Introduction

Despite the intensive research efforts over the past 30 years that have focused on HIV-1, there is still no vaccine for HIV-1 infection. However, tremendous progress in the understanding of the structure and biology of the HIV led to the development of safe and potent HIV-based transgene delivery vectors. These genetic vehicles are referred to as lentiviral vectors. HIV-1 and other lentiviruses (a genus of the *Retroviridae* family) have emerged as a powerful tool for transgene delivery into cells *in vitro* and *in vivo*, including for generating transgenic animals. Moreover, HIV-based vectors have reached the clinic as promising vehicles for anti-viral gene therapy or vaccine application. Viral vectors have gained their popularity in basic research and gene therapy application because of their high rates of gene transfer that are far superior to those achieved with non-viral methods (*Pluta K, Kacprzak MM.,2009*). Numerous types of virus-derived gene transfer system are now available; these genetic vehicles are based either on DNA viruses or on RNA viruses. In particular, DNA viruses-based vectors either integrate into the host genome or express their genetic information episomally, while RNA virus based vectors maintain their genome in the cytoplasm of the transduced cells. Retroviral vectors, due to their ability to integrate into the host DNA, offer stable and long term transgene expression; retroviral vectors derived from so-called gamma- or onco-retroviruses are widely used in cell biology and play an important role in modern biomedicine. More than 21% of gene therapy protocols approved worldwide employed retroviral vectors and, among all viral and non-viral methods, only adenovirus-based vectors are currently more popular (about 24%) *(http://www.wiley.com/legacy/wileychi/genmed/clinical/).* A serious drawback in using vectors based on gamma-retroviruses in clinical trials is their natural disposition to integrate near promoters and regulatory regions (*Daniel R, Smith JA., 2008*), and to induce insertional tumors due to the presence of potent transcriptional enhancer in the viral long terminal repeats (LTRs). In fact, such an adverse outcome has been reported earlier in two X-linked severe combined immunodeficiency (X-SCID1) patients treated with cytokine-activated hematopoietic stem/progenitor cells (HSPCs) transduced with murine leukemia virus (MLV). In these patients, blast cells harbour activating vector insertions near the LIM domain-only 2 (*LMO2*) and *BMI1* proto-oncogenes (*Hacein-Bey-Abina et al., 2003*). Another limitation to the application of gamma-retroviral vectors is that they can infect only dividing cells. These drawbacks can be avoided by using lentiviral vectors. Lentiviruses have evolved different strategies to interact with the host cell chromatin and do not

integrate preferentially into close proximity of transcriptional start sites but rather favour introns in chromosomal regions rich in expressed genes (*Wang et al., 2009*). Another advantage of lentivirus-based vectors, important for research and gene therapy applications, is their ability to infect both dividing and non-dividing cells. This feature, unique among retroviruses, substantially increases the range of cells available for gene transfer. Lentiviral vectors transduce quiescent cells, including primary hepatocytes, progenitor and stem cells (*Santoni de Sio & Naldini, 2009*), non-proliferating monocytes and macrophages (*Veron et al., 2009*), as well as postmitotic neurons (*Naldini et al., 1996a*). Although during the last two decades HIV-1 based vectors became accepted as the most promising gene delivery tool among lentiviruses, also other members of the genus serve as platforms of recombinant vector generation. To date, vectors based on genomes of the following lentiviruses have been developed: human immunodeficiency virus 2 (HIV-2), various simian immunodeficiency viruses (SIVs) (*Nègre et al., 2000*), feline immunodeficiency virus (FIV) (*Poeschla et al., 1998*), bovine immunodeficiency virus (BIV) (*Berkowitz et al., 2001a*), equine infectious anemia virus (EIAV) (*Olsen, 1998*), Jembrana disease virus (JDV) (*Metharom et al., 2000*) and Maedi-Visna virus (MVV) (*Berkowitz et al., 2001b*). Lentiviral vectors, due to their unique properties, including selection of "safe" integration sites in the host genome, efficient long-term gene delivery to both dividing and nondividing cells, relatively large cargo capacity (7-8 kb without affecting vector titer) and target specificity achieved by pseudotyping, are promising agent in research and biomedicine. However their origin from mammalian pathogens raises safety questions and this matter needs careful consideration.

Chapter 1

FROM LENTIVIRUS TO LENTIVIRAL VECTORS: TRANSFORMING INFECTIOUS PARTICLES INTO THERAPEUTICS

1.1 *HIV-1 genome*

Lentiviral genome consists of two linear positive-sense single stranded RNA molecules. The dimeric nature of the HIV-1 RNA genome is largely responsible for the high genetic variability of the viruses due to possible recombinations during reverse transcription. After its conversion into cDNA, it becomes integrated into chromosomal DNA as provirus (the size of HIV provirus is approximately 9.7 kb) and is transmitted through daughter cell generations upon cell division. As an independent transcription unit, it has its own regulatory elements and is transcribed by the cell's transcription system. In Figure 1 are shown the structural features of HIV-1 provirus.

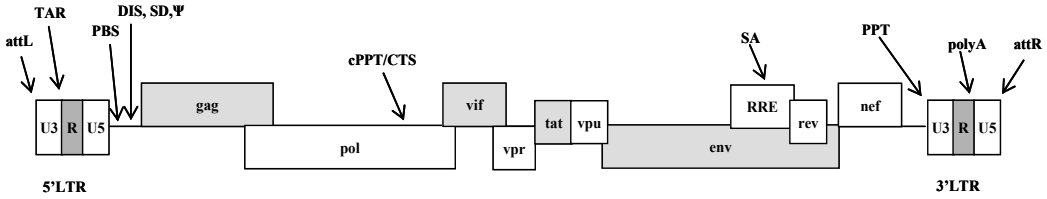

Figure 1. Schematic representation of HIV-1 genome features.
Abbreviations: LTR, long terminal repeat; attL, attR, left and right attachment sites; U3, 3' unique element; R, repeat element; U5, 5' unique element; TAR, transactivation response element; PBS, primer binding site; DIS, dimerization signal; SD, splice donor site; SA, splice acceptor site; Ψ, packaging signal; cPPT, central polypurine tract; CTS, central termination sequence; RRE, Rev response element; PPT, 3' polypurine tract; polyA, polyadenilation signal;

The protein encoding region are flanked by 5' e 3' LTRs, which consist of 3' unique elements (U3), repeat elements (R) and 5' unique elements (U5), and harbour some of the cis-acting elements. These cis elements contain signals important for provirus integration into host genome (*att* repeats, which are located at 5' and 3' ends of provirus DNA), enhancer/promoter sequences, transactivation response element (TAR) and polyadenilation signal (polyA). Besides the two LTRs there are other cis-acting sequences including the primer binding site (PBS); viral RNA packaging/dimerization signals (ψ and DIS); central polypurine tract (cPPT) and the central termination sequence (CTS), the latter leading to the formation, during reverse transcription, of a three-stranded DNA structure called central DNA Flap. In addition, there is the Rev response element (RRE) that is essential for post-transcriptional transport of unspliced and incompletely spliced viral mRNAs from

nucleus to cytoplasm and the purine-rich region (polypurine tract; PPT), which provides a second RNA primer for the initiation of plus strand DNA synthesis by virus-specific reverse transcriptase (*Srinivasakumar et al., 2001*). The *trans* elements of the HIV-1 provirus include nine open reading frames (ORFs) encoding for nineteen protein products (Table 1). The *gag-pol*, *gag* and *env* ORFs are encoding structural proteins and enzymes that are typical of all retroviruses. Additional ORFs code for essential regulatory proteins (*tat* and *rev* genes) and accessory proteins (genes: *vif*, *vpu*, *vpr* and *nef*). In almost all retroviruses, 5' LTR drives transcription of an initial genome-length RNA that also acts as an mRNA for translation of the viral Gag and Pol proteins. The Gag precursor protein (Pr55 is proteolytically cleaved by viral protease to yield the matrix (MA; p17), capsid (CA; p24), nucleocapsid (NC; p7), p6 and two spacer polypeptides: SP1 and SP2 (*Bukrinskaya, 2007*). The Gag-pol precursor protein (Pr160) is cleaved by the same protease resulting in the p6* transframe (TF) polypeptide and three viral enzymes: protease (PR; p11), reverse transcriptase with ribonuclease H (RNase H) activity (RT; heterodimer p66/p51 and RNase H; p15), and integrase (IN; p31). In more complex retoviruses including lentiviruses, this initial transcript is also processed into fully spliced transcripts encoding the Tat and Rev regulatory proteins as well as the auxiliary protein Nef. Alternatively, in HIV-1 this transcript can be processed into partially spliced mRNAs coding for three other accessory proteins: Vif, Vpu and Vpr (Vpx, a Vpr homolog present in HIV-2 and some SIVs is absent in HIV-1). These partially processed transcripts also contain the env ORF for envelope glycoprotein (Env, gp160). During maturation of the virus particle, cell proteases cleave the gp160 precursor to yield the gp120 surface subunit (SU) and the gp41 transmembrane subunit (TM) (*Bolinger & Boris-Lawrie, 2009*). The lipid-enveloped HIV virion measures approximately 80-120 nm in diameter and contains a number of host cellular proteins important for virus replication and pathogenesis.

Table 1. HIV-1 proteins and their functions (adapted from *Pluta K, Kacprzak MM, 2009*).

Category	ORF	Protein	Protein function
structural	*gag/* viral protease mediated cleavage of Gag polyprotein (Pr55) during virion maturation	matrix (MA; p17)	nuclear import of PIC, export of Gag-RNA complex to cytosol, membrane binding, virion and PIC formation
		capsid (CA; p24)	membrane binding and capsid formation
		nucleocapsid (NC; p7)	membrane binding, virion and PIC formation, RNA binding and cofactor of RT, genomic RNA selection, packaging and dimerization during virion assembly
		SP1	regulation of cleavage rate, membrane binding, virus budding and virion formation
		SP2	regulation of cleavage rate, membran binding, virus budding and virion formation
		p6	Vpr recruitment to virion, PTAP motif engaged in interactions with host proteins, membrane binding and virion formation
	gag-pol/ cleavage of Gag-pol polyprotein (Pr160) by PR during virion maturation	p6* (transframe polypeptide; TF)	stabilization of Gag-Gag-pol interactions and activation of PR-mediated cleavage
enzymes		reverse transcriptase and RNase H (RT: heterodimer p66/p51 and p15)	synthesis of viral cDNA (reverse transcriptase enzymatic activity with ribonuclease H activity located at C-terminus of p66 subunit
		Protease (PR; p11)	cleavage of p160, virion maturation
		Integrase (IN; p31)	nuclear import of viral PIC, viral cDNA integration into host genome
envelope glycoprotein	*env/* cleavage of Env (gp160) protein by cellular proteases	gp120 (SU)	interaction with host cell CD4 receptor and coreceptors
		gp41 (TM)	anchors gp120/gp41 complex in virus membrane, membrane fusion with host cell
regulatory protein	*tat*	transactivator (Tat; p16/p14)	initiation of transcription from viral LTR, transcriptional elongation, chromatin remodeling, regulation of apoptosis, modulator of host immune response
	rev	regulator of expression of viral protein (Rev; p19)	control of proviral DNA integration, stimulation of nuclear export of unspliced HIV RNAs, stimulation of translation of HIV RNAs
accessory protein	*nef*	negative factor (Nef; p27)	down-regulation of CD4 receptor, down regulation of MHC-I and MHC-II, signal transduction interference, enhancing infectivity (actin cytoskeleton rearrangement)
	vif	viral infectivity factor (Vif; p23)	inactivation of host immune response via binding apolipoprotein B mRNA editing enzyme (APOBEC3 proteins) and interferon regulatory factor, IRF3
	vpr	viral protein R (Vpr; p14)	Nuclear import of viral PIC, control of fidelity of reverse transcription, cell cycle arrest (block of G2/M transition), induction of apoptosis, transactivation of HIV LTR and host cell genes, modulation of host immune response, inhibition of cellular and viral pre-mRNA splicing
	vpu	viral protein U (Vpu; p16)	Degradation of CD4 receptor, release of viral particles, regulation of apoptosis

1.2 *HIV-1 life cycle*

The HIV replication cycle starts when the viral envelope glycoprotein gp120 binds to the CD4 receptor exposed on the surface of the host CD4-positive cells (Figure 2). After penetrating the cell membrane, the viral nucleoprotein core particle, is delivered into the cytoplasm, where the viral DNA is synthesized by reverse transcription of the viral RNA genome. Reverse transcription takes place in the reverse transcription complex (RTC), a nucleoprotein complex that is derived from the core of the infecting virion. Little is known about the properties and the component of the RTC. The RTC of HIV-1 contains the viral RNA genome, matrix protein (MA) and an accessory protein (Vpr), together with reverse transcriptase (RT) and integrase (IN) (*Iordanskiy et al, 2006*). The newly synthesized viral DNA remains associated with viral and cellular proteins in a large nucleoprotein complex called pre-integration complex (PIC). The PIC must be transported to the nuclear periphery and cross the nuclear envelope before the viral IN protein, one of the main component of the PIC, can catalyze the integration of the viral DNA into the host genome. The size of the PIC, that is estimated to be more than 50 nm, excludes passive diffusion as a viable mechanism for translocation, therefore active cellular transport mechanisms must be used for this task. In dividing cells, breaking and re-forming of the nuclear envelope during cell cycle allows an exchange of material between the nuclear and cytoplasmic compartments. Some retroviral PIC seem to take advantage of nuclear-envelope disassembly during cell division to gain access to chromatin. However, HIV-1 and other lentiviruses are able to infect non dividing cells (like macrophages and microglia cells) and must therefore cross an intact nuclear envelope. The ability to infect non-dividing cells makes lentiviruses attractive candidate for the development of retroviral gene vectors. The molecular mechanism by which retroviral PICs cross the nuclear envelope in non-dividing cells is still poorly understood. For HIV-1, a widely accepted model is that the nucleoprotein complex passes through the nuclear pore complex (NPC), that serves as a gate for traffic between the cytoplasm and nucleus. Many of the proteins of the HIV-1 PIC contain karyophilic[a] signals that cause them to be actively imported into the nucleus (*Suzuki, Craigie, 2007*). Understanding the mechanisms by which the PIC enters the nucleus, should provide new insights into cellular trafficking processes, the development of improved retroviral vectors for gene delivery and alternative strategies for blocking viral replication.

[a] Karyophilic proteins: proteins carrying signals that cause them to accumulate in the nucleus. This signal is usually a short peptide sequence called nuclear localization signal (NLS).

The late phase of the lentiviral life cycle starts when proviral DNA is transcribed by cellular RNA polymerase II (RNAP II). Before the integration step is completed, the infected cells are arrested in the G2 phase of the cell cycle. This favours viral replication because the promoter elements present in the LTR seem to be most active in G2 (*Gohn et al., 1998*). The block in the cell cycle progression is mediated by Vpr via indirect inactivation of the Cdc2 (CDK1)/cyclin B (CycB) kinase and Cdc25 phosphatase, that are regulators of the G2/M transition. For initiation of its transcription, HIV-1 usurps distinct cellular signaling pathways by using nuclear factor kappa-B (NF-kB) and nuclear factor of activated T cells (NFAT). NF-kB and NFAT binding sites along with binding sites for other cell-specific and constitutively expressed cellular transcription factors, are located in the U3 enhancer/promoter region of 5' LTR, upstream from the TATA box. The Tat protein is the main lentivirus-encoded transcription transactivator. Tat interacts with the *cis* elements of the virus located in the junction between the U3 and R region of the proviral 5' LTR, to increase the processivity of the RNAP II. Unlike typical transcriptional transactivators, Tat does not binds to DNA, but to an RNA bulge of a stem-loop structure present at the 5' end of all HIV-1 transcripts, known as the transactivation response element (TAR). Thus, Tat requires the initial transcription of TAR before it can stimulate further HIV transcription (*Barboric & Peterlin, 2005*). Transcription of viral RNA begins at the first nucleotide of the R region in the 5' LTR and polyadenylation occurs at the last nucleotide of R in the 3' LTR. HIV-1 uses the cellular splicing machinery to express its genes (Figure 2F). Cellular factors, in concert with both positive and negative *cis* elements within the viral genome, act to promote or repress splicing (*Jablonski & Caputi, 2009*). In addition to full-length genomic mRNA (about 9.2 kb in length), coding for the Gag and Gag-pol polyproteins, a number of various subgenomic species of mRNAs have been described in HIV-1-infected cells, with several slightly different mRNAs coding for the same protein. Of these, two classes of differentially spliced mRNAs can be detected in the cytoplasm: early, fully spliced (about 1.8 kb) and late, partially spliced transcripts (about 4 kb). The former encode the viral regulatory proteins Tat and Rev, along with the auxiliary protein Nef; the latter, accessory proteins: Vif, Vpu, and Vpr as well as the envelope protein Env. Under physiological conditions eukaryotic cells block the nuclear export of unspliced mRNAs. Since lentiviruses exploit a variety of differentially and incompletely spliced transcripts, including unspliced genomic RNA, for expression of the complete repertoire of their proteins, they have developed mechanisms to escape this tight export control. The key protein in this process is the regulator of expression of viral proteins, Rev.

Rev contains both a nuclear localization signal (NLS) and a nuclear export signal (NES), hence it shuttles between the nucleus and the cytoplasm (Figure 2G). Rev, produced from early transcripts, accumulates in the cell and binds to the Rev response element (RRE), a large (roughly 250 nucleotides in size) highly structured RNA element that is located in the *env* gene and is present in all unspliced and partially spliced HIV-1 RNAs (*Pond et al., 2009*). RRE constitutes a scaffold for Rev multimerization and the nascent directly binds CRM1, a member of the karyopherin family of nucleocytoplasmic transporters.

The synthesis of viral proteins is regulated by an interplay between host factors and cis-acting elements present in the viral transcript. For example, a weak Kozak sequence surrounding the HIV-1 *vpu* AUG start codon promotes translation of the downstream *env* gene from bicistronic *vpu/env* transcripts, a process referred to as leaky scanning transcripts. However, more recent studies showed that the different 5' untranslated regions (UTRs) present in alternatively spliced HIV-1 *vpu/env* mRNA isoforms affected Vpu synthesis, but not the downstream Env synthesis. Therefore, HIV-1 Env may also use an internal ribosome entry sequence (IRES) or the ribosome shunt translation mechanism to maximize Env production (*Anderson et al., 2007*). The sequences coding for Gag and Pol proteins, translated from a full-length unspliced RNA, are in different reading frames. HIV-1 requires a single frameshift to produce the Gag-pol polyprotein. The frameshifting occurs at an approximate rate of one Gag-pol for every twenty Gag molecules synthesized. The signal for programmed frameshifting is provided by viral sequences including an upstream slippery sequence. For HIV-1 this sequence is AAUUUUUU *(Watts et al., 2009)*. Another sequence engaged in the frameshifting is an RNA stem structure located downstream of the frameshift site. Interaction of tRNA with the ribosome, was postulated to be involved in this process (*Bolinger & Boris-Lawrie, 2009*). In addition to the translational mechanisms described above, retroviruses also employ cap-dependent translation enhancers. Post transcriptional control elements (PCEs), located in the 5' UTRs, have been found in many viruses, including HIV-1. PCEs are stem-loop RNA structures that specifically interact with RNA helicase A (RHA) and facilitate polysome loading and efficient viral protein production. Also other cellular proteins can bind viral RNA and affect its translation. A direct interaction between TAR and TAR-RNA binding protein (TRBP) enhances synthesis of viral proteins. The correlation between the organization of the RNA and sequences encoding inter-domain loops in HIV-1 proteins suggest that RNA structure modulates the speed of ribosome during elongation in order to promote proper

protein folding. Thus differences in RNA sequence accessibility represent additional elements to control the structure and function of HIV-1 proteins.

Production of HIV-1 particles by infected cells comprises a series of coordinated events: Gag dimerization and multimerization, binding of Gag complexes to genomic viral RNA (gRNA), transport of the Gag-RNA complexes, Gag-pol, Gag and Env to the site of assembly, and subsequent release of immature virions which later undergo maturation mediated by viral protease to become fully infectious (Figure 2H-K). The Gag precursor protein, Pr55, is responsible to balance translation of the full length primary transcript (pre-mRNA) and its encapsidation (*Anderson & Lever, 2006*). Viral structural proteins and the newly synthesized viral RNA molecules migrate to the site of HIV-1 assembly stabilized both by actin and tubulin cytoskeleton. Again, Pr55 plays a central role in virus assembly and is sufficient for production of non-infectious virus particles in absence of other viral proteins. Unlike other enveloped viruses, HIV-1 leaves the cell via one of two budding pathways depending on the cell type from which it is existing. HIV released from T lymphocytes predominantly buds through the cell membrane and acquires its envelope components from the plasma membrane, whereas virus replicating in macrophages usually exits the cell via vescicles of the endosomal network (*Bukrinskaya, 2007*). Lentiviruses, as many other enveloped viruses, complete their replication cycle by budding through the cellular membrane (Figure 2I). An alternative and very efficient way for virus spreading is cell-to-cell transmission. This transmission takes place at intercellular contact sites, so-called virological synapses (VS) that resemble immunological synapses. In the case of HIV-1, the VS is created by the recruitment of CD4 receptor, CXCR4 and CCR5 coreceptors, and lymphocyte function-associated antigen 1 (LFA-1) from the uninfected (target) cell and viral Env, Gag and adhesion molecules (intracellular adhesion molecule, ICAM-1) in polarized, lipid raft-like patches on the effector (donor) cell (*Wang et al, 2009b*). However, it has been demonstrated recently that HIV transmission between T cells occurs efficiently in absence of adhesion molecules (lack of LFA-1 and ICAM-1 interactions) and that Env-CD4 interactions are the main driving force of the VS creation between infected and non-infected CD4+ T cells (*Puigdomènech et al., 2008*). Initially, virus particles are released from the infected cell in an immature form; shortly after virus budding from the plasma membrane, viral protease cleaves the Gag and Gag-pol polyprotein precursors to generate the mature Gag and Pol proteins. This process takes place as an ordered cascade of cleavage reactions leading to the conversion of the immature virus particle into an infectious virion (Figure 2A).

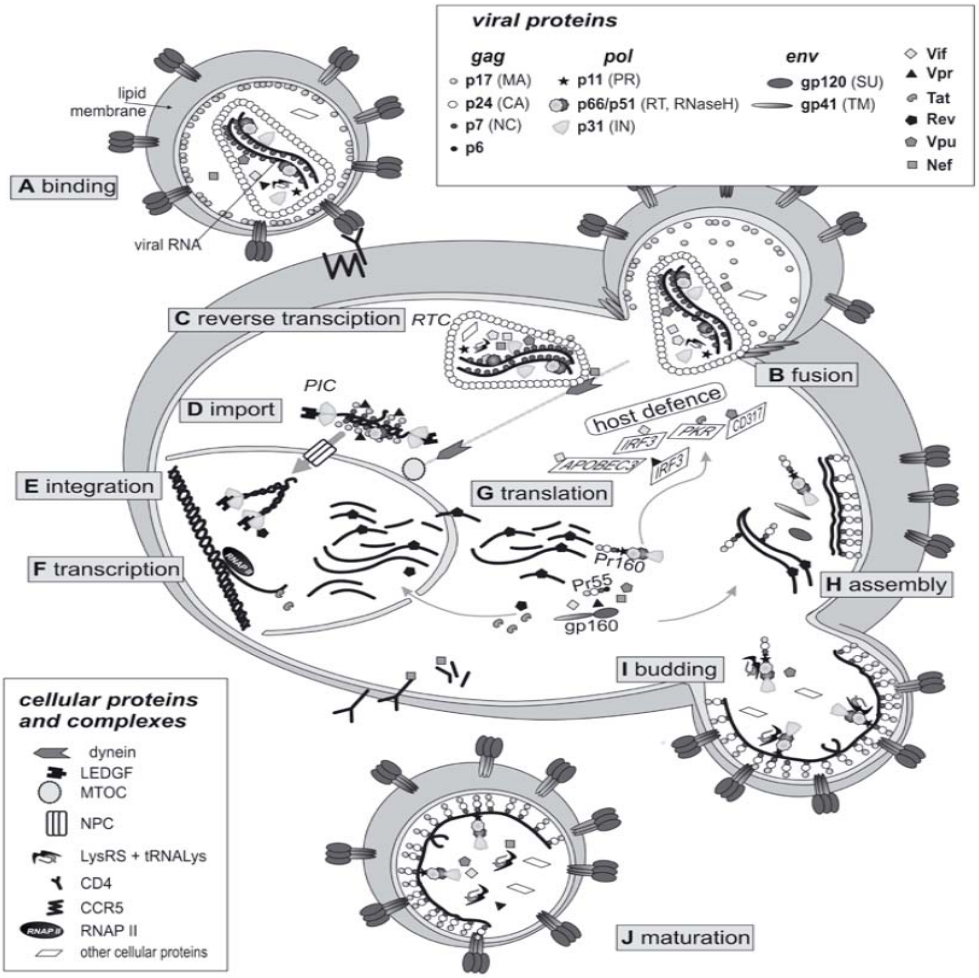

Figure 2. HIV-1 life cycle. Virus replication starts when the viral envelope glycoprotein gp120 binds to the CD4 receptor and a secondary receptor (CXCR4, CCR5) on the host cell surface (A). After binding to CD4, conformational changes in non-covalently associated gp41 subunit releases free energy sufficient to promote fusion of the virus particle with cell membrane. The viral core is then spilled into the cytoplasm (B). Core movement is facilitated by Nef protein which induces actin rearrangements. Reverse transcription of viral RNA occurs within CA capsid. Cellular protein dynein transports the reverse transcription complex (RTC) towards the nucleus along microtubules anchored to microtubule organizing center (MTOC) (C). When the reverse transcription is completed, provirus DNA and several viral proteins such as RT, IN, NC, Vpr, and MA, together with cellular proteins, form preintegration complex (PIC), which now can be actively transported into the nucleus via nuclear pore complex (NPC) (D). Proviral DNA is integrated into the host genome (E) by integrase, which cooperates with cellular protein LEDGF. Transcription of provirus is conducted by cellular RNAP II and enhanced by viral protein Tat. Unspliced and partially spliced transcripts are stabilized by viral protein Rev (F), which enables their export to the cytoplasm and subsequent translation (G). Some of the synthesized viral proteins participate in combat against host defense. Full-length RNA molecules, Gag, Gag-pol and Env proteins, start to assemble new virus particles via Gag dimerization and multimerization, binding of Gag complexes to genomic viral RNA (gRNA) and association with cellular membrane. Env precursor gp160 is cleaved by cellular proteases to give gp120 and gp41 subunits that are directed to the cell membrane by Gag (H). Initially, virus particles are released (I) from infected cell as an immature form (J). Digestion of polyproteins by viral protease leads to virus maturation (A). (Adapted from *Pluta K, Kacprzak MM, 2009*)

1.3 *Construction of safe HIV-1 based lentiviral vectors*

Because of their ability to transduce non dividing cells, lentivirus-based vectors have great potential for delivering therapeutic genes to cells. However, because of safety concerns due to the potential for horizontal and cross-species transmission of recombined chimeric lentiviruses, the design of such vectors should include fundamental safety principles. To address such safety issues several monitoring assays have been developed to test for the presence of replication competent lentiviruses (RCLs) in vector preparations, the possibility of vector mobilization from transduced cells, the persistence of vector-positive cells, and the abnormal clonal expansion of vector-modified cells (*Sastry & Cornetta, 2009*). Generally, to minimize risk associated with manufacturing and use of lentiviral vectors, all non-essential genes coding for accessory proteins and responsible for virulence, should be removed from the vector sequence. Additionally, the vector genome is usually split into several parts with limited sequence overlap in order to reduce to a minimum the possibility of recombination, vector mobilization and the generation of RCLs. The development of lentiviral vector systems is based on the concept of separating the *cis*-acting sequences that are essential for vector RNA synthesis, packaging, reverse transcription and cDNA integration, from the *trans* elements that encode viral enzymes as well as structural and accessory proteins. Hence, such a system typically consist of: a packaging expression cassette(s) (helper), a vector cassette (transfer vector), and an envelope expression cassette (Figure 3A-C).

1.3.1 *Packaging expression cassette(s)*

The packaging cassette expresses viral enzymes and structural proteins necessary for infectious particle formation, with the exception of Env. The original first-generation lentiviral vector system developed in 1996 contained the accessory proteins Vpu, Vpr, Vif, Nef, and the Rev and Tat regulatory proteins as functional components of the packaging plasmid. Further refinements included eliminating the LTRs as well as ψ and PBS sequences (*Naldini et al., 1996a; 1996b*). This prevented the packaging of full-length mRNA encoding trans elements into nascent vector particles. However, RRE and the 5' SD site remained unchanged, allowing normal mRNA processing and Rev-dependent export from the nucleus. In the absence of native LTRs, RNA synthesis was driven by promoters derived from other viruses, usually CMV (Cytomegalovirus) or RSV (Rous sarcoma virus), whereas the polyA signal was adopted from SV40 or insulin gene (Figure 3A). These packaging constructs were subsequently refined by eliminating all accessory

proteins that are associated with virulence and cytotoxicity (Table 1) and are not required for virus replication *in vitro* (*Gibbs et al., 1994*), eventually leading to the creation of the second-generation packaging cassette (*Zufferey et al., 1997; Kim et al., 1998;*). In this system only Rev and Tat proteins were expressed together with the Gag and Gag-pol polyproteins (Figure 3A).

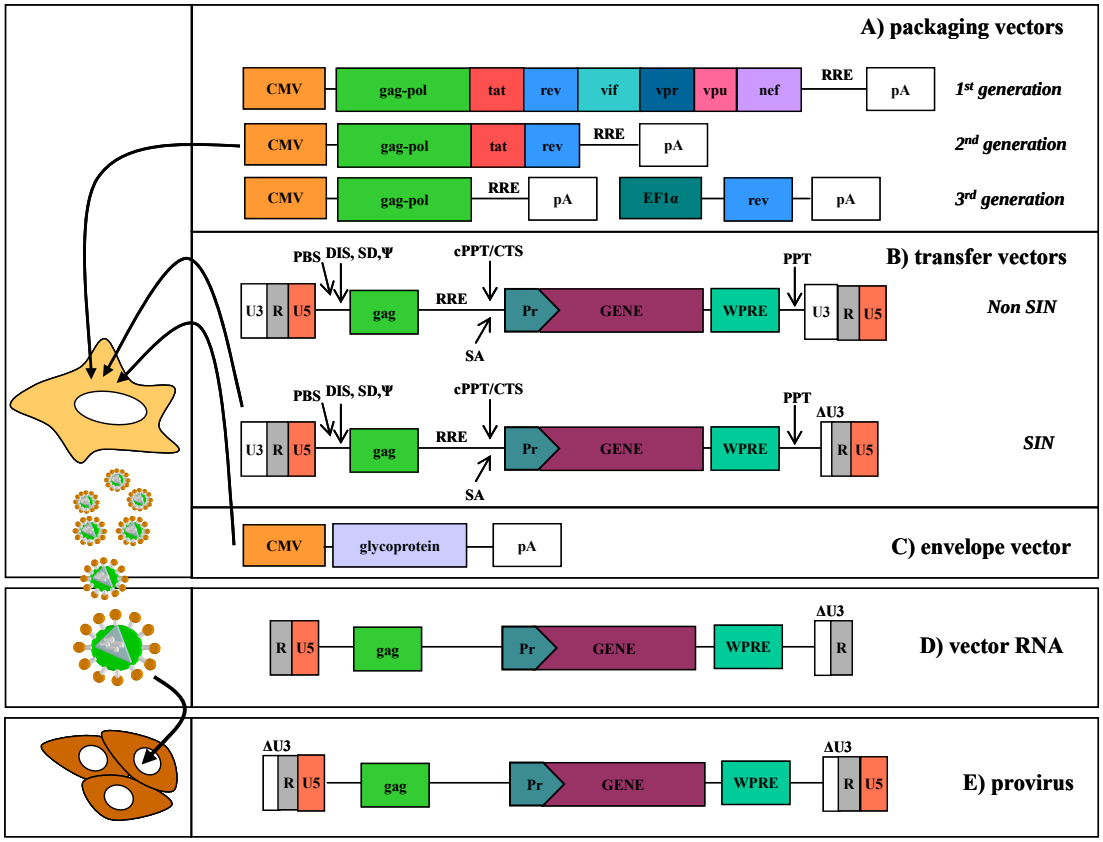

Figure 3. Lentivirus-based gene delivery system.
A-C. Plasmids used for transient transfection of producer cells. **A**. Packaging vector plasmid. First-generation vector contains all regulatory and accessory viral genes expressed from CMV promoter. Second-generation vector encodes only Tat and Rev proteins. Third generation packaging system consists of two plasmids: one encoding Gag and Gag-pol poly-proteins, the second - Rev protein. **B**. Transfer vector plasmids. In non self-inactivating (SIN) vectors viral RNA is expressed from intact 5'LTR. SIN vectors bears deletion in U3 region (ΔU3), which inactivates transcription of entire viral RNA after provirus integration. **C**. Envelope vector. Depending on glycoprotein used, different viral pseudotypes are formed. **D**. Schematic representation of vector genomic RNA. In presented setting, cotrasfection of either second-generation packaging vector or third generation vectors along with SIN transfer vector and envelope vector results in the formation of viral particles that contain dimeric RNA bearing ΔU3 mutation. **E**. After provirus integration into host cell DNA, transcription from mutated 5' LTR (duplicated 3' ΔU3) is abrogated. **Abbreviations**: CMV, cytomegalovirus immediate-early promoter; EF1α, human elongatin factor 1-α promoter; gag, 5' portion of *gag* gene containing dimerization/packaging signals; PBS, primer binding site; DIS, dimerization signal; SD, splice donor site; SA, splice acceptor site; ψ, packaging signal; cPPT, central polypurine tract; CTS, central termination sequence; RRE Rev response element; PPT, polypurine tract; pA, polyadenilation signal; ΔU3, SIN deletion in U3 region of 3' LTR; Pr, internal promoter for transgene expression; WPRE, woodchuck hepatitis virus (WHV) post-trasncriptional regulatory element.

In order to further improve the biosafety of the system, Rev was later placed on a second plasmid, while Tat was completely removed. The Tat function was replaced using modified 5' LTR enhancer/promoter elements containing strong, constitutive RSV or CMV-derived promoters in the corresponding vector constructs (Fig. 3A). This systems is referred to as a third-generation packaging cassette (*Dull et al., 1998*). Importantly, all these manipulations did not substantially affect vector production (titer) or infectivity. Moreover, the number of recombination events needed for the potential generation of an RCL increased considerably, and potential RCLs would not contain any proteins involved in wild-type virus virulence and pathogenicity. However, homologous recombination was still possible since there were overlapping HIV-1 sequences present in the helper and vector plasmids. These sequences encompass the RRE *cis*-element and a portion of the *gag* gene. RRE cannot be easily eliminated because it is necessary for efficient expression of HIV-1 *gag* and *pol* genes from the packaging construct. The reason for this requirement is low GC content and suboptimal codon usage in wild-type HIV mRNA. This causes instability of the RNA, which can be rescued from degradation by Rev binding. Codon optimization of the packaging cassette by changing the codon bias of the HIV-1 *gag-pol* to that of the human genome removed the homology with *gag* portion present in the transfer vector and led to efficient protein synthesis in the absence of Rev. This enhancement was shown to be mediated by increased mRNA stability and transport (*Ngumbela et al., 2008*). Thus, it allowed construction of packaging cassettes that do not contain an RRE. It was also shown that production of the Gag-pol polyprotein from the codon-optimized mRNA is Rev-independent and that this RNA does not use the CRM-1-mediated nuclear export pathway. However, proper vector cassette expression remains Rev-dependent due to the requirement of a portion of *gag* for efficient packaging of the vector RNA. A complete removal of Rev from the system, even if the helper plasmid lacked the RRE sequence, resulted in a significant (five-fold) reduction of vector titers when compared with the Rev/RRE containing systems (*Kotsopoulou et al., 2000*). Safety improvement in the design of vector packaging systems is still an area of intensive research.

1.3.2 *Transfer vector*

The vector cassette expresses full-length vector mRNA which contains all *cis*-acting elements necessary for efficient packaging, reverse transcription, nuclear import and integration into the host genomic DNA. Typically, the vector plasmid contains a transgene expression cassette with the gene of interest driven by an internal promoter, usually

positioned between the 3' Tat/Rev SA site and the 3' LTR (Figure 3B). The design of vector cassette has evolved in parallel with the development of the packaging cassette. Early constructs contained intact 5' and 3' LTRs, and, thus, transcription was Tat-dependent. (*Naldini et al., 1996a; 1996b*). Tat-independent transcription of the vector cassette was accomplished by replacement of the enhancer/promoter sequences in the U3 region of 5' LTR with a strong heterologous promoter derived from CMV (fig. 3B) (*Kim et al., 1998*). Such transfer vectors are now an integral part of the third-generation packaging systems lacking Tat. Deletion of the enhancer/promoter sequences in the U3 region of 3' LTR (120–400 bp; ΔU3) has led to the creation of so-called self-inactivating (SIN) vectors (fig. 3B) (*Dull et al., 1998; Iwakuma et al., 1999*). When viruses derived from such transfer constructs are used to infect cells, the deletion is reproduced in the 5' LTR during reverse transcription, resulting in the transcriptional inactivation of the provirus (Figure 3D, E). The most important improvement of viral vector safety offered by the SIN design is the reduction in expression of full-length transcripts available for packaging, which minimizes the possibility of vector mobilization upon superinfection with wild-type HIV-1 (*Bukovsky et al., 1999*). The chance of RCL formation is also reduced due to further elimination of sequences homologous to wild-type virus. Additionally, SIN deletion should make insertional mutagenesis less likely, preventing the transcriptional interference with promoter/enhancer elements present in the host genome. Actually, in the recently developed cell culture assay for insertional mutagenesis, SIN deletion of the promoter/enhancer region from HIV-based vector LTRs resulted in a complete abrogation of the insertional gene activation, frequently observed for both lentiviral and gamma-retroviral vectors (*Bokhoven et al., 2009*). Another possible source of viral genomes available for packaging are episomal forms of the vector DNA that circularizes and remains unintegrated in transduced cells. Such episomal forms of viral DNA, single- and double-LTR (1-LTR and 2-LTR) circles, are diluted upon cell division (*Van Maele et al., 2003*), but in nonproliferating, quiescent cells they are stable and transcriptionally competent (*Gillim-Ross et al., 2005*) and, thus, can considerably contribute to the availability of full-length gRNAs for packaging and mobilization.

1.3.3 *Envelope expressing cassette*

The last element of a lentiviral vector system is an envelope cassette, which provides nascent particles with envelope glycoprotein (Figure 3C). Since the genuine lentiviral envelope gene is removed from the system, an additional plasmid expressing a

heterologous glycoprotein is used during vector production. Such an approach, called pseudotyping, offers several significant advantages: 1) increase in vector safety due to elimination of sequence homology with wild-type virus, 2) expanded or selective specificity of pseudotyped vector tropism towards target cells, 3) improvement of particle stability allowing virus concentration and long-term virus stock storage. The glycoprotein of choice for the vast majority of vector systems currently in use is the VSV-G (vesicular stomatitis virus) envelope glycoprotein. It offers unequaled titers (up to 10^7 transducing units per ml, TU/ml, before concentration), excellent stability of the virions and allows virtually all mammalian cell types to be infected. The VSV-G envelope facilitates vector entry via the endocytic pathway, which diminishes the requirement for viral accessory proteins for full infectivity (i.e., Nef) (*Aiken, 1997*). However, the mechanism responsible for cell binding as well as cellular receptors for this glycoprotein remain unknown.

1.3.4 *Further vector improvement*

Since HIV-1 virus can effectively package approximately 9.2 kb of its genome, reduction of the length of the viral vector backbone should increase its payload carrying capacity. In the work published by Cui et al. (1999), subsequent mutations and deletions of the viral regulatory elements, including SD, most of gag, and RRE, have brought down the content of the original HIV-1 sequences in the vector DNA to less than 550 bp. This should allow accommodation of around 9 kb of foreign DNA. Yet the usefulness of such design was diminished by a decrease in vector titer, about half that of the wild-type construct. In another detailed study on lentiviral vector capacity, maximal vector size (including 1.6–2.2 kb of the virus backbone) has been estimated to approx. 13.5 kb (*Kumar et al., 2001*). Thus, the design of a safe and efficient lentiviral vector requires both deletions of non-necessary sequences from the backbone and insertions of elements that are proven to have a positive effect on vector titer or transgene expression (Figure 4). For example, the insertion of the cPPT (central polypurine tract) and CTS (central termination sequence) elements, that are implicated in HIV-1 reverse transcription and nuclear import of viral PIC, into HIV-based vectors strongly stimulates gene transfer efficiencies (up to ten-fold) in all cell types tested, both *in vivo* and ex vivo (*Van Maele et al., 2003*; *Ao et al., 2004*; *Logan et al., 2004*).

Another commonly used cis-acting element is the woodchuck hepatitis virus (WHV) post-transcriptional regulatory element (PRE), WPRE. Incorporation of this sequence (600 bp) in the 3' untranslated region of a transcript increases overall transgene expression by more than six-fold (*Oh et al., 2007*). One should note that the WPRE element contains an ORF

encoding a truncated peptide of the WHV X protein (WHX), which might be associated with hepatocellular carcinomas. However, neither expression of this peptide from a lentiviral vector nor its direct involvement in oncogenesis have been demonstrated. Besides, the potentially harmful sequences can be successfully eliminated from WPRE without negative effects on its functionality, which enables designing safe and efficient vectors for *in vivo* applications (*Schambach et al., 2006*). The exact role of WPRE in vector performance is unclear. One of its functions may be the reduction of viral readthrough transcription and, thus, improvement of transcript termination leading to higher titers and transgene expression (*Higashimoto et al., 2007*). Post-transcriptional regulatory elements may additionally enhance intronless transgene expression through stabilization of the 3' end of the transcript and improvement of the nucleocytoplasmic export of unspliced mRNAs. Also other than WPRE, different cis-acting elements (e.g., hepatitis B virus (HBV) PRE, HPRE) can be considered useful for *in vivo* applications (*Sun et al., 2009*).

Due to deletions in the U3 region containing polyadenylation "enhancers", SIN vectors suffer from leaky transcription termination increasing the probability of readthrough into cellular genes, potentially oncogenes. Hence, incorporation of strong polyA signals is of great importance for safe vector design. An attempt to solve this problem was made by Schambach et al. (2007). Insertion of a 100 bp sequence, comprising a recombinant direct repeat of the upstream polyadenylation enhancer element (or upstream sequence element, USE) derived from SV40, improved both the titer and gene expression from a viral vector. Additionally, this element suppressed readthrough more efficiently than WPRE did and was able to substitute for the WPRE functions to some extent. Furthermore, replacement of the original polyA signal in the R/U5 regions of 3' LTR with a bovine growth hormone polyadenylation (bGHpA) sequence significantly elevated the efficiency of SIN vectors (*Iwakuma et al., 1999*). The bGHpA signal was also successfully used to enhance expression of the second gene in vectors carrying multiple genes (*Osti et al., 2006*).

In order to minimize the potential of transcription of integrated SIN vectors from cryptic promoters, either within or upstream of the vector genome, a 1.2 kb fragment of the insulator element from the locus control region of the chicken β-globin gene, chicken hypersensitive site 4 (cHS4) sequence (*Chung et al., 1993*), was inserted into the vector LTRs. This resulted in a significant (to one-fourth of the control level) reduction in the full-length vector transcription (*Hanawa et al., 2009*). On the other hand, insertion of insulator sequences into vectors can help maintain long-term transgene expression by

suppression of chromosomal position effects (transgene silencing) resulting from integration into the host chromatin (*Hino et al., 2004; Arumugam et al., 2007*).

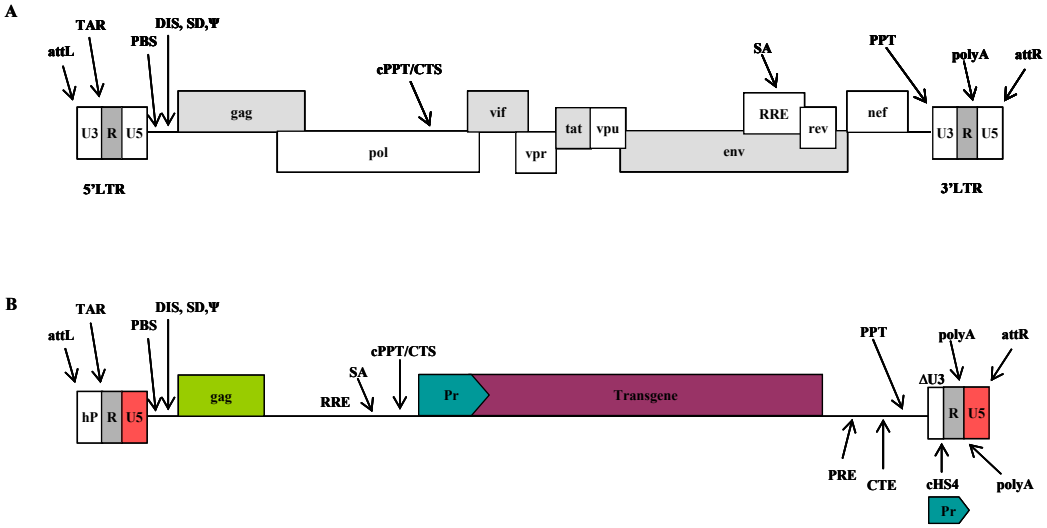

Figure 4. Comparison of wild-type HIV-1 provirus and optimized lentiviral vector.
A. Schematic representation of HIV-1 provirus. **B**. Typical cis-acting elements are marked above the vector. Below — examples of additional modifications introduced to Lentivirus-based vectors in order to elevate titer and/or transgene expression levels. After transfection, viral genomic RNA is synthesized using strong hybrid promoter placed at 5' end of vector. Elimination of weak promoter/enhancer from 5' LTR enables Tat-independent transcription and use of third-generation packaging system. Incorporation of post-transcriptional control element improves translation efficiency. In SIN (self-inactivating) vectors deletion in U3 region of 3' LTR (ΔU3) is duplicated during reverse transcription and integration, which results in abrogation of viral RNA transcription. ΔU3 site can be used for insertion of chromatin insulator or transgene cassette and, thus, these elements will be duplicated in provirus. To improve polyadenylation of SIN vector transcript, heterologous or synthetic polyadenylation signals can be inserted into R region of 3' LTR. Post transcriptional regulatory elements and RNA transport elements can be incorporated into 3' end of transgene to reinforce transcript 3' end formation, RNA stability and nuclear export. Abbreviations: hP, hybrid promoter; attL, attR, left and right attachment sites; TAR, transactivation response element; gag, 5' portion of gag gene containing dimerization/packaging signals; PBS, primer binding site; DIS, dimerization signal; SD, splice donor site; SA, splice acceptor site; ψ, packaging signal; cPPT, central polypurine tract; CTS, central termination sequence; RRE, Rev response element; PPT, polypurine tract; polyA, polyadenylation signal; ΔU3, SIN deletion in U3 region of 3' LTR; P, internal promoter for transgene expression; PCE, post-transcriptional control element; PRE - posttranscriptional regulatory element; CTE, constitutive transport element; cHS4, chicken β-globin insulator; polyA, polyA signal.

1.3.5 *Non integrating lentiviral vectors (NILVs)*

Long term and stable transgene expression, as a result of lentivector integration is very useful for diseases in which permanent cell correction is sought after. However, as already mentioned in the introduction, insertion of the gamma-retrovirus-based vector led to oncogene transactivation and leukaemia in some X-SCID children treated by gene therapy. Therefore, the development of integrase defective lentiviral vectors (IDLVs) has attracted much attention.

Accordingly, selected mutations within the IN-coding region of the packaging plasmid have been exploited to generate IDLV. These mutations eliminate IN activity without affecting reverse transcription and transport of the pre-integration complex to the nucleus. Then, the lentivector DNA remains in the nucleus as an episome, leading to sustained expression in post-mitotic cells and tissues such as retina, brain and muscle (Yanez-Munoz et al., 2006; *Apolonia et al., 2007*). Thanks to their episomal nature (integration frequencies are 3–4 logs below those of their integrating counterparts), the IDLVs have a greatly reduced risk of promoting insertional mutagenesis. Moreover, since active IN is required for the viral replication, nonintegrating vectors should not support the HIV-1-based replication if an unlikely recombination event generated an RCL.

As part of the nucleoprotein pre-integration complex, the IN bound to the attachment (*att*) regions of the LTRs mediates a covalent linkage of the viral double-stranded DNA to the cell genome through DNA cleavage and joining activities. There are at least two ways to impede integration: i) by mutating the integrase or ii) by mutating the att region of the LTRs. In 1999 Alan Engelman proposed distinguishing between two classes of integrase mutations: i) class I mutations that specifically block the integration, and ii) class II mutations with pleiotropic repercussions on the viral cycle (*Engelman A., 1999*). Thus, class I mutations do not affect reverse transcription of the viral genome or its translation into the nucleus, while class II mutants may display defects in other steps of the viral cycle, including for example, reverse transcription, nuclear import, viral genome transcription or particle assembly which further affects viral replication. Many different types of IN mutation cause class I phenotypes but with diverse consequences for the efficiency of transduction and most importantly for the rate of residual integration. The HIV-1 IN (288 residues) has three main functional domains, i) the N-terminal domain, formed by the first 60 residues interacts with the viral LTR, ii) the catalytic domain (from residues 61 to 212) contains the three conserved amino acid D-D35E (D64, D116 and E152 in HIV-1) and iii) the C-terminal region that has a non-specific DNA-binding function. Integrase deficient

vectors can be obtained by introducing mutations into any of these regions. Among the mutations studied, the D64V substitution (DDE) in the catalytic domain has been the most widely used because it has been well characterized in HIV-1 and shows the strongest inhibition of the IN without affecting viral DNA synthesis (*Sarkis et al., 2008*). It has been reported that in lentiviral vectors this mutation allows a transduction efficiency only slightly lower than integrative vectors but a residual integration that is about 1000-fold lower than an integrative vector at low vector doses. In another report a D116N mutation (DDE) results in residual integration about 2000 times lower than control vectors (*Leavitt et al., 1996*). However, it is difficult to establish whether D116N gives a more accentuated non-integrating phenotype than D64V and simultaneous studies of both these vectors suggest that they have similar phenotypes. Another way to prevent integration is to modify the attachment sequence of the IN to the virus. IN bound to the *att* regions of the LTRs first removes two bases from each 3' extremity of the provirus. The result is that at each end of the provirus there is a 5' cohesive end and a conserved 3' CA dinucleotide which is the substrate for viral attachment to recipient DNA. Mutating the CA dinucleotide in U3 and U5 of the LTRs (Δatt) inhibits viral integration. In HIV derived vectors, the Δatt mutation has been reported to affect integration as much as a D64V integrase mutation: integration was reduced by 1000 to 10000 times. However Apolonia et al. found that Δatt was about 10 times more integrative than D64V and 100 times less integrative than control integrative vectors (*Apolonia et al., 2007*). These studies agree, however that the two mutations were not, or only slightly, additive suggesting that the D64V mutation supersede the Δatt one.

Another issue for IDLVs concerns the strength of transgene expression. Non integrating vectors were initially considered to be non-functional because of a lack of transgene expression, unlike their integrating counterparts; this led to the view that episomal forms of lentiviral vectors could not be transcribed. More recent comparisons showed that transgene expression from IDLVs was 2 to 10 fold lower than that from integrated vectors in cell lines (*Apolonia et al., 2007; Cornu & Cathomen., 2007*) but equivalent *in vivo* (*Yanez-Munoz et al., 2006*). Furthermore, although IDLV episomal DNA expresses a transgene, this vector form dissipates with cell division. The duration of transgene expression varies depending on how rapidly transduced cells turn over. The reason why vector's episomes do not persist in dividing cells is because they lack an origin of replication (Ori). By including the SV40 Ori in IDLVs, transgene expression lasted 56 days, the final time point examined, in the SV40 T-antigen containing 293T cell line. This contrasted to parallel transductions in 293

cells, with no SV40 T-antigen, in which transgene expression was short lived (*Vargas et al., 2004*). Potentially, this approach could be used with any origin of replication derived from an episomal virus like BK and JC polyomavirus, human papillomavirus (HPV) and Epstein-Barr virus (EBV). For the purpose of the study described in this thesis, we used the latent origin of replication from EBV. Epstein Barr virus biology will be described in the next chapter.

Chapter 2

EPSTEIN BARR VIRUS BIOLOGY

2.1 *Epstein Barr virus Genome*

Epstein Barr virus (EBV), formally designated human herpesvirus 4 (HHV-4), is one of the eight known human herpesviruses. Like those of other herpesviruses, EBV virions have a double-stranded, linear DNA genome surrounded by a protein capsid. A protein tegument lies between the capsid and the envelope, which is embedded with glycoproteins that are important for cell tropism, host range, and receptor recognition. Mature virions are approximately 120 to 180 nm in diameter (*Oludare et al, 2011*). EBV preferentially infects B lymphocytes through the binding of the major viral envelope glycoprotein gp350 to the CD21 receptor on the surface of B cells, and through the binding of a second glycoprotein, gp42, to human leukocyte antigen (HLA) class II molecules as a co-receptor. Infection of other cell types (principally epithelial cells) is much less efficient and occurs through separate, as yet poorly defined, pathways. Importantly, EBV has the unique ability to transform resting B cells into permanent, latently infected lymphoblastoid cell lines (LCLs), an *in vitro* system that has provided an invaluable, albeit incomplete, model of the lymphomagenic potential of the virus. By contrast, infection of epithelial cells *in vitro* does not activate the full growth-transforming programme of the virus, and rarely — if ever — achieves full lytic replication. B-cell transformation to LCLs therefore remains the dominant *in vitro* model of infection. In EBV-transformed LCLs, every cell carries multiple extrachromosomal copies of the viral episome (Figure 5) and constitutively expresses a limited set of viral gene products, the so-called latent proteins, which comprise six EBV nuclear antigens (EBNAs 1, 2, 3A, 3B, 3C and -LP) and three latent membrane proteins (LMPs 1, 2A and 2B) (FIG. 1b). Transcripts from the *Bam*HIA region of the viral genome (so-called BART transcripts) are also detected in LCLs. In addition to the latent proteins, LCLs also show abundant expression of the small, non-polyadenylated (and therefore non-coding) RNAs, EBER1 and EBER2; the function of these transcripts is not clear, but they are consistently expressed in all forms of latent EBV infection (*Young & Rickinson, 2004*). This pattern of latent EBV gene expression, which appears to be activated only in B-cell infections, is referred to as ‘latency III’ (Figure 5).

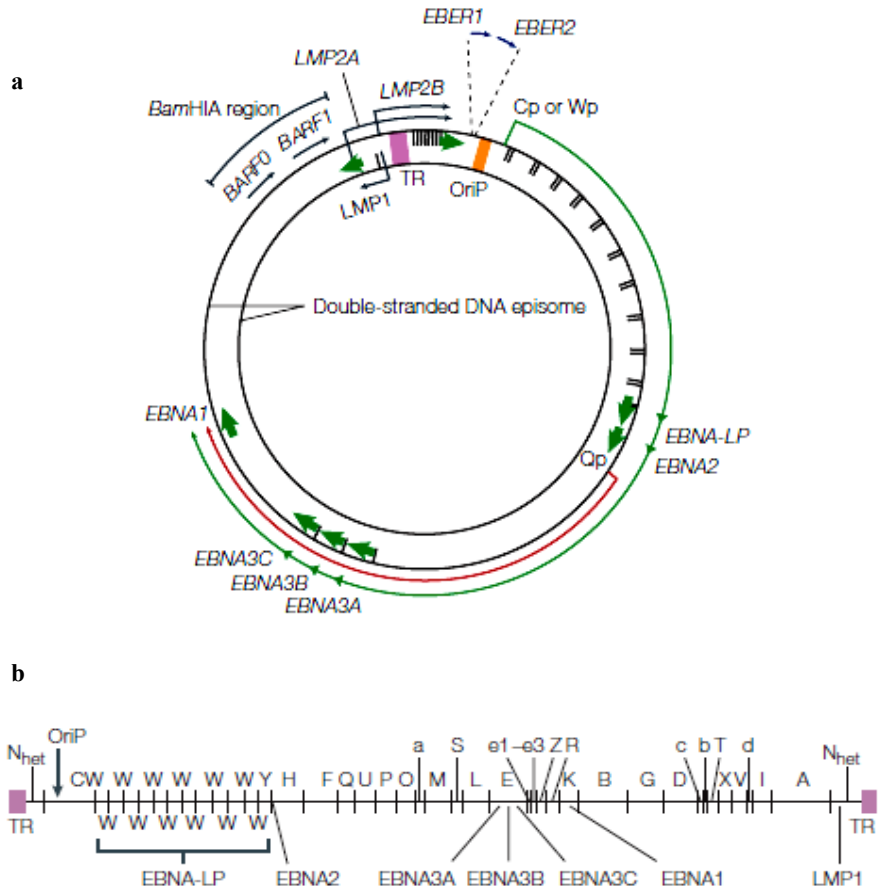

Figure 5. The Epstein–Barr virus genome. a. Diagram showing the location and transcription of the EBV latent genes on the double-stranded viral DNA episome. The origin of plasmid replication (*OriP*) is shown in orange. The large green solid arrows represent exons encoding each of the latent proteins, and the arrows indicate the direction in which the genes encoding these proteins are transcribed. The latent proteins include the six nuclear antigens (EBNAs 1, 2, 3A, 3B and 3C, and EBNA-LP) and the three latent membrane proteins (LMPs 1, 2A and 2B). EBNA-LP is transcribed from a variable number of repetitive exons. LMP2A and LMP2B are composed of multiple exons, which are located on either side of the terminal repeat (TR) region, which is formed during the circularization of the linear DNA to produce the viral episome. The blue arrows at the top represent the highly transcribed nonpolyadenylated RNAs EBER1 and EBER2; their transcription is a consistent feature of latent EBV infection. The long outer green arrow represents EBV transcription during a form of latency known as latency III (Lat III), in which all the EBNAs are transcribed from either the Cp or Wp promoter; the different EBNAs are encoded by individual mRNAs that are generated by differential splicing of the same long primary transcript. The inner, shorter red arrow represents the EBNA1 transcript, which originates from the Qp promoter during Lat I and Lat II. Transcripts from the *Bam*HIA region can be detected during latent infection, but no protein arising from this region has been definitively identified. The locations of the BARF0 and BARF1 coding regions are shown here. **b.** Location of open reading frames for the EBV latent proteins on the *Bam*HI restriction-endonuclease map of the prototype B95.8 genome. The BamHI fragments are named according to size, with A being the largest. Lowercase letters indicate the smallest fragments. Note that the LMP2 proteins are produced from mRNAs that splice across the terminal repeats (TRs) in the circularized EBV genome. This region is referred to as Nhet, to denote the heterogeneity in this region due to the variable number of TRs in different virus isolates and in different clones of EBV-infected cells. (*Adapted from Young and Rickinson, 2004*)

EBV-infected cells express a group of nuclear proteins that influence both viral and cellular transcription. EBNA1 is expressed in all virus-infected cells, in which its role in the maintenance and replication of the episomal EBV genome is achieved through

sequence-specific binding to the plasmid origin of viral replication, *OriP* (Figure 5 a,b). EBNA1 can also interact with certain viral promoters, thereby contributing to the transcriptional regulation of the EBNAs (including EBNA1 itself) and of LMP1. EBNA1 is separated into amino- and carboxy-terminal domains by a Gly-Ala repeat sequence, the main function of which seems to be to stabilize the mature protein, preventing its proteasomal breakdown. Gene-knockout studies indicate that EBNA1 does not have a crucial function in *in vitro*) B-cell transformation beyond the maintenance of the viral genome (*Humme et al., 2003*); on the other hand, a more direct involvement in oncogenesis is indicated by the ability of B-cell-directed EBNA1 expression to produce B-cell lymphomas in transgenic mice (*Wilson et al., 1996*), and by its possible contribution to the survival of Burkitt's lymphoma cells *in vitro* (*Kennedy et al., 2003*). The inability of one EBV strain P3HR-1, which carries a deletion of the gene that encodes EBNA2 and the last two exons of that for EBNA-LP, to transform B cells *in vitro* was the first indication of the crucial role of EBNA2 in the transformation process. Restoration of the EBNA2 gene in P3HR-1 has unequivocally confirmed the importance of EBNA2 in B-cell transformation and has allowed the functionally relevant domains of EBNA2 to be identified (*Cohen et al.,1989*). EBNA2 interacts with a sequence-specific DNA-binding protein, Jκ recombination-binding protein (RBP-Jκ), to transcriptionally activate cellular genes such as CD23 and the key viral genes LMP1 and LMP2A. EBNA-LP interacts with EBNA2 and is required for the efficient outgrowth of virus-transformed B cells *in vitro*. The transcriptional activation that is mediated by EBNA2 in conjunction with EBNA-LP is modulated by the EBNA3 family of proteins, which repress transactivation. An essential role for EBNA3A and EBNA3C in B-cell transformation *in vitro* has been shown using EBV recombinants. EBNA3C can cooperate with RAS in rodent-fibroblast transformation assays and disrupt cell-cycle checkpoints. These effects are partly explained by the interaction of EBNA3C with factors that modulate transcription (for example, histone deacetylase 1, nonmetastatic protein 23-homologue 1 and C-terminal binding protein) or influence cell-cycle progression (for example, cyclin A) (*Young & Rickinson, 2004*). Among the latent membrane protein (LMPs), LMP1 is the main transforming protein of EBV; it functions as a classic oncogene in rodent-fibroblast transformation assays and is essential for EBV-induced B-cell transformation *in vitro*. LMP1 has pleiotropic effects when it is expressed in cells, resulting in the induction of cell-surface adhesion molecules and upregulation of anti-apoptotic proteins (for example, BCL2 and A20). LMP1 functions as a constitutively activated member of the tumour necrosis factor receptor (TNFR)

superfamily, and activates several signalling pathways in a ligand-independent manner. Functionally, LMP1 resembles CD40, another member of the TNFR superfamily, and can partially substitute for CD40 *in vivo*, providing both growth and differentiation signals to B cells (*Uchida et al., 1999*). The LMP1 protein activates several downstream signalling pathways that contribute to the many phenotypic consequences of LMP1 expression, including the induction of various genes that encode anti-apoptotic proteins and cytokines. The LMP2 proteins, LMP2A and LMP2B, are not essential for EBV-induced B-cell transformation *in vitro*. However, expression of LMP2A in B cells in transgenic mice abrogates normal B-cell development, allowing immunoglobulin (Ig)-negative cells to colonize peripheral lymphoid organs. This indicates that LMP2A can drive the proliferation and survival of B cells in the absence of signalling through the B-cell receptor (BCR). LMP2A can transform epithelial cells and enhance their adhesion and motility, effects that might be mediated by the activation of the phosphatidylinositol-3-kinase–AKT pathway. In addition to these effects, LMP2A was found to induce expression of a range o genes that are involved in cell-cycle induction, inhibition of apoptosis and suppression of cell-mediated immunity (*Young & Rickinson, 2004*). In addition to the latent proteins, the two small non polyadenylated (non-coding) RNAs, EBER1 and EBER2, are expressed in all forms of latency. However, the EBERs are not essential for the EBV induced transformation of primary B lymphocytes. Expression of the EBERs in Burkitt's lymphoma cell lines has been found to increase tumorigenicity, promote cell survival and induce interleukin-10 (IL-10) expression (*Takada & Nanbo, 2001*). A group of abundantly expressed RNAs that are encoded by the *Bam*HIA region of the EBV genome were originally identified in nasopharyngeal carcinoma (NPC), but were subsequently found to be expressed in other EBV-associated malignancies, such as Burkitt's lymphoma, Hodgkin's lymphoma and nasal T-cell lymphoma, as well as in the peripheral blood of healthy individuals (*Deacon et al., 1993*). These highly spliced transcripts are commonly referred to as either *Bam*HIA rightward transcripts (BARTs) or complementary-strand transcripts (CSTs) (*Smith et al., 2000*). The protein products of these open reading frames remain to be conclusively identified. Another transcript that is generated from the *Bam*HIA region is BARF1, which encodes a 31-kDa protein that was originally identified as an early antigen expressed on induction of the EBV lytic cycle. Some studies have shown that BARF1 is a secreted protein that is expressed as a latent protein in EBV-associated NPC and gastric carcinoma (*Decaussin et al., 2000; zur Hausen et al., 2000*).

2.2 *Epstein Barr Virus life cycle*

Epstein-Barr virus is one of the viruses for which evidence of involvement in the development of human cancers was first presented. EBV is, however, a ubiquitous infectious agent, infecting more than 90% of the world's population. Although EBV infection is not generally related to cancer development, EBV has been linked to several specific human cancers such as Burkitt's lymphoma, nasopharyngeal carcinoma and some types of gastric cancers. Also, in immunocompromised patients, EBV causes lymphoid proliferation leading to the appearance of lymphomas. The lifecycle of EBV is quite distinct from those of other herpesviruses, such as herpes simplex virus type 1 (HSV-1) or cytomegalovirus (CMV). With the latter, full lytic replication can be accomplished by infection of certain cell types. Such an efficient lytic replication system, however, does not exist for EBV, which specifically infects resting B lymphocytes via CD21 and HLA class II molecules on the cell surface, inducing continuous proliferation of the infected cells and preventing them from undergoing apoptosis (*Tsurumi et al., 2005*). Primary infection (by oral transmission) is usually asymptomatic, but if it is delayed until adolescence it occasionally develops as infectious mononucleosis (IM). Patients with acute IM shed high titres of infectious virus in the throat from lytic infection at oropharyngeal sites. It is possible that this occurs in local mucosal B cells but, from evidence of virus replicative lesions that are seen in the oral mucosa of immunocompromised patients, it is likely that this also involves the oropharyngeal epithelium. At the same time, large numbers of latently infected B cells appear in tonsillar (and possibly other) lymphoid tissues. *In vitro*, both naive and memory B cells seem equally susceptible to EBV infection. This reservoir of infected cells is then stably maintained, and seems to be subject to the same physiological controls as the general mucosa-associated memory-B-cell pool. Such a strategy brings with it the possibility of fortuitous, antigen-driven recruitment of infected cells into germinal centres, leading to progeny that either re-enter the circulating memory pool or differentiate to become plasma cells that might migrate to mucosal sites. The different forms of latency that are seen in virus-associated malignancies might represent latency programmes that have evolved to accommodate such changes in host-cell physiology (*Young & Rickinson, 2004*). Germinal-centre transit therefore seems to activate a latency programme in which only the genome-maintenance protein EBNA1 is expressed, whereas exit from germinal centres is possibly linked to the transient activation of LMP1 and LMP2 expression. Similarly, a commitment to plasmacytoid differentiation is thought to trigger these cells to undergo lytic viral replication, providing a source of low-level virus

shedding into the oropharynx (fig. 6). There might also be circumstances in which infected cells in the reservoir can become reactivated to produce further proliferative latency III infections. Primary EBV infection elicits strong cellular immune responses that then bring the infection under control. The lymphocytosis that typifies acute IM therefore reflects the hyperexpansion of cytotoxic CD8+ T cells that are reactive to both lytic- and latent-cycle viral antigens, reactivities that are subsequently maintained in the CD8+ T-cell memory at levels that, collectively, might constitute up to 5% of the total circulating CD8+ T-cell pool (*Hislop et al., 2002*). This level of commitment to a single virus, which is apparent even in EBV carriers who have no prior history of IM, implies a crucial role for immune T-cell surveillance in controlling persistent EBV infection.

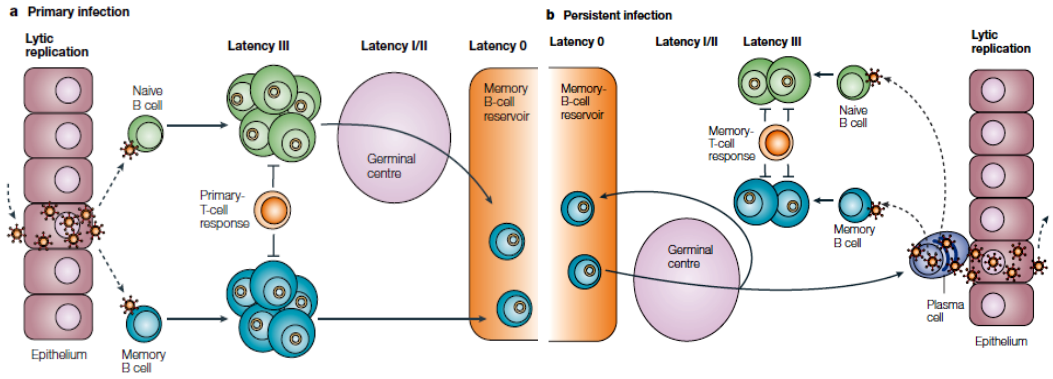

Figure 6. a. Primary infection. Incoming virus establishes a primary focus of lytic replication in the oropharynx (possibly in the mucosal epithelium), after which the virus spreads throughout the lymphoid tissues as a latent (latency III) growth-transforming infection of B cells. Many of these proliferating cells are removed by the emerging latent-antigen-specific primary-T-cell response, but some escape by downregulating antigen expression and establishing a stable reservoir of resting viral-genome positive memory B cells, in which viral antigen expression is mostly suppressed (latency 0). Different views of these events are shown. One view is that naive B cells are the main targets of new EBV infections *in vivo*. In this scenario, viral transformation drives naive cells into memory by mimicking the physiological process of antigen-driven memory-cell development in lymphoid tissues, a process involving somatic immunoglobulin-gene hypermutation during transit through a germinal centre. An alternative view therefore envisages infection of pre-existing memory cells as a direct route into memory; this is consistent with the above observations on IM tonsils, but still leaves unexplained the apparent disappearance of the infected naive cell population. **b. Persistent infection**. The reservoir of EBV infected memory B cells becomes subject to the physiological controls governing memory-B-cell migration and differentiation as a whole. Occasionally, these EBV-infected cells might be recruited into germinal-centre reactions, entailing the activation of different latency programmes, after which they might either re-enter the reservoir as memory cells or commit to plasma-cell differentiation, possibly moving to mucosal sites in the oropharynx and, in the process, activating the viral lytic cycle. Virions produced at these sites might initiate foci of lytic replication in permissive epithelial cells, allowing low-level shedding of infectious virus in the oropharynx, and might also initiate new growth transforming latency III infections of naive and/or memory B cells; these new infections might possibly replenish the B-cell reservoir, but are more likely to be efficiently removed by the now well established memory-T-cell response. (*Adapted from Young and Rickinson, 2004*).

2.2.1 *Latency*

Latency is the state of persistent viral infection without active viral production. EBV persists mostly in the memory B-cell compartment and possibly also in epithelial cells. Currently, it is thought that one in a million B cells carry the EBV genome in an individual after recovery from acute infection. There is limited expression of EBNA and latent membrane protein (LMP) gene products during latency (4). These include EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C, EBNA leader protein (EBNA-LP), LMP1, and LMP2. Characterization of gene expression patterns in different cell lines, like Burkitt's tumors and EBV-immortalized lymphoblastoid cell lines (LCLs), has determined that there are at least three different latency programs (*Thorley-Lawson & Gross, 2004*). By using different transcription programs, latent EBV genomes can multiply in dividing memory cells (type I), induce B-cell differentiation (type II), activate naïve B cells (type III), or completely restrict all gene expression in a context-specific manner. Only EBNA1 is expressed in the type I latency program, which is seen in Burkitt's lymphoma. CD8+ T cells specific for many EBV antigens arise during the immune response to natural infection, but not for EBNA1, which contributes to evasion during latency (*Blake at al., 1997*). EBNA1 and LMP1/2A are expressed in the type II latency program, which is observed in nasopharyngeal carcinoma and Hodgkin's lymphoma. LMP1 and LMP2 are responsible for B-cell activation and induction of a growth (proliferation) program (*Bornkamm and Hammerschmidt, 2001*) The type III latency program, in which all of the latency gene products are expressed, is often detected during acute infectious mononucleosis or in certain immunocompromised individuals.

EBV genomes in latently infected B cells exist as episomes (*Ambinder & Lin, 2005*), that are packaged in nucleosomal arrays with cellular histones, replicated once and only once during S phase, and partitioned faithfully into daughter cells during the mitotic phase. Only one viral *cis* element, *OriP*, and one viral protein, EBNA1, appear to be required for viral DNA maintenance (*Ariza et al., 2009*). *OriP* consists of two clusters of EBNA1 binding sequences; one is a family of repeats (FR) containing 20 sites, and the other a dyad symmetry (DS) containing four low-affinity EBNA1 binding sites (fig. 7) (*Lindner & Sugden, 2007*). It is generally agreed that EBNA1 permits the retention of replicated DNA in daughter cells by tethering the FR region to chromosomes. *OriP* functions as a replicator, for which the DS element is indispensable.

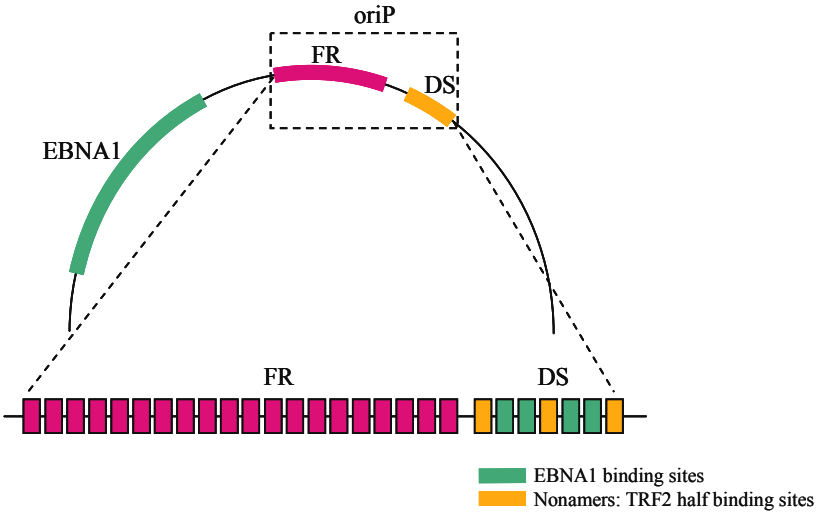

Figure 7. *OriP* structure. The origin of plasmid replication (*oriP*) is composed of the family of repeats (FR) and the dyad symmetry (DS) element, which both contain binding sites for EBNA1. FR is composed of 21 imperfect copies of a 30 bp repeat and contains 20 high affinity EBNA1-binding sites. When FR is bound by EBNA1, it both serves as a transcriptional enhancer of promoters in *cis* up to 10 kb away and contributes to the nuclear retention and faithful maintenance of FR containing plasmids. The efficient partitioning of *oriP* plasmids is also likely attributable to FR. The dyad symmetry element (DS) is sufficient for initiation of DNA synthesis in the presence of EBNA1. Initiation of DNA synthesis from DS is licensed to once-per-cell-cycle, and is regulated by the components of the cellular replication system. DS contains four EBNA1-binding sites, albeit with lower affinity than those found in FR. Flanking the pairs of EBNA1-binding sites of DS are three 9 bp elements that resemble telomeric repeats, and have been termed nonamers. These elements are bound by cellular proteins associated with the telomeres, including TTAGGG-repeat binding factors (TRF1, TRF2), the human repressor activator protein1 (hRap1), and tankyrase. Deletion or mutation of the nonamer elements decreases by two-fold the average number of plasmids maintained per cell. One protein found to bind these elements, TRF2, that may contribute to the enhanced replication of DNA by increasing the apparent affinity of EBNA1 for its binding sites. Despite this effect, the nonamer elements are not necessary for replicative function. The two pairs of EBNA1-binding sites of DS have an intra-pair spacing of 21 bp from the center of one site to the center of the paired site, and an inter-pair spacing of 33 bp between the centers of the adjacent binding sites. Nine base pair binding sites for the TTAGGG-repeat binding factor 2 (TRF2), also referred to as 'nonamers', flank the pairs of EBNA1-binding sites.

EBNA1 is the sole *trans*-acting element of EBV required for its plasmid replicon. EBNA1 lacks any enzymatic activities and must rely entirely on cellular factors to carry out these functions. Many proteins associated with EBNA1's replicative functions have been identified recently, while the cellular factors associated with its ability to modulate transcription are unknown. EBNA1-dependent DNA synthesis from DS is licensed to occur once-per-cell-cycle, and several groups have shown by chromatin immunoprecipitation (ChIP) that this process is mediated and regulated by the cellular replicative machinery (*Dhar et al., 2001; Ritzi et al., 2003; Wang et al., 2006*). When EBNA1 binds to a pair of binding sites of DS, the protein–DNA structure formed by this interaction recruits members of the pre Replicative Complex (pre-RC). The origin recognition complex (ORC1-6) and the putative replicative helicase (MCM2-7) have both been found to associate with these EBV origins in a similar cell-cycle dependent manner as is thought

to occur generally in human cellular origins (*Dhar et al., 2001; Ritzi et al., 2003; Wang et al., 2006*). Scott E. Lindner and Bill Sugden (*2007*) proposed a sequential model of the licensing process for DNA synthesis at an EBNA1-dependent origin: a pair of EBNA1-binding sites are bound and bent in the same direction by two dimers of EBNA1. EBNA1 binds to a pair of sites in DS interacting cooperatively with the TTAGGG-repeat binding factor 2 (TRF2), which binds to the adjacent 'nonamer'. EBNA1 and TRF2 associate with DS throughout the cell cycle, as do subunits two through six of the origin recognition complex (ORC2-6) (fig. 8a). ORC1 associates with DS in G1/S, as it does with cellular origins. The recruitment of ORC1 to DS does not require TRF2-binding sites. However, the association of the N-terminal basic domain of TRF2 with the N-terminal BAH domain of ORC1 enhances the association of ORC1 with DS. The regulatory protein Cdc6 is also likely recruited to DS in G1/S (fig. 8b). ORC1-6 and Cdc6 function as a clamp loader for the reiterative recruitment and loading of the putative replicative helicase MCM2-7 to yield a head-to-head double hexamer at DS. The association of MCM2-7 with DS is dependent upon the regulatory protein Cdt1, which possibly associates with DS in complex with MCM2-7 (fig. 8c). Taken together, this protein/DNA complex progresses from a pre-Replicative Complex (pre-RC) to a pre-initiation complex (pre-IC), allowing the efficient initiation of DNA synthesis to occur at or near DS in at least 84% of S-phases (*Lindner & Sugden, 2007*).

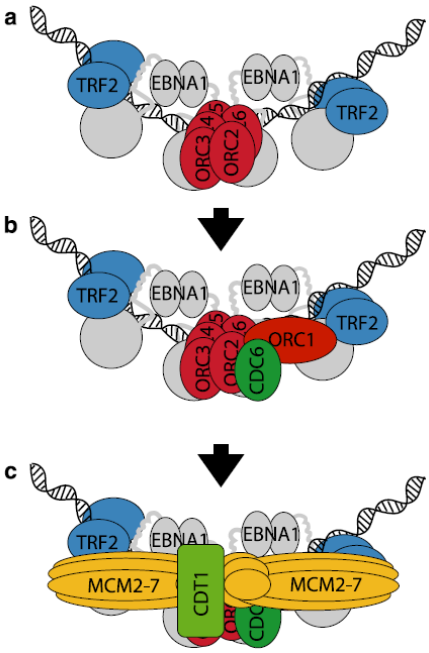

Figure 8. a. A pair of EBNA1 dimers binds to a pair of 16 bp binding sites with a precise 21 bp center-to-center spacing, placing the dimers on the same helical face of the DNA. The binding event bends the DNA toward the DNA binding and dimerization domain in an additive manner and in the same direction. The TTAGGG-repeat binding factor 2 (TRF2) is found as a dimer in the cell, and one monomer can bind to a 9 bp element that flanks the pairs of EBNA1-binding sites in DS. The binding of TRF2 to these elements (also termed nonamers) permits the cooperative binding of EBNA1 to its binding sites, thereby increasing EBNA1's apparent affinity. In addition, ORC2-6 were found to be bound to such an EBNA1-dependent origin throughout the cell cycle by chromatin immuno-precipitation (ChIP). **b.** ORC1 has been shown to interact with EBNA1-dependent origins in the presence and absence of the TRF2 binding sites, although ORC1's association with these origins is likely enhanced by the interaction of it's BAH domain with the N-terminal basic domain of TRF2. Cdc6 is also likely recruited to the origin in G1/S. **c.** Taken together, these assembled proteins at the EBNA1-dependent origin allow the subsequent recruitment of, and likely there iterative loading of, the putative replicative helicase MCM2-7 asa head-to-head double hexamer, possibly in a complex with theregulatory protein Cdt1. (*Adapted from Lindner SE & Sugden B, 2007*).

2.2.2 *Reactivation*

Latently infected B cells can occasionally be stimulated to reactivate EBV. This produces virus that can re-infect new B cells and epithelial cells, becoming a source of viral transmission (fig. 9). Although much is known about the molecular pathways involved in viral reactivation, what triggers reactivation *in vivo* is not known precisely. The presumption is that it occurs when latently infected B cells respond to unrelated infections, because B-cell receptor stimulation triggers reactivation in B-cell lines. It is also not known what fraction of EBV-infected cells is in the lytic or latent phase at any time. Lytic replication differs from the latent amplification state in that multiple rounds of replication are initiated within *oriLyt*, and the replication process has a greater dependence on EBV-encoded proteins (*Fixman et al., 1995*). Soon after induction of the lytic cycle, viral DNA is amplified to yield monomeric plasmid progeny DNA dependent on a functional *oriLyt* in *cis*. In the late phase of the viral productive cycle, the EBV genome is amplified 100- to 1000-fold. Intermediates of viral DNA replication are found as large head-to-tail concatemeric molecules, which are subsequently cleaved into unit length genomes and packaged into virions. Seven essential core EBV replication genes have been identified so far (BZLF1, BALF5, BMRF1, BALF2, BBLF4, BSLF1, BBLF2/3) to be necessary and sufficient for *oriLyt*-specific DNA replication. The BZLF1 protein acts as an *oriLyt*-binding protein in addition to its function as an immediate-early transactivator. The BALF5 gene encodes the DNA Pol catalytic subunit and the BMRF1 gene encodes the DNA Pol accessory subunit, that together form a heterodimer to function as the Pol holoenzyme, with high polymerase processivity, presumably synthesising both leading and lagging strands. A single-stranded DNA-binding protein is encoded by the BALF2 gene. The remaining three proteins encoded by the BBLF4, BSLF1 and BBLF2/3 genes form a tight complex and are predicted to act as helicase, primase and helicase-primase associated proteins, respectively, from sequence homology to the herpes simplex virus type 1 (HSV-1) UL5, UL52 and UL8 genes (*Tsurumi et al., 2005*). The multiple steps essential for DNA replication are catalysed by a number of proteins whose enzymatic reactions must be coordinated closely. It is likely that initiation of the lytic phase of EBV DNA replication involves the formation of an initiation complex at *oriLyt*, which consists of two essential upstream and downstream domains. Whereas the former contains several BZLF1 binding sites, the latter includes binding sites for several cellular proteins. The first step in this process would be the binding of the BZLF1 protein and two transcription factors, ZBP-89 and Sp1, to recognition sequences within *oriLyt* to form an initial complex. ZBP-89 and

Sp1 are known to stimulate replication (*Baumann et al., 1999*). BZLF1 interacts also with the BBLF4/BSLF1/BBLF2/3 complex, in order to recruit the viral helicase–primase complex to *oriLyt*, and with both BALF5 and BMRF1 proteins. The BALF2 ssDNA-binding protein then appears to interact with the BZLF1-BBLF4/BSLF1/BBLF2/3 pre-priming complex. These proteins together, therefore, would have the potential to open up the duplex DNA in the origin region and synthesise RNA primers. The interaction between the EBV Pol holoenzyme and the BBLF4/BSLF1/BBLF2/3 complex (*Fujii et al., 2000*) may play an important role in bringing the viral polymerase into the pre-priming complex to initiate DNA synthesis. It is possible that binding of the Pol holoenzyme to the BBLF4/BSLF1/BBLF2/3 complex reduces its affinity for BZLF1, allowing the polymerase–helicase–primase complex to migrate away from *oriLyt* to the replication forks. The BALF2 protein, EBV SSB, binds ssDNA templates and may function to melt out secondary structures in the viral displaced ssDNA of the replication fork, thereby facilitating movement of the EBV DNA Pol holoenzyme on the ssDNA template. As expected with the HSV-1 helicase/primase complex, the proposed EBV helicase and primase complex consisting of the BBLF4/BSLF1/BBLF2/3 proteins may bind to the lagging strand at the fork, translocate in the 5' to 3' direction and synthesise the RNA primer. Thus, the six viral replication proteins appear to all work at the replication fork as the replication machinery. Understanding how each gene product, whether lytic or latent, contributes to the pathogenesis of EBV-related diseases should lead to more rational and effective prevention and treatment strategies.

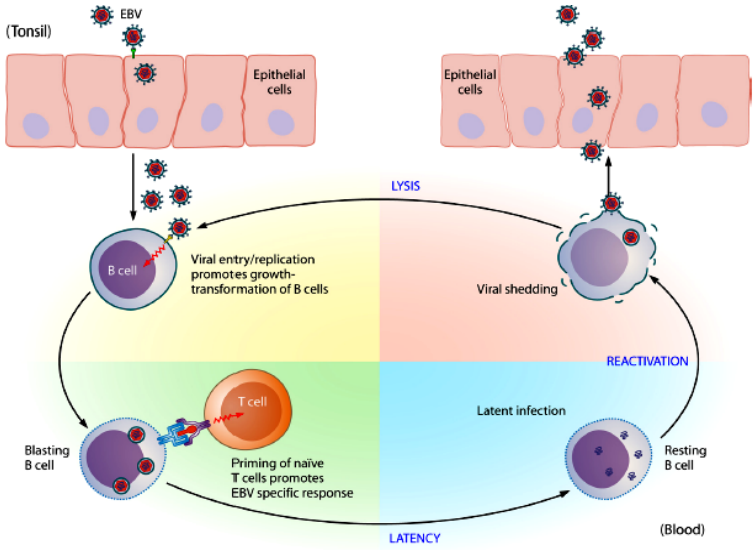

Figure 9. Viral entry results in transport of the EBV genome into the B-cell nucleus, where replication by cellular and viral DNA polymerases begins. EBV gene products activate the B-cell growth program, resulting in the proliferation of blasting B cells. Priming of naïve T cells by antigen-presenting cells occurs in parallel. Normally, these blasting B cells are destroyed by cytotoxic T lymphocytes. Once in the circulation, previously activated memory B cells may continue to undergo lytic replication or, if EBV shuts down most of its protein-encoding genes, latency occurs. At a later time, as cells recirculate between the oral and peripheral compartments, resting B cells may be activated, resulting in viral reactivation and shedding (*Adapted from Oludare et al., 2011*)

2.3 *EBV associated diseases*

The oncogenic properties of EBV have been appreciated for a long time. EBV is shown to be associated with a wide spectrum of human cancers, including both hematopoietic and epithelial tumors, most preferentially in posttransplant and AIDS patients (*Kuppers, 2005*). The primary neoplasms associated with EBV are B-cell lymphomas and nasopharyngeal carcinoma, reflecting the primary cellular targets of viral infection *in vivo*: B cells and tonsillar epithelium, respectively. The virus utilizes multiple mechanisms to promote neoplasm, including activation of the B-cell growth program, immune evasion, and inactivation of tumor suppressors (*Oludare et al., 2011*). Many lines of evidence suggest that ongoing immune control of EBV reactivation is critical to prevent transformation *in vivo* (*Münz & Moormann, 2008*). Combinations of environmental and genetic factors contribute to EBV-associated diseases, which include:

Infectious mononucleosis

A self-limiting lymphoproliferative disease that might arise if infection occurs in adolescence or later. At the height of the disease up to 50% of the T cells can be specific for the virus and up to 25% of the memory B cells are latently infected.

Burkitt's lymphoma (BL)

The tumour in which the virus was first discovered. The defining feature of BL is translocation of the *c-myc* proto-oncogene into one of the immunoglobulin loci leading to deregulated *myc* expression. EBV is present in most, but not all, BL, where it only expresses EBV nuclear antigen 1, but the role of the virus in pathogenesis is unclear.

Nasopharyngeal carcinoma

Extremely common tumour in certain high-risk groups such as the Southern Chinese and skimos. The undifferentiated form of nasopharyngeal carcinoma is always EBV-positive. Other cofactors must be involved, and both environmental (for example, salted, pickled fish) and genetic susceptibilities have been implicated (*Thorley-Lawson, 2001*).

Hodgkin's Disease (HD)

Infectious mononucleosis (IM) is a predisposing factor for the development of HD and the only time when the peculiar cells of HD (the Reed–Sternberg cells) are consistently observed outside of HD is during IM (*Kurth et al., 2000*). However, EBV is only present in 40–60% of cases 77, and IM is predisposing for both EBV-positive and EBV-negative HD.

This indicates that the disruption of IM, rather than EBV per se, might be the predisposing factor.

Immunoblastic lymphomas

A heterogeneous group of B-cell tumours, expressing the growth programme, that arise in individuals who are immunocompromised and unable to mount an affective cytotoxic T-lymphocyte response to these cells. Immunosuppression can arise iatrogenically (organ transplants), environmentally (human immunodeficiency virus infection) or genetically (X-linked lymphoproliferative disease) (*Thorley-Lawson, 2001*).

Autoimmune diseases

Epstein–Barr virus (EBV) has been suspected of involvement in the aetiology of various human chronic autoimmune diseases since the finding of elevated levels of antibodies to the virus in systemic lupus erythematosus (SLE). There is now a large body of evidence implicating the virus in several different autoimmune diseases, including SLE, multiple sclerosis (MS) (*Serafini et al, 2007*), Sjögren's syndrome, rheumatoid arthritis, autoimmune thyroiditis, autoimmune hepatitis, cryptogenic fibrosing alveolitis and pure red cell aplasia (*Pender et al., 2003).*

Chapter 3

DEVELOPMENT OF CONDITIONALLY REPLICATING INTEGRASE DEFECTIVE LENTIVIRALVECTORS FOR EPSTEIN BARR VIRUS GENE THERAPY

3.1 *Background and Aim of the work*

Lentiviral vectors (LV) are useful tool for gene transfer. From their first application to now, they have been strong developed in design, biosafety and in their ability of transgene expression into target cells. One of the major concern using viral vectors for gene therapy approach, is the risk of insertional mutagenesis due to viral DNA integration into the host genome. To circumvent insertional mutagenesis risks, integrase defective LV (IDLV) have been developed. IDLV episomes are a weaker template for effective gene expression and furthermore they are unstable in dividing cell types and so the duration of episomal vector gene expression is strictly dependent on how rapidly transduced cells divide (*Negri et al., 2011*). Giving this premise, the construction of an IDLV able to replicate autonomously into a target cell and to persist in the absence of integration, is desiderable. The instability of the IDLV in proliferating cells is mainly due to the lack of an origin of replication allowing vector's DNA replication. In the work described in this thesis, we evaluated if the inclusion of an origin of replication derived from Epstein-Barr virus into the backbone of an IDLV was able to allow selective vector replication and tethering into target cells expressing the origin trans-activating factor (EBNA1). Such a kind of vector, modified to express a toxic gene, such as thymidine kinase form herpes simplex virus 1 (HSV-1) (*Greco & Dach, 2001*), could be a useful tool of gene therapy against all EBV-mediated diseases.

3.2 **Materials and Methods**

3.2.1 Construction of modified lentiviral vectors

For the purposes of this work, we designed five different vectors, four containing the EBV origin of replication (*oriP*) and one control vector without *oriP*. In order to evaluate possible positional effects, the *oriP* has been inserted into two different conformations, either before the transcription unit or at the end of the transcription unit after the neomicine resistence gene. Furthermore, we used two different *oriP*, a wild type *oriP* (*oriP* wt) and a mutated one (*oriP* mut), kindly provided by Bill Sugden (*University of Winsconsis*) since the latter has been described by the author to initiate DNA synthesis 2.2-fold more efficiently than the wt *oriP* (*Lindner et al., 2008*). The basic features of the vectors used in this work are shown in Figure 10. Briefly, to produce pTY2-CMV-GFP-IRES2-Neo-OriPmut or pTY2-CMV-GFP-IRES2-Neo-OriPwt, the plasmid p3513 (*oriP* mut) or p3488 (*oriP* wt), kindly provided by Dr. B. Sugden, were digested with XbaI/KpnI to extract the modified or the wild type EBV *oriP*, respectively; The latter were subsequently cloned into pTY2-CMV-GFP-IRES2-Neo plasmid. To produce pTY2-OriPmut-CMV-GFP-IRES2-Neo or pTY2-OriPwt-CMV-GFP-IRES2-Neo, the modified or the wild type EBV *oriP* were extracted from p3513 or p3488 plasmid by XbaI/BamHI digestion and subsequently cloned into pTY2-CMV-GFP-IRES2-Neo plasmid.

3.2.2 Restriction and purification of the DNA fragments from agarose gel

Restriction enzyme (or restriction endonuclease) cuts double-stranded DNA at specific recognition nucleotide sequences known as restriction sites (usually palindromic sequences of 4-8 nucleotides).

2 µg of DNA are digested with 3-4 units of enzyme in a specific buffer, to a final volume of 30 µl, for 3 hour at optimal temperature (37°C for most restriction enzymes). The obtained DNA is purified through electrophoresis on agarose gel. The band corresponding to the fragment of interest is then excised from the gel and purified according to Qiagen gel extraction kit protocol.

3.2.3 Ligation of DNA fragment

The purified band is then cloned into the chosen vector through a reaction called ligation, catalyzed by T4 DNA ligase (derived from T4 bacteriophage) that catalyzes the joining of two strands of DNA between the 5´- phosphate and the 3´- hydroxyl groups of adjacent nucleotides. Vector DNA and insert DNA are ligated into a 1:3 molar ratio of vector/insert.

The amount of DNA needed from each fragment that has to be ligated, is then calculated according to this formula:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

Reaction mix contains: 10X ligation Buffer supplied with the enzyme (300mM Tris-HCl (pH 7.8), 100mM MgCl2, 100mM DTT and 10mM ATP), 1 Weiss unit of T4 DNA ligase and water to a final volume of a 10-15 µl. The reaction is then incubated overnight at 14°C.

3.2.4 Competent cells transformation

DH10B competent cells (Invitrogen) are transformed with the ligation product using the heat shock method. Briefly, 20 µl of DH10B are incubated 30 minutes on ice with 2 µl of ligated DNA. Competent cells are then heat-shocked at 42°C for 45 seconds and placed on ice for 2 minutes. After that 250µl of room temperature S.O.C. Medium (Invitrogen) is added and the cells are then shaken at 225 rpm (30°C) for 1 hour. Finally the cells are spread on LB (Luria Broth) agar plates with 100 µg/ml ampicillin. Presence of the ampicillin resistance gene in the plasmids allows for selective growth of transformed bacteria cells containing the plasmid (recombinants).

3.2.5 Colonies amplification and DNA extraction

Single colonies are picked from the selective plate and inoculated into 5 ml of LB medium containing 100 µg/ml ampicillin and incubated for 16 hours with vigorous shaking. Bacteria pellets are then harvested by centrifugation at 8000 rpm for 3 minutes and the DNA is extracted according to Qiagen miniprep kit protocol. An enzymatic digestion is then performed to verify the presence of the insert's DNA.

3.2.6 Maxiprep of plasmid DNA

To obtain higher amount of the newly constructed plasmid, 50 µl of the recombinant bacteria are re-incubated into 150 ml of LB medium overnight at 30°C with vigorous shaking. Bacteria pellet are then harvested by centrifugation at 4500 rpm for 30 minutes and the DNA is subsequently extracted following the manufacturer's instructions (Hi-speed Maxiprep kit, Qiagene). The concentration of DNA extracted from bacteria cells is then evaluated by spectrophotometer. The sample is diluted 1:100 in water and the absorbance at 260 nm (A260) is evaluated; an A260=1 correspond to 50µg/ml of DNA.

TRANSFER VECTORS

a. **pTY2 oriP wt CMV GFP IRES2 Neo** or **pTY2 CMV GFP IRES2 Neo oriP wt**

b. **pTY2 oriP mut CMV GFP IRES2 Neo** or **pTY2 CMV GFP IRES2 Neo oriP mut**

c. **pTY2 CMV GFP IRES2 Neo**

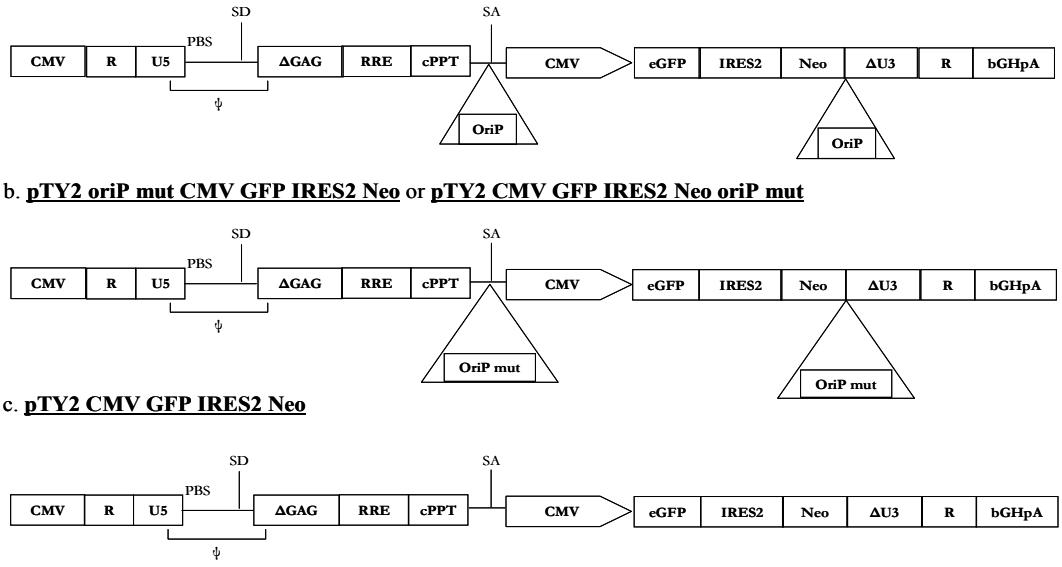

Figure 10. Schematic representation of the vectors used in this study. The splice donor site (SD), the splice acceptor (SA), the packaging signal (Ψ) and PBS are indicated. a) Vectors containing an EBV *oriP* wild type, inserted either before the CMV promoter or after the neomicine resistance gene. b) Vectors containing the modified *oriP* (*oriP* mut), inserted either before the CMV promoter or after the neomicine resistance gene. c) Control vector without *oriP*.

3.2.7 Lentiviral vectors production

Lentiviral vectors are generated in human kidney 293T cells by using a calcium phosphate transfection method. Briefly, 293T are maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% pennycillin-streptomycin-glutamine 100X (GIBCO). The day before the transfection, a total of $3.5x10^6$ 293T cells are plated per 10-cm plate in 10 ml of medium. At hour 0, the cells are transiently transfected with 1 ml of a calcium phosphate precipitation mixture (Promega) containing 6μg of the transfer plasmid, 4μg of the integrase defective packaging constructs pc-Help (IN-), expressing *gag*, *pol*, *tat* and *rev* genes from HIV-1, and 2μg of pCMV-VSV-G plasmid, expressing the VSV glycoprotein needed to pseudotype lentiviral particles. At hour 17, the medium is replaced with 8 ml of fresh medium. At hour 48, the medium was collected and spun at 1500 rpm for 15 min at 4°C. The medium is then filtered through a 0.45-mm pore size to remove cellular debris, and then stored at -80°C until use.

3.2.8 Concentration of lentiviral vector stocks

For cell transduction *in vitro*, crude (unconcentrated) vector stocks are often sufficient, whereas concentrated vector stocks are needed for *in vivo* applications and with some cell type that are less prone to infection. Concentration of lentiviral vector stocks are performed using ultracentrifugation approaches. This is a rapid and robust method for vector concentration at a small scale. Briefly, filtered vector-containing supernatant from the transfected cells are placed onto a 20% sucrose cushion into six sterile Ultra-clear SW28 centrifuge tubes (Beckman). All the six tubes are then placed into a Beckman SW28 ultracentrifuge rotor and the viral preparation is then centrifuged for 2:30 h at 23.000 r.p.m. (82,700g) and 4° C using an ultracentrifuge. The inclusion of the sucrose cushion during ultracentrifugation is beneficial in terms of reducing vector toxicity.

3.2.9 Lentiviral vector titration assays

For transduction experiments, supernatants can be normalized for cpm (counts per minute) content by RT (Reverse Transcriptase) activity assay, that measure the amount of incorporated radioactive-labeled nucleotides, providing the quantification of the total number of vector particles produced, or by citofluorimetry (FACS) titration, that instead provides the functional titers of the vectors preparation. For RT assay, 1 ml of viral preparation is ultracentrifuged at 100.000 rpm for 10 minutes at 4°C. The obtained pellet is then resuspended in 100 µl of TNE-TritonX-100 0,1% for 30 minutes on ice. 30 µl of lysate are then mixed to a reaction mixture (TrisCl 60mM, DTT 5,7mM, MgCl2 15mM, KCl 65mM, PolyA/Oligo dT 0,07mg/ml e 2 µCi of methyl 1',2'-3H-thymidine 5' triphosfate) and incubated for 1 hour at 37°C. The reverse transcription is then stopped by adding 500 µl of Na-Pyrophosphate 0,01M and 600 µl of 20% trichloroacetic acid (TCA). The mixture is then filtered on nitrocellulose membrane filters (Millipore) that are washed three times with 5% TCA. Finally, the filters are dissolved in 5 ml of a scintillation liquid (pseudocumene-based LSC cocktail) and read through a scintillation counter. The number of incorporated radioactive-labeled nucleotides is calculated according to this formula:

$$\text{cpm /ml} = (\text{cpm sample} - \text{cpm blank}) \times 10 / 3$$

For FACS titration, $5x10^4$ 293T cells are plated onto 12 wells plates the day before the infection with the newly produced lentiviral particles. Three dilutions (1:2, 1:10 and 1:50) of the vector preparation are then used to infect the cells. At day 3 post infection the

infected cells are collected, washed twice in PBS (phosphate saline buffer) and analysed for GFP (green fluorescence protein) expression by FACS. The titre (i.e. number of transducing unit [TU]) of the viral preparation is then calculated according to this formula:

TU = [dilution factor] x [number of plated cells] x [percentage of transduced cells]

3.2.10 Generation of EBNA1-expressing cell lines

In order to verify the ability of *oriP*-carrying plasmids to replicate in presence of the trans-acting factor, we have generated 293T and Hela cell lines expressing EBNA1 from EBV. EBNA1 is one of the nine latent viral proteins of EBV that activates the replication of latent EBV episomes and the transcription of EBV latency genes by binding to recognition sites in the *oriP*. Briefly, 4 µg of the plasmid pTY2-CMV-EBNA1-IRES-Puro were transfected into 293T and Hela cells and 2 days post-transfection these cells were put under puromycin selection (0.3 µg/mL) for 14 days. Puromycin selected cells were then collected and analyzed for EBNA1 expression by Western blot (Figure 11) using a monoclonal antibody to EBNA1 (Acris, Cat. BM 3167).

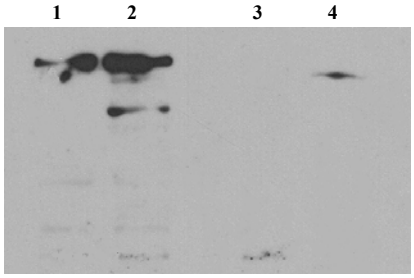

Figure 11. Western blot analysis of EBNA1 protein from cell lysates. Lanes: 1) Hela-EBNA1; 2) 293T-EBNA1; 3) 293T (negative control); 4) RAEL (EBV positive cell line, WB positive control)

3.2.11 Infection and measurement of GFP-positive cells

For transduction experiments 1×10^5 cells are plated on 6 well plates on day 0. The day after the cells are incubated with equal amounts of cpm or TU for 3 hours at 37 °C. The number of virions that are added per cell during infection is referred to as multiplicity of infection (MOI). At designated time points, 25% of the cells were harvested, washed twice in PBS, and assessed for GFP expression by fluorescence-activated cell sorting (FACS), using FACS Calibur flow cytometer analysis (Beckton Dickinson, Mountain View, CA). Data acquisition and analysis were done using CellQuest software (Becton Dickinson, Mountain View, CA).

Chapter 4

RESULTS

4.1 Single plasmid transfection and antibiotic selection

In order to assess the ability of the vectors containing the *oriP* from EBV, to selectively replicate in presence of EBNA1, we transfected 0,25 µg of these plasmids and of the control one (plasmid without *oriP*) into Hela and Hela-EBNA1 expressing cells; two days after transfection we put these cells under G418 (neomycin analogous) selection (800 µg/ml) for 14 days. Only EBNA1 positive cells transfected with the *oriP* containing plasmid were able to form a valuable number of colonies, while in absence of EBNA1 and in absence of *oriP* the number of colonies was too low. The positive colonies were analyzed for GFP expression by FACS analysis and the results are shown in Figure 12. The two vectors containing the *oriP* before the transgene cassette showed the higher percentage of GFP positive cells, indicating a possible positional effect for the *oriP*.

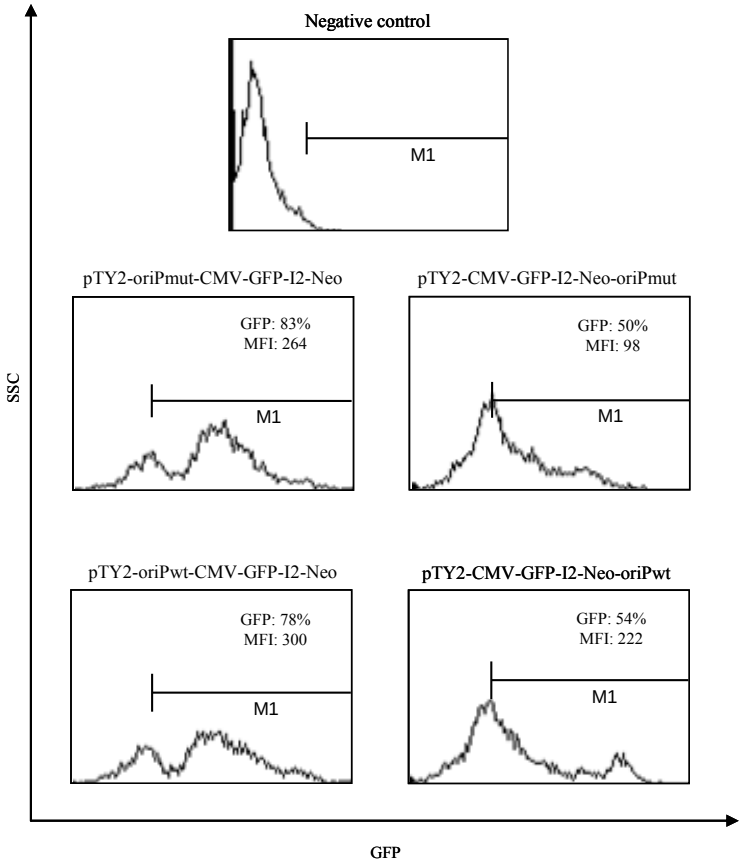

Figure 12. FACS analysis of Hela-EBNA1 cells transfected with the indicated plasmids. After 14 days of G418 selection, only Hela-EBNA1 cells transfected with *oriP* plasmids were able to form a valuable number of colonies, while most of the Hela cells died. Two of our vectors (the ones containing the *oriP* before the CMV promoter), presented the higher GFP expression.

We further repeated this experiment with different concentration of plasmid DNA (0,25 to 1 µg) to count the number of colonies formed by each of our vector. Plasmids bearing the *oriP* from EBV clearly showed a replicative advantage in EBNA1 expressing cells, in contrast with *oriP* lacking plasmid (Figure 13a). Furthermore, we have seen that vectors in which the *oriP* (either wt or mutated) was inserted before the transcriptional unit, seems to be able to form much more colonies compared to the vectors that have the *oriP* after the neomycin resistance gene confirming the positional effect of the origin of replication observed in the previous experiments. Furthermore, no difference in terms of number of resistant clones between the mutated or the wild type *oriP* has been observed. In EBNA1 negative cells the number of resistant clones was at least four times lower than in EBNA1 expressing cells. As expected, no difference between the *oriP* containing vectors and the control one has been observed (Figure 13b), clearly indicating the need of the *oriP*-EBNA1 interaction for vector replication and tethering.

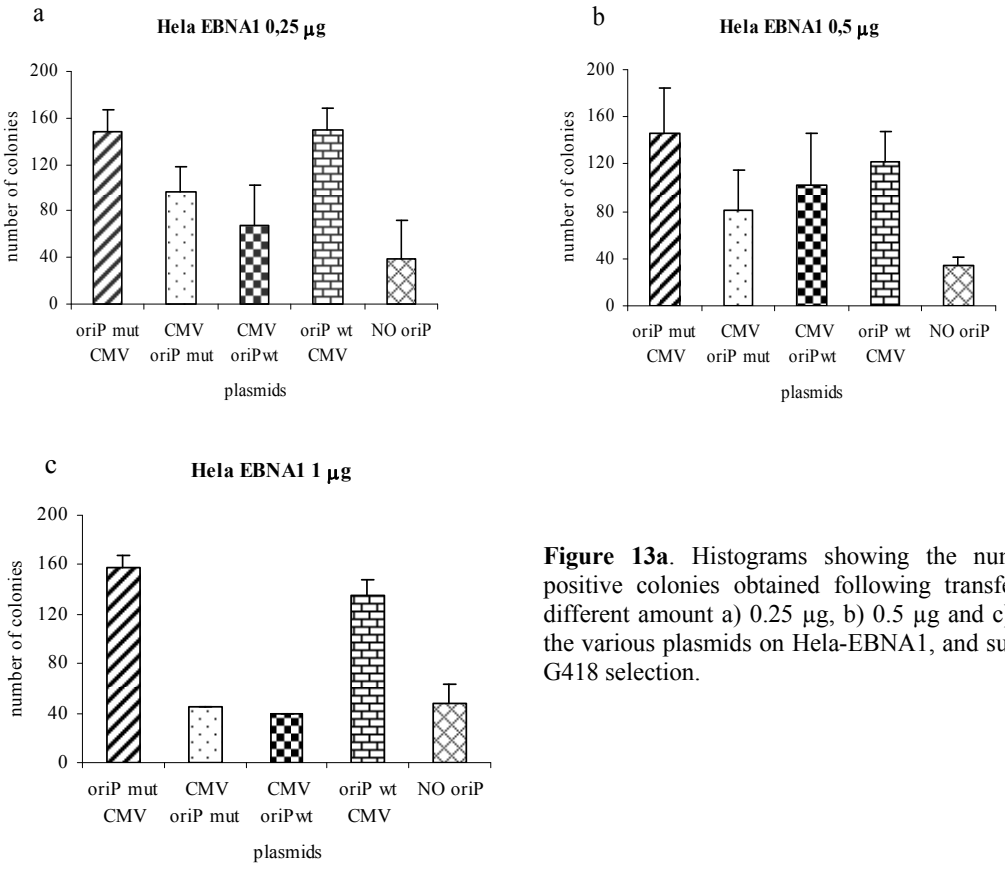

Figure 13a. Histograms showing the numbers of positive colonies obtained following transfection of different amount a) 0.25 µg, b) 0.5 µg and c) 1 µg of the various plasmids on Hela-EBNA1, and subsequent G418 selection.

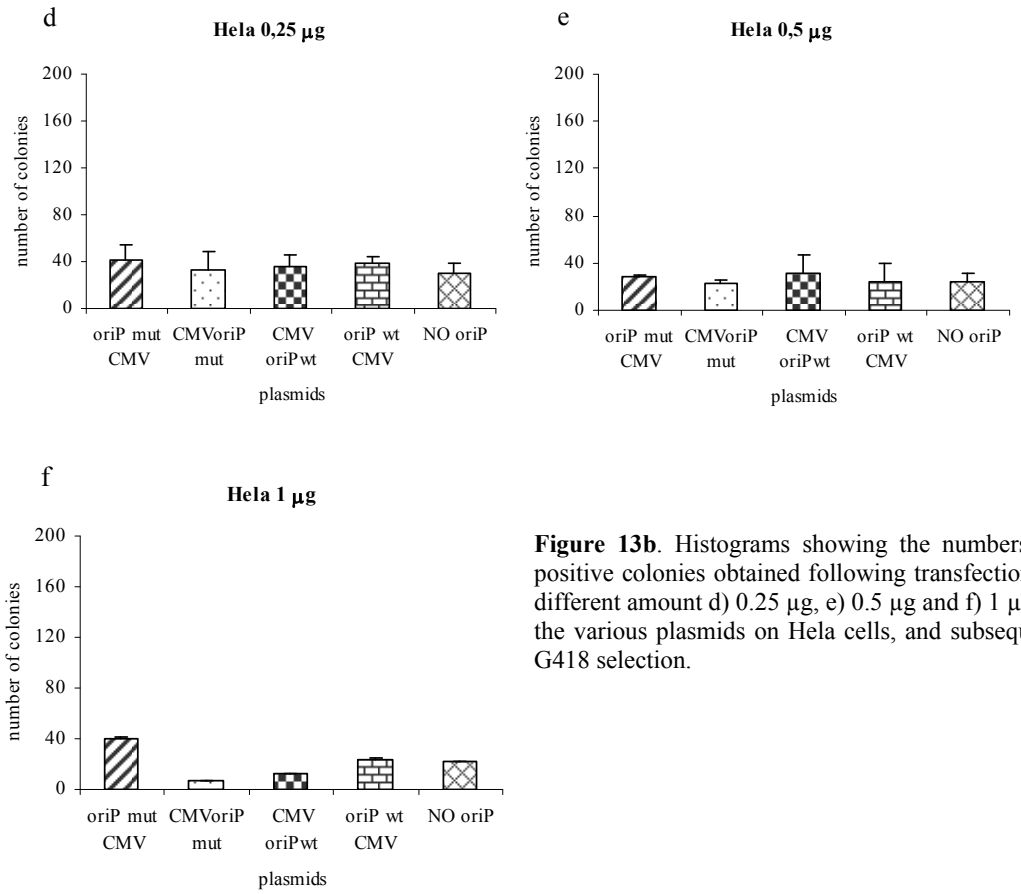

Figure 13b. Histograms showing the numbers of positive colonies obtained following transfection of different amount d) 0.25 µg, e) 0.5 µg and f) 1 µg of the various plasmids on Hela cells, and subsequent G418 selection.

4.2 Plasmid rescue from Hela-EBNA1 resistant clones

In order to assess the episomal status of the *oriP* containing plasmids in Hela-EBNA1 cells, the DNA of G418-resistant clones have been analyzed. The plasmids DNA are extracted from mammalian cells according to Qiagen miniprep kit protocol. DNA from miniprep extraction has been used to transform *E. coli* competent cells (DH5α, Invitrogen) and the obtained clones have been analysed by enzymatic digestion. The vectors restriction pattern of all clones was identical to that of the original vector. Figure 14 shows the restriction pattern of two transformed *E. coli* clones (one for each *oriP* configuration).

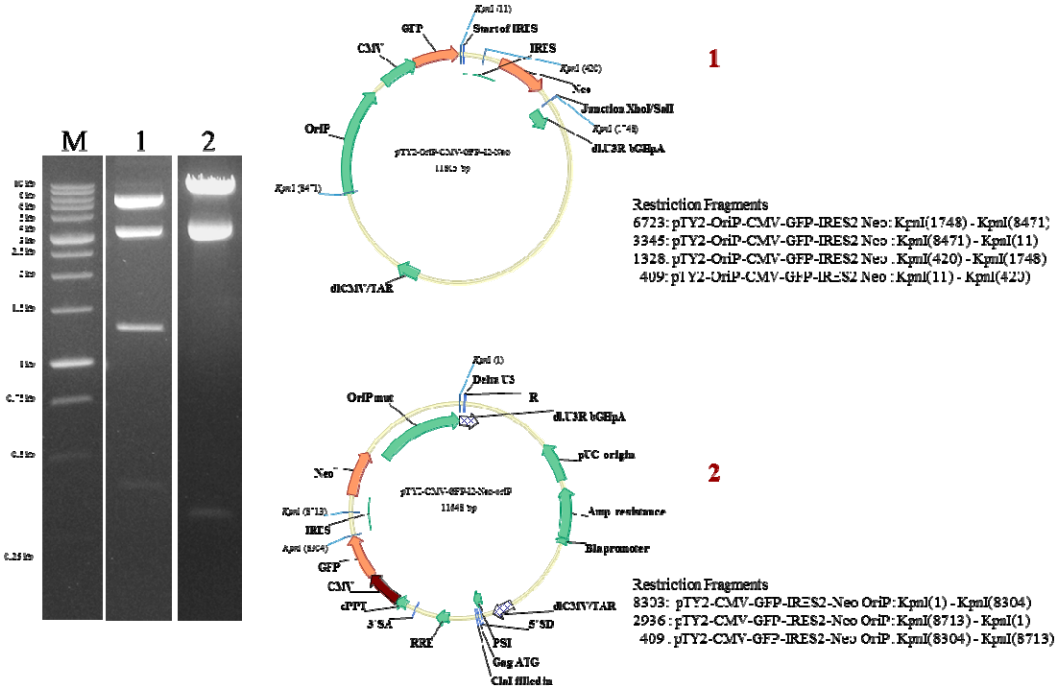

Figure 14. Plasmid rescue experiment in *E.coli* with miniprep extracts from Hela-EBNA1 cells transfected with pTY2-oriP-CMV-GFP-IRES2-Neo (1) and pTY-CMV-GFP-IRES2-Neo-oriP (2).

4.3 Lentiviral vectors transduction

After demonstrating that the *oriP* containing plasmids are able to selectively replicate in cells expressing EBNA1, our goal was to determine whether transgene expression from the *oriP* containing IDLVs would be mantained in the same cells. The transduction experiments have been conducted on different cell lines expressing the EBNA1 protein from EBV and their relative EBNA1 negative control that are: 293T-EBNA1/293T; Hela-EBNA1/Hela and RAEL/DG75. RAEL and DG75 cell lines are both Burkitt's lymphoma B cell lines, but while RAEL are EBV positive, DG75 are EBV negative and, consequently, do not express EBNA1. We first conducted transduction experiments on 293T cells; these cells are more prone to get infected compared to the other cell lines, and less amount of recombinant virus is required to get a good percentage of infected cells. Following transduction with unconcentrated IN-defective vectors (MOI 5 according to cpm content), percentage of GFP positive 293T-EBNA1 cells was assessed over-time by FACS analysis (Figure 15). Also in these experiments, we have seen that we gained the highest percentage of GFP positive cells after transduction with the two IN-defective vectors TY2-OriPmut-CMV-GFP-I2-Neo and TY2-OriPwt-CMV-GFP-I2-Neo, containing the EBV *oriP* before the transcriptional unit, resulting in an average of 75.6 % and 83.8 % GFP positive cells at early (day 3) time-point, respectively (Figure 15A and 15D, left panel); these vectors are

able to persist in EBNA1 expressing cells for two weeks, reaching an average of 15% GFP positive cells 10 days post infection, but failed to persist later, decreasing to background levels by day (Figure 15); parallel transduction of 293T cells (not expressing EBNA1) with the same two vectors, resulted in an average of 48% and 55% respectively (Figure 15A and 15D, right panel) at early timepoint (day 3), and rapidly decreased to background level seven days post infection. In the same way, transduction of EBNA1 expressing and not expressing cells with the vector without the *oriP* (i.e. TY2-CMV-GFP-I2-Neo) results in the rapid fall of GFP positive cells at an early time-point (Figure 15 E). These results suggest that the presence of the EBV *oriP* in the context of an IDLV, is able to confer a replicative advantage to the vector bearing it, with a detectable signal for two weeks. We decided to further continue our experiments using only the vector containing the *oriP*s before the transcriptional unit, since they presented the higher ability to replicate in EBNA1 positive cells. To confirm that the same EBNA1-dependent transgene expression occurs at a lower dose of vector, we repeated 293T-EBNA1 cells transduction using less amount of virus (MOI 0.5) according to FACS titration (functional titer) (Figure 16), to be sure that the effect observed was effectively due to EBNA1-*oriP* interaction. Cells transduced with the *oriP* negative vector showed maximal GFP expression at three days post infection with complete loss of expression by day 6 (Figure 16). The two *oriP*-containig vectors showed a similar expression profile in EBNA1-negative cells, with loss of GFP expression by day 6 post-infection. By contrast in EBNA1 positive cells GFP expression was maintained by the *oriP* vectors out to 14 days post-infection. Interestingly, in both set of experiments (Figures 15-16) we observed a higher percentage of cells expressing GFP from *oriP*-containing vectors in EBNA1 positive population (40%) compared to EBNA1-negative population of cells (10%). This could be the result of EBNA1-mediated transcriptional up-regulation of GFP from the *oriP* vectors soon after cells transduction, resulting in a four times higher GFP expression.

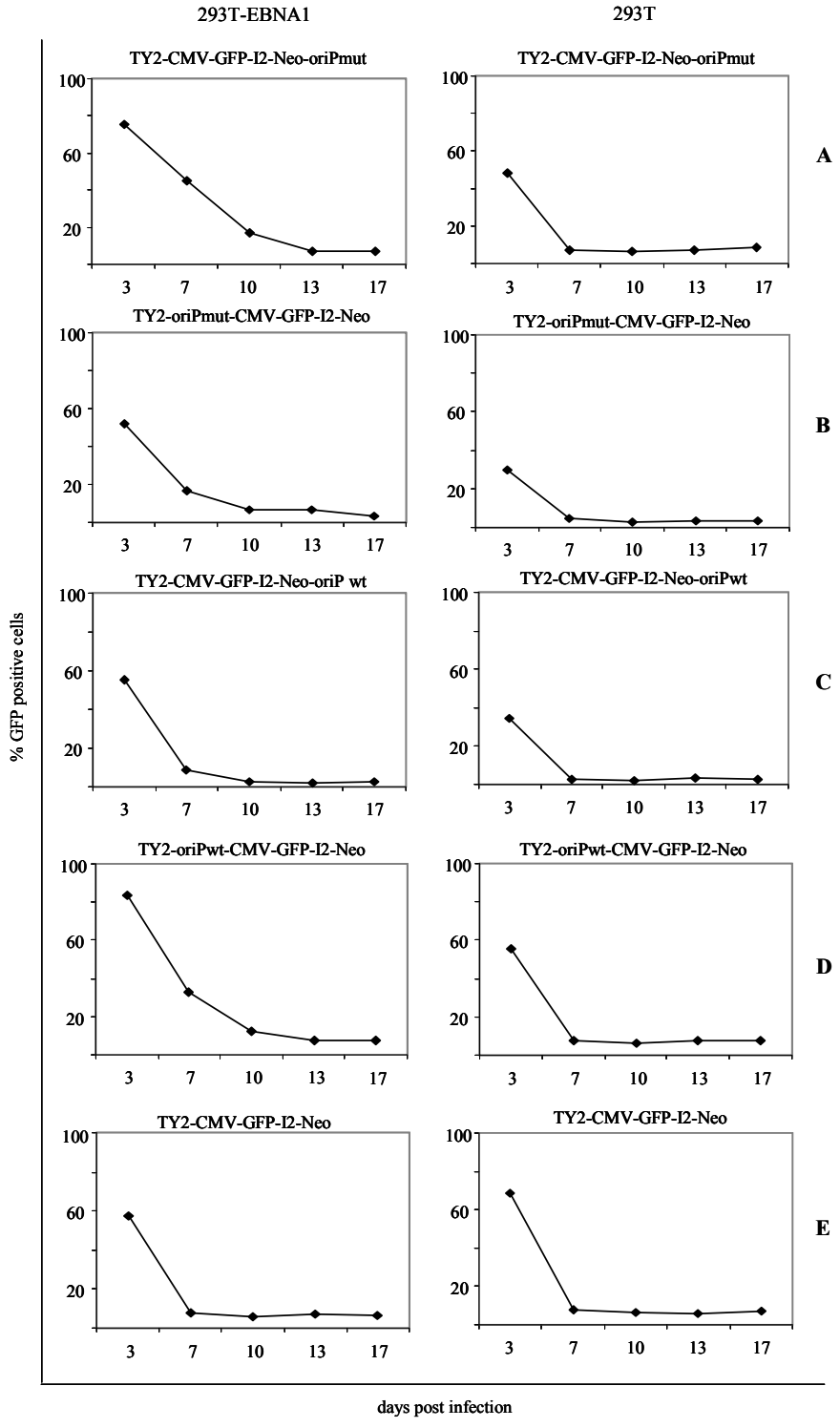

Figure 15. Time course analysis of GFP positive cells obtained after transduction with the indicated IDLVs. Experiments were done in duplicate using 5 MOI according to cpm content.

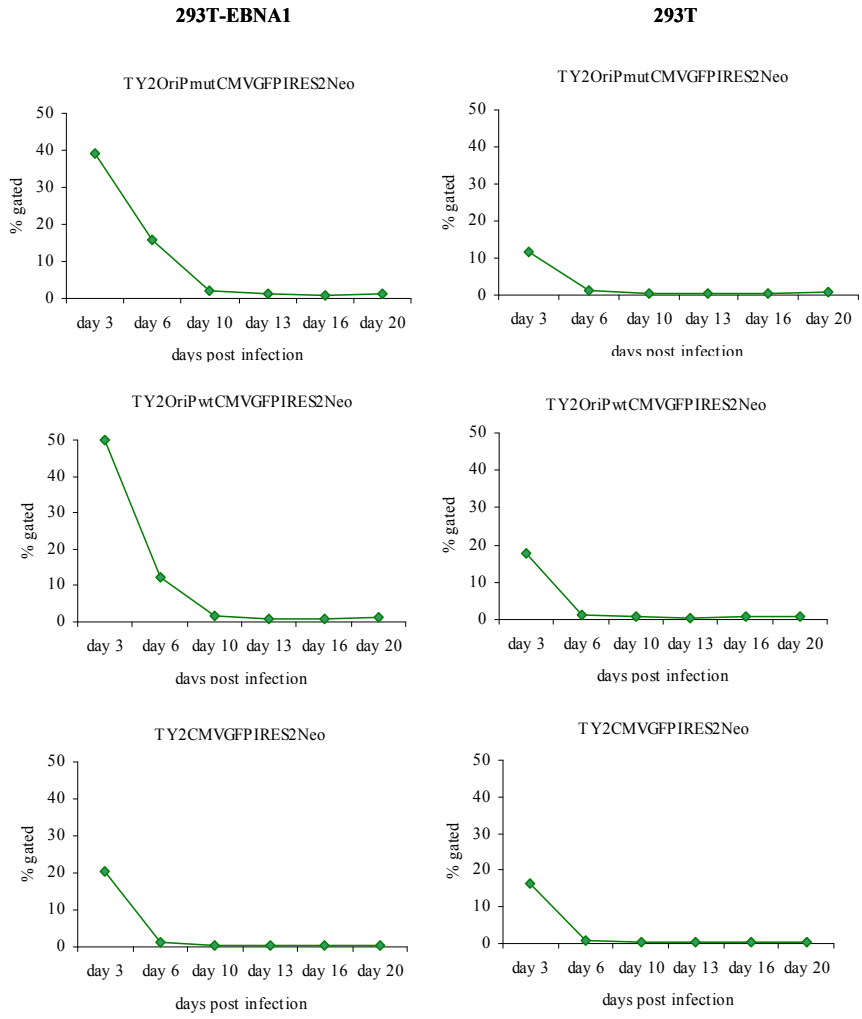

Figure 16. Time course analysis of GFP positive cells obtained after transduction with the indicated IDLVs. Experiments were done in duplicate using 0.5 MOI according to FACS titration.

For the transduction experiments on the other cell lines we used concentrated viruses. In particular, we made concentrated viral preparation of the TY2-oriPmut-CMV-GFP-IRES2-Neo and TY2-CMV-GFP-IRES2-Neo vectors. These viral preparations have been used to transduce RAEL B cells (EBV positive), DG75 B cells (EBV negative) and Hela-EBNA1 cells. The results of the transduction experiments on those cells are shown in Figures 17 and 18, respectively. We were able to appreciate also on these cells the replicative advantage that the *oriP* carrying vectors have in EBNA1 expressing cells, even if at a lower extent than in 293T-EBNA1 cells.

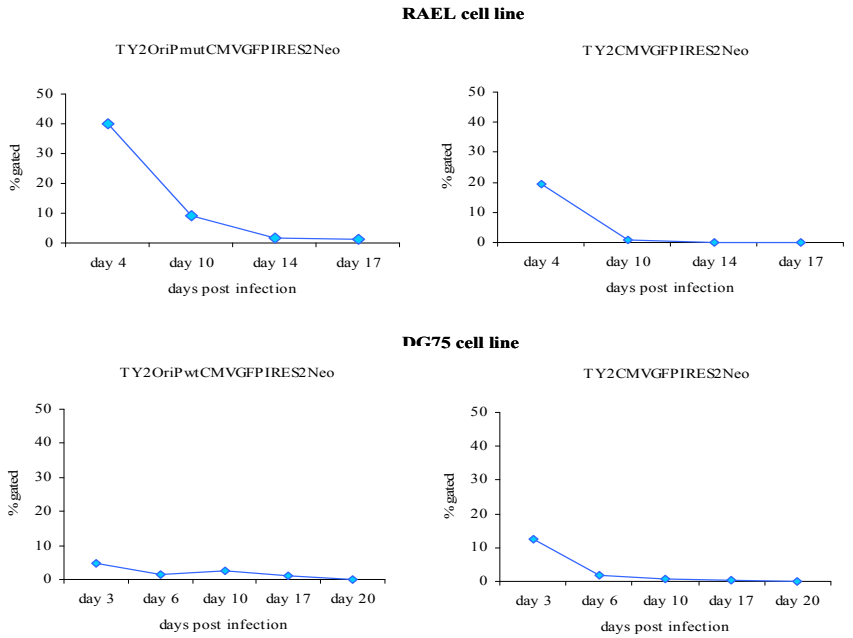

Figure 17. Time course analysis of GFP positive RAEL and DG75 cells obtained after transduction with the indicated IDLVs. For these experiments a concentrated viral preparation has been used (MOI5).

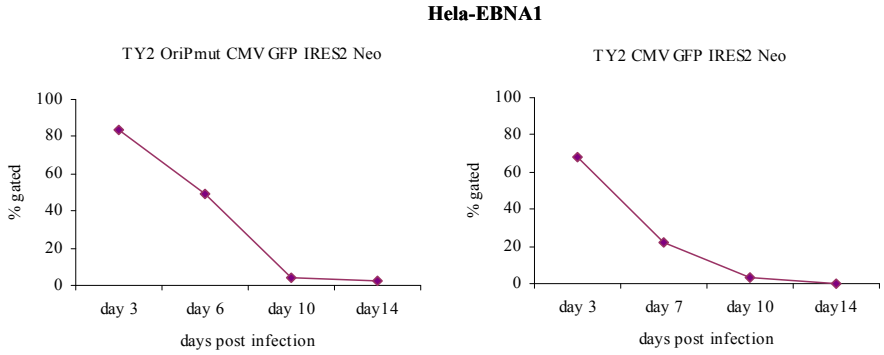

Figure 18. Time course analysis of GFP positive Hela-EBNA1 cells obtained after transduction with the indicated IDLVs. For these experiments a concentrated viral preparation has been used (MOI5).

Discussion

IDLVs show promise for achieving gene expression without integration, preserving some benefits of LVs, whereas reducing the potentially detrimental risk of insertional mutagenesis. Numerous reports supported the ability of these vectors to confer long-term gene expression in slowly dividing cell types for potentially corrective gene expression. These reports also highlighted additional applications of these vectors as delivery platforms for alternate integrative pathways, suicide gene therapy and vaccines (*Bayer et al., 2008; Cornu et al., 2007; Philippe et al., 2006; Negri et al, 2007*). Nevertheless, in cell culture systems the transgene expression duration of IDLVs is highly transient, because of the rapid cell turnover, limiting their use to non-dividing or slowly dividing cells. The instability of the IDLVs in proliferating cells is mainly due to the lack of an origin of replication allowing vector's DNA replication. In a work by Vargas et al., it has been demonstrated that if the transgene cassette carries a viral ORI, IDLV's transgene expression persists if the necessary *trans* protein is supplied. By including the SV40 ORI in IDLVs, transgene expression lasted 56 days, in the SV40 T-antigen containing 293T cell line. This contrasted to parallel transductions in 293 cells, with no SV40 T-antigen, in which transgene expression was short lived (*Vargas et al., 2008*). In the work described in this thesis we have engineered IDLVs to selectively transduce EBNA1-expressing cells by incorporating the EBV *oriP* into the IDLV genome. We have been able to demonstrate that the plasmids carrying the *oriP* selectively replicate in EBNA1 positive cells, showing a four time higher replication efficiency of the control vector lacking the *oriP* (Figures 13a-13b). We have also achieved EBNA1-dependent transcriptional up-regulation and prolonged expression of the IDLV-encoded transgene (GFP) in all the cell lines transduced with those vectors (Figures 15-18). Exploitation of EBNA1 to maintain DNA that contains the *oriP* has obvious use for gene therapy. Epstein–Barr virus (EBV) is widely associated with disease in both immunocompromised and immunocompetent hosts and has been implicated in approximately 1% of tumors worldwide. Current therapies are poorly effective and often toxic (*Carbone et al., 2008*), therefore, the development of an effective, tumor-specific treatment is of high priority. By engineering a suicide gene into one of those *oriP*-containing vectors, specific killing of EBNA1-expressing cells can be induced. The goal of IDLV-mediated suicide gene therapy of EBV-driven diseases would be to eliminate B cells that are latently infected with EBV using an IDLV that contains the EBV elements required for EBNA1-induced episome retention and transcriptional up-regulation. Other studies have previously engineered adenoviral vectors to specifically kill EBV

transformed nasopharyngeal carcinoma cells (*Li et al., 2002; Feng et al., 2002*) and, although these vectors effectively transduce epithelial cells, they would not be suitable for treatment of B cell diseases as a result of their low transduction efficiency of B cells. In the present study, we achieved effective gene transfer to B cells using IDLVs, making these vectors better candidates for gene therapy approach against EBV mediated B cell diseases. By using IDLVs that contain the EBV *oriP*, we successfully targeted EBNA1- positive cells because the IDLV genome is lost over time in EBNA1-negative dividing cells. To move these vectors towards *in vivo* testing, it will also be necessary to optimize IDLVs for B cells targeting (i.e. by incorporating a B cell-specific anti-CD20 single chain antibody into the vector envelope) (*Funke et al., 2008*). In conclusion, IDLVs represent a class of novel viral vectors with several established and emerging applications. These vectors support lower levels of gene expression, and efforts should be made to identify potential expression limiting *trans* acting factors and to boost episomal gene expression. A number of potential applications have been shown for IDLVs, but further *in vivo* testing is required to show the utility of these techniques.

Bibliography

- ✓ Adachi A, Ono N, Sakai H, Ogawa K, Shibata R, Kiyomasu T, Masuike H, Ueda S. *Generation and characterization of the human immunodeficiency virus type 1 mutants*. Arch Virol 1991; 117 (1-2): 45-58
- ✓ Aiken C. *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*. J Virol 1997; 71: 5871–5877.
- ✓ Ambinder RF and Lin L. *Mononucleosis in the laboratory*. J Infect Dis 2005; 192: 1503–1504.
- ✓ Anderson EC, Lever AM. *Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation*. J Virol 2006; 80: 10478-10486
- ✓ Anderson JL, Johnson AT, Howard JL, Purcell DF. *Both linear and discontinuous ribosome scanning are used for translation initiation from bicistronic human immunodeficiency virus type 1 env mRNAs*. J Virol 2007; 81: 4664-4676
- ✓ Ao Z, Yao X, Cohen EA. *Assessment of the role of the central DNA flap in human immunodeficiency virus type 1 replication by using a single-cycle replication system*. J Virol 2004; 78: 3170–3177
- ✓ Apolonia L, Waddington SN, Fernandes C, Ward NJ, Bouma G, Blundell MP, Thrasher AJ, Collins MK, Philpott NJ. *Stable gene transfer to muscle using non-integrating lentiviral vectors*. Mol Ther 2007; 15: 1947–1954.
- ✓ Arumugam PI, Scholes J, Perelman N, Xia P, Yee JK, Malik P. *Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element*. Mol Ther 2007; 15: 1863–1871
- ✓ Barboric M, Peterlin BM. *A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation*. PLoS Biol 2005; 3: e76.
- ✓ Baumann M, Feederle R, Kremmer E, Hammerschmidt W. *Cellular transcription factors recruit viral replication proteins to activate the Epstein-Barr virus origin of lytic DNA replication, oriLyt*. EMBO J 1999; 18: 6095–6105
- ✓ Bayer M, Kantor B, Cockrell A, Ma H, Zeithaml B, Li X, McCown T, Kafri T. *A large U3 deletion causes increased in vivo expression from a nonintegrating lentiviral vector*. Mol Ther 2008; 16(12): 1968-76
- ✓ Berkowitz R, Ilves H, Lin WY, Eckert K, Coward A, Tamaki S, Veres G, Plavec I. *Construction and molecular analysis of gene transfer systems derived from bovine immunodeficiency virus*. J Virol 2001a; 75: 3371-3382
- ✓ Berkowitz RD, Ilves H, Plavec I, Veres G. *Gene transfer systems derived from Visna virus: analysis of virus production and infectivity*. Virology 2001b; 279: 116-129
- ✓ Blake N, Lee S, Redchenko I, Thomas W, Steven N, Leese A, Steigerwald-Mullen P, Kurilla MG, Frappier L, Rickinson A. *Human CD8+ T cell responses to EBV EBNA1: HLA class I presentation of the (Gly-Ala)-containing protein requires exogenous processing*. Immunity 1997; 77: 791–802
- ✓ Bokhoven M, Stephen SL, Knight S, Gevers EF, Robinson IC, Takeuchi Y, Collins MK. *Insertional gene activation by lentiviral and gammaretroviral vectors*. J Virol 2009; 83: 283–294.
- ✓ Bolinger C, Boris-Lawrie. *Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome*. Retrovirology 2009; 6: 8

- ✓ Bornkamm GW and Hammerschmidt W. *Molecular virology of Epstein-Barr virus*. Philos Trans R Soc Lond B Biol Sci 2001; 356: 437–459.
- ✓ Bukovsky AA, Song JP, Naldini L. *Interaction of human immunodeficiency virus-derived vectors with wild-type virus in transduced cells*. J Virol 1999; 73: 7087–7092.
- ✓ Bukrinskaya A. *HIV-1 matrix protein: a mysterious regulator of the viral life cycle*. Virus Res 2007; 124: 1-11
- ✓ Carbone A, Gloghini a, Dotti G. *EBV associated lymphoproliferative disorders: classification and treatment*. Oncologist 2008; 13: 577-585
- ✓ Chung JH, Whiteley M, Felsenfeld G. *A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila*. Cell 1993; 74: 505–514.
- ✓ Cornu TI, Cathomen T. *Targeted genome modifications using integrase-deficient lentiviral vectors*. Mol Ther 2007; 15(12): 2107-13
- ✓ Cui Y, Iwakuma T, Chang LJ. *Contributions of viral splice sites and cis-regulatory elements to lentivirus vector function*. J Virol 1999; 73: 6171–6176.
- ✓ Daniel R, Smith JA. *Integration site selection by retroviral vectors: molecular mechanism and clinical consequences*. Hum Gene Ther 2008; 19: 557-568.
- ✓ Deacon EM, Pallesen G, Niedobitek G, Crocker J, Brooks L, Rickinson AB, Young LS. *Epstein–Barr virus and Hodgkin's disease: transcriptional analysis of virus latency in the malignant cells*. J. Exp. Med. 1993; 177: 339–349.
- ✓ Decaussin, G, Sbih-Lammali, F, de Turenne-Tessier, M, Bouguermouh, A & Ooka, T. *Expression of BARF1 gene encoded by Epstein–Barr virus in nasopharyngeal carcinoma biopsies*. Cancer Res. 2000; 60: 5584–5588
- ✓ Dhar SK, Yoshida K, Machida Y, Khaira P, Chaudhuri B, Wohlschlegel JA, Leffak M, Yates J, Dutta A. *Replication from oriP of Epstein–Barr virus requires human ORC and is inhibited by geminin*. Cell 2001; 106 (3): 287–296.
- ✓ Dull T, Zufferey R, Kelly M, Mandel RJ, Nguyen M, Trono D, Naldini L. *A third-generation lentivirus vector with a conditional packaging system*. J Virol 1998; 72: 8463–8471.
- ✓ Engelman A. *In vivo analysis of retroviral integrase structure and function*. Adv Virus Rev 1999; 52: 411-26
- ✓ Feng WH, Westphal E, Mauser A, Raab-Traub N, Gulley ML, Busson P, Kenney SC. *Use of adenovirus vectors expressing Epstein-Barr virus (EBV) immediate-early protein BZLF1 or BRLF1 to treat EBV-positive tumors*. J Virol. 2002; 76(21): 10951-9
- ✓ Fixman ED, Hayward GS, Hayward SD. *Replication of Epstein-Barr virus oriLyt: lack of a dedicated virally encoded origin-binding protein and dependence on Zta in cotransfection assays*. J Virol 1995; 69: 2998–3006
- ✓ Fujii K, Yokoyama N, Kiyono T, Kuzushima K, Homma M, Nishiyama Y, Fujita M, Tsurumi T. *The Epstein-Barr virus pol catalytic subunit physically interacts with the BBLF4-BSLF1-BBLF2/3 complex*. J Virol 2000; 74: 2550–2557
- ✓ Funke S, Maisner A, Mühlebach MD, Koehl U, Grez M, Cattaneo R, Cichutek K, Buchholz CJ. *Targeted cell entry of lentiviral vectors*. Mol Ther. 2008; 16(8): 1427-36

- ✓ Gibbs JS, Regier DA, Desrosiers RC. *Construction and in vitro properties of HIV-1 mutants with deletions in "nonessential" genes*. AIDS Res Hum Retroviruses 1994; 10: 343–350.
- ✓ Gillim-Ross L, Cara A, Klotman ME. *HIV-1 extrachromosomal 2-LTR circular DNA is long-lived in human macrophages*. Viral Immunol 2005; 18: 190–196
- ✓ Goh WC, Rogel ME, Kinsey CM, Michael SF, Fultz PN, Nowak MA, Hahn BH, Emerman M. *HIV-1 Vpr increases viral expression by manipulation of the cell cycle: a mechanism for selection of Vpr in vivo*. Nat Med 1998; 4: 65-71
- ✓ Greco O, Dachs GU. *Gene Directed Enzyme/Prodrug Therapy of Cancer: Historical Appraisal and Future Prospectives*. J Cell Physiol. 2001; 187(1): 22-36.
- ✓ Hacein-Bey-Abina S, von Kalle C Hacein-Bey-Abina S, von Kalle C, Schmidt M, Le Deist F, Wulffraat N, McIntyre E, Radford I, Villeval JL, Fraser CC, Cavazzana-Calvo M, Fischer A. *A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency*. N Engl J Med 2003; 348: 255-256.
- ✓ Hanawa H, Yamamoto M, Zhao H, Shimada T, Persons DA. *Optimized lentiviral vector design improves titer and transgene expression of vectors containing the chicken beta-globin locus HS4 insulator element*. Mol Ther 2009; 17: 667–674.
- ✓ Higashimoto T, Urbinati F, Perumbeti A, Jiang G, Zarzuela A, Chang LJ, Kohn DB, Malik P. *The woodchuck hepatitis virus post-transcriptional regulatory element reduces readthrough transcription from retroviral vectors*. Gene Ther 2007; 14: 1298–1304.
- ✓ Hino S, Fan J, Taguwa S, Akasaka K, Matsuoka M. *Sea urchin insulator protects lentiviral vector from silencing by maintaining active chromatin structure*. Gene Ther 2004; 11: 819–828.
- ✓ Hislop AD, Annels NE, Gudgeon NH, Leese AM & Rickinson AB. *Epitope-specific evolution of human CD8+ T cell responses from primary to persistent phases of Epstein–Barr virus infection*. J Exp Med 2002; 195: 893-905
- ✓ Humme S, Reisbach G, Feederle R, Delecluse HJ, Bousset K, Hammerschmidt W, Schepers A. *The EBV nuclear antigen1 (EBNA1) enhances B cell immortalisation several thousandfold*. Proc. Natl Acad. Sci. USA 2003; 100: 10989–10994
- ✓ Iordanskiy S, Berro R, Altieri M, Kashanchi F, Bukrinsky M. *Intracytoplasmic maturation of the human immunodeficiency virus type 1 reverse transcription complexes determines their capacity to integrate into chromatin.* Retrovirology 2006; 12; 3: 4.
- ✓ Iwakuma T, Cui Y, Chang LJ. *Self-inactivating lentiviral vectors with U3 and U5 modifications*. Virology 1999; 261: 120–132.
- ✓ Jablonski JA, Caputi M. *Role of cellular RNA processing factors in human immunodeficiency virus type 1 mRNA metabolism, replication, and infectivity.* J Virol 2009; 83: 981–992.
- ✓ Kennedy, G., Komano, J. & Sugden, B. *Epstein–Barr virus provides a survival factor to Burkitt's lymphomas*. Proc. Natl Acad. Sci. USA 2003; 100: 14269–14274
- ✓ Kim VN, Mitrophanous K, Kingsman SM, Kingsman AJ. *Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1*. J Virol 1998; 72: 811–816.
- ✓ Kotsopoulou E, Kim VN, Kingsman AJ, Kingsman SM, Mitrophanous KA. *A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene*. J Virol 2000; 74: 4839–4852.

- ✓ Kumar M, Keller B, Makalou N, Sutton RE. *Systematic determination of the packaging limit of lentiviral vectors*. Hum Gene Ther 2001; 12: 1893–1905.
- ✓ Kuppers R. *Mechanisms of B-cell lymphoma pathogenesis*. Nat Rev Cancer. 2005; 5(4): 251-62
- ✓ Leavitt AD, Robles G, Alesandro N, Vamus HE. *Human immunodeficiency virus type I integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection*. J Virol 1996; 70(2): 721-8
- ✓ JH, Chia M, Shi W, Ngo D, Strathdee CA, Huang D, Klamut H, Liu FF. *Tumor-targeted gene therapy for nasopharyngeal carcinoma*. Cancer Res. 2002; 62(1): 171-8
- ✓ Lindner SE, Sugden B. *The plasmid replicon of Epstein-Barr virus: mechanistic insights into efficient, licensed, extrachromosomal replication in human cells*. Plasmid 2007; 58(1): 1-12
- ✓ Lindner SE, Zeller K, Schepers A, Sugden B. *The affinity of EBNA1 for its origin of DNA synthesis is a determinant of the origin's replicative efficiency*. J Virol. 2008; 82(12): 5693-702
- ✓ Logan AC, Nightingale SJ, Haas DL, Cho GJ, Pepper KA, Kohn DB. *Factors influencing the titer and infectivity of lentiviral vectors*. Hum Gene Ther 2004; 15: 976–988.
- ✓ Metharom P, Takyar S, Xia HH, Ellem KA, Macmillan J, Sheperd RW, Wilcox GE, Wei MQ. *Novel bovine lentiviral vectors based on Jembrana disease virus*. J Gene Med 2000; 2: 176-185
- ✓ Naldini L, Blomer U, Gage FH, TronoD, Verma IM. *Efficient transfer, integration and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector*. Proc Natl Acad Sci USA 1996a; 93: 11382-11388
- ✓ Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D *In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector.* Science 1996b; 272: 263–267.
- ✓ Nègre D, Mangeot PE, Duisit G, Blanchard S, Vidalain PO, Leissner P, Winter AJ, Rabourdin-Combe C, Mehtali M, Moullier P, Darlix JL, Cosset FL. *Characterization of novel safe lentiviral vectors derived from simian immunodeficiency virus (SIVmac251) that efficiently transduce mature human dendritic cells*. Gene Ther 2000; 7: 1613-1623
- ✓ Negri DR, Michelini Z, Bona R, Blasi M, Filati P, Leone P, Rossi A, Franco M, Cara A. *Integrase-defective lentiviral vector based vaccine: a new vector for induction of T cell immunity.* Expert Opin Biol Ther. 2011; 11(6): 739-50.
- ✓ Ngumbela KC, Ryan KP, Sivamurthy R, Brockman MA, Gandhi RT, Bhardwaj N, Kavanagh DG. *Quantitative effect of suboptimal codon usage on translational efficiency of mRNA encoding HIV-1 gag in intact T cells.* PLoS ONE 2008; 3: e2356.
- ✓ Oh T, Bajwa A, Jia G, Park F. *Lentiviral vector design using alternative RNA export elements.* Retrovirology 2007; 4: 38.
- ✓ Olsen JC. *Gene trasnfer vectors derived from equine infectious anemia virus*. Gene Ther 1998; 5: 1481-1487
- ✓ Osti D, Marras E, Ceriani I, Grassini G, Rubino T, Vigano D, Parolaro D, Perletti G. *Comparative analysis of molecular strategies attenuating positional effects in lentiviral vectors carrying multiple genes*. J Virol Methods 2006; 136: 93–101.
- ✓ Philippe S, Sarkis C, Barkats M, Mammeri H, Ladroue C, Petit C, Mallet J, Serguera C. *Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo*. Proc Natl Acad Sci USA 2006; 103(47): 17684-9.

- ✓ Pluta K, Kacprzak MM. *Use of HIV as a gene transfer vector.* Acta Biochim Pol. 2009; 56(4): 531-95.
- ✓ Poeschla EM, Wong -Staal F, Looney DJ. *Efficient transduction of nondividing human cells by feline immunodeficiency virus lentiviral vectors*. Nat Med 1998; 4: 354-357
- ✓ Pond SJ, Ridgeway WK, Robertson R, Wang J, Millar DP. *HIV-1 Rev protein assembles on viral RNA one molecule at a time*. Proc Natl Acad Sci USA 2009; 106: 1404-1408
- ✓ Puigdomenech I, Massanella M, Izquierdo-Useros N, Ruiz-Hernandez R, Curriu M, Bofill M, Martinez-Picado J, Juan M, Clotet B, Blanco J. *HIV transfer between CD4 T cells does not require LFA-1 binding to ICAM-1 and is governed by interaction of HIV envelope glycoprotein with CD4*. Retrovirology 2008; 5: 32
- ✓ Ritzi M, Tillack K, Gerhardt J, Ott E, Humme S, Kremmer E, Hammerschmidt W, Schepers A. *Complex protein-DNA dynamics at the latent origin of DNA replication of Epstein–Barr virus*. J Cell Sci 2003; 116: 3971–3984
- ✓ Santoni de Sio F, Naldini L. *Short term culture of human CD34+ cells for lentiviral gene transfer*. Methods Mol Biol 2009; 506: 530-538
- ✓ Sarkis C, Philippe S. Mallet J, Serguera C. *Non-Integrating lentiviral Vectors*. Curr Gene Ther 2008; 8(6): 430-437
- ✓ Sastry L, Cornetta K. *Detection of replication competent retroviruses and lentiviruses*. Methods Mol Biol 2009; 506: 243-263
- ✓ Schambach A, Bohne J, Baum C, Hermann FG, Egerer L, von Laer D, Giroglou T. *Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression*. Gene Ther 2006; 13: 641– 645.
- ✓ Schambach A, Galla M, Maetzig T, Loew R, Baum C. *Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors*. Mol Ther 2007; 15: 1167–1173.
- ✓ Smith PR, de Jesus O, Turner D, Hollyoake M, Karstegl CE, Griffin BE, Karran L, Wang Y, Hayward SD, Farrell PJ. *Structure and coding content of CST (BART) family RNAs of Epstein–Barr virus*. J. Virol. 2000; 74: 3082–3092
- ✓ Srinivasakumar N. *HIV-1 vector systems*. Somat Cell Mol Genet 2001; 26: 51-81
- ✓ Stevenson M, Haggerty S, Lamonica CA, Meier CM, Welch SK, Wasiak AJ. *Integration is not necessary for expression of human immunodeficiency virus type I protein products*. J Virol 1990; 64(5): 2421-5
- ✓ Sun J, Li D, Hao Y, Zhang Y, Fan W, Fu J, Hu Y, Liu Y, Shao Y. *Post transcriptional regulatory elements enhance antigen expression and DNA vaccine efficacy*. DNA Cell Biol 2009; 28: 233–240
- ✓ Suzuki Y, Craigie R. *The road to chromatin - nuclear entry of retroviruses*. Nat Rev Microbiol. 2007; 5(3): 187-96
- ✓ Takada, K. & Nanbo, A. *The role of EBERs in oncogenesis*. Semin Cancer Biol 2001; 11, 461–467.
- ✓ Thorley-Lawson DA and Gross A. *Persistence of the Epstein-Barr virus and the origins of associated lymphomas*. N Engl J Med 2004; 350: 1328–1337.
- ✓ Tsurumi T, Fujita M, Kudoh A. *Latent and lytic Epstein-Barr virus replication strategies.* Rev Med Virol 2005; 15(1): 3-15
- ✓ Uchida J, Yasui T, Takaoka-Shichijo Y, Muraoka M, Kulwichit W, Raab-Traub N, Kikutani H. *Mimicry of CD40 signals by Epstein–Barr virus LMP1 in B lymphocyte responses*. Science 1999; 286: 300–303

- ✓ Van Maele B, De Rijck J, De Clercq E, Debyser Z. *Impact of the central polypurine tract on the kinetics of human immunodeficiency virus type 1 vector transduction*. J Virol 2003; 77: 4685–4694
- ✓ Vargas J Jr, Klotman ME, Cara A. *Conditionally replicating lentiviral-hybrid episomal vectors for suicide gene therapy*. Antiviral Res 2008; 80(3): 288-94
- ✓ Veron P, Boutin S, Martin S, Chaperot L, Plumas J, Davoust J, Masurier C. *Highly efficient transduction of human plasmacytoid dendritic cells without phenotypic and functional maturation*. J Transl Med 2009; 7:10
- ✓ Wang GP, Levine BL, Binder GK, Berry CC, Malani N, McGarrity G, Tebas P, June CH, Bushman FD. *Analysis of lentiviral vector integration in HIV+ study subjects receiving autologous infusions of gene modified CD4+ T cells*. Mol Ther 2009; 17: 844-850
- ✓ Wang J, Lindner SE, Leight ER, Sugden B**.** *Essential elements of a licensed, mammalian plasmid origin of DNA synthesis*. Mol Cell Biol 2006; 26 (3): 1124–1134
- ✓ Wang JH, Kwas C, Wu L. *Intercellular adhesion molecule 1 (ICAM-1), but not ICAM-2 and -3, is important for dendritic cell-mediated human immunodeficiency virus type 1 transmission*. J virol 2009b; 83: 4195-4204
- ✓ Wilson, J. B., Bell, J. L. & Levine, A. J. *Expression of Epstein–Barr virus nuclear antigen-1 induces B cell neoplasia in transgenic mice*. EMBO J. 1996; 15: 3117–3126
- ✓ Yáñez-Muñoz RJ, Balaggan KS, MacNeil A, Howe SJ, Schmidt M, Smith AJ, Buch P, MacLaren RE, Anderson PN, Barker SE, Duran Y, Bartholomae C, von Kalle C, Heckenlively JR, Kinnon C, Ali RR, Thrasher AJ. *Effective gene therapy with nonintegrating lentiviral vectors*. Nat Med 2006; 12(3): 348-53
- ✓ Young LS and Rickinson AB. Epstein-Barr virus: 40 years on. Nat Rev Cancer 2004; 4(10): 757-68
- ✓ Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D. *Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo*. Nat Biotechnol 1997; 15: 871–875.
- ✓ zur Hausen A, Brink AA, Craanen ME, Middeldorp JM, Meijer CJ, van den Brule AJ. *Unique transcription pattern of Epstein–Barr virus (EBV) in EBV-carrying gastric adenocarcinomas: expression of the transforming BARF1 gene*. Cancer Res 2000; 60: 2745–2748