A GENE TRANSFER SYSTEM DERIVED FROM HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1)

by

Maria Fuller

(B. Appl. Sc., M. Appl. Sc.)

Thesis submitted for the degree of

Doctor of Philosophy

in

The University of Adelaide

(Faculty of Medicine)

Lysosomal Diseases Research Unit,

Department of Chemical Pathology and Department of Paediatrics, Faculty of Medicine

Women's and Children's Hospital

South Australia

October 2001

Science knows no country, because knowledge belongs to humanity, and is the torch which illuminates the world. Science is the highest personification of the nation because that nation will remain the first which carries the furthest the works of thought and intelligence.

Louis Pasteur

TABLE OF CONTENTS

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

TABLE OF CONTENTS

CHAPTER TWO

MATERIALS AND METHODS

TABLE OF CONTENTS

CHAPTER TWO

MATERIALS AND METHODS

CHAPTER THREE

GENE TRANSFER FOR β-GLUCURONIDASE

TABLE OF CONTENTS

CHAPTER FOUR

HELPER PLASMIDS FOR HIV-1 DERIVED VECTORS

TABLE OF CONTENTS

CHAPTER FIVE

EVALUATION OF HIV-1 BASED GENE TRANSFER VECTORS

CHAPTER SIX

FINAL STATEMENTS

LIST OF ABBREVIATIONS

AAV	adeno-associated virus
AIDS	acquired immunodeficiency syndrome
ALV	avian leukemia virus
bp	base pair
CA	capsid protein
CHO	Chinese hamster ovary
CIP	calf intestinal phosphatase
CNS	central nervous system
cpm	counts per minute
CTE	constitutive transport element
dATP	deoxyadenosine triphosphate
DMEM	Dulbecco's Modified Eagles Medium
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotides
DTT	1,4-Dithiothreitol
dTTP	deoxythymidine triphosphate
Env	envelope
FCS	foetal calf serum
G418	Geneticin
h	hour (s)
HFV	human foamy virus
HIV-1	human immunodeficiency virus type 1
HSC	haematopoietic stem cell
HSV	herpes simplex virus
HTLV-1	human T cell leukemia virus
IN	integrase
IRES	internal ribosome entry site
ITR	inverted terminal repeats
kb	kilobase
kDa	kilo-Dalton
L	litre (s)
LTR	long terminal repeats
MA	matrix antigen
ml	millilitre

MLV	Moloney leukemia virus
min	minute (s)
MoMLV	Moloney murine leukemia virus
MPS	mucopolysaccharidosis
MPS VII	mucopolysaccharidosis type VII
mRNA	messenger RNA
NC	nucleocapsid
Neo^R	neomycin resistant
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PHSC	pluripotent haematopoietic stem cell
PPT	polypurine tract
pur^R	puromycin resistant
RER	rough endoplasmic reticulum
RNA	ribonucleic acid
RNase H	ribonuclease
RRE	Rev response element
RSV	Rous sarcoma virus
RT	reverse transcriptase
s	second (s)
SDS	sodium dodecyl sulphate
SIN	self-inactivating
sssDNA	strand strong stop deoxyribonucleic acid
SV40	simian virus early promoter
TAR	*tat* activation response element
VSV-G	vesicular stomatitis virus G glycoprotein

THESIS EPITOME

Gene therapy has been slow to yield its promised benefits, in many instances because of a lack of efficient gene transfer vehicles to enable delivery of the correct amount of gene product to target cells. Vectors derived from human immunodeficiency virus type 1 (HIV-1) appear an attractive option for many gene therapy applications. This is due to their ability to transduce non-cycling cell populations and integrate their genome into the chromosome of a host cell, thereby resulting in the stable genetic modification of the transduced cell. The initial work in this thesis was to construct an HIV-1 derived vector to transfer the lysosomal enzyme β-glucuronidase to mucopolysaccharidosis type VII (MPS VII) fibroblasts. These early attempts were unsuccessful. However, the publication of several HIV-1 based gene therapy reports that demonstrated the direct *in vivo* transduction of several tissues including the central nervous system, retinal and liver however, validated the basic concept. Nevertheless, the pathogenic nature of HIV-1 has raised considerable concerns about the safety of such vector systems.

In order to help address these concerns, each of the primary transcriptional units encoding the *trans* functions relevant for vector production was isolated. Coding sequences for the HIV-1 Tat (exon 1), Vpr, Vpu, Vif and Nef proteins were expressed from separate plasmid constructs. In each case a minimal reading frame was linked to heterologous transcriptional regulatory sequences. The native HIV-1 *gagpol* gene sequence was isolated from a proviral clone of HIV-1 YU2 and expressed in an individual plasmid. In addition a *gagpol* sequence codon-optimised for expression in mammalian cells, but retaining the HIV-1 translational frameshift for Pol expression and region of overlap between the *gag* and *pol* genes, was also constructed. The expression plasmid encoding the codon-optimised *gagpol* sequence resulted in efficient Rev/Rev response element (RRE) independent expression of Gag (p24) and Pol (reverse transcriptase, RT). Together with plasmids encoding the Rev and the vesicular stomatitis virus glycoprotein (VSV-G), and a HIV-1 derived vector, these constructs were able to support the generation of recombinant virus particles upon transient expression in 293T cells.

Separate expression plasmids for the Gag and GagPol polyproteins were subsequently made, again using codon-optimised sequence. In this instance neither sequence contains the HIV-1 translational frameshift signal. In combination these two plasmids can functionally replace the use of the codon-optimised *gagpol* expression construct that utilises the translational frameshift signal present in the native coding sequence. The use of separate expression plasmids for the Gag and GagPol polyproteins allows the removal of the frameshift signal in the *gagpol* gene from the virus production

system. As neither the Gag, nor the GagPol, expression construct alone can support the production of detectable virus titres, enhanced safety is achieved with such an approach. To demonstrate the safety of the HIV-1 gene transfer system, two assays were developed to measure the transfer of sequences functionally equivalent to the *gagpol* gene (ie capable of expressing Gag and GagPol polyproteins) to transduced cells.

The expression of the HIV-1 accessory proteins Vpr, Vif, Vpu and Nef was evaluated by western blot analysis. Only the expression of Vpr was detected. Inclusion of expression plasmids for these proteins in the virus production system caused a small, but significant increase in viral titre. Whilst these minor proteins are not considered necessary for the production of VSV-G pseudotyped recombinant virus in 293T cells, their role in both the production of virus of different pseudotypes, in different cell lines, and on virus phenotype, warrant closer attention.

In order to fully utilise the promise of HIV-1 it is important that the effects of viral *cis* sequence elements on vector function are fully delineated. To this end, the effect on vector function of various *cis* elements from the HIV-1 YU2 genome that have been implicated as either affecting vector performance, or HIV-1 replication, were evaluated. In addition the possibility of functionally substituting heterologous sequence elements in place of Rev/RRE that is otherwise required for efficient vector function, was assessed. Several HIV-1 derived vector constructs were made and transiently transfected into 293T cells along with a suitable helper plasmid and a plasmid encoding VSV-G. Titres of virus particles produced were examined on NIH3T3 cells. Apart from the minimal sequence elements required for effective vector function, a number of sequence elements did increase vector titre. These included the amount of *gag* sequence, the central polypurine tract, the 5' and 3' ends of the *env* gene and sequences immediately upstream from the polypurine tract.

STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this thesis copy, when deposited in the University Library, being available for loan and photocopying.

Maria Fuller

ACKNOWLEDGMENTS

Firstly and foremostly I would like to thank my supervisors, Dr Don Anson and Prof. John Hopwood. A great deal of credit for the breadth of the work covered in this thesis must go to Don Anson, for not only initiating and continuing this project but also for the large amount of experimentation that was performed by him which made this thesis possible. In particular he was responsible for the synthesis of the MPMV CTE sequence and the construction of the gagml, polml and gagpolfusionml sequences and the F4ΔF6, F3RRE, F4RRE, F6RRE, 5'envRRE, cppt, central RRE and CTE sequence elements. As such, I owe an enormous debt of gratitude to Don Anson for the completion of this PhD project and his unfailing belief in this work. To John Hopwood, many thanks for the support and encouragement, your advice will always remain with me.

Extended thanks to my friends and staff of the Department of Chemical Pathology at the Women's and Children's Hospital for creating such an enjoyable and stimulating work environment.

Many thanks to those people mentioned in the thesis for the provision of reagents.

Last but not least, my everlasting gratitude to my husband Peter. You know what you have done, words just do not suffice.

CHAPTER ONE

INTRODUCTION

1.1 HUMAN GENE THERAPY

1.1.1 WHERE ARE WE TODAY?

Gene therapy is defined as the introduction of normal genes into cells in place of defective or missing ones in order to correct genetic disorders (The Oxford English Reference Dictionary, 1996). In practice somatic cell gene therapy does not imply the replacement of a non-functional gene with a functional one but rather the complementation of a gene. This concept is not new, early exploits at human gene therapy appeared before the advent of recombinant DNA technology. Probably the first acknowledged attempt was by Stanfield Rogers who injected the Shope papilloma virus into patients suffering from arginase deficiency because earlier work had led him to believe that this virus contained an arginase gene (Rogers, 1959; Rogers *et al.*, 1973). Since then substantial progress in gene therapy has been attributed to both the ability to manipulate recombinant DNA and the advancements in techniques for transfection and selection systems for cultured cells. The definition of gene therapy has thus broadened to describe nucleic acid transfer, either DNA or RNA, to target cells for some therapeutic purpose. This expanded definition stands in contrast to the earlier held notion that gene therapy is aimed strictly at genetic disease. Many gene therapy protocols are currently directed at cancers, infectious and cardiovascular diseases, or at the stimulation of specific cellular actions such as cell killing, immune reactions (Rosenberg *et al.*, 1998), prodrug activation (Rigg and Sikora, 1997) and the production of molecular decoys required for viral replication (Lisziewicz *et al.*, 1993). There are essentially three classes of somatic cell gene therapy. The first is *ex vivo*, which involves removal of cells from the body followed by incubation with the desired gene and then the return of the genetically modified cells to the body. The second class is *in situ* which describes the direct administration of the gene transfer vehicle into affected cells or tissues. The final type, of which there are no examples, but is probably the most desirable scenario refers to *in vivo* gene therapy whereby the vector would be directly injected into the bloodstream, with subsequent targeting by virtue of the properties endogenous to the vector used (Anderson, 1998).

The first clinically approved therapeutic protocol for gene therapy started trials in 1990 for patients with adenosine-deaminase deficiency (severe combined immunodeficiency, SCID) (Blaese *et al.*, 1995). Despite the fact that the patients remained on enzyme replacement therapy, there appeared to be some long term benefit in several patients with SCID (Mullen *et al.*, 1996). Since then over 300 clinical regimes have been approved yet there is only conjectural evidence that a gene therapy procedure has

successfully treated a human disease[1]. Probably the most acute obstacle limiting the success of gene therapy is the inability to transfer the appropriate gene into the desired target cell such that sufficient amounts of the gene product required to treat the disease are synthesised and correctly delivered. For many gene therapy applications, the desired target cells are not easily transfected, are poorly defined or even unknown. A popular target, that exemplifies this problem, has been and still is, the haematopoietic stem cell (HSC) due to its capacity for self renewal as well as expansion and differentiation into all the cell lineages of the haematopoietic system. Furthermore, HSCs and their progeny interact with a number of other organ systems through circulation and migration providing access to the majority of the body. Nonetheless, the frequency of transduction by any method, of the HSC is very low (for an editorial review see Barranger, 1996).

1.1.2 GENE TRANSFER VECTORS

1.1.2.1 NON-VIRAL VECTORS

Gene transfer can be divided into two classes of vectors: viral and non-viral vectors. Naked DNA molecules cannot spontaneously penetrate into cells, non-viral methods of transfer try to achieve this so obviating the need for viral vectors. To some extent the efficiency of DNA transfer depends on the type of cell and its mitotic state. Integration of foreign DNA is a rare event and higher transformation efficiencies occur in rapidly dividing cell populations. Generally speaking non-viral methods of gene transfer have lower transduction efficiencies than their viral counterparts and do not permit the genes to integrate efficiently into the genome of target cells. Whilst this prevents the gene from being passed on to future cell generations, the genes delivered cannot replicate and there is no co-transfer of unwanted viral material to the cell.

1.1.2.1.1 CHEMICAL METHODS

More than 35 years ago it was shown that certain chemicals could facilitate the uptake of DNA into cells (Vaheri and Pagano, 1965). The basic principle is that negatively charged DNA molecules can be complexed with positively charged chemicals to make a more favourable product for endocytosis by the target cell. High salt and polycations have been used (Felgner, 1990) as well as DEAE-dextran (Vaheri and Pagano, 1965) and calcium phosphate (Graham and Eb, 1973).

[1]. ORDA Report: Human Gene Therapy Protocols (2/3/98) (Office of Recombinant DNA Activities, NIH, Bethesda, MD, 1998)

Of the non-viral gene delivery methods cationic lipid-based delivery systems are the most widely used. A cationic lipid called Lipofectin was developed and was one of the first chemical DNA delivery systems to be used in animals (Felgner *et al.*, 1987). Such liposomes are non-pathogenic, can be used for multiple treatments and unlike viral vectors they are relatively cheap and easy to prepare. There also appears to be no limit to the size of genes that can be delivered. Initial success with cationic lipids promoted further developments aimed at increasing transfection efficiency (Zhu *et al.*, 1993). Helper lipids have been incorporated into the liposome complex and an excess of lipid over DNA in the reaction seems to augment the transfer process (Felgner *et al.*, 1994). Still, lipid-based systems have important drawbacks, including the lack of targeting, the poorly understood structure of DNA-lipid complexes and variations that arise during fabrication (Filion and Phillips, 1998).

A further attempt at non-viral gene transfer has been to conjugate DNA to cationic peptides. For example, polylysine has been used with certain peptides to target DNA transfer to specific cell surface proteins and therefore to particular cell types (Feero *et al.*, 1997). Polylysine has been used to link DNA to inactivated adenovirus such that gene transfer can occur with the efficiency of adenoviral mediated uptake (Nguyen *et al.*, 1997). A combination of the polycationic peptide protamine sulphate has been used in combination with cationic lipids to enhance DNA delivery (Sorgi *et al.*, 1997). Short peptides (Schwartz *et al.*, 1999), fusion proteins (Paul *et al.*, 1997), as well as numerous other cationic polymers have been used to mediate uptake of DNA into cells (Luo and Saltzman, 2000). Biocompatible controlled release polymers, an advance on all the methods described above, enables long-term release without repeat administration, has also been reported. Matrices of poly D, L-lactide-co-glycolide (PLGA) were loaded with plasmid, which was subsequently released up to a month *in vitro*. Consequently, delivery of functional DNA to neighbouring cells was achieved (Shea *et al.*, 1999).

1.1.2.1.2 MECHANICAL METHODS

The simplest method of gene delivery is the direct injection of naked DNA into the nucleus of a target cell. While this is conceptually very simple it is not very practical. Various modifications of this principle have been applied including the use of high pressure (Mann *et al.*, 1999), and hydrodynamic injection (Zhang *et al.*, 1999). A procedure known as electroporation, which makes use of high voltage electric pulses to transiently open pores in cell membranes, has also been used to deliver DNA to a myriad of cell types (Neumann *et al.*, 1982; Aihara and Miyazaki, 1998).

Another physical approach to gene therapy has been *via* gene gun technology whereby the DNA is loaded onto tiny gold particles and fired into target cells *via* a helium gun. As with liposomes, the duration of gene expression is transient. The DNA/gold particles are cheap and easy to prepare but despite a claim of being innocuous in the target cells, their presence remains for a lengthy period of time. Plasmid coated gold beads have been used to transfer a viral replication inhibiting transdominant *rev* mutant into CD4+ T cells and shown to prolong target cell survival in HIV-1 infected individuals (Woffendin *et al.*, 1996). This technology has application for gene transfer to the skin (ie for immunisation) and would also show promise *ex vivo* where tissues could be targeted in culture and then returned to the patient. One major attraction to this technology is that it can be used to transfer genes to non-dividing cells, a prerequisite for many gene therapy applications, including HSCs.

In conclusion, non-viral methods of gene transfer are clearly advantageous when transient expression is the desired outcome, and their strength lies in their safety and ease of manufacture. The most notable drawback of the non-viral methods is that they are unable to offer stable gene expression.

1.1.2.2 VIRAL VECTORS

Viruses have evolved to become highly efficient in overcoming human barriers to infection and being able to insert their genetic material into human cells. For this reason they have become very popular in gene therapy applications with 80% of clinical protocols using viral vectors (Anderson, 1998). The ability of DNA and RNA tumour viruses to cause transformation *via* transfer of their genetic material has precipitated a plethora of studies investigating their use as gene transfer vehicles (Weiss *et al.*, 1982). Apart from the more established viral vectors, there are many different types of both DNA and RNA viruses that have been and are currently being explored including papovaviruses (eg SV40, Madzak *et al.*, 1992) and vaccinia virus (Moss, 1992). Clinical experience with these vectors remains limited, although it already appears that they suffer from many of the same problems affecting the more distinguished viral vectors that are described below.

1.1.2.2.1 ADENOVIRAL VECTORS

The adenoviridae is a large family of non-enveloped, double stranded linear DNA viruses comprising over 40 known human serotypes. Their genome is 36 kb in length with inverted terminal repeats and intrinsic promoter sets that transcribe early (E) and late genes. They have the potential to carry large DNA inserts up to the same size as their

genome (Ali *et al.*, 1994). They are human viruses but are usually only responsible for mild upper respiratory tract infections. Adenoviruses are able to transduce a large number of different kinds of human cells, including non-dividing cells, at high efficiency, approaching 100% in some cell types.

Adenoviruses express over 50 genes during the course of their life cycle. The first generation of adenoviral vectors was made by deleting the E1 region of the genome. This addressed a safety aspect because without a transactivating E1a protein the virus cannot replicate. To create space for insertion of the gene to be transferred a region (E3), which functions to suppress the host immune response during infection, was also eliminated. Initially these first generation adenoviral vectors showed promise in treating genetic disease, for example haemophilia (Kay *et al.,* 1994). Enthusiasm for such vectors was dampened when the production of viral antigens elicited an immune response (Ali *et al.*, 1994). Consequently approaches to delete those adenoviral genes that are most immunogenic ensued, but some genes related to the capsid antigen and the transgene itself may always be associated with a degree of immunity (Gao *et al.*, 1996). So called "gutless" vectors have now been developed where all of the viral genes have been removed (Haecker *et al.*, 1996). Unfortunately deleting all of these genes is not easy because of the size and complexity of the adenoviral genome. The fact that retaining some of these genes is potentially beneficial, for example E3 functions to suppress an immune response in the host, makes it necessary to strike a balance between the 'good' and 'bad' genes for a successful adenoviral vector (Schiedner *et al.*, 1998).

Aside from the issue of the immune response is the transient nature of gene expression from adenoviral vectors, especially in rapidly dividing cells. Following infection with adenovirus the viral DNA remains episomal, rather than integrating into the host cell chromosome so future cell generations will not contain the transgene. This limits the use of adenoviral vectors in some gene therapy protocols where long-term gene expression is the desired outcome. An adenoviral vector that can be assembled with a combination of genes that minimises any immunological problems will certainly have application in short term treatment protocols and those involving repeated doses of vector.

Due to their natural tropism for the epithelial cells of the respiratory tract, adenoviral vectors have been used to transfer the cystic fibrosis transmembrane receptor (CFTR) gene into these cells as a potential treatment for cystic fibrosis. Although adenoviral mediated gene transfer of CFTR has shown promise in correcting the genetic defect in cystic fibrosis *in vitro*, clinical trials in patients have been disappointing (Crystal

et al., 1994; Wilson, 1996). Another popular use of adenoviral vectors has been in the treatment of cancer where transient expression is desirable (Alemany *et al.*, 2000).

1.1.2.2.2 ADENO-ASSOCIATED VIRAL VECTORS

Another DNA virus is adeno-associated virus (AAV). AAV is a member of the parvovirus family and requires superinfection with a helper virus such as adenovirus or herpesvirus, in order to replicate. To date, the virus is not known to be associated with any human disease and is widespread, about 4/5 of the human population have antibodies to AAV. The most commonly studied isolate, AAV-2, has a broadly permissive host cell infectivity range. The viral genome is single stranded and small, only allowing room for the insertion of a foreign gene of up to about 5 kb. Original work with AAV suggested that its use as a gene therapy vector presented two advantages over adenoviral vectors. The first was that AAV could integrate into the host cell chromosome and the second was that such integration was specific, occurring on the short arm of chromosome 19 (Walsh *et al.*, 1992). Neither of these predictions has been borne out. Despite suggestions of AAV vectors integrating specifically (Linden and Berns, 1997), it would seem that the majority of AAV based vectors, unlike wild type AAV, integrate in a random manner (Kearns *et al.*, 1996).

In 1982, Samulski *et al.* demonstrated that the genome of AAV could be cloned into a plasmid and then rescued in cells by the addition of helper virus. This initiated the development of packaging lines to generate recombinant AAV based vectors. AAV encodes only two proteins *rep* (for replication) and *cap* (for encapsidation) which are replaced with the therapeutic gene in AAV derived vectors. This places the foreign sequence between the dual 145 bp inverted terminal repeats (ITRs) with little other AAV genomic sequence, thus minimising possible recombination events which could lead to the unwanted production of replication competent virus. The ITRs themselves lack functional promoter, termination and polyadenylation signals, which must therefore be added exogenously. Deletion of *rep* and *cap* results in the vector losing its ability to integrate specifically. To generate recombinant AAV, the vector is transfected into producing cells (commonly 293) along with a plasmid expressing *rep* and *cap* followed by infection with an E1 deleted adenovirus. During harvest of the recombinant AAV, contaminating adenovirus can largely be removed by differential centrifugation. Current methods of preparing AAV vectors are cumbersome and the number of infectious particles produced is unclear.

Nevertheless recombinant AAV vectors have been shown to efficiently transduce both muscle (Fisher *et al.*, 1997) and brain (Kaplitt *et al.*, 1994) resulting in long term stable gene expression. An encouraging report describing curative levels of clotting factor IX, subsequent to AAV mediated gene transfer, have been found in a mouse model that persisted for at least nine months with no evidence of toxicity (Snyder *et al.*, 1997). More recently, cognitive function in a murine model of mucopolysaccharidosis was shown to improve following intracranial injection of recombinant adeno-associated virus (Frisella *et al.*, 2001).

1.1.2.2.3 HERPES VIRUS VECTORS

A variety of human viruses fall into the family herpesviridae including epstein-barr virus, cytomegalovirus, varicella-zoster virus and herpes simplex virus (HSV). Most efforts for vector development have been focused on HSV which, has been used for gene transfer to neurons (Fink and Glorioso, 1997). HSV has a large (150 kb) double stranded linear DNA genome allowing insertion of at least 30 kb of foreign DNA sequence. HSV can infect non-dividing cells but remains episomal. It can establish latent infection that may last the entire life span of the cell, especially in the CNS (During *et al.*, 1994). Capitalising on this inherent natural tropism of HSV, the β-glucuronidase gene has been used for the correction of mucopolysaccharidosis type VII, with results showing long-term expression in latently infected neuronal cells (Wolfe *et al.*, 1992).

Development of recombinant vectors has been challenging because there are about 70 genes in the HSV genome, some still with unknown functions. Only transient virus production systems have been developed thus far which have low transduction efficiencies (Andersen *et al.*, 1992). Cellular toxicity is also apparent and this is further complicated by the fact that much of the population has herpes virus infections (Johnson *et al.*, 1992). Therefore, upon transduction with a recombinant HSV for gene therapy it may be possible to activate the latent virus. Clearly a better understanding of HSV is required before it can be a candidate for gene therapy trials.

1.1.2.2.4 RETROVIRAL VECTORS

The fact that the greatest number of active gene therapy trials involve the use of retroviral vectors suggests that they must, at least at some point, have been highly regarded for this purpose (Anderson, 1984). What makes the retrovirus so popular as a gene transfer vector is that, by nature, it integrates into the host cell genome, thus delivering the gene of interest not only to the target cell but also to all its progeny. The first vectors to be

developed were adapted from a murine retrovirus, the Moloney murine leukemia virus (MoMLV) (Anderson, 1992; Miller, 1992a). Since then this virus has been well characterised and it has been possible to disable MoMLV to generate safe and efficient packaging cell lines and recombinant vectors (Miller *et al.*, 1993). Clearly retroviruses have a number of features that make them the most promising candidates for gene therapy of genetic disorders whereby stable genetic modification of the target cell is the desired outcome. The following section is therefore dedicated to retroviruses and their development as gene transfer vehicles.

1.2 RETROVIRUSES

1.2.1 GENERAL PROPERTIES

The retroviridae family comprise a large class of enveloped viruses that contain a linear, diploid RNA genome composed of two identical single strands. The virion particles are spherical, 80-100 nm in diameter and the RNA genome is 7-12 kb in size. Their outer lipid envelope is acquired by budding from the host cell into which viral glycoproteins are inserted. Retroviruses were first remarked upon as unknown pathogenic agents in chickens in two independent reports at the turn into the twentieth century. The first, in 1908, by a Danish physician/veterinarian group, found that a lymphoma in chickens was caused by a virus (Ellermann and Bang, 1908). This virus is now known as avian leukemia virus (ALV), the simplest in the retroviridae family. In 1911, the transmission of a sarcoma in chickens was described as being "cell-free" by Peyton Rous in New York (Rous, 1911), which now bears his name, Rous sarcoma virus (RSV). Since then over 60 retroviruses have been discovered and the search for new animal retroviruses continues even today.

Retroviruses can be broadly classified into two classes; simple and complex, based on the nature and organisation of their genomes (Coffin, 1992). Common to both classes are three main coding domains. The first, *gag*, encodes the structural or core proteins of the virus that are proteolytically processed to form the matrix, the capsid and the nucleoprotein structures, and sometimes other proteins that are designated by their molecular weight sizes. Secondly, *pol*, which directs the synthesis of two viral enzymes, reverse transcriptase (RT), has both DNA polymerase and ribonuclease (RNase H) activity, and integrase, which acts during replication of the viral genome. The *env* coding domain contains information for the envelope protein, encoding both the surface glycoprotein (SU) and transmembrane (TM) components. These two polypeptides are

synthesised as a precursor molecule that is cleaved during transport to the cell surface. An additional fourth smaller sequence, found in all retroviruses, is termed *pro*, this spans the *gag* and *pol* coding domains and encodes a further enzyme, the viral protease, which is responsible for cleaving the primary translation products of *gagpol* into mature proteins.

Retroviruses replicate *via* a double stranded, chromosomally integrated DNA intermediate (section 1.2.2, Figure 1.1) and the structure of the genome is most often described in this form. The viral encoding domains are flanked by two identical sequences termed long terminal repeats (LTRs). Contained within the LTRs are signals required for synthesis and processing of viral RNA, organised in three elements. U3 and U5 are derived from RNA sequence unique to the 3' and 5' ends of the genomic RNA molecule respectively, and R is the sequence repeated at both ends of the RNA. U3 provides the major controlling elements for transcription, including promoter and enhancer elements. The U3, R and U5 regions contained within the 3' LTR provide all the necessary *cis* acting elements for post-transcriptional processing of the 3' end of the RNA. The site of transcription initiation is at the boundary between U3 and R and the junction of R and U5 is the site of poly (A) addition. Complex retroviruses have this same genetic organisation but with extra coding domains for additional proteins.

1.2.2 REPLICATION CYCLE

An overview of the life cycle of the retrovirus is shown in Figure 1.1. The prototypical feature of the retroviral life cycle is the reverse transcription reaction. As already mentioned retroviruses contain two identical copies of single stranded RNA, yet only one DNA molecule is produced from an infectious retroviral particle. This pseudodiploid feature, whilst essential for replication, may also help contribute to the genetic diversity of retroviruses and increase the likelihood of successful synthesis of its genome, as it is known that if one of the RNAs happens to be damaged, RT can switch to the other copy (Coffin, 1979).

1.2.2.1 VIRAL ENTRY

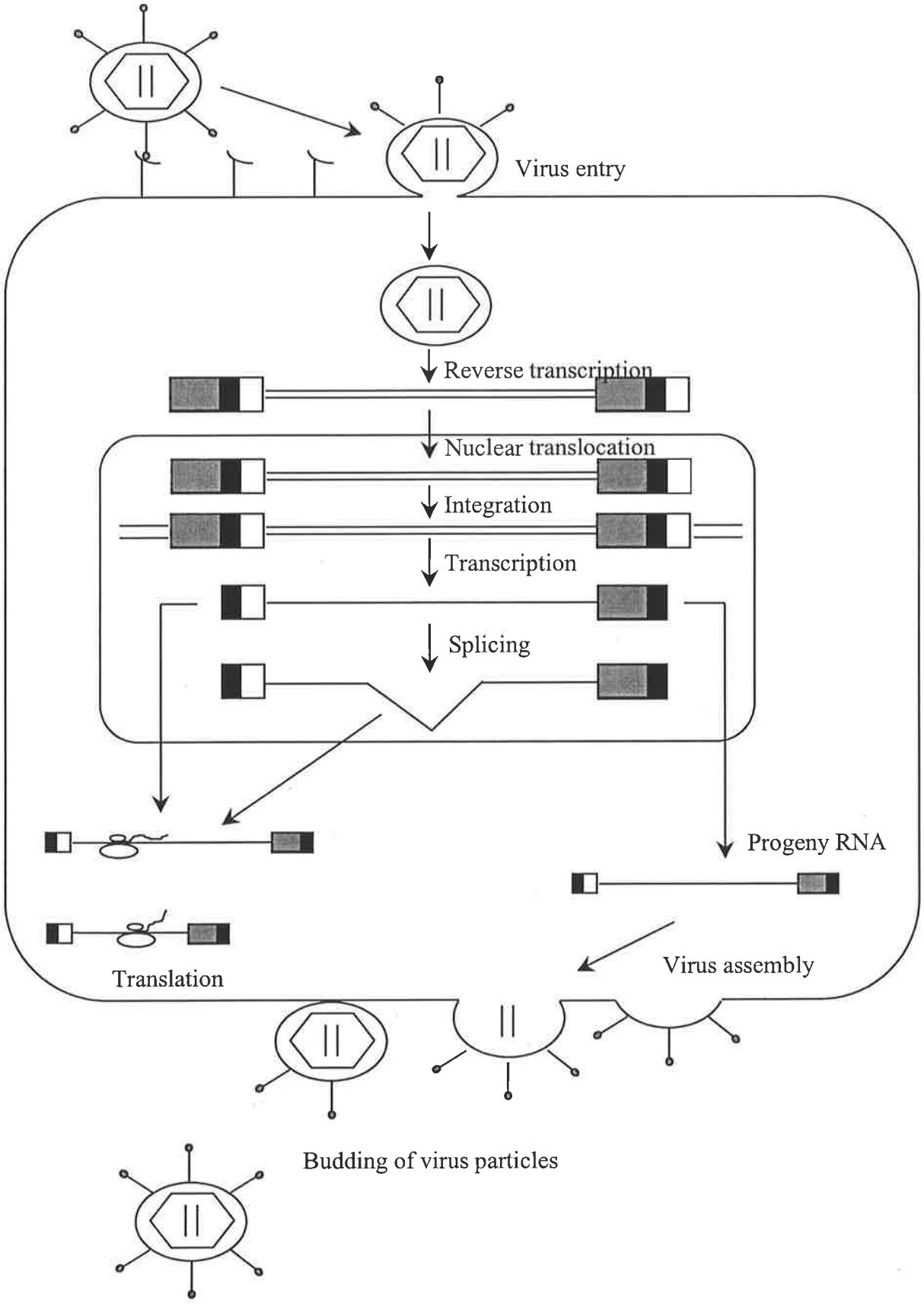
The first step in the life cycle of a retrovirus is the binding of the viral surface glycoproteins to receptors present on the target cells. As stated earlier the envelope protein is encoded by the *env* gene and is composed of a membrane spanning domain (TM) and the surface protein (SU). The latter interacts with host receptors, this is

FIGURE 1.1

Life cycle of the retrovirus.

The retrovirus attaches to a specific receptor on the surface of a host cell *via* the SU portion of the Env protein. The membranes then fuse and the core of the virus enters the cytoplasm. The genome of the virus is then reverse transcribed, transported into the nucleus and integrated into the host cell chromosome. Cellular RNA polymerase II proceeds to transcribe the integrated viral DNA producing RNA copies with the same terminal structures that appear in the parental genome. These copies are both full-length and spliced mRNAs as well as the progeny virion RNA. In the cytoplasm the viral messages are translated. Assembly of the virion RNA and proteins occurs and progeny virus is released by budding from the host cell membrane followed by maturation into infectious virus.

Virus adsorption to receptor



believed to abet fusion of the viral and host cell membranes by the TM domain. Exactly how this fusion process enables the entry of the virus particle into the host cell has not been elucidated, but it is believed to enter the cytoplasm as part of a nucleoprotein complex. The tropism (host range) of each particular retrovirus is defined by the interaction between SU and the corresponding host cell receptor. Virus particles may be considered complete but are clearly non-infectious without the envelope glycoproteins.

Once the viral RNA genome has entered the cytoplasm of the host cell the reverse transcription reaction begins. There is some evidence that DNA synthesis can be initiated within the virion prior to host cell entry (Zack *et al.*, 1992; Zhu and Cunningham, 1993), but there appears to be an activation event for the reaction upon entry of the viral core into the cytoplasm. Reverse transcription takes place within the formed nucleoprotein complex following internalisation and uncoating of the viral particle. Within this complex is the RNA genome, RT and other proteins encoded by the virus. The synthesis of retroviral DNA is governed by the DNA polymerase and RNase H of RT, but it is likely that other viral proteins augment the process.

1.2.2.2 REVERSE TRANSCRIPTION

The retroviral reverse transcription reaction is depicted in Figure 1.2. Retroviral DNA synthesis is primed by the 3' end of a partially unwound tRNA that anneals to the primer binding site (PBS) in the genomic RNA. The synthesis of this minus strand DNA continues towards the 5' end of genomic RNA, thereby generating minus strand strong stop DNA (-sssDNA). RNase H then degrades the R and U5 regions in the RNA strand of the RNA/-sssDNA duplex. Following degradation, -sssDNA undergoes what is known as the first strand transfer, to the 3' end of genomic RNA. The R region in -sssDNA then anneals to the complimentary R region in the 3' end of the RNA genome. Minus strand DNA synthesis is able to resume, accompanied by RNase H digestion of the template strand. A short oligoribonucleotide that is relatively resilient to RNase H activity, called the polypurine tract (PPT) persists. This primes plus strand DNA synthesis which is terminated when the portion of the tRNA primer that was originally annealed to the PBS is reverse transcribed, yielding a DNA fragment called plus strand strong stop DNA (+sssDNA). All retroviruses generate a plus strand primer from the PPT but some (eg HIV) also produce plus strand primers from the RNA genome. The tRNA primer is subsequently removed by RNase H digestion which leaves complimentary PBS sequence at each end of the replication intermediate exposed. Annealing of these PBS segments

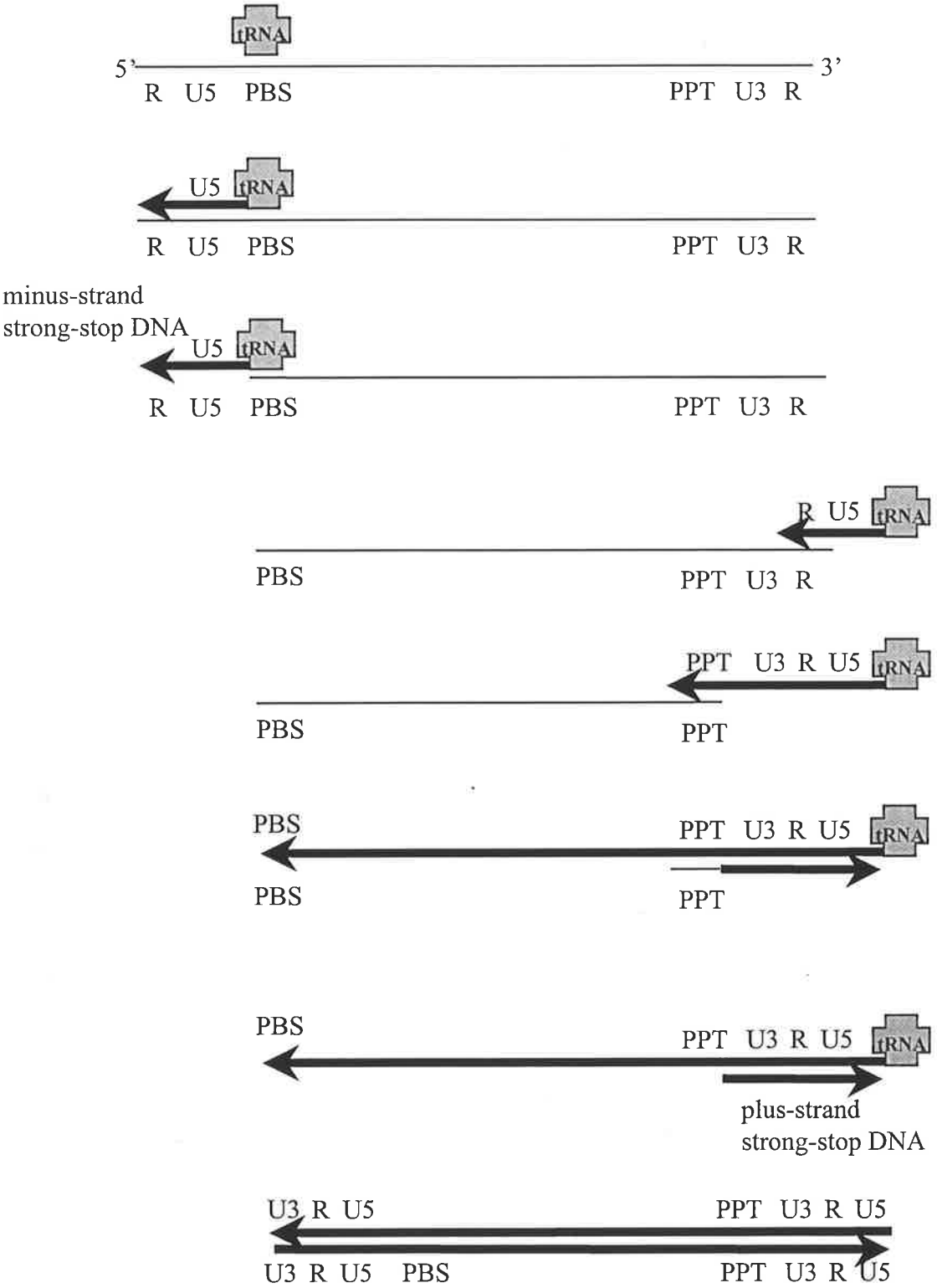


FIGURE 1.2

The retroviral reverse transcription reaction.

The thin black line represents RNA and the heavy black line denotes DNA.

See section 1.2.2.2 for a description of this process.

constitutes the second strand transfer. The synthesis of both the plus and minus strands can now be completed with each of the two DNA strands acting as template for each other. The completed double stranded DNA product is co-linear with its RNA template but contains two identical repeat sequences, the LTRs, consisting of U3, R and U5 that are not present in the viral RNA. This linear viral DNA is the precursor to the integrated provirus.

1.2.2.3 INTEGRATION

Viral DNA synthesis is initiated in the cytoplasm and may be completed before or after the nucleoprotein complex enters the nucleus, depending on the type of retrovirus. The retrovirus must also move through the cytoplasm from the periphery of the cell to the nucleus before integration can occur. The linear viral DNA remains associated with viral core proteins and probably some host cell proteins as well, and collectively are referred to as the preintegration complex. Those viral proteins necessary for nuclear entry and integration are carried into the cell by the virion.

Not long after the synthesis of viral DNA has been completed the viral enzyme cleaves both 3' ends, effectively removing two bases from either end. This provides 3'-OH groups for the attachment of the viral DNA to host DNA. This site for attachment is at a CA/TG dinucleotide pair at the internal boundaries of the viral LTRs. Integrase probably performs this function while the preintegration complex is still in the cytoplasm. The preintegration complex must now enter the nucleus, a step in the retroviral life cycle that has been the object of much investigation. Host cell division would appear to be the pivotal requirement for nuclear entry for the majority of retroviruses (Varmus *et al.*, 1977; Springett *et al.*, 1989). Integration, and therefore productive infection, by MLV is clearly dependent on the host cell undergoing mitosis (Roe *et al.*, 1993). It appears that the preintegration complex gains access to the host nuclear DNA during the breakdown of the nuclear membrane, an occurrence of mitosis. The preintegration complex remains resident in the nucleus when the nuclear membrane is reformed. In contrast to MLV, lentiviruses can transgress through the nuclear envelope obviating the requirement for mitosis (Lewis *et al.*, 1992; Lewis and Emerman, 1994). This was exemplified by the ability of HIV-1 to infect terminally differentiated macrophages *in vivo* (Bukrinsky *et al.*, 1993). This has been shown to be a direct consequence of karyophilic determinants contained in two virion proteins, the *gag* derived matrix antigen and Vpr. These proteins interact with the nuclear import machinery allowing the active transport of the preintegration complex through the nuclear membrane (von Schwedler *et al.*, 1994; Heinzinger *et al.*, 1994).

Once in the nucleus, the viral DNA complex binds, and is then joined to, host DNA by a reaction catalysed by integrase. The recessed 3'-OH moieties on the ends of the viral DNA react with the phosphodiester bonds on opposite strands of the host DNA forming an ester bond. The choice of integration sites in the host DNA is essentially random although highly bent DNA sites such as those in specific nucleosome positions are favourable (Pryciak and Varmus, 1992). The gaps that flank the junctions between the host and viral DNA are then repaired by viral RT with the possible help of other viral proteins. As to the events afterwards, nothing is yet known of this gap repair step or how the preintegration complex is finally disassembled.

1.2.2.4 SYNTHESIS OF VIRAL RNA AND PROTEINS

In the life cycle of the retrovirus, proviral DNA is the template for transcription and once integrated the provirus mimics a cellular gene. The viral genome must be able to direct the host cell machinery to express its own genes. It does this through the use of *cis* acting control elements present in the LTRs. The U3 region of the 5' LTR contains control elements that regulate transcriptional initiation by the host cell RNA polymerase II, the enzyme that mediates retroviral transcription. The 3' LTR contains *cis* acting elements that control post-transcriptional processing of the 3' end of the RNA product. In addition to control elements provided in *cis*, retroviruses utilise transcription factors provided by the host cell, which are often unique to a particular class of retrovirus. Complex retroviruses, for example HIV, encode their own transcriptional activators that further aid in controlling viral gene expression.

Viral transcription results in a full-length transcript, which will be packaged as the viral genome. Therefore, this fraction of RNA must reach the cytoplasm unspliced. The retroviral genome also contains open reading frames for viral proteins. Multiple protein products are synthesised from a single RNA species by frameshifting and processing of polyproteins by proteolysis. A portion of retroviral transcripts is spliced to generate smaller sized mRNAs. The retrovirus faces the formidable task of regulating the quantity and type of spliced and unspliced RNA.

The full-length transcript from the integrated provirus provides the polyprotein precursors for assembling the virus. These include Gag, Pro, Pol and Env, which, as described earlier, are common to all retroviruses. Gag is translated in the cytoplasm from the unspliced viral RNA and is central to the assembly process (Wills and Craven, 1991). The *pro* and *pol* genes are also derived from the unspliced viral RNA as a result of translational read-through. This translational read-through mechanism ensures that the

termination codon at the end of *gag* is bypassed allowing translation to continue into the adjacent *pro* and *pol* reading frames. In this instance Gag, Pro and Pol are not synthesised as individual polyproteins but as part of either a GagPol or a GagProPol fusion protein. These polyproteins are then directed towards the plasma membrane whereby they are processed and packaged along with two copies of viral RNA (a process nominally mediated by Gag). A spliced transcript is responsible for the synthesis of Env, which contains the SU and TM proteins. Env is synthesised and glycosylated on the membrane of the rough endoplasmic reticulum. These proteins are transported through the Golgi apparatus and then to the plasma membrane to the sites of budding. Here they are assembled onto particles containing Gag, GagProPol and the viral genome. These immature particles then bud from the cell surface. Virus production does not adversely affect cell function. The Gag and GagProPol polyproteins are cleaved by the viral protease resulting in mature virions which are available to infect other cells for another round of replication. It should be noted though, that not all processing occurs within the virion, some occurs before and some is coincidental with particle formation.

1.2.3 PRINCIPLES OF VECTOR DESIGN

Replication competent retroviral vectors have been described (Stuhlmann *et al.*, 1989; Dillon *et al.*, 1991) although their application for gene transfer is quite narrow. A more common approach is the design of replication incompetent vectors of which those derived from Moloney murine leukemia virus (MoMLV) has been the archetype. MoMLV was first described in 1960 (Moloney, 1960) and since then has become a well-studied and characterised retrovirus (Miller, 1992b; Miller *et al.*, 1993). Replication defective recombinant vectors have been successful principally because of the identification of sequences that act in *cis* and *trans* and their effective functional separation into complimentary vector and helper plasmids. This has allowed the production of recombinant retroviruses that were free from any wild type or infectious virus. Such recombinant retroviruses were replication defective but still able to transmit a therapeutic gene to a target cell.

The genes of the viral genome which are involved in the replication of MoMLV, ie *gag*, *pol* and *env*, are removed and replaced with a gene of interest. What remains are the *cis* replication elements, including a promoter and polyadenylation signal in the viral genome, a viral packaging (*psi*) signal to direct incorporation of vector RNA into virions, signals required for reverse transcription, such as a transfer RNA-binding site (PBS) and polypurine tract (PPT) for initiation of first- and second-strand DNA synthesis, and a

repeated (R) region at both ends of the viral RNA required for transfer of DNA synthesis between templates, and short partially inverted repeats located at the termini of the viral LTRs required for integration (Figure 1.3).

1.2.3.1 PACKAGING SEQUENCES

The packaging of viral genomic RNA into virus particles is coordinated by a specific *cis* acting sequence element known as *psi* (ψ). This sequence lies downstream from the splice donor site and extends into the *gag* reading frame. Initially the ψ sequence was identified as spanning 350 bp between the splice donor site and the initiating codon of *gag*. Deletion of this region produced viruses that were unable to package their own genome (Watanabe and Temin, 1983; Mann *et al.*, 1983). Insertion of this ψ sequence into other sites of the retroviral genome reinstated packaging of genomic RNA, including subgenomic RNA (Mann and Baltimore, 1985).

For MLV (Moloney leukemia virus) the optimal ψ sequence is now believed to cover 800 bp, just downstream from the splice donor and overlapping with the amino terminal sequences of the Gag protein (Bender *et al.*, 1987). This region has been shown to direct the efficient packaging of heterologous transcripts into virions (Adam and Miller, 1988). Between the splice donor site and the *gag* start codon, two hairpin structures have been unveiled (Konings *et al.*, 1992). The loops in these two hairpins have a conserved GACG motif that is able to form additional hydrogen bonds between bases in the loop, thereby contributing to the stability of the structure (Cheong *et al.*, 1990; Heus and Pardi, 1991).

1.2.3.2 PACKAGING CELL LINES

Packaging cell lines provide all the viral proteins in *trans* that are necessary for the assembly of virions. For MLV these packaging cell lines, originally derived from murine fibroblasts, supply the Gag, Pol and Env proteins but are incapable of packaging the transcripts encoding these proteins into virions. Early attempts at creating such packaging cells used a wild type proviral genome, which bore deletions of the ψ sequence. Murine fibroblasts transfected with this crippled virus produced all of the retroviral proteins but the particles were devoid of viral RNA. They were however, able to package other

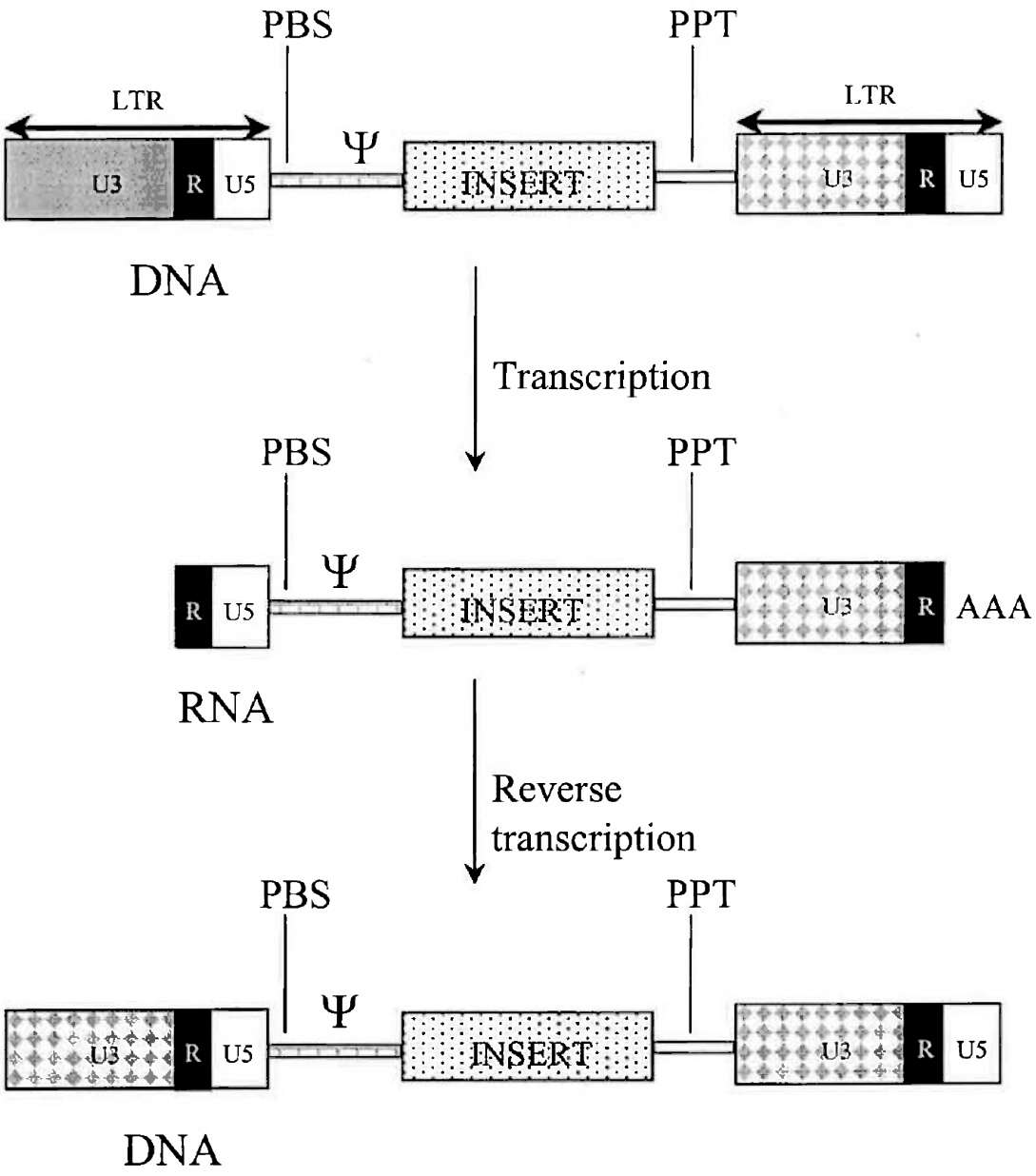


FIGURE 1.3

A retroviral vector and its replication cycle.

A retroviral vector containing an inserted gene to be transferred in place of the *gag*, *pol* and *env* genes is depicted. The 5' and 3' LTRs containing promoter and enhancer elements and the poly A site respectively, the polypurine tract (PPT), primer binding site (PBS) and packaging signal (Ψ) are shown. Forms of the vector are displayed following transcription and reverse transcription.

transcripts from recombinant retroviral vectors that contained the ψ sequence (Mann *et al.*, 1983; Cone and Mulligan, 1984; Miller *et al.*, 1985). The major problem facing these first generation packaging lines was that a single recombination event between the deleted retrovirus in the packaging cells and the retroviral vector could result in the production of wild type virus. Once formed this replication competent retrovirus can spread very efficiently (Otto *et al.*, 1994). The design of retroviral packaging cell lines has since evolved to address the undesirable spontaneous production of helper virus by minimising sequence homology between packaging and vector plasmids. Consequently packaging cell lines were constructed such that more recombination events were required making the formation of helper virus less likely. The next wave of packaging cell lines contained additional alterations in the genome of the helper virus, including deletions in the 5' LTR or replacement of the viral LTR promoter with other promoter sequences, and the substitution of the polyadenylation signal of SV40 for the 3' LTR (Sorge *et al.*, 1984; Miller and Buttimore, 1986). Two homologous recombination events are necessary to generate replication competent retrovirus in this generation of packaging cell lines. The latest development of packaging cells has been engineered with the *gagpol* and *env* coding regions separated over two expression plasmids (Markowitz *et al.*, 1988a). Additionally, cell lines free from endogenous retroviral sequences were used for the creation of the packaging cell lines (Cosset *et al.*, 1995). In this third generation of packaging cells three recombination events are required for wild type virus formation.

Such retroviral vector producer cell lines release replication incompetent recombinant retroviruses that are able to transduce target cells only once. Once internalised the retroviral vector can integrate into the target cell genome and the transgene is expressed as it if were a cellular gene. Thus, in theory at least, the goal of gene therapy has been achieved.

1.2.3.3 HOST RANGE

An important consideration in the selection and/or development of packaging cell lines is the host range of retroviral vectors that will result. As mentioned earlier, the host range is solely determined by the particular Env protein that is used in the construction of the packaging cell. The host range of retroviruses is generally divided into three classes, which were originally developed to describe murine retroviruses. The first of these comprise those viruses that can only infect murine cells, these are known as ecotropic viruses. Xenotropic viruses are unable to infect murine cells but can infect a wide range of other mammalian cell types. The last, and most versatile, class is referred to as the

amphotropic viruses. These can infect murine cells as well as a wide variety of mammalian species including human cells.

Viruses of each tropism infect cells *via* separate and distinct cellular receptors and some retroviruses can use more than one cell surface receptor for entry. For example, the 10A1 strain of MLV, from which packaging cell lines have been generated, can use either the amphotropic receptor Ram-1 or the gibbon ape leukemia virus (GALV) receptor Glv-1 for cell entry (Miller and Miller, 1994; Miller and Chen, 1996). For human gene therapy, the amphotropic recombinant vectors are typically used because they infect human cells. Pseudotyped viruses are those that have incorporated Env proteins from other viruses into their virions, thereby possessing altered tropism. The incorporation of the vesicular stomatitis virus G glycoprotein (VSV-G) as the Env protein has provided a virtually unlimited host range that has been termed pantropic (Burns *et al.*, 1993). VSV-G is supposed to mediate fusion between the viral envelope and cell membrane upon binding to phosphatidylserine, a lipid component of plasma membranes. Therefore, VSV-G pseudotyped viruses can infect a wide range of species, tissues and cell types, including those not previously susceptible to the murine amphotropic viruses, such as human hepatocytes (Yee *et al.*, 1994). If retroviral vectors can be packaged into heterologous viral envelopes with broadened host range, such as VSV-G, higher transduction levels can be achieved as a consequence of permitting viral entry *via* a wider spectrum of cell surface receptors. For example, the use of VSV-G pseudotyped MLV results in high-efficiency gene transfer into HSCs compared with virus carrying the amphotropic MLV envelope (Akkina *et al.*, 1996).

1.2.3.4 VECTOR CONFIGURATIONS

The most common application of recombinant retroviral vectors is for the production of a specific protein in transduced cells. The simplest approach is to remove the sequences encoding *gag*, *pol* and *env*, and insert a cDNA sequence encoding the gene of interest. The LTR can be used alone to control expression or insertions can be made in the enhancer and/or promoter regions of the LTR to facilitate expression of a transgene in the particular target cell. If the LTR cannot be modified to express the transgene adequately it is possible to express a single coding region from a chosen internal promoter.

Frequently it is desirable to express two genes from a retroviral vector, especially if one is a marker, which enables selection of transduced cells. Numerous retroviral vector designs are available to achieve this. It is possible to express two different promoters from alternatively spliced mRNAs transcribed from one promoter. This is modelled on the

MLV system where one protein mimics the polyprotein GagPol and is translated from the full-length viral mRNA and the other the Env protein, which is made from a spliced mRNA (Cepko *et al.*, 1984). Internal promoters have been widely used to express two genes and retroviral vectors expressing three different cDNAs, each from their own promoter have been engineered (Overell *et al.*, 1988). A promoter cDNA transcription unit can be inserted into the U3 region of the retroviral 3' LTR which results in two copies of this gene in the integrated proviral form (Hantzopoulos *et al.*, 1989). These have been termed "double copy" vectors and have also been created by placing the cDNA into the R region of the LTR (Adam *et al.*, 1995).

Multiple coding regions can be expressed from the one promoter using IRES (internal ribosome entry site) elements placed between the open reading frames (Adam *et al.*, 1991). These elements originate from the non-coding 5' ends of picornaviruses where they act to initiate the cap-independent translation of viral proteins (Jang *et al.*, 1990). IRES elements enable translation of the downstream open reading frame by promoting entry of the ribosome followed by the initiation of translation. Using a single RNA ensures coordinate expression of both proteins. Another option for the concordant expression of two proteins from a retroviral vector is to combine the function of both proteins using a chimeric protein. A chimeric protein can be created by an in frame fusion of the cDNAs encoding the proteins. For example, a chimeric combination of adenosine deaminase and a multidrug resistance gene has been generated and expressed *via* a retroviral vector (Germann *et al.*, 1990).

Self-inactivating (SIN) vectors have been created in an attempt to eliminate any effects of the LTRs on the transcription of internally placed genes. This is also of particular use in preventing the downstream transcription from the 3' LTR into the genomic DNA into which it is integrated in the host cell (Herman and Coffin, 1987). For gene therapy to be a safe mode of treatment it is imperative that there are no adverse affects such as the possible activation of an endogenous oncogene because of the integration of the retroviral vector. SIN vectors have been made by deletion of the transcriptional enhancer and/or promoter elements contained within the 3' LTR (Hawley *et al.*, 1987; Yee *et al.*, 1987). During reverse transcription the 3' LTR encompassing the deletion is used as the template for the synthesis of both LTRs in the newly formed proviral DNA. Thus after one cycle of replication the deletions will be incorporated within the 5' and 3' LTRs producing an inactive provirus.

1.2.3.5 DETECTION OF RECOMBINANT HELPER VIRUS

With most, if not all, retroviral vector systems described to date, it is possible that as a result of recombination between the helper and vector constructs, replication competent retrovirus can be produced. The frequency of such an event occurring appears to be related to the amount of homologous sequence overlap between the plasmids used to make the helper cells and the retroviral vector itself. Whilst this is the most likely scenario for the development of replication competent virus, it is conceivable that the defective helper and/or vector can recombine with endogenous sequence, or endogenous replication competent retroviruses present in the cells used for virus production could become activated, both resulting in the appearance of replication competent retrovirus production (Miller *et al.*, 1986).

Many methods have been developed in an attempt to cover every possible source of helper virus that may be generated during the production of (replication defective) recombinant retroviral particles. Probably the most sensitive method for detecting replication competent helper virus is a marker rescue assay. This involves exposing cell lines carrying a selectable retroviral vector to a large amount of the vector preparation to be tested for at least two weeks allowing enough time for any possible helper virus to spread. The medium from these cells is then tested for the ability to confer resistance to the selectable marker in naïve cells. If the cells are resistant then the selectable provirus was rescued due to the presence of replication competent virus in the initial vector preparation.

A similar technique for assaying helper virus is the S^+L^- assay (Bassin *et al.*, 1971; Peebles, 1975). This is another marker rescue assay, which is based on a defective oncogenic retrovirus. The transformed cells are prepared by infecting cells with a mixture of transforming virus (sarcomagenic) and helper virus (leukemogenic) at low multiplicity. Transformed clones that were maintained only by the defective sarcomagenic virus, are then isolated, ie sarcomagenic positive (S^+) leukomogenic negative (L^-). Subclones of these S^+L^- cells that display revertant flat morphology are then infected with potential helper virus. If present, the transforming virus spreads to nearby cells causing a transformed focus of cells in an otherwise flat layer. A modification of this assay exposes S^+L^- cells to vector preparations to be tested, with the subsequent addition of another non-transformed cell line to rescue the transforming virus (Miller *et al.*, 1985).

The major limitation of these helper assays described above is that they are only capable of end-point detection of helper virus. More general assays that would detect an earlier step in helper virus production, could include measuring the transfer of particular viral proteins, for example reverse transcriptase, and/or retroviral sequence.

1.2.3.6 LIMITATIONS

Some disadvantages of retroviral vectors, alluded to above, have precipitated investigations into different vector designs and packaging cell lines. Any sequence modifications made to the vector must be compatible with the retroviral life cycle. This includes any change to the backbone of the retrovirus, but is mainly relevant to the nature of foreign genes inserted into the retroviral vector. For example, the insertion of foreign introns in the sense orientation will be negated because the retroviral vector replicates *via* an RNA intermediate (McIvor, 1990). Another constraint is the insertion of sequences with a high AT content, which is often associated with cryptic polyadenylation signals. Such signals may obstruct transcription of the full-length viral genome. There has also been the occasional report of problematic sequences present in transgenes, for example, the cDNA for clotting factor VIII was shown to inhibit viral RNA production with a concomitant decrease in viral titre (Lynch *et al.*, 1993).

Retroviral vectors can only accommodate about 7-8 kb of foreign DNA. This perhaps is not a major drawback as most cDNAs would not exceed this size range but clearly larger genes will be precluded. The production of retroviral vectors can lead to mutations which has been estimated at 10^{-4} to 10^{-5} per base pair per replication cycle (Dougherty and Temin, 1988). However, this has largely been overcome by isolating vector preparations from packaging cells that contain a single non-rearranged integrated provirus.

Despite the numerous applications of retroviral mediated gene transfer using vectors based on MLV, there is one major drawback. The integration of MLV based vectors into the host cell genome will only occur in dividing cells. It seems that breakdown of the nuclear envelope that accompanies mitosis is necessary for the nuclear import of the retroviral preintegration complex (Miller *et al.*, 1990). The applicability of MLV vectors is therefore limited because many gene therapy targets have low or minimal mitotic activity and consequently are very poorly transduced with MLV based vectors (Roe *et al.*, 1993). This is particularly true for the pluripotent haematopoietic stem cell (PHSC) which is a popular gene transfer target because differentiated progeny, in both the bone marrow and peripheral blood, will theoretically express the new genetic material

(Barranger, 1996). Successful introduction of genes to the PHSC could thus alter the course of a wide variety of both inherited and acquired diseases. This shortcoming is actively being addressed with the construction of retroviral vectors derived from lentiviruses and other viruses such as the spumaviruses. Lentiviruses are a subfamily of retroviruses that have the unique ability to infect non-dividing cells (see section 1.3).

There is also evidence to suggest that retroviruses, in particular MLV, have a preference for active chromatin sites in order to integrate their genome (Coffin, 1990). This may help explain the phenomenon of transcriptional silencing of the transgene that is seen with MLV based vectors, especially in embryonic cells (Akgun *et al.*, 1991), presumably silencing occurs when the transduced cells return to a non-proliferating state and revise their pattern of chromatin expression. Both the retroviral LTRs and the tRNA-binding site have been implicated in this (Kempler *et al.*, 1993). It is conceivable that, again because lentiviruses are able to integrate into non-dividing cells, stable open chromatin sites could be chosen, thus selecting against the transcriptional shut-off of the transgene.

1.2.4 HUMAN FOAMY VIRUS VECTORS

The human foamy virus (HFV) is a member of the spumavirus genus of the retroviridae family. HFV was first isolated from a patient in 1971 (Achong *et al.*, 1971). Human cells are freely infected by HFV and can harbour it for extended periods (Yu *et al.*, 1996b). The name "foamy" is deserved from the foamy degeneration that these viruses induce in cell culture. Within the infected cell HFV are found in the endoplasmic reticulum and have a characteristic clubbed spikey appearance. These viruses are amphotropic, infecting cells of epithelial, fibroblastic and lymphoid origins in culture from a wide range of species, suggesting that receptors for HVF are widespread (McClure and Erlwein, 1995). HVF has not been shown to be associated with any natural human or animal infection (Weiss, 1988; Schweizer *et al.*, 1995). This stands in contrast to transgenic mouse models that have integrated HFV in their cellular chromosomes. These mice develop serious infection and die at a young age (Aguzzi, 1993). The primary effects are within the CNS suggesting a tropism for neural cells and the *Bet* protein of the virus seems to be the causative determinant (Tschopp *et al.*, 1996).

HFV has the largest genome of all retroviruses (11.67 kb) with the classic *gag*, *pol* and *env* genes flanked by LTRs (Rethwilm, 1995). Being a complex retrovirus it has accessory genes that are located between the *env* gene and the 3' LTR. These are called *bel* genes and are responsible for transcriptional activation, replication and other as yet

undefined functions. Unlike all other retroviruses a *gag-pro-pol* precursor cannot be seen in HFV infected cells, rather it seems that the *pro* and *pol* reading frames are expressed independently of the *gag* coding domain (Yu *et al.*, 1996a).

The broad host range, and apathogenic nature of HFV, has stimulated interest in the development of HFV as gene transfer vectors. Replication competent HFV vectors have been constructed by fusion proteins with the viral *bet* gene, or as independent translation products using inserted IRES (internal ribosome entry site) sequences (Schmidt and Rethwilm, 1995). Replication defective HFV vectors have been made by replacing either the *env* gene or the *bel1* and *bel2* genes with a reporter gene and providing the other required viral proteins by transfecting other plasmids or using helper virus (Russell and Miller, 1996; Bieniasz *et al.*, 1997). There was one suggestion that HFV may be able to infect non-dividing cells based on a calculation of transduction efficiency on dividing and stationary phase fibroblasts (Russell and Miller, 1996). This has been disputed, and HFV has subsequently been shown, in one particular cell line, to behave more like MLV and infect dividing cells only (Bieniasz *et al.*, 1995; Bienasz *et al.*, 1997). Packaging cell lines have not yet been created for HFV vectors and its development as a gene transfer vehicle is primitive in comparison with MLV.

1.3 LENTIVIRUSES

The lentiviruses comprise a genus within the retroviridae family that tend to be associated with slow progressive diseases affecting the immune system (Desrosiers and Letvin, 1987). Lentiviruses include primate viruses such as human immunodeficiency viruses HIV-1 and HIV-2 and simian immunodeficiency viruses (SIV), and non-primate viruses such as feline immunodeficiency virus (FIV), equine infectious anaemia virus (EIAV), caprine arthritis encephalitis virus (CAEV), bovine immunodeficiency virus (BIV) and maedi-visna virus (MVV). Lentiviruses can be distinguished from other retroviruses by morphology, with the mature virion having a characteristic cylindrical or cone shaped nucleoid. Furthermore, apart from the Gag, Pol and Env proteins expressed by all retroviruses, lentiviruses encode an array of accessory and regulatory genes that are involved in regulation of viral gene expression, assembly of viral particles and structural and functional alterations in the infected cell.

There are two subtypes of HIV, HIV-1 and HIV-2 that are both associated with progressive immunological deterioration. HIV-1 is the most prominent and is found throughout the world whereas HIV-2 is more prevalent in west Africa with only a few cases reported in other parts of the world. Epidemiological studies on the two isolates

suggest that the incubation period for disease development by HIV-2 is longer than that for its counterpart and that HIV-2 is not as easily transmitted perinatally as HIV-1 (Pepin *et al.*, 1991; Markovitz, 1993). HIV-1 is considered the prototype in the majority of studies on HIV and it is regarded as the most common cause of AIDS (acquired immunodeficiency syndrome).

1.3.1 HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1)

1.3.1.1 GENOME STRUCTURE

HIV-1 is classified as a complex retrovirus as its genome encodes not only Gag, Pol and Env, but also the accessory proteins Nef, Vif, Vpr, Vpu and two regulatory proteins Rev and Tat (Cullen, 1991). The life cycle of HIV-1 also involves multiple levels of regulation that are not apparent in the simpler class of retroviruses. The HIV-1 provirus is about 9.5 kb and is flanked by LTRs. Like all retroviruses the LTRs contain a constellation of *cis* acting elements essential for viral integration and transcription. By definition, RNA synthesis is initiated within the 5' LTR at the junction of the U3 and R regions. The 3' LTR directs the addition of poly A tails to viral RNA at the boundary between R and U5. The LTR can be divided into individual regions that control gene expression, many of which are common to the promoters of eukaryotic genes. These include the negative regulatory element (NRE), the enhancer, the basal promoter elements, the core promoter, the *tat* activation response element (TAR) and the inducer of short transcripts (IST). TAR and IST are both found downstream of the RNA start and appear to be unique to the lentiviruses. The IST is a positive downstream enhancer that stimulates the production of abortive transcripts. The TAR element is a positive enhancer that stimulates the synthesis of productive transcripts, however it is only functional as an RNA element. Between the LTRs lie the common retroviral genes *gag*, *pol* and *env* as well as the afore-mentioned accessory and regulatory genes (Figure 1.4).

The HIV-1 genome is transcribed to produce three classes of mRNA, encompassing full-length (~9 kb), partially spliced (4-5 kb) and multiply spliced (~2 kb) species. Several splice acceptor and splice donor sites have been found that enable HIV-1 to produce more than 30 differentially spliced messages (Schwartz *et al.*, 1990). The full-length viral transcript is used to produce the Gag and GagPol fusion proteins. Every other expressed open reading frame has a splice acceptor upstream of the initiating AUG giving

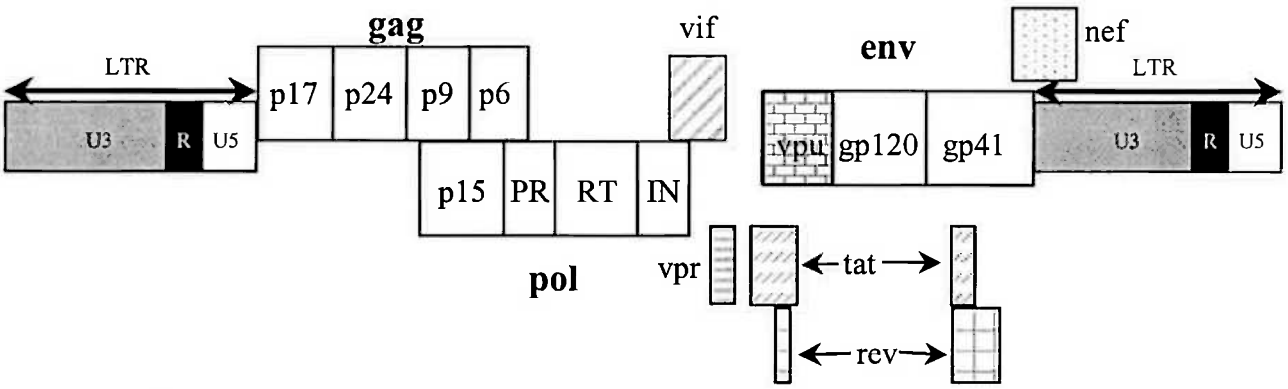


FIGURE 1.4

Genetic organisation of the HIV-1 genome.

Schematic diagram of the genome of HIV-1. The 5' and 3' LTRs are shown flanking the gene products. The three major genes of the virus *gag*, *pol* and *env*, common to all retroviruses, are depicted. The *gag* and *gapol* gene encodes the matrix (p17), the capsid (p24), the nucleocapsid (p15) and the catalytic proteins, protease (PR), reverse transcriptase (RT) and integrase (IN). The four accessory proteins, Vif, Vpr, Vpu and Nef as well as the two domains for each of the regulatory proteins Tat and Rev are also indicated. For the known functions of these gene products see section 1.3.1.

at least one unique message, the only exception is *env*. Different mRNA species predominate at various times of the HIV-1 replication cycle. Early in infection the multiply spliced ~2 kb mRNA molecules are present and high levels of Nef, Tat and Rev are produced. Later in infection, there is an increase in partially spliced and unspliced mRNA species with a concomitant decrease in the multiply spliced mRNAs. Thus, towards the end of the replication cycle the full-length RNA predominate and the viral structural and accessory proteins are synthesised (Kim *et al.*, 1989a).

1.3.1.2 PROTEINS DERIVED FROM GAG AND GAGPOL

The 1536 nucleotide long *gag* open reading frame is translated from full-length RNA into a polyprotein precursor of 55 kDa referred to as Pr55gag. The viral protease, p10 then cleaves Pr55gag to yield several smaller polypeptides. This is believed to occur at a rather late stage of maturation, either in the budding virion or inside the released viral particle (Kaplan and Swanstrom, 1991). The final protein products are the matrix protein (MA), the capsid protein (CA) and the nucleocapsid (NC). The MA (p17) is important in the early stages of viral assembly as it mediates aggregation of the polyprotein precursor under the plasma membrane. p17 may also be responsible for the recruitment of viral Env proteins to the surface of the host cell (Dorfman *et al.*, 1994). In infected cells, p17 is associated with the preintegration complex and has been shown to facilitate the transport of this complex into the nucleus of non-dividing host cells (von Schwedler *et al.*, 1994). The capsid protein (CA), p24, is a major core protein and forms a protein shell around the nucleocapsid (Gelderblom *et al.*, 1987). The interaction between p24 domains of the Gag and GagPol precursor molecules plays an essential role during viral assembly (Reicin *et al.*, 1995). The nucleocapsid (NC), p15, is a precursor polypeptide that is further processed to four peptides p1, p2, p6 and p7. The NC precursor contains zinc finger motifs that are recognised by specific nucleic acid binding proteins (Sheng and Erickson-Vitanen, 1994). The only form of nucleocapsid protein in detectable amounts in the mature virion is p7. Many functions of p7 have been identified including stimulation of reverse transcription, assistance in packaging genomic viral RNA and the promotion of viral RNA dimerisation (You and McHenry, 1994). The p6 protein has a role in the final steps of virus budding and in incorporation of the Vpr protein into the virion (Göttlinger *et al.*, 1991; Kondo *et al.*, 1995). Little is known about the function of the p1 and p2 proteins.

The *pol* open reading frame is 3045 nucleotides long and is translated as part of the precursor GagPol fusion protein referred to as Pr160, also from the full-length viral RNA.

Pr160 is produced by a translational frameshift mechanism as ribosomes read full-length genomic HIV-1 transcripts, bypassing the termination codon that defines the 3' terminus of *gag*. The efficiency of the frameshift defines the ratio of *gag* to *pol* gene products present in mature virions which is 20:1 (Jacks *et al.*, 1988). The *pol* precursor is cleaved to yield three catalytic proteins, protease, reverse transcriptase and integrase. The protease (PR), p10, is responsible for the proteolytic processing of the Gag and GagPol precursor molecules. Reverse transcriptase (RT), p66/51, a heterodimer, catalyses the reverse transcription reaction of the genomic viral RNA into double stranded DNA (described in section 1.2.2.2). While all reverse transcriptase enzymes are notorious for errors due to the lack of a proofreading activity, the HIV-1 RT appears to be significantly more error prone than other retroviral RTs. This may contribute to the high mutation rate of HIV-1 *in vivo* (Preston *et al.*, 1988; Roberts *et al.*, 1988). The integrase (IN), p32, mediates the integration of the viral genome into the host cell chromosome (described in section 1.2.2.3).

1.3.1.3 THE ENV ENCODED PROTEINS

The envelope glycoproteins are firstly synthesised as an 88 kDa precursor. In the endoplasmic reticulum high mannose N-linked carbohydrate chains are added and the molecule is folded into its appropriate tertiary structure (Fennie and Lasky, 1989). The carbohydrate chains are terminally modified in the Golgi apparatus. The signal peptide of about 30 amino acids is then cleaved from the amino terminus of the precursor molecule, gp160. Subsequent proteolytic processing produces the amino-terminal, gp120 (SU domain) and the carboxy-terminal, gp41 (TM domain) (Haseltine, 1991). The gp120 surface moiety plays a role in interacting with the CD4 receptor and possible cofactors on host cells. The transmembrane protein, gp41, is responsible for anchoring gp120 through non-covalent interactions thereby mediating the fusion process between viral particles and susceptible host cells (Lifson *et al.*, 1986; Alkhatib *et al.*, 1996; Deng *et al.*, 1996).

1.3.1.4 THE REGULATORY PROTEINS

The HIV-1 genome contains two regulatory genes, *tat* and *rev*, that encode transactivator proteins that are essential for virus replication. Along with Nef, Rev and Tat are the first proteins to be synthesised following integration of viral mRNA. Once synthesised, Rev and Tat augment the production of viral mRNAs. The intracellular concentration of Rev plays a role in the switch from the early phase of replication, marked by the abundance of the short transcripts, to the late phase of replication, characterised by

the presence of partially spliced and unspliced transcripts (Pomerantz *et al.*, 1992). In chronically infected cells, which show low levels of viral replication, an increase in the proportion of unspliced RNA appears to be associated with the induction of viral replication. This suggests that Rev is also, at least in part, responsible for viral latency (Michael *et al.*, 1991). There is some evidence that limiting concentrations of Tat also affect viral latency (Adams *et al.*, 1994), so both of the HIV-1 regulatory proteins are likely to orchestrate the conversion from latency to productive viral infection.

Rev is regarded as the regulator of viral protein expression. It is a 19 kDa protein of 116 amino acids that is found primarily in the nucleolus, but shuttles between the nucleus and cytoplasm (Felber *et al.*, 1989). Rev is expressed from two coding exons contained in a variety of multiply spliced mRNAs (Sadaie *et al.*, 1988). Two distinct domains have been identified in Rev. The first is a stretch of basic amino acids, part of which confers nuclear localisation (Zapp *et al.*, 1991). Removal of the amino-terminal portion of this basic region causes Rev to be maintained in the nucleus and excluded from the nucleolus. This results in an inactive Rev protein indicating that nucleolar localisation is crucial for its function (Venkatesh *et al.*, 1990; Cochrane *et al.*, 1990). Furthermore, this basic region is responsible for the interaction of Rev with its target sequence, the Rev response element (RRE). The RRE is a complex stem loop structure that is contained within the *env* coding sequence in HIV-1 (Heaphy *et al.*, 1990). The RRE is present in all Rev responsive viral mRNAs but is excluded, by splicing, from multiply spliced viral mRNAs. Consequently in the absence of Rev only multiply spliced mRNAs encoding the regulatory proteins accumulate in the cytoplasm of infected cells.

A leucine rich sequence constitutes the second important domain of Rev. This region is thought to be an activating domain, responsible for inducing nuclear export of RRE containing transcripts as a consequence of the binding of Rev to RRE (Ptashne, 1988). Following the binding of Rev with RRE, Rev interacts with the nuclear pore proteins to mediate the nuclear export of unspliced and singly spliced mRNAs that encode the viral structural proteins (Malim *et al.*, 1989). Thus the pivotal role of Rev is to permit the synthesis of new viral particles by facilitating the export of the aforesaid mRNAs.

The viral protein Tat is a transcriptional transactivator that is concentrated in the nucleus and nucleolus of infected cells. Tat binds to a stable stem loop structure located at the 5' end of all viral mRNAs, known as TAR (trans-activation responsive) (Berkhout *et al.*, 1989). The *tat* gene is encoded by two exons, one preceding the *env* gene coding for 72 amino acids and the other within the *env* gene encoding 29 amino acids (Kuppuswamy *et al.*, 1989). A smaller form of Tat comprising 86 amino acids exists for a few HIV-1

strains such as HXB2, LAI and pNL4-3. It has been shown that a single nucleotide change in the stop codon in this shorter Tat molecule converts the open reading frame to the 101 amino acid sequence. It is thus postulated that the 86 amino acid form of Tat is the result of a premature termination, possibly as a consequence of cell culture passage (Neuveut and Jeang, 1996).

Tat can be broadly divided into five physical domains, and of those, domain 4 remains the best-characterised (Jeang *et al.*, 1999). Domain 4 spans amino acids 49-72 and contains a basic motif. This motif is responsible for the binding of Tat to TAR (Weeks and Crothers, 1991; Chang and Jeang, 1992), is important for the nuclear localisation of Tat (Ruben *et al.*, 1989) and for the cellular uptake of the protein (Chang *et al.*, 1995). Tat is a most potent activator of gene expression, enhancing LTR directed gene expression several hundred-fold. Tat is considered essential for viral replication because mutations in the *tat* gene eliminate viral production (Fisher *et al.*, 1986). The regulation of viral transcription by Tat is at the level of elongation. When Tat binds to TAR, the stability of the RNA polymerase is increased allowing more efficient synthesis of full-length transcripts and an enhancement in the frequency of RNA initiation (Haseltine, 1991).

1.3.1.5 THE ACCESSORY PROTEINS

The accessory, secondary or auxiliary proteins of HIV-1 are those encoded by the retroviral genome in addition to the replicative genes *gag*, *pol* and *env* and the regulatory genes *tat* and *rev*. These proteins are often considered non-essential for virus replication in cell culture (Mustafa and Robinson, 1993; Balliet *et al.*, 1994). In the majority of instances the accessory proteins have only mild effects on viral replication. However, there are notable examples where expression of the accessory proteins is required to sustain the life cycle of the virus suggesting a critical role in either interaction with the host cell or another gene involved in replication (for review, see Trono, 1995). From what is known about the accessory and regulatory proteins it does seem that the interplay between them, as well as with cellular factors, affects the pattern of viral gene expression in different cell types in the infected patient.

1.3.1.5.1 VIF

The HIV-1 Vif (virion infectivity factor) protein is made from a singly spliced mRNA which accumulates late in infection (Garret *et al.*, 1991). The protein is 23 kDa and is encoded downstream from the *pol* gene (Lee *et al.*, 1986). In the infected cell Vif is

associated with membranes through interactions with carboxy-terminal basic regions (Goncalves *et al.*, 1994 and 1995). The protein has also been found to interact with the cytoskeleton (Karczewski and Strebel, 1996). The role of Vif seems to be during the assembly of infectious virus particles (Sakai *et al.*, 1993). Vif has been shown to facilitate the incorporation of the Env protein into virions and to promote viral DNA synthesis in new rounds of infection (Borman *et al.*, 1995; von Schwedler *et al.*, 1993). Only traces of Vif protein were originally thought to be present in virions, therefore it was suggested that the effects of the protein on viral assembly were indirect (Trono, 1995). However, Vif has subsequently been detected in viral particles at levels similar to those of the *pol* gene products, and Vif can also be incorporated into MLV virions (Camaur and Trono, 1996). From all the evidence presented so far, how Vif exerts its effects to enhance infectivity of the virions still remains rather elusive.

Vif appears to be necessary for the production of infectious virions only in some cell types, which further adds to the enigma of the protein. Vif is not essential for the replication of HIV-1 in permissive cell lines such as HeLa-CD4+ (Gabuzda *et al.*, 1992). For cells that are natural targets of HIV-1 infection, such as CD4+ T lymphocytes and macrophages, Vif is required for infectious virus production (von Schwedler *et al.*, 1993; Fouchier *et al.*, 1996). An explanation for this has been put forward suggesting that non-permissive cells contain an endogenous inhibitor of HIV-1 viral production that is overcome by the Vif protein (Madani and Kabat, 1998).

1.3.1.5.2 VPR

Vpr (viral protein R) is a short basic protein associated with the virion and thus could be considered a structural component of the virion (Cohen *et al.*, 1990a). Vpr was the first retroviral protein identified that was neither part of the capsid nor the capsid precursor but rather was included in the virion itself. The protein is packaged into virions through an interaction with the carboxy-terminal of *gag* and is present at concentrations analogous to those of the *gag* proteins. A repeat sequence within p6 is responsible for the packaging of Vpr and transfer of this motif to MLV confers the ability for Vpr incorporation into the virion (Lu *et al.*, 1995).

In infected cells Vpr accumulates in the nucleus, a property driven by the amino-terminal domain of the protein (Lu *et al.*, 1993; Yao *et al.*, 1995). The first function of Vpr to be documented was as a weak transcriptional transactivator of the viral LTR. Vpr was shown to activate HIV-1 LTR specific gene expression 2-3-fold suggesting that the protein may enhance gene expression during early phases of future viral infections (Cohen

et al., 1990b). This activation is now thought to be *via* the glucocorticoid-receptor complex, which may interact with a potential binding site in the HIV-1 promoter (Refaeli *et al.*, 1995). Since then, and partly because of its nuclear localisation, a number of functions have been assigned to Vpr. Vpr assists in the transport of the viral preintegration complex from the cytoplasm into the nucleus of non-dividing cells (Heinzinger *et al.*, 1994). This is not seen as a vital role because the mechanism is redundant, as MA and integrase also facilitate nuclear import of the preintegration complex. The carboxy-domain of Vpr appears capable of disrupting the cell cycle, arresting cells at the G2/M phase (Jowett *et al.*, 1995). The role of this cell cycle arrest is unclear but it may influence the activity of transcription factors. Little is really known about the implications of Vpr in viral infection and replication.

1.3.1.5.3 VPU

Vpu (viral protein U) is a late gene product expressed from the same singly spliced mRNA as the Env glycoprotein (Strebel *et al.*, 1988). Vpu is a small integral protein of 81 amino acids. It is phosphorylated and is associated with the cytoplasmic membrane of infected cells (Strebel *et al.*, 1989). At least two functions of Vpu have been identified which appear to be regulated by the phosphorylation state of the protein. Vpu enhances the release of viral particles from infected cells, a function not dependent on its phosphorylation (Schubert and Strebel, 1994). The mechanism by which Vpu enhances virion release remains obscure, especially considering related viruses such as HIV-2 are able to bud effectively without an equivalent protein.

The second function of Vpu is the down modulation of CD4, the HIV-1 receptor on target cells. In an infected cell it is believed that the co-expression of CD4 and the gp160 precursor results in a complex of these proteins that is retained within the endoplasmic reticulum. Vpu is responsible for the dissociation of this complex, which leads to its subsequent degradation (Willey *et al.*, 1992a and b). Phosphorylation of Vpu is mandatory for CD4 degradation. Degradation of CD4 allows transport of Env to the surface of the cell and its incorporation into virions. A third function has been attributed to Vpu, the down regulation of the major histocompatibility complex (MHC) class 1 molecules on the surface of infected cells (Kerkau *et al.*, 1997). This prevents recognition of HIV-1 infected cells by cytotoxic T lymphocytes.

1.3.1.5.4 NEF

Nef (negative factor) is a 27 kDa, myristylated phosphoprotein associated with cytoplasmic membrane structures in infected cells (Hammes *et al.*, 1989). The association of Nef with the viral particle seems to be non-specific, but it is known that myristylation is required and some of Nef is cleaved by the viral protease (Bukovsky *et al.*, 1997). This protein accumulates even earlier than do Tat and Rev (Kim *et al.*, 1989b). Many functions have been assigned to Nef but the mechanisms of action of the protein remain unknown. Named to imply a negative action, Nef was originally thought to suppress transcription from the viral LTR (Ahmad and Venkatesan, 1988; Niederman *et al.*, 1989). This property of Nef is no longer supported and it now appears that its action may be quite the contrary.

Nef appears to be essential for viral infectivity *in vivo* but not *in vitro* (Kestler *et al.*, 1991). Nef has a role in the assembly and the release of viral particles *via* a number of activities within the infected cell. Firstly, Nef reduces the interactions between intracellular CD4 and the Env protein by inducing the internalisation and subsequent degradation of CD4 (Piguet *et al.*, 1999). Secondly, Nef like Vpu, can also downregulate the cell surface expression of MHC class I molecules thereby protecting infected cells from the destructive effects of cytotoxic T cells (Collins *et al.*, 1998). Furthermore, independently of the effects on CD4, Nef has been shown to enhance the infectivity of viral particles (Miller *et al.*, 1995; Goldsmith *et al.*, 1995). This enhancement of activity occurs at the level of proviral DNA synthesis, early in the replication cycle (Schwartz *et al.*, 1995). It appears that like Vif, Nef affects the capability of the viral particle to carry out DNA synthesis in a subsequent round of infection (Chowers *et al.*, 1995; Aiken and Trono, 1995).

1.3.2 HIV-1 BASED GENE TRANSFER

1.3.2.1 WHY A PATHOGENIC VIRUS FOR A THERAPY?

The principle reason for HIV-1 based vector development for gene therapy, is the ability of this virus, as well as other lentiviruses, to infect non-dividing cells, in combination with all the other advantages of retroviral vectors in general (Lewis *et al.*, 1992; Lewis and Emerman, 1994). Vectors derived from other lentivirus species have also been developed (section 1.3.3) but the lack of understanding of the biology of these viruses has so far impinged on the construction of efficient and safe gene transfer vehicles. In addition, it is unclear whether the natural tropism of other lentivirus species is of much, if any, relevence in the context of highly reconstructed vector systems. The applicability of

lentiviral vectors is broad because many relevant targets for gene therapy are non-dividing cells such as HSCs, neurons, hepatocytes and myocytes. Lentiviral vectors also retain the property of retroviruses - chromosomal integration, so the stability of the transgene is potentially indefinite.

Lentiviral vectors offer opportunity for the treatment of a wide variety of disorders including genetic and metabolic deficiencies, viral infection and cancer. In fact, many HIV-1 based vectors have already been developed for such conditions as cystic fibrosis (Goldman *et al.*, 1997), retinitis pigmentosa (Miyoshi *et al.*, 1997), and haematopoietic disorders where bone marrow cells, progenitors for red blood cells, lymphocytes and macrophages have been the target cells (Akkina *et al.*, 1996; Corbeau *et al.*, 1998; Sutton *et al.*, 1998).

1.3.2.2 EARLY HIV-1 VECTORS

The first HIV-1 vector to be described carried a chloramphenicol acetyltransferase gene and was replication competent (Terwilliger *et al.*, 1989). Such vectors are not considered safe for gene therapy applications but have some use in studying the biology of HIV-1 and possible disease treatment. Further HIV-1 based vectors were constructed specifically for looking at the infectivity of HIV-1 (Page *et al.*, 1990; Landau *et al.*, 1991). They were made by co-transfection of two plasmids, the transfer construct and an envelope plasmid. The transfer construct was simply an envelope defective HIV-1 genome with a guanine phosphoribosyltransferase (*gpt*) gene in place of *env*. The envelope plasmid either expressed HIV-1 Env or a heterologous envelope gene such as the amphotropic and ecotropic MoMLV, and human T cell leukemia virus type 1 (HTLV-1) glycoprotein, giving the viral vectors expanded host cell tropism. The viral titres obtained were in the order of 2×10^5 IU/ml.

Subsequent HIV-1 based vectors were constructed from the same envelope defective HIV-1 genome but the reporter gene was inserted in place of the *nef* open reading frame (Camerini *et al.*, 1994; Planelles *et al.*, 1995). Again the construct was co-transfected with a plasmid expressing various envelope glycoproteins. These HIV-1 vectors were used for investigating the host range and effects of chromosomal integration of the virus.

The first attempt at disabling the transfer vectors was by splitting the HIV-1 genome into two plasmids and providing *env* in *trans* (Buchschacher and Panganiban, 1992). The transfer plasmid contained the reporter gene and sequences believed to be required for replication and gene expression. The packaging construct expressed all the

viral proteins except Env. Env would be provided in *trans* for the production of infectious viral particles. Viral titres were 10^2 IU/ml, too low to be considered useful for gene therapy applications.

1.3.2.3 FIRST GENERATION HIV-1 VECTORS

Knowledge regarding the *cis* and *trans* functions of HIV-1 was applied to generate the first HIV-1 vectors that were intended for use in gene therapy (Parolin *et al.*, 1994). The viral genes required in *trans* were segregated onto two different plasmids. The *gag*, *pol*, *tat* and *vif* genes were expressed on one plasmid in which the 5' LTR was replaced with the immediate early promoter of human cytomegalovirus (CMV). The second plasmid contained *env* and *rev*, which were expressed from the 5' LTR. In both plasmids the packaging signal (a region between the 5' splice donor site and the initiation codon of *gag*) had been deleted and the 3' LTRs had been replaced with the polyadenlyation sequence from SV40. A series of vector constructs was made containing sequence elements believed to be required in *cis*, as well as sequences derived from the *gag*, *pol* and *env* genes that may or may not be required for vector function. Viral titres were low, in the range of $10^{0.7}$ to $10^{2.2}$, but the conceptual basis for the development of future vectors had been demonstrated.

The first generation of high titre replication defective HIV-1 gene transfer vehicles was composed of a transfer, packaging and envelope construct (Naldini *et al.*, 1996a). The transfer vector contained LTRs, and contiguous *gag* sequence to base 350, the RRE along with some *env* sequence, an internal promoter with the transgene and the HIV-1 LTRs. The packaging construct contained virtually the entire HIV-1 genome under heterologous transcriptional control. It lacked the packaging region and adjacent sequences in the 5' untranslated region, and was defective in both the *env* and *vpr* genes. The envelope glycoprotein from the vesicular stomatitis virus (VSV-G) was supplied in a third plasmid (Burns *et al.*, 1993). Titres of 4×10^5 were achieved and the vectors successfully transduced growth arrested HeLa cells and rat 208F fibroblasts. A further deletion of 1.4 kb of *env* sequence was made in the packaging construct with no loss of titre (Naldini *et al.*, 1996b). The viral vectors were able to mediate stable *in vivo* transfer into adult rat neurons.

The *vpr* and *vif* genes were the next to be removed from the vector system without adversely affecting titre (Kafri *et al.*, 1997). Long term expression of the transgene was observed at more than 5 months in the liver, and 8 weeks in muscle tissue, of mice.

1.3.2.4 SECOND, THIRD.....? GENERATION HIV-1 VECTORS

The next step in the goal of generating safe and effective HIV-1 vectors came with the evaluation of the requirement for the accessory genes of HIV-1. From what is known of the functions of the accessory proteins in the retroviral life cycle (section 1.3.1.5), removal of these sequences would significantly improve vector safety. To this end, packaging constructs were created with deletions in all of the accessory genes (Zufferey *et al.*, 1997; Uchida *et al.*, 1998; Mochizuki *et al.*, 1998). High titre virus particles in the order of 10^7 IU/ml were achieved with HIV-1 vector systems deleted in all four of the accessory genes (*nef*, *vif*, *vpr* and *vpu*). These multiply attenuated vectors were shown to transduce dividing and growth arrested 293T cells, human HSCs, cardiac myocytes and post-mitotic rat cerebellar neurons. *In vivo* transduction of adult rat neurons was also demonstrated. It was thereby concluded that the HIV-1 accessory proteins are dispensable for gene transduction in non-dividing cells, and an effective packaging construct needs only five of the nine HIV-1 genes. However, one positive effect of the accessory proteins was noted, a 50 % reduction in transduction of macrophages in the absence of *vpr* (Zufferey *et al.*, 1997). Though not essential, Vpr may augment gene transfer in particular cell types. Moreover, because Vpr as well as Nef are encapsidated in the viral particles, further gains in terms of a decrease in cell toxicity may be accomplished by their deletion.

A third generation of HIV-1 vector systems has been attempted by removing the dependence on Tat. In terms of the vector, *tat* independence was achieved by replacing the U3 region in the 5' LTR with the CMV immediate early promoter, thereby creating a Tat independent LTR. In one report this resulted in a three-fold decrease in viral titre with the deletion of *tat* in the packaging plasmid (Kim *et al.*, 1998). Another report claims a less than two-fold reduction in titre (Dull *et al.*, 1998). In this latter report the defect in Tat function was created by the introduction of two mutations, mutation of the initiation codon and a premature stop at codon 46. The requirement for Tat was overcome by replacing the U3 region of the 5' LTR with the RSV or CMV early promoter. Furthermore, an improved safety feature was added by deleting *rev* from the packaging construct and providing it on a separate plasmid. Consequently, co-transfection of the Tat negative vector, the VSV-G envelope plasmid, the *gagpol* plasmid and a plasmid expressing Rev, produced high titre viral particles comparable to earlier less defective systems.

In other attempts to improve the bio-safety of an HIV-1 based system, Rev has been deleted from the packaging system leaving only GagPol and Tat (Kim *et al.*, 1998; Gasmi *et al.*, 1999). Rev, as described in section 1.3.1.4 is an essential player in the

replication of HIV-1, was replaced with constitutive transport elements (CTE) whose presence stabilises the HIV-1 transcripts (Zolotukhin *et al.*, 1994). Replacement of RRE in the packaging construct with the CTE from the Mason-Pfizer monkey virus allowed Rev/RRE independent expression of *gag*. The titre on 293T cells however was ten-fold lower than that of the Rev/RRE vector indicating that the CTE element could only partially substitute for the function of Rev/RRE.

As described in section 1.2.3.4 for MLV based vectors, self-inactivating (SIN) vectors have also been developed for HIV-1. One SIN HIV-1 based vector was created by deletion of 133 bp in the 3' LTR encompassing the TATA box, Sp1 and NF-kB sites (Miyoshi *et al.*, 1998). A second SIN vector was made with a deletion (400 bp) of almost the entire U3 region in the 3' LTR and no significant decrease in viral titre was observed (Zufferey *et al.*, 1998). As opposed to the MLV based SIN vectors, SIN vectors for HIV-1 lack any LTR driven transcription activity. Furthermore, inactivation of the 5' LTR prevents any promoter interference with the downstream heterologous promoter of the transgene.

Conditionally replicating vectors based on HIV-1 have been developed aimed at inhibiting HIV-1 infection. A vector was constructed with the reporter gene upstream of the major 5' splice donor such that it was expressed directly from the LTR under the control of *tat* (Buchschacher and Panganiban, 1992). This vector was able to propagate in HeLa T4 cells expressing *tat* but not in native HeLa cells, opening up the possibility to limit expression of the marker gene and viral replication in HIV-1 infected cells. Another conditionally replicating vector was capable of inhibiting HIV-1 infection (Dropulic *et al.*, 1996). This vector had deletions in all of the viral genes and encoded a U5 specific hammerhead ribosyme that selectively cleaved HIV-1 RNA. The U5 region in the vector was rendered ribosyme resistant by the substitution of critical oligonucleotides at the cleavage site. The vector was therefore only able to replicate in the presence of live HIV-1 and, when co-transfected into Jurkat cells, inhibited viral replication.

1.3.2.5 PACKAGING CELL LINES

The HIV-1 based vectors described above were all produced in transient expression systems because of the toxicity of proteins like VSV-G (Burns *et al.*, 1993) and the HIV-1 protease (Kaplan and Swanstrom, 1991; Kräusslich *et al.*, 1993). Despite relatively high titres (10^6 - 10^7) of replication incompetent virus being produced, transient transfection methods have particular drawbacks. Firstly, viral titres are subject to an inherent batch to batch variation, secondly, there is a limitation on the volume of vector particles that can be

easily generated, and thirdly, they are more susceptible to recombination events due to the high copy numbers of DNA molecules. Therefore for large scale production of replication defective retrovirus, stable packaging cell lines, providing all the viral proteins required in *trans*, analogous to those for MLV, are desirable.

The first stable HIV-1 based packaging cell line was made by a deletion of the major packaging site in the HIV-1 proviral clone (Corbeau *et al.*, 1996). This deletion began 6 nucleotides downstream from the 5' major splice donor to 7 nucleotides upstream of the beginning of *gag*, encompassing a total of 37 nucleotides. This packaging plasmid was then co-transfected with a plasmid carrying a selectable marker and a resistant clone was identified. This clonal cell line was able to synthesise HIV-1 proteins, however, virion like particles produced by these cells were non-infectious because the genomic viral RNA could not be packaged. For evaluation of gene transfer this packaging cell line was transfected with an HIV-1 vector that contained the packaging region along with the first 653 bp of *gag* and the RRE (Parolin *et al.*, 1994). The cell line produced virions capable of transducing CD4+ HeLa T4 cells with a maximum titre of 10^5 IU/ml. There is a potential problem with this type of packaging cell line because one single recombination event between HIV-1 sequences in the packaging cells and the transfer vector could lead to the undesirable production of replication competent virus.

A split genome approach was applied with the view to minimise the chances of homologous recombination (Srinivasakumar *et al.*, 1997). The packaging cell line was created by transfecting five separate plasmids. One of the plasmids expressed *gagpol* and *vif* and the remaining plasmids expressed one of the other genes, *nef*, *tat*, *rev* and *env*. The 5' sequences required for packaging were deleted from each construct. When the packaging cell line was transfected with the transfer vector, virus particles were produced with a titre of up to 10^4 IU/ml.

The constitutive expression of some of the viral proteins required in *trans*, namely HIV-1 *env* (Sodroski *et al.*, 1986), VSV-G (Burns *et al.*, 1993), protease (Kaplan and Swanstrom, 1991; Kräusslich *et al.*, 1993) and *vpr* (Planelles *et al.*, 1995; Rogel *et al.*, 1995) have been shown to be cytotoxic. To overcome this problem inducible packaging cell lines have been developed (Yu *et al.*, 1996; Kaul *et al.*, 1998). Both of these packaging cell lines produce viral proteins only after the removal of tetracycline from the culture medium and yielded low viral titres of up to 10^3 upon transfection with the transfer vector. Another tetracycline inducible HIV-1 packaging cell line was developed that produced pseudotyped VSV-G viral particles at a titre of more than 10^6 IU/ml (Kafri *et al.*, 1999). A stable inducible packaging line based on the doxycycline-repressible expression

of HIV-1 Rev/Gag/Pol and of VSV-G has also been described (Klages *et al.*, 2000). Upon induction, this cell line produced 1 to 20 HeLa-transducing units per cell per day for approximately one week. This compared favourably to yields produced from transient systems and the virions were able to transduce non-mitotic cell targets. More recently, a tetracycline-regulated packaging cell line has been developed that produces self-inactivating (SIN) HIV-1 derived vector particles (Xu *et al.*, 2001). The U3 transcription regulatory elements in the vector were replaced with the Tet-responsive element, which allowed vector production only in cells expressing the synthetic Tet-regulated transactivator. Following transduction of a Tet-regulated HIV-1 packaging cell line with the SIN vector, 10^6 IU/ml of viral particles could be produced, which were able to efficiently transduce terminally differentiated neurons in the brain of rats.

To increase expression from packaging plasmids and to ensure greater and more predictable safety, synthetic *gagpol* genes have been constructed. Synthetic *gagpol* genes of HIV-1 and SIV-1 (simian immunodeficiency virus) have been created in which the codon usage had been optimised for expression in human cells without altering the amino acid sequences (Wagner *et al.*, 2000). The synthetic *gagpol* genes were efficiently expressed in the absence of Rev and could mediate transduction of a vector construct into non-dividing cells with a titre of about 10^6 transducing units per ml. Homologous recombination events between the synthetic *gagpol* plasmid and the vector were shown to be undetectable using a biological assay. A split packaging system has also been reported which segregates the *gag* and *gagpol* into two parts (Wu *et al.*, 2000). One construct expressed Gag/Gag-pro and the other expressed reverse transcriptase and integrase as fusion partners of Vpr. Using an assay developed to detect recombinant HIV-1 mobilisation, it was demonstrated that such a split system approach prevented the generation of recombinants that contained a functional *gagpol* gene structure.

1.3.3 OTHER LENTIVIRAL VECTORS

The use of HIV-1 based gene therapy transfer systems faces a perceived major hurdle because HIV-1 is a deadly human pathogen. The fervour to develop a system absolutely guaranteed to be free from replication competent virus continues, but in the meantime, efforts have also been put towards developing vectors based on other lentiviruses. HIV-2 based vectors have only a marginal advantage over HIV-1 due to its decreased pathogenicity. The *cis* acting regions of HIV-2 responsible for encapsidation have been identified (Poeschla *et al.*, 1998a). A packaging plasmid based on an apathogenic strain of HIV-2 was made by deletion of 61 bp from between the major 5'

splice donor and the *gag* initiation codon. This prevented HIV-2 packaging and replication, but not expression of the viral genes. Co-transfection of this packaging plasmid with VSV-G and a transfer vector produced viral particles at titres of up to 10^6 IU/ml that were suitable for transducing non-dividing cells. Unfortunately transduction of CD34+ HSCs could not be demonstrated. This packaging plasmid was then used to make a stable cell line capable of expression of the HIV-2 viral proteins and devoid of replication competent virus (Arya *et al.*, 1998).

Feline immunodeficiency virus (FIV) is a non-primate lentivirus that does not productively infect human or other known non-feline cells. An FIV gene transfer system has been developed capable of transducing primary human macrophages and post-mitotic neurons with titres of 10^5 IU/ml (Poeschla *et al.*, 1998b). The packaging plasmid was constructed by replacing the 5' LTR with the CMV early promoter and the 3' LTR with the bovine growth hormone polyadenylation signal, and deleting 875 nucleotides from the *env* gene. In the transfer vector, the U3 region of the 5' LTR was replaced by fusing the CMV promoter to the R repeat and a frameshift mutation was introduced into the *gag* gene. This demonstrated that the U3 region in the 5' LTR of FIV limits the productive phase of virus replication in human cells and that the substitution of this region with a heterologous promoter allows expression of FIV proteins in human cells. Since then subsequent efforts have been directed at optimising FIV derived vectors. Second- and third- generation three-plasmid FIV vector systems have now been developed using the minimum amount of FIV viral proteins. Optimal vectors have been designed that transduce non-dividing cells with titres of 1 x 10^6 IU/ml (Curran *et al.*, 2000).

Attempts to generate recombinant vectors from EIAV demonstrated transduction of human cells, but titres were low (Olsen, 1998). Less fruition came from a CAEV based vector whose very poor transduction efficiency was due to a combination of defective RNA packaging and an instability of the viral particles (Mselli-Lakhal *et al.*, 1998). However, successful gene transfer vectors have been developed from a non-primate bovine lentivirus with titres of 10^6 and were shown to transduce a variety of primary and transformed cells (Metharom *et al.*, 2000).

RNA pseudotypes have been developed whereby retroviral particles can package a heterologous RNA (Rizvi and Panganiban, 1993). An HIV-1/HIV-2 pseudotyped vector was generated that was capable of transducing non-dividing cells with titres of 10^4 IU/ml (Corbeau *et al.*, 1998). Recently, an SIV/HIV-1 pseudotyped vector, with titres up to 10^6 IU/ml was reported (White *et al.*, 1999). These vectors were free from replication competent virus and were able to transduce a wide variety of cell types.

Whilst the vast amount of information that has been gathered about all aspects of HIV-1 has facilitated the development of HIV-1 based vectors, others have sought to develop vectors from lentiviruses that are not known to be pathogenic in humans. Nonetheless, without a detailed understanding of the mechanisms that prevent these viruses causing disease in humans, the real gains in safety afforded by this approach cannot be properly assessed.

1.4 A PERSPECTIVE

Since its inception in 1990, the Human Genome Project continues to provide knowledge regarding human genetics (Guyer and Collins, 1995), culminating in the sequence of the entire human genome. As more genes are identified the scope for human gene therapy widens. The general line of thought is that gene therapy is a natural extension of the basic sciences into therapeutic applications, but these therapeutic applications have been a long time in coming. Gene therapy has the potential to revolutionise modern medicine, but will this ever happen?

The future of gene therapy is tightly connected to the development of safe and efficient vector systems that can be used like conventional drugs. For lentiviral vectors there are two immediate stumbling blocks to be overcome. Firstly, prior to testing in humans, detection systems for helper virus must be in place. A standard assay for the detection of replication competent retrovirus for lentiviral vectors does not exist at the moment and it is hard to understand how this may be accomplished. Secondly, the use of lentiviral vectors in human gene therapy will require the development of large-scale production methods, which are currently unavailable. Once these obstacles are overcome, controlled modifications of the human genome using lentiviral vectors may be unlimited.

1.5 AIM

The overall aim of this project was to design and construct a replication defective viral vector system that can transfer a gene of interest into non-dividing cells. A safe and effective HIV-1 based vector was assembled using the principles learnt during the development of MLV based systems. The vector system was centred on HIV-1 as apposed to other lentiviruses because HIV-1 must be one of the most intensely studied viruses as a consequence of the AIDS pandemic. The focus was to engineer such a vector system that could potentially be used for gene therapy for the group of lysosomal storage disorders known as the mucopolysaccharidoses. Much research is being done in the area of

lysosomal storage disorders, in particular the mucopolysaccharidosis (MPS), here in the Department of Chemical Pathology at the Women's and Children's Hospital. These disorders are also considered good candidates for gene therapy approaches. In order to achieve this, the aim is divided into specific objectives, some of which were only developed during the course of the work as results became available.

1.5.1 SPECIFIC OBJECTIVES

1. To construct a hybrid Moloney leukemia virus (MLV) and human immunodeficiency virus type 1 (HIV-1) retroviral vector.

2. To evaluate the HIV-1 vector system's ability to transfer the lysosomal enzyme β-glucuronidase to MPS VII deficient fibroblasts.

3. To create an HIV-1 based vector and evaluate the effects of particular *cis* sequence elements on the performance of the vector.

4. To generate helper plasmids to provide HIV-1 *trans* functions necessary for recombinant virus production.

5. To identify necessary HIV-1 gene sequences and proteins required in the helper system for HIV-1 derived vectors.

6. To develop packaging cell lines for HIV-1 based vectors.

7. To evaluate the safety of the HIV-1 gene transfer system.

CHAPTER TWO

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 COMMON MATERIALS

The following materials were obtained from:

DE52 paper	Whatman International Ltd. (Maidstone, England, UK)
Ministart 0.2 μm filters	Sartorious AG (Göttingen, Germany)
Nitrocellulose membrane	Bio-Rad Laboratories (Hercules, Calif., USA)
Positive Land film, type 667	Polaroid Corp. (Bedford, Mass., USA)
X-ray film	DuPont Photo Products (Sydney, NSW, Australia)

2.1.2 REAGENTS, CHEMICALS AND KITS

The following reagents were obtained from:

Acrylamide solution (LiquiGel)	Gradipore (Pyrmont, NSW, Australia)
Agarose, DNA grade	Progen Industries Ltd. (Darra, QLD, Australia)
Ammonium persulphate	Bio-Rad Laboratories
Amplify solution	Amersham Life Science Ltd. (Buckinghamshire, England)
BCIP/NBT tablets	Sigma Chemical Co. (St. Louis, MO, USA)

Bromophenol blue	BDH Chemicals Ltd. (Poole, Dorset, England)
4-chloro-1-napthol	Sigma Chemical Co.
Deoxyribonucleotides	Boehringer Mannheim (Mannheim, Germany)
Ethidium bromide	Amresco (Solon, Ohio, USA)
Expand™	Boehringer Mannheim
fmol DNA sequencing kit	Promega (Madison, Wis., USA)
Geneclean II kit	BIO 101 (Vista, Calif., USA)
HIV-1 p24 ELISA kit	NEN™ Life Science Products Inc. (Boston, Mass., USA)
Lipofectamine transfection reagent	Gibco-BRL (Glen Waverley, Vic., Australia)
Mark12™ Wide-range protein standards (2.5-200kDa)	Novex (San Diego, Calif., USA)
4-methyumbelliferyl β-D-glucuronide	Sigma Chemical Co.
Naphthol AS-BI β-D-glucuronide	Sigma Chemical Co
Optiphase Hisafe scintillation fluid	Fisons Chemicals (Homebush, NSW, Australia)
Pararosaniline hydrochloride	Calbiochem (La Jolla, Calif., USA)

Plasmid and PCR purification kits	Qiagen (Hilden, Germany) and Geneworks (Adelaide, SA, Australia)
Propidium iodide	Sigma Chemical Co.
Protein A Sepharose	Amersham Pharmacia (Uppsala, Sweden)
Proteinase K	Boehringer Mannheim
Reverse Transcriptase Assay, non-radioactive kit	Boehringer Mannheim
SequaGel XR	National Diagnostics (Atlanta, Georgia, USA)
Sequenase Version 2.0 DNA sequencing kit	Amersham Life Science Ltd.
SppI *Eco*RI, pUC19 *Hpa*II size markers	Geneworks
Taq DyeDeoxy™ Terminator Cycle Sequencing Kit	Applied Biosystems (Foster City, Calif., USA)
TEMED (N,N,N',N'-tetramethyl ethylenediamine)	Bio-Rad Laboratories
Triton X-100	Ajax Chemicals

All other fine chemicals not listed here were of analytical grade and purchased from Ajax Chemicals, BDH Chemicals Ltd., Bio-Rad Laboratories, Boehringer Mannheim and Sigma Chemical Co..

All restriction endonucleases and other enzymes used for DNA preparations were obtained from Boehringer Mannheim and New England Biolabs (Beverly, Mass., USA).

2.1.3 RADIOCHEMICALS

The following radiochemicals were from:

[α-^{32}P]dTTP (500 Ci/mmol)	Geneworks
[γ^{32}P]-dATP (10 mCi/ml)	Geneworks
Sodium [^{35}S]sulphate (550 mCi/mmol)	NEN^{TM} Life Science Products Inc.
[^{35}S]-dATP (10 mCi/ml)	Geneworks
EXPRESS [^{35}S] protein labelling mix	DuPont NEN Research products (Boston, Mass., USA)

2.1.4 BUFFERS AND SOLUTIONS

The buffers and solutions routinely used were as follows:

10 x DNA loading buffer	50 % (v/v) glycerol, 1 % (w/v) SDS, 100 mM EDTA, pH 8.0, 0.1 % (w/v) bromophenol blue
LiCl lysis buffer	2.5 M LiCl, 50 mM Tris-HCl, pH 8.0, 62.5 mM EDTA, 4 % (v/v) Triton X-100
2 x HeBS	0.28 M NaCl, 0.05 M HEPES, 1.5 mM Na_2HPO_4, pH 7.05
Phosphate buffered saline (PBS)	130 mM NaCl, 10 mM $NaHPO_4$, 10 mM NaH_2PO_4, pH 7.2
RIPA buffer	150 mM NaCl, 1 % (v/v) NP-40, 0.5 % (w/v) DOC, 0.1 % (w/v) SDS, 50 mM Tris-HCl pH 8

SDS-PAGE electrophoresis buffer	25 mM Tris-HCl, 0.192 M glycine, 0.1 % (w/v) SDS pH 8.3
SDS-PAGE sample buffer	1 % (w/v) SDS, 4 M urea, 80 mM Tris-HCl, pH 6.8, 0.1 % (w/v) bromophenol blue
20 x SSC	3 M NaCl, 0.3 M tri-sodium citrate, pH 7
TBE	89 mM Tris-base, 89 mM boric acid, 2 mM EDTA, pH 8.3
TE	10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA
TN	20 mM Tris-HCl pH 7, 0.5 M NaCl
Transformation buffer	50 mM $CaCl_2$, 10 mM PIPES-HCl pH 6.6, 15 % (v/v) glycerol (solution sterile filtered)
Western blot:	
Block solution	7.5 % (w/v) glycine, 5 % (w/v) non-fat dry milk, 1 % (w/v) ovalbumin, 5 % (v/v) FCS in PBS.
Transfer buffer	10 mM Caps, pH 11, 10 % (v/v) methanol
Wash buffer	0.01 % (w/v) non-fat dry milk, 0.01 % (w/v) ovalbumin, 1 % (v/v) FCS in PBS

2.1.5 BACTERIAL STRAINS AND MEDIA

E. coli K12 strain MC1061	F^- araD139 Δ(ara-leu)7696galE15galK16 Δ(lac)X74rpsL (Str^r) hsdR2 ($r_k^- m_k^+$) mcrA mcrB1

E. coli K12 strain GM48	F^- thr leu thi lacY galK galT ara fhuA tsx dam dcm supE44
E. coli K12 strain DH10B	F^- mcrA Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara,leu)7697 araD139 galU galK nupG rpsL λ^-
L-broth	1 % (w/v) Bacto tryptone, 0.5 % (w/v) Bacto yeast extract, 1 % (w/v) NaCl pH 7.5.
L-Agar	L-broth, 1.5 % (w/v) Bacto agar.
2 x YT	1.6 % (w/v) Bacto tryptone, 1 % (w/v) Bacto yeast extract, 0.5 % (w/v) NaCl pH 7.5.

All media was autoclaved at 121°C for 20 minutes.
Bacto tryptone and bacto agar were obtained from DIFCO Laboratories (Detroit, Mich., USA).

2.1.6 ANTIBIOTICS

Antibiotics used in this study were obtained from the following companies:

Ampicillin	Boehringer Mannheim
Dexamethasone	Sigma Chemical Co.
Penicillin/Streptomycin sulphate	CSL Ltd. (Parkville, Vic., Australia)
Puromycin	Sigma Chemical Co.
G418 sulphate (Neomycin) Geneticin	Sigma Chemical Co. and Gibco-BRL.

2.1.7 CELL LINES AND CELL CULTURE PRODUCTS

All cell lines and culture materials were from the following:

Cell Lines:	
A549	American Type Culture Collection (ATCC) CCL 185
Chinese Hamster Ovary (CHO-K1)	ATCC, CRL 9618
Cos1	ATCC, CRL 1650
293	ATCC, CRL 1573
293T	ATCC, SD 3515
m3521	Imortalised murine MPS VII fibroblasts (gift from W. Sly)
NIH3T3	ATCC, CRL 1658
PA317	ATCC, CRL 9078 (Miller and Buttimore, 1986)
SF4681	Control skin fibroblasts (Chemical Pathology, Women's and Children's Hospital)
SF382	MPS VII patient skin fibroblasts ATCC, GM 00151
Materials:	
Dulbecco's Modified Eagle Medium (DMEM)	Gibco-BRL

Dulbecco's Phosphate Buffered Saline (PBS) without calcium and magnesium	CSL Ltd.
Foetal calf serum (FCS)	CSL Ltd. and Gibco-BRL

Dialysed FCS was prepared by dialysis of 50 ml of FCS with 3 changes of 500 ml of PBS over a 24 h period at 4°C.

Ham's F12	Gibco-BRL
Methionine/cysteine free DMEM	Gibco-BRL
Sterile pyrogen free water	Baxter Healthcare (Toongabbie, Vic., Australia)
Trypsin/versene solution (0.1 % (w/v) trypsin, 0.02 % (w/v) versene)	CSL Ltd.

All disposable cell culture plasticware was supplied from Costar (Corning, NY, USA) and Greiner labortechnik (Germany).

2.1.8 PLASMIDS

All plasmids were obtained from the following:

pBluescript II SK (+)	Stratagene (La Jolla, Calif., USA)
pBluescriptpur^R	from D. Anson (unpublished)
pcDNA3MLVgagpol	from D. Anson (unpublished)
pHIVext5SV40Neoppt+RRE	from D. Anson (personal communication)
pATX	Kawaoka *et al.*, 1998

pcDNA3.1	Invitrogen (Gröningen, The Netherlands)
pSP70	Promega
pEYFP	Clontech Laboratories Inc. (Palo Alto, Calif., USA)
penvAm	Markowitz *et al.*, 1988b
pCMVrev	from the National Institute of Allergy and Infectious Diseases AIDS Research and Reference Reagent Program, McKesson BioServices, catalogue number 1443 (Rockville, Md., USA) (Lewis *et al.*, 1990)
pLNCX and pLNSX	gift from A. D. Miller (Miller and Rosman, 1989)
pCMVΔRnr	gift from I. M. Verma (Kafri *et al.*, 1997; Naldini *et al.*, 1996b).
pHIV-YU2	gift from S. K. Ghosh (Li *et al.*, 1992, Genbank accession number M93258)
pCMV-G	gift from T. Friedmann (Yee *et al.*, 1994).
pGRE5.1	gift from J. White (Mader and White, 1993).
pGEM4βgluc	gift from A. J. Reuser (Oshima *et al.*, 1987).
signalpIg	Novagen Inc. (Madison, WI, USA)
pCITE4a (+)	Novagen Inc.

pWHV8 ATCC 45097D (Girones *et al.*, 1989)

2.1.9 DNA OLIGONUCLEOTIDES

DNA oligonucleotides were manufactured by Gibco-BRL or Sigma Chemical Co. Oligonucleotides used for PCR, linkers for cloning and constructing coding sequences are listed below. All sequences are 5' to 3'.

PCR primers for vector and helper plasmids:

LTRMLVF1 GCGCGCATGCGGCCGCAATGAAAGACCCCACC

LTRMLVR1 GCGACTAGTCGGATGCAACTGCAAG

LTRMLVF2 GCGCGCGCTGCAGACCTGTAGGTTTGGCAA

LTRMLVR2 GCGCGCAAGCTTGGCAGGGGTCTCCCGA

5'LTRF ATACGAACTGGCGGCCGCTGGAAGGGCTAATTCACTCCCAAC

5'LTRR CTAGCTACTAGTTAAAGCACTCAAGGCAAGCTTTAT

3'LTRF ACGCTGGATCCAAAAGAAAAGGGGGGACTGGAAGG

3'LTRR GACCCATCGATTGCTAGAGATTTTCCACACTGACT

5'HIVext1R ACGCTGGATCCTCTCTCCTTCTAGCCTCCGCTAG

5'HIVext2R ACGCTGGATCCTATGTTTTAATCTATATTGTTTCT

5'HIVext3R ACGCTGGATCCTGGCTACTGTATTATATAATGATC

5'HIVext4R ACGCTGGATCCAATCTTGTGGGGTGGCTCCTTCTG

5'HIVext5R	ACGCTGGATCCTAAAATTGCCTCTCTGCATCATTA
3'LTRF	ACGCTGGATCCAAAAGAAAAGGGGGGACTGGAAGG
3'LTRF2	ACAGGTCTGAATTCCAGCTATGGATCTTAGCCACTTTT
3'LTRF3	ACAGGTCTGAATTCGTTTCAGACCCACCTCCCAGCTCA
3'LTRF4	ACAGGTCTGAATTCCCCACCTCCCAGCTCGGAGGGGAC
3'LTRF6	ACAGGTCTGAATTCAAAGAGCTTTTAGAGCTGTTCTTC
cpptClaI	GGGCCCATCGATTCATCCACAATTTTAAAAGAAAAG
cpptBstBI	GGGCCCTTCGAATGTCCCTGTAATAAACCCGAAAATTTTG
IRESF	CTAGGCCTGGATCCGGTTATTTTCCACCATATTGCC
IRESR	GGGCATATGGATCCATTATCATCGTGTTTTTCAAAG
F6RevMfeI	GGGCCCAATTGTAATATTTCTATAACCC
RREF	ATACGAACTGGCGGCCGCAGGAGCTTTGTTCCTTGGG TTCTT
RRER	CTAGCTAGGTCTAGAAGGAGCTGTTGATCCCTTAGGTAT
H2A5'	CCGCCGGGATCCTGGAGCTGGCGGGCAAC
H2A3'	CCGCCGAATTCCAGCAACTTGTTTAGCTCCTCG
WHVERVF	CCGGGATATCAACCTCTGGATTACAAAAT
WHVPVUR	CCGGGCAGCTGCAGGCGGGGAGGCGGCCCAAA
tatF	ACGCTGAATTCGCCACCATGGAGCCAGTAGATCCTAACCTA

tatR	CTATGCTCTAGATTACTGCTTTGATAGAGAACTTTG
vifF	CGGGAATTCGCCACCATGGAAAACAGATGGCAGGTGATG
vifR	ACGCGGATCCCTAGTGTCCATTCATTGTGCGGCT
vpuF	GCGCATGCCATGGCCACCATGCAATCTTTACAAGTATTAGCA
vpuR	ATAAGAATGCGGCCGCTACAGATCATCAACATCCCAAGG
vprF	CGGCTGCAGGCCACCATGGAACAAGCCCCAGAAGACCAA
vprR	ACGCGGATCCTTAGGATCTACTGGCTCCATTTCT
nefF	GCGCATGCCATGGCCACCATGGGTGGCAAGTGGTCAAAACGT
nefR	ATAAGAATGCGGCCGCTCAGTTCTTGTAGTACTCCGGATG
gagpolF	GAAAGCATAGTAATATGGGGAAA
gagpolR	TGCTGTCCCTGTAATAAACCGGA
amp3.5N	GGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACAC
295F	TCCTGCAGCCCGGGGGATCCCCGAGTGGGAGTT CGTGCACAACCCCCCCCTGGTGAAGCTGTGGT
370R	TGGCCGGTCTCGGCGGGGATCACC
370F	GGTGATCCCCGCCGAGACCGGCCA
pB/SF	GTAAAACGACGGCCAGT
pB/SR	GGAAACAGCTATGACCATG

amp215F	GGCACCAGGTACCAGTACAACGTG
amp215R	CACGTTGTACTGGTACCTGGTGCC
amp2.5'	AGGCCAGGGTGCTGGCCGAGGCCA
amp2.3'	CTCCAGCTGGTACCACAGCTTCACC
260F	TGTGCAAGCTGCTGAGGGGCACCA
260R	TGGTGCCCCTCAGCAGCTTGCACA
R1	ACTGAACGCTGAAGGTGCAGTTGTA
IN5'	CCCGGGATCCAGGTAAGTCGATCCGGAGGGAGGGTGTCTGC
IN3'	CCCGGGAATTCTCTGCAGAGAGAAGATTGGGAGAGAC
RREXba	GGGCCTCTAGAGCTTTGTTCCTTGGGTTCTTG
RREApa	GGGTCGGGCCCAAATCCCTAGGAGCTGTTG
135F	AGCAGACCAGAGCCAACAGC
135R	GCTGTTGGCTCTGGTCTGCT
ramp	TCCTGCCGATGATGTTCACGGGGGT
polHindIII	GGGCCCAAGCTTACCATGTTTTTTAGGGAAGATCTGGCC
FusFamp	ATCGCCAAGAACTGCAGGGCCCCCA
Fusramp	TCCTGCCGATGATGTTCACGGGGGT

pB/ST7 GTAATACGACTCACTATAGGGC

HIV-1 and other coding sequences:

HIV1Fn CTAGTGTGTGCCCGTCTGTTGTGACTCTGGTAACTAGAGA
TCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA

HIV2Fn GTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGAAAACC
AGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAG

HIV1Rn CGCGCTTCAGCAAGCCGAGTCCTGCGTCGAGAGAGCTCCTCTG

HIV2Rn GTTTTCCTTTCGCTTTCAAGTCCCTGTTCGGGCG
CCACTGCTAGAGATTTTCCACACTGACTAAAAGGG

HIV3Rn TCTGAGGGATCTCTAGTTACCAGAGTCACAACAGACGGGCACACA

HIV3Fn CGCGCACGGCAAGAGGCGAGGGGCGGCGACTGCTCAGT
AGGGCAAAAAATTTTTGACTAGCGGAGGCTAGAAGGAG

HIV4Fn AGAGTAGGGTGCGAGAGCGTCAGTATTAAGTGCGGGGGA
TTAAGATAAGTGGGAAAAAATTCGGTTAAGGCCAGGGGG

HIV5Fn AAAGAAACAATATAGATTAAAACATAGGATCCAAAAGAA
GAAGGGGGACTGGAAGGGCTAATTCACTCCCAACTGCA

HIV4Rn GTTGGGAGTGAATTAGCCCTTCCAGTCCCCCCTTTTCTTT

HIV5Rn TGGATCCTATGTTTTAATCTATATTGTTTCTTTCCCC
CGGTTTCCAACCGAATTTTTTCCCACTTATCTAATT

HIV6Rn CCCCCGCACTTAATACTGACGCTCTCGCACCCTACT
CTCTCCTTCTAGCCTCCGCTAGTCAAAAATTTTTTG

HIV7Rn CCCTACTGAGCAGTCGCCGCCCCTCGCCTCTTGCCGTG

gag1F	AATTCAACCATCAGCAAGCAGGTCATTGTTCCAACAT
	CGGGGCTCGTGCTAGCGTTCTGTCCGCTGGTGAGCTG
gag2F	GACAAATGGGAGAAGATCCGTCTGCGCCCGGGCGGTA
	AAAAGCAGTACCGTCTGAAGCACATCGTTTGGGCTTC
gag3F	TCGTGAACTGGAGCGCTTTGCTGTAGACCCGGGTCTG
	TCGAGGACCAGCGAGGGTTGCCGTCAGATCCTGGGCC
gag4F	AACTGCAGCCATCTCTGCAAACTGGTTCCGAGGAACT
	GTCGCCCCTGTACAACACCGTTGCGACTCTGTACTGT
gag1R	GTACACAGTACAGAGTCGCAACGGTGTTGTACAGGGA
gag2R	GCGCAGTTCCTCGGAACCAGTTTGCAGAGATGGCTGC
	GAGTTGCCCAGGATCTGACGGCAACCCTCGCTGGTCT
gag3R	CCAGCAGACCCGGGTCTACAGCAAAGCGCTCCAGTTC
	AGAGCAAGCCCAAACGATGTCTTCAGACGGTACTGC
gag4R	TTTTTACCGCCCGGGCGCAGACGGATCTTCTCCCATT
	TTGACCGCTCACCAGCGGACAGAACGCTAGCACGAGC
gag5R	GCCCATGTTGGAACAATGACCTGCTTGCTGATGGTTG
CTEF1	ATCCTCCCCTGTGAGCTAGACTGGACAGCCAATGACGGGTAAGAGAGTG
CTEF2	ACATTTCTCACTAACCTAAGACAGGAGGGCCGTCAAAGCTACTGCCTAA
CTEF3	TCCAATGACGGGTAATAGTGACAAGAAATGTATCACTCCAACCTAAGAC
CTEF4	AGGCGCAGCCTCCGAGGGATGTGTGAT
CTER1	ATCACACATCCCTCGGAGGCTGCGCCTGTCTTAGGTTGGAGTGATA
	CATTT
CTER2	CTTGTCACTATTACCCGTCATTGGATTAGGCAGTAGCTTTGACGGCCCT

CTER3 CCTGTCTTAGGTTAGTGAGAAATGTCACTCTCTTACCCGTCATTGGCTG

CTER4 TCCAGTCTAGCTCACAGGGGAGTAG

Linkers for cloning:

BssHIIF AATTCACTAGTTTCGGCGCGCCAATATCCTGTTCCTGCAGG

BssHIIR AATTCCTGCAGGAACAGGATATTGGCGCGCCGAAACTAGTG

BamHIF GATCCCAGGCCTAAGCTTGGAATTC

BamHIR GATCGAATTCCAACCTTAGGCCTGG

BstEIIlinkerF AATTGGGTGACCGCGCG

BstEIIlinkerR GATCCGCGCGGTCACCC

2.1.10 ANTISERUM

Antisera to HIV-1 proteins were all obtained from the National Institute of Allergy and Infectious Diseases AIDS Research and Reference reagent Program:

HIV-1_{NL4-3} *vpr* antiserum	catalogue # 3252
HIV-1_{NL4-3} *vpu* antiserum	catalogue # 969
HIV-1_{HXB2} *vif* antiserum	catalogue # 2221
HIV-1 *nef* antiserum	catalogue # 2949
anti rabbit IgG conjugated to alkaline phosphatase	Sigma Chemical Co.
anti human IgG conjugated to horseradish peroxidase	Silenus Labs (Melbourne, Vic., Australia)

HIV-1 positive human serum was kindly provided by L. Peng (Institute of Medical and Veterinary Science (IMVS), Adelaide, Australia).

2.2 GENERAL METHODS

2.2.1 PLASMID DNA PREPARATIONS

2.2.1.1 MINIPREPS

Single colonies were inoculated into 1 ml 2 x YT medium containing 100 μg/μl ampicillin and grown 4 h to overnight with aeration. The cells were sedimented by microcentrifugation of 0.5 ml of this culture for 5 min at maximum speed. Pelleted cells were resuspended in 100 μl of LiCl lysis buffer, and extracted with an equal volume of TE saturated phenol/chloroform (1:1) and the supernatant transferred to a clean eppendorf tube containing 60 μl of isopropanol. The DNA was precipitated by microcentrifugation at maximum speed for 20 minutes, washed with 70 % (v/v) ethanol, air dried and finally resuspended in 20 μl TE buffer.

2.2.1.2 LARGE SCALE PLASMID PREPARATIONS

Single colonies of recombinant clones were grown 6 - 8 h in 5 ml of 2 x YT medium containing 100 μg/μl ampicillin with aeration. Overnight cultures of 50 - 200 ml were then grown from 0.5 ml of the earlier culture in the same medium. The culture was then centrifuged to pellet the bacterial cells and plasmid was purified using commercial plasmid preparation kits according to the manufacturer's instructions.

2.2.2 PHENOL/CHLOROFORM EXTRACTION AND ETHANOL PRECIPITATION OF DNA

Crude DNA samples were extracted by adding an equal volume of TE saturated phenol/chloroform (1:1). The mixture was then vortexed, centrifuged for 5 min in a microcentrifuge and the aqueous phase collected into a clean tube.

DNA samples were precipitated by the addition of a 1/10 volume of 3 M sodium acetate, pH 5.5 and 2.5 volumes of 97 - 100 % (v/v) ethanol for 1 h to overnight at -20°C. Following microcentrifugation, DNA pellets were rinsed with 70 % (v/v) ethanol, air dried and then resuspended in either water or TE buffer.

2.2.3 RESTRICTION ENDONUCLEASE DIGESTION OF DNA

Restriction endonuclease digestion of all DNA preparations was carried out using the conditions appropriate for each enzyme as detailed by the manufacturer.

2.2.4 BLUNT-ENDING OF DNA FRAGMENTS

Following restriction enzyme digestion, 5' overhanging ends were filled in with 1 unit of DNA polymerase (Klenow fragment) and 1 mM dNTPs at room temperature for 30 min. The buffer used for the reaction was the same as that for the restriction endonuclease digestion. The reaction was then heated at 75°C for 10 min to inactivate the polymerase activity.

For 3' overhanging ends, 1 unit of T4 DNA polymerase and 1 mM dNTPs were added to the restricted DNA. The volume of the reaction was made up to 1 x T4 polymerase reaction buffer (Boehringer Mannheim) by the addition of 10 x buffer and water. The reaction was incubated at 11°C for 20 min and then terminated by heating to 75°C for 10 min.

2.2.5 AGAROSE GEL ELECTROPHORESIS OF DNA

Electrophoresis of DNA was performed in Kodak Biomax QS710 tanks for horizontal submerged gel electrophoresis. Samples were mixed with 10 x DNA loading buffer and electophoresed in agarose of 0.8 - 4 % (w/v) in 1 x TBE until the bromophenol blue marker dye had migrated a sufficient distance to ensure adequate separation of the desired DNA fragments. The gel was then stained briefly in a 10 μg/ml solution of ethidium bromide in 1 x TBE buffer to allow visualisation of the DNA. The gels were then rinsed in deionised water and photographed using an ultraviolet transilluminator (ULTRA-LUM Inc, Carson, Calif., USA) with a Polaroid MP-4 Land Camera containing 667 film. For preparative gels the required DNA fragment was excised from the gel and then purified using a Geneclean II kit.

2.2.6 PLASMID VECTOR PREPARATION

The plasmid was firstly digested to completion with the appropriate restriction endonuclease as assessed by agarose gel electrophoresis. The 5'-terminal phosphate

groups were then removed with calf intestinal phosphatase (CIP) by incubation for 1 h at 37°C, in the same buffer used for the restriction enzyme digest but with the addition of 15 mM EDTA to inhibit endonuclease activity. The restricted and phosphatased DNA was then extracted with TE saturated phenol/chloroform and precipitated with 2 volumes of isopropanol in the presence of 2.5 M ammonium acetate.

2.2.7 LIGATION OF DNA

Ligation of DNA with overhanging ends was performed in 1 x ligation buffer (Boehringher Mannheim) with 1 unit of T4 DNA ligase for 4 - 24 h at room temperature. For the ligation of blunt-ended DNA fragments, 2 units of T4 DNA ligase were used and the reaction proceeded for a minimum of 16 h at room temperature. Generally, the DNA insert and vector preparations were semi-quantitated by agarose gel electrophoresis, and combined in a 3:1 molar ratio in a final reaction volume of 20 µl.

2.2.8 PREPARATION OF COMPETENT *E.COLI*

Three strains of *E.coli*, MC1061, DH10B and GM48 were made competent for use in transformation. Overnight cultures were obtained for each strain from inoculation with a single colony in 5 ml of 2 x YT broth. For MC1061, 1 L of 2 x YT was inoculated with 1/100 of the overnight culture and grown with aeration until the OD_{600nm} was 0.4. The culture was then cooled on ice and the cells pelleted at 2,500 x g for 10 min at 4°C. The cells were resuspended in 500 ml of ice cold transformation buffer and incubated on ice for 30 min. Again the cells were pelleted and resuspended in 50 ml ice cold transformation buffer and incubated on ice for at least 1 h. Cells were either used directly or aliquoted into 300 µl lots, snap frozen in liquid N_2 and stored at -80°C until needed.

For DH10B and GM48, cells were prepared directly as required. A 50 ml culture of 2 x YT broth was inoculated with 0.5 ml of a fresh overnight culture and grown until OD_{550nm} reached 0.3. The cells were pelleted at 2,500 x g for 10 min at 4°C and resuspended in 25 ml of ice cold 50 mM $CaCl_2$. Following a 20 min incubation on ice the cells were pelleted again and resuspended in 5 ml of ice cold 50 mM $CaCl_2$. The cells were left on ice for at least 1 h before use and were viable for up to 3 days if kept at 0°C.

2.2.9 TRANSFORMATION OF *E.COLI*

For DH10B and GM48, 200 µl of cells and 5 µl of the ligation reaction were mixed and incubated on ice for 30 min. The transformation reaction was then heat shocked at

42°C for 2 min and immediately returned to ice and 200 µl of 2 x YT added. Cells were incubated at 37°C for 40 min and then plated onto L-Amp plates and incubated overnight at 37°C. Competent MC1061 was thawed on ice and 5 µl of the ligation reaction was mixed with 100 µl of cells and incubated on ice for 10 min. The cells were then heat shocked for 5 min at 37°C, recovered with the addition of an equal volume of 2 x YT and then incubated for a further 40 min at 37°C. This transformation reaction was finally plated onto L-Amp plates and incubated overnight at 37°C.

2.2.10 DNA SEQUENCING

The majority of sequencing performed was automated, using a Taq DyeDeoxy™ Terminator Cycle Sequencing Kit according to the protocol included with the kit. The samples were then sent to the IMVS central sequence facility and loaded onto an Applied Biosystems 373A DNA Sequencer according to the instructions in the User's Manual.

Some sequence data were obtained by manual sequencing using either a Sequenase Version 2.0 DNA sequencing kit or a *fmol* DNA seqencing kit. Reactions were performed as described by the manufacturers of the sequencing kits using [^{35}S]-dATP as the labelled nucleotide. Chain termination products were resolved by electrophoresis on 6 % (w/v) SequaGel XR followed by autoradiography.

2.2.11 POLYMERASE CHAIN REACTION (PCR)

2.2.11.1 ISOLATION OF SEQUENCES

PCRs were performed to amplify specific regions of the HIV-1 genome and other plasmids for sub-cloning. Generally 2 units of Expand™ polymerase were used for each amplification reaction in a final 1 x concentration of the supplied buffer, 200 µM of each dNTP, 250 µM of each primer and 1 µg of template. Between 10 - 20 cycles of amplification were performed depending on the quantity of resultant product. A typical cycle consisted of denaturation at 95°C for 15 sec, annealing at 55 - 65°C for 20 sec and 68°C elongation based on a synthesis rate of 1 kb per min. The amplified product was purified using a PCR purification kit, and cloned usually *via* restriction endonuclease sites incorporated into the PCR primers. Resultant clones were sequenced to confirm error-free sequence.

2.2.11.2 DETECTION OF GAGPOL SEQUENCE

A PCR was also developed for the detection of *gagpol* plasmid sequence. Two primers gagpolF and gagpolR were used at concentrations of 250 μM in a final reaction mixture of 100 μl, containing 1 x Taq PCR buffer, 2 units of Taq and 200 μM of each dNTP. To evaluate the reaction, 1 ng of plasmid or 1000 cells was used and amplified for 40 cycles at 94°C for 15 sec, 58°C for 20 sec and 72°C for 45 sec. Reaction products were then analysed by agarose gel electrophoresis.

2.2.12 OLIGONUCLEOTIDE ASSEMBLY

Overlapping oligonucleotides were assembled by a modification of the method described by Stemmer *et al.* (1995) which uses a polymerase to build increasingly longer DNA fragments during assembly. The reaction was performed in two steps, firstly assembling all the oligonucleotides and then amplifying the entire fragment with two outer primers. In the first stage all oligonucleotides were individually dissolved in water at a final concentration of 100 μM. Equal volumes of all the oligonucleotide solutions were mixed and then diluted 100 fold in a 20 μl PCR reaction mix containing 5 units of Expand™ in 1 x reaction buffer and 200 μM each dNTP and then subjected to 55 cycles at 94°C for 30 sec, 52°C for 30 sec and 68°C for 30 sec. This assembly reaction was then diluted 40 fold in a 100 μl PCR mix as for the assembly reaction along with the addition of 100 μM of two outer amplification primers and amplified for 23 cycles at 94°C for 30 sec, 52°C for 30 sec and 68°C for 60 sec. The final assembled and amplified product was run on a preparative agarose gel, purified using a Geneclean II kit and cloned. Resultant clones were sequenced for confirmation.

2.2.13 REVERSE TRANSCRIPTASE ASSAYS

The majority of reverse transcriptase assays were performed using a Reverse transcriptase assay kit for non-radioactive detection (Boehringer Mannheim), as described by the manufacturer. Initial reverse transcriptase measurements however, were determined by monitoring the incorporation of radioactive dTTP based on a method described for assaying murine reverse transcriptase (Goff *et al.*, 1981). Viral pellets were prepared from 2 ml of viral supernatant by ultracentrifugation at 50,000 x g in a swing out T75 rotor in a Beckman ultracentrifuge. The viral pellet was then resuspended in 50 μl of 50 mM Tris-Cl pH 8.3, 10 mM DTT, 1 mM $MnCl_2$, 0.15 M NaCl, 5 μg/ml oligo(dT), 5 μg/ml poly(rA) 10 μM dTTP, 0.05 % (v/v) NP-40 and 25 Ci/mmol [α-^{32}P]dTTP and then incubated for 1

h at 37°C. A 10 µl sample of each reaction was then spotted onto DE52 paper and washed in 2 x SSC. The paper was soaked for 1 min in 95 % ethanol, air dried and then counted in a scintillation counter.

2.2.14 SCINTILLATION COUNTING

Samples containing radioisotopes (usually less than 1 ml) were counted following the addition of 4 ml Optiphase HiSafe scintillation fluid in a Wallac 1409 liquid scintillation counter (Wallac Oy, Turku, Finland).

2.2.15 CELL CULTURE

All cell lines used were maintained in growth medium, usually DMEM except for CHO-K1 cells, which were grown in Ham's F12 medium, supplemented with 10 % (v/v) FCS at 37°C in a 5 % CO_2 atmosphere. Serum free Ham's F12 medium was used for $^{35}SO_4$ labelling of MPS VII fibroblasts (see section 2.2.19). For transfections using calcium phosphate precipitation, the culture media was supplemented with penicillin (100 U/ml) and streptomycin sulphate (100 µg/ml). Primary cultures of human skin fibroblasts were harvested by washing once in PBS followed by trypsin/versene treatment, all other cell types were also washed once with PBS but harvested with a 10 fold diluted solution of trypsin/versene in PBS.

Unless indicated otherwise, cell lysates required for further analysis were prepared by centrifugation of the trypsinised cells followed by two washes with PBS. Cells were then resuspended in 200 µl of TN buffer and subjected to six cycles of freezing and thawing. The cell debris was then removed by centrifugation and the supernatant transferred to a new tube.

2.2.16 CELL COUNTING

Cells to be counted were harvested with trypsin/versene, washed once in PBS and then resuspended in typically 1 ml of PBS. A 10 µl aliquot of cells was then mixed with an equal volume of 0.1 % (w/v) trypan blue in 0.9 % (w/v) NaCl. The cell number was then determined using a heamocytometer counting chamber (Weber Scientific Int. Ltd., Lancing, Sussex, UK). Viable cells (not stained blue) were counted in the central square of the counting chamber, and the number of cells per ml determined by multiplying the cell count by 2 (dilution factor) and 10^4.

2.2.17 DNA TRANSFECTION

Lipofectamine transfection reagent was used for transfecting DNA into Cos1, 293 and NIH3T3 cells. For a 10 cm culture dish, plasmids (up to a total of 10 μg of DNA) were combined in 100 μl of sterile water and added to 900 μl serum free DMEM. This was then added to a separate tube containing 60 μl of Lipofectamine and 900 μl serum free DMEM, mixed, and left to stand at room temperature for 30 min. This Lipofectamine/DNA/DMEM mixture was added to another 4 ml of serum free DMEM and then added to a dish of cells (60 - 80 % confluent) that had been washed once with serum free DMEM. The cells were incubated with the DNA/Lipofectamine mixture for 16 h after which time the medium was replaced with DMEM supplemented with 10 % (v/v) FCS. Cells were either placed into selective medium 24 h later or virus containing medium was collected 48 post transfection.

For 293T cells calcium phosphate mediated transfection was used (Pear *et al.*, 1993; Jordan *et al.*, 1996). Cells were continuously sub-cultured 1 in 2 for at least 2 consecutive days such that they did not reach confluency. The cells were then harvested, counted and seeded at 8.0 x 10^5 cells per 35 mm dish ready for transfection the following day (ie 16 h later). Just prior to transfection the medium was replenished with fresh medium. Plasmids were made up to a volume of 54 μl with sterile water to which 6 μl of 2.5 M $CaCl_2$ was added. This DNA/$CaCl_2$ mixture was then added drop-wise to 60 μl of 2 x HeBS with continuous vortexing and the entire precipitate was added directly to the 35 mm dish of cells. The transfection medium was left on the cells for 24 h and was then replaced with fresh growth medium.

Electroporation was used to introduce plasmid DNA into CHO-K1 cells. CHO-K1 cells were grown to about 80 % confluency in a T75 cell culture flask, harvested with 10 % (v/v) trypsin/versene, washed once in PBS and then resuspended in 800 μl of PBS. The cells were transferred into a disposable electroporation chamber along with the plasmid DNA (20 μg) to be transfected and equilibrated on ice for 15 min. The cells were then electroporated with a pulse of 275 V/330 μF using a BRL Cell Porator. Electroporated cells were removed and placed in Ham's F12 growth medium supplemented with 10 % (v/v) FCS for 48 h. They were then subcultured 1:5 into the same medium containing 0.8 mg/ml G418.

2.2.18 VIRUS PRODUCTION AND TRANSDUCTION

For virus production 293T cells were used and transfected as described above. Following a 24 h incubation with transfection medium, each 35 mm dish of 293T cells was replenished with 2 ml of medium. Viral supernatants were collected 48 h later, centrifuged at low speed (2,000 x g) to remove debris and then filtered through 0.2 μm filters. Viral titres were determined on NIH3T3 cells or MPS VII fibroblasts. Confluent cultures of MPS VII fibroblasts were established in 35 mm grid dishes for transduction. For NIH3T3 transduction, cells were counted, plated at 1.2 x 10^6 cells/35 mm dish and then allowed to adhere for 3 - 4 h. Overnight transductions with virus stock in culture medium supplemented with 4 μg/ml polybrene, typically in a volume of 2 ml, were then performed after which time the viral containing medium was removed and replenished with fresh growth medium. Transduced cells were sub-cultured 1 in 3 after another 24 h, and 72 h post-transduction were assessed for virus titre by staining for β-glucuronidase activity (MPS VII fibroblasts) (section 2.2.19) or FACScan analysis for EYFP expression (NIH3T3) (section 2.2.21).

Using highly concentrated VSV-G pseudotyped viral particles might lead to pseudotransduction artefacts. The *lacZ* reporter gene has been shown to produce a pseudotransduction phenomena under conditions where concentrated MoMLV/LacZ vectors were used (Liu *et al.*, 1996). Pseudotransduction was ruled out with the HIV-1 derived vector system described in this thesis by measuring reporter gene expression (EYFP by FACScan analysis) up to10 days post-transduction indicating also that viral titres are stable over this time frame.

293T and A549 cells were also transduced for the "*gag*" and "*gagpol*" transfer assays as described in section 2.2.25. For transducing these cells virus stocks were concentrated by a combination of ultrafiltration (Amicon stirred cell with ZM500 membrane) and ultracentrifugation at 50,000 x g in a swing out T75 rotor in a Beckman ultracentrifuge, at 4°C.

2.2.19 β-GLUCURONIDASE DETECTION

The demonstration of β-glucuronidase in transduced NIH3T3 cells was based on a simultaneous coupling reaction (Hayashi *et al.*, 1964). The cells were washed in PBS and then fixed in formal calcium (40 % (v/v) formaldehyde, 10 % (w/v) $CaCl_2$) at 4°C for a minimum of 1 h. A substrate solution was prepared containing 0.5 mM naphthol AS-BI β-

D-glucuronide in 50 mM $NaHCO_3$ and 0.1 M acetate buffer, pH 5.0. For the diazo reagent, 1 g of pararosaniline hydrochloride was dissolved in 20 ml distilled water and 5 ml concentrated HCl with gentle warming. The incubating reagent (0.25 mM substrate and 1.8 mM diazo reagent) consisted of 0.3 ml each of the pararosaniline hydrochloride solution and a 4 % (w/v) sodium nitrite solution (freshly prepared) that was firstly mixed together, 10 ml of the substrate solution was then added. The pH was adjusted to 5.2 with 1 M NaOH, the volume made up to 20 ml with distilled water and the final incubating solution was filtered prior to use. Enough incubating solution was added to cover the dish of cells, which were then placed at 37°C for 1 - 2 h. The incubating solution was washed off with distilled water and stained cells (pale pink to a dark red colour) were counted on 40 x magnification under a light microscope. Cells were counted with the aid of a grid on the 35 mm dish upon which the MPS VII fibroblasts were plated. In instances of low numbers of stained cells, the number of pink to red cells was counted in every 10 squares of the grid.

2.2.20 LABELLING OF MPS VII FIBROBLASTS

MPS VII skin fibroblasts were grown to confluence and then labelled with $^{35}SO_4$ essentially as described by Anson *et al.*, (1992). Firstly DMEM/10 % (v/v) FCS growth medium was replaced with Ham's F12/10 % (v/v) dialysed FCS. The following day the medium was replenished with the inclusion of 10 μCi/ml $Na_2[^{35}S]SO_4$ and cultured for 72 h. This medium was then removed, the cells rinsed in DMEM/10 % (v/v) FCS and re-fed with DMEM/10 % (v/v) FCS. After another 72 h, the cells were harvested by trypsinisation, washed by resuspension in PBS and lysed by six cycles of freeze/thaw in TN buffer and then clarified by microcentrifugation. The cell lysates were assayed for β-glucuronidase activity using 4-methyumbelliferyl β-D-glucuronide (Glaser and Sly, 1973) as substrate. Fluorescence of the 4-methylumbelliferone (4MU) was measured spectrophotometrically using a Perkin Elmer spectrofluorometer at the emission wavelength of 446 nm (E_{446nm}) and an excitation wavelength of 366 nm. One relative fluorogenic unit of activity was defined as the amount of enzyme required to produce 1 nmole of (4MU) per min; 1 nmol of 4MU was equivalent to an E_{446nm} signal of 10. The incorporation of ^{35}S into glycosaminoglycans was measured by scintillation counting of the cell extract.

2.2.21 FACSCAN ANALYSIS

Cells for analysis were harvested by trypsinisation, washed in PBS and then resuspended in PBS containing 10 μg/ml propidium iodide. Cells were measured for expression of EYFP (FL2, FITC channel) and propidium iodide (FL1). Cells were analysed using the Cellquest software (version 3.0 1f, Becton Dickinson) after gating on live cells as determined by low propidium iodide fluorescence and forward/side scatter plots. Histogram markers for EYFP positive cells were used such that untransduced NIH3T3 cells gave a false positive rate of 0.26 % ± 0.10 (n = 10). For all experiments a false positive background rate of 1 % was then assumed and deducted from the experimental value. Virus titre was calculated by multiplying the percentage of positive cells by the number of NIH3T3 cells used (generally 1.2 x 10^6) with suitable adjustments incorporated for the volume of virus added.

2.2.22 METABOLIC LABELLING AND IMMUNOPRECIPITATION

The medium from transfected CHO-K1 cells (T75) was aspirated and the cells washed twice with PBS. Cells were then incubated for 30 min with 3 ml of methionine/cysteine free DMEM containing 10 % (v/v) dialysed FCS. Following equilibration the medium was aspirated and replaced with another 3 ml of the same medium containing 100 μCi/ml EXPRESS [^{35}S] protein labelling mix. The cells were labelled for 5 h at 37°C in 5 % CO_2, the medium removed and the cell monolayer washed three times with PBS. The cells were harvested from the plastic culture flask by scraping with a rubber policeman (cell scraper), washed twice with PBS and resuspended in 100 μl RIPA buffer and freeze/thawed six times. To the cell lysate 100 μl of protein A Sepharose slush (~ 50 μl of beads) was added and incubated on a rotator overnight at 4°C. The Sepharose was sedimented by centrifugation and the supernatant transferred to a new tube. The relevant antibody was then added to this and incubated for 24 h on a rotator at 4°C. After this time, 100 μl of protein A Sepharose slurry was again added followed by a further overnight incubation on a rotator at 4°C. The protein A Sepharose was sedimented by centrifugation and washed by resuspenson and centrifugation four times with RIPA buffer. Antigen and antibody complexes were eluted in 100 μl of SDS-PAGE sample buffer, of which 10 μl aliqouts were loaded onto 15 % (w/v) polyacrylamide mini SDS-PAGE gels. These mini-gels were 82 mm wide by 73 mm long and were prepared and run using the Hoeffer Mighty Small Electrophoresis units. Electrophoresis was for

approximately 1 h at 200 V until the bromophenol blue was at the bottom of the gel. The gel was then incubated in Amplify solution overnight and subjected to fluorography.

2.2.23 WESTERN BLOTTING

For analysis of the HIV-1 accessory proteins, Vif, Vpr, Vpu and Nef, transfected Cos1 cells were harvested with trypsin/versene, washed in PBS and then resuspended in 20 µl RIPA buffer. Cell lysates were prepared by six cycles of freeze/thaw. A 5 µl (10 µg) aliquot of the cell lysate was mixed with 15 µl SDS-PAGE sample buffer and solubilised at room temperature for a minimum of 1 h. Samples were then separated on 15 % (w/v) polyacrylamide mini SDS-PAGE gels, again using the Hoeffer Mighty Small Electrophoresis units, for approximately 1 h at 200 V until the bromophenol blue was at the bottom of the gel. Gels were then equilibrated in transfer buffer and transferred onto nitrocellulose at 250 mA for 1 h. Non-specific protein sites were masked by incubating the membrane overnight at 4°C in block solution. The membrane was washed three times in wash buffer and then incubated with the required antibody overnight at 4°C at a 1 in 1000 dilution. The blot was again washed three times and incubated with a secondary antibody (anti rabbit IgG conjugated to alkaline phosphatase) for 1 h at room temperature. Another three washes were performed, followed by a final wash in 0.1 M Tris-HCl, pH 9.5, 0.1 M NaCl, 5 mM $MgCl_2$ and the blot was developed with BCIP/NBT dissolved tablets. After sufficient colour development, the reaction was stopped by rinsing the membrane with water.

For analysis of p24 protein in 293T cells transfected with plasmids expressing Gag and GagPol, cells were harvested and cell lysates were prepared as described in section 2.2.15. One hundred micrograms of protein from each cell lysate was mixed with an equal volume of SDS-PAGE sample buffer, denatured at 100°C for 5 min, electrophoresed on a 12.5 % (w/v) polyacrylamide SDS-PAGE gel using a Bio-Rad Laboratories system (16 x 20 cm) and transferred onto a nitrocellulose membrane. The membrane was incubated for 1 h at room temperature in block solution and subsequently hybridised with heat inactivated (56°C for 1 h) HIV-1 positive human serum. Signals were detected with a horseradish peroxidase secondary antibody (anti human IgG conjugated to horseradish peroxidase) and 4-chloro-1-napthol.

2.2.24 STATISTICAL ANALYSIS

Unless otherwise indicated results are given as mean ± standard deviation. In cases of n = 2 the mean and individual values are reported. Where appropriate statistical analysis was performed using a two tailed Students t-test to compare data. Results were considered significantly different at $p < 0.05$.

2.2.25 ASSAY FOR TRANSFER OF GAG AND GAGPOL SEQUENCES

High titre stocks of virus were prepared using either pCMVΔRnr, pcDNA3gagpolml or pcDNA3gagml/pcDNA3gagpolfusionml (plasmids are described in Chapter four) and concentrated to titres of 2.0×10^7, 2.7×10^7 and 0.8×10^7 NIH3T3 transducing units per ml respectively. Virus was prepared, concentrated and the titre determined on NIH3T3 cells as described in section 2.2.18. In 6 well dishes, 293T cells were plated at 1.5×10^6 cells per well and 6 h later transduced with virus in the presence of 4 μg/ml polybrene in a total volume of 2 ml. Following transduction the cells were expanded and, for each assay, 2×10^6 were plated in a 60 mm dish. Untransduced 293T cells were used as a control. Twenty-four hours after plating, the cells were transfected with a mixture of plasmids encoding either (i) the vector construct pHIVext5SV40Neoppt+RRE (this vector is analogous to pHIVext5SV40EYFPppt+RRE (section 5.3) except that it transduces the neomycin resistance gene), VSV-G, Tat, Rev or (ii) VSV-G, Tat, Rev and pHIVext5SV40Neoppt+RRE. A description of the plasmids encoding these helper functions is provided in Chapters three and four. Conditioned medium was collected, 0.2 μm filtered and assayed on A549 cells (these cells were used because they yield approximately ten- to twenty- times higher titres than NIH3T3 cells) for Neo^R transducing units as follows. The A549 cells were plated at 2×10^6 per 60 mm dish and transduced 5 h later in the presence of 4 μg/ml polybrene, with the entire volume of conditioned medium. Twenty-four hours afterwards the cells were subcultured into 3 x 100 mm dishes and after a further 24 h selected with 1 mg/ml G418 for two weeks after which time G418 resistant colonies were enumerated. In the first of these assays (ie transfection with the gagpolfusionml-containing plasmid) the production of Neo^R transducing virus particles will depend on the transfer of biologically active sequences encoding Gag to the 293T cells *via* the virus stock that is being assessed. This has been referred to as the "*gag* transfer" assay. In the second instance, production of Neo^R transducing virus particles will depend on the transfer of biologically active sequences

capable of substituting for the HIV-1 *gagpol* gene sequence (ie capable of expressing both the Gag and GagPol polyproteins). This has been termed the "*gagpol* transfer" assay.

CHAPTER THREE

GENE TRANSFER FOR β-GLUCURONIDASE

3.1 INTRODUCTION

The main safety criterion for retroviral gene vector systems of any sort (including human immunodeficiency virus type 1 (HIV-1)), is the probability of contamination with replication competent virus. Replication competent viruses can be generated by recombination of the constituents of the virus production system (ie vector and packaging constructs) with each other or with endogenous viral sequences in the cell lines used to generate virus. The probability of such recombination occurring can be minimised by separating the packaging (viral *trans*) functions onto several different plasmids and by reducing homology between the different constituents of the vector production system as far as possible. The power of designing virus production systems in this way has been well-demonstrated over the years during the development of vector/packaging systems from oncogenic viruses such as Moloney leukemia viruses (MLV) (Cosset *et al.*, 1995; Rigg *et al.*, 1996).

A vector/packaging system based on HIV-1 was designed based on the principles and the lessons learnt from the construction and modification of MLV based gene transfer systems. The advantages of using HIV-1 as a gene transfer vector have been described in section 1.3.2.1. The approach taken was to separate the *cis* and *trans* functions of HIV-1 and pseudotype the virus particles with an MLV amphotropic envelope protein. HIV-1 genomes have been shown to be packaged into heterologous retroviral envelopes, resulting in a broadened host range (Spector *et al.*, 1990). An HIV-1 based vector designed to express the lysosomal enzyme β-glucuronidase was constructed along with a panel of helper plasmids for the viral proteins of HIV-1. The deficiency of β-glucuronidase causes a classical lysosomal storage disorder known as Mucopolysaccharidosis type VII (MPS VII) or Sly syndrome (Sly *et al.*, 1973). Central nervous system manifestations in this disorder are severe with the accumulation of storage vacuoles in all cell types. MPS VII, among other lysosomal storage disorders is a candidate for gene therapy. Retroviral vector mediated gene transfer of β-glucuronidase into the bone marrow of MPS VII mice has demonstrated low level expression of the enzyme and clearance of accumulated storage material in the liver and spleen (Wolfe *et al.*, 1992). Such retroviral gene transfer is unlikely to affect the lesions in the central nervous system because of the requirement for cell division for transduction of the haematopoietic stem cell (HSC). However, it is understood that this should be overcome with an HIV-1 based vector. For this reason β-glucuronidase was inserted into the gene transfer vector, for preliminary *in vitro* evaluation in MPS VII skin fibroblasts. A practical advantage of using this system resides in the

ability to detect β-glucuronidase activity both by a simple enzyme assay (Glaser and Sly, 1973) and histochemically (Hayashi *et al.*, 1964).

The design of the proposed vector was to combine MLV U3 and R LTR sequences with the HIV-1 U5 and other *cis* signals from HIV-1. This was proposed as a way to make the LTR constitutive, removing the dependence of the HIV-1 LTR on the Tat protein. The rationale behind this was to try and generate a vector without the *trans*-activating mechanisms used by HIV-1 and one in which the MLV LTR could also be used to express the heterologous sequence. Some additional sequence changes were made, including mutagenesis of firstly, the *gag* initiation codon in the vector packaging signal to prevent translation initiation, and secondly, the splice donor to prevent any possible splicing to cryptic splice acceptor sites within the entire vector sequence. Complimentary changes were made to preserve the theoretical secondary structures of the packaging and dimerisation signal (Harrison and Lever, 1992). When this work began there were very few reports of HIV-1 derived gene transfer vectors and no systems specifically designed for gene therapy applications. The challenge was to put together a vector/packaging gene transfer system that worked and would achieve the ultimate aim of gene transfer to non-cycling cell populations. This was based on theoretical considerations of vector design and the knowledge that has been gained from the continued development of MLV based vectors which has established some basic principles for vector performance.

RESULTS AND DISCUSSION

3.2 β-GLUCURONIDASE VECTOR CONSTRUCTION

Three vectors were constructed including an MLV based vector to be used as a control, a hybrid HIV-1/MLV vector and an entirely HIV-1 based vector, all of which were designed to express the β-glucuronidase cDNA. The β-glucuronidase cDNA (2191 bp) was excised from pGEM4βgluc (Oshima *et al.*, 1987) with *Eco*RI, made blunt-ended by filling in with Klenow (section 2.2.4) and cloned into the *Stu*I site of pLX. The pLX vector is a simplified MLV vector derived from pLNCX (Miller and Rosman, 1989) but with both the neomycin resistance gene and cytomegalovirus (CMV) immediate early promoter removed. This created a MLV based retroviral vector for gene transfer of the lysosomal enzyme β-glucuronidase, pLXβgluc (Figure 3.1).

The initial rationale for the development of an HIV-1 based vector was to create a transfer vector that was independent of the viral protein Tat (described in section 1.3.1.4).

The proposal was to use a hybrid of MLV and HIV-1 that incorporated the U3 and R regions from the long terminal repeats (LTRs) of MLV, and a 390 bp sequence derived from HIV-1, that effectively replaced the TAR (trans-activation responsive) sequence, the site of action for Tat. This latter sequence extended from 60 bp 5' of the HIV-1 U5 integration signal and included, the primer binding site, dimerisation and packaging signals, and the 5' end of the *gag* coding sequence with a mutagenised ATG followed by a *Bam*HI site for the cloning of reporter genes (β-glucuronidase in this case), the polypurine tract and U3 integration signal. Additional changes encompassed mutation of the splice donor site sequence and two, second nucleotide alterations necessary to maintain the secondary structure of this region of the HIV-1 genome (Figure 3.1).

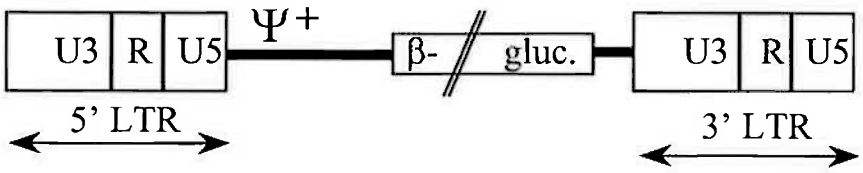
To construct this vector, regions from the long terminal repeat (LTR) sequences from Moloney leukemia virus (MLV) were amplified by PCR using pLNCX (Genbank accession number M28247) as template. Primers LTRMLVF1 and R1 were used for the 5' LTR segment (bases 3666 - 4189) which was cloned 5' *Not*1/3' *Spe*1 into pBluescript II SK (+). The 3' LTR region (bases 3679 - 4292), amplified with LTRMLVF2 and R2, was cloned 5' *Pst*1/3' *Hind*III. A 390 bp oligonucleotide fragment (described above and see Figure 3.1) was assembled from 12 overlapping oligonucleotides that was assembled in two halves. The 5' half encompassing oligonucleotides HIV1Fn, HIV2Fn, HIV1Rn, HIV2Rn and HIV3Rn was cloned 5' *Spe*I/3' *Bss*HII. Oligonucleotides HIV3Fn, HIV4Fn, HIV5Fn, HIV4Rn, HIV5Rn, HIV6Rn and HIV7Rn comprise the 3' end which was cloned 5' *Bss*HII/3' *Pst*I. For cloning, the oligonucleotides were heated to 95°C and then annealed by cooling slowly to room temperature. Both these fragments were cloned into pSP70BssHII, a modification of pSP70 containing a polylinker with a *Bss*HII restriction site. This was made by annealing and cloning two oligonucleotides, BssHIIF and BssHIIR, 5' *Spe*I/3' *Pst*I into pSP70. Several clones of this 390 bp fragment were sequenced but all had deviations from the predicted sequence. One clone was identified that contained only a single base difference from the expected sequence, a C residue in place of a G residue (see Figure 3.1). This base change was not considered likely to be significant as it was in the *gag* coding region, 3' of the major packaging signal sequence. Following sequencing the 390 bp fragment was excised 5' *Spe*I/ 3' *Pst*I and then cloned into the pBluescript II SK (+) vector that contained the MLV LTRs. Lastly, the β-glucuronidase cDNA was inserted as an *Eco*RI blunt-ended fragment from pGEM4βgluc into an end-filled *Bam*HI site, to create the final vector pHIV/MLVβgluc (Figure 3.1).

FIGURE 3.1

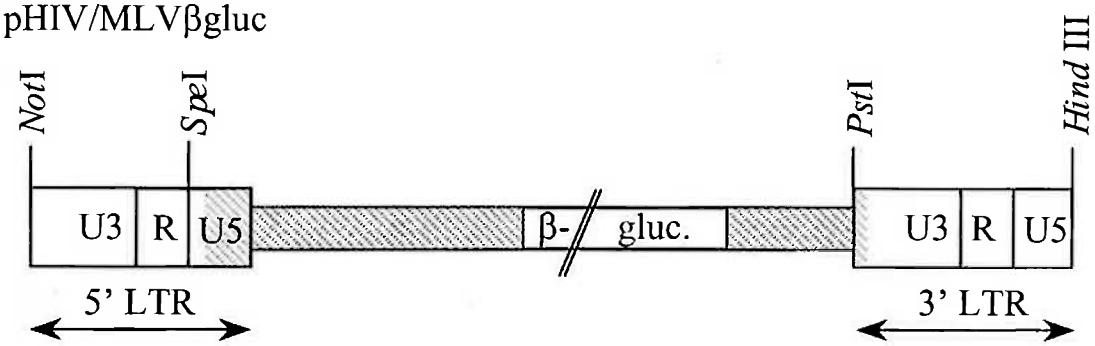
Diagrammatic representation of vectors expressing β-glucuronidase.

Detailed construction of the three vectors, pLXβgluc, pHIV/MLVβgluc and pHIVSV40βgluc is provided in section 3.2. The pLXβgluc vector contains the β-glucuronidase cDNA in the *Stu*I site of pLX, the extended packaging signal (Ψ+) is shown. The 390 bp HIV-1 sequence is shown in full underneath the three vector diagrams. The integration signal is shown in italics, the primer binding site is underlined, splice donor shown in bold, the major packaging signal is in bold and italicised and the polypurine tract is underlined and italicised. Additional nucleotide changes are shown in brackets. The change in the expected sequence from a G to a C is indicated in lower case (see text for details). The *gag* initiation codon mutagenised to TAG is shown is shown in bold. The open region of the LTRs in pHIV/MLVβgluc is derived from MLV and shaded regions denote HIV-1 sequence. The pHIVSV40βgluc vector contains LTRs from pHIV-YU2, thus have no MLV sequence.

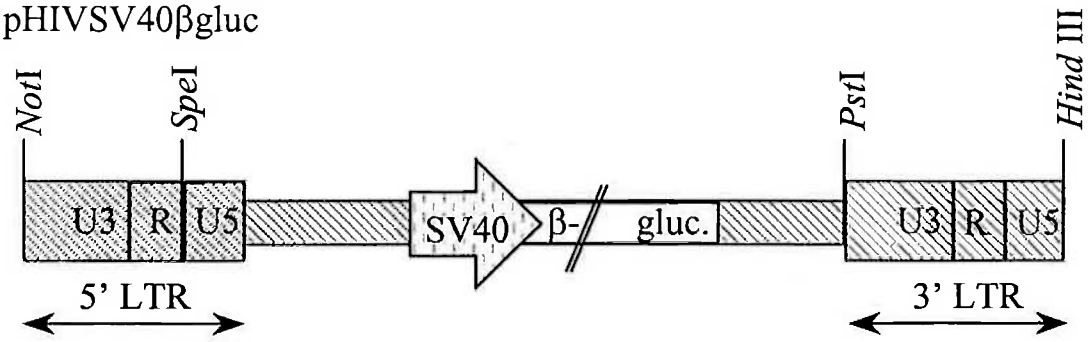
pLXβgluc



pHIV/MLVβgluc



pHIVSV40βgluc



5' CTAGTGTGTGCCCGTCTGTTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTA

GT*CAGTGTGGAAAATCTCTAGCA*GTGGCGCCCGAACAGGGACTTGAAAGCGAAAGCGAA
integration signal — primer binding site

AGGAAAACCAGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAA

GAGGCGAGGGGCGGCGACTG(C)T(C)AGTA(G)G(G)C***AAAAATTTTTGACTAGCGGA***
splice donor — packaging signal

GGC**TAG**AAGGAGAGAGTAGGGTGCGAGAGCGTCAGTATTAAGTGCGGGcGAATTAGATA
gag

AGTGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAACAATATAGATTAAAACATAGG
*Bam*HI

ATCC*AAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCC*AACTGCA3'
polypurine tract — integration signal

A third vector was created by replacing the LTRs of pHIV/MLVβgluc with HIV-1 LTRs and including the Simian Virus immediate early promoter (SV40) as an internal promoter to drive β-glucuronidase expression. This vector therefore contains no MLV sequences. The use of an internal promoter was necessary because the HIV-1 LTR will not function as a promoter without the *trans*-activating protein Tat. This (HIV-1 based) vector, pHIVSV40βgluc, was produced by amplification of the HIV-1 LTRs by PCR from pHIV-1 YU2. The 5' LTR was amplified with primers 5'LTRF and 5'LTRR and cloned into pBluescript II SK (+) 5'*Not*1/3'*Spe*I. Primers 3'LTRF and 3'LTRR were used to generate the 3' LTR which was cloned 5'*Bam*HI/3'*Cla*I. The same 390 bp oligonucleotide synthesised fragment from the above pSP70 clone, was then cloned 5' *Spe*I/3' *Pst*I between the two LTRs. A linker incorporating the restriction sites (5' to 3') *Bam*HI, *Stu*I, *Hind*III and *Eco*RI, made by annealing two oligonucleotides *Bam*HIF and *Bam*HIR, was then cloned into the *Bam*HI site to give pB/SHIVLTR. The SV40 promoter was removed *Bam*HI (5') and *Hind*III (3') from pLNSX (Miller and Rosman, 1989), end-filled, and then blunt-cloned into the end-filled *Bam*HI site of the vector to create pHIVSV40. Finally, the β-glucuronidase cDNA was excised from pGEM4βgluc with *Eco*RI and cloned into the vector, as already described, to create pHIVSV40βgluc (Figure 3.1).

3.3 HELPER PLASMIDS

The MLV amphotropic envelope protein was provided by penvAm (Markowitz *et al.*, 1988b). For the MLV *gagpol* sequence pcDNA3MLVgagpol (D. Anson, unpublished) was used. The HIV-1 *gagpol* sequence was isolated from pHIV-YU2 as a 5' *Hae*II/3' *Nde*I fragment (bases 638 to 5122), made blunt-ended with Klenow (section 2.2.4) and cloned into the *Eco*RV site of pcDNA3.1. Correct orientation of the *gagpol* insert was determined by restriction enzyme analysis. This construct was named pcDNA3HIV10gagpol.

The *tat* (exon 1), *vif*, *vpr*, *vpu* and *nef* reading frames were isolated by PCR using primers encompassing the ATG initiation codon and the relevant stop codon for each protein from pHIV-YU2 (Genbank accession number M93258) with no extraneous (ie 5' or 3' of these codons, respectively) HIV-1 sequence included. A Kozak translation initiation consensus sequence (GCCACC) was added upstream of the initiating methionine. For the first exon of *tat* (bases 5831 - 6049 of pHIV-YU2), the primer pair tatF and tatR, flanked by a 5' *Eco*RI and a 3' *Xba*I restriction site was used. The PCR product was then cloned as an *Eco*RI/*Xba*I fragment into pcDNA3.1 to create pcDNA3tat.

In the first instance bicistronic expression vectors were made for two pairs of the secondary proteins in each plasmid. For *vif* (pHIV-YU2 bases 5039-5617) the primer pair used was vifF and vifR to create a PCR product that was cloned into pBluescript II SK (+) as an *Eco*RI/*Bam*HI fragment, to create pB/Svif. The primers for *vpu* (pHIV-YU2 bases 6062-6307) were vpuF and vpuR and the resulting PCR product was cloned *Nco*I/*Not*I into pB/Svif. To provide an ATG initiation codon for *vpu*, base 6062 was changed from a C in the pHIV-YU2 sequence to an A. The IRES (internal ribosome entry site) sequence was amplified from pCITE4a using primers IRESF and IRESR, both of which incorporate a *Bam*HI site. This PCR product was then cloned into the *Bam*HI site of pB/Svifvpu and the orientation of the IRES fragment was determined by *Apa*I/*Not*I digestion. The entire vifIRESvpu sequence was then restricted from this clone with an *Eco*RI/*Not*I digest and transferred into the same two sites of pcDNA3.1 thus producing the final expression plasmid pcDNA3vifIRESvpu.

The *nef* sequence (bases 8758-9402 of pHIV-YU2) was amplified using primers nefF and nefR to generate a PCR product that was cloned *Nco*I/*Not*I into pBluescript II SK (+) to give pB/Snef. The *vpr* reading frame (pHIV-YU2 bases 5557-5850) was also amplified by PCR with primers vprF and vprR and cloned into the *Pst*I/*Bam*HI sites. The IRES sequence was then cloned into the *Bam*HI site as described above. The transcription unit encompassing nefIRESvpr was then restricted with *Pst*I, 3' end-filled, then digested with *Not*I and cloned into the *Eco*RV and *Not*I sites of pcDNA3.1 to give the final expression vector pcDNA3nefIRESvpr. All sequences isolated by PCR were sequenced to confirm their fidelity.

3.4 TRANSDUCTION OF MPS VII FIBROBLASTS

The three vectors described in section 3.2, along with the appropriate helper plasmids (section 3.3), were transfected into 293T cells using Lipofectamine as described in section 2.2.17. For every transfection 1 μg of each plasmid was transfected. Medium from the transfected 293T cells was collected and used to transduce both human MPS VII skin fibroblasts (SF382) and immortalised murine MPS VII fibroblasts (m3521). To assess the efficiency of the two HIV-1 based vectors compared to the conventional MLV vector, transduced MPS VII fibroblasts were stained histochemically 72 hr after transduction as described in section 2.2.19 for the presence of β-glucuronidase. Low amounts of enzyme activity resulted in the cytoplasm staining pink grading through to a

deep red colour for larger amounts of β-glucuronidase. Pink/red coloured cells were counted.

Table 3.1 shows that the MLV system worked well and successfully transduced the target cells (immortalised murine MPS VII fibroblasts (m3521)) as noted by the dark red cytoplasmic staining. In contrast, m3521 cells transduced with the HIV-1 based vector (pHIVSV40βgluc) only stained a weak pink colour. The hybrid MLV/HIV vector was clearly negative as evidenced by a lack of pink colour seen in the cytoplasm of any of the m3521 cells. The plasmid combinations used in Table 3.1 were also used to transduce human MPS VII skin fibroblasts (SF382) and Figure 3.2 shows an example of the histochemical staining of SF382 that have been transduced with the HIV-1 derived vector as compared with some untreated SF382. The results from the transductions of SF382 have not been tabulated but support those presented in Table 3.1.

The inclusion of expression constructs for the secondary proteins in the transfection did not seem to have any positive effect on the number of MPS VII fibroblasts that expressed β-glucuronidase. The number of transduced MPS VII fibroblasts actually decreased in the presence of the secondary proteins from 32 to 8 positive cells counted in 10 squares. This negative effect observed with the secondary proteins may be a consequence of the transfection conditions. Due to the fact that 1 μg of each plasmid was used in every transfection inclusion of the secondary protein constructs increased the total amount of DNA transfected by 2 μg.

Of particular interest was the 13 positive cells seen when pHIVSV40βgluc was transfected with the *gagpol* sequence derived from MLV. Upon the inclusion of an expression construct for Tat, the number of transduced MPS VII fibroblasts increased to 29 with the MLV *gagpol* sequence. This was unexpected because it was not thought possible that the structural and enzymatic retroviral components from MLV could package an HIV-1 genome. However, it later became obvious that there was an alternative explanation for this observation other than transduction (see the following section).

3.5 BIOLOGICAL EVALUATION

The minimal plasmid combination required for detectable transduction as measured by β-glucuronidase staining in MPS VII fibroblasts, consisting of pHIVSV40βgluc, pcDNA3tat, pcDNA3HIV10gagpol and penvAm was tested for the ability to correct both the enzyme defect and storage of glycosaminoglycans in human MPS VII skin fibroblasts (SF382) and immortalised murine MPS VII fibroblasts (m3521). As no positive influence

could be attributed to the inclusion of the secondary protein expression constructs, these were omitted. Medium from 293T cells transfected with the plasmid combinations shown in Table 3.2 was used to transduce murine and human MPS VII fibroblasts that had been metabolically labelled with $^{35}SO_4$ as described in section 2.2.20. Studies with cultured fibroblasts from patients with a variety of MPS disorders have shown that active enzyme inside the cell is transported to the lysosome and results in the normalisation of substrate turnover (Neufeld and Muenzer, 1995).

It was evident that storage could be cleared from the fibroblasts as shown by enzyme activities and the reduction in ^{35}S counts in the cell lysates of the target cells (Table 3.2). The HIV-1 derived vector pHIVSV40βgluc seemed to be more effective than the MLV vector pLXβgluc. The amount of ^{35}S storage in human MPS VII fibroblasts was reduced from 2937 cpm to 225 cpm with pHIVSV40βgluc. This correlated well with the amount of labelled material in the control skin fibroblasts (118 cpm ^{35}S). However, it was also shown that medium removed from 293T cells that had been transfected with only the vector (pHIVSV40βgluc), without any packaging functions, contained β-glucuronidase activity (data not shown). Sufficient β-glucuronidase produced from pHIVSV40βgluc was taken up from the medium of the transfected 293T cells allowing enzymatic correction of MPS VII fibroblasts. Consequently β-glucuronidase was detected, and radiolabelled storage material was cleared from the target cells. Furthermore these levels were analogous to those seen with MPS VII fibroblasts transduced with medium from 293T cells transfected with the vector and the helper plasmids (Table 3.2).

It is well known that lysosomal enzymes, including β-glucuronidase can enter cells *via* mannose-6-phosphate receptors and correct the disease phenotype present in such cells (Neufeld and Muenzer, 1995, Lagunoff *et al.*, 1973). Enzyme replacement therapy for these disorders was founded on such a concept. Reappraisal of the results shown in Table 3.1 and Figure 3.2, suggests that all of the histochemical staining present in the MPS VII fibroblasts can be attributed to β-glucuronidase expressed from the vector, pHIVSV40βgluc, in the 293T cells used for virus production. The LTR of MLV was unable to produce enough β-glucuronidase from the hybrid MLV/HIV vector to enable enzymatic correction of MPS VII fibroblasts. In contrast, the very dark histochemical staining and the numerous numbers of MPS VII fibroblasts that stained dark red are suggestive of true retroviral mediated gene transfer from the MLV system. If the HIV-1 based system had worked, enzyme uptake *via* the MPS VII fibroblasts would not have posed an insurmountable problem because the expected observation from gene transfer would be dark red staining as seen with the MLV system. This would allow subjective

TABLE 3.1

Histochemical staining of transduced MPS VII fibroblasts.

Viral supernatant from 293T cells transfected with the following plasmid combinations was used to transduce murine MPS VII fibroblasts (m3521) as described in sections 2.2.18. Positive cells were visualised histochemically by staining for β-glucuronidase as described in section 2.2.19 under a light microscope (40 x magnification). Cells with pink to red cytoplasm were counted and recorded. In cases of few positive cells 10 squares of the 35 mm grid dish were counted as indicated above. (n = 3).

Plasmids transfected	Positive cells counted
pLXβgluc pcDNA3MLVgagpol penvAm	too many positive cells to count, at least 50% of the cells in each square were positive (dark red cytoplasm)
pHIV/MLVβgluc pcDNA3MLVgagpol penvAm	negative
pHIVSV40βgluc pcDNA3MLVgagpol penvAm	13 ± 3 positive cells in 10 squares (pale pink)
pLXβgluc pcDNA3MLVgagpol penvAm pcDNA3tat	too many positive cells to count, 25% of the cells in each square were positive (dark red cytoplasm)
pHIV/MLVβgluc pcDNA3MLVgagpol penvAm pcDNA3tat	negative
pHIVSV40βgluc pcDNA3MLVgagpol penvAm pcDNA3tat	29 ± 4 positive cells in 10 squares (pale pink)
pLXβgluc pcDNA3HIV10gagpol penvAm	negative
pHIV/MLVβgluc pcDNA3HIV10gagpol penvAm	negative
pHIVSV40βgluc pcDNA3HIV10gagpol penvAm	3 ± 2 positive cells in 10 squares (pale pink)

Plasmids transfected	Positive cells counted
pLXβgluc pcDNA3HIV10gagpol penvAm pcDNA3tat	negative
pHIV/MLVβgluc pcDNA3HIV10gagpol penvAm pcDNA3tat	negative
pLXβgluc pcDNA3HIV10gagpol penvAm pcDNA3viflRESvpu pcDNA3neflRESvpr	negative
pHIV/MLVβgluc pcDNA3HIV10gagpol penvAm pcDNA3viflRESvpu pcDNA3neflRESvpr	negative
pHIVSV40βgluc pcDNA3HIV10gagpol penvAm pcDNA3viflRESvpu pcDNA3neflRESvpr	negative
pLXβgluc pcDNA3HIV10gagpol penvAm pcDNA3tat pcDNA3viflRESvpu pcDNA3neflRESvpr	negative
pHIVSV40βgluc pcDNA3HIV10gagpol penvAm pcDNA3tat	32 ± 4 positive cells in 10 squares (pale pink)
pHIVSV40βgluc pcDNA3HIV10gagpol penvAm pcDNA3tat pcDNA3viflRESvpu pcDNA3neflRESvpr	8 ± 2 positive cells in 10 squares (pale pink)

FIGURE 3.2

Transduced MPS VII fibroblasts.

The top photograph depicts human MPS VII skin fibroblasts (SF382) stained histochemically for the lysosomal enzyme β-glucuronidase as described in section 2.2.19. Medium from 293T cells transfected with pHIVSV40βgluc, pcDNA3HIV10gagpol, penvAm and pcDNA3tat was used to transduce the MPS VII fibroblasts. The fibroblasts were stained for β-glucuronidase activity 72 h following transduction.

The bottom photograph shows SF382 stained for β-glucuronidase activity that had been transduced with medium from untransfected 293T cells (negative control).

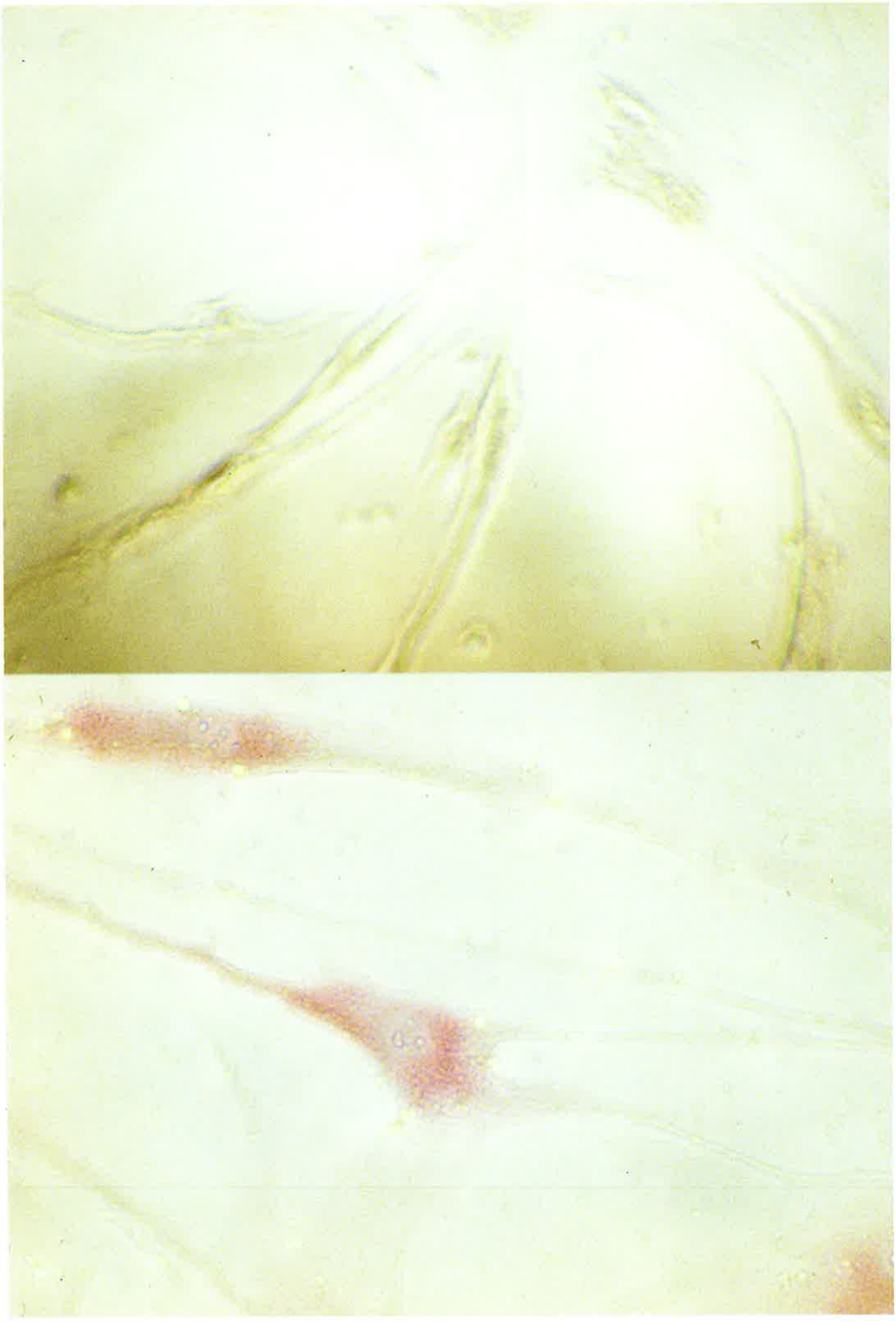


TABLE 3.2

Labelled storage material and β-glucuronidase activity in transduced enzyme-deficient fibroblasts.

Human MPS VII skin fibroblasts (SF382) and immortalised murine MPS VII fibroblasts (m3521) were metabolically labelled with $^{35}SO_4$ as described in section 2.2.20. Viral supernatant from 293T cells transfected with the above plasmid combinations was used to transduce these fibroblasts as described in sections 2.2.18. Cell lysates prepared from the transduced cells were then counted for ^{35}S containing storage material and assayed for β-glucuronidase activity as described in section 2.2.20. Where possible, control fibroblasts (SF4681) were included for comparative purposes.

Plasmids transfected	Cell line transduced	^{35}S cpm	β-glucuronidase nmol/min/ml
none	Control skin fibroblasts (SF4681)	118	12.5
none	m3521	35558	0.1
none	SF382	2937	0.03
pHIVSV40βgluc pcDNA3tat pcDNA3HIV10gagpol penvAm	SF382	225	7.4
pHIVSV40βgluc pcDNA3tat pcDNA3HIV10gagpol penvAm	m3521	1304	5.0
pHIVSV40βgluc	SF382	298	5.8
pLXβgluc pcDNA3MLVgagpol penvAm	SF382	323	2.3
pLXβgluc pcDNA3MLVgagpol penvAm	m3521	2912	22.3

identification of transduced cells on the assumption that the SV40 promoter or MLV/HIV hybrid LTR are as transcriptionally active as the MLV LTR.

Any virus particles that may possibly have been produced in the HIV-1 vector system could conceivably have been masked by the β-glucuronidase in the medium from the 293T cells. For this reason reverse transcriptase (RT) measurements were performed on all the viral supernatants used in Table 3.1 by monitoring the incorporation of radioactive dTTP as described in section 2.2.13. RT was only detected with the MLV system (^{32}P incorporation ranged from 3700 to 4500 cpm with pcDNA3MLVgagpol, compared to 120 to 300 cpm with pcDNA3HIV10gagpol, the latter being indistinguishable from background). Therefore, there is no evidence that pcDNA3HIV10gagpol is functional as one of the gene products, reverse transcriptase, could not be detected.

3.6 CONCLUDING DISCUSSION

The lysosomal storage disorder, MPS VII, theoretically provided an attractive experimental system. Viral titre can be analysed by histochemical staining for β-glucuronidase in transduced cells and gene expression of the enzyme can be measured by a simple fluorimetric enzyme assay. Gene therapy for MPS VII is biologically relevant, and it is easy to measure correction of MPS VII cells by the clearance of ^{35}S labelled substrate storage. However, the expression and transfer of β-glucuronidase from the 293T cells as observed, was not anticipated.

The SV40 promoter is inactive in 293T cells due to trans-repression by large T antigen (Dubridge *et al.*, 1987). Therefore in the pHIVSV40βgluc vector, expression of β-glucuronidase must be from the LTR and sufficient to enable enzyme transfer from the 293T cells *via* the medium to the MPS VII fibroblasts. This allowed enzymatic correction of the MPS VII fibroblasts which mimics the situation that would occur in enzyme replacement therapy. The hybrid MLV/HIV-1 vector, pHIV/MLVβgluc, did not seem to express β-glucuronidase at levels high enough to be transferred to the MPS VII fibroblasts *via* the medium, as evidenced by the lack of histochemical staining. Nevertheless, the fact that it may have expressed the reporter gene, albeit at weaker levels, cannot be ruled out.

The MLV system appeared to be working because of the brightly stained MPS VII fibroblasts. There is still the possibility that this is a consequence of enzyme transfer because virus was not directly demonstrated in the supernatant from the 293T cells. However, reverse transcriptase measurements show that virus particles were in fact being produced. This would be anticipated with such a well-established and efficient system.

The lack of the viral enzyme reverse transcriptase in the medium of the transfected 293T cells with either the pHIV/MLVβgluc or pHIVSV40βgluc vectors, together with any of the combinations of helper plasmids, suggested no viral particles were produced. Alternatively, perhaps very low numbers of virus particles were produced but below limits of detection for reverse transcriptase activity. Some of the transfected 293T cells transfected with the minimal plasmid combination (section 3.5) were examined under electron microscopy to see if there was any evidence of viral particles budding from the surface of the transfected 293T cells. The electron micrographs were hard to interpret but no virus-like particles could be identified (data not shown). This avenue of experimentation was not pursued.

There are several possible reasons as to why the HIV-1 vector system was not working. Pseudotyping the HIV-1 particles should not be the cause of the problem as this has been successfully shown previously (Spector *et al.*, 1990). The HIV-1 LTR seems to be operational, at least in terms of its promoter function, as demonstrated by efficient β-glucuronidase expression. The hybrid MLV/HIV-1 LTR seems to be less efficient although it remains unknown as to whether or not it is detrimental to virus production. The absence of reverse transcriptase in the medium indicates that the pcDNA3HIV10gagpol plasmid is unable to express and/or process the GagPol polyprotein protein efficiently. The two bicistronic expression plasmids (pcDNA3vifIRESvpu and pcDNA3nefIRESvpr) were not tested for their ability to express any of the secondary proteins of HIV-1. One or more of the helper plasmids may simply not be expressing a viral protein that is essential for virus production. Alternatively, there may be a fault in a regulatory or *cis* function of the virus.

Limited sequence analysis performed on the pHIV-YU2 clone, has identified some polymorphisms (data not shown). This has raised the question as to whether or not the pHIV-YU2 clone has been misidentified. To answer this, the pHIV-YU2 clone should be tested for replication competency. This could not be done however, because the facilities for such an experiment were not available. In conclusion, from the experiments conducted in this chapter it is not possible to ascertain what component of the virus production system is preventing the formation of virus particles. The results in Chapter four suggest that the *gagpol* sequence is functional, and probably both the vector and packaging system were inefficient, resulting in very low viral titres that could not be detected with this experimental system.

In July 1996 Naldini *et al.* published an HIV-1 based gene transfer system that successfully transduced non-dividing cells. The packaging plasmid pCMVΔRnr (Kafri *et*

al., 1997; Naldini *et al.*, 1996b) used in this study was obtained to eliminate the many variables of the packaging system described in this chapter. It should be noted that, at this point, it would have been possible to test both the pHIV/MLVβgluc and pHIVSV40βgluc vectors with the pCMVΔRnr packaging plasmid. Due to the other limitations of the MPS VII system outlined in this chapter, it was decided to pursue a simpler system. Such an approach was aimed at developing and assessing the helper and vector virus functions independently. Furthermore, a more versatile and robust envelope protein from the vesicular stomatitis virus G glycoprotein (VSV-G) provided by pHCMV-G (Yee *et al.*, 1994) was used. Using pCMVΔRnr and pHCMV-G a direct assessment of vector function could be made without the complicating factors of all the helper plasmids. The vector was further simplified by inserting the EYFP (enhanced yellow fluorescence protein) coding sequence as the reporter gene, enabling transduction to be measured by FACScan analysis. The evaluation of a packaging system for HIV-1 based vectors is described in Chapter four and different vector constructions in Chapter five.

CHAPTER FOUR

HELPER PLASMIDS FOR HIV-1 DERIVED VECTORS

4.1 INTRODUCTION

The approach taken here for engineering HIV-1 (human immunodeficiency virus type 1) based packaging plasmids and cell lines has been modelled on that for MLV (Moloney leukemia virus) (Cosset *et al.*, 1995; Rigg *et al.*, 1996). Gene delivery systems based on MLV have been designed to enable the production of helper free recombinant virus containing a therapeutic gene. This has been made possible by the separation of the *cis* and *trans* genetic functions of the retrovirus into a recombinant vector plasmid and constructs expressing the packaging functions (Morgan *et al.*, 1993). The rationale has been one of enhanced safety with maximal separation of the *trans* functions onto different plasmids, by minimising sequence homology between the plasmids used to make the virus, and with endogenous viral sequences in the cell lines used for virus production. Therefore, the probability of recombination events occurring between the constituents of the virus producing system, leading to the production of unwanted replication competent virus, is reduced. The first generation of MLV packaging cell lines were susceptible to the production of replication competent virus by the mechanisms suggested above (Miller, 1990; Lynch and Miller, 1991; Otto *et al.*, 1994). Later versions have minimised homology between all constituents of the system and require at least three non-homologous recombination events to produce replication competent virus (Markowitz *et al.*, 1988a; Rigg *et al.*, 1996). Such systems have been shown to be highly resistant to the production of replication competent virus even under conditions designed to facilitate their generation.

The continued development of MLV packaging and vector systems has established the basic principles of a workable design as well as some criteria for maximal safety. This has led to the development of HIV-1 based gene delivery systems (Naldini *et al.*, 1996a), as well as some other lentiviruses such as HIV-2 (Sadaie *et al.*, 1998) and feline immunodeficiency virus (Poeschla *et al.*, 1998b). However, due to the complexity of the genome, the construction of HIV-1 based packaging cell lines is technically more difficult than MLV. Furthermore, the pathogenic nature of HIV-1 raises heightened safety concerns. Following in the light of MLV, HIV-1 based packaging systems have evolved through successive alterations to the original proviral clone (Carroll *et al.*, 1994; Reiser *et al.*, 1996; Corbeau *et al.*, 1996; Naldini *et al.*, 1996a; Srinivasakumar *et al.*, 1997). This has included the use of heterologous promoters and polyadenylation signals, deletions of the proposed packaging signal and non-essential coding sequences such as *env* and the accessory proteins. Attempts have also been made to remove particular HIV-1 genes and

substitute them with heterologous sequence elements that can function in their place (Kim *et al.*, 1998; Gasmi *et al.*, 1999; Srinivasakumar and Schuening, 1999). All these modifications have been aimed at preventing the generation of replication competent virus.

Although transient transfections are now relatively common for the production of lentiviral vectors, such methods limit the amount of vector particles that can easily be produced. Bio-safety is further compromised in transient expression systems because recombination events appear more likely than in stable packaging cell lines. Ideally, stable HIV-1 based packaging cell lines, providing the viral proteins in *trans,* are needed. Like first generation MLV systems, a common problem is that stable HIV-1 packaging lines have used a single partially defective HIV-1 genome to express *gag*, *pol* and the accessory and regulatory proteins (Carroll *et al.*, 1994; Corbeau *et al.*, 1996; Kafri *et al.*, 1999). This design presents unacceptable bio-safety risks because the packaging constructs are too close to the parental virus. Additional problems have been low vector titres and in most cases because the cognate viral envelope was used, tropism has been limited to CD4+ cells (Yu *et al.*, 1996; Kaul *et al.*, 1998). More recently a multiply attenuated, inducible HIV-1 derived packaging construct expressing only Rev and GagPol, pseudotyped with the vesicular stomatitis virus G envelope, has been created (Klages *et al.*, 2000). This inducible cell line yielded titres comparable with transient systems for approximately 1 week after induction.

This chapter is devoted to the development of a packaging system for HIV-1 based vectors. Most of the HIV-1 derived vector production systems described to date, consist of the co-transfection of three plasmid constructs (i) a packaging plasmid containing all HIV-1 genes except the *env* gene, (ii) a plasmid expressing the G glycoprotein of vesicular stomatitis virus (VSV-g) to confer a broad host range to the vector, and (iii) an HIV-1 based vector containing the gene to be transferred to target cells. The aim of this work was to develop a helper cell line for HIV-1 based vectors (described in Chapter five) designed primarily to maximise bio-safety but also to optimise the titre obtained from such vectors. Initially the creation of a stable inducible packaging cell line was attempted to provide the HIV-1 viral proteins required in *trans*. However, this proved unsuccessful. Subsequently a transient system was developed by dissecting the HIV-1 genome into individual plasmids to provide all the necessary *trans* functions of the virus. Individual expression constructs were made for the expression of *gagpol*, *tat* (exon 1), *vif*, *vpr*, *vpu* and *nef*. For each of these plasmids a minimal open reading frame was linked to heterologous transcription regulatory sequences. Rev was provided by pCMVrev (Lewis *et al.*, 1990) in which the Rev protein is expressed from the cytomegalovirus (CMV) promoter.

In order to maximise expression and reduce sequence homology between the vector and packaging constructs the codon usage of *gagpol* was modified. The entire *gagpol* open reading frame was resynthesised using optimised mammalian codon usage. HIV-1 uses rare codons for the majority of its amino acid sequence, which may be important for temporal regulation of gene expression. These rare codons probably result in RNA instability and thus poor expression in the absence of Rev/Rev response element (RRE). By optimising the DNA sequence to use preferred codons for the respective amino acids it was hoped that expression of the GagPol protein may be increased and the dependence on Rev/RRE removed. Codon-optimisation will also minimise sequence homology with the vector sequence in which 5' *gag* sequence is deemed necessary for efficient packaging (McBride *et al.*, 1997). The idea of maximal separation of the *trans* functions was then extended by attempting to express the two polyproteins encoded by the *gagpol* transcriptional unit *via* two separate plasmids, again using codon-optimised sequence. This was unsuccessful primarily because the *pol* encoded proteins need to be provided in the form of a *gagpol* precursor molecule. Therefore two codon-optimised plasmids for the expression of Gag and GagPol were created. The helper functions for virus production were able to be provided from these separate codon-optimised expression constructs meaning the *gagpol* translational frameshift sequence could be removed from the virus production system.

As outlined in Chapter one, Gag molecules must be generated in large numbers to serve as precursors to the structural proteins of the virus particles. The enzymes encoded by the *pol* gene are usually needed in smaller amounts. By using the same initiation codon in the same mRNA to express the *gag* and *pol* genes, co-regulated expression of these proteins is achieved. In general between 10 and 20 Gag molecules are synthesised for every molecule of GagPol. To produce the *gagpol* precursor the *gag* termination codon is bypassed, a process that occurs by one of two mechanisms. The first mechanism, used by MLV and related viruses is read-through suppression, whereby the *gag* termination codon is occasionally misread as a sense codon. Translation can then proceed past the stop codon and into the *pol* reading frame.

The second way to bypass the *gag* termination codon is known as ribosomal frameshifting. Ribosomal frameshifting is used by most retroviruses including HIV-1. The mechanism involves a ribosome to occasionally slip backward one nucleotide ie -1 in the 5' direction (-1 frameshift), during the translation of *gag*. As such, the ribosome leaves the *gag* reading frame, with its downstream termination codon, and shifts into an overlapping portion of the *pol* reading frame (Jacks *et al.*, 1988). It would appear from

mutational analysis of the frameshift region that the -1 frameshift occurs after a particular codon has been read (Jacks *et al.*, 1988; Wilson *et al.*, 1988). Following the translocation, the nucleotide in the third position of the last codon read in the original reading frame (0) becomes the first codon in the new reading frame (-1). This -1 frameshift signal is made up of two elements. A seven nucleotide region encompassing the last two codons in the original reading frame and the preceding nucleotide. This heptanucleotide is known as the slippage or shifty sequence and is where the ribosome actually shifts. This frameshift region is depicted in Figure 4.1. The second element is a downstream stem loop structure that functions as a stimulator of frameshifting.

To remove this framshift region and express only the GagPol precursor protein, a construct termed pcDNA3gagpolfusionml was made. In this construct the frameshift/overlap region was replaced with codon-optimised sequence encoding the GagPol precursor polypeptide. This effectively abolishes the regulated control of the GagPol precursor polyproteins and synthesis of the Gag precursor molecules. When pcDNA3gagpolfusioml was combined with a plasmid expressing only Gag, viral particles were produced at comparable levels to a *gagpol* plasmid containing a functional frameshift region.

To increase safety of vector systems based on HIV-1, most of the accessory genes have been deleted from the packaging construct (Zufferey *et al.*, 1997; Mochizuki *et al.*, 1998; Gasmi *et al.*, 1999), thereby minimising the risk of producing wild type HIV-1. Whilst the minor proteins are generally not considered essential for production of VSV-G pseudotyped recombinant virus, they have been shown to have a relatively small influence on virus production. For example, Srinivasakumar and Schuening (1999) reported that Vpr and the combination of Vpr and Vpu increased virus particle production. Additionally, Chinnasamy *et al.* (2000) have shown that the transduction of primary resting lymphocytes with HIV-1 based vectors requires the presence of the viral accessory proteins. In Chapter three bicistronic expression plasmids for the accessory proteins were included in the gene transfer system. There was however, no indication that any of the proteins were actually expressed from the constructs. In this chapter, individual expression constructs were made for the accessory proteins, evaluated by western blot analysis and assessed for their contribution in virus production.

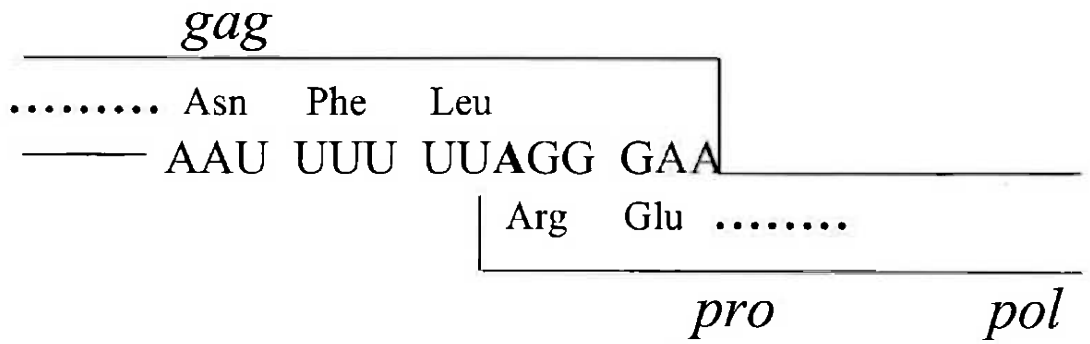


FIGURE 4.1

Frameshift suppression in the synthesis of HIV-1 GagPol.

The nucleotide sequences at the HIV-1 frameshift site and the amino acids that are encoded in the Gag and GagPol precursor molecule are shown. The amino acid sequence shown above the sequence is read by the *gag* reading frame and the amino acid sequence indicated below the sequence is read by the *gagpol* reading frame. The boxes demarcate the two reading frames for *gag* and *gagpol* and the A residue (shown in bold) represents the nucleotide that is read in both reading frames.

The chapter is divided into three parts dealing with (1) the core viral proteins contained within the *gagpol* gene, (2) the regulatory proteins and finally (3) the accessory proteins. For determinations of viral titre the vector used was pHIVext5SV40EYFPppt+RRE (see section 5.3) which transduces the enhanced yellow fluorescent protein (EYFP) under the transcriptional control of the SV40 early promoter. While this is not a fully optimised vector design it was the best one available at the time. HIV-1 viral particles were pseudotyped with VSV-G provided by pHCMV-G (Yee *et al.*, 1994). This envelope protein expands the host range of viral particles produced so as not to limit tropism to CD4+ cells as would be the case if the cognate HIV-1 envelope was used. An additional advantage is the robust property of VSV-G conferring stability to pseudotyped HIV-1 viral particles.

RESULTS AND DISCUSSION

PART ONE

4.2 STABLE CELL LINE FOR GAGPOL

A construct for the stable expression of *gagpol* was generated using a glucocorticoid inducible promoter. This promoter is contained in the pGRE5.1 expression plasmid and is composed of five high affinity glucocorticoid response elements placed upstream of the adenovirus 2 major late promoter "TATA" region and can be induced by dexamethasone (Mader and White, 1993). Firstly, the 5' region of the *gagpol* sequence was altered to reduce sequence homology with 5' *gag* sequence present in the vector construct. A 1.625 kb *Bgl*II fragment was isolated from pHIV-YU2 (bases 470 – 2095) (Li *et al.*, 1992) and cloned into pSP70. The first 300 bp of *gag* and 30 bp of the rat preproinsulin 5' untranslated region were synthesised as nine overlapping oligonucleotides (gag1F to gag5R). This *gag* sequence encoded the same amino acids for the HIV-1 *gag* gene with alternative codon usage (Australian National Genome Information Service (ANGIS) ecohigh.cod). This sequence was cloned 5' *Eco*RI/3' *Bsr*GI such that it replaced endogenous sequence up to the *Bsr*GI site (base 1049 in pHIV-YU2) in the 1.625 kb *Bgl*II *gagpol* fragment. The remaining 3' end of *gagpol* (bases 2095 – 5122) was isolated from pHIV-YU2 as a 5' *Bgl*II/3' *Nde*I fragment and cloned 5' *Bgl*II/3' *Afl*II to complete the *gagpol* coding sequence. This entire sequence was then cloned 5' *Eco*RI/3' *Not*I into pGRE5.1. The GRE promoter element, modified *gagpol* sequence and poly A site were

removed as an *Xba*I fragment and cloned into the *Spe*I site of pBluescriptpur^R (D. Anson, unpublished), to create pB/Spur^{R}GRE5.1gagpol.

pB/Spur^{R}GRE5.1gagpol was then transfected into both 293 and NIH3T3 cells using Lipofectamine and cells were selected with 2 μg/ml puromycin. Selection of NIH3T3 cells worked well, 48 clones were isolated and expanded into T25 flasks. The selection of 293 proved more difficult. Despite the construction of a killing curve, which identified a concentration of 2 μg/ml puromycin as sufficient to kill 293 cells, no transformants were identified. Lowering the concentration of puromycin to either 1.5 μg/ml or 1 μg/ml was insufficient to kill non-transfected 293 cells. Consequently transfected 293 cells were selected for five days in 2 μg/ml puromycin and then the puromycin was removed and the cells allowed to recover. Using this method, 20 clones of puromycin resistant 293 cells were isolated.

To induce expression of *gagpol*, 25 nM dexamethasone (final concentration) was added to the culture medium of the 48 pur^R NIH3T3 clones and the 20 293 pur^R clones. The cells were induced for 48 h after which reverse transcriptase assays were performed on 10 μl samples of the culture medium and filter papers were counted for ^{32}P incorporation (sections 2.2.13 and 14). For a positive control, medium from PA317 cells, which is a stable packaging cell line for MLV vectors, was used with and without exposure to dexamethasone. The results in Table 4.1 show that no reverse transcriptase activity was detected in any of the pur^R NIH3T3 or 293 clones, in contrast to the PA317 cell line. It was interesting to observe that addition of dexamethasone to the PA317 cell line attenuated the reverse transcriptase activity by approximately 58 %, although no obvious cell damage was observed. The PA317 cells growing in 25 nM dexamethasone appeared to display the same phenotype and growth characteristics as those growing in the absence of dexamethasone.

A PCR assay was used to assess the presence of any *gagpol* plasmid sequence in the transfected 293 and NIH3T3 clones. For each of the clonal cell lines, the cells were harvested and counted. One thousand cells were incubated overnight in 1 x PCR buffer in the presence of 30 μg/ml proteinase K at room temperature. The proteinase K was then heat inactivated at 95°C for 5 min prior to the PCR. Primers gagpolF and gagpolR were used to create a 1240 bp product and the PCR was performed as described in section 2.2.11.2. Figure 4.2 shows that only one of the 293 clones demonstrated a possible weak PCR product of the correct size. Photographs of agarose gel electrophoresis of PCRs on all the NIH3T3 clones and the remainder of the 293 clones, positive and negative controls

TABLE 4.1

Measurements of ^{32}P corresponding to reverse transcriptase activity.

Cell lines constructed in 293 and NIH3T3 cells transfected with the construct pB/Spur^{R}GRE5.1gagpol. Cells were selected in 2 μg/ml puromycin and induced for GagPol expression with 25 nM dexamethasone. Reverse transcriptase was measured in the supernatant and is expressed as ^{32}P counts per minute (cpm) detected on filter paper following the enzyme reaction (sections 2.2.13 and 2.2.14). PA317 stably expresses the Moloney leukemia virus (MLV) GagPol.

Clone type and number	^{32}P cpm
293 #1	813
293 #2	663
293 #3	675
293 #4	682
293 #5	647
293 #6	551
293 #7	533
293 #8	587
293 #9	558
293 #10	606
293 #11	997
293 #12	589
293 #13	578
293 #14	622
293 #15	678
293 #16	591
293 #17	813
293 #18	763
293 #19	826
293 #20	780

Clone type and number	^{32}P cpm
NIH3T3 #1	730
NIH3T3 #2	728
NIH3T3 #3	763
NIH3T3 #4	935
NIH3T3 #5	637
NIH3T3 #6	662
NIH3T3 #7	635
NIH3T3 #8	678
NIH3T3 #9	614
NIH3T3 #10	590
NIH3T3 #11	673
NIH3T3 #12	643
NIH3T3 #13	730
NIH3T3 #14	755
NIH3T3 #15	764
NIH3T3 #16	701
NIH3T3 #17	722
NIH3T3 #18	687
NIH3T3 #19	654
NIH3T3 #20	665
NIH3T3 #21	779
NIH3T3 #22	869
NIH3T3 #23	783
NIH3T3 #24	761
NIH3T3 #25	658
NIH3T3 #26	598
NIH3T3 #27	881
NIH3T3 #28	653
NIH3T3 #29	687
NIH3T3 #30	566
NIH3T3 #31	576
NIH3T3 #32	535

Clone type and number	^{32}P cpm
NIH3T3 #33	654
NIH3T3 #34	614
NIH3T3 #35	656
NIH3T3 #36	817
NIH3T3 #37	828
NIH3T3 #38	678
NIH3T3 #39	809
NIH3T3 #40	850
NIH3T3 #41	672
NIH3T3 #42	654
NIH3T3 #43	567
NIH3T3 #44	584
NIH3T3 #45	705
NIH3T3 #46	658
NIH3T3 #47	635
NIH3T3 #48	717
PA317	10211
induced PA317	4870
NIH3T3	689
induced NIH3T3	663
293	1021
induced 293	623
blank	151

(plasmid DNA only) are not depicted. A repeat reverse transcriptase measurement was performed on this one PCR positive 293 clone, but no activity was detected.

Induction by dexamethasone is possibly causing some problems and affecting the expression of reverse transcriptase especially in light of its action on the PA317 cells. Some doubt as to the integrity of the *gagpol* sequence could have been introduced *via* the modifications that were incorporated at the 5' end. Although the correct sequence had been confirmed there may be something in the nature of the 5' modification affecting transcription and/or translation of the GagPol polyprotein. Furthermore, as cited in Chapter three there was an element of doubt as to the integrity of the pHIV-YU2 clone (see section 3.6). Future experimentation in section 4.3 however, shows that the native gagpol sequence from the pHIV-YU2 clone does produce p24 and is able to package virus (see next section). Apart from the one possible 293 clone that looked to contain the *gagpol* sequence, the remaining negative PCR results suggest that selected clones were either artifactual or had deleted *gagpol* sequence. More analysis is clearly required to accurately determine this. One other avenue would have been to express the native HIV-YU2 *gagpol* sequence in an alternative inducible expression vector and use a different inducing agent such as tetracycline. Such a system has been described previously for the expression of the *trans* functions of HIV-1, albeit with low titres (Yu *et al.*, 1996). Nevertheless, for the reasons outlined above this did not seem worthy of pursuing at the time. Thus it was decided that the optimisation of helper and vector constructs in a transient system would be more feasible. Even though a stable packaging cell line for HIV-1 based vectors is the desired outcome, their development can hopefully be revisited following demonstration of an optimal system using transient expression.

4.3 TRANSIENT EXPRESSION OF GAGPOL

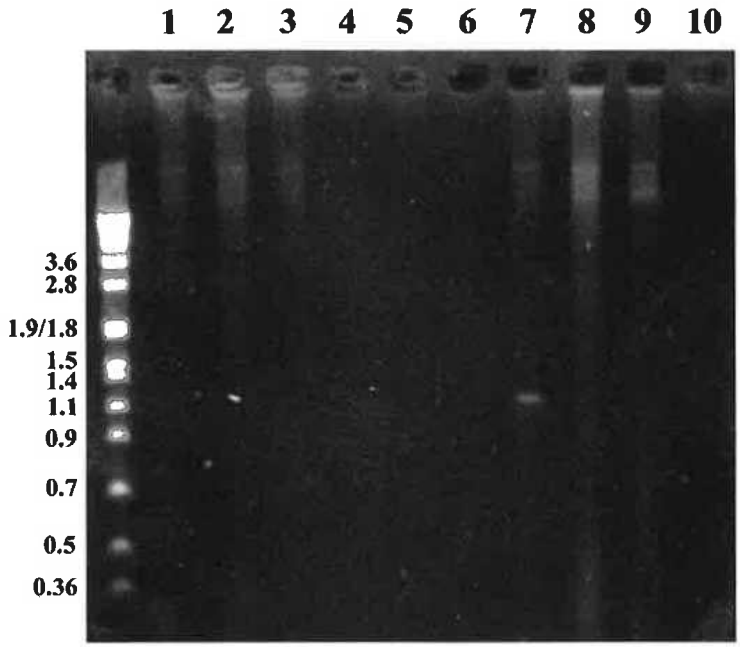
4.3.1 NATIVE SEQUENCE

Initially three constructs for the expression of *gagpol* were made. The native GagPol sequence encompassing bases 638 - 5122 from pHIV-YU2 was isolated as a 5' *Hae*II/3' *Nde*I fragment, end-filled with Klenow and cloned into the *Eco*RV site of pcDNA3.1, creating pcDNA3HIV10gagpol. Correct orientation of the *gagpol* sequence was determined by restriction enzyme analysis. This fragment contains the entire GagPol

FIGURE 4.2

Agarose gel electrophoresis of PCRs from 293 clones.

The *gagpol* expression construct pB/Spur^{R}GRE5.1gagpol, was transfected into 293 cells. Cells were selected in 2 μg/ml puromycin and induced for GagPol expression with 25 nM dexamethasone. Primers gagpolF and gagpolR were used to amplify a 1240 bp PCR product from one thousand cells. On the left of the figure are markers generated from pUC19/HpaII run on a 3 % (w/v) agarose gel. Lane 7 shows the presence of a weak PCR product at 1240 bp in one of the 293 clones. No PCR product is evident in any of the other lanes.


1 2 3 4 5 6 7 8 9 10
3.6
2.8
1.9/1.8
1.5
1.4
1.1
0.9
0.7
0.5
0.36

coding sequence, 153 bp of 5' non-coding sequence and 31 bp of 3' non-coding sequence. Two derivatives of this native sequence were then made. The 220 bp Rev response element (RRE) sequence was amplified by PCR from pHIV-YU2 using RREF and RRER. A 5' *Not*I and 3' *Xba*I site were introduced into this sequence *via* the PCR primers and the product was cloned into the same two sites of pcDNA3HIV10gagpol, placing it in the 3' non-coding region, to give pcDNA3HIV10gagpolrre. The last construct contained an intron sequence as well as the RRE. The intron sequence was isolated by PCR from signalpIG as a *Bam*HI/*Eco*RI fragment using the primers IN5' and IN3'. The resulting PCR product was then cloned into the *Bam*HI and *Eco*RI sites of pcDNA3HIV10gagpolrre, placing it in the 5' non-coding region, to produce pcDNA3HIV10gagpolrreintron.

4.3.2 CODON-OPTIMISATION

In order to produce a codon-optimised GagPol sequence the entire coding sequence for pHIV-YU2 GagPol was back-translated using optimal codon usage for mammalian gene expression (Australian National Genome Information Service (ANGIS) mam_h.cod) The only exception was from bases 2060 - 2294 of the YU2 sequence which encompasses the translational frameshift signal for expression of the GagPol polyprotein and the region of overlap between the *gag* and *pol* reading frames. A short 5' non-translated sequence containing a Kozak consensus sequence was added at the 5' end of the Gag coding sequence. The entire sequence was flanked by a 5' *Hind*III and a 3' *Xba*I restriction site, and cloned into pcDNA3.1 to produce pcDNA3gagpolml. The sequence was synthesised as 175 overlapping 50 base oligonucleotides with two 25 base oligonucleotides for the sequence overlap at both ends. The entire sequence is shown in Figure 4.3. The fragment was assembled as described in section 2.2.12. It was not possible to put together the entire sequence in one piece despite varying the assembly and amplification conditions.

Consequently it was decided to split the fragment into three separate segments, each to be assembled and then joined through restriction sites at each end. The first fragment, as defined by the 5' start, was 1.108 kb flanked by a 5' *Hind*III and a 3' *Bst*EII site (fragment 1), the middle fragment of 2.193 kb was defined by a 5' *Bst*EII and a 3' *Bam*HI site (fragment 2), and the final fragment of 1.315 kb completed the *gagpol* sequence beginning with a 5' *Bam*HI and ending with a 3' *Xba*I site (fragment 3). It was necessary to make some extra oligonucleotides to amplify each of the three fragments.

These are shown in Figure 4.3. Each fragment was assembled as described in section 2.2.12 and then cloned into pBluescript II SK (+) using the appropriate restriction sites.

The 1.108 kb fragment 1 was successfully assembled and amplified, and was cloned into the *Hind*III and *Bst*EII sites of pB/sBstEII. pB/SBstEII was created by cloning a *Bst*EII linker (using oligonucleotides BstEIIlinker F and R) into the *Eco*RI and *Bam*HI sites of pBluescript II SK (+). Fifteen clones of fragment 1 were sequenced with the pB/SF and pB/SR primers and one was found to be completely correct. Thus no further manipulation of fragment 1 was required.

Fragment 2 could not be isolated, despite manipulation of the assembly and amplification conditions, only small products were obtained. Therefore the fragment was split into two halves that were cloned and sequenced separately. Two primers amp215F and amp215R were made that span an internal *Kpn*I site dividing fragment 2 into a 1.108 kb 5' piece and a 1.085 kb 3' piece (see Figure 4.3). The 5' half was assembled and then amplified using amp2.5' and amp215R. The 1.108 kb product was cloned 5' *Bst*EII/3' *Kpn*I into pBluescript II SK (+) creating pB/Sfragment25'. The 3' half was assembled and then amplified using amp2.3' and amp215F and then cloned 5' *Kpn*I/3' *Bam*HI into pBluescript II SK (+), to produce pB/Sfragment23'.

Twelve clones of pB/Sfragment23' were sequenced but none of them were found to be correct. However, it was possible to obtain a correct sequence by joining the 5' end of one clone, pB/Sfragment23'#4, with the 3' end of another, pB/Sfragment23'#7. The errors in pB/Sfragment23'#4 included an extra G residue at position 2811, a T residue in place of a G residue at position 2660 and the omission of a C residue at position 2607. In pB/Sfragment23'#7 the mistakes identified were the omission of an A residue at position 2259, a G residue at position 2356 and a G residue at position 2609. The primers 260F and 260R (Figure 4.3) that were used for the amplification of these two fragments rectified the mistakes at both position 2607 and 2609. The 5' end of pB/Sfragment23'#4 was amplified with 10 cycles of PCR using primers 260R and pB/SF and the 3' end of pB/Sfragment23'#7 was amplified with 260F and pB/SR. These two PCR products were purified from an agarose gel and then placed together in another PCR. In this reaction the overlap between the two PCR products allows synthesis of a full-length sequence which could then be amplified using the external primers pB/SF and pB/SR. To enable the PCR products to denature and re-anneal creating a full-length template for amplification, the original set of conditions was modified. Cycle 1 was 95°C for 15 sec, 40°C for 30 sec with a 60 min ramp time and 72°C for 1 min. Cycle 2 was 95°C for 15 sec, 45°C for 30 sec with a 30 min ramp time and 72°C for 1 min, which was then followed by 8 cycles of 95°C

for 15 sec, 50°C for 30 sec and 72°C for 1 min. The reaction mix was as described in section 2.2.12 with the two individual PCR products added as template. The concentrations of these were estimated by agarose gel electrophoresis to enable the addition of approximately 100 ng of each to the reaction. The required fragment was purified away from other non-specific bands by agarose gel electrophoresis and cloned into the 5' *Kpn*I/3' *Bam*HI sites of pBluescript II SK (+). Six pB/S clones containing this fragment were sequenced and all were error-free.

Ten clones of pB/Sfragment25' were sequenced with pB/SF and pB/SR. All ten clones were found to have errors. However, again two clones were identified that could be aligned together to form a contiguous error-free sequence. The 5' end of pB/Sfragment25'#2 had correct sequence from the beginning of the fragment at the *Bst*EII site up to base 1400. This region was amplified with the primer pair pB/SF and 135F. The 3' end of pB/Sfragment25'#7 had no errors from position 1320 until the end of the clone at the *Kpn*I site and was amplified with pB/SR and 135R. These two PCR products were purified from an agarose gel and then added to a subsequent amplification reaction with the addition of the outer primers pB/SF and pB/SR and amplified with the same conditions as those described above for the 3' end of fragment 2. The required fragment was purified away from other non-specific bands by agarose gel electrophoresis and cloned into the 5' *Bst*EII/3' *Kpn*I/3' sites of pBluescript II SK (+). Six pB/S clones containing this fragment were sequenced, again all clones were free from errors.

For fragment 3 the full-length product of 1.315 kb was successfully isolated from the assembly and amplification reaction. The fragment was cloned into the *Bam*HI and *Xba*I sites of pBluescript II SK (+). Six clones were sequenced but all of them contained errors. However, two clones, pB/Sfragment3#5 and pB/Sfragment#8, were identified that could be joined by PCR to reconstitute the fragment free from errors. Sequencing pB/Sfragment3#5 revealed error-free sequence from the 5' end of the clone at the *Bam*HI site up to position 3750, with the exception of two base changes, a C residue in place of a G residue at position 3009 and a T in place of a G at position 3027. Both of these base changes were corrected with the sequence of the amplification primer 295F. The 5' end of clone 5 was amplified by 10 cycles of PCR using two primers 295F and 370R. For the pB/Sfragment3#8 error-free sequence was determined from position 3651 up to the 3' end at the *Xba*I site. The 3' end of clone 8, using 370F and R1 primers was also amplified by 10 cycles of PCR. These two PCR products were then mixed in equal concentrations and amplified with amp3.5N and pB/SR primers. The same amplification conditions as those

FIGURE 4.3

Complete sequence of codon-optimised *gagpol.*

The entire *gagpol* sequence was resynthesised as 175 overlapping oligonucleotides of 50 bases each, with the exception of the amplification oligonucleotides which are listed in section 2.1.9. The restriction sites that demarcate the three fragments that were cloned separately, sequenced and then aligned together to form a contiguous sequence are indicated below the italicised sequence. Internal primers used to put together fragments of the correct sequence are also illustrated. The frameshift region is shown in bold.

```
0     gggcccactcacagaagcttaaccatcagcaagcaggtcattgtgccac
                      HindIII
      MetGlyAlaArgAlaSerValLeuSerAlaGlyGluLeuAspLysTrpGl
1     ATGGGCGCCAGGGCCAGCGTGCTGAGCGCCGGCGAGCTGGACAAGTGGGA

      uLysIleArgLeuArgProGlyGlyLysLysGlnTyrArgLeuLysHisI
51    GAAGATCAGGCTGAGGCCCGGCGGCAAGAAGCAGTACAGGCTGAAGCACA

      leValTrpAlaSerArgGluLeuGluArgPheAlaValAspProGlyLeu
101   TCGTGTGGGCCAGCAGGGAGCTGGAGAGGTTCGCCGTGGACCCCGGCCTG

      LeuGluThrSerGluGlyCysArgGlnIleLeuGlyGlnLeuGlnProSe
151   CTGGAGACCAGCGAGGGCTGCAGGCAGATCCTGGGCCAGCTGCAGCCCAG

      rLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsnThrValAlaT
201   CCTGCAGACCGGCAGCGAGGAGCTGAGGAGCCTGTACAACACCGTGGCCA

      hrLeuTyrCysValHisGlnLysIleGluValLysAspThrLysGluAla
251   CCCTGTACTGCGTGCACCAGAAGATCGAGGTGAAGGACACCAAGGAGGCC

      LeuGluLysIleGluGluGluGlnAsnLysSerLysLysLysAlaGlnGl
301   CTGGAGAAGATCGAGGAGGAGCAGAACAAGAGCAAGAAGAAGGCCCAGCA

      nAlaAlaAlaAspThrGlyAsnSerSerGlnValSerGlnAsnTyrProI
351   GGCCGCCGCCGACACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCA

      leValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArg
401   TCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCAGCCCCAGG

      ThrLeuAsnAlaTrpValLysValValGluGluLysAlaPheSerProGl
451   ACCCTGAACGCCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCAGCCCCGA

      uValIleProMetPheSerAlaLeuSerGluGlyAlaThrProGlnAspL
501   GGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGCGCCACCCCCCAGGACC

      euAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet
551   TGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATG
```

```
     LeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisPr
601  CTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACAGGCTGCACCC

     oValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlyS
651  CGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGAGGGAGCCCAGGGGCA

     erAspIleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMet
701  GCGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATG

     ThrAsnAsnProProIleProValGlyGluIleTyrLysArgTrpIleIl
751  ACCAACAACCCCCCCATCCCCGTGGGCGAGATCTACAAGAGGTGGATCAT

     eLeuGlyLeuAsnLysIleValArgMetTyrSerProThrSerIleLeuA
801  CCTGGGCCTGAACAAGATCGTGAGGATGTACAGCCCCACCAGCATCCTGG

     spIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe
851  ACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTC

     TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMe
901  TACAAGACCCTGAGGGCCGAGCAGGCCAGCCAGGAGGTGAAGAACTGGAT

     tThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleL
951  GACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCC

     euLysAlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCys
1001 TGAAGGCCCTGGGCCCCGCCGCCACCCTGGAGGAGATGATGACCGCCTGC

     GlnGlyValGlyGlyProGlyHisLysAlaArgValLeuAlaGluAlaMe
1051 CAGGGCGTGGGCGGCCCCGGCCACAAGGCCAGGGTGCTGGCCGAGGCCAT

     tSerGlnValThrAsnSerAlaThrIleMetMetGlnArgGlyAsnPheA
1101 GAGCCAGGTGACCAACAGCGCCACCATCATGATGCAGAGGGGCAACTTCA
               BstEII
     rgAsnGlnArgLysThrValLysCysPheAsnCysGlyLysGluGlyHis
1151 GGAACCAGAGGAAGACCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCAC

     IleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGl
1201 ATCGCCAAGAACTGCAGGGCCCCCAGGAAGAAGGGCTGCTGGAAGTGCGG
                          ApaI
     yLysGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheL
                                                  pol  PhePh
1251 CAAGGAGGGCCACCAGATGAAGGATTGTACTGAGAGACAGGCTAATTTTT

     euGlyLysIleTrpProSerHisLysGlyArgProGlyAsnPheLeuGln
     eArgGluAspLeuAlaPheProGlnGlyLysAlaArgLysPheSerSerG
1301 TAGGGAAGATCTGGCCTTCCCACAAGGGAAGGCCAGGAAATTTTCTTCAG

     SerArgProGluProThrAlaProSerGluGluSerValArgPheGlyGl
     luGlnThrArgAlaAsnSerProIleArgArgGluArgGlnValTrpArg
1351 AGCAGACCAGAGCCAACAGCCCCATCAGAAGAGAGCGTCAGGTTTGGAGA
          135F and 135R
```

uGluThrThrThrProSerGlnLysGlnGluProIleAspLysGluLeuT
ArgAspAsnAsnSerLeuSerGluAlaGlyAlaAspArgGlnGlyThrVa
1401 AGAGACAACAACTCCCTCTCAGAAGCAGGAGCCGATAGACAAGGAACTGT

yrProLeuAlaSerLeuArgSerLeuPheGlySerAspProSerSerGln
lSerPheSerPheProGlnIleThrLeuTrpGlnArgProLeuValThrI
1451 ATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAA

***
leLysIleGlyGlyGlnLeuLysGluAlaLeuLeuAspThrGlyAlaAsp
1501 TAAAGATCGGCGGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCGCCGAC

AspThrValLeuGluGluMetAsnLeuProGlyArgTrpLysProLysMe
1551 GACACCGTGCTGGAGGAGATGAACCTGCCCGGCAGGTGGAAGCCCAAGAT

tIleGlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGlnIleP
1601 GATCGGCGGCATCGGCGGCTTCATCAAGGTGAGGCAGTACGACCAGATCC

roIleGluIleCysGlyHisLysAlaIleGlyThrValLeuValGlyPro
1651 CCATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGGT*GGGCCC*C
***Apa*I**
ThrProValAsnIleIleGlyArgAsnLeuLeuThrGlnIleGlyCysTh
1701 ACCCCCGTGAACATCATCGGCAGGAACCTGCTGACCCAGATCGGCTGCAC

rLeuAsnPheProIleSerProIleGluThrValProValLysLeuLysP
1751 CCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGC

roGlyMetAspGlyProLysValLysGlnTrpProLeuThrGluGluLys
1801 CCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAG

IleLysAlaLeuValGluIleCysThrGluMetGluLysGluGlyLysIl
1851 ATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGAT

eSerLysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleL
1901 CAGCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCA

ysLysLysAspSerThrLysTrpArgLysLeuValAspPheArgGluLeu
1951 AGAAGAAGGACAGCACCAAGTGGAGGAAGCTGGTGGACTTCAGGGAGCTG

AsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisPr
2001 AACAAGAGGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCC

oAlaGlyLeuLysLysLysLysSerValThrValLeuAspValGlyAspA
2051 CGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACG

laTyrPheSerValProLeuHisGluAspPheArgLysTyrThrAlaPhe
2101 CCTACTTCAGCGTGCCCCTGCACGAGGACTTCAGGAAGTACACCGCCTTC
***Kpn*I**
ThrIleProSerIleAsnAsnGluThrProGlyThrArgTyrGlnTyrAs
2151 ACCATCCCCAGCATCAACAACGAGACCCCC<u>GGCACCA*GGTACC*AGTACAA</u>
amp215F and amp 215R
nValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerM
2201 <u>CGTG</u>CTGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCAGCA

```
    etThrThrIleLeuGluProPheArgLysGlnAsnProAspLeuValIle
2251 TGACCACCATCCTGGAGCCCTTCAGGAAGCAGAACCCCGACCTGGTGATC

    TyrGlnTyrMetAspAspLeuTyrValGlySerAspLeuGluIleGlyGl
2301 TACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCA

    nHisArgThrLysIleGluGluLeuArgGlnHisLeuLeuArgTrpGlyP
2351 GCACAGGACCAAGATCGAGGAGCTGAGGCAGCACCTGCTGAGGTGGGGCT

    heThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMet
2401 TCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATG

    GlyTyrGluLeuHisProAspLysTrpThrValGlnProIleValLeuPr
2451 GGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGTGCTGCC

    oGluLysAspSerTrpThrValAsnAspIleGlnLysLeuValGlyLysL
2501 CGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGC

    euAsnTrpAlaSerGlnIleTyrAlaGlyIleLysValArgGlnLeuCys
2551 TGAACTGGGCCAGCCAGATCTACGCCGGCATCAAGGTGAGGCAGCTGTGC

    LysLeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGl
2601 AAGCTGCTGAGGGGCACCAAGGCCCTGACCGAGGTGATCCCCCTGACCGA
          260F and 260R
    uGluAlaGluLeuGluLeuAlaGluAsnArgGluIleLeuLysGluProV
2651 GGAGGCCGAGCTGGAGCTGGCCGAGAACAGGGAGATCCTGAAGGAGCCCG

    alHisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGln
2701 TGCACGGCGTGTACTACGACCCCAGCAAGGACCTGATCGCCGAGATCCAG

    LysGlnGlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGluProPheLy
2751 AAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAA

    sAsnLeuLysThrGlyLysTyrAlaArgThrArgGlyAlaHisThrAsnA
2801 GAACCTGAAGACCGGCAAGTACGCCAGGACCAGGGGCGCCCACACCAACG

    spValLysGlnLeuThrGluAlaValGlnLysIleAlaThrGluSerIle
2851 ACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGAGCATC

    ValIleTrpGlyLysThrProLysPheLysLeuProIleGlnLysGluTh
2901 GTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGAC

    rTrpGluThrTrpTrpThrGluTyrTrpGlnAlaThrTrpIleProGluT
2951 CTGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGT
          amp3.5N                              BamHI
    rpGluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGlu
3001 GGGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAG
                295F
    LysGluProIleIleGlyAlaGluThrPheTyrValAspGlyAlaAlaAs
3051 AAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAA

    nArgGluThrLysLeuGlyLysAlaGlyTyrValThrAsnLysGlyArgG
3101 CAGGGAGACCAAGCTGGGCAAGGCCGGCTACGTGACCAACAAGGGCAGGC
```

```
      lnLysValValSerLeuThrAspThrThrAsnGlnLysThrGluLeuGln
3151  AGAAGGTGGTGAGCCTGACCGACACCACCAACCAGAAGACCGAGCTGCAG

      AlaIleTyrLeuAlaLeuGlnAspSerGlyLeuGluValAsnIleValTh
3201  GCCATCTACCTGGCCCTGCAGGACAGCGGCCTGGAGGTGAACATCGTGAC

      rAspSerGlnTyrAlaLeuGlyIleIleGlnAlaGlnProAspArgSerG
3251  CGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAGGAGCG

      luSerGluLeuValSerGlnIleIleGluGlnLeuIleLysLysGluLys
3301  AGAGCGAGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAG

      ValTyrLeuAlaTrpValProAlaHisLysGlyIleGlyGlyAsnGluGl
3351  GTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCA

      nValAspLysLeuValSerAlaGlyIleArgLysValLeuPheLeuAspG
3401  GGTGGACAAGCTGGTGAGCGCCGGCATCAGGAAGGTGCTGTTCCTGGACG

      lyIleAspLysAlaGlnGluGluHisGluLysTyrHisSerAsnTrpArg
3451  GCATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCACAGCAACTGGAGG

      AlaMetAlaSerAspPheAsnLeuProProValValAlaLysGluIleVa
3501  GCCATGGCCAGCGACTTCAACCTGCCCCCCGTGGTGGCCAAGGAGATCGT

      lAlaSerCysAspLysCysGlnLeuLysGlyGluAlaMetHisGlyGlnV
3551  GGCCAGCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGG

      alAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGly
3601  TGGACTGCAGCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGC

      LysValIleLeuValAlaValHisValAlaSerGlyTyrIleGluAlaGl
3651  AAGGTGATCCTGGTGGCCGTGCACGTGGCCAGCGGCTACATCGAGGCCGA

      uValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysL
3701  GGTGATCCCCGCCGAGACCGGCCAGGAGACCGCCTACTTCCTGCTGAAGC
            370F and 370R
      euAlaGlyArgTrpProValThrThrIleHisThrAspAsnGlySerAsn
3751  TGGCCGGCAGGTGGCCCGTGACCACCATCCACACCGACAACGGCAGCAAC

      PheThrSerAlaThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGl
3801  TTCACCAGCGCCACCGTGAAGGCCGCCTGCTGGTGGGCCGGCATCAAGCA

      nGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGluSerM
3851  GGAGTTCGGCATCCCCTACAACCCCCAGAGCCAGGGCGTGGTGGAGAGCA

      etAsnLysGluLeuLysLysIleIleGlyGlnValArgAspGlnAlaGlu
3901  TGAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGAGGGACCAGGCCGAG

      HisLeuLysThrAlaValGlnMetAlaValPheIleHisAsnPheLysAr
3951  CACCTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACTTCAAGAG

      gLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArgIleValAspIleI
4001  GAAGGGCGGCATCGGCGGCTACAGCGCCGGCGAGAGGATCGTGGACATCA
```

```
    leAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleThrLysIle
4051 TCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATC

    GlnAspPheArgValTyrTyrArgAspSerArgAspProLeuTrpLysGl
4101 CAGAACTTCAGGGTGTACTACAGGGACAGCAGGGACCCCCTGTGGAAGGG

    yProAlaLysLeuLeuTrpLysGlyGluGlyAlaValValIleGlnAspA
4151 CCCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGATCCAGGACA

    snSerAspIleLysValValProArgArgLysAlaLysIleIleArgAsp
4201 ACAGCGACATCAAGGTGGTGCCCAGGAGGAAGGCCAAGATCATCAGGGAC

    TyrGlyLysGlnMetAlaGlyAspAspCysValAlaGlyArgGlnAspGl
4251 TACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGGCAGGCAGGACGA

    uAsp***
4301 GGACTAGacatctagattgttggattacaactgcaccttcagcgttcagt
```

***Xba*I**

described above for the 3' end of fragment 2 were used to amplify this region. The relevant fragment was isolated from non-specific products by agarose gel electrophoresis and then cloned into the *Bam*HI and *Xba*I sites of pBluescript II SK (+). Eight clones were sequenced and one was found to be of the correct sequence.

The final step was to put these four fragments together to form a contiguous sequence. The 5' and 3' halves of fragment 2 were joined by a tripartite ligation into the *BstE*II and *Bam*HI sites of pB/SBstEII. Fragment 1 was then cloned into the *Hind*III and *Bst*EII sites, placing it in front of fragment 2. Fragment 3 completed the *gagpol* sequence and was cloned into the *Bam*HI and *Xba*I sites of pB/SBstEII already containing fragments 1 and 2, to create pB/Sgagpolml. The entire codon-optimised *gagpol* fragment was then removed from pB/SBstEII with a 5' *Hind*III and 3' *Xba*I restriction digest and cloned into the same two sites of the expression vector pcDNA3.1 to give pcDNA3gagpolml.

A derivative of this plasmid, pcDNA3gagpolmlrre, was made which included the RRE in the 3' non-coding sequence.

4.3.3 EXPRESSION OF GAGPOL FROM SINGLE TRANSCRIPTION UNITS

Expression of GagPol from these five plasmids described above was tested by transient expression in 293T cells. Each of the GagPol expressing plasmids, pcDNA3HIV10gagpol, pcDNA3HIV10gagpolrre, pcDNA3HIV10gagpolrreintron, pcDNA3gagpolml and pcDNA3gagpolmlrre were co-transfected with the vector pHIVext5SV40EYFPppt+RRE (described in section 5.3), pCMVrev (Lewis *et al.*, 1990) pcDNA3tat and pHCMV-G (Yee *et al.*, 1994) into 293T cells as described in section 2.2.17. A constant molar quantity (100 fmoles) of each of the GagPol expression constructs (653 to 850 ng, depending on the construct) was transfected along with 1 μg of pHIVext5SV40EYFPppt+RRE (section 5.3) and 0.5 μg each of pCMVrev, pcDNA3tat and pHCMV-G. Carrier DNA in the form of pcDNA3.1 was added where necessary to make the quantity of DNA the same for each transfection. The pCMVΔRnr (Kafri *et al.*, 1997) construct (used in Chapter five) was included here as a positive control, replacing the GagPol, Tat and Rev expression constructs. Viral supernatant was collected and assayed on NIH3T3 cells as described in section 2.2.18. Reverse transcriptase (RT) and p24 measurements were also performed on the viral supernatant. Measurement of p24 allows direct assessment of Gag expression. p24 is derived from the Gag precursor and is

a major structural core component of HIV-1 viral particles. This 24 kDa protein is immunologically distinct from most proteins in other retroviruses. Viral titres are often reported per quantity of p24 as a measure of packaging efficiency as it equates to viral particles (ie 'viable' viral particles) (Wehrly and Chesebro, 1997). The enzyme reverse transcriptase (RT) is derived from the *gagpol* precursor and provides another measurement of GagPol expression.

Table 4.2 shows that the lowest virus titre was obtained with the native pHIV-YU2 *gagpol* sequence. The pcDNA3HIV10gagpol expression plasmid yielded a titre of 3.8 x 10^3 EYFP transducing units (tU)/ml, more than two orders of magnitude less than pCMVΔRnr which was taken as a standard point of reference. This packaging plasmid expresses all HIV-1 *trans* functions with the exception of Vpr, Vif and Env (Kafri *et al.*, 1997). A seven-fold increase in viral titre was observed with the addition of the RRE to the GagPol expression plasmid. The Rev protein and RRE are necessary for the efficient nuclear export of the HIV-1 unspliced transcript into the cytoplasm. The viral titre could be doubled again with the inclusion of an intron yielding a titre of 5.8 x 10^4 tU/ml. However, as a titre of 1.2 x 10^4 tU/ml equates to a transduction rate of 1 %, these results should be interpreted with some caution. Later reassessment of these *gagpol* constructs using a more efficient vector (data not shown) confirmed the above results, and thus clarifies the integrity of the pHIV-YU2 clone, which was discussed in section 3.6.

In contrast, the plasmid with the codon-optimised sequence, pcDNA3gagpolml, produced higher viral titres compared with those obtained using the native pHIV-YU2 *gagpol* sequence. In the case of the codon-optimised GagPol expression plasmid the addition of RRE did not further increase viral titre or p24 expression. The use of a codon-optimised expression plasmid for GagPol therefore eliminates the dependence of *gagpol* expression on Rev and the RRE. Although the sequences that prevent efficient nuclear export of the *gagpol* transcript in the absence of Rev and RRE remain loosely defined, they are presumably destroyed by codon-optimisation. Moreover, optimising the codons for *gagpol* expression effectively removes all the AUUUA RNA destabilisation *cis* sequence elements in the *gagpol* sequence. A reasonable assumption is that the mere optimisation of codons *per se*, would also augment the translational efficiency of the transcript.

Measurement of RT was performed on viral supernatant taken from the codon optimised *gagpol* expression construct and pCMVΔRnr. Enzyme activities were similar, 3.3 (± 0.9) and 2.2 (± 0.9) ng/ml (n = 4) respectively. Despite comparable RT and p24 levels from pCMVΔRnr and the codon-optimised *gagpol* expression construct, a three-fold

TABLE 4.2

Titres and p24 levels of *gagpol* plasmids.

Different constructs designed to provide in *trans* the *gagpol* functions for transient virus production were used. Viral titres were determined on NIH3T3 cells from viral supernatant collected from 293T cells co-transfected with an equivalent molar amount (100 fmoles) of the relevant GagPol expression plasmid (ng quantity in parenthesis), 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, 0.5 μg of pcDNA3tat and 0.5 μg of pCMVrev. Transduced cells were measured for EYFP expression by FACScan analysis and viral titres were determined by the percentage of fluorescent cells from the total number plated. Measurements of p24 were performed on the viral supernatant collected from the transfected 293T cells using a standard ELISA kit.

***gagpol* plasmid**	**p24 ng/ml**	**Titre EYFP tu/ml**
pcDNA3HIV10gagpol (653)	12 ± 4	$3.8 \pm 1.4 \times 10^3$
pcDNA3HIV10gagpolrre (671)	115 ± 21	$2.6 \pm 0.3 \times 10^4$
pcDNA3HIV10gagpolrreintron (678)	237 ± 24	$5.8 \pm 0.5 \times 10^4$
pcDNA3gagpolml (653)	477 ± 35	$2.4 \pm 0.3 \times 10^5$
pcDNA3gagpolmlrre (671)	392 ± 12	$1.9 \pm 0.01 \times 10^5$
pCMVΔRnr (850)	597 ± 103	$6.6 \pm 1.0 \times 10^5$

n = 4 for all values.

higher titre is still achieved with pCMVΔRnr. This may be a consequence of the complex regulatory mechanisms for HIV-1 which, may be attenuated when the proviral clone is disassembled to create individual packaging plasmids. The pCMVΔRnr helper construct expresses Rev, Tat, Vpu and Nef, in addition to GagPol and this may have important implications for HIV-1 based vectors. Optimal expression from HIV-1 derived transcripts may be dependent upon the context in which they are expressed. In this particular instance, pCMVΔRnr expresses Tat and Rev in the one plasmid and segregating these two functions of HIV-1 into separate plasmids may interfere with what is probably a very coordinated and systematic series of regulated expression events. Additionally the stoichiometry of all the proteins in this transient expression system has not been optimised. At this point no attempt had been made to optimise the ratio of *gagpol*, *rev* and *tat* constructs in this system. Furthermore, Vpu and Nef, which are expressed in pCMVΔRnr are absent in the helper system outlined here. The effect of these proteins on viral titre is described in section 4.5.3.

4.3.4 SEPARATE TRANSCRIPTION UNITS FOR GAGPOL

An attempt was made to further disassemble the HIV-1 functions by expressing Gag and Pol from two separate expression plasmids. A codon-optimised *gag* construct was made by replacing the frameshift and *pol* overlap sequence in pB/Sgagpolml with sequence to complete the *gag* reading frame (Figure 4.4a). This oligonucleotide fragment was flanked by a 5' *Apa*I and a 3' *Xba*I site. Following sequencing, this fragment was re-cloned 5' *Hind*III/3' *Xba*I into pcDNA3.1, producing pcDNA3gagml. The *pol* construct was made by replacing the *gag* sequence up to the first codon of the *pol* reading frame in pB/Sgagpolml, with sequence encompassing a *Hind*III site, a short non-coding region and an ATG initiation codon that was in-frame with the *pol* reading frame. These changes were encompassed within a PCR product by incorporation of the relevant changes in the PCR primer polHindIII. The PCR product made with primers polHindIII and ramp from pcDNA3gagpolml was cloned *Hind*III/*Apa*I into pBluescript II SK (+) and sequenced. This fragment was then sub-cloned by co-ligating this *Hind*III/*Apa*I fragment with an *Apa*I/*Bam*HI (rest of *pol*) fragment into the *Hind*III and *Bam*HI sites of pcDNA3gagpolml thus creating pcDNA3polml.

Due to the fact that the Pol proteins are normally made *via* a GagPol polyprotein, it seemed unlikely that the sequence encoding *pol* would enable the synthesis of the Pol gene

FIGURE 4.4a

Sequence for codon-optimised *gag*.

The frameshift region and the *pol* overlap sequence from the codon-optimised *gagpol* construct was replaced with overlapping oligonucleotides 50 bases in length to complete the *gag* coding sequence. The bold C identifies the start of the overlap sequence and is marked below with an asterisk. The numbering of the sequence on the left-hand side identifies these oligonucleotides by an ML suffix. This new fragment was inserted as an *ApaI*/*Xba*1 fragment thereby replacing the *gagpol* region shown in Figure 4.3.

```
0      gggcccactcacagaagcttaaccatcagcaagcaggtcattgtgccac
                        HindIII
       MetGlyAlaArgAlaSerValLeuSerAlaGlyGluLeuAspLysTrpGl
1      ATGGGCGCCAGGGCCAGCGTGCTGAGCGCCGGCGAGCTGGACAAGTGGGA

       uLysIleArgLeuArgProGlyGlyLysLysGlnTyrArgLeuLysHisI
51     GAAGATCAGGCTGAGGCCCGGCGGCAAGAAGCAGTACAGGCTGAAGCACA

       leValTrpAlaSerArgGluLeuGluArgPheAlaValAspProGlyLeu
101    TCGTGTGGGCCAGCAGGGAGCTGGAGAGGTTCGCCGTGGACCCCGGCCTG

       LeuGluThrSerGluGlyCysArgGlnIleLeuGlyGlnLeuGlnProSe
151    CTGGAGACCAGCGAGGGCTGCAGGCAGATCCTGGGCCAGCTGCAGCCCAG

       rLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsnThrValAlaT
201    CCTGCAGACCGGCAGCGAGGAGCTGAGGAGCCTGTACAACACCGTGGCCA

       hrLeuTyrCysValHisGlnLysIleGluValLysAspThrLysGluAla
251    CCCTGTACTGCGTGCACCAGAAGATCGAGGTGAAGGACACCAAGGAGGCC

       LeuGluLysIleGluGluGluGlnAsnLysSerLysLysLysAlaGlnGl
301    CTGGAGAAGATCGAGGAGGAGCAGAACAAGAGCAAGAAGAAGGCCCAGCA

       nAlaAlaAlaAspThrGlyAsnSerSerGlnValSerGlnAsnTyrProI
351    GGCCGCCGCCGACACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCA

       leValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArg
401    TCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCAGCCCCAGG

       ThrLeuAsnAlaTrpValLysValValGluGluLysAlaPheSerProGl
451    ACCCTGAACGCCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCAGCCCCGA

       uValIleProMetPheSerAlaLeuSerGluGlyAlaThrProGlnAspL
501    GGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGCGCCACCCCCCAGGACC

       euAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet
551    TGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATG

       LeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisPr
601    CTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACAGGCTGCACCC
```

```
          oValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlyS
651       CGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGAGGGAGCCCAGGGGCA

          erAspIleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMet
701       GCGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATG

          ThrAsnAsnProProIleProValGlyGluIleTyrLysArgTrpIleIl
751       ACCAACAACCCCCCCATCCCCGTGGGCGAGATCTACAAGAGGTGGATCAT

          eLeuGlyLeuAsnLysIleValArgMetTyrSerProThrSerIleLeuA
801       CCTGGGCCTGAACAAGATCGTGAGGATGTACAGCCCCACCAGCATCCTGG

          spIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe
851       ACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTC

          TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMe
901       TACAAGACCCTGAGGGCCGAGCAGGCCAGCCAGGAGGTGAAGAACTGGAT

          tThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleL
951       GACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCC

          euLysAlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCys
1001      TGAAGGCCCTGGGCCCCGCCGCCACCCTGGAGGAGATGATGACCGCCTGC

          GlnGlyValGlyGlyProGlyHisLysAlaArgValLeuAlaGluAlaMe
1051      CAGGGCGTGGGCGGCCCCGGCCACAAGGCCAGGGTGCTGGCCGAGGCCAT

          tSerGlnValThrAsnSerAlaThrIleMetMetGlnArgGlyAsnPheA
1101      GAGCCA*GGTGACC*AACAGCGCCACCATCATGATGCAGAGGGGCAACTTCA
                    ***BstE*II**
          rgAsnGlnArgLysThrValLysCysPheAsnCysGlyLysGluGlyHis
1151      GGAACCAGAGGAAGACCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCAC

          IleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGl
1201      ATCGCCAAGAACTGCA*GGGCCC*CCAGGAAGAAGGGCTGCTGGAAGTGCGG
                              ***Apa*I**
          yLysGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheL
1251      CAAGGAGGGCCACCAGATGAAGGA**C**TGCACCGAGAGGCAGGCCAACTTCC
                                    *
          euGlyLysIleTrpProSerHisLysGlyArgProGlyAsnPheLeuGln
1301ML    TGGGCAAGATCTGGCCCAGCCACAAGGGCAGGCCCGGCAACTTCCTGCAG

          SerArgProGluProThrAlaProSerGluGluSerValArgPheGlyGl
1351ML    AGCAGGCCCGAGCCCACCGCCCCCAGCGAGGAGAGCGTGAGGTTCGGCGA

          uGluThrThrThrProSerGlnLysGlnGluProIleAspLysGluLeuT
140ML     GGAGACCACCACCCCCAGCCAGAAGCAGGAGCCCATCGACAAGGAGCTGT

          yrProLeuAlaSerLeuArgSerLeuPheGlySerAspProSerSerGln
145ML     ACCCCCTGGCCAGCCTGAGGAGCCTGTTCGGCAGCAACCCCAGCAGCCAG

          ***
150ML     TAGTAG*TCTAGA*CTGTGAACGACCG
                    ***Xba*I**
```

products. For this reason a construct designed to only express the GagPol polyprotein was also made. This was done by replacing both the *gagpol* overlap sequence and the frameshift signal with sequence that was codon-optimised for the GagPol polyprotein (see Figure 4.4b). Overlapping oligonucleotides for this region, flanked with a 5' *Hind*III and 3' *Xba*I site, were assembled and then amplified with FusFamp and Fusramp and then cloned into pBluescript II SK (+). The region was sequenced and then sub-cloned as an *Apa*I fragment into pcDNA3gagpolml to create pcDNA3gagpolfusionml.

Additional expression constructs for Gag and GagPol containing the RRE and Mason-Pfizer monkey virus constitutive transport element (MPMV-CTE) (Pasquinelli *et al.*, 1997) in the 3' non coding sequence were made by re-cloning the gagml and gagpolfusionml fragments from the pBluescript II SK (+) clones into pcDNA3rre and pcDNA3CTE, producing pcDNA3gagmlrre, pcDNA3gagmlCTE, pcDNA3gagpolfusionmlrre and pcDNA3gagpolfusionmlCTE. To construct pcDNA3rre, the RRE (bases 7734 - 7974) was isolated from pHIV-YU2 by PCR as a 5' *Xba*I/3' *Apa*I fragment using primers RREXbaI and RREApaI. The RRE fragment was then cloned between respective sites in pcDNA3.1 to give pcDNA3rre. The MPMV-CTE fragment was synthesised by heating oligonucleotides CTEF1, CTEF2, CTEF3, CTEF4, CTER1, CTER2, CTER3 and CTER4 at 95°C for 1 min and then cooling slowly to room temperature. This annealed fragment was subsequently cloned into the *Eco*RV site of pBluescript II SK (+). DNA sequencing was used to identify a clone of the correct sequence which was then isolated as an *Eco*RV fragment and sub-cloned into an *Apa*I (end-filled with Klenow) site in pcDNA3.1 to create pcDNA3CTE.

4.3.5 ASSESSMENT OF SEPARATE GAGPOL TRANSCRIPTION UNITS

Viral titres obtained using the helper plasmids described in section 4.3.4 were measured as described previously by substituting various combinations of the helper constructs for the expression of GagPol. Included in the transfection was 0.5 μg each of pCMVrev, pcDNA3tat and pHCMV-G as well as 1 μg of the vector pHIVext5SV40EYFPppt+RRE (section 5.3). The amount of *gag* plasmid in each transfection was standardised to 100 fmoles and various amounts of *pol* and *gagpolfusion* containing plasmids were added ranging from 0 to 100 fmoles. In each case carrier DNA (pcDNA3.1) was added such that the total amount of DNA in each transfection was the same.

FIGURE 4.4b

Codon-optimisation region for the GagPol polyprotein.

The *gagpol* overlap region from the codon-optimised construct was replaced with the sequence between the two *Apa*I sites shown here. This sequence was synthesised as overlapping 50 base oligonucleotides that have been identified with a "fus" suffix. The region of normal overlap for the native pHIV-YU2 GagPol precursor polyprotein is shown in bold.

```
                IleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGl
1201            ATCGCCAAGAACTGCAGGGCCCCCAGGAAGAAGGGCTGCTGGAAGTGCGG
                                    ApaI
                yLysGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheL
1251            CAAGGAGGGCCACCAGATGAAGGACTGCACCGAGAGGCAGGCCAACTTCC

1301fus         euArgGluAspLeuAlaPheProGlnGlyLysAlaArgLysPheSerSer
                TGAGGGAGGACCTGGCCTTCCCCCAGGGCAAGGCCAGGAAGTTCAGCAGC

                GluGlnThrArgAlaAsnSerProIleArgArgGluArgGlnValTrpAr
1351fus         AGCAGACCAGGGCCAACAGCCCCATCAGGAGGGAGAGGCAGGTGTGGAGG

                gArgAspAsnAsnSerLeuSerGluAlaGlyAlaAspArgGlnGlyThrV
1401fus         GAGGGACAACAACAGCCTGAGCGAGGCCGGCGCCGACAGGCAGGGCACCG

                alSerPheSerPheProGlnIleThrLeuTrpGlnArgProLeuValThr
1451fus         TGAGCTTCAGCTTCCCCCAGATCACCCTGTGGCAGAGGCCCCTGGTGACC

                IleLysIleGlyGlyGlnLeuLysGluAlaLeuLeuAspThrGlyAlaAs
1501fus         ATCAAGATCGGCGGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCGCCGA

                pAspThrValLeuGluGluMetAsnLeuProGlyArgTrpLysProLysM
1551            CGACACCGTGCTGGAGGAGATGAACCTGCCCGGCAGGTGGAAGCCCAAGA

                etIleGlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGlnIle
1601            TGATCGGCGGCATCGGCGGCTTCATCAAGGTGAGGCAGTACGACCAGATC

                ProIleGluIleCysGlyHisLysAlaIleGlyThrValLeuValGlyPr
1651            CCCATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGGTGGGCCC
                                                        ApaI
                oThrProValAsnIleIleGlyArgAsnLeuLeuThrGlnIleGlyCys
1701            CACCCCCGTGAACATCATCGGCAGGAACCTGCTGACCCAGATCGGCTGC
```

Neither of the *gag*, *pol* or *gagpolfusion* plasmids could on their own support the synthesis of viral particles (Tables 4.3, 4.4, 4.5 and 4.6). Because the proteins derived from the *pol* gene are normally made from the GagPol precursor polyprotein it is not surprising that, the combination of pcDNA3gagml and pcDNA3polml was also unable to produce any detectable virus particles, although they were capable of expressing Gag (p24) and reverse transcriptase (RT) respectively (Table 4.3). This presumably is due to the production of so called pseudovirions because all of the structural proteins provided by the *gag* gene are able to be synthesised by the transfected cell. HIV-1 Gag proteins produced in the absence of all other viral gene products in a variety of cell types have been shown to elicit the formation of virus-like particles (Karacostas *et al.*, 1989; Gheysen *et al.*, 1989; Shioda and Shibuta, 1990; Ratner *et al.*, 1991; Spearman *et al.*, 1994). However, it is thought that the Gag portion of the GagPol polyprotein is required for its incorporation into virus particles, making the presence of RT in the medium of these transfected 293T cells harder to explain. The observation of large amounts of extracellular RT with pcDNApolml alone suggests that the enzyme may not necessarily be associated with virions even when Gag is included (Table 4.3).

In contrast, when transfected together pcDNA3gagml and pcDNA3gagpolfusionml were able to direct synthesis of viral particles. Tables 4.4, 4.5 and 4.6 display viral titres for the three sets of Gag and GagPol plasmids (unmodified or containing the RRE or CTE). Neither construct on its own supports the generation of detectable viral titre because in each case essential viral coding sequence is absent. For pcDNA3gagml this is the entire *pol* gene, and for pcDNA3gagpolfusionml the p6 protein, which would normally be encoded in the region of overlap between the Gag and Pol reading frames. However, the viral titre for the pcDNA3gagml/pcDNA3gagpolfusionml combination was less than half that obtained when GagPol was expressed as a single transcriptional unit, as was the case with pcDNA3gagpolml. One possible explanation for this discrepancy is that translation of the two polyproteins from the same RNA results in their co-localisation, which could facilitate the assembly of virion particles. Alternatively, empty virion particles produced by pcDNA3gagml may be interfering with EYFP transduction. There is some evidence to suggest that virion assembly is localised within the cell (Rhee and Hunter, 1991). Further experimentation is needed to explore this issue. It is interesting to note that a three- to five-fold decrease in titre has been reported by others when Gag and Pol are expressed separately using a different approach (Wu *et al.*, 2000).

TABLE 4.3

Expression of Gag and Pol from separate transcription units.

Two separate plasmids for the expression of the Gag and Pol proteins. For each experiment 455 ng of pcDNA3gagml (100 fmoles) was co-transfected with the appropriate amount of pcDNA3polml (56, 140 or 560 ng, 10 to 100 fmoles) and 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, 0.5 μg of pcDNA3tat and 0.5 μg of pCMVrev into 293T cells. Supernatant from the transfected 293T cells was assayed for p24 and RT using standard ELISA kits. Viral titres were measured by the transfer of enhanced yellow fluorescent protein (EYFP) expression to NIH3T3 cells.

Molar ratio of pcDNA3gagml: pcDNA3polml	p24 (ng/ml)	RT (ng/ml)	Titre EYFP tu/ml
1:0	568 ± 41	ND	ND
10:1	1205 ± 95	23 ± 7	ND
4:1	1376 ± 74	118 ± 4	ND
1:1	1573 ± 116	283 ± 38	ND
0:1	ND	316 ± 20	ND

ND not detected. n = 3.

TABLE 4.4

Viral titres, RT and p24 measurements using pcDNA3gagml and pcDNA3gagpolfusionml packaging plasmids.

Segregation of the *gag* and *gagpol* packaging functions into two plasmids. For each experiment 455 ng of pcDNA3gagml (100 fmoles) was co-transfected with the appropriate amount of pcDNA3gagpolfusionml (32.7 to 327 ng, 5 to 50 fmoles) and 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, 0.5 μg of pcDNA3tat and 0.5 μg of pCMVrev into 293T cells. Reverse transcriptase (RT) and p24 levels were determined on the viral supernatant from transfected 293T cells using standard ELISA kits. Viral titres were calculated from the percentage of transduced NIH3T3 cells that expressed enhanced yellow fluorescent protein (EYFP).

Molar ratio of pcDNA3gagml: pcDNA3gagpolfusionml	p24 (ng/ml)	RT (ng/ml)	Titre EYFP tu/ml
1:0	96 ± 9	ND	ND
20:1	607 ± 8	1.4 (0.97, 1.73)	$7.5 \pm 0.4 \times 10^4$
10:1	590 ± 39	4.0	$8.9 \pm 0.7 \times 10^4$
5:1	567 ± 22	2.5 (2.01, 1.99)	$1.2 \pm 0.1 \times 10^5$
3:1	535 ± 60	1.5 (1.70, 1.20)	$9.5 \pm 1.1 \times 10^4$
2.5:1	590 ± 53	2.3 (2.23, 2.42)	$9.6 \pm 0.3 \times 10^4$
2:1	493 ± 82	2.4 (1.88, 2.86)	$7.7 \pm 0.1 \times 10^4$
0:1	87 ± 26	2.10	ND

ND not detected

n = 3 for all p24 and titre results. However, for RT only duplicates are reported except "10:1" and "0:1", n = 1.

TABLE 4.5

Assessment of Gag and GagPol expression constructs containing the RRE.

Segregation of the *gag* and *gagpol* packaging functions into two plasmids including the Rev response element (RRE). For each experiment 455 ng of pcDNA3gagmlRRE (100 fmoles) was co-transfected with the appropriate amount of pcDNA3gagpolfusionmlRRE (33 to 330 ng, 5 to 50 fmoles) and 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, 0.5 μg of pcDNA3tat and 0.5 μg of pCMVrev into 293T cells. Supernatant from the transfected 293T cells was assayed for p24 using an ELISA kit. Viral titres were calculated from the percentage of transduced NIH3T3 cells that expressed enhanced yellow fluorescent protein (EYFP).

Molar ratio of pcDNA3gagmlRRE: pcDNA3gagpolfusionmlRRE	**p24 (ng/ml)**	**Titre EYFP tu/ml**
1:0	52 ± 21	ND
20:1	592 ± 59	$1.1 \pm 0.1 \times 10^5$
10:1	728 ± 22	$1.4 \pm 0 \times 10^5$
5:1	572 ± 12	$1.3 \pm 0.1 \times 10^5$
3:1	493 ± 37	$8.9 \pm 0.1 \times 10^4$
2.5:1	437 ± 17	$1.0 \pm 0 \times 10^5$
2:1	430 ± 26	$6.4 \pm 0.2 \times 10^4$
0:1	102 ± 22	ND

ND not detected. n = 3.

TABLE 4.6

Assessment of Gag and Gagpol packaging constructs containing the MPMV-CTE.

Segregation of the *gag* and *gagpol* packaging functions into two plasmids including the Mason-Pfizer monkey virus constitutive transport element (MPMV-CTE). For each experiment 455 ng of pcDNA3gagmlCTE (100 fmoles) was co-transfected with the appropriate amount of pcDNA3gagpolfusionmlCTE (33 to 330 ng, 5 to 50 fmoles) and 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, 0.5 μg of pcDNA3tat and 0.5 μg of pCMVrev into 293T cells. Supernatant from the transfected 293T cells was assayed for p24 using an ELISA kit. Viral titres were calculated from the percentage of transduced NIH3T3 cells that expressed enhanced yellow fluorescent protein (EYFP).

Molar ratio of pcDNA3gagmlCTE: pcDNA3gagpolfusionmlCTE	p24 (ng/ml)	Titre EYFP tu/ml
1:0	66 ± 12	ND
20:1	557 ± 29	$9.5 \pm 0.1 \times 10^4$
10:1	725 ± 88	$1.1 \pm 0 \times 10^5$
5:1	643 ± 46	$1.3 \pm 0.2 \times 10^5$
3:1	467 ± 48	$8.4 \pm 0.2 \times 10^4$
2.5:1	288 ± 13	$6.2 \pm 0.1 \times 10^4$
2:1	335 ± 54	$4.3 \pm 0.1 \times 10^4$
0:1	119 ± 14	ND

ND not detected. n = 3.

Because the inhibitory *cis* sequences had been removed *via* the codon-optimisation process, addition of the RRE or the MPMV-CTE into the *gag* and *gagpol* expression constructs resulted in no significant gain in terms of viral titre or p24 expression. This strongly supports the notion that the codon-optimisation process has removed the dependence on the Rev/RRE mechanism. In contrast, when the native HIV-1 GagPol sequence was used, incorporation of the RRE into the expression construct increased the viral titre seven-fold (Table 4.2).

Varying the ratio of the *gagpol* to *gag* expression plasmids from 2.5:1 through to 20:1 had no effect on viral titre or on the ratio of p24:RT in the viral supernatants (Table 4.4). This was unexpected because in the natural life cycle of HIV-1 between 10 and 20 Gag molecules are synthesised for every molecule of GagPol and this is regulated by a frameshifting mechanism as outlined in the introduction to this chapter. In the *gagpolfusion* construct the frameshift signal has been replaced with optimised codons for the expression of the GagPol precursor polyprotein. Thus it was postulated that the mechanism to control the amount of Gag relative to that of GagPol would be provided by the quantities supplied in the transfection. However, this was found not to be the case. It is conceivable that the p24 levels would remain the same because p24 is a protein translated from the Gag precursor molecule and the amount of *gag* plasmid in each transfection remained constant at 100 fmoles. The levels of RT theoretically would be directly proportional to the quantity of the *gagpolfusion* plasmid because RT is a product of this precursor polypeptide. This would only be true for the RT levels not associated with virion production, as the latter is dependent on Gag. It is interesting to see that pcDNA3pol produces at least ten times more RT than pcDNA3gagpolfusion. The ratio of p24:RT in the virion is presumably determined by the assembly process, although if RT is limiting then virions without RT could conceivably be produced. The higher RT levels seen with pcDNA3pol may reflect quantities that are not associated with viral particles. This could have been tested by sedimenting viral particles and measuring RT levels both in the supernatant and sediment. Due to time restrictions and the fact that the central aim of this project was end-point viral titre determination such an experiment was not performed. It is also worthy to note that whilst virus could not be produced from pcDNA3gagpolfusionml, p24 was detected. This is easily explained because the *gagpolfusion* sequence encodes p24 and the protease to enable processing of the polyprotein precursor.

4.3.6 PROCESSING OF GAGPOL

It is conceivable that the process of codon-optimisation and/or the separation of the two reading frames could have an effect on the processing of the GagPol polyprotein. To test this, western blot analysis was performed on cell lysates prepared from 293T cells transfected with either pCMVΔRnr, pcDNA3gagpolml or a combination of pcDNA3gagml and pcDNA3gagpolfusionml. Heat inactivated (56°C for 1 h) human serum from an HIV-1 positive individual was used as antiserum to p24 (Gag/p24 is one of the most immunoreactive HIV-1 proteins). Figure 4.5 shows the presence of p24 protein in cell lysates transfected with pCMVΔRnr, pcDNA3gagpolml and pcDNA3gagml/pcDNA3gagpolfusionml. No apparent difference is seen between these three samples demonstrating that neither codon-optimisation nor the separation of the two reading frames has a significant effect on the processing of Gag.

4.3.7 TRANSFER OF SEQUENCES ENCODING GAG AND GAGPOL

To demonstrate that codon-optimising the *gagpol* sequence, and expressing Gag and GagPol separately, resulted in improved safety, two assays were developed to measure the transfer of HIV-1 *gag* and *gagpol* sequences to transduced cells. These have been termed the "*gag*" and "*gagpol* transfer" assays and are described in section 2.2.25. Stocks of virus were prepared using either pCMVΔRnr, pcDNA3gagpolml or pcDNA3gagml/pcDNA3gagpolfusionml and concentrated to titres of 2.0×10^7, 2.7×10^7 and 0.8×10^7 NIH3T3 transducing units per ml, respectively. Samples of 2×10^5 and 1×10^6 NIH3T3 transducing units of the virus prepared with pCMVΔRnr were then tested in the "*gag*" and "*gagpol*" transfer assays using a 24 h transduction period for the 293T cells. Similarly, 1×10^6 and 5×10^6 NIH3T3 transducing units of the virus made with pcDNA3gagpolml and 1×10^6 and 2.5×10^6 NIH3T3 transducing units of virus made with pcDNA3gagml/pcDNA3gagpolfusionml were also tested. The results of these assays are shown in Table 4.7. The untransduced 293T control indicated the presence of a low background (ie one colony in the "*gagpol*" transfer assay). Both quantities of virus produced with pCMVΔRnr generated a high read-out in each assay, which was roughly proportional to the original dose of virus tested. With virus preparations generated using pcDNA3gagpolml, a read-out was detected in both the "*gag*" and "*gagpol*" transfer assays although at a much reduced level, approximately 0.1 to 1 % of that seen with pCMVΔRnr. The results of the "*gagpol*" transfer assay, however, were not entirely consistent because

the number of Neo^R colonies obtained with the higher dose of virus was less than that with the lower dose. With virus produced using pcDNA3gagml/pcDNA3gagpolfusionml, only the "*gag*" transfer assay yielded a read-out above the observed background, and this was only 3 % of that generated from virus produced with pCMVΔRnr.

In this first experiment it was observed that, exposure of 293T cells to highly concentrated virus for 24 h had a negative effect on cell growth. For this reason the "*gag*" and "*gagpol*" transfer assays were repeated using a 3 h transduction period, additionally, the transfections of the transduced 293T cells were performed in duplicate. In this experiment duplicate samples of 1 x 10^5 NIH3T3 transducing units of the virus prepared using pCMVΔRnr and 1 x 10^6 NIH3T3 transducing units of the virus prepared with pcDNA3gagpolml and pcDNA3gagml/pcDNA3gagpolfusionml were tested. The results are shown in Table 4.8 and are consistent with the results obtained in the previous experiment (Table 4.7). Again, the virus prepared with pCMVΔRnr produced a high read-out in both assays. Using a ten-fold larger sample, virus generated with pcDNA3gagpolml also gave a read-out in both assays but this was five to eight-fold less than that observed with pCMVΔRnr. The virus prepared using pcDNA3gagml/pcDNA3gagpolfusionml gave a result that was above background only in the "*gag*" transfer assay. Thus, while transfer of "*gagpol*" can clearly occur with pCMVΔRnr and pcDNA3gagpolml, there is no evidence for transfer with the pcDNA3gagml/pcDNA3gagpolfusionml system.

PART TWO

4.4 REGULATORY PROTEINS

The two regulatory proteins Tat and Rev were provided by two simple monocistronic expression constructs. The HIV-1 Tat protein is 101 amino acids long in pHIV-YU2 and the coding region is separated over two exons. The first exon codes for 72 amino acids and almost all of the trans-activating function of Tat is encoded in this first exon (Garcia *et al.*, 1988; Ott *et al.*, 1997). The first exon of *tat* was cloned into the expression vector pcDNA3.1 to produce pcDNA3tat as described in section 3.3. Rev was expressed from a previously described construct pCMVrev (Lewis *et al.*, 1990) that was obtained from the NIH AIDS research and reference program (catalogue number 1443). To examine the effect of Tat and Rev expression on viral titre 293T cells were transiently transfected with 653 ng (100 fmoles) of pcDNA3gagpolml, 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G and/or 0.5 μg each of pcDNA3tat,

FIGURE 4.5

Western blot analysis of p24 expression.

Cell lysates were prepared from 293T cells 72 h after transfection with pCMVΔRnr, pcDNA3gagpolml and pcDNA3gagml/pcDNA3gagpolfusionml. Western blot analysis using antiserum from an HIV-1 positive individual was used to detect p24 expression. Lane 1, control (mock transfected) cells; lane 2, pCMVΔRnr transfected cells; lane 3, pcDNA3gagpolml transfected cells; lane 4, pcDNA3gagml/pcDNA3gagpolfusionml transfected cells. The sizes of the molecular weight standards are indicated to the left of the figure. An arrow to the right indicates the position of p24.

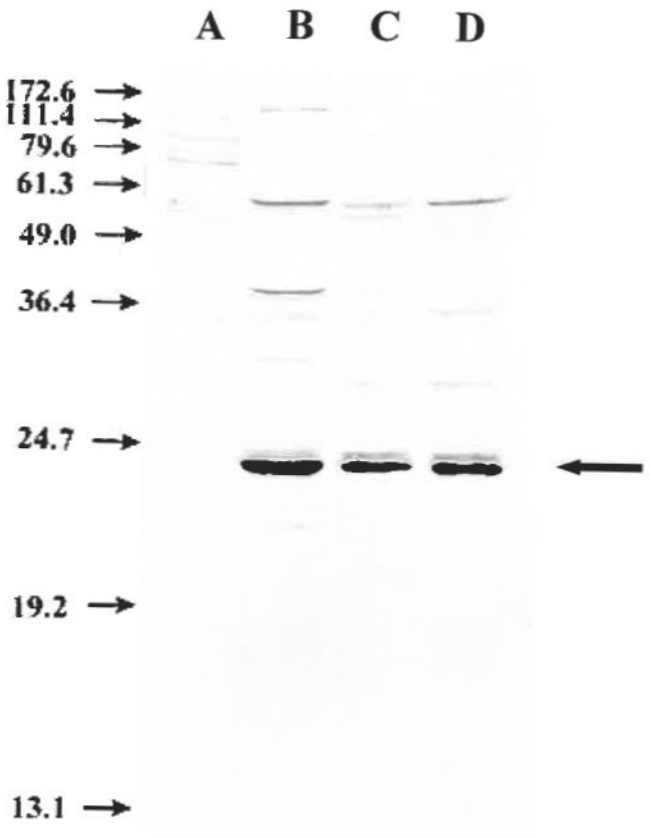

A
B
C
D
172.6
111.4
79.6
61.3
49.0
36.4
24.7
19.2
13.1

TABLE 4.7

Transfer of *gag* and *gagpol* sequences *via* recombinant virus stocks.

Samples of virus prepared with either the pCMVΔRnr, pcDNA3gagpolml or the pcDNA3gagml/pcDNA3gagpolfusionml packaging systems were used to transduce 293T cells which were then assayed for expression of biologically active Gag and Gag/GagPol coding sequences as described in section 2.2.25. For each sample the number of transducing units (tu) assayed and the number of Neo^R colonies produced are reported.

***gagpol* helper construct**	**NIH3T3 tu**	***gag* assay Neo^R colonies**	***gagpol* assay Neo^R colonies**
pCMVΔRnr	0.2×10^6	11	195
pCMVΔRnr	1×10^6	816	1140
pcDNA3gagpolml	1×10^6	1	10
pcDNA3gagpolml	5×10^6	20	1
pcDNA3gagml/pcDNA3gagpolfusionml	1×10^6	26	1
pcDNA3gagml/pcDNA3gagpolfusionml	2.5×10^6	25	0
293T control	-	0	1

TABLE 4.8

Transfer of *gag* and *gagpol* sequences *via* recombinant virus stocks.

Duplicate samples of virus prepared with either the pCMVΔRnr, pcDNA3gagpolml or the pcDNA3gagml/pcDNA3gagpolfusionml packaging systems were used to transduce 293T cells. Each resulting cell culture was assayed in duplicate for transfer of biologically active Gag and Gag/GagPol coding sequences as described in section 2.2.25. For each sample the number of transducing units (tu) assayed and the number of Neo^R colonies produced is reported (total (duplicate 1, duplicate 2)).

***gagpol* helper construct**	**NIH3T3 tu**	***gag* assay Neo^R colonies**	***gagpol* assay Neo^R colonies**
pCMVΔRnr	0.1×10^6	51 (49, 2)	281 (193, 88)
pCMVΔRnr	0.1×10^6	53 (9, 44)	156 (111, 45)
pcDNA3gagpolml	1×10^6	12 (8, 4)	16 (9, 7)
pcDNA3gagpolml	1×10^6	8 (5, 3)	36 (19,17)
pcDNA3gagml/pcDNA3gagpolfusionml	1×10^6	12 (1, 11)	0 (0, 0)
pcDNA3gagml/pcDNA3gagpolfusionml	1×10^6	107 (53, 54)	1 (1, 0)
293T control	-	1 (1, 0)	1 (1, 0)

0.5 μg of pCMVrev/ 0.5 μg of carrier DNA (pcDNA3.1). The supernatant was collected from the transfected 293T cells and assayed for viral titre on NIH3T3 cells as previously described.

The results in Table 4.9 show that addition of Tat increases viral titre almost three-fold ($p < 0.001$) with a concomitant increase in p24. Therefore, Tat is required for maximal viral titre. This finding is in agreement with other reports that show that Tat increases the gene transfer efficiency of vectors that contain the HIV-1 LTR as the promoter (Kim *et al.*, 1998; Mochizuki *et al.*, 1998; Srinivasakumar *et al.*, 1997; Srinivasakumar and Schuening, 1999). The primary function of Tat is in the enhancement of transcription from the HIV-1 LTR. Tat has also been shown to be necessary for optimal virus particle production (Iwakuma *et al.*, 1999) and possibly also for maximal reverse transcriptase activity (Harrich *et al.*, 1997). Therefore, the increase in viral titre observed in the presence of Tat could be due to any one or a number of these processes. However, the three-fold increase in viral production in the presence of Tat is much less than the level of *trans*-activation from the viral LTR by Tat which is hundreds- to thousands-fold (Dayton *et al.*, 1986; Fisher *et al.*, 1986). This disparity may be a reflection of the experimental conditions, if other factors are limiting differences in vector RNA expression will be of less consequence in viral titre estimations.

The expression of the Rev protein is absolutely required for virus production (Table 4.10) although there is little apparent difference in p24 levels. This was expected because the codon optimised GagPol expression construct is not Rev/RRE dependent. The expression of p24 is therefore independent of Rev, but the vector construct, pHIVext5SV40EYFPppt+RRE remains Rev dependent making viral titres dependent on Rev.

TABLE 4.9

Viral titres and p24 with and without Tat.

For each experiment 293T cells were co-transfected with 653 ng (100 fmoles) of pcDNA3gagpolml, 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, and 0.5 μg of pCMVrev, either with or without 0.5 μg of pcDNA3tat. Supernatant from the transfected 293T cells was assayed for p24 using an ELISA kit. Viral titres were determined on NIH3T3 cells as described in sections 2.2.18 and 2.2.21.

	p24 (ng/ml)	**Titre EYFP tu/ml**	**Titre EYFP tu/ng p24**
with pcDNA3tat	473 ± 145	$1.9 \pm 0.6 \times 10^5$	409 ± 78
without pcDNA3tat	299 ± 110	$7.0 \pm 0.2 \times 10^4$	246 ± 44

ND not detected. n = 9.

TABLE 4.10

Viral titres and p24 with and without Rev.

For each experiment 293T cells were co-transfected with 653 ng (100 fmoles) of pcDNA3gagpolml, 1 μg of pHIVext5SV40EYFPppt+RRE, 0.5 μg of pHCMV-G, and 0.5 μg of pcDNA3tat, either with or without 0.5 μg of pCMVrev. Supernatant from the transfected 293T cells was assayed for p24 using an ELISA kit. Viral titres were determined on NIH3T3 cells as described in sections 2.2.18 and 2.2.21.

	p24 (ng/ml)	**Titre EYFP tu/ml**	**Titre EYFP tu/ng p24**
with pCMVrev	473 ± 145	$1.9 \pm 0.6 \times 10^5$	409 ± 78
without pCMVrev	287	ND	ND

ND not detected. n = 9, except without pCMVrev, n = 1.

PART THREE

4.5 SECONDARY PROTEINS

4.5.1 BICISTRONIC EXPRESSION

In the first instance the bicistronic expression constructs, pcDNA3vifIRESvpu and pcDNA3vprIRESnef, for the secondary proteins were made as described in section 3.3. In section 3.4 these expression constructs were transfected into 293T cells along with other helper plasmids and a vector construct. No detectable virus was produced from these 293T cells and the reasons for this have been outlined in the previous chapter. As the role of the HIV-1 minor proteins in virus production was to be investigated it was important to determine whether or not these proteins were being produced in such a system. Cell lysates were prepared from the transfected 293T cells in section 3.4, to look for evidence of gene expression from pcDNA3vifIRESvpu and pcDNA3vprIRESnef. This was performed by western blot analysis as described in section 2.2.23 using antisera to each of the four secondary proteins; HIV-1_{NL4-3} vpr, HIV-1_{NL4-3} vpu, HIV-1_{HXB2} vif and HIV-1 nef (see section 2.1.10). The results depicted non-specific banding on each of the blots (data not shown). Furthermore, cell lysates prepared from untransfected 293T cells, that served as a control, showed no difference in antisera reactions with any of the 293T cells that had been transfected with either pcDNA3vifIRESvpu or pcDNA3vprIRESnef.

Therefore a different experimental system was used to assess expression. pcDNA3vifIRESvpu and pcDNA3vprIRESnef were transfected stably into CHO-K1 cells by electroporation as described in section 2.2.17. Following 10 days of selection of the transfected CHO-K1 cells in 0.8 mg/ml G418, resistant colonies could be seen. These were trypsinised, pooled and grown together as a mass culture. This mass culture was passaged twice in 0.8 mg/ml G418 and then the selection pressure was removed. A cell lysate was prepared from each of the CHO-K1 mass cultures that had been transfected with pcDNA3vifIRESvpu and pcDNA3vprIRESnef, ie CHOvifvpu and CHOvprnef respectively. Western blot analysis was then performed on both lysates with all four antibodies listed above, such that each one behaved as a control for the other. Once again the blots showed a high level of non-specific binding (data not shown). There appeared to be a lot more reactive cellular material compared to that seen in the transiently transfected 293T cells. Of particular note was high molecular weight material that seemed to cross react with all four antibodies.

Immunoprecipitation of the radiolabelled cells was then used as an alternative analysis. G418 resistant mass cultures of CHOvifvpu and CHOvprnef were radiolabelled with 100 μCi/ml EXPRESS [^{35}S] protein labelling mix and cell lystaes immunoprecipitated as described in section 2.2.22. All four antibodies to each of the secondary proteins Vif, Vpr, Vpu and Nef were tested in this way. No specific bands could be seen (data not shown). The fluorograph was smeared with cross reacting material, which again tended to be most prominent in the higher molecular weight range.

In conclusion, it was not possible to show that the two bicistronic expression vectors actually expressed any of the secondary proteins that they were designed to, using the available reagents. There are several potential explanations for this. First, is the nature of the antibodies, all of which are polyclonal antisera directed toward the HXB2 strain of HIV-1 and therefore may display reduced reactivity with the YU2 strain from which the secondary protein expression constructs were made. Second, stable cell lines may pose a problem because some of these proteins, most obviously Vpr may be cytotoxic. Finally, the most simple explanation may be that the level of expression is actual less than the limit of detection. Therefore, individual expression plasmids for each of the secondary proteins were constructed on the assumption that these would be more efficient than the bicistronic constructs.

4.5.2 SINGLE EXPRESSION CASSETTES

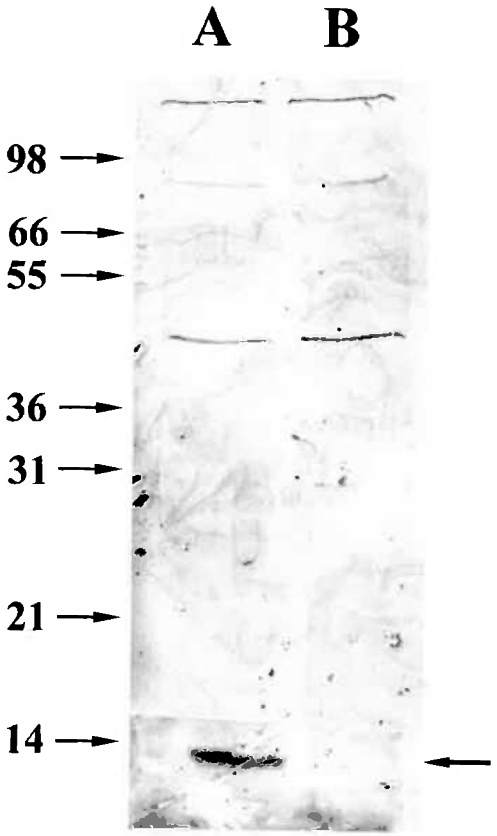
Individual expression vectors for each of the secondary proteins Vif, Vpr, Vpu and Nef were made in pcDNA3.1. The open reading frames were restricted from the pBluescript II SK (+) clones described in section 3.3 and each was individually cloned into pcDNA3.1, to create four expression constructs; pcDNA3Nef, pcDNA3Vif, pcDNA3Vpr and pcDNA3Vpu. These plasmids were separately transiently transfected into Cos1 cells *via* Lipofectamine as described in section 2.2.17. As Cos1 cells express high levels of the SV40 large tumour (T) antigen, plasmids containing the SV40 origin of replication, such as pcDNA3.1, can be amplified to high copy number significantly enhancing expression. Seventy-two hours post transfection the Cos1 cells were harvested with a rubber policeman (cell scraper) and cell lysates prepared in RIPA buffer. Western blot analysis, as above (and section 2.2.23), was performed on each of the four transfected Cos1 cells with mock transfected Cos1 cells serving as a negative control in each case.

Figure 4.6 shows a western blot analysis of the Vpr protein. A band with a molecular weight of 14 kDa is clearly visible in Cos1 cells transiently transfected with

FIGURE 4.6

Western blot analysis of Vpr expression.

Cos1 cells were transfected with pcDNA3Vpr and cell lysates prepared 72 h later. Western blot analysis was performed using antiserum to Vpr (NIH AIDS research and reference reagent program cat. No. 3252) as described in section 2.2.23 Lane A, lysate from pcDNA3Vpr transfected cells; lane B, lysate prepared from mock transfected Cos1 cells. Molecular weight standards are indicated to the left of the figure. The arrow on the right shows the position of the protein in the pcDNA3Vpr transfected cells.


A
B
98
66
55
36
31
21
14

pcDNA3Vpr, indicating successful expression. It was not possible to demonstrate expression of the other three minor proteins, Vpu, Nef or Vif. As was seen before, all three antibodies displayed a high level of cross reactivity with other proteins in the Cos1 cells (data not shown). There are two plausible explanations for the lack of a positive reaction with these antibodies. First, as mentioned in the previous section, all four antibodies, were directed toward the HXB2 strain of HIV-1. Second, because of the high cross reactivity toward cellular proteins displayed with these antibodies any positive reaction may have been masked. Further work is needed here. Alternatively northern analysis using probes to look for gene expression would have been of some interest, but due to time constraints this avenue of experimentation was not pursued.

4.5.3 EFFECT OF SECONDARY PROTEINS ON VIRAL TITRE

The preceding section showed that it was not possible to directly demonstrate the expression of the HIV-1 minor proteins with the exception of Vpr. Nevertheless due to the simplicity of the constructs, the assumption was made that Vif, Vpu and Nef were being expressed. Expression of HIV-1 BH10 Nef from pcDNA3.1 has been described by others (Cooke *et al.*, 1997) supporting this view. Therefore these plasmids were used to assess the effect of the secondary proteins on viral titre. 293T cells were transiently transfected with the following plasmid combinations; as a negative control 1.4 μg of carrier DNA (pcDNA3.1) was transfected along with 0.4 μg each of pHIVext5SV40EYFPppt+RRE, pcDNA3gagpolml, pcDNA3tat, pCMVrev and 0.2 μg of pHCMV-G. The secondary proteins, 0.3 μg each of pcDNA3Vif, pcDNA3Vpr, pcDNA3Vpu and pcDNA3Nef, were used in place of the carrier DNA for the test sample. Virus was collected, and titred on NIH3T3 cells as previously described. Table 4.11 shows that addition of the minor proteins caused a small (26 %) increase in viral titre ($p < 0.05$) and, in terms of titre per ng of p24, nearly a two-fold increase ($p < 0.001$).

There is published evidence that both supports and contradicts this finding. Some investigators have reported that removal of the minor proteins from the packaging system has no detrimental effect on viral titre (Zufferey *et al.*, 1997, Kim *et al.*, 1998). Other reports have shown that Vpr and the combination of Vpr and Vpu have increased viral particle production (Srinivasakumar and Schuening, 1999). However, while the effect on titre in this analysis is relatively small, this does not deny the possibility that inclusion of

TABLE 4.11

Viral titres and p24 with and without expression constructs for the minor proteins Vpr, Vif, Vpu and Nef.

For each experiment 293T cells were co-transfected with 0.4 μg each of pcDNA3gagpolml, pHIVext5SV40EYFPppt+RRE, pCMVrev, pcDNA3tat and 0.2 μg of pHCMV-G, either with or without 0.3 μg of each of the minor protein expression constructs pcDNA3Vpr, pcDNA3Vif, pcDNA3Vpu and pcDNA3Nef. Supernatant from the transfected 293T cells was assayed for p24 using an ELISA kit. Viral titres were determined on NIH3T3 cells as described in sections 2.2.18 and 2.2.21.

	p24 (ng/ml)	**Titre EYFP tu/ml**	**Titre EYFP tu/ng p24**
with minor proteins	211 ± 28	$1.9 \pm 0.4 \times 10^5$	895 ± 150
without minor proteins	321 ± 36	$1.5 \pm 0.2 \times 10^5$	477 ± 80

For both experiments n = 9.

some, or if not all, of the accessory proteins may be important or absolutely essential in some circumstances. For example, the transduction of quiescent skin fibroblasts and hepatocytes does require the function of the secondary proteins (Gasmi *et al.*, 1999). Chinnasamy *et al.* (2000) have shown that transduction of primary resting lymphocytes with HIV-1 based vectors also requires the presence of the secondary proteins.

4.6 CONCLUDING DISCUSSION

The development of helper systems for HIV-1 based vectors is complicated because HIV-1 is a complex retrovirus. Its genome encodes not only the *gag*, *pol* and *env* genes common to all retroviruses, but also contains genes for the accessory proteins Nef, Vif, Vpr, Vpu and two regulatory proteins Tat and Rev (Cullen, 1991). To produce HIV-1 vector particles the *trans* functions need to be provided. To do this, the pHIV-YU2 *trans* functions have been dissassembled with the exception of Env, which is substituted with the VSV-G protein. The YU-2 strain of HIV-1 was chosen as a fully replication-competent, non-cell culture-adapted strain capable of infecting macrophages. Each of the relevant reading frames has been isolated as a minimal or near minimal transcription unit and cloned into separate expression constructs. In addition the *gagpol* gene was codon-optimised for high-level expression in mammalian cells.

Expression of GagPol from the native pHIV-YU2 coding sequence was shown to be dependent on Rev/RRE and was increased further with the inclusion of an intron. A codon-optimised expression plasmid for GagPol effectively eliminated the dependence on Rev/RRE, and the inclusion of the constitutive transport element (CTE) from the Mason-Pfizer monkey virus, which has been shown to be able to partially substitute for the action of Rev/RRE for the expression of GagPol (Gasmi *et al.*, 1999), had no effect. This was despite the fact that the frameshift signal and the region of overlap between the *gag* and *pol* reading frames were unaltered. Rev counteracts the nuclear retention of viral RNA that is controlled by splice sites and *cis* repressive sequences to enable nuclear export of RRE containing RNA. These inhibitory sequence elements have been located within the splice donor site (Borg *et al.*, 1997; Hammarskjöld *et al.*, 1994; Purcell and Martin, 1993), p17 of *gag* (Berthold and Maldarelli, 1996; Schwartz *et al.*, 1992a and b), *pol* (Cochrane *et al.*, 1991; Maldarelli *et al.*, 1991; Olsen *et al.*, 1992) and *env* (Nasioulas *et al.*, 1994). Presumably those sequences in *gag* and *pol* that prevent efficient nuclear transport of the native *gagpol* sequence are destroyed by codon-optimisation. For example, codon-

optimisation removes all the AUUUA RNA destabilisation *cis* sequence elements in the *gagpol* sequence and this is likely to be one mechanism through which expression is enhanced. Kotsopoulou *et al.* (2000) have also reported a codon-optimised coding sequence for HIV-1 GagPol and showed that codon-optimisation results primarily in increased mRNA levels. Codon-optimisation itself would also be expected to directly enhance the translational efficiency of the transcript.

A second important consequence of codon-optimisation is that homology with the native *gagpol* sequence of HIV-1 is reduced to 75% overall so reducing in turn the probability of recombination events occurring between vectors containing *gag* sequence or HIV-1 itself and *gagpol* expressing helper constructs. Furthermore, codon-optimisation destroys the *vif* reading frame overlapping the 3' end of the native *gagpol* sequence. Although the levels of p24 and RT expression from pcDNA3gagpolml (codon-optimised sequence) are comparable to pCMVΔRnr (being 80 % and 150 % respectively, of the values for pCMVΔRnr) the resulting viral titre is only 36 % of that obtained with pCMVΔRnr. One possible explanation is that the expression of Tat and Rev as two separate constructs may not be optimal as their expression is no longer coordinately regulated with expression of GagPol. A second is that, pCMVΔRnr expresses Nef and Vpu and results presented here suggest that the accessory proteins can influence viral titre. These issues need to be further investigated.

Having achieved Rev/RRE independent expression of GagPol, the idea of further separation of the *trans* functions for HIV-1 was explored. The coding sequence for the entire *gagpol* region was split over two plasmids. First, separate codon optimised expression constructs were created for *gag* and *pol*. Not surprisingly, given the coordinated and orderly processing of the Gag and GagPol polyproteins and the functional linkage of these processes to their incorporation into the virion, the use of separate reading frames for *gag* and *pol* did not result in the production of detectable levels of virus. Conversely, separate codon optimised plasmids for expression of the Gag protein and the GagPol polyprotein (pcDNA3gagml and pcDNA3gagpolfusionml) in combination did result in the efficient synthesis of p24 and RT, and in high viral titres.

Segregating the reading frames for the Gag and GagPol polyproteins presents significant safety advances toward the development of safe packaging systems for HIV-1 based vectors. It enables the removal of a vital HIV-1 *cis* acting element, the *gagpol* translational frameshift signal, from the virus production system. Second, because neither plasmid on their own can provide for the synthesis of viral particles, the probability of recombination leading to the unwanted generation of replication competent virus is further

reduced. It is hard to envisage how recombination events alone could create replication competent virus when separate constructs are used for the expression of Gag and GagPol.

A somewhat similar system has been described by Wu *et al.* (2000), where Gag-Pro (Protease) and Vpr-RT-IN (Integrase) polyproteins are expressed separately. In this instance RT-IN is expressed as a fusion protein with Vpr to facilitate its incorporation into the virion. They have not used separate plasmids for expressing Tat, Rev or Vif (the only secondary protein expressed by their system) and retain the frameshift signal for expression of Pro. Of note is a region of overlap between the Gag-Pro and RT-IN constructs, although the RT and IN reading frames are blocked by stop codons in the Gag-Pro expression construct.

The results of the "*gag*" and "*gagpol*" transfer assay demonstrate that the use of a codon-optimised reading frame decreases, but does not completely abolish, the transfer of biologically active sequences encoding either the Gag polyprotein alone, or sequences the equivalent of the *gagpol* gene. This is not surprising because the codon-optimised reading frames still have short stretches of absolute sequence homology with the native *gagpol* gene sequence that will allow homologous recombination between the vector *gag* sequence and the codon-optimised helper sequence. Others (Wagner *et al.*, 2000) have also shown that codon-optimisation improves safety but in this instance an assay for replication-competent virus was used with negative results. This suggests that the "*gag*" and "*gagpol*" transfer assays described in section 4.3.7 provide a more stringent assessment of safety (at least in terms of measuring transfer of sequences expressing HIV-1 Gag/GagPol polyproteins) and that it is therefore a more appropriate assay to use for comparative and absolute assessment of different HIV-1 packaging systems. With the split helper system utilising separate plasmids for the expression of the Gag and GagPol polyproteins, no transfer of sequences capable of substituting for the *gagpol* gene (ie expressing Gag and GagPol) using the "*gagpol*" transfer assay was detected. However, it still remains to be determined whether such an assay is capable of detecting single events (ie transfer of a single intact reading frame). As HIV-1 is diploid, coincident transfer of the reading frames for *gag* and *gagpolfusion* is theoretically possible.

It now seems clear that the accessory proteins are not essential for the efficient production of recombinant HIV-1 based vectors from transfected 293T cells (Kafri *et al.*, 1997; Zufferey *et al.*, 1997; Kim *et al.*, 1998). Additionally, Vif, Vpr, Vpu and Nef have been shown to be dispensable for HIV-1 replication in immortalised cell lines (Miller and Sarver, 1997). The four minor proteins have been expressed separately here, in individual

constructs, and can therefore be included in any combination that may be required to successfully transduce certain cell types (see section 4.5.3).

The activity of the Nef protein appears to be directly related to the route that the provirus takes to the nucleus (Luo *et al.*, 1998). HIV-1 enters cells by direct fusion with the viral envelope protein and the CD4 receptor on target cells. Nef has been shown to increase viral titre 5-10-fold in HIV-1 based vectors that use the native envelope protein (Srinivasakumar *et al.*, 1997). Pseudotyping HIV-1 with VSV-G targets entry of virions to an endocytic pathway and suppresses the requirement for Nef (Aiken, 1997). Such pseudotyping also broadens considerably the host range of the recombinant viral vectors potentially enabling gene transfer into a wide variety of cell types. The activities of Vif and Vpu would appear to be cell specific and neither proteins appear to be required in 293T cells. The effects of Vif seem to be limited to lymphocytes and macrophages (Gabuzda *et al.*, 1992; von Schwedler *et al.*, 1993), and although Vpu has been shown to enhance virion release in some cell types (Klimkait *et al.*, 1990; Göttlinger *et al.*, 1993), it does not appear to have any influence on the release of virions from 293T cells (Zufferey *et al.*, 1997). Due to the fact that 293T cells have a very high efficiency of transfection, large amounts of the vector genome and HIV-1 encoded proteins are probably synthesised thus making the effect of Vpu negligible.

The Vpr protein contains a nucleophilic determinant that enables nuclear transport of the viral capsid and replication in non-dividing cells (Heinzinger *et al.*, 1994). The observation that Vpr is not required is most likely due to the redundancy of mechanisms permitting nuclear localisation, which is also directed by the karyophilic properties of the matrix antigen (MA) and integrase (Bukrinsky *et al.*, 1993; Gallay *et al.*, 1997). In the presence of these two proteins the role of Vpr appears only additive and noticeable mainly in terminally differentiated macrophages (Zufferey *et al.*, 1997). Of particular relevance in the development of stable packaging cell lines is the cytostatic nature of Vpr, making it an attractive candidate to remove from the packaging system.

The western blot analysis performed in this chapter could only demonstrate the successful expression of the Vpr protein. As mentioned previously the high cross reactivity, the nature of the antibodies and the unknown efficiency of expression may explain the lack of a positive reaction with Vpu, Vif and Nef. Further experimentation is clearly needed to see which of the secondary proteins and/or combination is causing an increase in viral titre when transducing NIH3T3 cells. No other cell types were transduced in this chapter, and it is quite likely that the effect of these minor proteins is cell specific. For the transduction of some specific cellular targets, one or more of the HIV-1 accessory

proteins may be essential. As more information regarding the action of the secondary proteins becomes available, their role in HIV-1 based gene transfer systems will become clearer.

The main mode of action of Tat is transcription from the HIV-1 LTR. This requirement for Tat in HIV-1 based vectors can be circumvented with the use of promoters other than the HIV-1 LTR to drive expression of the vector genome (Kim *et al.*, 1998; Miyoshi *et al.*, 1998; Iwakuma *et al.*, 1999). However, there is some evidence that Tat is also required in the HIV-1 life cycle at a post-entry step within the host cell and may enhance infection of the target cells (Huang *et al.*, 1994; Harrich *et al.*, 1997). The findings presented here also suggested that the expression of Tat in a packaging cell line can increase vector titres as seen by a three-fold reduction in viral titre in the absence of Tat. The relevance of these observations to HIV-1 derived vector systems, nevertheless, is uncertain.

The other regulatory protein of HIV-1, Rev, binds to RRE within the *env* gene present in viral mRNAs enabling their transportation from the nucleus to the cytoplasm (Emerman *et al.*, 1989; Malim *et al.*, 1989). This is an essential function for HIV-1 replication but other transport elements have been shown to substitute for RRE and Rev (Huang and Liang, 1993; Bray *et al.*, 1994). However, in the experiments reported here, Rev is required for virus production because of the vector used, pHIVext5SV40EYFPppt+RRE remains Rev/RRE dependent.

Taken together, codon-optimisation of *gagpol* and the removal of the frameshift, minimises considerably sequence homology between HIV-1 based vectors containing *gag* sequence and the packaging plasmids. In addition by manipulating both the codon usage and structure of the *gag* and *gagpol* reading frames it has been possible to express them independently from Rev and RRE. Moreover, the Gag and GagPol polyproteins can be expressed from separate codon optimised expression constructs thereby removing the *gagpol* translational frameshift from the virus production system. The construction and evaluation of these plasmids provide the basis for the production of recombinant HIV-1 vectors in transient expression systems with increased safety as has been demonstrated by measuring the incidence of transfer of biologically active sequences encoding Gag/GagPol to transduced cells.

To further develop this transient packaging system an obvious requirement is to optimise expression of each individual protein and subsequently to titrate the optimal amount of each plasmid against each and every other plasmid. This process will also be

important for creating reagents required to develop sensitive assays for the transfer of biologically active HIV-1 helper protein reading frames as discussed.

Furthermore, additional variables are introduced as soon as additional plasmids are included in the transfection procedure. Packaging cell lines for the stable expression of the *trans* functions of HIV-1 are likely to be a way around this problem, even if only some plasmids could be stably transfected at this stage. It is important to remember that there may be a trade off for a lower viral titre but a safer, in terms of possible recombination events, system. Despite the fact that high titre HIV-1 based vector systems have been produced by transient transfection of 293T cells, such a method is just not suitable for large scale vector production. The establishment of stable packaging cell lines for HIV-1 based vectors will be far more amenable to mass production and will also make helper virus generation less likely because multiple homologous recombination events are required. Such events will occur more frequently before stable transfection and integration take place. The next step in evolving packaging cell lines for HIV-1 based vectors is to make stable cell lines. Now that the Gag and GagPol functions can be split over two plasmids perhaps the first place to start would be a stable cell line for the expression of Gag.

CHAPTER FIVE

EVALUATION OF HIV-1 BASED GENE TRANSFER VECTORS

5.1 INTRODUCTION

The advantages of using human immunodeficiency virus type 1 (HIV-1) as a gene transfer vector have been described in Chapter one. These advantages indicate that it is likely that HIV-1 will replace the archetypical Moloney leukemia virus (MLV) as the vector of choice for gene transfer into many cell populations, including haematopoietic stem cells, hepatocytes, myofibres and neurones (Naldini *et al.*, 1996a; Reiser *et al.*, 1996; Zufferey *et al.*, 1997; Sutton *et al.*, 1998; Mochizuki *et al.*, 1998; Miyoshi *et al.*, 1999). A gene transfer vector based on HIV-1 would ideally incorporate the minimum amount of viral elements required in *cis* for high efficiency transfer into non-dividing cells, and the gene to be transferred along with its associated expression signals. At first sight these must presumably include (i) a packaging signal (Ψ) to expedite packaging of viral RNA into progeny particles, (ii) signals required for reverse transcription, encompassing the repeated (R) region necessary for DNA strand transfer and two regions for the priming of DNA synthesis, the PBS (RNA-binding site) that binds the minus-strand tRNA primer and the polypurine tract (PPT) for priming of plus strand synthesis, (iii) the LTRs (long terminal repeats) which contain signals for synthesis and processing of viral RNA including promoter and enhancer elements and polyadenylation signals and (iv) sequences required for integration.

Other *cis* active sequences that have been identified in the HIV-1 genome are the Rev response element (RRE), and the central polypurine tract (Charneau *et al.*, 1992). Additionally, and more vaguely, there are sequences that have been suggested to play a significant, but not vital, role in HIV-1 replication and also a number of sequences that have been implicated in affecting HIV-1 vector performance. Despite numerous reports describing HIV-1 vectors (Naldini *et al.*, 1996a; Reiser *et al.*, 1996; McBride *et al.*, 1997; Kim *et al.*, 1998; Mochizuki *et al.*, 1998; Chang *et al.*, 1999; Cui *et al.*, 1999; Gasmi *et al.*, 1999), uncertainty still exists regarding the extent of *cis*-acting viral sequences required for optimal vector function. The main parameter of concern has been virus titre, but transgene expression and safety considerations are also important.

In contrast to MLV, it has been difficult to define a sequence from the 5' region of the HIV-1 genome that is sufficient for the packaging of a heterologous sequence (Berkowitz *et al.*, 1995). It would appear that the inclusion of some 5' *gag* sequence is beneficial for packaging of RNA (McBride *et al.*, 1997). HIV-1 vectors with deletions of most of the structural genes are successfully packaged into viral particles (Poznansky *et al.*, 1991; Shimada *et al.*, 1991; Buchschacher and Panganiban, 1992). Arguments have been put forward to suggest that portions of the *env* and *gag* genes enhance packaging

efficiency (Richardson *et al.*, 1993; Parolin *et al.*, 1994, Kaye *et al.*, 1995) and a region contained within the first 40 nucleotides of *gag* has also been shown to be important (Luban and Goff, 1994). The incorporation of RNA into budding virions is reduced, albeit at different levels, as a consequence of sequence deletions between the splice donor site and the *gag* initiation codon (Lever *et al.*, 1989; Aldovini and Young, 1990; Clavel and Orenstein, 1990). Sequences within U5 have also demonstrated their importance in efficient RNA packaging (Kim *et al.*, 1994; Vicenzi *et al.*, 1994). To fully utilise the promise of HIV-1 as a vector for gene therapy it is important to define as exactly as possible the effects of *cis* elements on vector function, allowing the cost versus benefit (ie safety versus titre) of their inclusion to be properly assessed.

As more HIV-1 based vectors were described in the literature, it became possible to take a more systematic approach to vector design. The work reported in this chapter attempted to evaluate *cis* sequences from the HIV-1 genome that may or may not be required for optimal vector performance with the combined aim of maximising viral titre and improving bio-safety. The results obtained in Chapter three showed that some initial ideas about HIV-1 based vector design were correct in principle, but also highlighted many problems associated with the over-complicated approach adopted. In this chapter a more systematic angle was taken using a reporter gene and packaging system of known efficiency. The vectors constructed were designed to express the enhanced yellow fluorescent protein (EYFP), a reporter gene that allows the easy measurement of transduction by FACScan analysis. This was chosen to overcome the inherent difficulties that had become apparent with the transfer of the lysosomal enzyme β-glucuronidase described in Chapter three. Furthermore, to eliminate some of the variables in the virus production system an established packaging construct pCMVΔRnr (Kafri *et al.*, 1997; Naldini *et al.*, 1996b) was used. This HIV-1 packaging construct expresses all HIV-1 *trans* functions with the exception of Vpr, Vif and Env. The HIV-1 based vector was pseudotyped with the G glycoprotein of vesicular stomatitis virus (VSV-G). The use of VSV-G results in a pantropic pseudotype and a more stable virion (Yee *et al.*, 1994). Additional analysis included estimating the number of HIV-1 viral particles produced by measuring p24 antigen, a virion protein encoded by the HIV-1 *gag* gene (Dorfman *et al.*, 1993).

To date, from the number of studies describing HIV-1 vectors, it is clear that parts of *env* and much of the middle of the HIV-1 genome, such as the 3' region of *gagpol* can be deleted without significant deleterious effects on vector function. Relatively simple vectors have been made that are packaged reasonably efficiently (Chang *et al.*, 1999; Cui

et al., 1999) and it has been demonstrated that certain sequence elements, such as the promoter/enhancer region in the 5' LTR and the polyadenylation signal in the 3' LTR, can be replaced with heterologous sequences (Miyoshi *et al.*, 1998; Iwakuma *et al.*, 1999). In this chapter a number of HIV-1 *cis* sequences in general, and various specific *cis* sequences of defined function, were evaluated for their effects on vector performance. These sequences included i) regions from the 5' end of the *gag* gene that have been implicated in the packaging function of HIV-1 (Parolin *et al.*, 1994), ii) regions from the 5' and 3' ends of the *env* gene (Richardson *et al.*, 1993, Kaye *et al.*, 1993), iii) sequences immediately 5' from the polypurine tract (Ilyinskii and Desrosiers, 1998), iv) the central polypurine tract (Charneau and Clavel, 1991; Charneau *et al.*, 1992) and v) splicing signals (Cui *et al.*, 1999). In addition, some heterologous sequence elements were tested for their ability to substitute for the Rev/RRE function.

RESULTS AND DISCUSSION

5.2 BASIC STARTING VECTOR CONSTRUCTION

A vector backbone was designed to incorporate some basic sequences deemed necessary for vector function to begin evaluating specific sequence elements within the HIV-1 genome that may effect vector performance. The incorporation of the aforesaid sequences was based on the idea developed in Chapter three and what was in the literature at the time. The first two HIV-1 derived EYFP (enhanced yellow fluorescent protein) transfer vectors that were made contained the following elements; the 5' long terminal repeat (LTR) and contiguous sequence up to 100 (ext2) or 1150 (ext5) bp of the HIV-1 YU2 *gag* gene, the SV40 immediate early promoter, a polylinker sequence containing, from 5' to 3', a *Stu*I, *Hind*III and an *Eco*RI site, into the latter of which the cDNA for EYFP was cloned, the polypurine tract (ppt) and lastly the 3' LTR. These two vectors, pHIVext2SV40EYFPppt and pHIVext5SV40EYFPppt, were developed to allow other elements to be analysed in the context of vectors containing two different lengths of *gag* ie 100 and 1150 bp of *gag* respectively. The sequence elements present in pHIVext2SV40EYFPppt and pHIVext5SV40EYFPppt are depicted in Figure 5.1.

These two constructs were made as follows, firstly two PCR primers (5'HIVext2R and 5'HIVext5R) were synthesised to allow amplification of sequences to base 887 and 1937 of pHIV-YU2 respectively. A PCR product (5' *Not*I/3' *Bam*HI) was generated with

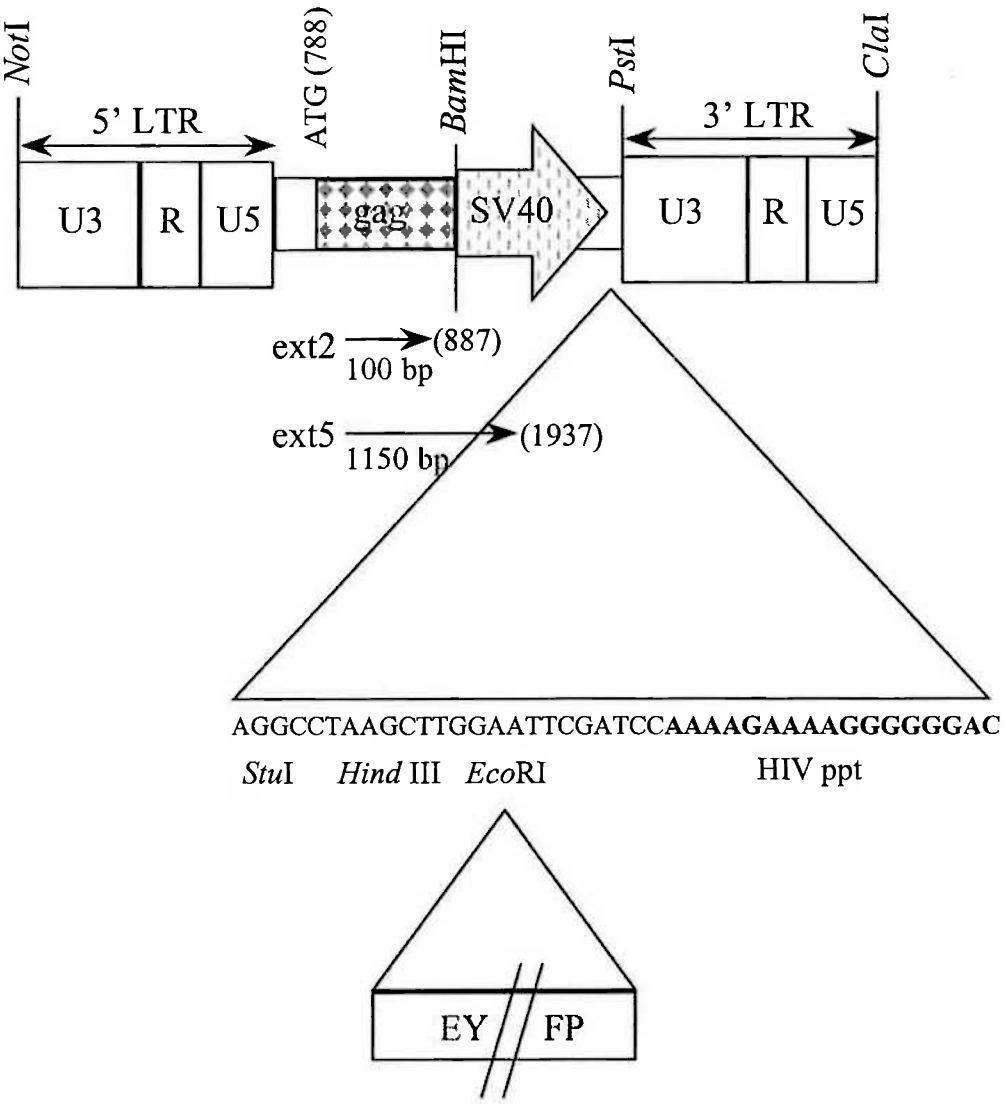


FIGURE 5.1

pHIVext2SV40EYFPppt and pHIVext5SV40EYFPppt.

Detailed construction of these two vectors is provided in section 5.2. The vectors contain from 5' to 3' the following sequence elements; HIV-1 5' YU-2 LTR and contiguous sequence up to base 887 in the ext2 derivative, or 1937 in the ext5, the SV40 early promoter, a polylinker containing a *Stu*I, *Hind*III and *Eco*RI site (the EYFP gene cloned into the *Eco*RI site), the polypurine tract (ppt, sequence in bold) and the 3' YU-2 LTR.

each of these primers in combination with the primer 5'LTRF, and these were then cloned into pB/SHIVLTR (section 3.2) such that they replaced the 5' LTR and the 390 bp oligonucleotide fragment in this clone (ie up to the *Bam*HI site). These two constructs were termed pHIVext2 and pHIVext5 respectively. The SV40 promoter was isolated from pLNSX (Miller and Rosman, 1989) with *Bam*HI (5') and *Hind*III (3'), end-filled and then cloned into the end-filled *Bam*HI site of both the ext2 and ext5 constructs. To complete these two vectors, the EYFP sequence was removed from pEYFP (Clontech) as an *Xba*I fragment, end-filled and then cloned into the end-filled *Eco*RI site of the two vectors. This produced two vectors, pHIVext2SV40EYFPppt and pHIVext5SV40EYFPppt.

5.3 ADDITION OF EXTRA HIV-1 SEQUENCES

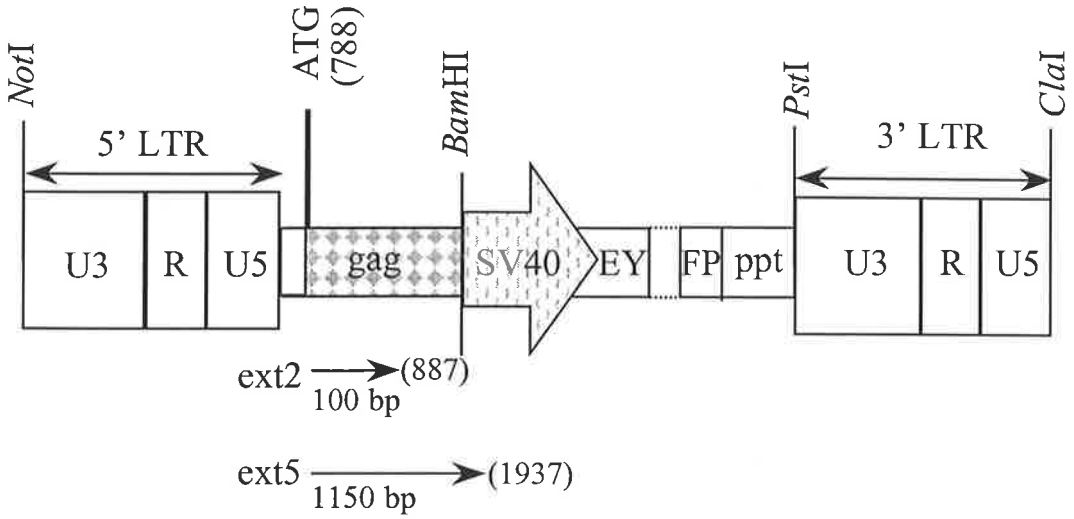
Various additional sequences from pHIV-YU2 were then incorporated into both the pHIVext2SV40EYFPppt and pHIVext5SV40EYFPppt vector constructs to assess their effect on viral titre. A series of vector pairs were so created which contained either 100 (ext2) or 1150 bp (ext5) of *gag* sequence allowing any effect seen to be related to the length of *gag* in the vector.

The sequence immediately 5' of the polypurine tract has been shown to be essential for simian immunodeficiency virus replication (Ilyinskii and Desrosiers, 1998), and may therefore influence the efficiency of HIV-1 packaging. To assess this sequence in the context of HIV-1 vector function, 25 bp of HIV-1 YU2 sequence 3' of the polypurine tract was incorporated into the ext2 and ext5 vector pair, creating pHIVext2SV40EYFPppt+ and pHIVext5SV40EYFPppt+. Further additions of sequence from the 3' end of the *env* gene in pHIV-YU2 were then added. Sequences within the *env* gene have been suggested to contribute to the packaging of HIV-1 based vectors (Richardson *et al.*, 1993). The first vector pair pHIVSAext2SV40EYFP and pHIVSAext5SV40EYFP, contained contiguous sequence from immediately 5' of the splice acceptor site for the second exon of the *tat* and *rev* genes (base 8331 in pHIV-YU2) to the 3' LTR. The vectors, pHIVF4ext2SV40EYFP and pHIVF4ext5SV40EYFP, contained envelope sequence immediately 3' of the same splice acceptor site (base 8341 in pHIV-YU2), contiguous with the 3' LTR and pHIVF6ext2SV40EYFP and pHIVF6ext5SV40EYFP, encompassed a smaller region of HIV-1 envelope sequence extending to base 8685 in pHIV-YU2, again contiguous with the 3' LTR (see Figure 5.2). These vectors were made using a series of 3'LTRF primers, 3'

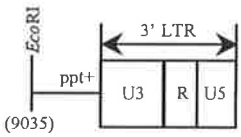
FIGURE 5.2

Addition of extra HIV-1 sequence elements.

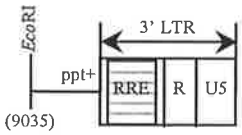
The two basic vector constructs pHIVext2SV40EYFPppt and pHIVext5SV40EYFPppt are shown at the top of the figure (see also Figure 5.1). Underneath the HIV-1 sequence variations that were incorporated into the basic vectors are depicted. The ppt+ sequence encompasses HIV-1 sequence from base 9035 of YU-2 until the end of the 3' LTR. The ppt+RRE is the same as the previous sequence but also includes the Rev response element (RRE) from pHIV-YU2 cloned between the *Eco*RV and *Pvu*II sites in the 3' LTR. The SA contains HIV-1 envelope sequence 5' of the splice acceptor site for the second exon of the *tat* and *rev* genes (base 8331 in pHIV-YU2) contiguous to the end of the 3' LTR. Similarly the F4 region comprises HIV-1 envelope sequence starting immediately 3' of the same splice acceptor site (base 8341 in pHIV-YU2) while F6 extends to base 8685 in pHIV-YU2. These latter two fragments are also contiguous with the 3' LTR. The 5' *env* sequence encompasses bases 5742-7030 of pHIV-YU2 and is joined to the ppt+RRE fragment from base 9035 until the end of the 3' LTR. The F2Sal fragment covers sequence from base 7583 to 8458, thereby including the RRE, which is joined to the ppt+ fragment from base 9035 until the end of the 3' LTR. It was not possible to create an ext5 vector derivative containing the F2Sal fragment. Dotted lines in the figure represent sequence that is absent.



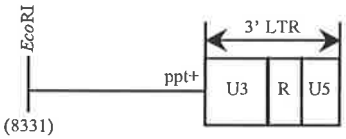
ppt+



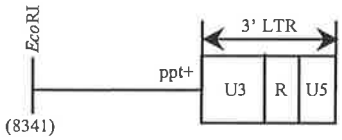
ppt+RRE



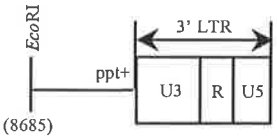
SA



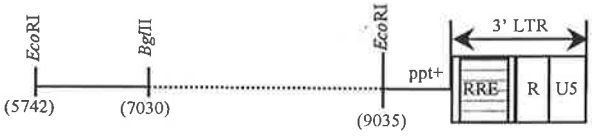
F4



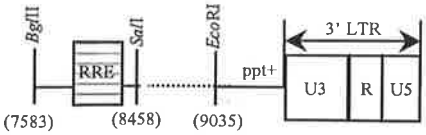
F6



5'env



F2Sal



LTRF2 (ppt+), 3' LTRF3 (SA), 3' LTRF4 (F4) and 3' LTRF6 (F6) in combination with the 3' LTRR primer to create PCR products that were cloned 5' *Eco*RI/3' *Cla*I, thereby replacing the original ppt and 3' LTR in pHIVext2 and pHIVext5. The SV40 promoter was cloned into the end-filled *Bam*HI site as described in section 5.2. For the ppt+ vector derivative EYFP was cloned into the *Eco*RI site and for the SA, F4 and F6 vector variants the EYFP sequence was cloned into the *Stu*I site of the polylinker.

The Rev response element (RRE) from pHIV-YU2 (bases 7730 - 7963, see section 4.3.4) was isolated rom pHIV-YU2 using RREF and RRER. The resulting PCR product was then blunt-end cloned between the 5' *Eco*RV and 3' *Pvu*II sites in the 3' LTR of both pHIVext2SV40EYFPppt+ and pHIVext5SV40EYFPppt+, replacing the U3 sequence. The EYFP sequence was cloned into the *Eco*RI site of the polylinker to create pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE. In these constructs the strategy for placing the RRE in the U3 region of the 3' LTR was to facilitate the incorporation of further sequence elements into the vector. However, this arrangement has the additional advantage of making the vector self-inactivating because the U3 sequence that is replaced by the RRE includes transcriptional control elements such as AP1, NFκβ and SP1 binding sites and the TATAA box sequence.

Another construct was made that included sequence from the HIV-1 envelope region encompassing the RRE. A *Sal*I fragment was isolated by PCR from pHIVext2SV40EYFPppt+ using the primers 3'LTRF2Sal and pB/ST7. Following a *Sal*I digest this sequence represented pHIV-YU2 from base 9035 to the *Sal*I site at the end of the 3' LTR. A *Bgl*II/*Sal*I restriction fragment from pHIV-YU2 (bases 7583-8458) was then isolated, and both fragments were then cloned into pHIVext2 such that they replaced the pptLTR sequence (Figure 5.2). The SV40 promoter and the EYFP cDNA (cloned into the *Stu*I site) were then incorporated as described above to create the vector pHIVF2Salext2SV40EYFP. It was not possible to make the analogous construct with the ext5 vector derivative.

In addition HIV-1 envelope sequence 5' of the RRE was tested. The envelope sequence from the *Eco*RI site at base 5742 of pHIV-YU2, to the *Bgl*II site at base 7030 was isolated and cloned into the *Eco*RI (end-filled) site of pHIVext2 and pHIVext5. The SV40 promoter was cloned into the end-filled *Bam*HI site and EYFP into the *Stu*I site both as previously described. Both of these vectors, pHIVext2SV40EYFPppt+RRE5'env and pHIVext5SV40EYFPppt+RRE5'env also contain the RRE in the 3' LTR and the 25 bp of

sequence that lie immediately upstream of the HIV-1 polypurine tract (ppt+). The sequence composition of all of these vectors is shown in Figure 5.2.

5.4 PRIMARY VECTOR ASSESSMENT

All EYFP containing vectors described in this chapter was evaluated by their ability to transduce NIH3T3 and express EYFP as measured by FACScan analysis. In this way the effect of each sequence addition on vector performance could be determined. Viral supernatant was collected from 293T cells that were co-transfected with 1 μg each of vector, pCMVΔRnr (Naldini *et al.*, 1996b; Kafri *et al.*, 1997) and pHCMV-G (Yee *et al.*, 1994) unless stated otherwise. Where indicated measurements of p24 antigen were performed on the viral supernatant collected from the transfected 293T cells. This HIV-1 core protein is produced by the *gag* gene and provides an estimate of HIV-1 viral particles (Dorfman *et al.*, 1993). All procedures were as described in Materials and methods (Chapter two).

Table 5.1 displays the viral titres obtained with each of the vector constructs described in sections 5.2 and 5.3. The first observation is that, in every instance, the longer length of *gag* sequence (1150 bp) in the ext5 vector derivative results in a higher viral titre than the ext2 variant (100 bp). The effect of *gag* length on vector titre is explored in more detail in section 5.5. From all of the sequences to be incorporated into the vector pairs, only two specific regions appear to increase viral titre. The first of these is the extra 25 bp that encompass the five T residues immediately 5' of the ppt (ie ppt+). This sequence increases viral titre 20-fold when comparing pHIVext2SV40EYFPppt to pHIVext2SV40EYFPppt+, and 12-fold for the ext5 counterpart. Maximal viral titres were noted with the inclusion of the RRE in combination with the ppt+ sequence, the highest average titre being 8.7 x 10^5 for pHIVext5SV40EYFPppt+RRE. The addition of other sequences in the vector did not appear to augment viral titres. In actual fact, with one exception, viral titres were generally seen to decrease when envelope elements additional to the ppt+ and RRE were placed in the vector construct. The exception was pHIVF2Salext5SV40EYFP, in which viral titres were comparable to those obtained from vector constructs containing both the ppt+ and RRE elements. It was hard to understand why additional sequence present in the transfer vectors would result in a decrease in viral titre. One difference in the creation of these vectors had been the restriction site used for insertion of the EYFP sequence. Initially EYFP had been cloned into the *Eco*RI site in the polylinker and the derivatives encompassing the ppt+ and RRE sequence elements had

TABLE 5.1

Titres obtained with a pair of vectors incorporating different HIV-1 sequences.

Two vectors with either 100 bp (ext2) or 1150 bp (ext5) of HIV-1 *gag* sequence were modified to include specific and more general HIV-1 sequences. One exception was pHIVF2Salext2SV40EYFP, for which no ext5 counterpart was made. For details of the vectors see sections 5.2 and 5.3. Viral titres were determined on NIH3T3 cells using viral supernatant collected from 293T cells co-transfected with each vector construct, pCMVΔRnr and pHCMV-G, by the percentage of EYFP positive cells as described in sections 2.2.18 and 2.2.21.

Vector	Titre EYFP tu/ml x 10^5
pHIVext2SV40EYFPppt	0.1 ± 0.1
pHIVext5SV40EYFPppt	0.1 ± 0
pHIVext2SV40EYFPppt+	1.4 ± 0.2
pHIVext5SV40EYFPppt+	1.6 ± 0.4
pHIVSAext2SV40EYFP*	0.4 ± 0.1
pHIVSAext5SV40EYFP*	1.5 ± 0.2
pHIVF4ext2SV40EYFP*	0.9 ± 0.7
pHIVF4ext5SV40EYFP*	1.8 ± 0.4
pHIVF6ext2SV40EYFP*	0.3 ± 0
pHIVF6ext5SV40EYFP*	1.6 ± 0.1
pHIVext2SV40EYFPppt+RRE	5.1 ± 0.1
pHIVext5SV40EYFPppt+RRE	8.7 ± 0.3
pHIVF2Salext2SV40EYFP*	6.5 ± 0.5
pHIVext2SV40EYFPppt+RRE5'env*	0.5 ± 0
pHIVext5SV40EYFPppt+RRE5'env*	0.6 ± 0.1

n = 3. * EYFP cloned into the *Stu*I site in these constructs.

also been generated in this manner. However, to facilitate cloning procedures, the EYFP sequence had been blunt-end cloned into the *Stu*I site of the polylinker in the "SA", "F4", "F6", "F2Sal" and "5'env" vector derivatives.

To investigate whether insertion of EYFP in the *Stu*I sitedid actually reduce viral titre, pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE were re-made but with the EYFP gene cloned into the *Stu*I site rather than the *Eco*RI site. By directly comparing viral titres obtained from both these sets of vectors, the effect of cloning EYFP into the *Stu*I site could be evaluated directly. These two vectors were denoted pHIVext2SV40EYFPppt+RRE(StuI) and pHIVext5SV40EYFPppt+RRE(StuI). The viral titres obtained (from a single experiment) for these two vectors were considerably reduced. For pHIVext2SV40EYFPppt+RRE(StuI) only 2 % of NIH3T3 cells were shown to be transduced compared to 26 % for pHIVext2SV40EYFPppt+RRE. The same trend was observed for the ext5 derivative, pHIVext5SV40EYFPppt+RRE(StuI) transduced 4 % of NIH3T3 cells whereas pHIVext5SV40EYFPppt+RRE resulted in 46 % transduction. The region flanking the end of the SV40 promoter and the entire polylinker extending into the EYFP coding region was sequenced in both pHIVext5SV40EYFPppt+RRE(StuI) and pHIVext5SV40EYFPppt+RRE to see if any aberrant base changes were present with negative results, the sequence of both being as predicted for EYFP cloned into the *Stu*I and *Eco*RI sites respectively. The basis for this effect was not further investigated and for all future experiments, vectors were made with the EYFP gene cloned into the *Eco*RI site of the polylinker.

However, several conclusions could still be drawn from this work. First, the ppt+ sequence, which contains an additional 25 residues 5' of the defined polypurine tract located at the junction of the 3' LTR, does increase viral titre. Second, the inclusion of RRE also appears to be required for maximal vector function and this is analysed in more detail in sections 5.10 and 5.11. Third, it appears that the longer *gag* sequence of 1150 bp (ext5) is more beneficial than the shorter 100 bp (ext2), with higher viral titres in every instance. This is followed up in more detail in the next section.

5.5 INCORPORATION OF 5' GAG SEQUENCE

A series of vectors were made containing, either no (ext1), 100 (ext2), 250 (ext3), 550 (ext4), or 1150 (ext5) bp of *gag* sequence. Two initial constructs

pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE have been described above. Another three vectors were then generated by using a common 5' primer, 5'LTRF, with the following specific primers 5'HIVext1R, 5'HIVext3R and 5'HIVext4R, to amplify the 5' LTR with corresponding *gag* sequence from pHIV-YU2, to produce ext1 (no *gag*), ext3 (250 bp of *gag*) and ext4 (650 bp of *gag*) respectively. These fragments were cloned 5' *Not*I/3' *Bam*HI into pB/SHIVLTR (section 3.2). The 3' LTR was then replaced with a 5' *Eco*RI/3' *Cla*I fragment from pHIVext2SV40EYFPppt+RRE followed by the SV40 promoter (cloned into the *Bam*HI site) and EYFP (into the *Eco*RI site) as described above. Therefore, in each case the vectors all contained the "ppt+" sequence and the RRE, both of which were shown to be required for maximal virus titres in section 5.4. The five resultant vectors are shown diagrammatically in Figure 5.3.

Comparing vector constructs in Table 5.2, there is a concomitant increase in viral titre for *gag* sequence up to 650 bp (ext4), with levels of p24 remaining fairly constant. The inclusion of another 500 bp of *gag* sequence (ext5) failed to increase viral titre. This is consistent with many other reports that the first 21-653 nucleotides of *gag* are important for packaging (Berkowitz *et al.*, 1995; Clever *et al*, 1995; Cui *et al.*, 1999; Luban and Goff, 1994; McBride *et al.*, 1997; Parolin *et al.*, 1994). Inclusion of *gag* sequences in these vectors was designed in an attempt to delineate the relationship between *gag* sequence content and vector performance. Obviously, from a bio-safety point of view, the less *gag* sequence present in the vector the less the chance of homologous recombination with helper plasmids expressing GagPol. Therefore, the amount of *gag* sequence necessary for optimal vector function needs to be balanced against the risk of homologous recombination with *gag* sequence present in packaging constructs.

There is no real difference in p24 levels obtained with the different length of *gag* sequences in the vector (Table 5.2). The ratio between functional virus titre and the amount of p24 in the supernatant of the transfected 293T cells is also displayed in Table 5.2. This ratio reflects the number of particles that actually carry a viral genome with respect to the total number of virions synthesised and secreted to the medium. This ratio increases with the length of *gag* sequence in the vector suggesting that the increase in packaging efficiency of the viral genome is directly related to the length of *gag* sequence in the vector construct. It is generally assumed that p24 is directly proportional to the number of virus particles. Nevertheless it must be noted that because the measurements of p24 were made on the supernatant of the transfected 293T cells, there is no proof that the protein is actually contained within viral particles. To show that there is no extracellular free p24, the supernatant from the transfected 293T cells needs to be subjected to ultra-

TABLE 5.2

Viral titres from a series of vectors incorporating increasing lengths of *gag* sequence.

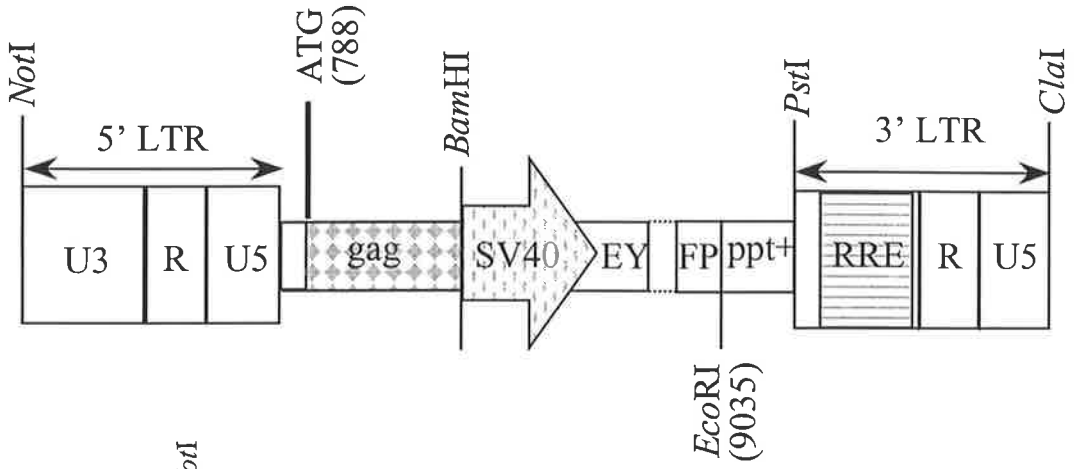
Vectors were constructed containing no (ext1), 100 (ext2), 250 (ext3), 550 (ext4) or 1150 (ext5) bp of *gag* sequence (Figure 5.3). Viral titres were determined on NIH3T3 cells using viral supernatant collected from 293T cells co-transfected with each vector construct, pCMVΔRnr and pHCMV-G. Transduced NIH3T3 cells were assayed for EYFP expression by FACScan analysis as described in section 2.2.21. Measurements of p24 were made on the viral supernatant.

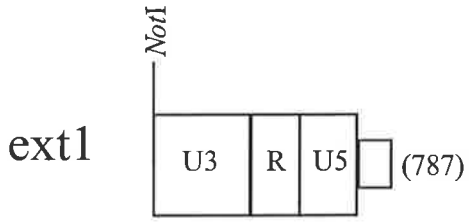
Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext1SV40EYFPppt+RRE	2.6 ± 0.2	330 ± 43	787 ± 85	5
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext3SV40EYFPppt+RRE	3.6 ± 0.6	383 ± 65	955 ± 177	6
pHIVext4SV40EYFPppt+RRE	7.6 ± 2.8	464 ± 250	1845 ± 635	12
pHIVext5SV40EYFPppt+RRE	7.5 ± 0.9	336 ± 74	2329 ± 542	12

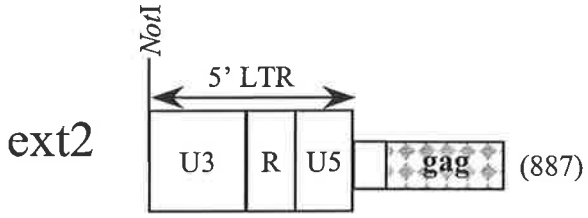
FIGURE 5.3

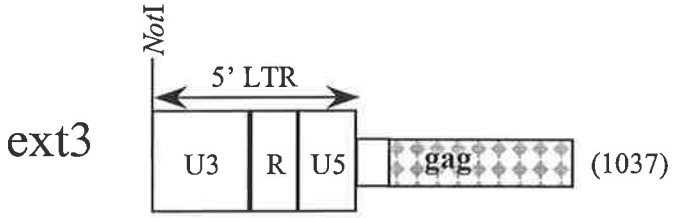
Incorporation of different lengths of *gag* sequence.

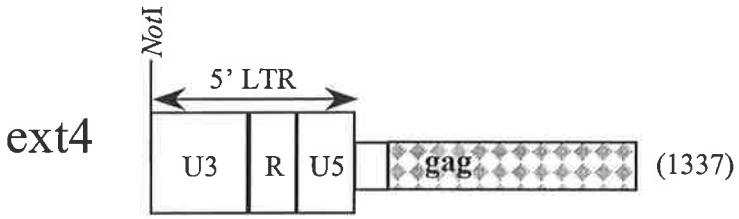
The *gag* sequence elements in the ext1-5 constructs is shown underneath the top complete vector construct. In each case the sequence was joined to the SV40 promoter, EYFP reporter gene, polypurine tract + (ppt+) and the 3' LTR containing the Rev response element (RRE) in the U3 region. The number in parenthesis indicates the base number in the pHIV-YU2 sequence that the *gag* sequence extends to.

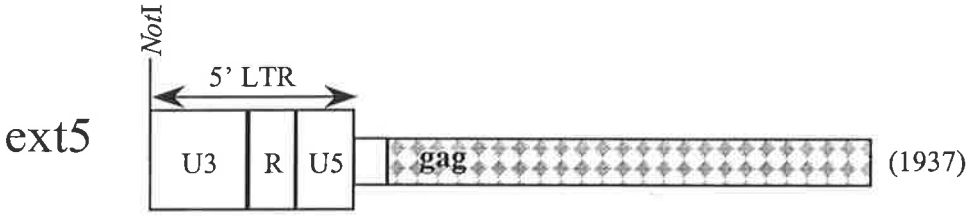

NotI
ATG (788)
BamHI
PstI
ClaI
5' LTR
3' LTR
U3
R
U5
gag
SV40
EY
FP
ppt+
RRE
R
U5
EcoRI (9035)


ext1
NotI
U3
R
U5
(787)


ext2
NotI
5' LTR
U3
R
U5
gag
(887)


ext3
NotI
5' LTR
U3
R
U5
gag
(1037)


ext4
NotI
5' LTR
U3
R
U5
gag
(1337)


ext5
NotI
5' LTR
U3
R
U5
gag
(1937)

centrifugation to pellet the viral particles and the sediment lysed, and assayed for p24. This line of experimentation was not explored further.

5.6 MUTAGENESIS OF THE MAJOR SPLICE DONOR AND GAG INITIATION CODON

The 390 bp oligonucleotide fragment described in section 3.2 (see also Figure 3.1) that was used for the construction of the β-glucuronidase vector contains sequence changes from the native pHIV-YU2 sequence that were designed to improve bio-safety. The major splice donor sequence (underlined) was mutated from GGTGAGTACGCC to GCTCAGTA*GGG*C, the additional alterations shown in italics were designed to maintain the secondary structure of the region as proposed by Harrison and Lever (1992) on the basis that the secondary structure of this region is functionally important. Additionally, the *gag* initiation codon was mutagenised to a stop (TAG). This 390 bp oligonucleotide fragment was inserted into the ext1 and ext2 EYFP containing vectors described in section 5.5. For the ext1 vector derivative a PCR product was generated from pHIVSV40βgluc (section 3.2) using primers 5'HIVext1 and pB/SR. This fragment was then restricted with *Bam*HI and *Not*I, gel purified and cloned into the same two sites of pB/SHIVLTR (section 3.2). The 3' LTR was subsequently replaced with a 5' *Eco*RI/3' *Cla*I fragment from pHIVext2SV40EYFPppt+RRE to include the ppt+ and RRE sequences. The SV40 promoter was then cloned into the end-filled *Bam*HI site and subsequently the EYFP coding sequence was cloned into the *Eco*RI site, both as detailed earlier (section 5.2). For the ext2 vector derivative, a *Stu*I/*Sal*I fragment was isolated from pHIVext2SV40EYFPppt+RRE and cloned into the same two sites of pHIVSV40βgluc (section 3.2). This effectively replaced the 3' sequence in this vector between the *Stu*I and *Sal*I sites encompassing sequence from the EYFP reporter gene until the end of the 3' LTR. The two constructs produced were termed pHIVext1chSV40EYFPppt+RRE and pHIVext2chSV40EYFPppt+RRE. It should be noted that only the latter encompasses the mutagenised *gag* initiation codon, the ext1 sequence terminates immediately 5' from the *gag* ATG.

Table 5.3 shows viral titres obtained from these two vectors as compared with the analogous two vectors containing native pHIV-YU2 sequence (ie the splice donor and *gag*

TABLE 5.3

Mutagenesis of the major splice donor site and the *gag* initiation codon in the "ext1" and "ext2" vector derivatives.

The two vectors containing no (ext1) or 100 (ext2), bp of *gag* sequence was altered from their native pHIV-YU2 to create "ext1ch" and "ext2ch". Such sequence modifications included mutagenesis of both the major splice donor site and the *gag* initiation codon (for details see text). Viral titres were determined on NIH3T3 cells using viral supernatant collected from 293T cells co-transfected with each vector construct, pCMVΔRnr and pHCMV-G, as described in sections 2.2.18 and 2.2.21. Measurements of p24 were made on the viral supernatant.

Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext1SV40EYFPppt+RRE	2.6 ± 0.2	330 ± 43	787 ± 85	5
pHIVext1chSV40EYFPppt+RRE	0.49 ± 0.1	402 ± 89	120 ± 21	6
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext2chSV40EYFPppt+RRE	2.6 ± 0.2	440 ± 114	629 ± 127	6

ATG), pHIVext1SV40EYFPppt+RRE and pHIVext2SV40EYFPppt+RRE. From this table it is evident that the introduction of the splice donor mutation into the ext1 vector decreased viral titre five-fold although the quantity of p24 remained the same. In the ext2 vector the combination of both the splice donor and *gag* initiation codon mutations had no effect, viral titres and p24 levels were similar to those of pHIVext2SV40EYFPppt+RRE. It has been demonstrated previously that mutations in the splice donor along with the deletion of the splice acceptor site, resulted in a small drop in viral titre (10 %), despite northern analysis showing full-length viral RNA expression was reduced by 50 % (Cui *et al.*, 1999). It is note-worthy that these vectors all contained *gag* sequence. The drop in titre seen with the ext1 derivative may therefore be a consequence of the absence of *gag* sequence. As the levels of p24 remain similar, the number of virions present is unlikely to be altered. One possibility is that the genomic RNA is inefficiently packaged or expression of the transgene inhibited because of the mutation of the splice donor. However, the latter is unlikely because the expression cassette for EYFP is actually downstream from the mutation. Interestingly, it has been reported that the splice donor site imposes a suppressive effect on the action of the polyadenlyation site in the 5' LTR. (Ashe *et al.*, 1997). Northern analysis would give some indication as to whether the amount of full-length viral RNA seen in the ext1 vector containing the splice donor mutation was reduced.

This experiment has shown that the addition of the first 100 bp of *gag* into the vector did not increase viral titre, although the longer *gag* sequence did enable the introduction of the splice donor and ATG mutations without a decrease in titre. However, as noted earlier, increasing the length of *gag* sequence up to 650 bp (ext4 vector, see Table 5.2) did result in a higher viral titre. In order to develop an optimum vector, mutagenesis of the major splice donor site and *gag* initiation codon needs to be incorporated into the ext4 vector derivative to see what effect, if any, is observed. Such changes are clearly beneficial from a bio-safety perspective.

5.7 INCORPORATION OF ENVELOPE REGIONS

5.7.1 3' ENVELOPE SEQUENCES

Regions of the HIV-1 envelope sequence were incorporated into vector constructs to assess whether any effect on viral titre could be attributed to such elements. Sequences from the 3' end of the *env* gene were included in PCR products generated using the

primers 3' LTRF3, 3' LTRF4 and 3' LTRF6 as described in section 5.3. These were the same pHIV-YU2 *env* sequences that were incorporated into the initial vectors in section 5.3. However, the EYFP coding sequence was cloned into the *Eco*RI site in these vectors. These *env* regions were incorporated into vectors containing 100 bp (ext2) and 1150 bp (ext5) of *gag* sequence, the SV40 promoter, EYFP gene and RRE. A fragment termed F4ΔF6 (pHIV-YU2 bases 8341 - 8687, ie sequence between the F4 and F6 primers) was generated by PCR using the 3'LTRF4 primer and F6RevMfeI. This fragment was then cloned using flanking *Eco*RI and *Mfe*I sites incorporated by the PCR primers into *Eco*RI restricted pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE. The EYFP sequence was re-cloned back into these two vectors into a reformed *Eco*RI site at the 5' end of the F4ΔF6 fragment. This created two vectors pHIVext2SV40EYFPF4ΔF6ppt+RRE and pHIVext5SV40EYFPF4ΔF6ppt+RRE which are the same as pHIVext2SV40EYFPF4RRE and pHIVext5SV40EYFPF4RRE respectively except with the sequence between the 3'LTRF2 and 3'LTRF6 primers (ie bases 8687 - 9035 in pHIV-YU2) deleted. The sequence composition of the vectors in this section is shown in Figure 5.4.

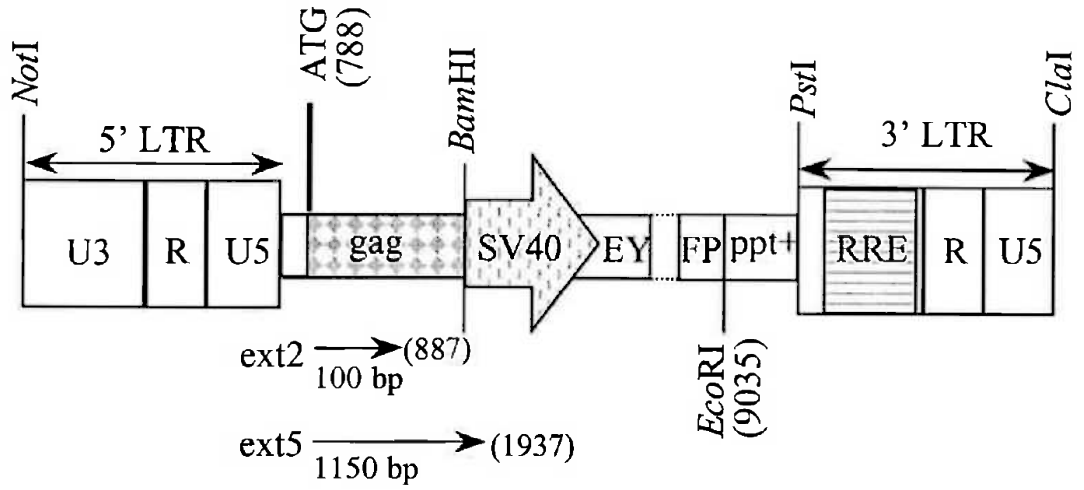
Table 5.4 shows the viral titres obtained for all these vectors as directly compared with the ext2 and ext5 derivatives which contain no *env* sequence other than the RRE, ie pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE. From this table it is evident that the insertion of the F4 fragment increases viral titre which is more notable in the ext2 vector, and levels of p24 are unchanged. The region designated as F4 comprises 3' HIV-1 envelope sequence from base 8341 of pHIV-YU2 to the end of the 3' LTR. The splice acceptor site for the second exon of the *tat* and *rev* genes is therefore not included. On the other hand, the F3 fragment contains the same envelope region but extends further 5' to base 8331 in pHIV-YU2 to include the splice acceptor site. The difference in titre observed between these two vector constructs would suggest that the splice acceptor site has a negative effect on viral titre. However, as this construct lacks the lariat sequence associated with the splice acceptor it is unclear whether the site is functional. Northern analysis would be required to assess if splicing was in fact occurring. Levels of p24 antigen remained similar, suggesting approximately similar levels of virion production with or without inclusion of the splice acceptor site.

The inclusion of the F6 fragment into the vector did not improve vector function. The F6 envelope region is also included in the F4 fragment so a third fragment was generated containing the F4 envelope sequence without the F6 region to try and delineate

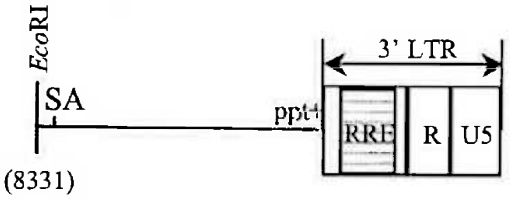
FIGURE 5.4

Addition of 3' *env* sequence elements.

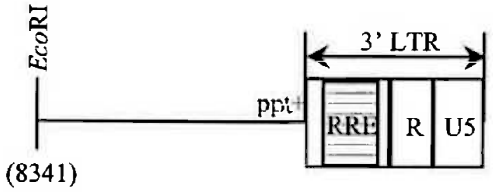
The basic starting vectors pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE are shown in full. Underneath each of the 3' *env* sequences that has been incorporated is depicted. The F3 fragment contains HIV-1 envelope sequence 5' of the splice acceptor site for the second exon of the *tat* and *rev* genes (base 8331 in pHIV-YU2) contiguous to the end of the 3' LTR. The F4 region comprises HIV-1 envelope sequence immediately 3' of the same splice acceptor site (base 8341 in pHIV-YU2) and F6 extends to base 8685 in pHIV-YU2. These latter two fragments are also contiguous with the 3' LTR. The F4ΔF6 contains HIV-1 *env* sequence from bases 8341 - 8687 in pHIV-YU2, and this fragment is the same as F4 except bases 8685 - 9035 are deleted. Dotted lines in the figure represent sequence that is absent.



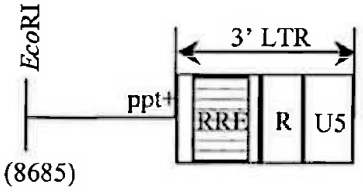
F3



F4



F6



F4ΔF6

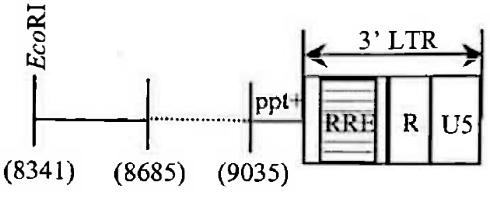


TABLE 5.4
Incorporation of 3' *env* sequences into the transfer vector.

The two vectors containing 76 (ext2) or 1150 (ext5), bp of *gag* sequence were modified to incorporate 3' *env* regions from pHIV-YU2 (Figure 5.4). The F3 derivative contains 3' *env* sequence up to base 8331 of pHIV-YU2, which is immediately 5' of the splice acceptor site for *tat* and *rev*. Immediately 3' of the same splice acceptor (base 8341 of pHIV-YU2) is included in the F4 fragment and F6 includes up to base 8685. Bases 8341 - 8687 of the pHIV-YU2 3' *env* sequence have been incorporated into the fragment termed F4ΔF6 (for details see text). Viral titres were determined on NIH3T3 cells using viral supernatant collected from 293T cells co-transfected with each vector construct, pCMVΔRnr and pHCMV-G, as described in sections 2.2.18 and 2.2.21. Measurements of p24 were made on the viral supernatant.

ext2 Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext2SV40EYFPF3RRE	2.7 ± 0.8	381 ± 139	757 ± 226	9
pHIVext2SV40EYFPF4RRE	5.1 ± 0.7	431 ± 145	1322 ± 411	9
pHIVext2SV40EYFPF6RRE	2.2 ± 0.7	463 ± 169	502 ± 81	9
pHIVext2SV40EYFPF4ΔF6ppt+RRE	3.1 ± 0.8	193 ± 60	1630 ± 183	9

ext5 Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext5SV40EYFPppt+RRE	7.5 ± 0.9	336 ± 74	2329 ± 542	12
pHIVext5SV40EYFPF3RRE	4.7 ± 1.3	363 ± 150	1647 ± 1014	9
pHIVext5SV40EYFPF4RRE	9.7 ± 2.7	330 ± 178	3760 ± 1657	9
pHIVext5SV40EYFPF6RRE	6.9 ± 2.6	544 ± 162	1304 ± 377	9
pHIVext5SV40EYFPF4ΔF6ppt+RRE	6.6 ± 2.5	205 ± 48	3144 ± 544	9

more specifically what part of the HIV-1 envelope region is having a positive effect on viral titre. The vector containing this fragment (F4ΔF6) did not cause the same increase in vector performance as the F4 region.

Interestingly, the inclusion of this F4ΔF6 fragment caused a decrease in viral production as measured by p24 antigen in both the ext2 and ext5 vector derivatives. Despite the reduction in viral particles when the F4ΔF6 fragment is included in the vector, the titres are similar to those obtained with pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE. The titre/ng of p24 measures the packaging efficiency of the retroviral genome into the virion particle, assuming that genomic RNA levels are the same and that post-transduction events are of equal efficiency. To evaluate this more fully, the levels of vector RNA in virions and in the cytoplasm of the transfected 293T cells could be measured by RNase protection assays. This can measure the amounts of properly initiated vector transcripts and distinguish the vector RNA from any contaminating vector DNA from the transfection. The ratio of virion to cytoplasmic RNA can then be equated to packaging efficiency. Northern blot analysis of RNA from transfected 293T cells could also be used to measure levels of full-length vector transcripts. However, for the development of a vector construct the most meaningful measurement is still viral titre. In conclusion, the region of HIV-1 envelope sequence from the 3' end to base 8341 of pHIV-YU2 (designated as F4) does have a moderate influence on vector function and this is most likely to be a consequence of more efficient viral packaging (ie titre/ng of p24) of full-length genomic viral RNA, or perhaps increased genomic RNA expression.

5.7.2 5' ENVELOPE SEQUENCES

Two vectors containing either 100 (ext2) or 1150 (ext5) bp of *gag* were constructed that contained sequence from the 5' end of the envelope gene of pHIV-YU2. An *Eco*RI/*Bgl*II fragment encompassing the 5' end of the *env* gene from pHIV-YU2 (bases 5742 - 7030) was cloned into pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE immediately 3' of the EYFP sequence to give pHIVext2SV40EYFP5'envppt+RRE and pHIVext5SV40EYFP5'envppt+RRE. The sequence incorporated extended from base 5742 to 7030 and was joined to the 3' LTR that included the ppt+ sequence and the RRE. Viral titres obtained with these two vectors pHIVext2SV40EYFP5'envppt+RRE and pHIVext5SV40EYFP5'envppt+RRE were directly compared to those from the same two vectors without the 5'env sequence and the results are shown in Table 5.5. Inclusion of this 5' *env* sequence significantly increases viral titre in both the ext2 and ext5 derivative ($p < 0.01$ in both cases), and again a larger

TABLE 5.5

Incorporation of 5' *env* sequences into the transfer vector.

The two vectors containing 100 (ext2) or 1150 (ext5), bp of *gag* sequence were modified to incorporate a 5' *env* region from pHIV-YU2 (bases 5742 - 7030). These two vectors termed pHIVext2SV40EYFP5'envRRE and pHIVext5SV40EYFP5'envRRE were directly compared with their counterparts without the 5' *env* sequence. Viral titres were determined on NIH3T3 cells using viral supernatant collected from 293T cells co-transfected with each vector construct, pCMVΔRnr and pHCMV-G, as described in sections 2.2.18 and 2.2.21. Measurements of p24 were made on the viral supernatant.

Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext2SV40EYFP5'envRRE	4.2 ± 0.8	446 ± 116	983 ± 209	9
pHIVext5SV40EYFPppt+RRE	7.5 ± 0.9	336 ± 74	2329 ± 542	12
pHIVext5SV40EYFP5'envRRE	9.0 ± 0.9	392 ± 126	2525 ± 778	9

effect is observed with the ext2 vector. Small increases in both p24 and titre/ng p24 are also seen, however these are not significant.

5.8 INCLUSION OF THE CENTRAL POLYPURINE TRACT

The vectors described so far all contain a polypurine tract sequence that is located at the 5' border of the 3' LTR. HIV-1 and other lentiviruses contain another polypurine tract/termination signal within the *pol* coding region. It is likely that this central polypurine tract (cppt) is used as a second origin for plus-strand viral DNA synthesis (Charneau and Clavel, 1991; Charneau *et al.*, 1992).

To determine whether the presence of the cppt in the transfer vector improves function, bases 4769 - 4915 of pHIV-YU2 that encompass the cppt and associated termination sequence, were included. A PCR product was isolated as a 5' *Cla*I/ 3' *Bst*BI from pHIV-YU2 using PCR primers cpptClaI and cpptBstBI, and cloned into pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE immediately 5' of the SV40 promoter. This created two vector constructs pHIVext2cpptSV40EYFPppt+RRE and pHIVext5cpptSV40EYFPppt+RRE that contain both HIV-1 polypurine tracts.

Comparing these constructs in Table 5.6 shows that inclusion of the cppt increases viral titres in both the ext2 and ext5 vector derivatives ($p < 0.001$ in both instances). This increase in viral titre attributed to the cppt has also been reported by other investigators who have shown that this sequence not only increases viral titres but also plays a pivotal role in the transduction of non-dividing cells (Mautino *et al.*, 2000; Zennou *et al.*, 2000). The levels of p24 antigen are not significantly different in the ext2 derivative but are significantly increased (although to a lesser degree than the viral titres) in the ext5 containing vector ($p < 0.005$). The increase in titre/ng of p24 upon inclusion of the cppt is significant in the ext2 vector construct ($p < 0.001$) but not significant in the ext5 variant. These findings would indicate that although the cppt clearly augments viral titres, the extent of the effect is influenced by the length of *gag* sequence in the vector construct. This is in agreement with most other sequences tested displaying a greater effect with the ext2 derivative than with the ext5, which is probably just because the ext2 vector is less efficient in the first place making any effect more obvious.

TABLE 5.6

Inclusion of the central polypurine tract.

The central polypurine tract (cppt) present in HIV-1 from bases 4769 - 4915 (pHIV-YU2) was incorporated into the ext2 and ext5 vector derivatives. Within these vectors was also the extended polypurine tract (ppt+) that is located at the junction of the 3' LTR in HIV-1. Vector constructs containing both of these ppt sequences were compared to their counterparts with only one ppt (ie ppt+). Viral supernatant collected from 293T cells co-transfected with each of these vectors, pCMVΔRnr and pHCMV-G was used to transduce NIH3T3 cells. Viral titres were measured by FACScan analysis as determined by the percentage of cells expressing EYFP (section 2.2.21). Viral supernatant was also assayed for p24.

Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext2cpptSV40EYFPppt+RRE	5.7 ± 1.5	376 ± 119	1579 ± 358	9
pHIVext5SV40EYFPppt+RRE	7.5 ± 0.9	336 ± 74	2329 ± 542	12
pHIVext5cpptSV40EYFPppt+RRE	10.4 ± 1.2	484 ± 125	2296 ± 688	9

5.9 THE RRE AND ITS POSITIONING

The Rev-response element (RRE) is a *cis*-acting sequence located within the *env* gene of HIV-1 that is required for Rev to regulate the export of structural mRNAs from the nucleus to the cytoplasm (for more details see section 1.3.1.4). To determine whether the presence of the RRE in the vector influenced the efficiency of gene transfer, the RRE was deleted from three of the HIV-1 vectors described above. The *Eco*RI/*Sal*I region (ie ppt+RRELTR) of pHIVext1SV40ppt+RRE, pHIVext2SV40ppt+RRE and pHIVext5SV40ppt+RRE was removed and replaced with an *Eco*RI/*Sal*I fragment from pHIVext2SV40EYFPppt+ (section 5.3) containing only the ppt+ sequence and the 3'LTR. The EYFP coding sequence was then blunt-end cloned into the *Eco*RI site (end-filled), creating the final constructs; pHIVext1SV40EYFPppt+, pHIVext2SV40EYFPppt+ and pHIVext5SV40EYFPppt+.

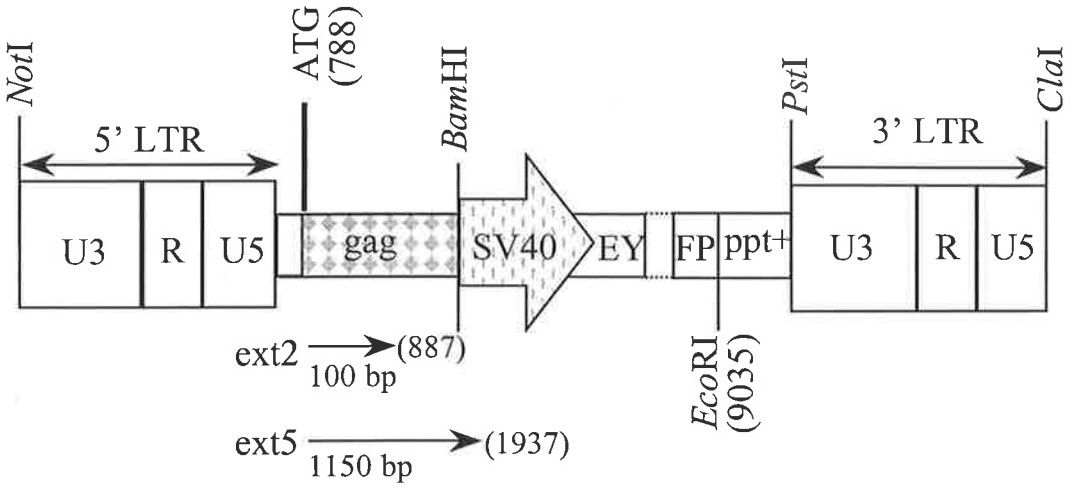
The original vectors containing the RRE did so in the U3 region of the 3' LTR. Deletion of this U3 region in the 3' LTR does not alter the phenotype of HIV-1 derived vectors. This has been shown by the characterisation of self-inactivating HIV-1 based vectors (Zufferey *et al.*, 1998; Iwakuma *et al.*, 1999), although it should be noted that Iwakuma *et al.* (1999) did observe a small decrease in titre with larger U3 deletions. Another vector was constructed with the RRE cloned immediately 5' of the SV40 sequence (pHIVext2cRRESV40EYFPppt+). These vectors are depicted in Figure 5.5. An ext5 counterpart for the RRE in this central position was not constructed. As expected Table 5.7 shows that deletion of the RRE from the vector led to a marked reduction in vector titre, although the levels of p24 remain unchanged. This is consistent with the known RRE effect on cytoplasmic viral RNA expression. Previous observations have shown that removal of RRE from HIV-1 derived vectors abolishes cytoplasmic localisation of full-length viral RNA (Kaye *et al.*, 1995; Parolin *et al.*, 1994). An alternative effect of RRE on vector function independent of the known contribution of RRE on cytoplasmic viral RNA expression has been suggested by Cui *et al.* (1999) who demonstrated similar amounts of cytoplasmic full-length RNA in the absence or presence of the RRE. In the absence of any northern analysis on the transfected 293T cells in these experiments it is impossible to say how the RRE is exerting its effect in the vectors described above.

The results shown here also reveal that the effect of RRE on the efficiency of gene transfer is dependent on the length of *gag* sequences included in the vector. This is in agreement with Parolin *et al.* (1994) who showed a positive influence of RRE on gene transfer to be dependent on the amount of *gag* sequence present in the vector. This

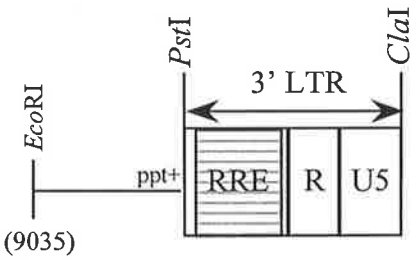
FIGURE 5.5

Incorporation of the Rev response element (RRE).

The vector pHIVext1SV40EYFPppt+ is shown at the top of the figure. The extra length of *gag* sequence incorporated in pHIVext2SV40EYFPppt+ and pHIVext5SV40EYFPppt+ is shown directly underneath. The ppt+RRE 3' modification is shown in a) depicting the incorporation of the Rev response element (RRE) into the *Eco*RV and *Pvu*II sites in the 3' LTR. For the ext2 vector derivative the RRE was cloned in a more central position in the vector (ext2cRRE) which was immediately 5' of the SV40 early promoter, shown in b).



a) **ppt+RRE**



b) **ext2cRRE**

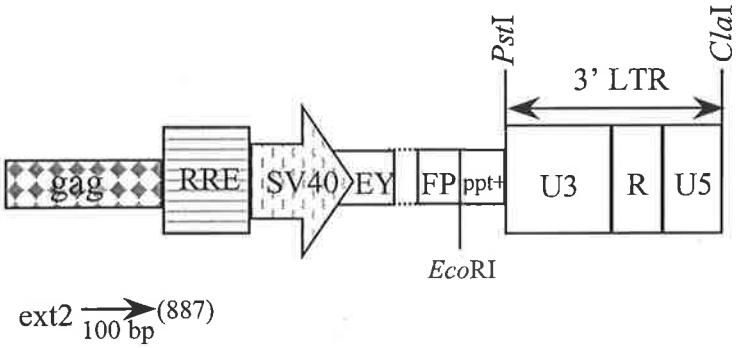


TABLE 5.7
Incorporation of the Rev-response element.

The HIV-1 Rev-response element (RRE) was incorporated into the U3 region of the 3' LTR in the ext1, 2 and 5 vector deivatives. For the ext2 containing vector the RRE was cloned immediately 5' of the SV40 promoter instead of in the 3' LTR (Figure 5.5). Viral supernatant collected from 293T cells co-transfected with each of these vectors, pCMVΔRnr and pHCMV-G was used to transduce NIH3T3 cells. This medium was also assayed for p24. Gene transfer to NIH3T3 cells was measured by FACScan analysis as described in section 2.2.21.

Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext1SV40EYFPppt+	0.2 ± 0	355 ± 8	47 ± 4	3
pHIVext1SV40EYFPppt+RRE	2.6 ± 0.2	330 ± 43	787 ± 85	5
pHIVext2SV40EYFPppt+	0.5 ± 0.2	346 ± 103	139 ± 40	6
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext2cRRESV40EYFPppt+	4.6 ± 2.2	359 ± 123	1297 ± 162	12
pHIVext5SV40EYFPppt+	0.1 ± 0	304 ± 18	33 ± 9	3
pHIVext5SV40EYFPppt+RRE	7.5 ± 0.9	386 ± 36	1984 ± 206	12

positive effect apparently resulted from the effect of Rev on the accumulation of unspliced, packaged vector RNA in the cytoplasm rather than a direct effect of the RRE on packaging efficiency. The quantity of vector RNA in the cytoplasm of the transfected 293T cells described here however, has not been determined. Incorporation of the RRE in a more central position in the vector (ie 5' of SV40) did seem to have an advantage over its inclusion in the 3' LTR, at least in terms of the ext2 derivative (comparison of pHIVext2SV40EYFPppt+RRE and pHIVext2cRRESV40EYFPppt+).

5.10 SUSTITUTION OF THE RRE WITH CONSTITUTIVE NUCLEAR EXPORT SIGNALS

The requirement for Rev in *trans* and the RRE in *cis* has been shown to be at least partially substituted for by other *cis*-acting elements known as RNA constitutive transport elements (CTE) (Bray *et al.*, 1994; Zolotukhin *et al.*, 1994). CTEs are stem-loop RNA structures that can rescue RNA from splicing (Ernst *et al.*, 1997). They interact with the nuclear export machinery to facilitate the export of unspliced RNA to the cytoplasm (Tang *et al.*, 1997; Gruter *et al.*, 1998). It may be possible, therefore, to design a Rev/RRE independent HIV-1 virus production system by replacing the RRE in the vector with a CTE. To this end, three different constitutive transport elements were incorporated into the ext2 and ext5 vector derivatives.

The mouse histone H2A nuclear export signal (H2A, Huang and Carmichael, 1997, Genbank accession number X16495, bases 189 - 298) was isolated by PCR from total murine genomic DNA using primers H2A5' and H2A3'. The PCR product was then cloned as a 5' *Bam*HI/3' *Eco*RI fragment into pBluescript II SK (+) *via* homologous restriction sites incorporated into the PCR primers. The fragment was sequenced for verification, excised with *Bam*HI/*Eco*RI and then end-filled. The woodchuck hepatitis virus post-transcriptional regulatory element (WHVPRE, Donello *et al.*, 1998, Genbank accession number JO4514, bases 1094 - 1684) was generated by PCR from pWHV8 using the primers WHVERVF and WHVPVUR. The PCR product was isolated as a 5' *Eco*RV/ 3' *Pvu*II fragment due to homologous restriction sites incorporated *via* the PCR primers. The Mason-Pfizer monkey virus constitutive transport element (MPMVCTE, Pasquinelli *et al.*, 1997 Genbank accession number AF033815, bases 7386 - 7554) was synthesised and cloned into the *Eco*RV site of pBluescript SK II (+) as described in section 4.3.4. These three elements (H2A, WHVPRE and MPMVCTE) were then cloned into

pATXppt+LTR. This shuttle vector was created by cloning an *Eco*RI/*Sal*I fragment from pHIVext2SV40EYFPppt+ into pATX. The H2A was cloned as an *Eco*RI/*Bam*HI fragment, the WHVPRE as an *Eco*RV/*Pvu*II fragment and the MPMCTE as an *Eco*RV fragment. Where necessary fragments were end-filled with Klenow and cloned between the *Eco*RV and *Pvu*II sites within the 3' LTR of pATXppt+LTR replacing the U3 sequence between these two sites. An *Eco*RI/*Sal*I fragment was subsequently isolated from each of these three clones and cloned into the same two sites of pHIVext2SV40ppt+RRE and pHIVext5SV40ppt+RRE, thereby replacing the *Eco*RI/*Sal*I fragment encompassing the ppt+RRELTR sequence in these two vectors. The blunt-ended *Xba*I fragment representing the EYFP sequence was then blunt-end cloned into *Eco*RI site of the resulting constructs to give pHIVext2SV40EYFPppt+WHVPRE, pHIVext5SV40EYFPppt+WHVPRE, pHIVext2SV40EYFPppt+H2A, pHIVext5SV40EYFPppt+H2A, pHIVext2SV40EYFPppt+CTE and pHIVext5SV40EYFPppt+CTE. These vectors are shown in Figure 5.6.

Table 5.8 shows viral titres obtained with these vectors as compared to vectors containing the RRE. It is evident that, in each case, viral titre is reduced demonstrating that none of these constitutive transport elements can substitute effectively for the function of RRE/Rev. This supports earlier findings that have shown that although HIV-1 constructs containing a CTE in place of the RRE can replicate in culture, they do so with reduced efficiency (Bray *et al.*, 1994; Zolotukhin *et al.*, 1994). There are many conflicting reports as to the efficiency of CTEs which appears to depend on the assay system. For example, a stable HIV-1 packaging system independent of RRE/Rev has been described using the MPMV CTE (Srinivasakumar *et al.*, 1997). However, Bray *et al.* (1994) and Kim *et al.* (1998) supported the observation described here of incomplete RRE/Rev replacement with MPMV CTE. The apparent discrepancy may be due to the different CTE containing fragments used, in the study by Srinivasakumar *et al.* (1997) both the MPMV CTE and its associated polyadenylation signal were used. Another report (Gasmi *et al.*, 1999) showing that the CTE derived from the MPMV genome could partially substitute for RRE/Rev suggested that because the Rev protein is supplied in the transfection, large amounts of it are available to efficiently transport RRE-containing transcripts from the nucleus to the cytoplasm. In comparison, the CTE from MPMV relies on the interaction with endogenous factors that may be in limited supply. On the other hand, the RRE and Rev may just be more efficient than the CTE to transport HIV-1

TABLE 5.8

Substitution of RRE with constitutive transport elements.

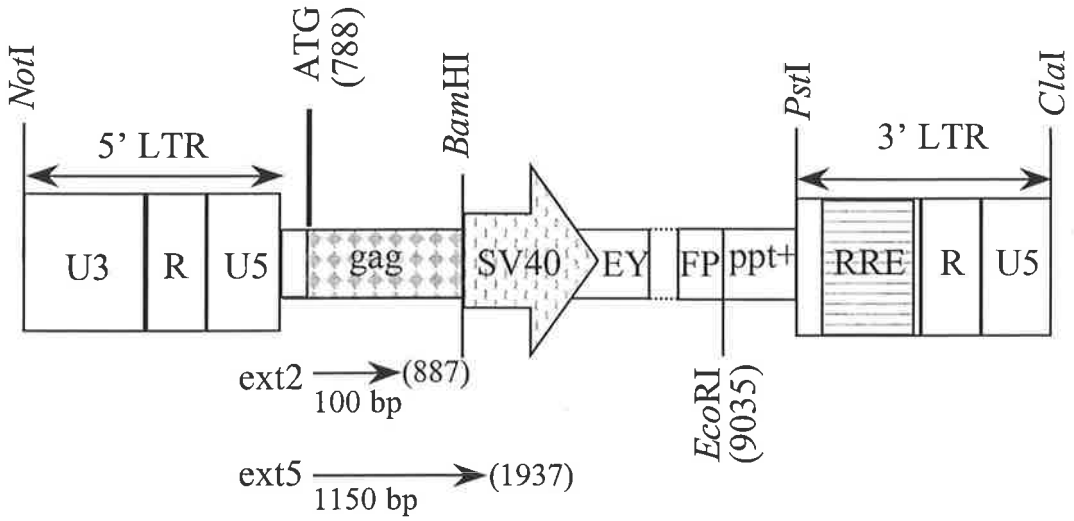
The constitutive transport elements, woodchuck hepatitis virus post-transcriptional regulatory element (WHVPRE), the Mason-Pfizer monkey virus constitutive transport element (CTE) and the mouse histone H2A gene nuclear export signal (H2A), were substituted for the Rev-response element (RRE) in the ext2 and ext5 vector derivatives (Figure 5.6). Viral supernatant collected from 293T cells co-transfected with each of these vectors, pCMVΔRnr and pHCMV-G was used to transduce NIH3T3 cells. This medium was also assayed for p24. Gene transfer to NIH3T3 cells was measured by FACScan analysis as described in section 2.2.21.

Vector	Titre EYFP tu/ml x 10^5	p24 ng/ml	Titre/ng p24	n=
pHIVext2SV40EYFPppt+RRE	2.6 ± 1.2	316 ± 132	835 ± 144	12
pHIVext2SV40EYFPppt+WHVPRE	0.7 ± 0.1	411 ± 78	176 ± 45	6
pHIVext2SV40EYFPppt+CTE	0.2 ± 0.1	262 ± 64	77 ± 41	6
pHIVext2SV40EYFPppt+H2A	0.3 ± 0.2	367 ± 189	71 ± 37	6
pHIVext5SV40EYFPppt+RRE	7.5 ± 0.9	336 ± 74	2329 ± 542	12
pHIVext5SV40EYFPppt+WHVPRE	0.2 ± 0.1	276 ± 20	73 ± 23	3
pHIVext5SV40EYFPppt+CTE	0.1 ± 0	228 ± 34	32 ± 1	3
pHIVext5SV40EYFPppt+H2A	0.1 ± 0	201 ± 31	45 ± 14	3

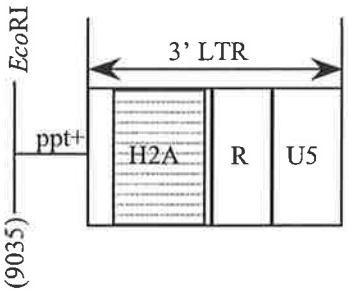
FIGURE 5.6

Incorporation of alternative transport elements.

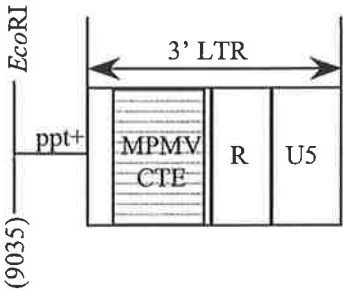
The basic starting vectors pHIVext2SV40EYFPppt+RRE and pHIVext5SV40EYFPppt+RRE are shown at the top of the figure. The mouse histone H2A gene nuclear export signal (H2A), Mason-Pfizer monkey virus constitutive transport element (MPMVCTE) and the woodchuck hepatitis virus post-transcriptional regulatory element (WHVPRE) were cloned into the *Eco*RV and *Pvu*II sites in the 3' LTR thereby replacing the RRE in the two initial vectors.



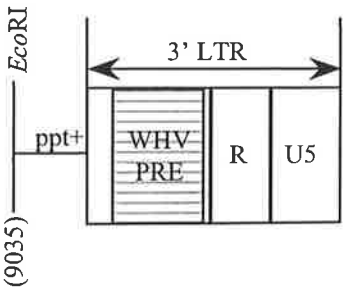
H2A



MPMVCTE



WHVPRE



encoded messages. It would be possible to elucidate the efficiency of RRE and CTE to transport HIV-1 encoded messages by using stable cell lines expressing the HIV-1 Rev protein.

It is also plausible that Rev has an additional function to the already understood Rev/RRE interaction that is responsible for the export of unspliced transcripts from the nucleus to the cytoplasm. Northern analysis would be required to see if genomic RNA is more abundant in the presence of the Rev protein. Secondly, the level of EYFP expression in the transfected 293T cells could be measured to again determine if the Rev protein augments this. Such analysis would effectively provide an estimation of cytoplasmic RNA.

5.11 CONCLUDING DISCUSSION

Evaluating the performance of vector constructs based on the inclusion of specific sequence elements as shown in this chapter, demonstrates that a functional, although not very efficient, vector can be produced that contains a relatively small amount of HIV-1 sequence. However, the efficiency of the vector construct can be enhanced by the addition of several sequence elements. The minimum amount of HIV-1 sequence required for a functional vector encompasses the 5' LTR and contiguous sequence up to, but not including, the Gag initiation codon (ext1), the polypurine tract (ppt) and the 3' LTR. These sequences obviously include several vital *cis* acting elements for many steps in HIV-1 replication. The terminal repeat (R) and the unique 5' sequence (U5) within the 5' LTR, as well as the primer binding site are important for the initiation and translocation steps of reverse transcription (Berkhout *et al.*, 1995; Harrich *et al.*, 1996; Isel *et al.*, 1995; Wakefield and Morrow; 1996; Wakefield *et al.*, 1996). Sequences throughout the 5' untranslated region are believed to be required in the dimerisation process (Sakaguchi *et al.*, 1993; Sundquist and Heaphy, 1993), particularly a stem-loop (SL1) located between the primer binding site and the major splice donor (Muriaux *et al.*, 1995; Paillart *et al.*, 1996; Skripkin *et al.*, 1994). For expression of a transgene in target cells the SV40 early promoter and EYFP coding sequence has also been included.

Several sequences in HIV-1 have been implicated by numerous deletion experiments in the dimerisation and encapsidation of viral RNA. Unlike MLV, where the packaging signal has been located in the 5'-untranslated leader downstream of the major splice donor site, the role of this analogous region in HIV-1 appears to be more responsible for discriminating genomic from spliced transcripts than to promote efficient encapsidation (Lever *et al.*, 1989; Aldovini and Young, 1990; Luban and Goff, 1994).

However, many reports have demonstrated that this region located between the major splice donor site and the initiation codon of the *gag* gene does form part of the HIV-1 packaging signal (Lever *et al.*, 1989; Clavel and Orenstein, 1990; Aldovini and Young, 1990; Poznansky *et al.*, 1991; Hayashi *et al.*, 1992), but the 5' and 3' extent of this are not known. Sequences in the transcribed LTRs and 5' leader sequences upstream of the major splice donor site have also been implicated in RNA packaging (Kim *et al.*, 1994; Vicenzi *et al.*, 1994; Geigenmuller and Linial, 1996; Paillart *et al.*, 1996; McBride and Panganiban, 1996; McBride and Panganiban, 1997). Additional sequences in the 5' end of the *gag* gene and even the 3' end of the *env* gene, have also been implicated in efficient packaging of HIV-1 (Richardson *et al.*, 1993; Kaye *et al.*, 1995). In reality, it has not been possible to define a minimal essential packaging signal by juxtaposition of a proposed sequence to heterologous RNA, therefore it is likely to be multipartite or highly sequence specific (Berkowitz *et al.*, 1995; Kaye *et al.*, 1995).

Although no *gag* sequence is essential for packaging of the HIV-1 derived vectors described in this chapter, as shown by the ext1 construct, the inclusion of up to 650 bp (ext4) of *gag* sequence increased viral titres over two-fold. The exact portion of *gag* that is required for encapsidation has not been fully elucidated, but this finding is in general agreement with others (Buchschacher and Panganiban; 1992; Parolin *et al.*, 1994; Luban and Goff, 1994; McBride *et al.*, 1997). The increase in gene transfer attributed to *gag* sequences is partly due to an increase in packaging of the vector RNA into virions, but this region likely contributes to functions besides packaging, such as dimerisation (Sundquist and Heaphy, 1993). Interestingly, Parolin *et al.* (1994) showed that inclusion of more than 653 bp of *gag* sequence actually decreased the ability of these vectors to be propagated. The inclusion of *gag* sequences in MLV vectors, although not absolutely required for packaging or gene transfer, has also been reported to increase viral titres (Armentano *et al.*, 1987; Bender *et al.*, 1987).

In addition to the Rev response element (RRE), the incorporation of the extended polypurine tract (ppt+) and the central polypurine tract (cppt) also increased viral titres. The inclusion of HIV-1 *env* sequences increased viral titre and packaging efficiency. The region of HIV-1 *env* sequence from the 3' end, up to but not including the splice acceptor site (F4), was clearly beneficial, a 5' *env* region from bases 5742 to 7030 (pHIV-YU2). It should be noted though, that some of these additional sequences include intact reading frames, which is contrary to the desire to separate the *cis* and *trans* genetic functions of HIV-1 to the highest possible degree. The 5' end of the *env* gene contains intact reading frames for Vpu, and exon 1 of both Tat and Rev. The F4 sequence encompasses the *nef*

gene and in combination with the 3' LTR, the entire *nef* reading frame. These sequences could be blocked by mutagenesis provided no significant loss in viral titre resulted.

The RRE was absolutely required for efficient vector function. In the absence of RRE it is predicted that the efficient cytoplasmic accumulation of unspliced, packageable vector RNA would be prevented. Inclusion of the RRE may also be warranted in HIV-1 vectors given the improved efficiency associated with vectors containing lengths of *gag* sequence inclusive of *cis*-acting repressor sequences (Rosen *et al.*, 1988; Schwartz *et al.*, 1992a and b). Attempts to make the vector system independent of Rev/RRE were not successful. It is clear from other reports that CTEs function with varying efficiency depending on their context (Bray, *et al.*, 1994; Rizvi *et al.*, 1996; Tabernero *et al.*, 1996; Zolotukhin *et al.*, 1994). An interesting report has shown that exchanging the RRE with the CTE of the simian retrovirus type 1 (SRV-1) increased levels of unspliced cytoplasmic RNA but that this did not correlate with viral titre. Despite higher mRNA levels with the SRV-1 CTE, viral titres were 5- to 10-fold lower (Mautino *et al.*, 2000). It has been suggested that the CTE and Rev/RRE export pathways are in actual fact different (Saavedra *et al.*, 1997), and this may have consequences for different cytoplasmic localisation's of mRNA which may equate to both lower packaging efficiencies and viral titres. One factor involved in the interaction of the CTE from MPMV has been identified as ATP-dependent RNA helicase A (Tang *et al.*, 1997). As suggested earlier, the relative abundance of the Rev protein compared to interactive factors for the CTEs may effect the efficiency with which HIV-1 encoded messages can be transported from the nucleus to the cytoplasm. Another approach to using CTEs in place of Rev/RRE may be to increase their copy number within the vector, especially in transient expression systems when large amounts of virus is made. A recent study has shown that polymerisation of the MPMV CTE increased its efficiency as a nuclear export signal (Wodrich *et al.*, 2000).

As already mentioned, assaying for p24 HIV-1 core antigen production is indicative of virus particle formation but does not discriminate between particles with and particles devoid of an RNA genome. The ratio of viral titre to p24 can be taken as an indicator of packaging efficiency if equivalence of steady-state levels of genomic RNA in the cytoplasm of producer cells, and of post-transduction events in the cells used for titre determination, are assumed. To determine more unambiguously if different vector designs were affecting the efficiency of packaging the vector RNA into virions, the steady-state levels of vector RNA in virions and in the cytoplasm of the transfected 293T cells will need to be directly measured. Secondly, the design of the vector construct may influence the expression of the transgene in the target cells. This could be evaluated by directly

transfecting the target cells or virus producing cells, in this instance NIH3T3 or 293T cells, with the vector construct. The transfected cells would then be lysed a few days later, and similar amounts of cell extracts used to measure the level of reporter gene expression. In this chapter because the transgene, EYFP, was measured by FACScan analysis the mean fluorescence intensity of various vector constructs could be compared. This was performed with a few of the original constructs and no difference in fluorescent intensity was observed, suggesting that this is not affecting the results (data not shown).

Apart from p24 assays as an indicator of viral particle formation, the function of the vectors described in this chapter have been restricted to end-point determination of gene transfer *in vitro*. For their use in gene therapy this is a logical test for vector performance. However, to directly assess the effect of particular sequence elements in HIV-1 derived vectors it may be more useful to look at the individual role such an element may have. Practically, this is very difficult to do. The continual addition of specific sequence elements has shown that viral titre is affected by the sequence composition of the vector, but that these viral titres reach a maximal limit. This was based on the finding that the inclusion of every sequence element that was shown to have a positive influence on vector function, did not cause an additive response in viral titres. The next question to be solved is to ascertain what is limiting the formation of viral particles?

From the work in this chapter some additional HIV-1 sequence elements can be identified and used to design a theoretical optimal vector without unduly compromising safety. They would include 650 nucleotides of *gag* sequence (ext4), the five conserved T residues immediately upstream of the polypurine tract (ppt+) and the RRE. Further vector development would involve mutagenesis of the *gag* initiation codon, splice acceptor and splice donor sites in such a vector. The effects of such mutations in the ext1, ext2 and ext5 vectors suggests that their introduction is likely to be context dependent. Also of note is that there are clearly a number of different ways to achieve an efficient vector. Other factors may influence the design of a vector construct. For example, from a safety point of view, the nature of the helper constructs used could help determine what *cis* elements may or may not be included in the vector to minimise the chance of recombination events leading to the generation of unwanted replication competent virus. The zenithal analysis of safety will depend as much on the systems used to provide the packaging functions and the detection of replication competent virus, as it will on vector design. The next step to evaluate some of these vector designs will be to evaluate them in conjunction with the helper plasmids in Chapter four. However, the analysis that was done in Chapter four with the pHIVext5SV40EYFPppt+RRE vector shows that a safe system can be developed and

that presumably this can only be improved by careful vector optimisation. At this stage such analysis would involve looking for the transfer of biologically active sequences that can substitute for the function of Gag and GagPol in transduced cells as described in section 4.3.7.

CHAPTER SIX

FINAL STATEMENTS

6.1 CONCLUDING REMARKS

Gene therapy is viewed by many as a panacea for genetic disease. However, a decade has passed since the first clinical gene therapy trials commenced and progress has been tardy. The challenge has been to deliver the therapeutic gene intact to the target cell, allow persistent levels of transgene expression to correct the disease phenotype without any unwanted side-effects as a consequence of vector exposure. To date, no such vector system satisfies all such criteria. Gene therapy and the means to promote it depend heavily on the development and improvement of new gene vector systems.

The advantages of HIV-1 as a vector system for gene therapy applications (outlined in section 1.3.2) are such that it continues to be developed with the aim of eventual clinical application. The work in this thesis has created a series of constructs for the expression of the HIV-1 proteins required in *trans* for the production of recombinant HIV-1 vectors. In combination with a suitable vector construct recombinant virus has been produced in a transient expression system. Two packaging systems have been developed that separate the helper functions over a panel of plasmids. A codon-optimised *gagpol* construct in combination with individual expression plasmids for Tat and Rev, was shown to be significantly better from a safety perspective (see section 4.3.7) than pCMVΔRnr which expresses GagPol, Tat, Rev, Nef and Vpu. Further improvements in safety were obtained by expressing the GagPolfusion polyprotein and Gag protein separately. The price paid for the increase in safety was a reduction in viral titre. It remains to be seen why this is so and if it can be overcome.

A series of HIV-1 derived vector constructs were made to evaluate the effect of *cis* sequences on vector performance. Whilst some sequences clearly increased the viral titre, the exact role of such sequences has not been elucidated. There also appears to be a fair degree of redundancy, as similar virus titres can be achieved with different sequence elements. The observation that individual sequence elements on their own did increase titre but such effects were not additive, suggests that other factors may be limiting. For example, one or more of the *trans* virus functions may be exhausted or the virus production system may be operating at saturation for any number of (unknown) reasons. Until the way in which many of the *cis* elements of HIV-1 exert their effects on virus titre is defined, perhaps the most beneficial rationale for a vector at this present time would be one of safety. Design of vectors for gene therapy will almost invariably involve compromises between the (usually conflicting) aims of maximal safety and maximal efficiency. The evaluation of *cis* sequence elements in this thesis does not pretend to define what may be considered an optimal vector design but shows that there are a number of sequences in pHIV-YU2 that effect the

efficiency of virus production. At least some of these sequences could be incorporated into a vector design without compromising the safety of the vector system to an unacceptable degree. This could be evaluated by marrying together several different vector constructs developed in Chapter five with the helper systems developed in Chapter four and making a safety assessment using the "*gag*" and "*gagpol*" transfer assays (section 4.3.7).

It is conceivable, that all the modifications made to the *trans* and *cis* functions of HIV-1 in the design of helper plasmids and vectors have altered the phenotype of the virus such that it no longer infects non-dividing cells. This is however, unlikely because the karyophilic determinants provided by the *gag* derived matrix antigen and Vpr are believed to be redundant (von Schwedler *et al.*, 1994). The matrix protein alone is an efficient nuclear localisation signal and this remains unaltered in the virus production system. In the final assessment of the HIV-1 gene transfer system developed in this thesis, transduction of non-dividing cells needs to be addressed. Simple *in vitro* transduction of non-dividing cells can be measured by treating target cells with aphidicolin which will growth arrest cells at the G_1-S boundary of the cell cycle (Huberman, 1981). Target cells can also be growth arrested at the G_2 phase by gamma irradiation (Kastan *et al.*, 1992). The ultimate test for the HIV-1 gene transfer system to successfully transduce a non-dividing cell population would be a biological cell that exhibits low or no mitotic activity. Such a cell is the haematopoietic stem cell (HSC). HSCs in the bone marrow have long been considered an ideal target for gene therapy, particularly in the interests of treating lysosomal storage disorders. This is due to their ability for self-renewal as well as expansion and differentiation into the cell lineages of the haematopoietic system. Revisiting the initial construction of an HIV-1 derived vector to transfer the lysosomal enzyme β-glucuronidase to MPS VII fibroblasts outlined in Chapter three would be the next appropriate step in the evolution of this HIV-1 vector system. If *in vitro* evaluation in MPS VII fibroblasts was successful, transduction of bone marrow could then be tried in an animal model of MPS VII (Birkenmeier *et al.*, 1989). The outcome of such an experiment would be a true indication of the success of the HIV-1 gene transfer system described in this thesis. The use of HIV-1 vectors provides a previously unexplored basis for the study of human gene therapy with the use of HSCs. The use of an HIV-1 derived vector to directly transduce neurons, has already led to the reversal of the pathology in the brain of MPS VII mice (Bosch *et al.*, 2000).

6.2 PHILOSOPHICAL CONSIDERATIONS

If HIV-1 derived vectors are to be considered for human gene therapy it is imperative that the possibility of unintended generation of replication competent virus is eliminated. Since the majority of HIV-1 vector systems have been produced *via* transient transfection of the relevant plasmids into 293T cells, then homologous recombination among these plasmids during transfection is the most likely way to generate unwanted replication competent virus. The crux of the safety issue resides in eliminating those viral genetic elements involved in the generation of replication competent virus without impairing the production of vector particles. Furthermore, removing genes that are not absolutely required for vector production and viral infectivity will facilitate the development of stable packaging cell lines for HIV-1 vectors.

To test the safety of an HIV-1 based vector production system presented here, as well as any other approach rests solely on the ability for the unitary detection of replication competent virus. Perhaps a more progressive approach would be the development of systems capable of detecting precursors to the formation of replication competent virus. It is important that such techniques make as few assumptions as possible as to the speculated phenotype of replication competent virus. In this regard assays for HIV-1 replication are probably unsuitable. Currently the method most frequently cited for evaluating the existence of replication competent virus is for the transfer of Gag expression to transduced cells by p24. Random batches of virus produced by the helper systems described in this thesis have been assayed for the transfer of p24 to transduced cells by ELISA, with consistently negative results (data not shown). However the fact that the more complete proviral packaging construct pCMVΔRnr (Kafri *et al.*, 1997) also gave negative results suggests that this approach is inappropriate for comparative assessment of the safety of different virus production systems. The assay described by Wu *et al.* (2000) that detects recombinant lentiviral DNA mobilisation exemplifies the type of test required. However, because their method relies on the Gag and Tat reading frames being linked it is unsuited for assessing the safety of the approach outlined in this thesis. The "*gag*" and "*gagpol*" transfer assays described in section 4.3.7 appear to be more sensitive. Using this assay is was possible to compare the difference in transfer of sequences capable of substituting for GagPol for different helper systems. However, the sensitivity and appropriateness of this assay to measure bio-safety for HIV-1 derived vector systems still needs to be evaluated before the clinical use of such systems is considered.

Once the safety issues of an HIV-1 vector system have been fully resolved there still remains a catechism. Stable packaging systems are clearly needed for HIV-1 vectors

and as yet are unavailable. The problems associated with transient transfections have already been alluded to and the clinical use of HIV-1 gene therapy protocols will be restricted until stable helper cell lines like those available for MLV are developed. As stated above, stable packaging cell lines also present significant safety advantages. In order to obtain high-titre virus stocks high degrees of (volume) concentration are required which can result in noticeable toxicity on some cell types after prolonged exposure. This indicates that more refined virus purification methods are needed for the preparation of high-titre virus stocks for some applications.

The public perception of the stigma associated with using HIV-1 as a treatment modality cannot be ignored. The "science" versus "emotive" debate will continue to be deliberated. A congenital disease with a poor prognosis may very well be amenable to a potential treatment that carries a risk and such instances are probably best decided by individual choice. The ethics of such a scenario will argue the risk/benefit dilemma a patient contemplating the treatment will face. Science will continue to go forward to present the facts, and hopefully the ethics and emotions involved will become easier as a result. HIV-1 contains a perceived safety risk, but more so than using another retrovirus? Does the entire safety issue rest with the generation of replication competent virus? Recombination events occurring within the virus production system would not produce wild-type HIV-1 because of the codon-optimised *gagpol* sequence and the VSV-G envelope. As mentioned, other lentiviral species that are non-pathogenic in humans are also being developed. Nonetheless, given the lack of understanding of these viruses, especially in regard to why they cannot infect human cells, it could be argued as to whether or not their development is really worth while. With all this in mind, one can continue to live in hope that with technological advances, gene therapy for treating a wide variety of diseases will become a reality within the early part of this century.

BIBLIOGRAPHY

Achong, B., Mansel, P. W. A., Epstein, M. A. and Clifford, P. (1971) An unusual virus in cultures from a human nasopharyngeal carcinoma. *J. Natl. Cancer Inst.* **46,** 299-302.

Adam, M. A. and Miller, A. D. (1988) Identification of a signal in a murine retrovirus that is sufficient for packaging of non-retroviral RNA into virions. *J. Virol.* **62,** 3802-6.

Adam, M. A., Osborne, W. R. and Miller, A. D. (1995) R-region cDNA inserts in retroviral vectors are compatible with virus replication and high-level protein synthesis from the insert. *Hum. Gene Ther.* **6,** 1169-76.

Adam, M. A., Ramesh, N., Miller, A. D. and Osborne, W. R. A. (1991) Internal initiation of translation in retroviral vectors carrying picornavirus 5' nontranslated regions. *J. Virol.* **65,** 4985-90.

Adams, M., Sharmeen, L., Kimpton, J., Romeo, J. M., Garcia, J. V., Peterlin, B. M., Groudin, M. and Ermerman, M. (1994) Cellular latency in human immunodeficiency virus infected individuals with high CD4 levels can be detected by the presence of promoter-proximal transcripts. *PNAS* **91,** 3862-6.

Aguzzi, A. (1993) The foamy virus family: molecular biology, epidemiology and neuropathology. *Biochim. Biophys. Acta* **1155,** 1-24.

Ahmad, N. and Venkatesan, S. (1988) *Nef* protein of HIV-1 is a transcriptional repressor of HIV-1 LTR. *Science* **241,** 1481-5.

Aihara, H. and Miyazaki, J. (1998) Gene transfer into muscle by electroporation in vivo. *Nat. Biotech.* **16,** 867-70.

Aiken, C. (1997) Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for nef and the sensitivity to cyclosporin A. *J. Virol.* **71,** 5871-7.

Aiken, C. and Trono, D. (1995) Nef stimulates human immunodeficiency virus type 1 proviral DNA synthesis. *J. Virol.* **69,** 5048-56.

Akgun, E., Ziegler, M. and Grez, M. (1991) Determinants of retrovirus gene expression i embryonal carcinoma cells. *J. Virol.* **65,** 382-8.

Akkina, R. K., Walton, R. M., Chen, M. L., Li, Q. X., Planelles, V. and Chen, I. S. (1996 High efficiency gene transfer into CD34+ cells with a human immunodeficiency virus typ 1 based retroviral vector pseudotyped with vesicular stomatitis virus envelope glycoprotei G. *J. Virol.* **70,** 2581-5.

Aldovini, A. and Young, R. A. (1990) Mutations and protein sequences involved in humar immunodeficiency virus type 1 packaging result in production of non-infectious virus. *J. Virol.* **64,** 1920-26.

Alemany, R., Balague, C. and Curiel, D. T. (2000) Replicative adenoviruses for cance therapy. *Nat. Biotech.* **18,** 723-7.

Ali, M., Lemoine, R. and Ring, J. A. (1994) The use of DNA viruses as vectors for gene therapy. *Gene Ther.* **1,** 367-84.

Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M. and Berger, E. A. (1996) CC CKR5: A rantes, MIP-1α, MIP-1β receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* **272,** 1955-8.

Andersen, J. K., Garber, D. A., Meaney, C. A. and Breakfield, X. A. (1992) Gene transfer into mammalian central nervous system using herpes virus vectors: extended expression of bacterial LacZ in neurons using the neuron-specific enolase promoter. *Hum. Gene. Ther.* **3,** 487-99.

Anderson, W. F. (1984) Prospects for human gene therapy. *Science* **266,** 401-9.

Anderson, W. F. (1992) Human gene therapy. *Science* **256,** 808-13.

Anderson, W. F. (1998) Human gene therapy. *Nature* **392,** 25-30.

Anson, D. S., Taylor, J. A., Bielicki, J., Harper, G. S., Peters, C., Gibson, G. J. and Hopwood, J. J. (1992) Correction of human mucopolysaccharidosis type VI fibroblasts with recombinant N-acetylgalactosamine 4-sulphatase. *Biochem. J.* **284,** 789-94.

Armentano, D., Yu, S. F., Kantoff, P. W., von Ruden, T., Anderson, W. F. and Gilboa, E. (1987) Effect of internal viral sequences on the utility of retroviral vectors. *J. Virol.* **61,** 1647-50.

Arya, S. K., Zamini, M. and Kundra, P. (1998) Human immunodeficiency virus type 2 lentivirus vectors for gene transfer: expression and potential for helper virus-free packaging. *Hum. Gene Ther.* **9,** 1371-80.

Ashe, M. P., Pearson, L. H. and Proudfoot, N. J. (1997) The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J.* **16,** 5752-63.

Balliet, J. W., Kolson, D. L., Eiger, G., Kim, F. M., McGann, K. A., Srinivasan, A. and Collman, R. (1994) Distinct effects in primary macrophages and lymphocytes of the human immunodeficiency virus type 1 accessory genes *vpr*, *vpu* and *nef*: Mutational analysis of a primary HIV-1 isolate. *Virology* **200,** 623-31.

Barranger, J. A. (1996) Haematopoietic stem cell gene transfer (editorial) *Gene Ther.* **3,** 379-80.

Bassin, R. H., Tuttle, N. and Fischinger, P. J. (1971) Rapid cell culture assay technique for murine leukemia viruses. *Nature* **229,** 564-6.

Bender, M. A., Palmer, T. D., Gelinas, R. E. and Miller, A. D. (1987) Evidence that the packaging signal of Moloney murine leukemia virus extends into the gag region. *J. Virol.* **6,** 1639-46.

Berkhout, B., Silverman, R. H. and Jeang, K. T. (1989) Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* **59,** 273-82.

Berkhout, B., van Wamel, J. and Klaver, B. (1995) Requirements for DNA strand transfe during reverse transcription in mutant HIV-1 virions. *J. Mol. Biol.* **252,** 59-69.

Berkowitz, R. D., Hammarskjöld, M. L., Helga, M. C., Rekosh, D. and Goff, S. P. (1995 5' regions of human immunodeficiency virus type 1 RNAs are not sufficient fo encapsidation: Implications for the human immunodeficiency virus type 1 packagin signal. *Virology* **212,** 718-23.

Berthold, E. and Maldarelli, F. (1996) Cis acting elements in human immunodeficienc virus type 1 RNAs direct viral transcripts to distinct intranuclear locations. *J. Virol.* **70** 4667-82.

Bieniasz, P. D., Erlwein, O., Aguzzi, A., Rethwilm, A. and McClure, M. O. (1997) Gene transfer using replication defective human foamy virus vectors. *Virology* **235,** 65-72.

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

Birkenmeier, E. H., Davisson, M. T., Beamer, W. G., Ganschow, R. E., Vogler, C. A., Gwynn, B., Lyford, K. A., Maltais, L. M. and Wawrzyniak, C. J. (1989) Murine mucopolysaccharidosis type VII. Characterisation of a mouse with β-glucuronidase deficiency. *J. Clin. Invest.* **83,** 1258-66.

Blaese, R. M., Culver, K. W., Miller, A. D., Carter, C. S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P., Greenblatt, J. J., Rosenberg, S. A., Klein, H., Berger, M., Mullen, C. A., Ramsey, W. J., Muul, L., Morgan, R. A. and Anderson, W. F. (1995) T lymphocyte-directed gene therapy for ADA-SCID: initial trial results after 4 years. *Science* **270,** 475-80.

Borg, K. T., Favaro, J. P. and Arrigo, S. J. (1997) Involvement of human immunodeficiency virus type 1 splice sites in the cytoplasmic accumulation of viral RNA. *Virology* **236,** 95-103.

Borman, A. M., Quillent, C., Charneau, P., Dauguet, C. and Clavel, F. (1995) Human immunodeficiency virus type 1 Vif-mutant particles from restrictive cells: Role of *vif* in correct particle assembly and infectivity. *J. Virol.* **69,** 2058-67.

Bosch, A., Perret, E., Desmaris, N., Trono, D. and Heard, J. M. (2000) Reversal of pathology in the entire brain of mucopolysaccharidosis type VII mice after lentivirus-mediated gene transfer. *Hum. Gene Ther.* **11,** 1139-50.

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

Buchschacher, G. L. and Panganiban, A. T. (1992) Human immunodeficiency virus vectors for inducible expression of foreign genes. *J. Virol.* **66,** 2731-9.

Bukovsky, A. A., Dorfman, T., Weimann, A. and Göttlinger, H. G. (1997) *Nef* association with human immunodeficiency virus type 1 virions and cleavage by the viral protease. *J. Virol.* 71, 1013-8.

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

Burns, J. C., Friedmann, T., Driever, W., Burrascano, M. and Yee, J. K. (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titre and efficient gene transfer into mammalian and nonmammalian cells. *PNAS* **90,** 3033-7.

Camaur, D. and Trono, D. (1996) Characterisation of human immunodeficiency virus type I *vif* particle incorporation. *J. Virol.* **70,** 6106-11.

Camerini, D., Planelles, V. and Chen, I. (1994) CD26 antigen and HIV fusion? *Science* **264,** 1160-1.

Carroll, R., Lin, J-T., Dacquel, E. J., Mosca, J. D., Burke, D. S. and St. Louis, D. C. (1994 A human immunodeficiency virus type 1 (HIV-1)-based retroviral vector system utilisin stable HIV-1 packaging cell lines. *J. Virol.* **68,** 6047-51.

Cepko, C. L., Roberts, B. E. and Mulligan, R. C. (1984) Construction and applications of highly transmissible murine retrovirus shuttle vector. *Cell* **37,** 1053-62.

Chang, H-K., Gallo, R. C. and Ensoli, B. (1995) Regulation of cellular gene expressio and function by the human immunodeficiency virus type 1 Tat protein. *J. Biomed. Sci.* **2** 189-202.

Chang, L-J., Urlacher, V., Iwakuma, T., Cui, Y. and Zucali, J. (1999) Efficacy and safet analyses of a recombinant human immunodeficiency virus type 1 derived vector system *Gene Ther.* **6,** 715-28.

Chang, Y. N. and Jeang, K. T. (1992) The basic RNA-binding domain of HIV-2 Ta contributes to preferential trans-activation of a TAR2-containing LTR. *Nucl. Acids Res.* **20,** 5465-72.

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

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

Cheong, C., Varani, G. and Tinoco, I. J. (1990) Solution structure of an unusually stable RNA hairpin, 5' GGAC(UUCG)GUCC. *Nature* **346,** 680-2.

Chinnasamy, D., Chinnasamy, N., Enriquez, M. J., Otsu, M., Morgan, R. A. and Candotti, F. (2000) Lentiviral-mediated gene transfer into human lymphocytes: role of HIV-1 accessory proteins. *Blood* **96,** 1309-16.

Chowers, M. Y., Pandori, M. W., Spina, C. A., Richman, D. D. and Guatelli, J. C. (1995) The growth advantage conferred by HIV-1 *nef* is determined at the level of viral DNA formation and is independent of CD4 downregulation. *Virology* **212,** 451-7.

Clavel, F.and Orenstein, J. M. (1990) A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal partical morphology. *J. Virol.* **64,** 5230-4.

Clever, J., Sassetti, C. and Parslow, T. G. (1995) RNA secondary structure and binding sites for *gag* gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J. Virol.* **69,** 2101-2109.

Cochrane, A. W., Perkins, A. and Rosen, C. A. (1990) Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. *J. Virol.* **64,** 881-5.

Cochrane, A. W., Jones, K. S., Beidas, S., Dillon, P. J., Skalka, A. M. and Rosen, C. A. (1991) Identification and characterisation of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J. Virol.* **65,** 5305-13.

Coffin, J. M. (1979) Structure, replication and recombination of retrovirus genomes: Some unifying hypotheses. *J. Gen. Virol.* **42,** 1-26.

Coffin, J. M. (1990) Retroviridae and their replication. In *Virology* (ed. B. N. Fields *et al.*) pp. 1437-500. Raven Press, New York.

Coffin, J. M. (1992) Structure and classification of retroviruses. In *The retroviridae* (ed. J. A. Levy), pp 19-49. Plenum Press, New York.

Cohen, E. A., Dehni, G., Sodroski, J. G. and Haseltine W. A. (1990a) Human immunodeficiency virus *vpr* product is a virion-associated regulatory protein. *J. Virol.* **64,** 3097-9.

Cohen, E. A., Terwilliger, E. F., Jalinoos, Y., Proulx, J., Sodroski, J. G. and Haseltine, W. A. (1990b) Identification of HIV-1 *vpr* product and function. *J. Acquir. Immune Defic. Syndr.* **3,** 11-8.

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

Cone, R. D. and Mulligan, R. C. (1984) High efficiency gene transfer into mammalian cells: Generation of helper free recombinant retrovirus with broad mammalian host range *PNAS* **81,** 6349-53.

Cooke, S. J., Coates, K., Barton, C. H., Biggs, T. E., Barrett, S. J., Cochrane, A., Oliver K., McKeating, J. A., Harris, M. P. G. and Mann, D. A. (1997) Regulated expression vectors demonstrate cell-type-specific sensitivity to human immunodeficiency virus type 1 Nef-induced cytostasis. *J. Gen. Virol.* **78,** 381-92.

Corbeau, P., Kraus, G. and Wong-Staal, F. (1996) Efficient gene transfer by a human immunodeficiency virus type 1 (HIV-1)-derived vector utilizing a stable HIV packaging cell line. *PNAS* **93,** 14070-5.

Corbeau, P., Kraus, G. and Wong-Staal, F. (1998) Transduction of human macrophages using a stable HIV-1/HIV-2 derived gene delivery system. *Gene Ther.* **5,** 99-104.

Cosset, F. L., Takeuchi, Y., Battini, J. L., Weiss, R. A. and Collins, M. K. (1995) High titre packaging cells producing recombinant retroviruses resistant to human serum. *J. Virol.* **69,** 7430-6.

Crystal, R. G., McElvaney, N. G., Rosenfeld, M. A., Chu, C. S., Mastrangeli, A., Hay, J. G., Brody, S. L., Jaffe, H. A., Eissa, N. T. and Danel, C. (1994) Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat. Genetics* **8,** 42-51.

Cui, Y., Iwakuma, T. and Chang, L. J. (1999) Contributions of viral splice sites and *cis*-regulatory elements to lentivirus vector function. *J. Virol.* **73,** 6171-6.

Cullen, B. R. (1991) Regulation of human immunodeficiency virus replication. *Annu. Rev. Microbiol.* **45,** 219-50.

Curran, M. A., Kaiser, S. M., Achacoso, P. L. and Nolan, G. P. (2000) Efficient transduction of nondividing cells by optimised feline immunodeficiency virus vectors. *Mol. Ther.* **1,** 31-8.

Dayton, A. I., Sodroski, J. G., Rosen, C. A., Goh, W. C. and Haseltine, W. A. (1986) The *trans*-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* **44,** 941-7.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di-Marzio, P., Marmon, S., Sutton, R. E., Hill, C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R. and Landau, N. R. (1996) Identification of a major co-receptor for primary isolates of HIV-1 [see comments]. *Nature* **381,** 661-6.

Desrosiers, R. C. and Letvin, N. L. (1987) Animal models for acquired immunodeficiency syndrome. *Rev. Infect. Dis.* **9,** 438-46.

Dillon, P. J., Lenz, J. and Rosen, C. A. (1991) Construction of a replication competent murine retrovirus vector expressing the human immunodeficiency virus type 1 tat transactivator protein. *J. Virol.* **65,** 4490-3.

Donello, J. E., Loeb, J. E. and Hope, T. J. (1998) Woodchuck hepatitis virus contains a tripartite posttranscriptional regulatory element. *J. Virol.* **72,** 5085-92.

Dorfman, T. F., Luban, J., Goff, S. P., Haseltine, W. A. and Göttlinger, H. G. (1993) Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* **67,** 6159-6169.

Dorfman, T. F., Mammano, F., Shoeman, R. L. and Göttlinger, H. G. (1994) Role of matrix protein in the virion association of the HIV-1 envelope protein. *J. Virol.* **68,** 1689-96.

Dougherty, J. P. and Temin, H. M. (1988) Determination of the rate of base pair substitution and insertion mutations in retrovirus replication. *J. Virol.* **62,** 2817-22.

Dropulic, B., Hermankova, M. and Pitha, P. M. (1996) A conditionally replicating HIV-vector interferes with wild type HIV-1 replication and spread. *PNAS* **93,** 11103-8.

DuBridge, R. B., Tang, P., Hsia, H. C., Leong, P. M., Miller, J. H. and Calos, M. P. (1987 Analysis of mutations in human cells by using an Epstein-Barr virus shuttle system. *Mol Cell. Biol.* **7,** 379-87.

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D. and Naldini, L (1998) A third generation lentivirus vector with a conditional packaging system. *J. Virol* **72,** 8463-71.

During, M.J., Naegele, J. R., O'Malley, K. L. and Geller, A. I. (1994) Long-term behavioural recoverey in Parkinson rats by an HSV vector expressing tyrosine hydroxylase. *Science* **266,** 1399-403.

Ellermann, V. and Bang, O. (1908) Experimentelle Leukämie bei Hühnern. *Zentralbl. Bakteriol. Parasitenkd. Infectionskr. Hyg. Abt. Orig.* **46,** 595-609.

Emerman, M., Vazeux, R. and Peden, K. (1989) The *rev* gene product of the human immunodeficiency virus affects envelope specific RNA localisation. *Cell* **57,** 1155-65.

Ernst, R. K., Bray, M., Rekosh, D. and Hammarskjöld, M. L. (1997) Secondary structure and mutational analysis of the Mason-Pfizer monkey virus constitutive transport element. *RNA* **3,** 210-22.

Feero, W. G., Li, S., Rosenblatt, J. D., Sirianni, N., Morgan, J. E., Partridge, T. A., Huang, L. and Hoffman, E. P. (1997) Selection and use of ligands for receptor mediated gene delivery to myogenic cells. *Gene Ther.* **4,** 664-7.

Felber, B. K., Hadzopoulou-Cladaras, M., Copeland, C. and Pavlakis, G. N. (1989) Rev protein of HIV-1 affects the stability and transport of viral mRNA. *PNAS* **86,** 1495-9.

Felgner, J. H., Kumar, R., Sridhar, C. N., Wheeler, C. J., Tsai, Y. J., Border, R., Ramsey, P., Martin, M. and Felgner, P. L. Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. (1994) *J. Biol. Chem.* **269,** 2550-61.

Felgner, P. L. Particulate systems and polymers for in vitro and in vivo delivery of polynucleotides. (1990) *Adv. Drug Deliv. Rev.* **5,** 163-87.

Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. and Danielsen, M. (1987) Lipofection: a highly efficient, lipid mediated DNA transfection procedure. *PNAS* **84,** 7413-7.

Fennie, C. and Lasky, L. A. (1989) Model for intracellular folding of the human immunodeficiency virus type 1 gp120. *J. Virol.* **63,** 639-46.

Filion, M. C. and Phillips, N. C. (1998) Major limitations in the use of cationic liposomes for DNA delivery. *Int. J. Pharmaceut.* **162,** 159-70.

Fink, D. J. and Glorioso, J. C. (1997) Engineering herpes simplex virus vectors for gene transfer to neurons. *Nature Med.* **3,** 357-9.

Fisher, A. G., Feinberg, M. B., Josephs, S. F., Harper, M. E., Marselle, L. M., Reyes, G., Gonda, M. A., Aldovini, A., Debouk, C., Gallo, R. C. and Wong-Staal, F. (1986) The *trans* activator of HTLV-III is essential for virus replication. *Nature* **320,** 367-71.

Fisher, K. J., Jooss, K., Alston, J., Yang, Y., Haecker, S. E., High, K., Pathak, R., Raper, S. E. and Wilson, J. M. (1997) Recombinant adeno-associated virus for muscle directed gene therapy. *Nat. Med.* **3,** 306-12.

Fouchier, R. A., Simon, J. H., Jaffe, A. B. and Malim, M. H. (1996) Human immunodeficiency virus type 1 *vif* does not influence expression or virion incorporation of gag-, pol-, and env- encoded proteins. *J. Virol.* **70,** 8263-9.

Frisella, W. A., O'Connor, L. H., Vogler, C. A., Roberts, M., Walkley, S., Levy, B., Daly, T. M. and Sands, M. S. (2001) Intracranial injection of recombinant adeno-associated virus improves cognitive function in a murine model of mucopolysaccharidosis type VII. *Mol. Ther.* **3,** 351-8.

Gabuzda, D. H., Lawrence, K., Langhoff, E., Terwilliger, E., Dorfman, T., Haseltine, W A. and Sodroski, J. (1992) Role of *vif* in replication of human immunodeficiency viru type 1 in CD4+ T lymphocytes. *J. Virol.* **66,** 6489-95.

Gallay, P., Hope, T., Chin, D. and Trono, D. (1997) HIV-1 infection of nondividing cell through the recognition of integrase by the importin/karyopherin pathway. *PNAS* **94,** 9825 30.

Gao, G. P., Yang, Y. and Wilson, J. M. (1996) Biology of adenovirus vectors with E1 anc E4 deletions for liver directed gene therapy. *J. Virol.* **70,** 8934-43.

Garcia, J. A., Harrich, D., Pearson, L., Mitsuyasu, R. and Gaynor, R. B. (1988) Functional domains required for tat induced transcriptional activation of the HIV-1 long terminal repeat. *EMBO J.* **7,** 3143-7.

Garret, E., Tiley, L. and Cullen, B. (1991) *Rev* activates expression of the HIV-1 *vif* and *vpr* gene products. *J. Virol.* **65,** 1653-7.

Gasmi, M., Glynn, J., Jin, M. J., Jolly, D. J., Yee, J. K. and Chen, S. T. (1999) Requirements for efficient production and transduction of human immunodeficiency virus type 1 based vectors. *J. Virol.* **73,** 1828-34.

Geigenmuller, U. and Linial, M. L. (1996) Specific binding of human immunodeficiency virus type 1 (HIV-1) Gag-derived proteins to a 5' HIV-1 genomic RNA sequence. *J. Virol.* **70,** 667-71.

Gelderblom, H. R., Hausman, E. H. S., Ozel, M., Pauli, G. and Koch, M. A. (1987) Fine structure of human immunodeficiency virus (HIV) and immunolocalization of structural proteins. *Virology* **156,** 171-6.

Germann, U. A., Chin, K. V., Pastan, I. and Gottesman, M. M. (1990) Retroviral transfer of a chimeric multidrug resistance adenosine deaminase gene. *FASEB J.* **4,** 1501-7.

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

Girones, R., Cote, P. J., Hornbuckle, W. E., Tennant, B. C., Gerin, J. L., Purcell, R. H. and Miller, R. H. (1989) Complete nucleotide sequence of a molecular clone of woodchuck hepatitis that is infectious in the natural host. *PNAS* **86,** 1846-9.

Glaser, J. H., Sly, W. S. (1973) β-glucuronidase deficiency mucopolysacchardidosis: methods for enzymatic diagnosis. *J. Lab. Clin. Med.* **82,** 969-77.

Goff, S., Trakman, P. and Baltimore, D. (1981) Isolation and properties of murine leukemia virus mutants: Use of a rapid assay for release of virion reverse transcriptase. *J. Virol.* **38,** 239-48.

Goldman, M. J., Lee, P. S., Yang, J. S. and Wilson, J. M. (1997) Lentiviral vectors for gene therapy of cystic fibrosis. *Hum. Gene Ther.* **8,** 803-15.

Goldsmith, M. A., Warmerdam, M. T., Atchison, R. E., Miller, M. D. and Greene, W. C. (1995) Dissociation of the CD4 downregulation and viral infectivity enhancement functions of human immunodeficiency virus type 1 *nef. J. Virol.* **69,** 4112-21.

Goncalves, J., Jallepalli, P. and Gabuzda, D. H. (1994) Subcellular localisation of the *vif* protein of human immunodeficiency virus type 1. *J. Virol.* **68,** 704-12.

Goncalves, J., Shi, B., Yang, X. and Gabuzda, D. (1995) Biological activity of human immunodeficiency virus type 1 *vif* requires membrane targeting by C-terminal basic domains. *J. Virol.* **69,** 7196-204.

Göttlinger, H. G., Dorfman, T., Sodroski, J. and Haseltine, W. (1991) Effect of mutations affecting the p6 gag protein on HIV particle release. *PNAS* **88,** 3195-9.

Göttlinger, H. G., Dorfman, T., Cohen, E. and Haseltine, W. A. (1993) Vpu protein of human immunodeficiency virus type 1 enhances the release of capsids produced by gag gene constructs of widely divergent retroviruses. *PNAS* **90,** 7381-5.

Graham, F. L. and Eb, A. J. V. D. (1973) A new technique for the assay of infectivity o human adenovirus 5 DNA. *Virology* **52,** 456-67.

Gruter, P., Tabernero, C. von Kobbe, C., Schmidt, C., Saavedra, C., Bachi, A., Wilm, M. Felber, B. K. and Izaurralde, E. (1998) TAP, the human homolog of Mex67p, mediate CTE-dependent RNA export from the nucleus. *Mol. Cell* **1,** 649-59.

Guyer, M. S. and Collins, F. S. (1995) How is the Human Genome Project doing, and wha have we learned so far? *PNAS* **92,** 10841-8.

Haecker, S. E., Stedman, H. H., Balice-Gordon, R. J., Smith, D. B. J., Greelish, J. P. Mitchell, M. A., Wells, A., Sweeney, H. L. and Wilson, J. M. (1996) *In vivo* expression of full-length human dystrophin from adenoviral vectors deleted of all viral genes. *Hum. Gene Ther.* **7,** 1907-14.

Hammarskjöld, M. L., Li, H., Rekosh, D. and Prasad, S. (1994) Human immunodeficiency virus env expression becomes Rev independent if the env region is not defined as an intron. *J. Virol.* **68,** 951-8.

Hammes, S. R., Dixon, E. P., Malim, M. H. Cullen, B. R. and Greene, W. C. (1989) *Nef* protein of human immunodeficiency virus type 1: evidence against its role as a transcriptional inhibitor. *PNAS* **86,** 9549-53.

Hantzopoulos, P. A., Sullenger, B. A., Ungers, G. and Gilboa, E. (1989) Improved gene expression upon transfer of the adenosine deaminase minigene outside the transcription unit of a retroviral vector. *PNAS* **86,** 3519-23.

Harrich, D., Ulich, C., Garcia-Martinez, L. F. and Gaynor, R. B. (1997) Tat is required for efficient HIV-1 reverse transcription . *EMBO* **16,** 1224-35.

Harrich, D., Ulich, C. and Gaynor, R. B. (1996) A critical role for the TAR element in promoting efficient human immunodeficiency virus type 1 reverse transcription. *J. Virol.* **70,** 4017-27.

Harrison, G. P. and Lever, A. M. (1992) The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure. *J. Virol.* **66,** 4144-53

Haseltine, W. A. (1991) Molecular biology of HIV-1. *FASEB J.* **5,** 2349-60.

Hawley, R. G., Covarrubias, L., Hawley, T. and Mintz, B. (1987) Handicapped retroviral vectors efficiently transduce foreign genes into heamatopoietic stem cells. *PNAS* **84,** 2406-10.

Hayashi, M., Nakajima, Y. and Fishman, W. H. (1964) The cytologic demonstration of β-glucuronidase employing naphthol AS-BI glucuronide and hexazonium pararosanalin; a preliminary report. *J. Histochem. Cytochem.* **12,** 293-7

Hayashi, T., Shioda, T., Iwakura, Y. and Shibuta, H. (1992) RNA packaging signal of human immunodeficiency virus type 1. *Virology* **188,** 590-9.

Heaphy, S., Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Karn, J., Lowe, A. D., Singh, M. and Skinner, M. A. (1990) HIV-1 regulator of virion expression (rev) binds to an RNA stem loop structure located within the rev responsive element region. *Cell* **60,** 685-93.

Heinzinger, N. K. Bukinsky, M. I., Haggerty, S. A., Ragland, A. M., Kewalramani, V., Lee, M. A., Gendelman, H. E., Ratner, L., Stevenson, M. and Emerman, M. (1994) The *vpr* protein of human immunodeficiency virus type 1 influences nuclear localisation of viral nucleic acids in nondividing host cells. *PNAS* **91,** 7311-5.

Herman, S. A. and Coffin, J. M. (1987) Efficient packaging of readthrough RNA in ALV: Implications for oncogene transduction. *Science* **236,** 845-8.

Heus, H. A. and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **253,** 191-4.

Huang, J. and Liang, T. J. (1993) A novel hepatitis B virus (HBV) genetic element wit Rev response element like properties that is essential for expression of HBV gen products. *Mol. Cell. Biol.* **13,** 1416-86.

Huang, L. M., Joshi, A., Willey, R., Orenstein, J. and Jeang, K. T. (1994) Humai immunodeficiency viruses regulated by alternative trans-activators: genetic evidence for a novel non-transcriptional function of Tat in virion infectivity. *EMBO* **13,** 2886-96

Huang, Y. and Carmichael, G. G. (1997) The mouse histone H2A gene contains a smal element that facilitates cytoplasmic accumulation of intronless gene transcripts and o unspliced HIV-1-related mRNAs. *PNAS* **94,** 10104-9.

Huberman, J. A. (1981) New views of the biochemistry of eucaryotic DNA replicatior revealed by aphidicolin, an unusual inhibitor of DNA polymerase alpha. *Cell* **23,** 647-8.

Ilyinskii, P. O. and Desrosiers, R. C. (1998) Identification of a sequence element immediately upstream of the polypurine tract that is essential for replication of simian immunodeficiency virus. *EMBO J.* **17,** 3766-74.

Isel, C., Ehresmann, C., Keith, G., Ehresmann, B. and Marquet, R. (1995) Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/tRNA(Lys3) (template/primer). *J. Mol. Biol.* **247,** 236-50.

Iwakuma, T., Cui, Y. and Chang, L-J. (1999) Self-inactivating lentiviral vectors with U3 and U5 modifications. *Virol.* **261,** 120-132.

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

Jang, S. K., Pestova, T. V., Hellen, C. U., Witherell, G. W. and Wimmer, E. (1990) Cap-independent translation of picornavirus RNAs: Structure and function of the internal ribosomal entry site. *Enzyme* **44,** 292-309.

Jeang, K-T., Xiao, H. and Rich, E. A. (1999) Multifaceted activities of the HIV-1 transactivator of transcription, tat. *J. Biol. Chem.* **274,** 28837-40.

Johnson, P. A., Miyanohura, A., Levind, F., Cahill, T. and Friedmann, T. (1992) Cytotoxicity of a replication-defective mutant of herpes simplex virus type 1. *J. Virol.* **66,** 2952-65.

Jordan, M., Schallhorn, A. and Wurm, F. M. (1996) Transfecting mammalian cells: optimisation of critical parameters affecting calcium-phosphate preciptitation formation. *Nucl. Acids Res.* **4,** 596-601

Jowett, J. B., Planelles, V., Poon, B., Shah, N. P., Chen, M. L. and Chen, I. S. (1995) The human immunodeficiency virus type 1 *vpr* gene arrests infected T cells in the G2 + M phase of the cell cycle. *J. Virol.* **69,** 6304-13.

Kafri, T., Blömer, U., Peterson, D. A., Gage, F. H. and Verma, I. M. (1997) Sustained expression of genes delivered directly into liver and muscle by lentiviral vectors. *Nat. Genet.* **17,** 314-17.

Kafri, T., van Praag, H., Ouyang, L., Gage, F. H. and Verma, I. M. (1999) A packaging cell line for lentivirus vectors. *J. Virol.* **73,** 576-84.

Kaplan, A. H. and Swanstrom, R. (1991) The HIV-1 gag precursor is processed via two pathways: implications for cytotoxicity. *Biomed. Biochim. Acta* **50,** 647-53.

Kaplitt, M. G., Leone, P., Samulski, R. J., Xiao, X., Pfaff, D. W., O'Malley, K. L. and During, M. J. (1994) Long-term gene expression and phenotypic correction using adeno-associated virus vectors in the mammalian brain. *Nat. Genet.* **8,** 148-54.

Karacostas, V., Nagashima, K., Gonda, M. A. and Moss, B. (1989) Human immunodeficiency virus like particles produced by a vaccinia virus expression vector. *PNAS* **86,** 8964-7.

Karczewski, M. K. and Strebel, K. (1996) Cytoskeleton association and virio incorporation of the human immunodeficiency virus type 1 *vif* protein. *J. Virol.* **70,** 494 507.

Kastan, M. B., Zhan, Q., el Deiry, W. S., Carrier, F., Jacks, T., Walsh, W. V., Plunkett, B S., Vogelstein, B. and Fornace, A. J. Jr. (1992) A mammalian cell cycle checkpoin pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell* **71,** 587-97

Kaul, M., Yu, H., Ron, Y. and Dougherty, J. P. (1998) Regulated lentiviral packaging cel line devoid of most viral cis acting sequences. *Virology* **249,** 167-74.

Kawaoka, Y., Bean, W. J. and Webster, R. G. (1989) Evolution of the heamagglutin of equine H3 influenza viruses. *Virology* **169,** 283-92.

Kay, M. A., Landen, C. N., Rothenberg, S. R., Taylor, L. A., Leland, F., Wiehle, S., Fang, B., Bellinger, D., Finegold, M., Thompson, A. R., Read, M., Brinkhous, K. M. and Woo, S. L. C. (1994) In vivo hepatic gene therapy: complete albeit transient correction of factor IX deficiency in hemophilia B dogs. *PNAS* **91,** 2353-7.

Kaye, J. F., Richardson, J. H. and Lever, A. M. L. (1995) *cis*-Acting sequences involved in human immunodeficiency virus type 1 RNA packaging. *J. Virol.* **69,** 6588-92.

Kearns, W. G., Afione, S. A., Fulmer, S. B., Pang, M. C., Erikson, D., Egan, M., Landrum, M. J., Flotte, T. R. and Cutting, G. R. (1996) Recombinant adeno-associated virus (AAV-CFTR) vectors do not integrate in a site specific fashion in an immortalized epithelial cell line. *Gene Ther.* **3,** 748-55.

Kempler , G., Freitag, B., Berwin, B., Nanassy, O. and Barklis, E. (1993) Characterisation of the Moloney murine leukemia virus stem cell specific repressor binding site. *Virology* **193,** 690-9.

Kerkau, T., Bacik, I., Bennink, J. R., Yewdell, J. W., Hunig, T., Schimpl, A. and Schubert, U. (1997) The human immunodeficiency virus type 1 (HIV-1) *vpu* protein interferes with an early step in the biosynthesis of major histocompatibiliy complex (MHC) class I molecules. *J. Exp. Med.* **185,** 1295-305.

Kestler, H. W., Ringler, D. J., Panicaly, D. L., Sehgal, P. K., Daniel, M. D. and Desrosiers, R. C. (1991) Importance of the *nef* gene for maintenance of high virus load and for development of AIDS. *Cell* **65,** 651-62.

Kim, H. J., Lee, K. and O'Rear, J. J. (1994) A short sequence upstream of the 5' major splice site is important for encapsidation of human immunodeficiency virus type 1 genomic RNA. *Virology* **198,** 336-40.

Kim, S., Byrn, R., Groopman, J. E. and Baltimore, D. (1989a) Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J. Virol.* **63,** 3708-13.

Kim, S., Ikeuchi, K., Byrn, R., Groopman, J. E. and Baltimore, D. (1989b) Lack of a negative influence on viral growth by the *nef* gene of HIV-1. *PNAS* **86,** 9544-8.

Kim, V. N., Mitrophanous, K., Kingsman, S. M. and Kingsman, A. J. (1998) Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1. *J. Virol.* **72,** 811-6.

Klages, N., Zufferey, R. and Trono, D. (2000) A stable system for the high-titre production of multiply attenuated lentiviral vectors. *Mol. Ther.* **2,** 170-6.

Klimkait, T., Strebel, K., Hoggan, M. D., Martin, M. A. and Orenstein, J. M. (1990) The human immunodeficiency virus type 1 specific protein vpu is required for efficient virus maturation and release. *J. Virol.* **64,** 621-9.

Kondo, E., Mammano, F., Cohen, E. A. and Göttlinger, H. (1995) The p6 gag domain of HIV-1 is sufficient for the incorporation of *vpr* into heterologous particles. *J. Virol.* **69,** 2750-64.

Konings, D. A., Nash, M. A., Maizel, J. V. and Arlinghaus, R. B. (1992) Novel GACG-hairpin pair motif in the 5' untranslated region of type C retroviruses related to murine leukemia virus. *J. Virol.* **66,** 632-40.

Kotsopoulou, E., Kim, V. N., Kingsman, A. J., Kingsman, S. M. and Mitrophanos, K. A (2000) A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vectc that exploits a codon optimised HIV-1 gag-pol gene. *J. Virol.* **74,** 4839-52.

Kräusslich, H. G., Ochsenbauer, C., Traenckner, A. M., Mergener, K., Facke, M Gelderblom, H. R. and Bosch, V. (1993) Analysis of protein expression and virus-lik particle formation in mammalian cell lines stably expressing HIV-1 *gag* and *env* gen products with or without active HIV proteinase. *Virology* **192,** 605-17.

Kuppuswamy, M., Subramanian, T., Srinivasan, A. and Chinnadurai, G. (1989) Multipl functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis *Nucl. Acids Res.* **17,** 3551-61.

Lagunoff, D., Nicol, D. M. and Pritzl, P. (1973) Uptake of β-glucuronidase by deficien human fibroblasts. *Lab. Invest.* **29,** 449-453.

Landau, N. R., Page, K. A. and Littman, D. R. (1991) Pseudotyping with human T cell leukemia virus type 1 broadens the human immunodeficiency virus host range. *J. Virol.* **65,** 162-9.

Lee, T. H., Coligan, J. E., Allan, J. S., McLane, M. F., Groopman, J. E. and Essex, M. (1986) A new HTLV-III/LAV protein encoded by a gene found in cytopathic retroviruses. *Science* **231,** 1546-9.

Lever, A. M. L., Göttlinger, H., Haseltine, W. and Sodroski, J. (1989) Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 into virions. *J. Virol.* **63,** 4085-7.

Lewis, N., Williams, J., Rekosh, D. and Hammarskjöld, M. L. (1990) Identification of a cis-acting element in a human immunodeficiency virus type 2 (HIV-2) that is responsive to the HIV-1 rev and human T-cell leukemia virus type I and II rex proteins. *J. Virol.* **64,** 1690-7.

Lewis, P. and Emerman, M. (1994) Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.* **68,** 510-516.

Lewis, P., Hensel, M. and Emerman, M. (1992) Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J.* **11,** 3053-8.

Li, Y., Hui, H., Burgess, C. J., Price, R. W., Sharp, P. M., Hahn, B. H. and Shaw, G. M. (1992) Complete nucleotide sequence, genome organisation, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation. *J. Virol.* **66:** 6587-600.

Lifson, J. D., Feinberg, M. R., Reyes, G. R., Rabin, L., Basiripour, B., Chakrabarti, S., Moss, B., Wong-Stahl, F., Steimer, K. S. and Engleman, E. G. (1986) Induction of CD-4 dependent cell fusion by HALVE/LA envelope glycoprotein. *Nature* **323,** 725-8.

Linden, R. M. and Berns, K. I. (1997) Site-specific integration by adeno-associated virus: a basis for a potential gene therapy vector. *Gene Ther.* **4,** 4-5.

Lisziewicz, J., Sun, D., Smythe, J., Lusso, P., Lori, F., Louie, A., Markham, P., Rossi, J., Reitz, M. and Gallo, R. C. (1993) Inhibition of human immunodeficiency virus type 1 replication by regulated expression of a polymeric Tat activation response RNA decoy as a strategy for gene therapy in AIDS. *PNAS* **90,** 8000-4.

Liu, M. L., Winther, B. L. and Kay, M. A. (1996) Pseudotransduction of hepatocytes by using concentrated pseudotyped vesicluar stomatits virus G glycoprotein (VSV-G)-Moloney murine leukemia virus-derived retrovirus vectors: comparison of VSV-G and amphotropic vectors for hepatic gene transfer. *J. Virol.* **70,** 2497-502.

Lu, Y. L., Bennet, R. P., Wills, J. W., Gorelick, R. and Ratner, L. (1995) A leucine triplet repeat sequence (LXX) 4 in $p6^{gag}$ is important for *vpr* incorporation into human immunodeficiency virus type 1 particles. *J. Virol.* **69,** 7630-8.

Lu , Y. L., Spearman, P. and Ratner, L. (1993) Human immunodeficiency virus type 1 viral protein R localisation in infected cells and virions. *J. Virol.* **67,** 6542-50.

Luban, J. and Goff, S. P. (1994) Mutational analysis of *cis*-acting packaging signals in human immunodeficiency virus type 1 RNA. *J. Virol.* **68,** 3784-93.

Luo, T., Douglas, J. L., Livingston, R. L. and Garcia, J. V. (1998) Infectivity enhancemer by HIV-1 nef is dependent on the pathway of virus entry: Implications for HIV-1 base gene transfer systems. *Virology* **241,** 224-33.

Luo, D. and Saltzman, W. M. (2000) Synthetic DNA delivery systems. *Nat. Biotech.* **18** 33-7.

Lynch, C. and Miller, A. D. (1991) Production of high titre helper virus free retrovira vectors by cocultivation of packaging cells with different host ranges. *J. Virol.* **65,** 3887 3890.

Lynch, C. M., Israel, D. I., Kaufman, R. J. and Miller, A. D. (1993) Sequences in the coding region of clotting factor VIII act as dominant inhibitors of RNA accumulation and protein production. *Hum. Gene Ther.* **4,** 259-72.

Madani, N. and Kabat, D. (1998) An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral *vif* protein. *J. Virol.* **72,** 10251-5.

Mader, S. and White, J. H. (1993) A steroid-inducible promoter for the controlled overexpression of cloned genes in eukaryotic cells. *PNAS* **90,** 5603-7.

Madzak, C., Margot, A. and Sarasin, A. (1992) Efficient in vivo encapsidation of a shuttle vector into pseudo-simian virus 40 virions using a shuttle virus as helper. *J. Gen. Virol.* **73,** 1533-6.

Maldarelli, F., Martin, M. A. and Strebel, K. (1991) Identification of post transcriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: novel level of gene regulation. *J. Virol.* **65,** 5732-43.

Malim, M., Hauber, J., Les, Y., Maizel, J. V. and Cullen, B. R. (1989) The HIV-1 rev transactivate through a structural target sequence to activate nuclear export of unspliced mRNA. *Nature* **338,** 254-7.

Mann, M. J., Gibbons, G. H., Hutchinson, H., Poston, R. S., Hoyt, E. G. Robbins, R. C. and Dzau, V. J. (1999) Pressure mediated oligonucleotide transfection of rat and human cardiovascular tissues. *PNAS* **96,** 6411-6.

Mann, R. and Baltimore, D. (1985) Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. *Virology* **54,** 401-7.

Mann, R. S., Mulligan, R. C. and Baltimore, D. (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33,** 871-9.

Markovitz, D. M. (1993) Infection with the human immunodeficiency virus type 2. *Ann. Intern. Med.* **118,** 211-8.

Markowitz, D., Goff, S. and Bank, A. (1988a) A safe packaging line for gene transfer, separating viral genes on two different plasmids. *J. Virol.* **62,** 1120-4.

Markowitz, D., Goff, S. and Bank, A. (1988b) Construction and use of a safe and efficient amphotropic packaging cell line. *Virology* **167,** 400-6.

Mautino, M. R., Ramsey, W. J., Reiser, J. and Morgan, R. A. (2000) Modified human immunodeficiency virus-based lentiviral vectors display decreased sensitivity to trans-dominant Rev. *Hum. Gene. Ther.* **11,** 895-908.

McBride, M. S. and Panganiban, A. (1996) The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures [published erratum appears in *J. Virol.* 1997 Jan; 71 (1):858] *J. Virol.* **70,** 2963-73.

McBride, M. S. and Panganiban, A. T. (1997) Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J. Virol.* **71,** 2050-8.

McBride, M. S., Schwartz, M. D. and Panganiban, A. T. (1997) Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterisation of *cis* elements required for encapsidation. *J. Virol.* **71,** 4544-54.

McClure, M. and Erlwein, O. (1995) Foamy viruses-Pathogenicity or potential? *Rev. Mec Virol.* **5,** 229-37.

McIvor, R. S. (1990) Deletion in a recombinant retroviral vector resulting from a crypti splice donor signal in the Moloney leukemia virus envelope gene. *Virology* **176,** 652-5.

Metharom, P., Takyar, S., Xia, H. H., Ellem, K. A., Macmillan, J., Shepherd, R. W. Wilcox, G. E. and Wei, M. Q. (2000) Novel bovine lentiviral vectors based on Jembran disease virus. *J. Gene Med.* **2,** 176-85.

Michael, N. L., Morrow, P., Mosca, J., Vahey, M., Burke, D. and Redfield, R. R. (1991 Induction of human immunodeficiency virus type 1 expression in chronically infected cells is associated primarily with a shift in RNA splicing patterns. *J. Virol.* **65,** 1291-303.

Miller, A. D. (1990) Retrovirus packaging cells. *Hum. Gene Ther.* **1,** 5-14.

Miller, A. D. (1992a) Human gene therapy comes of age. *Nature* **357,** 455-60.

Miller, A. D. (1992b) Retroviral vectors. *Curr. Top. Microbiol. Immunol.* **158,** 1-24.

Miller, A. D. and Buttimore, C. (1986) Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Mol. Cell Biol.* **6,** 2895-02.

Miller, A. D. and Chen, F. C. (1996) Retrovirus packaging cells based on 10A1 murine leukemia virus for production of vectors that use multiple receptors for cell entry. *J. Virol.* **70,** 5564-71.

Miller, A. D. and Rosman, G. J. (1989) Improved retroviral vectors for gene transfer and expression. *BioTechniques* **7,** 980-90.

Miller, A. D., Law, M. F. and Verma, I. M. (1985) Generation of helper free amphotropic retroviruses that transduce a dominant-acting, methotrexate-resistant dihydrofolate reductase gene. *Mol. Cell. Biol.* **5,** 431-7.

Miller, A. D., Miller, D. G., Garcia, J. V. and Lynch, C. M. (1993) Use of retroviral vectors for gene transfer and expression. *Methods Enzymol.* **217,** 581-99.

Miller, A. D., Trauber, D. R. and Buttimore, C. (1986) Factors involved in production of helper virus free retrovirus vectors. *Somatic Cell Mol. Genet.* **12,** 175-83.

Miller, D. G. and Miller, A. D. (1994) A family of retroviruses that use related phosphate transporters for cell entry. *J. Virol.* **68,** 8270-6.

Miller, D. G., Adam, M. A. and Miller, A. D. (1990) Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol. Cell. Biol.* **10,** 4239-42.

Miller, M. D., Warmerdam, M. T., Page, K. A., Feinberg, M. B. and Greene, W. C. (1995) Expression of the human immunodeficiency virus type 1 (HIV-1) *nef* gene during HIV-1 production increases progeny particle infectivity independently of gp160 or viral entry. *J. Virol.* **69,** 579-84.

Miller, R. H. and Sarver, N. (1997) HIV accessory proteins as therapeutic targets. *Nat. Med.* **3,** 389-94.

Miyoshi, H., Blömer, U., Takahashi, M., Gage, F. H. and Verma, I. M. (1998) Development of a self-inactiviating lentivirus vector. *J. Virol.* **72,** 8150-7.

Miyoshi, H., Takahashi, M., Gage, F. H. and Verma, I. M. (1997) Stable and efficient gene transfer into the retina using an HIV-1 based lentiviral vector. *PNAS* **94,** 10319-23.

Mochizuki, H., Schwartz, J. P., Tanaka, K., Brady, R. O. and Reiser, J. (1998) High-titre human immunodeficiency virus type 1 based vector systems for gene delivery into nondividing cells. *J. Virol.* **72,** 8873-83.

Moloney, J. B. (1960) Biological studies on a lymphoid leukemia virus extracted from sarcoma S.37. I. Origin and introductory investigations. *J. Natl. Cancer Inst.* **24,** 933-51.

Morgan, J. R., Tompkins, R. G. and Yarmush, M. L. (1993) Advances in recombinar retroviruses for gene delivery. *Ad. Drug Del. Rev.* **12,** 143-158.

Moss, B. (1992) Vaccinia and other poxvirus expression vectors. *Curr. Top. Microbio Immunol.* **158,** 25-38.

Mselli-Lakhal, L., Favier, C., Da Siva Teixeira, M. F., Chettab, K., Legras, C., Ronfort, C Verdier, G., Mornex, J. F. and Cheblоune, Y. (1998) Defective RNA packaging i responsible for low transduction efficiency of CAEV-based vectors. *Arch. Virol.* **143,** 681 95.

Mullen, C. A., Snitzer, K., Culver, K. W., Morgan, R. A., Anderson, W. F. and Blaese, R M. (1996) Molecular analysis of T lymphocyte-directed gene therapy for adenosin deaminase deficiency: long-term expression in vivo of genes introduced with a retrovira vector. *Hum. Gene Ther.* **7,** 1123-9.

Muriaux, D., Girard, P. M., Bonnet-Mathoniere, B. and Paoletti, J. (1995) Dimerization o HIV-1Lai RNA at low ionic strength. An autocomplementary sequence in the 5' leade region is evidenced by an antisense oligonucleotide. *J. Biol. Chem.* **270,** 8209-16.

Mustafa, F. and Robinson, H. L. (1993) Context dependent role of human immunodeficiency virus type 1 auxiliary genes in the establishment of chronic virus producers. *J. Virol.* **67,** 6909-15.

Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M. and Trono, D. (1996a) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272,** 263-7.

Naldini, L., Blömer, U., Gage, F. H., Trono, D. and Verma, I. M. (1996b) Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *PNAS* **93,** 11382-8.

Nasioulas, G., Zolotukhin, A. S., Tabernero, C., Solomin, L., Cunningham, C. P., Pavlakis, G. N. and Felber, B. K. (1994) Elements distinct from human immunodeficiency virus type 1 splice sites are responsible for the Rev dependence of *env* mRNA. *J. Virol.* **68,** 2986-93.

Neufeld, E. F. and Meunzer, J. (1995) in The Metabolic Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S. and Valle, D., eds.), 7th edn., chapter 78, pp. 2465-2494, McGraw-Hill, New York.

Neumann, E., Schaefer-Rider, M., Wang, Y. and Hofschneider, P. H. (1982) Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* **1,** 841-5.

Neuveut, C. and Jeang, K. T. (1996) Recombinant human immunodeficiency virus type 1 genomes with tat unconstrained by overlapping reading frames reveal residues in Tat important for replication in tissue culture. *J. Virol.* **70,** 5572-81.

Nguyen, D. M., Wiehle, S. A., Roth, J. A. and Cristiano, R. J. (1997) Gene delivery into malignant cells in vivo by a conjugated adenovirus DNA complex. *Cancer Gene Ther.* **4,** 183-90.

Niederman, T. M. J., Thielan, B. J. and Ratner, L. (1989) Human immunodeficiency virus type 1 negative factor is a transcriptional silencer. *PNAS* **86,** 1128-32.

Olsen, J. C. (1998) Gene transfer vectors derived from equine infectious anaemia virus. *Gene Ther.* **5,** 1481-7.

Olsen, H. S., Cochrane, A. W. and Rosen, C. (1992) Interaction of cellular factors with intragenic cis-acting repressve sequences within the HIV genome. *Virology* **191,** 709-15.

Oshima, A., Kyle, J. W., Miller, R. D., Hoffmann, J. W., Powell, P. P., Grubb, J. H., Sly, W. S., Tropak, M., Guise, K. S. and Gravel, R. A. (1987) Cloning, sequencing, and expression of cDNA for human β-glucuronidase. *PNAS* **84,** 685-9

Ott, M., Emiliani, S., Van Lint, C., Herbein, G., Lovett, J., Chirmule, N., McCloskey, T Pahwa, S. and Verdin, E. (1997) Immune hyperactivation of HIV-1 infected T cel mediated by Tat and the CD28 pathway. *Science* **275,** 1481-5.

Otto, E., Jones-Trower, A., Vanin, E. F., Stambaugh, K., Mueller, S. N., Anderson, W. F and McGarrity, G. (1994) Characterisation of a replication competent retrovirus resultin from recombination of packaging and vector sequences. *Hum. Gen. Ther.* **5,** 567-75.

Overell, R. W., Weisser, K. E. and Cosman, D. (1988) Stably transmitted triple promote retroviral vectors and their use in transformation of primary mammalian cells. *Mol. Cell Biol.* **8,** 1803-8.

Page, K. A., Landau, N. R. and Littman, D. R. (1990) Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. *J. Virol.* **64,** 5270-76.

Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C. and Marquet, R. (1996) A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA *PNAS* **93,** 5572-7.

Parolin, C., Dorfman, T., Palu, G., Göttlinger, H. and Sodroski, J. (1994) Analysis in human immunodeficiency virus type 1 vectors of *cis* acting sequences that affect gene transfer into human lymphocytes. *J. Virol.* **68,** 3888-95.

Pasquinelli, A. E., Ernst, R. K., Lund, E., Grimm, C., Zapp, M. L., Rekosh, D., Hammarskjöld, M. L. and Dahlberg, J. E. (1997) The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO J.* **16,** 7500-10.

Paul, R. W., Weisser, K. E., Loomis, A., Sloane, D. L., LaFoe, D., Atkinson, E. M. and Overell, R. W. (1997) Gene transfer using a novel fusion protein, GAL4/invasin. *Hum. Gene Ther.* **8,** 1253-62.

Pear, W. S., Nolan, G .P., Scott, M. L. and Bultimore, D. (1993) Production of high-titre helper-free retroviruses by transient transfection. *PNAS* **90,** 8392-6.

Peebles, P. T. (1975) An in vitro focus-induction assay for xenotropic murine leukemia virus, feline leukemia virus C and the feline primate viruses RD-114/CCC/M-7. *Virology* **67,** 288-91.

Pepin, J., Morgan, G., Dunn, D., Gevao, S., Mendy, M., Gaye, I., Scollen, N., Tedder, R. and Whittle, H. (1991) HIV-2 induced immunosuppression among asymptomatic west African prostitutes: Evidence that HIV-2 is pathogenic, but less so than HIV-1. *AIDS* **4,** 1165-72.

Piguet, V., Gu, F., Foti, M., Demaurex, N., Gruenberg, J., Carpentier, J. L. and Trono, D. (1999) *Nef* induced CD4 degradation: a diacidic-based motif in *nef* functions as a lysosomal targeting signal through the binding of beta-COP in endosomes. *Cell* **97,** 63-73.

Planelles, V., Bachelerie, F., Jowett, J. B. M., Haislip, A., Xie, Y., Banooni, P., Masuda, T. and Chen, I. S. Y. (1995) Fate of the human immunodeficiency virus type 1 provirus in infected cells: a role for *vpr*. *J. Virol.* **69,** 5883-9.

Poeschla, E., Gilbert, J., Li, X., Huang, S., Ho, A. and Wong-Staal, F. (1998a) Identification of a human immunodeficiency virus type 2 (HIV-2) encapsidation determinant and transduction of nondividing human cells by HIV-2 based lentivirus vectors. *J. Virol.* **72,** 6527-36.

Poeschla, E., Wong-Staal, F. and Looney, D. J. (1998b) Efficient transduction of nondividing human cells by feline immunodeficiency virus lentiviral vectors. *Nat. Med.* **4,** 354-7.

Pomerantz, F. J., Seshamma, T. and Trono, D. (1992) Efficient replication of human immunodeficiency virus type 1 requires a threshold level of Rev: potential implications for latency. *J. Virol.* **66,** 1809-13.

Poznansky, M., Lever, A., Bergeron, L., Haseltine, W. and Sodroski, J. (1991) Gene transfer into human lymphocytes by a defective human immunodeficiency virus type 1 vector. *J. Virol.* **65,** 532-6.

Preston, B. D., Poisez, B. J. and Loeb, L. A. (1988) Fidelity of HIV-1 reverse transcriptase *Science* **242,** 1168-71.

Pryciak, P. M. and Varmus, H. E. (1992) Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69,** 769-80.

Ptashne, M. (1988) How eukaryotic transcriptional activators work. *Nature* **355,** 683-9.

Purcell, D. F. and Martin, M. A. (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication and infectivity. *J. Virol* **67,** 6365-78.

Ratner, L., Vander-Heyden, N., Garcia, J., Polinsky, M., Westervelt, P., Becich, M. (1991) Formation of non infectious HIV-1 virus particles lacking a full length envelope protein *AIDS Res. Hum. Retroviruses* **7,** 287-94.

Refaeli, Y., Levy, D. N. and Weiner, D. B. (1995) The glucocorticoid receptor type II complex is a target of the HIV-1 *vpr* gene product. *PNAS* **92,** 3621-5.

Reicin, A. S., Paik, S., Berkowitz, R. D., Luban, J., Lowy, I. and Goff, S. P. (1995) Linker insertions in the HIV-1 gag gene: effects on virion particle assembly, release and infectivity. *J. Virol.* **69,** 642-50.

Reiser, J., Harmison, G., Kluepfel-Stahl, S., Brady, R. O., Karlsson, S. and Schubert, M. (1996) Transduction of nondividing cells using pseudotyped defective high-titre HIV type 1 particles. *PNAS* **93,** 15266-71.

Rethwilm, A. (1995) Regulation of foamy virus gene expression. *Curr. Top. Microbiol. Immunol.* **193,** 1-23.

Rhee, S. S. and Hunter, E. (1991) Amino acid substitutions within the matrix protein of type D retroviruses affect assembly, transport and membrane association of a capsid. *EMBO J.* **10,** 535-46.

Richardson, J. H., Child, L. A. and Lever, A. M. L. (1993) Packaging of human immunodeficiency virus type 1 RNA requires *cis*-acting sequences outside the 5' leader region. *J. Virol.* **67,**3997-4005.

Rigg, R. J., Chen, J., Dando, J. S., Forestell, S. P., Plavec, I. and Bohnlein, E. (1996) A novel human amphotropic packaging cell line: high titre, complement resistance and improved safety. *Virology* **218,** 290-5.

Rigg, A. and Sikora, K. (1997) Genetic prodrug activation therapy. *Mol. Med. Today* **3,** 359-66.

Rizvi, T. A. and Panganiban, A. T. (1993) Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. *J. Virol.* **67,** 2681-8.

Rizvi, T. A., Schmidt, R. D., Lew, K. A. and Keeling, M. E. (1996) Rev/RRE-independent Mason-Pfizer monkey virus constitutive transport element-dependent propagation of SIVmac239 vectors using a single round of replication assay. *Virology* **222,** 457-463.

Roberts, J. D., Bebenek, K. and Kunkel, T. A. (1988) The accuracy of reverse transcriptase from HIV-1. *Science* **242,** 1171-3.

Roe, T., Reynolds, T. C., Yu, G. and Brown, P. O. (1993) Integration of murine leukemia virus depends on mitosis. *EMBO J.* **12,** 2099-108.

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

Rogers, S. (1959) Induction of arginase in rabbit epithelium by the Shope rabbit papilloma virus. *Nature* **183,** 1815-6.

Rogers, S., Lowenthal, A., Terheggen, H. G. and Columbo, J. P. (1973) Induction of arginase activity with the Shope papilloma virus in tissue culture cells from an argininemic patient. *J. Exp. Med.* **137,** 1091-6.

Rosen, C. A., Terwilliger, E., Dayton, A., Sodroski, J. and Haseltine, W. A. (1988 Intragenic *cis*-acting ART gene-responsive sequences of the human immunodeficienc virus. *PNAS* **85,** 2071-5.

Rosenberg, S. A., Zhai, Y. F., Yang, J. C., Schwartzentruber, D. J., Hwu, P., Marincola, F M., Topalian, S. L., Restifo, N. P., Seipp, C. A., Einhorn, J. H., Roberts, B. and White, D E. (1998) Immunizing patients with metastatic melanoma using recombinant adenoviruses encoding MART-1 or gp 100 melanoma antigens. *J. Natl. Cancer Inst.* **90,** 1894-900.

Rous, P. (1911) A sarcoma of the fowl transmissible by an agent separable from the tumour cells. *J. Exp. Med.* **13,** 397-411.

Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., Haseltine, W. A. and Rosen, C. A. (1989) Structural and functional characterisation of human immunodeficiency virus tat protein. *J. Virol.* **63,** 1-8.

Russell, D. W. and Miller, A. D. (1996) Foamy virus vectors. *J. Virol.* **70,** 217-22.

Saavedra, C., Felber, B. and Izaurralde, E. (1997) The simian retrovirus-1 constitutive transport element, unlike the HIV-1 RRE, uses factors required for cellular mRNA export. *Curr. Biol.* **7,** 619-28.

Sadaie, M. R., Rappaport, J., Benter, T., Josephs, S. F., Willis, R. and Wong-Staal, F. (1988) Missense mutations in an infectious human immunodeficiency viral genome: functional mapping of tat and identification of the rev splice acceptor. *PNAS* **85,** 9224-28.

Sakaguchi, K., Zambrano, N., Baldwin, E. T., Shapiro, B. A., Erickson, J. W., Omichinski, J. G., Clore, G. M., Gronenborn, A. M. and Appella, E. (1993) Identification of a binding site for the human immunodeficiency virus type 1 nucleocapsid protein. *PNAS 90,* 5219-23.

Sakai, H., Shibata, R., Sakuragi, J., Sakuragi, S., Kawamura, M. and Adachi, A. (1993) Cell-dependent requirement of human immunodeficiency virus type 1 *vif* protein for maturation of virus particles. *J. Virol.* **67,** 1663-6.

Samulski, R. J., Berns, K. I., Tan, M. and Muzyczka, N. (1982) Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. *PNAS* **79,** 2077-81.

Schiedner, G., Morral, N., Parks, R. J., Wu, Y., Koopmans, S. C., Langston, C., Graham, F. L., Beaudet, A. L. and Kochanek, S. (1998) Genomic DNA transfer with a high-capacity adenovirus vector results in improved in vivo gene expression and decreased toxicity. *Nat. Genet.* **18,** 180-3.

Schmidt, M. and Rethwilm, A. (1995) Replicating foamy virus-based vectors directing high level expression of foreign genes. *Virology* **210,** 167-78.

Schubert, U. and Strebel, K. (1994) Differential activities of the human immunodeficiency virus type 1-encoded *vpu* protein are regulated by phosphorylation and occur in different cellular compartments. *J. Virol.* **68,** 2260-71.

Schwartz, B., Ivanov, M. A., Pitard, B., Escriou, V., Rangara, R., Byk, G., Wils, P., Crouzet, J. and Scherman, D. (1999) Synthetic DNA-compacting peptides derived from human sequence enhance cationic lipid mediated gene transfer in vitro and in vivo. *Gene Ther.* **6,** 282-92.

Schwartz, O., Maréchal, V., Danos, O. and Heard, J. M. (1995) HIV-1 *nef* increases the efficiency of reverse transcription in the infected cells. *J. Virol.* **69,** 4053-9.

Schwartz, S., Campbell, M., Nasioulas, G., Harrison, J., Felber, B. K. and Pavlakis, G. N. (1992a) Mutational inactivation of an inhibitory sequence in human immunodeficiency virus type 1 results in Rev independent gag expression. *J. Virol.* **66,** 7176-82.

Schwartz, S., Felber, B. K. and Pavlakis, G. N. (1992b) Distinct RNA sequences in the *gag* region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. *J. Virol.* **66,** 150-9.

Schwartz, S., Felber, B. K., Benko, D. M., Fenyo, E. M. and Pavlakis, G. N. (1990) Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *J. Virol.* **64,** 2519-29.

Schweizer, M., Turek, R., Hahn, H., Schliephake, A., Netzer, K. O., Eder, G., Reinhardt M., Rethwilm, A. and Neumann Haefelin, D. (1995) Markers of foamy virus infections in monkeys, apes and accidentally infected humans: appropriate testing fails to confirm suspected foamy virus prevalence in humans. *AIDS Res. Hum. Retroviruses* **11,** 161-170.

Shea, L. D., Smiley, E., Bonadio, J. and Mooney, D. J. (1999) DNA delivery from polyme matrices for tissue engineering. *Nat. Biotechnol.* **17,** 551-4.

Sheng, N. and Erickson-Vitanen, S. (1994) Cleavage of p15 in vitro by HIV-1 protease is RNA dependent. *J. Virol.* **68,** 6207-14.

Shimada, T., Fujii, H., Mitsuya, H. and Nienhuis, A. W. (1991) Targeted and highly efficient gene transfer into CD4+ cells by a recombinant human immunodeficiency virus retroviral vector. *J. Clin. Invest.* **88,** 1043-7.

Shioda, T. and Shibuta H. (1990) Production of human immunodeficiency virus (HIV) like particles from cells infected with recombinant vaccinia viruses carrying the gag gene of HIV. *Virology* **175,** 139-48.

Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B. and Ehresmann, C. (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *PNAS* **91,** 4945-9.

Sly, W. S., Quinton, B. A., McAlister, W. H. and Rimoin, D. L. (1973) Beta glucuronidase deficiency: a report of clinical, radiologic, and biochemical features of a new mucoploysaccharidosis. *J. Pediatr.* **82,** 249-57

Snyder, R. O., Miao, C. H., Patjin, G. A., Spratt, S. K., Danos, O., Nagy, D., Gown, A. M., Winther, B., Meuse, L., Cohen, L. K., Thompson, A. R. and Kay, M. A. (1997) Persistent and therapeutic concentrations of human factor IX in mice after hepatic gene transfer of recombinant AAV vectors. *Nat. Genet.* **16,** 270-6.

Sodroski, J., Goh, W. C., Rosen, C., Campbell, K. and Haseltine, W. A. (1986) Role of the HTLV-III/LAV envelope in syncytium formation and cytopathicity. *Nature* **322,** 470-4.

Sorge, J., Wright, D., Erdman, V. D. and Cutting, A. E. (1984) Amphotropic retrovirus vector system for human cell gene transfer. *Mol. Cell. Biol.* **4,** 1730-7.

Sorgi, F. L., Bhattacharya, S. and Huang, L. (1997) Protamine sulphate enhances lipid mediated gene transfer. *Gene Ther.* **4,** 961-8.

Spearman, P., Wang, J. J., Vander-Heyden, N. and Ratner, L. (1994) Identification of human immunodeficiency virus type 1 Gag protein domains essential to membrane binding and particle assembly. *J. Virol.* **68,** 3232-42.

Spector, D. H., Wade, E., Wright, D. A., Koval, V., Clark, C., Joquish, D. and Spector, S. A. (1990) Human immunodeficiency virus pseudotypes with expanded cellular and species tropism. *J. Virol.* **64,** 2298-308.

Springett, G. M., Moen, R. C., Anderson, R. M., Blaese, R. M. and Anderson, W. F. (1989) Infection efficiency of T lymphocytes with amphotropic retroviral vectors is cell cycle dependent. *J. Virol.* **63,** 3865-9.

Srinivasakumar, N., Chazal, N., Helga-Maria, C., Prasad, S., Hammarskjöld, M. L. and Rekosh, D. (1997) The effect of viral regulatory protein expression on gene delivery by human immunodeficiency virus type 1 vectors produced in stable packaging cell lines. *J. Virol.* **71,** 5841-8.

Srinivasakumar, N. and Schuening, F. G. (1999) A lentivirus packaging system based on alternative RNA transport mechanisms to express helper and gene transfer vector RNAs and its use to study the requirement of accessory proteins for particle formation and gene deleivery. *J. Virol.* **73,** 9589-98.

Stemmer, W. P. C., Crameri, A., Ha, K. D., Brennan, T. M. and Heyneker, H. L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164,** 49-53.

Strebel, K., Klimkait, T. and Martin, M. A. (1988) A novel gene of HIV-1, *vpu*, and its 16 kDa product. *Science* **241,** 1221-3.

Strebel, K., Klimkait, T., Maldarelli, F. and Martin, M. A. (1989) Molecular an biochemical analyses of human immunodeficiency type 1 *vpu* protein. *J. Virol.* **63,** 3784 91.

Stuhlmann, H., Jaenisch, R. and Mulligan, R. C. (1989) Construction and properties o replication competent murine retroviral vectors encoding methotrexate resistance. *Mol Cell. Biol.* **9,** 100-8.

Sundquist, W. I. And Heaphy, S. (1993) Evidence for interstrand quadruplex formation ir the dimerisation of human immunodeficiency virus 1 genomic RNA. *PNAS* **90,** 3393-7.

Sutton, R. E., Wu, H. T., Rigg, R., Bohnlein, E. and Brown, P. O. (1998) Human immunodeficiency virus type 1 vectors efficiently transduce human haematopoietic stem cells. *J. Virol.* **72,** 5781-8.

Tabernero, C., Zolotukhin, A. S., Valentin, A., Pavlakis, G. N. and Felber, B. K. (1996) The posttranscriptional control element of the simian retrovirus type 1 forms an extensive RNA secondary structure necessary for its function. *J. Virol.* **70,** 5998-6011.

Tang, H., Gaietta, G. M., Fischer, W. H., Ellisman, M. H. and Wong-Staal, F. (1997) A cellular cofactor for the constitutive transport element of type D retrovirus. *Science* **276,** 1412-5.

Terwilliger, E. F., Godin, B., Sodroski, J. G. and Haseltine, W. A. (1989) Construction and use of a replication competent human immunodeficiency virus (HIV-1) that expresses the chloramphenical acetyltransferase enzyme. *PNAS* **86,** 3857-61.

Trono, D. (1995) HIV accessory proteins: Leading roles for the supporting cast. *Cell* **82,** 189-92.

Tschopp, R. R., Brandner, S., Marino, S., Bothe, K., Horak, I., Rethwilm, A. and Aguzzi, A. (1996) Analysis of the determinants of neurotropism and neurotoxicity of HFV in transgenic mice. *Virology* **216,** 338-46.

Uchida, N., Sutton, R. E., Friera, A. M., He, D., Reitsma, M. J., Chang, W. C., Veres, G., Scollay, R. and Weissman, I. L. (1998) HIV, but not murine leukemia virus, vectors mediate high efficiency gene transfer into freshly isolated G0/G1 human haematopoietic stem cells. *PNAS* **95,** 11939-44.

Vaheri, A. and Pagano, J. S. (1965) Infectious poliovirus RNA: a sensitive method of assay. *Virology* **27,** 434-6.

Varmus, H. E., Padgett, T., Heasley, S., Simon, G. and Bishop, J. M. (1977) Cellular functions are required for the synthesis and integration of avian sarcoma virus-specific DNA. *Cell* **11,** 307-19.

Venkatesh, L. K., Mohammed, S. and Chinnadurai, G. (1990) Functional domains of the HIV-1 rev gene required for trans-regulation and subcellular localization. *Virology* **176,** 39-47.

Vicenzi, E., Dimitrov, D. S., Engelman, A., Migone, T. S., Purcell, D. F., Leonard, J., Englund, G. and Martin, M. A. (1994) An integration defective U5 deletion mutant of human immunodeficiency virus type 1 reverts by eliminating additional long term repeat sequences. *J. Virol.* **68,** 7879-90.

von Schwedler, U., Kornbluth, R. S. and Trono, D. (1994) The nuclear localisation signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent T lymphocytes. *PNAS* **91,** 6992-6.

von Schwedler, U., Song, J., Aiken, C. and Trono, D. (1993) *Vif* is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *J. Virol.* **67,** 4945-55.

Wagner, R., Graf, M., Bieler, K., Wolf, H., Grunwald, T., Foley, P. and Uberla, K. (2000) Rev-independent expression of synthetic gag-pol genes of human immunodeficiency virus type 1 and simian immunodeficiency virus: Implication for the safety of lentiviral vectors. *Hum. Gene Ther.* **11,** 2403-2413.

Wakefield, J. K., Kang, S. M. and Morrow, C. D. (1996) Construction of a type 1 huma immunodeficiency virus that maintains a primer binding site complementary to $tRNA^{His}$. . *Virol.* **70,** 966-75.

Wakefield, J. K. and Morrow, C. D. (1996) Mutations within the primer binding site of th human immunodeficiency virus type 1 define sequence requirements essential for revers transcription. *Virology* **220,** 290-8.

Walsh, C. E., Liu, J. M., Xiao, X., Young, N. S., Nienhuis, A. W. and Samulski, R. J (1992) Regulated high level expression of a human gamma-globin gene introduced intc erythroid cells by an adeno-associated virus vector. *PNAS* **89,** 7257-61.

Watanabe, S. and Temin, H. M. (1983) Construction of a helper cell line for aviar reticuloendotheliosis virus cloning vectors. *Mol. Cell. Biol.* **3,** 2241-9.

Weeks, K. M. and Crothers, D. M. (1991) RNA recognition by Tat-derived peptides: interaction in the major groove? [published erratum appears in *Cell* 1992 Sep 18;70(6):following 1068] *Cell* **66,** 577-88.

Wehrly, K. and Chesebro, B. (1997) p24 Antigen capture assay for quantification of human immunodeficiency virus using readily inexpensive reagents. *Meth. Enzymol.* **12,** 288-93.

Weiss, R. A. (1988) Foamy retroviruses. A virus in search of a disease [news]. *Nature* **333,** 497-8.

Weiss, R., Teich, N., Varmus, H. and Coffin, J. (1982) Molecular biology of tumor viruses. In: *RNA Tumor Viruses*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

White, S. M., Renda, M., Nam, N. Y., Klimatcheva, E., Zhu, Y., Fisk, J., Halterman, M., Rimel, B. J., Federoff, H., Pandya, S., Rosenblatt, J. D. and Planelles, V. (1999) Lentivirus vectors using human and simian immunodeficiency virus elements. *J. Virol.* **73,** 2832-40.

Willey, R. L., Maldarelli, F., Martin, M. A. and Strebel, K. (1992a) Human immunodeficiency virus type I *vpu* protein regulates the formation of intracellular gp160-CD4 complexes. *J. Virol.* **66,** 226-34.

Willey, R. L., Maldarelli, F., Martin, M. A. and Strebel, K. (1992b) Human immunodeficiency virus type 1 *vpu* protein induces rapid degradation of CD4. *J. Virol.* **66,** 7193-200.

Wills, J. W. and Craven, R. C. (1991) Form, function, and use of retroviral gag proteins. *AIDS* **5,** 639-54.

Wilson, J. M. (1996) Adenoviruses as gene-delivery vehicles. *New Eng. J. Med.* **334,** 1185-9.

Wilson, W., Braddock, M., Adams, S. E., Rathgen, P. D., Kingsman, S. M. and Kingsman, A. J. (1988) HIV expression strategies: Ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* **55,** 1159-69.

Wodrich, H., Schambach, A. and Kräusslich, H. G. (2000) Multiple copies of the Mason-Pfizer monkey virus constitutive RNA transport element lead to enhanced HIV-1 Gag expression in a context-dependent manner. *Nucl. Acids Res.* **28,** 901-10

Woffendin, C., Ranga, U., Yang, Z. Y., Xu, L. and Nabel, G. J. (1996) Expression of a protective gene prolongs survival of T cells in human immunodeficiency virus-infected patients. *PNAS* **93,** 2889-94.

Wolfe, J. H., Deshmane, S. L. and Fraser, N. W. (1992) Herpesvirus vector gene transfer and expression of β-glucuronidase in the central nervous system of MPS VII mice. *Nat. Genet.* **1,** 379-84.

Wolfe, J. H., Sands, M. S., Barker, J. E., Gwynn, B., Rowe, L. B., Vogler, C. A. and Birkenmeier, E. H. (1992) Reversal of pathology in murine mucoploysaccharidosis type VII by somatic cell gene transfer. *Nature* **360,** 749-53.

Xu, K., Hong, M., McCown, T. J., Verma, I. M. and Kafri, T. (2001) Generation of stable cell line producing high-titre self-inactivating lentiviral vectors. *Mol. Ther.* **3,** 97 104.

Wu, X., Wakefield, J. K., Liu, H., Xiao, H., Kralovics, R., Prchal, J. T. and Kappes, J. C (2000) Development of a novel trans-lentiviral vector that affords predictable safety. *Mol Ther.* **2,** 47-55.

Yao, X. J., Subbramanian, R. A., Rougeau, N., Boisvert, F., Bergeron, D. and Cohen, E. A (1995) Mutagenic analysis of human immunodeficiency virus type 1 *vpr*: Role of a predicted N-terminal α-helical structure in vpr nuclear localisation and virior incorporation. *J. Virol.* **69,** 7032-44

Yee, J. K., Miyanohara, A., LaPorte, P., Bouic, K., Burns, J. C. and Friedmann, T. (1994) A general method for the generation of high-titre, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. *PNAS* **91,** 9564-8.

Yee, J. K., Moores, J. C., Jolly, D. J., Wolff, J. A., Respess, J. G. and Friedmann, T. (1987) Gene expression from transcriptionally disabled retroviral vectors. *PNAS* **84,** 5197-201.

You, J. C. and McHenry, C. S. (1994) HIV nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. *J. Biol. Chem.* **269,** 31491-5.

Yu, H., Rabson, A. B., Kaul, M., Ron, Y. and Dougherty, J. P. (1996) Inducible human immunodeficiency virus type 1 packaging cell lines. *J. Virol.* **70,** 4530-7.

Yu, S. F., Baldwin, D. N., Gwynn, S. R., Yendapalli, S. and Linial, M. L. (1996a) Human foamy virus replication: A pathway distinct from that of retroviruses and hepadnaviruses. *Science* **271,** 1579-82.

Yu, S. F., Stone, J. and Linial, M. L. (1996b) Productive persistent infection of hematopoietic cells by human foamy virus. *J. Virol.* **70,** 1250-4.

Zack, J. S., Haislip, A. M., Krogstad, P. and Chen, I. S. Y. (1992) Incompletely reverse transcribed human immunodeficiency virus type 1 genomes in quiescent cells can function as intermediates in the retroviral life cycle. *J. Virol.* **66,** 1717-25.

Zapp, M. L., Hope, T. J., Parslow, T. G. and Green, M. R. (1991) Oligomerisation and RNA binding domains of the type 1 human immunodeficiency virus *rev* protein: A dual function for an arginine-rich binding motif. *PNAS* **88,** 7734-8.

Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. and Charneau, P. (2000) HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101,** 173-85.

Zhang, G., Budker, V. and Wolff, J. A. (1999) High levels of foreign gene expression in hepatocytes after tail vein injections of naked plasmid DNA. *Hum. Gene Ther.* **10,** 1735-7.

Zhu, J. and Cunningham, J. M. (1993) Minus-strand DNA is present within murine type C ecotropic retroviruses prior to infection. *J. Virol.* **67,** 2385-8.

Zhu, N., Liggit, D. and Debs, R. (1993) Systemic gene expression after intravenous DNA delivery into adult mice. *Science* **261,** 209-11.

Zolotukhin, A. S., Valentin, A., Pavlakis, G. N. and Felber, B. K. (1994) Continuous propagation of RRE (-) and Rev (-) RRE (-) human immunodeficiency virus type 1 molecular clones containing a *cis* acting element of simian retrovirus type 1 in human peripheral blood lymphocytes. *J. Virol.* **68,** 7944-52.

Zufferey, R., Dull, T., Mandel, R. J., Bukovsky, A., Quirox, D., Naldini, L. and Trono, D. (1998) Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* **72,** 9873-80.

Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L. and Trono, D. (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotech.* **15,** 871-5.

PUBLICATIONS ARISING FROM THIS THESIS

Fuller, M. and Anson, D. S. (2001) Helper plasmids for production of HIV-1 derived vectors. *Hum. Gene Ther.* In Press.

Anson, D. S. and Fuller, M. (2001) Evaluation of the effect of HIV-1 sequence elements on vector performance. *Hum. Gene Ther.* Submitted.