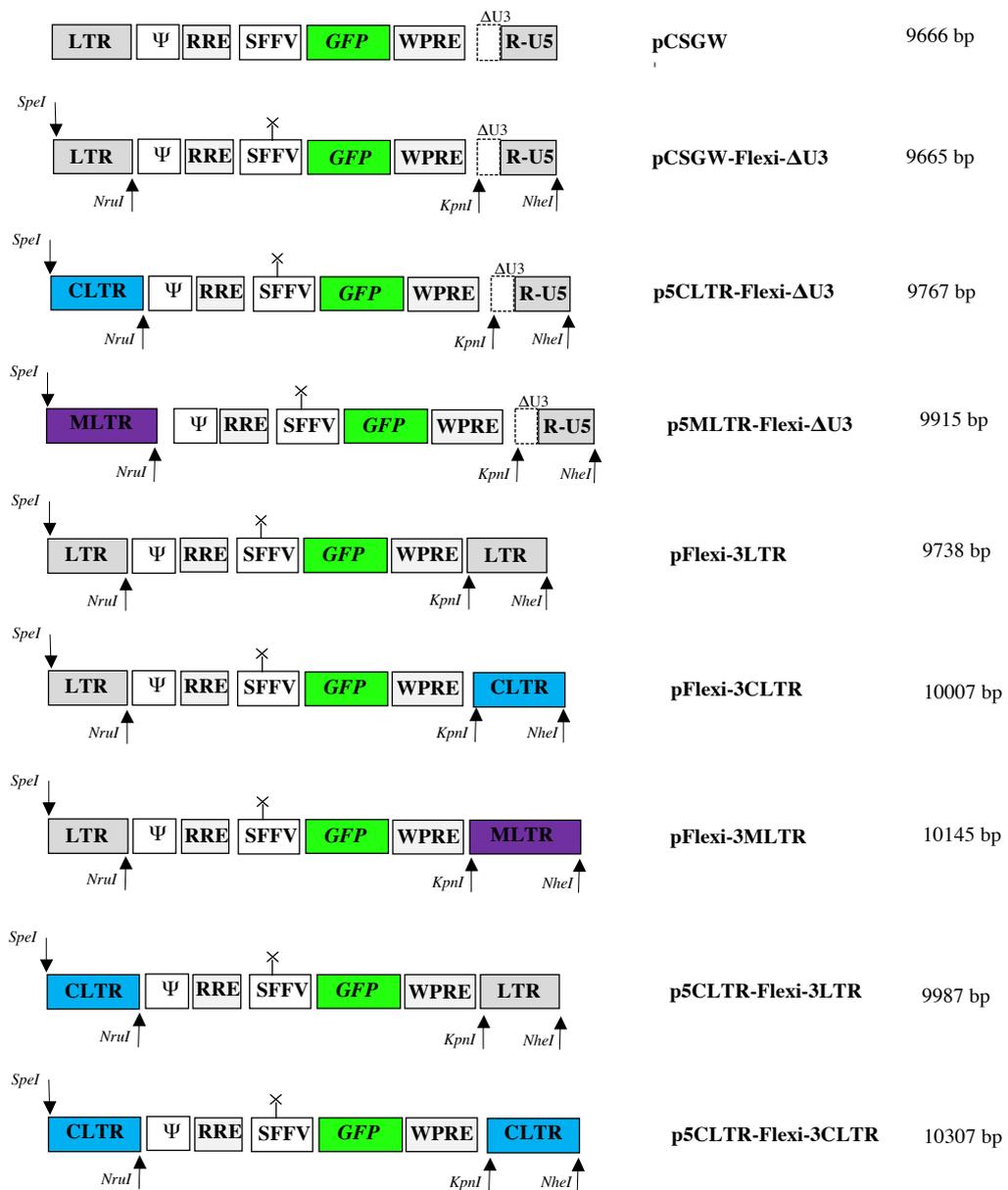


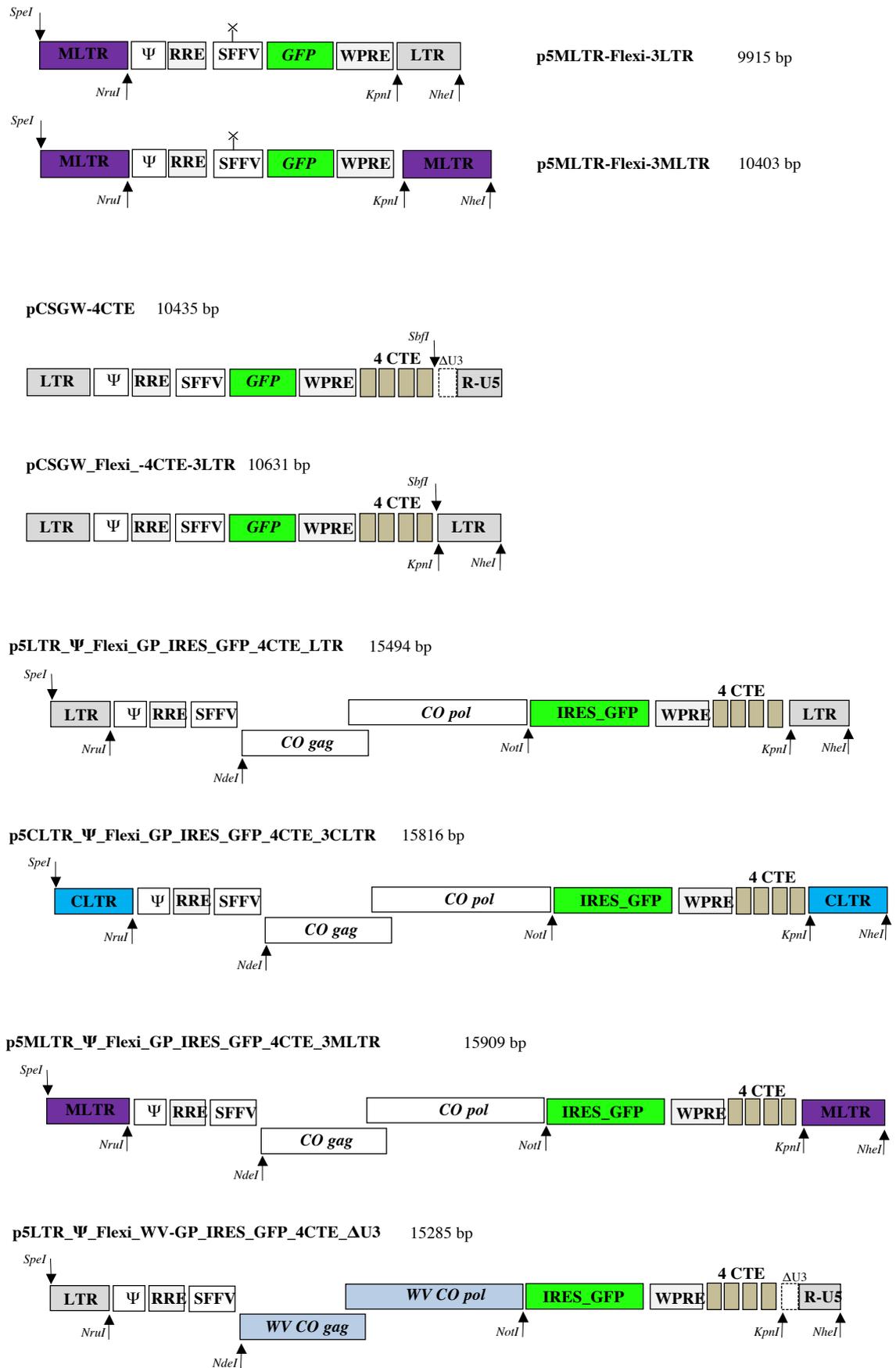




### SYNGP clones



**5CLTR-SYNGP-Ψ-IRES-EGFP-3CLTR** 11545 bpCSGW clones



All constructs developed during this thesis were successfully sequenced to confirm that their sequences were correct.

CLTR	CMV-R-U5 chimeric promoter (R and U5 regions from HIV LTR)
<i>CO gag</i>	codon-optimized <i>gag</i> gene for humans
<i>CO pol</i>	codon-optimized <i>pol</i> gene for humans
CTE	constitutive transport element from Mason-Pfizer monkey virus (MPMV)
<i>env</i>	envelope
<i>gag</i>	group-specific antigen (MA, CA, p2, NC, p1 and p6)
<i>GFP</i>	green fluorescent protein
<i>IRES_GFP</i>	cassette of an internal ribosomal entry site preceding a GFP gene
LTR	long terminal repeat
MLTR	MLTR-R-U5 chimeric promoter (R and U5 regions from HIV LTR)
<i>pol</i>	polymerase (PR, RT, RNase H and IN)
poly A	polyadenylation signal of SV40
R	R region from HIV LTR
<i>rev</i>	regulator of the expression of virion proteins
RRE	Rev response element from HIV
SFFV	spleen focus forming virus promoter
<i>tat</i>	<i>trans</i> -activator of transcription
U5	U5 region from HIV LTR
<i>vif</i>	virion infectivity factor
<i>vpr</i>	viral protein R
<i>vpu</i>	viral protein U
WPRES	Woodchuck hepatitis virus post-transcriptional regulatory element
<i>WV CO gag</i>	codon-optimized <i>gag</i> gene for humans with reduced CpG number
<i>WV CO pol</i>	codon-optimized <i>pol</i> gene for humans with reduced CpG number
$\Delta rev$	complete deletion of <i>rev</i> exon 1
$\Delta tat-D$	deletion of 43 nucleotides in <i>tat</i> exon 1
$\Delta tat-F$	frameshift mutation in <i>tat</i> exon 1
$\Delta U3$	deletion in HIV U3 region
$\Psi$	packaging signal of HIV-1

### 2.2.5 Antibodies

The primary antibody used in this thesis was mouse anti-p24 (183-H12-5C, 2mg/ml stock concentration; 1:50 final concentration) produced from an hybridoma of Balb/c mouse splenocyte x SP2-0 myeloma from the National Institutes for Health AIDS Reagents Program (catalog number 1513). The fluorescently labelled secondary antibody was GAM goat anti-mouse [IRDye800 Thermo Scientific, 0.5 mg/ml stock concentration; 1:10000 final concentration]. Both antibodies were used for Western blot assays.

## 2.3 Chemicals and buffers

### 2.3.1 Cell culture

1X DMEM – Dulbecco’s modified Eagle medium 1X High glucose (4500mg/L) with sodium pyruvate (110 mg/L), [Invitrogen 41966-029], 10% fetal calf serum (FCS), Gentamicin (20µg /ml final concentration).

1X RPMI - Roswell Park Memorial Institute medium 1640 with L-Glutamine [Invitrogen 21875-034], 10% fetal calf serum (FCS), Gentamicin (20µg /ml).

1000x Ampicillin (100 mg/ml) – Melford Laboratories

1000x AMP (100 mg/ml) - Melford Laboratories

1000x G418 (100 mg/ml) – Gentamicin, Melford Laboratories

Puromycin – Melford Laboratories

Trypsin 0.05% EDTA – purchased from Sigma-Aldrich

Formal saline fixative (FS) - Sigma-Aldrich, UK. 4% formaldehyde final concentration

LB Broth - provided from EO Labs, pH 7.0 REF BM5300, 350ml

LB Agar - provided from EO Labs, pH 7.0 REF BM5280, 300ml

SOC recovery medium - 20g Bacto Tryptone, 5g Bacto Yeast Extract, 2ml of 5M NaCl, 2.5ml of 1M KCl, 10ml of 1M MgCl<sub>2</sub>, 10ml of 1M MgSO<sub>4</sub>, 20ml of 1M glucose (1 litre final volume)

Phosphate Buffered Saline (PBS) - MgCl<sub>2</sub> and CaCl<sub>2</sub> free; Sigma-Aldrich

Terrific broth (TB) buffer for competent cells preparation- PIPES 10mM, KCl 250mM, CaCl<sub>2</sub> 15mM, 1mM MnCl<sub>2</sub>; pH 6.7, [1L final volume]

Polyethylenimine (PEI-25000) - provided by Polysciences, USA

### 2.2.2 Protein analysis

4x NuPage Sample Buffer - Tris 3.4g, SDS (8%), EDTA dehydrate 0.024g, glycerol

(40%), blue bromophenol; pH 8.5;  $\beta$ -mercaptoethanol (1.5%) added before use. NuPage Novex BT GEL 4-12% 1.0MM - (Invitrogen)

### 2.3.2 DNA analysis

TAE Buffer - Tris 245.375g, Acetic acid (17.5%), EDTA dehydrate (18.5g)

Agarose – 0.8-1.5% (w/v) diluted in 1x TAE (A9539 Sigma-Aldrich)

DNA loading buffer - 0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0); Promega.

P1 buffer\* - Tris 3.03g, EDTA dihydrate 1.86mg, pH 8.0. RNase was added at 100 $\mu$ g/ml

P2 buffer\* - SDS (20%), NaOH 4g

P3 buffer\* - Potassium acetate 147.25g, and Acetic acid until reach pH 5.5

Neutralisation solution – provided by Pure Yield Plasmid Midiprep System (Promega) and Plasmid Midi Kit (Qiagen,12143).

\*Qiagen Midi Kit buffers were used to purify our plasmids from *E.coli* DH10B strain

### 2.3.3 RNA extraction

Viral RNA was extracted and purified from virions using a QIAamp Viral RNA Kit (Qiagen-52904) following the manufacturer's instructions.

### 2.3.4 Western blot reagents

20x MOPS - MOPS 104.6g, Tris 60.6 g, SDS (2%), EDTA Free Acid (0.6%); pH 7.7 [500ml final volume]

10x Transfer buffer - Tris 120g, glycine 576.9 g. Prior to use add absolute ethanol (20%)

PBS Tween-20 – washing buffer 1% final concentration (Sigma-Aldrich)

Nitrocellulose Membranes – Invitrogen, 77012

## 2.4 Methods

### 2.4.1 Cell culture

Adherent mammalian cell lines were cultured in 10cm plates (Corning) and split every 3-4 days at 90-100% confluency. Firstly the monolayer was washed gently once with PBS, and then trypsin 0.05% EDTA (ethylenediaminetetraacetic acid) was added to detach the adherent cells through incubation at 37°C for 5-8 minutes. Cells were re-suspended in 5 ml of fresh medium and centrifuged at 1200 rpm for 5 minutes. The supernatant was removed and cells were resuspended in 10 ml of fresh medium, split into new 10 cm plates and maintained at 37°C with 5% CO<sub>2</sub>.

Suspension cell lines were maintained in 75cm<sup>2</sup> and Split every 3-4 days, splitting into a new 75cm<sup>2</sup> flask with fresh media.

### 2.4.2. Transfection of human and murine cells

Human and murine cell line transfections were carried out when the cells were at ~80% confluency. Transfections were made in 6 well plates/10 cm plates using PEI-25000 (polyethylene glycol) as a transfection agent (4µl PEI / µg of DNA). Transfection volumes are shown in Table 2.4.

**Table 2.4.** Transfection volumes

Cell culture vessel	Cell culture media volume (ml) [each well]	Surface area for each well cm <sup>2</sup>	Transfection mix volume (µl)
6 well plate	2	10	200
12 well plate	2	4	200
24 well plate	1	2	200
10 cm <sup>2</sup> plate	12	60	500

### 2.4.3 Assays for viral particle production and infectivity production and infectivity

Single cycle of replication viruses were generated by co-transfection of multiple plasmids in HEK-293Ts and NIH-3T3 cells. Cells were seeded at 1x10<sup>6</sup> per well in 6-well plate 24 hours prior to transfection. The transfection solution was mix by pulse-vortexing 20 times during 20 seconds and added to the culture 25-30 minutes later. Cell culture media was changed 24 hours post-transfection, and the supernatant containing infectious viral particles was harvested at 48 or 72 hours post-transfection using 0.45 µm 25mm diameter syringe filters (Sartorius *Minisart*®) and inoculated onto target cells. The infectivity assays were developed in human and murine cell lines by serial dilutions (1:3) made out in 96

well plates to determine the titre of the viruses measuring GFP expression by flow cytometry (Guava® EasyCyte, Merck Millipore) (Figure 2-2).



**Figure 2-2.** Experiment protocol to generate replication competent HIV-1 particles or single cycle of replication viruses in HEK-293T cells. GFP expression can be detected 24 hours post-transfection (h.p.t.). Viral particles are harvested 48 h.p.t. and inoculated to a target cell culture (AA-2 or PM1). Human lymphocytes are spinoculated for 1 hour and then incubated at 37°C. FACS analysis is carried out 48 hours post-infection (h.p.i.).

## 2.4.4 Cloning

### 2.4.4.1 NHG clones

The backbone used to develop these clones was pNHG, an HIV-1 replication competent genome (Accession number *JQ585717*) that carries a *GFP* gene instead of *nef*. Firstly, a polylinker containing unique sites for *XhoI*, *NotI* and *XbaI* sites was cloned preceding the 3'LTR. The sites *XhoI* and *NotI* were used to insert all LTRs: LTRwt (U3-R-U5) was amplified using the oligos WV-001, WV-002. The CMV-R-U5 (CLTR) was generated using pCRV1-NLGP as template and the oligos WV-002 and WV-004. The chimeric HIV-MLV LTR (MLTR) was cloned by PCR using pNHG and pCNCG as templates, with the oligos WV-050, WV-094, WV-098, WV-051, WV-052, WV-053, WV-054, WV-095, WV-127. Latter oligos present different restriction sites that were used to clone the MLTR into other plasmids.

The mutations in *tat* exon 1 ( $\Delta$ *tat*) were made following two different strategies: a) Removing the *tat* start codon and introducing stop codons along the sequence to generate a frameshift (*tat-F*); b) Deletion of 43 nucleotides to impede a reversion of the *tat* expression (*tat-D*). Both mutations were cloned by PCR by cleavage with *NdeI* with the oligos: WV-006, WV-007, WV-008, WV-009, WV-012 and WV-013, using pNHG as template. The mutation in *rev* ( $\Delta$ *rev*) was developed by PCR to completely remove exon 1. This strategy also removed 76 nucleotides of the *tat* exon 1 due to overlapping sequences. The oligos for

*Δrev* were WV-006, WV-141 and WV-157 and WV-158 prior to insertion by cleavage with *EcoRI* and *BamHI*. Prior to insertion of the CTE into NHG we cloned by PCR a polylinker containing *MluI* and *PacI* sites with the oligos WL-191, WV-159 and WV160 digesting with *BamHI* and *XhoI*. Then four copies of CTE were inserted using *PacI*.

#### 2.4.4.2 SYNGP clones

The 5'CMV promoter already present in pSYNGP was replaced by the CLTR from pCRV1-NLGP using the oligos WV-018 and WV-019 and digested with *SnaBI* (Promega R6791) and *EcoRI* (Promega R6011). The 3'CLTR was cloned using the oligos WV-019 and WV-020 by digestion with *NotI* and *HpaI* (R6301).

The HIV-1 packaging signal ( $\Psi$ ) was cloned into SYNGP constructs by PCR using the following oligos: WV-025, WV-026, WV-027 and WV-028, using pNL4-3 as a template by cleavage with *NdeI* and *MluI* (R6381). Afterwards, the IRES-GFP cassette was introduced by PCR from pIRES2-GFP with the oligos WV-041 and WV-042 digesting with *XhoI* and *NotI*.

Four copies of the Constitutive Transport Element (CTE) were amplified by PCR using Psiv-MPCG (MPMV proviral clone) as template and the oligos WV-058, WV-059. One copy was cloned into the pDEST40 expression vector by cleavage with *Sall* (Promega R6051). The resultant vector, DEST40-CTE was digested with *BglII* (Promega R6081) and *NheI* (Promega R6501) to insert one copy of the CTE digested with *BamHI* and *NheI*. This step was repeated twice to generate 4 copies of the CTE *in cis* in DEST40. The 4CTEs were cloned into SYNGP digesting with *SbfI*.

Several restriction sites were introduced in SYNGP following the last copy of the CTE (preceding the 3'LTR) with a polylinker synthesized by annealing the oligos WV-056 and WV-057. The annealing reaction was composed of 10 $\mu$ l of each primer, 2.5 $\mu$ l 2M NaCl and 50 $\mu$ l H<sub>2</sub>O to 50 $\mu$ l final volume. Then, the reaction was boiled for 5 minutes and cool down slowly at room temperature and cloned into our constructs digesting with *XhoI*.

#### 2.4.4.3 CSGW-flexi clones

The system developed to determine the functionality of the LTRs in this project is called Flexi, which used pCSGW as a template. Several modifications were made in pCSGW to

allow us to test different LTRs in an easy way. Accordingly, it was necessary to remove the *NheI* site located in the SFFV promoter by PCR using the oligos: WV-082, WV-083, WV-084 and WV-085. A unique *SpeI* site was introduced at the beginning of the 5'LTR by PCR using the oligos WV-086, WV-087, WV-088 and WV-089 by cleavage with *BamHI* and *EcoRI*. An additional *BclI* site was also introduced immediately before the HIV-1 packaging signal. At this point, the several 5'LTRs were cloned into CSGW-flexi digesting with *HpaI* (Promega) and *NruI* (Promega R7091); about the 3'LTRs cloning, the digestion was made with *KpnI* (Promega R6341) and *NheI*. In order to introduce the different LTRs at the 5' end, the CLTR was amplified by PCR with oligos WV-094 and WV-095; about the MLTR, the oligos employed were WV-094 and WV-096. At the 3'end, the CLTR was amplified by PCR using the oligos WV-097 and WV-101; the oligos used for the MLTR were WV-098 and WV-099.

#### 2.4.4.4 CSGW\_GP clones

The SFFV-*gag-pol* cassette was subcloned from 5CLTR\_Ψ\_SYNGP\_IRES\_GFP\_4CTE\_3CLTR into CSGW\_4CTE by cleavage with *EcoRI* and *NotI*. The CLTR and MLTR promoters were cloned in the same manner as CSGW\_Flexi clones, using the unique restriction sites located at both ends.

#### 2.4.4.5 DEST40 clones

HIV-1 *tat* and *rev* genes were cloned into this expression vector. The *tat* gene was cloned by PCR with the oligos WV-010 and WV-011 by cleavage with *SfiI* (DEST40-*tat*). The *rev* gene cloning followed the same strategy as before, but the oligos were WV-114 and WV-115 (DEST40-*rev*). Both strategies removed the introns of both genes. The Constitutive Transport Element (CTE) was cloned by PCR using pSARM4 as a template with the oligos WV-155 and WV-156, which contain several restriction sites. DEST40 vector was digested with *BamHI* and *NheI*, while the CTE was digested with *BglII* and *NheI*. This strategy was repeated twice to achieve four copies in total (DEST40-4CTE).

#### 2.4.4.6 MLV based retroviral vectors for ex vivo gene transfer

HIV-1 receptor CD4 was cloned into pLPSG by PCR using the oligos WV-161 and WV-162 by cleavage with *SfiI*. To clone the SFFV-CXCR4 cassette into pLNCX2, we

introduced a *BglIII* site in by PCR using the following oligos: WV-184, WV-185 WV-186, WV-188, WV-175 and WV-178. Then, the SFFV-CXCR4 cassette was digested with *BglIII*, generating the pLNSX-CXCR4. Viral particles were generated following the protocol described in Figure 2-2. The *gag-pol* genes are provided by pMLV-GP (1µg) and *envelope* by pCMV-VSV-G (100ng) (1:100 ratio). After 48 hours post infection the p/m40 AK were subjected to selection with the antibiotics described in section 2.1.2.

#### 2.4.5 Ligation reactions

After vector digestion the plasmid was dephosphorylated using CIAP (Calf Intestinal Alkaline Phosphatase, Promega M1821). Restriction digestion volumes were 40 µl total, adding 5 µl of 10X calf intestinal alkaline phosphatase buffer (CIAP) and 4.5 µl of CIAP enzyme. The reactions were incubated for 2 hours at 37°C following an incubation at 55°C during 10 minutes. Samples were purified from an agarose gel after electrophoresis and transferred into gel extraction columns (Purelink Quick Gel Extraction Kit, Invitrogen K210012) following manufacturers instructions. Ligation reactions were carried out adding 5 µl of 2x Rapid DNA Ligation buffer (Promega, C671) and 0.5 µl T4 DNA Ligase (Promega, C1263) per reaction. Vector/insert ratio used was 1:3.5, and the reactions were incubated for 1-3 hours before the transformation using DH10B.

#### 2.4.6 Bacteria transformation

Bacteria were incubated with the plasmid of interest on ice for 30 minutes and subjected to a heat shock for 45 seconds at 42 °C. Then, bacteria were cultured in liquid phase SOC media during 45 minutes prior to transferal to a solid phase LB agar plate (100µg/ml final concentration of ampicillin/kanamycin) that was incubated overnight.

##### 2.4.6.1 Chemically competent *Escherichia coli* DH10B bacteria preparation

*E. coli* DH10B competent bacteria were prepared (by Walter Vera) using the following protocol:

1. Streak the host bacteria on an agar plate without selection.
2. Incubate at 37°C overnight.
3. Pick a single colony to inoculate several starter cultures (20ml) without selection.

4. Shake at 37°C for 8 hours.
5. Add starter culture to 1.4L of LB in a conical flask and shake at 18°C until O.D. reaches 0.6 (~16-24h).
6. Put culture on ice for 10 min. Then split into 6 centrifuge bottles (200x6 and leftover) and pellet the bugs at 2000g for 10 min at 4°C.
7. Discard supernatant.
8. Resuspend all pellets in 6x75 ml of ice-cold TB (The six bottles must be combined into two at this stage). Incubate for 5 minutes.
9. Spin at 2000g for 10 min at 4°C.
10. Discard supernatant.
11. Resuspend the pellet in 40ML 7% TB-DMSO and transfer all volume into one. Sit the culture on ice for another 10 min.
12. Snap-freeze the bugs in liquid nitrogen/ethanol ice bath.
13. Aliquot the bugs e.g. 250-300ul/tube bath and store at -80°C.

#### 2.4.7 Miniprep cultures, DNA extraction and restriction digestion

Miniprep cultures were performed to determine the presence of the insert within the plasmid. Bacterial colonies were picked and deposited in a replica agar plate with antibiotic resistance and resuspended in 1ml LB with Ampicillin or Kanamycin. Then cultures were incubated at 30°C at 240 rpm for 10-12 hours. Plasmid extraction from bacteria was performed using the following protocol:

1. Spin down minicultures for 5-10 minutes at 13.000 rpm
2. Aspirate LB growth medium
3. Add 200 µl of buffer P1 to each tube. Resuspend pellet.
4. Add 200 µl of buffer P2 to each tube. Heat SDS prior to use to ensure it is in solution. Mix by inversion following addition of P2.
5. Incubate for 10-15 minutes at room temperature.
6. Add 250 µl of buffer P3 and mix by inversion.
7. Centrifuge at 13000 rpm for 10-15 minutes and add 1 ml ethanol into new 1.5 ml tubes.
8. Carefully remove the supernatant from the centrifugation avoiding the pellet and transfer this supernatant to the ethanol. Shake the tube to mix and place at -20 °C

for 15-30 minutes.

9. Spin at 13000 rpm for 10 minutes, aspirate supernatant alcohol and wash with 200  $\mu$ l of 70% ethanol.
10. Centrifuge at 13.000 rpm for 2-5 minutes
11. Aspirate supernatant to remove all ethanol, air dry the tubes until totally dry and resuspend in 50  $\mu$ l of MiliQ water.

The extracted DNA was digested with restriction enzymes to determine the digestion pattern of the plasmid including the insert. Samples were prepared in 10  $\mu$ l final volume and incubated for 3-4 hours at 37 or 50°C (depending of the enzyme features). Digested samples were run into agarose gels as described in section 2.4.12.

#### 2.4.8 Midiprep cultures

Midiprep cultures were carried out to generate plasmids stocks containing the insert of interest. Colonies were picked from the replica plate and inoculated into 250 ml of LB with Ampicillin/Kanamycin and cultured at 240 rpm and 30°C for 10-12 hours. DNA extractions from the bacteria were performed using the Plasmid Midi Kit (Qiagen,12143) following manufacturer instructions.

#### 2.4.9 RNA extraction and purification

Supernatant containing viral particles were DNase treated for 1 hour at 37°C to remove plasmid background from the transfection. Then, viral particles were purified using 20% sucrose cushions in 1.5ml tubes and centrifuged at 14000 rpm for 4 hours at 4°C. The RNA extraction was carried out using a QIAamp Viral RNA Kit (Qiagen-52904) following the manufacturer instructions.

#### 2.4.10 Reverse transcription and qPCR assays

The synthesis of cDNA was carried out using a SUPERScript III Kit (Invitrogen 18080044); the RNase Out Ribonuclease was provided from Invitrogen-10777.019. The random oligos were provided from Promega (Oligo(dt15) Primer-C1101). The qPCR design was performed following a protocol developed to determine the titre of lentiviral vectors: Comparison of lentiviral vector titration methods; *BMC Biotechnology* 2006, 6:34

doi:10.1186/1472-6750-6-34 (2006). The oligos and the GFP probe design were identical: 5'- GGAGCGCACGATCTTCTTCA-3' and 5'-AGGGTGTCGCCCTCGAA-3'; the 5'Taq-Man probe 5'-FAM-CTACAAGACCCGCGCCGAGGTG-TAMRA-3'. Q-PCRs were developed using Brilliant III Ultra-Fast QPCR Master Mix (Agilent-600880).

#### qPCR reactions

Forward primer (10 $\mu$ M)	0.5 $\mu$ l per reaction	
Probe (5 $\mu$ M)	0.5 $\mu$ l per reaction	+ 2 $\mu$ l of cDNA
ROX/dye (1:500)	0.3 $\mu$ l per reaction	

#### 2.4.11 PCR

PCRs were carried out using Pfu Turbo DNA Polymerase® (Agilent Technologies, 600250) in 50 $\mu$ l reaction volumes containing the following components:

5 $\mu$ l of 10x Reaction buffer
1 $\mu$ l of dNTPs (10 $\mu$ M)
1.5 $\mu$ l forward oligonucleotide
1.5 $\mu$ l reverse oligonucleotide
1 $\mu$ l DNA template
39 $\mu$ l H <sub>2</sub> O
1 $\mu$ l Pfu polymerase

Initial denaturation	96°C 5 minutes
Denaturation	96°C 30 seconds
Annealing	55°C 30 seconds/1 minute
Elongation	72°C 1 minute/kb
Extra elongation	72°C 2 minutes
Hold	4°C

Primer annealing temperature was determined by primer T<sub>m</sub>. Extension time was determined by the length of the amplicon (1 minute/kb using Pfu Turbo polymerase). DMSO was added to the reaction when GC contents were 60-70% to avoid secondary structures formation in oligonucleotides.

#### 2.4.12 Agarose gel electrophoresis

This technique was performed to separate DNA fragments for either visual analysis or gel purification. Agarose solution was prepared in TAE buffer, transferred into a horizontal gel tank and fully covered with TAE as running buffer. Loading dye was added to samples (Promega) in 1:6 ratio dye/simple. Either 1kb DNA ladder was used to determine the size of DNA fragments (Promega, G5711 ) or 100bp DNA ladder (Promega, G2101). Agarose

gels were run at 120 V until DNA separation was achieved. DNA fragments were observed using a UV transilluminator and purified from the gel using a Purelink Quick Gel Extraction Kit, Invitrogen K210012) following manufacturer instructions.

#### 2.4.13 Western blot analysis

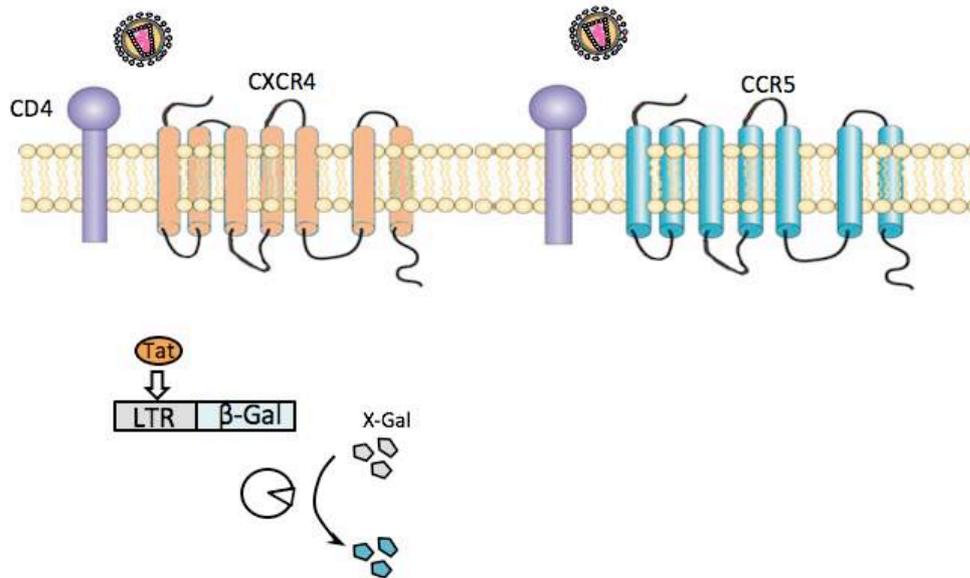
Transfected HEK-293T cells were treated in protein sample buffer and sonicated for 45 seconds at 4 kg/m<sup>2</sup>. Filtered supernatants containing viruses were purified by centrifugation on sucrose cushions (20%) in 1.5 ml tubes and centrifuged at 13000 rpm during 4 hours at 4°C. Viruses were resuspended in 50 µl of protein sample buffer. All samples were heated at 70°C for 5 minutes and resolved on 4-12% Bis-Tris acrylamide gradient gels (Life Technologies, UK). Proteins were separated by electrophoresis at 140 V for 1.5 hours in MOPS buffer and transferred onto nitrocellulose membranes (Invitrogen™ Nitrocellulose Membranes) for liquid blotting transfer technique overnight. The membranes were blocked with 50% heat inactivated bovine calf serum-PBS solution (50%) for 30 minutes shaking at 20 oscillations/minute at room temperature. Then membranes were probed with antibodies diluted in 50% FCS-PBS. The membranes were washed with PBS-Tween solution (1%) after each antibody incubation step and prior revealing the results using a LiCor Odyssey Scanner with Image Studio Software.. The secondary antibody was GAM goat anti-mouse [IRDye800 Thermo Scientific (1:10000)].

#### 2.4.14 TZM-bl assays

This cell line contains integrated reporter genes for firefly luciferase and *E. coli* β-galactosidase under control of an HIV-1 LTR. The aim of this assay is to detect HIV-1 Tat protein expression, which binds to the LTR stimulating transcription of the *lacZ* gene, originating the α-fragment. The β-galactosidase binds to the α-fragment and hydrolyses the X-Gal generating a blue precipitate where Tat has been expressed. (Figure 2-3). TZM-bl cells were seeded in 96 well plate (1x10<sup>5</sup> cells/well) and infected with viruses generated in HEK-293Ts. The plate was incubated for 48 hours and fixed in 2.5% glutaraldehyde (Sigma-Aldrich, UK). The fixative was removed and wells were washed twice with PBS before adding the X-Gal staining solution. Finally, the plate was incubated at 37°C in the dark for 4 hours.

X-Gal solution

5.3 ml PBS  
 300  $\mu$ l 100mM  $K_4Fe(CN)_6$   
 300  $\mu$ l 100mM  $K_3Fe(CN)_6$   
 6  $\mu$ l  $MgCl_2$  (1M)  
 60  $\mu$ l X-Gal (50mg/ml)



**Figure 2-3.** Diagram of a standard TZM-bl assay. TZM-bl cells are a reporter cell line presenting an *E. coli*  $\beta$ -galactosidase gene under control of an HIV-1 LTR (Tat dependent) to measure Tat expression after a single round of infection. TZM express HIV-1 receptor and co-receptor (CD4 and CXCR4 respectively). Diagram adapted from Akrit Sodhi, Silvia Montaner & J. Silvio Gutkind, *Nature Reviews Molecular Cell Biology* 5, 9981012 (December 2004).

#### 2.4.15 Cell preparation for flow cytometric analysis

Suspension cells were transferred into 96 well plates in an equal volume (100:100) of 2.5% glutaraldehyde (Sigma-Aldrich, UK) and 10.000 events were collected using a Guava EasyCyte flow cytometer (Merck Millipore) measuring the percentage of GFP-positive infected cells. To analyse adherent cells, the media was removed and replaced by trypsin 0.05% EDTA (Sigma-Aldrich) and then transferred into 96 well plates in an equal volume (100:100) of 2.5% glutaraldehyde.

## **Chapter 3**

## 3.1 Results

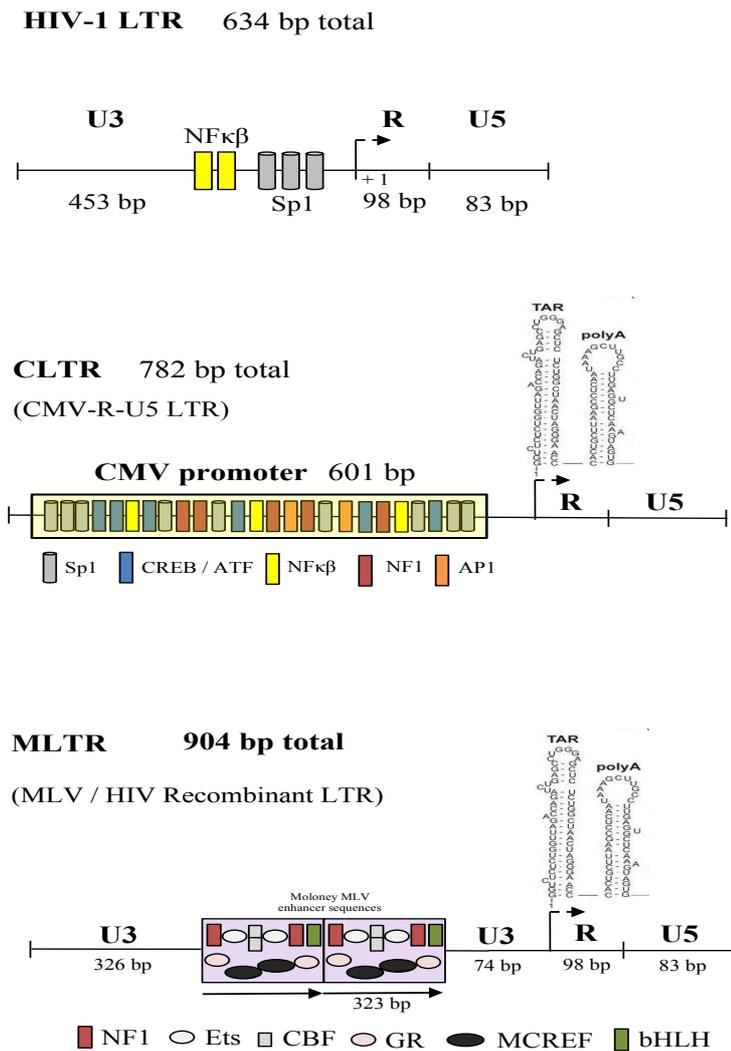
### 3.1.1 CLTR and MLTR chimeric promoters support transcription and reverse transcription

In order to free HIV-1 from Tat-dependence, we needed to generate Tat-independent LTRs. Several approaches have been taken previously to enhance viral gene expression in lentiviral vectors, such as pCVR1\_NLGP (Figure 3-1), which encodes a CLTR chimeric promoter (CMV-R-U5). This construct is potentially Tat-independent and engenders high levels of expression due to the replacement of the U3 region in 5'LTR with a CMV promoter (Figure 3-2). This idea was followed for the generation of the MLTR (MLV-R-U5) by inserting heterologous enhancer elements into the HIV-1 LTR (Chang et al. 1993). The MLTR encodes enhancer sequences derived from the MLV LTR promoter that stimulates transcription in a Tat-independent manner (Figure 3-1).

The aim of this project was to generate an HIV-1 clone able to replicate within murine cells, and the MLTR was a good candidate to pursue because MLV is adapted to the murine host. In summary, the CLTR and MLTR promoters should be able to enhance viral gene expression in the absence of Tat, leading to the generation of sufficient viral RNA levels to express viral proteins and genomic RNA necessary to complete the viral replication cycle. We were able to generate lentiviral vectors in our lab by co-transfecting the essential viral genes into HEK-293Ts to produce high amounts of virions. It is necessary to provide the *gag*, *pol* and *env* genes in *trans* alongside a lentiviral genome including LTRs that can be reverse transcribed and subsequently integrated into the host genomic DNA. These vectors were originally designed for gene therapy requiring efficient transgene expression.

HIV-1 structural proteins are provided during transfection from pCVR1\_NLGP (*gag-pol* expression plasmid derived from NL4-3 analogous to p8.9 commonly used in lentiviral vector generation), pCMV\_VSV-G (a VSV-G envelope expression plasmid) and a lentiviral vector genome encoding plasmid CSGW. HIV-1 Tat and Rev proteins are encoded in pCVR1\_NLGP together with the structural HIV-1 *gag-pol* genes.

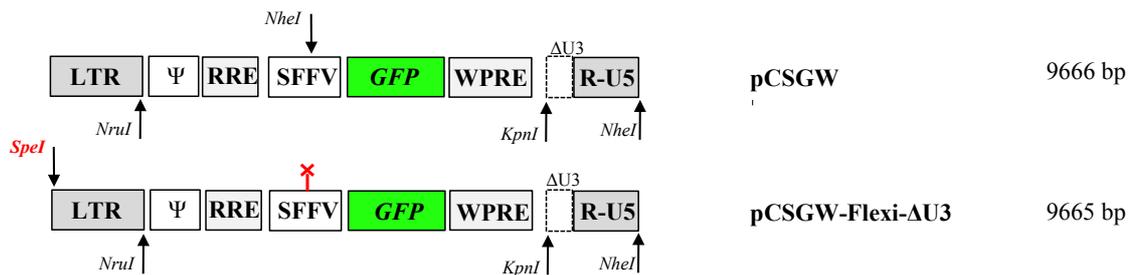
(Figure 3-2) (Demaison et al. 2002).



**Figure 3-2.** HIV-1 LTR, CLTR and MLTR composition. B) Composition of the different LTRs employed in this thesis. The CLTR contains a CMV promoter replacing the U3 region, and the MLTR encodes the Moloney enhancer sequences from the MLV promoter into the U3 region. Adapted from M Martin J. *Virology* 1993, 67(2):743 and Morita E, Arii J, Christensen D, Votteler J, Sundquist WI. *Biotechnology* 2012 Aug ;0(0):1-5. doi : 10.2144/000113909 and Diagram adapted from *Retroviruses*. Coffin JM, Hughes SH, Varmus HE, editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1997.



The CSGW plasmid was modified to test the CLTR and MLTR, prior to including them in replication competent viruses (due to lengthy refurbishment of the MacRobert CL3 laboratories of the Richard Elliott Biosafety Laboratory CL3, it was necessary to, where possible, conduct experiments at CL2). Thus, modification of CSGW facilitated examination of chimeric LTRs at CL2. The “Flexi” CSGW vector is a derivative of the parental vector CSGW, where restriction sites were modified at the 5′ and 3′ end (of each LTR) to enable manipulation of LTRs at both ends of the vector (Figure 3-3).



**Figure 3-3.** Location of the additional *SpeI* site at the beginning of 5′LTR in CSGW\_Flexi\_ΔU3, which is absent in CSGW. We removed the *NheI* site at position 7842 (within the SFFV internal promoter) in CSGW and preserve the *NheI* site at the end of the 3′LTR, unique after this mutagenesis to originate to generate the CSGW\_Flexi\_ΔU3 vector.

There is evidence suggesting that enhancer elements can replace long terminal repeats in lentiviral vectors efficiently, and transcription from chimeric LTRs had no effects on titre, infectivity, and genomic stability (Lotti et al. 2002). In this study, they used a very similar system to modify the HIV-1 LTR by replacing most of the U3 region with an autoregulatory enhancer (HS2) of the erythroid-specific GATA-1 transcription factor, which controls the expression of the gene of interest (GFP). An internal promoter (PGK instead of SFFV) drives the expression of a constitutive reporter gene, generating an efficient transduction of 70% in human and murine cells. The modification of the HIV-1 LTR provides a more efficient vector design, employing several transcription enhancer elements to stimulate viral transcription of the gene of interest. These data also support the hypothesis that generating a Tat-independent virus by including our CLTR and MLTR chimeric promoters into CSGW vector is potentially possible.

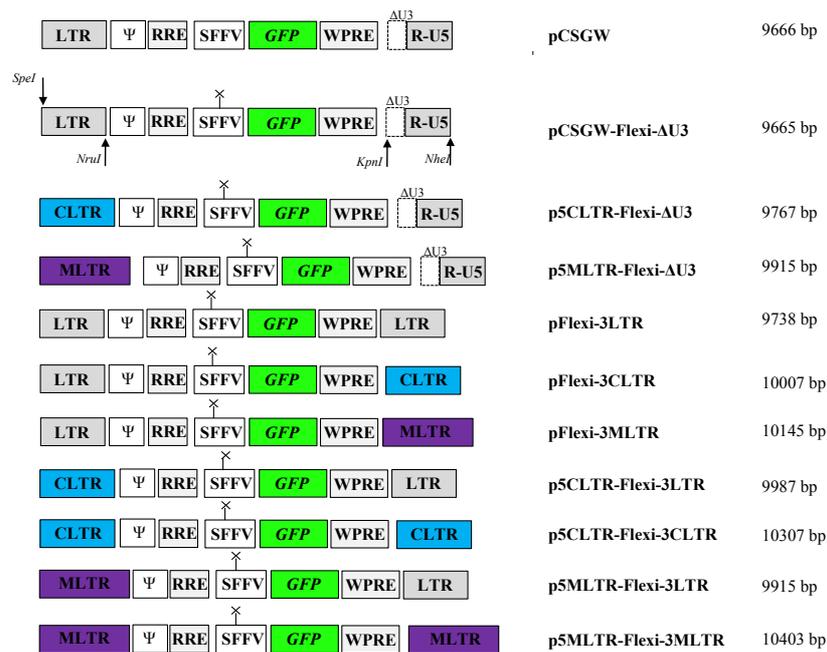
In pCSGW ‘flexi’, the 5′LTR can be replaced using unique *SpeI* and *NruI* sites to include the CLTR or MLTR, to examine the ability of the modified promoters to drive transcription. The ΔU3 3′LTR in CSGW is effectively copied in the 5′ LTR after reverse

transcription and integration of the  $\Delta$ U3 provirus results in the transcriptional inactivation of the proviral LTR (self inactivation). A similar strategy was used at the 3' end, where unique *KpnI* and *NheI* sites flank the 3' LTR. By inserting LTRs at the 3' end allows chimeric LTRs to be examined for their ability to support reverse transcription (necessary for completion of the first cycle of replication) (Figure 3-4, panel A). Crucially, transgene expression is driven by the internal SFFV promoter and should be relatively immune to variation in the nature of the 5' LTR.

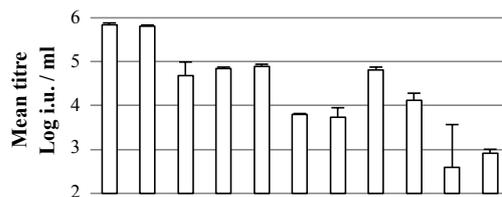
The CSGW\_Flexi clone had a similar titre compared to the parental CSGW construct, highlighting that a G→A substitution in position 7845 to remove the *NheI* site, and the addition of the *SpeI* site in position 5530 (G→C) had little effect on the fitness of this vector. Importantly, the CLTR and MLTR promoter supported transcription and reverse transcription efficiently in the CSGW\_Flexi clone (Figure 3-4, B, lanes 1 and 2). However, the viral titre was decreased when the 5' and 3' LTRs were modified in CSGW\_Flexi clones (Figure 3-4, B, lanes 3 to 11). The HIV-1 structural proteins were provided in *trans* transfecting CMV\_NLGP vector (Figure 3-2) and the Gag/p55 expression was not affected by the various CSGW\_Flexi clones in transfected HEK-293T cells (Figure 3-4, C).

In all cases, VLPs were efficiently generated in the presence of all the clones, suggesting that the infectivity reduction could be related to packaging of the vector genome within the nascent particles, or inefficiency in completing any of the early steps of the viral replication cycle (such as reverse transcription, chromosomal integration or transgene transcription) that precede reporter gene expression (Figure 3-4, D) (Verdin et al. 1993; Bokhoven et al. 2009).

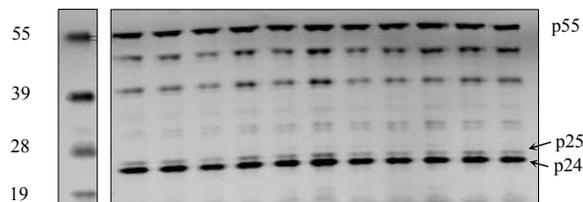
A



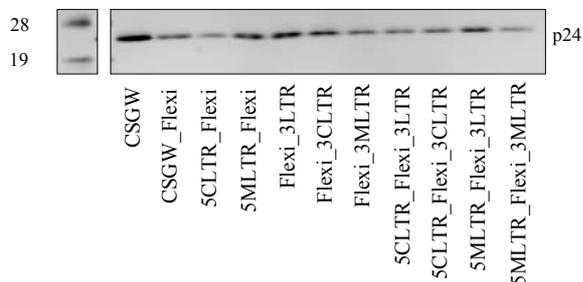
B



C



D



**Figure 3-4.** Multiple CSGW clones encoding HIV LTRs, or chimeric CLTR and MLTR. Molecular weight is shown in kDa. (A) CSGW clones used in this experiment. The ‘X’ represents a deletion of a *NheI* site, and the ‘arrows’ indicate restriction sites employed to modify the 5’ and 3’ LTRs. The envelope is provided by pCMV-VSV-G and the *gag-pol* HIV-1 genes by pCVR1-NLGP. (B) Mean titre of CSGW clones in PM1 cells 48 h.p.i. The DNA transfected amounts were 1  $\mu$ g of each CSGW vector, 1  $\mu$ g of pCMV\_NLGP and 100ng pCMV\_VSV-G. The error bars show the standard deviation for each titre (triplicates) (C) WB analysis (p55 and p24) of cell lysates from transfected HEK-293T cells 48 h.p.t. (D) WB analysis for p24 from purified VLPs collected after centrifugation of sucrose cushions from the supernatant of transfected HEK-293Ts 48 h.p.t. The primary antibody was mouse anti-p24.

### 3.1.2 A 118 base pair fragment preceding the 3' end of CSGW is involved in viral infectivity

Because substituting the 3'ΔU3 LTR with a complete LTR led to a greater than 10-fold reduction in infectious titre (Figure 3-4, panel B), we examined the cloning strategy used to generate these clones to ensure this reduction was caused by a defect in the cloning strategy. Reviewing the cloning strategy to modify the 5' and 3' LTRs revealed a deletion of 118 base pairs preceding the 3'ΔU3 LTR when CSGW\_Flexi was generated. This sequence does not appear to encode any relevant genetic features (Figure 3-4, panel A). However, removing this 118 nucleotides could explain why CSGW\_Flexi clones had reduced infectivity, suggesting that this region plays a role in virus production, perhaps improving the efficiency of minus strand transfer in infected cells (Velpandi et al. (1992), Piekna-Przybylska et al. 2008).

We decided to introduce this 118 bp sequence back into all the CSGW\_Flexi clones encoding the chimeric LTRs to determine if viral fitness was restored by repeating the same experiment as Figure 3-4. The viral titre of the 5'CLTR\_Flexi clone was practically equal compared to titres derived from the CSGW and CSGW\_Flexi clones (Figure 3-6, B, lanes 1 to 3), meanwhile the 5'MLTR titre was decreased by ~1 log compared to CSGW and 5'CSGW\_Flexi (controls). The 5'MLTR\_Flexi titre was also reduced by ~1 log compared to the 5'CLTR\_Flexi clone (Figure 3-6, B, lanes 3 and 4).

The addition of a competent 3'LTR replacing the 3'ΔU3 LTR did not affect reverse transcription (Figure 3-6, panel B, lanes 5,6 and 7). A reduction of GFP expression in target cells could be related to a reduction in reverse transcription in CSGW\_Flexi clones, but we have not determined the RT levels for each vector in our experiments. However, the 5'CLTR\_Flexi\_3CLTR titre was decreased by ~1 log (Figure 3-6, panel B, lane 9). The 5'MLTR\_Flexi\_3LTR and 5'MLTR\_Flexi\_3MLTR had the lowest titre (Figure 3-6, panel B, lanes 10 and 11), a predicted result considering the titre reduction of 5'MLTR\_Flexi clone, where the MLTR drove transcription more poorly than the 5'LTR and 5'CLTR (Figure 3-6, panel B, panel B). In summary, the CSGW\_Flexi clones including the 118 bp fragment possessed similar titres to CSGW and CSGW\_Flexi clones (excepting the 5MLTR\_Flexi, 5CLTR\_Flexi\_3CLTR, 5MLTR\_Flexi\_3LTR and 5MLTR\_Flexi\_3MLTR clones) compared to the clones lacking the 118 bp (Figure 3-6, panel B).

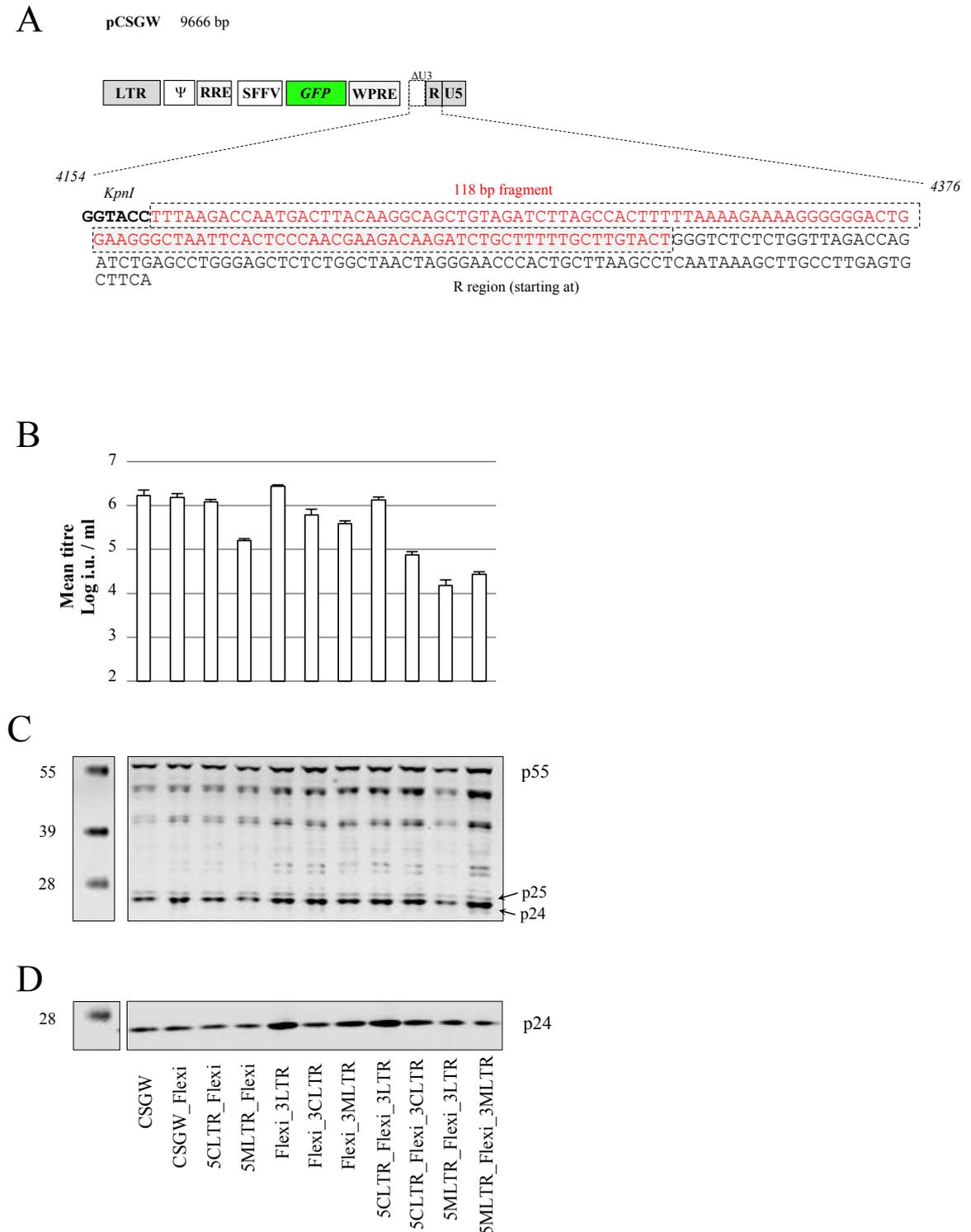


determine Gag intermediates generated after HIV-1 protease cleavage (Figure 3-6, panel C).

We used CVR1\_NLGP vector again to provide the HIV-1 structural proteins in this experiment (Figure NLGP). Gag expression was similar for all CSGW\_Flexi clones including the 118 bp fragment (Figure 3-6, panel C). Some Gag intermediate products were observed in the cell lysates from transfected HEK-293T cells, especially when the 3' end was modified (Figure 3-6, panel C, lanes 5-11). In addition, capsid expression was very similar in all CSGW\_Flexi clones, and p24 expression was higher than p25 (bottom and upper band respectively in CA lane) (Figure 3-6, lanes 1 to 11). The p25 detectable in HIV-1 infected cells is composed of four species with related isoelectric points and related to the phosphorylated state of p25. The VLPs collected from the supernatant of transfected HEK-293T cells 48 hours post-transfection were filtered and purified using 20% sucrose cushions to pellet the virions. Then, we used western blotting to detect the amount of VLPs generated in the presence of each CSGW\_Flexi clone, which were very similar for all CSGW\_Flexi clones (Figure 3-6, panel D).

These experiments shown the ability of the CLTR and MLTR chimeric promoters to carry out transcription and reverse transcription efficiently, but we still observe substantial reductions in the viral titres compared to the HIV-1 LTR (Figure 3-6, panel B). Thus, we have not determined reverse transcription levels and these results can not confirm that RT levels are not affected in CSGW\_Flexi clones. However, we did not measure directly the reverse transcription levels for each CSGW clone, so we could only determine the efficiency of transcription and reverse transcription based on GFP expression after transfection and GFP expression in the target cells (PM1). We could have performed reverse transcription assays to determine the reverse transcription levels in our clones, and also investigate the RNA species within the transfected cells. Unfortunately, these experiments will consume much time and we decided to continue our investigations with the CLTR and MLTR to achieve Tat Independence in human cells before we move to murine cells. The presence of *tat* and *rev* genes within pCVR1\_NLGP plasmid obscures our ability to determine the ability of the CLTR and MLTR to drive transcription and reverse transcription in a Tat and Rev independent manner. It is necessary to remove the HIV-1 regulatory genes from the experiment to investigate the real ability of the CLTR and MLTR. We can solve this issue by replacing pCVR1\_NLGP for pSYNGP, which encodes the HIV-1 *gag* and *pol* genes codon-optimized for humans and lacks *tat* and *rev*

genes, so we will be able to repeat these experiments again within a Tat and Rev free environment.



**Figure 3-6.** CSGW clones encoding HIV LTR, CLTR, MLTR and the 118 bp region from CSGW backbone. Molecular weight is shown in kDa (A) Location of the 118 bp region in CSGW. (B) Mean titre of CSGW clones encoding the 118 bp region in PM1 cells 48 h.p.i. Three plasmids were transfected to obtain the virions: pCSGW clone (wt-CSGW as control), pCMV\_VSV-G to generate the envelope and pCVR1\_NLGP to provide *gag-pol* HIV-1 genes. The DNA transfected amounts were 1 $\mu$ g of each CSGW vector, 1 $\mu$ g of pCMV\_NLGP and 100ng pCMV\_VSV-G. The viral titre was determined by flow cytometry measuring GFP expression in PM1 cells. The error bars show the standard deviation for each titre (triplicates) (C) WB analysis (p55 and p24) of cell lysates from transfected HEK-293T cells 48 h.p.t. (D) WB analysis for p24 from purified VLPs collected after centrifugation of sucrose cushions from the supernatant of transfected HEK-293Ts 48 h.p.t. The primary antibody was mouse anti-p24.

### **3.1.3 Viral particles derived from SYNGP clones are efficiently made in human and murine cells. Codon optimization and chimeric CLTR facilitate Tat and Rev independent VLP production from human and murine cells**

The aim of this project was to generate infectious particles in mice in a Tat and Rev independent manner. To achieve this ambitious aim, it is necessary to replace HIV-1 Rev function, involved in RNA stability and transport (Pollard et al. 1998). HIV-1 regulatory proteins cannot exert their function in murine cells, and viral RNA levels are insufficient to complete the viral life cycle (Barbara et al. 1993). The codon-optimized HIV-1 *gag-pol* genes (encoded in SYNGP) increase the amount of *gag-pol* RNAs within human cells due to the adaptation of HIV *gag* and *pol* genes to the human codon bias, leading to high expression levels of *gag-pol* viral RNAs generating infectious progeny in the absence of Rev (Kotsopoulou et al. 2000). Previous results demonstrated that chimeric CLTR and MLTR LTRs were functional in CSGW\_Flexi clones when HIV-1 *gag-pol*, *tat* and *rev* genes were provided by pCVR1\_NLGP, and the envelope by pCMV\_VSV-G (Figure 3-6).

Following this result we decided to substitute pCVR1\_NLGP by pSYNGP to determine if single cycle of replication viruses are generated in a Tat and Rev independent manner (in both human and murine cells). The single cycle of replication viruses were generated by transfecting pCSGW (vector genome), pCMV\_VSV-G (VSV-G envelope expression plasmid) and SYNGP (a GagPol expression plasmid which replaces pCVR1\_NLGP) in human HEK-293T and murine NIH-3T3 cells. We reasoned that the codon-optimized *gag-pol* genes and the CLTR could replace Rev and Tat function and permit the generation of infectious particles from murine cells.

To attempt this experiment, the CMV promoter located at the 5' end of the parental pSYNGP vector was replaced with the CLTR by cleavage with *SnaBI* and *EcoRI* to introduce the CLTR. Another CLTR was cloned at the 3' end of pSYNGP, causing the effective deletion of the poly A signal in pSYNGP during the cloning process (Figure 3-7, panel A). The lack of *tat* and *rev* genes in pSYNGP made it possible to investigate the viability of these promoters in a Tat and Rev free environment using this 5CLTR\_SYNGP\_3CLTR clone (Figure 3-7, panel A). Gag/CA expression levels in transfected human HEK-293T cells were higher for the control (NLGP), but SYNGP also expressed Gag/CA efficiently (Figure 3-7, upper panel B). However, in murine NIH 3T3

cells, Gag synthesis was barely detectable and correspondingly, very little capsid expression was detected when *gag-pol* genes were provided by NLGP (Figure 3-7, bottom panel B, lane 1). Regardless of whether NLGP or SYNGP was used, CA/p24 expression was higher compared to p55 in human and murine cell lines (Figure 3-7, left panel, lanes 1-4). The human codon-optimized *gag-pol* genes from pSYNGP were able to rescue Gag and Capsid expression in transfected NIH 3T3 cells (Figure 3-7, bottom panel B, lanes 2-4).

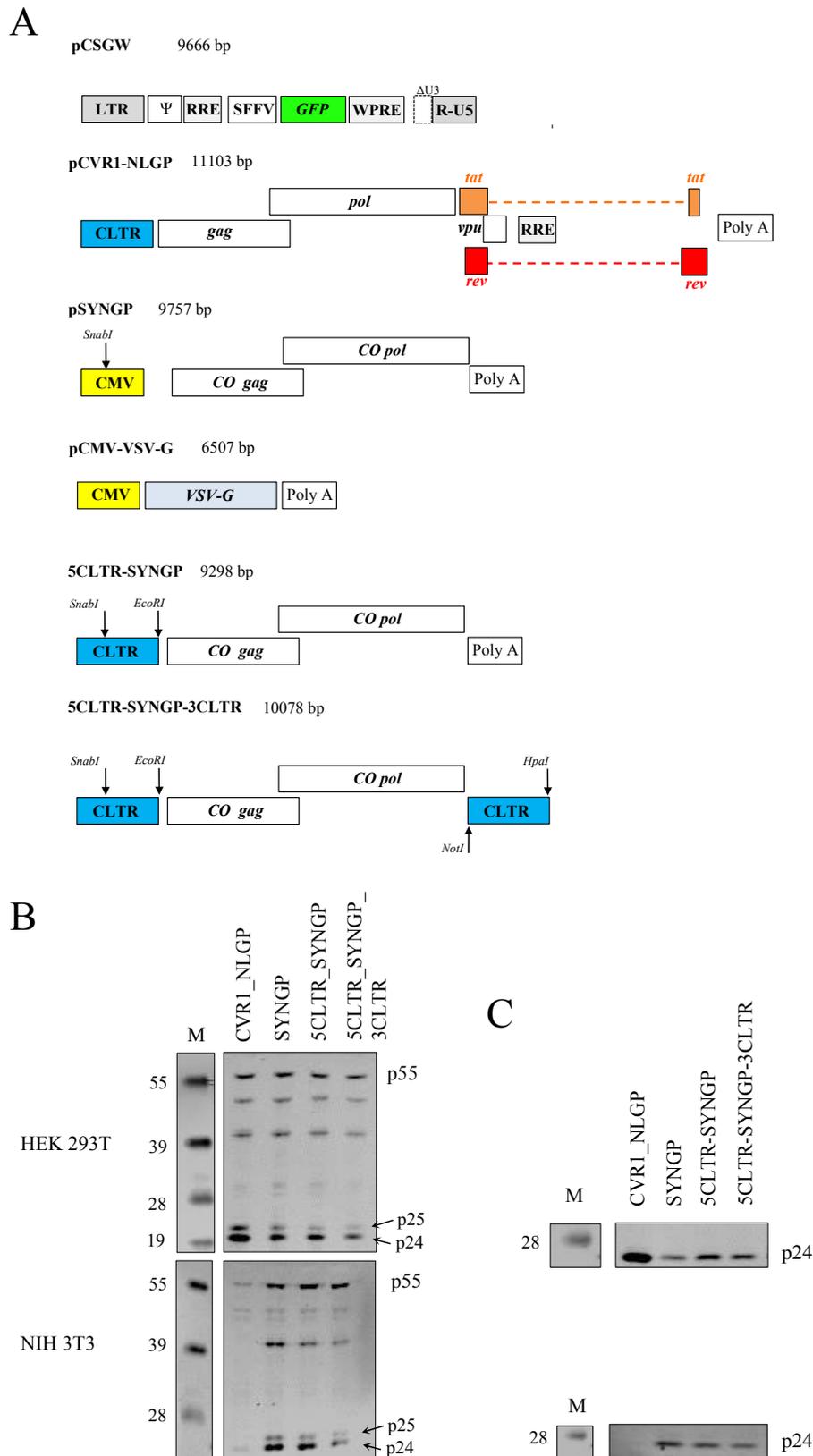
The amount of VLPs collected from transfected HEK-293T and generated using SYNGP was lower compared to the amount of VLPs detected using NLGP (Figure 3-7 right panel C, lane 1-4). Importantly however, NIH 3T3 cells were able to generate codon-optimized HIV-1 particles in a Tat and Rev independent manner (Figure 3-7, right bottom panel C, lanes 2-4) whereas wild type *gag-pol* HIV-1 genes encoded in pCVR1\_NLGP were unable to express sufficient Gag to enable the assembly of nascent VLPs in NIH 3T3 cells (Figure 3-7, panel C). The SYNGP construct was able to generate HIV particles in murine cells meanwhile pCVR1\_NLGP failed because of the lack of codon-optimized *gag* and *pol* genes in this vector. The presence of *tat* and *rev* genes encoded in pCVR1\_NLGP contributed to the more efficient generation of VLPs than pSYNGP in human cells, but no marked differences were observed in Gag expression in HEK-293T transfected with NLGP and SYNGP (Figure 3-7, left panel B, lanes 1-4).

Tat and Rev proteins are unable to exert their function in murine cells, and consequently *gag-pol* genes were not expressed from pCVR1\_NLGP (which lacks codon optimization) in murine cells. The codon-optimization from SYNGP contributes to increase the amount of *gag-pol* RNAs within the host cell, and subsequently Gag polyprotein expression is rescued in NIH-3T3 murine cells (Figure 3-7, panel B and C). However, wild type HIV-1 *gag* and *pol* genes encoded in CMV\_NLGP vector (lacking codon-optimization) were unable to generate sufficient *gag-pol* RNAs in murine cells, impairing Gag production (Figure 3-7, panel B and C).

We have demonstrated that codon optimization rescued *gag-pol* expression in human and murine cells, and VLPs can be generated efficiently. However, no infection was detected after the inoculation of VLPs from 5CLTR\_SYNGP and 5CLTR\_SYNGP\_3CLTR clones onto PM1 cells. Gag polyprotein was generated in both cell types, and it was processed successfully to form CA protein p24 (Figure 3-7). Perhaps the absence of the HIV-1

---

packaging signal in 5CLTR\_SYNGP and 5CLTR\_SYNGP\_3CLTR clones could be the major factor impairing viral replication within human and murine cells, and the 5CLTR\_SYNGP and 5CLTR\_SYNGP\_3CLTR genomes could be unable to be packaged into the nascent viral particles (Clever et al. 2002).



**Figure 3-7.** (A) Diagram of the constructs used in this experiment. The SYNGP clones encode a CLTR at the 5'end, 3'end or both. (B) WB for capsid (p24) of transfected human HEK-293Ts and murine NIH 3T3 cells 48 hours post transfection. (C) WB analysis (p55 and p24) of VLPs produced from the previous cell lines 48 h.p.t. The primary antibody was mouse anti-p24. It is remarkable that pCVR1\_NLGP was unable to generate VLPs in murine cells but SYNGP clones could rescue p55 and p24 expression. The DNA transfected amounts were 1 $\mu$ g of CSGW vector, 1 $\mu$ g of pCMV\_NLGP/pSYNGP vector and 100ng pCMV\_VSV-G.

### 3.1.4 Introducing the HIV-1 packaging signal ( $\Psi$ ) in SYNGP clones did not increase the efficiency of particle production

The viral particles generated by SYNGP clones could be lacking the viral genome due to the absence of the HIV-1 packaging signal ( $\Psi$ ), which is essential to package the genome within the nascent particles (Jouvenet et al. 2009). Secondary structure studies have determined that the HIV-1 packaging signal contains several RNA loops (SL1 to SL4) involved in packaging the viral genome into the particles through interactions with Gag (Zeffman et al. 2000).

Despite the absence of the HIV-1 packaging signal in 5CLTR\_SYNGP and 5CLTR\_SYNGP\_3CLTR clones, they efficiently generated viral particles when CSGW and VSV envelope were provided in *trans*, as the CSGW has an intact  $\Psi$  and would be packaged into the particles, but no infection was detected when they were inoculated onto human lymphocytes PM1 (Figure 3-7). Perhaps the lack of  $\Psi$  in pSYNGP is affecting viral RNA stability-folding necessary for Gag-RNA interaction, causing the production of VLPs lacking the viral genome (Kuzembayeva et al. 2014). The HIV-1 packaging signal adopts a highly branched secondary structure, which is essential to achieve the entrance of the viral genome within the nascent particles. The absence of  $\Psi$  in 5CLTR\_SYNGP and 5CLTR\_SYNGP\_3CLTR could be determinant in the absence of viral infectivity and also impair the the appropriate RNA folding necessary to perform Gag-RNA interaction.

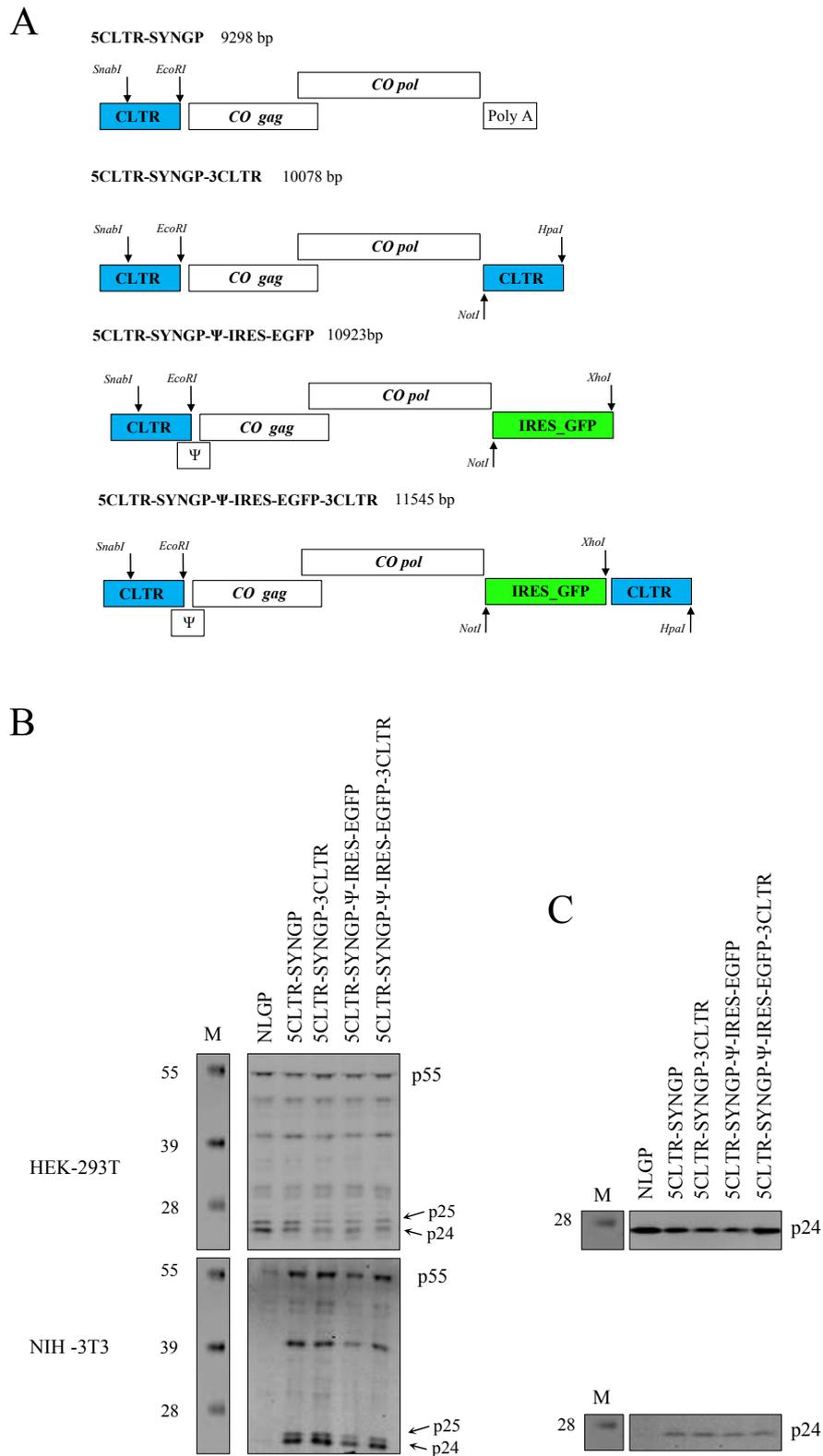
An internal ribosome entry site (IRES) was cloned into the 5CLTR\_SYNGP and 5CLTR\_SYNGP\_3CLTR clones, followed by a GFP reporter gene by cleavage with *NotI* and *XhoI* (Figure 3-8). HIV-1 harbours an IRES located within 5'UTR and the *gag* coding region, which can drive 40S ribosomal subunit recruitment and positioning on the mRNA either at or upstream of the start codon (Vallejos et al. 2012) and contributing to increase translation of the viral proteins in HIV-1 infected cells (Gendron et al. 2011). Then, we decided to clone the HIV-1 packaging signal into the 5CLTR\_SYNGP\_IRES\_GFP and 5CLTR\_SYNGP\_IRES\_GFP\_3CLTR clones to determine if the viral particles become infectious. We deoptimized the previously codon-optimized 21 nucleotides at the beginning of *gag* to instead preserve the RNA structure of the SL4 loop (Keane et al. 2015). We followed this strategy to preserve as much as possible the original position that  $\Psi$  has within an HIV-1 genome: starting at the end of the U5 region from the 5'LTR until the first nucleotides of *gag* (Figure 3-8, panel A). (Olson et al. 2015; Paillart et al. 2014).



The resulting vectors were named 5CLTR\_Ψ\_SYNGP\_IRES-GFP and 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_3CLTR. We repeated the same experiment as shown in Figure 3-7, but using these new clones instead to determine if the addition of the Ψ was able to rescue the infectivity of the VLPs.

Gag expression was not increased or reduced either in transfected HEK-293T cells or NIH-3T3 cells (Figure 3-9, panel B). In addition, the introduction of the Ψ did not increase the amount of VLP production collected from the supernatant 48 hours post-transfection in both cell types, generating similar amounts to SYNGP clones lacking Ψ (Figure 3-9, panel B). We detected Gag intermediates in transfected human HEK-293T and murine NIH-3T3 as well (Figure 3-9, panel B). However, the VLPs generated from 5CLTR\_Ψ\_SYNGP\_IRES-GFP and 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_3CLTR constructs within both cell types were not infectious despite the successful generation of viral particles. The codon-optimized *gag-pol* genes from SYNGP contributed to the generation of more Gag protein within murine cells, meanwhile the wild type *gag-pol* genes from NLGP were not expressed from pNLGP in murine cells (Figure 3-9, panel B). The percentage of transfected HEK-293T cells was much higher than NIH 3T3 cells (Figure 3-9, panel D), but codon-optimization efficiently rescued Gag expression in murine cells despite the low percentage of transfection (about 30-50% approximately).

Western blots for Gag/CA demonstrated the potential of murine NIH-3T3 cells to generate HIV-1 Gag protein until its maturation to create the viral Capsid, suggesting that the CLTR was efficiently driving the transcription of codon-optimized *gag-pol* RNAs (Gao et al. 2003). HIV Gag protein was successfully cleaved and processed to generate p24 protein and subsequently viral particles. The formation of VLPs suggest that viral proteins encoded within codon-optimized *pol* gene could be successfully generated, but further investigations need to be done to determine if VLPs produced from 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_3CLTR clone had an adequate morphology because we can not determine that by western blot assays. Based on the results obtained from the CSGW\_Flexi system, the 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_3CLTR clone should be able to support efficient transcription and reverse transcription, potentially generating infectious viral particles in human cells. However, the lack of infection in the target cells using this construct suggests that problems exist with this strategy.



**Figure 3-9.** (A) Composition of SYNGP clones encoding CLTRs, the HIV-1 packaging signal ( $\Psi$ ) and an IRES\_GFP. (B) WB analysis (p55 and p24) using mouse anti-p24 antibody from cell lysates collected from HEK-293T (human) and NIH 3T3 (murine) cells 48 hours post transfection. (C) WB analysis for p24 from purified VLPs collected after centrifugation of sucrose cushions from the supernatant of transfected HEK-293Ts and NIH-3T3 cell lines 48 h.p.t. The primary antibody was mouse anti-p24. Molecular weight marker in kDa. The DNA transfected amounts were 1 $\mu$ g of CSGW vector, 1 $\mu$ g of each pSYNGP vector and 100ng pCMV\_VSV-G.

### 3.1.5 DEST40-*tat* drives Tat expression efficiently in TZM-bl cells

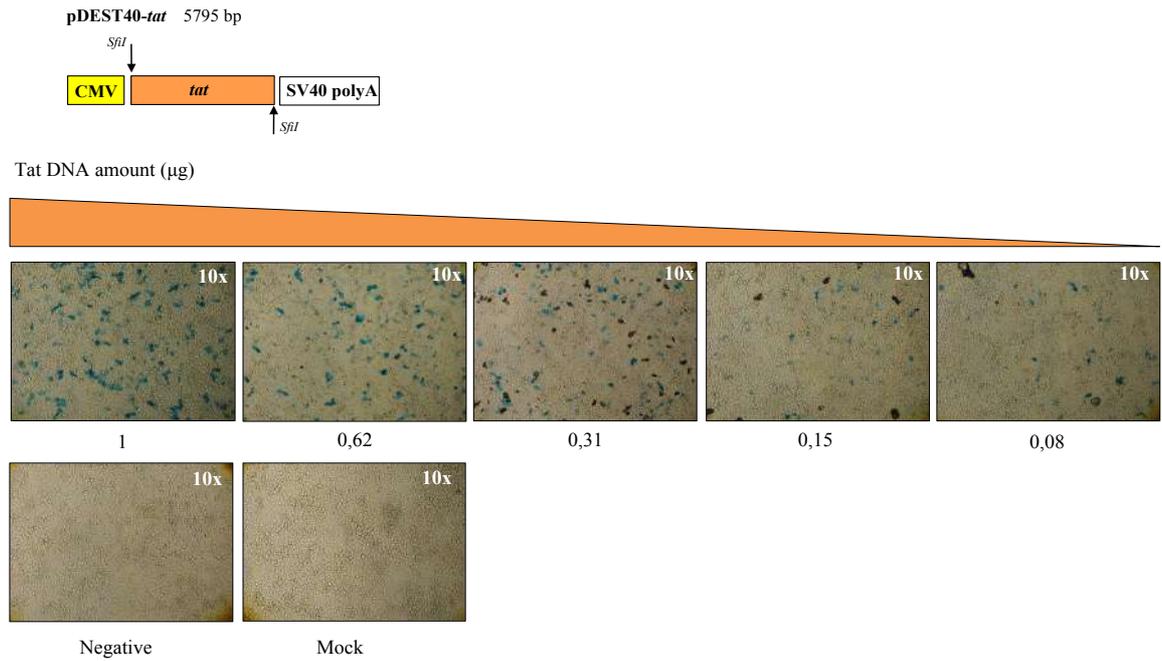
We wanted to investigate the absence of infectivity in 5CLTR\_Ψ \_SYNGP\_IRES-GFP\_3CLTR clone by including Tat into the system to check if its infectivity could be rescued. It was necessary to clone a *tat* gene into another vector to facilitate Tat expression *in trans*. Tat exon 1 and 2 were cloned into the expression vector DEST40 by cleavage with *SfiI* (introns were removed). The CMV promoter encoded in DEST40 will drive transcription of Tat after transfection.

First, we wanted to test if pDEST40-*tat* actually was able to express functional levels of Tat in TZM-bl cells, HeLa cells that have been modified to detect *tat* expression by TZM-bl assay (Montefiori et al. 2006). TZM-bl cells have been genetically modified to encode a  $\beta$ -galactosidase gene under the control of an HIV-1 LTR promoter. After transfection, Tat will stimulate transcription of the LTR, leading to production of RNA and subsequently the synthesis of the  $\beta$ -galactosidase enzyme after translation, which will hydrolyze the X-Gal (lactose analogue), creating a visible blue precipitate.

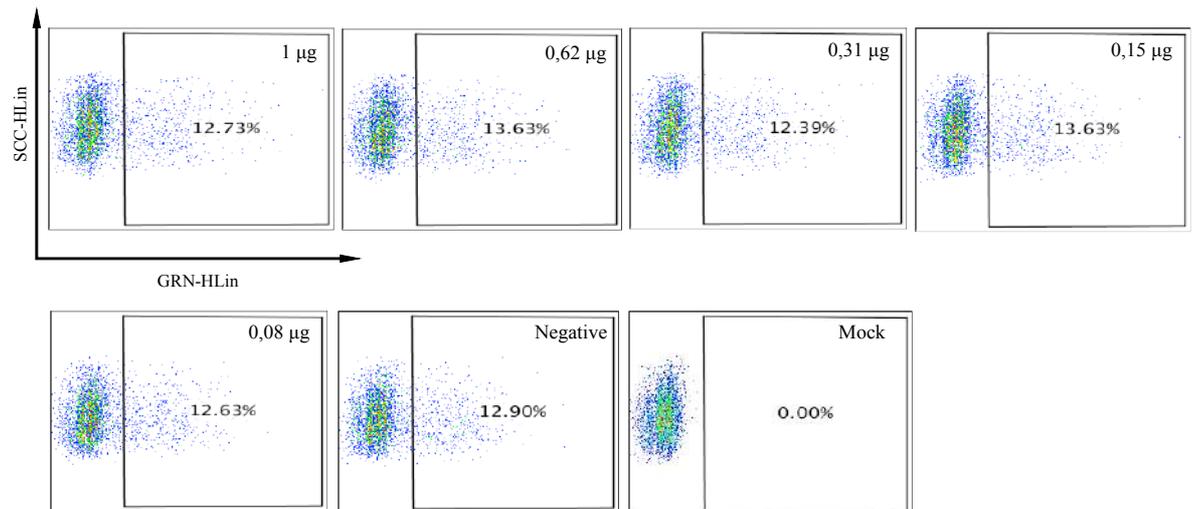
The pDEST40-*tat* was transfected at increasing doses to determine the best concentration for Tat expression. Higher concentrations of pDEST40-*tat* resulted in more blue precipitates (Figure 3-10, panel A). Some crystals were observed in all transfections, even in the mock control (Figure 3-10, panel A), likely originated from the X-Gal solution employed in the experiment. We used the pIRES2\_GFP vector as transfection control for all conditions (Figure 3-10, panel B), and blue foci per field were quantified to determine more accurately the amount of functional Tat expression (Figure 3-10, panel C).

Transfection efficiency was consistent at 70-80% under all conditions in this experiment (Figure 3-10, panel B), demonstrating that the decrease in blue foci was not due to reduced transfection. Higher amounts of DEST40-*tat* plasmid correlated to more blue precipitates observed, demonstrating that DEST40-*tat* expressed Tat efficiently and can be employed in further experiments.

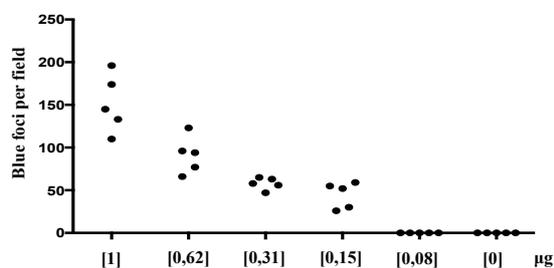
A



B



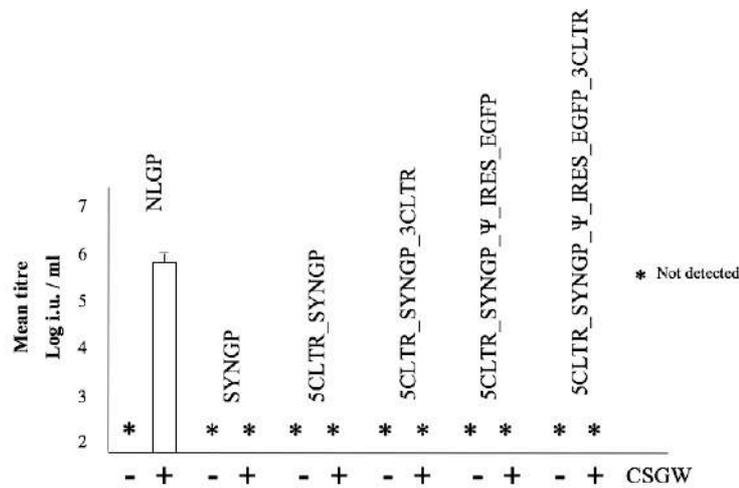
C



**Figure 3-10.** Tat expression driven from pDEST40-*tat* vector. (A) TZM-bl cells transfected with increasing amounts of pDEST40-*tat* DNA. (B) Flow cytometric analysis of TZM-bl transfected with pDEST40-*tat* and pCRV1-EGFP (transfection control). The SCC-HLin (side scatter) determined the size and granularity of the cell population selected for this experiment. The GRN-HLin axis determined the intensity of GFP expression in TZM-bl transfected with pCVR1-GFP (control). (C) Blue cells foci per field. The measurement of blue foci per field was performed in 5 wells for each condition.

### 3.1.6 HIV-1 codon-optimized VLPs are infectious in a Rev independent manner when CSGW and DEST40\_ *tat* are provided *in trans*

Once we demonstrated the ability of DEST40-*tat* to generate Tat efficiently in TZM-bl cells, we wanted to transfect this vector together with 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_3CLTR clone to determine if its infectivity could be rescued. We performed the same experiment as shown in Figure 3-9, but this time Tat was added *in trans*. No infectivity was detected in PM1 cells in the absence of CSGW from VLPs generated in human and murine cells, despite providing Tat *in trans* (Figure 3-11).



**Figure 3-11.** Mean titre of SYNGP clones and NLGP (positive control) in PM1 cells. The SYNGP clones were transfected into HEK-293T cells and produced VLPs were collected and purified from the supernatant by centrifugation of sucrose cushions 48 h.p.t. Then, VLPs were inoculated into PM1 cells to determine the viral titre. CSGW vector was added *in trans* with pCMV\_VSV-G to determine if CSGW could rescue SYNGP clones infectivity. We transfected the same amount of DNA as previous experiments: [1μg of CSGW: 1μg of SYNGPwt/SYNGP vector: 100ng pCMV\_VSV-G]

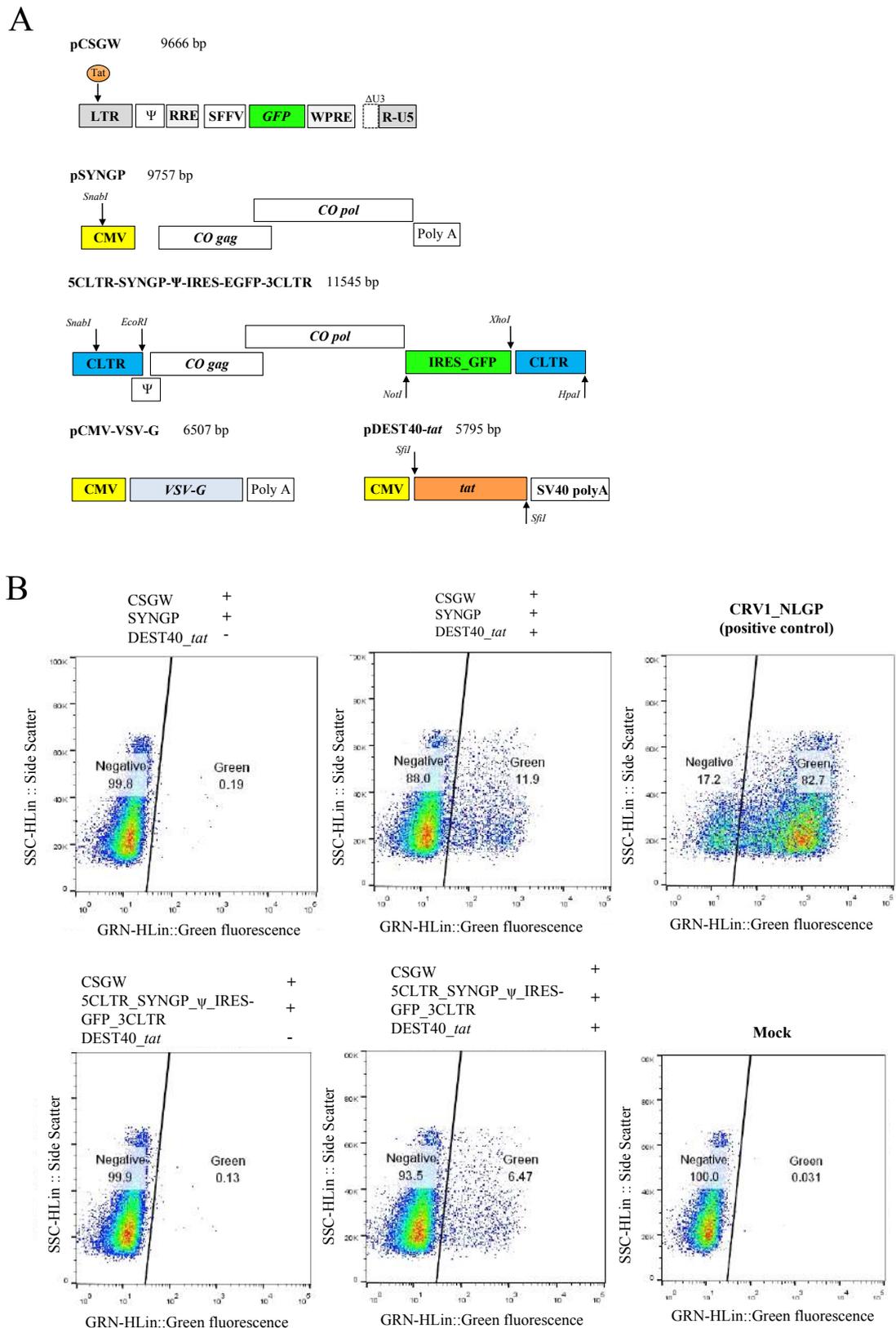
However, transfecting CSGW together with SYNGP and DEST40-*tat* in HEK-293T cells generated infectious viral particles (12% of cells were infected) (Figure 3-12, panel B). The addition of Tat *in trans* rescued the infectivity of SYNGP, so we wanted to determine if we could obtain the same result using 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone instead SYNGP. When 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone was co-transfected together with CSGW and DEST40-*tat*, they efficiently generated infectious VLPs that

infected around the 6.47% of PM1 cells (target cells) (Figure 3-12, panel B). However, no infectious particles were generated in the absence of pCSGW when p5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR, pDEST40-*tat* and pCMV\_VSV-G were co-transfected together (Figure 3-12, panel B). The fitness of these viruses was much lower than NLGP viruses employed in this experiment as positive control (82.7% of infected cells) (Figure 3-12, panel B). It is not surprising considering the presence of the regulatory proteins encoded in pCVR1\_NLGP. These data demonstrated that the structural genes expressed from 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR were capable of packaging a competent genome in *trans* (CSGW) and be infectious (delivering transgene expression). The structural proteins encoded by 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR, could, in principle, package the RNA expressed from the clone, which encodes two competent LTRs, HIV-1 gag-pol codon optimized genes and the packaging signal. Unfortunately, the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone appeared to still be not infectious and more investigation needs to be done to determine which stage(s) of the virus lifecycle fails to be successfully completed.

The CLTRs located at both ends of the vector-genome should be able to drive transcription and reverse transcription, but unknown problems are impairing replication of 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR when CSGW is not provided in *trans* (Figure 3-12). Co-transfecting CSGW and DEST40-*tat* with the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone resulted in infectious viral progeny, suggesting that CSGW was being packaged into the viral particles generated by 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR. This result demonstrated that *gag-pol* were efficiently generated, but the full-length viral RNAs from 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR could not be exported to the cytoplasm in the absence of Rev, remaining in the nucleus and unable to be packaged into nascent particles. Another possibility is that full-length RNAs were not generated within the host cell, or perhaps their stability was compromised in the absence of Rev. However, the absence of Rev did not affect the infectivity of the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone, which was infectious when we added with pCSGW and pDEST40-*tat* together with pCMV\_VSV-G (envelope) during transfection. The amount of cells infected by 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR in this experiment was 6.47% (Figure 3-12, panel B). This lower percentage could be related to the absence of Rev in the experiment. The addition of Tat in *trans* provided sufficient viral RNA genomes to be packaged within

the particles, and human codon-optimization for *gag-pol* genes contributed to increase *gag* and *pol* RNA transcripts.

It has been demonstrated that Rev function can be replaced using codon optimized Gag-Pol (SYNGP) instead of wt Gag Pol (NLGP), but the CLTR did not replace Tat function yet, which still needs Tat to generate infectious viral progeny that is not produced in its absence (Figure 3-7 and 3-9). Published evidence suggests that Tat is not only involved in viral transcription, possessing additional functions such as influencing HIV-1 RNA splicing, capping or reverse transcription (Jablonski et al. 2010; Chiu et al. 2002; Charnay et al. 2009). Replacing Tat function is a difficult aim due to its diverse contributions to viral replication. (Das et al. 2011).



**Figure 3-12.** (A) Diagram of the SYNGP clones employed in this experiment. pCSGW, pCMV\_VSV-G and pDEST40-tat were also included during transfection into HEK-293T cells. VLPs were purified from the supernatant 48 h.p.t. by sucrose cushion centrifugation and inoculated into PM1 cells to determine the viral titre. (B) Flow cytometric analysis of PM1 cells infected with SYNGP clones 48 h.p.i. to determine the viral titre/infectivity by measuring GFP expression. The SSC-HLin (side scatter) determined the size and granularity of the cell population selected for this experiment. The GRN-HLin axis determined the intensity of GFP expression in PM1 infected cells. The DNA transfected amounts were 1µg of CSGW vector, 1µg of pCMV\_NLGP/pSYNGP, 500ng of DEST40\_tat and 100ng pCMV\_VSV-G.

### 3.2 Discussion chapter 3

In order to free HIV-1 from Tat-dependence, we needed to generate Tat-independent LTRs. The aim of this project was to generate an HIV-1 clone able to replicate within murine cells, and the MLTR was a good candidate to pursue because MLV is adapted to the murine host. In summary, the CLTR and MLTR promoters should be able to enhance viral gene expression in the absence of Tat, leading to the generation of sufficient viral RNA levels to express viral proteins and genomic RNA necessary to complete the viral replication cycle.

The genome length increase of the CSGW\_Flexi clones including the CLTR or MLTR should not affect viral genome packaging (Figure 3-4, panel A), but could introduce novel RNA secondary structures impairing proper folding of the genome and hence packaging into the particles, or creating more unstable viral RNAs (Kumar et al. 2001). In addition, previous studies have demonstrated that these promoters reach high levels of infection in human cells (such as pCVR1\_NLGP), suggesting that the reduced viral fitness found in CSGW\_Flexi clones is likely related to something else (Hill et al. 2001). The 5' MLTR\_Flexi clone did not drive transcription as efficiently as the 5' CLTR\_Flexi clone, suggesting that the MLTR could require further modification/adaptation to reach the same transcription levels as the 5' CLTR or 5' LTR (Figure 3-6, panel B, lane 4). Reverse transcription could be also affected in the Flexi\_3MLTR clone as well, and when both 5' and 3' MLTRs were encoded within the same vector (5MLTR\_Flexi\_3MLTR), it was severely attenuated ~10-fold (Figure 3-6, panel B, lane 11). Restoring the 118 bp fragment confirmed the potential of the CLTR and MLTR as possible candidates to replace the HIV-1 LTR, which is Tat dependent (Figure 3-6).

Despite the existence of reported studies that have evaluated the ability of enhancer elements to replace long terminal repeats in lentiviral vectors efficiently (Lotti et al. 2002), the transcription driven from CLTR and MLTR had effects on titre, infectivity, and maybe genomic stability. We also observed a reduction in the viral titre when we placed a 3' CLTR/MLTR into CSGW\_Flexi, suggesting a defect during reverse transcription in the target cells (Powell et al. 2015). We could have performed reverse transcription assays to quantify the reverse transcription levels in our CSGW\_Flexi and SYNGP clones and investigate the RNA species within transfected cells to determine viral transcription more accurately. Unfortunately, these experiments will consume much time and we decided to

continue our investigations with the CLTR and MLTR to achieve Tat Independence in human cells.

However, the ability of the MLTR to drive transcription and reverse transcription could be restricted because unknown murine co-factors could be absent in human cells (Studamire et al. 2010). The RNA levels in the absence of HIV-1 regulatory proteins contributed negatively to gene expression (Karn et al. 2012). The mechanism of function of these proteins is related to transcription elongation control (Tat) and RNA stability and export (Rev). Tat and Rev are essential to generate HIV-1 infectious particles *in vitro*, remaining important targets for drug discovery (Balachandran et al. 2017). In summary, HIV-1 replication fails in the absence of both regulatory proteins, which are unique for HIV.

The HIV Tat protein is the main regulator of viral gene expression through control of elongation by RNA polymerase II, and in its absence, the polymerase engages poorly and disengages from the DNA template prematurely, severely affecting viral transcription (Karn 1999). The reduction in the viral titres derived from CSGW\_Flexi clones persisted even in the presence of Tat and Rev proteins and when the CLTR and MLTR were located at the 5' end and 3' end simultaneously (Figure 3-6, panel B), suggesting that the lower titres were not related to the function of the regulatory proteins, and are likely due to defects related to viral gene expression or the appearance of aberrant RNA species originating from the codon-optimization of *gag* and *pol* genes (Knoepfel et al. 2013). Even low amounts of Gag processing intermediates can display a trans-dominant effect on HIV particle maturation (Müller et al. 2009), which are more abundant in CSGW\_Flexi clones encoding chimeric LTRs at the 3' end (Figure 3-6, panel B, lanes 5-11). The VLPs generated were almost equal for all CSGW\_Flexi clones, revealing that the 118 bp sequence was involved somehow in viral gene expression, but not in Gag maturation (Figure 3-6, panel D). In addition, the CLTR was not sufficient to generate efficient levels of transcription in a Tat independent manner, suggesting that the CLTR still needs Tat to enhance viral gene expression to complete the viral life cycle (Figure 3-12). We know already that the CLTR is able to generate sufficient transcription levels when Tat is provided in *cis* by NLGP (Figure 3-4 and Figure 3-6), but the 5CLTR\_Ψ \_SYNGP\_IRES-GFP and 5CLTR\_Ψ \_SYNGP\_IRES-GFP\_3CLTR clones lack Tat in *cis* and they are not infectious, suggesting the Tat dependence of the CLTR.

Secondary structure studies have determined that the HIV-1 packaging signal contains several RNA loops (SL1 to SL4) involved in packaging the viral genome into the particles by Gag interactions that could be altered in the 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_3CLTR clone (Greatorex et al. 2002). The codon-optimization of *gag-pol* genes contributes to the generation of more RNA species within human cells, but maybe the modifications made to reach this aim introduced novel RNA loops into the viral genome that could impair their packaging into the nascent particles. The compactly folded tertiary structure of viral RNA is essential for the viability of the virus at distinct steps along the replication cycle, such as reverse transcription, viral transcription or packaging (Berkhout et al. 2000). The RNA secondary structures are essential for the packaging of the viral genome within the viral particles. This aspect suggests that the absence of infectivity could be related to the disruption in the RNA folding in the viral genome introduced by codon optimization of *gag-pol* genes, or the higher number of CpGs in SYNGP compared to NLGP (226 to 23 respectively), which could impair viral replication kinetics (Atkinson et al. 2014; Gaunt et al. 2016). Our results suggest that RNA secondary structures in the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone could be affecting the packaging of the viral genome into the nascent particles or affecting RNA stability along the viral life cycle (Watts et al. 2009; Sükösd et al. 2015). However, further investigations need to be done to contrast these hypothesis.

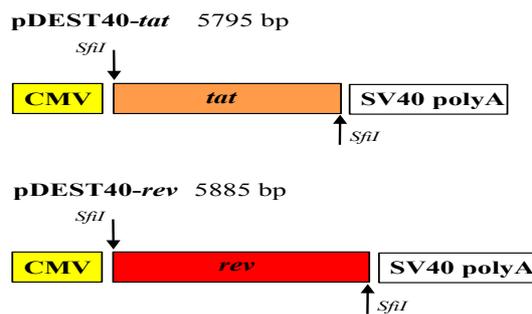
Finally, the absence of Rev could be affecting full length viral RNA transport towards the cytoplasm, which could remain within the nucleus. Despite the codon-optimized Gag being efficiently expressed and assembled into nascent viral particles, full-length viral RNAs need Rev and co-factors such as CRM1/exportin 1 to be exported to the cytoplasm and be packaged into the nascent viral particles (Taniguchi et al. 2014). The RNA stability of full-length viral and single spliced RNAs in the absence of Rev also could contribute to their degradation (Blissenbach et al. 2010).

# Chapter 4

## 4.1 Results

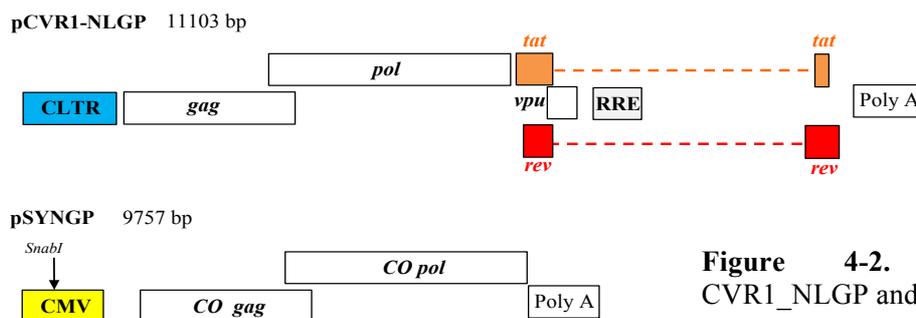
### 4.1.1 HIV-1 codon-optimized particles containing CSGW\_Flexi clones were infectious in human cells when Tat, Rev or both were provided in trans

Following our previous results where we achieved infectious viral particles from 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR construct when CSGW and DEST40-*tat* were supplied in *trans* (Figure 3-12), we wanted to determine the potential of the CSGW\_Flexi clones to generate infectious particles providing Tat and Rev proteins in *trans* during transfection. We transfected the CSGW\_Flexi clones together with pDEST40-*tat*, pDEST40-*rev*, pCMV\_VSV-G and pSYNGP instead of pCVR1\_NLGP. We created the vector DEST40-*rev* cloning the HIV-1 *rev* exon 1 and 2 together by PCR, removing the introns following the same overlap extension PCR strategy that we used to generate the DEST40-*tat* vector (Figure 4-1).



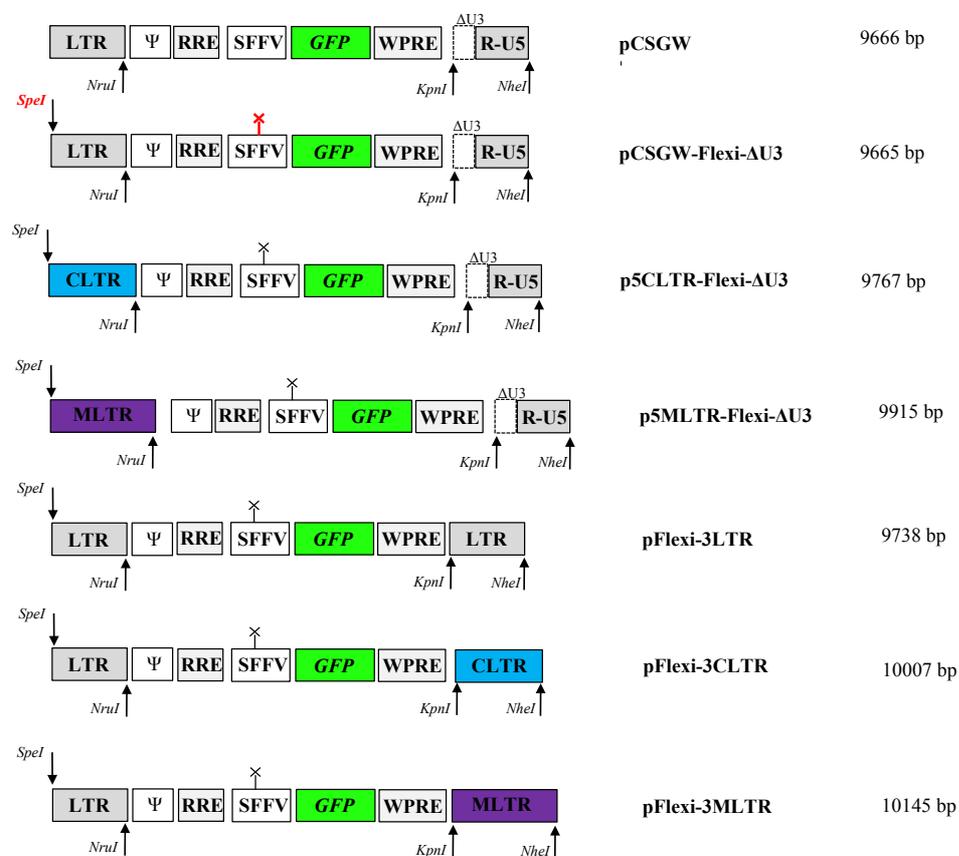
**Figure 4-1.** Diagram of DEST40-*tat* and DEST40-*rev* expression cassettes. We used the *SfiI* restriction sites to introduce *tat* and *rev* genes lacking introns.

The CMV\_NLGP vector encodes HIV-1 *tat* and *rev* genes, and we expected to determine if both proteins supplied in *trans* could rescue viral infectivity of CSGW\_Flexi clones cotransfected with SYNGP (Figure 3-4, panel A). Importantly, the SYNGP vector lacks Tat and Rev proteins, and encodes human codon-optimized HIV-1 *gag-pol* genes (Figure 4-2).



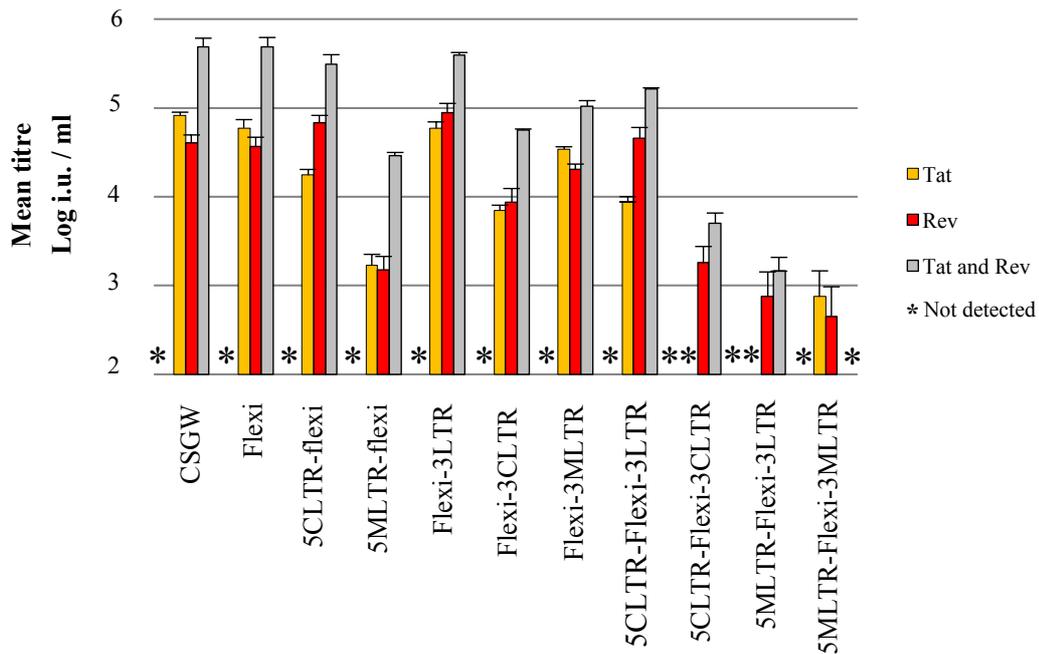
**Figure 4-2.** Diagram of CVR1\_NLGP and SYNGP vectors

We transfected the CSGW\_Flexi clones from the experiment shown in Figure 4-3 together with pSYNGP, pDEST40\_tat, pDEST40\_rev and pCMV\_VSV-G into human HEK-293T cells. Including Tat and Rev proteins in this experiment would determine if they were able to generate infectious viral progeny in a Tat or Rev independent manner in human cells. We harvested supernatant from the transfected cells 48 hours post-transfection and inoculated the produced virions onto human AA-2 lymphocyte cells, which are particularly permissive to HIV-1 replication, to measure GFP expression and determine the viral infectivity/titre (Figure 4-4).



**Figure 4-3.** Diagram of CSGW\_Flexi clones employed in the following experiments. The ‘X’ represents a deletion of a *NheI* site, and the ‘arrows’ indicate restriction sites employed to modify the 5’ and 3’ LTRs. We can substitute the 5’ LTR using *SpeI* and *NruI* restriction enzymes, and also substitute the 3’ LTR using *KpnI* and *NheI* enzymes.

Either Tat or Rev in *trans* supported the generation of infectious progeny from nearly all of the CSGW\_Flexi vector-genomes (Figure 4-3). The exceptions were the 5CLTR\_Flexi\_3CLTR and 5MLTR\_Flexi\_3LTR clones, which were not infectious when only Tat was provided in *trans* (Figure 4-4, lanes 9 and 10), with respect to these observations it is important to note that these were amongst the most attenuated vector systems and reductions in titre of a similar magnitude to those obtained in other clones could reduce infectivity to undetectable levels. Importantly however, no infection was detected in the absence of both proteins for any vector genome (Figure 4-4). Interestingly, the 5MLTR\_Flexi titre was ~100-fold lower (when Tat or Rev were supplied in *trans*) and ~10-fold lower when both Tat and Rev were supplied in *trans* than the 5CLTR\_Flexi and CSGW titres (Figure 4-4, lane 4). This pattern was also observed when a competent LTR was located at the 3' end in 5MLTR\_Flexi\_3LTR clone (Figure 4-4, lane 10), but this titre reduction was not observed for the 5CLTR\_Flexi\_3LTR clone (Figure 4-4, lane 8).



**Figure 4-4.** Mean titre of CSGW\_Flexi clones in AA-2 cells. The CSGW\_Flexi clones were transfected into HEK-293T with pCVR1\_NLGP (HIV *gag* and *pol*) and pCMV\_VSV-G (envelope), and HIV regulatory proteins Tat and Rev were added in *trans*. The VLPs were collected and purified from the supernatant by centrifugation of sucrose cushions 48 h.p.t. and inoculated into AA-2 cells to determine the viral titre of each CSGW\_Flexi clone. We transfected the same amount of DNA as previous experiments: [1µg of pCSGW: 1µg of pCVR1\_NLGP:: 100ng pCMV\_VSV-G]

Indeed, The CLTR attenuated the titre most when inserted at the 3' end whereas the MLTR caused the biggest decrease in titre when inserted at the 5' end. The 5CLTR\_Flexi titre was ~10-fold higher than the 5MLTR\_Flexi (Figure 4-4, panel A, lanes 3 and 4), but the Flexi\_3CLTR titre was still reduced ~1 log compared to the Flexi\_3MLTR (Figure 4-4, panel A, lanes 6 and 7). However, the Flexi\_3LTR and Flexi\_3MLTR titres were very similar (~ 5-fold difference Figure 4-4, panel A, lanes 5 and 7).

#### **4.1.2 CLTR could be better at driving transcription whereas the MLTR efficiently supports reverse transcription**

We wanted to determine the ability of CSGW\_Flexi clones to produce VLPs and also if adding Tat and Rev during transfection could rescue viral infectivity. Moreover, we repeated the same experiment shown in figure 4-4 (CSGW\_Flexi clones, pSYNGP, pCMV\_VSV-G, pDEST40\_tat and pDEST40\_rev) and performed western blot analysis for cell lysates (transfected HEK-293T) and VLPs purified from the supernatant 48 hours post-transfection by sucrose cushion centrifugation (produced in transfected HEK-293T).

The CLTR and MLTR promoters were able to generate infectious particles in most CSGW\_Flexi clones when only Tat was provided in *trans* (Figure 4-5, panel A). However, this was not the case for 5CLTR\_Flexi\_3CLTR and 5MLTR\_Flexi\_3LTR clones, which were not infectious under these conditions (Figure 4-5, panel A, lanes 9 and 10). As mentioned above, these vectors had such a low titre in the presence of Tat and Rev that a reduction in viral titre of 10-20-fold would reduce the titre to below the level of detection. Disappointingly, although the CLTR should be able to drive transcription in the absence of Tat, all the CSGW\_Flexi clones possessed a higher titre in the presence of Tat (when added in *trans*) (Figure 4-5, panel A).

The western blot analysis (using mouse anti-p24 antibody) for CSGW\_Flexi clones when only Tat was supplied in *trans* (Figure 4-5, panel A) shown no differences in Gag/CA expression from transfected HEK-293T, and Gag intermediates were generated for all CSGW\_Flexi clones (Figure 4-5, bottom panel A). The production of VLPs was similar for all CSGW\_Flexi clones, even for the 5CLTR\_Flexi\_3CLTR and 5MLTR\_Flexi\_3LTR

clones that were not infectious (Figure 4-5, bottom panel A). This was not unexpected as pSYNGP lacks elements that would be suspected to be influenced by Tat and Rev (Figure 4-2).

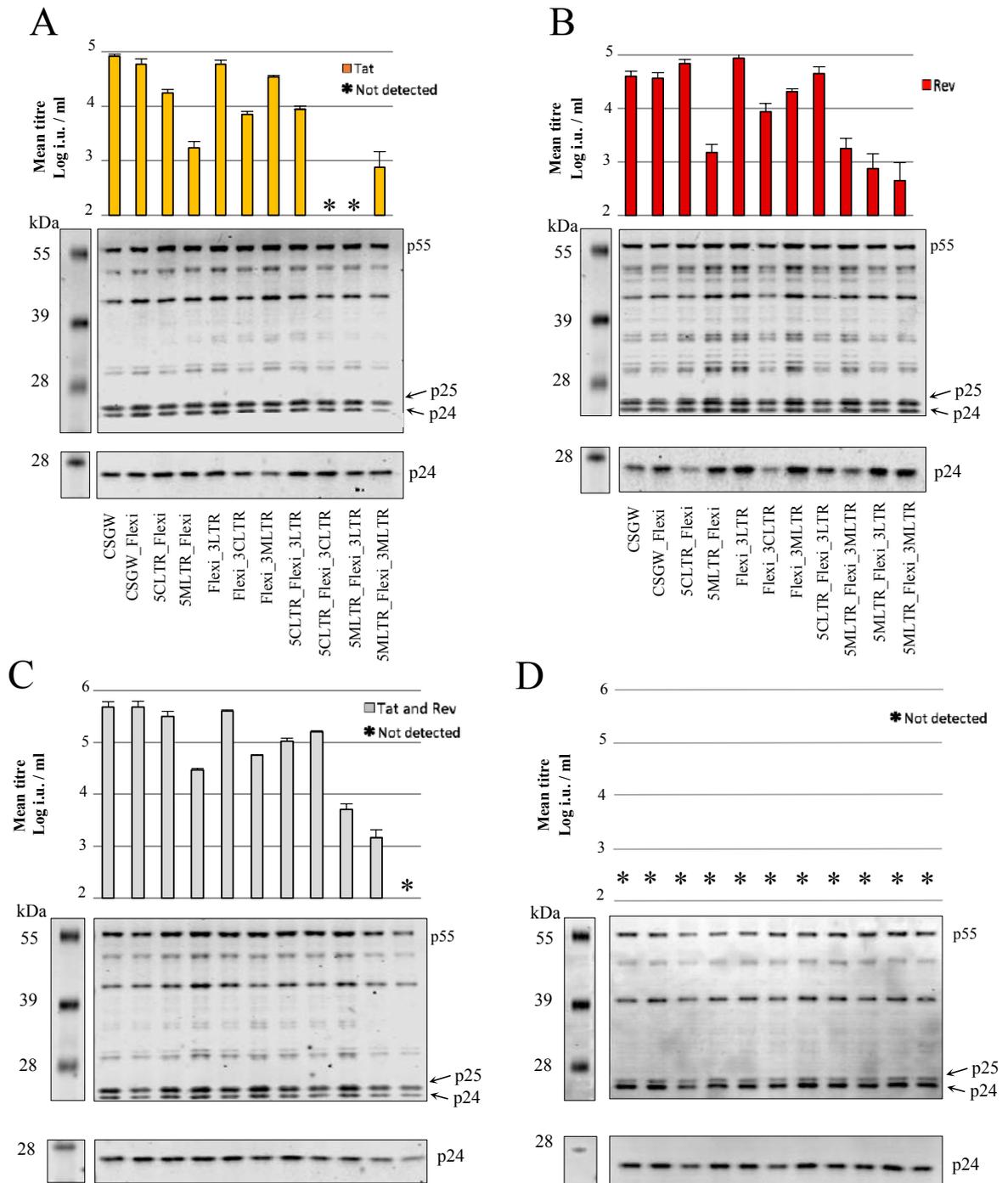
Somewhat surprisingly, all the CSGW\_Flexi clones were infectious when only Rev was added to the transfection (Figure 4-5, panel B). The 5MLTR\_Flexi titre was reduced ~100-fold compared to the 5CLTR\_Flexi (Figure 4-5, panel B, lanes 3 and 4). However, the modifications at the 3' end in Flexi\_3CLTR and Flexi\_3MLTR clones resulted in a ~10-fold titre reduction compared to CSGW and Flexi titres when only Rev was added in *trans* (Figure 4-5, panel B, lanes 6 and 7). The 5CLTR\_Flexi\_3LTR, 5CLTR\_Flexi\_3CLTR, 5MLTR\_Flexi\_3LTR and 5MLTR\_Flexi\_3MLTR titres were ~100-fold reduced when only Rev was added in *trans* compared to CSGW and CSGW\_Flexi (Figure 4-5, panel B, lanes 9-11). Unsurprisingly, western blot analysis did not show any differences regarding Gag/CA expression in transfected HEK-293T cells, and VLPs amounts collected from the supernatant were also similar between all CSGW\_Flexi clones generated when only Rev was added in the transfection (Figure 4-5, bottom panel B). The observation that all the vector genomes (CSGW\_Flexi clones) were expressed and incorporated into nascent virions in the absence of Tat was somehow unexpected as without Tat the transcriptional activity of the LTR is predicted to be low.

Finally, when both HIV regulatory proteins were supplied during transfection the CSGW\_Flexi clones were infectious (Figure 4-5, panel C). However, the 5MLTR\_Flexi\_3MLTR was not infectious (Figure 4-5, panel C, lane 11). Importantly the infectivity of this latter vector genome examined under all Tat and Rev permutations was quite close to background (using flow cytometry). The 5CLTR\_Flexi\_3CLTR and 5MLTR\_Flexi\_3LTR titres were ~10 fold and ~100 fold reduced respectively compared to wild-type CSGW and CSGW\_Flexi (Figure 4-5, panel C, lanes 10 and 11).

The 5MLTR\_Flexi titre was ~10 fold reduced compared to 5CLTR\_Flexi, following the same pattern previously observed when only Tat or only Rev were transfected (Figure 4-5, panel C, lane 3 and 4). The western blot analysis under these conditions (Tat and Rev together) revealed that Gag and CA expression was similar for all CSGW\_Flexi clones in transfected HEK-293T cells, and the VLPs production was very similar as well (Figure 4-5, bottom panel C). Thus, no differences were observed in VLP production between

CSGW\_Flexi clones when Tat and Rev were not transfected and none of them were infectious (Figure 4-5, panel D).

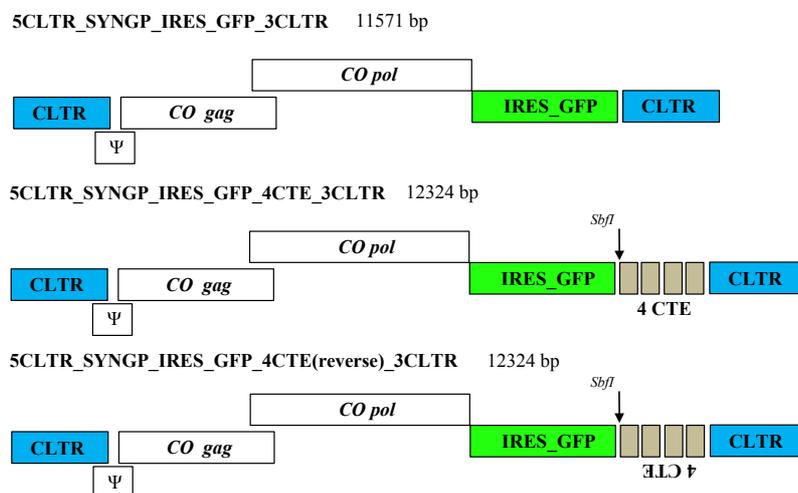
In summary, the HIV regulatory proteins Tat and Rev were able to rescue the infectivity of most of CSGW\_Flexi clones when supplied during transfection. However, the presence or absence of both proteins did not affect the amount of VLPs in all conditions produced from SYNGP vector. Unfortunately, no infection was observed in the absence of Tat and Rev and further investigations need to be done to determine the infectivity problems in CSGW\_Flexi clones under these conditions. We could have performed RNA studies for all CSGW\_Flexi clones (from transfected HEK-293T cells) or investigate the reverse transcription levels to provide new insights. Moreover, we could observe high amounts of p25 capsid intermediate in all western blots from cell lysates (transfected HEK-293T). The high expression of p25 could be related to the codon-optimized *gag* and *pol* genes encoded in SYNGP because the p25 expression was normal when wild type *gag* and *pol* genes were transfected in previous experiments using pCVR1\_NLGP instead SYNGP (Figure 3-4, panel C). Perhaps the higher amounts of viral RNAs for *gag* and *pol* genes are contributing to generate more p25 intermediate compared to wild type HIV *gag-pol* genes.



**Figure 4-5.** Mean titre of CSGW\_Flexi clones obtained from GFP reporter expression from infected AA-2 cells. Western blot analysis (mouse anti-p24 antibody) from CSGW\_Flexi clones transfected in HEK-293T cells and VLPs collected from the supernatant and purified by sucrose cushion centrifugation 48 h.p.t. The CSGW\_Flexi clones were generated by transfection using pSYNGP, pCMV\_VSV-G and the regulatory HIV proteins Tat (pDEST40<sub>tat</sub>) and Rev (pDEST40<sub>rev</sub>). We transfected the same amount of DNA as previous experiments [ $1\mu\text{g}$  of CSGW\_Flexi clone:  $1\mu\text{g}$  of pSYNGP:  $100\text{ng}$  pCMV\_VSV-G:  $500\text{ng}$  DEST40<sub>tat</sub> and  $500\text{ng}$  DEST40<sub>rev</sub>]; the viral titre was quantified 48 h.p.i. in AA-2 cells. (A) Mean titre of CSGW\_Flexi clones when only Tat was added in *trans* during transfection and VLPs production from transfected HEK-293T cells 48 h.p.t. (B) Mean titre of CSGW\_Flexi clones when only Rev was added in *trans* and VLPs production from transfected HEK-293T cells 48 h.p.t (C) Mean titre of CSGW\_Flexi clones when Tat and Rev were added in *trans* and VLPs production from transfected HEK-293T cells 48 h.p.t (D) Mean titre of CSGW\_Flexi clones when none of HIV regulatory proteins were added in *trans* and VLPs production from transfected HEK-293T cells 48 h.p.t

### 4.1.3 The CLTR is Tat responsive

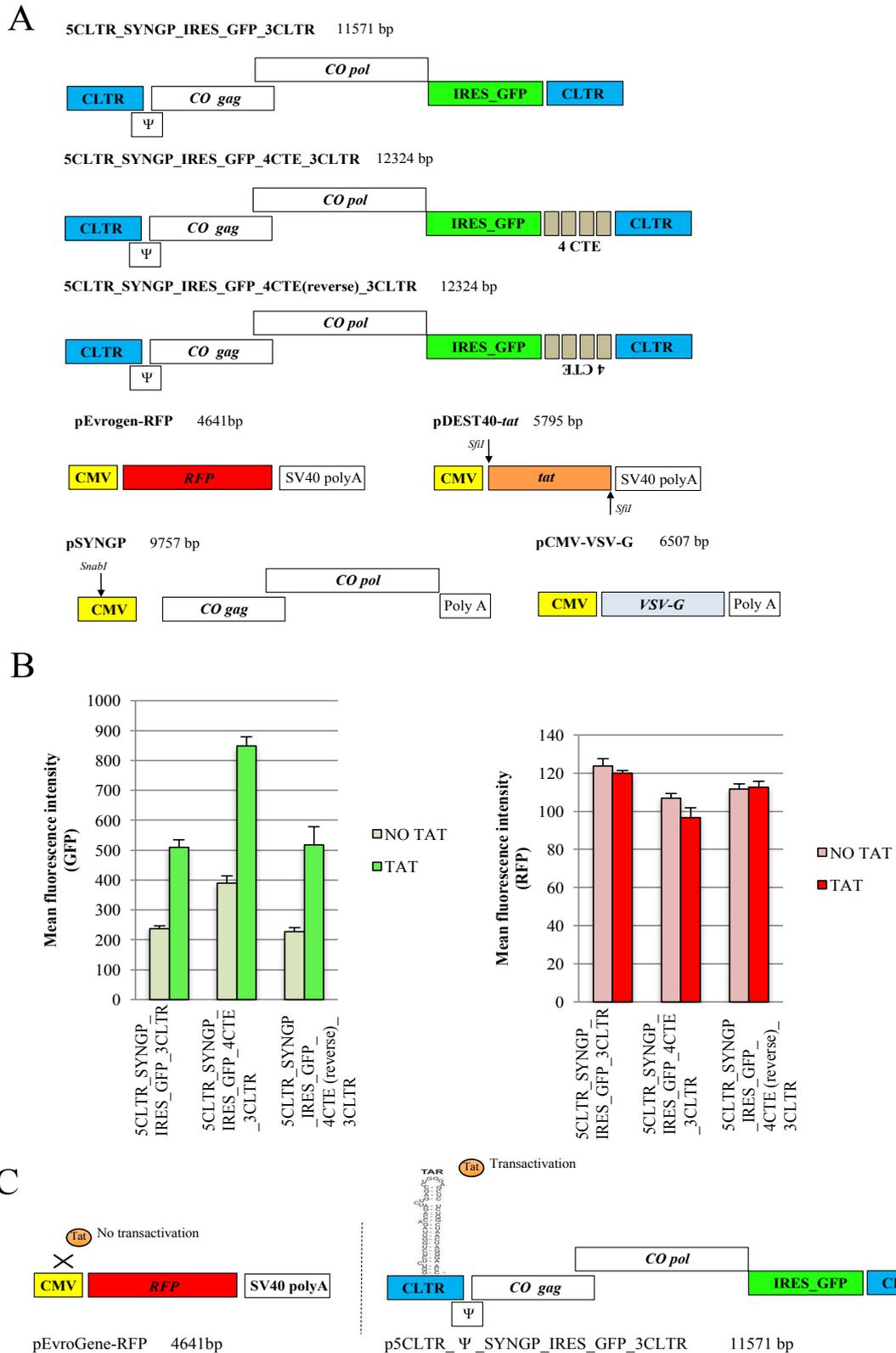
Since many all of the 5CLTR-containing vectors had higher titres when produced in the presence of Tat, we wanted to determine the Tat responsiveness of the CLTR. We investigated this using the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone by measuring the mean fluorescence intensity (MFI) from the GFP reporter gene (expressed under control of the CLTR) in transfected human HEK-293T cells. Comparing the MFI in the presence or absence of Tat would help us to determine the difference in transcriptional activity. Prior to attempting this experiment, we also cloned 4 copies of the constitutive transport element (CTE) preceding the 3'LTR, where they exert a similar function as Rev protein in the Mason-Pfizer monkey virus (MPMV) (Wodrich et al. 2000), to compensate for the lack of Rev function. The addition of 4CTEs will enhance viral RNA transport and contribute to its stability. We cloned 4CTEs into 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone at position 6603 by cleavage with *SbfI*, preceding the 3'CLTR, generating the vector called SYNGP\_4CTE (Figure 4-6). During the cloning process we also generated a 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone encoding 4CTEs in the opposite orientation (SYNGP\_4CTE(reverse)) (Figure 4-6).



**Figure 4-6.** Diagram of SYNGP\_4CTE clones. The orientation of the CTEs is different in 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_4CTE\_3CLTR and 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_4CTE(reverse)\_3CLTR.

These new SYNGP\_4CTE clones were transfected together with pDEST40\_ *tat* and pCMV\_VSV-G in HEK-293T cells. We also transfected the EvroGen TagRFP plasmid as a transfection control. Then, we determined GFP expression 48 hours post-transfection by flow cytometry (Figure 4-7, panel B). By using this experiment set-up we were able to select only the transfected cells expressing RFP and GFP, and submit them to FACS analysis to compare their MFI in the presence and absence of Tat (Figure 4-7, panel B). Crucially, TagRFP levels were very similar for all of the clones tested (5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_4CTE\_3CLTR and 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_4CTE(reverse)\_3CLTR clones and also the control 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_CLTR) in the presence and absence of Tat (Figure 4-7, panel B (right)). This indicated that Tat did not globally influence transcription of transfected DNA. In contrast, the GFP expression was higher when Tat was provided in *trans*, demonstrating the Tat responsiveness of the CLTR encoded in SYNGP\_4CTE clones (Figure 4-7, panel B). In addition, the highest MFI was observed in the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_4CTE\_3CLTR clone, which encodes 4 copies of the CTE in the correct orientation (Figure 4-7, panel B, lane 2). However, the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_4CTE(reverse)\_3CLTR clone MFI was very similar compared to the control 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR even in the presence of Tat (Figure 4-7, panel B, lane 3).

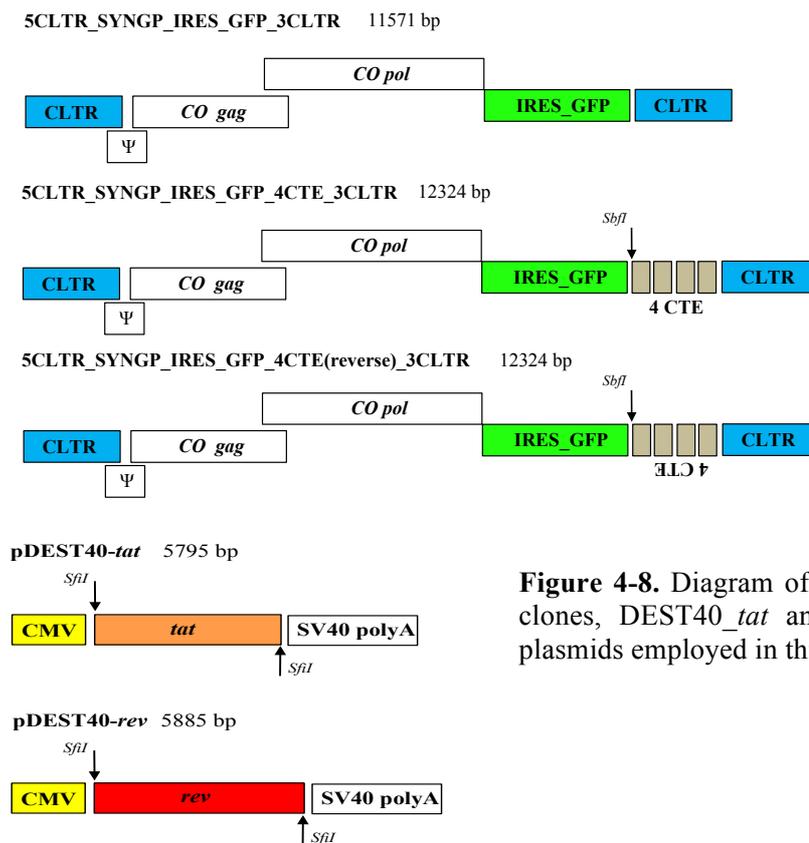
The observation that the CMV promoter encoded in EvroGene TagRFP vector did not show Tat responsiveness (Figure 4-7, panel C) was expected because the stimulation of viral transcription is carried out by the interaction of Tat with the Trans-acting response element (TAR) located in the R region from HIV-1, which is absent in pEvroGene TagRFP (Roy et al. 1990).



**Figure 4-7.** Mean fluorescence intensity (48 h.p.t.) in transfected HEK-293T cells with SYNGP clones and DEST40-*tat*. The envelope was provided in *trans* with pCMV-VSV-G as usual. We transfected the following amount of DNA: [1 $\mu$ g of each SYNGP clone : 1 $\mu$ g of pCSGW: 100ng pCMV\_VSV-G: 500ng DEST40-*tat*]. (A) Diagram of SYNGP clones (encoding CLTRs, the HIV-1 packaging signal, codon-optimized HIV-1 *gag-pol* genes, an IRES-GFP cassette and 4 copies of the CTE), the pCMV\_VSV-G and pEvrogen-RFP (transfection control) vectors. (B) Mean fluorescence intensity (MFI) in transfected HEK-293T cells; RFP fluorescence was used as transfection control. The error bars show the standard deviation for each titre (triplicates) (C) Diagram of Tat responsiveness in the CLTR encoded in 5CLTR\_Ψ SYNGP IRES GFP 3CLTR.

#### 4.1.4 5CLTR\_Ψ\_SYNGP\_IRES-GFP\_4CTE\_3CLTR clone was not infectious in the absence of Tat and Rev. However, both proteins rescued viral infectivity

After confirming the Tat responsiveness of the CLTR encoded in SYNGP\_4CTE clones (Figure 4-7), we wanted to investigate if they were able to package a competent genome (CSGW) provided in *trans* and become infectious. In addition, we supplied Tat and Rev proteins during transfection to determine if they could rescue the viral infectivity (in the absence or presence of CSGW). The 4CTEs from SYNGP\_4CTE clones would replace Rev function, so we were expecting to achieve Rev independent infectious particles. We co-transfected pCSGW, SYNGP\_4CTE clones and pCMV\_VSV-G (VSV envelope) together with Tat, Rev or both into HEK-293T cells to produce single cycle of replication particles as usual. The supernatant was harvested at 48 hours post transfection and used to inoculate AA-2 cells to determine the viral titre (Figure 4-10).



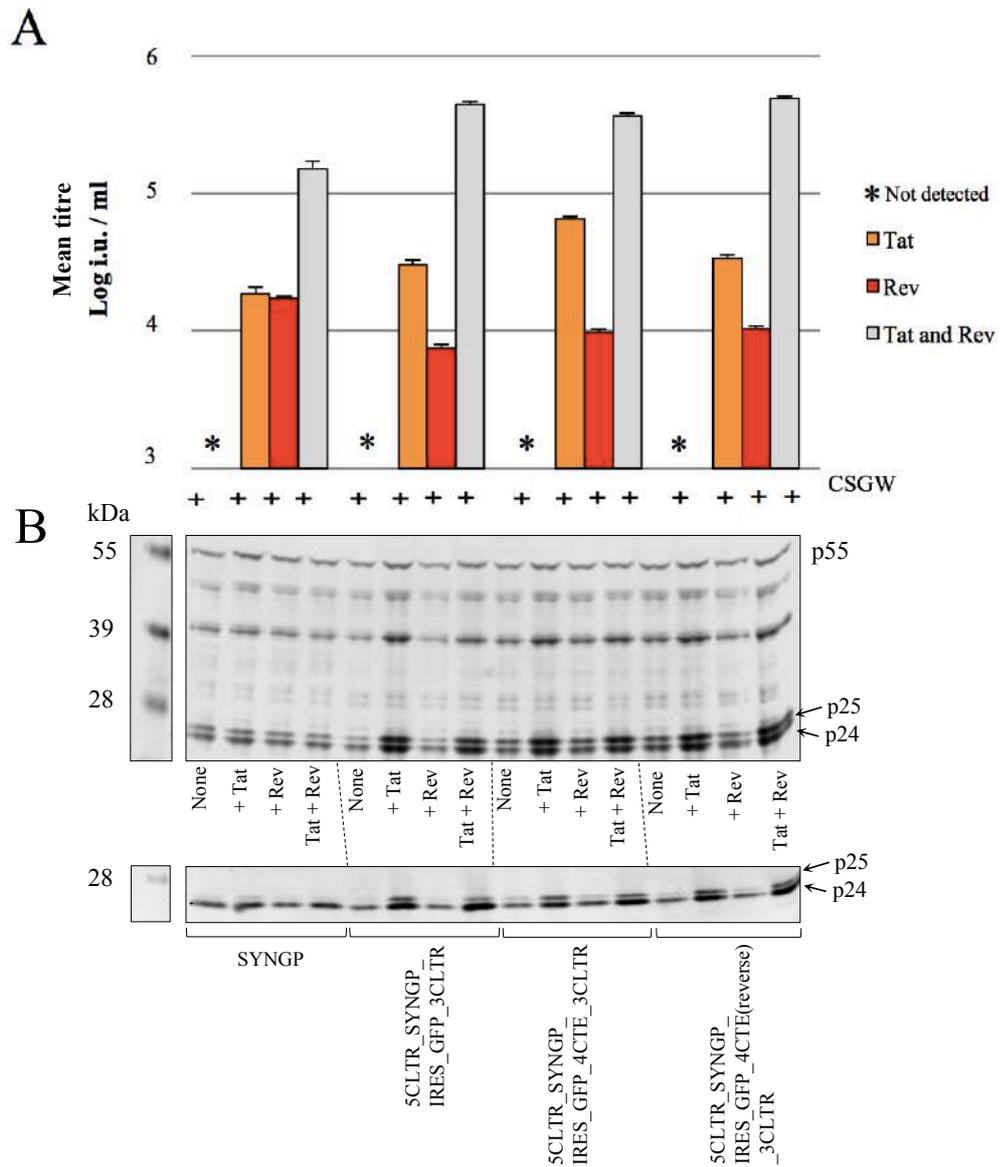
**Figure 4-8.** Diagram of SYNGP\_4CTE clones, DEST40\_*tat* and DEST40\_*rev* plasmids employed in this experiment.

The addition of Tat, Rev or both in *trans* rescued infectivity in SYNGP\_4CTE clones, but no infection was detected in the target cells in their absence even when CSGW was provided (Figure 4-10, panel A). The mean titre for SYNGP\_4CTE clones was similar

compared to the control 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR (which lacks 4CTEs in *cis*), when Tat and Rev were supplied together in *trans*. Viral infectivity was recovered when Tat or Rev were provided in *trans*, generating infectious particles in a Tat/Rev independent manner (Figure 4-10, panel A). However, the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR and SYNGP\_4CTE titres were ~5-fold higher when only Tat was supplied compared when only Rev was supplied (Figure 4-10, panel A, orange and red columns). The 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR and SYNGP\_4CTE clones titres were ~5-fold higher compared when only Tat or only Rev were provided during transfection (Figure 4-10, panel A, grey columns). The 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR and SYNGP\_4CTE clones had similar titres under each condition (Figure 4-10, panel A, lanes 3-5). However, CA expression levels in transfected HEK-293T cells were higher in SYNGP\_4CTE clones compared to 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR transfection, but Gag expression was very similar in all cases under each condition (Figure 4-10, upper panel B). The amount of VLPs produced from 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR and SYNGP\_4CTE clones correlated to more capsid expression in transfected HEK-293T when only Tat or both proteins were added in *trans* (Figure 4-10, panel B, bottom lane). The CA P25 band was particularly prominent in the purified VLPs (Figure 4-10, panel B, bottom lane), the precise reason for this is unclear.

No infection was observed when CSGW was not provided in *trans*, despite the addition of Tat and Rev proteins (Figure 4-9). However, the SYNGP\_4CTE clones were able to package CSGW when provided in *trans* in the presence of Tat and Rev, but they failed to be infectious in their absence. Despite the addition of 4CTEs from SYNGP\_4CTE clones did not rescue the viral infectivity in the absence of CSGW and the regulatory proteins, suggesting that Tat and Rev were still essential for these viruses. When both regulatory proteins were added in *trans* together with CSGW, the viral titre reached about  $10^{5.5}$  log i.u./ml for SYNGP\_4CTE clones.





**Figure 4-10.** GFP reporter expression of SYNGP clones in AA-2 cells and WB analysis for Gag and p24 (using mouse anti-p24 antibody). The SYNGP clones were transfected in HEK-293T cells [ $1\mu\text{g}$  of each SYNGP clone:  $1\mu\text{g}$  of pCSGW:  $100\text{ng}$  pCMV\_VSV-G:  $500\text{ng}$  DEST40\_tat:  $500\text{ng}$  DEST40\_rev]. (A) Mean viral titre of SYNGP clones in AA-2 cells 48 h.p.i. The viral titre was determined by flow cytometry measuring GFP expression in PM1 cells and the error bars show the standard deviation for each titre (triplicates). (B) Western blot analysis for Gag/p24 expression in transfected HEK-293T cells 48 h.p.t. (D) VLPs collected and purified by sucrose cushion centrifugation from the supernatant of transfected HEK-293Ts (48 h.p.t.).

## 4.2 Discussion chapter 4

The 5CLTR was more efficient at driving transcription than the 5MLTR, reaching higher levels of infection under all conditions (adding Tat, Rev or both) (Figure 4-4). This result suggested that the CLTR is still Tat dependent and fails to drive transcription efficiently in its absence. In order to determine the ability of the CLTR to drive transcription in a Tat independent manner we could remove the R region of the 5' end that encodes the TAR loop, and investigate the viral RNAs produced within the transfected cells. However, this will generate an aberrant virus unable to carry out reverse transcription and subsequently be unable to complete the viral replication cycle (Berkhout et al. 2001; Kleiman et al. 2002). This fact highlights the limitations of studying viral transcription in the absence of Tat protein, which could be also involved in another functions not related to viral transcription.

The viral titre was severely affected when the 5' and 3' ends in CSGW\_Flexi clones were modified CLTRs or MLTRs (Figure 4-4 and 4-5). The CLTR supported transcription when it was located in the 5' end, but other aspects, most likely reverse transcription were affected in the Flexi\_3CLTR clone compared to Flexi\_3LTR, Flexi\_3MLTR and the controls (CSGW and CSGW\_Flexi) when only Tat was added in *trans* (Figure 4-4, panel B, lane 6). Perhaps the CLTR requires more time to complete reverse transcription than the wild type HIV-1 LTR. In addition, the titre was decreased ~1 log when the CLTR or MLTR were located at the 3' end under each conditions (none, Tat, Rev or both) (Figure 4-4, lane 6 and 7). These results suggested that non-transcriptional aspects of the LTR such as reverse transcription could be affected in Flexi\_3CLTR, Flexi\_3MLTR clones. Because the CSGW\_Flexi\_3'LTR clone did not show this pattern and its titre was similar compared to CSGW and CSGW\_Flexi controls (Figure 4-4, lanes 5-7). This suggests that the CLTR and MLTR constructs support reverse transcription less efficiently than the wild type HIV-1 LTR, but are still able to generate infectious viruses. The MLTR chimeric promoter was designed for the murine host, and we only developed this experiment in human cells, perhaps explaining why the MLTR is not as efficient as the HIV-1 LTR is driving transcription and reverse transcription. Introducing the CLTR and MLTR probably altered RNA secondary structures in the wild type LTR, which could have affected reverse transcription in 5CLTR\_Flexi and 5MLTR\_Flexi clones (Figure 4-4). Our evidence demonstrated that viral RNA are able to express Gag and p24 successfully in transfected HEK-293T and generating VLPs.

However, the VLPs could lack the viral genome, so we could have performed RNA extractions from transfected cells to investigate the differences between RNA species for each CSGW\_Flexi and SYNGP clones. Thus, we could have also investigated the RNA content within the VLPs to detect if the viral genome is able to be packaged into nascent particles or perhaps can not interact with the packaging signal. Several experiments were performed to investigate the RNA content in VLPs by qPCR, but none of the results were conclusive and are not included in this thesis. Moreover, other investigations have determined the essential role of the HIV-1 RNA folding, encoding highly variable regions but also high conserved structures with unknown functions (Wilkinson et al. 2008; Westerhout et al. 2005).

Gag expression was similar for all conditions for CSGW\_Flexi vector genomes when Tat or Rev were transfected in *trans* (Figure 4-4, panels A, B, C and D). However, we obtained similar levels of Gag for CSGW\_Flexi clones when the regulatory proteins were not provided in *trans* (Figure 4-5, bottom panel D). The western blot analysis for Gag/CA expression demonstrated that the infectivity problems in CSGW\_Flexi clones were not related to Gag protein expression or its processivity within the transfected HEK-293T cells (Figure 4-4). VLPs produced from CSGW\_Flexi vector genomes and SYNGP vector were very similar under all conditions, even when Tat and Rev proteins were not provided in *trans* (Figure 4-4, panel D).

However, the presence of the CLTR or MLTR in the 5' end should be able to drive viral transcription in CSGW\_Flexi clones, as the enhancers in the CMV promoter are widely regarded to be relatively efficient at driving transcription. No evidence was observed to suggest issues with VLP formation for SYNGP\_4CTE clones, but they could lack the viral genome or could be aberrant in some other way (Lu et al. 2011). Moreover, we detected some p25 intermediate protein in western blot analysis of VLPs from all SYNGP clones (Figure 4-10, panel B). We could have investigated if the generated VLPs had a proper morphology because it is not possible to determine this feature by western blot analysis, so we could have prepared some purified VLPs for each CSGW\_Flexi and SYNGP clones to submit to electron microscopy and approach this hypothesis. Thus, to elucidate the likely problems with full-length RNA packaging/production we need to investigate the RNA production within transfected cells by RNA extraction and subsequently northern blot analysis.

Notably, when CSGW was provided *in trans*, Gag expression from SYNGP\_4CTE clones was capable of supporting a functional single round of infection when Tat or Rev were provided *in trans* together with CSGW vector (Figure 4-10). This result suggests that transcription levels and viral RNA export were not sufficient to complete the first cycle of replication in SYNGP\_4CTE clones in the absence of the regulatory proteins and CSGW vector (Figure 4-10, panel A, lanes 2-4). The SYNGP\_4CTE clones in the context of a viral genome can be expressed at appropriate levels and package a competent genome. Indeed adding Rev alone did not increase the abundance of nascent VLPs (Figure 4-10, panel B) suggesting that the Rev dependency of this system could be as much to do with the dependency of CSGW on Rev as a requirement for Rev to support the expression of Gag and subsequent VLP assembly. Perhaps the modifications introduced by introducing the CLTR and MLTR into CSGW\_Flexi clones are affecting the viral RNA secondary structure, which is required to interact with the Gag protein to package the viral genome into the nascent particles (Amarasinghe et al. 2001). Alternatively, the viral genome is generated within the transfected cell, but it could be unable to be packaged within the nascent viral particles (Mailler et al. 2016).

The addition of 4CTEs into 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR clone (SYNGP\_4CTE) could have contributed to the viral RNA export to the cytoplasm, and their orientation seems relevant somehow (Figure 4-7, panel B, lane 2 and 3). In addition, the mean fluorescence intensity observed in transfected HEK-293T cells with 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR and SYNGP\_4CTE clones was similar in SYNGP\_4CTE clone in the absence of Tat compared to SYNGP\_4CTE(reverse) clone with Tat *in trans* (Figure 4-7, panel B), suggesting that the absence of 4CTEs in *cis* could attenuate viral transcription somehow, perhaps because viral RNAs could remain inside the nucleus and not be exported to the cytoplasm (Cmarko et al. 2002). However, no infection was detected from VLPs generated from SYNGP\_4CTE clones in the absence of CSGW vector and the regulatory proteins (Figure 4-10, panel B), but the addition of CSGW, Tat and Rev rescued the viral infectivity of SYNGP\_4CTE viruses (Figure 4-10). These results suggest that maybe the 4CTEs are unable to interact with the viral RNAs and export them to the cytoplasm, and viral transcripts remain stuck inside the nucleus (Rizvi et al. 1997). To investigate if the viral RNAs could remain in the nucleus we could have done RNA extractions from transfected cells with SYNGP clones (with 4CTEs) and compare the amounts of full-length genomes produced (northern blot) or include a fluorescent dye

in viral RNAs to detect the amount of RNA that remain in the nucleus compared to the cytoplasm in transfected cells. The SYNGP\_4CTE clones still defective at some step following viral integration, despite the higher MFI detected in transfected HEK-293T cells shown in Figure 4-7, panel B. We could detect high levels of GFP expression after integration using SYNGP\_4CTE clones, but they are defective afterwards because they are unable to generate infectious viruses. Our evidence demonstrated that Gag is efficiently expressed in SYNGP\_4CTE clones, and capable producing VLPs, suggesting that viral transcription is not affected and the problems could be likely related to RNA stability/transport (Zhou et al. 2005).

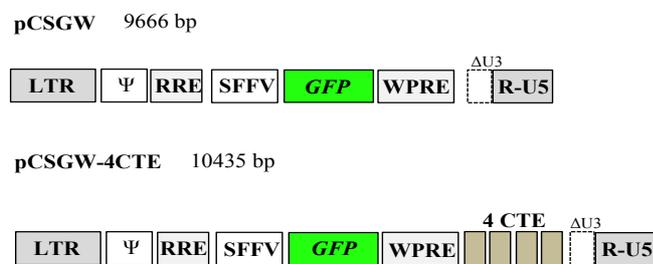
To really understand the major problem in SYNGP\_4CTE clones we will need to investigate each step along the viral life cycle, such as viral RNA production or the packaging of the viral genome into the nascent particles, which will take much intensive research not possible in this PhD due to time constraints. The aim of this project was to achieve a Tat and Rev independent HIV-1. At this point we had two possibilities: determine the causes that impair viral replication within human cells in SYNGP\_4CTE clones, or proceed with development of another single cycle of replication virus using CSGW vector as template instead of SYNGP.

## **Chapter 5**

## 5.1 Results

### 5.1.1 CSGW containing 4 copies of the Constitutive Transport Element (CTE) is able to generate infectious particles in a Tat-Rev independent manner in human cells

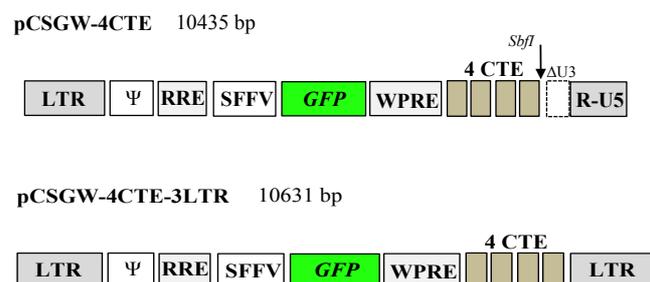
SYNGP\_4CTE clones failed to generate infectious viral progeny in a Tat and Rev independent manner, and both proteins were always required for the viability of these viruses (Figure 4-10). There are several studies showing the ability of constitutive transport elements encoded in lentiviral vectors that increase the viral titre, contributing to the export and stability of the viral RNA (Nasioulas et al. 1994). These studies demonstrate that the absence of infectivity from SYNGP\_4CTE could be related to viral RNA transport. We suspected that viral RNAs were not produced efficiently in the absence of Tat from SYNGP\_4CTE clones, so we transferred the 4CTEs into CSGW to investigate whether they could generate infectious Rev independent particles as the parental vector SYNGP did before (Figure 5-1). The 4CTEs were cloned into CSGW using the natural restriction site *SbfI* at position 9030, immediately after the GFP gene and preceding the 3'ΔU3 LTR (Figure 5-1).



**Figure 5-1.** Diagram of CSGW and CSGW encoding 4 copies of the CTE from Mason-Pfizer monkey virus (CSGW\_4CTE)

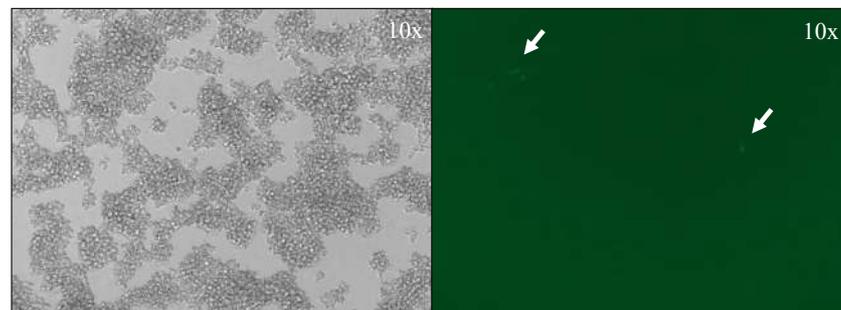
We expected to rescue Rev function in human cells by adding the 4CTEs into CSGW. The Tat and Rev ‘independence’ was achieved by CSGW\_4CTE construct, (which encodes a Tat dependent LTR) (Figure 5-2) when HEK-293Ts were transfected using pCSGW\_4CTE, pSYNGP and pCMV\_VSV-G. Following this promising result, we wanted to determine whether the CSGW\_4CTE clone was able to infect murine cells. In order to carry out this experiment we used murine macrophages RAW 264.7 as target cells, which will be infected with our pseudotyped CSGW\_4CTE virus (Figure 5-3).

Unfortunately, the CSGW\_4CTE clone was unable to generate the same level of infection in RAW cells compared to highly permissive human AA-2 cells, but some infected cells were detected (Figure 5-3). The CSGW\_4CTE construct was transfected together with pSYNGP, pCMV\_VSV-G, pDEST40\_ *tat* and pDEST40\_ *rev* in HEK-293T cells to produce viral particles as usual (Figure 5-4, panel A). Surprisingly, the CSGW\_4CTE construct generated infectious viruses in a Tat and Rev independent manner (Figure 5-4, panel B). The CSGW\_4CTE clone was also infectious when Tat, Rev or both together were supplied in *trans* (Figure 5-4, panel B). The CSGW\_4CTE titre in the absence of the HIV-1 regulatory proteins was similar compared to the titre when only Rev was added in *trans* (Figure 5-4, panel B, white and red columns). However, the mean titre from CSGW\_4CTE was slightly reduced for each condition compared to the control CSGW (Figure 5-4, panel B, lane 1 and 2). The CSGW\_4CTE titre did not increase when Rev was added in *trans* despite the presence of 4 copies of the CTE in *cis* (Figure 5-4, panel B). No differences were observed in Gag/CA expression in transfected HEK-293Ts, or in the VLPs collected from the supernatant between CSGW\_4CTE and CSGW\_4CTE\_3LTR (Figure 5-4, bottom panel B). We wanted to generate a replication competent entity by replacing the  $\Delta$ U3 region (located at the 3' end in CSGW\_4CTE vector) for a wild type U3 region. The 5'LTR needs to be functional after reverse transcription, and the wild type LTRs located at both ends of the vector genome CSGW\_4CTE provide us a more realistic model to work with at the CL2 lab. This modification could also produce a higher viral titre during the first cycle of replication (Figure 5-2).



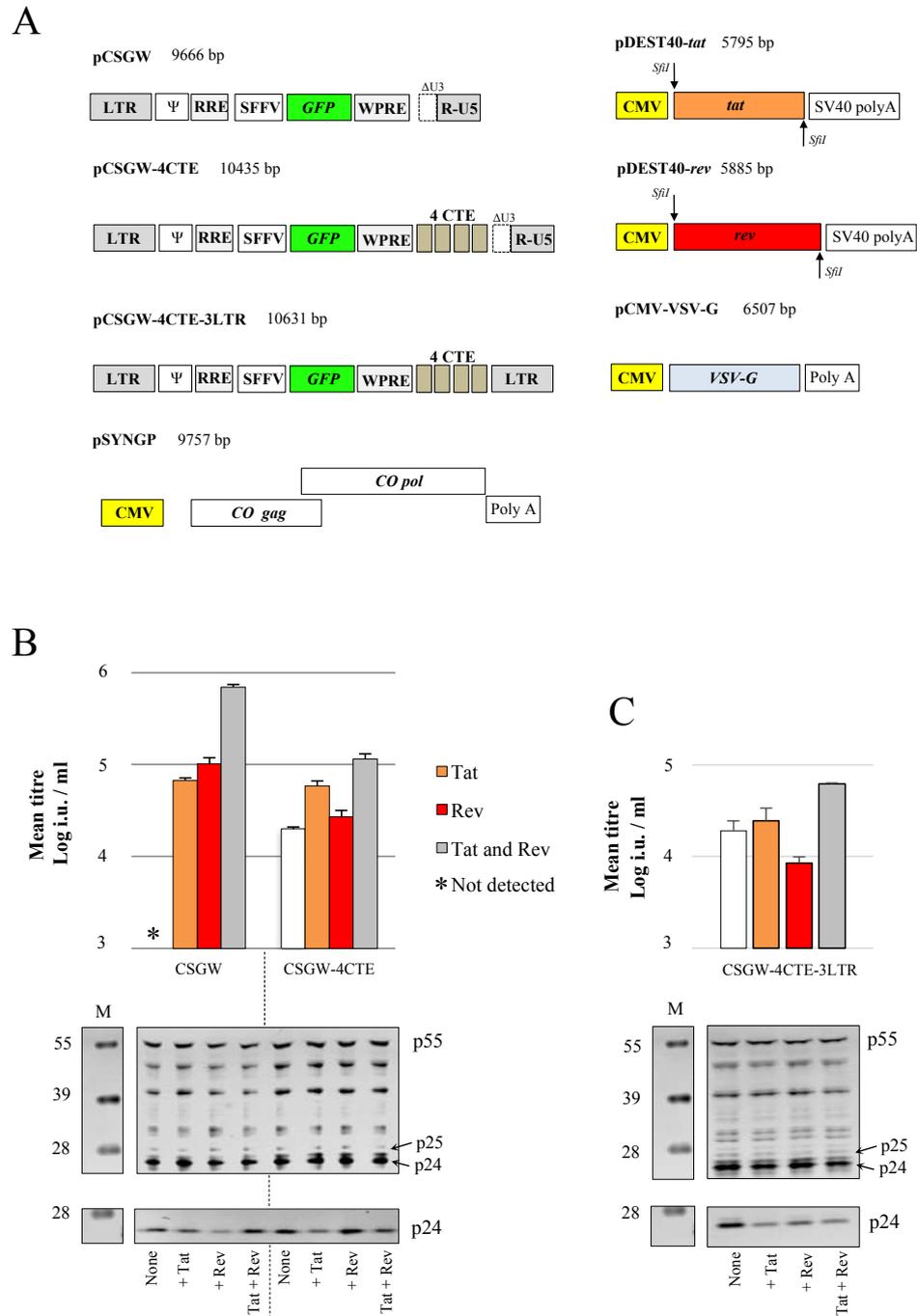
**Figure 5-2.** Diagram of CSGW\_4CTE and CSGW\_4CTE encoding a competent HIV-1 LTR (CSGW\_4CTE\_3LTR)

The addition of a competent LTR at the 3' end of CSGW\_4CTE did not increase viral titre under all conditions, resulting in a very similar titre compared to the CSGW\_4CTE containing a 3'ΔU3 LTR (CSGW\_4CTE)(Figure 5-4, panel C). The CSGW\_4CTE\_3LTR was also able to generate infectious viral particles in a Tat and Rev independent manner, similar to the CSGW\_4CTE (Figure 5-4, panel C). The addition of the HIV-1 regulatory proteins (in *trans*) did not increase Gag or CA expression, and Gag intermediates were observed for all conditions in CSGW\_4CTE and CSGW\_4CTE\_3LTR clones (Figure 5-4, panel B and C).



**Figure 5-3.** Murine RAW 264.7 cells infected with CSGW\_4CTE viruses produced in HEK-293Ts. Cells were incubated for 72 hours after the inoculation.

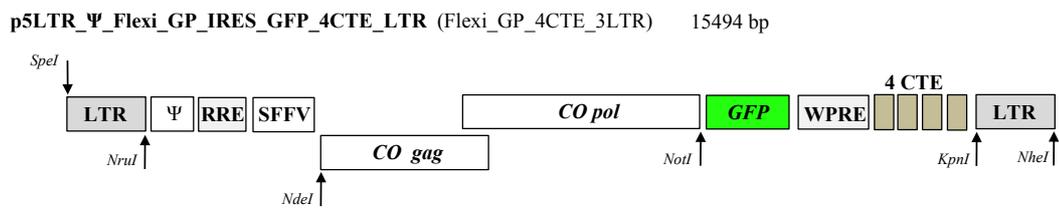
These results demonstrated the ability of the CTE to rescue the infectivity in CSGW\_4CTE and CSGW\_4CTE\_3LTR clones by replacing Rev function and supporting the viral RNA export and stability within human cells (Figure 5-4, panel B and C). We also demonstrated that pseudotyped HIV-1 particles generated using CSGW including 4 copies of the CTE in *cis* (CSGW\_4CTE) contributed to the generation of a Tat and Rev independent viruses (Figure 5-4, panel B). These results were the first evidence that supported the development of an HIV-1 Tat and Rev independent clone, which seems to be achievable despite the previous negative results obtained in this thesis. However, these experiments were performed in human cells only (HEK-293T and AA-2), and more investigation needs to be done prior to attempting to reproduce the same results in murine cells. Importantly, the lack of the HIV-1 regulatory proteins severely attenuates the ability of CSGW\_4CTE to replicate efficiently in their absence, and would have to be much fitter before we could attempt to reproduce the same result in murine cells.



**Figure 5-4.** Experiment to determine the ability of CSGW\_4CTE and CSGW\_4CTE\_3LTR to generate infectious VLPs in human cells by transfection. We also determined the viral titre infecting human AA-2 afterwards. The amount of transfected DNA was 1 $\mu$ g of CSGW\_4CTE/CSGW\_4CTE\_3LTR, 1 $\mu$ g of pSYNGP, 100ng of pCMV\_VSV-G, 500ng of DEST40\_tat and 500ng of DEST40\_rev. (A) pCSGW clones including 4 copies of the CTE. HIV-1 Codon-optimized *gag-pol* genes were provided by pSYNGP and the envelope by pCMV-VSV-G. (B) Mean titre of CSGW\_4CTE clones in AA-2 cells. The error bars show the standard deviation for each titre (triplicates). DEST40\_tat and DEST40\_rev were added *in trans* (orange and red bars respectively); Western blot for CSGW\_4CTE (using mouse anti-p24 antibody) from transfected HEK-293T and viral particle production (VLPs) collected and purified from the supernatant by sucrose cushion centrifugation 48 hours post transfection (bottom panel B). (C) Mean titre of CSGW\_4CTE encoding a competent LTR at the 3' end (CSGW\_4CTE\_3LTR) in AA-2 cells; western blot for CSGW\_4CTE\_3LTR (using mouse anti-p24 antibody) from transfected HEK-293T and VLPs collected and purified from the supernatant by sucrose cushion centrifugation 48 hours post transfection (bottom panel C). Molecular weight is shown in kDa.

### 5.1.2 CSGW\_4CTE clones encoding codon-optimized *gag-pol* HIV-1 genes in *cis* efficiently generated Gag, but VLPs produced were not infectious

The successful generation of single cycle of replication HIV-1 particles through co-transfecting CSGW\_4CTE (together with SYNGP and VSV-G in *trans*) suggested that the generation of independent HIV-1 Tat and Rev particles could be achieved with HIV-1 *gag-pol* genes provided in *cis*. As we had already developed the CSGW\_4CTE construct that lacks the HIV-1 *gag-pol* genes in *cis*, and we needed to provide in *trans* using SYNGP during transfection. We transferred the human codon-optimized *gag-pol* genes from SYNGP using *NdeI* and *NotI* sites encoded in CSGW\_4CTE\_3LTR (Figure 2-1) to generate the vector genome CSGW\_GP\_4CTE\_3LTR, which encodes the *gag-pol* genes after the SFFV promoter (Figure -5-5).



**Figure 5-5.** Diagram of CSGW\_4CTE\_3LTR encoding HIV-1 codon-optimized *gag-pol* genes (Flexi\_GP\_4CTE\_3LTR)

Introducing HIV-1 *gag* and *pol* genes (4307 bp) into CSGW\_4CTE considerably increased the size of the vector (exceeding 15kb), but there are studies showing that viral titre is not affected even when the proviral length is larger than 18 kb (Yacoub et al. 2007). Previous results demonstrated the ability of CSGW\_4CTE to generate infectious particles in a Tat and Rev independent manner, so we expected to obtain infectious Tat and Rev independent particles again using the Flexi\_GP\_4CTE\_3LTR construct (Figure 5-5). We created additional Flexi\_GP clones encoding CLTRs and MLTRs at both ends of the genome using the same strategy that we used to generate the CSGW\_Flexi clones (Figure 5-6, panel A). We co-transfected the Flexi\_GP clones together with pCMV\_VSV-G into human HEK-293T cells to produce viral particles that were inoculated onto AA-2 cells. The codon-optimized *gag-pol* genes will be provided by these clones (in *cis*).

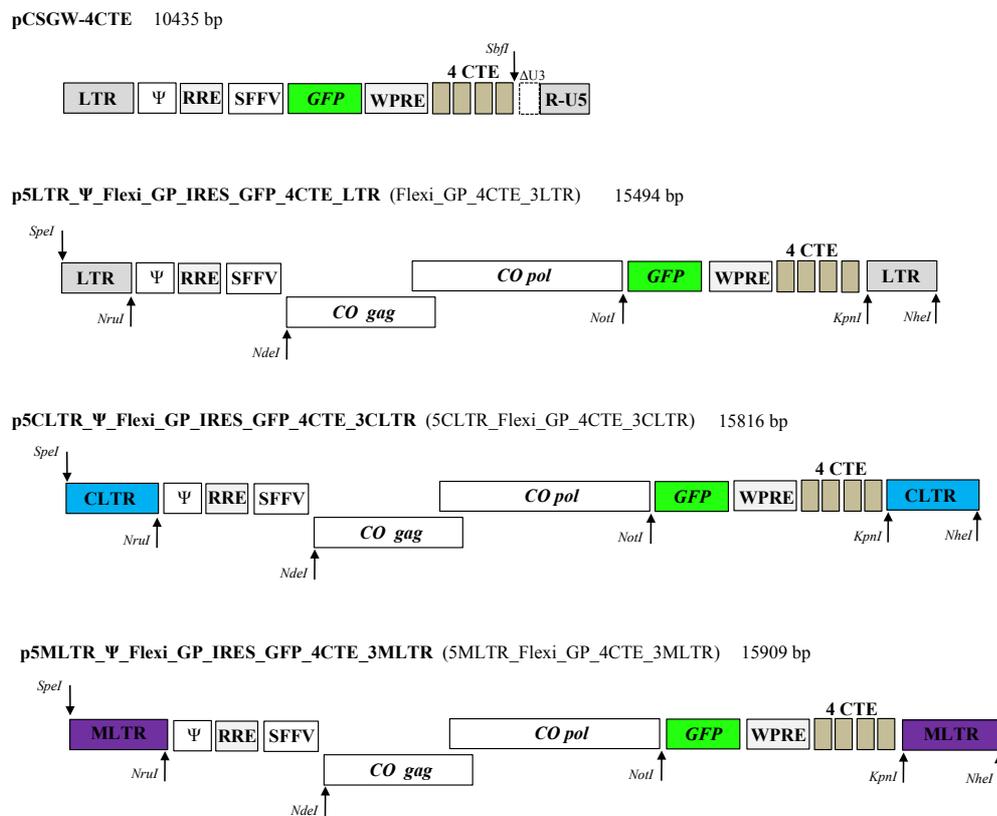
The CSGW\_GP\_4CTE\_3LTR clone was unable to generate Tat and Rev independent particles, but Gag/CA expression was similar compared to CSGW\_4CTE positive control (Figure 5-6, panel B and C). The 5CLTR\_Flexi\_GP\_4CTE\_3CLTR and 5MLTR\_Flexi\_GP\_4CTE\_3MLTR clones were not infectious in the absence of regulatory proteins, and Gag/CA expression was similar compared to Flexi\_GP\_4CTE\_3LTR and CSGW\_4CTE control (Figure 5-6, panel B and C). Even the addition of Tat and Rev in *trans* did not rescue the infectivity for Flexi\_GP\_4CTE\_3LTR, 5CLTR\_Flexi\_4CTE\_GP\_3CLTR and 5MLTR\_Flexi\_4CTE\_GP\_3MLTR clones (Figure 5-6, panel B). Western blots determined that Gag production (p55) was similar for all clones under all conditions, except for the 5MLTR\_Flexi\_4CTE\_GP\_3MLTR clone, which produced less pelletable CA protein compared to Flexi\_GP\_4CTE\_3LTR and 5CLTR\_Flexi\_4CTE\_GP\_3CLTR constructs (Figure 5-6, panel C). Gag intermediates were observed in Flexi\_GP\_4CTE\_3LTR, 5CLTR\_Flexi\_4CTE\_GP\_3CLTR and 5MLTR\_Flexi\_4CTE\_GP\_3MLTR transfected cells, but the 5CLTR\_Flexi\_4CTE\_GP\_3CLTR and CSGW\_4CTE produced more CA within transfected cells than the other constructs (Figure 5-6, panel C). Transferring the HIV-1 codon-optimized *gag-pol* genes into CSGW\_4CTE resulted in no detectable infectious progeny (Figure 5-6). We only provided a viral envelope in *trans* (using pCMV\_VSV-G as usual), but the *gag-pol* genes were not supplied in *trans* for this experiment. However, the CSGW\_GP\_4CTE\_3LTR construct was not infectious despite the addition of Tat and Rev proteins in *trans* (Figure 5-6, panel B).

The CSGW\_4CTE construct generated infectious particles in the absence of Tat and Rev when HIV-1 *gag-pol* genes were provided in *trans* by SYNGP (Figure 5-4), but including the *gag* and *pol* genes in *cis* into CSGW\_4CTE\_3LTR clone (CSGW\_GP\_4CTE\_3LTR) totally abolished viral replication (Figure 5-6). The CSGW\_GP\_4CTE\_3LTR construct was able to generate viral particles, but they were not infectious despite the successful generation of Gag within transfected cells, and CA from the VLPs collected from the supernatant (Figure 5-6, panel C). The high CpG content in codon-optimized *gag-pol* genes from SYNGP compared to wt *gag-pol* genes (227 to 24 respectively) could have disrupted the viral RNA secondary structures essential for the viability of CSGW\_4CTE construct, so we decided to re-design these genes to reduce the CpG content and investigate if we could rescue viral infectivity. However, we did not measure transfection levels directly in AA-2 cells from Flexi\_GP\_4CTE clones, but the GFP expression detected in transfected HEK-293Ts were

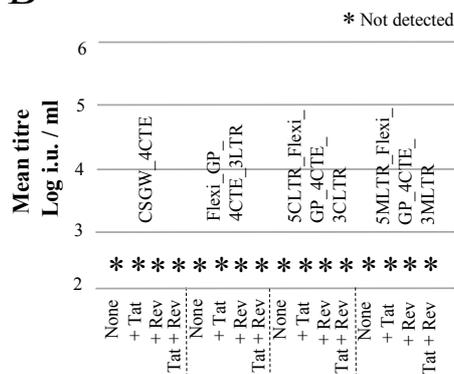
---

similar compared to CSGW\_4CTE (control). Perhaps the AA-2 cells could generate infectious VLPs when transfected, so we could attempt this experiment and determine the differences between transfected HEK-293Ts and AA-2 for p55 and VLPs production in cell lysates and from the supernatant and maybe reveal new insights.

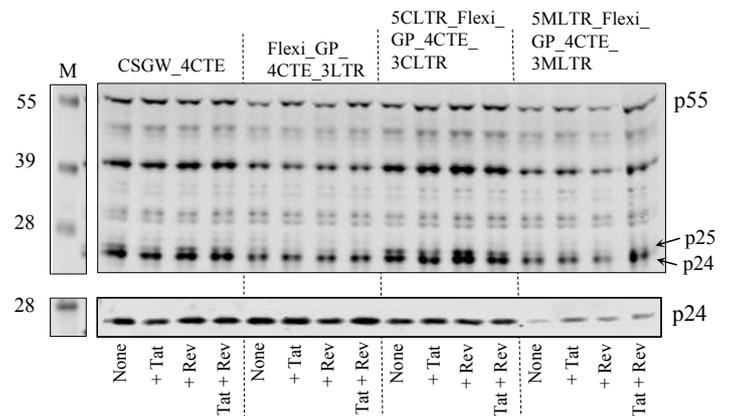
A



B



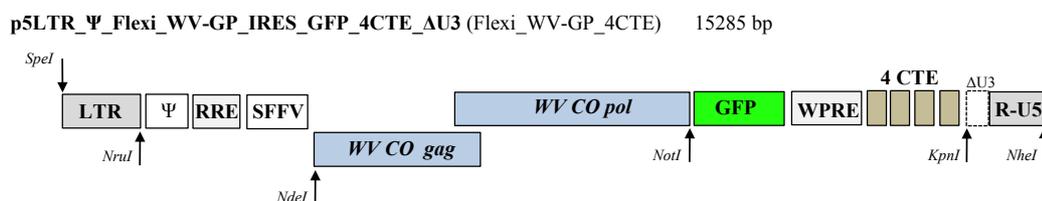
C



**Figure 5-6.** Experiment to determine the ability of Flexi\_GP\_4CTE clones to generate infectious VLPs in human cells by transfection. These clones were not infectious in AA-2 cells even in the presence of HIV regulatory proteins Tat and Rev. The amount of transfected DNA was 1 $\mu$ g for each Flexi\_GP\_4CTE clone, 100ng of pCMV\_VSV-G, 500ng of DEST40<sub>tat</sub> and 500ng of DEST40<sub>rev</sub>. (A) Diagram of the Flexi\_GP\_4CTE clones encoding the HIV-1 *gag-pol* codon-optimized genes from pSYNGP, the HIV-1 LTR, CLTR or MLTR and 4 copies of the CTE. (B) Mean titre of Flexi\_GP clones in AA-2 cells. None of them were infectious despite the addition of Tat and Rev *in trans* during transfection. (C) Western blot for Flexi\_GP\_4CTE clones (using mouse anti-p24 antibody) from transfected HEK-293T and viral particle production (VLPs) collected and purified from the supernatant by sucrose cushion centrifugation 48 hours post transfection (bottom panel C). Molecular weight is shown in kDa.

### 5.1.3 Codon-optimized HIV-1 *gag-pol* with reduced number of CpGs was unable to generate HIV-1 particles

The huge difference in the CpG content in codon-optimized *gag-pol* genes from SYNGP compared to wild type *gag-pol* genes from CMV\_NLGP (227 to 24 respectively) could be the major problem affecting the viability of our SYNGP\_4CTE and Flexi\_4CTE\_GP clones. Higher CpG content is related to pathogenicity and replication attenuation in RNA viruses (Atkinson et al. 2014; Gaunt et al. 2016). In order to reduce the number of CpGs within the human codon-optimized HIV-1 *gag-pol* genes from SYNGP, we re-designed these genes to encode fewer CpGs. This modification severely reduced the CpG content from 227 to 24, preserving the same amino acid sequence and the slippery sequence at the end of *gag* (1222-1503 position starting from *gag* ATG), which is essential to produce the required ratio of the Gag and Gag-Pol polyproteins (Biswas et al. 2004). The HIV-1 slippery sequence is a well conserved sequence involved in the ribosomal frameshift, essential for viral replication (Mathew et al. 2015). The amino acid sequence remains the same in the new genes called WV\_ *gag-pol*, (which was synthesized by Genewiz Company into pUC57- Amp vector). We cloned the WV\_ *gag-pol* genes into Flexi\_GP\_4CTE using *NdeI* and *NotI* to introduce them downstream of the SSFV promoter (Figure 5-7).



**Figure 5-7.** Diagram of CSGW\_4CTE encoding HIV-1 codon-optimized *gag-pol* genes with reduced CpG number (Flexi\_WV-GP\_4CTE)

Perhaps by reducing the CpG content from HIV-1 codon-optimised *gag-pol* genes to achieve the same number of CpG as NGLP *gag-pol* genes could restore the infectivity for CSGW\_GP\_4CTE\_3LTR construct.

We transfected the Flexi\_WV-GP\_4CTE construct together with pCMV\_VSV-G into HEK-293T cells, but we also wanted to determine the ability of this construct to package a competent genome in *trans*, such as CSGW, so we included this vector in the experiment as well (Figure 5-8). In addition, we provided DEST40\_ *tat* to compare how the Flexi\_WV-

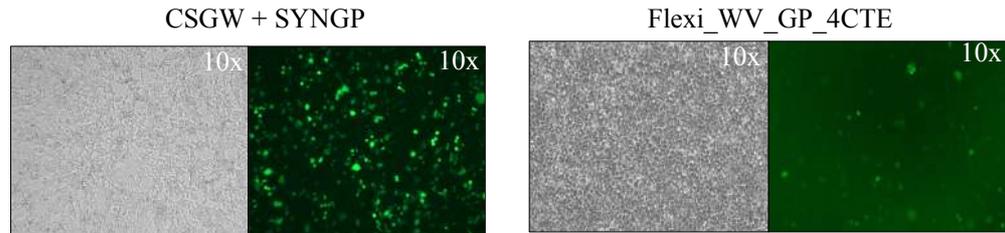
GP\_4CTE expressed Gag under different conditions. The number of transfected cells was lower using Flexi\_WV-GP\_4CTE construct compared to SYNGP control (Figure 5-8, panel B). The Gag expression from CSGW\_4CTE encoding WV\_*gag-pol* genes (CSGW\_WV-GP\_4CTE) was very poor compared to Gag produced in pSYNGP transfected HEK-293T cells (Figure 5-8, panel C). The addition of CSGW in *trans* did not affect Gag expression, but including Tat in *trans* together with CSGW resulted in a slight increase in p55-Gag expression, but no difference was observed in CA expression (Figure 5-8, panel C). No VLPs were generated from CSGW\_4CTE\_WV\_GP construct, and subsequently no infection was detected in AA-2 cells, despite the addition of Tat and CSGW in *trans* (Figure 5-8, panel D). We did not attempt to provide codon-optimized *gag-pol* genes from SYNGP (in *trans*) together with CSGW\_4CTE\_WV\_GP. Despite the barely expression from WV\_GP, perhaps we could rescue Gag expression co-transfecting SYNGP with CSGW\_4CTE\_WV\_GP, but this would not solve the additional problems that we likely introduced by reducing the CG content.

These results suggest that the modifications introduced in *gag-pol* genes severely affected Gag expression. Reducing the CG number in HIV-1 *gag-pol* genes did not solve the infectivity problem when these genes were located in *cis* into CSGW\_4CTE (CSGW\_GP\_4CTE\_3LTR), and maybe introduced additional issues that abolished VLPs production (Figure 5-8, panel C).

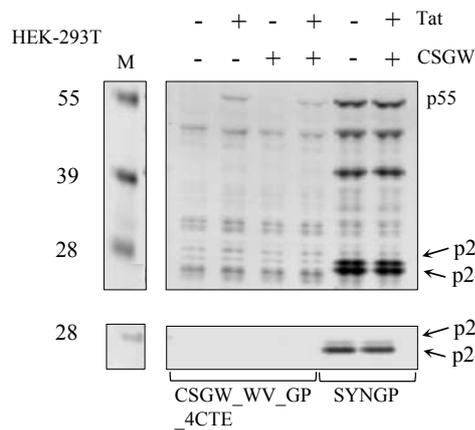
A

	HIV- <i>gag-pol</i> (pCVR1-NLGP)	SYNGP <i>gag-pol</i>	WV <i>gag-pol</i>
CpG Number	23	227	24
Length (nt)	4307	4307	4307

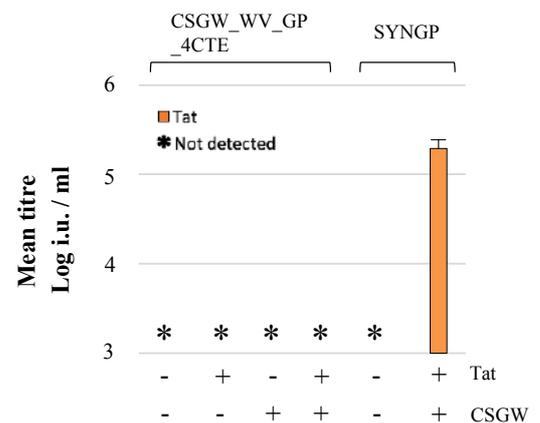
B



C



D



**Figure 5-8.** Experiment to determine the effect reduced CpG number in codon-optimized HIV-1 *gag* and *pol* genes by using WV *gag-pol* (22CpGs) instead SYNGP (227 CpGs). We transfected HEK-293T cells with 1 $\mu$ g of pCSGW, 1 $\mu$ g of pSYNGP/pFlexi\_WV\_GP\_4CTE, 100ng of pCMV\_VSV-G and 500ng of DEST40\_ *tat*. The produced VLPs were collected and purified by sucrose cushion centrifugation 48 h.p.t. and inoculated into AA-2 cells to determine the viral titre. (A) Table of CpG content from HIV-1 *gag-pol* genes, HIV-1 codon-optimized *gag-pol* genes and HIV-1 codon-optimized *gag-pol* genes with reduced CpG number. (B) GFP expression from transfected HEK-293T cells. The GFP expression was lower in Flexi\_WV\_GP\_4CTE compared to SYNGP (control). Thus, the percentage of transfection was also reduced for Flexi\_WV\_GP\_4CTE. (C) Western blot analysis of Flexi\_WV\_GP\_4CTE clone (using mouse anti-p24 antibody) from transfected HEK-293T and viral particle production (VLPs) collected and purified from the supernatant by sucrose cushion centrifugation 48 hours post transfection (bottom panel C). Molecular weight is shown in kDa. (D) Mean titre (48 h.p.i.) of VLPs collected from the supernatant 48 h.p.t. The Flexi\_WV\_GP\_4CTE clone was not infectious despite the addition of Tat in *trans* during transfection.

## 5.2 Discussion chapter 5

We successfully generated infectious particles using the CSGW\_4CTE construct (when pSYNGP and pCMV\_VSV-G were provided in *trans*) in a Tat and Rev independent manner from HEK-293T cells (Figure 5-4). When we tried to infect murine macrophages (RAW) with these viruses, which were generated in human cells, we only observed a few infected cells 72 hours post-infection (Figure 5-3). We have used the VSV envelope instead HIV-1 envelope in all experiments due to its ability to enhance viral infectivity and also to avoid incoming issues (Aiken 1997). Firstly we need to develop an HIV-1 clone able to replicate in murine cells before we replace the VSV envelope in our experiments. Thus, there are reports suggesting that pseudotyped HIV-1 particles using VSV-G protein enhance viral infectivity (Bartz et al. 1997; Aiken 1997), so we can totally focus in viral replication at this point. This enhanced viral infectivity by VSV envelope usage could explain why we saw a lower number of murine cells infected with CSGW\_4CTE vector in the absence of Tat and Rev (produced in human HEK-293T cells by transfection adding pSYNGP and pCMV\_VSV-G in *trans*), suggesting that even replacing Rev function by 4 copies of the CTE is not sufficient to rescue the infectivity within murine cells (Figure 5-3).

There are reports showing that increasing the length of lentiviral vectors could severely affect packaging and subsequently the viral titre (Kumar et al. 2001). However, the length of Flexi\_GP\_4CTE clones should not impair viral replication, and the viral RNA genome of Flexi\_GP\_4CTE clones should be able to be packaged as CSGW\_4CTE does. Surprisingly, the addition of Tat and Rev proteins in *trans* did not rescue viral infectivity from CSGW\_GP\_4CTE\_3LTR, 5CLTR\_CSGW\_GP\_4CTE\_3CLTR and 5MLTR\_CSGW\_GP\_4CTE\_3MLTR, suggesting that the viral genome must be unable to interact with HIV-1 regulatory proteins due to an aberrant RNA folding, impairing vital interactions along the viral replication cycle, such as Gag with full length genomes, or highly conserved regions in viral RNA could be also affected by introducing the *gag-pol* genes into Flexi\_GP\_4CTE clones (Le et al. 1990; Watts et al. 2009). In addition, reducing the CG content from HIV-1 codon-optimized genes could have affected RNA stability or folding, generating novel RNA secondary structures that impair viral replication. We unintentionally introduced more unknowns by reducing the CPG content in the codon-optimized *gag-pol* genes.

This hypothesis could explain why the addition of Rev protein did not rescue the viral infectivity in Flexi\_GP-4CTE clones (Figure 5-6). HIV-1 Rev protein interacts with the Rev Response Element (RRE) located within the *env* gene of full-length viral genomes, which are exported to the cytoplasm to be packaged into the nascent particles (Ping et al. 1997). The Flexi\_GP-4CTE vector encodes the RRE outside the *env* gene near the 3' end (its natural position), but the ability of wild-type CSGW to produce infectious viruses when SYNGP/CVR1\_NLGP and CMV\_VSV-G are provided during transfection supports the idea that the position of the RRE is not essential for RNA transport or stability. Thus, the RRE from CSGW\_4CTE clone is not located at the 3' end, and previous experiments demonstrated the ability of CSGW\_4CTE to generate infectious viral particles in the absence of Tat and Rev (Figure 5-4). Perhaps the position of the RRE in Flexi\_GP\_4CTE clones is affecting viral RNA transport and stability somehow, but we do not have further evidence to prove this hypothesis. However, we could transfer the RRE into its correct position (at the end of 3' transcripts) and repeat the same experiments to determine if this modification rescues viral infectivity. The viral RNA folding is essential for Rev-RRE interaction, likely non-existent in our Flexi\_GP\_4CTE clones, suggesting the likely generation of aberrant viral RNAs within the host cell (Sherpa et al. 2015). There is evidence suggesting that viral non-coding RNAs (ncRNAs) are relevant epigenetic modulators for viral transcription, which could be absent or aberrant in Flexi\_GP\_4CTE clones (Saayman et al. 2014). The higher number of CpGs present in *gag-pol* genes from SYNGP compared to wild type HIV-1 *gag-pol* genes (226-23 respectively) could also impede the generation of non-infectious particles, inhibiting transcription by RNA methylation (Bednarik et al. 1990; Blazkova et al. 2009). There is evidence to support CG dinucleotide suppression enabling antiviral defence targeting non-self RNA, which could be affecting our CSGW\_GP\_4CTE constructs (Takata et al. 2017). In addition, the usage of VSV envelope to generate the Flexi\_GP\_4CTE viruses instead using HIV-1 envelope could also affect the generation of viral transcripts within the host cell, impairing viral replication (Ran et al. 2017).

The potential of CSGW\_4CTE construct represents the best achievement we obtained at this point, generating infectious particles in a Tat and Rev independent manner. The CSGW\_4CTE is a great candidate to develop an HIV-1 clone able to replicate in mice, but we need to investigate how this clone overcomes the absence of the regulatory proteins. We could have performed RNA extractions from transfected HEK-293T with CSGW\_4CTE clone to investigate the RNA expression levels compared to wild-type

---

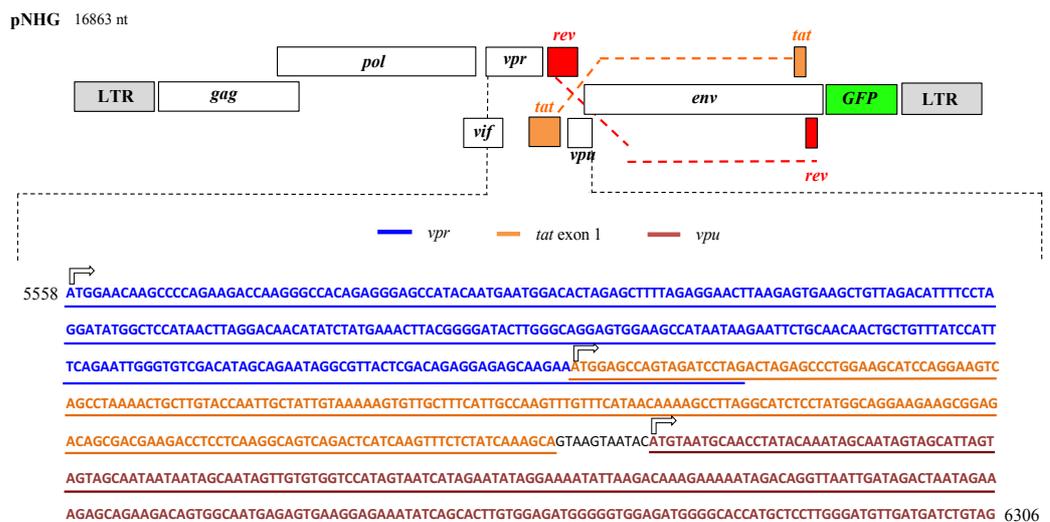
CSGW and also in VLPs by qPCR assays. Moreover, we tried several qPCR assays to investigate this hypothesis but the presence of background in all samples including the negative control discarded these results and still not clear yet if the 4CTEs encoded in CSGW\_4CTE is substituting Rev efficiently as we suggest with our previous results in this chapter.

## **Chapter 6**



Firstly, we introduced a polylinker encoding *NotI* and *XhoI* restriction sites at position 9568 (immediately before the 3'LTR). To determine the viability of the CLTR and MLTR, we generated mutations in the *tat* gene to abrogate its function and test if these chimeric promoters were able to carry out transcription and reverse transcription along several cycles of infection. We performed two kinds of mutations in the first exon of *tat* because the second exon overlaps with *env*. (Figure 6-2, panel A). The *tat-F* mutation is a frameshift modification along the first exon, the start codon was removed as well and we also introduced several stop codons (Figure 6-2, panel B). The *tat-D* mutation is a deletion of 43 nucleotides to severely impede reversion events, which could restore Tat expression after several cycles of replication in the Tat-F constructs. The NHG vectors were called NHG\_Δ*tat-F* and NHG\_Δ*tat-D* (Figure 6-2, panel B). We decided to develop two mutants of Tat to increase the possibilities to achieve replication competent virus able to replicate in its absence.

A

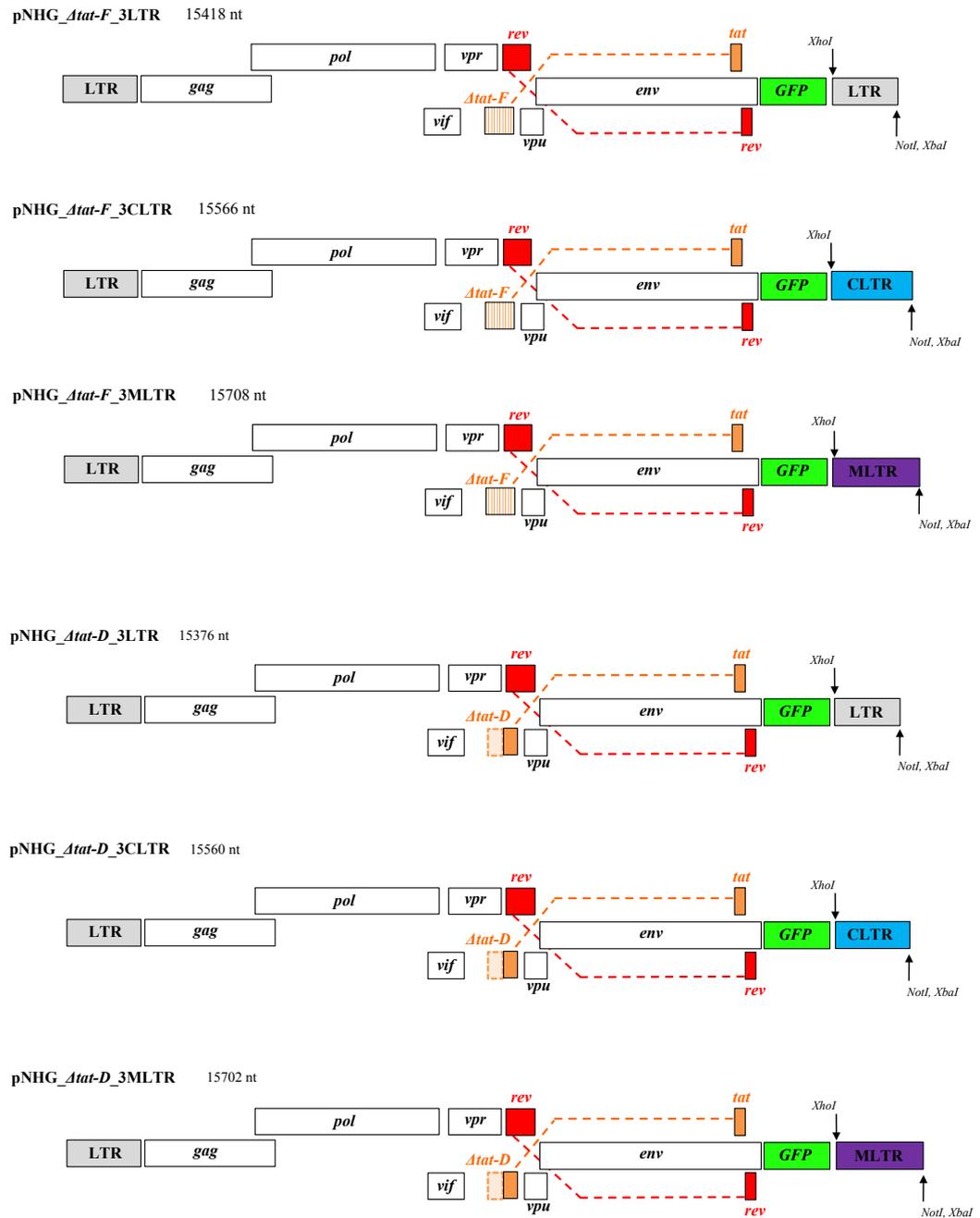


B



**Figure 6-2.** (A) Tat exon 1 location in pNHG. (B) Mutations made in *tat* exon 1: *tat-F* is a frameshift mutation introducing several stop codons and the start codon was also removed; *tat-D* is a deletion of 43 nucleotides to impede recombination events that could restore *tat* expression.

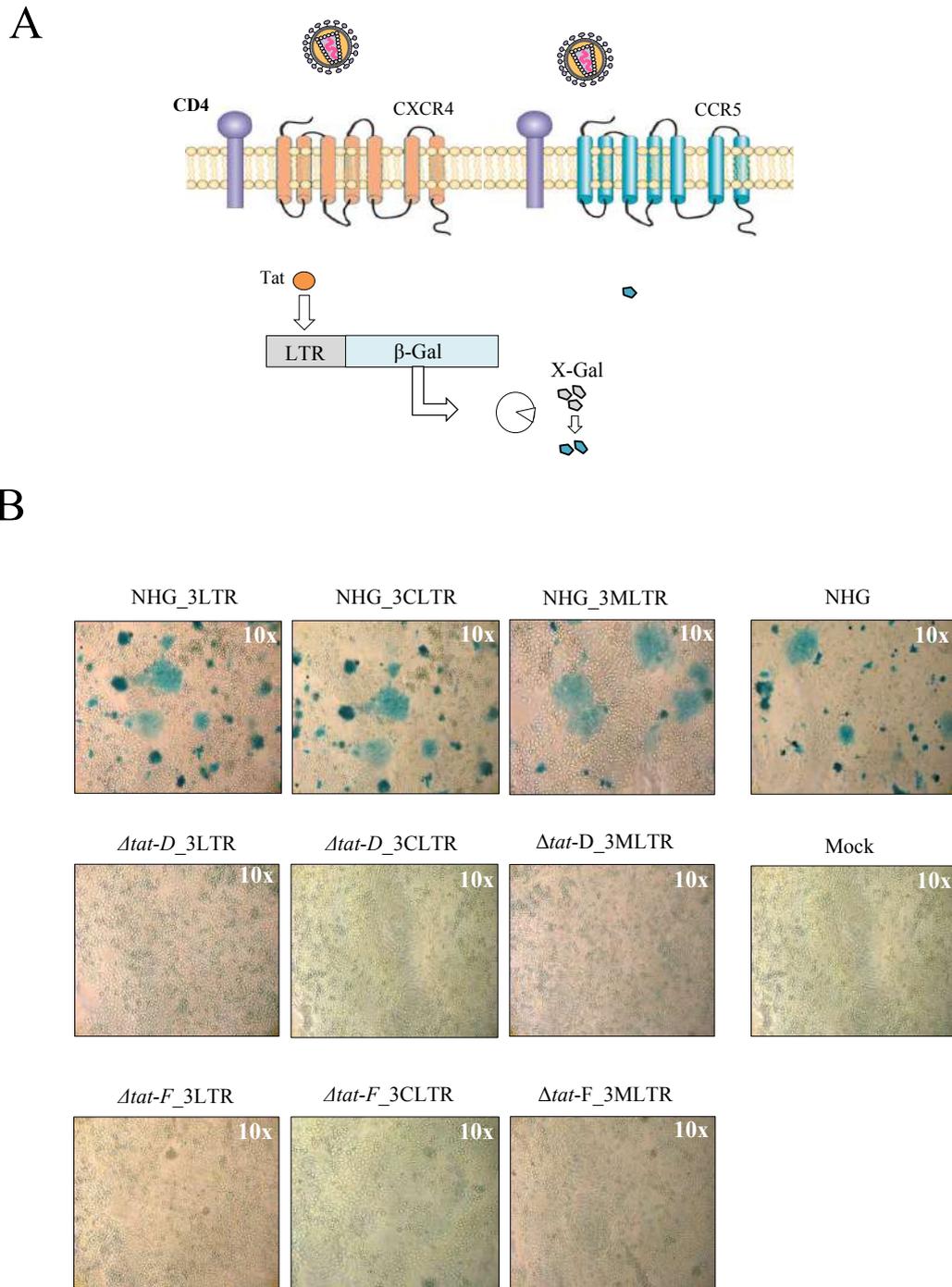
Then, we generated NHG\_Δ*tat-F* and NHG\_Δ*tat-D* constructs encoding LTR, CLTR or MLTR at the 3' end. Finally, we created 6 different NHG\_Δ*tat-F* and NHG\_Δ*tat-D* viruses (Figure 6-3).



**Figure 6-3.** Diagram of NHG\_Δ*tat-F* and NHG\_Δ*tat-D* clones encoding a 3'LTR (control), CLTR and MLTR. The *tat-F* mutation is a frameshift that introduced several stop codons along the first exon, and the start codon was also removed; *tat-D* is a deletion of 43 nucleotides to impede recombination events that could restore *tat* expression.

We performed TZM assays to determine the absence of functional Tat protein in NHG\_Δ*tat* clones. TZM-bl cells are a reporter cell line encoding an *E. coli* β-galactosidase gene under control of an HIV-1 LTR promoter (Tat dependent) to measure Tat expression after a single round of infection. TZM-bl cells also express HIV-1 CD4 receptor and co-receptors (CXCR4 and CCR5). (Figure 6-4, panel A). In the absence of Tat, the LTR is unable to drive transcription and the β-galactosidase enzyme is not generated. However, HIV-1 Tat stimulates the HIV-1 LTR from the TZM-bl cells, leading to the generation of the β-galactosidase enzyme, which hydrolyzes the X-Gal (lactose analogue), creating blue precipitates that we can observe and quantify (Figure 6-4, panel B).

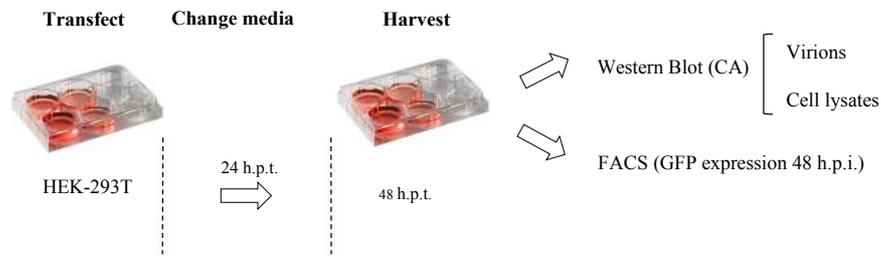
Firstly we transfected the 6 NHG\_Δ*tat* clones from (Figure 6-3) into TZM-bl cells, and then prepared a fresh X-Gal solution to add onto the TZM-bl cells 48 hours post-transfection to detect Tat activity. The NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR clones developed blue precipitates upon viral infection, but none of NHG\_Δ*tat* clones (NHG\_Δ*tat*\_3LTR, NHG\_Δ*tat*\_3CLTR and NHG\_Δ*tat*\_3MLTR) formed blue precipitates (Figure 6-4, panel B). These TZM assays demonstrated that *tat* expression was not detected following transfection of NHG\_Δ*tat* clones, confirming that our Δ*tat-F* and Δ*tat-D* mutations efficiently abrogated Tat expression during the first cycle of replication.



**Figure 6-4.** (A) Diagram of a standard TZM-bl assay. TZM-bl cells are a reporter cell line presenting an *E. coli*  $\beta$ -galactosidase gene under control of an HIV-1 LTR (Tat dependent) to measure Tat expression after a single round of infection. TZM-bl also express HIV-1 receptor and co-receptor (CD4 and CXCR4 respectively). Diagram adapted from Akrit Sodhi, Silvia Montaner & J. Silvio Gutkind, *Nature Reviews Molecular Cell Biology* 5, 9981012 (December 2004). (B) Transfection of NHG constructs into TMZ-bl cells to detect Tat expression. The amount of transfected DNA was  $1\mu\text{g}$  of each NHG construct and  $500\text{ng}$  of DEST40-*tat*]. Some blue crystals can be observed, consistent with Tat activity.

### 6.1.2 Tat expressed *in trans* rescues NHG-*tat* mutants. The CLTR and MLTR support reverse transcription efficiently

We previously demonstrated that NHG\_Δ*tat* mutants do not express Tat protein, so the next step was to generate these replication competent viruses within the CL3 facilities following the same procedure as we did with the SIN vectors in CL2 labs (Figure 6-5). We transfected the NHG\_Δ*tat* clones (Figure 6-3) into human HEK-293T cells, and collected the viruses from the supernatant 48 hours post transfection to inoculate them onto PM1 human lymphoid cell line to determine the viral titre (Figure 6-6, panel A). In order to generate NHG\_Δ*tat* clones it was necessary to provide DEST40\_Δ*tat* in *trans* during transfection to stimulate transcription levels for the first cycle of replication.



**Figure 6-5.** Standard experiment to generate replication competent HIV-1 particles in human HEK-293T cells. The virions were collected 48 hours post-transfection and inoculated onto PM1 or AA-2 cells (human lymphoid cell lines) to determine the viral titre.

All NHG\_Δ*tat* viruses were infectious in PM1 cells when Tat was supplied in *trans* during transfection (Figure 6-6). The NHG\_3LTR virus had a similar titre to the parental NHG virus, demonstrating that the addition of the polylinker near the 3'LTR did not affect viral fitness (Figure 6-6, panel A, lanes 1 and 2). However, the NHG\_3CLTR and NHG\_3MLTR titre decreased ~10-fold compared to NHGwt and NHG\_3LTR viruses (Figure 6-6, panel A, lanes 3 and 4). None of the NHG\_Δ*tat* mutants were infectious when Tat was not provided in *trans* during transfection, but their infectivity was rescued when Tat was added for the first cycle of replication (Figure 6-6, panel A, lanes 5 to 10). However, the titre from NHG\_Δ*tat* clones was reduced ~10-fold compared to NHGwt, NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR viruses (Figure 6-6, panel A).

We performed western blot analysis in the transfected cells to detect Gag/CA expression, and we observed that Gag expression was slightly higher in NHG clones encoding *tat* in *cis*

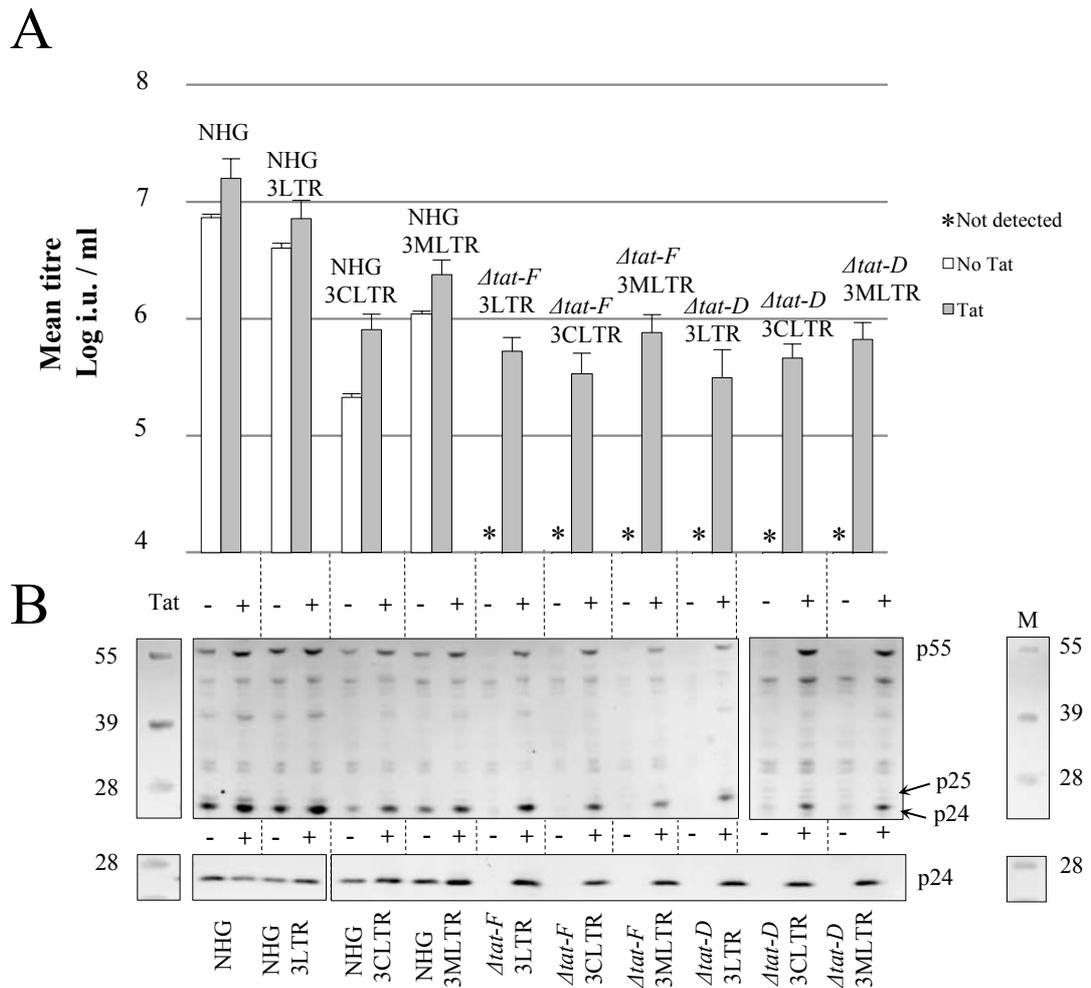
(NHGwt, NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR). We did not observe p55 or p24 expression in transfected HEK-293T cells with NHG\_Δtat constructs due to low transcription levels in the absence of Tat (when DEST40\_tat was not provided in *trans* during transfection). Thus, infectious particles were not generated by the NHG\_Δtat constructs in the absence of exogenous Tat, which correlates with the absence of p55 and p24 in transfected cells (Figure 6-6, panel B, lanes 9, 11, 13 and 15). It is essential to provide Tat protein during the first cycle of replication to enhance viral transcription and produce enough p55 and p24 proteins to allow the formation of nascent particles in transfected HEK-293T cells.

The NHG\_Δtat clones produced p55/p24 successfully when Tat was added in *trans* during transfection (Figure 6-6, panel B). The amount of Gag protein produced was very similar for parental NHG, NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR viruses compared to NHG\_Δtat clones when Tat was added in *trans* (Figure 6-6, panel B, bottom panel). After producing NHG\_Δtat clones in human HEK-293T cells, we collected and filtered the supernatant containing these viruses and used them to inoculate a human lymphoid cell line (AA-2 cells). We intended to perform viral passages in an attempt to increase the viral fitness of these NHG\_Δtat viruses. However, prior to attempting this experiment we decided to validate that the infectious viruses were truly *tat*-deficient. We inoculated the NHG\_Δtat viruses onto TZM-bl cells to perform a TZM assay 48 hours post-infection (Figure 6-7, panel A). We observed blue precipitates in NHGwt, NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR viruses. However, the NHG\_Δtat\_3LTR, NHG\_Δtat\_3CLTR and NHG\_Δtat\_3MLTR viruses did not produce blue precipitates (Figure 6-7, panel A). Crucially, all NHG\_Δtat viruses were infectious (GFP expression) when Tat was provided in *trans* during transfection (Figure 6-7, panel B) indicating that the absence of blue precipitate was not merely due to an absence of infected cells.

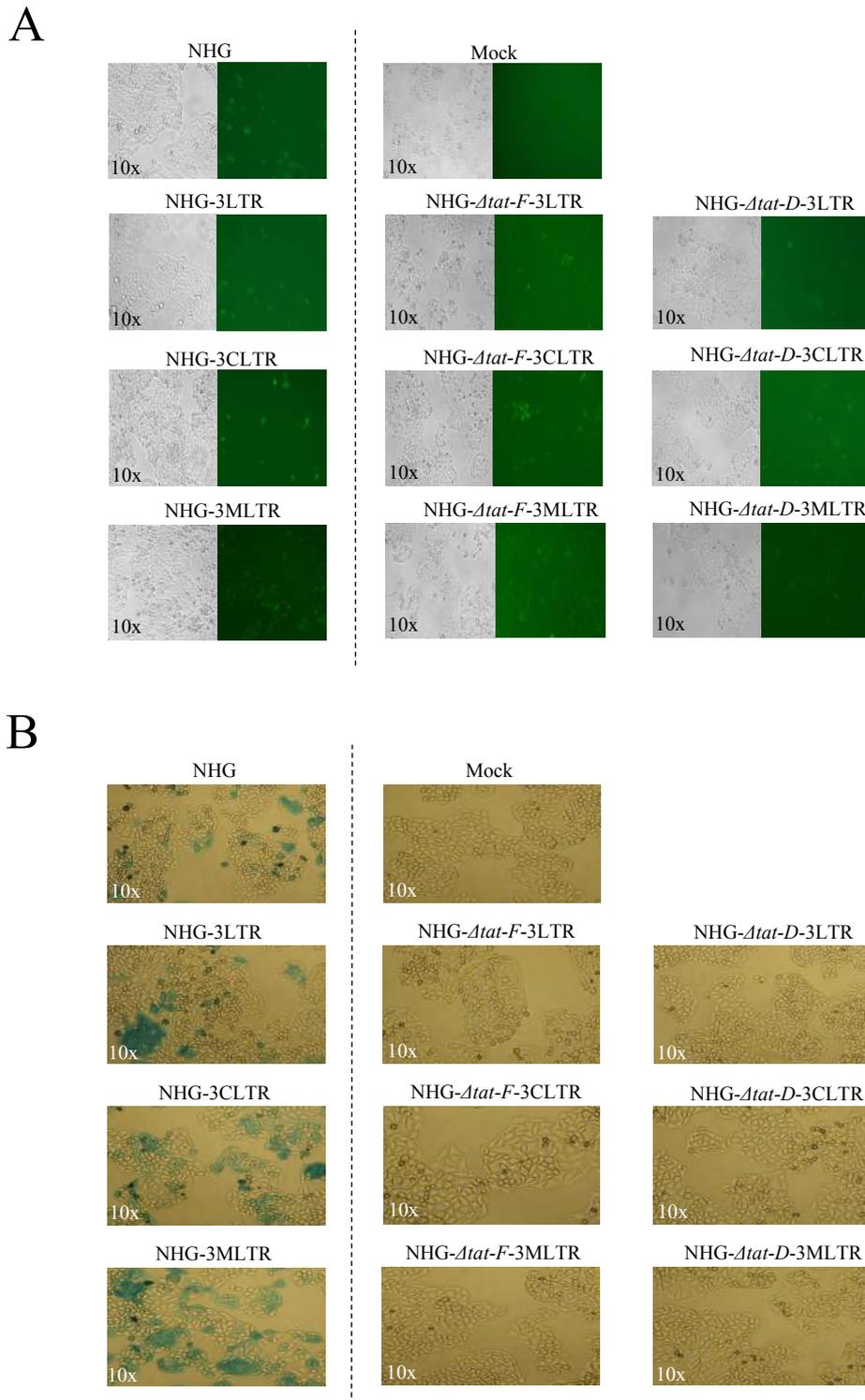
The NHG\_Δtat clones supported efficient transcription and reverse transcription during the first cycle of replication when Tat was provided in *trans*. These replication competent viruses will allow us to investigate their ability to complete subsequent cycles of replication in a Tat-independent manner within a more representative environment. These viruses could, in principle, adapt to human AA-2 cells increasing replicative fitness through repeated viral passage, and we should be able to detect (by direct sequencing) the

---

modifications selected in NHG\_Δ*tat* clones that increase the viral fitness after multiple cycles of replication. We have sequenced all viruses employed in these experiments by sequencing.



**Figure 6-6.** Experiment to determine the ability of NHG\_Δtat constructs to generate infectious particles in PM1 by transfection. HEK-293T cells were transfected with NHG\_Δtat constructs and viral particles were collected and purified by sucrose cushion centrifugation 48 h.p.t. The amount of transfected DNA was 1 μg for each NHG\_Δtat construct and 500 ng of DEST40\_tat. (A) Mean titre of NHG clones encoding *tat* in *cis* (left), NHG\_Δtat-F and NHG\_Δtat-D clones in PM1 cells. Tat was provided *in trans* by DEST40\_tat for the first cycle of replication. No infection was observed in NHG\_Δtat constructs when DEST40\_tat was not supplied *in trans* during transfection. The error bars show the standard deviation for each titre (triplicates). (B) Western blot analysis for p55 and p24 (using mouse anti-p24 antibody) of transfected HEK-293T cells and viral particle production from the supernatant 48 hours post transfection (upper panel). Viral particles produced were collected and purified by sucrose cushion centrifugation 48 h.p.t. (bottom panel). Molecular marker shown in kDa.



**Figure 6-7.** Experiment to determine Tat expression of NHG  $\Delta$ tat constructs in TSM-bl cells after the first cycle of replication. We inoculated the same amount of viruses using the mean titre of each virus from Figure 6-6 to obtain same levels of infections for all constructs NHG. (A) GFP expression from infected TSM cells (48 h.p.i.) with NHG clones encoding *tat* in *cis* (left column) and NHG  $\Delta$ tat constructs (NHG  $\Delta$ tat-*F* and NHG  $\Delta$ tat-*D*). (B) TZM assay to detect Tat expression in our NHG  $\Delta$ tat constructs 48 h.p.i. Blue precipitates were generated when Tat was present in the system.

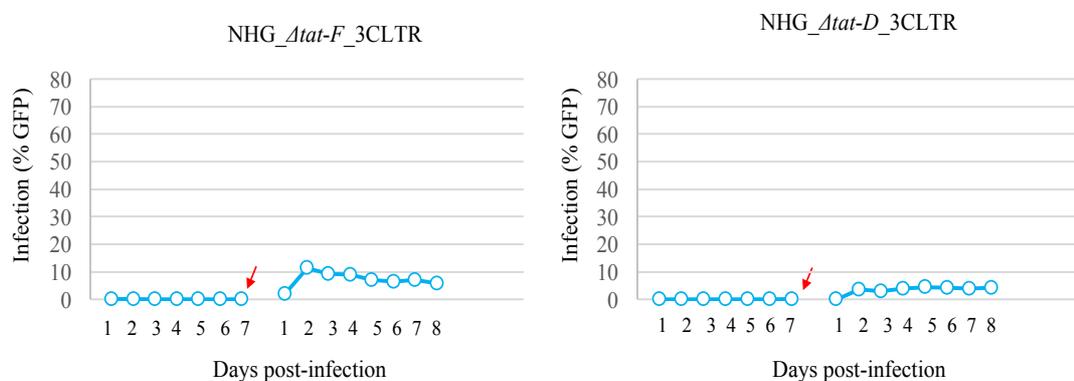
### 6.1.3 NHG\_Δtat clones were able to complete subsequent cycles of replication in a Tat-independent manner

Once we determined that all NHG\_Δtat viruses were infectious during the first cycle of replication we proceeded to perform viral passage experiments to increase the viral fitness. The levels of infection in a Tat free environment were expected to be low, but the NHG\_Δtat viruses should be able to overcome this situation following several rounds of replication (Bordería et al. 2010), allowing the virus to evolve and adapt within human AA-2 cells prior attempting to infect murine cells. In summary, the resultant virus would be able to replicate efficiently in a Tat-independent manner predisposing such a virus to replicate within murine cells. The modifications introduced through viral passage could be transferred into our CSGW\_Flexi and SYNGP\_4CTE systems to finally acquire a Tat independent virus within a CL2 lab, which is an easier and also safer way to investigate the molecular basis of these adaptations.

We passaged all the viruses shown in Figure 6-3. The NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR viruses all displayed similar growth kinetics compared to the NHG wild type virus. These viruses killed all the cells in the culture within a week. Specifically, the NHG\_3MLTR only needed 3 days (Figure 6-9) whereas the NHG\_Δtat-F\_3LTR and NHG\_Δtat-D\_3LTR mutants needed 6 and 7 days respectively to reach same levels of infection as NHG\_3LTR, NHG\_3CLTR and NHG\_3MLTR viruses. The NHG\_Δtat-F\_3MLTR and the NHG\_Δtat-D\_3MLTR displayed similar growth curves (comparing both), but the NHG\_Δtat-F\_3CLTR and NHG\_Δtat-D\_3CLTR viruses propagated very slowly or not at all as the level of infection was stable (~15% of cells infected/GFP positive) after 13 days (Figure 6-9). The next passage of all NHG\_Δtat viruses determined that viral fitness had increased or perhaps achieved higher titre with growth kinetics resembling wild type growth curves (Figure 6-9). However, the NHG\_Δtat-F\_3CLTR and NHG\_Δtat-D\_3CLTR remained the same, and we did not observe any enhancement in the replicative fitness of these viruses. Unfortunately, the NHG\_Δtat-F\_3CLTR and NHG\_Δtat-D\_3CLTR passages were interrupted due to a protracted CL3 shutdown (red arrows), suspending the adaptation of these viruses for a considerable period of time.

No differences were observed in the growth curves from NHG\_Δtat-F\_3LTR and NHG\_Δtat-D\_3LTR viruses, suggesting that neither mutagenesis strategy affected viral replication. The incredibly rapid adaptation of NHG\_Δtat-F\_3LTR and NHG\_Δtat-D\_3LTR viruses to

the culture suggested a likely contamination event; same hypothesis could be applied to the NHG\_ *Atat-F*\_3MLTR and NHG\_ *Atat-D*\_3MLTR viruses. The NHG\_ *Atat-F*\_3CLTR and NHG\_ *Atat-D*\_3CLTR growth curves remained consistent along 13 days, confirming that viral adaptation is a slow process and the virus needs many cycles of replication before generate mutations to increase its fitness (Locher et al. 2003). Unfortunately, the viral titre from the NHG\_ *Atat* viruses was so low as to be undetectable by flow cytometry. In addition, another shutdown was carried out for refurbishment within the CL3 facilities, and experiments were interrupted once again for approximately 6 months. Once the CL3 lab was available again, some viruses that were frozen prior the shutdown were inoculated onto fresh AA-2 cells, but they did not infect any cells. The frozen NHG\_ *Atat* viruses could not be recovered and we speculate the titre was so low that freeze-thawing effectively eliminated infections virus from the culture. Thus, we decided to generate the NHG\_ *Atat-F*\_3CLTR and NHG\_ *Atat-D*\_3CLTR viruses once again. The new growth curves from NHG\_ *Atat-F*\_3CLTR and NHG\_ *Atat-D*\_3CLTR viruses had a very similar pattern as the previous experiments, where the percentage of infection remained similar for ~8 days, but these samples were frozen due to CL3 fumigation/housekeeping duties by this time (Figure 6-8).



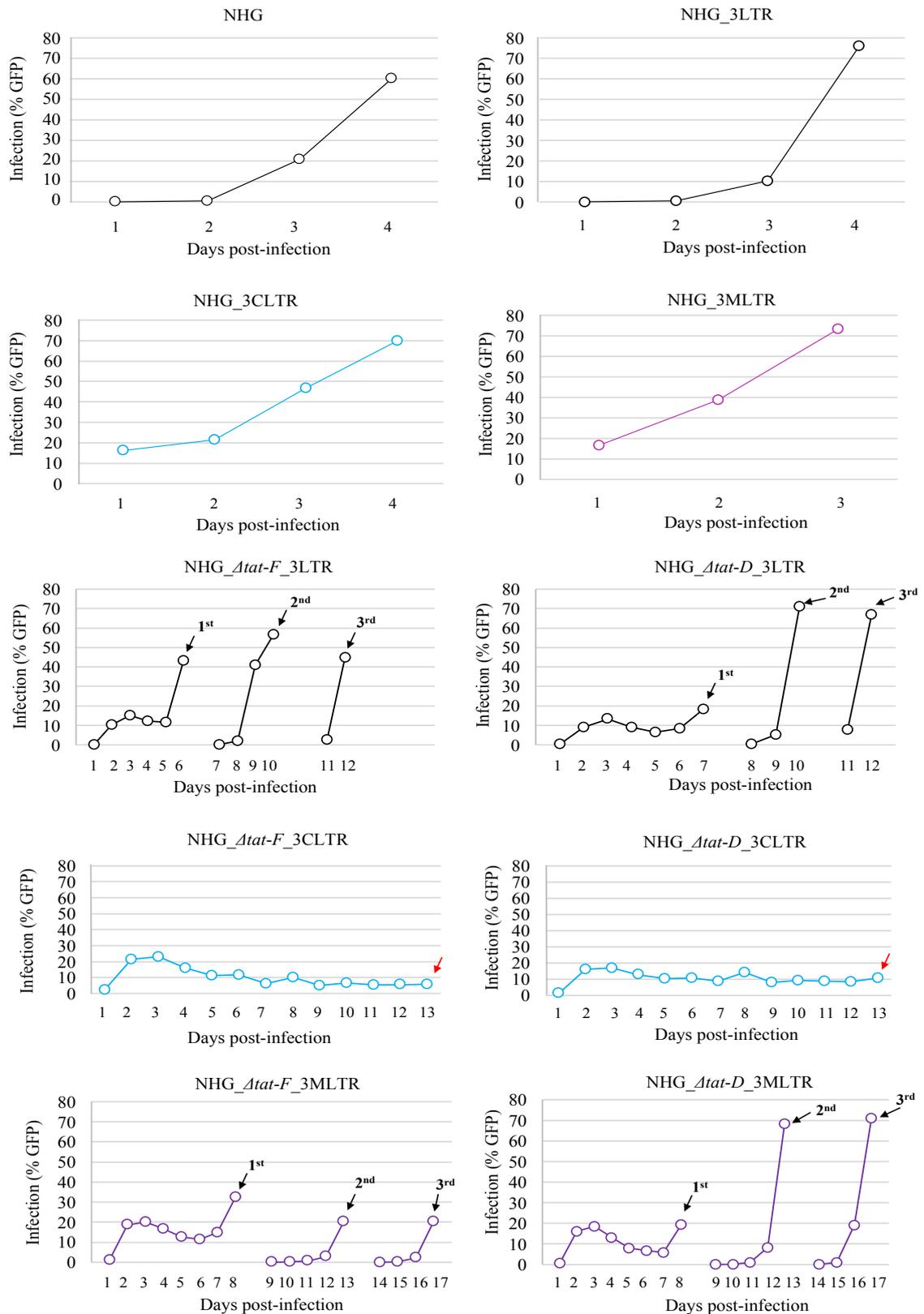
**Figure 6-8.** NHG\_ *Atat-F*\_3CLTR and NHG\_ *Atat-D*\_3CLTR viruses adaptation in human AA-2 cells. The growth curves of. Red arrows indicate shutdown period in the CL3 facilities and passages were interrupted. The right curve represents the growth of a fresh virus after the shutdown period that did not spread due to the interruption of the culture (frozen stock remove any infectivity from this virus).

To determine if the NHG\_ *tat* mutants had recovered functional Tat expression, we carried out a TZM-bl assay, which would also detect contamination events in NHG\_ *tat* mutants cultures. The rapid adaptation of NHG\_ *Atat-F*\_3LTR, NHG\_ *Atat-D*\_3LTR, NHG\_ *Atat-F*\_3MLTR and NHG\_ *Atat-D*\_3MLTR viruses to human AA-2 cells culture was achieved

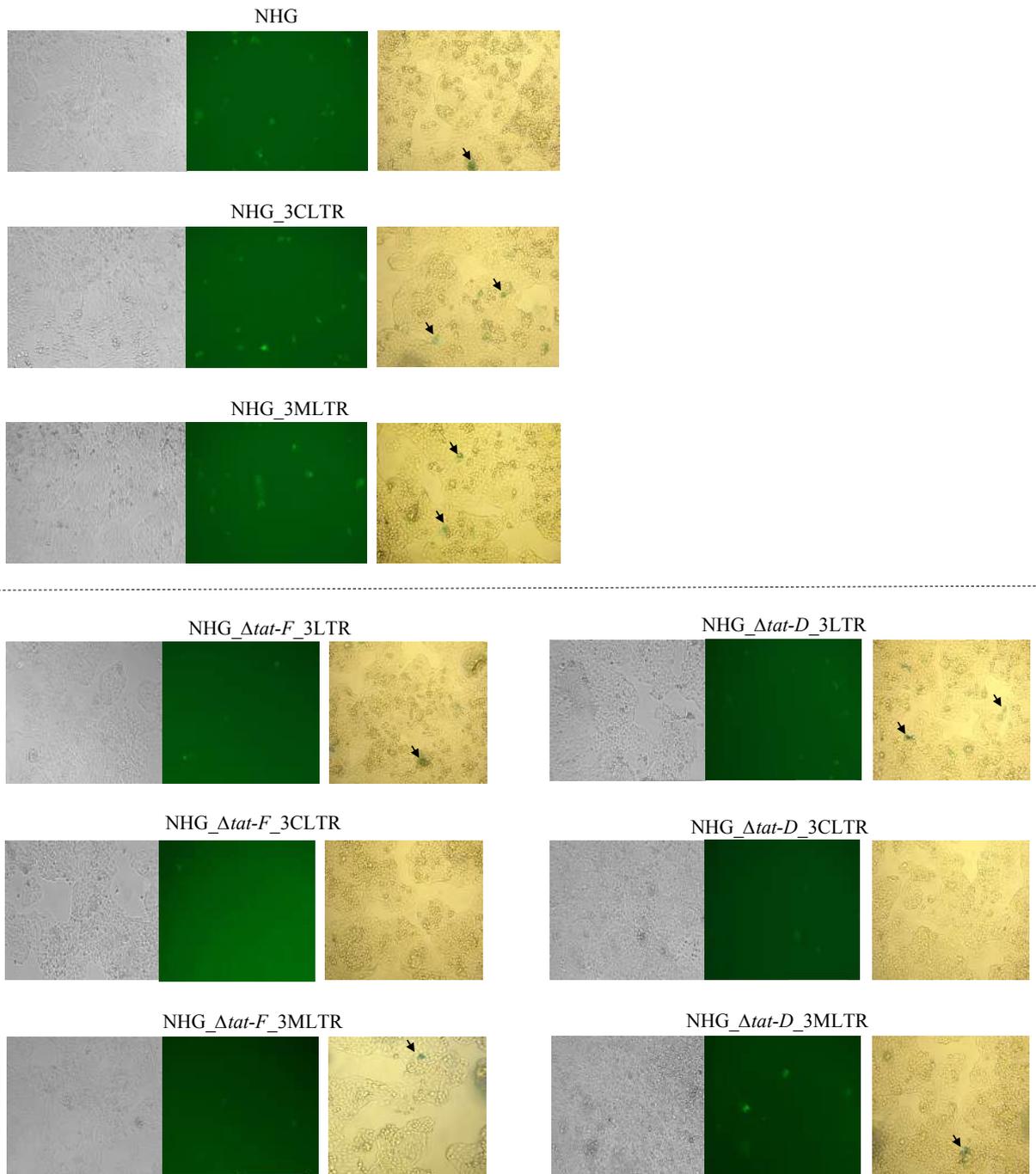
by cross contamination of a wild type virus, which was determined by PCR analysis. Following this surprising results we decided to sequence *tat* 1 exon of infected cells from NHG\_Δ*tat-F*\_3LTR, NHG\_Δ*tat-D*\_3LTR, NHG\_Δ*tat-F*\_3MLTR and NHG\_Δ*tat-D*\_3MLTR cultures, so we centrifuged each culture to collect the infected cells, and we extracted their genomic DNA and sequenced the first exon of *tat* (the site that we originally mutated). The sequencing results from NHG\_Δ*tat-F*\_3LTR, NHG\_Δ*tat-D*\_3LTR, NHG\_Δ*tat-F*\_3MLTR and NHG\_Δ*tat-D*\_3MLTR infected cells determined that *tat* exon 1 was wild type (NHG *tat*) for all of them. We also performed TZM assays of all NHG\_Δ*tat* viruses, which confirmed the presence of functional Tat within the cultures (Figure 6-10). However, the NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR viruses were the only ones not contaminated.

Despite the issues related to contamination events and extended and unpredictable shutdown periods in the CL3 lab, these results demonstrated that the CLTR was able to drive transcription and reverse transcription efficiently after subsequent rounds of infection in a Tat independent manner (Figure 6-9 and 6-10), resulting in the ‘best’ candidate to develop an HIV-1 Tat independent clone in human cells. Moreover, TZM assays did not detect Tat expression for NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR viruses after 11 days post-infection, but they still infectious (figure 6-10). However, the NHG\_Δ*tat-F*\_3LTR and NHG\_Δ*tat-F*\_3MLTR, NHG\_Δ*tat-D*\_3LTR and NHG\_Δ*tat-D*\_3MLTR viruses were expressing Tat 11 days post-infection (Figure 6-10)

Conclusions derived from the NHG\_Δ*tat-F*\_3LTR and NHG\_Δ*tat-F*\_3MLTR, NHG\_Δ*tat-D*\_3LTR and NHG\_Δ*tat-D*\_3MLTR viruses were not reliable due to contamination events, and more experiments need to be performed to investigate their potential.



**Figure 6-9. HIV-1 adaptation in human AA-2 cells.** The NHG\_*Atata-F* and NHG\_*Atata-D* viruses were passaged in 6 ml cultures of AA-2 until the viruses totally spread (most of cell expressing GFP and cytopathic effects). Then, we prepared a new culture with fresh AA-2 cells and inoculated the virus from the 1<sup>st</sup> passage (previously filtered with 0.45  $\mu$ m 25mm diameter syringe filters). The inoculum for each virus into the new cultures was 500 $\mu$ l. To quantify GFP expression from the cultures we submitted to flow cytometry analysis the infected cells from each culture. Black arrows indicate the end of a passage, and red arrows a CL3 facility shutdown.



**Figure 6-10.** TZM-bl assays to detect Tat expression from NHG\_Δtat-F and NHG\_Δtat-D viruses from AA-2 passages. We suspected that NHG\_Δtat-F viruses will revert Tat function easily, but NHG\_Δtat-D viruses should be unable to express Tat. The TZM-bl cells were inoculated with NHG\_3LTR, NHG\_3CLTR, NHG\_3MLTR and NHG\_Δtat viruses 11 days post-infection and incubated for 72 hours prior the addition of X-gal. Some Tat expression was recovered (black arrows). NHG\_3LTR, NHG\_3CLTR, NHG\_3MLTR virus inoculum was diluted 1:100 (control). The level of infection of NHG\_Δtat-F and NHG\_Δtat-D viruses were too low to determine the viral titre by flow cytometry.

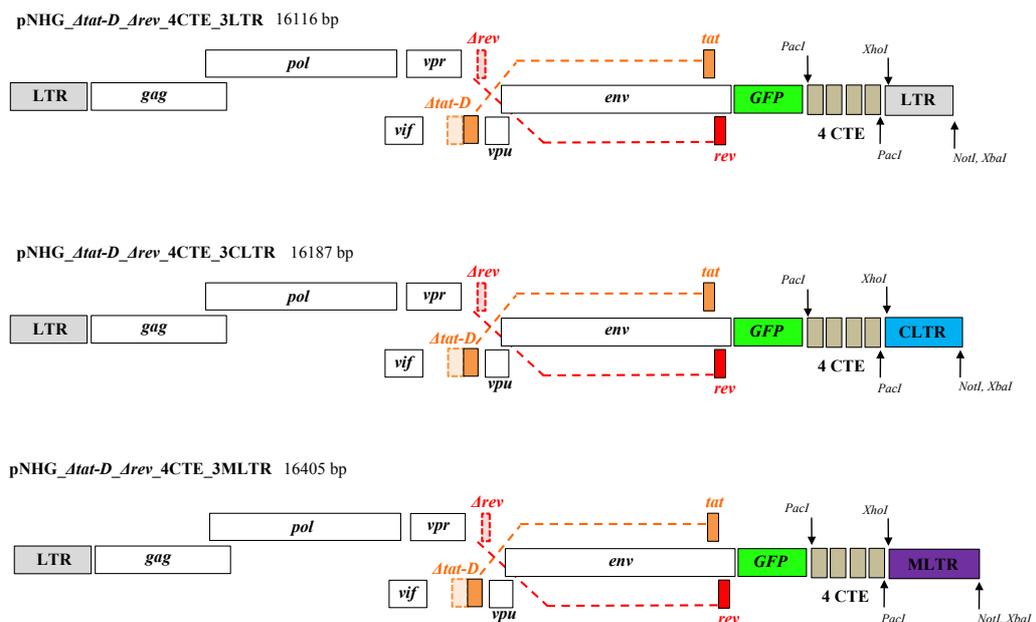
#### **6.1.4 NHG-*tat-rev*-4CTE clones were able to complete subsequent cycles of replication in a Tat and Rev independent manner, but infectivity was lost along the passages**

Our previous results obtained from NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR passages demonstrated that they could still be infectious after 11 days post-infection in a Tat independent manner (Figure 6-9). The TZM assays to detect Tat expression for NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR viruses did not detect Tat expression, and these viruses were infectious after 11 days post-infection (Figure 6-7). These data support the concept that a Tat-independent virus encoding the CLTRs was possible, therefore we considered the NHG\_Δ*tat-F*\_3MLTR and NHG\_Δ*tat-D*\_3MLTR viruses as potential candidates to eventually replicate efficiently within a Tat independent environment. Further adaptation would be necessary prior to attempting to infect murine cells. In order to achieve a Tat and Rev independent HIV-1 clone able to infect murine cells it is necessary to remove Rev-dependency. We attempted to select Rev-independent viruses by deleting *rev* exon 1 from NHG\_Δ*tat* viruses to generate NHG\_Δ*tat*\_Δ*rev* vectors that are unable to express HIV Tat and Rev proteins, so we could introduce 4 copies of the CTE and chimeric promoters CLTR and MLTR into NHG\_Δ*tat*\_Δ*rev* vectors to investigate the ability of this vector to replicate in the absence of the regulatory proteins. However, we could have introduced mutations in the HIV RRE to abolish Rev function or removing entirely the RRE from these constructs instead modifying the *rev* exon 1. However, the modifications in the RRE could alter the viral envelope structure (RRE is located within the *env* gene) and we decided to remove *rev* exon 1 that does not overlap with other genes. Same hypothesis was followed when we introduced mutations in the *tat* exon 1 in NHG\_Δ*tat* constructs, which did not alter other sequences along the genome. We could have modified the TAR element in NHG\_Δ*tat* clones, but altering the secondary structure of HIV LTR could severely impede the efficiency of viral replication.

Therefore, we completely removed the first exon of *rev*, which does not overlap with any other viral genes (Δ*rev*). The viral RNA stability and transport should have been severely attenuated, but our previous results carried out in the CL2 system using 4CTEs (encoded into CSGW vector genome) demonstrated that the 4CTEs were partially able to replace Rev function (Figure 5-4). We followed this strategy in our NHG\_Δ*tat-D*\_Δ*rev* viruses and transferred 4 copies of the CTE into them, generating 4 new viruses encoding Δ*tat-D* and

$\Delta rev$  mutations and also the chimeric CLTR and MLTR promoters at the 3' end (Figure 6-11). The addition of 4CTEs should contribute to the adaptation of NHG\_ $\Delta tat$ - $\Delta rev$ \_4CTE viruses along several cycles of replication in AA-2 cells within a Tat and Rev free environment. Conceptually the modified LTRs were intended to eventually result in Tat independence and the CTEs to promote Rev-independence.

The *rev* exon 2 was intact in NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ -4CTE viruses because it overlaps with the *env* gene. The NHG- $\Delta tat$ - $F$  mutants were excluded from this modification for their ability to easily revert to functional Tat expression (Figure 6-10). We transfected the NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ \_4CTE viruses (NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ \_4CTE\_3LTR, NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ \_4CTE\_3CLTR and NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ \_4CTE\_3MLTR) into HEK-293T cells to generate replication competent HIV-1 particles, which were used to infect human AA-2



**Figure 6-11.** NHG\_ $\Delta tat$ - $D$ - $\Delta rev$  viruses encoding 4CTEs at the 3' end. The LTR, CLTR and MLTR promoters are also located at the 3' end such as NHG\_ $\Delta tat$  mutants.

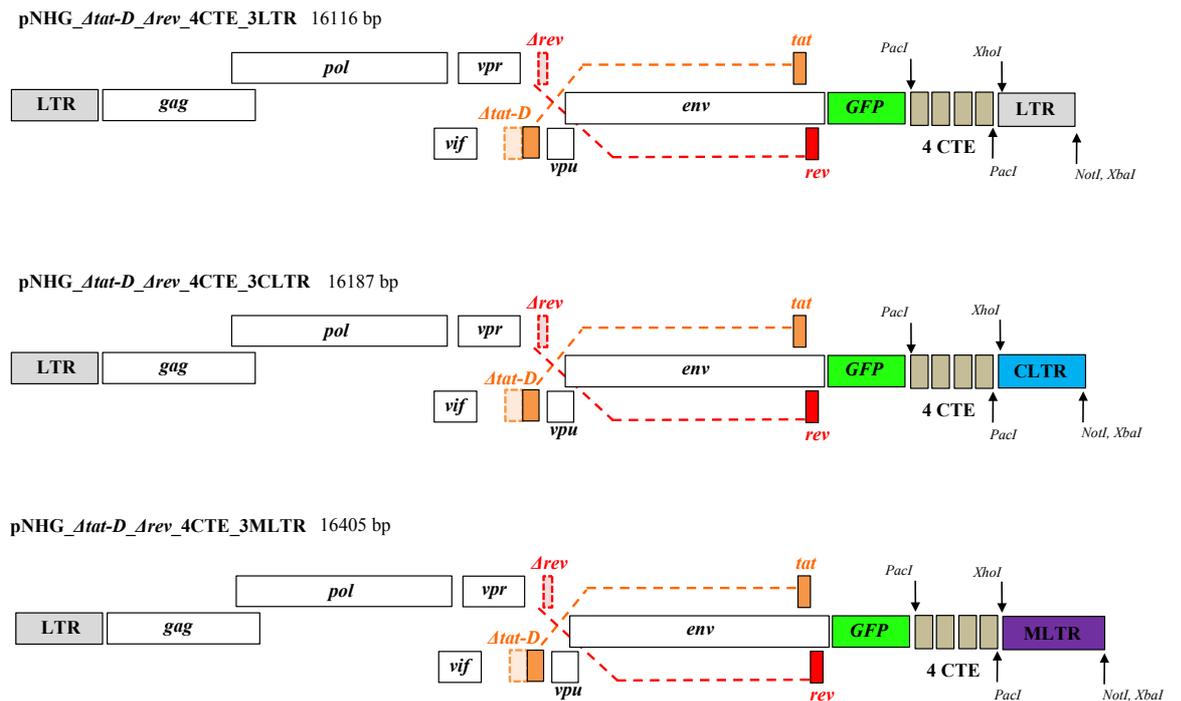
cells to perform viral passages. It was necessary to initially provide *tat* and *rev* genes in *trans* by co-transfection (using DEST40\_*tat* and DEST40\_*rev* respectively) to generate infectious NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ \_4CTE viruses in the first cycle of replication. When these proteins were not provided together in *trans*, no infectious NHG\_ $\Delta tat$ - $D$ - $\Delta rev$ \_4CTE viruses were produced (Figure 6-5).

We transfected the NHG\_Δtat-D-Δrev\_4CTE constructs into HEK-293T cells and collected the viruses 48 hours post-transfection. Prior to attempting the infection of fresh AA-2 cells we filtrated the supernatant from HEK-293T cells with 0.45 μm 25mm diameter syringe filters and then inoculated the viruses into AA-2 cells. We collected a sample of each culture every 24 hours post-infection, and changed the cell culture media to keep cells healthy. This protocol was repeated until the culture was entirely infected and therefore we collected and filtered the viruses from these cultures to inoculate them into fresh AA-2 cultures.

The NHG\_Δtat-D-Δrev-4CTE\_3LTR virus was not infectious despite the addition of the regulatory proteins supplied in *trans*, but NHG\_Δtat-D-Δrev-4CTE\_3CLTR and NHG\_Δtat-D-Δrev-4CTE\_3MLTR were infectious 48 h.p.i. (Figure 6-12, panel B). The infectivity was consistent after 9 d.p.i., but no spreading was observed by this time. We reasoned that it was possible that the initial viral inoculum did not reach sufficient levels of infection to have adequate complexity from which to select fitter variants, which could have affected the spreading potential of NHG\_Δtat-D-Δrev-4CTE\_3CLTR and NHG\_Δtat-D-Δrev-4CTE\_3MLTR viruses. We decided to generate these latter viruses in 10 cm dishes (in duplicate) by transfection of HEK-293T cells to generate a more diverse HIV-1 population and increase the viral titre before performing spreading/passage experiments. By generating more viruses, we would increase the diversity, which could potentially facilitate the selection of fitter derivatives. The NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus was concentrated ~20 times by benchtop centrifugation and inoculated onto AA-2 cells, producing a higher percentage of infected cells than the previous experiment (Figure 6-12, panel C). In addition, GFP positive cells were detected 9 d.p.i. with the NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus (Figure 6-13, panel A). However, without further experimentation we are unable to conclude whether this represents initial Tat and Rev-dependent infection (from the inoculum) or resulted from limited spreading of the virus. Thus, the level of infection was too low to determine the viral titre or p24 expression from infected cells.

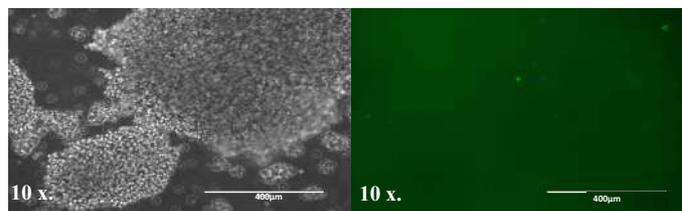
It remains possible that the NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus was able to inefficiently replicate in a Tat and Rev independent manner in human cells, but we observed a decay in GFP expression over time indicating the virus was either replication incompetent or severely attenuated (Figure 6-13, panel B).

A



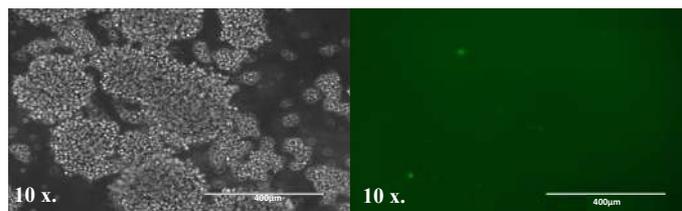
B

NHG\_Δtat-D\_Δrev\_4CTE\_3CLTR



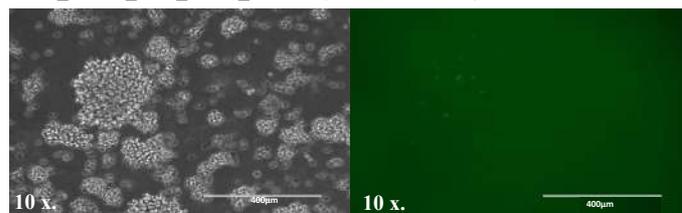
4 d.p.i.

NHG\_Δtat-D\_Δrev\_4CTE\_3MLTR



C

NHG\_Δtat-D\_Δrev\_4CTE\_3CLTR (20x concentrated)

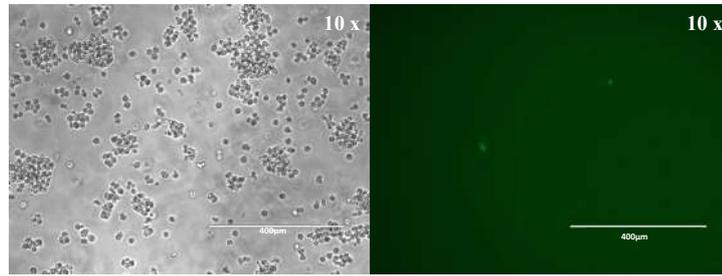


9 d.p.i.

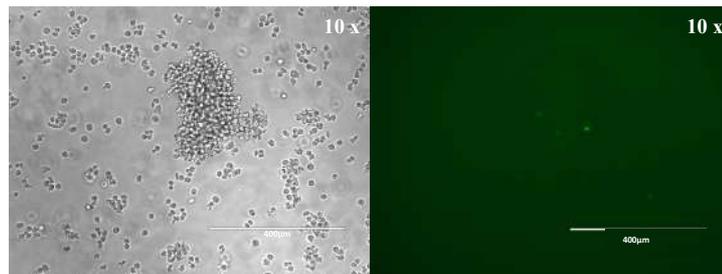
**Figure 6-12.** (A) Maps of the replication competent constructs NHG\_Δtat-D\_Δrev\_4CTE viruses encoding a 43 nt deletion in *tat* exon 1 (Δtat-D) and a deletion of the first exon of *rev* (Δrev), 4CTEs near the 3' end and chimeric LTRs (CLTR or MLTR). (B) GFP expression of NHG\_Δtat-D\_Δrev\_4CTE viruses in AA-2 cells 4 d.p.i. These viruses were generated in HEK-293T cells by transfection (1 μg) and collected and purified by sucrose cushion centrifugation 48 h.p.t. Then, the resultant viruses were inoculated into AA-2 cells to start the passages (1 ml for each virus). Panel C shows the NHG\_Δtat-D\_Δrev\_4CTE\_3CLTR virus infection when concentrated 20 times by centrifugation before the inoculum into AA-2 cells. However, the level of infection of NHG\_Δtat-D\_Δrev\_4CTE viruses were too low to submit the AA-2 culture to flow cytometry analysis.

A

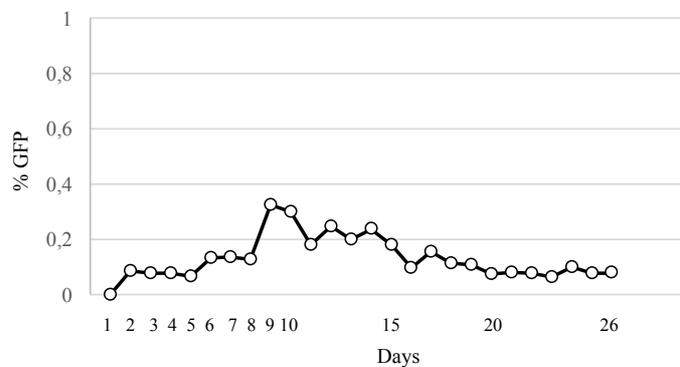
NHG\_Atat-D\_Arev\_4CTE\_3CLTR



12 d.p.i.



B



DAY	GFP%
Day 1	0
Day 2	0,08648649
Day 3	0,07619462
Day 4	0,0764693
Day 5	0,06713662
Day 6	0,13345196
Day 7	0,13480117
Day 8	0,12653859
Day 9	0,32447359
Day 10	0,29943935
Day 11	0,18034265
Day 12	0,24673082
Day 13	0,19936672
Day 14	0,23670794
Day 15	0,18098288
Day 16	0,0973236
Day 17	0,15560812
Day 18	0,11251572
Day 19	0,10731432
Day 20	0,07397362
Day 21	0,07868705
Day 22	0,07654263
Day 23	0,06408016
Day 24	0,09992359
Day 25	0,07820019
Day 26	0,07854985

**Figure 6-13. NHG\_Atat-D\_Arev\_4CTE\_3CLTR adaptation in AA-2 culture.** (A) GFP expression in NHG\_Atat-D\_Arev\_4CTE\_3CLTR virus 12 d.p.i. Few cells were observed, but the level of infection was too low to submit the culture to flow cytometry analysis. (B) Growth curve of NHG\_Atat-D\_Arev\_4CTE\_3CLTR virus along 24 days in culture. GFP percentages are shown in the panel at the right. The timepoint curve from figure B was performed by Victor Iliev, a graduate student working in our lab. This passage was finished after day 26 due to the lost of GFP expression from the culture.

## 6.2 Discussion chapter 6

The NHG\_Δ*tat-F* and NHG\_Δ*tat-D* mutants were unable to generate sufficient transcription levels when Tat was not provided in *trans* during transfection, and we did not find differences in virion production (compared to NHGwt). The infectivity was reduced for all NHG\_Δ*tat-F* and NHG\_Δ*tat-D* clones (Figure 6-6). This result suggested that perhaps the viral RNAs were poorly synthesized (likely due to decreased LTR promoter activity in the absence of Tat-transactivation) in NHG\_Δ*tat-F* and NHG\_Δ*tat-D* mutants and demonstrated that the CLTR and MLTR chimeric promoters did not rescue Tat function during the first cycle of replication. This was entirely expected as in the plasmid, the CLTR and MLTRs are only present at the 3' end and cannot drive transcription in the first cycle of replication. However, the ability of the CLTR and MLTR to drive transcription and reverse transcription efficiently (when Tat was provided in *trans* during transfection) was expected from previous results using the CSGW\_Flexi system (Figure 6-6, panel A). Moreover, the NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR clones replicated efficiently in the absence of Tat after 12 days post-infection, suggesting that the CLTR could be a good candidate to achieve Tat independence (Figure 6-13).

The AA-2 cultures infected with NHG\_Δ*tat-F* viruses (frameshift) could easily revert and generate the *tat* exon 1 wild type sequence. However, the NHG\_Δ*tat-D* (deletion) should not revert as the frameshift mutation (the 43 nucleotides deletion should impair *tat* expression for several cycles of replication). The initial NHG\_Δ*tat-F* and NHG\_Δ*tat-D* viruses employed to perform the passages were sequenced before attempting any experiment, but our results obtained from the TZM assays detected Tat expression in NHG\_Δ*tat-D*\_3LTR, NHG\_Δ*tat-D*\_3CLTR and NHG\_Δ*tat-D*\_3MLTR cultures, which should not express Tat after the 43nt deletion we made (Figure 6-2). These findings suggest that the Tat expression could have originated from cross-contamination from other cultures during passage. However, the NHG\_Δ*tat-F* viruses recovered Tat expression after few rounds of replication and sequencing results confirmed this hypothesis. The frameshift mutation was easy to overcome for the virus, but this should not occur for the deletion mutation from NHG\_Δ*tat-D* clones.

Moreover, several viral passages attempts performed afterwards to avoid cross-contamination events, but unfortunately these passages yielded very similar results,

suggesting that the plasmids employed to generate the NHG\_Δtat-F and NHG\_Δtat-D viruses could be the origin of this contamination even although we did not detect any Tat expression when they were transfected into TZM cells (Figure 6-4). It is remarkable that the NHG\_Δtat-F\_3CLTR and NHG\_Δtat-D\_3CLTR viruses were the only ones where the contamination events were not observed, suggesting that the results obtained from these passages could still be reliable (Figure 6-9). The NHG\_Δtat-F\_3CLTR and NHG\_Δtat-D\_3CLTR clones caused similar levels of infection after several days, suggesting that perhaps the adaptation could be achieved with more time (Van Opijnen et al. 2007). The protracted and unpredictable shutdown periods within the CL3 facilities severely affected the consistency and potential of these viruses, and experiments were cancelled for long periods of time. However, we could resolve this problem by repeating the passages during longer periods of time with no interruptions in the near future and determine if NHG\_Δtat-F\_3LTR, NHG\_Δtat-F\_3MLTR, NHG\_Δtat-D\_3LTR and NHG\_Δtat-D\_3MLTR viruses were able to replicate in a Tat free environment.

The deletion of the first exon of *rev* (Δrev) and the addition of 4CTEs into NHG\_Δtat-F and NHG\_Δtat-D viruses was not as successful as we had hoped from the results obtained using CSGW\_4CTE vector genome in the CL2 laboratory (Figure 5-4). The NHG\_Δtat-D-Δrev-4CTE clones (Figure 6-12, panel A) were infectious at early steps of infection, but unfortunately none of them could progress and spread within the AA-2 cultures. The NHG\_Δtat-D-Δrev-4CTE\_3MLTR was infectious for few days only, suggesting that the MLTR failed to support efficient replication following the first round of infection. This could be due to a failure to drive transcription, reverse transcription, maintain RNA stability or other problems due to the absence of Tat and Rev. However, the NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus was infectious after 9 d.p.i., but did not progress as we expected (Figure 6-13, panel B). Unfortunately, we did not have time to perform cell free passages and determine whether this virus propagated at all.

The initial levels of infection observed from NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus suggested that it may be possible to develop an HIV-1 virus that is able to replicate in a Tat and Rev independent manner in human cells, but there could be several problems likely related to viral transcription and RNA stability that we could not determine. The CLTR seemed to partially overcome Tat function, but the issues related to CL3 refurbishment impaired the progress of the passages and further experiments need to be performed to demonstrate the CLTR ability to replace Tat function. However, the 4CTEs seemed to not

contribute to the RNA stability and transport in NHG\_Δtat-D-Δrev-4CTE\_3LTR, NHG\_Δtat-D-Δrev-4CTE\_3CLTR and NHG\_Δtat-D-Δrev-4CTE\_3MLTR viruses as we observed in CSGW\_4CTE vector genomes (Figure 5-4). The mechanistic basis of these problems are yet to be determined, but future investigation of RNA production and viral transcription could determine what is affecting the potential of our NHG\_Δtat-D-Δrev\_4CTE viruses to adapt and evolve within the culture. We could check this hypothesis developing reverse transcription assays to investigate the reverse transcription levels in NHG\_Δtat-D-Δrev\_4CTE viruses to determine the RNA species within transfected (HEK-293T) and infected cells (AA-2). We could extract the viral RNA from infected AA-2 cells during the passages or adding a reverse transcriptase inhibitor to determine if the virus is spreading in culture or not. The RNA extractions could be also performed for viral particles produced during the first cycle of replication in HEK-293T cells to detect if NHG\_Δtat-D-Δrev\_4CTE viruses are able to package the viral RNA successfully or whether packaging is affected somehow. Moreover, to investigate if the viral RNAs export and stability is affected in NHG\_Δtat-D-Δrev\_4CTE viruses we could have performed northern blot analysis of viral RNAs extracted from transfected cells and compare the amounts of full-length genomes produced in all NHG\_Δtat-D-Δrev\_4CTE viruses. Thus, we could have included a fluorescent dye in viral RNAs to detect the amount of RNA that remain in the nucleus compared to the cytoplasm in transfected cells.

Thus, the absence of *nef* gene in NHG constructs could be also affecting viral replication. Nef protein has well-characterized effects on host membrane trafficking and receptor downregulation, increasing the intrinsic infectivity of HIV-1 virions in a host-cell-dependent manner. However, the addition of *nef* into NHG constructs will increase the size of the genome and we could introduce additional RNA secondary structures that could disrupt viral RNA packaging. The GFP reporter gene is replacing *nef* in all NHG constructs, therefore we need to preserve this gene to measure viral infectivity and *in vitro* evolution during the passages. We could include *nef* gene in NHG constructs to enhance viral infection in human or murine cells to facilitate the *in vitro* adaptation (passages) of our NHG\_Δtat-D-Δrev-4CTE constructs. Published reports demonstrate that the negative effects of Nef expression are not present when Nef is expressed in T cells from transgenic mice, which moderate positive T cell activation process (Rhee et al. 1994).

Unfortunately, the NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus was not sufficient to generate Tat and Rev independent progeny, but the findings from this thesis suggest that this

ambitious aim could be achievable. Moreover, we could investigate the Tat and Rev independence using CL3 clones independently: investigate Tat independence in  $\Delta tat$  mutants encoding a CLTR (NHG\_ $\Delta tat$ -D-3CLTR) and investigating Rev independence using  $\Delta rev$  mutants encoding 4CTEs (NHG\_ $\Delta tat$ -D- $\Delta rev$ -4CTE\_3CLTR). The HIV-1 regulatory genes (*tat* and *rev*) are essential for viral survival, and overcoming both the function of both genes together could be impossible unless we understand better how we can overcome their relevant function with our CL3 viruses. Perhaps the aim of this thesis was too ambitious to be achieved in the period of time that we had. However, the findings obtained from this thesis suggest that we could focus in transcriptomics and RNA expression from NHG\_ $\Delta tat$ -D- $\Delta rev$ -4CTE\_3CLTR and NHG\_ $\Delta tat$ -D viruses, which could reveal new insights and maybe identify the factors that impair viral replication. In conclusion, supporting more intense research to investigate why the NHG\_ $\Delta tat$ -D- $\Delta rev$ -4CTE\_3CLTR failed to spread in human cells could overcome the issues that impair NHG\_ $\Delta tat$ -D- $\Delta rev$ -4CTE\_3CLTR replication.

# **Chapter 7**

## **Final reflections**

## 7.1 Final reflections

### 7.1.1 Summary

The primary aim of this project was to develop an HIV-1 clone able to replicate in mice, with the future aim of the generating an HIV-1 clone to assist in antiviral drug development. The potential of CLTR and MLTR promoters to generate infectious pseudotyped HIV-1 particles in human cells was successful, but attempts to reproduce this result in murine cells failed and opened up many questions that need to be investigated.

The absence of Rev protein in our CSGW\_Flexi and SYNGP constructs could be affecting full length viral RNA transport towards the cytoplasm, which could remain within the nucleus. The RNA stability of full-length viral and single spliced RNAs in the absence of Rev could contribute to their degradation (Blissenbach et al. 2010). Perhaps the modifications introduced by introducing the CLTR and MLTR into CSGW\_Flexi clones are affecting the viral RNA secondary structure, which is required to interact with the Gag protein to package the viral genome into the nascent particles (Amarasinghe et al. 2001). Alternatively, the viral genome is efficiently generated within the transfected cell, but it could be unable to be packaged within the nascent viral particles (Mailler et al. 2016). The 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR construct generated VLPs in human cells but they were not infectious. The RNA secondary structures in the 5CLTR\_SYNGP\_Ψ\_IRES\_EGFP\_3CLTR construct introduced by the addition of the CLTRs and other elements could be affecting the RNA stability and affecting the packaging of the viral genome into the nascent particles (Watts et al. 2009; Sükösd et al. 2015). Thus, the SYNGP\_4CTE clones were defective and unable to generate infectious viruses despite the higher MFI detected in transfected HEK-293T cells shown in Figure 4-7, panel B. No evidence was observed to suggest issues with VLP formation for SYNGP\_4CTE clones, but they could lack the viral genome or could be aberrant in some other way (Lu et al. 2011). Moreover, we detected some p25 intermediate protein in western blot analysis of VLPs from all SYNGP\_4CTE clones (Figure 4-10, panel B). In addition, reducing the CpG content from HIV-1 codon-optimized genes (from SYNGP vector) could have affected RNA stability or folding, generating novel RNA secondary structures that impair viral replication.

The length of Flexi\_GP\_4CTE clones should not impair viral replication, and the viral RNA genome of Flexi\_GP\_4CTE clones should be able to be packaged as CSGW\_4CTE

does. Surprisingly, the addition of Tat and Rev proteins in *trans* did not rescue viral infectivity from CSGW\_GP\_4CTE\_3LTR, 5CLTR\_CSGW\_GP\_4CTE\_3CLTR and 5MLTR\_CSGW\_GP\_4CTE\_3MLTR, suggesting that the viral genome must be unable to interact with HIV-1 regulatory proteins due to an aberrant RNA folding, impairing vital interactions along the viral replication cycle, such as Gag with full length genomes, or highly conserved regions in viral RNA could be also affected by introducing the *gag-pol* genes into Flexi\_GP\_4CTE clones (Le et al. 1990; Watts et al. 2009). However, we successfully generated infectious particles using the CSGW\_4CTE construct (when pSYNGP and pCMV\_VSV-G were provided in *trans*) in a Tat and Rev independent manner from HEK-293T cells (Figure 5-4). When we tried to infect murine macrophages (RAW) with these viruses, which were generated in human cells, we only observed a few infected cells 72 hours post-infection (Figure 5-3). These data suggest that replacing Rev function by 4 copies of the CTE is not sufficient to rescue the infectivity within murine cells (Figure 5-3). The Flexi\_GP\_4CTE vector encodes the RRE outside the *env* gene near the 3' end (its natural position), but the ability of wild-type CSGW to produce infectious viruses when SYNGP/CVR1\_NLGP and CMV\_VSV-G are provided during transfection supports the idea that the position of the RRE is not essential for RNA transport or stability.

Finally, the infectivity was reduced for all NHG\_Δ*tat-F* and NHG\_Δ*tat-D* clones compared to wild-type NHG (Figure 6-6). These data suggested that perhaps the viral RNAs were poorly synthesized (likely due to decreased LTR promoter activity in the absence of Tat-transactivation) in NHG\_Δ*tat-F* and NHG\_Δ*tat-D* mutants and demonstrated that the CLTR and MLTR chimeric promoters did not rescue Tat function during the first cycle of replication. Moreover, the NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR clones replicated efficiently in the absence of Tat after 12 days post-infection, suggesting that the CLTR could be a good candidate to achieve Tat independence (Figure 6-13). The NHG\_Δ*tat-F*\_3CLTR and NHG\_Δ*tat-D*\_3CLTR clones caused similar levels of infection after several days, suggesting that perhaps the adaptation could be achieved with more time (Berkhout et al. 2007). The deletion of the first exon of *rev* (Δ*rev*) and the addition of 4CTEs into NHG\_Δ*tat-F* and NHG\_Δ*tat-D* viruses was not as successful as we had hoped from the results obtained using CSGW\_4CTE vector genome in the CL2 laboratory (Figure 5-4). The NHG\_Δ*tat-D*-Δ*rev*-4CTE clones (Figure 6-12, panel A) were infectious at early steps of infection, but unfortunately none of them could progress and spread within

the AA-2 cultures. The NHG\_Δtat-D-Δrev-4CTE\_3MLTR was infectious for few days only, suggesting that the MLTR failed to support efficient replication following the first round of infection. This could be due to a failure to drive transcription, reverse transcription, maintain RNA stability or other problems due to the absence of Tat and Rev. The mechanistic basis of these problems are yet to be determined, but future investigation of RNA production and viral transcription could determine what is affecting the potential of our NHG\_Δtat-D-Δrev\_4CTE viruses to adapt and evolve within the culture.

Unfortunately, the NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus was not sufficient to generate Tat and Rev independent progeny, but the findings from this thesis suggest that this ambitious aim could be achievable and perhaps the aim of this thesis was too ambitious to be achieved in the period of time that we had. Unfortunately, the NHG\_Δtat-D-Δrev-4CTE\_3CLTR virus was not sufficient to generate Tat and Rev independent progeny and failed after several days of culture (and perhaps only generated one round of infectious particles from transient transfection when Tat was added in *trans*). The findings from this thesis suggest that Tat and Rev independent virus could be achievable and perhaps, with modifications, the NHG\_Δtat-D-Δrev-4CTE\_3CLTR clone could become a future candidate for HIV research.

### 7.1.2 Potential for future research

Moving forward, we could investigate the Tat and Rev independence using the CL3 clones independently: investigate Tat independence in *tat* mutants encoding a CLTR (NHG\_Δtat-D\_3CLTR) and investigating *rev* independence using Rev mutants encoding 4CTEs (NHG\_Δtat-D-Δrev-4CTE\_3CLTR). The HIV-1 regulatory genes (*tat* and *rev*) are essential for viral survival, and overcoming both gene functions together could be a big genetic barrier to overcome simultaneously. Unless we understand how we can overcome their relevant function within our CL3 viruses. Perhaps the aim of this thesis was too ambitious to be achieved by the period of time that we had and also for the challenging aim of the proposal. For example, several avenues were not examined due to time constraints. Although HIV-1 cannot ‘utilise’ murine cyclin T1 (Tat) or mCRM1(Rev) it is likely that HIV-1 could adapt to the murine orthologue if pressured to do so. With more time, I would have knocked out these factors in AA2 cells using CRISPR CAS9 technology and reconstituted their function with a (CRISPR resistant) murine orthologue. Perhaps HIV-1

could adapt to utilise these individual murine orthologues and this could be a platform to investigate further blocks to HIV-1 replication in murine cells.

The potential for future research could approach investigations to determine the essential role of the HIV-1 RNA folding, encoding highly variable regions but also highly conserved structures with unknown functions. Secondary structure studies have determined that the HIV-1 packaging signal contains several RNA loops (SL1 to SL4) involved in packaging the viral genome into the particles by Gag interactions that could be altered in the 5'CLTR\_Ψ\_SYNGP\_IRES-GFP\_3'CLTR clone (Greatorex et al. 2002). The codon-optimization of *gag-pol* genes contributes to the generation of more RNA species within human cells, but maybe the modifications made to reach this aim introduced novel RNA loops into the viral genome that could impair their packaging into the nascent particles. The compactly folded tertiary structure of viral RNA is essential for the viability of the virus at distinct steps along the replication cycle, such as reverse transcription, viral transcription or packaging (Berkhout et al. 2000). The RNA secondary structures are essential for the packaging of the viral genome within the viral particles. We unintentionally introduced more unknowns by reducing the CpG content in the codon-optimized *gag-pol* genes that affected viral replication.

This aspect suggests that the absence of infectivity could be also related to the disruption in the RNA folding in the viral genome introduced by codon optimization of *gag-pol* genes, or the higher number of CpGs in SYNGP compared to NLGP (226 to 23 respectively), which could impair viral replication kinetics (Atkinson et al. 2014; Gaunt et al 2016). We could have performed reverse transcription assays to quantify the reverse transcription levels in our CSGW\_Flexi and SYNGP clones and investigate the RNA species within transfected cells to determine viral transcription more accurately. Thus, we could have investigated if the generated VLPs had a proper morphology because it is not possible to determine this feature by western blot analysis, so we could have prepared some purified VLPs for each CSGW\_Flexi and SYNGP clones to submit to electron microscopy and approach this hypothesis. To elucidate the likely problems with full-length RNA packaging/production we could investigate the RNA production within transfected cells by RNA extraction and subsequently northern blot analysis.

Our hypothesis will focus on viral RNA production, stability and transport within transfected cells to determine the ability of the constructs developed in this thesis to identify

the problems of infectivity or packaging. We could investigate whether the viral RNAs remain stucked in the nucleus performing RNA extractions from transfected cells with SYNGP\_4CTE clones (with 4CTEs in *cis*) and compare the amounts of full-length genomes produced (northern blot) or include a fluorescent dye in viral RNAs to detect the amount of RNA that remain in the nucleus compared to the cytoplasm in transfected cells. The VLPs generated from SYNGP\_4CTE clones could lack the viral genome, so we could have performed RNA extractions from transfected cells to investigate the differences between RNA species for each CSGW\_Flexi and SYNGP clones. Thus, we could have also investigated the RNA content within the VLPs to detect if the viral genome is able to be packaged into nascent particles or perhaps can not interact with the packaging signal. Several experiments were performed to investigate the RNA content in VLPs by qPCR.

We could have performed RNA extractions from transfected HEK-293T with CSGW\_4CTE clone to investigate the RNA expression levels compared to wild-type CSGW and also in VLPs by qPCR assays. Moreover, we tried several qPCR assays to investigate this hypothesis but the presence of background in all samples including the negative control discarded these results and still not clear yet if the 4CTEs encoded in CSGW\_4CTE construct is substituting Rev efficiently as we suggest with our results. However, the potential of CSGW\_4CTE construct represents the best achievement we obtained at this point, generating infectious particles in a Tat and Rev independent manner. The CSGW\_4CTE is a great candidate to develop an HIV-1 clone able to replicate in mice, but we need to investigate how this clone overcomes the absence of the regulatory proteins. We should be able to reproduce the results of CSGW\_4CTE clone generating infectious particles in the absence of Tat and Rev in murine cells, but urther investigations need to be done to determine the features of this construct in human cells. The RRE from CSGW\_4CTE clone is not located at the 3' end, but previous experiments demonstrated the ability of CSGW\_4CTE to generate infectious viral particles in the absence of Tat and Rev (Figure 5-4). However, we could transfer the RRE into its correct position (at the end of 3' transcripts) in CSGW\_4CTE and repeat the same experiments to determine if this modification enhances viral infectivity or maybe contributes to the RNA transport in a Rev-independent manner.

The reverse transcription assays could be also developed to investigate the reverse transcription levels in NHG\_Δtat-D-Δrev\_4CTE viruses to determine the RNA species within transfected (HEK-293T) and infected cells (AA-2). We could extract the viral RNA

from infected AA-2 cells during the passages or adding a reverse transcriptase inhibitor to determine if the virus is spreading in culture or not. The RNA extractions could be also performed for viral particles produced during the first cycle of replication in HEK-293T cells to detect if NHG\_Δtat-D-Δrev\_4CTE viruses are able to package the viral RNA successfully or whether packaging is affected somehow. Moreover, to investigate if the viral RNAs export and stability is affected in NHG\_Δtat-D-Δrev\_4CTE viruses we could have performed northern blot analysis of viral RNAs extracted from transfected cells and compare the amounts of full-length genomes produced in all NHG\_Δtat-D-Δrev\_4CTE viruses. Thus, we could have included a fluorescent dye in viral RNAs to detect the amount of RNA that remain in the nucleus compared to the cytoplasm in transfected cells.

Nonetheless, although Tat-responsivity was maintained, CLTR and MLTR containing viruses were viable (partially substituting for Tat function) and the MPMV CTEs improved the single cycle HIV-1 systems, our findings suggest that assisted evolution could one day pave the way for an HIV-1 virus that replicates in murine cells. However, the success about HIV-1 curative therapies involves the usage of animal models, but their current limitations severely impede the potential of vaccine development to cure HIV infection. Perhaps we could transfer the discoveries from the ‘cured’ patient into animal models to investigate and reveal new insights about a cure that we could export to humans. The current antiretroviral treatments fail to eradicate HIV infection in patients despite the combination of different drugs targeting several proteins of the virus. We could solve these limitations at long-term if we investigate the CSGW\_4CTE construct ability to produce infectious particles in the absence of HIV regulatory proteins. We could transfer the features from this construct to a replication competent vector such as the NHG\_Δtat-D-Δrev\_4CTE\_3CLTR construct to adapt this virus to human cultures before attempting murine cells. By developing a HIV-1 clone able to replicate in murine cells in the absence of Tat and Rev proteins will generate a huge contribution of the usage of this virus in animal models that won’t need to be ‘humanized’ (humanized mice) and adapted to HIV-1 infection (current macaque animal models). Thus, extrapolating results obtained in current animal models to human disease is often not straightforward, and HIV-1 is highly specific for humans and faces multiple barriers to replication in most other animals. The development of strains that more accurately reflect the features of the virus responsible for the AIDS pandemic would overcome many of the remaining limitations of current models, and it is possible that no animal model will ever fully capture all of the features of human HIV-1 infection (Hatzioannou et al. 2012).

---

Developing an adapted HIV-1 clone for mice will be a great achievement to investigate curative therapies to eradicate HIV-1 from infected patients. However, there is still a long way to go and much that needs to be performed to achieve this aim.

# **Chapter 8**

## **References**

## 8.1 References

- Aiken, C. (1997). Pseudotyping Human Immunodeficiency Virus Type 1 ( HIV-1 ) by the Glycoprotein of Vesicular Stomatitis Virus Targets HIV-1 Entry to an Endocytic Pathway and Suppresses both the Requirement for Nef and the Sensitivity to Cyclosporin A, *71*(8), 5871–5877.
- Akkina R. (2013). New generation humanized mice for virus research: comparative aspects and future prospects. *Virology*. 2013 January 5; 435(1): 14–28. doi:10.1016/j.virol.2012.10.007.
- Aligeti, M., Behrens, R. T., Pocock, G. M., Schindelin, J., Dietz, C., Eliceiri, K. W., ... Sherer, N. M. (2014). Cooperativity among Rev-Associated Nuclear Export Signals Regulates HIV-1 Gene Expression and Is a Determinant of Virus Species, *88*(24), 14207–14221. <https://doi.org/10.1128/JVI.01897-14>
- Allers, K., Hu, G., Lodenkemper, C., Rieger, K., Thiel, E., & Schneider, T. (2018). Evidence for the cure of HIV infection by CCR5 Δ32 / Δ32 stem cell transplantation, *117*(10), 2791–2800. <https://doi.org/10.1182/blood-2010-09-309591>.An
- Amarasinghe, G., & Zhou, J. (2001). Stem-loop SL4 of the HIV-1 Ψ RNA packaging signal exhibits weak affinity for the nucleocapsid protein. Structural studies and implications for genome recognition. *Journal of Molecular Biology*, *314*(5), 961–970. <https://doi.org/10.1006/jmbi.2001.5182>
- Ambrose Z. and Aiken C. (2014). HIV-1 Uncoating: Connection to Nuclear Entry and Regulation by Host Proteins. *Virology*. 2014 April ; 0: 371–379. doi:10.1016/j.virol.2014.02.004.
- Anderson, J. L. & Hope, T. J. (2008). APOBEC3G restricts early HIV-1 replication in the cytoplasm of target cells. *Virology* *375*, 1–12. (doi: 10.1016/j.virol.2008.01.042)
- Andrew Rambaut, David Posada, Keith A. Crandall & Edward C. Holmes (2004). The causes and consequences of HIV evolution. *Nature Reviews Genetics* *5*, 52-61 (January 2004) | doi:10.1038/nrg1246
- Arhel NJ, Souquere-Besse S, Munier S, Sou- que P, Guadagnini S, Rutherford S, Prevost MC, Allen TD, Charneau P (2007). HIV-1 DNA Flap formation promotes uncoating of the pre- integration complex at the nuclear pore. *EMBO Journal* 2007, *26*:3025-3037.
- Atkinson, N. J., Witteveldt, J., Evans, D. J., & Simmonds, P. (2014). The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. *Nucleic Acids Research*, *42*(7), 4527–4545. <https://doi.org/10.1093/nar/gku075>
- Balachandran, A., Wong, R., Stoilov, P., Pan, S., Blencowe, B., Cheung, P., ... Cochrane, A. (2017). Identification of small molecule modulators of HIV-1 Tat and Rev protein accumulation. *Retrovirology*, *14*(1), 7. <https://doi.org/10.1186/s12977-017-0330-0>

- Baltimore D. (1970). Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses. *Nature* 226, 1209 - 1211 (27 June 1970); doi:10.1038/2261209a0
- Barbara J. Winslow and Didier Trono (1993). The Blocks to Human Immunodeficiency Virus Type 1 Tat and Rev Functions in Mouse Cell Lines Are Independent. *Journal of Virology*, 2349-2354
- Barbara Müller, Maria Anders, Hisashi Akiyama, Sonja Welsch, Bärbel Glass, Krisztina Nikovics, Francois Clavel, Hanna-Mari Tervo, Oliver T. Keppler, and Hans-Georg Kräusslich (2009). HIV-1 Gag Processing Intermediates Trans-dominantly Interfere with HIV-1 Infectivity. *Journal of Biological Chemistry*, 284(43), 29692–29703. <https://doi.org/10.1074/jbc.M109.027144>
- Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L.(1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983 May 20;220(4599):868-71.
- Bartee E, McCormack A, Fruh K. (2006). Quantitative membrane proteomics reveals new cellular targets of viral immune modulators. *PLoS Pathogens* 2: e107. doi: 10.1371/journal.ppat.0020107.
- Barton F. Haynes, Dennis R. Burton (2017). Developing an HIV vaccine. *Science* 17 Mar 2017:Vol. 355, Issue 6330, pp. 1129-1130 DOI: 10.1126/science.aan0662
- Bartz SR and Vodicka MA (1997). Production of high-titer human immunodeficiency virus type 1 pseudotyped with vesicular stomatitis virus glycoprotein. *Methods*. 1997 Aug;12(4):337-42.
- Bednarik DP., Cook JA . and PM Pitha (1990). Inactivation of the HIV LTR by DNA CpG methylation: evidence for a role in latency. *EMBO Journal* vol. 9 no.4 pp. 1157-1164.
- Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Pac, J., Burt, A., & Tristem, M. (2004). Long-term reinfection of the human genome by endogenous retroviruses. *Proc Natl Acad Sci U S A*. 2004 Apr 6;101(14):4894-9. Epub 2004 Mar 25.
- Belzile JP, Duisit G, Rougeau N, Mercier J, Finzi A, Cohen EA (2007) HIV-1 Vpr-mediated G2 arrest involves the DDB1-CUL4AVPRBP E3 ubiquitin ligase. *PLoS Pathogens* 2007, 3:e85.
- Bentham M, Mazaleyrat S and Harris M (2006). Role of myristoylation and N-terminal basic residues in membrane association of the human immunodeficiency virus type 1 Nef protein. *Journal of General Virology*. 2006 Mar;87(Pt 3):563-71.
- Berkhout B., Gatignol A., Rabson AB, KT Jeang (1990). TAR-independent activation of the HIV-1 LTR: Evidence that Tat requires specific regions of the promoter. *Cell*, Volume 62, Issue 4, p757–767. [http://dx.doi.org/10.1016/0092-8674\(90\)90120-4](http://dx.doi.org/10.1016/0092-8674(90)90120-4)
- Berkhout, B. E. N., & Wamel (2000). The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure, *RNA*, 282–295.

- Berkhout, B., Vastenhouw, N. L., Klasens, B. I., & Huthoff, H. (2001). Structural features in the HIV-1 repeat region facilitate strand transfer during reverse transcription. *RNA*, 7(8), 1097–1114. <https://doi.org/10.1017/S1355838201002035>
- Bester, A. C., Kafri, M., Maoz, K., & Kerem, B. (2013). Infection with retroviral vectors leads to perturbed DNA replication increasing vector integrations into fragile sites. *Scientific Reports*, 3. <https://doi.org/10.1038/srep02189>
- Bieniasz, P. D., Grdina, T. A., Bogerd, H. P., & Cullen, B. R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. *The EMBO Journal*, 17(23), 7056–65. <https://doi.org/10.1093/emboj/17.23.7056>
- Bif, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T. *et al.* (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341: 1233158.
- Biswas P., Jiang X., Annmarie L. Pacchia, Joseph P. Dougherty and SW. Peltz (2004). The Human Immunodeficiency Virus Type 1 Ribosomal Frameshifting Site Is an Invariant Sequence Determinant and an Important Target for Antiviral Therapy. *Journal of Virology* 78(4): 2082–2087.
- Blazkova, J., Trejbalova, K., Halfon, P., Guiguen, A., Verdin, E., Olive, D., ... Hirsch, I. (2009). CpG Methylation Controls Reactivation of HIV from Latency, 5(8). *Science* . 2016 May 13; 352(6287): 828–833. doi:10.1126/science.aae0474.
- Blissenbach, M., Grewe, B., Hoffmann, B., Brandt, S., & Klaus, U. (2010). Nuclear RNA Export and Packaging Functions of HIV-1 Rev Revisited, *Journal of Virology*, 84(13), 6598–6604. <https://doi.org/10.1128/JVI.02264-09>
- Bohne, J., Schambach, A., & Zychlinski, D. (2007). New way of regulating alternative splicing in retroviruses: the promoter makes a difference. *Journal of Virology*, 81(7), 3652–6. <https://doi.org/10.1128/JVI.02105-06>
- Bokhoven, M., Stephen, S. L., Knight, S., Gevers, E. F., Robinson, I. C., Takeuchi, Y., & Collins, M. K. (2009). Insertional Gene Activation by Lentiviral and Gammaretroviral Vectors. *Journal of Virology*, 83(1), 283–294. <https://doi.org/10.1128/JVI.01865-08>
- Bokhoven, M., Stephen, S. L., Knight, S., Gevers, E. F., Robinson, I. C., Takeuchi, Y., & Collins, M. K. (2009). Insertional gene activation by lentiviral and gammaretroviral vectors. *Journal of Virology*, 83(1), 283–94. <https://doi.org/10.1128/JVI.01865-08>
- Booth, D. S., Cheng, Y., & Frankel, A. D. (2014). The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. *eLife*, 3, e04121. <https://doi.org/10.7554/eLife.04121>
- Bordería, A. V., Lorenzo-Redondo, R., Pernas, M., Casado, C., Alvaro, T., Domingo, E., & Lopez-Galindez, C. (2010). Initial fitness recovery of HIV-1 is associated with quasispecies heterogeneity and can occur without modifications in the consensus sequence. *PLoS ONE*, 5(4). <https://doi.org/10.1371/journal.pone.0010319>

- Borkow G. (2005) Mouse Models for HIV-1 Infection. *IUBMB Life*, 57(12): 819 – 823, December 2005
- Bordería, A. V., Lorenzo-Redondo, R., Pernas, M., Casado, C., Alvaro, T., Domingo, E., & Lopez-Galindez, C. (2010). Initial fitness recovery of HIV-1 is associated with quasispecies heterogeneity and can occur without modifications in the consensus sequence. *PLoS ONE*, 5(4). <https://doi.org/10.1371/journal.pone.0010319>
- Bozzette SA, Ake CF, Tam HK, Chang SW, Louis TA (2003). Cardiovascular and cerebrovascular events in patients treated for human immunodeficiency virus infection. *New England Journal of Medicine*. 2003;348(8):702–710.
- Brehm MA, Shultz LD, Luban J. and Greiner DL (2013). Overcoming Current Limitations in Humanized Mouse Research. *Journal of Infectious Diseases* 2013;208(S2):S125–30: doi: 10.1093/infdis/jit319
- Briggs DR, Tuttle DL, Sleasman JW, Goodenow MM (2000). Envelope V3 amino acid sequence predicts HIV-1 phenotype (co-receptor usage and tropism for macrophages). *AIDS* 2000; 14:2937–9.
- Bukrinsky MI, Haffar OK (1998). HIV-1 nuclear import: matrix protein is back on center stage, this time together with Vpr. *Molecular Medicine* 1998, 4:138- 143.
- Canard, B., Sarfati, R., & Richardson, C. C. (1997). Binding of RNA template to a complex of HIV-1 reverse transcriptase/primer/template. *Proc Natl Acad Sci USA*, 94(21), 11279–11284. <https://doi.org/10.1073/pnas.94.21.11279>
- Carthagena L., Parise MC, Ringeard M., Chelbi-Alix MK, Hazan U. and Nisole S. (2008) Implication of TRIMalpha and TRIMCyp in interferon-induced anti-retroviral restriction activities. *Retrovirology* 2008, 5:59 doi:10.1186/1742-4690-5-59
- Cesana, D., Ranzani, M., Volpin, M., Bartholomae, C., Duros, C., Artus, A., ... Montini, E. (2014). Uncovering and Dissecting the Genotoxicity of Self-inactivating Lentiviral Vectors In Vivo, 22(4), 774–785. <https://doi.org/10.1038/mt.2014.3>
- Clapham, P. R., Blanc, D., and Weiss, R. A. (1991) Specific cell surface requirements for the infection of CD4-positive cells by human immunodeficiency virus types 1 and 2 and by Simian immunodeficiency virus. *Virology* 181, 703–715.
- Clever, J. L., Miranda, D., & Parslow, T. G. (2002). RNA structure and packaging signals in the 5' leader region of the human immunodeficiency virus type 1 genome. *Journal of Virology*, 76(23), 12381–12387.
- Cmarko, D., Bøe, S., Scassellati, C., Szilvay, A. M., Davanger, S., Fu, X. Dong, Haukenes, G., Kalland, KH. And S. Fakan, S. (2002). Rev Inhibition Strongly Affects Intracellular Distribution of Human Immunodeficiency Virus Type 1 RNAs. *Journal of Virology*, 76(20), 10473–10484. <https://doi.org/10.1128/JVI.76.20.10473>
- Cochrane AW, Perkins A, Rosen CA. (1990) Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. *Journal of Virology* 1990; 64: 881-5.

- Coffin JM (1979). Structure, Replication, and Recombination of Retrovirus Genomes: Some Unifying Hypotheses. (1977). *Cancer Research*, 1–26.
- Cohen CJ, Lock WM, Mager DL. (2009). Endogenous retroviral LTRs as promoters for human genes: a critical assessment. *Gene* 448: 105–14.
- Conte, M. R., & Matthews, S. (1998). Retroviral matrix proteins: a structural perspective. *Virology*, 246(2), 191–198. <https://doi.org/10.1006/viro.1998.9206>
- Conti L, Rainaldi G, Matarrese P, Varano B, Rivabene R, Columba S, Sato A, Belardelli F, Malorni W, Gessani S (1998). The HIV-1 vpr protein acts as a negative regulator of apoptosis in a human lymphoblastoid T cell line: possible implications for the pathogenesis of AIDS. *Journal of Experimental Medicine* 198, 187:403-413.
- Courgnaud, V., M. Salemi, X. Pourrut, E. Mpoudi-Ngole, B. Abela, P. Auzel, F. Bibollet-Ruche, B. Hahn, A.-M. Vandamme, E. Delaporte, and M. Peeters (2002). Characterization of a novel simian immunodeficiency virus with a *vpu* gene from greater spot-nosed monkeys (*Cercopithecus nictitans*) provides new insights into simian/human immunodeficiency virus phylogeny. *Journal of Virology* 76:8298–8309.
- Cron RQ, Bartz SR, Clausell A, Bort SJ, Klebanoff SJ, Lewis DB. (2000). NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells. *Clinical Immunology* 2000;94:179–91. [PubMed: 10692237]
- Cronin, J., Zhang, X.-Y., & Reiser, J. (2005). Altering the tropism of lentiviral vectors through pseudotyping. *Current Gene Therapy*, 5(4), 387–398. <https://doi.org/10.2174/1566523054546224>
- Cullen, B. R. (2003). Nuclear mRNA export: insights from virology. *Biochemical Science* 28:419–424.
- Czernilofsky, A. P., A. D. Levinson, H. E. Varmus, J. M. Bishop, E. Tischer, and H. M. Goodman (1983). Corrections to the nucleotide sequence of the *src* gene of Rous sarcoma virus. *Nature* 301:736–738.
- Chang LJ, McNulty E and M Martin (1993). Human immunodeficiency viruses containing heterologous enhancer/promoters are replication competent and exhibit different lymphocyte tropisms. *Journal of Virology*, 67(2):743
- Chang LJ, McNulty E. and M Martin (1993). Human immunodeficiency viruses containing heterologous enhancer/promoters are replication competent and exhibit different lymphocyte tropisms. *Journal of Virology*, 67(2):743
- Charnay, N., et al. (2009). Mechanism of HIV-1 Tat RNA translation and its activation by the Tat protein. *Retrovirology* 6:74.
- Checkley, M. A., Luttge, B. G., & Freed, E. O. (2011). *Journal of Molecular Biology* 2011 July 22; 410(4): 582–608. doi:10.1016/j.jmb.2011.04.042.
- Chen, J., Powell, D., & Hu, W.-S. (2006). High frequency of genetic recombination is a common feature of primate lentivirus replication. *Journal of Virology*, 80(19), 9651–8. <https://doi.org/10.1128/JVI.00936-06>

- Cherepanov, P., & Engelman, A. (2014). Retroviral Integrase Structure and DNA Recombination Mechanism. *Microbiology Spectrum*, 2(6), 1–22. <https://doi.org/10.1128/microbiolspec.MDNA3-0024-2014>
- Chiu, Y. L., et al. (2002). Tat stimulates cotranscriptional capping of HIV mRNA. *Molecular Cell* 10:585–597
- Dahiya, S., Liu, Y., Pirrone, V., & Michael, R. and B. Wigdahl (2014). Role of Downstream Elements in Transcriptional Regulation of the HIV-1 Promoter, 1(2), 1–7. <https://doi.org/10.15406/jhvr.2014.01.00006>
- Dahlberg JH (1988). An overview of retrovirus replication and classification. *Adv Vet Sci Comp Med*. 1988;32:1-35.
- Dannull, J., Surovoy, A., Jung, G., & Moelling, K. (1994). Specific binding of HIV-1 nucleocapsid protein to PSI RNA in vitro requires N-terminal zinc finger and flanking basic amino acid residues. *The EMBO Journal*, 13(7), 1525–33. doi <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=394981&tool=pmcentrez&rendertype=abstract>
- Das, A. T., Harwig, A., & Berkhout, B. (2011). The HIV-1 Tat Protein Has a Versatile Role in Activating Viral Transcription. *Journal of Virology*, 85(18), 9506–9516. <https://doi.org/10.1128/JVI.00650-11>
- Das, A. T., Vrolijk, M. M., Harwig, A., & Berkhout, B. (2012). Opening of the TAR hairpin in the HIV-1 genome causes aberrant RNA dimerization and packaging. *Retrovirology*, 9(1), 59. <https://doi.org/10.1186/1742-4690-9-59>
- Dayton, A. I. (2004). Within you, without you: HIV-1 Rev and RNA export. *Retrovirology*, 1, 35. <https://doi.org/10.1186/1742-4690-1-35>
- Debaisieux, S., Rayne, F., Yezid, H., & Beaumelle, B. (2012). The Ins and Outs of HIV-1 Tat. *Traffic*, 13(3), 355–363. <https://doi.org/10.1111/j.1600-0854.2011.01286.x>
- Demaison, Christophe & Parsley, Kathryn & Brouns, Gaby & Scherr, Michaela & Battmer, Karin & Kinnon, Christine & Grez, Manuel & J Thrasher, Adrian. (2002). High-Level Transduction and Gene Expression in Hematopoietic Repopulating Cells Using a Human Immunodeficiency Virus Type 1-Based Lentiviral Vector Containing an Internal Spleen Focus Forming Virus Promoter. *Human gene therapy*. 13. 803-13. 10.1089/10430340252898984.
- Denton, P. W., & Garcia, J. V. (2009). Novel humanized murine models for HIV research. *Current HIV/AIDS Reports*, 6(1), 13–19. <https://doi.org/10.1007/s11904-009-0003-2>
- Deptala A, Bedner E, Gorczyca W, and Z Darzynkiewicz. Activation of nuclear factor kappa B (NF-kB) assayed by laser scanning cytometry (LSC) (1998). *Cytometry*. 1998 November 1; 33(3): 376–382.

- Dong X, Biswas A, Süel KE, Jackson LK, Martinez R, Gu H and Chook YM (2009). Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature*. 2009 April 30; 458(7242): 1136–1141. doi:10.1038/nature07975.
- Du, L., Zhao, Y., Chen, J., Yang, L., Zheng, Y., Tang, Y., Shen, X., Jiang, H., (2008). D77, one benzoic acid derivative, functions as a novel anti-HIV-1 inhibitor targeting the interaction between integrase and cellular LEDGF/p75. *Biochem. Biophys. Res. Commun.* 375 (1), 139–144.
- Campbell EM and T. J. Hope (2015). HIV-1 Capsid: the multifaceted key player in HIV-1 infection. *Nature Reviews Microbiology* 13, 471–483 doi:10.1038/nrmicro3503
- Domingo E. and J.Holland, (2000) “RNA virus mutations and fitness for survival,” *Annual Review of Microbiology*, vol. 51, pp. 151–178, 1997.
- Ernst,R.K., Bray,M., Rekosh,D. and Hammarskjöld,M.-L. (1997) Secondary structure and mutational analysis of the Mason–Pfizer monkey virus constitutive transport element. *RNA*, 3, 210–222.
- Ernst,R.K., Bray,M., Rekosh,D. and Hammarskjöld,M.-L. (1997) A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Molecular Cell Biology*, 17, 135–144.
- Escors, D., & Breckpot, K. (2011). Lentiviral vectors in gene therapy: Their current status and future potential. *Archivum Immunologiae et Therapiae Experimentalis*, 58(2), 107–119. <https://doi.org/10.1007/s00005-010-0063-4>.Lentiviral
- Feinberg MB and Moore JP (2002). AIDS vaccine models: challenging challenge viruses. *Nature Medicine*, 2002 Mar;8(3):207-10.
- Feng Y, Broder C.C., Kennedy P.E. and Berger E.A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*. 1996 May 10;272(5263):872-7.
- Fernandes JD, Jayaraman B. and Frankel AD. (2012). The HIV-1 Rev response element. *RNA Biology* 9:1, 6-11; Landes Bioscience.
- Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, et al. (1999) Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV- 1, even in patients on effective combination therapy. *Nature Medicine* 5: 512-517.
- Fisher KJ, Joos K., Alston J., Yang Y., Haecker SE., High K., Pathak R., Raper SE. And Wilson JM (1997). Recombinant adeno-associated virus for muscle directed gene therapy. *Nature Medicine*, 1997 Mar;3 (3):306-12
- Fisher, K. J., G. P. Gao, M. D. Weitzman, R. DeMatteo, J. F. Burda, and J. M. Wilson. (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *Journal of Virology* 70:520–532.

- Fukuda M, Asano S, Nakamura T, *et al.* (1997). CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* 1997; 390: 308-11.
- Fun, A., Wensing, A. M. J., Verheyen, J., & Nijhuis, M. (2012). Human Immunodeficiency Virus gag and protease : partners in resistance. *Retrovirology*, 9:63; <http://www.retrovirology.com/content/9/1/63>
- Gage, F. H., & Verma, I. M. (1998). Development of a Self-Inactivating Lentivirus Vector, *Journal of Virology*, 72(10), 8150–8157.
- Gallay P, Swingler S, Song J, Bushman F, Trono D (1995). HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. *Cell* 1995, 83:569-576.
- Gallay, P. A., Hope, T. J., Chin, D., & Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 94(18), 9825–30. <https://doi.org/10.1073/pnas.94.18.9825>
- Ganser-Pornillos BK, Yeager M, Sundquist WI (2008) The structural biology of HIV assembly. *Curr Opin Struct Biol* 18: 203–217.
- Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH (1999). Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*, 4;397(6718):436-41.
- Gao F, Li Y, Decker JM, Peyerl FW, Bibollet-Ruche F, Rodenburg CM, Chen Y, Shaw DR, Allen S, Musonda R, Shaw GM, Zajac AJ, Letvin N, Hahn BH (2003). Codon usage optimization of HIV type 1 subtype C gag, pol, env, and nef genes: in vitro expression and immune responses in DNA-vaccinated mice. *AIDS Res Hum Retroviruses*. 2003 Sep;19(9):817-23.
- Garber, M. E., Wei, P., KewalRamani, V. N., Mayall, T. P., Herrmann, C. H., Rice, A. P., ... Jones, K. A. (1998). The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes and Development*, 12(22), 3512–3527. <https://doi.org/10.1101/gad.12.22.3512>
- Gaunt, E., Wise, H. M., Zhang, H., Lee, L. N., Atkinson, N. J., Nicol, M. Q., ... Simmonds, P. (2016). Elevation of CpG frequencies in influenza a genome attenuates pathogenicity but enhances host response to infection. *eLife*, 5, 1–19. <https://doi.org/10.7554/eLife.12735>
- Gendron, K., Ferbeyre, G., Heveker, N., & Brakier-Gingras, L. (2011). The activity of the HIV-1 IRES is stimulated by oxidative stress and controlled by a negative regulatory element. *Nucleic Acids Research*, 39(3), 902–912. <https://doi.org/10.1093/nar/gkq885>
- Gibbs JS., Lackner AA, Lang SM, Simon MA, Sehgal PK, Daniel MD and Desroisiers RC (1995). Progression to AIDS in the absence of a gene for *vpr* or *vpx*. *Journal of Virology*, Apr. 1995, p. 2378–2383 Vol. 69, No. 40022-538X/95/\$04.0010

- Gonçalves, J., Moreira, E., Sequeira, I. J., Rodrigues, A. S., Rueff, J., & Brás, A. (2016). Integration of HIV in the Human Genome: Which Sites Are Preferential A Genetic and Statistical Assessment. *International Journal of Genomics*, 2016, 8–13. <https://doi.org/10.1155/2016/2168590>
- Greatorex J, Gallego J, G. Varani and A. Lever (2002). Structure and stability of wild-type and mutant RNA internal loops from the SL-1 domain of the HIV-1 packaging signal. *Journal of Molecular Biology*, 20;322(3):543-57.
- Greatorex, J. (2004). The retroviral RNA dimer linkage: different structures may reflect different roles. *Retrovirology*, 1, 22. <https://doi.org/10.1186/1742-4690-1-22>
- Gringhuis, S. I., Hertoghs, N., Kaptein, T. M., Zijlstra-Willems, E. M., Sarrami-Forooshani, R., Sprockholt, J. K., ... Geijtenbeek, T. B. H. (2016). HIV-1 blocks the signaling adaptor MAVS to evade antiviral host defense after sensing of abortive HIV-1 RNA by the host helicase DDX3. *Nature Immunology*, 18, 225. Retrieved from <http://dx.doi.org/10.1038/ni.3647>
- Haim, H., Salas, I., & Sodroski, J. (2012). Proteolytic Processing of the Human Immunodeficiency Virus, 87(3), 1884–1889. <https://doi.org/10.1128/JVI.02765-12>
- Hallenberger S, Bosch V, Angliker H, Shaw E, Klenk HD, Garten W.(1992). Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature*. 1992 Nov 26;360(6402):358-61.
- Harrich, D., Garcia, J., Wu, F., Mitsuyasu, R., Gonzalez, J., & Gaynor, R. (1989). transcriptional regulation of the human Role of SPi-Binding Domains in In Vivo Transcriptional Regulation of the Human Immunodeficiency Virus Type 1 Long Terminal Repeat, 63(6), 2585–2591.
- Harrich, D., Hooker, C. W., & Parry, E. (2000). The human immunodeficiency virus type 1 TAR RNA upper stem-loop plays distinct roles in reverse transcription and RNA packaging. *Journal of Virology*, 74(12), 5639–5646. <https://doi.org/10.1128/JVI.74.12.5639-5646.2000>
- Hatzijioannou, T., Ambrose, Z., Chung, N. P. Y., Piatak, M., Yuan, F., Trubey, C. M., ... Bieniasz, P. D. (2009). A macaque model of HIV-1 infection, 106(11), 4425–4429.
- Hatzijioannou T. and Evans DT (2012). Animal models for HIV/AIDS research. *Nature Review Microbiology*. 2012 Dec; 10(12): 852-867.doi: 10.1038/nrmicro2911
- Hemelaar, J., Gouws, E., Ghys, P. D., & Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS*, 20(16), 13–23. <https://doi.org/10.1097/01.aids.0000247564.73009.bc>
- Hill, M. K., Hooker, C. W., Harrich, D., Crowe, S. M., & Mak, J. (2001). Gag-Pol Supplied in trans Is Efficiently Packaged and Supports Viral Function in Human Immunodeficiency Virus Type 1 Gag-Pol Supplied in trans Is Efficiently Packaged and Supports Viral Function in Human Immunodeficiency Virus Type 1. *Society*, 75(15), 6835–6840. <https://doi.org/10.1128/JVI.75.15.6835>

- Holder WD Jr, Wells SA Jr. Antibody reacting with the murine mammary tumor virus in the serum of patients with breast carcinoma: a possible serological detection method for breast carcinoma. *Cancer Research* 1983;43:239–44.
- Hofmann, H., Logue, E. C., Bloch, N., Daddacha, W., Polsky, S. B., Schultz, M. L., & Kim, B. (2012). The Vpx lentiviral accessory protein targets SAMHD1 for degradation in the nucleus, *Journal of Virology*, 86(23), 12552–12560. <https://doi.org/10.1128/JVI.01657-12>
- Hrecka K, et al. (2011). Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474:658 – 661
- Huang, Y., Yario, T.A., Steitz, J.A (2004). A molecular link between SR protein dephosphorylation and mRNA export. *Proc. Natl. Acad. Sci. USA* 2004, 101, 9666–9670.
- Hulme AE, Perez O, Hope TJ. (2011). Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc Natl Acad Sci USA*. 2011;108(24):9975–80.
- Hung, M., Patel, P., Davis, S., & Green, S. R. (1998). Importance of ribosomal frameshifting for human immunodeficiency virus type 1 particle assembly and replication. *Journal of Virology*, 72(6), 4819–24. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=110024&tool=pmcentrez&rendertype=abstract>
- Huthoff, H. & Malim, M. H. (2007). Identification of amino acid residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation. *Journal of Virology* 81, 3807–3815. (doi:10.1128/JVI. 02795-06)
- Hütter G., Nowak D., Mossner M., Ganepola S., Arne Müßig, Kristina Allers, Thomas Schneider, Jörg Hofmann, Claudia Kücherer, Olga Blau, Igor W. Blau, Wolf K. Hofmann, and Eckhard Thiel (2009). Long-term control of HIV by CCR5 delta32/delta32 stem-cell transplantation. *New England Journal of Medicine* 2009; 360:692-698 doi: 10.1056/NEJMoa0802905.
- Iordanskiy, S., & Bukrinsky, M. (2007). Reverse transcription complex: the key player of the early phase of HIV replication. *Future Virology*, 2(1), 49–64. <https://doi.org/10.2217/17460794.2.1.49>
- Isbel L, Whitelaw E. 2012. Endogenous retroviruses in mammals: an emerging picture of how ERVs modify expression of adjacent genes. *BioEssays* 34: 734–8.
- J.P. Stoye (2012). Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nature Reviews Microbiology* 10, 395–406 (2012) doi:10.1038/nrmicro2783
- Jablonski, J. A., A. L. Amelio, M. Giacca, and M. Caputi. (2010). The transcriptional transactivator Tat selectively regulates viral splicing. *Nucleic Acids Research* 38:1249–1260

- Jacques, D. A., McEwan, W. A., Hilditch, L., Price, A. J., Towers, G. J., & James, L. C. (2016). HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature*, 536(7616), 349–353. <https://doi.org/10.1038/nature19098>
- Jern P, Coffin JM (2008). Effects of retroviruses on host genome function. *Annual Review of Genetics* 42: 709–32.
- Joag,S.V.;Li,Z.;Foresman,L.;Stephens,E.B.;Zhao,L.J.;Adany,I.;Pinson,D.M.;McClure,H.M Narayan, O. (1996). Chimeric simian/human immunodeficiency virus that causes progressive loss of cd4+ t cells and aids in pig-tailed macaques. *Journal of Virology* 1996, 70, 3189–3197.
- Joseph SB, Swanstrom R, Kashuba AD, Cohen MS. (2015) Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nature Review Microbiology* 2015;13(7):414–25.
- Jouvenet, N., Simon, S. M., & Bieniasz, P. D. (2009). Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45), 19114–9. <https://doi.org/10.1073/pnas.0907364106>
- Jouvenet, N., Simon, S. M., & Bieniasz, P. D. (2011). Visualizing HIV-1 assembly. *Journal of Molecular Biology*, 410(4), 501–511. <https://doi.org/10.1016/j.jmb.2011.04.062>
- Karn J (1999). Tackling Tat. *Journal of Molecular Biology* 22;293(2):235-54
- Karn, J., & Stoltzfus, C. M. (2012). Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor Perspectives in Medicine*, 2(2), 1–17. <https://doi.org/10.1101/cshperspect.a006916>
- Kashanchi, F., Piras, G., Radonovich, M. F., Duvall, J. F., Fattaey, A., Chiang, C.-M., Roeder, R. G. & Brady, J. N. (1994). Direct interaction of human TFIID with the HIV-1 transactivator Tat. *Nature*, 367, 295-299.
- Keane, S. C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., ... Summers, M. F. (2015). Structure of the HIV-1 RNA packaging signal. *Science*, 348(6237), 917–921. <https://doi.org/10.1126/science.aaa9266>
- Keating, C. P., Hill, M. K., Hawkes, D. J., Smyth, R. P., Isel, C., Le, S., ... Mak, J. (2009). The A-rich RNA sequences of HIV-1 pol are important for the synthesis of viral cDNA, 37(3), 945–956. <https://doi.org/10.1093/nar/gkn1015>
- Khan Mohammad A, Aberham Claudia, Kao Sandra, Akari Hirofumi, Gorelick Robert, Bour Stephan, and Strebel. K. (2001). Human Immunodeficiency Virus Type 1 Vif Protein Is Packaged into the Nucleoprotein Complex through an Interaction with Viral Genomic RNA Human Immunodeficiency Virus Type 1 Vif Protein Is Packaged into the Nucleoprotein Complex through an Interaction with. *Journal of Virology*, 75(16), 7252–7265. <https://doi.org/10.1128/JVI.75.16.7252>
- Kimata, J. T. (2014). Stepping toward a Macaque Model of HIV-1 Induced AIDS, 3643–3651. <https://doi.org/10.3390/v6093643>

- Kinoshita S, Su L, Amano M, Timmerman LA, Kaneshima H, Nolan GP. (1997). The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells. *Immunity* 1997;6:235–44. [PubMed: 9075924]
- Kjems J, Brown M, Chang DD, Sharp PA. Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element (1991). *Proc Natl Acad Sci USA* 1991; 88:683-7; PMID:1992459; <http://dx.doi.org/10.1073/pnas.88.3.683>
- Kjems J, Calnan BJ, Frankel AD, Sharp PA. (1992). Specific binding of a basic peptide from HIV-1 Rev. *EMBO Journal* 1992; 11: 1119-29.
- Klaver, B., & Berkhout, B. (1994). Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. *Journal of Virology*, 68(6), 3830–40. From <http://www.ncbi.nlm.nih.gov/pubmed/8189520> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC236888>
- Kleiman, L. (2002). The Primer tRNA for Reverse Transcription in HIV-1. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)*, 53(2), 107–114. <https://doi.org/10.1080/15216540211469>
- Knight, S., Bokhoven, M., Collins, M., & Takeuchi, Y. (2010). Effect of the internal promoter on insertional gene activation by lentiviral vectors with an intact HIV long terminal repeat. *Journal of Virology*, 84(9), 4856–9. <https://doi.org/10.1128/JVI.02476-09>
- Knoepfel, S. A., & Berkhout, B. (2013). On the role of four small hairpins in the HIV-1 RNA genome, *RNA Biology*, 10(4), 540–552.
- Kong, R., Xu, K., Zhou, T., Acharya, P., Lemmin, T., Ozorowski, G., ... Mascola, J. R. (2016). *Science*, 352(6287), 828–833. <https://doi.org/10.1126/science.aae0474>. Fusion
- Kotsopoulou E, Kim VN, Kingsman AJ, Kingsman SM, Mitrophanous KA. (2000) A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene. *Journal of Virology*, p. 4839–4852 0022-538X/00/\$04.00 0, *American Society for Microbiology*
- Kubota S, Pomerantz RJ. (1998). A cis-acting peptide signal in human immunodeficiency virus type I Rev which inhibits nuclear entry of small proteins. *Oncogene* 1998; 16: 1851-61.
- Kumar M, Keller B., Makalou N. and Sutton RE (2001) Systematic Determination of the Packaging Limit of Lentiviral Vectors. *Human Gene Therapy*; 12(15): 1893-1905. <https://doi.org/10.1089/104303401753153947>
- Kutky M., KA. Hudak (2017). Expression of an RNA glycosidase inhibits HIV-1 transactivation of transcription. *Biochemical Journal*, BCJ20170353; doi: 10.1042/BCJ20170353

- Kuzembayeva, M., Dilley, K., Sardo, L., & Hu, W. S. (2014). Life of psi: How full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology*, 454–455(1), 362–370. <https://doi.org/10.1016/j.virol.2014.01.019>
- Lai, M., Wang, S., Cheng, L., Tarn, W., Tsai, S., & Sun, H. S. (2013). Human DDX3 Interacts with the HIV-1 Tat Protein to Facilitate Viral mRNA Translation, 8(7), 1–14. <https://doi.org/10.1371/journal.pone.0068665>
- Lalezari, JP; Eron, JJ; Carlson, M; Cohen, C; Dejesus, E; Arduino, RC; Gallant, JE; Volberding, P; et al. (2003). "A phase II clinical study of the long-term safety and antiviral activity of enfuvirtide-based antiretroviral therapy". *AIDS (London, England)*. 17 (5): 691–8. doi:10.1097/01.aids.0000050825.06065.84
- Le Rouzic, E., & Benichou, S. (2005). The Vpr protein from HIV-1: distinct roles along the viral life cycle. *Retrovirology*, 2, 11. <https://doi.org/10.1186/1742-4690-2-11>
- Le, S., Malim, M. H., Cullen, B. R., & Maizel, J. V. (1990). A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency viruses. *Nucleic Acids Research*, Vol 18(6), 1613–1623.
- Leblanc, J., Weil, J., & Beemon, K. (2013). Posttranscriptional regulation of retroviral gene expression: Primary RNA transcripts play three roles as pre-mRNA, mRNA, and genomic RNA. *Wiley Interdisciplinary Reviews: RNA*, 4(5), 567–580. <https://doi.org/10.1002/wrna.1179>
- Letvin, N. L. et al. (1985). Induction of AIDS-like disease in macaque monkeys with T-cell tropic retrovirus STLV-III. *Science* 230, 71–73. (doi:10.1126/science.2412295)
- Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., ... Bushman, F. D. (2006). Retroviral DNA Integration: Viral and Cellular Determinants of Target-Site Selection. *PLOS Pathogens*, 2(6), e60. Retrieved from <https://doi.org/10.1371/journal.ppat.0020060>
- Liberatore RA, Bieniasz PD (2011) Tetherin is a key effector of the antiretroviral activity of type I interferon in vitro and in vivo. *Proc Natl Acad Sci USA* 108:18097–18101. doi:10.1073/pnas.1113694108
- Locker, N., Chamond, N., & Sargueil, B. (2011). A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. *Nucleic Acids Research*, 39(6), 2367–2377. <https://doi.org/10.1093/nar/gkq1118>
- Locher, C. P., Witt, S. A., Herndier, B. G., Abbey, N. W., Tenner-Racz, K., Racz, P., Kiviat, N B., Murthy, K K., Brasky, K., Leland, M. and Levy, J. A. (2003). Increased virus replication and virulence after serial passage of human immunodeficiency virus type 2 in baboons. *Journal of Virology*, 77(1), 77–83. <https://doi.org/10.1128/JVI.77.1.77-83.2003>
- Lotti, F., Menguzzato, E., Rossi, C., Naldini, L., Ailles, L., Mavilio, F., & Ferrari, G. (2002). Transcriptional Targeting of Lentiviral Vectors by Long Terminal Repeat Enhancer Replacement. *Journal of Virology*, 76(8), 3996–4007. <https://doi.org/10.1128/JVI.76.8.3996>

- Love, D. C., Sweitzer, T. D., & Hanover, J. A. (1998). Reconstitution of HIV-1 rev nuclear export: independent requirements for nuclear import and export. *Proc Natl Acad Sci U S A*, 95(18), 10608–10613. <https://doi.org/10.1073/pnas.95.18.10608>
- Lower R, Tonjes RR, Korbmacher C, Kurth R, Lower J. (1995). Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K. *Journal of Virology* 69:141–149.
- Lu K., Heng X., and MF Summers (2011). Structural Determinants and Mechanism of HIV-1 Genome Packaging. *Journal of Molecular Biology* . 2011 July 22; 410(4): 609–633. doi:10.1016/j.jmb.2011.04.029.
- Llano M, Saenz DT, Meehan A, Wongthida P, Peretz M, Walker WH, Teo W, Poeschla EM (2006) An Essential Role for LEDGF/p75 in HIV Integration. *Science* 314: 461–464.
- MA Endsley, AD Somasunderam, L. Guagyu, N. Oezguen, V. Thiviyanathan, JL Murray, DH Rubin, TW Hodge, WA O'Brien, B. Lewis and MR Ferguson (2014). *Virology*. 2014 April ; 0: 60–66. doi:10.1016/j.virol.2014.02.006.
- Mackay, C. R. (1996). Chemokine receptors and T cell chemotaxis. *Journal of Experimental Medicine* 184, 799 – 802.
- Magin, C., Löwer, R., & Löwer, J. (1999). cORF and RcRE, the Rev/Rex and RRE/RxRE homologues of the human endogenous retrovirus family HTDV/HERV-K. *Journal of Virology*, 73(11), 9496–507. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10516058%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC112984>
- Mahboobi, S. H., Javanpour, A. A., & Mofrad, M. R. K. (2015). The Interaction of RNA Helicase DDX3 with HIV-1 Rev-CRM1-RanGTP Complex during the HIV Replication Cycle. <https://doi.org/10.1371/journal.pone.0112969>
- Mailler, E., Bernacchi, S., Marquet, R., Paillart, J. C., Vivet-Boudou, V., & Smyth, R. P. (2016). The life-cycle of the HIV-1 gag–RNA complex. *Viruses*, 8(9), 1–19. <https://doi.org/10.3390/v8090248>
- Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, Spindler J, Ferris AL, Mellors JW, Kearney MF, Coffin JM, Hughes SH. (2014). HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* 345:179 – 183.
- Maldarelli, F., Kearney, M., Palmer, S., Stephens, R., Mican, J., Polis, M. A., ... Coffin, J. M. (2013). HIV Populations Are Large and Accumulate High Genetic Diversity in a Nonlinear Fashion. *Journal of Virology*, 87(18), 10313–10323. <https://doi.org/10.1128/JVI.01225-12>
- Malim MH, Bohnlein S, Hauber J, Cullen BR. (1989). Functional dissection of the HIV-1 Rev trans-activator-derivation of a trans-dominant repressor of Rev protein. *Cell* 1989; 58: 205-14

- Mamede, J. I., Cianci, G. C., Anderson, M. R., & Hope, T. J. (2017). Early cytoplasmic uncoating is associated with infectivity of HIV-1. *Proceedings of the National Academy of Sciences*, 201706245. <https://doi.org/10.1073/pnas.1706245114>
- Mao Y., Wang L., Gu C., Herschhorn A., Xiang SH, Haim H., Yang X. and Sodroski J. (2012). Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer. *Nature Structural & Molecular Biology* 19, 893–899 (2012) doi:10.1038/nsmb.2351
- Margolis L. and Shattock R. (2006). Selective transmission of CCR5-utilizing HIV-1: the ‘gatekeeper’ problema resolved?. *Nature Review Microbiology*, 2006 Apr;4(4): 312-7
- Marshall HM, Ronen K, Berry C, Llano M, Sutherland H, Saenz D, Bickmore W, Poeschla E, Bushman FD. (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS One* 2:e1340. <http://dx.doi.org/10.1371/journal.pone.0001340>.
- Masso, M., & Vaisman, I. I. (2010). Accurate and efficient gp120 V3 loop structure based models for the determination of HIV-1 co-receptor usage. *BMC Bioinformatics* 2010, 11:494 <http://www.biomedcentral.com/1471-2105/11/494>
- Mathew, S. F., Crowe-McAuliffe, C., Graves, R., Cardno, T. S., McKinney, C., Poole, E. S., & Tate, W. P. (2015). The highly conserved codon following the slippery sequence supports - 1 frameshift efficiency at the HIV-1 frameshift site. *PLoS One*, 10(3), 1–24. <https://doi.org/10.1371/journal.pone.0122176>
- Matthias T. Dittmar, Áine McKnight, Graham Simmons, Paul R. Clapham, Robin A. Weiss & Peter Simmonds (1997) HIV\_ tropism and co-receptor use. *Nature* 385, 495–496 (06 February 1997) doi:10.1038/385495a0
- McCune JM, Rabin LB, Feinberg MB, Lieberman M, Kosek JC, Reyes GR, Weissman IL. (1988). Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell*. 1988 Apr 8;53(1):55-67.
- Mertz, J.A., Simper, M.S., Lozano, M.M., Payne, S.M. and Dudley, J.P. (2005) Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. *Journal of Virology*, 79, 14737–14747.
- Meyerhans A., Cheynier R., J. Albert et al. (1989) “Temporal uctua- tions in HIV quasispecies in vivo are not re ected by sequential HIV isolations,” *Cell*, vol. 58, no. 5, pp. 901–910, 1989.
- Mishra, A., Lai, G. H., Schmidt, N. W., Sun, V. Z., Rodriguez, A. R., Tong, R., ... Wong, G. C. L. (2011). Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions. *Proceedings of the National Academy of Sciences*, 108(41), 16883–16888. <https://doi.org/10.1073/pnas.1108795108>
- Mlcochova, P., Apolonia, L., Kluge, S. F., Sridharan, A., Kirchhoff, F., Malim, M. H., ... Gupta, R. K. (2015). Immune evasion activities of accessory proteins Vpu, Nef and Vif are conserved in acute and chronic HIV-1 infection. *Virology*, 482, 72–78. <https://doi.org/10.1016/j.virol.2015.03.015>

- Montano, M. a, Kripke, K., Norina, C. D., Achacoso, P., Herzenberg, L. a, Roy, a L., & Nolan, G. P. (1996). NF-kappa B homodimer binding within the HIV-1 initiator region and interactions with TFII-I. *Proceedings of the National Academy of Sciences of the United States of America*, 93(22), 12376–12381. <https://doi.org/10.1073/pnas.93.22.12376>
- Montefiori et al. (2006). Protocol for Measuring Neutralizing Antibodies Against HIV-1, SIV and SHIV Using a Luciferase Reporter Gene Assay in TZM-BL Cells
- Morita, E., Sandrin, V., Chung, H. Y., Morham, S. G., Gygi, S. P., Rodesch, C. K., & Sundquist, W. I. (2007). Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO Journal*, 26(19), 4215–4227. <https://doi.org/10.1038/sj.emboj.7601850>
- Mostowy, R., Kouyos, R. D., Fouchet, D., & Bonhoeffer, S. (2011). The role of recombination for the coevolutionary dynamics of HIV and the immune response. *PLoS One*, 6(2), 1–10. <https://doi.org/10.1371/journal.pone.0016052>
- Mu, M., Salmons, B., Gu, W. H., & Indik, S. (2008). Identification of the Rem-responsive element of mouse mammary tumor virus, 36(19), 6284–6294. <https://doi.org/10.1093/nar/gkn608>
- Mueller, N., van Bel, N., Berkhout, B., & Das, A. T. (2014). HIV-1 splicing at the major splice donor site is restricted by RNA structure. *Virology*, 468, 609–620. <https://doi.org/10.1016/j.virol.2014.09.018>
- Müller, B., Anders, M., Akiyama, H., Welsch, S., Glass, B., Nikovics, K., Clavel F., Tervo H-M., Keppler O. and Kräusslich, H. G. (2009). HIV-1 gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. *Journal of Biological Chemistry*, 284(43), 29692–29703. <https://doi.org/10.1074/jbc.M109.027144>
- Nagai-Fukataki M, Ohashi T, Hashimoto I, Kimura T, Hakata Y, Shida H. (2011). Nuclear and cytoplasmic effects of human CRM1 on HIV-1 production in rat cells. *Genes to Cells* 16:203–216. doi: 10.1111/j.1365-2443.2010.01476.x.
- Nasioulas G, Zolotukhin A, Tabernero C, Solomin L, Cunningham C, Pavlakis G, Felber B (1994). Elements Distinct from Human Immuno- deficiency Virus Type 1 Splice Sites Are Responsible for the Rev Dependence of env mRNA. *Journal of Virology* 1994, 68:2986-2993.
- Ndungu UT, Lu Y., Renjifo B., Touzjian N., Kushner N. Pena-cruz, V., ... Essex, M. A. X. (2001). Infectious Simian / Human Immunodeficiency Virus with Human Immunodeficiency Virus Type 1 Subtype C from an African Isolate : Rhesus Macaque Model. *Journal of Virology*, 75(23), 11417–11425. <https://doi.org/10.1128/JVI.75.23.11417>
- Neil SJ., Zang T. and Bieniasz PD. (2008). Tetherin inhibits retrovirus release and is atagonized by HIV-1 Vpu. *Nature*, 2008, Jan 24;451 (7177):425-30. Doi: 10.1038/nature06553.

- Nikolaitchik OA, Galli A, Burdick RC, Levine L, Li K, Rein A, Pathak VK, Hu WS (2017). Interactions between HIV-1 Gag and Viral RNA Genome Enhance Virion Assembly. *Journal of Virology* 27;91(16). pii: e02319-16. doi: 10.1128/JVI.02319-16.
- Nitahara-Kasahara Y, Kamata M, Yamamoto T, Zhang X, Miyamoto Y, Muneta K, Iijima S, Yoneda Y, Tsunetsugu-Yokota Y, Aida Y (2007). Novel nuclear import of Vpr promoted by importin alpha is crucial for human immunodeficiency virus type 1 replication in macrophages. *Journal of Virology* 2007, 81:5284-5293.
- Nitta, T., Hofacre, A., Hull, S., & Fan, H. (2009). Identification and Mutational Analysis of a Rej Response Element in Jaagsiekte Sheep Retrovirus RNA, *Journal of Virology* 83(23), 12499–12511. <https://doi.org/10.1128/JVI.01754-08>
- O’Keeffe, B., Fong, Y., Chen, D., Zhou, S., & Zhou, Q. (2000). Requirement for a kinase-specific chaperone pathway in the production of a Cdk9/cyclin T1 heterodimer responsible for P-TEFb-mediated Tat stimulation of HIV-1 transcription. *Journal of Biological Chemistry*, 275(1), 279–287. <https://doi.org/10.1074/jbc.275.1.279>
- O’Neil, P. K., Sun, G., Yu, H., Ron, Y., Dougherty, J. P., & Preston, B. D. (2002). Mutational analysis of HIV-1 long terminal repeats to explore the relative contribution of reverse transcriptase and RNA polymerase II to viral mutagenesis. *Journal of Biological Chemistry*, 277(41), 38053–38061. <https://doi.org/10.1074/jbc.M204774200>
- Ogert, R. A., Lee, L. H., & Beemon, K. L. (1996). Avian Retroviral RNA Element Promotes Unspliced RNA Accumulation in the Cytoplasm, *Journal of Virology*, 70(6), 3834–3843.
- Ohi, Y., & Clever, J. L. (2000). Sequences in the 5’ and 3’ R elements of human immunodeficiency virus type 1 critical for efficient reverse transcription. *Journal of Virology*, 74(18), 8324–8334.
- Olson, E. D., Cantara, W. A., & Musier-Forsyth, K. (2015). New Structure Sheds Light on Selective HIV-1 Genomic, *Viruses* 4826–4835. <https://doi.org/10.3390/v7082846>
- Paillart, J., Marquet, R., Hijnen, M., Jossinet, F., Mak, J., JC Paillart and R. Marquet (2014). Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nature communications*, 5:4304, <https://doi.org/10.1038/ncomms5304>
- Paul R Clapham and Áine McKnight (2001). HIV receptors and cell tropism. *British Medical Bulletin* 2001;58: 43–59
- Peterlin, B. M., & Price, D. H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. *Molecular Cell*, 23(3), 297–305. <https://doi.org/10.1016/j.molcel.2006.06.014>
- Piekna-Przybylska D., Dykes C, Demeter LM, and Bambara RA (2008). Sequences in the U3 region of Human Immunodeficiency Virus 1 improve efficiency of minus strand transfer in infected cells. *Virology*, 86(12), 3279–3288. <https://doi.org/10.1007/s11103-011-9767-z.Placid>

- Ping YH., Liu Y., Wang X., Neenhold HR. And TM Rana (1997). Dynamics of RNA-protein interactions in the HIV-1 Rev-RRE complex visualized by 6-thioguanosine-mediated photocrosslinking. *RNA* 1997, *Cambridge University Press* 3:850-860.
- Pollard VW, Malim MH. (1998) The HIV-1 Rev protein. *Annual Review of Microbiology*; 52:491-532
- Powell, S. K., Rivera-Soto, R., & Gray, S. J. (2015). Viral expression cassette elements to enhance transgene target specificity and expression in gene therapy. *Discovery Medicine*, 19(102), 49–57. <https://doi.org/10.1530/ERC-14-0411>. Persistent
- Ramezani, A, Hawley, TS and Hawley, RG (2008). Combinatorial incorporation of enhancer-blocking components of the chicken beta-globin 5'HS4 and human T-cell receptor alpha/delta BEAD-1 insulators in self-inactivating retroviral vectors reduces their genotoxic potential. *Stem Cells* 26: 3257–3266.
- Ran, X., Ao, Z., Trajtman, A., Xu, W., Kobinger, G., & Keynan, Y. (2017). HIV-1 envelope glycoprotein stimulates viral transcription and increases the infectivity of the progeny virus through the manipulation of cellular machinery. *Scientific Reports*, (February), 1–16. <https://doi.org/10.1038/s41598-017-10272-7>
- Rhee SS and JW Marsh (1994). HIV-1 Nef activity in murine T cells. CD4 modulation and positive enhancement. *Journal of Immunology*, May 15, 1994, 152 (10) 5128-5134.
- Richter, S., Ping, Y.-H., & Rana, T. M. (2002). TAR RNA loop: a scaffold for the assembly of a regulatory switch in HIV replication. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), 7928–7933. <https://doi.org/10.1073/pnas.122119999>
- Riddler SA, Haubrich R, DiRienzo AG, et al (2008). AIDS Clinical Trials Group Study A5142 Team. Class-sparing regimens for initial treatment of HIV-1 infection. *New England Journal of Medicine* 2008;358(20):2095–2106.
- Riedel, C., Vasishtan, D., Siebert, C. A., Whittle, C., Lehmann, M. J., Mothes, W., & Grünewald, K. (2017). Native structure of a retroviral envelope protein and its conformational change upon interaction with the target cell. *Journal of Structural Biology*, 197(2), 172–180. <https://doi.org/10.1016/j.jsb.2016.06.017>
- Rihn, S. J., Wilson, S. J., Loman, N. J., Alim, M., Bakker, S. E., Bhella, D., ... Bieniasz, P. D. (2013). Extreme Genetic Fragility of the HIV-1 Capsid. *PLOS Pathogens*, 9(6), e1003461. Retrieved from <https://doi.org/10.1371/journal.ppat.1003461>
- Rizvi, T. A., Schmidt, R. D., & Lew, K. A. (1997). Mason – Pfizer Monkey Virus ( MPMV ) Constitutive Transport Element ( CTE ) Functions in a Position-Dependent Manner. *Virology*, 129(236), 118–129.
- Rizvi, T. A., Schmidt, R. D., Lew, K. A., & Keeling, M. E. (1996). Rev/RRE-independent Mason-Pfizer monkey virus constitutive transport element-dependent propagation of SIVmac239 vectors using a single round of replication assay. *Virology*, 222(2), 457–463. <https://doi.org/10.1006/viro.1996.0444>

- Roberts JD, Bebenek K, Kunkel TA. (1988). The accuracy of reverse transcriptase from HIV-1. *Science*. 1988 Nov 25;242(4882):1171-3.
- Roy, S., Delling, U., & Rosen, C. A. And N. Sonenberg (1990). A bulge structure in HIV-I TAR RNA is required for Tat binding and Tat- mediated trans-activation, *Gen Dev*, 1365–1373.
- Saayman, S., Ackley, A., Turner, A. W., Famiglietti, M., Bosque, A., Clemson, M., ... Morris, K. V. (2014). An HIV-Encoded Antisense Long Noncoding RNA Epigenetically Regulates Viral Transcription. *Molecular Therapy*, 22(6), 1164–1175. <https://doi.org/10.1038/mt.2014.29>
- Sheehy AM, Gaddis NC, Choi JD, Malim MH (August 2002). "Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein". *Nature*. 418 (6898): 646–50. [doi:10.1038/nature00939](https://doi.org/10.1038/nature00939).
- Sakuma, T., Barry, M. A., & Ikeda, Y. (2012). Lentiviral vectors: basic to translational. *Biochemical Journal*, 443(3), 603–618. <https://doi.org/10.1042/BJ20120146>
- Salter, J. D., Morales, G. A., & Smith, H. C. (2014). Structural insights for HIV-1 therapeutic strategies targeting Vif. *Trends in Biochemical Sciences*, 39(9), 373–380. <https://doi.org/10.1016/j.tibs.2014.07.001>
- Sawaya, B. E., Khalili, K., Gordon, J., Taube, R., & Amini, S. (2000). Cooperative interaction between HIV-1 regulatory proteins Tat and Vpr modulates transcription of the viral genome. *Journal of Biological Chemistry*, 275(45), 35209–35214. <https://doi.org/10.1074/jbc.M005197200>
- Scarlata, S., & Carter, C. (2003). Role of HIV-1 Gag domains in viral assembly. *Biochimica et Biophysica Acta - Biomembranes*, 1614(1), 62–72. [https://doi.org/10.1016/S0005-2736\(03\)00163-9](https://doi.org/10.1016/S0005-2736(03)00163-9)
- Sharma, S. K., deVal, N., Bale, S., Guenaga, J., Tran, K., Feng, Y., ... Wyatt, R. T. (2015). Cleavage-Independent HIV-1 Env Trimers Engineered as Soluble Native Spike Mimetics for Vaccine Design. *Cell Reports*, 11(4), 539–550. <https://doi.org/10.1016/j.celrep.2015.03.047>
- Sharp, P. M., & Hahn, B. H. (2010). The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1552), 2487–2494. <https://doi.org/10.1098/rstb.2010.0031>
- Sharp, P. M., & Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harbor Perspectives in Medicine*, 1(1), 1–22. <https://doi.org/10.1101/cshperspect.a006841>
- Shen N, Jetté L, Liang C, Wainberg MA and M Laughrea (2000). Impact of Human Immunodeficiency Virus Type 1 RNA Dimerization on Viral Infectivity and of Stem-Loop B on RNA Dimerization and Reverse Transcription and Dissociation of Dimerization from Packaging. *Journal of Virology*, 0022-538X/00/\$04.0010 June 2000, p. 5729–5735.

- Sherer NM, Swanson CM, Hué S., Roberts RG, Julien R. C. Bergeron, MH Malim (2011). Evolution of a Species-Specific Determinant within Human CRM1 that Regulates the Post-transcriptional Phases of HIV-1 Replication. *PLOS pathogens*, , Volume 7, Issue 11, e1002395.
- Sherpa, C., Rausch, J. W., Le Grice, S. F. J., Hammarskjold, M. L., & Rekosh, D. (2015). The HIV-1 Rev response element (RRE) adopts alternative conformations that promote different rates of virus replication. *Nucleic Acids Research*, 43(9), 4676–4686. <https://doi.org/10.1093/nar/gkv313>
- Simpson, S. B., Zhang, L. E. I., & Craven, R. C. (1997). Rous Sarcoma Virus Direct Repeat cis Elements Exert Effects at Several Points in the Virus Life Cycle, 71(12), 9150–9156.
- Southgate, C. D., & Green, M. R. (1991). The HIV-1 Tat protein activates transcription from an upstream DNA- binding site: implications for Tat function. *Genes Dev*, 5(12B), 2496–2507.
- Spasic, A., Assmann, S. M., Bevilacqua, P. C., & Mathews, D. H. (2018). Modeling RNA secondary structure folding ensembles using SHAPE mapping data, 46(1), 314–323. <https://doi.org/10.1093/nar/gkx1057>
- Spiegelman S, Axel R, Schlom J. (1972). Virus-related RNA in human and mouse mammary tumors. *Journal of National Cancer Institute* 1972;48:1205-11
- Strebel K. (2003). Virus-host interactions: role of HIV proteins Vif, Tat and Rev. *AIDS*, 17 (suppl4): S25-S34
- Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J.(2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature*. 2004 Feb 26;427(6977):848-53.
- Stroud JC, Oltman A., Han A., Bates DL and Chen L. (2009). Structural basis of HIV-1 activation by NF-kappaB - a higher-order complex of p50: RelA bound to the HIV-1 LTR. *Journal of Molecular Biology* 2009 October 16; 393(1): 98–112. doi:10.1016/j.jmb.2009.08.023.
- Studamire, B., & Goff, S. P. (2010). Interactions of host proteins with the murine leukemia virus integrase. *Viruses*, 2(5), 1110–1145. <https://doi.org/10.3390/v2051110>
- Su, K., Wang, D., Ye, J., Kim, Y. C., & Chow, S. A. (2009). Site-specific integration of retroviral DNA in human cells using fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc-finger protein E2C. *Methods*, 47(4), 269–276. <https://doi.org/10.1016/j.ymeth.2009.01.001>.Site-Specific
- Suhasini, M., & Reddy, T. R. (2009). Cellular proteins and HIV-1 Rev function. *Current HIV Research*, 7(1), 91–100. <https://doi.org/10.2174/157016209787048474>

- Sükösd, Z., Andersen, E. S., Seemann, S. E., Jensen, M. K., Hansen, M., Gorodkin, J., & Kjems, J. (2015). Full-length RNA structure prediction of the HIV-1 genome reveals a conserved core domain. *Nucleic Acids Research*, *43*(21), 10168–10179. <https://doi.org/10.1093/nar/gkv1039>
- Summa, V., Petrocchi, A., Bonelli, F., Crescenzi, B., Donghi, M., Ferrara, M., Fiore, F., Gardelli, C., Gonzalez Paz, O., Hazuda, D.J., Jones, P., Kinzel, O., Laufer, R., Monteagudo, E., Muraglia, E., Nizi, E., Orvieto, F., Pace, P., Pescatore, G., Scarpelli, R., Stillmock, K., Witmer, M.V., Rowley, M., (2008). Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. *Journal of Medicinal Chemistry* *51* (18), 5843–5855.
- Sun, M., Grigsby, I. F., Gorelick, R. J., Mansky, L. M., & Musier-Forsyth, K. (2014). Retrovirus-Specific Differences in Matrix and Nucleocapsid Protein-Nucleic Acid Interactions: Implications for Genomic RNA Packaging. *Journal of Virology*, *88*(2), 1271–1280. <https://doi.org/10.1128/JVI.02151-13>
- Sundquist, W. I., & Kräusslich, H. G. (2012). HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspectives in Medicine*, *2*(7), 1–24. <https://doi.org/10.1101/cshperspect.a006924>
- Sunshine, S., Kirchner, R., Amr, S. S., Mansur, L., Shakhbatyan, R., Kim, M., ... Li, Z. (2016). HIV Integration Site Analysis of Cellular Models of HIV Latency with a Probe-Enriched Next-Generation Sequencing Assay, *90*(9), 4511–4519. <https://doi.org/10.1128/JVI.01617-15>. Editor
- Takata, M. A., Gonçalves-carneiro, D., Zang, T. M., Soll, S. J., York, A., Blanco-melo, D., & Bieniasz, P. D. (2017). defence targeting non-self RNA. *Nature*. <https://doi.org/10.1038/nature24039>
- Taniguchi, I., Mabuchi, N., & Ohno, M. (2014). NAR Breakthrough Article HIV-1 Rev protein specifies the viral RNA export pathway by suppressing TAP / NXF1 recruitment, *42*(10), 6645–6658. <https://doi.org/10.1093/nar/gku304>
- Triques, K., Bourgeois, A., Saragosti, S., Vidal, N., Mpoudi-Ngole, E., Nzilambi, N., ... Peeters, M. (1999). High diversity of HIV-1 subtype F strains in Central Africa. *Virology*, *259*(1), 99–109. <https://doi.org/10.1006/viro.1999.9720>
- Truant R, Cullen BR. (1999) The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. *Molecular Cell Biology* 1999; 19: 1210-7.
- Valentin, A, Lu, W., Rosati, M., Schneider, R., Albert, J., Karlsson, a, & Pavlakis, G. N. (1989). Dual effect of interleukin 4 on HIV-1 expression: implications for viral phenotypic switch and disease progression. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(July), 8886–8891. <https://doi.org/10.1073/pnas.95.15.8886>
- Vallejos, M., Carvajal, F., Pino, K., Navarrete, C., Ferres, M., Huidobro-Toro, J. P., Sargeil B. and López-Lastra, M. (2012). Functional and structural analysis of the internal ribosome entry site present in the mrna of natural variants of the hiv-1. *PLoS One*, *7*(4), 1–12. <https://doi.org/10.1371/journal.pone.0035031>

- Van Damme N, Goff D, Katsura C, Jorgenson RL, Mitchell R, Johnson MC, Stephens EB, Guatelli J. (2008). The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* 3: 245–252.
- Van Opijnen, T., de Ronde, A., Boerlijst, M. C., & Berkhout, B. (2007). Adaptation of HIV-1 depends on the host-cell environment. *PLoS One*, 2(3), 1–9. <https://doi.org/10.1371/journal.pone.0000271>
- Velpandi A1, Nagashunmugam T, Otsuka T, Cartas M, Srinivasan A. (1992). Structure-function studies of HIV-1: influence of long terminal repeat U3 region sequences on virus production. *DNA Cell Biology* 1992 Jun;11(5):369-76.
- Verdin, E., Paras, P., & Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *The EMBO Journal*, 12(8), 3249–3259.
- Veselinovic M, Yang KH, LeCureux J, et al. (2014). HIV pre- exposure prophylaxis: Mucosal tissue drug distribution of RT inhibitor Tenofovir and entry inhibitor Maraviroc in a humanized mouse model. *Virology* 2014;464–465:253– 263.
- Viphakone, N.; Hautbergue, G.M.; Walsh, M.; Chang, C.T.; Holland, A.; Folco, E.G.; Reed, R.; Wilson, S.A (2012). TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nature Communications*. 2012, 3, doi:10.1038/ncomms2005
- Vogt, P. K. (2012). Retroviral Oncogenes: A Historical Primer. *Nature Reviews Cancer* 2012 September ; 12(9): 639–648. doi:10.1038/nrc3320.
- Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Jr, W. B., Swanstrom, R., Burch L.C. and Weeks M. (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, 460(7256), 711–716. <https://doi.org/10.1038/nature08237.Architecture>
- Westerhout, E. M., Ooms, M., Vink, M., Das, A. T., & Berkhout, B. (2005). HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Research*, 33(2), 796–804. <https://doi.org/10.1093/nar/gki220>
- Whitney JB, Ruprecht RM (2004) Live attenuated HIV vaccines: pitfalls and prospects. *Curr Opin Infect Dis* 17: 17-26.
- Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H., ... Weeks, K. M. (2008). High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biology*, 6(4), 883–899. <https://doi.org/10.1371/journal.pbio.0060096>
- Wodrich, H., Schambach, A., & Kräusslich, H. G. (2000). Multiple copies of the Mason-Pfizer monkey virus constitutive RNA transport element lead to enhanced HIV-1 Gag expression in a context-dependent manner. *Nucleic Acids Research*, 28(4), 901–10. <https://doi.org/10.1093/nar/28.4.901>

- Wodrich, H., Bohne, J., Gumz, E., Welker, R. and Krausslich, H.G. (2001) A new RNA element located in the coding region of a murine endogenous retrovirus can functionally replace the Rev/Rev-responsive element system in human immunodeficiency virus type 1 Gag expression. *Journal of Virology*, 75, 10670–10682.
- Woodward CL, Prakobwanakit S, Mosessian S, Chow SA (2009). Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. *Journal of Virology* 2009, 83:6522-6533.
- Wyatt R, Sodroski J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280:1884–1888
- Yacoub al N., Romanowska M., Haritonova N. And Foerster J. (2007). Optimized production and concentration of lentiviral vectors containing large inserts. *Journal of Gene Medicine*. 2007 Jul;9(7):579-84, <https://doi.org/10.1002/jgm.1052>
- Yamashita M, Emerman M (2004). Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. *Journal of Virology* 2004, 78:5670- 5678.
- Yang, J. I. N., Bogerd, H. A. L., Le, S., & Cullen, B. R. (2000). The human endogenous retrovirus K Rev response element coincides with a predicted RNA folding region. *RNA* (2000), 6 :1551–1564+ Cambridge1551–1564.
- Yedavalli, V. S. R. K., Benkirane, M., & Jeang, K. T. (2003). Tat and trans-activation-responsive (TAR) RNA-independent induction of HIV-1 long terminal repeat by human and murine cyclin T1 requires Sp1. *Journal of Biological Chemistry*, 278(8), 6404–6410. <https://doi.org/10.1074/jbc.M209162200>
- Yedavalli VS, Neuveut C, Chi YH, Kleiman L, Jeang KT (2004) Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. *Cell* 119: 381–392. PMID: [15507209](https://pubmed.ncbi.nlm.nih.gov/15507209/)
- Younis I. and PL. Green (2005). The human t-cell leukemia virus Rex protein. *Bioscience* 2005 Jan 1;10:431-45.
- Yu H., Borsotti C., Schickel JN, Zhu S., Strowig T., Eynon EE, Frleta D., Gurer C., Murphy AJ, Yancopoulos GD, Meffre E., Manz MG and Flavell RA (2017). A novel humanized mouse model with significant improvement of class-switched, antigen-specific antibody production. *Blood*, 23 February 2017, volumen 129, number 8.
- Yu, S., Ruden, T. V. O. N., Kantofft, P. W., Garbert, C., Seiberg, M., Ruthert, U., ... Gilboa, E. L. I. (1986). Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells, 83(May), 3194–3198.
- Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P. & Yu, X. F. (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302, 1056–1060. (doi:10.1126/science.1089591)
- Zeffman A., Hassard S., Varani G. and Lever A. (2000). Major HIV-1 Packaging Signal is an Extended Bulged Stem Loop whose Structure is Altered on Interaction with the Gag Polyprotein. *Journal of Molecular Biology* (2000) 297, 877±893

- Zhou C. and T. Rana (2002). A Bimolecular Mechanism of HIV-1 Tat Protein Interaction with RNA Polymerase II Transcription Elongation Complexes. *Journal of Molecular Biology*, Volume 320, Issue 5. [https://doi.org/10.1016/S0022-2836\(02\)00556-9](https://doi.org/10.1016/S0022-2836(02)00556-9)
- Zhou, Y., Zhang, H., Siliciano, J. D., Siliciano, R. F., & Irol, J. V. (2005). Kinetics of Human Immunodeficiency Virus Type 1 Decay following Entry into Resting CD4<sup>+</sup> T Cells, *79*(4), 2199–2210. <https://doi.org/10.1128/JVI.79.4.2199>
- Zhu, H., Jian, H., & Zhao, L. J. (2004). Identification of the 15FRFG domain in HIV-1 Gag p6 essential for Vpr packaging into the virion. *Retrovirology*, *1*, 1–5. <https://doi.org/10.1186/1742-4690-1-26>
- Zuber, G., McDermott, J., Karanjia, S., Zhao, W., Schmid, M. F., & Barklis, E. (2000). Assembly of retrovirus capsid-nucleocapsid proteins in the presence of membranes or RNA. *Journal of Virology*, *74*(16), 7431–7441. <https://doi.org/10.1128/JVI.74.16.7431-7441.2000>
- Zufferey, R., Donello, J. E., Trono, D., & Hope, T. J. (1999). Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *Journal of Virology*, *73*(4), 2886–92. <https://doi.org/10.1046/J.0962-1075.2001.00312.X>
- Zufferey, R., Dull, T., Mandel, R. J., Bukovsky, A., Quiroz, D., Naldini, L., & Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of Virology*, *72*(12), 9873–80. <https://doi.org/99030895>