STUDIES USING PSEUDOTYPED RETROVIRAL VECTORS

Catherine H Mahoney

A thesis submitted for the degree of Doctor of Philosophy

Department of Veterinary Pathology

Faculty of Veterinary Medicine

University of Glasgow

November, 1999

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Summary

The host range of a retrovirus depends largely upon the properties of the envelope glycoprotein, Env (Hunter, 1997). As a result this host range can be changed by replacing the envelope protein with that from another virus. The pseudotypes formed in this process possess the host range and fusion properties of the virus donating the envelope protein (Závada, 1982). As retroviruses integrate into the genome of the infected host cell they have proved useful for the development of vectors for gene delivery (Miller, 1997). The host ranges of these vectors are restricted to cells that express the specific receptor recognised by the particular envelope protein. Several pseudotyped vectors have therefore been developed in order to increase host range or to target the vectors to specific cell types (Burns *et al.*, 1993; Emi *et al.*, 1991; Hatziioannou *et al.*, 1998; Miller *et al.*, 1991; Spiegel *et al.*, 1998). The production of retroviral vectors pseudotyped with the envelope glycoprotein of the rhabdovirus vesicular stomatitis virus (VSV) was a significant advance in this field (Burns *et al.*, 1993; Emi *et al.*, 1991). Not only did these vectors display the wide host range of VSV, but they were also very stable, allowing concentration to high titres. Unfortunately the VSV envelope protein causes cell fusion when expressed at high levels, making continual production of these vectors very difficult. In this study an attempt was made to produce similar vectors pseudotyped with the envelope glycoprotein of another rhabdovirus, rabies virus. These vectors displayed the host range properties of rabies, but they were produced at very low titres. The glycoprotein was capable of inducing cell fusion at low pH, indicating that it was functional. Therefore the low titres suggest that the rabies envelope protein was not incorporated efficiently into the retroviral vectors. Although these pseudotyped vectors did not prove to be useful as gene delivery vectors, the observations described here may provide interesting information on the processes of virion assembly and envelope incorporation.

The second part of this study involved investigations into the infectivity of the porcine endogenous retroviruses. Concern has been raised recently over the potential infectivity of these viruses following proposals to use porcine organs and tissues in human transplants. Endogenous retroviruses are integrated into the host germ line (Boeke and Stoye, 1997) and would therefore be present in all transplanted cells. Initial cocultivation experiments indicated that these viruses could infect human cells. No obvious interference was

observed with other mammalian C-type retroviruses, indicating that the porcine viruses recognise a novel receptor. Subsequent sequence analysis has revealed that at least three viral subgroups exist in the porcine genome. The determination of the *env* sequence of one of these viruses is reported here. The encoded protein has a very similar structure to the Env proteins of other mammalian C-type retroviruses, but shows significant differences in the predicted receptor binding regions. The porcine virus subgroups also differ from one another within these regions, suggesting that these viruses recognise different receptors, and may therefore have different host ranges.

Another potential application of the ability of retroviral vectors to incorporate foreign envelope proteins is in studies of the infectivity and host range of new viruses (Rong *et al.*, 1997; Wool-Lewis and Bates, 1998). In order to characterise these properties of the porcine endogenous retroviruses, pseudotyped vectors were produced, bearing the Env proteins of the three subgroups. Infected cells were then simply identified from the expression of a marker gene inserted in the vector genome. Analysis of the vectors revealed that at least two of the subgroups can infect human cells, and could therefore present a risk of infection to transplant recipients. Consequently it is important that further studies of these viruses are carried out in order to assess their tissue tropism and potential pathogenicity. These pseudotyped vectors should also be useful for such investigations of Env function and receptor usage of the porcine endogenous retroviruses.

Contents

List of figures

List of tables

Acknowledgements

I would firstly like to thank my supervisors: Professor David Onions for all his support and advice, particularly with the PoERV work, and Dr Simon Tucker for his guidance with the rabies G pseudotyped vectors and encouragement during the first months. I am also very grateful to all those in the Department of Veterinary Pathology who have given of their time to help me with this study. In particular, Matt Golder and Mike MacDonald in the Feline Virus Unit, Ross Blackley and Colin Wilson in the Electron Microscopy Unit, Alan May for all his photography work, and Linda Scobie for all her advice and time spent proof-reading. I would like to thank Yuko Soneoka for the HIT system plasmids which enabled production of the various pseudotyped vectors. Many thanks are also due to everyone, especially those in the Gene Therapy and MRC Retrovirus Research laboratories, who answered my questions and offered invaluable advice on many occasions.

A big thank you to all those friends, both within and without the department, who made my time in Glasgow enjoyable, and helped me to stay sane! I would especially like to thank Mike for all his encouragement and patience while I was writing this thesis. Finally, a special thank you to my family for their continuous love and support throughout it all.

This work was funded by the BBSRC and Q-One Biotech Ltd.

Declaration

The studies described in this thesis were carried out in the Department of Veterinary Pathology at the University of Glasgow, between October 1995 and June 1999. The author is responsible for all results, apart from the following. The ELISAs for FeLV Gag described in Chapters 3 and 6, and the virus isolation assay described in Chapter 6 were performed by Matt Golder and Mike MacDonald in the Feline Virus Unit, University of Glasgow Veterinary School. The Raji 4 cells described in Chapter 4 were infected with PoERV by staff at Q-One Biotech Ltd., Glasgow. Infection of Raji 4 cells with X-MLV, FeLV-B and FeLV-C, described in Chapter 4, was also performed by Matt Goulder. The electron microscopy reported in Chapter 4 was carried out by Ross Blackley and Colin Wilson in the Electron Microscopy Unit, University of Glasgow Veterinary School. The phylogenetic trees presented in Chapters 4 and 6 were constructed by Douglas Hart, Department of Veterinary Pathology, University of Glasgow.

No part of this thesis has been presented previously to any university, but it has been reproduced in part in the following publication:

Onions, D., Hart, D., Mahoney, C., Galbraith, D., and Smith, K. (1998). Endogenous retroviruses and the safety of porcine xenotransplantation. *Trends Microbiol.* **6**, 430-431.

List of abbreviations

αBtx	α-bungarotoxin
αGal	galactose-α(1-3)-galactose
Ach	acetylcholine
ALV	avian leukosis virus
A-MLV	amphotropic MLV
Amp	ampicillin
Amp^r	ampicillin resistance
AMV	avian myeloblastosis virus
APS	ammonium persulphate
Arg	arginine
ASLV	avian sarcoma leukosis virus
Asn	asparagine
Asp	aspartic acid
β-gal	β-galactosidase
BaEV	baboon endogenous virus
bp	base pair
bcfu	blue cell colony forming units
BSA	bovine serum albumin
°C	degrees celsius
CA	capsid
CAT	chloramphenicol acetyltransferase
CD	circular dichroism
cDNA	complementary DNA
CER	chicken embryo-related
cfu	colony forming units
CHO	Chinese hamster ovary
CIAP	calf intestinal alkaline phosphatase
CMV	cytomegalovirus
CNBr	cyanogen bromide
CNS	central nervous system
Cys	cysteine
dGTP	deoxyguanosine triphosphate
dH_2O	deionised water
DMEM	Dulbecco's Modified Eagle's medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphospate
DSP	dithiobis-succinimidyl propionate
DTT	dithiothreitol
ECACC	European Collection of Cell Cultures
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme linked immunosorbent assay
EM	electron microscope
E-MLV	ecotropic MLV
Env	retroviral envelope protein
ERV	endogenous retrovirus
FCS	foetal calf serum
FeLV	feline leukaemia virus
g	gram
GALV	gibbon ape leukaemia virus
GLA	glycophospho-lipid-anchoring
Glu	glutamic acid
Gly	glycine

HA	hemagglutinin
HAR	hyperacute rejection
HBS	HEPES buffered saline
HEPES	*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulphonic acid
His	histidine
HIV	human immunodeficiency virus
HTLV	human T-cell leukaemia virus
IF	immunofluorescence
Ig	immunoglobulin
IN	retroviral integrase
IRES	internal ribosome entry site
kb	kilobase
kD	kilodalton
LB	Lennox L Broth
LDL	low density lipoprotein
Leu	leucine
LTR	long terminal repeat
µl	microlitre
µM	micromolar
M	molar
M	matrix protein (rhabdoviruses)
MA	matrix protein (retroviruses)
mAb	monoclonal antibody
MDCK	Madin-Darby canine kidney
MEM	minimum essential medium
Met	methionine
mg	milligram
MHC	major histocompatibility complex
min	minute
MLV	murine leukaemia virus
mM	millimolar
MMTV	mouse mammary tumour virus
MoMLV	Moloney MLV
mRNA	messenger RNA
NA	neuraminidase
NaB	sodium butyrate
nAchR	nicotinic acetylcholine receptor
NC	nucleocapsid
NCAM	neural cell adhesion molecule
neo	neomycin phosphotransferase gene
nm	nanometre
NMJ	neuromuscular junctions
oligo(dT)	oligodeoxythymidylic acid
ori	origin of replication
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PE	phosphatidylethanolamine
PI	phosphatidylinositol
P-MLV	polytropic MLV
PMSF	phenylmethylsulphonyl fluoride
PoERV	porcine endogenous retrovirus
poly A	polyadenylation
PPT	polypurine tract
PR	retroviral protease
PRR	proline-rich region
Pro	proline
PS	phosphatidylserine
RER	rough endoplasmic reticulum

RIPA	radio-immune precipitation assay
RNA	ribonucleic acid
RNAse	ribonuclease
rpm	revolutions per minute
RPMI	Rosewell Park Memorial Institute
RSV	Rous sarcoma virus
RT	reverse transcriptase
RT-PCR	reverse transcription PCR
–sssDNA	minus-strand strong-stop DNA
+sssDNA	plus-strand strong-stop DNA
s	second
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
Ser	serine
SIV	simian immunodeficiency virus
SPF	specific pathogen free
SPIT	sodium selenite, sodium pyruvate, insulin, transferrin (Sigma, UK)
SSAV	simian sarcoma associated virus
Staph. A.	*Staphylococcus aureus*
SV40	simian virus 40
SU	surface glycoprotein of Env
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TBS	Tris buffered saline
TEMED	N, N, N', N'-tetramethylethylenediamine
Thr	threonine
TM	transmembrane protein of Env
Tris	tris(hydroxymethyl) aminomethane
TRITC	tetramethylrhodamine isothiocyanate
tRNA	transfer RNA
Trp	tryptophan
ts	temperature sensitive
UV	ultraviolet
Val	valine
VRA	variable region A
VRB	variable region B
VSV	vesicular stomatitis virus
X-gal	5-bromo-4-chloro-3-indoyl β-D-galactopyranoside
X-MLV	xenotropic MLV

CHAPTER 1

Chapter 1 General Introduction

1.1 Retroviruses

The retroviruses are a large family of enveloped RNA viruses, approximately 80-100nm in diameter. Their genome normally consists of two identical single-stranded RNA molecules 7-12 kb in length, enclosed within a protein core. The outer lipid envelope contains the viral glycoproteins that are involved in receptor recognition and cell infection. These viruses are associated with a variety of disorders including malignancies, anaemia, osteopetrosis, neurological disease and immunodeficiencies, as well as lifelong viraemia without obvious illness (Coffin, 1990; Rosenberg and Jolicoeur, 1997).

1.1.1 History

Retroviruses were first discovered in association with neoplastic diseases in chickens (Ellermann and Bang, 1908; Rous, 1911). These were originally thought to be restricted to birds and it was not until several years later that similar tumour causing viruses were recognised in mammals (reviewed by Levy, 1986; Vogt, P.K., 1997). In 1936, Bittner described a filterable 'active milk-influence' that could transmit mammary carcinoma in mice (Bittner, 1936; Bittner, 1942). This 'influence' was later shown to be a retrovirus, now known as mouse mammary tumour virus (MMTV). The first mammalian leukaemia virus, Gross murine leukaemia virus, was reported in 1951 (Gross, 1951). Following this discovery an enthusiastic search for virally induced tumours in mammals began. This resulted in the identification of a variety of viruses in mice, cats, cows, hamsters, rats and monkeys (reviewed by Levy, 1986; Vogt, P.K., 1997). Retroviruses have now been described in birds, mammals, amphibians, reptiles and fish (Herniou *et al.*, 1998). They may also be present in invertebrates such as molluscs (Oprandy and Chang, 1983) and insects (Finnegan, 1994; Hajek and Freisen, 1998; Kim *et al.*, 1994; Song *et al.*, 1994).

An early problem facing researchers in this field was the mechanism of retroviral replication. Other single-stranded RNA viruses had been shown to replicate via RNA-dependent RNA synthesis, creating double stranded replicative forms. However, no similar intermediate forms could be found in retrovirus-infected cells (reviewed by Vogt, P.K.,

1997). In 1964, Temin put forward the provirus hypothesis (Temin, 1964b), which suggested that a DNA copy of the viral genome is synthesised, and then integrated into the host cell genome. This hypothesis was based on the observations that infection could be prevented by inhibitors of DNA synthesis and that viral RNA appeared to have homology with DNA from infected but not uninfected cells (Bader, 1965; Duesberg and Vogt, 1969; Temin, 1964a). Integration of genetic information would also help to explain how retroviruses can stably transform cells. The proposal went against the established 'central dogma' of one-way transfer of information from DNA to RNA to protein and, therefore, it was generally disregarded. It was not until the revolutionary discovery of a viral RNA-dependent DNA polymerase (Baltimore, 1970; Temin and Mizutani, 1970) that proof was finally obtained, and Temin's hypothesis accepted. It is now known that this type of DNA polymerase, reverse transcriptase (RT), is not restricted to retroviruses, but is present in a variety of viral and nonviral genetic elements (reviewed by Vogt, P.K., 1997).

1.1.2 Classification

Retroviruses are broadly divided into two main groups depending on their genomic organisation. The simple viruses possess only the major coding regions: *gag*, encoding structural proteins; *pro*, the viral protease; *pol*, the viral enzymes including reverse transcriptase; and *env*, the envelope glycoproteins. Complex viruses also carry additional genes encoding regulatory proteins (reviewed by Vogt, P.K., 1997). Traditionally, retroviruses have subsequently been divided into three subgroups based on pathogenicity. These groups are oncovirinae, lentivirinae and spumavirinae. The oncovirinae consist of the viruses originally discovered in association with tumours, and other closely related viruses. They are present in all classes of vertebrates. The lentivirinae are also found in many species. They are associated with diseases that have a long incubation period, and often act by killing or causing loss of function in certain cell types. Relatively little is known about the spumavirinae, which have not, as yet, been associated with any disease (reviewed by Coffin, 1990; Levy, 1986).

The use of evolutionary relationships and morphology has resulted in a further division of these groups into seven genera, described in Table 1.1 (reviewed by Coffin, 1990; Vogt, P.K., 1997). These genera are based mainly on the sequence relatedness of the RT enzyme.

Subgroup	Genus	Genome	Morphology	Example
Oncovirinae	Avian sarcoma and leukosis viruses	simple	C-type particles	Rous sarcoma virus
	Mammalian B-type viruses	simple	B-type particles	mouse mammary tumour virus
	Mammalian C-type viruses	simple	C-type particles	Moloney murine leukaemia virus
	Human T-cell leukaemia–bovine leukaemia viruses	complex	central, spherical core	human T-cell leukaemia virus
	D-type viruses	simple	D-type particles	Mason-Pfizer monkey virus
Lentivirinae	Lentiviruses	complex	cone-shaped core	human immunodeficiency virus
Spumavirinae	Spumaviruses	complex	central, spherical core	simian foamy virus

Table 1.1. Classification of retroviruses.

In general, viruses in the same genus show identity at more than two thirds of the amino acid residues within the most conserved region of this enzyme. Little genomic sequence similarity is observed between genera, even amino acid identity is low except in the RT (reviewed by Vogt, V.M., 1997). A new nomenclature for retroviral classification has recently been proposed that encompasses novel viruses such as those found in fish (D.Onions, pers. com.). However, in order to avoid confusion, the system described in Table 1.1 will be used throughout this thesis.

Several different virion morphologies have been identified by thin section electron microscopy. A-type particles were first described as the immature, cytoplasmic precursors to B- and D-type particles. They now also include the endogenous intracisternal type-A particles (IAPs) found budding into and within the endoplasmic reticulum of several cell lines derived from rodents. Since these have budded through a membrane they are enveloped, unlike the cytoplasmic particles. A-type particles have an electron-lucent centre surrounded by one or two concentric electron-dense rings, giving a hollow, double-walled appearance. Newly budded B-type particles resemble A-type, but quickly mature into particles with an acentric round core and prominent envelope proteins. C-type particles include the majority of the avian and mammalian oncovirinae. Just after budding, these particles have a hollow core. This matures into a central, electron-dense, round or slightly angular structure. The lentiviruses resemble C-type particles while budding but the core then matures into a cone-shape. D-type particles have an acentric bar-shaped core. B- and D-type viruses and the spumaviruses differ from the majority of retroviruses in that they do not assemble at the plasma membrane. The cores of these viruses assemble within the cytoplasm, and are then transported to the plasma membrane for envelopment and release. The assembly and budding of most retroviruses appear to occur simultaneously (reviewed by Coffin, 1990; Vogt, V.M., 1997).

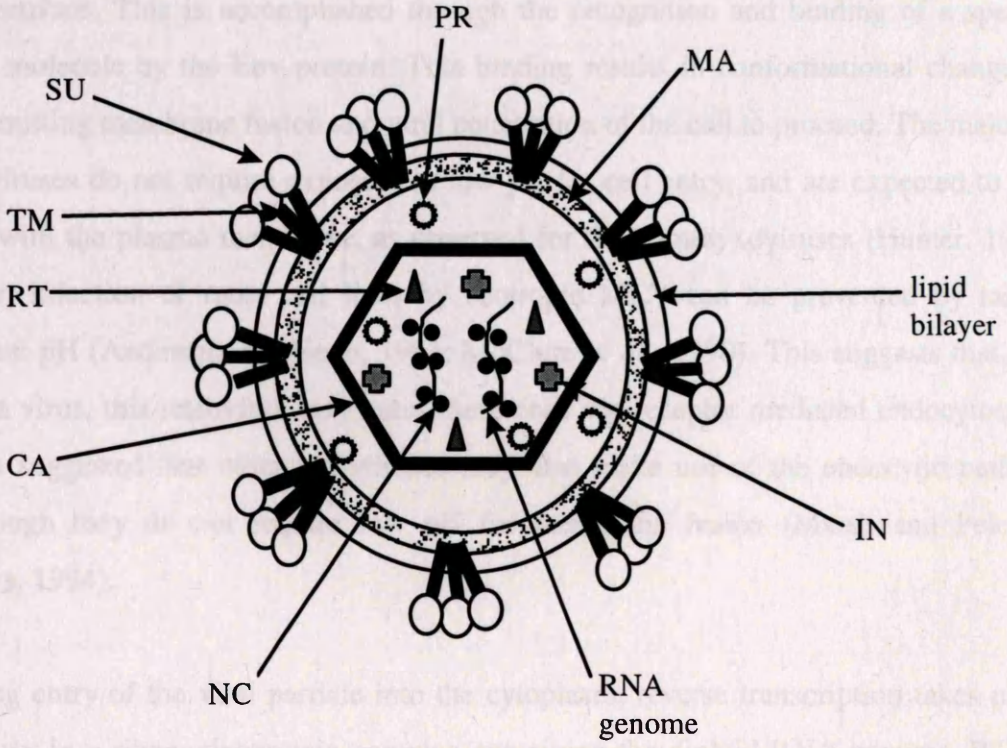
1.1.3 Viral structure

Retroviruses carry their genome as RNA but replicate via a DNA intermediate, the provirus. The viral genes encoded by the provirus, *gag, pro, pol* and *env*, are flanked by the long terminal repeats (LTRs). These are identical sequences formed during reverse transcription and composed of three regions, U3, R and U5. U3 is obtained from a unique

sequence found at the 3' end of the genomic RNA, U5 from a unique sequence found at the 5' end, and R from a sequence repeated at both ends. The lengths of these three regions vary significantly between different retroviruses. The majority of the transcriptional control elements of the provirus are found within U3. These include the promoter and several enhancer sequences. The particular enhancer sequences present are important determinants of the tissue specificity of viral replication (reviewed by Vogt, V.M., 1997).

The structural proteins of retroviruses are encoded by the *gag* gene. These are the matrix (MA), capsid (CA) and nucleocapsid (NC) proteins. The MA protein is believed to line the lipid membrane. MA from avian sarcoma leukosis virus (ASLV) and murine leukaemia virus (MLV) can be chemically crosslinked to the phospholipid phosphatidylethanolamine (Pepinsky and Vogt, 1979). The protein is myristylated in the majority of retroviruses. This involves the cotranslational addition of a 14-carbon fatty acid, and is a modification common to many cellular proteins associated with membranes. MA is thought to be involved in the budding process of the virus, and may also direct incorporation of the envelope proteins. CA is the largest of the Gag proteins. It is thought to form a shell between the MA and the nucleoprotein core, although its function is unknown at present. The NC protein is a small basic protein bound to the genomic RNA within the core of the virion. It promotes binding of the primer tRNA and stabilisation of the dimeric genomic RNA. The core of the virion also contains the viral enzymes RT and integrase (IN), encoded by the *pol* gene. These enzymes control the reverse transcription and integration processes necessary for viral replication. The position of the protease (PR) within the virion is not yet known. It is required for the cleavage of the Gag and Pol proteins after budding of the virion from the cell (reviewed by Vogt, V.M., 1997).

The viral lipid membrane contains the envelope spike glycoproteins encoded by *env*. The spikes contain two proteins held together by noncovalent interactions, and occasionally by disulphide bonds. These are the surface glycoprotein (SU) and the transmembrane protein (TM). The Env proteins are a major determinant of the host range of the virus as they recognise and interact with a specific cellular receptor. SU contains the receptor binding domains while TM contains the fusion peptide and is required for the process of membrane fusion (reviewed by Hunter, 1997; Vogt, V.M., 1997).



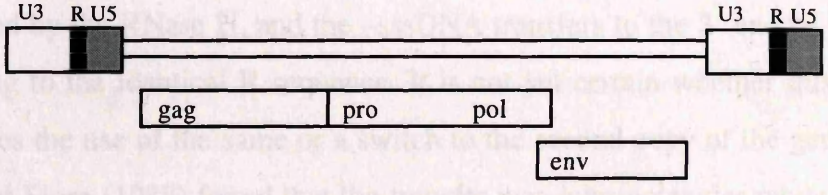


Fig. 1.1 Schematic representation of a cross-section through a retroviral virion, and a mammalian C-type provirus. SU, surface envelope glycoprotein; TM, transmembrane envelope protein; MA, matrix protein; CA, capsid protein; NC, nucleocapsid protein; PR, viral protease; RT, reverse transcriptase; IN, integrase.

1.1.4 Life cycle

The retroviral life cycle is illustrated in Fig. 1.2. The first step in this cycle is attachment to the cell surface. This is accomplished through the recognition and binding of a specific receptor molecule by the Env protein. This binding results in conformational changes in Env, permitting membrane fusion and viral penetration of the cell to proceed. The majority of retroviruses do not require exposure to low pH for cell entry, and are expected to fuse directly with the plasma membrane, as observed for the paramyxoviruses (Hunter, 1997). However, infection of most cell lines by ecotropic MLV can be prevented by raising endosomal pH (Andersen and Nexo, 1983; McClure *et al.*, 1990). This suggests that, like influenza virus, this retrovirus may enter these cells via receptor mediated endocytosis. It has been suggested that other retroviruses may also make use of the endocytic pathway even though they do not require low pH for membrane fusion (Marsh and Pelchen-Matthews, 1994).

Following entry of the viral particle into the cytoplasm, reverse transcription takes place. This occurs in a ribonucleoprotein complex containing the diploid RNA genome, RT and other viral proteins. RT possesses two enzymatic activities, DNA polymerase and ribonuclease H (RNase H), both of which are required for the reverse transcription process. The current model for this process is illustrated in Fig. 1.2. Initially, the virus-specific primer tRNA binds to the primer binding site and minus-strand DNA synthesis progresses until the 5' end of the RNA genome is reached. This DNA intermediate is known as the minus-strand strong stop DNA (–sssDNA). The RNA strand of the RNA:–sssDNA duplex is then degraded by the RNase H, and the –sssDNA transfers to the 3' end of the genomic RNA, annealing to the identical R sequence. It is not yet certain whether this first strand transfer involves the use of the same or a switch to the second copy of the genomic RNA. Panganiban and Fiore (1988) found that the transfer was intermolecular while Jones *et al.* (1994) found that it was intramolecular. A combination of both forms of transfer has also been observed (Hu and Temin, 1990). Minus-strand DNA synthesis then continues, accompanied by RNase H digestion of the RNA template. Within this template there is a short polypurine tract (PPT) which is resistant to RNase H digestion. This sequence acts as a primer for plus-strand DNA synthesis. This synthesis proceeds until a portion of the primer tRNA is reverse-transcribed, producing another DNA intermediate termed the plus-

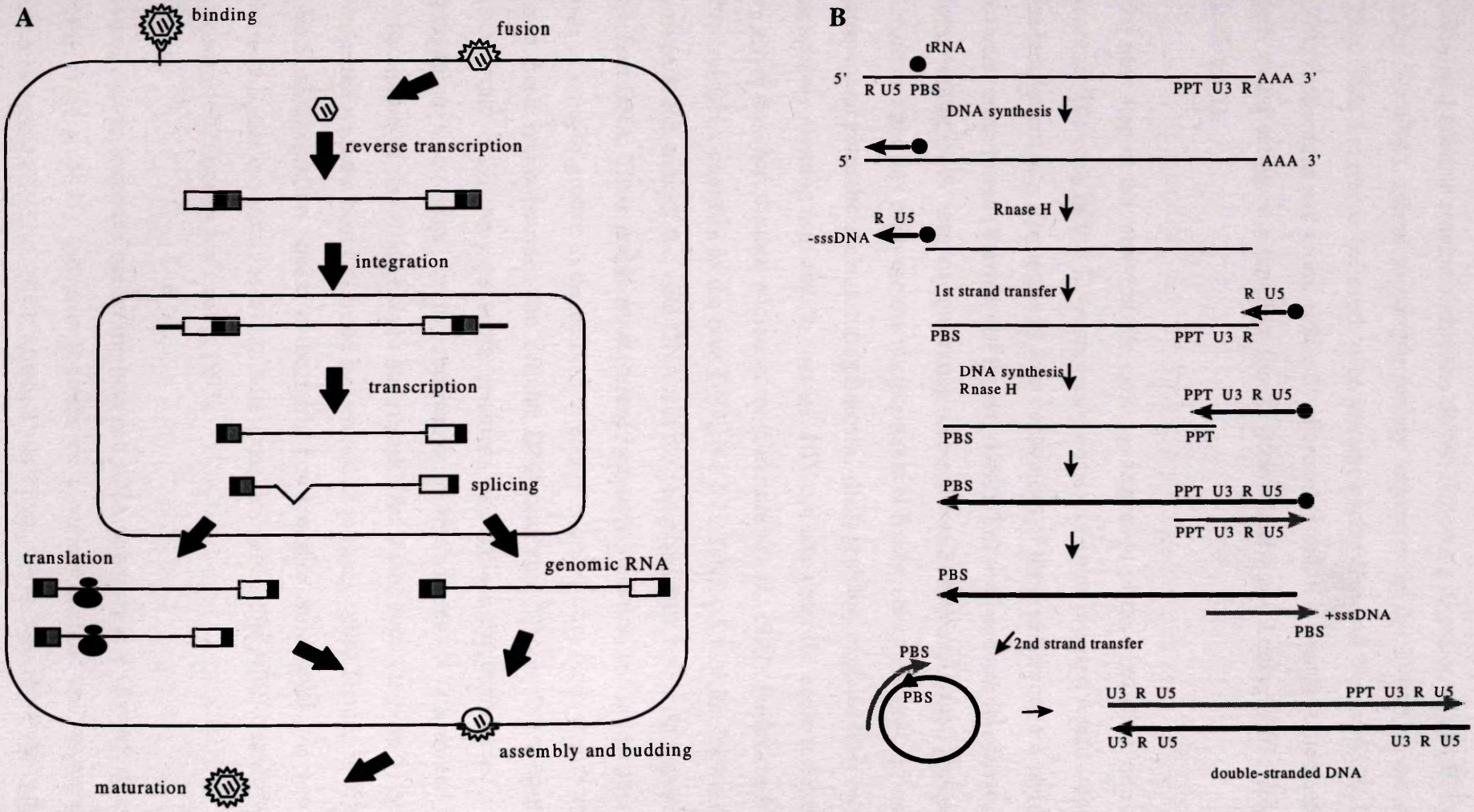


Fig. 1.2. A) Retroviral life cycle (modified from Vogt, P.K., 1997). B) Reverse transcription (modified from Telesnitsky and Goff, 1997). Thin black line, RNA; thick black line, minus-strand DNA; thick grey line, plus-strand DNA; PBS, primer binding site; PPT, polypurine tract; -sssDNA, minus-strand strong stop DNA; +sssDNA, plus-strand strong stop DNA.

strand strong-stop DNA (+sssDNA). The termination point for the +sssDNA is a methylated adenine residue within the tRNA. Following digestion of the tRNA, sequences in the +sssDNA anneal to complementary sequences in the 3' end of the minus-strand DNA. This transfer is believed to be intramolecular (Hu and Temin, 1990; Jones *et al.*, 1994; Panganiban and Fiore, 1988). Synthesis of both DNA strands is then completed with each strand acting as a template for the other (Luciw and Leung, 1993; Telesnitsky and Goff, 1997).

The next step in the retroviral life cycle is integration of the viral DNA into the host cell genome. The ends of this DNA molecule consist of short inverted repeats that function as the attachment site. The terminal four nucleotides of these sequences are highly conserved between retroviruses (Varmus and Brown, 1989). The viral enzyme IN cleaves 2 or 3 bases from each of the 3' termini, generating recessed free 3' hydroxyl (-OH) groups. Either just before or after this cleavage the nucleoprotein complex enters the host cell nucleus. Most retroviruses enter the nucleus during mitosis, which probably explains their requirement for an actively dividing host cell. In contrast, HIV can also enter the nucleus during interphase by active transport through nuclear pores (Bukrinsky *et al.*, 1992). Following binding of the IN:viral DNA complex to the host DNA, the IN catalyses a reaction between the 3' -OH groups at the ends of the viral DNA and the phosphodiester bonds on opposite strands of the host DNA. These bonds are separated by four to six bases, and are positioned on either side of a major groove in the helix. The energy released from the broken bonds is used to form fresh bonds between the cellular DNA and viral 3'ends. DNA synthesis is then carried out to fill the gaps at the junctions and replace the mismatched viral 5' ends. Integration is essential for the replication of retroviruses. It is required for efficient transcription of the viral genome and of mRNAs. Also, once the virus is integrated it is replicated with the host DNA and transmitted to progeny cells (Temin, 1976). The sites of retroviral integration appear to be fairly non-specific. As a result the process is often mutagenic due to insertions within host genes or in areas that affect transcriptional control (Brown, 1997; Luciw and Leung, 1993).

Transcription is carried out by the host cell RNA polymerase II (Sawadogo and Sentenac, 1990; Young, 1991) and also depends on a variety of host cell transcription factors (Johnson and McKnight, 1989; Mitchell and Tjian, 1989). These include both constitutive

and tissue-specific factors, and are specific to the particular retrovirus. While the simple retroviruses rely entirely on these cellular factors (Majors, 1990), the complex retroviruses also encode transactivators involved in the control of transcription and RNA processing (Green and Chen, 1990; Haseltine, 1991). The particular transcription factors employed by a virus can affect tissue tropism and may also have a role in pathogenesis (Luciw and Leung, 1993; Rabson and Graves, 1997). The primary transcripts from the provirus resemble cellular mRNA. They are capped at the 5' end, polyadenylated at the 3' end, and methylated at several internal sites. These transcripts can be divided into two groups, one unspliced and one spliced. The unspliced group consists of genomic RNA and mRNA for *gag, pro* and *pol*. The spliced group forms the mRNA for *env*. The splicing process removes the *gag, pro* and *pol* genes as an intron, and joins the 5' end of the RNA to the *env* gene. The ratio of these two groups is controlled by *cis*-acting sequences in simple retroviruses, and transactivators in complex retroviruses (reviewed by Luciw and Leung, 1993; Vogt, V.M., 1997).

Translation of the Gag, Pro and Pol proteins takes place on free polyribosomes. This translation produces both Gag and Gag-Pro-Pol polyproteins. Approximately 10 to 20 times more Gag is made than Gag-Pro-Pol. The full length polyprotein is produced when the *gag* termination codon is bypassed. There are two mechanisms by which this is achieved. The mammalian C-type retroviruses employ readthrough suppression, which allows the amber termination codon to be occasionally misread as a sense codon. For example, the *pro* coding region of Moloney MLV (MoMLV) and feline leukaemia virus (FeLV) starts four codons upstream of the *gag* termination codon. This amber UAG termination codon is occasionally read as the glutamine codon CAG, and translation is continued (Yoshinaka *et al.*, 1985a; Yoshinaka *et al.*, 1985b). This suppression mechanism is regulated by viral *cis*-acting sequences. Most other retroviruses use ribosomal frameshifting. This occurs when the downstream genes are in a different reading frame to *gag,* and requires a seven nucleotide sequence, known as the slippery or shifty sequence, within the region of overlap (Jacks *et al.*, 1988a; Jacks *et al.*, 1988b). When two tRNAs are bound to this sequence, they are able to slip back one nucleotide, thereby leaving the *gag* reading frame to follow that of *pro-pol* (Swanstrom and Wills, 1997).

As mentioned above, the Env polyprotein is translated from a spliced mRNA. It is initially translated on free ribosomes until the hydrophobic leader sequence is made. This sequence enables binding to a signal recognition particle which targets the protein to the membrane of the rough endoplasmic reticulum (RER). Here the polyprotein is further translated, passing through the membrane into the lumen of the RER. It remains anchored in the membrane via a hydrophobic region close to the carboxyl terminal (reviewed by Vogt, V.M., 1997). During translation the leader sequence is removed by a cellular protease and the Env protein is glycosylated. Folding and oligomerisation of Env also occur within the RER and are required for transport to the cell surface. Cleavage of the polyprotein into the SU and TM proteins takes place within the Golgi apparatus and is probably carried out by a cellular protease, such as furin (reviewed by Swanstrom and Wills, 1997).

Once all the viral proteins are synthesised, virion assembly can proceed. The budding process, which relies on the Gag protein, is not well understood. The cores of B- and D-type retroviruses assemble in the cytoplasm and then bud through the plasma membrane which contains the Env proteins. The core formation and budding of other retroviruses occur simultaneously. It is thought that high affinity lipid-particle interactions allow coating and then separation of a growing spherical structure built from the interacting Gag proteins. Intermediate budding particles have also been observed on narrow stalks, which could suggest an alternative mechanism. The Gag-Pro-Pol is incorporated as a polyprotein, and is probably targeted to the assembly site by the Gag portion (reviewed by Luciw and Leung, 1993; Swanstrom and Wills, 1997). The mechanism of Env protein incorporation is not yet understood. It is generally thought to involve an interaction with a Gag protein, probably MA, although this is not required for virion budding (Dickson *et al.*, 1982; Kawai and Hanafusa, 1973). This is discussed in greater detail in section 1.5.1. Selective packaging of viral genomic RNA is controlled both by signals present in the viral proteins, particularly NC (Berkowitz *et al.*, 1995), and a signal in the RNA (ψ). The ψ signal of MLV is approximately 300 bases long and situated between the splice donor site and the start codon for *gag* (Mann *et al.*, 1983). The subgenomic *env* mRNA is not packaged since this signal is removed during splicing. Following budding, maturation of the viral particle takes place. Gag and Gag-Pro-Pol are cleaved by the viral protease, PR, into their respective proteins. This results in an observable change in the structure and morphology of the virion. This mature structure is less stable and is required for infectivity. The Env

protein of at least some retroviruses is also cleaved by PR after budding. This involves removal of a carboxy-terminal peptide, known as the R peptide. This cleavage renders the Env protein fusion active (Luciw and Leung, 1993; Swanstrom and Wills, 1997).

1.1.5 Genetic diversity and recombination of retroviruses

As a consequence of the nature of their replication cycle, retroviruses are able to evolve rapidly, allowing adaptation to changing environments. Viruses, such as HIV-1, that have high rates of replication tend to exist as polymorphic populations within the host (Coffin, 1995). Neither the host RNA polymerase II nor the viral RT have a proof-reading activity, and are therefore both likely to contribute to the introduction of errors into the genome. The error rate of RT is estimated to be approximately 10^{-4} errors per base *in vitro* (Telesnitsky and Goff, 1997). This is equivalent to one error per genome in each replication cycle. Depending on the system used, error rates *in vivo* have been estimated as 5×10^{-3}, 2×10^{-5} and 1.4×10^{-4} per base (Dougherty and Temin, 1986; Dougherty and Temin, 1988; Leider *et al.*, 1988; Monk *et al.*, 1992).

During the process of reverse transcription a single DNA provirus is made from a virion containing a diploid genome (Hu and Temin, 1990; Panganiban and Fiore, 1988). Although the RNA molecules are usually identical, heterozygous virions can be released from a cell infected by two different viruses. Reverse transcription of heterozygous genomes can result in the production of a recombinant provirus containing portions of both sequences. Hu and Temin (1990) estimated that homologous recombination between two markers 1kb apart occurs in about 4% of the DNAs produced in one cycle of reverse transcription by such virions. The majority of recombinations are thought to occur through intermolecular strand transfer during minus-strand DNA synthesis, although it may also occur during plus-strand synthesis (Telesnitsky and Goff, 1997). Recombination between genomes derived from different retroviruses constitutes a form of sexual replication, and therefore contributes to the genetic diversity and adaptation of viral populations (Temin, 1993).

1.1.6 Endogenous retroviruses

As mentioned above (section 1.1.4), an important step in the life cycle of infectious exogenous viruses is the permanent integration of the provirus into the host cell genome. As a consequence of this process, insertion of proviruses into the host germ line has occurred on several occasions throughout evolution. These proviruses are then inherited as a mendelian trait, and are known as endogenous retroviruses (ERVs). During the evolutionary process these viruses have adapted to this form of transmission. The selective pressure to produce infectious viral particles is removed resulting in the accumulation of mutations in the viral genes. Such defective proviruses are less pathogenic to the host and therefore more likely to be maintained in the genome. Many are also transcriptionally silent as a result of host methylation of the viral sequences. ERV sequences have been found in all vertebrate species so far studied. They are often found in groups of highly homologous sequences that differ in their site of integration in the genome (Boeke and Stoye, 1997; Patience *et al.*, 1997b). It is believed that each group of closely related ERVs resulted from an initial cross-species infection by an exogenous retrovirus. The different proviruses in a group were produced by reinfection, or perhaps by intracellular retrotransposition, and are separated by only a few rounds of replication (King *et al.*, 1987; Stoye and Coffin, 1987). Generation of new proviruses ceases when the accumulation of mutations renders them noninfectious, or when the host develops mechanisms to prevent further infection, such as the selection of a receptor allele not recognised by the virus (Boeke and Stoye, 1997).

Occasionally the retention of a functional ERV gene can confer a selective advantage. For example, the human ERV-3 virus possesses a functional *env* gene. This is expressed in the syncytiotrophoblast within the placenta where it may have a role in cell fusion or immunosuppression (Boyd *et al.*, 1993; Harris, 1991; Venables *et al.*, 1995). It has even been suggested that the acquisition of such an ERV enabled evolution of placental mammals to proceed (Harris, 1991; Villareal, 1997). Infection of early embryos may have resulted in cell fusion and the formation of an invasive syncytial plate that expressed immune-suppressing viral Env protein. ERV expression may also provide the host with resistance to infection by related exogenous viruses (Boeke and Stoye, 1997). For example, endogenous Env proteins can block common receptors, causing receptor interference (described in section 1.2.1). The murine *Fv-1* gene provides resistance to certain strains of

MLV. It is believed to have evolved from an endogenous *gag* gene (Best *et al.*, 1996), and probably acts by preventing proviral integration (Brown, 1997).

The presence of ERVs can alter the structure and functioning of the host genome. Every proviral integration has the potential to result in insertional mutagenesis or changes in gene expression. As a consequence, the germ line insertions of ERVs can have a significant effect on the evolution of the host genome. An interesting example of this is the insertion of a C-type ERV upstream of a duplicated pancreatic amylase gene of an ancestor of humans, Old World monkeys and apes (Meisler and Ting, 1993; Samuelson *et al.*, 1990; Ting *et al.*, 1992). This insertion appears to have resulted in expression of amylase in saliva, and may therefore reflect the importance of starch in the human diet. The presence of these retroviral sequences can also influence evolution of infecting exogenous retroviruses, since their expression provides material for recombination. This may also have consequences for the host since the resulting recombinant viruses often show increased pathogenicity (Ott *et al.*, 1990; Stewart *et al.*, 1986).

ERVs have entered genomes at various times during evolution, and are usually described as ancient or modern (Boeke and Stoye, 1997). Ancient viruses were acquired prior to speciation and are found in the same location in the genomes of all individuals of a species, and often of related species. A few species possess modern viruses that were acquired relatively recently in evolution and may still retain the ability to release infectious virions. The production of these virions can occur spontaneously or be induced with chemicals such as 5-aza-cytidine or 5-bromo-2-deoxyuridine that reverse DNA methylation (Patience *et al.*, 1997b). These modern ERVs are closely related to exogenous retroviruses, and often vary in their insertion site in individual hosts. Many of these viruses are unable to infect cells of the host organism. This is thought to be the result of selection of mutations in the cellular receptor which protect the host from reinfection and the potentially pathogenic effects of ERV viraemia. The receptor for xenotropic MLV (X-MLV) is expressed by several strains of wild mice but is absent from inbred mice, suggesting that a non-functional receptor has been selected (Kozak, 1985). Other organisms will not have evolved such resistance mechanisms and therefore can be susceptible to infection (Boeke and Stoye, 1997). This is illustrated by the infection of human tumour xenografts with murine and feline ERVs (Tralka *et al.*, 1983; McAllister *et al.*, 1972). It seems possible that

transfer of ERVs from the xenograft to the host could also occur. This may prove to be a problem if animal organs and tissues are used in the treatment of human patients.

It has been proposed that retroviruses have evolved from transposable elements, mobile DNA sequences that replicate via an RNA intermediate (Doolittle *et al.*, 1989; Doolittle *et al.*, 1990; Doolittle and Feng, 1992; Temin, 1980; Temin, 1985). If this is the case then it seems likely that at least some defective ERV sequences are not products of ancient infections, but instead represent stages in the evolution of retroviruses. This evolutionary pathway could have begun with the juxtaposition of a reverse transcriptase gene with a simple transposable element. This element could then become gradually more complex through the acquisition of further coding sequences that would have evolved into *gag*-like genes. Finally, the incorporation of an *env* gene would allow infection of other cells, and perhaps other species. However, it is also possible that transposable elements are the remnants of ancient viral infections (Finnegan, 1983). As there have been many cases of exogenous retroviruses entering genomes and endogenous retroviruses 'escaping' to infect alternative hosts throughout evolution it is very difficult to discern which was the original mode of transmission.

1.2 Host range and receptors

The host range of a virus is largely dependent upon the interaction between a specific cell surface receptor molecule and the envelope proteins. For many viruses, these receptors are not only necessary for viral attachment, but are also involved in the process of infection. A large variety of molecules appear to be able to act as receptors, including proteins, carbohydrates and lipids. Some are also involved in recognition and adhesion processes between cells. While the interaction with a virus often has little effect on the normal function of the receptor, this is not always the case. For example, the viral protein may block the normal interactions of the receptor or mimic the activity of a ligand. The distribution of receptors determines the range of cells susceptible to infection and is therefore important in pathogenesis. Consequently, the identification of viral receptors is a necessary step towards understanding virus-host relationships. Knowledge of receptors and their tissue distribution is also useful when designing and choosing viral vectors for gene transfer (Weiss, 1993; Haywood, 1994).

1.2.1 Retroviral receptors

Most animal viruses within the same genus show little, if any, variation in receptor usage. However, closely related retroviruses have been found to use different receptors. The viruses in the avian and mammalian C-type genera both show a high degree of polymorphism both in receptor usage and receptor expression within the host (Hunter, 1997). These viruses have been placed into subgroups on the basis of host range, neutralisation and receptor interference (Weiss, 1993). The murine C-type viruses have been classified into four subgroups. Ecotropic viruses are restricted to murine and rat cells, while xenotropic viruses will only infect cells from other species. Amphotropic and polytropic viruses infect both cell groups but do not interfere with one another indicating that they use different receptors. The FeLVs have been divided into three subgroups. FeLV-A infects only feline cells while FeLV-B and -C both have wider host ranges (Jarett *et al.*, 1973; Sarma and Log, 1973; Sarma *et al.*, 1975).

Receptor interference occurs when the receptors of an infected cell are blocked or downregulated, thereby preventing superinfection by another virus recognising the same

receptor. The actual mechanism involved is not yet fully understood, but it probably involves a direct interaction between the Env protein and the receptor. This interaction could occur on the cell surface or within the RER (Hunter, 1997; Weiss, 1993). Infection interference can also occur as a result of the expression of endogenous *env* sequences. The murine *Fv-4* gene which confers resistance to Friend-MLV has been found to express an ecotropic Env (Ikeda and Sugimura, 1989). This form of infection interference is likely to be advantageous to the host, providing protection from infection by exogenous virus and also prevention of ERV viraemia. As a consequence, such genes would be expected to be preserved within the genome. Several defective ERV proviruses retain a functional *env* gene (Hunter, 1997). Receptor interference is usually reciprocal, however there are a few notable exceptions. For example, the murine 10A1 virus blocks infection by the related amphotropic MLV, but is able to infect cells releasing this virus (Rein and Schultz, 1984). The MLVs also have unexpected interference patterns when tested on *Mus dunni* cells. Certain xenotropic strains are blocked by preinfection with ecotropic and amphotropic viruses (Chesebro and Wehrly, 1985). In addition, the viruses placed within the same interference group may display small differences in host range. Takeuchi *et al.* (1992b) have shown that simian sarcoma associated virus (SSAV) will infect rat cells, which are resistant to FeLV-B.

Analysis of interference between retroviruses infecting human cells has revealed the existence of eight receptor groups (Sommerfelt and Weiss, 1990). These are listed in Table 1.2. To date the receptors corresponding to three of these interference groups have been identified. These are Ram-1 (GLVR-2 or Pit-2) for group 2: amphotropic MLV (A-MLV, Miller *et al.*, 1994; van Zeijl *et al.*, 1994); GLVR-1 (Pit-1) for group 5: gibbon ape leukaemia virus (GALV, O'Hara *et al.*, 1990), the closely related SSAV (Takeuchi *et al.*, 1992b), and FeLV-B (Takeuchi *et al.*, 1992b); a 70kD membrane glycoprotein for group 6: bovine leukaemia virus (BLV, Ban *et al.*, 1993); and CD4 for group 8: the human and simian immunodeficiency viruses (HIV and SIV, Dalgleish *et al.*, 1984; Hoxie *et al.*, 1988; Klatzman *et al.*, 1984; Sattentau *et al.*, 1988). The ecotropic MLV (E-MLV) receptor mCAT-1 (Albritton *et al.*, 1989), and the avian sarcoma and leukosis virus (ASLV) -A receptor, a protein related to the low density lipoprotein (LDL) receptor (Young *et al.*, 1993; Young *et al.*, 1994), have also been cloned. The known receptors of the mammalian C-type viruses are described below.

Receptor group	Virus	Receptor
1	RD114	unknown
	BaEV	
	SRV-1	
	SRV-2	
	MPMV	
	SRV-4	
	SRV-5	
	PO-1-Lu	
	SMRV	
2	A-MLV	Ram-1 (GLVR-2/Pit-2)
3	X-MLV	unknown
4	FeLV-C	unknown
5	FeLV-B	GLVR-1 (Pit-1)
	SSAV	
	GALV	
6	BLV	70kD membrane glycoprotein
7	HTLV-1	unknown
	HTLV-2	
	ChTLV	
	STLV	
8	HIV-1	CD4
	HIV-2	
	SIV_{mac}	
	SIV_{smm}	
	SIV_{agm}	

Table 1.2. Receptor interference groups of retroviruses infecting human cells (based on Sommerfelt and Weiss, 1990)

The E-MLV receptor was identified by transfecting resistant human cells with murine DNA (Albritton *et al.*, 1989). These cells were then tested for their susceptibility to infection with a recombinant E-MLV virus carrying a neomycin resistance gene. Cells expressing the murine receptor were infected by this virus and could then be selected by their resistance to neomycin. The DNA clone obtained by this process was found to encode a 622 amino acid membrane protein that contains at least 14 transmembrane domains connected by short hydrophilic loops. Expression of this protein, now termed mCAT-1, in *Xenopus* oocytes produces an increase in sodium-independent uptake of L-arginine, L-lysine and L-ornithine, indicating that it is a cationic amino acid transporter (Kim *et al.*, 1991; Wang *et al.*, 1991). Chronic infection of mink fibroblasts with E-MLV has been shown to result in only a 50-70% reduction in transporter activity (Wang *et al.*, 1992), suggesting that bound mCAT-1 remains capable of transporting amino acids. Substitutions of amino acid sequences with those from the nonpermissive human homologue have identified a region of 29 residues within the third extracellular domain of the receptor that is necessary for the interaction with E-MLV (Albritton *et al.*, 1993). This is illustrated in Fig. 1.3. Sequence variations within this region can affect strains of E-MLV differently. For example, a single amino acid change in the *Mus dunni* receptor prevents infection by MoMLV, but has no effect on the Rauscher or Friend E-MLV strains (Eiden *et al.*, 1993).

A similar method to that of Albritton *et al.* (1989) was used to identify and clone the receptor for GALV (O'Hara *et al.*, 1990). This receptor contains 679 amino acids and a series of 10 potential transmembrane domains connected by hydrophilic loops. The cross interference between GALV, SSAV and FeLV-B implied that these viruses all use GLVR-1 as a receptor (Sommerfelt and Weiss, 1990; Takeuchi *et al.*, 1992b). Takeuchi *et al.* (1992b) have shown that expression of GLVR-1 in resistant murine cell lines confers susceptibility to infection by SSAV and FeLV-B, indicating that these viruses do all recognise this receptor. The sequence between residues 550 and 558 of GLVR-1, illustrated in Fig. 1.3, has been shown to be necessary for infection by each of these viruses (Johann *et al.*, 1993; Tailor *et al.*, 1993). This region is highly divergent between species, and the different viruses exhibit different sequence requirements (Tailor *et al.*, 1993). These differences are thought to reflect the small variations in host cell range.

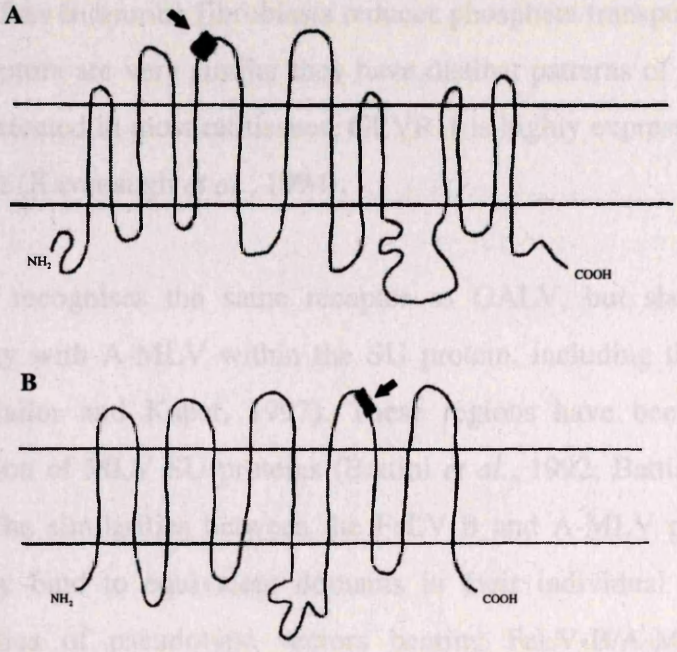


Fig. 1.3. Schematic representations of the structures of mammalian C-type retrovirus receptors (modified from Weiss, 1993). A) mCAT-1. B) GLVR-1. The regions thought to be required for viral infection are indicated by the arrows.

In an attempt to isolate further retroviral receptors, van Zeijl *et al.* (1994) screened a human cDNA library for clones related to GLVR-1. They found that expression of one such clone, GLVR-2, allowed infection of nonpermissive Chinese hamster fibroblasts with A-MLV. This 652 amino acid protein shares 62% overall homology with GLVR-1. The rat homologue of this receptor has been identified by Miller *et al.* (1994) and designated Ram-1. Certain hybrid molecules between the two related receptors can function for both GALV and A-MLV (Miller and Miller, 1994). These hybrids contain the amino-terminal of GLVR-2/Ram-1 and the carboxy-terminal of GLVR-1, including the critical region 550-558. The 10A1 virus, a recombinant between 1504A A-MLV and an endogenous murine retrovirus, can use both receptors (Miller and Miller, 1994). The expanded host range of 10A1 can be obtained with only a small number of changes in the amino acid sequence of A-MLV Env, presumably a reflection of the homology between these receptors.

The GALV and A-MLV receptors both function as sodium-dependent phosphate symporters (Kavanaugh *et al.*, 1994; Olah *et al.*, 1994). Kavanaugh *et al.* found that expression of the receptors in *Xenopus* oocytes or rat fibroblasts resulted in an uptake of phosphate in the presence of sodium ions. They were also able to show that expression of

A-MLV Env in murine fibroblasts reduces phosphate transport by 50-60%. Although these two receptors are very similar they have distinct patterns of tissue expression. While both can be detected in most rat tissues, GLVR-1 is highly expressed in bone marrow and Ram-1 in heart (Kavanaugh *et al.*, 1994).

FeLV-B recognises the same receptor as GALV, but shows a much higher level of homology with A-MLV within the SU protein, including the variable regions VRA and VRB (Tailor and Kabat, 1997). These regions have been implicated in the receptor recognition of MLV SU proteins (Battini *et al.*, 1992; Battini *et al.*, 1998; Morgan *et al.*, 1993). The similarities between the FeLV-B and A-MLV proteins therefore suggest that they may bind to equivalent domains in their individual receptors. By comparing the infectivities of pseudotype vectors bearing FeLV-B/A-MLV SU chimaeras on cells expressing GLVR-1/Ram-1 chimaeras, Tailor and Kabat were able to identify such domains. The VRA of each SU protein was found to bind to the extracellular loops 4 and 5 of the respective receptors, and VRB to loop 2. The critical region defined by Johann *et al.* (1993) between residues 550 and 558 of GLVR-1 is situated on loop 4. It therefore seems likely that this loop is also involved in the binding of GALV and SSAV.

The similarities between the receptors so far described for mammalian C-type retroviruses suggest that only a limited number of cell surface proteins are suitable for this function. This suitability could be due to tissue distribution or an ability to permit virus entry into the cell. It is also possible that the interaction between the virus and its receptor confers a growth advantage to the cell, thereby enhancing viral spread (Takeuchi *et al.*, 1992b). These restrictions may be peculiar to the viral genus since the CD4 receptor recognised by HIV and SIV is unrelated to the transporter molecules described here. This protein is a member of the immunoglobulin (Ig) superfamily, containing a single transmembrane domain (Maddon *et al.*, 1985). Although the receptors of the C-type viruses are related they vary in their distribution. This is likely to explain why relatively minor changes in Env proteins can result in significant differences in host range. It seems quite possible that other mammalian C-type retroviruses will also be found to recognise similar receptors (Cunningham and Kim, 1994). It is tempting to speculate that the receptors for FeLV-A and -C are related to the GLVR-1 recognised by FeLV-B. The FeLV-B Env appears to have arisen as a result of recombination of FeLV-A with endogenous *env* sequences (Stewart *et*

al., 1986), and the FeLV-C from a mutation of FeLV-A (Neil *et al.*, 1991). Although the host range of FeLV-C is much wider than that of FeLV-A, the receptor binding regions of the two Env proteins differ only slightly. The sensitivity of mammalian C-type viruses to minor differences in their receptors can at least partly explain the unexpected patterns of receptor interference observed among MLVs (Chesebro and Wehrly, 1985; Rein and Schultz, 1984). Although two viruses may recognise the same receptor in certain cell types, species specific changes in the receptor sequence could differentially affect viral recognition.

1.2.2 Rhabdoviral receptors

Other viruses, such as rhabdoviruses, have less restricted host ranges than the retroviruses. The envelope protein of vesicular stomatitis virus (VSV) is thought to interact with a phospholipid constituent of the cell membrane. Schlegel *et al.* (1983) tested solubilised cell membrane components for their ability to inhibit VSV binding. They found that the phospholipid phosphatidylserine (PS) inhibited the specific binding of the virus to Vero monkey cells. VSV binding to chicken embryo-related (CER) cells is also inhibited by PS and, to a lesser extent, sphingomyelin (Superti *et al.*, 1984b). Partial inhibition is seen with phosphatidylethanolamine (PE) and phosphatidylinositol (PI). The apparent ability to enter cells without relying on the presence of a specific protein receptor allows VSV to infect a very broad host cell range, including fish and insect cells.

The receptor usage of rabies virus, another rhabdovirus, has also not yet been fully defined. This virus has a very wide host range *in vitro,* but has a strong tropism for the nervous system *in vivo*. It has been found to concentrate at neuromuscular junctions (NMJ), areas rich in nicotinic acetylcholine receptors, nAchRs (Lentz *et al.*, 1982). Detection of bound rabies virus using immunofluorescence staining revealed that the pattern of binding at NMJ and on cultured chick myotubes is very similar to that of α-bungarotoxin (αBtx), an antagonist of the AchR. The presence of antagonists such as αBtx and *d*-Tubocurarine inhibits infection of myotube cells, presumably by competing for binding to AchR. Rabies G is also capable of blocking the binding of αBtx (Bracci *et al.*, 1988). In addition, a region of homology has been found between the nAchR binding site of curaremimetic neurotoxins and a segment of the G protein. Further supporting evidence for this receptor comes from

the observation that a monoclonal anti-idiotype-antibody directed against a neutralising anti-G antibody will also bind to purified nAchR, NMJ and some cells of the central nervous system (CNS, Hanham *et al.*, 1993). It seems likely that the antigen binding site of the anti-idiotype-antibody mimics the neutralising epitope of rabies G, and that this epitope is involved in the binding of nAchR. This antibody itself elicited a response in mice which, although low in titre, was protective against virus challenge.

Although these studies all indicate that nAchR is a receptor for rabies virus a wide range of cell lines are susceptible to infection *in vitro*, many of which do not express this receptor (Reagan and Wunner, 1985; Seganti *et al.*, 1990; Thoulouze *et al.*, 1998; Wunner *et al.*, 1984). While rabies G is able to fully block binding of αBtx, the reverse is not observed (Bracci *et al.*, 1988). Also, the neutralising antibody used by Hanham *et al.* to raise the anti-idiotype did not prevent rabies infection of culture cells that lacked the nAchR. Therefore, it appears that rabies virus utilises multiple receptors. It has been suggested that, like VSV, rabies virus binds to a lipid component of the cell membrane (Seganti *et al.*, 1990; Superti *et al.*, 1984a; Superti *et al.*, 1984b; Wunner *et al.*, 1984). Binding to specific, saturable sites on hamster BHK-21 cells has been observed (Wunner *et al.*, 1984), which is not affected by predigestion of the cells with trypsin or chymotrypsin. Protease digestion also does not prevent binding to CER cells (Superti *et al.*, 1984a; Superti *et al.*, 1984b). In contrast, phospholipase treatment does have an effect, although only phospholipase A2 acts in a dose-response manner (Seganti *et al.*, 1990; Superti *et al.*, 1984a; Superti *et al.*, 1984b; Wunner *et al.*, 1984). Membrane lipid extracts will also inhibit rabies virus binding. Like VSV, binding is prevented by PS, and to a lesser extent by PE and PI (Superti *et al.*, 1984b). However sphingomyelin, or treatment of cells with sphingomyelinase, has no effect. Seganti *et al.* (1990) found no noticeable difference between the inhibitory effects of phospholipids and total membrane lipids. In this study the most significant inhibition of rabies virus binding to CER cells was observed in the presence of the glycolipid fraction from mosquito cells. Digestion with neuraminidase at 5-20U/ml reduces binding to most cells, apart from a mosquito line (Seganti *et al.*, 1990; Superti *et al.*, 1984a). Wunner *et al.* (1984) were unable to reduce binding to BHK-21 cells, but this is likely to be due to the low concentration of enzyme (1U/ml) used. Treatment of the membrane lipid extracts with this enzyme abolishes their inhibitory effect with all cell lines tested.

It has been suggested that gangliosides could act as receptors for rabies virus (Superti *et al.*, 1986). These are a group of glycosphingolipids that carry one or more sialic acid residues in their oligosaccharide chain. They are present on membranes of eukaryotic cells, and are particularly abundant on nervous cells. Superti *et al.* found that exogenous addition of gangliosides can restore susceptibility to neuraminidase treated CER cells. Certain gangliosides are more efficient than others, depending on the number and position of sialic acid residues. The most significant restoration of susceptibility is seen with combinations of different gangliosides.

The ability of rabies virus to infect a wide variety of cell types *in vitro* via an interaction with a lipid does not explain its clear neurotropism *in vivo*. It seems likely that a high affinity receptor(s) is expressed on neural cells. Although nAchR does appear to be such a receptor, it has been shown that *in vivo* spread also involves non-cholinergic neurons (Kucera *et al.*, 1985; Thoulouze *et al.*, 1998; Tuffereau *et al.*, 1998). In addition, Tuffereau *et al.* (1998) found that insect cells expressing rabies G on their surface will bind specifically to several cholinergic and adrenergic neuronal cell lines. Infection of the target cells with rabies blocks binding, presumably by downregulation or masking of a receptor. The presence of a monoclonal antibody raised against nAchR did not prevent binding in this study suggesting that another receptor is involved. Escape mutants involving substitution of Arg 333 have been shown to be non-pathogenic, and display a different host cell range to the wild type (Seif *et al.*, 1985). While Tuffereau *et al.* were unable to affect binding activity by mutating this residue, a double mutation also involving residue 330 did suppress cell binding.

The neural cell adhesion molecule (NCAM) has been proposed as another receptor for rabies virus (Thoulouze *et al.*, 1998). This is a cell adhesion glycoprotein of the Ig superfamily. It has been found to be expressed on the surfaces of highly susceptible cell lines, but not on more resistant cells. Specific blocking of NCAM with antibody or ligand significantly reduces viral infection, as does preincubation of the virus with soluble NCAM. Cells transfected with the *NCAM* gene show enhanced susceptibility to rabies infection. Thoulouze *et al.* also found that infection of NCAM deficient mice resulted in restricted brain invasion by the virus, and a slower disease progression. While it is possible that nAchR and NCAM could act together as receptors, the majority of evidence appears to

suggest that they act independently. Rabies can infect nAchR negative cells expressing NCAM, and can also infect brain tissue in the absence of NCAM. The regions infected in the NCAM negative mice include those known to be binding sites for αBtx. It seems possible that rabies virus uses at least two high affinity receptors to spread *in vivo*, while a lower affinity receptor, such as a ganglioside, allows infection of a wider range of cell types *in vitro*. Another interpretation is that the virus uses the same ganglioside mediated mechanism to enter cells, but the ability to bind high affinity receptors such as nAchR and NCAM allows targeting of particular cell types.

1.3 Virus-cell fusion

The cell entry of all enveloped viruses involves membrane fusion. This procedure is effected by the viral fusion protein. These proteins consist of one or two subunits and are often arranged as trimers. They are usually also responsible for binding to the cell surface receptor. Most fusion proteins are thought to undergo a conformational change, exposing a hydrophobic fusion peptide which interacts with the target membrane. For some viral fusion proteins, including those of influenza virus and the rhabdoviruses, this conformational change is triggered by low pH. The fusion peptides that have been so far identified consist of short hydrophobic sequences, often arranged as an α-helix with a highly hydrophobic face. They tend to show high conservation within but not between virus families (White, 1992). Most of the information available on the viral fusion process has been obtained from the well studied influenza virus.

1.3.1 Retroviral cell fusion

Very little is known about the actual mechanisms involved in retroviral cell fusion. However, these mechanisms are expected to be similar to those employed by influenza virus. Although there are some significant differences between the processes employed by retroviruses and influenza, for example the majority of retroviruses do not require a low pH step, there appear to be sufficient similarities for this model to be a useful aid to understanding retroviral fusion. The hemagglutinin (HA) protein of influenza is functionally equivalent to the Env of retroviruses. Each protein is composed of two subunits: HA1 which is equivalent to SU, and HA2 which is equivalent to TM. These complexes are arranged as trimers on the surface of the virion. The receptor binding domain is formed from folds of the HA1 polypeptide chain, and binds with high specificity but low affinity to the sialic acid residues present on most glycoproteins and glycolipids. Following receptor binding the virus enters the cell via endocytosis. The drop in pH inside the endosome induces a conformational change within the protein, initiating the fusion process. The HA2 protein contains a long carboxy-terminal α-helix and a shorter amino-terminal α-helix. These lie near each other, connected by a loop, in a hairpin-like structure. Just before the shorter α-helix, near the amino-terminal, there is a hydrophobic, fusogenic peptide, thought to be directly involved in membrane insertion. In the inactive state this

peptide is positioned inside the stem of the complex, approximately 35Å from the viral membrane and 100Å from the receptor binding domain. It is held by a system of hydrogen bonds. During the pH-dependent conformational change the hairpin is predicted to straighten out to form a single long α-helix, which along with the other molecules in the trimer creates a triple helix coiled-coil. This brings the fusogenic peptide out into a position where it is able to insert into the target cell membrane. At the same time the binding domains of HA1 dissociate from one another (Bullough *et al.*, 1994; Hunter, 1997; White, 1992). The low pH also causes seven residues near the carboxy-terminus of the coiled-coil to refold into a 180° turn, bending the remainder of the carboxy-terminal end of the helix and three short β-strands up towards the fusion peptide end (Bullough *et al.*, 1994). A representation of the structure of HA is shown in Fig. 1.4.

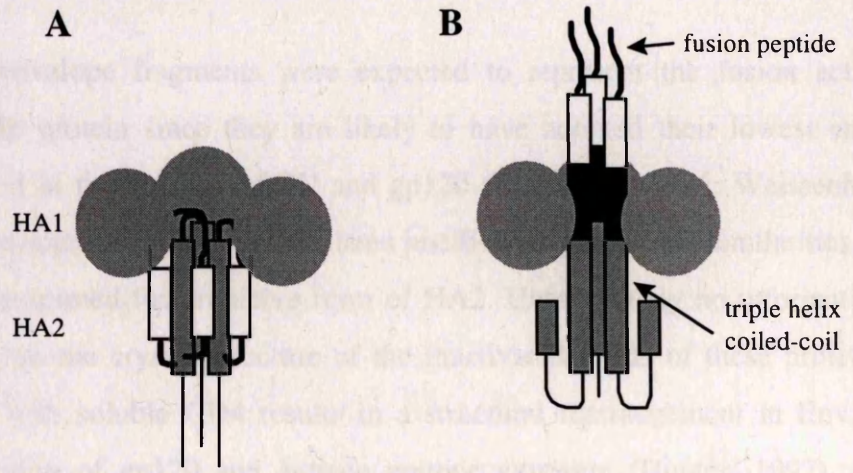


Fig. 1.4. Schematic representation of the structure of the influenza virus HA protein (modified from Hunter, 1997). A) Neutral pH form. B) Activated low pH form.

The ectodomain of the retroviral TM protein is also predicted to contain two α-helices (Zhao *et al.*, 1998). Analysis of the crystal structure of a 55 residue fragment of the MoMLV TM (Fass *et al.*, 1996) has shown that the α-helix immediately downstream of the fusion peptide forms a trimeric coiled-coil, similar to that formed by the equivalent α-helix in the low pH form of HA2. The carboxy-terminal residues at the end of this fragment reverse direction, and extend back up towards the amino-terminus. The crystal structure of a larger fragment of the HIV-1 TM (gp41) ectodomain is also available (Weissenhorn *et al.*, 1997). This structure contains the proposed leucine zipper region and the carboxyl-

terminal helical region. Like HA2 and the MoMLV fragment, the zipper region forms a trimeric coiled-coil. The second α-helix forms an outer layer, running antiparallel to the core, approximately following the groove between two helices. The complex is stabilised by a system of hydrogen bonds and salt bridges. This structure is very similar to that described for the low pH conformation of influenza HA2. The reversal of direction of the polypeptide chain also agrees with the change in direction observed with the MoMLV fragment. The turn is at the same position, and is followed by a glycine rich stretch of 8 amino acids and a disulphide loop in both proteins. These similarities strongly suggest that the two α-helices of the MoMLV ectodomain will also run antiparallel to each other in a similar trimeric structure. Such a structure would be compatible with the theory that a functional interaction occurs between these two α-helical regions during the fusion process (Zhao *et al.*, 1998).

These envelope fragments were expected to represent the fusion active forms of the complete protein since they are likely to have adopted their lowest energy states when produced in the absence of SU and gp120 (Fass *et al.*, 1996; Weissenhorn *et al.*, 1997). This assumption appears to have been justified judging by the similarities with the structure of the presumed fusion active form of HA2. Unfortunately no information is available at present on the crystal structure of the inactivated forms of these proteins. Incubation of HIV-1 with soluble CD4 results in a structural rearrangement in Env, characterised by dissociation of gp120 and antigen epitope exposure (Hunter, 1997). A conformational change can also be induced in the ASLV Env by the binding of a soluble form of its receptor (Hernandez *et al.*, 1997). This change renders the ectodomain hydrophobic, allowing it to bind liposomes. The binding has been shown to be mediated by the TM protein. These receptor-induced changes appear to be analogous to the low pH-induced changes in HA, although further investigations of the process should be carried out before definite conclusions can be made.

The processes by which these proteins actually cause membrane fusion are unknown at present. The structures described above are all trimeric α-helical rods, organised so that the fusion domain and membrane anchor are positioned at the same end of the rod. Similar configurations have been described for the fusion proteins of SV5, a paramyxovirus (Joshi *et al.*, 1998), and the Ebola virus, a filovirus (Weissenhorn *et al.*, 1998), suggesting that a

common mechanism may be employed by these viruses. One possibility is that the positioning of the fusion domain and membrane anchor allows close apposition of the cellular and viral membranes, thereby promoting fusion. The rod-shaped molecules may cluster together at their hydrophobic ends, forming an oligomeric pore (Weissenhorn *et al.*, 1997; Weissenhorn *et al.*, 1998). The cellular SNARE protein complexes involved in vesicle membrane fusion are also composed of α-helical rods with the membrane anchors of the vesicle and plasma membrane proteins localised at one end. The viral fusion peptide may be equivalent to the membrane anchor of the plasma membrane SNARE protein (Skehel and Wiley, 1998). A second possibility is that the observed structures of the viral proteins represent a stable postfusion state (Joshi *et al.*, 1998; Weissenhorn *et al.*, 1998). Insertion of the fusion peptide into the cellular membrane could occur while the ectodomain is in a transient state, and the membranes are then brought together as the rod conformation is formed. The disulphide loop-containing linker present in the MoMLV fragment (Fass *et al.*, 1996) and predicted for gp41 (Weissenhorn *et al.*, 1997) is also found in the Ebola virus protein (Weissenhorn *et al.*, 1998). The conservation of this structure implies that it has a function beyond reversing the polypeptide chain direction, and it has been proposed to act as a hinge between the two α-helices (Weissenhorn *et al.*, 1998). This model is supported by the fact that peptides corresponding to both α-helices of gp41 will inhibit fusion, probably by preventing the protein folding into the rod-shaped conformation. This inhibition only occurs when the peptides are present at the time of fusion. However, other studies have suggested that these peptides exert their effects by disrupting the rod after it has formed (Chan *et al.*, 1997; Rimsky *et al.*, 1998; Weissenhorn *et al.*, 1997), or by preventing pore assembly (Munoz-Barroso *et al.*, 1998). Peptides corresponding to the equivalent α-helices of SV5 will also inhibit fusion activity of the envelope protein (Joshi *et al.*, 1998).

1.3.2 Rabies virus cell fusion

Relatively little is known about the fusion process of rhabdoviruses. Like retroviral Env proteins, rabies G protein undergoes a conformational change which promotes membrane fusion. However, this change results from exposure to low pH during endocytosis of the virus. The rabies G protein is thought to adopt three different conformational states: native (N), active (A) and inactive (I) (Gaudin *et al.*, 1991a; Gaudin *et al.*, 1993). The N state is

that normally observed on the surface of virions. Shortly after acidification (pH 6.7) the A state can be detected by an increase in hydrophobicity (Gaudin *et al.*, 1993). This state is transitory, and is thought to be the conformation required for fusion activity. The I state is reached following prolonged incubation at low pH. In this conformation the protein is 32Å longer than the N state, it becomes sensitive to proteases such as bromelain, and differs antigenically from the other states. While in the I conformation G is unable to promote membrane fusion. However, reincubation at pH 7.15 restores fusogenicity indicating that the conformational changes are reversible (Gaudin *et al.*, 1991a). The I and N states are in a pH-dependent equilibrium. After incubation at pH 6.7 for two hours half of the G molecules can be seen by electron microscope to be in the I state. In this situation virions are unable to fuse with liposome membranes, indicating that several G trimers are required for the fusion process (Gaudin *et al.*, 1993). Gaudin *et al.* (1995) suggest that G is transported in the I state to prevent non-specific fusion activity while in the acidic Golgi apparatus. The reversibility would be required to obtain active G on the surface of the virion.

The rabies G protein does not possess a typical fusion peptide. A putative fusion domain, shown in Fig. 3.1, has been identified between residues 84 and 160 (Durrer *et al.*, 1995). The region between residues 360 and 366 has also been suggested to be involved in the fusion process, due to its homology with other viral fusion peptides (Morimoto *et al.*, 1992). These regions are discussed in more detail in section 3.1.5.

1.4 Retroviral vectors

1.4.1 Vectors

The ability of retroviruses to integrate stably into genomic DNA and be transmitted to progeny cells makes them ideal candidates for the production of vectors for gene transfer. The processes of viral entry, reverse transcription and integration do not require viral protein synthesis. Therefore all viral genes in the vector genome can be replaced with foreign sequences. As a result there is no production of viral proteins after infection, and no subsequent spread of vector. This makes these vectors suitable for use in gene therapy. Vectors can also be used to study retroviral replication since the insertion of a selectable marker often simplifies detection of viral replication and infection (Miller, 1997).

The majority of vectors in use have been derived from the murine and avian retroviruses, although other viruses may also be used. While all viral genes can be removed from these vectors, several *cis*-acting sequences are required. These include a promoter, a polyadenylation signal, the packaging signal ψ, the signals necessary for reverse transcription such as the primer binding site, PPT and the R sequences, and the partially inverted repeats at the ends of the LTRs that are required for integration (Miller, 1997). It has been shown that sequences at the 5' end of *gag* increase the packaging of MLV-based vectors (Adam and Miller, 1988; Armentano *et al.*, 1987; Bender *et al.*, 1987). This larger signal has been designated ψ^+. Its retention allows production of high vector titres of around 10^7 colony forming units (cfu)/ml. The original viral promoter present in the LTR can be used for the expression of the inserted genes. Alternatively, this can be replaced by an internal promoter that may alter the level or tissue distribution of expression (Miller, 1997). Vectors can also be used for the expression of more than one protein. These can be expressed from alternatively spliced mRNAs transcribed by one promoter (Cepko *et al.*, 1984), from separate mRNAs transcribed from multiple promoters (Chen *et al.*, 1992; Hantzopoulos *et al.*, 1989; Overell *et al.*, 1988), or from a single mRNA containing internal ribosome entry site (IRES) elements which allow translation of downstream open reading frames (Adam *et al.*, 1991; Chen *et al.*, 1993; Koo *et al.*, 1992). Two proteins can also be expressed together as a fusion protein that displays functions of both (Germann *et al.*, 1990; Lupton *et al.*, 1991).

1.4.2 Packaging cell lines

In order to package the vector genome, the viral proteins must be provided *in trans* in packaging cells. In early experiments this was achieved by transfecting cells with a vector containing the gene of interest and then infecting the same cells with a wild type helper virus. However, this procedure results in the release of contaminating replication competent retrovirus (Harris and Sikora, 1993). The release of this virus should be avoided if the packaging signal ψ is removed from the helper genome. Mann *et al.* (1983) constructed a MoMLV packaging mutant by deleting 351 base pairs between the left LTR and the start codon for *gag*. Three stable cell lines expressing this construct, ψ-1, ψ-2, and ψ-3, were obtained, and only one of these ψ-3, was found to release infectious virus. Viral RNA from ψ-2 cells was packaged into viral particles with an efficiency of less than 1% that of wild type viral RNA. The cell lines were able to package a vector genome that did not encode any functional viral proteins, ψ-2 cells producing a titre of 10^4 cfu/ml. However, during these transfection experiments high levels of RT activity were detected in the culture medium 9 days after transfection. Mann *et al.* suggested that this was a result of recombination between the mutant genome and endogenous retroviral sequences present in the cellular genome.

There are three major mechanisms that give rise to the production of replication competent viruses from packaging cells lines. These are recombination between the packaging sequences and the vector, recombination of either of these with endogenous retroviral sequences, and the activation of functional endogenous retroviruses (Miller, 1997). The first of these mechanisms is thought to be involved in most instances (Miller and Buttimore, 1986; Miller *et al.*, 1986). When a vector with homologous sequences 3' of the deletion in the defective packaging genome is used in a cell line such as ψ-2, only one recombination event is required to produce wild type virus. Deletion of the ψ signal does not fully prevent packaging of the helper genome. Therefore it is important to reduce homology between the packaging construct and the vector in order to lower the risk of recombination (Miller, 1997).

Later packaging cell lines contain further deletions and mutations. In the PA317 cell line (Miller and Buttimore, 1986) the ψ signal and the 5' end of the 5' LTR have been deleted, and the 3' LTR and second strand initiation site replaced with the polyadenylation signal from simian virus 40 (SV40). As a result any RNA that is packaged cannot be reverse transcribed or integrated into a host genome. Two recombination events are required to produce infectious helper virus. A further modification involved the separation of the *gag/pro/pol* and *env* genes onto two different plasmids, both of which had deletions of the ψ signal and 3' LTR (Danos and Mulligan, 1988; Markowitz *et al.*, 1988). In this system three recombination events are required to produce replication competent virus. However, as the end of the *pol* gene overlaps with the beginning of *env*, recombination between the two constructs can occur. This can be reduced by using the *env* gene from another virus that contains little sequence homology. An early example of this is the DAN amphotropic packaging cell line (Dougherty *et al.*, 1989). This contains the *gag/pro/pol* genes from spleen necrosis virus and the *env* gene from 4070A A-MLV. Although these cells are very effective at preventing helper virus production the vector titre (4×10^4 cfu/ml) is lower than that released from other similar packaging cell lines. More recently, similar packaging cell lines releasing high titre pseudotyped vectors based on MoMLV and bearing the Env proteins of 4070A (Cosset *et al.*, 1995), RD114 (Cosset *et al.*, 1995) or 10A1 (Miller and Chen, 1996) have been produced.

Although it is possible to carry out many experimental protocols with low titre vector, high titre stocks are required for several applications, including *in vivo* gene therapy. The infection of certain cell types, such as human hematopoietic progenitor cells, depends on high titres (Hock and Miller, 1986; Hogge and Humphries, 1987). Titres of around 10^7 cfu/ml can now be obtained with several packaging cell lines. Some of these lines have now been used for human gene therapy trials. Their use indicates that helper virus production occurs at a very low rate in these cell lines, and that it is possible to make large stocks of helper-free vectors (Miller, 1997). It has been shown that prolonged passage of packaging lines results in a reduction in titre (Bender *et al.*, 1987). However, the use of early passage stocks not only prevents this problem, but also reduces the risk of recombination leading to the production of replication competent helper virus.

It is now possible to produce high titre retroviral vector in transient transfection systems (Landau and Littman, 1992; Soneoka *et al.*, 1995). These systems rely on the use of cells such as Cos-7 or 293T that express the SV40 large T antigen. This antigen induces episomal replication of the plasmids, which contain the SV40 origin of replication. Following cotransfection of Cos-7 cells with *gag/pol*, *env* and vector plasmids, Landau and Littman (1992) achieved vector titres of 10^4 cfu/ml, 30 to 40 fold higher than those previously obtained by transient transfection. Soneoka *et al.* (1995) increased this to 10^7 cfu/ml by using the cytomegalovirus (CMV) promoter to drive expression of the packaging components and the vector. The use of this system is illustrated in Fig. 1.5.

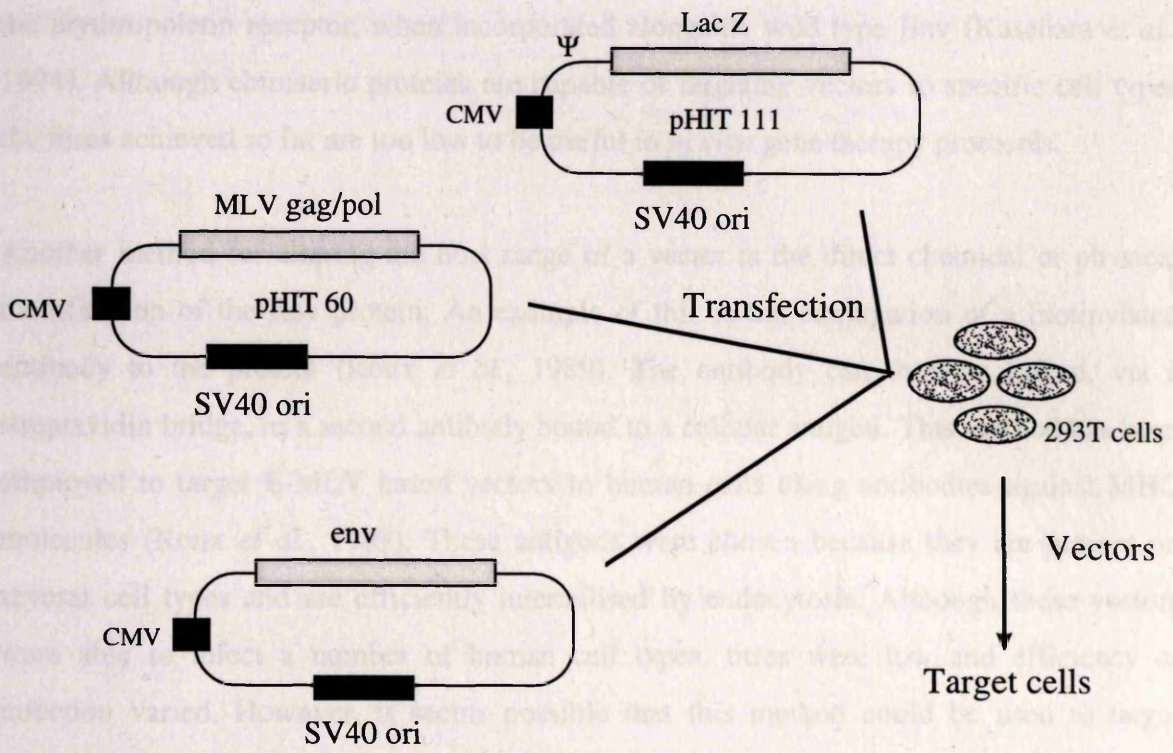


Fig. 1.5. Use of the HIT system (Soneoka *et al.* 1995). The three plasmid constructs, encoding the viral structural proteins and the vector genome, are cotransfected into 293T cells. Transient expression of these constructs then results in the release of high titre vector.

1.4.3 Alteration of the host range of retroviral vectors

Retroviral vectors are restricted to the host cell range of the parental virus. This can be changed by making alterations to the Env protein, either to target the vector to a specific cell type or to expand the host range. One approach involves the production of chimaeric

proteins containing an antibody fragment or ligand that will promote specific binding to the cell type of interest. These sequences are inserted into the amino-terminal region of the SU protein, thought to be involved in normal receptor binding. Antibody fragments have been used to target vectors to a tumour antigen (Chu and Dornburg, 1995), the human LDL receptor (Somia *et al.*, 1995), and human major histocompatibility complex (MHC) class I molecules (Marin *et al.*, 1996). The first two of these studies found that wild type Env was also required. Marin *et al.* reported that their vectors were functional in the absence of wild type Env, but titres were only around 50 cfu/ml. The incorporation of wild type Env also appears to be required for the majority of ligand chimaeras. For example a chimaeric E-MLV Env containing erythropoietin has been shown to target vectors to cells expressing the erythropoietin receptor, when incorporated alongside wild type Env (Kasahara *et al.*, 1994). Although chimaeric proteins are capable of targeting vectors to specific cell types the titres achieved so far are too low to be useful in *in vivo* gene therapy protocols.

Another method for altering the host range of a vector is the direct chemical or physical modification of the Env protein. An example of this is the conjugation of a biotinylated antibody to the protein (Roux *et al.*, 1989). The antibody can then be linked, via a streptavidin bridge, to a second antibody bound to a cellular antigen. This method has been employed to target E-MLV based vectors to human cells using antibodies against MHC molecules (Roux *et al.*, 1989). These antigens were chosen because they are present on several cell types and are efficiently internalised by endocytosis. Although these vectors were able to infect a number of human cell types, titres were low and efficiency of infection varied. However, it seems possible that this method could be used to target several different antigens. Neda *et al.* (1991) chemically coupled lactose residues to E-MLV, which allowed binding to hepatocyte-specific receptors. The modified viruses were able to infect these cells, perhaps via receptor mediated endocytosis. The fact that the modified virus was capable of introducing its genome into the hepatocyte cells suggests that the chemical modification did not hinder the normal fusion of the viral membrane with that of the endosome, prior to destruction in lysosomes.

An antibody or ligand can also be directly fused to the Env protein. Russell *et al.* (1993) have shown that a hybrid protein consisting of an antibody fragment and the MoMLV Env will bind specifically to the targeted antigen. The vectors used in this study were infectious,

although it is not known whether membrane fusion was mediated by this protein as wild type Env was also incorporated. Nilson *et al.* (1996) produced a functional targeted protein by fusing epidermal growth factor to Env via a linker containing a protease recognition site. After binding to the receptor the ligand is proteolitically cleaved, allowing the Env to interact with its normal receptor.

The host cell range of a vector can also be altered by replacing the envelope protein with a protein from another virus. For example, Rous sarcoma virus (RSV) vectors bearing the hemagglutinin (HA) protein of influenza virus have been produced (Dong *et al.*, 1992). This envelope protein extended the host range of the vectors, though titres were reduced. MLV based vectors bearing a variety of envelope proteins have also been constructed. These include GALV Env (Miller *et al.*, 1991), Sendai virus glycoproteins (Spiegel *et al.*, 1998), fowl plague virus HA (Hatziioannou *et al.*, 1998), and VSV G (Burns *et al.*, 1993; Emi *et al.*, 1991).

1.5 Viral pseudotypes

Phenotypic mixing has been observed to occur naturally in cells doubly infected with related or unrelated enveloped viruses (Závada, 1982). The pseudotypes formed in this process have incorporated envelope protein of the coinfecting virus. As a result they display the host range and neutralisation properties of this virus. The formation of pseudotypes indicates that during the budding process, viral particles are able to non-selectively incorporate viral envelope proteins while excluding the majority of cellular membrane proteins (Simons and Garoff, 1980; Závada, 1982). The mechanisms that allow this to occur remain unclear, and much of the evidence appears contradictory. In 1982, Závada put forward two possible models for enveloped virus assembly, illustrated in Fig. 1.6. In model A the viral matrix protein attaches to the lipid membrane where it forms a crystal-like lattice. The internal domains of viral glycoproteins are accepted into this structure while nonviral proteins are rejected. The matrix lattice is essential for viral budding and can function in the absence of envelope protein. In model B the attachment of the matrix protein results in an alteration of the ultrastructure of the lipid membrane, causing cellular glycoproteins to become relatively less soluble and the viral proteins relatively more soluble. As a consequence, all viral envelope proteins concentrate around the area of bound matrix and cellular proteins are excluded.

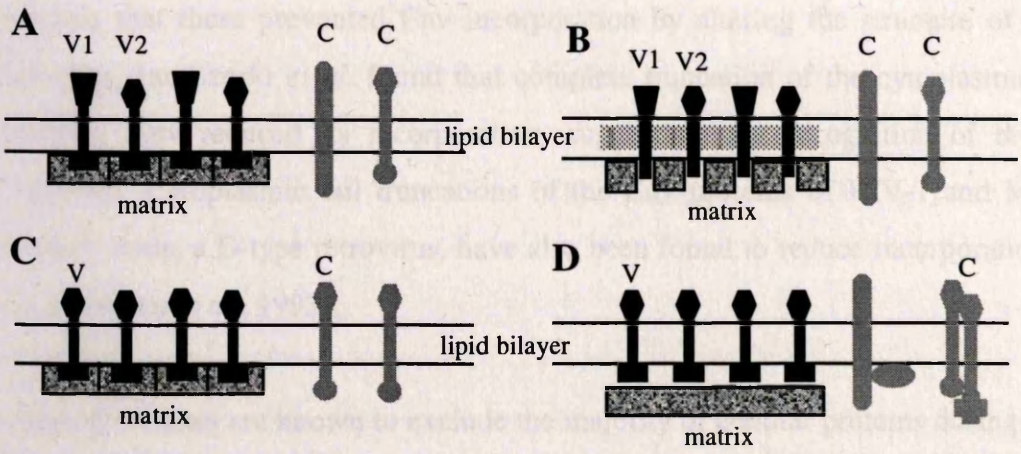


Fig. 1.6. Models of viral envelope incorporation, described in the text. A) Model A (Závada, 1982). Viral proteins are incorporated through an interaction with the matrix protein. B) Model B (Závada, 1982). The matrix protein alters the structure of the lipid bilayer, causing viral proteins to become more soluble than cellular. C) Active incorporation of retroviral Env proteins through a specific interaction with the matrix protein. D) Passive incorporation of retroviral Env proteins. Cellular proteins are prevented from moving into areas of budding by their interactions with other cellular proteins.
V1 and V2, envelope proteins from different viruses; V, viral envelope protein; C, cellular membrane protein.

1.5.1 Envelope protein incorporation by retroviruses

The assembly models proposed by Závada generally correspond to the active and passive models of retroviral Env incorporation (Swanstrom and Wills, 1997). These are also shown in Fig. 1.6. In the active model, incorporation is driven by a specific interaction between the Env cytoplasmic domain and the viral MA proteins underneath the membrane. In the passive model, Env proteins are free to move within the membrane to sites of budding, while cellular proteins are immobilised by interactions with other cellular proteins and are therefore excluded. Model A and the active packing model imply that recognition of the cytoplasmic domain of the Env protein by MA is required for incorporation into the virion. However, it has been shown that RSV *env* mutants lacking this domain yield similar levels of infectious virus as the wild type (Perez *et al.*, 1987). Perez *et al.* claimed that preliminary data implied that Env proteins with deletions within the transmembrane domain are inefficiently incorporated even though they are expressed at the host cell surface. This could mean that this domain, not the cytoplasmic domain, is recognised by the matrix protein. It has also been suggested that sequences within the transmembrane domain may be involved in the incorporation of MoMLV Env (Ragheb and Anderson, 1994), although other studies (Granowitz *et al.*, 1991; Gray and Roth, 1993; Januszeski *et al.*, 1997) have implicated regions of the cytoplasmic tail. The majority of these cytoplasmic domain mutants contained insertions or internal deletions. It is therefore possible that these prevented Env incorporation by altering the structure of the protein. However, Januszeski *et al.* found that complete truncation of the cytoplasmic domain of MoMLV Env reduced its incorporation, suggesting that recognition of this region is important. Cytoplasmic tail truncations of the Env proteins of HIV-1 and Mason-Pfizer monkey virus, a D-type retrovirus, have also been found to reduce incorporation (Brody *et al.*, 1994; Yu *et al.*, 1993).

Although viruses are known to exclude the majority of cellular proteins during the budding process the incorporation of human CD4 by avian leukosis virus (ALV) has been reported (Young *et al.*, 1990). Two chimaeric proteins, consisting of the extracellular region of CD4 and the transmembrane and cytoplasmic domains of the Env protein of either ALV or MLV, were transiently expressed in quail QT6 cells that had been previously transfected with an ALV vector. These proteins were incorporated into the virions, but the control wild

type CD4 protein was incorporated 5 to 10 times more efficiently. Young *et al.* suggested that the quail cells may lack the proteins that would normally interact with CD4, therefore allowing the molecule to move freely within the membrane. This implies that incorporation of host cell surface proteins is normally prevented by their interactions with other membrane associated and cytoplasmic proteins, as proposed in the passive model (Swanstrom and Wills, 1997). The lower level of incorporation of the chimaeric CD4 proteins may have been a result of competition or other interactions with the retroviral Env protein. Young *et al.* claimed that unpublished data show that an avian sarcoma virus that lacks *env* will incorporate both the wild type and chimaeric CD4 with equal efficiencies.

Dong *et al.* (1992) found that the influenza HA protein could be inserted into the envelope of RSV as efficiently as a chimaeric HA containing the membrane anchor and cytoplasmic domain of RSV Env. They concluded that this result indicated that recognition of either of these domains is not required for envelope incorporation. Dong *et al.* also found that the presence of influenza neuraminidase (NA) increased the incorporation of HA into RSV particles by approximately 70% relative to Gag. They suggested that the action of this enzyme may have released the HA protein from interactions with cell surface proteins, allowing it to move freely into regions of budding. However, as implied in model A, it is possible that a recognition signal is present in both viral envelope glycoproteins. This would allow the incorporation of the foreign viral protein at a comparable efficiency. In contrast, Lindemann *et al.* (1997) found that the replacement of the majority of the cytoplasmic domain of human foamy virus Env with that of MLV Env increased its incorporation into MLV-based vectors.

If there is a specific interaction between Env and MA it would be expected that mutations in MA could also affect Env incorporation. This has been found to be the case for certain deletions and point mutations within the amino-terminus of HIV-1 MA (Freed and Martin, 1996; Lee *et al.*, 1997; Yu *et al.*, 1992). While this seems to indicate an interaction between the two proteins, it is also possible that a specific structure of MA is required in order to contain the long cytoplasmic domain of HIV-1, and that this is disrupted by these mutations. Interestingly, at least some MA mutants are able to incorporate HIV-1 Env proteins with truncated cytoplasmic domains (Freed and Martin, 1995; Mammano *et al.*, 1995). These mutants are also able to incorporate A-MLV Env, which has a much shorter

cytoplasmic domain than HIV-1 Env. The difference in the length of this domain in the two proteins may explain why HIV-1 Env is not incorporated by MoMLV (Wilson *et al.*, 1989). Mammano *et al.* (1997) have shown that truncation of the long cytoplasmic domain does permit the incorporation of HIV-1 Env by this virus.

In infected polarised epithelial cells budding of enveloped viruses often occurs specifically at either the apical or basolateral membrane. Budding of the avian and mammalian C-type retroviruses and of HIV-1 occurs at the basolateral membrane (Lodge *et al.*, 1994; Owens *et al.*, 1991; Roth *et al.*, 1983). Following the expression of HIV-1 Gag in epithelial cells in the absence of Env, viral maturation and budding is spread evenly between the two membranes. In the presence of Env expression the polarisation of budding is restored (Owens *et al.*, 1991). This restriction of budding is due to a basolateral membrane targeting signal present in the cytoplasmic domain of Env (Lodge *et al.*, 1997a; Lodge *et al.*, 1997b). The ability of Env to direct the budding of virions in this way supports the theory that there is a specific interaction between Env and MA, and suggests that such an interaction could occur within the cell, prior to assembly at the cell surface. An interaction within the cell is also implied by the observation that certain MA mutants affect the cellular transport of HIV-1 Env in $CD4^+$ T lymphocytes (Lee *et al.*, 1997). While it may be suggested that the high levels of protein expression in these experiments could in itself be influencing the intracellular transport of Gag, it has been observed that mutations in MA that normally affect Env incorporation can prevent Env-directed polarised budding (Lodge *et al.*, 1994).

The Env proteins of human T-cell leukaemia virus type 1 (HTLV-1) and MoMLV have been shown to direct the polarised budding of Env-negative HIV-1 virions in epithelial cells (Lodge *et al.*, 1997a), suggesting that the putative incorporation signal is common to all Env proteins. However, a higher level of expression was required to achieve polarisation compared with that required with the HIV-1 Env. This implies that incorporation of the heterologous Env proteins is less efficient. Heterologous envelope proteins may also be incorporated less efficiently by MoMLV. Suomalainen and Garoff (1994) found that Semliki Forest virus spike protein or a truncated mutant of the human transferrin receptor could only be incorporated efficiently when expressed at high density on the cell surface. In contrast, MoMLV Env was efficiently incorporated when expressed at low density. Further evidence of a virus-specific interaction between Env and the Gag

proteins comes from the work of Weclewicz *et al.* (1998). This study found that when expressed separately in rat neurons, Env is restricted to the somatic region of the cells while Gag is found in both the axonal and somatic regions. When the proteins are coexpressed they are both found only in the somatic region. Weclewicz *et al.* were able to show that this was true for both MLV and HIV. However, expression of the Gag proteins with the heterologous Env did not restrict the localisation of Gag, indicating that the interaction is virus-specific. It seems possible that such an interaction allows the virus to preferentially incorporate its own Env protein. However, since this interaction is not essential for virion release (Owens *et al.*, 1991; Ruta *et al.*, 1979; Smith *et al.*, 1990), budding virions are able to incorporate foreign envelope proteins either passively or via a low affinity interaction.

1.5.2 Envelope protein incorporation by rhabdoviruses

Much of the evidence obtained from studies of VSV suggests that efficient incorporation of the G envelope glycoprotein depends on an interaction with the matrix (M) or other internal viral proteins. Several studies investigating the incorporation of VSV G have made use of the temperature sensitive mutant *ts*045 (Flamand, 1970; Lafay, 1974). These viruses possess a mutation in the G gene which results in the accumulation of the protein in the endoplasmic reticulum at the nonpermissive temperature, 40°C. Whitt *et al.* (1989) tested a series of mutated G proteins for their ability to rescue this mutant at the nonpermissive temperature. G proteins that retained 9 or 18 of the membrane proximal cytoplasmic amino acids rescued *ts*045 at levels 3 to 4 fold lower than the wild type. Those that retained only 1 or 3 of these residues produced just a slight increase in titre compared to the mock-transfected control. This suggests that an interaction between this membrane proximal region and the internal VSV proteins may be necessary for G incorporation. G proteins with internal deletions in the cytoplasmic domain rescued *ts*045 at near wild type levels, indicating that the membrane distal region may also be important. G proteins with foreign cytoplasmic domains did not rescue to a significant level. Whitt *et al.* suggested that the low level of rescue obtained with some of the G protein mutants may be the result of random incorporation. The cytoplasmic tail of the rabies G protein also appears to contain a signal that promotes its incorporation into virions. Mebatsion *et al.* (1996) found that removal of this domain significantly reduces incorporation of G by rabies virus. In addition

the virion particles varied in their G protein content, suggesting that incorporation of this truncated protein is non-specific.

Avian myeloblastosis virus (AMV) has been shown to be capable of rescuing VSV *ts*045 at the nonpermissive temperature (Závada, 1972). The virus particles released had the neutralisation, host range and interference specificity of the retrovirus indicating that the Env protein had been incorporated. Weiss and Bennett (1980) were also able to show rescue of *ts*045 by AMV. However, rescue attempts with MoMLV were much less successful. As this discrepancy did not appear to be the result of the cell type used the authors suggested that these differences were due to the strain of retrovirus. If this is the case it is possible that the cytoplasmic domains of the different envelope proteins have different affinities for the M protein. This would be in partial agreement with model A (Závada, 1982, Fig. 1.6), since it implies that the different viral envelope glycoproteins interact with matrix proteins, albeit with different efficiencies. It would be expected that the native proteins would display the highest affinity.

Witte and Baltimore (1977) reported that VSV pseudotypes containing the envelope protein of MoMLV could only be produced at permissive and semi-permissive temperatures with various G *ts* mutants, or with wild type VSV. They interpreted this to mean that at least some G protein was required for the budding process. Schnell *et al.* (1996) have shown that CD4 or measles virus envelope proteins can be incorporated into VSV alongside the G protein. As no reduction in the levels of G was observed, the authors suggested that this protein interacts specifically with M while the foreign proteins are passively trapped into the budding virions. The spikeless *ts*045 mutant virions do contain some truncated G protein, consisting of the transmembrane and cytoplasmic domains (Metsikkö and Simons, 1986). This may explain why these viruses are able to bud at the nonpermissive temperature. Pattnaik and Wertz (1991) also concluded that G is required for budding of VSV. In this study the five VSV proteins were expressed in cells in order to test their ability to rescue defective interfering particles. These particles are unable to replicate independently because they do not code for any functional VSV proteins. The expression of the nucleocapsid proteins in the absence of either M or G did not result in viral budding, indicating that both proteins are required. However, culture media were only tested for the presence of virus following a second infection. Any virus lacking G that was

produced during the first infection would be unable to infect and replicate in the second set of cells, and therefore would not have been detected. The related rabies virus has been shown to be able to bud in the absence of its G protein, although particle production is reduced approximately 30 fold (Mebatsion *et al.*, 1996).

Owens and Rose (1993) were unable to rescue *ts*045 by the expression of the HIV-1 Env protein in infected cells. However, the expression of a chimaeric protein, in which the cytoplasmic domain had been replaced by that of VSV G, resulted in the release of infectious virus. These viruses were specific for $CD4^+$ human HeLa cells and could be neutralised by anti-HIV serum. The ability of this chimaeric protein to rescue the VSV mutant suggests that the cytoplasmic tail of G provides a positive signal for the incorporation of the envelope protein into the virion. Similar results have also been obtained with rabies virus. Mebatsion and Conzelmann (1996) have shown that HIV-1 Env is only incorporated into a G-negative rabies mutant when the cytoplasmic domain is replaced with that of rabies G, indicating that this region is required. Of course, it is also possible that replacement of the long HIV-1 Env cytoplasmic domain removed a structure that was incompatible with incorporation into rhabdovirus particles. Owens and Rose (1993) found that a protein containing part of the HIV-1 Env cytoplasmic domain as well as that of VSV G was not incorporated by the *ts*045 virus. Rabies virus lacking G protein can incorporate the G protein of Mokola virus, another *Lyssavirus*, although the titre of this pseudotype is reported to be approximately 30 fold lower than that of the wild type rabies virus (Mebatsion *et al.*, 1995). The cytoplasmic domain of Mokola G shares little homology with that of rabies G but it is also short, and may therefore be more easily accommodated than HIV-1 Env. Schnell *et al.* (1998) found that recombinant VSV expressing a chimaeric G protein containing the cytoplasmic domain of CD4 are able to bud efficiently, while those containing a cytoplasmic domain of only 1 amino acid did not. Following passage of this latter mutant, a revertant was selected that had acquired an 8 amino acid cytoplasmic domain. These observations imply that an envelope protein with a short cytoplasmic domain is necessary for efficient budding of VSV, but that there is no strong sequence requirement within this domain.

1.5.3 Uses of viral pseudotypes

The ability of enveloped viruses to incorporate foreign envelope proteins can be exploited for several experimental and clinical applications. For example, as mentioned above (section 1.4.3), pseudotyping can be employed to alter the properties of viral vectors used for gene transfer. The production of vectors containing the VSV G glycoprotein was an important advance in the development of retroviral vectors (Burns *et al.*, 1993; Emi *et al.*, 1991). These vectors have a wide host range, and can be concentrated to high titres. Pseudotyped viruses can also be used to easily define the host range of viruses for which no simple infection assays are available. These pseudotypes bear the envelope protein of the virus of interest, and therefore will only infect cells that express the receptor recognised by this protein. Pseudotypes of VSV have been used to study the receptor usage of several viruses, including hepatitis C virus and HIV (Boettiger, 1979; Dalgleish *et al.*, 1984; Lagging *et al.*, 1998; Sommerfelt and Weiss, 1990). Infection can be detected by the formation of visible plaques. Pseudotyped vectors can also be used for receptor studies as the opportunity for the insertion of marker genes into the vector genome allows simple detection of infected target cells. Retroviral vectors have been used to investigate the function of several viral glycoproteins, for example those of ASLV-A and Ebola (Rong *et al.*, 1997; Wool-Lewis and Bates, 1998).

1.6 Aims

The aims of this study were two-fold, both of which made use of the possibility of producing pseudotyped retroviral vectors that display the host range properties of the virus donating the envelope protein.The first part involved attempts to produce a retroviral vector bearing the envelope protein of rabies virus. Retroviral vectors bearing the envelope protein of VSV have been shown to display increased stability and the wider host range of VSV (Burns *et al.*, 1993). Unfortunately high level expression of the VSV protein causes cell fusion. It was hoped that the equivalent rabies protein would be more suitable for high expression in cell culture, and that its incorporation into a retroviral vector would confer similar properties as the VSV protein.

The second part involved analysis of the host ranges of endogenous retroviruses of pigs. This is of interest due to the proposed use of porcine organs and tissues for human transplants, and the potential risk these viruses pose in such a process. Although host range can be investigated through the transfer of infection from virus-releasing cells to target cells, the use of pseudotyped vectors allows individual analysis of the Env proteins of each ERV strain.

CHAPTER 2

Chapter 2 Materials and methods

2.1 Materials

2.1.1 Cell lines

Cos-7 is an African green monkey kidney fibroblast line, transformed by an origin-defective mutant of SV40. These cells express the SV40 large T antigen (Gluzman, 1981).

293 is a human embryonic kidney line, transformed with sheared human adenovirus 5 DNA (Graham *et al.*, 1977; Graham *et al.*, 1978; Harrison *et al.*, 1977).

293T is an SV40 large T antigen expressing line derived from 293 cells (DuBridge *et al.*, 1987).

293-GP is a FeLV Gag/Pol expressing line derived from 293 cells (a gift from Derek Bain, Vet Pathology, University of Glasgow)

BHK TK$^-$ is a Syrian hamster kidney fibroblast line, derived from BHK21 (clone 13) and rendered TK negative by selection in BUdR (Littlefield and Basilico, 1966; Stoker and Macpherson, 1964).

AH927 is a feline embryonic fibroblast line (Rasheed and Gardner, 1980).

NIH 3T3 is a murine fibroblast line, established from a NIH Swiss mouse embryo (Andersson *et al.*, 1979; Copeland and Cooper, 1979; Jainchill *et al.*, 1969).

PK15 is a porcine kidney epithelial line (Armstrong *et al.*, 1971).

MPK is a minipig kidney fibroblast line, established from a 118 day old foetal minipig (ECACC ref. no. 87032604).

ST is a porcine testicular line established from the trypsinised testis of a normal 80-90 day old porcine foetus (McClurkin and Norman, 1966).

Raji is a human Burkitt lymphoma line established from an 11 year old male (Pulvertaft, 1964).

SupT1 is a human T cell lymphoblastomic lymphoma line (Smith *et al.*, 1984).

J.Jhans is a human T cell line.

RK13 is a rabbit kidney fibroblast line established from a 5 week old rabbit (Beale *et al.*, 1963).

Tb1Lu is a bat lung epithelial line established from an adult female bat (ECACC ref. no. 90020805).

2.1.2 Cell culture media

All media and supplements were purchased from Gibco, UK, unless otherwise stated. Media described as 10% contained 10% foetal calf serum (FCS). Cos-7, AH927, NIH 3T3, ST, and RK13 cells were all grown in 10% DMEM (Dulbecco's Modified Eagle's Medium). 293, 293T, 293-GP, PK15, MPK, and Tb1Lu were all grown in 10% DMEM supplemented with 1X non-essential amino acids. 400µg/ml G418 (Calbiochem, UK) was added to stock flasks of 293T cells in order to select for neomycin resistant cells expressing the SV40 large T antigen. 50µg/ml zeocin (Invitrogen, UK) was added to stock flasks of 293-GP cells in order to select for zeocin resistant cells expressing FeLV Gag/Pol. BHK TK^- cells were grown in Glasgow MEM (Minimum Essential Medium) supplemented with 5% FCS. Raji, SupT1 and J.Jhans cells were grown in 10% RPMI (Rosewall Park Memorial Institute) 1640 medium. All media were also supplemented with 2mM glutamine, 100 U/ml penicillin and 10µg/ml streptomycin.

2.1.3 Bacterial strains

E. coli DH5α: subcloning efficiency DH5α competent cells (Gibco, UK). F^- ϕ80d*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1.

E. coli HB101: subcloning efficiency HB101 competent cells (Gibco, UK). F^- *mcr*B *mrr* *hsd*S20(r_B^-, m_B^-) *rec*A13 *sup*E44 *ara*14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20(Sm^r) *xyl*5 λ^- *leu mtl*1.

E. coli DM1: subcloning efficiency DM1 competent cells (Gibco, UK). Absence of *dam* and *dcm* methylases prevents methylation at GATC and CC(A/T)GG sites. F^- *dam*$^-$ 13::Tn9(Cm^R) *dcm-mcr*B *hsd*R^-M^+ *gal*1 *gal*2 *ara*$^-$ *lac*$^-$ *thr*$^-$ *leu*$^-$ *ton*^{R} *tsx*^{R} *su*^{O} λ^-.

2.1.4 Plasmids

The eukaryotic expression vectors used were pcDNA3 (Invitrogen, UK) and pCI-neo (Promega, UK).

pcDNA3 carries ampicillin resistance (Amp^r), neomycin phosphotransferase (*neo*) which confers resistance to G418, the human CMV immediate-early promoter, the bovine growth hormone polyadenylation (poly A) signal, and the SV40 origin of replication (*ori*) which induces transient, episomal replication of the vector in cells expressing the SV40 large T antigen.

pCI-neo also carries Amp^r, *neo*, the human CMV immediate-early promoter, and the SV40 *ori*. In addition it carries the SV40 poly A signal, and a chimaeric intron downstream of the CMV promoter region, believed to increase the level of gene expression.

The HIT system plasmids, pHIT60, pHIT456 and pHIT111 (Soneoka *et al.*, 1995), were kindly provided by Yuko Soneoka (University of Oxford). These are MLV based retroviral constructs with the CMV promoter and SV40 *ori*.

pHIT60 encodes the *gag/pol* gene.

pHIT456 encodes the amphotropic MLV *env* gene.

pHIT111 is a vector construct which carries the packaging signal ψ, *neo* and the *LacZ* gene encoding β-galactosidase.

pMV12 was a gift from Gavin Wilkinson (University of Wales College of Medicine, Cardiff). This plasmid carries the *LacZ* gene under the control of the CMV (strain AD196) promoter (position -299 to +69) and immediate-early poly A signal.

pGT5 was a gift from John Oldfield (Veterinary Pathology, University of Glasgow). This plasmid encodes the FeLV *gag/pol* gene.

2.1.5 Reagents and stock solutions

All reagents and chemicals were obtained from Sigma, BDH, Boehringer Mannheim, Pharmacia, Fischer Scientific or Gibco unless otherwise stated, and were of Analar quality.

Agarose gel (TBE): 0.8 or 1% (w/v) agarose, 50ml 1X TBE buffer, 0.5μg/ml ethidium bromide.

Agarose gel (TAE): 4% (w/v) agarose, 50ml 1X TAE buffer, 0.5μg/ml ethidium bromide.

Ammonium persulphate (APS): 10% (w/v) in deionised water (dH_2O). Stored at 4°C for up to a week.

Ampicillin: 50mg/ml stock in dH_2O. Filtered through a 0.2μm filter and stored at -20°C.

Denaturing polyacrylamide gel: 4% (w/v) acrylamide (Sequagel XR, National Diagnostics), 7M urea, 1.2% (v/v) TBE in dH_2O. 400μl of 10% APS and 40μl of TEMED (N, N, N', N'-tetramethylethylenediamine) was added for each 60ml sequencing gel.

DNA size markers: ϕX174 RF DNA/HaeIII fragments and λDNA/HindIII fragments.

Dithiothreitol (DTT): 1M stock in dH_2O. Stored in 1ml aliquots at -20°C.

Ethidium bromide: 10mg/ml stock in dH_2O. Stored at room temperature protected from light.

G418: 50mg/ml stock in dH_2O. Filtered through a 0.2µm filter. Stored at -20°C.

Gel loading buffer: 50% (v/v) glycerol, 20mM EDTA, 0.5% (w/v) bromophenol blue, 0.5% (w/v) xylene cyanol. Stored at room temperature. Used at a 1/10 dilution.

2X HBS: 280mM NaCl, 50mM HEPES, 1.5mM Na_2HPO_4, pH 7.12. Filtered through a 0.2µm filter, and stored at 4°C.

Laemmli sample buffer: 2% (w/v) SDS, 10% (v/v) glycerol, 60mM Tris pH 6.8, 0.005% (w/v) bromophenol blue. Stored at room temperature. 100mM DTT was added immediately prior to use from a 1M stock.

Laemmli running buffer (10X): 2.5M Tris, 19M glycine, 1% (w/v) SDS, pH 8.3. Stored at room temperature.

LB: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl in dH_2O. Sterilised by autoclave. Stored at room temperature.

LB-agar: 1.5% (w/v) agar in LB. Sterilised by autoclave. Stored at room temperature.

Low TE buffer: 1mM Tris pH 7.5, 50µM EDTA. Filtered through a 0.2µm filter. Stored at 4°C or at room temperature.

4% paraformaldehyde: 0.1M PBS, 0.12M sucrose, 4% (w/v) paraformaldehyde, pH 7.4. Stored at 4°C.

PBS: 0.1M Na_2HPO_4, 0.1M NaH_2PO_4, 0.6% (w/v) NaCl, pH 7.4. Sterilised by autoclave, and stored at room temperature.

Phenylmethylsulphonyl fluoride (PMSF): 100mM stock in ethanol. Stored at -20°C.

Poly-L-lysine: 1mg/ml in 0.1M borate buffer, pH 8.5. Filtered through a 0.2µm filter. Stored at 4°C.

Polyacrylamide gel (SDS-PAGE): 10% (w/v) acrylamide mix (Scotlab, UK), 390mM Tris pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) APS, 0.04% TEMED.

PRO-MIX™: 70% L-[^{35}S]Met, 30% L-[^{35}S]Cys. (Amersham Life Science, UK)

Protein markers: Rainbow high molecular weight markers (Amersham Life Science, UK).

RIPA lysis buffer: 1% (w/v) NP-40, 150mM NaCl, 50mM Tris pH7.4, 20mM EDTA pH8. Stored at room temperature. 1mM PMSF was added immediately prior to use from a 100mM stock.

TAE (50X): 2M Tris-HCl pH 8.15, 1.5M NaOAc, 1M NaCl, 0.1M EDTA. Stored at room temperature.

TBE (10X): 0.9M Tris-HCl, 0.9M boric acid, 25mM EDTA. Stored at room temperature.

TBS: 137mM NaCl, 3mM KCl, 25mM Tris base, 1.5% phenol red, pH 7.4.

TE buffer: 10mM Tris, 1mM EDTA, pH 8. Sterilised by autoclave or by filtration through a 0.2µm filter. Stored at room temperature.

Versene: 800µg/ml NaCl, 20µg/ml KCl, 115µg/ml Na_2HPO_4, 20µg/ml EDTA, 1% phenol red. Sterilised by autoclave. Stored at room temperature.

X-gal staining solution: 2.1mg/ml potassium ferrocyanide, 1.6mg/ml potassium ferricyanide, 0.186mg/ml magnesium chloride. Stored at 4°C for up to a week. 1mg/ml X-gal was added immediately prior to use from a 40mg/ml stock.

X-gal: 40mg/ml stock in DMSO.

2.2 Methods

2.2.1 Eukaryotic cell culture

2.2.1.1 Maintenance of cells

Cells were grown at 37°C in an atmosphere of 5% CO_2 in air, in sterile plastic flasks. Adherent cells were subcultured by washing in phosphate buffered saline (PBS) and incubating for 1 to 5mins in trypsin/EDTA (Gibco, UK). The cells were then washed off and resuspended in fresh medium. MPK, PK15 and ST cells were washed in versene prior to incubation in trypsin/EDTA. Non-adherent cells were recovered by centrifugation and resuspended in fresh medium.

2.2.1.2 Cell storage

PK15 cells were stored in 5% glycerol, 10% FCS, 85% DMEM. MPK cells were stored in 5% dimethyl sulphoxide (DMSO), 10% FCS, 85% DMEM. All other cell lines were stored in 10% DMSO, 10% FCS, 80% medium. Cells were either frozen to -100°C in a controlled rate cell freezer (Planer Kryo 10, Series II), or to -70°C in a polystyrene container overnight, before being transferred to liquid nitrogen vapour phase.

2.2.1.3 Transfection of cells

Most transfections were carried out using calcium phosphate precipitation (Graham and van der Eb, 1973). Cells were seeded the day before transfection to give approximately 40% confluency in 25cm^2 flasks. 5µg of each plasmid was added to low TE buffer to a final volume of 210µl. 30µl of 2M $CaCl_2$ was added drop-wise and the solution was mixed thoroughly. This was then mixed with 240µl 2X HEPES buffered saline (HBS), left to form a precipitate for 30min, and then added to the cells in fresh medium. 16-18 hours later the transfection medium was removed, the cells were washed in PBS and fresh medium was added.

The other methods of transfection used were lipofection and DEAE-dextran. Lipofection was carried out using LIPOFECTAMINE™ (Gibco, UK). This is a 3:1 (w/w) liposome formulation of 2,3-dioleyloxy-N-[2(sperminecarboxamido)ethyl]-N,N-dimethyl-1-propanaminium trifluoroacetate, and dioleoyl phosphatidylethanolamine in water. Approximately 1 x 10^5 cells were seeded in 6 well plates 24 hours prior to transfection. 1µg of plasmid in 100µl serum free DMEM was added to 6µl LIPOFECTAMINE™ in 100µl serum free DMEM, and incubated at room temperature for 45min. This was then mixed with 800µl serum free DMEM and added to the cells. All medium used was antibiotic free. After incubation at 37°C for 5 hours the cells were washed and fresh, complete medium was added. For the transfection of cells in 25cm^2 flasks, figures were increased by a factor of 2.5.

When the DEAE-dextran transfection method was used, 6µg of plasmid was mixed with 2.5µl 50mg/ml DEAE-dextran/Tris buffered saline (TBS) in 500µl serum free DMEM. The cells to be transfected were washed twice with serum free DMEM and then incubated with the DNA/DEAE-dextran mixture for 4 hours at 37°C. The transfection medium was then removed and 10% DMSO in TBS was added for 2min. The cells were washed with PBS and complete medium was added.

2.2.2 Preparation of DNA

2.2.2.1 Preparation of plasmids

Small scale preparations of plasmids were obtained from 2ml overnight cultures using the Quantum Prep Plasmid Miniprep Kit (Bio-Rad, UK), according to the manufacturer's instructions. Briefly the cells were lysed using alkaline lysis, cellular debris removed by sodium dodecyl sulphate (SDS) -precipitation, and the remaining solution mixed with a DNA-binding matrix. This was washed, and the DNA then eluted with dH_2O or TE buffer, and stored at 4°C.

Large scale DNA preparations were obtained from 100-500ml overnight cultures using the Qiagen maxi-prep kit (Qiagen Ltd., UK), according to the manufacturer's instructions.

Briefly the cells were lysed using alkaline lysis, and RNA degraded by RNase A. Cellular debris was removed by SDS-precipitation. Plasmid DNA was then bound to an anion-exchange resin under low salt conditions, washed with a medium salt buffer to remove contaminants, and eluted by a high salt buffer. The DNA was desalted by isopropanol precipitation, washed with 70% ethanol, air-dried and redissolved in TE buffer. The concentration of DNA was calculated on the basis that 50µg/ml double stranded DNA has an OD_{260} measurement of 1. DNA solutions were stored at 4°C or -20°C.

2.2.2.2 Preparation of genomic DNA

Genomic DNA was prepared using the Nucleon® Genomic DNA Extraction Kit (Scotlab, UK), according to the manufacturer's instructions. Briefly, approximately 5 x 10^6 cells were lysed and then deproteinised with sodium percolate. The DNA was extracted with chloroform while a resin was added to bind and remove any cell impurities. The DNA was then ethanol precipitated, washed in 70% ethanol, air-dried, and dissolved in low TE buffer. The concentration of DNA was calculated on the basis that 50µg/ml double stranded DNA has an OD_{260} measurement of 1. DNA solutions were stored at 4°C or -20°C.

2.2.2.3 Restriction endonuclease digestion

Restriction endonucleases were used with the buffers supplied, according to the manufacturer's instructions. In general, 10 units of enzyme was used to digest 1µg of DNA at 37°C for 2 hours. The enzyme was then heat inactivated by incubation at 70°C for 10min.

2.2.2.4 Gel electrophoresis

DNA samples were separated by electrophoresis on 0.8 - 4% agarose gels in 1X TBE or 1X TAE buffer. 50ml gels were poured in perspex tanks and wells were cast using the appropriate combs. Gel loading buffer was added to the samples which were then loaded

into the wells and run through the gel at 70 to 100V. Gels were viewed on a short wave UV transilluminator (UVP Inc.).

2.2.2.5 Gel purification of DNA

DNA samples were run out on 0.8% agarose TAE gels. The DNA was visualised under ultra-violet light and the appropriate band was then excised using a clean scalpel. The DNA was purified from the gel using the QIAquick™ Gel Extraction kit (Qiagen Ltd., UK), according to the manufacturer's instructions. Briefly, the gel was dissolved in a high salt buffer and the DNA was then bound to a silica membrane. After washing, the DNA was eluted from the membrane with dH_2O.

2.2.3 Cloning of hybrid DNA molecules

2.2.3.1 Ligation of DNA fragments

Fragments of DNA generated by restriction endonuclease digestion were ligated into vectors by the following procedure. Firstly, vector plasmid was linearised with an appropriate restriction enzyme. To prevent recircularisation of the vector the 5' phosphate groups at the digested ends were hydrolysed by incubation with calf intestinal alkaline phosphatase (CIAP, Gibco, UK) at 37°C for 45mins. This enzyme was then heat inactivated for 10min at 70°C. The linearised vector and DNA fragment were gel purified (section 2.2.2.5) and then ligated using T4 DNA ligase (Gibco, UK) either at room temperature for 4 hours or at 14°C overnight. The ligation reaction was then used to transform *E. coli*.

2.2.3.2 Transformation of *E. coli*

The majority of plasmids were grown in subcloning efficiency DH5α competent cells (Gibco, UK). pHIT60 was grown in HB101 competent cells (Gibco, UK) to avoid deletion of the plasmid. 1-10ng of DNA was added to 50µl of cells and incubated on ice for 30min. The cells were heat-shocked at 37°C for 20s and then returned to ice for a further 2min.

950µl of Lennox L Broth (LB, Gibco, UK) was added and the cells were incubated for 1 hour at 37°C in an orbital incubator (Stuart Scientific, S1 50). The cultures were then spread onto LB agar (Gibco, UK) plates containing 50µg/ml ampicillin (Amp, Sigma, UK) and incubated overnight at 37°C. Amp resistant colonies were picked with sterile toothpicks and grown up in LB containing 50µg/ml Amp. Clones containing the plasmid of interest were identified by small scale DNA purification (section 2.2.2.1) and restriction digestion (section 2.2.2.3).

2.2.4 Production of pseudotyped vectors

2.2.4.1 Transient production of vectors

Vectors were produced using the HIT system as described previously (Soneoka *et al.*, 1995). 293T cells were transfected with pHIT60, pHIT111, and the appropriate envelope protein construct, using calcium phosphate precipitation (section 2.2.1.3). One flask was transfected with pHIT111 alone as a negative control, and one with pMV12 as a control for transfection. Following transfection fresh DMEM was added to the cells. In order to measure vector titre, the medium was removed, filtered through 0.45µm filters, and added to target cells in 1ml aliquots in the presence of 8µg/ml Polybrene (hexadimethrine bromide, Sigma, UK). Samples from pHIT456 transfections were added as a 1/1000 dilution. The cells were incubated at 37°C for 2 hours and then fresh medium was added. 48 hours later the cells were stained for β-galactosidase (β-gal) activity. Staining of pMV12 transfected cells was carried out when the samples were taken.

2.2.4.2 Detection of β-gal activity

Cells were washed in warm PBS and fixed for 15min in 4% paraformaldehyde. They were then washed twice before being permeabilised with 0.3% Triton X-100/PBS for 5min. After a further 2 washes staining solution containing 1mg/ml X-gal (5-bromo-4-chloro-3-indoyl β-D-galactopyranoside, Biogene Ltd., UK) was added and the cells were incubated at 37°C overnight. Vector titres were then determined by counting the number of blue cell colonies.

CHAPTER 3

Chapter 3 Production of retroviral vectors pseudotyped with the rabies virus G glycoprotein

3.1 Introduction

3.1.1 VSV G pseudotyped vectors

A limitation in the use of retroviral vectors for gene transfer is their restricted host cell range. This can be altered through the production of pseudotypes. For example, MoMLV based vectors bearing the Env protein of GALV are able to infect bovine and hamster cells (Miller *et al.*, 1991). Another advantage in the use of vectors pseudotyped with a foreign envelope protein is that it reduces the risk of recombination between packaging constructs. Since retroviruses have also been shown to be able to form pseudotypes with paramyxoviruses, orthomyxoviruses, togaviruses and rhabdoviruses (Dong *et al.*, 1992; Závada, 1982), any one of these could be chosen for the production of a pseudotyped vector. However, both togaviruses and paramyxoviruses require two proteins for membrane binding and fusion. Also the incorporation of influenza virus HA into RSV particles is increased by approximately 70% in the presence of NA, an enzyme normally found in the envelope of influenza virions (Dong *et al.*, 1992). Rhabdoviruses such as VSV possess only one envelope protein, the glycoprotein G (Wagner, 1990). Emi *et al.* (1991) showed that this protein can be incorporated into MoMLV-based vectors in the absence of any other VSV proteins. These vectors were infectious and could be neutralised by anti-VSV antibodies.

A major advance in retroviral vector technology was the production of high titres of retroviral vectors pseudotyped with this VSV envelope glycoprotein (Burns *et al.*, 1993). These vectors were produced in 293 cells stably expressing the MoMLV *gag* and *pol* genes, which were further transfected with a vector encoding VSV G. The high titres (>10^9cfu/ml) that can be achieved with these pseudotypes are mainly due to the fact that they can be concentrated by ultracentrifugation. Attempts to concentrate retroviruses by such means results in a significant loss of infectivity. This is thought to be due to the instability of the Env protein (Burns *et al.*, 1993; Yee *et al.*, 1994). Since the SU protein is

noncovalently attached to TM it is prone to detachment. The VSV G protein is a homotrimer (Doms *et al.*, 1987; Doms *et al.*, 1988) and, therefore, more resistant to the shearing forces met during ultracentrifugation. This ability to concentrate vector preparations is advantageous for many procedures requiring the infection of large numbers of cells, including *in vivo* gene therapy.

VSV enters cells via receptor-mediated endocytosis. The reduction in pH within the endosome triggers a conformational change in the envelope glycoprotein, resulting in membrane fusion (Wagner, 1990). As discussed in section 1.2.2 the G protein is thought to interact with a phospholipid constituent of the cell membrane (Schlegel *et al.*, 1983; Superti *et al.*, 1984b). As a result the virus has a very broad host range, and this is conferred to the pseudotyped vectors (Burns *et al.*, 1993; Yee *et al.*, 1994).

Unfortunately VSV G is toxic to cells and high level expression causes cell fusion (Burns *et al.*, 1993). Stable expression of G has been achieved, but only at levels too low for efficient vector production. Florkiewicz *et al.* (1983) isolated a murine C127 cell line in which more than 95% of the cells were found to be positive for G expression using immunofluorescence. However, there was a cell to cell variation in expression of at least 30 fold. Extra G-related proteins were also produced by these cells, possibly as a result of aberrant splicing of the mRNA. Madin-Darby canine kidney (MDCK) cells stably expressing G have also been obtained (Roman *et al.*, 1988). The proportion of positive cells was enriched from 2 to 74% through the use of immunoselection. However, after 5 to 10 passages this level decreased to between 13 and 20%, probably because of the toxicity of the G glycoprotein. Burns *et al.* produced their vectors by making use of a period of high level expression of VSV G, before accumulation of the protein at the cell surface led to syncytium formation and cell death. Since high titres are produced this method is sufficient for many experimental techniques. Nonetheless, stable expression would be preferred for the large scale production of vectors for gene therapy.

3.1.2 Rabies G

It seems possible that a G protein from another rhabdovirus may also be suitable for the production of high titre pseudotyped vector. Like VSV, rabies virus has a very wide host

range. Although it has a strong neurotropism *in vivo*, it can infect a large variety of cells *in vitro*, including mammalian, avian, fish, reptile and insect cells (Seganti *et al.*, 1990). VSV and rabies virus belong to two different rhabdovirus genera, the *Vesiculoviruses* and the *Lyssaviruses* respectively. Viruses from both genera infect vertebrates. The *Vesiculoviruses* also infect insects which may be involved in infection transmission. These viruses infect epithelial cells causing vesicular disease. The *Lyssaviruses* infect the central nervous system causing fatal encephalitis (Baer *et al.*, 1990; Wagner, 1990). *Vesiculovirus* infection is characterised by a rapid course, cytopathology, a high density of viral budding from plasma membranes, and extracellular accumulation of virus. In contrast, *Lyssavirus* infection follows a slower course. There is a progressive accumulation of inclusion bodies in infected cells, and viral budding occurs from internal membranes as well as the plasma membrane. Laboratory adapted strains of *Lyssaviruses* appear to bud mainly from the plasma membrane (Murphy and Harrison, 1980). Despite these differences, VSV and rabies virus show many similarities in morphology and genomic organisation. The two G proteins are approximately the same size; rabies G consists of 524 amino acids, and VSV G of 511. They have an overall sequence homology of 20%, with some regions, including a glycosylation site, having more than 50%. In addition, 8 out of the 15 cysteine residues present in VSV G align with 8 of the 17 in rabies G. These similarities suggest that the two glycoproteins have similar secondary structures (Rose *et al.*, 1982).

Rabies G has a slower rate of transport through the cell than VSV G. This can be measured by the time taken for acquisition of endoglycosidase H-resistant sugars. The presence of these sugars indicates arrival to the medial cisternae of the Golgi complex (Dunphy and Rothman, 1985). Rabies G has a half-time of approximately 50 minutes (Whitt *et al.*, 1991), which compares to 15-20 minutes for VSV G (Rose and Bergmann, 1983).

An advantage of the rabies G protein is that it has a lower pH threshold for fusion activity. HeLa cells transiently expressing VSV G have been shown to form syncytia at pH 6.3 and below, while those expressing rabies G did not form syncytia above pH 6.1 (Whitt *et al.*, 1991). In addition, hemolysis of human erythrocytes by rabies virus almost completely disappears above pH 6.2, while a low level can be detected with VSV even at neutral pH (Mifune *et al.*, 1982). This property suggests that rabies G may be more suitable for long term expression in cell culture. Stable expression of this protein has been reported. Burger

et al. (1991) have achieved expression in Chinese hamster ovary (CHO) cells without any observed cytopathogenicity. The level of G expression remained stable for over 6 months in culture.

3.1.3 Structure of rabies G

The rabies G glycoprotein is also believed to form homotrimers on the surface of mature virions. Although these trimers cannot be purified under conditions used for VSV G, they can be cross-linked with reagents such as dithiobis-succinimidyl propionate (DSP, Whitt *et al.*, 1991). Further evidence for such a structure comes from electron microscopy of negatively stained rabies virions (Gaudin *et al.*, 1992). In micrographs the G protein appears as a head on a thin stalk, approximately 8.3nm in length. In an end on view the proteins look like triangles containing three dots, which is highly suggestive of a trimeric configuration. These trimers are not very stable. Unlike VSV G they are not stabilised by low pH, and they seem to be sensitive to most detergents apart from 1%CHAPS, a zwitterionic detergent (Gaudin *et al.*, 1992; Whitt *et al.*, 1991).

The initial 19 amino-terminal residues of the transcribed protein, which are predominantly hydrophobic, act as a signal peptide (Anilionis *et al.*, 1981). Although the site of proteolytic cleavage of this peptide aligns with that of VSV, the two sequences are not very homologous. It seems likely that the main constraint is maintenance of hydrophobicity (Rose *et al.*, 1982). Another stretch of 22 hydrophobic amino acids near the carboxy-terminus corresponds to the transmembrane domain. This leaves a sequence of 44 residues to form the cytoplasmic tail (Anilionis *et al.*, 1981; Conzelmann *et al.*, 1990). These two regions are highly conserved among some, though not all, strains of rabies virus (Conzelmann *et al.*, 1990).

The rabies G protein is palmitoylated (Gaudin *et al.*, 1991b). This involves the post translational addition of long chain fatty acids, and it is frequently observed with the glycoproteins of enveloped RNA viruses. These additions are usually linked to threonine, serine or cysteine residues on the cytoplasmic face of membrane binding or transmembrane domains. Gaudin *et al.* believe that cysteine 460 and/or 461 are the linkage sites for the palmitoylation of rabies G. As would be expected these residues are located at the

cytoplasmic end of the transmembrane domain. The function of palmitoylation is not yet known, although roles in membrane fusion, and the assembly and budding of viral particles have been suggested.

There are three potential sites for N-linked glycosylation within the ectodomain of G (Anilionis *et al.*, 1981; Conzelmann *et al.*, 1990; Shakin-Eshleman *et al.*, 1992, Fig. 3.1). This type of modification involves the cotranslational addition of oligosaccharides to asparagine residues by enzymes present in the lumen of the RER. The recognition sequence is Asn-X-Ser/Thr, where X is any amino acid apart from proline or tryptophan (Myers, 1997; Rademacher *et al.*, 1988). N-linked glycosylation appears to be important in intracellular transport, immunogenicity, stability and biological activity (Shakin-Eshleman *et al.*, 1992). Inhibition by tunicamycin prevents expression of G on the cell surface (Burger *et al.*, 1991). It may act through affecting the structure of the protein, providing hydrophilicity which helps to stabilise correct folding. The N-linked oligosaccharide groups may also function in interactions with calnexin, a molecular chaperone involved in the folding process (Gaudin, 1997). Studies have suggested that two forms of G are present in intact virions of some strains. These forms are glycosylated at one or two of the three potential sites (Wunner *et al.*, 1985). It has been suggested that the site at Asn 37 is not accessible to glycosylation as it is within a relatively hydrophobic region of the protein. However, it has been shown that this site can also be glycosylated although inefficiently. This alone is sufficient to support cell surface expression (Shakin-Eshleman *et al.*, 1992). Mutational analysis of this site has revealed that the amino acid following the recognition sequence can affect the efficiency of glycosylation, positively charged residues generally being the most favourable (Mellquist *et al.*, 1998). This is less important when the Ser residue is replaced with Thr.

The patterns of cross-neutralisation by monoclonal antibodies (mAbs) with selected G protein mutants have been used to define antigenic sites (Lafon *et al.*, 1983). The majority of these mAbs recognise sites II and III, shown in Fig. 3.1 (Benmansour *et al.*, 1991). The main antigenic site, site II, is discontinuous, located between residues 34-42 and 198-200 (Prehaud *et al.*, 1988). Since most antibodies can be used to select mutants in both regions, it seems very likely that they are positioned close together in the mature protein. Site III is linear and positioned between residues 330-340 (Seif *et al.*, 1985). Escape mutants

signal peptide

-19 **MVPQALLFV PLLVFPLCFG** KFPIYTIPDK LGPWSPIDIH HLSCPNNLVV

site II

31 EDEGCTNLSG FSYMELKVGY ILAIKVNGFT CTGVVTEAET YTNFVGYVTT

*

section containing putative fusion peptide

81 TFK**RKHFRPT PDACRAAYNW KMAGDPRYEE SLHNPYPDYR WLRTVKTTKE**

131 **SLVIISPSVA DLDPYDRSLH SRVFPSGKCS** GVAVSSTYCS TNHD**YTIWMP**

predicted nAchR binding site

181 **ENPRLGMSCD IFTNSRGKRA SKG**SETCGFV DERGLYKSLK GACKLKLCGV

site II

231 LGLRLMDGTW VSMQTSNETK WCPPDKLVNL HDFRSDEIEH LVVEELVRKR

*

281 EECLDALESI MTTKSVSFRR LSHLRKLVPG FGKAYTIFNK TLMEADAHYK

*

site III ?fusion site

331 SVRTWNEILP SKGCLRVGGR CHPHVNGVF**F NGIILG**PDGN VLIP**EMQSSL**

α-helix α-helix

381 **LQQHMELLES SVI**PLVHP**LA DPSTVFKDGD EAEDF**VEVHL PDVHNQVSGV

TM region

431 DLGLPNWGK**Y VLLSAGALTA LMLIIFLMTC C**RRVNRSEPT QHNLRGTGRE

481 VSVTPQSGKI ISSWESHKSG GETRL

Fig. 3.1. Amino acid sequence of SAD B19 rabies G (Genbank accession number M31046, Conzelmann *et al.*, 1990). Features mentioned in the text are in bold type. The positions of the HP hinge and antigenic sites are underlined. *, potential site for N-linked glycosylation.

involving substitution of Arg 333 have been shown to be nonpathogenic, and display a different host cell range to the wild type.

The high conservation of cysteine residues and glycosylation sites between rhabdovirus glycoproteins implies that these proteins have similar structures. The crystal structure of the protein has not yet been obtained, although a hypothetical structural model of the glycoprotein head region has been proposed (Walker and Kongsuwan, 1999). This model is based on the pattern of conserved cysteine residues, positioning of antigenic sites, predicted secondary structural features, and experimental data such as those from cyanogen bromide (CNBr) cleavage studies. This head region is believed to contain the domains involved in attachment and membrane fusion. Antigenic sites II and III have been predicted to interact with the receptor(s) and are thought to be positioned closely within the glycoprotein head (Gaudin, 1997). This model brings the two identified elements of site II into close alignment.

3.1.4 Cell attachment

As discussed in section 1.2.2, the receptor usage of rabies virus is not yet fully understood. Like VSV, rabies virus is also able to infect a wide variety of cells *in vitro*. This is again thought to be due to an interaction between the G protein and a phospholipid component of the cell membrane. In addition, both the nAchR and NCAM have been proposed as neural specific receptors. The receptor binding regions of G have not been defined, although a region of homology has been identified between rabies G and the snake venom curaremimetic neurotoxins. These toxins bind with high affinity to nAchR. Conserved residues that are thought to be involved in this binding are located within the second loop of the neurotoxin peptides. These residues are also conserved in the region between residues 175 to 203 in rabies G, shown in Fig. 3.1 (Lentz *et al.*, 1984). This indicates that this region may function as a nAchR binding site. A section of the discontinuous antigenic site II is within this putative binding region. The majority of neutralising mAbs are directed against this site (Benmansour *et al.*, 1991). Some site II escape mutants, especially those with a substitution of residue 198, are less pathogenic (Prehaud *et al.*, 1988). This residue is one of the few that are conserved between all of the neurotoxins as well as G.

Donnelly-Roberts and Lentz (1991) found that two polyclonal antibodies raised against a synthetic peptide representing this region of homology showed some cross reactivity with αBtx and native G. This suggests that the two putative binding sites have similar conformations. Bracci *et al.* (1988) found a similar result in a separate study. Although they were unable to demonstrate binding of their peptide to the nAchR, certain mAbs raised against it were able to recognise, and inhibit the binding of both rabies G and αBtx. Such structural similarity was also predicted by circular dichroism (CD) spectroscopy of peptides corresponding to the two sites (Donnelly-Roberts and Lentz, 1991). The major structure predicted was a β-sheet, which is in agreement with studies undertaken with intact toxin proteins. Rustici *et al.* (1993) suggest that the NSRG sequence within this region of the G protein (Fig. 3.1) is responsible for binding to nAchR. This sequence was common between a set of peptides which were recognised by monoclonal antibodies, along with neurotoxins, native G and acetylcholine (Ach). It corresponds to a peak in the hydrophilicity of G, and is therefore likely to be located on the protein surface (Bracci *et al.*, 1988). Rustici *et al.* propose that the side chains of the Arg196 and Asn194 residues mimic the structure of Ach, in a manner similar to that suggested for the Asp and Arg residues of the binding site of the snake neurotoxins.

3.1.5 Membrane fusion

Like VSV, rabies virus is believed to enter cells by adsorptive endocytosis. Incubation of the virus with human erythrocytes results in high hemolysis and cell fusion at low pH (Mifune *et al.*, 1982). Also, the lysosomotropic agents chloroquine and ammonium chloride will inhibit infection of CER cells (Superti *et al.*, 1984a), and rabies virus has been shown to colocalise with endosome tracers and an acidic probe shortly after virus adsorption to IMR-32 human neuroblastoma cells (Lewis *et al.*, 1998). Adsorptive endocytosis involves internalisation of the virus in clathrin-coated pits that form intracellular vesicles. These then fuse with endosomes to create an acidic interior. The acidic environment produces a conformational change within the G protein promoting membrane fusion, and delivery of the virus into the cytoplasm (Gaudin *et al.*, 1991a; Gaudin *et al.*, 1993).

As has been described in section 1.3.2, the rabies G protein can adopt three conformational states: native (N), active (A) and inactive (I). The transition between these states is believed to be necessary for the fusion process, and therefore protein function. These transitions can be studied through the characterisation of mutants that escape neutralisation by antibodies specifically targeted to a single conformational state of G. Mutations affecting amino acids 10, 13 and 15 have been found to alter the epitope targeted by a group of antibodies that specifically recognise the I state (Gaudin *et al.*, 1996). This epitope is not accessible in the N state on the virion surface, but becomes so when the molecule is solubilised. This suggests that it is situated under the head of the structure. The same study found that mutations in amino acids 44 and 282 alter the kinetics of the change to I, and in 392 and 396 affect the thermodynamic equilibrium between I and N. Residues 392 and 396 are close to His 397 and Pro 398 which are highly conserved among rhabdoviruses. These residues are located between two putative amphipathic α-helices and may act as a hinge during the conformational changes. The α-helices are located downstream of a highly conserved glycine residue which is though to mark the start of the stalk region. Specific mutations within the equivalent α-helical region of VSV G influence the fusogenic activity of the protein, possibly by interfering with the necessary conformational changes (Li *et al.*, 1993; Shokralla *et al.*, 1998).

Neither VSV G nor rabies G possess a typical fusion peptide. A putative internal fusion peptide has been identified between residues 123 and 137 of VSV G (Zhang and Ghosh, 1994). This peptide is not hydrophobic, but contains a sequence of uncharged amino acids. It is highly conserved between vesiculoviruses. Mutations within this region can abolish fusogenicity or reduce the pH threshold for fusion (Fredericksen and Whitt, 1995; Fredericksen and Whitt, 1996; Li *et al.*, 1993; Zhang and Ghosh, 1994). Certain mutations of the acidic residues Asp 137 and Glu 139 at the carboxy-terminal end of the fusion domain will lower the pH threshold. Fredericksen and Whitt (1996) suggest that these mutations increase the level of protonation required to induce conformational changes, and thereby stabilise the fusion-inactive conformation of the protein. The Glu 139 residue is conserved in lyssaviruses where it may act in a similar fashion.

Cleverley and Lenard (1998) have shown that a Gly residue within the transmembrane domain of VSV G is required for fusion. This amino acid is found in many other viral

fusion proteins, including rabies G. The authors suggest that the Gly may allow the protein to bend during the fusion process, thereby promoting the formation or enlargement of a fusion pore.

A region of rabies G considered to be a putative fusion domain is situated between residues 84 and 160 (Durrer *et al.*, 1995). This region was identified by photolabeling of exposed hydrophobic domains following incubation of whole protein or CNBr fragments with unilamellar vesicles at low pH. Part of this region does align with the predicted fusion peptide of VSV G, although there is little sequence homology. A higher degree of homology is found between Val 125 and Trp 178. The sequence between residues 84 and 160 does not contain an extended hydrophobic domain. Durrer *et al.* suggest that the region which interacts with the target membrane may adopt a sided α-helix or β structure. The sequence between positions 360 and 366 of rabies G (FNGIILG) shows a high degree of homology with the N-terminals of paramyxovirus F_1 and HIV-1 gp41. It has therefore been suggested that this region may also be involved in membrane fusion (Morimoto *et al.*, 1992).

3.1.6 Rabies G pseudotyped vectors

The properties that rabies G shares with VSV G suggest that this glycoprotein could also be useful for producing pseudotyped vectors for gene delivery. Both proteins are homotrimers, indicating that they are robust structures. This property of VSV G allows concentration of pseudotyped vectors to the high titres required for many applications. Both proteins are also able to mediate entry into a wide range of host cells. This allows infection of many cell types not normally susceptible to retroviral infection. The fact that rabies enters cells via a different route to the majority of retroviruses is unlikely to pose a problem. It has already been shown that VSV G pseudotyped retroviral vectors can survive the acidic conditions encountered during endocytosis.

The experiments described in this chapter were carried out in order to test the suitability of rabies G for vector production. The availability of efficient transient packaging systems such as the HIT system (Soneoka *et al.*, 1995) allows rapid analysis of these pseudotyped vectors. In such a way, a proposed envelope protein can be tested without the need for

stable packaging cell lines. If experiments are successful these cell lines could then be produced.

3.2 Materials and methods

3.2.1 Cloning of the rabies G glycoprotein

The plasmid pT7T-G (Conzelmann and Schnell, 1994), encoding the rabies virus SAD B19 G glycoprotein (nucleotides 3311-5138), was a gift from Karl-Klaus Conzelmann (FRCVDA, Tübingen). The rabies G gene was subcloned into the eukaryotic expression vector, pcDNA3. pT7T-G and pcDNA3 were both digested with Eco RI (Gibco, UK) at 37°C for 2 hours, and the relevant fragments were then ligated as described in section 2.2.3.1. The ligation reaction was used to transform DH5α cells as described previously (section 2.2.3.2). The clone obtained was designated pcDNA3.G. A second clone, containing G in the reverse orientation was also acquired (pcDNA3.GR). The orientation of the rabies G gene in these clones was verified by Hind III (Gibco, UK) digestion.

3.2.2 Cloning of pcDNA3.CAT

The plasmid pUC119.CAT was provided by Takayuki Miyazawa (Veterinary Pathology, University of Glasgow). This construct contained the chloramphenicol acetyltransferase (CAT) gene cloned into the plasmid pUC119, between the Bam HI and Hind III restriction sites. pUC119.CAT and pcDNA3 were digested with Bam HI and Hind III (Gibco, UK) and the relevant fragments were then ligated as described previously (section 2.2.3.1). The ligation reactions were used to transform DH5α cells (section 2.2.3.2) and DNA was purified from the resulting colonies. The orientation of the CAT gene in the final clones was verified by Eco RI digestion.

3.2.3 Cloning of a chimaeric rabies G glycoprotein

The construct pcDNA3.GMAenv was produced by replacing the cytoplasmic domain of rabies G with that of amphotropic MLV Env. pHIT456 and pcDNA3.G were both digested with the restriction enzymes Xba I (Gibco, UK) and Cla I (Gibco, UK). Prior to digestion pcDNA3.G was grown up in the *dam*$^-$ cells DM1 (Gibco, UK) in order to avoid methylation of the Cla I recognition site. The method for transformation of these cells

differs slightly from DH5α cells. 0.5ng of plasmid was mixed with 100µl of cells and incubated on ice for 30min. The cells were then heat shocked at 42°C for 45s and incubated on ice for a further 2min. 900µl of SOC medium was added and the cells were incubated at 37°C for 1 hour in an orbital incubator. The cultures were then spread onto LB agar plates containing 100µg/ml Amp and incubated at 37°C overnight. Amp resistant colonies were then grown up in LB broth containing Amp, as described for DH5α cells (section 2.2.3.2). The digests were gel purified (section 2.2.2.5) and the relevant fragments were ligated using T4 DNA ligase at 14°C overnight. The ligation reaction was then used to transform DH5α cells. The orientation of the insert was tested by Bam HI (Gibco, UK) digestion and the sequence was verified by automated sequencing (section 5.2.2.2), using primers targeted to the T7 (sense) and SP6 (antisense) promoters.

3.2.4 CAT enzyme linked immunosorbent assay (ELISA)

CAT expression was detected using a commercial CAT ELISA kit (Boehringer Mannheim), according to the manufacturer's instructions. Cell lysates were diluted with sample buffer (PBS and blocking reagents) to give a protein concentration of 250µg/ml. Concentrations were determined using Bradford reagent (Sigma, UK). The lysates were then added to a microtiter plate provided with the kit. This had been precoated with sheep polyclonal anti-CAT antibody. The wells were washed with PBS/Tween 20 and anti-CAT-digoxigenin was added. After incubating at 37°C for 1 hour, the wells were washed and anti-digoxigenin-peroxidase was added. All antibodies were supplied with the kit. Following incubation at 37°C for another hour the wells were washed and the peroxidase substrate was added for 15min until a green colour developed. Absorbance was measured at 405nm.

3.2.5 Immunofluorescence (IF) staining

Glass coverslips were sterilised in 100% ethanol and then coated with 1mg/ml poly-L-lysine (Sigma, UK) in 0.1M borate buffer, pH 8.5. This was done by dropping the solution onto the coverslip, removing excess fluid, and leaving to dry overnight in a laminar flow

hood. The coverslips were then washed twice for two hours with sterile dH_2O. 0.8-1 x 10^5 cells were then plated onto the coverslips in 6 well plates.

When the cells were ready for staining they were washed twice in warm PBS before fixing in 4% paraformaldehyde for 15min. They were then washed 4 times in PBS, permeabilised in 0.3% Triton X-100/PBS for 5min, washed 3 times in PBS, and incubated for 30min to 1 hour in a blocking solution consisting of 10% swine serum in PBS. After blocking the cells were washed once with 1% swine serum/PBS and then incubated for 1 hour with the primary antibody RV1C5 (anti-rabies virus, monoclonal mouse IgG2a, Immunogen International Ltd., UK) diluted 1/500 in1% swine serum/PBS. They were then washed 6 times and incubated for 1 hour with the secondary antibody, rabbit-anti-mouse TRITC (tetramethylrhodamine isothiocyanate, Sigma, UK) diluted 1/100 in 1% swine serum/PBS. Finally the cells were washed 6 times in PBS and then once in dH_2O. The coverslips were mounted on slides and viewed under a fluorescence microscope (40x objective).

3.2.6 Testing for cell fusion at low pH

Cos-7 cells were plated onto coverslips as described above and were transfected with either pcDNA3.G or with pcDNA3.GR using DEAE-dextran (section 2.2.1.3). 48 hours after transfection the cells were incubated for 1 min in PBS at various pH values (pH 7.0, 6.0, 5.0, 4.5), and then for 1 hour in 10% DMEM. All cells were then fixed and labelled as described above, except that 0.25µg/ml Hœchst stain was added with the secondary antibody. Photographs were taken through the 100x objective of the fluorescence microscope.

3.2.7 Radiolabelling of cells

Medium was removed from cells and replaced with Met-free Eagle's MEM (Gibco, UK). After 1 hour's incubation at 37°C this was replaced with 500µl of medium containing approximately 100µCi L-[^{35}S]Met (PRO-MIX^{TM}, Amersham Life Science, UK), and the cells were incubated for a further hour. The labelling medium was then removed and retained for analysis. Cells were lysed on ice with lysis buffer and the lysates were centrifuged for 10min to pellet debris.

3.2.8 Radio-immune precipitation assay (RIPA)

50µl of fixed *Staph. A.* (Sigma, UK), preloaded with normal serum was added to cell lysates and incubated for 1 hour. This was then pelleted in a centrifuge and the supernatant was transferred to a fresh tube. Antibody was added to the samples, both cell lysates and labelling media, and incubated at 4°C for 1 hour. Rabies G samples were incubated with RV1C5 (1/500), and the CAT samples with the anti-CAT-digoxigenin (1/250) from the ELISA kit. The samples were centrifuged for 10min to remove any aggregates that may have formed, and then incubated at 4°C for 1 hour with 10µl of either a 50% protein A/agarose (Sigma, UK) slurry or anti-digoxigenin-agarose (Sigma, UK). The complexes were pelleted by centrifugation at 3000 rpm for 1min, and then washed 4 times with lysis buffer and once with 50mM Tris pH 6.8 on ice. 50µl of Laemmli sample buffer was added and the samples were boiled for 5min. 10µl of each sample was run on a Tris/glycine SDS-polyacrylamide (SDS-PAGE) gel, according to the method of Laemmli (1970). The gel was then fixed in 25% isopropanol/10% acetic acid for 30min, soaked in fluorographic reagent (Amplify™, Amersham, UK) for 15min, vacuum dried and exposed to film (Amersham, UK) at -70°C.

3.2.9 Production of vectors

Vectors were produced using the HIT system (Soneoka *et al.*, 1995) as described in section 2.2.4.1. 293T cells in 25cm^2 flasks were transfected with pHIT60, pHIT111, and either pHIT456, pcDNA3.G or pcDNA3.GMAenv using calcium phosphate precipitation (section 2.2.1.3). One flask was transfected with pHIT111 alone as a negative control, and one with pMV12 as a control for transfection. Following transfection fresh DMEM containing 10mM HEPES pH 7.4 was added to the cells. Vector titres were determined by staining infected cells for β-gal activity, as described in section 2.2.4.2.

3.2.10 Optimisation of the HIT system

Several conditions were tested in order to optimise the production of pseudotyped vectors. The target cell types were BHK TK$^-$, AH927 and NIH 3T3. The medium conditions tested

were complete DMEM, serum free DMEM, serum free DMEM supplemented with SPIT (Sigma, UK), serum free DMEM containing 0.2% bovine serum albumin (BSA), and serum free DMEM supplemented with SPIT and 0.2% BSA. SPIT is a commercial medium supplement containing insulin from bovine pancreas, partially iron saturated human transferrin, sodium selenite and sodium pyruvate in Earle's Balanced Salt Solution. pcDNA3.G was added at 0.5μg, 5μg and 10μg per transfection. The effect of sodium butyrate (NaB) induction was tested by adding 10mM NaB to the culture medium after the end of transfection. Following incubation for 8 hours the cells were washed in PBS, and fresh complete DMEM was added. Incubation of the transfected cells at 32°C for 24 and 48 hours was also tested. Titrations of samples taken at 24, 48 and 72 hours after transfection were also carried out in the absence of 8μg/ml Polybrene. Vector production over time was observed by measuring titre at various points on a time scale. Samples were taken from cells at 0, 2, 4, 6, 24, 48 and 72 hours after the end of transfection.

3.2.11 Antibody neutralisation

Vectors bearing A-MLV Env, rabies G or GMAenv were produced by transient transfection of 293T cells as described above. 1ml of each supernatant was incubated with RV1C5 (1/250 dilution) at 37°C for 1 hour in an orbital incubator. 1ml samples were also incubated in the absence of antibody. Vectors were then titred on AH927 cells as before.

3.2.12 Titration of G418 sensitivity of 293 cells

Prior to transfection, 293 cells were tested for their sensitivity to G418. Three 96 well plates were set up with 0.4×10^4 cells in each well. The following day G418 was added to the cells at concentrations of 4mg/ml, 3mg/ml, 2mg/ml, 1.5mg/ml, 1mg/ml, 0.75mg/ml, 0.5mg/ml, 0.375mg/ml, 0.25mg/ml, 0.188mg/ml, 0.125mg/ml, and 0.094mg/ml. 8 wells of cells in a fourth plate were also set up in the absence of G418. After 4, 8, and 12 days cell viability was determined by MTT (thiazoyl blue, Sigma, UK) staining of the cells on one of the plates. MTT dissolved in DMEM was added to the wells to give a final concentration of 1 mg/ml, and the plate was incubated at 37°C for 3 hours. The medium was removed

and the purple formazan crystals were solubilised in 100% isopropanol for 15min. Absorbance was then measured at 570nm.

3.2.13 Cloning of a permanent rabies G expressing cell line

Approximately 1 x 10^5 293 cells were seeded onto 6cm dishes. 24 hours later the cells were transfected with pcDNA3.G using calcium phosphate precipitation (section 2.2.1.3). After 2 days 400µg/ml G418 was added to the growth medium and this medium was replaced every 3-4 days for a month, until selection of resistant colonies was possible.

3.2.14 RIPA of pelleted vector

A 75cm^2 flask of 293T cells was transfected with pHIT60 and pcDNA3.G using calcium phosphate precipitation (section 2.2.1.3). A second flask of cells was transfected with pHIT60 alone. The cells were then radiolabelled as described above (section 3.2.7), but were incubated in the labelling medium containing 100µCi/ml PRO-MIX™ overnight. This medium was then removed and centrifuged at 10 x g. The cells were lysed and immune precipitated as before. Vector was pelleted from the medium samples by centrifuging through a sucrose cushion. 1ml of 30% sucrose/TBS was pipetted into a 35ml clear polypropylene Beckman centrifuge tube. This was overlaid with 1ml of 20% sucrose/TBS and then the medium containing the vector. The remainder of the tube was filled with TBS. The samples were centrifuged at 35,000 rpm for 1 hour in a SW41 rotor in a Beckman L8-70M ultracentrifuge. The pellets were then resuspended in 250µl of lysis buffer and immune precipitated.

3.2.15 Time scale comparison of rabies G and Gag expression

293T cells were seeded onto 15 coverslips. 14 of these were transfected with pGT5, pHIT111 and pcDNA3.G. pGT5 is a FeLV based *gag/pol* construct and was a gift from John Oldfield (Veterinary Pathology, University of Glasgow). Cells were also seeded into one well of a six well plate and transfected with pMV12 as a transfection control. At 0, 2, 4, 6, 24, 48, and 72 hours after the end of transfection, medium was removed from 2 coverslips, filtered through a 0.45µm-pore filter, and stored at -70°C. The cells were then

analysed by IF staining as described above (section 3.2.5), although only one from each pair was permeabilised by incubation with 0.3% Triton X-100. The non-transfected sample was taken at 48 hours and was permeabilised. ELISA screening for Gag expression was carried out by the Feline Virus Unit, Veterinary Pathology, University of Glasgow.

3.3 Results

3.3.1 Transient expression of the rabies G glycoprotein

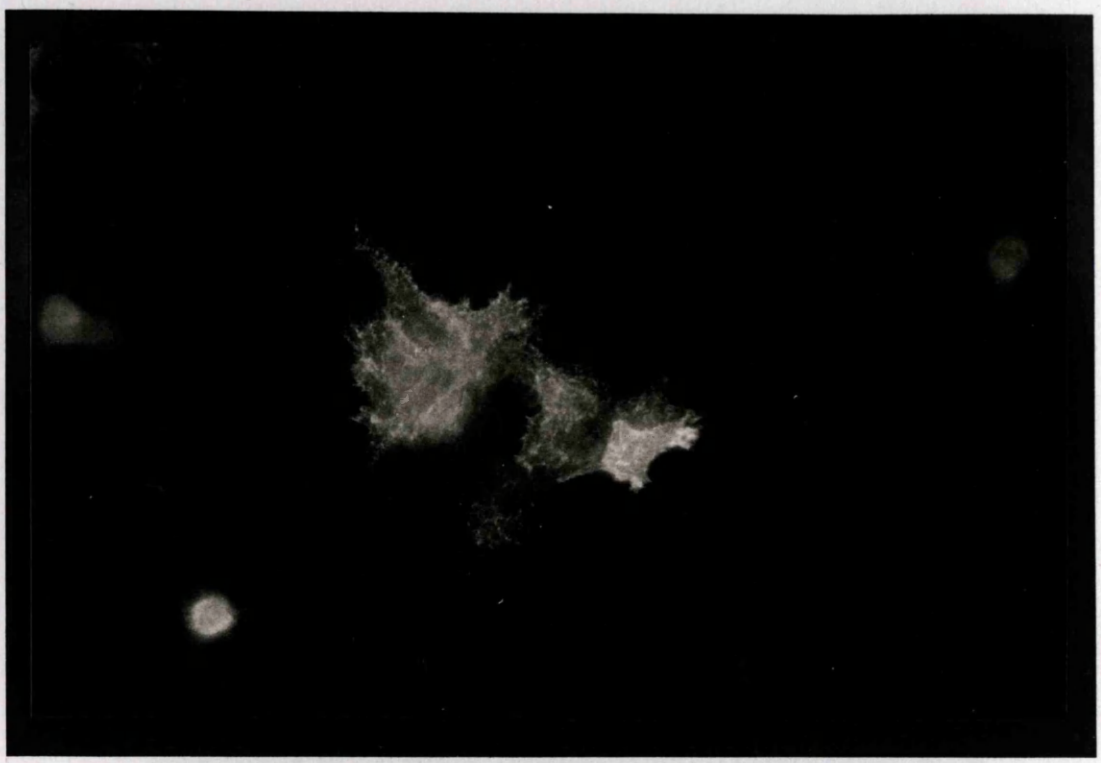
Expression of the G protein from pcDNA3.G was verified by transient transfection of Cos-7 cells. As shown in Fig. 3.2, IF staining detected expressed G protein in the transfected cells. The stronger staining intensity around the cell outline implies that G is localised on the cell surface, as would be expected for an envelope protein. No staining of cells transfected with pcDNA3.GR was observed, indicating that the positive signal is specific for G. Transfection efficiency was tested by transfecting cells in parallel with pcDNA3.CAT. The CAT expression level in these cells was consistently above the highest standard provided with the ELISA kit, indicating a concentration greater than 0.797ng/ml of protein. An example of the typical results obtained in the CAT ELISA is shown in Table 3.1.

Sample		**Absorbance (405nm)**	
Standard	0.797ng/ml	0.259	0.335
	0.399ng/ml	0.166	0.133
	0.200ng/ml	0.133	0.290
	0.100ng/ml	0.109	0.087
Cell lysate	24 hours	1.381	1.283
	48 hours	1.599	1.281
	72 hours	1.192	1.661

Table 3.1. Typical result obtained from ELISA of lystes taken from cells transfected with pcDNA3.CAT. Lysates were taken at 24, 48 and 72 hours after the end of transfection.

G protein expression was also detected by RIPA, as shown in Fig. 3.3. Cos-7 cell lysate samples were taken at 24, 48 and 72 hours after transfection and immune precipitated with the antibody RV1C5. A protein of approximately 67kD was detected in lysates taken from cells 24 and 48 hours after transfection with pcDNA3.G (lanes D and E), but not in those from cells transfected with pcDNA3.GR (lanes A, B and C) or pcDNA3.CAT (lane H), nor in non-transfected cells (lane G). This protein is the expected size for rabies G.

A



B

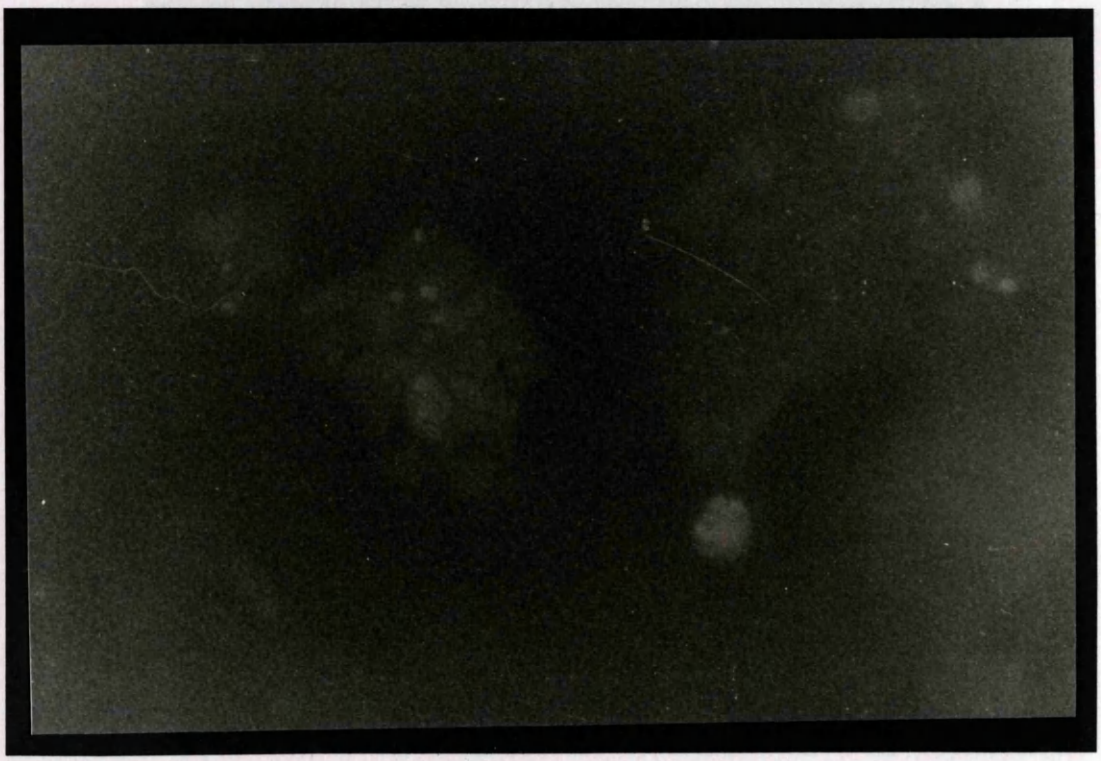


Fig. 3.2. IF staining of transfected Cos-7 cells using RV1C5 antibody. A) Cells transfected with pcDNA3.G. B) Cells transfected with pcDNA3.GR.

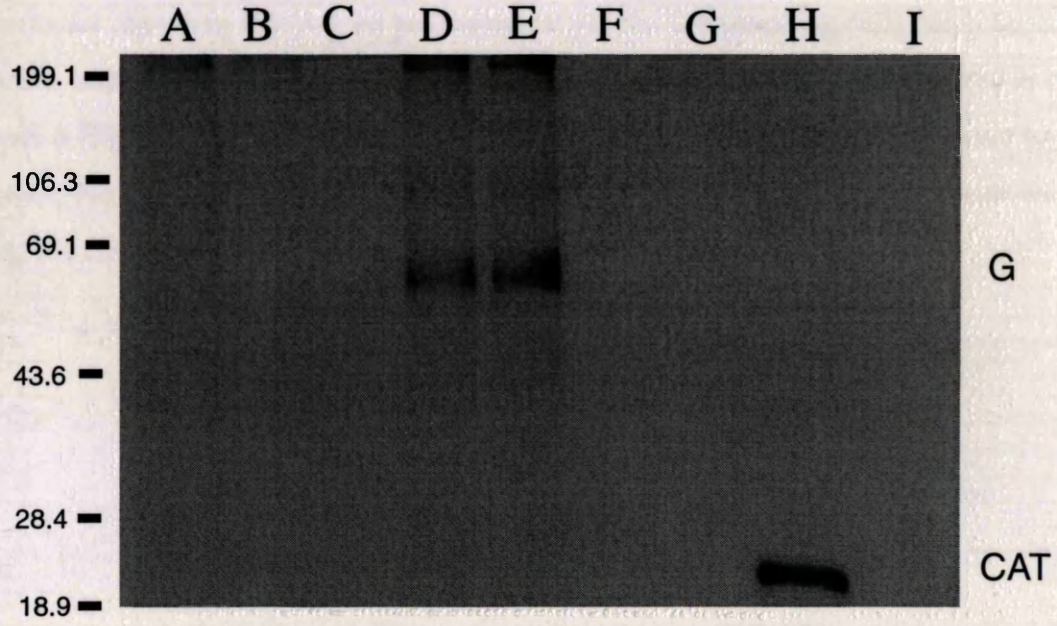


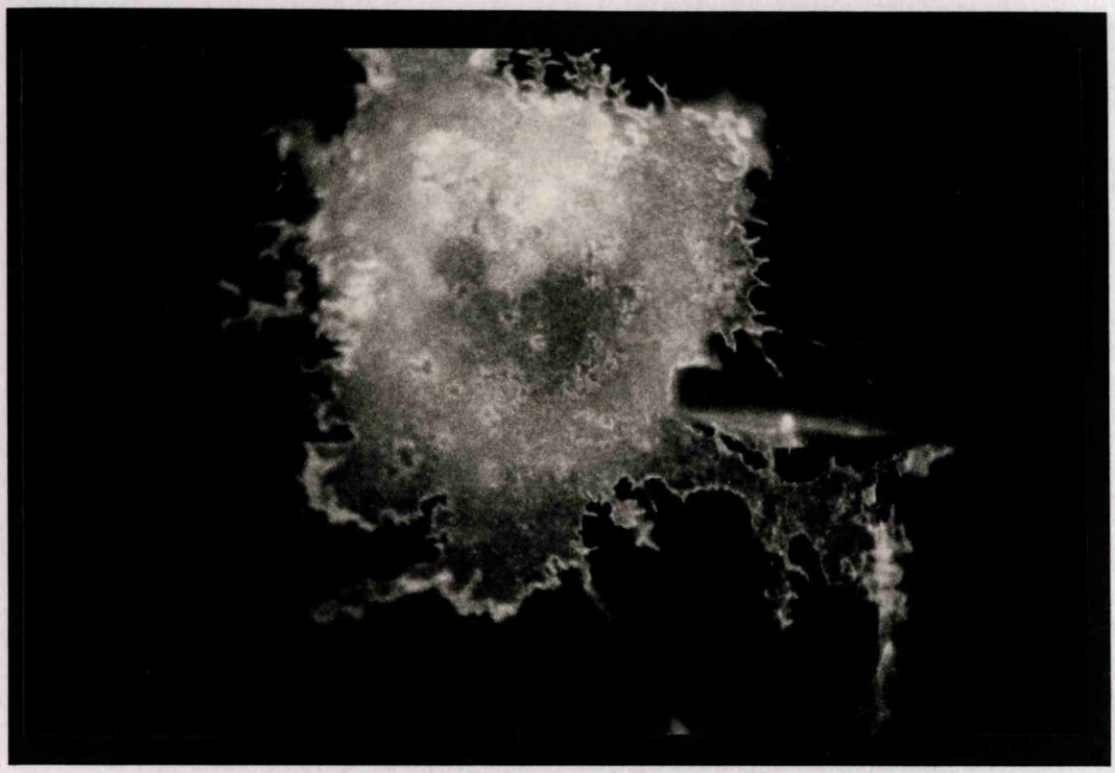
Fig. 3.3. RIPA of Cos-7 cells using RV1C5. A-C) Lysates of cells transfected with pcDNA3.GR, taken at 24, 48 and 72 hours. D-F) Lysates from cells transfected with pcDNA3.G, taken at 24, 48 and 72 hours. G) Lysate of non-transfected cells. H) Lysate of cells transfected with pcDNA3.CAT, precipitated with anti-CAT antibody. I) Medium sample taken at 24 hours

(Witkor *et al.*, 1973; Wunner *et al.*, 1985). G protein could not be detected in samples of culture media (lane I).

3.3.2 Fusion activity at low pH

In order to show that the G protein expressed on the cell surface was functional, transfected Cos-7 cells were incubated in a series of low pH buffers, and then analysed by IF staining. The results are shown in Fig. 3.4. At physiological pH the G-expressing cells could be seen as separate, intact cells (Fig. 3.4, A and B). Large, multinucleate cells were observed at and below pH 6 (Fig. 3.4, C-H), indicating syncytia formation. These cells were positive for G expression. No fusion of control cells transfected with pcDNA3.GR was observed under identical pH conditions.

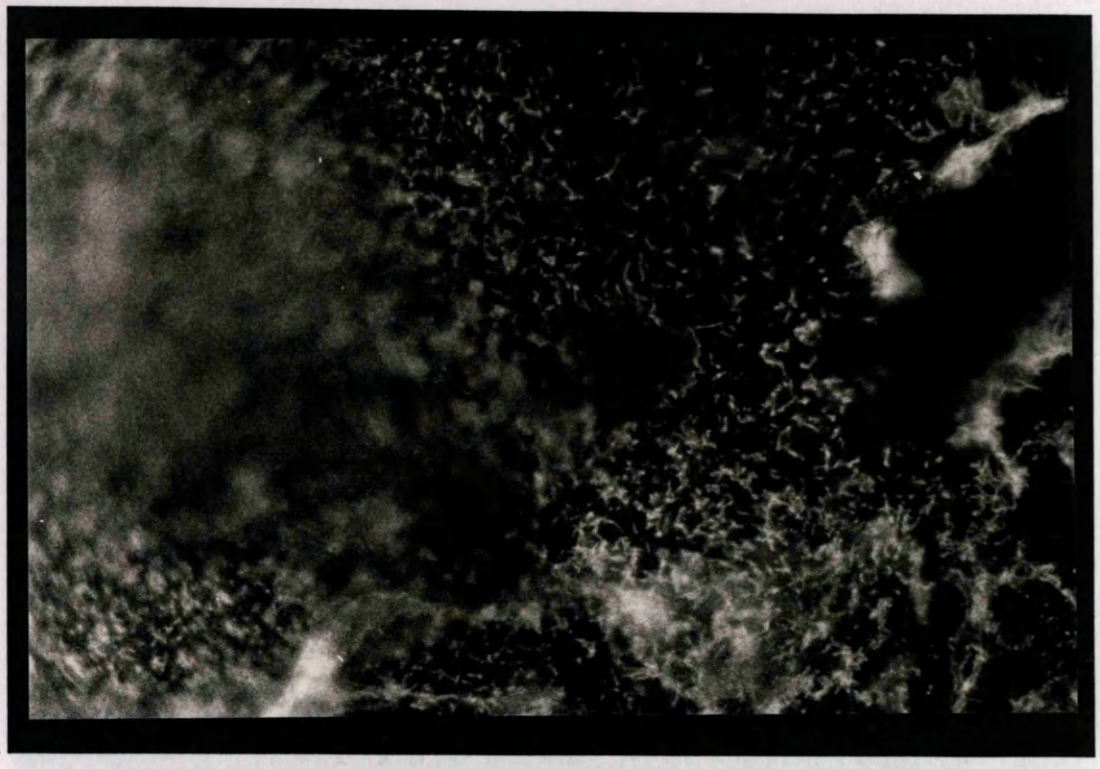
A



B

Fig. 3.4. Cos-7 cells transfected with pcDNA3.G and then incubated in a series of low pH buffers. A) IF staining of untreated cells. B) Hœscht staining of the same untreated cells.

C



D

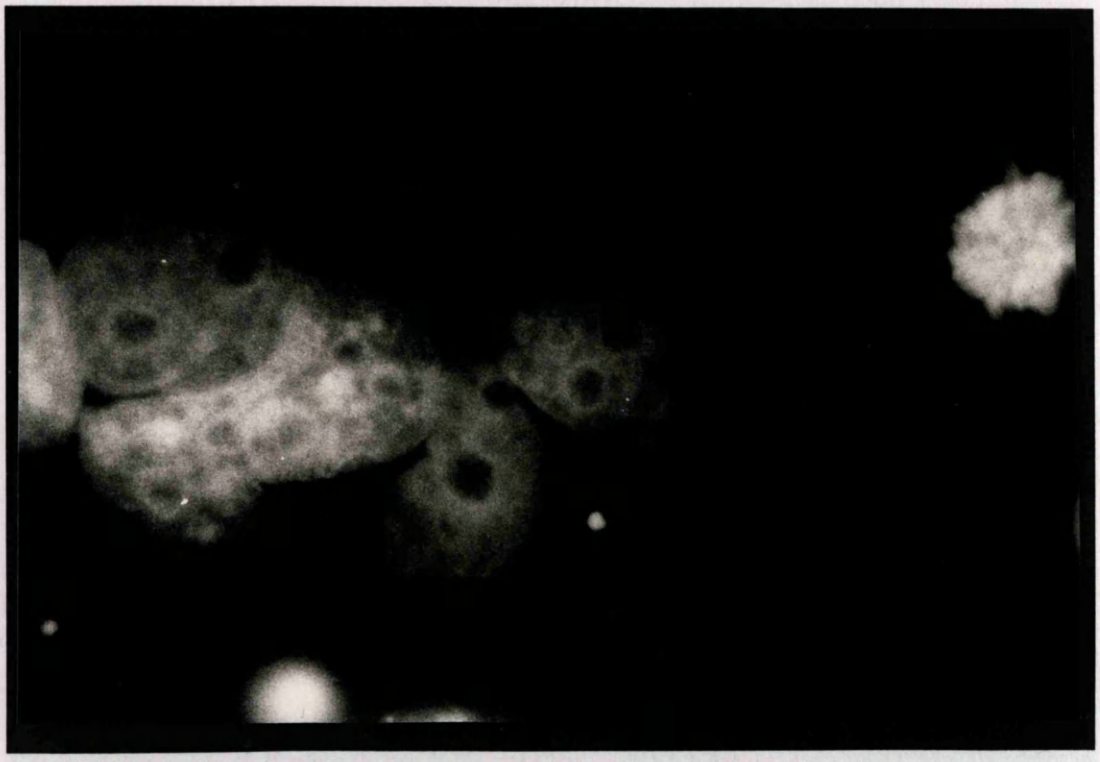


Fig. 3.4. Cont. C) IF staining of cells incubated at pH6. D) Hœscht staining of the same cells incubated at pH6.

E

F

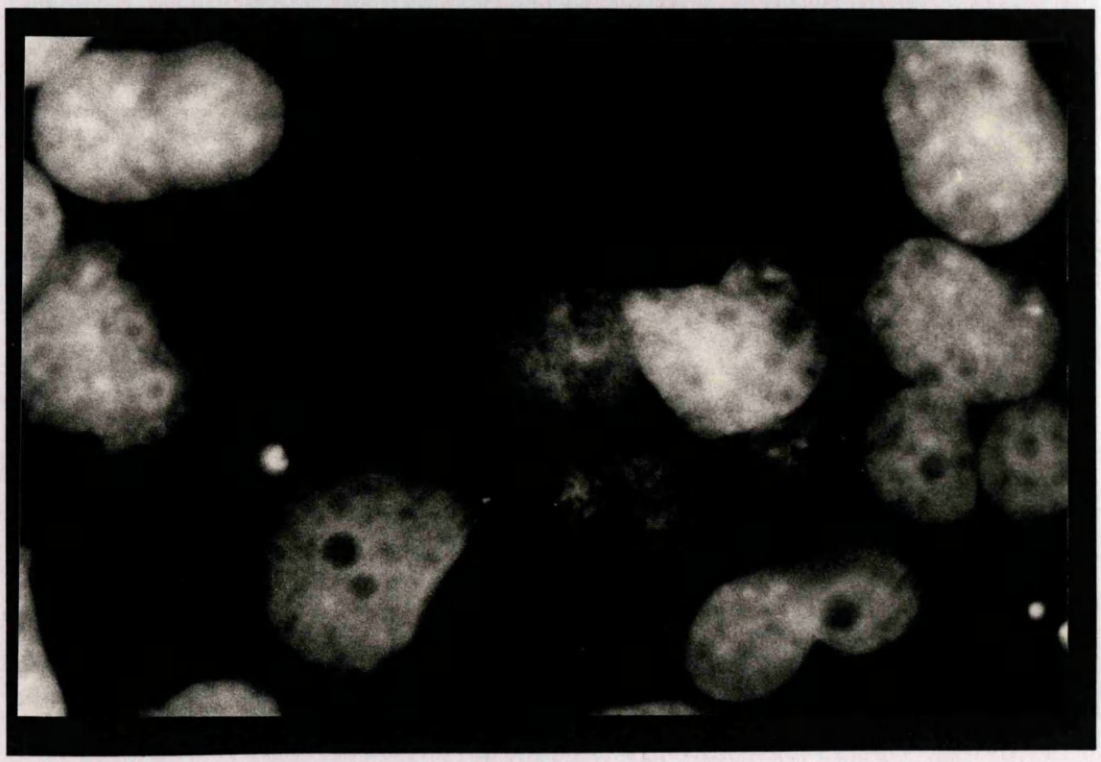
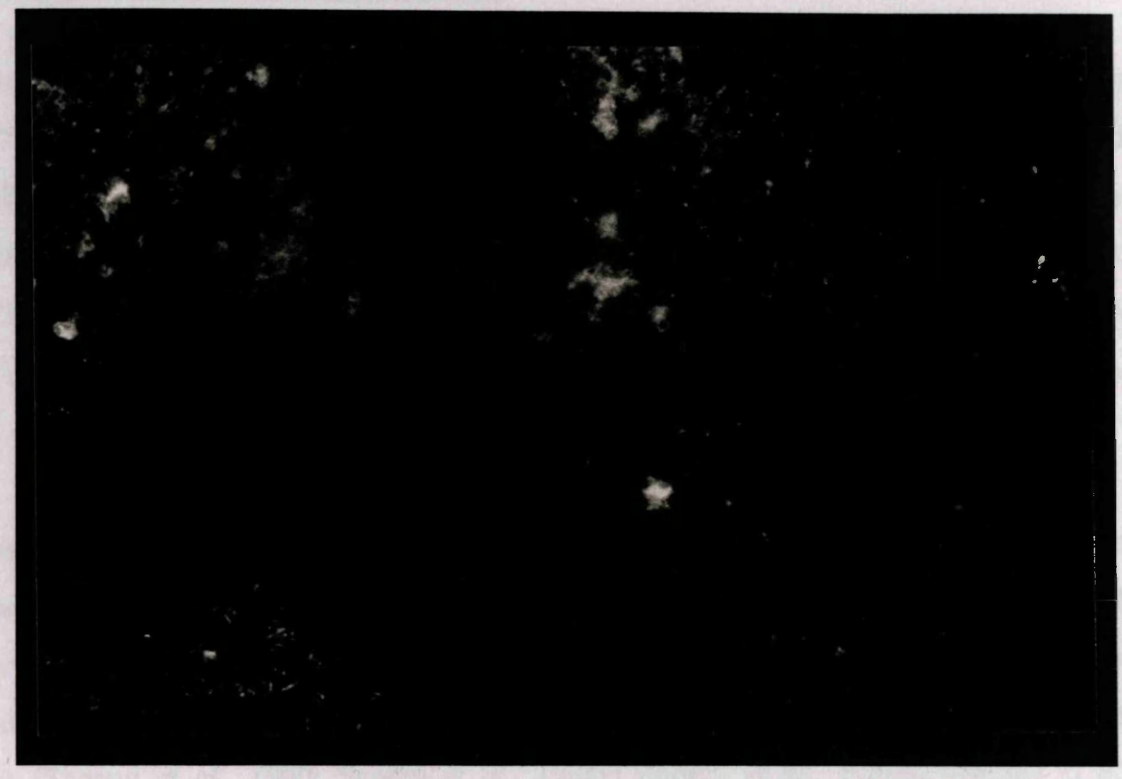


Fig. 3.4. Cont. E) IF staining of cells incubated at pH5. F) Hœscht staining of the same cells incubated at pH5.

G



H

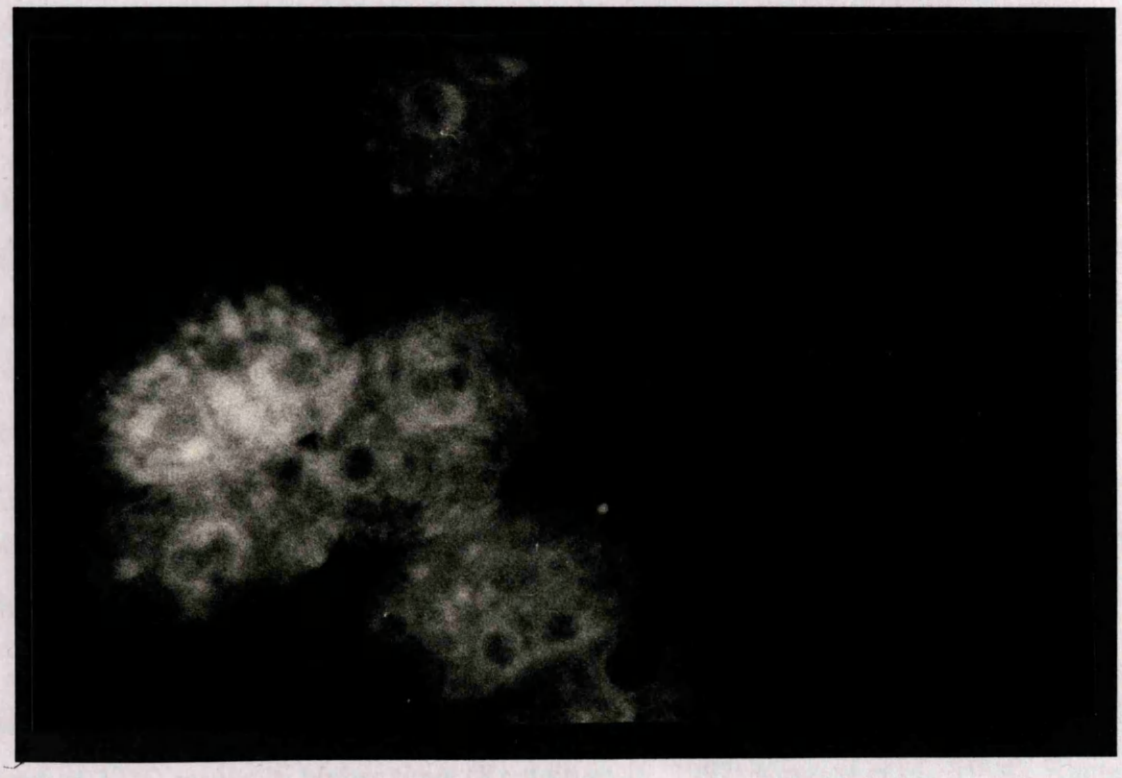


Fig. 3.4. Cont. G) IF staining of cells incubated at pH4.5. H) Hœscht staining of the same cells incubated at pH4.5.

3.3.3 Production of pseudotyped vectors

To test the ability of the rabies G protein to form pseudotypes with MLV-based vectors, 293T cells were co-transfected with 5μg each of pcDNA3.G and the HIT system plasmids pHIT60 and pHIT111, described in 2.1.4. Transfections were also carried out replacing pcDNA3.G with pHIT456, encoding the A-MLV *env*. Control cells were transfected with pHIT111 alone, and no titre was obtained with these in any experiment. Titres for the A-MLV vector were 1-2 x 10^4 blue cell colony forming units (bcfu)/ml on NIH 3T3 cells, and 0.4-1 x 10^5 bcfu/ml on AH927 cells.

In an attempt to optimise the titre of pseudotyped vector several experimental conditions were altered (Table 3.2). Of the three target cell types tested the feline fibroblast AH927 line was the most susceptible to infection by the pseudotyped vector. The replacement of serum with 0.2% BSA was tested as this is often used in studies involving the use of rabies virus (Seganti *et al.*, 1990; Superti *et al.*, 1984a; Superti *et al.*, 1984b; Wunner *et al.*, 1984). This change did result in a slightly higher titre, as did replacement with the SPIT supplement. Over the range tested, the optimal concentration of pcDNA3.G used in the transfection was 5μg per flask, equal to that of the cotransfected plasmids pHIT60 and pHIT111.

Treatment of the transfected cells with 10mM NaB has been reported to increase titres obtained using the HIT system (Soneoka *et al.*, 1995). NaB has been found to increase the proportion of cells expressing exogenous DNA (Gorman *et al.*, 1983), although the mechanism has not been fully defined. It has also been shown to activate certain eukaryotic promoters including the CMV promoter (Radsak *et al.*, 1989; Tanaka *et al.*, 1991). Therefore the effect of NaB induction on the titre of these pseudotyped vectors was tested. Cells were incubated in 10mM NaB for 8 hours after the end of transfection. However, no improvement in titre was detected.

It has been shown that harvesting of retroviral vector from packaging cell lines at 32°C instead of 37°C results in higher titres (Kaptein *et al.*, 1997; Kontani *et al.*, 1994). Kaptein *et al.* found that vector half-life was greatly increased at this lower temperature (37-39 hours compared to 9 hours), and that production rate was approximately 3 to 4 fold higher.

A

Condition	Target Cell Type	Titre (bcfu/ml)
standard (5µg pcDNA3.G, 10% FCS, 37°C, 24hr)	BHK TK$^-$	3.1×10^1
	AH927	7.0×10^1
	NIH 3T3	1.0×10^0
serum free medium	BHK TK$^-$	0
serum free + 0.2% BSA	BHK TK$^-$	7.1×10^1
serum free + SPIT	BHK TK$^-$	1.0×10^2
serum free + SPIT + BSA	BHK TK$^-$	4.4×10^1
0.5µg pcDNA3.G	BHK TK$^-$	1.0×10^0
10µg pcDNA3.G	BHK TK$^-$	1.3×10^1
sodium butyrate	BHK TK$^-$	1.2×10^1
32°C 24hr	BHK TK$^-$	4.0×10^0
48hr	BHK TK$^-$	1.0×10^0

B

Condition		Titre (bcfu/ml) 1 on BHK TK$^-$	Titre (bcfu/ml) 2 on BHK TK$^-$
Polybrene	24hr	3.3×10^1	2.7×10^1
	48hr	8.6×10^0	3.9×10^0
	72hr	0	0
no Polybrene	24hr	0.7×10^0	3.0×10^0
	48hr	4.8×10^0	9.0×10^0
	72hr	0	0

C

Time point	Titre (bcfu/ml) on BHK TK$^-$	Titre (bcfu/ml) on AH927
0-6hr	ND	0
24hr	3.1×10^1	7.0×10^1
48hr	7.0×10^0	1×10^1
72hr	0	2.5×10^0

Table 3.2. Optimisation of HIT system vectors pseudotyped with rabies G. A) Effects of target cell type, media composition, pcDNA3.G concentration, NaB, temperature of incubation of transfected cells. B) Effects of Polybrene (8µg/ml) and length of incubation. As duplicate titres were taken in this experiment both results are shown. C) Effect of length of incubation of transfected cells. (ND - not done). Unless otherwise stated, infections were carried out in the presence of 8µg/ml Polybrene).

They suggest that this higher production rate is due to more efficient protein folding, which results in increased surface expression of Env. In this case, the pseudotyped vector produced from cells incubated at 32°C had a lower titre than that from cells at 37°C at both time points tested.

The polycation Polybrene is used in the majority of retroviral vector protocols. Its presence improves efficiency of infection, probably by neutralising electrostatic repulsion between the virus and cell via an interaction with the cell membrane (Toyishama and Vogt, 1969). It is not believed to be able to mediate infection in the absence of an envelope-receptor interaction (Porter *et al.*, 1998). Since Burns *et al.* (1993) found that infection efficiency of VSV G pseudotyped vectors is optimal in the presence of 8μg/ml Polybrene this concentration was also used with these rabies G pseudotypes. As can be seen in Table 3.2B, the addition of Polybrene does improve titres on BHK TK^- cells.

Vector titres were taken at various time points after transfection from cultures incubated at 32°C or 37°C, in the presence or absence of Polybrene, and titrated on either BHK TK^- or AH927 cells. On each occasion, the highest titres were obtained from samples taken 24 hours after the end of transfection.

The highest titre obtained in these experiments was 1×10^2 bcfu/ml. This was on BHK TK^- cells in the presence of Polybrene, and was with vector produced by cells grown for 24 hours after the end of transfection, in serum free medium supplemented with SPIT. These conditions were therefore used in subsequent experiments. AH927 cells were used as target cells since a higher titre was obtained on these than on BHK TK^- cells. However, it is clear that none of the conditions tested produced a significant increase in the titre of the pseudotyped vector. The low titres are unlikely to be due to poor transfection efficiencies since parallel transfections with pMV12 showed a consistent efficiency of approximately 50 to 70%.

3.3.4 Specific neutralisation of vectors pseudotyped with rabies G

To ensure that the titres obtained with the pseudotyped vectors were due to the incorporation of rabies G, medium samples from transfected cells were incubated with the antibody RV1C5. After incubation for 1 hour the samples were titred on AH927 cells as before. As shown in Table 3.3, this antibody specifically neutralised the vector pseudotyped with G.

Vector	Conditions	Titre (bcfu/ml)
A-MLV	no incubation	2.2×10^5
	1hr at 37°C	2.7×10^5
	1hr at 37°C with RV1C5	2.5×10^5
G pseudotype	no incubation	2.6×10^1
	1hr at 37°C	3.1×10^1
	1hr at 37°C with RV1C5	0

Table 3.3. Specific neutralisation of vectors pseudotyped with rabies G, by incubation with RV1C5.

3.3.5 Cloning of a permanent rabies G expressing cell line

Burns *et al.* (1993) increased titres of their VSV G pseudotyped vector by selecting cell clones expressing high levels of the vector proteins. It therefore appears possible that the use of a cell line expressing rabies G along with the Gag and Pol proteins would also improve vector titres. As a first step towards producing such a packaging cell line, 293 cells were transfected with pcDNA3.G. The 293 cells had been found to be sensitive to 400µg/ml G418 and therefore this concentration was used for selection. Following selection with G418 for one month, two resistant colonies were picked, expanded for 5-6 passages, and then analysed for G expression by IF staining. Only single, isolated, G positive cells were found, approximately 1% of the total cell number. The number of positive cells was not increased by passage in 800µg/ml G418. No cell fusion was

observed. As the number of positive cells was so low these cell lines were not pursued any further.

3.3.6 Efficiency of G incorporation

A possible explanation for the low titres of pseudotyped vector is that G is not incorporated into the envelope very efficiently. In an attempt to look at this, radiolabelled vector released from transfected 293T cells was purified through a sucrose cushion and then analysed by immune precipitation with RV1C5. As can be seen in Fig. 3.5 a protein of a similar size to rabies G is present in this sample (lane A), and absent from the negative control (lane B). However, the signal was very weak - this result was obtained after 10 days' exposure. This suggests that G may be poorly incorporated into the vectors.

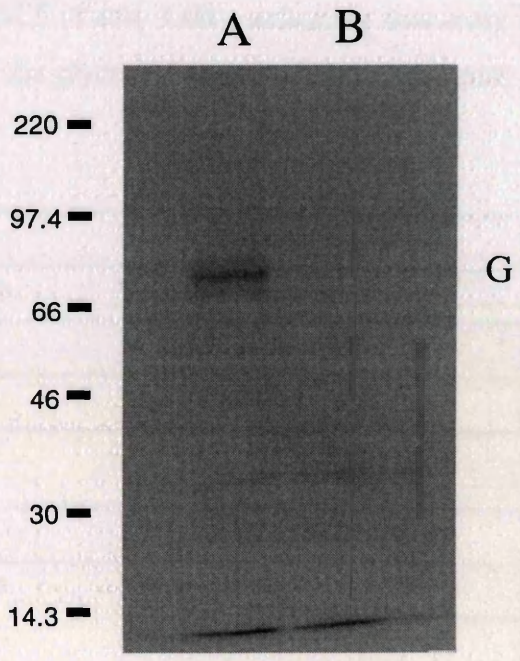


Fig. 3.5. RIPA of vector pellets using RV1C5. A) Cells transfected with pHIT60 and pcDNA3.G. B) Cells transfected with pHIT60 alone.

There is some evidence that signals in the cytoplasmic tail of retroviral Env proteins are involved in their incorporation into virion envelopes (section 1.5.1). The construct GMAenv was produced to test whether the replacement of the cytoplasmic domain of G with that of A-MLV Env could improve its incorporation. The vector titres obtained with this envelope construct, 2-3 x 10^0 bcfu/ml, were consistently lower than those obtained with wild type rabies G (Table 3.4A). These pseudotyped vectors were also neutralised by incubation with RV1C5 (Table 3.4B), indicating that entry of these vectors was mediated by the G domains of the glycoprotein, as would be expected.

A

Vector	Titre (bcfu/ml) 1	Titre (bcfu/ml) 2
GMAenv pseudotype	2 x 10^0	3 x 10^0
G pseudotype	1 x 10^1	2.6 x 10^1
A-MLV	5.4 x 10^4	6.6 x 10^3

B

Vector	Conditions	Titre (bcfu/ml)
GMAenv pseudotype	no incubation	3 x 10^0
	1hr at 37°C	2 x 10^0
	1hr at 37°C with RV1C5	0

Table 3.4. Titres of vectors pseudotyped with GMAenv. A) Comparison of titres obtained with vectors pseudotyped with GMAenv, G or A-MLV Env, on AH927 cells. B) Specific neutralisation of vectors pseudotyped with GMAenv, by incubation with RV1C5.

3.3.7 Time scale comparison of rabies G and Gag expression

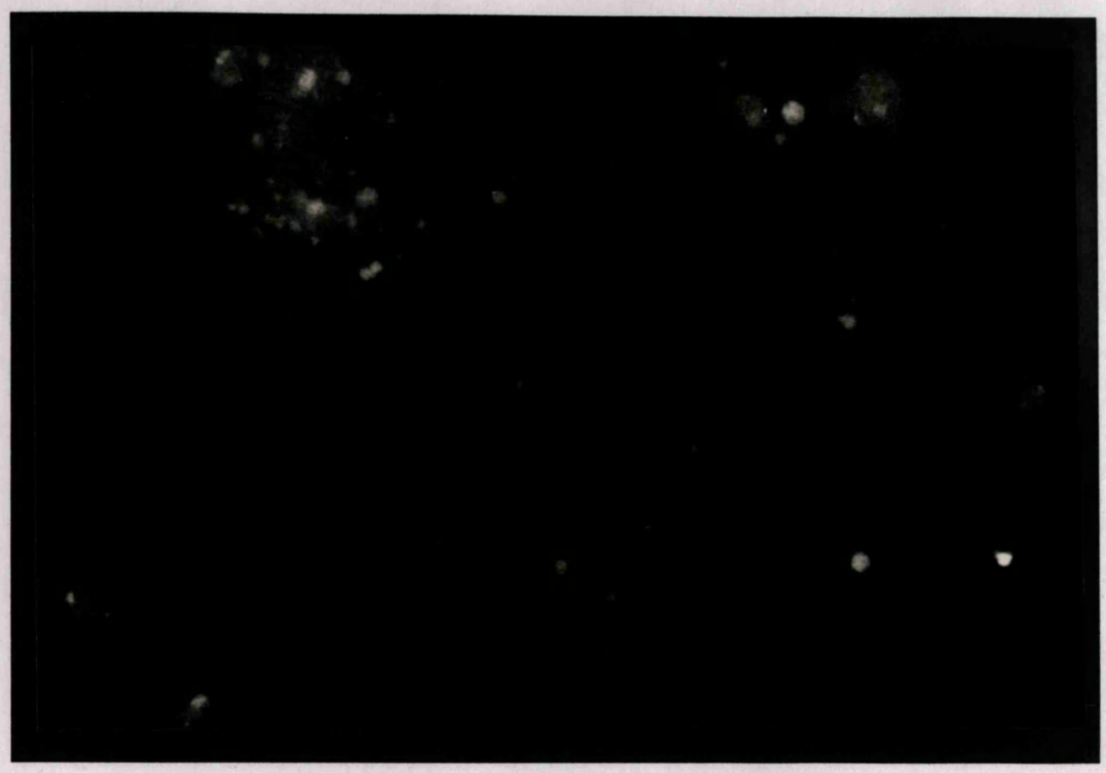
It is possible that the poor incorporation of G results from a low level of glycoprotein on the cell surface during the budding process. It has been reported (Whitt *et al.*, 1991) that rabies G has a slower rate of transport through the cell than VSV G. This implies that the rabies protein accumulates less rapidly on the plasma membrane. If this rate of surface accumulation is also significantly slower than that of the Gag and Pol proteins, incorporation of G by the vectors would be expected to be affected. This could be tested by comparing the expression of these proteins at various time points after transfection.

293T cells were co-transfected with pcDNA3.G, pGT5 and pHIT111. At 0, 2, 4, 6, 24, 48 and 72 hours after the end of transfection, these were analysed for G expression by IF staining, and for Gag expression by ELISA. pGT5 was used because an ELISA test for FeLV Gag was readily available. G protein could be detected in cells immediately after the end of transfection (Fig. 3.6, B) and was present on the cell surface by 2 hours (Fig. 3.6, C). The IF staining was most intense at 6 and 24 hours (Fig. 3.6, G-J), and began to fade at later time points (Fig. 3.6, K-N). Some possible localised cell fusion was observed at 72 hours. The ELISA results obtained in this experiment were very poor (Table 3.5). The readings taken at 24 to 72 hours were slightly higher than the nontransfected control, but they remained approximately 10 fold lower than the positive control consisting of pooled serum/plasma samples from FeLV-positive cats. This low level of detected Gag may have been the result of a reduced transfection efficiency of 20-30%.

Sample	**Absorbance (450nm)**	
0 hours	0.048	0.064
2 hours	0.055	0.060
4 hours	0.045	0.053
6 hours	0.070	0.067
24 hours	0.096	0.091
48 hours	0.101	0.065
72 hours	0.116	0.131
nontransfected cells	0.070	
positive control	1.223	

Table 3.5. ELISA of FeLV Gag expression at various time points after the end of transfection.

A



B

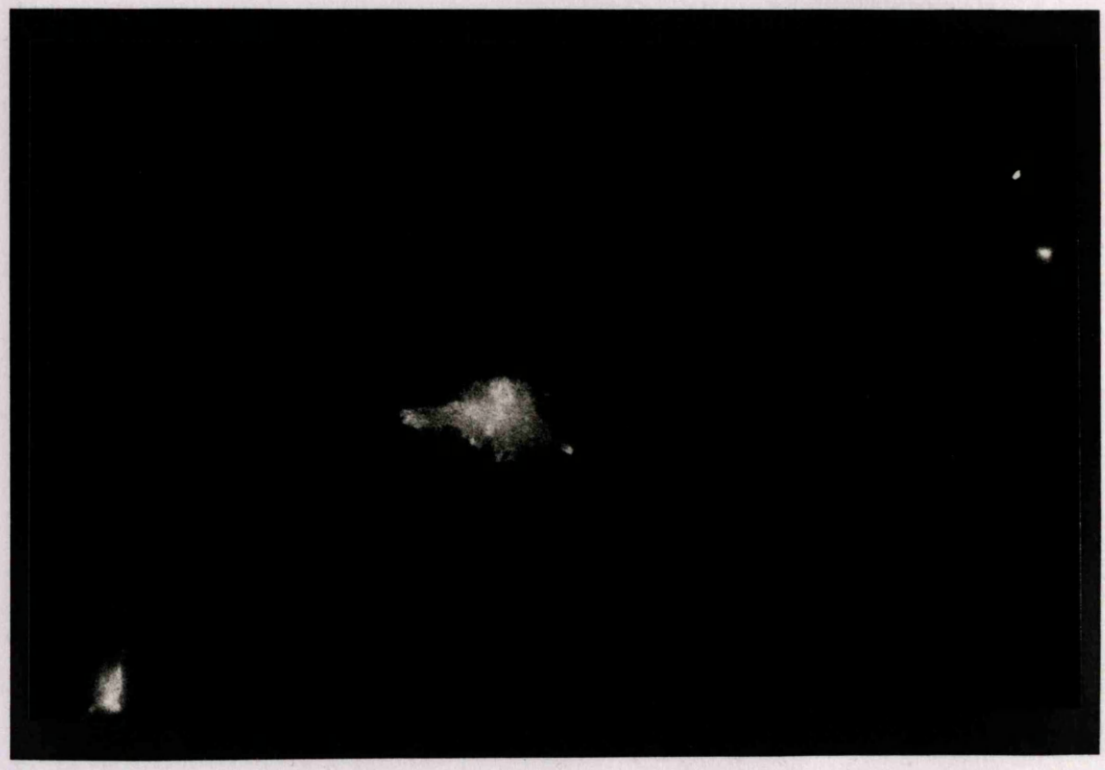


Fig. 3.6. Expression of rabies G by 293T cells at various time points after transfection with pcDNA3.G. A) IF staining of nontransfected cells. B) IF staining of permeabilised cells 0 hours after transfection.

C

D

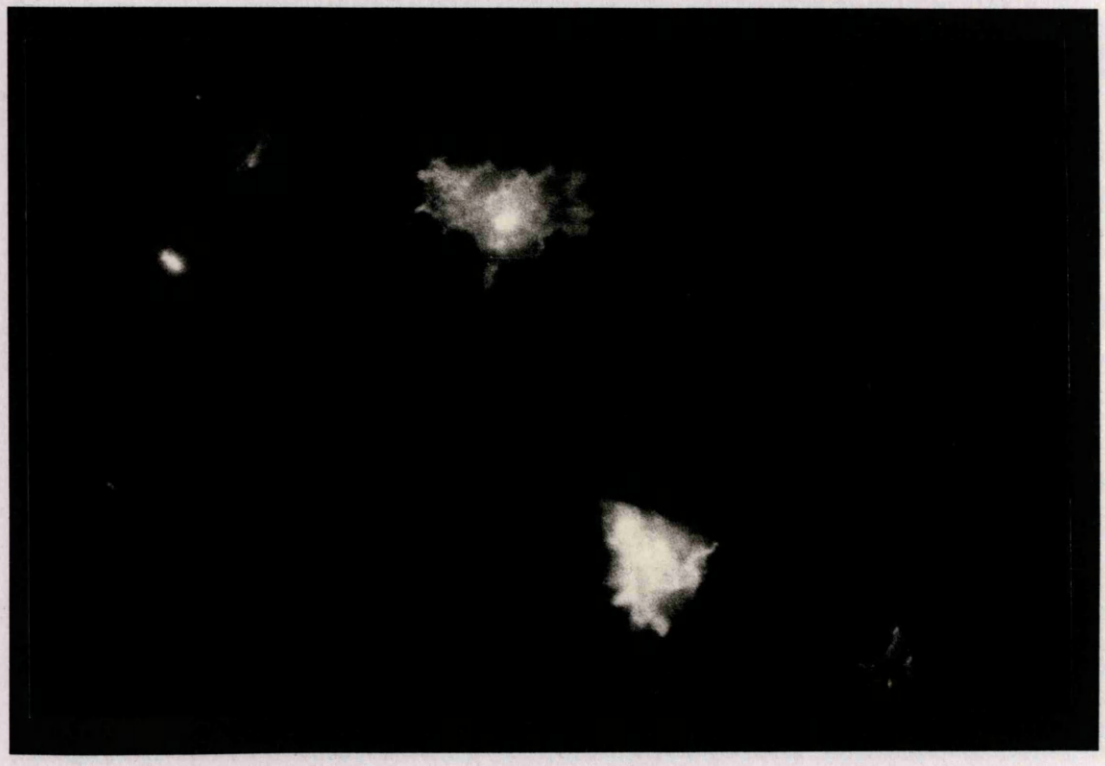
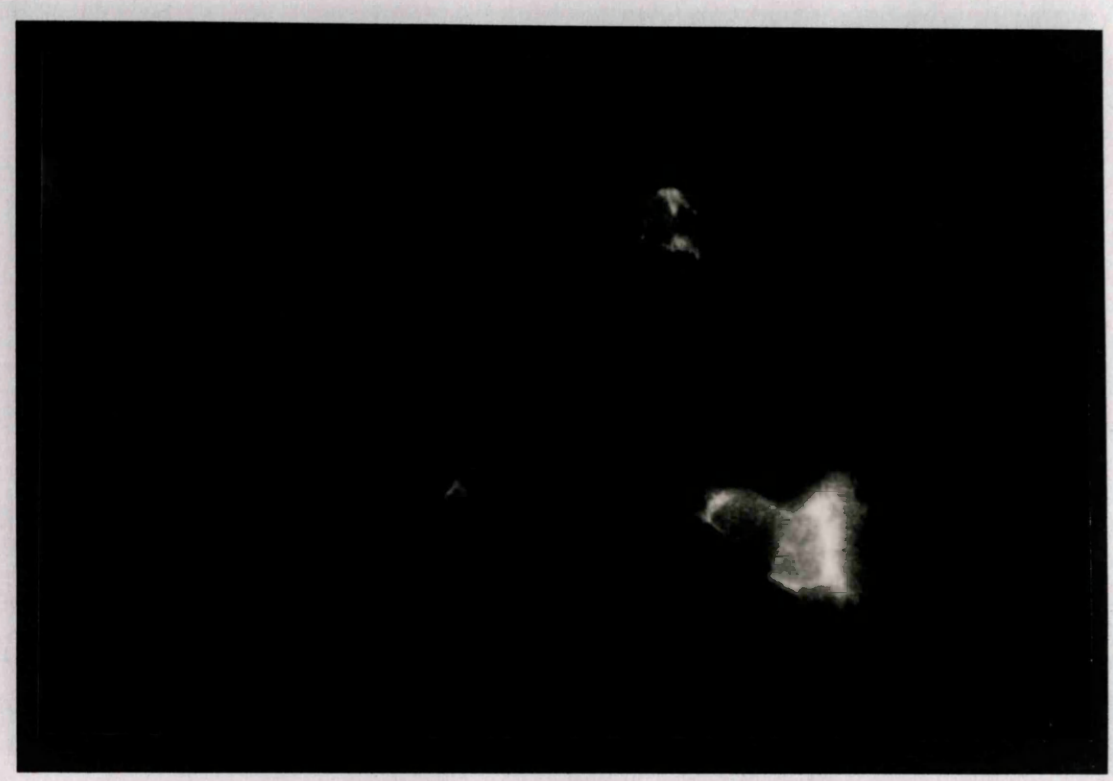


Fig. 3.6. Cont. C) IF staining of cells 2 hours after transfection. D) IF staining of permeabilised cells 2 hours after transfection.

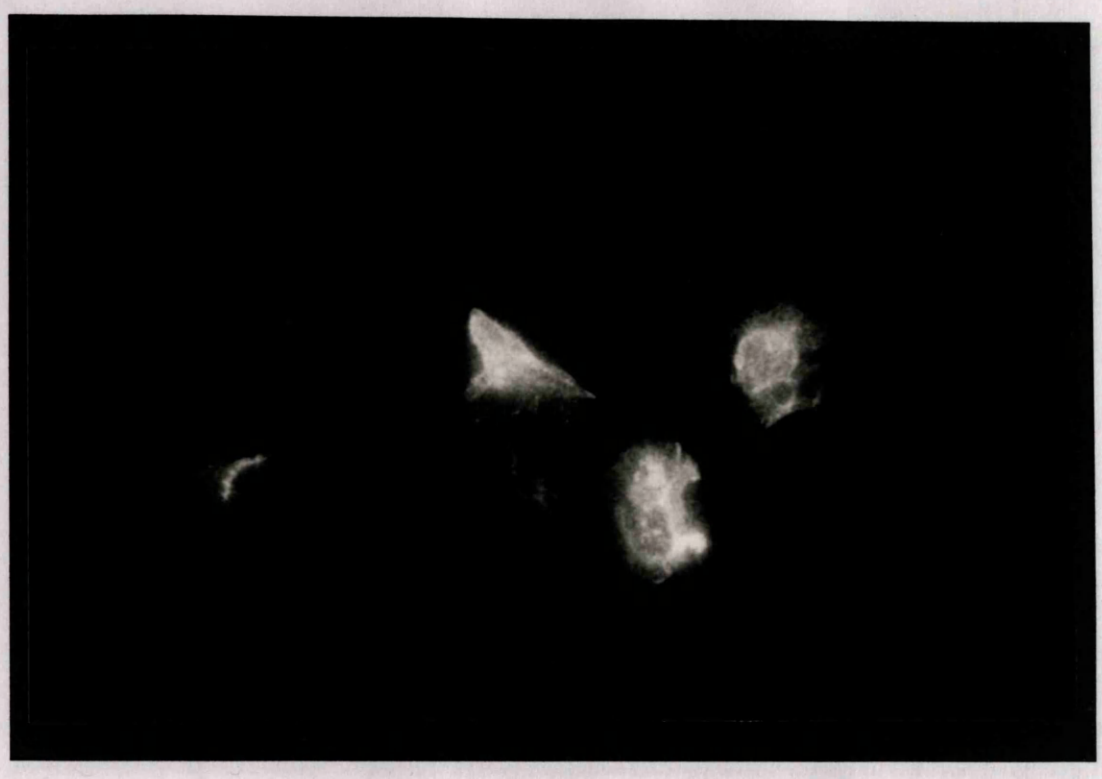
E



F

Fig. 3.6. Cont. E) IF staining of cells 4 hours after transfection. F) IF staining of permeabilised cells 4 hours after transfection.

G



H

Fig. 3.6. Cont. G) IF staining of cells 6 hours after transfection. H) IF staining of permeabilised cells 6 hours after transfection.

I

J

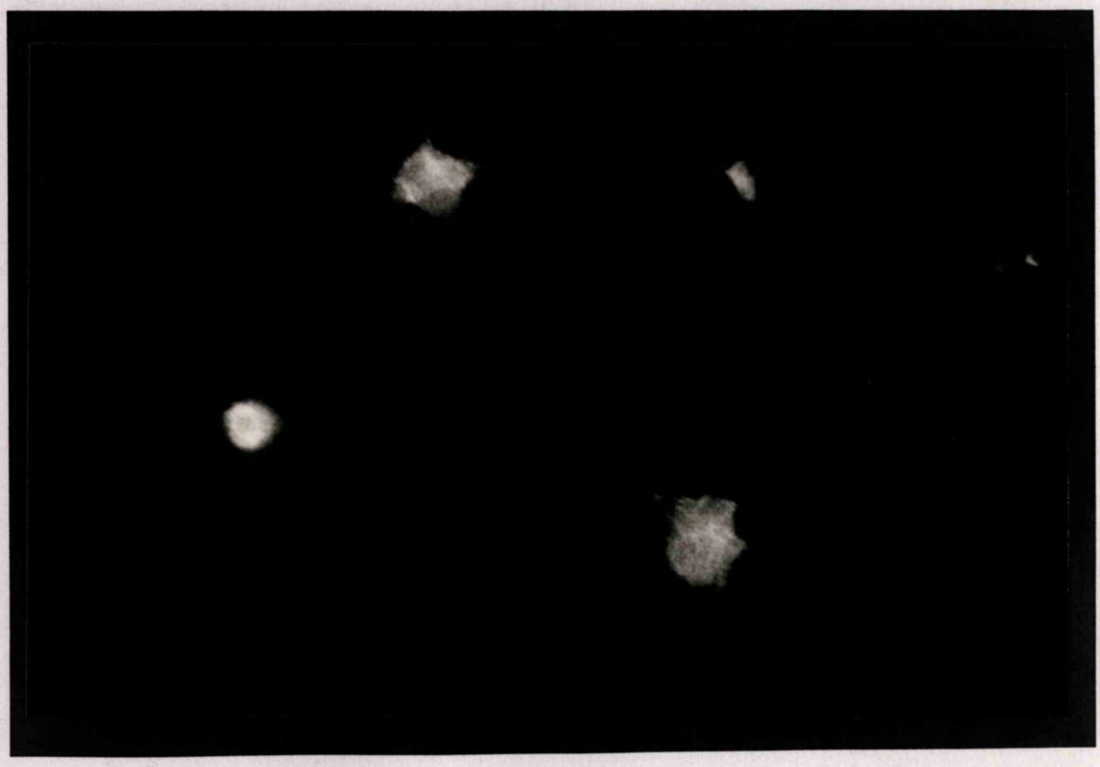


Fig. 3.6. Cont. I) IF staining of cells 24 hours after transfection. J) IF staining of permeabilised cells 24 hours after transfection.

K

L

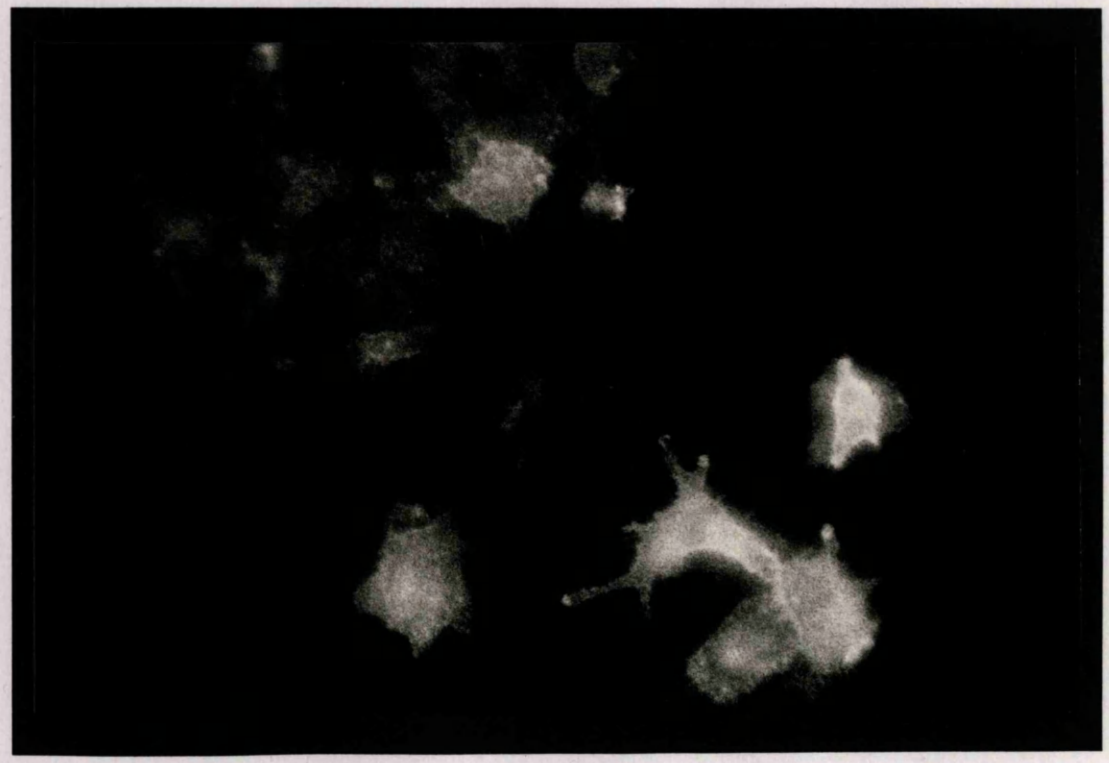
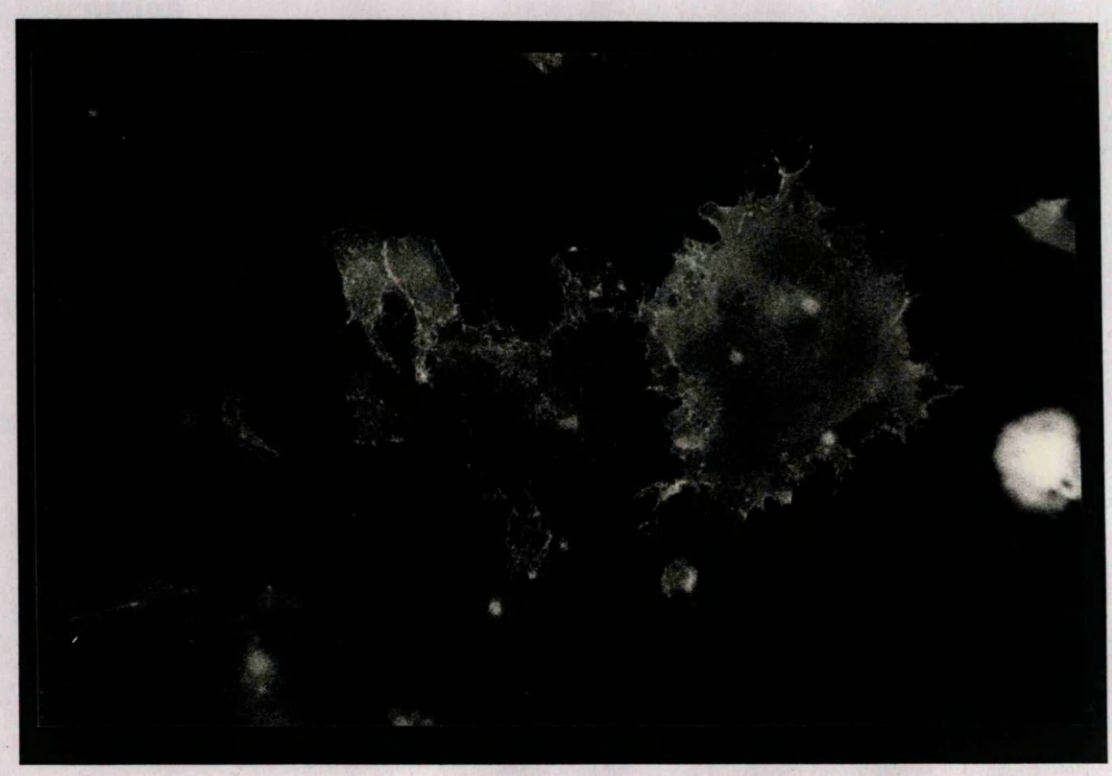


Fig. 3.6. Cont. K) IF staining of cells 48 hours after transfection. L) IF staining of permeabilised cells 48 hours after transfection.

M



N

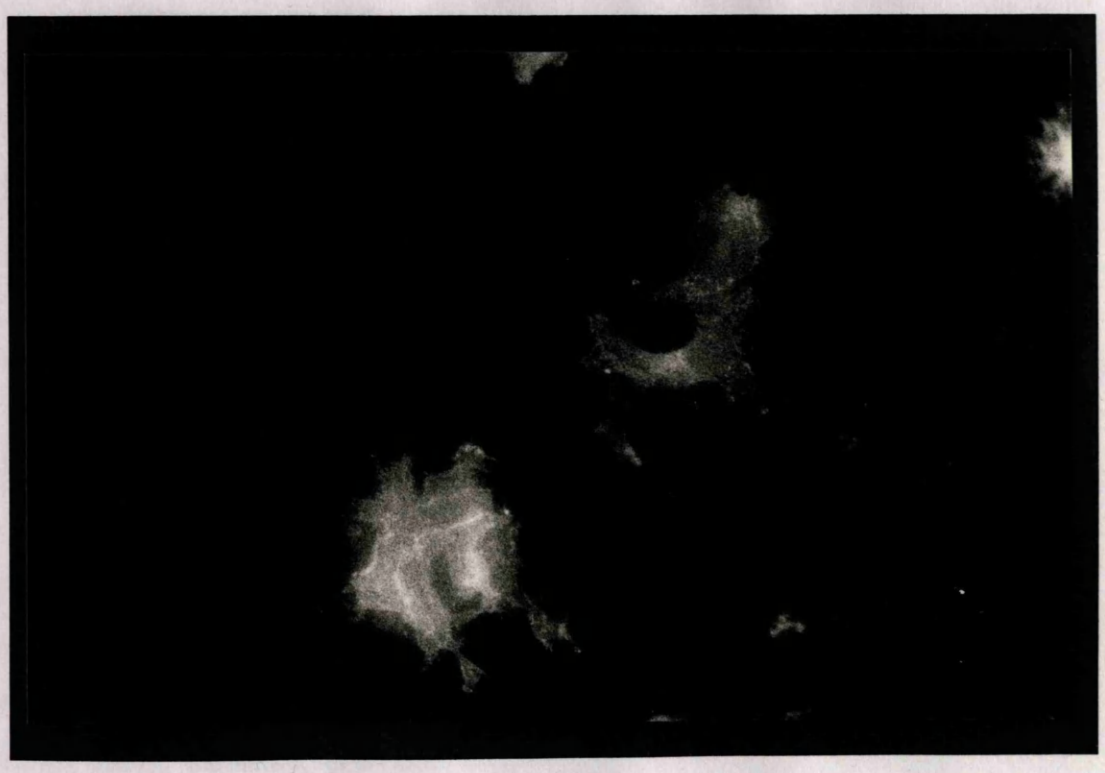


Fig. 3.6. Cont. M) IF staining of cells 72 hours after transfection. N) IF staining of permeabilised cells 72 hours after transfection.

3.4 Discussion

The major conclusion that can be drawn from these experiments is that rabies G can be incorporated into retroviral vectors. The infectivity of these vectors was dependent on the incorporation of the foreign glycoprotein as demonstrated by their specific neutralisation by an antibody directed against the rabies G protein. Unfortunately, despite confirmation of the cell surface expression and fusion activity of G, these pseudotypes had very low titres. These titres were $\leq 10^2$ bcfu/ml, while those for the A-MLV Env vector were 10^4-10^5 bcfu/ml. In contrast, Emi *et al.* (1991) found that transiently produced VSV G pseudotyped vectors had titres comparable to those of similar vectors containing A-MLV Env.

The attempts to optimise these titres, described in Table 3.2, show that they cannot easily be improved, at least in this transient system. The replacement of serum in the culture medium did appear to increase titres marginally. This could be due to the removal of proteins which promote aggregation of the vectors. Despite the fact that NaB is reported to increase expression from the HIT system plasmids (Soneoka *et al.*, 1995), its addition did not increase the titres of the pseudotyped vectors. Altering the dosage of pcDNA3.G in the transfections appeared to decrease vector titre, suggesting that the favoured concentration is one that is equal to those of the other two plasmids. However, as only three alternatives were tried it is possible that another concentration would be optimal. The failure to increase titres of G pseudotyped vector by harvesting at 32°C suggests that the factors affected by this incubation temperature are not relevant to the production of these pseudotypes. However, Kaptain *et al.* were also able to show that the improved vector stability observed at 32°C is not seen at 30°C nor 34°C. Therefore it is possible that poor temperature control could have prevented the potential improvement in titre.

Although the RIPA analysis of the vector was not quantitative, the weakness of the G signal in Fig. 3.5 suggests that little of the protein was present. This weak signal was only detected after samples were taken from cultures grown in 75cm^2 flasks rather than 25cm^2. It seems likely, therefore, that the low titres of pseudotyped vector were due to poor incorporation of the rabies glycoprotein, although a control in which Env or VSV G was precipitated from vector would have made this conclusion more reliable. Replacement of the cytoplasmic domain of rabies G with that of A-MLV Env did not increase titres. This

could indicate that any possible recognition sequences present in the retroviral Env cytoplasmic domain are not sufficient to promote incorporation of the chimaeric protein. However, it is also possible that the presence of this domain somehow interferes with the functioning of G. Although rabies G can mediate infection in the absence of a cytoplasmic domain (Mebatsion *et al.*, 1996), the addition of a foreign sequence could disrupt secondary structure or alter interactions between the monomers.

One probable cause of inefficient G incorporation is the low rate of transport of the protein within the cell. This implies that rabies G accumulates at the plasma membrane more slowly than the VSV protein and is probably, therefore, present on the cell surface at a lower concentration. If the level of rabies G protein on the cell surface can be shown to be lower than that of VSV G, data presented here would appear to support a model of pseudotype formation in which foreign proteins are passively incorporated into the virion envelope (section 1.5.1). If this is the case it would be expected that increasing the level of expression of G with respect to Gag would increase the production of pseudotyped vectors. Rose and Bergmann (1983) have shown that certain mutations of the VSV G cytoplasmic domain result in a reduced rate of acquisition of endoglycosidase H-resistant sugars, implying a reduced rate of transport through the cell. It is possible that GMAenv is affected in a similar fashion, thereby further reducing cell surface expression and, therefore, incorporation.

Despite the fact that the optimised conditions were used in subsequent experiments titres were not always as high as 10^2 bcfu/ml. This could mean that the production of this pseudotype is very sensitive to small variations in experimental procedure, such as transfection efficiency, cell number and rate of growth. This might be expected if there is no interaction between the retroviral proteins and the G protein. Incorporation would be passive and, therefore, relatively random.

Further support for the theory that the low level of surface expression of G produces the low titres of pseudotype comes from the observations of Whitt *et al.* (1991). They found that the titre of VSV ts045 rescued by rabies G is 50 to 100 fold lower than that rescued by VSV G. Whitt *et al.* suggest that this is due to the low rate of transport of the rabies protein since they found similar loss of titre in a previous study with certain mutant VSV G

proteins that are transported more slowly than wild type (Whitt *et al.*, 1989). Another possibility is that the rabies protein is incorporated less efficiently. Several studies have indicated that the cytoplasmic domains of the VSV and rabies G proteins contain a signal that promotes incorporation (Mebatsion *et al.*, 1996; Owens and Rose, 1993; Whitt *et al.*, 1989). The cytoplasmic domain of rabies G shares little homology with the corresponding sequences of VSV G, suggesting that this signal is virus-specific. VSV G proteins with certain foreign cytoplasmic tails have been shown to be incapable of rescuing VSV ts045 (Whitt *et al.*, 1989).

It is difficult to conclude much from the data obtained from the comparison of rabies G and Gag expression over time. Although G can be detected on the surface of cells soon after infection, it is impossible to tell whether the expressed levels are high enough to support pseudotype formation. The intensity of the IF staining does suggest that G expression peaks at around 24 hours and then falls. A more quantitative approach such as an ELISA would provide a better picture of this pattern of expression, and would also facilitate comparison with that of the Gag proteins. The failure of the Gag ELISA on this occasion appears to have been the result of a poor transfection efficiency. However, without repeating the experiment this cannot be concluded with any certainty.

The analysis of cell clones permanently transfected with pcDNA3.G revealed that only a very few cells expressed rabies G. This is perhaps surprising since others have reported stable expression without any obvious cytotoxicity (Burger *et al.*, 1991). However, these studies were carried out with different cell lines and G proteins from different strains of rabies virus. It may be possible that cytotoxicity is increased in this system. Possible cell fusion was observed 72 hours after the end of transfection (Fig. 3.6), and Mebatsion *et al.* (1999) have reported that high accumulation of the SAD B19 G protein on the cell surface following infection by a mutant rabies virus results in syncytia formation. Also, it has recently been reported that expression of G is associated with rabies virus induced apoptosis of lymphocytes, possibly as a result of oxidative stress caused by accumulation of the protein in intracytoplasmic structures (Thoulouze *et al.*, 1997). The low level of expressing cells could also be due to the expression plasmid used. If this is the case, the use of a plasmid that contains the G protein and selector genes in the same reading frame may be more appropriate. In such a system expression of the selector gene occurs as a result of

continued transcription from the G gene, reducing the risk of downregulation of G expression.

If it can be assumed from the work of Emi *et al.* (1991) that a VSV G pseudotyped vector would have a titre comparable to that of the A-MLV Env vector if used in this system, this would be up to 1000 fold greater than that achieved with the rabies G pseudotype. This is a much larger difference than would be expected from the variation in rate of protein transport. Another possible factor influencing the abilities of these proteins to form pseudotypes with retroviruses involves their role in virion budding. Mebatsion *et al.* (1996) found that the presence of the G protein results in an approximately 30 fold increase in the budding efficiency of rabies virions. Their conclusion that this indicates an exocytotic activity of G is supported by the work of Rolls *et al.* (1994). This study showed that the expression of VSV or rabies G from a Semliki Forest virus replicon RNA results in the release of self-propagating infectious vesicles containing the glycoprotein and the RNA replicon. Presumably the G protein induces vesiculation of the plasma membrane and the high concentration of the vector in the cell allows it to be randomly incorporated. Although these infectious vesicles were produced with both VSV and rabies G, those containing the rabies protein had titres at least 100 fold lower. Rolls *at al.* suggested that this could be explained by the slower transport rate of rabies G. However, Mebatsion *et al.* claim that their unpublished data indicate that the efficiency of the exocytotic activities of the two glycoproteins differ. If this is the case it could be proposed that the efficient activity of the VSV G protein promotes budding of retroviral particles in the same region of the plasma membrane, thereby producing high titres of pseudotyped vectors. In contrast, the weaker activity of rabies G has little effect on retroviral budding, and therefore the protein is only incorporated by chance. It would be interesting to test whether pseudotyping with VSV G can increase the titre of rabies virus, as would be predicted by this model.

Although the G protein is the only VSV protein required to produce functional pseudotyped retroviral vectors (Emi *et al.*, 1991) this does not necessarily hold true for rabies G. Infection of cells with a mutant rabies virus lacking the M protein results in cell surface accumulation of G protein, presumably as a result of a low level of incorporation into budding virions (Mebatsion *et al.*, 1999). This suggests that rabies G requires M for efficient incorporation. Data from the same study reveal that a specific interaction between

these two proteins can be detected. This interaction may also be required for the efficient functioning of the G protein. It is thought that the VSV matrix protein M stabilises the G trimers (Lyles *et al.*, 1992). It seems possible that a similar interaction occurs between the M and G proteins of rabies, and as the trimers of rabies G are less stable (Gaudin *et al.*, 1992; Whitt *et al.*, 1991) the requirement for such an interaction may be increased. Although the G protein appeared to be functional in the cell fusion assay, this might not be truly relevant to the fusion of the viral and cellular membranes.

This work was carried out in the hope that rabies G would provide a useful alternative to VSV G in the production of stable, high titre pseudotyped retroviral vector. These results clearly show that this is not the case. However, further work with these pseudotypes may provide interesting insights into the processes governing virion budding and the incorporation of viral envelope proteins.

CHAPTER 4

Chapter 4 Infectivity of porcine endogenous retroviruses

4.1 Introduction

The genomes of many, if not all vertebrates contain endogenous retroviral sequences (section 1.1.6). Some of these proviral sequences still retain the capacity to produce infectious virus capable of replicating in the cells of other species. This property of ERVs has recently become of interest due to the problems it poses for xenotransplantation. This involves the transplantation of organs or tissues from one species to another, and has the potential to provide an ideal environment for the transmission of these viruses. It has therefore become important to identify infectious ERVs present in the genomes of proposed donor animals, and to determine their host range.

4.1.1 Xenotransplantation

Transplantation is currently the best, if not the only, treatment for several conditions such as end-stage organ failure and hepatic disease. Consequently the demand for organs is continually rising. According to UNOS there are currently 65,204 patients waiting for a transplant in the USA. At the end of 1988 this number was 16,026. In contrast to this increasing demand the most reliable sources of organ donors are declining, mainly through improvements in road safety and in the treatment of strokes. As a result only 20,967 transplants were performed in the USA in 1998. (Data published by UNOS [www.unos.org]).

One potential solution to this problem is xenotransplantation, the transplantation of an organ, tissue or cells from another species. Such transplants could be used either as a permanent treatment or as a bridging therapy until a suitable human organ becomes available. The use of animal cells or tissues may also allow the development of new procedures for treating illnesses such as AIDS, diabetes and Parkinson's disease (Groth *et al.*, 1994; Isacson and Breakefield, 1997; Nowak, 1994). The breeding of animals specifically for transplants would provide almost unlimited numbers of organs. In addition,

the use of quarantined, specific pathogen free (SPF) animals would improve the quality control of organs and tissues.

Of course there are ethical objections to the use of animals in this way. Many people disapprove of the practice of breeding animals specifically for their organs, especially when this involves raising in an unnatural, sterile environment and the possibility of genetic manipulation. However, the large demand from patients, coupled with the potential benefits of xenotransplantation and probable large commercial returns, has provided strong motivation for research in this area.

The first attempts at xenotransplantation used baboon organs (Bailey *et al.*, 1985; Starzl *et al.*, 1993; reviewed by Taniguchi and Cooper, 1997). This approach has now been virtually abandoned due to strong ethical objections, the high risk of disease transmission including human tropic baboon endogenous retrovirus (BaEV), and the impossibility of breeding the large numbers that would be required (Allan, 1996; Butler, 1999). Most work now concentrates on the use of pigs, since their organs are also approximately the same size as those of humans. These animals are thought to present a lower risk of infection than baboons, due to their being less closely related to humans. Further advantages of pigs over baboons are that they breed quickly and have been domesticated for hundreds of years (Sachs, 1994).

4.1.2 Porcine endogenous retroviruses

Although the breeding of SPF animals is intended to remove infectious agents, this process is not infallible. The pathogens most likely to pose a potential risk of infection from xenotransplantation are those that are undetectable, or impossible to eliminate from the donor animals. ERVs are carried in the germ line and therefore cannot be removed from the animals. Since some of these proviral sequences still code for functional virus, it is plausible that they could be transmitted as a xenozoonosis (Chapman *et al.*, 1995; Smith, 1993; Stoye and Coffin, 1995). Xenozoonosis is the term used to describe the transfer of infections through xenotransplantation (Michaels and Simmons, 1994). A potential danger in the use of any transplant is the introduction of infectious agents, and the transmission of pathogens via allotransplantation often occurs. In this situation no vector is required for

disease transmission and migration of infected cells throughout the body can occur. In addition, the use of immunosuppression can result in the absence of the usual symptoms of inflammation, thereby delaying diagnosis (Bach *et al.*, 1998). The risks associated with xenotransplantation are probably higher. Novel, as yet unidentified, organisms could be present, and the higher levels of immune suppression and graft rejection are likely to increase activation of any latent pathogens (Bach *et al.*, 1998). The prolonged close contact between the animal and human tissues, in the absence of the normal barriers to infection, will probably greatly increase the possibility of pathogen transmission (Brown *et al.*, 1998). In addition, organisms present in the graft may not be known human pathogens, but may be xenotropic and therefore non-pathogenic in the pig (Bach *et al.*, 1998). Many ERVs are xenotropic, and it has already been shown that viruses from baboons, cats and mice are able to infect and replicate in human cells (Stoye and Coffin, 1995). The feline ERV RD114 was first detected in a human tumour xenograft in a foetal cat (McAllister *et al.*, 1972). Although ERVs tend not to be harmful to their hosts, and are likely to be selected at least partly on the basis of low pathogenicity, they cannot be assumed to remain harmless if transferred to a new host (Boeke and Stoye, 1997). While the majority of these viruses cannot replicate in their natural host they are often able to grow to high titres on cells of other animals (Stoye and Coffin, 1995). There are also documented cases of cross species infections by ERVs. RD114 is believed to have originated from the transfer of BaEV, and GALV from an endogenous C-type retrovirus found in mice (Benveniste and Todaro, 1974; Lieber *et al.*, 1975b). Pathogenic effects of retroviral infection often appear to result from replication in an unnatural host where the normal evolved balance between viral replication and host cell damage is lost (Coffin *et al.*, 1997).

Cultured porcine cell lines, such as kidney PK15, MPK and lymphoma V 38A-1, have been known for some time to release C-type retroviral particles which are thought to be endogenous (Armstrong *et al.*, 1971; Lieber *et al.*, 1975a; Moennig *et al.*, 1974; Todaro *et al.*, 1974). Evidence from DNA hybridisation experiments suggests that this porcine retrovirus (PoERV) evolved from an endogenous murine virus introduced into the pig genome 5 to 10 million years ago (Benveniste and Todaro, 1975). Transmission is thought to have occurred prior to the divergence of the murine species *Mus musculus* and *Mus caroli*, and the original murine ERV sequences no longer appear to code for functional virus. Early reports (Lieber *et al.*, 1975a; Todaro *et al.*, 1974) suggested that PoERV had a

narrow ecotropic host range, since productive infection was only achieved in the porcine ST-IOWA cell line. However, rescue of Kirsten murine sarcoma virus from nonproducer cell lines could be attained by cocultivation. Both serological (Moennig *et al.*, 1974) and sequence (Q-One Biotech Ltd. and Imutran Ltd., 1997) data show that PoERV is closely related to GALV. This is illustrated by the phylogenetic tree in Fig. 4.1, constructed from the alignment of Pol amino acid sequences. GALV is horizontally transmitted between gibbons, and causes lymphoid and myeloid malignancies (Kawakami *et al.*, 1972; Kawakami *et al.*, 1980). It is thought to have recently evolved from an endogenous retrovirus of the Asian mouse *Mus caroli* or a closely related species (Lieber *et al.*, 1975b). This cross species transmission will have occurred under less permissive conditions than those presented by xenotransplantation (Brown *et al.*, 1998).

Both RD114 and GALV have been shown to be in the same interference groups, with the majority of cell types, as their ancestral viruses, BaEV and the *Mus caroli* ERV respectively (Sommerfelt and Weiss, 1990; Lieber *et al.*, 1975b). This indicates that the viruses recognise the same receptors (section 1.2.1). Therefore, it is possible that the PoERVs recognise a receptor related to that of the original parent murine ERV, and could display a host range wider than that so far observed. As a result, the PoERVs must be considered as a potential risk in xenotransplant procedures. As described in section 1.1.5, retroviruses are able to evolve rapidly, adapting to an altered host environment or changing their tissue tropism. This implies that, during exposure to human cells, a PoERV with low infectivity could adapt to replicate in the new host, either through mutation or by recombination with a human retrovirus. Such changes could also increase the pathogenicity of the virus. Both FeLV-B and -C are more pathogenic than the FeLV-A from which they are derived, by recombination and mutation respectively (Neil *et al.*, 1991; Stewart *et al.*, 1986). The highly oncogenic murine 10A1 virus is the result of a recombination event between an A-MLV and a MLV-related endogenous *env* sequence (Ott *et al.*, 1990).

The transplantation of an organ carrying endogenous viruses has been likened to the contamination of gene therapy vectors with replication competent virus (Stoye and Coffin, 1995). An example illustrating the potential dangers of such contamination is the case of 10 immunosuppressed rhesus monkeys that received autologous transplants of bone marrow transduced with an A-MLV retroviral vector preparation containing replication

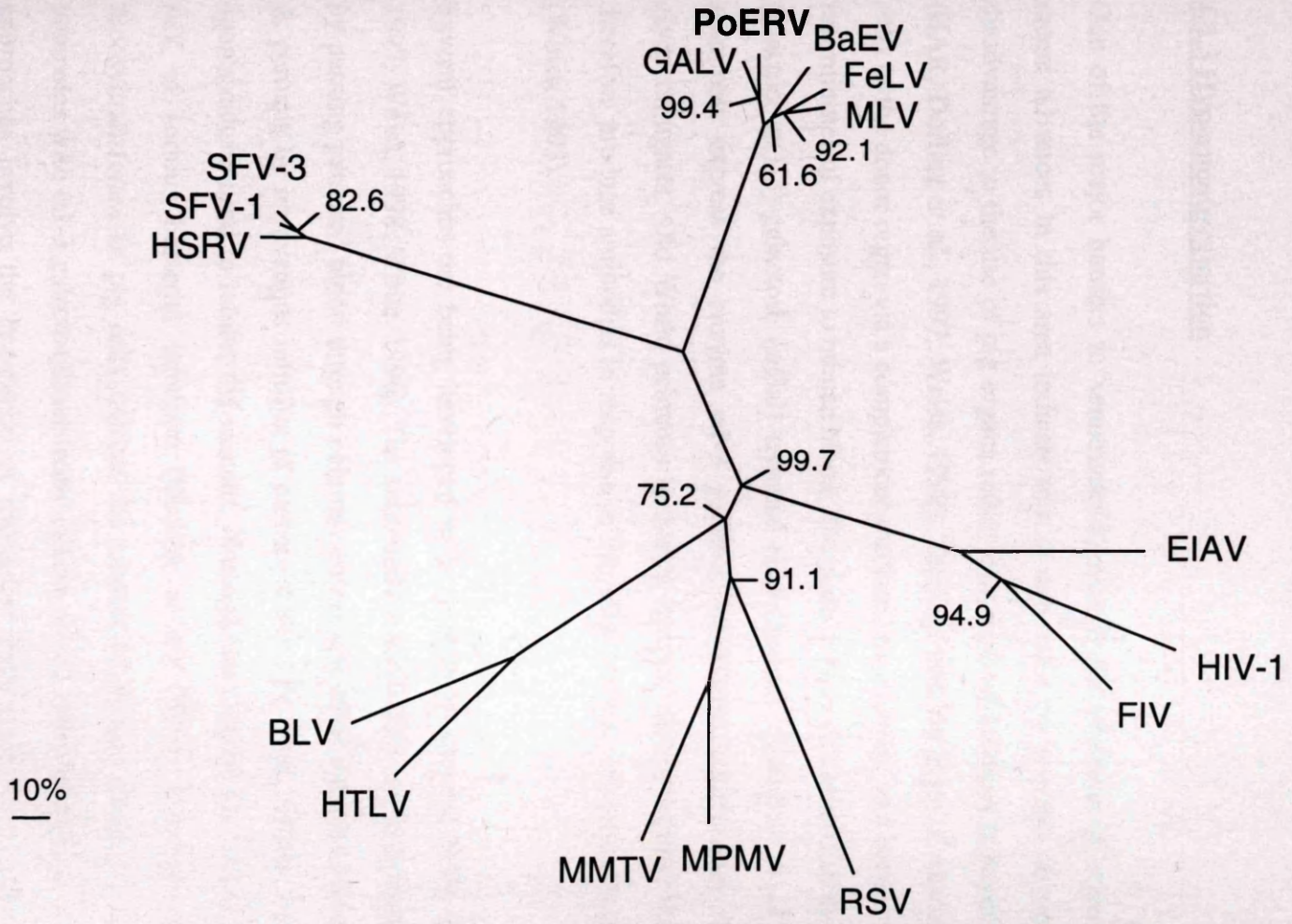


Fig.4.1. Unrooted phylogenetic tree of retroviruses based on alignment of Pol amino acid sequences, showing PoERV placed within the mammalian C-type group. The tree was constructed from an alignment of 16 representative retrovirus sequences corresponding to amino acids 178 to 740 of the predicted PoERV Pol polyprotein (reverse transcriptase). The values given at the branch forks are the bootstrap percentages for 2000 samplings, with missing values all being 100%. Sequences were aligned using the CLUSTAL W program and the PHYLIP package was used for protein distance calculation and tree construction. The solid line beside the diagram represents 10% divergence. Constructed by Douglas Hart (Veterinary Pathology, University of Glasgow).

competent virus (Donahue *et al.*, 1992; Vanin *et al.*, 1994). Although this virus was believed to be non-pathogenic three of the monkeys developed a rapidly progressive T-cell lymphoma, probably as a result of insertional mutagenesis. In contrast to the seven unaffected animals these three appeared to lack any immunological response to the virus. This observation highlights the increased risk of a pathogenic outcome of retroviral infection in an immune-suppressed transplant patient.

4.1.3 Hyperacute rejection

One of the major barriers to xenotransplantation is the problem of rejection. Although recent advances in this area indicate that it may soon be possible to minimise this, a disadvantage in the use of pig organs rather than those of baboons is hyperacute rejection (HAR, Dorling *et al.*, 1997; Weiss, 1998). This involves the lysis of vascular endothelial cells of the donor organ via a complement-mediated mechanism, and occurs within the first few minutes of exposure to human blood. The main antigen recognised in this response is a galactose-α(1-3)-galactose (αGal) terminal sugar residue (Vaughan *et al.*, 1994). Most mammals express the enzyme α1-3 galactosyltransferase which adds this residue to glycoconjugates. Old World primates including humans do not express this enzyme and therefore produce antibodies in response to the αGal antigen present on bacteria in the gut (Weiss, 1998).

Several approaches are being developed in an attempt to prevent HAR (Dorling *et al.*, 1997; Weiss, 1998; White, 1996). The xenoreactive antibodies can be temporarily removed by passing patients' blood through columns containing αGal linkages (Good *et al.*, 1992) or perhaps by intravenous infusion of carbohydrates (Ye *et al.*, 1994). Anti-complement agents can be used to inhibit the cascade, although the clinical use of these increases the risk of serious bacterial infection (Dorling *et al.*, 1997). Overexpression of α1,2-fucosyltransferase in pig cells reduces the amount of surface antigen since this enzyme competes with α1-3 galactosyltransferase (Koike *et al.*, 1996; Sandrin *et al.*, 1995). Other approaches involve the generation of transgenic pigs expressing human complement-regulatory proteins. These regulatory proteins are CD55 (or decay accelerating factor, DAF), CD46 (or membrane cofactor protein, MCP-1) and CD59 (or protectin). They act by

inhibiting steps in the complement cascade. Several transgenic herds expressing one or more of these proteins have been developed, and the approach appears promising (Cozzi and White, 1995; Diamond *et al.*, 1996; Fodor *et al.*, 1994; Pino-Chavez *et al.*, 1996; Rosengard *et al.*, 1995; Zaidi *et al.*, 1998).

The strategies being undertaken to combat HAR could also confer resistance to viruses released from the graft (Weiss, 1998). Enveloped viruses released from animal cells are susceptible to lysis by human complement. This has been shown to be due to the presence of αGal residues on the viral envelope (Takeuchi *et al.*, 1996). If the same viruses are grown in human cells they become resistant to this lysis. It therefore seems very possible that viruses grown in transgenic tissues will also be resistant. The presence of CD46, CD55 and CD59 in the HIV-1 envelope has been shown to be protective (Montefiori *et al.*, 1994; Saifuddin *et al.*, 1995). The suppression of the complement cascade is also likely to leave the patient more vulnerable to viral infection. Another risk arising from the use of transgenic animals is that CD46 and CD55 are known to act as receptors for human viruses and would render the pig tissues susceptible to infection by these viruses. It is also possible that related viruses that recognise the porcine equivalent of these proteins could adapt to use the human protein, thereby becoming potential human pathogens (Weiss, 1998).

4.1.4 Are PoERVs a potential risk?

It is very difficult to assess the risk to a patient since the pathogenicity of an infection can change unpredictably when introduced into a new species (Chapman *et al.*, 1995). This can be seen in naturally occurring zoonoses. For example, the rodent hantaviruses produce no detectable mortality or morbidity in their natural hosts. However, human infection can result in mortality rates of up to 50% (Chapman and Khabbaz, 1994). The measles virus can cause severe, often fatal disease in marmosets, but infection of rhesus monkeys results in mild disease comparable to that normally seen in humans (Kalter and Heberling, 1976).

If a xenozoonotic infection is restricted to the transplant recipient, then the risk of such an infection can be simply balanced against the benefits of the procedure. However it is quite possible that disease could be passed on to other individuals, in which case the risk to the population as a whole would need to be taken into account. There are several well known

cases of zoonotic viruses causing outbreaks of disease in humans. In 1967 the importation into Germany of vervet monkeys infected with Marburg virus resulted in human-to-human transmission involving 31 individuals (Martini, 1969). Several outbreaks of Ebola virus infection in Sudan and Zaire have resulted in multiple generations of transmission with fatality rates of 80-90% (WHO, 1978). Although such cases are very dramatic, the outbreaks are short lived and fairly contained. Widespread infections are most efficiently achieved by the transmission of an agent that establishes long term infection with delayed onset of disease (Brown *et al.*, 1998). Such infections are characterised by retroviruses. The human immunodeficiency viruses are generally believed to have resulted from the adaptation of simian retroviruses following zoonotic infection. HIV-1 is thought to have originated in chimpanzees (Gao *et al.*, 1999) and HIV-2 in sooty mangabeys (Chen *et al.*, 1997; Gao *et al.*, 1992). Following initial infection of humans transmission continued for well over a decade before AIDS was identified (Chapman *et al.*, 1995).

In order for PoERV to pose a threat to public health several conditions need to be met (Stoye, 1998). The virus must first be present in the germline of the pigs bred for transplantation and be expressed by the transplanted cells. It must also be able to infect and replicate in the recipient's cells and thereby cause disease. Finally, transmission of the virus to other individuals must occur. These viruses are thought to be present in all pigs (Lieber *et al.*, 1975a; Todaro *et al.*, 1974), and it is quite likely that they will be released from transplanted cells. Murine ERVs can be activated through graft-versus-host reactions (Hirsch *et al.*, 1970; Levy *et al.*, 1977). Therefore the next logical step would be to investigate the infectivity of PoERVs in various human cell cultures. It is very difficult to experimentally predict pathogenic potential. However, the host range and tissue tropism of a virus is an important determinant of this (Hunter, 1997). An environment where cell to cell transmission can occur also greatly increases the chance of viral insertion near a proto-oncogene (Coffin *et al.*, 1997). If PoERV can be placed into a known interference group it will facilitate prediction of this host range and the potential receptor usage of the virus.

4.2 Materials and methods

4.2.1 Materials

4.2.1.1 Raji 4 cells

The Raji 4 cell clone was a gift from Q-One Biotech Ltd. Raji 4 had been derived from human Raji cells cocultivated with irradiated PK15 cells which are reported to release PoERV (Armstrong *et al.*, 1971; Lieber *et al.*, 1975a; Todaro *et al.*, 1974). This clone tested positive for RT activity, indicating that it had been infected with PoERV released from the PK15 cells (Q-One Biotech Ltd. and Imutran Ltd., 1997).

4.2.2 Methods

4.2.2.1 Preparation of Raji 4 cells for electron microscope (EM) study

The Raji 4 cell suspension was mixed 1:1 with paraformaldehyde/glutamine fix and immediately centrifuged at 2000 rpm for 10min in a bench top centrifuge (Beckman). The supernatant was removed and fresh fix added. After 10min the pellet was gently loosened and left for a further 2 hours. The sample was then given to the Electron Microscopy Unit, Veterinary Pathology, University of Glasgow, for further processing and EM analysis.

4.2.2.2 Polymerase chain reaction (PCR) analysis of Raji 4 cells

Genomic DNA was prepared from Raji 4, SupT1 and J.Jhans cells as described in section 2.2.2.2. Each PCR was performed with 500ng of template, 30pmol of each primer, 1.5mM $MgCl_2$, 0.125mM of each dNTP, and 1U *Taq* DNA polymerase (Perkin Elmer). The primers, which were targeted to the *pol* region of PoERV, were a gift from Christine Haworth (Q-One Biotech Ltd.). The cycling parameters used were 30 cycles of 1min at 95°C, 1min at 55°C, and 1min at 72°C; and a final extension step of 10min at 72°C. Cycling reactions were carried out in 0.2ml reaction tubes in a programmable thermal

cycler (Perkin Elmer DNA Thermal Cycler). 5µl of each product was analysed by agarose gel electrophoresis.

4.2.2.3 Infection interference studies

Approximately 1 x 10^6 Raji or Raji 4 cells were mixed with 1ml of virus stock in the presence of 8µg/ml Polybrene. The viruses used were A-MLV, X-MLV, FeLV-B and FeLV-C. Equal numbers of cells were mock infected with just growth medium and Polybrene. After incubation at 37°C for 90min, 4ml fresh medium was added to the flasks, and the cells were passaged every 3-4 days. Genomic DNA was extracted as described in section 2.2.2.2 and analysed by PCR.

Each PCR was performed with 500ng of template DNA, 20pmol of each primer, 0.125mM of each dNTP, and 2.5U *Pfu* DNA polymerase (Stratagene, UK). 5% DMSO was also added to the FeLV-C reactions to increase the stringency of primer annealing. The primers used to amplify the specific envelope regions were as follows:

A-MLV	sense	5′ - CCT CGT TGT CTG AAG TAG TCC - 3′
	antisense	5′ - CAT GGC TCG TAC TCT ATG GGT - 3′
X-MLV	sense	5′ - GGA CGA TGA CAG ACA CCT TCC CTA - 3′
	antisense	5′ - CAA CCA TCC CCA ATA TCG GGT TCT - 3′
FeLV-B/-C	sense	5′ - CGT GGA ACT TAG TGT TTC TG - 3′
FeLV-B	antisense	5′ - GAA TGG TTC CTG ATC TGA AG - 3′
FeLV-C	antisense	5′ - ATA ACG TGC CCA AGA TCT TG - 3′

The cycling parameters used for A-MLV and X-MLV were 5min at 94°C; 35 cycles of 50s at 94°C, 50s at 55°C, and 50s at 72°C; and a final extension step of 7min at 72°C. The parameters used for FeLV-B and FeLV-C were 5min at 94°C; 30 cycles of 30s at 94°C, 45s at 50°C, and 30s at 72°C; and a final extension step of 7min at 72°C.

4.3 Results

4.3.1 Presence of PoERV in the Raji 4 culture

Electron microscopy of the Raji 4 cells revealed the presence of virus-like particles roughly 80nm in diameter, as shown in Fig. 4.2. Recently budded particles can be seen close to the plasma membrane, indicated by the small arrows. Mature extracellular particles are also visible in B, indicated by the large arrow. These particles have the typical appearance of C-type retroviruses (Coffin, 1990; Vogt, P.K., 1997), consisting of a well defined envelope surrounding an angular electron dense core. A less condense core is visible within the immature particles at the plasma membrane. Env protein spikes are just discernible in B, again indicated by the large arrow.

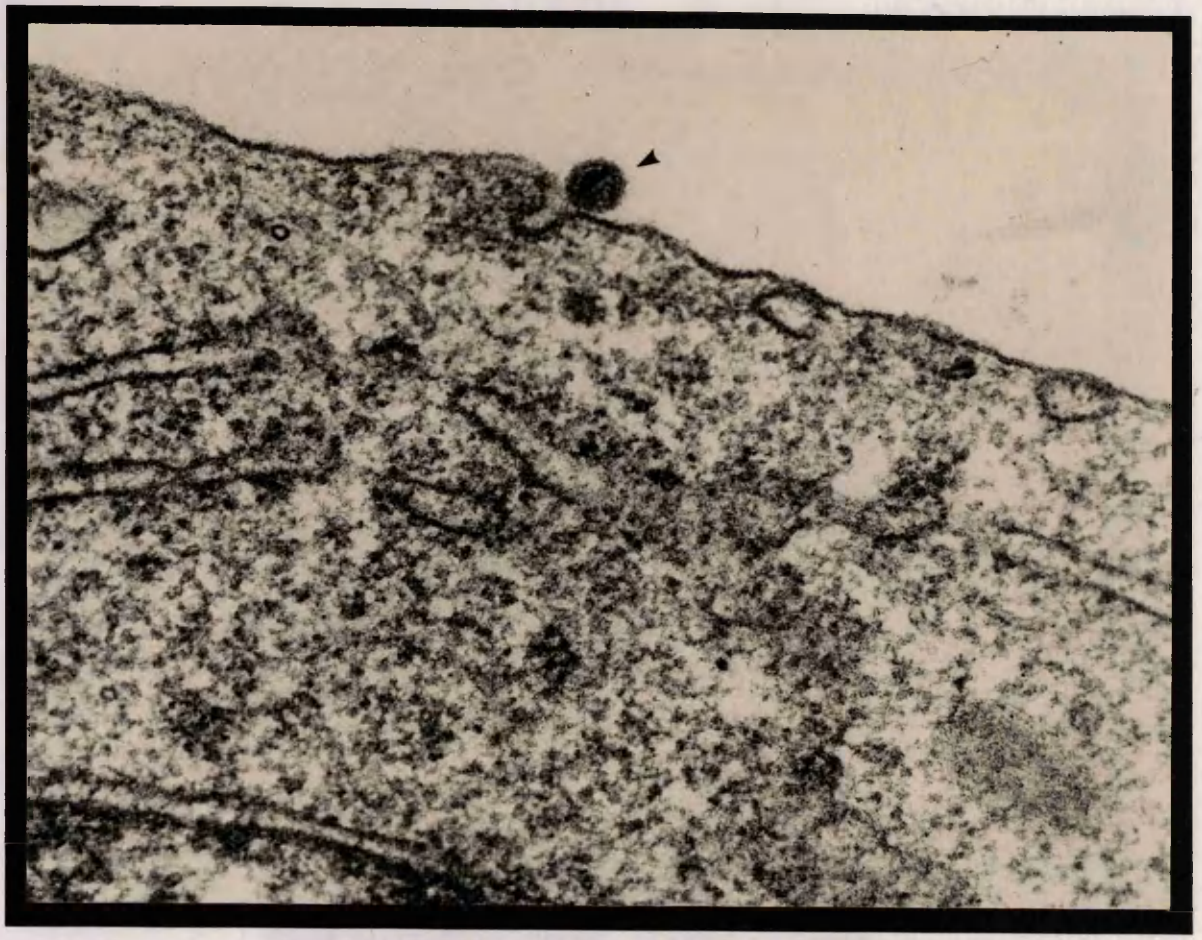
To demonstrate the presence of PoERV proviral DNA in the Raji 4 cells, PCR analysis was carried out on genomic DNA prepared from these cells. PCR analysis using PoERV *pol* based primers was expected to produce a band at 468bp in the presence of this provirus. As can be seen in Fig. 4.3, a band at this size was obtained from Raji 4 genomic DNA (lane D), but not from DNA taken from the other, uninfected, human lymphocyte cell lines SupT1 and J.Jhans. This suggests that the virus present in the Raji 4 culture may be PoERV.

4.3.2 Infection interference

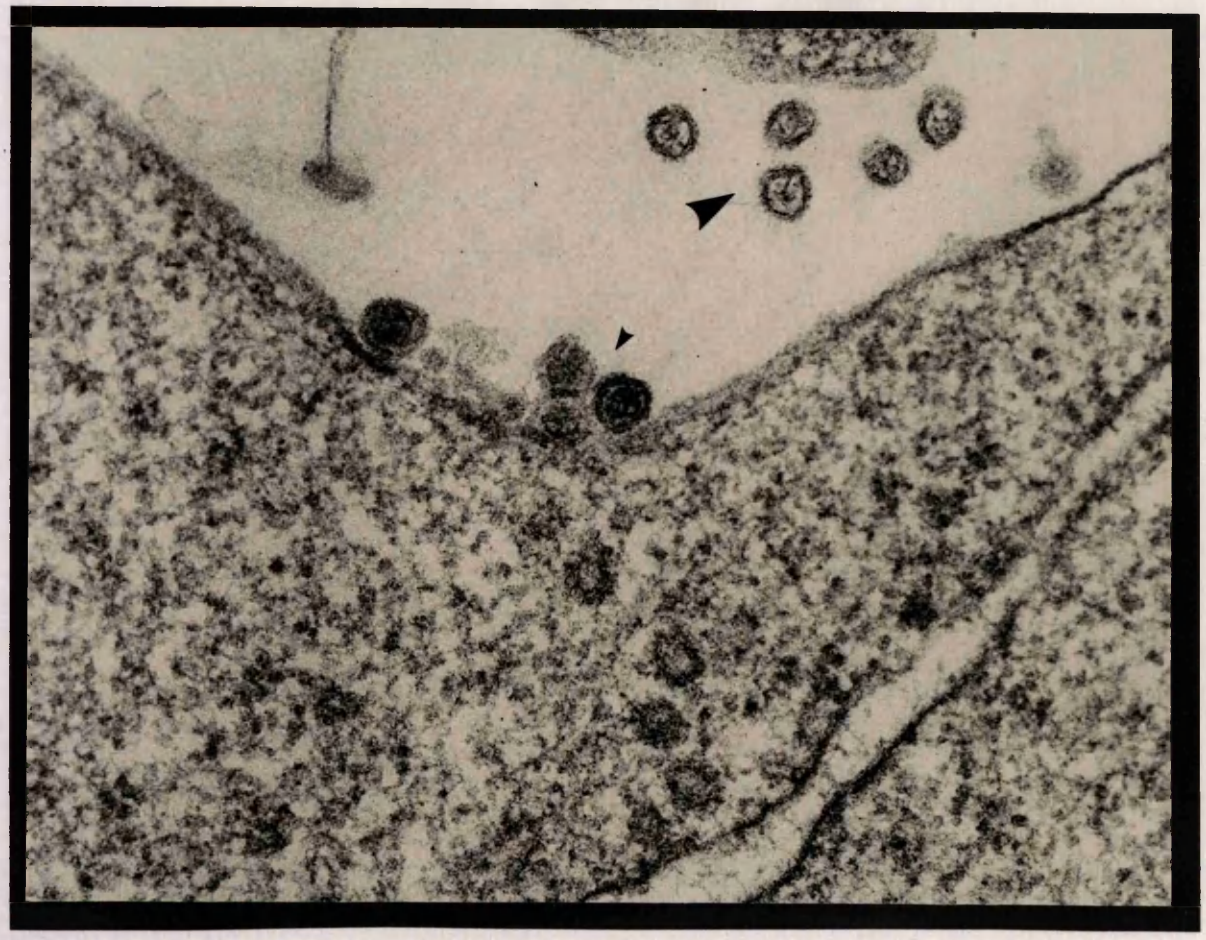
It is possible that PoERV recognises the same receptor as another retrovirus. As mentioned previously (section 1.2.1), this would result in infection interference. As an initial attempt to place PoERV into an interference group, Raji and Raji 4 cells were both tested for their susceptibility to infection by other human-tropic C-type retroviruses. If PoERV is in the same interference group as one of these viruses, its presence in the Raji 4 culture would be expected to prevent infection.

Raji and Raji 4 cells were incubated with each of the four viruses, A-MLV, X-MLV, FeLV-B and FeLV-C. Infection was then detected by PCR, using primers targeted to *env* regions specific to each virus. The results of these PCRs are shown in Fig. 4.4 and

A



B



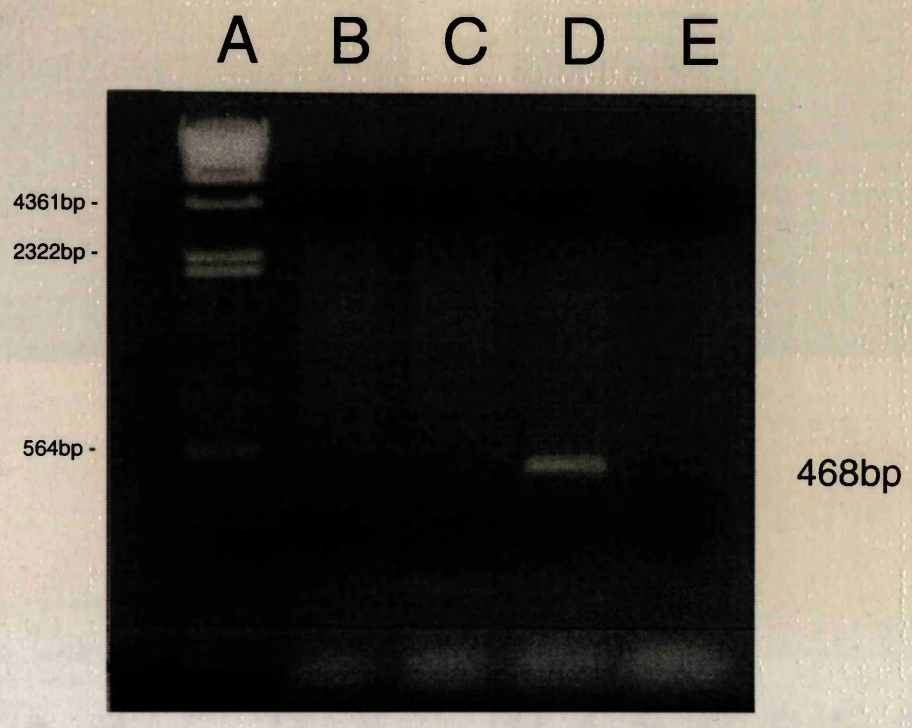


Fig. 4.3. PCR analysis of human cell lines using primers targeted to the *pol* region of PoERV. A) λHindIII DNA markers (Gibco, UK). B) SupT1 cells. C) J.Jhans cells. D) Raji 4 cells. E) Water control.
Samples were run on a 1% agarose gel.

Fig. 4.2. Previous page. EM photographs of cells taken from the Raji 4 culture (X45,000). A) Recently budded virus particle, indicated by the small arrow. B) Recently budded particles, indicated by the small arrow, and mature extracellular virus particles, indicated by the large arrow.

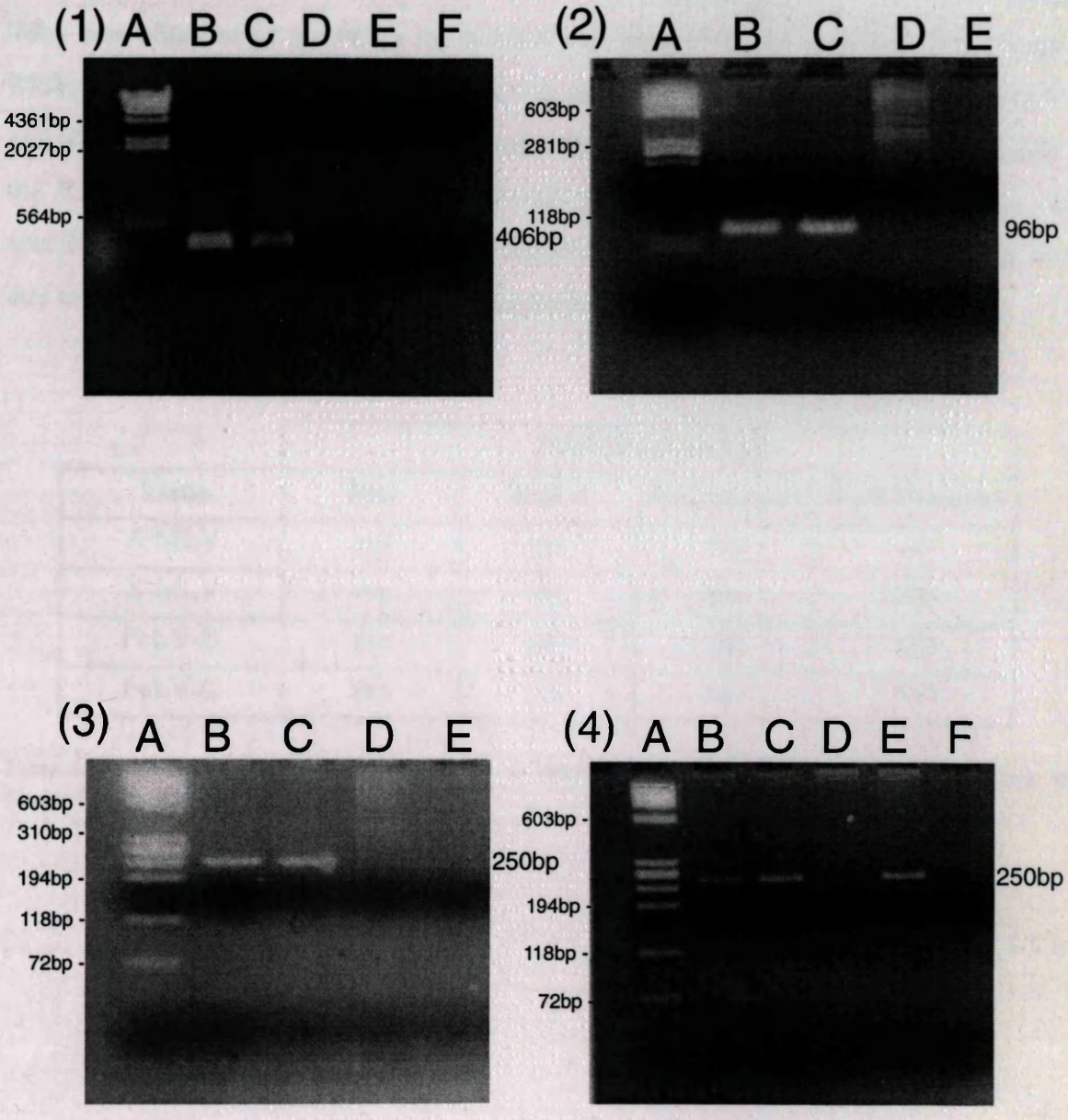


Fig. 4.4. PCR analysis of Raji and Raji 4 cells infected with mammalian C-type retroviruses.
(1) PCR with the A-MLV *env* primers. A) λHindIII DNA markers (Gibco, UK). B) Raji 4 cells infected with A-MLV. C) Raji cells infected with A-MLV. D) Noninfected Raji 4 cells. E) Noninfected Raji cells. F) Water control.
(2) PCR analysis with the X-MLV *env* primers. A) ϕX174 DNA markers (Gibco, UK). B) Raji 4 cells infected with X-MLV. C) Raji cells infected with X-MLV. D) Noninfected Raji cells. E) Water control.
(3) PCR analysis with the FeLV-B *env* primers. A) ϕX174 DNA markers. B) Raji 4 cells infected with FeLV-B. C) Raji cells infected with FeLV-B. D) Noninfected Raji cells. E) Water control.
(4) PCR analysis with the FeLV-C *env* primers. A) ϕX174 DNA markers. B) Raji 4 cells infected with FeLV-C. C) Raji cells infected with FeLV-C. D) Noninfected Raji cells. E) Positive control plasmid containing FeLV-C *env*. F) Water control
The A-MLV samples were run on a 1% agarose gel, the rest on 4% agarose gels.

summarised in Table 4.1. For each virus, a band at the expected size was obtained with DNA taken from infected Raji and Raji 4 cells, but not from uninfected cells. A band at 406bp was obtained in the PCRs for A-MLV *env* (Fig. 4.4.1 lanes B and C). A band at 96bp was obtained in the PCRs for X-MLV *env* (Fig. 4.4.2 lanes B and C). A band at 250bp was obtained in the PCRs for FeLV-B *env* (Fig. 4.4.3 lanes B and C) and FeLV-C *env* (Fig. 4.4.4 lanes B and C). This indicates that all four viruses were able to replicate in the Raji 4 cell culture, revealing no obvious interference with PoERV. Since no virus specific sequences were detected in the negative controls these results are unlikely to be due to DNA contamination. Each PCR was repeated at least three times.

	Infection detected			
Virus	**Raji**	**Raji 4**	**Raji (mock)**	**Raji 4 (mock)**
A-MLV	yes	yes	no	no
X-MLV	yes	yes	no	ND
FeLV-B	yes	yes	no	ND
FeLV-C	yes	yes	no	ND

Table 4.1. Susceptibility of Raji and Raji 4 cells to infection by mammalian C-type viruses. Infection was detected by PCR as described in the text.

4.4 Discussion

The EM photographs presented in Fig. 4.2 and PCR results presented in Fig. 4.3 imply that the Raji 4 cell culture is infected with PoERV. The particles shown in the EM photographs have C-type morphology, as expected for this virus (Armstrong *et al.*, 1971; Moennig *et al.*, 1974). The presence of extracellular virions indicate that the virus is transcriptionally active in these cells, confirming the previously recorded RT activity (Q-One Biotech Ltd. and Imutran Ltd., 1997). Although it is possible that contaminating PK15 cells are still present in the cell culture this seems unlikely as these cells were irradiated before the coculture was set up. In addition PK15 cells are adhesive and would not be expected to grow well in the suspension cell culture. The amplification of a band at the expected size during PCR analysis of Raji 4 DNA suggests that PoERV sequences are present. The absence of this band in the reactions carried out in parallel with DNA from other human lymphocyte cell lines indicates that this result was not due to the primers binding to human DNA sequences. The conclusion that PoERV can infect and replicate in human cells has been confirmed by studies carried out by other groups. In addition to the Raji cell line, PoERV infection of human kidney, lung, muscle and T-cell lines by cocultivation with PK15 cells has been demonstrated (Patience *et al.*, 1997a). PoERV has also been obtained from mitogenically activated primary blood mononuclear cells taken from miniature pigs (Wilson *et al.*, 1998). This virus was reported to be capable of infecting human kidney cells in a coculture.

All of the retroviruses tested in the interference assay (Table 4.1) were able to infect both the Raji and the Raji 4 cells. As the Raji 4 cells are believed to be infected with PoERV, this result indicates that PoERV cannot be placed in any of the four interference groups represented by these viruses. However, it is possible that PoERV replication occurs at too low a level in the human cells to produce significant receptor interference. A fall in the infection efficiency of the superinfecting virus may not have been detected since the PCR analysis used here was not quantitative. This experiment would also have been unable to detect any non-reciprocal interference as no attempt was made to infect other virus producing cells with PoERV. However, the conclusion that PoERV lies outside these four interference groups has also been reached by others. Takeuchi *et al.* (1998) have recently shown that MLV based vectors bearing the PoERV Env are able to infect 293 cells infected

with these four viruses. The same result was obtained with cells infected with RD114, GALV or FeLV-C. This lack of interference suggests that PoERV recognises a novel receptor(s) on 293 cells not used by other mammalian C-type retroviruses. It still remains possible that PoERV will belong to an interference group containing viruses from other genera since the RD114 group contains both C- and D-type viruses (Sommerfelt and Weiss, 1990).

It is perhaps surprising that no interference is observed with FeLV-B, since this virus recognises the same receptor as GALV, a virus that shares close homology with PoERV (Q-One Biotech Ltd. and Imutran Ltd., 1997). It is possible that GALV and PoERV evolved from the same ancestral murine ERV. However, closely related avian and mammalian C-type retroviruses often recognise different receptors (Hunter, 1997). As described in section 1.2.1, FeLV-B and A-MLV have highly homologous Env proteins that bind to different receptors, the phosphate symporters GLVR-1 and Ram-1 respectively (Tailor and Kabat, 1997). As the receptor molecules recognised by mammalian C-type retroviruses are often very similar proteins (section 1.2.1), small changes in the viral Env proteins can result in a change in receptor recognition, and therefore a potentially significant change in host range and pathogenicity. The differences in generation time between mice and pigs imply that viruses in the respective genomes will evolve at different rates (Benveniste and Todaro, 1975). The porcine ERV is likely to be much more homologous to the ancestral virus than the ERVs that remained in the murine genome at the time of the cross species infection that gave rise to GALV. It is also likely that, subsequent to the initial cross species transfer and germ line insertion, PoERV has evolved to adapt to its new host. This may have resulted in a change in receptor recognition. Of course it is possible that GALV and PoERV evolved from different, related viruses that recognised different receptors.

The observation that PoERV can infect and replicate in human cells *in vitro* indicates that these viruses may indeed pose a threat to xenotransplant recipients. As mentioned in section 4.1, such patients are more vulnerable to infections as a result of the transplant procedure and the possible suppression of normal complement-dependent viral lysis. Although no evidence of infection has yet been found in patients exposed to pig tissues in medical procedures (Heneine *et al.*, 1998; Paradis *et al.*, 1999; Patience *et al.*, 1998), it still

remains possible that rare infections will occur. The risk of infection is likely to be increased with actual organ transplants, as the PoERV releasing tissue will remain in close contact with human cells for long periods of time, perhaps years. In addition, none of the patients included in these studies had been exposed to transgenic pig tissues, which are likely to pose a greater risk of virus transmission. It is therefore important to determine the host cell range and pathogenic potential of these viruses. Even if PoERV is non-pathogenic and does not replicate to high titres in the human body, the ability to enter cells raises the possibility of recombination with human exogenous and endogenous retroviral sequences. In this way, these viruses could adapt to human tissues and perhaps increase their pathogenicity.

CHAPTER 5

Chapter 5 Sequencing of a novel PoERV *env* gene

5.1 Introduction

The retroviral envelope glycoprotein recognises and binds to the receptor molecule, and mediates fusion of the viral and host cell membranes initiating the infection cycle. The binding of the specific receptor molecule is thought to lead to a conformational change within the Env protein which enables the fusion process to proceed. Although viruses may bind weakly to nonpermissive cells, fusion does not occur in the absence of a specific receptor. This means that the host range and tissue tropism of a retrovirus is highly dependent upon its envelope glycoprotein. (Hunter, 1997; Vogt, V.M., 1997).

Inbred strains of mice possess closely related MLV-like ERVs that differ in their host ranges (Boeke and Stoye, 1997). As would be expected, the major differences between these viruses are found within the *env* gene. It seems possible that porcine genomes also contain several strains of PoERV, differing in *env* and host range. It is therefore important to identify and analyse each of these for their potential risk in xenotransplantation.

The function of a protein is dictated by its primary structure. This can be easily predicted from the nucleotide sequence of the gene. However, it is difficult to determine the three-dimensional structure or activity of a protein just from its amino acid composition. The presence of hydrophobic residues can indicate membrane spanning regions, and cysteine residues disulphide bonding. Several computer-assisted algorithms such as the Chou-Fasman (Chou and Fasman, 1978a; Chou and Fasman, 1978b) have been developed in order to predict secondary structure motifs including the α-helix and β-sheet. Unfortunately such predictions are not highly accurate (Myers, 1997). However, obtaining the sequence of a protein does allow comparisons to be made with related proteins for which structural and functional information has already become available. The most highly characterised Env proteins of the mammalian C-type viruses are those of the MLV-related viruses, particularly the ecotropic MoMLV.

5.1.1 Env structure

As mentioned in section 1.1.4, Env is initially translated as a precursor polyprotein. The signal peptide at the amino-terminal of the protein binds to a signal recognition particle which then targets further translation to the RER. Following cleavage of the signal peptide the protein is transferred to the Golgi apparatus. Here the polyprotein is cleaved by cellular proteases to form the surface glycoprotein SU and the transmembrane protein TM (Vogt, V.M., 1997).

While within the RER Env is cotranslationally N-glycosylated by cellular enzymes present in the lumen. These enzymes attach mannose-rich oligosaccharides to the asparagine residues in the consensus motifs Asn-X-Ser and Asn-X-Thr (Hunter, 1997), where X is any residue except proline and tryptophan (Rademacher *et al.*, 1988). The number and position of these residues can vary significantly between virus groups, although a high level of conservation may occur within a group. Mutational analysis of the Friend-MLV SU protein has shown that only one out of eight glycosylation sites is necessary for viral growth in cell culture (Kayman *et al.*, 1991). This site is highly conserved among mammalian C-type viruses, and its removal probably interferes with the processing of the precursor protein. Glycosylation may have a role in the structure of the protein, providing hydrophilicity which helps to stabilise correct folding. It may also reduce immunogenicity and proteolytic digestion by masking susceptible amino acid sequences. This would explain why sites are conserved even though they appear unnecessary for growth *in vitro*. Some retroviruses, including HIV and MLV are also modified by the addition of O-linked oligosaccharides. The role of this is unknown (Hunter, 1997).

The Env proteins are arranged in oligomers on the surface of the virion. These are probably trimers, although dimers and tetramers have also been suggested (Hunter, 1997; Tucker *et al.*, 1991). The crystal structure of a fragment of the MoMLV TM reveals this protein to be trimeric (Fass *et al.*, 1996). The oligomers form in the RER via domains which appear to be within the ectodomain of the precursor protein (Hunter, 1997; Tucker *et al.*, 1991). The TM proteins of ASLV and HIV will form oligomers when expressed in the absence of SU. The formation of oligomers appears to have a role in the intracellular transport of the protein, presumably by enclosing hydrophobic sequences within the core of the structure.

Oligomerisation is probably also important for providing mechanical rigidity, and for achieving membrane fusion (Hunter, 1997).

5.1.2 Structure and function of the SU protein

The positions of cysteine residues in the SU proteins of related viruses tend to be conserved (Hunter, 1997). The amino-terminal domains of the SU proteins of mammalian C-type viruses all contain eight conserved cysteines, suggesting a common structure (Battini *et al.*, 1992). Disulphide-bonding patterns have been deduced for Friend-MLV (Linder *et al.*, 1992) and Friend mink cell focus-inducing virus (Linder *et al.*, 1994) in which the bonds are found to form between cysteines reasonably close together in the primary sequence. There are no bonds linking the amino- and carboxy-terminals. A similar pattern has been observed with HIV-1 gp120 (Hunter, 1997). The loops that are formed by the disulphide-bonding tend to contain highly variable amino acid sequences. These regions are thought to be involved in the receptor specificity of the mammalian C-type viruses.

Within the amino-terminal region of the mammalian C-type SU proteins there are two hypervariable regions, VRA and VRB, believed to be involved in receptor binding. VRA is approximately 50 amino acid residues from the amino-terminal, and 35 to 80 amino acids long. It contains a peptide of variable length, followed by a hydrophilic region containing two conserved cysteine residues. VRB is situated about 40 amino acids after VRA, and is 22 to 29 amino acids long. It contains one conserved cysteine which is followed by a sequence of polar residues (Battini *et al.*, 1992; Hunter, 1997). VRA can be further subdivided into three segments consisting of two disulphide-linked loops and an interloop domain (Han *et al.*, 1998). The first of these loops is highly variable; the ecotropic MLV proteins possess two extra disulphide bonds within this region. VRB forms just one disulphide loop in the majority of SU proteins, although those of A-MLV, 10A1 and FeLV-B contain two more cysteine residues which may form another small loop (Han *et al.*, 1998).

The evidence that these regions are involved in receptor interactions comes mainly from studies of Env chimaeras. By swapping various lengths of sequences from this region between A- and E-MLV SU proteins, Morgan *et al.* (1993) were able to show that the

critical sequences for host range are within the first 157 amino acids of the A-MLV SU, and the first 88 of the E-MLV SU. These chimaeras lacked full activity, suggesting that other regions may also be involved to some extent. Battini *et al.* (1995) also found that the amino-terminal regions of the A- and E-MLV SUs are sufficient for receptor binding and interference, although additional sequences are required by the polytropic MLV (P-MLV) and X-MLV proteins. In another study Battini *et al.* (1992) formed chimaeras by swapping the VRA, VRB and remaining carboxy-terminal halves of the MLV proteins. They found that host range was determined primarily, but not solely, by the specific VRA. A-MLV proteins containing a foreign VRB were able to bind to receptors on murine cells but were unable to recognise the allelic form present on canine cells. The authors suggest this shows that although VRA is the major determinant of host range for A-MLV, VRB also contributes to the stability of the receptor recognition complex. This is supported by the observation that chimaeras containing only the VRA of A-MLV were inactive. Also, although substitutions of segments in the variable regions of A-MLV SU with an epitope tag from VSV-G only abolished receptor recognition when they were within the first loop of VRA, substitutions in the remainder of VRA and in VRB also affected binding (Battini *et al.*, 1998). While VRB appears to be important for A-MLV, sequences in the carboxy-terminal half seem to have a role in the function of the xenotropic and polytropic Env proteins. When these regions were replaced with that of the amphotropic protein, the resulting chimaeras were inactive. Receptor choice of xenotropic and polytropic based chimaeras was unaffected by swapping VRB (Battini *et al.*, 1992). These accessory segments may be required for additional binding interactions, or for appropriate folding to allow optimal receptor contact (Battini *et al.*, 1998). Mutations within the ectodomain of MoMLV TM can affect receptor binding, presumably by influencing the conformation of the binding regions of the SU protein (Berkowitz and Goff, 1993).

Analysis of the crystal structure of the receptor binding domain of Friend-MLV SU has shown that VRA and VRB form antiparallel α-helices (Fass *et al.*, 1997). The close association of these two regions provides some explanation for the inability to produce functional chimaeras between ecotropic and non-ecotropic Env proteins. Large differences in length and sequence would not be easily tolerated in such a structure.

Between the amino-terminal receptor binding domain and the more conserved carboxy-terminal of SU proteins there is a proline-rich region (PRR), approximately 50 amino acids long. This consists of a fairly conserved domain of 14 or 15 amino acids, followed by a region that is highly variable both in length and sequence. The positions of the proline residues vary between viruses, but they make up about one third of the residues in this domain (Hunter, 1997; Wu *et al.*, 1998). Limited proteolysis of P- and E-MLV SU glycoproteins leads to cleavage at both ends of this region, suggesting that it is not just a link but a functional domain (Linder *et al.*, 1994). The presence of the proline residues may permit the formation of contiguous β-turn helices (Lavillette *et al.*, 1998). It has been suggested that the PRR is necessary for the stabilisation of the receptor binding domain (Battini *et al.*, 1992; Battini *et al.*, 1995; Tailor and Kabat, 1997). The presence of this region has been shown to be required for the binding of P-MLV and E-MLV SUs (Battini *et al.*, 1995) and to increase infectivity of vectors bearing A-MLV SU chimaeras (Tailor and Kabat, 1997). Although it is possible that the PRR also interacts with a receptor site it seems more likely that it has a role in the induction of appropriate structure of the amino-terminal binding domain (Battini *et al.*, 1995).

Mutation analyses of the PRR have shown that this region is not only important for stability of receptor interactions but is involved in the function of the whole Env complex. Although the amphotropic PRR can substitute for the differing ecotropic region, and truncations of up to 29 amino acids in the variable region have little effect on vector titre, further truncations result in decreased levels of precursor cleavage into SU and TM (Wu *et al.*, 1998). These mutants also show low levels of cell-cell fusion and transduction efficiency. Deletion of the conserved segment of the PRR produces a protein that is not incorporated into virions at any detectable level, despite normal precursor processing. High levels of SU are shed into the culture medium suggesting a reduced stability in the interaction between SU and TM. This is supported by the observation that the mutant is temperature dependent and low levels of protein could be detected on virions produced at 32°C.

The PRR also appears to have a role in the fusion process. Lavillette *et al.* (1998) have shown that various modifications in this region can result in either high cell-cell fusion activity accompanied by decreased SU-TM stability, or low cell-cell fusion and high virus-

cell fusion accompanied by SU-TM stability. In the same study it was found that chimaeric SU proteins in which the PRR is paired with the corresponding binding domain or downstream regions are more stable. The authors suggest a model in which the two functions of the PRR are to stabilise the structure of the SU, and to mediate the changes involved in fusion initiation. The variation among PRRs would therefore be expected to reflect the differences in structure required to interact with different receptors. The predicted β-turn helical structure is compatible with an ability to transmit the small deformations resulting from receptor interactions to the carboxy-terminal fusion domain, thereby triggering the necessary conformational changes.

5.1.3 Structure and function of the TM protein

The TM subunit of Env provides the membrane anchor for the protein, and is required for fusion. It is connected to SU via noncovalent interactions and, in some viruses such as ASLV, by disulphide bonds (Hunter, 1997). The TM protein has several predicted structural features. The first 35 amino acids are thought to form the fusion peptide (Jones and Risser, 1993). Mutations within this stretch of hydrophobic residues in the TM of MoMLV have been shown to influence fusogenicity (Jones and Risser, 1993; Zhao *et al.*, 1998). The carboxy-terminal end of this peptide is rich in glycine and threonine residues. This is followed by a heptad repeat region containing a putative leucine zipper motif (Delwart *et al.*, 1990), and then a region containing a CX_6CC motif highly conserved among TM proteins. These cysteine residues are also present in the envelope proteins of the filovirus family (Weissenhorn *et al.*, 1998). Three regions of the protein are predicted to form α-helices (Zhao *et al.*, 1998). These are the heptad repeat sequence, a region towards the carboxy-terminal end of the ectodomain, and the transmembrane domain with the first 16 amino acids of the cytoplasmic tail.

As mentioned previously (section 1.1.4) the TM protein is cleaved by the viral protease near the carboxyl-terminus, releasing the 16 amino acid R-peptide. This occurs during and after viral budding, and both forms of TM coexist in the virion (Green *et al.*, 1981; Januszeski *et al.*, 1997). R-peptide cleavage renders the protein fusion competent, and has been shown to enhance cell-cell fusion by MoMLV Env (Januszeski *et al.*, 1997; Rein *et al.*, 1994). Cleavage of the cytoplasmic tails of the proteins from other retroviruses, for

example Mason-Pfizer monkey virus (Brody *et al.*, 1994), also enhances fusogenicity. Removal of the R-peptide is necessary for function as removal of the cleavage recognition site renders the protein almost completely defective (Rein *et al.*, 1994). It seems likely that Env is synthesised with this peptide intact in order to prevent premature activation of fusion potential prior to viral release. A fusogenic protein may be toxic to the host cell or may be inactivated by inappropriate interaction with receptor molecules.

Both the ectoplasmic and transmembrane regions of TM are necessary for membrane fusion and cell transduction (Ragheb and Anderson, 1994). Mutant proteins in which the transmembrane domain has been replaced with glycophospho-lipid-anchoring (GLA) sequences can be rescued by the presence of wild type TM in the oligomers. However this is not true of mutants lacking the ectodomain, suggesting that necessary intramolecular interactions may take place within this region, though it should be noted that the constructs used in this study also had altered SU carboxy-terminals.

More detailed analyses of the functional domains of the TM protein have involved the use of point mutations and small deletions mapped to specific regions. A number of mutations of hydrophobic residues in the leucine zipper region, even conservative changes, have been shown to abolish infectivity and cell fusion activity of E-MLV Env (Ramsdale *et al.*, 1996). The mutation of hydrophilic residues reduces infectivity but has little or no effect on rat XC cell fusion. It seems likely that the hydrophobic residues are necessary for the formation of a structure required for the fusion process, perhaps forming a fusion pore within the trimer (Ramsdale *et al.*, 1996; White, 1992). Zhao *et al.* (1998) found that fusion-defective proteins containing mutations within the predicted helical regions could by complemented by a protein with a mutation in the glycine/threonine (GT) rich area, and to a lesser extent by one with a mutation in the fusion peptide. However the helix mutants and the amino-terminal mutants were unable to complement themselves, indicating that these are two separate complementation groups. The authors interpret this as indicative of a functional interaction between the three α-helices. Given the ability of R-peptide cleavage to enhance fusogenicity in *trans*, R-less versions of these mutants were tested for their ability to produce syncytia when coexpressed with wild type protein. Only the GT rich region mutant produced syncytia under these conditions. The same R-less mutant was also able to enhance fusion activity of the helix mutants, though reciprocal complementation

was not observed. This suggests that the *trans* enhancement of fusion by R-peptide cleavage has specific *cis* requirements. Zhao *et al.* believe that this cleavage may enhance fusion by removing a conformational constraint of the cytoplasmic tail which is then transmitted to the rest of the molecule, probably via an interaction between the two ectodomain α-helical regions.

After removal of the R-peptide further truncations of the carboxy-terminal end of the MoMLV TM protein, even as far as complete removal of the cytoplasmic tail, have very little effect on the ability to cause cell-cell fusion (Januszeski *et al.*, 1997; Ragheb and Anderson, 1994). In contrast, internal deletions in the membrane proximal region of the cytoplasmic tail result in reduced fusogenicity (Januszeski *et al.*, 1997). This suggests that residues in this area may be important in the fusion process, particularly since more distal deletions can enhance fusion. Januszeski *et al.* propose that the residues Leu 602 and Val 603 have an important role. Several substitutions and deletions in this area reduce fusogenicity of the full length protein. The fusion defect in some of these could not be restored by R-peptide cleavage.

Although truncation of the cytoplasmic tail has little effect on cell-cell fusion, this is not always true of vector infectivity (Januszeski *et al.*, 1997; Ragheb and Anderson, 1994). This reduction in titre could be due to reduced expression and/or association with virus, or it may reflect a role for this region in post fusion events. Certain mutations in the cytoplasmic tail, especially in the region amino-terminal to the R-peptide cleavage site, permit extensive cell-cell fusion but significantly reduce vector titre (Januszeski *et al.*, 1997). A group of these mutants in which membrane proximal residues have been deleted or substituted are not efficiently incorporated into virions although they are expressed on the cell surface. This suggests that the cytoplasmic tail of MoMLV TM has a role in Env incorporation (discussed in section 1.5.1). It is also possible that these alterations result in a secondary structure that prevents the protein associating with the vector particle.

The observations that several mutations and truncations can be introduced into the cytoplasmic tail of MoMLV TM without abolishing infectivity (Granowitz *et al.*, 1996; Januszeski *et al.*, 1997; Ragheb and Anderson, 1994) suggests that this domain is not absolutely required. However, such mutants have been shown to rapidly revert to the wild

type during replication (Granowitz *et al.*, 1996). This implies that the wild type sequence is favoured for rapid replication in cell culture.

Several retroviruses, including MoMLV, possess a basolateral membrane targeting signal in the cytoplasmic tail of TM (Lodge *et al.*, 1997a). The consensus for this signal is Y-X-X-aliphatic/aromatic. Its conservation may reflect the importance of polarised budding in the viral life cycle. Basolaterally targeted viruses also tend to bud at sites of intercellular contact. This signal may therefore aid the cell-to-cell transmission of virus.

5.1.4 Sequencing of PoERV *env*

At the time that this work was undertaken the sequence of only one porcine endogenous virus was available. This sequence was obtained at Q-One Biotech Ltd from virions released from PK15 cells (Q-One Biotech Ltd. and Immutran Ltd., 1997). As it is quite likely that other ERVs are present in the porcine genome, it is important that each potentially infectious virus is identified. Since Env is the main determinant of host range the cloning and sequencing of PoERV *env* genes is a useful aid to the assessment of the infection risk of these viruses. Sequence analysis also enables determination of functional and evolutionary relationships between viral proteins. Although it is very difficult, if not impossible at present, to predict receptor usage from a sequence, identification and cloning of authentic *env* genes allows host ranges to be determined using procedures such as pseudotyped vector analysis.

5.2 Materials and methods

5.2.1 Materials

5.2.1.1 Primers

The primers used for cloning and sequencing were as follows:

U5657	5′ - CGA AGG CGG CCG CAC CTG GAT CCA TGC ATC - 3′
xhis	5′ - CTA ATC TTA GAA CGC GGC CGC CTA GAG GTC AGT TT - 3′
T7	5′ - TAA TAC GAC TCA CTA TAG G - 3′
T3	5′ - ATT AAC CCT CAC TAA AGG GAA G - 3′
621FOR	5′ - GTG GAT AAA TGG TAT GAG CT - 3′
1869FOR	5′ - ACA GTT GGG CCT TGC TTA AT - 3′
1316REV	5′ - GAC CAT GAC ACA GAA ATC TT - 3′

5.2.2 Methods

5.2.2.1 Cloning of the *env* gene from Raji 4 cells

Genomic DNA was extracted from Raji 4 cells as described previously (section 2.2.2.2). Primers were designed, based on the PoERV sequence obtained by Q-One Biotech, to bind to sequences spanning the start and stop signals for *env* transcription. These primers were U5657 (sense primer) and xhis (antisense primer). Synthetic Not I restriction sites are underlined. Each PCR was performed with 500ng of template, 20pmol of each primer, 0.125mM of each dNTP and 2.5U *Pfu* DNA polymerase. The cycling parameters used were 4 mins at 96°C; 30 cycles of 30s at 96°C, 45s at 55°C, and 5 mins at 72°C; and a final extension step of 7 mins at 72°C. The PCR product obtained was then gel purified.

The PCR product and the eukaryotic expression vector pCI-neo were both digested with Not I (New England Biolabs). The vector was dephosphorylated with CIAP as described previously (section 2.2.3.1). Both digests were gel purified (section 2.2.2.5) and ligated using T4 DNA ligase at 14°C overnight. The ligation reaction was then used to transform DH5α cells as described previously (see 2.2.3.2).

5.2.2.2 Sequencing of the *env* gene

Sequencing was carried out using the dideoxy chain termination method (Sanger *et al.*, 1977). The reactions were prepared using the Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech, UK), and primers labelled with the fluorescent dye IRD-41. Initial sequence was obtained with primers targeted to the T7 and T3 RNA polymerase promoters positioned either side of the insert in pCIneo. The cycling parameters used were 5 mins at 95°C, and 25 cycles of 30s at 95°C, 30s at 55°C, and 30s at 72°C. The sequence thus obtained was then used to design three more primers which allowed the rest of the gene to be sequenced. These were sense primers 621FOR and 1869FOR, and antisense primer 1316REV. The cycling parameters used for 621FOR and 1869FOR were identical to those for the T7 and T3 primers. The parameters used for 1316REV were 5mins at 95°C, and 25 cycles of 30s at 98°C, 30s at 50°C, and 30s at 72°C. The reactions were run on a 60ml denaturing polyacrylamide gel and sequenced using an automated fluorescent sequencer (LI-COR DNA Sequencer model 4000L, MWG-Biotech).

5.2.2.3 Sequence analysis

Sequence alignments and translations were made using the Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisc. The *env* sequences used for the protein alignment were obtained from Genbank.

5.3 Results

5.3.1 Cloning of the *env* gene from Raji 4 cells

In order to obtain an *env* clone for sequencing, PCR of Raji 4 genomic DNA was carried out using primers based on the PoERV sequence already available. This resulted in two products, approximately 2kb and 1kb in size. As the *env* product was expected to be around 2kb, the larger band was gel purified and inserted into pCI-neo.

5.3.2 Sequencing of the *env* gene

Initial sequencing of the two ends of the insert showed that it was a retroviral *env* gene. Therefore internal primers were designed and the full insert sequenced. The resulting DNA sequence is shown in Fig. 5.1. The assembled sequencing fragments are provided in the Appendices. Comparison with the Q-One Biotech sequence (Fig. 5.2) reveals that these *env* genes are very similar, but that this new sequence contains a 2 base pair insertion at position 572 and an insertion and substitution at 637. These differences do not appear to be due to PCR or sequencing errors. The *Pfu* enzyme used in the PCR has a 3' to 5' exonuclease proofreading activity, and therefore has the lowest error rate among the thermostable DNA polymerases used in PCR (Andre *et al.*, 1997; Cline *et al.*, 1996). The same sequence was obtained from a further 2 clones taken from separate PCR reactions. The insertions result in two frameshift events altering the predicted amino acids from position 192 to 215 in the SU protein (Fig. 5.2B). This area is downstream of the region corresponding to the VRB of MLV Env.

```
   1  TAGAGTCGAC CCGGGCGGCC GCACCTGGCT CCATGCATCC CACGTTAAGC
  51  TGGCGCCACC TCCCGACTCG GGGTGGAGAG CCGAAAAGAC TGAGAATCCC
 101  CTTAAGCTTC GCCTCCATCG CCTGGTTCCT TACTCTAACA ATAACTCCCC
 151  AGGCCAGTAG TAAACGCCTT ATAGACAGCT CGAACCCCCA TAGACCTTTA
 201  TCCCTTACCT GGCTGATTAT TGACCCTGAT ACGGGTGTCA CTGTAAATAG
 251  CACTCGAGGT GTTGCTCCTA GAGGCACCTG GTGGCCTGAA CTGCATTTCT
 301  GCCTCCGATT GATTAACCCC GCTGTTAAAA GCACACCTCC CAACCTAGTC
 351  CGTAGTTATG GGTTCTATTG CTGCCCAGGC ACAGAGAAAG AGAAATACTG
 401  TGGGGGTTCT GGGGAATCCT TCTGTAGGAG ATGGAGCTGC GTCACCTCCA
 451  ACGATGGAGA CTGGAAATGG CCGATCTCTC TCCAGGACCG GGTAAAATTC
 501  TCCTTTGTCA ATTCCGGCCC GGGCAAGTAC AAAATGATGA AACTATATAA
 551  AGATAAGAGC TGCTCCCCAT CAGACTTAGA TTATCTAAAG ATAAGTTTCA
 601  CTGAAAAAGG AAAACAGGAA AATATTCAAA AGTGGATAAA TGGTATGAGC
 651  TGGGGAATAG TTTTTTATAA ATATGGCGGG GGAGCAGGGT CCACTTTAAC
 701  CATTCGCCTT AGGATAGAGA CGGGGACAGA ACCCCCTGTG GCAATGGGAC
 751  CCGATAAAGT ACTGGCTGAA CAGGGGCCCC CGGCCCTGGA GCCACCGCAT
 801  AACTTGCCGG TGCCCCAATT AACCTCGCTG CGGCCTGACA TAACACAGCC
 851  GCCTAGCAAC AGTACCACTG GATTGATTCC TACCAACACG CCTAGAAACT
 901  CCCCAGGTGT TCCTGTTAAG ACAGGACAGA GACTCTTCAG TCTCATCCAG
 951  GGAGCTTTCC AAGCCATCAA CTCCACCGAC CCTGATGCCA CTTCTTCTTG
1001  TTGGCTTTGT CTATCCTCAG GGCCTCCTTA TTATGAGGGG ATGGCTAAAG
1051  AAAGAAAATT CAATGTGACC AAAGAGCATA GAAATCAATG TACATGGGGG
1101  TCCCGAAATA AGCTTACCCT CACTGAAGTT TCCGGGAAGG GGACATGCAT
1151  AGGAAAAGCT CCCCCATCCC ACCAACACCT TTGCTATAGT ACTGTGGTTT
1201  ATGAGCAGGC CTCAGAAAAT CAGTATTTAG TACCTGGTTA TAACAGGTGG
1251  TGGGCATGCA ATACTGGGTT AACCCCCTGT GTTTCCACCT CAGTCTTCAA
1301  CCAATCCAAA GATTTCTGTG TCATGGTCCA AATCGTCCCC CGAGTGTACT
1351  ACCATCCTGA GGAAGTGGTC CTTGATGAAT ATGACTATCG GTATAACCGA
1401  CCAAAAAGAG AACCCGTATC CCTTACCCTA GCTGTAATGC TCGGATTAGG
1451  GACGGCCGTT GGCGTAGGAA CAGGGACAGC TGCCCTGATC ACAGGACCAC
1501  AGCAGCTAGA GAAAGGACTT GGTGAGCTAC ATGCGGCCAT GACAGAAGAT
```

Fig. 5.1. DNA sequence of the *env* clone obtained from Raji 4 cells. Vector sequences are in bold type and underlined.

1551 CTCCGAGCCT TAAAGGAGTC TGTTAGCAAC CTAGAAGAGT CCCTGACTTC

1601 TTTGTCTGAA GTGGTTCTAC AGAACCGGAG GGGATTAGAT CTGCTGTTTC

1651 TAAGAGAAGG TGGGTTATGT GCAGCCTTAA AAGAAGAATG TTGCTTCTAT

1701 GTAGATCACT CAGGAGCCAT CAGAGACTCC ATGAACAAGC TTAGAAAAAA

1751 GTTAGAGAGG CGTCGAAGGG AAAGAGAGGC TGACCAGGGG TGGTTTGAAG

1801 GATGGTTCAA CAGGTCTCCT TGGATGACCA CCCTGCTTTC TGCTCTGACG

1851 GGGCCCCTAG TAGTCCTGCT CCTGTTACTT ACAGTTGGGC CTTGCTTAAT

1901 TAATAGGTTT GTTGCCTTTG TTAGAGAACG AGTGAGTGCA GTCCAGATCA

1951 TGGTACTTAG GCAACAGTAC CAAGGCCTTC TGAGCCAAGG AGAAACTGAC

2001 CTCTAG**GCGG CCGC**

Fig. 5.1. Cont.

```
          1                         →                                  50
q1env     AGGAATCCCC ACCTGGATCC ATGCATCCCA CGTTAAGCCG GCGCCACCTC
r4env     ~~GCGGCCGC ACCTGGCTCC ATGCATCCCA CGTTAAGCTG GCGCCACCTC

          51                                                        100
q1env     CCGACTCGGG GTGGAGAGCC GAAAAGAcTG AGAATCCCCT TAAGCTTCGC
r4env     CCGACTCGGG GTGGAGAGCC GAAAAGACTG AGAATCCCCT TAAGCTTCGC

          101                                                       150
q1env     CTCCATCGCC TGGTTCCTTA CTCTAACAAT AACTCCCCAG GCCAGTAGTA
r4env     CTCCATCGCC TGGTTCCTTA CTCTAACAAT AACTCCCCAG GCCAGTAGTA

          151                                                       200
q1env     AACGCCTTAT AGACAGCTCG AACCCCCATA GACCTTTATC CCTTACCTGG
r4env     AACGCCTTAT AGACAGCTCG AACCCCCATA GACCTTTATC CCTTACCTGG

          201                                                       250
q1env     CTGATTATTG ACCCTGATAC GGGTGTCACT GTAAATAGCA CTCGAGGTGT
r4env     CTGATTATTG ACCCTGATAC GGGTGTCACT GTAAATAGCA CTCGAGGTGT

          251                                                       300
q1env     TGCTCCTAGA GGCACCTGGT GGCCTGAACT GCATTTCTGC CTCCGATTGA
r4env     TGCTCCTAGA GGCACCTGGT GGCCTGAACT GCATTTCTGC CTCCGATTGA

          301                                                       350
q1env     TTAACCCCGC TGTTAAAAGC ACACCTCCCA ACCTAGTCCG TAGTTATGGG
r4env     TTAACCCCGC TGTTAAAAGC ACACCTCCCA ACCTAGTCCG TAGTTATGGG

          351                                                       400
q1env     TTCTATTGCT GCCCAGGCAC AGAGAAAGAG AAATACTGTG GGGGTTCTGG
r4env     TTCTATTGCT GCCCAGGCAC AGAGAAAGAG AAATACTGTG GGGGTTCTGG

          401                                                       450
q1env     GGAATCCTTC TGTAGGAGAT GGAGCTGCGT CACCTCCAAC GATGGAGACT
r4env     GGAATCCTTC TGTAGGAGAT GGAGCTGCGT CACCTCCAAC GATGGAGACT

          451                                                       500
q1env     GGAAATGGCC GATCTCTCTC CAGGACCGGG TAAAATTCTC CTTTGTCAAT
r4env     GGAAATGGCC GATCTCTCTC CAGGACCGGG TAAAATTCTC CTTTGTCAAT

          501                                                       550
q1env     TCCGGCCCGG GCAAGTACAA AATGATGAAA CTATATAAAG ATAAGAGCTG
r4env     TCCGGCCCGG GCAAGTACAA AATGATGAAA CTATATAAAG ATAAGAGCTG

          551                                                       600
q1env     CTCCCCATCA GACTTAGATT ATCTAAAGAT AAGTTTCACT G..AAAGGAA
r4env     CTCCCCATCA GACTTAGATT ATCTAAAGAT AAGTTTCACT GAAAAAGGAA

          601                                                       650
q1env     AACAGGAAAA TATTCAAAAG TGGATAAATG GTATGAGCTG GGGAATAGTT
r4env     AACAGGAAAA TATTCAAAAG TGGATAAATG GTATGAGCTG GGGAATAGTT

          651                                                       700
q1env     TTTTAT.TAT ATGGCGGGGG AGCAGGGTCC ACTTTAACCA TTCGCCTTAG
r4env     TTTTATAAAT ATGGCGGGGG AGCAGGGTCC ACTTTAACCA TTCGCCTTAG

          701                                                       750
q1env     GATAGAGACG GGGACAGAAC CCCCTGTGGC AATGGGACCC GATAAAGTAC
r4env     GATAGAGACG GGGACAGAAC CCCCTGTGGC AATGGGACCC GATAAAGTAC
```

Fig. 5.2A. Comparison of the PoERV *env* sequences. The insertions and substitution are boxed. q1env, sequence obtained by Q-One Biotech; r4env, sequence obtained from Raji 4 cells.

```
        751                                                        800
q1env   TGGCTGAACA GGGGCCCCCG GCCCTGGAGC CACCGCATAA CTTGCCGGTG
r4env   TGGCTGAACA GGGGCCCCCG GCCCTGGAGC CACCGCATAA CTTGCCGGTG

        801                                                        850
q1env   CCCCAATTAA CCTCGCTGCG GCCTGACATA ACACAGCCGC CTAGCAACAG
r4env   CCCCAATTAA CCTCGCTGCG GCCTGACATA ACACAGCCGC CTAGCAACAG

        851                                                        900
q1env   TACCACTGGA TTGATTCCTA CCAACACGCC TAGAAACTCC CCAGGTGTTC
r4env   TACCACTGGA TTGATTCCTA CCAACACGCC TAGAAACTCC CCAGGTGTTC

        901                                                        950
q1env   CTGTTAAGAC AGGACAGAGA CTCTTCAGTC TCATCCAGGG AGCTTTCCAA
r4env   CTGTTAAGAC AGGACAGAGA CTCTTCAGTC TCATCCAGGG AGCTTTCCAA

        951                                                       1000
q1env   GCCATCAACT CCACCGACCC TGATGCCACT TCTTCTTGTT GGCTTTGTCT
r4env   GCCATCAACT CCACCGACCC TGATGCCACT TCTTCTTGTT GGCTTTGTCT

        1001                                                      1050
q1env   ATCCTCAGGG CCTCCTTATT ATGAGGGGAT GGCTAAAGAA AGAAAATTCA
r4env   ATCCTCAGGG CCTCCTTATT ATGAGGGGAT GGCTAAAGAA AGAAAATTCA

        1051                                                      1100
q1env   ATGTGACCAA AGAGCATAGA AATCAATGTA CATGGGGGTC CCGAAATAAG
r4env   ATGTGACCAA AGAGCATAGA AATCAATGTA CATGGGGGTC CCGAAATAAG

        1101                                                      1150
q1env   CTTACCCTCA CTGAAGTTTC CGGGAAGGGG ACATGCATAG GAAAAGCTCC
r4env   CTTACCCTCA CTGAAGTTTC CGGGAAGGGG ACATGCATAG GAAAAGCTCC

        1151                                                      1200
q1env   CCCATCCCAC CAACACCTTT GCTATAGTAC TGTGGTTTAT GAGCAGGCCT
r4env   CCCATCCCAC CAACACCTTT GCTATAGTAC TGTGGTTTAT GAGCAGGCCT

        1201                                                      1250
q1env   CAGAAAATCA GTATTTAGTA CCTGGTTATA ACAGGTGGTG GGCATGCAAT
r4env   CAGAAAATCA GTATTTAGTA CCTGGTTATA ACAGGTGGTG GGCATGCAAT

        1251                                                      1300
q1env   ACTGGGTTAA CCCCCTGTGT TTCCACCTCA GTCTTCAACC AATCCAAAGA
r4env   ACTGGGTTAA CCCCCTGTGT TTCCACCTCA GTCTTCAACC AATCCAAAGA

        1301                                                      1350
q1env   TTTCTGTGTC ATGGTCCAAA TCGTCCCCCG AGTGTACTAC CATCCTGAGG
r4env   TTTCTGTGTC ATGGTCCAAA TCGTCCCCCG AGTGTACTAC CATCCTGAGG

        1351                                                      1400
q1env   AAGTGGTCCT TGATGAATAT GACTATCGGT ATAACCGACC AAAAAGAGAA
r4env   AAGTGGTCCT TGATGAATAT GACTATCGGT ATAACCGACC AAAAAGAGAA

        1401                                                      1450
q1env   CCCGTATCCC TTACCCTAGC TGTAATGCTC GGATTAGGGA CGGCCGTTGG
r4env   CCCGTATCCC TTACCCTAGC TGTAATGCTC GGATTAGGGA CGGCCGTTGG

        1451                                                      1500
q1env   CGTAGGAACA GGGACAGCTG CCCTGATCAC AGGACCACAG CAGCTAGAGA
r4env   CGTAGGAACA GGGACAGCTG CCCTGATCAC AGGACCACAG CAGCTAGAGA
```

Fig. 5.2A. Cont.

```
        1501                                                    1550
q1env   AAGGACTTGG TGAGCTACAT GCGGCCATGA CAGAAGATCT CCGAGCCTTA
r4env   AAGGACTTGG TGAGCTACAT GCGGCCATGA CAGAAGATCT CCGAGCCTTA

        1551                                                    1600
q1env   AAGGAGTCTG TTAGCAACCT AGAAGAGTCC CTGACTTCTT TGTCTGAAGT
r4env   AAGGAGTCTG TTAGCAACCT AGAAGAGTCC CTGACTTCTT TGTCTGAAGT

        1601                                                    1650
q1env   GGTTCTACAG AACCGGAGGG GATTAGATCT GCTGTTTCTA AGAGAAGGTG
r4env   GGTTCTACAG AACCGGAGGG GATTAGATCT GCTGTTTCTA AGAGAAGGTG

        1651                                                    1700
q1env   GGTTATGTGC AGCCTTAAAA GAAGAATGTT GCTTCTATGT AGATCACTCA
r4env   GGTTATGTGC AGCCTTAAAA GAAGAATGTT GCTTCTATGT AGATCACTCA

        1701                                                    1750
q1env   GGAGCCATCA GAGACTCCAT GAACAAGCTT AGAAAAAAGT TAGAGAGGCG
r4env   GGAGCCATCA GAGACTCCAT GAACAAGCTT AGAAAAAAGT TAGAGAGGCG

        1751                                                    1800
q1env   TCGAAGGGAA AGAGAGGCTG ACCAGGGGTG GTTTGAAGGA TGGTTCAACA
r4env   TCGAAGGGAA AGAGAGGCTG ACCAGGGGTG GTTTGAAGGA TGGTTCAACA

        1801                                                    1850
q1env   GGTCTCCTTG GATGACCACC CTGCTTTCTG CTCTGACGGG GCCCCTAGTA
r4env   GGTCTCCTTG GATGACCACC CTGCTTTCTG CTCTGACGGG GCCCCTAGTA

        1851                                                    1900
q1env   GTCCTGCTCC TGTTACTTAC AGTTGGGCCT TGCTTAATTA ATAGGTTTGT
r4env   GTCCTGCTCC TGTTACTTAC AGTTGGGCCT TGCTTAATTA ATAGGTTTGT

        1901                                                    1950
q1env   TGCCTTTGTT AGAGAACGAG TGAGTGCAGT CCAGATCATG GTACTTAGGC
r4env   TGCCTTTGTT AGAGAACGAG TGAGTGCAGT CCAGATCATG GTACTTAGGC

        1951                                                    2000
q1env   AACAGTACCA AGGCCTTCTG AGCCAAGGAG AAACTGACCT CTAGCCTTCC
r4env   AACAGTACCA AGGCCTTCTG AGCCAAGGAG AAACTGACCT CTAGGCGGCC

        2001
q1env   CA
r4env   GC
```

Fig. 5.2A. Cont.

```
          1                                                  SP|SU  50
q1env   MHPTLSRRHL PTRGGEPKRL RIPLSFASIA WFLTLTITPQ ASSKRLIDSS
r4env   MHPTLSWRHL PTRGGEPKRL RIPLSFASIA WFLTLTITPQ ASSKRLIDSS

         51                                                   100
q1env   NPHRPLSLTW LIIDPDTGVT VNSTRGVAPR GTWWPELHFC LRLINPAVKS
r4env   NPHRPLSLTW LIIDPDTGVT VNSTRGVAPR GTWWPELHFC LRLINPAVKS

         101                                                  150
q1env   TPPNLVRSYG FYCCPGTEKE KYCGGSGESF CRRWSCVTSN DGDWKWPISL
r4env   TPPNLVRSYG FYCCPGTEKE KYCGGSGESF CRRWSCVTSN DGDWKWPISL

         151                                                  200
q1env   QDRVKFSFVN SGPGKYKMMK LYKDKSCSPS DLDYLKISFT ERKTGKYSKV
r4env   QDRVKFSFVN SGPGKYKMMK LYKDKSCSPS DLDYLKISFT EK..GKQENI

         201                                                  250
q1env   DKW...YELG NSFLLYGGGA GSTLTIRLRI ETGTEPPVAM GPDKVLAEQG
r4env   QKWINGMSWG IVFYKYGGGA GSTLTIRLRI ETGTEPPVAM GPDKVLAEQG

         251                                                  300
q1env   PPALEPPHNL PVPQLTSLRP DITQPPSNST TGLIPTNTPR NSPGVPVKTG
r4env   PPALEPPHNL PVPQLTSLRP DITQPPSNST TGLIPTNTPR NSPGVPVKTG

         301                                                  350
q1env   QRLFSLIQGA FQAINSTDPD ATSSCWLCLS SGPPYYEGMA KERKFNVTKE
r4env   QRLFSLIQGA FQAINSTDPD ATSSCWLCLS SGPPYYEGMA KERKFNVTKE

         351                                                  400
q1env   HRNQCTWGSR NKLTLTEVSG KGTCIGKAPP SHQHLCYSTV VYEQASENQY
r4env   HRNQCTWGSR NKLTLTEVSG KGTCIGKAPP SHQHLCYSTV VYEQASENQY

         401                                                  450
q1env   LVPGYNRWWA CNTGLTPCVS TSVFNQSKDF CVMVQIVPRV YYHPEEVVLD
r4env   LVPGYNRWWA CNTGLTPCVS TSVFNQSKDF CVMVQIVPRV YYHPEEVVLD

         451          | TM                                    500
q1env   EYDYRYNRPK REPVSLTLAV MLGLGTAVGV GTGTAALITG PQQLEKGLGE
r4env   EYDYRYNRPK REPVSLTLAV MLGLGTAVGV GTGTAALITG PQQLEKGLGE

         501                                                  550
q1env   LHAAMTEDLR ALKESVSNLE ESLTSLSEVV LQNRRGLDLL FLREGGLCAA
r4env   LHAAMTEDLR ALKESVSNLE ESLTSLSEVV LQNRRGLDLL FLREGGLCAA

         551                                                  600
q1env   LKEECCFYVD HSGAIRDSMN KLRKKLERRR REREADQGWF EGWFNRSPWM
r4env   LKEECCFYVD HSGAIRDSMN KLRKKLERRR REREADQGWF EGWFNRSPWM

         601                                                  650
q1env   TTLLSALTGP LVVLLLLLTV GPCLINRFVA FVRERVSAVQ IMVLRQQYQG
r4env   TTLLSALTGP LVVLLLLLTV GPCLINRFVA FVRERVSAVQ IMVLRQQYQG

         651
q1env   LLSQGETDL
r4env   LLSQGETDL
```

Fig. 5.2B. Comparison of the predicted PoERV Env amino acid sequences. The variable region is boxed. SP, signal peptide; SU, surface glycoprotein; TM, transmembrane protein; q1env, sequence obtained by Q-One Biotech; r4env, sequence obtained from Raji 4 cells.

5.3.3 Features of the Env sequence

Alignment of the amino acid sequence with those of several other mammalian C-type retroviruses reveals several conserved domains and motifs. These are illustrated in Fig. 5.3. As would be predicted, the major sequence differences are seen in the regions corresponding to VRA, VRB and the variable PRR. Like the Q-One Biotech sequence, greatest homology is seen with that of GALV. The PoERV Env protein contains five potential N-glycosylation sites within the SU. Two of these are conserved among the other mammalian C-type proteins, and this includes the site reported to be necessary for Friend-MLV growth in cell culture (Kayman *et al.*, 1991). Six of the eight cysteine residues normally conserved within the amino terminal of SU are present. The cysteine residues throughout the rest of the Env protein are conserved. The VRA is relatively short, containing only 28 amino acids. The first cysteine residue usually present in this region is missing, although an extra residue is found adjacent to the cysteine towards the carboxy-terminal end of VRA. The cysteine residue normally found just after VRB is also missing.

The TM proteins of the mammalian C-type retroviruses are highly homologous. This appears to hold true for the PoERV TM. The putative fusion peptide and leucine zipper motifs, as well as the conserved CX_6CC sequence and basolateral targeting signal, all appear to be present. The greatest variation is seen within the fusion peptide. In this region the PoERV sequence shows closest homology with GALV. Although the majority of features described for TM proteins are conserved, the leucine and glutamine residues in the membrane proximal region of the cytoplasmic tail are not present. Residues in this region have been reported to be important for fusion and incorporation (section 5.1.3).

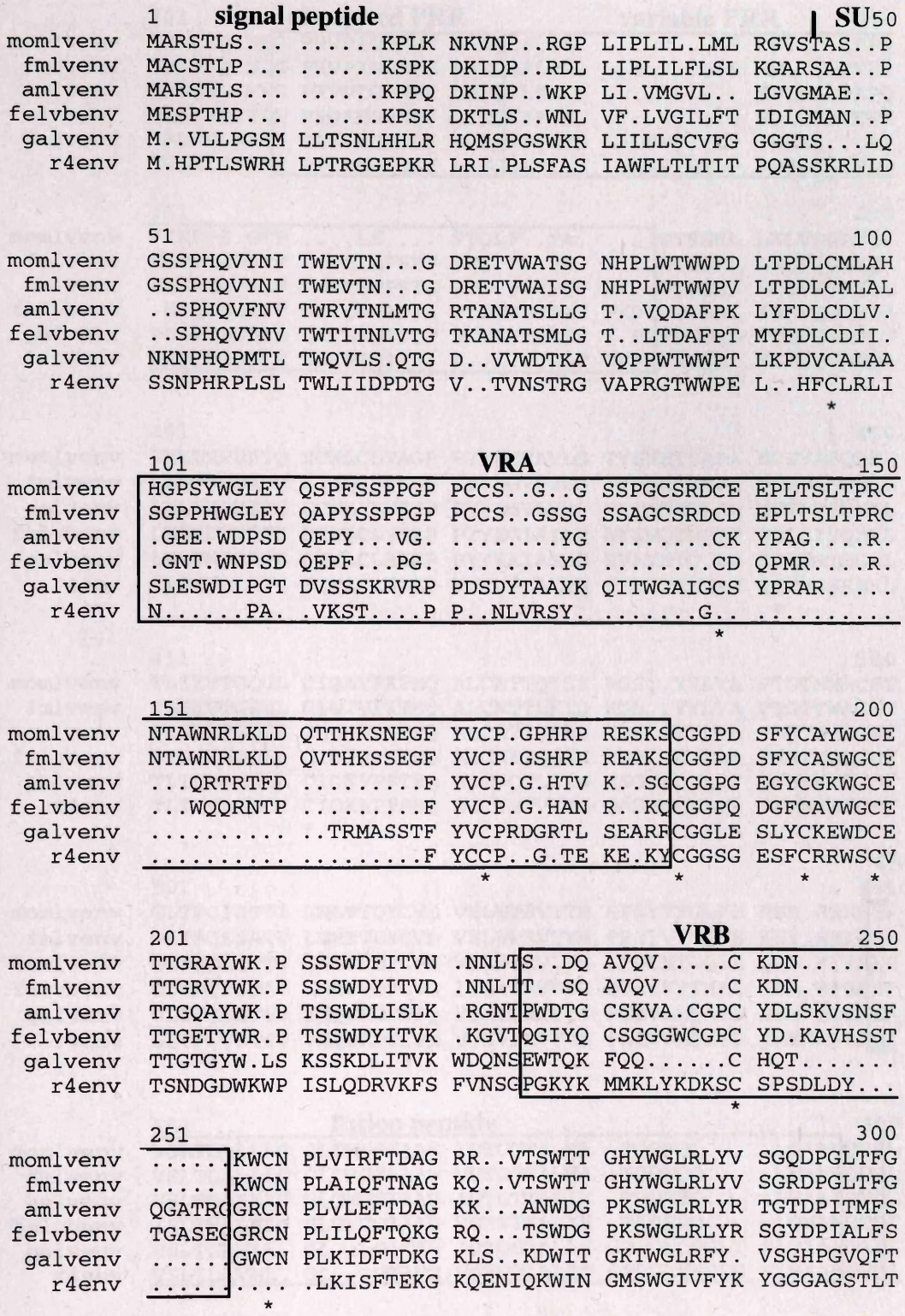


Fig. 5.3. Alignment of mammalian C-type Env peptides. Structural features predicted for ecotropic MLV are boxed. The conserved N-linked glycosylation sites in SU and the basolateral targetting signal are in bold type. SU, surface glycoprotein; TM transmembrane protein; * conserved cysteine residue; momlvenv, Moloney MLV ENV (Shinnick *et al.*, 1981; Van Beveren *et al.*, 1981); fmlvenv, Friend MLV Env (Koch *et al.*, 1983); amlvenv, 4070A amphotropic MLV Env (Ott *et al.*, 1990); felvbenv, Gardner-Arnstein FeLV-B Env (Wunsch *et al.*, 1983); galvenv, GALV SEATO Env (Ting *et al.*, 1998); r4env, sequence obtained from Raji 4 cells.

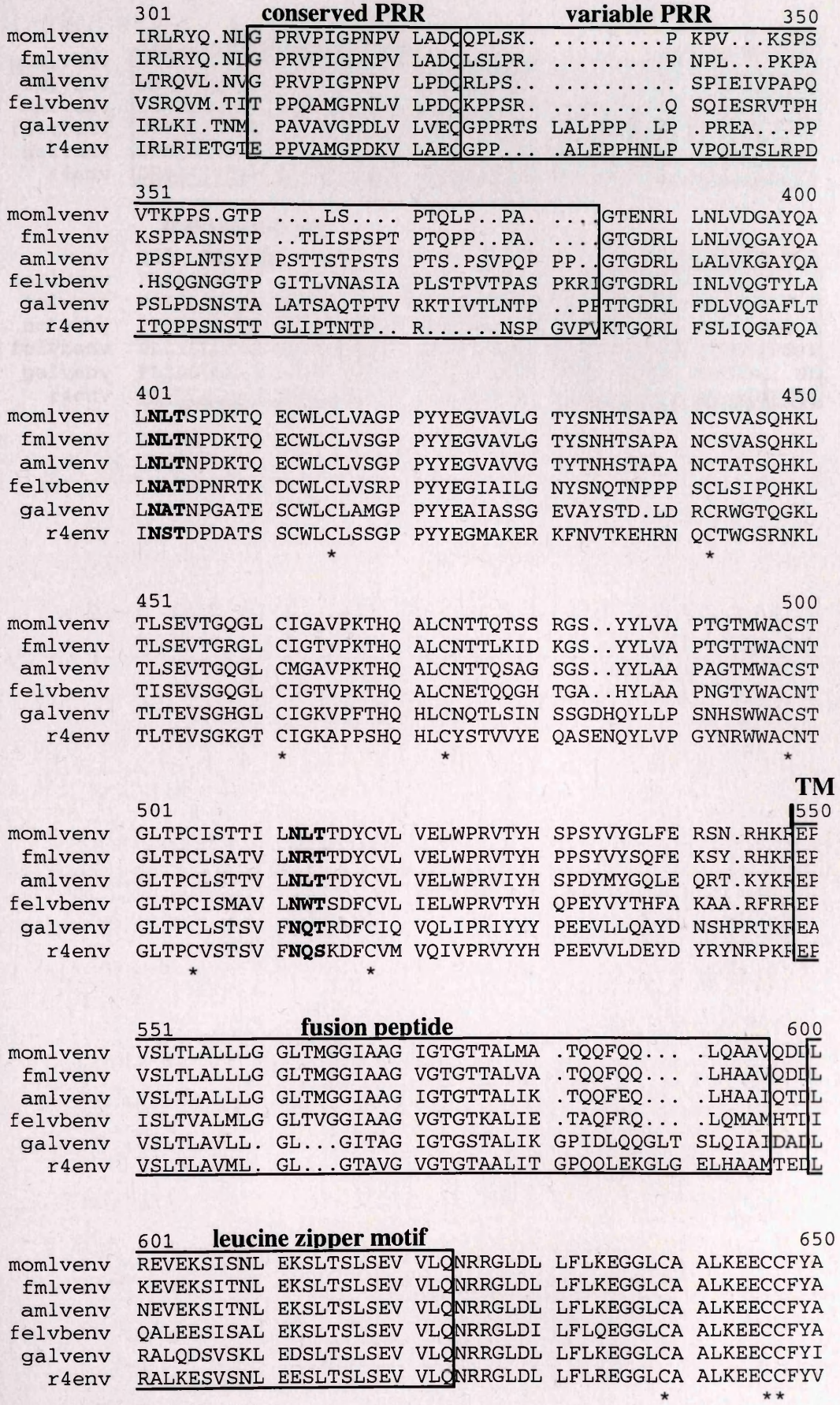

conserved PRR
variable PRR
fusion peptide
leucine zipper motif
TM
momlvenv
fmlvenv
amlvenv
felvbenv
galvenv
r4env

Fig. 5.3. Cont.

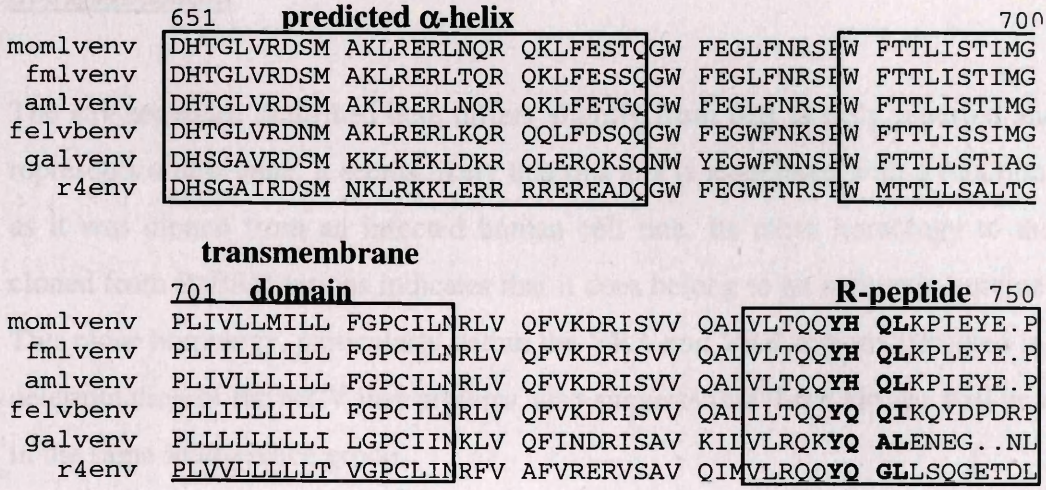

651
predicted α-helix
700
momlvenv DHTGLVRDSM AKLRERLNQR QKLFESTQGW FEGLFNRSPW FTTLISTIMG
fmlvenv DHTGLVRDSM AKLRERLTQR QKLFESSQGW FEGLFNRSPW FTTLISTIMG
amlvenv DHTGLVRDSM AKLRERLNQR QKLFETGQGW FEGLFNRSPW FTTLISTIMG
felvbenv DHTGLVRDNM AKLRERLKQR QQLFDSQQGW FEGWFNKSPW FTTLISSIMG
galvenv DHSGAVRDSM KKLKEKLDKR QLERQKSQNW YEGWFNNSPW FTTLLSTIAG
r4env DHSGAIRDSM NKLRKKLERR RREREADQGW FEGWFNRSPW MTTLLSALTG
transmembrane
701
domain
R-peptide
750
momlvenv PLIVLLMILL FGPCILNRLV QFVKDRISVV QAIVLTQQYH QLKPIEYE.P
fmlvenv PLIILLLILL FGPCILNRLV QFVKDRISVV QAIVLTQQYH QLKPLEYE.P
amlvenv PLIVLLLILL FGPCILNRLV QFVKDRISVV QAIVLTQQYH QLKPIEYE.P
felvbenv PLLILLLILL FGPCILNRLV QFVKDRISVV QAIILTQQYQ QIKQYDPDRP
galvenv PLLLLLLLLI LGPCIINKLV QFINDRISAV KIIVLRQKYQ ALENEG..NL
r4env PLVVLLLLLT VGPCLINRFV AFVRERVSAV QIMVLRQQYQ GLLSQGETDL
*

Fig. 5.3. Cont.

5.4 Discussion

The *env* sequence described here differs slightly from that already reported and therefore represents a new gene. It seems likely that this *env* is associated with a functional provirus as it was cloned from an infected human cell line. Its close homology to the *env* gene cloned from PoERV virions indicates that it does belong to an authentic porcine retrovirus. This close homology, particularly within the VRA and VRB regions involved in host range determination of the MLV Env proteins, also suggests that these viruses will be found to be in the same interference group.

Several similarities, as well as some differences, exist between this Env protein and the other mammalian C-type sequences shown in Fig. 5.3. The first of the conserved N-linked glycosylation sites is probably also important for growth of this virus in cell culture. As discussed in section 5.1.2, this site is thought to affect processing of the precursor Env protein (Kayman *et al.*, 1991). The conservation of the majority of the cysteine residues within the SU suggests that this protein has a similar structure to the others. However, the few changes indicate that there are some differences. The positions of the cysteine residues within the region corresponding to VRA seem to imply that the interloop domain is absent. The loss of the residue following the VRB region prevents the formation of a disulphide-linked loop in this area. These structural differences may indicate that this Env recognises a receptor that is unrelated to the transporter molecules recognised by other mammalian C-type Env proteins.

As expected, the TM protein shows a high degree of homology to the other proteins compared here. This strongly indicates that these proteins function in a similar manner. The presence of a basolateral targeting signal suggests that this virus may also use this pathway for virion release and cell to cell spread.

It has been reported that certain deletions and substitutions of amino acids in the membrane proximal region of the cytoplasmic tail of MoMLV TM reduce fusion activity and incorporation of the protein (Januszeski *et al.*, 1997). These residues are highly conserved among the C-type viruses. It is therefore perhaps surprising that the Leu and Gln residues are not conserved in the PoERV sequence. The Leu is believed by Januszeski *et al.* to be

particularly important in the fusion process. The same study also found that deletion of Leu and the adjacent Val, or substitution of Gln with Ala and Phe with Met reduced incorporation into virions. While the substitution of Leu with Phe in the PoERV TM is a conservative change, that of Gln with Ala is not. However, Januszetski *et al.* suggested that the fusion process requires hydrophobic residues within this region. These substitutions result in an overall increase in hydrophobicity. The mutations that affect incorporation of the MoMLV protein may be related to structural features that differ in the PoERV protein. Although most of the changes are conservative, the sequence of the cytoplasmic tail of the PoERV TM shares only approximately 45% homology with those of the other proteins.

The cloning and sequencing of this *env* gene confirms that more than one transcriptionally active strain of PoERV exists. Although this *env* is highly homologous to the gene sequenced previously, it is possible that the two proteins recognise different receptors. This can only be tested through infectivity assays. It is now possible to use this *env* clone in pseudotyping experiments utilising the HIT system (Soneoka *et al.* 1995). Vectors bearing this Env protein will be expected to exhibit the host range of the PoERV from which it is derived.

CHAPTER 6

Chapter 6 Determination of the host range of PoERV

6.1 Introduction

6.1.1 PoERV subgroups

Several closely related ERVs are now known to be present in the porcine genome (Akiyoshi *et al.*, 1998; Le Tissier *et al.*, 1997; Q-One Biotech Ltd. and Imutran Ltd., 1997). The complete nucleotide sequence of two of these viruses are now available: the PoERV-B1 sequence obtained from virions released from PK15 cells (Q-One Biotech Ltd. and Imutran Ltd., 1997), and the PERV-MSL sequence obtained from a proviral clone derived from miniature swine lymphocytes (Akiyoshi *et al.*, 1998). These viruses share 90.1% homology within the *gag* gene, and 96.8% within the *pol* gene (Onions *et al.*, 1998). The major sequence differences are found within the SU protein, suggesting that they recognise different receptors. As discussed in section 1.2.1, closely related retroviruses are placed into subgroups on the basis of host range, neutralisation and receptor interference analyses (Weiss, 1993). Although these properties have not been fully defined for the PoERVs, these viruses have been tentatively placed into three subgroups, PoERV-A, -B and -C (Onions *et al.*, 1998). The PoERV-A subgroup Env sequence was first described by Le Tissier *et al.* (1997). A second, related sequence, designated PoERV-A1 has also been placed into this subgroup (G.Langford, pers. com). The PoERV-B subgroup contains both PoERV-B1 and the *env* sequence obtained from the infected Raji 4 cells, described in the previous chapter. This sequence was also obtained by Le Tissier *et al.* Both PoERV-A and PoERV-B specific sequences have been detected in PK15 cells and PoERV-infected human 293 cells (Le Tissier *et al.*, 1997). Subgroup C contains PERV-MSL, the Tsukuba-1 virus isolated from a porcine lymphoma cell line (Suzuka *et al.*, 1986; The General Hospital Corporation, 1997), and also a sequence obtained recently, designated PoERV-C1 (P.Banerjee, pers. com.). Hybridisation analyses have revealed that normal pig tissues carry between 10 and 23 copies of PoERV-A sequences and between 7 and 12 of PoERV-B (Le Tissier *et al.*, 1997). PoERV-C is present in 8 to 15 copies in the tissues of miniature swine (Akiyoshi *et al.*, 1998). This subgroup does not appear to be present in all pigs, and has not

been detected in ST or PK15 cells (D.Hart, pers. com.). It is not yet known how many of these proviral sequences encode functional virus.

Laboratory mice are the only other animals in which groups of closely related infectious ERVs with different *env* genes have been described (Boeke and Stoye, 1997). The genomes of all inbred strains of mice contain several xenotropic and polytropic endogenous MLV sequences (Hoggan *et al.*, 1983; O'Neill *et al.*, 1986), and some also contain ecotropic sequences (Jenkins *et al.*, 1982). Kozak and O'Neill (1987) have found that the different endogenous *env* sequences are restricted to specific taxonomic groups among wild mice. This indicates that they were acquired independently. The presence of the different ERV sequences within the laboratory strains is likely to be the result of interbreeding between mice of different geographical and taxonomic origins. Ecotropic sequences have only been found in a small number of wild and laboratory mouse strains (Jenkins *et al.*, 1982; Lafon *et al.*, 1983), suggesting that these were acquired more recently. It seems possible that the different subgroups of PoERV were also acquired independently. Although the majority of domestic pigs are believed to have been derived from Asian species, European wild boars are also likely to have had an influence (Epstein and Bichard, 1984). Later interbreeding between pig breeds would have allowed the different PoERVs to become spread throughout the population. PoERV-C sequences were first isolated from miniature pig cells and do not appear to be present in all pigs (Akiyoshi *et al.*, 1998; Wilson *et al.*, 1998; D.Hart, pers. com.). This suggests that, like the endogenous E-MLV, this virus was acquired more recently than the other subgroups. It is also feasible that the differences in the PoERV *env* genes arose from recombination with other, as yet unidentified, ERV sequences in the porcine genome.

Multiple copies of ERV proviruses are believed to have resulted from additional germline infections following replication of the original provirus (King *et al.*, 1987; Stoye and Coffin, 1987). The heterogeneity observed in PoERV proviral copy numbers (Le Tissier *et al.*, 1997; Akiyoshi *et al.*, 1998; D.Hart, pers. com.) is likely to reflect the lack of significant inbreeding (Epstein and Bichard, 1984). It is possible that proviral acquisition is still an ongoing process. Studies of inbred strains of laboratory mice have revealed a continual movement of endogenous E-MLV sequences between generations (Boeke and Stoye, 1997). Novel insertions have been found in the progeny of viraemic females

(Jenkins and Copeland, 1985; Rowe and Kozak, 1980), probably as a consequence of oocyte infection (Lock *et al.*, 1988). This process may also occur with PoERV since these viruses are capable of infecting at least some porcine cells (Lieber *et al.*, 1975a; Patience *et al.*, 1997a; Wilson *et al.*, 1998), and some pigs may have a low level viraemia – RT activity has been detected in serum although no transfer of infectivity to cultured cells has yet been observed (Onions *et al.*, 1999).

6.1.2 PoERV Env proteins

The relationship between the PoERV Env proteins is depicted in the phylogenetic tree in Fig 6.1 and a comparison of the four sequences is shown in Fig. 6.2. The four TM proteins are highly homologous, although PoERV-C has a divergent R-peptide sequence. PoERV-B and B1 Env proteins differ only in the region between residues 192 and 215. The PoERV-A and -C proteins show a greater degree of homology with PoERV-B within this region. However, PoERV-A and -C show differences within the areas corresponding to VRA, VRB and the PRR. As described in section 5.1.2, VRA and VRB are believed to be involved in the receptor binding of Env proteins. The most significant areas of divergence among the PoERV proteins are within VRB and the variable PRR. It is possible that these differences reflect variation in receptor usage, and therefore host range. Five potential N-linked glycosylation sites are conserved between all the SU proteins, while PoERV-A and -C possess another three. Sixteen conserved cysteine residues are present in all four proteins. These include 6 of the 8 residues normally present in the amino-terminal region of mammalian C-type SU proteins, and the highly conserved CX_6CC motif of the TM protein (section 5.1). PoERV-A and -C possess only one of the two cysteine residues normally conserved within VRA. This indicates further differences in the structure of the receptor binding region.

6.1.3 Analysis of host range and receptor usage of retroviruses

The predicted differences in receptor usage may prove very important when assessing the risk the PoERV subgroups pose through xenotransplantation. This is because the receptor used by a virus not only determines its host range but also has a large influence on the

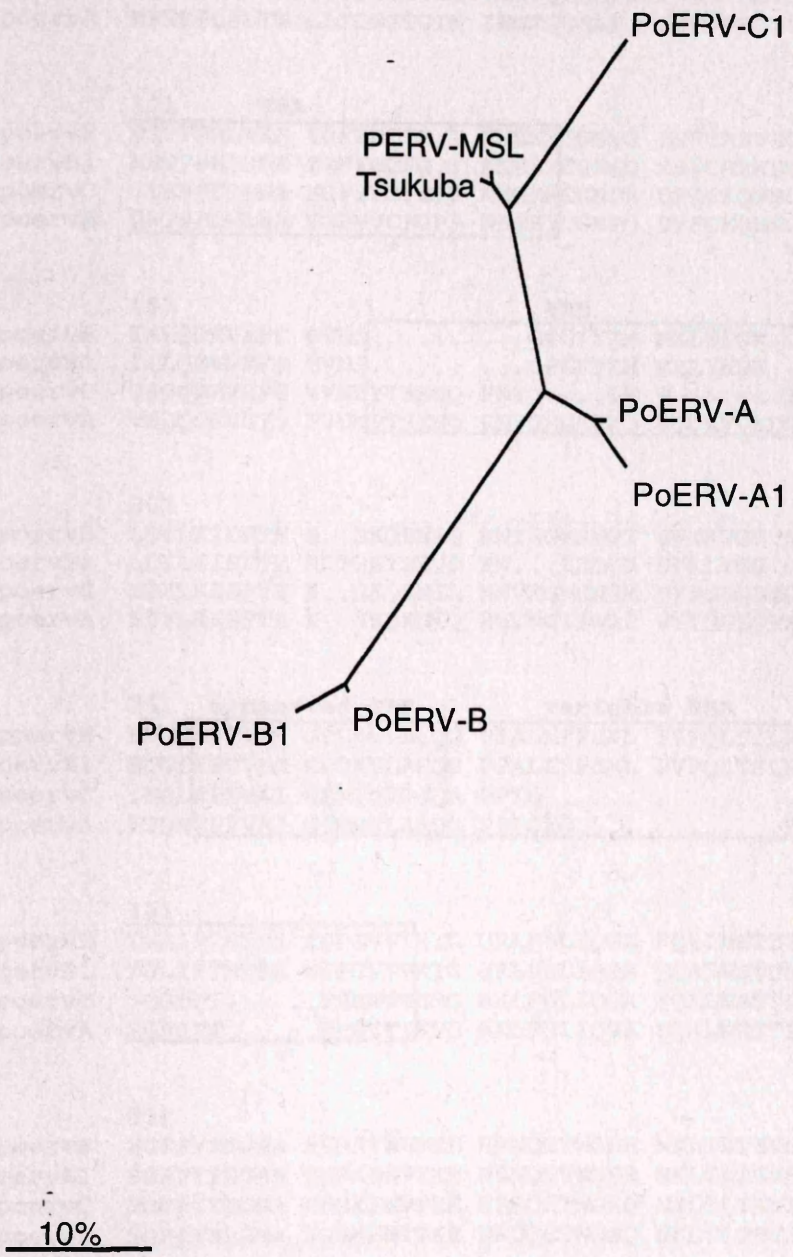


Fig. 6.1. Unrooted phylogenetic tree of the PoERV retroviruses based on alignment of SU glycoprotein amino acid sequences. The PHYLIP package was used for protein distance calculation and tree construction. The solid line below the diagram represents 10% divergence. Constructed by Douglas Hart (Veterinary Pathology, University of Glasgow).

```
          1 signal peptide  |SU                                          50
 poervB   MHPTLSWRHL PTRGGEPKRL RIPLSFASIA WFLTLTITPQ ASSKRLIDSS
poervB1   MHPTLSRRHL PTRGGEPKRL RIPLSFASIA WFLTLTITPQ ASSKRLIDSS
 poervC   MHPTLNRRHL PIRGGKPKRL KIPLSFASIA WFLTLSITSQ TNGMRIGDSL
 poervA   MHPTLSRRHL PIRGGKPKRL KIPLSFASIA WFLTLSITPQ VNGKRLVDSP

          51                                                      100
 poervB   NPHRPLSLTW LIIDPDTGVT VNSTRGVAPR GTWWPELHFC LRLINPAV..
poervB1   NPHRPLSLTW LIIDPDTGVT VNSTRGVAPR GTWWPELHFC LRLINPAV..
 poervC   NSHKPLSLTW LITDSGTGIN INNTQGEAPL GTWWPDLYVC LRSVIPSL..
 poervA   NSHKPLSLTW LLTDSGTGIN INSTQGEAPL GTWWPELYVC LRSVIPGLND
                                                        *

          101      VRA                                            150
 poervB   KSTPPNLVRS YGFYCCPG.T EKEKYCGGSG ESFCRRWSCV TSNDGDWKWP
poervB1   KSTPPNLVRS YGFYCCPG.T EKEKYCGGSG ESFCRRWSCV TSNDGDWKWP
 poervC   .TSPPDILHA HGFYVCPGPP NNGKHCGNPR DFFCKQWNCV TSNDGYWKWP
 poervA   QATPPDVLRA YGFYVCPGPP NNEEYCGNPQ DFFCKQWSCI TSNDGNWKWP
                          *          *      *     *

          151                          VRB                        200
 poervB   ISLQDRVKFS FVNS..... ...GPGKYKM MKLYKDK... ....SCSPSD
poervB1   ISLQDRVKFS FVNS..... ...GPGKYKM MKLYKDK... ....SCSPSD
 poervC   TSQQDRVSFS YVNTYTSSGQ FNY.....LT W.......IR TGSPKCSPSD
 poervA   VSQQDRVSYS FVNNPTSYNQ FNYGHGRWKD WQQRVQKDVR NKQISCHSLD
                                                          *

          201                                                     250
 poervB   LDYLKISFTE K..GKQENIQ KWINGMSWGI VFYKYGG.GA GSTLTIRLRI
poervB1   LDYLKISFTE RKTGKYSKVD KW...YELGN SFLLYGG.GA GSTLTIRLRI
 poervC   LDYLKISFTE K..GKQENIL KWVNGMSWGM VYYGGSGKQP GSILTIRLKI
 poervA   LDYLKISFTE K..GKQENIQ KWVNGISWGI VYYGGSGRKK GSVLTIRLRI

          251  conserved PRR          variable PRR                300
 poervB   ETGTEPPVAM GPDKVLAEQG PPALEPPHNL PVPQLTSLRP DITQPPSNST
poervB1   ETGTEPPVAM GPDKVLAEQG PPALEPPHNL PVPQLTSLRP DITQPPSNST
 poervC   .NQLEPPMAI GPNTVLTGQR PPTQ...... .......... GPGPSS.NIT
 poervA   ETQMEPPVAI GPNKGLAEQG PPIQEQ.... ........RP SPNPSDYNTT

          301                                                     350
 poervB   TGLIPTNTPR NSPGVPVKTG QRLFSLIQGA FQAINSTDPD ATSSCWLCLS
poervB1   TGLIPTNTPR NSPGVPVKTG QRLFSLIQGA FQAINSTDPD ATSSCWLCLS
 poervC   SGSDPT.... .ESNSTIKMG AKLFSLIQGA FQALNSTTPE ATSSCWLCLA
 poervA   SGSVPT.... .EPNITIKTG AKLFSLIQGA FQALNSTTPE ATSSCWLCLA
                                                      *  *

          351                                                     400
 poervB   SGPPYYEGMA KERKFNVTKE HRNQCTWGSR NKLTLTEVSG KGTCIGKAPP
poervB1   SGPPYYEGMA KERKFNVTKE HRNQCTWGSR NKLTLTEVSG KGTCIGKAPP
 poervC   SGPPYYEGMA RRGKFNVTKE HRDQCTWGSQ NKLTLTEVSG KGTCIGKVPP
 poervA   SGPPYYEGMA RGGKFNVTKE HRDQCTWGSQ NKLTLTEVSG KGTCIGMVPP
                                                        *
```

Fig. 6.2. Alignment of the PoERV Env peptides so far described. The features indicated are described in section 5.1. Predicted structural features are boxed. The conserved N-linked glycosylation signals and the basolateral targeting signal are in bold type. SU, surface glycoprotein; TM, transmembrane protein; *, conserved cysteine residue; poervB, PoERV-B Env (Chapter 5, Le Tissier *et al.*, 1997); poervB1, PoERV-B1 Env (Q-One Biotech Ltd, Immutran Ltd, 1997); poervC, PoERV-C Env (Akiyoshi *et al.*, 1998); poervA, PoERV-A Env (Le Tissier *et al.*, 1997).

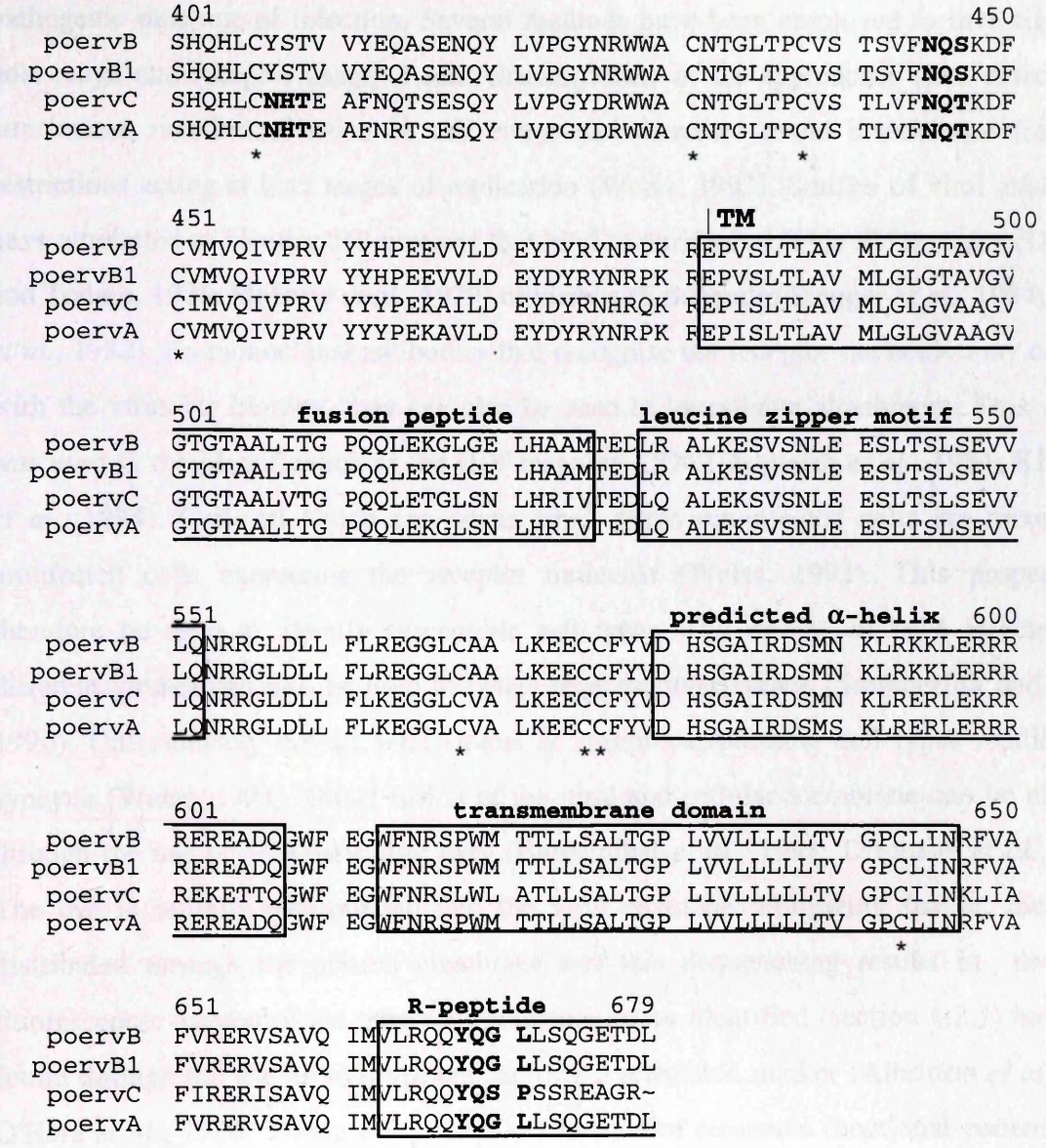

401 450
poervB SHQHLCYSTV VYEQASENQY LVPGYNRWWA CNTGLTPCVS TSVFNQSKDF
poervB1 SHQHLCYSTV VYEQASENQY LVPGYNRWWA CNTGLTPCVS TSVFNQSKDF
poervC SHQHLCNHTE AFNQTSESQY LVPGYDRWWA CNTGLTPCVS TLVFNQTKDF
poervA SHQHLCNHTE AFNRTSESQY LVPGYDRWWA CNTGLTPCVS TLVFNQTKDF
451 TM 500
poervB CVMVQIVPRV YYHPEEVVLD EYDYRYNRPK REPVSLTLAV MLGLGTAVGV
poervB1 CVMVQIVPRV YYHPEEVVLD EYDYRYNRPK REPVSLTLAV MLGLGTAVGV
poervC CIMVQIVPRV YYYPEKAILD EYDYRNHRQK REPISLTLAV MLGLGVAAGV
poervA CVMVQIVPRV YYYPEKAVLD EYDYRYNRPK REPISLTLAV MLGLGVAAGV
501 fusion peptide leucine zipper motif 550
poervB GTGTAALITG PQQLEKGLGE LHAAMTEDLR ALKESVSNLE ESLTSLSEVV
poervB1 GTGTAALITG PQQLEKGLGE LHAAMTEDLR ALKESVSNLE ESLTSLSEVV
poervC GTGTAALVTG PQQLETGLSN LHRIVTEDLQ ALEKSVSNLE ESLTSLSEVV
poervA GTGTAALITG PQQLEKGLSN LHRIVTEDLQ ALEKSVSNLE ESLTSLSEVV
551 predicted α-helix 600
poervB LQNRRGLDLL FLREGGLCAA LKEECCFYVD HSGAIRDSMN KLRKKLERRR
poervB1 LQNRRGLDLL FLREGGLCAA LKEECCFYVD HSGAIRDSMN KLRKKLERRR
poervC LQNRRGLDLL FLKEGGLCVA LKEECCFYVD HSGAIRDSMN KLRERLEKRR
poervA LQNRRGLDLL FLKEGGLCVA LKEECCFYVD HSGAIRDSMS KLRERLERRR
601 transmembrane domain 650
poervB REREADQGWF EGWFNRSPWM TTLLSALTGP LVVLLLLLTV GPCLINRFVA
poervB1 REREADQGWF EGWFNRSPWM TTLLSALTGP LVVLLLLLTV GPCLINRFVA
poervC REKETTQGWF EGWFNRSLWL ATLLSALTGP LIVLLLLLTV GPCIINKLIA
poervA REREADQGWF EGWFNRSPWM TTLLSALTGP LVVLLLLLTV GPCLINRFVA
651 R-peptide 679
poervB FVRERVSAVQ IMVLRQQYQG LLSQGETDL
poervB1 FVRERVSAVQ IMVLRQQYQG LLSQGETDL
poervC FIRERISAVQ IMVLRQQYQS PSSREAGR~
poervA FVRERVSAVQ IMVLRQQYQG LLSQGETDL

Fig. 6.2. Cont.

pathogenic outcome of infection. Several methods have been employed to investigate the host range and receptor usage of retroviruses. Many of the approaches used detect virus attachment, membrane fusion or cell entry, and therefore avoid interference from any restrictions acting at later stages of replication (Weiss, 1993). Studies of viral attachment have attempted to identify cell proteins that bind to purified, soluble SU proteins (DeLarco and Todaro, 1976; Moldow *et al.*, 1979) or virions (Krichbaum-Stenger *et al.*, 1984; Notter *et al.*, 1982). As monoclonal antibodies that recognise the receptor molecule may compete with the virus for binding, they can also be used to investigate attachment. This method was used in the identification of the HIV receptor, CD4 (Dalgleish *et al.*, 1984; Klatzman *et al.*, 1984). Cell-cell fusion can occur when retrovirus-infected cells are mixed with uninfected cells expressing the receptor molecule (Weiss, 1993). This property can therefore be used to identify susceptible cell types. The mixing of cells infected with different viruses can also be used to detect receptor interference (Sommerfelt and Weiss, 1990). Unfortunately not all viral strains or receptor-expressing cell types readily form syncytia (Weiss, 1993). Direct fusion of the viral and cellular membrane can be observed through the use of self-quenching dyes (Blumenthal *et al.*, 1991; Dimitrov *et al.*, 1991). The dye is initially incorporated into the viral envelope. Following fusion, the dye is distributed through the plasma membrane and this dequenching results in detectable fluorescence. Several of the retrovirus receptors so far identified (section 1.2.1) have been found through the use of viral vectors carrying a selectable marker (Albritton *et al.*, 1989; O'Hara *et al.*, 1990; Young *et al.*, 1993). This system measures functional penetration of the cell, which can be simply detected by the subsequent expression of the selectable marker gene. An adaptation of this method involves the use of pseudotyped viruses and vectors bearing the Env protein of the virus of interest (Hunter, 1997; Weiss, 1993).

6.1.4 The use of pseudotypes in host range and receptor analysis

Pseudotypes of VSV have been used to study the receptor usage of several viruses, often those for which no *in vitro* assay was available (Boettiger, 1979; Dalgleish *et al.*, 1984; Lagging *et al.*, 1998; Sommerfelt and Weiss, 1990). These pseudotypes bear the envelope protein of the virus of interest, and therefore will only infect cells that express the receptor recognised by this protein. Following cell entry the VSV genome replicates as normal, forming visible plaques. VSV is normally able to replicate in a wider range of cell types

than the virus used to form the pseudotype. Therefore, these pseudotypes can be used to determine whether host range restrictions occur at the level of cell entry, or at a later stage of the infection cycle (Weiss, 1980). For example, the ecotropic MLV virus strains differ in their ability to infect the NIH 3T3 and BALB 3T3 cell lines (Hartley *et al.*, 1970; Pincus *et al.*, 1975). Experiments with VSV pseudotypes gave the first indication that these differences were due to factors acting at a post penetration step (Weiss, 1980).

Pseudotyped retroviral vectors have also been used in studies of host range and receptor usage. Such vectors have been used to investigate the function of several viral glycoproteins, for example those of ASLV-A and Ebola (Rong *et al.*, 1997; Wool-Lewis and Bates, 1998). Their use avoids the need for infectious virus, and the opportunity for the insertion of marker genes into the vector genome allows simple detection of infected target cells. The transient packaging systems described in section 1.4.2 provide a convenient method of producing high titre pseudotyped vectors by replacing the envelope protein construct. Since the use of a pseudotype system requires only the *env* genes to be cloned, it enables rapid analysis of a wide variety of cell lines for their susceptibility to infection. These vectors also permit investigation of the specific host range of each PoERV Env individually. Viral sequences detected in cells infected by wild type virus could have been carried by virions bearing the Env protein of another subgroup, or as genomic pseudotypes since these have been shown to form between mammalian C-type retroviruses (Takeuchi *et al.*, 1992a). Although it is possible that the pseudotyped vectors will be able to infect cells that would not normally support replication of the PoERVs, the ability to enter cells may still be relevant when assessing the risks posed by these viruses.

6.2 Materials and methods

6.2.1 Cell free infection

Fresh DMEM medium was added to 75cm^2 flasks of subconfluent MPK and PK15 cells, and incubated overnight at 37°C. Medium was then removed and filtered through 0.45µm-pore filters. 1ml aliquots were added to 25cm^2 flasks of target cells, seeded the previous day at 5 x 10^5 cells per flask, in the presence of 8µg/ml Polybrene. The cells were incubated at 37°C for 4 hours and 5ml fresh medium was added. The target cells used were AH927, 293 and RK-13. MPK medium was also added to ST and PK15 cells. The target cells were then subcultured, and after the third and fifth passages, cells were removed and genomic DNA was extracted as described previously (section 2.2.2.2). This DNA was then analysed by PCR for PoERV *env* sequences.

Each PCR was performed with 500ng of template DNA, 20pmol of each primer, 2.5mM $MgCl_2$, 0.125mM of each dNTP and 1U *Taq* DNA polymerase. The primers used were designed by Douglas Hart (Veterinary Pathology, University of Glasgow) and were as follows:

PV2USU	5′ - ATA AGA GCT GCT CCC CAT CAG AC - 3′
PV2DSU	5′ - GCT AGG CGG CTG TGT TAT GTC AG - 3′
PV3USU	5′ - CTG ACC TGG ATT AGA ACT GGA AG - 3′
PV3DSU	5′ - TAT GTT AGA GGA TGG TCC TGG TC - 3′
PV4USU	5′ - GGA GAT GGA AAG ATT GGC AAC AG - 3′
PV4DSU	5′ - CAG AGG TTG TAT TGT AAT CAG AG - 3′

PV2USU and PV2DSU are specific for PoERV-B *env*, PV3USU and PV3DSU for PoERV-C *env*, and PV4USU and PV4DSU for PoERV-A *env*. The cycling parameters used with the PV2 and PV3 primers were 5 mins at 94°C; 25 cycles of 30s at 94°C, 30s at 65°C, and 1 min at 72°C; and a final extension step of 7 mins at 72 °C. The parameters used with the PV4 primers were 5 mins at 94°C; 25 cycles of 30s at 94°C, 30s at 60°C, and 1 min at 72°C; and a final extension step of 7 mins at 72°C.

6.2.2 Cloning of the PoERV-A and -MSL *env* genes

The PoERV-A *env* gene was obtained by reverse transcription PCR (RT-PCR) of PK15 mRNA. mRNA was prepared from PK15 cells using the QuickPrep mRNA Purification Kit (Pharmacia Biotech, UK), according to the manufacturer's instructions. Briefly, cells from a confluent $75cm^2$ flask were homogenised in a buffer containing a high concentration of guanidinium thiocyanate, ensuring rapid inactivation of endogenous RNase activity. The extract was then diluted, centrifuged, and added to an oligo(dT)-cellulose matrix in a spun column, supplied with the kit. The mRNA was allowed to bind to the matrix, washed with a high salt and then a low salt buffer, and then eluted at 65°C. The sample was concentrated by precipitation using potassium acetate, glycogen and 95% ethanol. Reverse transcription was then carried out using the First-Strand cDNA Synthesis Kit (Pharmacia Biotech, UK), according to the manufacturer's instructions. The mRNA solution was heated to 65°C for 10 mins and then placed on ice. To this was added the Bulk First-Strand cDNA Reaction Mix (containing MoMLV RT, BSA and dNTPs), DTT, and an oligo-dT primer. The mixture was incubated at 37°C for 1 hour, then heated to 95°C for 5 mins to denature the mRNA-cDNA duplex, and chilled on ice. PCR was then carried out using the U5657 and xhis primers as described for PoERV-B *env* (section 5.2.2.1). The PCR product was then Not I digested and cloned into pCI-neo as before. The clone obtained was identified by automated sequencing as described in section 5.2.2.2, using the same T7 and T3 primers, along with the sense primer PA878F and antisense primer PA1742R shown below:

PA878F	5′ - GGT ATA TCT TGG GGA ATA GT - 3′
PA1742R	5′ - GGA GAT CTT CCG TTA CAA TT - 3′

The resulting construct was designated pCIneo.penvA.

The PERV-MSL (PoERV-C) *env* clone (Akiyoshi *et al.*, 1998) in the pNEB193 vector was a gift from Papia Banerjee (BioTransplant Inc., Massachusetts, USA). The *env* gene was removed from the vector by digestion with the restriction enzymes Xba I (Gibco, UK) and

Sal I (Gibco, UK), and then cloned between the equivalent sites in pCI-neo (section 2.2.3). The resulting construct was designated pCIneo.mslenv.

6.2.3 Transient production of pseudotyped vectors

Pseudotyped vectors were produced using the HIT system as described in section 2.2.4.1, with the *env* constructs pCIneo.penvA, pCIneo.penvB (containing the *env* clone described in Chapter 5), and pCIneo.mslenv. Following transfection, serum free DMEM supplemented with SPIT was added to the cells (section 3.3). The vectors were titred 24 hours after transfection on the 293, AH927, Tb1Lu, RK13, ST, PK15 and MPK cell lines. Infection was again detected by staining for β-gal activity (section 2.2.4.2).

6.2.4 Production of permanent cell lines releasing pseudotyped vectors

The 293-GP cells were a gift from Derek Bain (Veterinary Pathology, University of Glasgow). These cells permanently express FeLV Gag, Pro and Pol. 293-GP cells were transfected with a PoERV *env* construct, either pCIneo.penvA, pCIneo.penvB or pCIneo.mslenv, using calcium phosphate precipitation (section 2.2.1.3). 48 hours after transfection 500μg/ml G418 was added to the culture medium. Following selection for a month the resistant colonies were pooled and passaged in G418 medium. In order to titrate pseudotyped vectors released from these cell lines, approximately 2 x 10^6 cells were seeded in 25cm^2 flasks and then transiently transfected with pHIT111 the following day. 48 hours after transfection the medium was removed, filtered and added to target cells as described for the HIT system (section 2.2.4.1). In this experiment 8μg/ml Polybrene was added to the target cells 24 hours prior to infection. Titres were determined by staining for β-gal activity (section 2.2.4.2).

6.3 Results

6.3.1 Cell free infection

In order to further characterise the host range of PoERV released from PK15 and MPK cells, a variety of cell lines from different species were incubated with medium taken from cultures of these cells. PoERV infection was then detected by amplification of *env*-specific sequences from target cell genomic DNA. Cells incubated with PK15 medium were tested for PoERV-A and -B sequences with the PV4 and PV2 primers respectively, and those incubated with MPK medium were tested for PoERV-C sequences with the PV3 primers. The ST and MPK cell lines were not incubated with PK15 medium in this experiment since they are reported to contain A- and B-specific sequences which could have interfered with PCR analysis, although they do not appear to encode infectious virus (Patience *et al.*, 1997a; Le Tissier *et al.*, 1997; D.Hart, pers. com.). The results of the PCRs are shown in Fig. 6.3 and summarised in Table 6.1. As can be seen in Fig. 6.3, bands at the expected size were amplified from the positive control DNA samples taken from untreated PK15 and MPK cells, indicating that the primers were able to recognise the PoERV sequences. A band at 336bp was amplified from PK15 DNA using the PV4 primers (Fig. 6.3.1 lane H). A band at 302bp was amplified from PK15 DNA using the PV2 primers (Fig. 6.3.2 lane H). A band at 282bp was amplified from MPK DNA using the PV3 primers (Fig. 6.3.3 lanes F and L). The PV3 primers also amplified a faint band at 282bp from DNA taken from ST cells incubated with MPK medium (Fig. 6.3.3 lane K), but not from the control, untreated ST cells (Fig. 6.3.3 lane E). This indicates that PoERV released from MPK cells was able to infect the ST cells. No bands were amplified from the DNA samples taken from the other cell lines, and therefore no evidence of infection was detected. No bands were amplified from the water control samples (Figs. 6.3.1 lane I, 6.3.2 lane I and 6.3.3 lane M). Each PCR reaction was carried out at least twice.

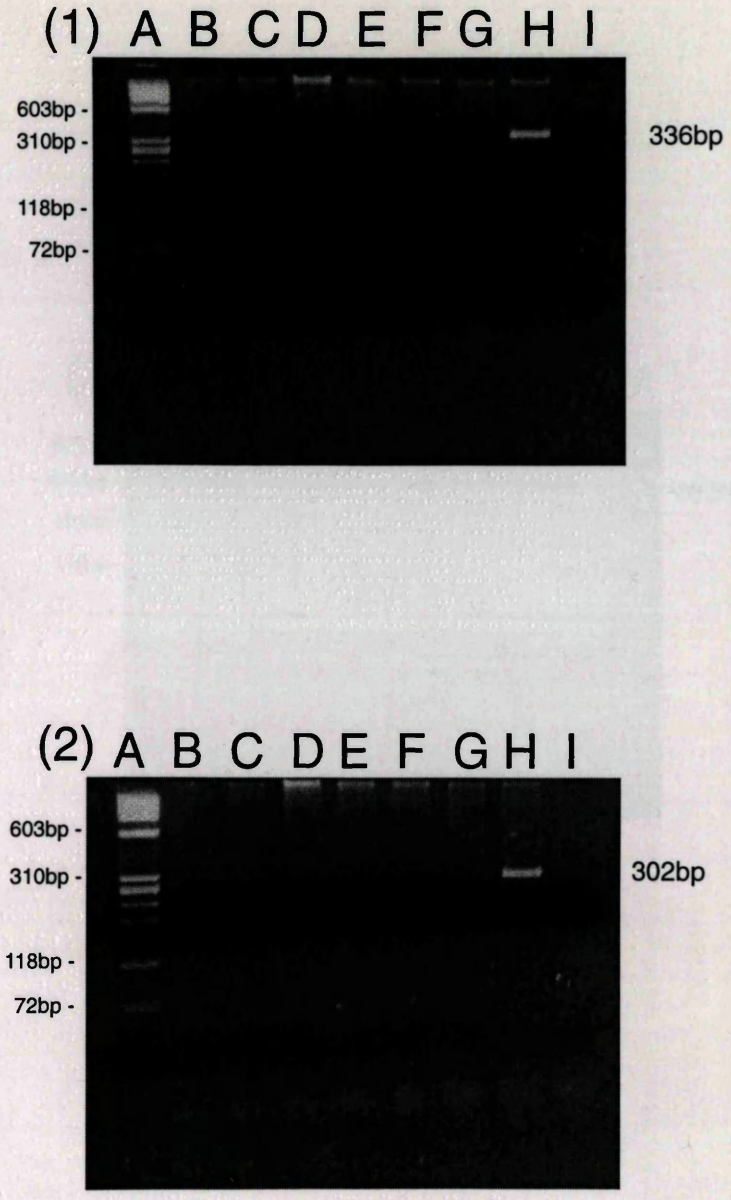


Fig. 6.3. PCR analysis of cells incubated with medium taken from PK15 or MPK cell cultures. All samples were run on 4% agarose gels.
(1) PCR analysis with the PV4 primers. A) ϕX174 DNA markers (Gibco, UK). B) Untreated RK-13 cells. C) Untreated AH927 cells. D) Untreated 293 cells. E) RK-13 cells incubated with PK15 medium. F) AH927 cells incubated with PK15 medium. G) 293 cells incubated with PK15 medium. H) PK15 cells. I) Water control.
(2) PCR analysis with the PV2 primers. A) ϕX174 DNA markers. B) Untreated RK-13 cells. C) Untreated AH927 cells. D) Untreated 293 cells. E) RK-13 cells incubated with PK15 medium. F) AH927 cells incubated with PK15 medium. G) 293 cells incubated with PK15 medium. H) PK15 cells. I) Water control.

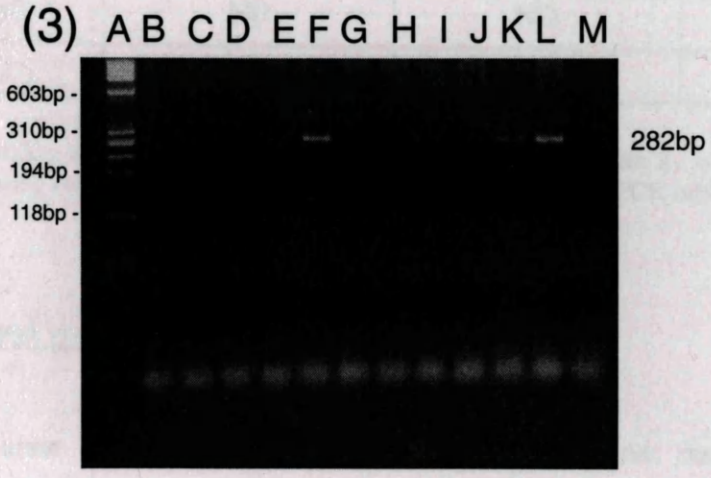


Fig. 6.3. Cont.
(3) PCR analysis with the PV3 primers. A) ϕX174 DNA markers. B) Untreated RK-13 cells. C) Untreated AH927 cells. D) Untreated 293 cells. E) Untreated ST cells. F) MPK cells. G) PK15 cells incubated with MPK medium. H) RK-13 cells incubated with MPK medium. I) AH927 cells incubated with MPK medium. J) 293 cells incubated with MPK medium. K) ST cells incubated with MPK medium. L) MPK cells. M) Water control

	Infection detected		
	PK15 medium		MPK medium
Cell type	**PoERV-A**	**PoERV-B**	**PoERV-C**
RK-13 (rabbit)	no	no	no
AH927 (feline)	no	no	no
293 (human)	no	no	no
PK15 (porcine)	ND	ND	no
ST (porcine)	ND	ND	yes

Table 6.1. Susceptibility of various cell lines to infection with PoERV strains by cocultivation with media from the virus releasing cell lines PK15 and MPK. Infection was detected by PCR as described in the text

6.3.2 Pseudotyped vectors

As a more accurate and sensitive method of defining the host range of the PoERV subgroups, pseudotyped vectors bearing the different PoERV Env proteins were produced. This was initially achieved through the use of the HIT system (Soneoka *et al.*, 1995), previously used to produce the rabies G pseudotyped vectors described in Chapter 3. The titres of vectors produced from the HIT system are expected to be higher than those of ERVs released from cultured cells. This method also allows investigation of the susceptibility of ST and MPK cells to infection with PoERV-A and -B. Vectors bearing the Env proteins of PoERV-A, PoERV-B, PERV-MSL or A-MLV were titred on a variety of cell lines from different species. Titres were determined by staining for β-gal activity, as described in section 2.2.4.2. The results are shown in Table 6.2.

Both PoERV-A and -B pseudotyped vectors were able to infect 293 cells, although the titres of 1-6 x 10^0 bcfu/ml were approximately 10^5 times lower than those obtained with the control vectors bearing the A-MLV Env, 0.5-2.5 x 10^5 bcfu/ml. Although neither of these vectors appeared to infect porcine cell lines, titres of around 10^3 bcfu/ml were obtained with the PERV-MSL pseudotype on ST and PK15 cells. However, this vector did not infect the MPK cells. None of the other cell types tested were susceptible to infection by any of the PoERV Env pseudotyped vectors. In addition, no titre was observed on the porcine

	Titres of pseudotyped vectors (bcfu/ml)				
Cell type	**PoERV-A**	**PoERV-B**	**PERV-MSL**	**A-MLV**	**control**
293 (human)	0-1 x 10^0	0-6 x 10^0	ND	0.5-2.5 x 10^5	0
AH927 (feline)	0	0	ND	0.4-6.8 x 10^5	0
ST (porcine)	0	0	3 x 10^3	0	0
PK15 (porcine)	0	0	1.1-1.2 x 10^3	0	0
MPK (porcine)	0	0	0	0	0
Tb1Lu (bat)	0	0	ND	4.4 x 10^4	0
RK13 (rabbit)	0	0	ND	1.5 x 10^5	0

Table 6.2. Titres of pseudotyped vectors obtained using the HIT system. Infection was detected by staining for β-gal activity. Where a range is given, titrations were carried out four times.

cell lines with the vectors bearing A-MLV Env. Control cells were transfected with the vector plasmid pHIT111 alone, and no titre was obtained from these on any target cell line.

In an attempt to increase the titres of the pseudotyped vectors, and to obtain a continuous supply for host range studies, stably transfected cell lines were made. Cells from the 293-GP cell line stably transfected with FeLV *gag/pro/pol* were subsequently transfected with the expression plasmid containing the *env* gene of PoERV-A (pCIneo.penvA), PoERV-B (pCIneo.penvB), or PERV-MSL (pCIneo.mslenv), and cells expressing both constructs were selected with 500µg/ml G418. The resulting cell lines were designated 293-GP/A, 293-GP/B and 293-GP/MSL respectively. Following transient transfection with pHIT111 the vectors were titred on human 293T cells. The titres obtained in this way are shown in Table 6.3. As expected, these are higher than those obtained in the transient system. A titre of 3.5×10^2 bcfu/ml was obtained with the PoERV-A pseudotype, 8.6×10^2 bcfu/ml with the PoERV-B pseudotype, and 2.8×10^1 bcfu/ml with the PERV-MSL pseudotype. However, a titre of 1.2×10^3 bcfu/ml was also obtained with the negative control medium taken from 293-GP cells. These cells were expected to release Env-free vector which would be unable to infect the 293T cells. No blue cells were detected in uninfected 293T cells, incubated with the X-Gal staining solution. This suggests that the vector producing cells were contaminated with a functional *env* gene or an infectious virus. A sample of medium taken from the 293-GP cells gave a negative result in a virus isolation assay, carried out by the Feline Virus Unit, University of Glasgow. This assay is designed to detect retroviruses that infect feline cells and therefore a negative result does not rule out the possibility of contamination with virus possessing a different host range. However, the most likely source of contamination of this cell line is FeLV.

Titres on 293T cells (bcfu/ml)			
PoERV-A	**PoERV-B**	**PERV-MSL**	**Env-free vector**
3.5×10^2	8.6×10^2	2.8×10^1	1.2×10^3

Table 6.3. Titres obtained on 293T cells with PoERV Env pseudotyped vectors produced from permanent cell lines.

6.3.3 Receptor interference between subgroups of PoERV

As an initial investigation into the patterns of receptor interference between the PoERV subgroups, pseudotyped vectors were titred on the vector producing cell lines 293-GP/A, 293-GP/B and 293-GP/MSL. These titres are shown in Table 6.4. It was expected that the Env protein produced in the cell would block its receptor, preventing infection by pseudotyped vectors that recognise the same receptor. A lack of interference would imply that the Env proteins of the different subgroups recognise different receptors. As shown in Table 6.4, infection was detected for all combinations of vectors and cell lines. No obvious receptor interference was observed, not even when vector was titred on its equivalent pseudotype producing cell line, expressing the same Env protein.

	Titres of pseudotyped vectors (bcfu/ml)		
Cell line	**PoERV-A**	**PoERV-B**	**PERV-MSL**
293-GP/A	4.4×10^1	1.5×10^2	2.0×10^1
293-GP/B	2.4×10^1	7.6×10^1	2.0×10^1
293-GP/MSL	4.0×10^1	1.4×10^2	2.0×10^1

Table 6.4. Titres obtained on permanent vector producing cell lines with PoERV pseudotyped vectors produced from these cells.

6.4 Discussion

The results obtained in the cell free infection experiment, shown in Fig. 6.3 and Table 6.1, provide little evidence of PoERV infectivity. Transferred PoERV-specific sequences were only detected in the ST cells incubated with MPK medium. It is unlikely that the 293 cells were resistant to infection with PoERV since Patience *et al.* (1997a) have reported infection of 293 and 293T cells incubated with PK15 medium. It seems possible that the number of cells infected with PK15-derived virus was too low to be detected in this experiment. This would suggest that these viruses are expressed at a low level or that they are unstable and require close contact between virus-releasing and target cells for efficient infection. It is also possible that the PoERV Env proteins bind with low affinity to the receptor on human cells. The infection of ST cells with PoERV-C indicates that these viruses are adapted to infect pig cells. The ability to detect this infection may also imply that PoERV-C is more stable, perhaps reflecting a more recent integration into the porcine genome (section 6.1.1).

The pseudotyped vectors were produced as a more sensitive method of investigating the host range of the PoERV subgroups. The results presented in Table 6.2 show that a wider range of PoERV infectivity can be detected by this method. Both PoERV-A and -B pseudotypes were able to infect 293 cells, and the PERV-MSL (PoERV-C) pseudotype was able to infect PK15 and ST cells. The titres obtained with these pseudotyped vectors were much lower than those obtained with the vectors bearing A-MLV Env. This suggests that the PoERV Env proteins are less efficient at promoting cell entry, or that they have a low affinity for the receptor alleles present on these target cell types. It is also possible that these Env proteins are incorporated into the MLV-based vectors with lower efficiency than A-MLV Env. No infectivity was detected with the other target cell types tested, as shown in Table 6.2. Takeuchi *et al.* (1998) reported a low susceptibility of Tb1Lu cells to a similar PoERV-B pseudotyped vector. However, this was achieved with higher titre vector produced from a stably transfected cell line. For comparison, this cell line releases vector at titres of 1-5 x 10^7 infectious units/ml when transfected with an A-MLV Env construct (Cosset *et al.*, 1995).

Although the PERV-MSL (PoERV-C) pseudotype was able to infect ST and PK15 cells, no titre was obtained on MPK cells. It seems likely that this was the result of infection interference since MPK cells, but not ST or PK15, contain PoERV-C proviral sequences. The Env protein expressed by these sequences would be expected to block the receptors on the MPK cells. The lack of titre obtained with the PoERV-A and -B pseudotypes on the porcine cell lines may also be the result of infection interference. ST and MPK cells do not release infectious PoERV-A and -B, but the sequences present in these cells may encode Env capable of blocking receptors. It is also possible that these vectors are unable to infect these cell lines due to a lack of the necessary receptor. Either situation would imply that PERV-MSL recognises a different receptor from the other two subgroups. Although no infection of porcine cells by the PoERV-A and -B pseudotypes was detected here, Takeuchi *et al.* (1998) reported that these pseudotypes could infect porcine cell lines with variable efficiencies. This could be an effect of the system used for vector production, or it could reflect a variability in the level of available receptor, perhaps as a result of changes in PoERV or receptor expression by the target cells. Although Takeuchi *et al.* achieved a titre of 2.1 x 10^3 infectious units/ml for the PoERV-B pseudotype on PK15 cells, these cells were resistant in an earlier experiment.

No titre was obtained on the porcine cell lines with the control vectors bearing the A-MLV Env. This lack of infectivity was not due to a block acting post entry since vectors pseudotyped with PERV-MSL Env were able to infect ST and PK15 cells. This could indicate that these cells lack a functional receptor for A-MLV. However, this virus is able to infect cells from a large number of species (Miller, 1997). It therefore seems possible that infection was prevented by receptor interference with PoERV Env proteins. As discussed above, all three cell lines may express PoERV-A and -B Env proteins capable of blocking their receptors. Although no interference between PoERV and A-MLV was detected in the experiment described in section 4.3.2, this was only tested in human cells, perhaps only infected with PoERV-B. It is possible that the interference occurs with PoERV-A, or that it is affected by species-specific differences in the receptor. Certain xenotropic MLV strains are unable to infect *Mus dunni* cells preinfected with ecotropic and amphotropic viruses (Chesebro and Wehrly, 1985). It would be interesting to test whether the porcine homologue of the A-MLV vector, Ram-1, can render non-susceptible cell lines sensitive to infection by PoERVs.

It is difficult to conclude very much from the results obtained with the pseudotyped vectors produced by the permanent expression cell lines (Tables 6.3 and 6.4). Although the titres were higher than those obtained with the HIT system, a titre of 1 x 10^3 bcfu/ml was obtained from the negative control cells. These control cells were expected to express only the FeLV Gag/Pol construct and the vector plasmid pHIT111, and therefore to release Env-free vectors. The addition of Polybrene is unlikely to have been the cause of this unexpected titre since it is reported to have no effect on the infectivity of Env-free virus (Porter *et al.*, 1998). Although it is possible that the pHIT111 plasmid may have been transferred to the target cells, this has not been observed in any previous experiments. The titre obtained from these cells indicates that they were contaminated with an envelope protein capable of being incorporated into the vector particles, or a C-type virus capable of packaging the vector genome, although no replication-competent virus could be detected by virus-isolation assay. It should be noted that the assay used was designed to detect feline retroviruses, however these are the most likely source of contamination since they are studied in the same laboratory. It appears to be possible that the contamination is derived from an endogenous retrovirus present in the genome of the 293 cells used to produce the original clones. If this is shown to be the case it will have serious implications for the use of such cells for the manufacture of retroviral packaging cell lines. Unless the cause of this result can be determined the titres obtained with the pseudotyped vectors in this experiment cannot be interpreted. If neutralising antibodies become available these could be used to test the specificity of the pseudotype titres, as described in section 3.3.4.

It was expected that the pseudotyped vectors would show reduced infectivity on their own producing cell lines, due to the expression of Env protein blocking receptors. However, as can be seen in Table 6.4, no significant reduction in infectivity was observed. This lack of interference suggests that these samples were contaminated in the same way as the negative control vectors. However, it is also possible that the Env expression level of these cell lines was too low to significantly block the receptors. These cell lines were not derived from single cell clones, and no antibodies have yet become available to analyse the level of Env expression. Although no interference was observed between the PoERV subgroups these results are not reliable. However, Takeuchi *et al.* (1998) have also found no evidence of interference between these viruses. This supports the prediction that PoERV-A, -B and -C recognise different receptors, and can be placed into different interference groups.

Both PoERV-A and -B pseudotypes have been shown to be able to infect a large number of human cell lines in addition to the 293 cells reported here, while the PoERV-C pseudotype has only so far been shown to infect HT1080 fibrosarcoma cells (Takeuchi *et al.*, 1998). This suggests that the A and B subgroups may pose the greater risk in xenotransplantation, particularly since they have been found in all pigs so far tested (Le Tissier *et al.*, 1997; Martin *et al.*, 1998; D.Hart, pers.com.). It is possible that the PoERVs may not be able to establish a productive infection in all the cells susceptible to these pseudotyped vectors. However, the ability of the virus to enter human cells may still be significant since even partial expression may allow recombination with human retroviral sequences. While it is important to define the cell types susceptible to PoERV, resistant cell lines may also prove useful for receptor identification. Following transfection with porcine or human DNA such cells could be tested for their susceptibility to the vectors pseudotyped with PoERV Env, in a method similar to that used by Albritton *et al.* (1989) and O'Hara *et al.* (1990) described in section 1.2.1. Identification of the receptors recognised by the PoERV subgroups would be a useful step in the ongoing investigation into the risk posed by these viruses in xenotransplantation.

CHAPTER 7

Chapter 7 General discussion

These studies made use of the ability of retroviruses to incorporate the envelope proteins of other viruses, and adopt the host range and cell fusion properties of the virus donating the envelope protein. In the first section of this work, the process of pseudotype formation was exploited in an attempt to produce stable gene delivery vectors with an expanded host range, by pseudotyping MLV-based vectors with the rabies envelope glycoprotein G. Although high titre VSV G pseudotyped vectors had been produced previously (Burns *et al.*, 1993), the vectors produced in this system had low titres, probably as a result of inefficient incorporation of the rabies G protein.

As discussed in section 3.4, a likely factor in the inefficient incorporation of rabies G is its low rate of transport through the cell compared to the VSV glycoprotein (Whitt *et al.*, 1991), resulting in a lower level of expression on the cell surface. This strongly suggests that there is no interaction between the rabies glycoprotein and the MLV MA protein, since MLV is able to efficiently incorporate its own Env protein when this is expressed at low levels on the cell surface (Suomalainen and Garoff, 1994). It seems more likely that the G protein is incorporated passively, perhaps as a consequence of alterations in the structure of the lipid bilayer, as implied by the model B put forward by Závada (1982), or because it is not prevented from moving into areas of viral budding by interactions with host cellular proteins (Dong *et al.*, 1992; Swanstrom and Wills, 1997; Young *et al.*, 1990). Differences in the assembly and budding mechanisms of retroviruses and rhabdoviruses may also influence the incorporation of rabies G by the MLV-based vector. These are discussed further in section 3.4.

In the second section, pseudotyped vectors were used to study the host ranges of the PoERV subgroups. The ability of PoERV to infect human cells has been demonstrated by cocultivation of virus-releasing cells with target cells (Chapter 4; Patience *et al.*, 1997a). However, this was difficult to achieve by exposure to medium taken from cultures of PoERV-releasing cells (Chapter 6), indicating that these viruses have low infectivity or are unstable. The cloning of the Env genes (Chapters 5 and 6) allowed production of pseudotyped vectors with the host ranges of the different PoERV subgroups. These vectors provided a more sensitive method of investigating infection and allowed rapid detection of

infected cells. Their use also allowed individual analysis of the different subgroup Env proteins. The results described in section 6.3 show that the vectors pseudotyped with the PoERV-A and -B Env proteins are able to infect human 293 cells, implying that these viruses may pose a risk to xenotransplant recipients. The vector pseudotyped with PERV-MSL (PoERV-C) Env was only observed to infect porcine cell lines, suggesting that these may have an ecotropic host range. However, as discussed in section 6.4, other human cell lines may be susceptible to this virus.

Although the pseudotyped vectors were found to be a useful method of investigating PoERV host range, the titres achieved were, like those of the rabies G pseudotype, much lower than those of the control vector bearing the A-MLV Env protein (section 6.3, Table 6.2). It may be expected that an envelope protein from a closely related virus would be more efficiently incorporated by the retroviral vector than the rabies glycoprotein. Like MLV, the PoERVs are mammalian C-type retroviruses and, as described in Chapter 5, MLV and PoERV Env proteins have very similar structures. It therefore seems likely that the low titres were due to the binding and fusion properties of the PoERV Env proteins, or the level of available receptors on the target cells. Many ERVs are defective due to the lack of selective pressure to produce infectious viral particles (section 1.1.6). As mentioned above, PoERVs released from cultured porcine cell lines do not easily infect target cells. In addition, Patience *et al.* (1997a) found that infected 293 and muscle RD cells released PoERV at low titres of less than 500 infectious units/ml. While this may be the result of a low level of expression, it could also indicate that these viruses have poor infectivity. As Env is the only PoERV protein present in the pseudotyped vectors, it seems likely that poor infectivity could, at least partly, be due to a defect in Env function. However, it is also possible that these Env proteins were inefficiently incorporated.

Several studies have suggested that retroviral Env proteins contain a signal in the cytoplasmic domain which is required for efficient incorporation (Brody *et al.*, 1994; Granowitz *et al.*, 1991; Gray and Roth, 1993; Januszeski *et al.*, 1997; Yu *et al.*, 1993). Although the PoERV and MLV Env proteins are very similar they do not have identical cytoplasmic domains. This may indicate that different sequences are required for the specific interaction with the homologous MA protein. As discussed in section 5.4, two of the intracytoplasmic residues, Leu 602 and Gln 604, reported to be involved in the efficient

incorporation of MoMLV Env (Januszeski *et al.*, 1997) are not conserved in the PoERV-B Env protein. It is therefore possible that the PoERV Env proteins lack the sequence required for specific incorporation into MLV-based vectors. Although both PoERV-A and -C Env proteins also lack Gln 604, PoERV-C does contain the conserved Leu residue. It may be interesting to investigate whether this residue increases incorporation of this protein by the MLV-based vector. These residues are both conserved in the GALV Env protein (section 5.3). MLV vectors pseudotyped with this protein have been reported to have similar titres to vector bearing A-MLV Env, implying that GALV Env is efficiently incorporated (Miller *et al.*, 1991). If the PoERV Env proteins do lack a sequence required for an interaction with the MLV MA it would be expected that a high level of expression would be required for their efficient incorporation into the MLV-based vector (Suomalainen and Garoff, 1994). Takeuchi *et al.* (1998) observed variations in the titres of MLV-based vectors pseudotyped with PoERV Env proteins, plated on human and porcine cells. While it is quite likely that these variations were due to changes in levels of available receptor, it could also be the result of changes in Env expression and incorporation. In the absence of mAbs directed against the PoERV Env proteins, it is not possible to determine the level of expression of these proteins. Therefore it is possible that the low titres reported here were partly due to low levels of Env expression resulting in poor incorporation. Such inefficient incorporation of the Env proteins could prevent detection of infection of cell lines with low susceptibility.

The establishment of cell lines releasing high titres of vector pseudotyped with the PoERV Env proteins would be useful for further characterisation of the host ranges of these viruses, and for the characterisation of potential receptors. The use of a stable cloned packaging cell line, such as the TELCeB6 line used by Takeuchi *et al.* (1998), would be expected to result in increased vector titres. If the low titres are a reflection of inefficient incorporation of PoERV Env, these should be improved by increasing the expression of Env relative to Gag. This could perhaps be achieved by transfecting a higher concentration of the *env* construct, or through the use of a plasmid that contains the *env* and selector genes in the same reading frame. Since expression of the selector gene in such a construct occurs as a result of continued expression from *env*, this should reduce any downregulation of Env expression by the cell.

In conclusion, the results described in this study demonstrate that the pseudotyping of retroviral vectors can be used both to target vectors to a defined host range, or to investigate the unknown host range of a new virus. Unfortunately, pseudotyping with the rabies G protein proved to be inefficient and the resulting vectors are unlikely to be useful for gene delivery and gene therapy protocols. However, analysis of the factors influencing the production of these pseudotypes should aid investigations into the processes of virion assembly and envelope incorporation. The vectors pseudotyped with the PoERV Env proteins also had very low titres. However, they still proved to be useful for detection of cell lines susceptible to infection by these vectors. The production of higher titre stocks of these pseudotypes would allow more confident detection of non susceptible cell lines, and also increase the utility of these pseudotyped vectors in receptor studies. As it has been established that these viruses are able to infect human cells, further investigation of their host range and infectivity is required in order to assess the risk posed to the human recipients of xenotransplants.

REFERENCES

List of references

Adam, M.A. and Miller, A.D. (1988). Identification of a signal in a murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. *J.Virol.* **62**, 3802-3806.

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**, 4985-4990.

Akiyoshi, D.E., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P., and Fishman, J.A. (1998). Identification of a full-length cDNA for an endogenous retrovirus of miniature swine. *J.Virol.* **72**, 4503-4507.

Albritton, L.M., Kim, J.W., Tseng, L., and Cunningham, J.M. (1993). Envelope-binding domain in the cationic amino acid transporter determines the host range of ecotropic murine retroviruses. *J.Virol.* **67**, 2091-2096.

Albritton, L.M., Tseng, L., Scadden, D., and Cunningham, J.M. (1989). A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell* **57**, 659-666.

Allan, J.S. (1996). Xenotransplantation at a crossroads: Prevention versus progress. *Nature Med.* **2**, 18-21.

Andersen, K.B. and Nexo, B.A. (1983). Entry of murine retroviruses into mouse fibroblasts. *Virology* **125** , 85-98.

Andersson, P., Goldfarb, M.P., and Weinberg, R.A. (1979). A defined subgenomic fragment of in vitro synthesized Moloney sarcoma virus DNA can induce cell transformation upon transfection. *Cell* **16**, 63-75.

Andre, P., Kim, A., Khrapko, K., and Thilly, W.G. (1997). Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. *Genome Res.* **7**, 843-852.

Anilionis, A., Wunner, W.H., and Curtis, P.J. (1981). Structure of the glycoprotein gene in rabies virus. *Nature* **294**, 275-278.

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.* **70**, 2581-2585.

Armstrong, J.A., Porterfield, J.S., and De Madrid, A.T. (1971). C-type virus particles in pig kidney cell lines. *J.gen.Virol.* **10**, 195-198.

Bach, F.H., Fishman, J.A., Daniels, N., Proimos, J., Anderson, B., Carpenter, C.B., Forrow, L., Robson, S.C., and Fineberg, H.V. (1998). Uncertainty in xenotransplantation: Individual benefit versus collective risk. *Nature Med.* **4**, 141-144.

Bader, J.P. (1965). The requirement for DNA synthesis in the growth of Rous sarcoma and Rous associated viruses. *Virology* **26**, 253-261.

Baer, G.M., Bellini, W.J., and Fishbein, D.B. (1990). Rhabdoviruses. *Virology* (Fields, B.N., Knipe, D.M., *et al.*, Eds.) Second edition. Raven Press, New York. 883-930.

Bailey, L.L., Nehtseu-Canarella, S.L., Conepcion, W., and Joley, W.B. (1985). Baboon to human cardiac xenotransplantation in a neonate. *JAMA* **254**, 3321

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **226**, 1209-1211.

Ban, J., Portetelle, D., Altaner, C., Horion, B., Milan, D., Krchnak, V., Burny, A., and Kettmann, R. (1993). Isolation and characterization of a 2.3-kilobase-pair cDNA fragment encoding the binding domain of the bovine leukemia virus cell receptor. *J.Virol.* **67**, 1050-1057.

Battini, J.-L., Danos, O., and Heard, J.M. (1995). Receptor-binding domain of murine leukemia virus envelope glycoproteins. *J.Virol.* **69**, 713-719.

Battini, J.-L., Danos, O., and Heard, J.M. (1998). Definition of a 14-amino-acid peptide essential for the interaction between the murine leukemia virus amphotropic envelope glycoprotein and its receptor. *J.Virol.* **72**, 428-435.

Battini, J.-L., Heard, J.M., and Danos, O. (1992). Receptor choice determinants in the envelope glycoproteins of amphotropic, xenotropic, and polytropic murine leukemia viruses. *J.Virol.* **66**, 1468-1475.

Beale, A.J., Christofinis, C.G., and Furminger, I.G.S. (1963). Rabbit cells susceptible to rubella virus. *Lancet* **2**, 640-641.

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.* **61**, 1639-1646.

Benmansour, A., Leblois, H., Coulon, P., Tuffereau, C., Gaudin, Y., Flamand, A., and Lafay, F. (1991). Antigenicity of rabies virus glycoprotein. *J.Virol.* **65**, 4198-4203.

Benveniste, R.E. and Todaro, G.J. (1974). Evolution of C-type viral genes: Inheritance of exogenously acquired viral genes. *Nature* **252**, 456-459.

Benveniste, R.E. and Todaro, G.J. (1975). Evolution of type C viral genes: Preservation of ancestral murine type C viral sequences in pig cellular DNA. *Proc.Natl.Acad.Sci.USA* **72**, 4090-4094.

Berkowitz, R.D. and Goff, S.P. (1993). Point mutations in Moloney murine leukemia virus envelope protein: Effects on infectivity, virion association, and superinfection resistance. *Virology* **196**, 748-757.

Berkowitz, R.D., Ohagen, Å., Höglund, S., and Goff, S.P. (1995). Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric gag polyproteins during RNA packaging in vivo. *J.Virol.* **69**, 6445-6456.

Best, S., Le Tissier, P., Towers, G., and Stoye, J.P. (1996). Positional cloning of the mouse retrovirus restriction gene, *Fv1*. *Nature* **382**, 826-829.

Bittner, J.J. (1936). Some possible effects of nursing on the mammary gland tumor incidence in mice. *Science* **84**, 162-162.

Bittner, J.J. (1942). The milk-influence of breast tumors in mice. *Science* **95**, 462-463.

Blumenthal, R., Puri, A., and Dimitrov, D.S. (1991). Initial steps of enveloped virus entry into animal cells. *Biotechnology of cell regulation* (Verna, R. and Nishizuka, Y., Eds.) Raven Press, New York. 115-133.

Boeke, J.D. and Stoye, J.P. (1997). Retrotransposons, endogenous retroviruses, and the evolution of retroelements. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 343-435.

Boettiger, D. (1979). Animal virus pseudotypes. *Prog.med.Virol.* **25**, 37-68.

Boyd, M.T., Bax, C.M., Bax, B.E., Bloxam, D.L., and Weiss, R.A. (1993). The human endogenous retrovirus ERV-3 is upregulated in differentiating placental trophoblast cells. *Virology* **196**, 905-909.

Bracci, L., Antoni, G., Cusi, M.G., Lozzi, L., Niccolai, N., Petreni, S., Rustici, M., Santucci, A., Soldani, P., Valensin, P.E., and Neri, P. (1988). Antipeptide monoclonal antibodies inhibit the binding of rabies virus glycoprotein and alpha-bungarotoxin to the nicotinic acetylcholine receptor. *Mol.Immunol.* **25**, 881-888.

Brody, B.A., Rhee, S.S., and Hunter, E. (1994). Postassembly cleavage of a retroviral glycoprotein cytoplasmic domain removes a necessary incorporation signal and activates fusion activity. *J.Virol.* **68**, 4620-4627.

Brown, J., Matthews, A.L., Sandstrom, P.A., and Chapman, L.E. (1998). Xenotransplantation and the risk of retroviral zoonosis. *Trends Microbiol.* **6**, 411-415.

Brown, P.O. (1997). Integration. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 161-203.

Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T.L., Bukrinskaya, A.G., Haggerty, S., and Stevenson, M. (1992). Active nuclear import of human immunodeficiency virus type 1 preintegration complex. *Proc.Natl.Acad.Sci.USA* **89**, 6580-6584.

Bullough, P.A., Hughson, F.M., Skehel, J.J., and Wiley, D.C. (1994). Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37-43.

Burger, S.R., Remaley, A.T., Danley, J.M., Moore, J., Muschel, R.J., Wunner, W.H., and Spitalnik, S.L. (1991). Stable expression of rabies virus glycoprotein in Chinese hamster ovary cells. *J.gen.Virol.* **72**, 359-367.

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 titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc.Natl.Acad.Sci.USA* **90**, 8033-8037.

Butler, D. (1999). FDA warns on primate xenotransplants. *Nature* **398**, 649

Cepko, C.L., Roberts, B.E., and Mulligan, R.C. (1984). Construction and applications of a highly transmissable murine retrovirus shuttle vector. *Cell* **37**, 1053-1062.

Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263-273.

Chapman, L.E., Folks, T.M., Salomon, D.R., Patterson, A.P., Eggerman, T.E., and Noguchi, P.D. (1995). Xenotransplantation and xenogeneic infections. *N.Engl.J.Med.* **333**, 1498-1501.

Chapman, L.E. and Khabbaz, R.F. (1994). Etiology and epidemiology of the Four Corners hantavirus outbreak. *Infect.Agents Dis.* **3**, 234-244.

Chen, B.F., Hsieh, C.L., Chen, D.S., and Hwang, L.H. (1992). Improved gene expression by a U3-based retroviral vector. *Biochem.Biophys.Res.Commun.* **184**, 330-337.

Chen, B.F., Hwang, L.H., and Chen, D.S. (1993). Characterization of a bicistronic retroviral vector composed of the swine vesicular disease virus internal ribosome entry site. *J.Virol.* **67**, 2142-2148.

Chen, Z., Luckay, A., Sodora, D.L., Telfer, P., Reed, P., Gettie, A., Kanu, J.M., Sadek, R.F., Yee, J., Ho, D.D., Zhang, L., and Marx, P.A. (1997). Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J.Virol.* **71**, 3953-3960.

Chesebro, B. and Wehrly, K. (1985). Different murine cell lines manifest unique patterns of interference to superinfection by murine leukemia viruses. *Virology* **141**, 119-129.

Chou, P.Y. and Fasman, G.D. (1978a). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv.Enzymol.* **47**, 45-148.

Chou, P.Y. and Fasman, G.D. (1978b). Empirical predictions of protein conformations. *Annu.Rev.Biochem.* **47**, 251-276.

Chu, T.H. and Dornburg, R. (1995). Retroviral vector particles displaying the antigen-binding site of an antibody enable cell-type-specific gene transfer. *J.Virol.* **69**, 2659-2663.

Cleverley, D.Z. and Lenard, J. (1998). The transmembrane domain in viral fusion: Essential role for a conserved glycine residue in vesicular stomatitis virus G protein. *Proc.Natl.Acad.Sci.USA* **95**, 3425-3430.

Cline, J., Braman, J.C., and Hogrefe, H.H. (1996). PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nuc.Acids Res.* **24**, 3546-3551.

Coffin, J.M. (1990). Retroviridae and their replication. *Virology* (Fields, B.N., Knipe, D.M., *et al.*, Eds.) Second edition. Raven Press, New York. 1437-1500.

Coffin, J.M. (1995). HIV population dynamics *in vivo*: Implications for genetic variation, pathogenesis and therapy. *Science* **267**, 483-489.

Coffin, J.M., Hughes, S.H., and Varmus, H.E. (1997). The interactions of retroviruses and their hosts. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 335-341.

Conzelmann, K.-K., Cox, J.H., Schneider, L.G., and Thiel, H.-J. (1990). Molecular cloning and complete nucleotide sequence of the attenuated rabies virus SAD B19. *Virology* **175**, 485-499.

Conzelmann, K.-K. and Schnell, M. (1994). Rescue of synthetic genomic RNA analogs of rabies virus by plasmid-encoded proteins. *J.Virol.* **68**, 713-719.

Copeland, N.G. and Cooper, G.M. (1979). Transfection by exogenous and endogenous murine retrovirus DNAs. *Cell* **16**, 347-356.

Cosset, F.-L., Takeuchi, Y., Battini, J.-L., Weiss, R.A., and Collins, M.K.L. (1995). High-titer packaging cells producing recombinant retroviruses resistant to human serum. *J.Virol.* **69**, 7430-7436.

Cozzi, E. and White, D.J.G. (1995). The generation of transgenic pigs as potential organ donors for humans. *Nature Med.* **1**, 964-966.

Cunningham, J.M. and Kim, J.W. (1994). Cellular receptors for type C retroviruses. *Cellular receptors for animal viruses* (Wimmer, E., Ed.) Cold Spring Harbor Press, New York. 49-59.

Dalgleish, A.G., Beverley, P.C.L., Clapham, P.R., Crawford, D.H., Greares, M.F., and Weiss, R.A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763-767.

Danos, O. and Mulligan, R.C. (1988). Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc.Natl.Acad.Sci.USA* **85**, 6460-6464.

DeLarco, J. and Todaro, G.J. (1976). Membrane receptors for murine leukemia viruses: Characterization using purified envelope glycoprotein, gp71. *Cell* **8**, 365-371.

Delwart, E.L., Mosialos, G., and Gilmore, T. (1990). Retroviral envelope glycoproteins contain a 'leucine zipper' like repeat. *AIDS Res.Hum.Retroviruses* **6**, 703-706.

Diamond, L.E., McCurry, K.R., Martin, M.J., McClellan, S.B., Oldham, E.R., Platt, J.L., and Logan, J.S. (1996). Characterization of transgenic pigs expressing functionally active human CD59 on cardiac endothelium. *Transplantation* **61**, 1241-1249.

Dickson, C., Eisenman, R., Fan, H., Hunter, E., and Teich, N. (1982). Protein biosynthesis and assembly. *RNA tumour viruses* (Weiss, R., Teich, N., Varmus, H., and Coffin, J., Eds.) Cold Spring Harbor Laboratory Press, New York. 513-648.

Dimitrov, D.S., Golding, H., and Blumenthal, R. (1991). Initial stages of HIV-1 envelope glycoprotein-mediated cell fusion monitored by a new assay based on redistribution of fluorescent dyes. *AIDS Res.Hum.Retroviruses* **7**, 799-805.

Doms, R.W., Keller, D.S., Helenius, A., and Balch, W.E. (1987). Role for adenosine triphosphate in regulating the assembly and transport of vesicular stomatitis virus G protein trimers. *J.Cell Biol.* **105**, 1957-1969.

Doms, R.W., Ruusala, A., Machamer, C., Helenius, J., Helenius, A., and Rose, J.K. (1988). Differential effects of mutations in three domains on folding, quaternary structure, and intracellular transport of vesicular stomatitis virus G protein. *J.Cell Biol.* **107**, 89-99.

Donahue, R.E., Kessler, S.W., Bodine, D., McDonagh, K., Dunbar, C., Goodman, S., Agricola, B., Byrne, E., Raffeld, M., Moen, R., Bacher, J., Zsebo, K.M. and Nienhuis, A.W. (1992). Helper virus induced T cell lymphoma in nonhuman primates after retroviral mediated gene transfer. *J.Exp.Med.* **176**, 1125-1135.

Dong, J., Roth, M.G., and Hunter, E. (1992). A chimeric avian retrovirus containing the influenza virus hemagglutinin gene has an expanded host range. *J.Virol.* **66**, 7374-7382.

Donnelly-Roberts, D.L. and Lentz, T.L. (1991). Structural and conformational similarity between synthetic peptides of curaremimetic neurotoxins and rabies virus glycoprotein. *Mol.Brain Res.* **11**, 107-113.

Doolittle, R.F. and Feng, D.-F. (1992). Tracing the origin of retroviruses. *Curr.Top. Microbiol.Immunol.* **176**, 195-211.

Doolittle, R.F., Feng, D.-F., Johnson, M.S., and McClure, M.A. (1989). Origins and evolutionary relationships of retroviruses. *Q.Rev.Biol.* **64**, 1-30.

Doolittle, R.F., Feng, D.-F., McClure, M.A., and Johnson, M.S. (1990). Retrovirus phylogeny and evolution. *Curr.Top.Microbiol.Immunol.* **157**, 1-18.

Dorling, A., Riesbeck, K., Warrens, A., and Lechler, R. (1997). Clinical xenotransplantation of solid organs. *Lancet* **349**, 867-871.

Dougherty, J.P. and Temin, H.M. (1986). High mutation rate of a spleen necrosis virus-based retrovirus vector. *Mol.Cell Biol.* **6**, 4387-4395.

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**, 3084-3091.

Dougherty, J.P., Wisniewski, R., Yang, S., Rhode, B.W., and Temin, H.M. (1989). New retrovirus helper cells with almost no nucleotide sequence homology to retrovirus vectors. *J.Virol.* **63**, 3209-3212.

DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.M., Miller, J.H., and Calos, M.P. (1987). Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol.Cell Biol.* **7**, 379-387.

Duesberg, P.H. and Vogt, P.K. (1969). On the role of DNA synthesis in avian tumor virus infection. *Proc.Natl.Acad.Sci.USA* **64**, 939-946.

Dunphy, W.G. and Rothman, J.E. (1985). Compartmental organization of the Golgi stack. *Cell* **42**, 13-21.

Durrer, P., Gaudin, Y., Ruigrok, R.W.H., Graf, R., and Brunner, J. (1995). Photolabeling identifies a putative fusion domain in the envelope glycoprotein of rabies and vesicular stomatitis viruses. *J.Biol.Chem.* **270**, 17575-17581.

Eiden, M.V., Farrell, K., Warsowe, J., Mahan, L.C., and Wilson, C.A. (1993). Characterization of a naturally occurring ecotropic receptor that does not facilitate entry of all ecotropic murine retroviruses. *J.Virol.* **67**, 4056-4061.

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

Emi, N., Friedmann, T., and Yee, J.-K. (1991). Pseudotype formation of murine leukemia virus with the G protein of vesicular stomatitis virus. *J.Virol.* **65**, 1202-1207.

Epstein, H. and Bichard, M. (1984). Pig. *Evolution of domesticated animals* (Mason, I.L., Ed.) Longman, London and New York. 145-162.

Fass, D., Davey, R.A., Hamson, C.A., Kim, P.S., Cunningham, J.M., and Berger, J.M. (1997). Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution. *Science* **277**, 1662-1666.

Fass, D., Harrison, S.C., and Kim, P.S. (1996). Retrovirus envelope domain at 1.7 Å resolution. *Nature Struct.Biol.* **3**, 465-469.

Finnegan, D.J. (1983). Retroviruses and transposable elements - which came first? *Nature* **302**, 105-106.

Finnegan, D.J. (1994). Retroviruses and transposons. Wandering retroviruses? *Curr.Biol.* **4**, 641-643.

Flamand, A. (1970). Étude genetique du virus de la stomatite vésiculaire: Classement de mutants thermosensible spontanés en groups de complementation. *J.gen.Virol.* **8**, 187-195.

Florkiewicz, R.Z., Smith, A., Bergmann, J.E., and Rose, J.K. (1983). Isolation of stable mouse cell lines that express cell surface and secreted forms of the vesicular stomatitis virus glycoprotein. *J.Cell Biol.* **97**, 1381-1388.

Fodor, W.L., Williams, B.L., Matis, L.A., Madri, J.A., Rollins, S.A., Knight, J.W., Velander, W., and Squinto, S.P. (1994). Expression of a functional human complement inhibitor in a transgenic pig as a model for the prevention of xenogeneic hyperacute organ rejection. *Proc.Natl.Acad.Sci.USA* **91**, 11153-11157.

Fredericksen, B.L. and Whitt, M.A. (1995). Vesicular stomatitis virus glycoprotein mutations that affect membrane fusion activity and abolish virus infectivity. *J.Virol.* **69**, 1435-1443.

Fredericksen, B.L. and Whitt, M.A. (1996). Mutations at two conserved acidic amino acids in the glycoprotein of vesicular stomatitis virus affect pH-dependent conformational changes and reduce the pH threshold for membrane fusion. *Virology* **217**, 49-57.

Freed, E.O. and Martin, M.A. (1995). Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix. *J.Virol.* **69**, 1984-1989.

Freed, E.O. and Martin, M.A. (1996). Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. *J.Virol.* **70**, 341-351.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., Sharp, P.M., and Hahn, B.H. (1999). Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**, 436-441.

Gao, F., Yue, L., White, A.T., Pappas, P.G., Barchue, J., Hanson, A.P., Greene, B.M., Sharp, P.M., Shaw, G.M., and Hahn, B.H. (1992). Human infection by genetically diverse SIV_{SM}-related HIV-2 in west Africa. *Nature* **358**, 495-499.

Gaudin, Y. (1997). Folding of rabies virus glycoprotein: Epitope acquisition and interaction with endoplasmic reticulum chaperones. *J.Virol.* **71**, 3742-3750.

Gaudin, Y., Raux, H., Flamand, A., and Ruigrok, R.W.H. (1996). Identification of amino acids controlling the low-pH-induced conformational change of rabies virus glycoprotein. *J.Virol.* **70**, 7371-7378.

Gaudin, Y., Ruigrok, R.W.H., Knossow, M., and Flamand, A. (1993). Low-pH conformational changes of rabies virus glycoprotein and their role in membrane fusion. *J.Virol.* **67**, 1365-1372.

Gaudin, Y., Ruigrok, R.W.H., Tuffereau, C., Knossow, M., and Flamand, A. (1992). Rabies virus glycoprotein is a trimer. *Virology* **187**, 627-632.

Gaudin, Y., Tuffereau, C., Segretain, D., Knossow, M., and Flamand, A. (1991a). Reversible conformational changes and fusion activity of rabies virus glycoprotein. *J.Virol.* **65**, 4853-4859.

Gaudin, Y., Tuffereau, C., Benmansour, A., and Flamand, A. (1991b). Fatty acylation of rabies virus proteins. *Virology* **184**, 441-444.

Gaudin, Y., Tuffereau, C., Durrer, P., Flamand, A., and Ruigrok, R.W.H. (1995). Biological function of the low-pH, fusion-inactive conformation of rabies virus glycoprotein (G): G is transported in a fusion-inactive state-like conformation. *J.Virol.* **69**, 5528-5534.

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**, 1501-1507.

Gluzman, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182.

Good, A.H., Cooper, D.K., Malcolm, A.J., Ippolito, R.M., Koren, E., Neethling, F.A., Ye, Y., Zuhdi, N., and Lamontagne, L.R. (1992). Identification of carbohydrate structures that bind human antiporcine antibodies: implications for discordant xenografting in humans. *Transplant.Proc.* **24**, 559-562.

Gorman, C.M., Howard, B.H., and Reeves, R. (1983). Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nuc.Acids Res.* **11**, 7631-7648.

Graham, F.L., Harrison, T., and Williams, J. (1978). Defective transforming capacity of adenovirus type 5 host-range mutants. *Virology* **86**, 10-21.

Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J.gen.Virol.* **36**, 59-74.

Graham, F.L. and van der Eb, A.J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456-467.

Granowitz, C., Berkowitz, R.D., and Goff, S.P. (1996). Mutations affecting the cytoplasmic domain of the Moloney murine leukemia virus envelope protein: rapid reversion during replication. *Virus Res.* **41**, 25-42.

Granowitz, C., Colicelli, J., and Goff, S.P. (1991). Analysis of mutations in the envelope gene of Moloney murine leukemia virus: Separation of infectivity from superinfection resistance. *Virology* **183**, 545-554.

Gray, K.D. and Roth, M.J. (1993). Mutational analysis of the envelope gene of Moloney murine leukemia virus. *J.Virol.* **67**, 3489-3496.

Green, N., Shinnick, T.M., Witte, O.N., Ponticelli, A., Sutcliffe, J.G., and Lerner, R.A. (1981). Sequence-specific antibodies show that maturation of Moloney leukemia virus envelope polyprotein involves removal of a COOH-terminal peptide. *Proc.Natl.Acad.Sci.USA* **78**, 6023-6027.

Green, P.L. and Chen, I.S.Y. (1990). Regulation of human T cell leukemia virus expression. *FASEB J.* **4**, 169-175.

Gross, L. (1951). "Spontaneous" leukemia developing in C3H mice following inoculation, in infancy, with Ak-leukemic extracts, or Ak-embryos. *Proc.Soc.Exp.Biol.Med.* **76**, 27-32.

Groth, C.G., Korsgren, O., Tibell, A., Tollemar, J., Moller, E., Bolinder, J., Ostman, J., Reinholt, F.P., Hellerstrom, C., and Andersson, A. (1994). Transplantation of porcine fetal pancreas to diabetic patients. *Lancet* **344**, 1402-1404.

Hajek, K.L. and Freisen, P.D. (1998). Proteolytic processing and assembly of *gag* and *gag-pol* proteins of TED, a baculovirus-associated retrotransposon of the Gypsy family. *J.Virol.* **72**, 8718-8724.

Han, J.-Y., Zhao, Y., Anderson, W.F., and Cannon, P.M. (1998). Role of variable regions A and B in receptor binding domain of amphotropic murine leukemia virus envelope protein. *J.Virol.* **72**, 9101-9108.

Hanham, C.A., Zhao, F., and Tignor, G.H. (1993). Evidence from the anti-idiotypic network that the acetylcholine receptor is a rabies virus receptor. *J.Virol.* **67**, 530-542.

Hantzopoulos, P.A., Sullenger, B.A., Ungers, G., and Gilboa, E. (1989). Improved gene expression upon transfer of the adenosine deaminase minigene outside the transcriptional unit of a retroviral vector. *Proc.Natl.Acad.Sci.USA* **86**, 3519-3523.

Harris, J. and Sikora, K. (1993). Human genetic therapy. *Molec.Aspects Med.* **14**, 451-546.

Harris, J.R. (1991). The evolution of placental mammals. *FEBS Letters* **295**, 3-4.

Harrison, T., Graham, F., and Williams, J. (1977). Host-range mutants of adenovirus type 5 defective for growth in HeLa cells. *Virology* **77**, 319-329.

Hartley, J.W., Rowe, W.P., and Huebner, R.J. (1970). Host-range restrictions of murine leukemia viruses in mouse embryo cell cultures. *J.Virol.* **5**, 221-225.

Haseltine, W.A. (1991). Molecular biology of the human immunodeficiency virus type 1. *FASEB J.* **5**, 2349-2360.

Hatziioannou, T., Valsesia-Wittman, S., Russell, S.J., and Cosset, F.-L. (1998). Incorporation of fowl plague virus hemagglutinin into murine leukemia virus particles and analysis of the infectivity of the pseudotyped retroviruses. *J.Virol.* **72**, 5313-5317.

Haywood, A.M. (1994). Virus receptors: Binding, adhesion strengthening, and changes in viral structure. *J.Virol.* **68**, 1-5.

Heneine, W., Tibell, A., Switzer, W.M., Sandstrom, P., Rosales, G.V., Mathews, A., Korsgren, O., Chapman, L.E., Folks, T.M., and Groth, C.G. (1998). No evidence of infection with porcine endogenous retrovirus in recipients of porcine islet-cell xenografts. *Lancet* **352**, 695-699.

Hernandez, L.D., Peters, R.J., Delos, S.E., Young, J.A.T., and Agard, D.A. (1997). Activiation of a retroviral membrane fusion protein: Soluble receptor-induced liposome binding of the ALSV envelope glycoprotein. *J.Cell Biol.* **139**, 1455-1464.

Herniou, E., Martin, J., Miller, K., Cook, J., Wilkinson, M., and Tristem, M. (1998). Retroviral diversity and distribution in vertebrates. *J.Virol.* **72**, 5955-5966.

Hirsch, M.S., Black, P.H., Tracy, G.S., Leibowitz, S., and Schwartz, R.S. (1970). Activation of leukemia viruses by graft-versus-host and mixed lymphocyte reactions *in vitro*. *Proc.Natl.Acad.Sci.USA* **69**, 1069-1072.

Hock, R.A. and Miller, A.D. (1986). Retrovirus-mediated transfer and expression of drug resistance genes in human haematopoietic progenitor cells. *Nature* **320**, 275-277.

Hoggan, M.D., Buckler, C.E., Sears, J.F., Rowe, W.P., and Martin, M.A. (1983). Organization and stability of endogenous xenotropic murine leukemia virus proviral DNA in mouse genomes. *J.Virol.* **45**, 473-477.

Hogge, D.E. and Humphries, R.K. (1987). Gene transfer to primary normal and malignant human hemopoietic progenitors using recombinant retroviruses. *Blood* **69**, 611-617.

Hoxie, J.A., Haggarty, B.S., Bonser, S.E., Rackowski, J.L., Shan, H., and Kanki, P.J. (1988). Biological characterization of a simian immunodeficiency virus-like retrovirus (HTLV-IV): Evidence for CD4-associated molecules required for infection. *J.Virol.* **62**, 2557-2568.

Hu, W.-S. and Temin, H.M. (1990). Retroviral recombination and reverse transcription. *Science* **250**, 1227-1233.

Hunter, E. (1997). Viral entry and receptors. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 71-119.

Ikeda, H. and Sugimura, H. (1989). *FV-4* resistance gene: A truncated endogenous murine leukemia virus with ecotropic interference properties. *J.Virol.* **63**, 5405-5412.

Isacson, O. and Breakefield, X.O. (1997). Benefits and risks of hosting animal cells in the human brain. *Nature Med.* **3**, 964-969.

Jacks, T., Madhani, H.D., Masiarz, F.R., and Varmus, H.E. (1988a). Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. *Cell* **55**, 447-458.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., and Varmus, H.E. (1988b). Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. *Nature* **331**, 280-283.

Jainchill, J.L., Aaronson, S.A., and Todaro, G.J. (1969). Murine sarcoma and leukemia viruses: assay using clonal lines of contact-inhibited mouse cells. *J.Virol.* **4**, 549-553.

Januszeski, M.M., Cannon, P.M., Chen, D., Rozenberg, Y., and Anderson, W.F. (1997). Functional analysis of the cytoplasmic tail of Moloney murine leukemia virus envelope protein. *J.Virol.* **71**, 3613-3619.

Jarett, O., Laird, H.M., and Hay, D. (1973). Determinants of the host range of feline leukaemia viruses. *J.gen.Virol.* **20**, 169-175.

Jenkins, N.A. and Copeland, N.G. (1985). High frequency germline acquisition of ecotropic MuLV proviruses in SWR/J-RF/J hybrid mice. *Cell* **43**, 811-819.

Jenkins, N.A., Copeland, N.G., Taylor, B.A., and Lee, B.K. (1982). Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus*. *J.Virol.* **43**, 26-36.

Johann, S.V., van Zeijl, M., Cekleniak, J., and O'Hara, B. (1993). Definition of a domain of GLVR1 which is necessary for infection by gibbon ape leukemia virus and which is highly polymorphic between species. *J.Virol.* **67**, 6733-6736.

Johnson, P.F. and McKnight, S.L. (1989). Eukaryotic transcriptional regulatory proteins. *Annu.Rev.Biochem.* **58**, 799-839.

Jones, J.S., Allan, R.W., and Temin, H.M. (1994). One retroviral strand is sufficient for synthesis of viral DNA. *J.Virol.* **68**, 207-216.

Jones, J.S. and Risser, R. (1993). Cell fusion induced by the murine leukemia virus envelope glycoprotein. *J.Virol.* **67**, 67-74.

Joshi, S.B., Dutch, R.E., and Lamb, R.A. (1998). A core trimer of the paramyxovirus fusion protein: parallels to influenza virus hemagglutinin and HIV-1 gp41. *Virology* **248**, 20-34.

Kalter, S.S. and Heberling, R.L. (1976). Virology. *Handbook of laboratory animal science Vol 3* (Melby, E.C.J. and Altman, N.H., Eds.) CRC Press, Cleveland. 655

Kaptein, L.C.M., Greijer, A.E., Valerio, D., and van Beusechem, V.W. (1997). Optimized conditions for the production of recombinant amphotropic retrovirus preparations. *Gene Ther.* **4**, 172-176.

Kasahara, N., Dozy, A.M., and Kan, Y.W. (1994). Tissue-specific targeting of retroviral vectors through ligand-receptor interactions. *Science* **266**, 1373-1376.

Kavanaugh, M.P., Miller, D.G., Zhang, W., Law, W., Kozak, S.L., Kabat, D., and Miller, A.D. (1994). Cell-surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium-dependent phosphate symporters. *Proc.Natl.Acad.Sci.USA* **91**, 7071-7075.

Kawai, S. and Hanafusa, H. (1973). Isolation of defective mutant of avian sarcoma virus. *Proc.Natl.Acad.Sci.USA* **70**, 3493-3497.

Kawakami, T.G., Huff, S.D., Buckley, P.M., Dungworth, D.L., Snyder, S.P., and Gilden, R.V. (1972). C-type virus associated with gibbon lymphosarcoma. *Nature New Biol.* **235**, 170-171.

Kawakami, T.G., Kollias, G., and Holmberg, C. (1980). Oncogenicity of gibbon type C myelogenous leukemia virus. *Int.J.Cancer* **25**, 841-846.

Kayman, S.C., Kopelman, R., Projan, S., Kinney, D.M., and Pinter, A. (1991). Mutational analysis of N-linked glycosylation sites of Friend murine leukemia virus envelope protein. *J.Virol.* **65**, 5323-5332.

Kim, A.I., Terzian, C., Santamaria, P., Pélisson, A., Prud'homme, N., and Bucheton, A. (1994). Retroviruses in invertebrates: The *gypsy* retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc.Natl.Acad.Sci.USA* **91**, 1285-1289.

Kim, J.W., Closs, E.I., Albritton, L.M., and Cunningham, J.M. (1991). Transport of cationic amino acids by the mouse ecotropic retrovirus receoptor. *Nature* **352**, 725-728.

King, S.R., Horowitz, J.M., and Risser, R. (1987). Nucleotide conservation of endogenous ecotropic murine leukemia proviruses in inbred mice: Implications for viral origin and dispersal. *Virology* **157**, 543-547.

Klatzman, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C., and Montagnier, L. (1984). T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**, 767-768.

Koch, W., Hunsmann, G., and Friedrich, R.W. (1983). Nucleotide sequence of the envelope gene of Friend murine leukemia virus. *J.Virol.* **45**, 1-9.

Koike, C., Kannagi, R., Takuma, Y., *et al.* (1996). Introduction of α(1-2)-fucosyltransferase and its effect on αgal epitopes in transgenic pig. *Xenotransplantation* **3**, 81-86.

Kontani, H., Newton, P.B.3., Zhang, S., Chiang, Y.L., Otto, E., Weaver, L., Blaese, R.M., Anderson, W.F., and McGainty, G.J. (1994). Improved methods of retroviral vector transduction and production for gene therapy. *Hum.Gene Ther.* **5**, 19-28.

Koo, H.-M., Brown, A.M.C., Kaufman, R.J., Prorock, C.M., Ron, Y., and Dougherty, J.P. (1992). A spleen necrosis virus-based retroviral vector expresses two genes from a dicistronic mRNA. *Virology* **186**, 669-675.

Kozak, C.A. (1985). Susceptibility of wild mouse cells to exogenous infection with xenotropic leukemia viruses: Control by a single dominant locus on chromosome 1. *J.Virol.* **55**, 690-695.

Krichbaum-Stenger, K., Poiesz, B.J., Keller, P., Ehrlich, G., and Gavalchin, J. (1984). Specific adsoption of HTLV-I to various target human and animal cells. *Blood* **70**, 1303-1311.

Kucera, P., Dolivo, M., Coulon, P., and Flamand, A. (1985). Pathways of the early propagation of virulent and avirulent rabies strains from the eye to the brain. *J.Virol.* **55**, 158-162.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lafay, F. (1974). Envelope proteins of vesicular stomatitis virus: Effects of temperature-sensitive mutants in complementation groups III and V. *J.Virol.* **14**, 1220-1228.

Lafon, M., Wiktor, T.J., and Macfarlan, R.I. (1983). Antigenic sites on the CVS rabies virus glycoprotein: Analysis with monoclonal antibodies. *J.gen.Virol.* **64**, 843-851.

Lagging, L.M., Meyer, K., Owens, R.J., and Ray, R. (1998). Functional role of hepatitis C virus chimeric glycoproteins in the infectivity of pseudotyped virus. *J.Virol.* **72**, 3539-3546.

Landau, N.R. and Littman, D.R. (1992). Packaging system for rapid production of murine leukemia virus vectors with variable tropism. *J.Virol.* **66**, 5110-5113.

Lavillette, D., Maurice, M., Roche, C., Russell, S.J., Sitbon, M., and Cosset, F.-L. (1998). A proline-rich motif downstream of the receptor binding domain modulates conformation and fusogenicity of murine retroviral envelopes. *J.Virol.* **72**, 9955-9965.

Le Tissier, P., Stoye, J.P., Takeuchi, Y., Patience, C., and Weiss, R.A. (1997). Two sets of human-tropic pig retrovirus. *Nature* **389**, 681-682.

Lee, Y.-M., Tang, X.-B., Cimakasky, L.M., Hildreth, J.E.K., and Yu, X.-F. (1997). Mutations in the matrix protein of human immunodeficiency virus type 1 inhibit surface expression and virion incorporation of viral envelope glycoproteins in $CD4^+$ T lymphocytes. *J.Virol.* **71**, 1443-1452.

Leider, J.M., Palese, P., and Smith, F.I. (1988). Determination of the mutation rate of a retrovirus. *J.Virol.* **62**, 3084-3091.

Lentz, T.L., Burrage, T.G., Smith, A.L., Crick, J., and Tignor, G.H. (1982). Is the acetylcholine receptor a rabies virus receptor? *Science* **215**, 182-184.

Lentz, T.L., Wilson, P.T., and Speicher, D.M. (1984). Amino acid sequence similarity between rabies virus glycoprotein and snake venom curaremimetic neurotoxins. *Science* **226**, 847-848.

Levy, J.A. (1986). The multifaceted retrovirus. *Cancer Res.* **46**, 5457-5468.

Levy, J.A., Datta, S.K., and Schwartz, R.S. (1977). Recovery of xenotropic virus but not ecotropic virus during graft-versus-host reaction in mice. *Clin.Immunol.Immunopathol.* **7**, 262-268.

Lewis, P., Fu, Y., and Lentz, T.L. (1998). Rabies virus entry into endosomes in IMR-32 human neuroblastoma cells. *Exp.Neurol.* **153**, 65-73.

Li, Y., Drone, C., Sat, E., and Ghosh, H.P. (1993). Mutational analysis of the vesicular stomatitis virus glycoprotein G for membrane fusion domains. *J.Virol.* **67**, 4070-4077.

Lieber, M.M., Sherr, C.J., Benveniste, R.E., and Todaro, G.J. (1975a). Biologic and immunologic properties of porcine type C viruses. *Virology* **66**, 616-619.

Lieber, M.M., Sherr, C.J., Todaro, G.J., Benveniste, R.E., Callahan, R., and Coon, H.G. (1975b). Isolation from the Asian mouse *Mus caroli* of an endogenous type C virus related to infectious primate type C viruses. *Proc.Natl.Acad.Sci.USA* **72**, 2315-2319.

Lindemann, D., Bock, M., Schweizer, M., and Rethwilm, A. (1997). Efficient pseudotyping of murine leukemia virus particles with chimeric human foamy virus envelope proteins. *J.Virol.* **71**, 4815-4820.

Linder, M., Linder, D., Hahnen, J., Schott, H.-H., and Strim, S. (1992). Localization of the intrachain disulfide bonds of the envelope glycoprotein 71 from Friend murine leukemia virus. *Eur.J.Biochem.* **203**, 65-73.

Linder, M., Wenzel, V., Linder, D., and Stirm, S. (1994). Structural elements in glycoprotein 70 from polytropic Friend mink cell focus-inducing virus and glycoprotein 71 from ecotropic Friend murine leukemia virus, as defined by disulfide-bonding pattern and limited proteolysis. *J.Virol.* **68**, 5133-5141.

Littlefield, J.W. and Basilico, C. (1966). Infection of thymidine kinase-deficient BHK cells with polyoma virus. *Nature* **211**, 250-252.

Lock, L.F., Keshet, E., Gilbert, D.J., Jenkins, N.A., and Copeland, N.G. (1988). Studies on the mechanism of spontaneous germline ecotropic provirus acquisition in mice. *EMBO J.* **7**, 4169-4177.

Lodge, R., Delamarre, L., Lalonde, J.-P., Alvarado, J., Sanders, D.A., Dokhélar, M.-C., Cohen, É.A., and Lemay, G. (1997a). Two distinct oncornaviruses harbor an intracytoplasmic tyrosine-based basolateral targeting signal in their viral envelope glycoprotein. *J.Virol.* **71**, 5696-5702.

Lodge, R., Göttlinger, H., Gabuzda, D., Cohen, É.A., and Lemay, G. (1994). The intracytoplasmic domain of gp41 mediates polarized budding of human immunodeficiency virus type 1 in MDCK cells. *J.Virol.* **68**, 4857-4861.

Lodge, R., Lalonde, J.P., Lemay, G., and Cohen, É.A. (1997b). The membrane-proximal intracytoplasmic tyrosine residue of HIV-1 envelope glycoprotein is critical for basolateral targeting of viral budding in MDCK cells. *EMBO J.* **16**, 695-705.

Luciw, P.A. and Leung, N.J. (1993). Mechanisms of retrovirus replication. *The retroviridae Vol 1* (Levy, J.A., Ed.) Plenum Press, New York. 159-298.

Lupton, S.D., Brunton, L.L., Kalberg, V.A., and Overell, R.W. (1991). Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. *Mol.Cell Biol.* **11**, 3374-3378.

Lyles, D.S., McKenzie, M., and Parce, J.W. (1992). Subunit interactions of vesicular stomatitis virus envelope glycoprotein stabilized by binding to viral matrix protein. *J.Virol.* **66**, 349-358.

Maddon, P.J., Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L., and Axel, R. (1985). The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. *Cell* **42**, 93-104.

Majors, J. (1990). The structure and function of retroviral long terminal repeats. *Curr.Top.Microbiol.Immunol.* **157**, 49-52.

Mammano, F., Kondo, E., Sodroski, J., Bukovsky, A., and Göttlinger, H.G. (1995). Rescue of human immunodeficiency virus type 1 matrix protein mutants by envelope glycoproteins with short cytoplasmic domains. *J.Virol.* **69**, 3824-3830.

Mammano, F., Salvatori, F., Indraccolo, S., De Rossi, A., Chieco-Bianchi, L., and Göttlinger, H.G. (1997). Truncation of the human immunodeficiency virus type 1 envelope glycoprotein allows efficient pseudotyping of Moloney murine leukemia virus particles and gene transfer into $CD4^+$ cells. *J.Virol.* **71**, 3341-3345.

Mann, R., Mulligan, R.C., and Baltimore, D. (1983). Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**, 153-159.

Marin, M., Noel, D., Valsesia-Wittman, S., Brockly, F., Etienne-Julan, M., Russell, S., Cosset, F.-L., and Piechaczyk, M. (1996). Targeted infection of human cells via major histocompatibility complex class I molecules by Moloney murine leukemia virus-derived viruses displaying single-chain antibody fragment-envelope fusion proteins. *J.Virol.* **70**, 2957-2962.

Markowitz, D., Goff, S., and Bank, A. (1988). A safe packaging line for gene transfer: Separating viral genes on two different plasmids. *J.Virol.* **62**, 1120-1124.

Marsh, M. and Pelchen-Matthews, A. (1994). The endocytic pathway and virus entry. *Cellular receptors for animal viruses* (Wimmer, E., Ed.) Cold Spring Harbor Laboratory Press, New York. 215-240.

Martin, U., Kiessig, V., Blusch, J.H., Haverich, A., von der Helm, K., Herden, T., and Steinhoff, G. (1998). Expression of pig endogenous retrovirus by primary porcine endothelial cells and infection of human cells. *Lancet* 352, 692-694.

Martini, G.A. (1969). Marburg agent disease: in man. *Trans.R.Soc.Trop.Med.Hyg.* **63**, 295-302.

McAllister, R.M., Nicholson, M.O., Gardner, M.B., Rongey, R.W., Rasheid, S., Sarma, P.S., Huebner, R.J., Hatanaka, M., Oroszlan, S., Gilden, R.V., Kabigting, A., and Vernon, L. (1972). C-type virus released from cultured human rhabdomyosarcoma cells. *Nature* **235**, 3-6.

McClure, M.O., Sommerfelt, M.A., Marsh, M., and Weiss, R.A. (1990). The pH independence of mammalian retrovirus infection. *J.gen.Virol.* **71**, 767-773.

McClurkin, A.W. and Norman, J.O. (1966). Studies on transmissible gastroenteritis of swine. II. Selected characteristics of a cytopathogenic virus common to five isolates from transmissible gastroenteritis. *Can.J.Comp.Med.Vet.Sci.* **30**, 190-198.

Mebatsion, T. and Conzelmann, K.-K. (1996). Specific infection of $CD4^+$ target cells by recombinant rabies virus pseudotypes carrying the HIV-1 envelope spike protein. *Proc.Natl.Acad.Sci.USA* **93**, 11366-11370.

Mebatsion, T., König, M., and Conzelmann, K.-K. (1996). Budding of rabies virus particles in the absence of the spike glycoprotein. *Cell* **84**, 941-951.

Mebatsion, T., Schnell, M.J., and Conzelmann, K.-K. (1995). Mokola virus glycoprotein and chimeric proteins can replace rabies virus glycoprotein in the rescue of infectious defective rabies virus particles. *J.Virol.* **69**, 1444-1451.

Mebatsion, T., Weiland, F., and Conzelmann, K.-K. (1999). Matrix protein of rabies virus is responsible for the assembly and budding of bullet-shaped particles and interacts with the transmembrane spike glycoprotein G. *J.Virol.* **73**, 242-250.

Meisler, M.H. and Ting, C.N. (1993). The remarkable evolutionary history of the human amylase genes. *Crit.Rev.Oral Biol.Med.* **4**, 503-509.

Mellquist, J.L., Kasturi, L., Spitalnik, S.L., and Shakin-Eshleman, S.H. (1998). The amino acid following an Asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochem.* **37**, 6833-6837.

Metsikkö, K. and Simons, K. (1986). The budding mechanism of spikeless vesicular stomatitis virus particles. *EMBO J.* **5**, 1913-1920.

Michaels, M.G. and Simmons, R.L. (1994). Xenotransplant-associated zoonoses: Strategies for prevention. *Transplantation* **57**, 1-7.

Mifune, K., Ohuchi, M., and Mannen, K. (1982). Hemolysis and cell fusion by rhabdoviruses. *FEBS Letters* **137**, 293-297.

Miller, A.D. (1997). Development and applications of retroviral vectors. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 437-473.

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**, 2895-2902.

Miller, A.D. and Chen, F. (1996). Retovirus packaging cells based on 10A1 murine leukemia virus for production of vectors that use multiple receptors for cell entry. *J.Virol.* **70**, 5564-5571.

Miller, A.D., Garcia, J.V., von Suhr, N., Lynch, C.M., Wilson, C., and Eiden, M.V. (1991). Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J.Virol.* **65**, 2220-2224.

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**, 175-183.

Miller, D.G., Edwards, R.H., and Miller, A.D. (1994). Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *Proc.Natl.Acad.Sci.USA* **91**, 78-82.

Miller, D.G. and Miller, A.D. (1994). A family of retroviruses that utilize related phosphate transporters for cell entry. *J.Virol.* **68**, 8270-8276.

Mitchell, P.M. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378.

Moennig, V., Frank, H., Hunsmann, G., Ohms, P., Schwarz, H., Schäfer, W., and Strandstrom, H. (1974). C-type particles produced by a permanent cell line from a leukemic pig. *Virology* **57**, 179-188.

Moldow, C.F., Reynolds, F.H., Lake, J., Lundberg, K., and Stephenson, J.R. (1979). Avian sarcoma virus envelope glycoprotein (gp85) specifically binds chick embryo fibroblasts. *Virology* **97**, 448-453.

Monk, R.J., Malik, F.G., Stokesberry, D., and Evans, L.H. (1992). Direct determination of the point mutation rate of a murine retrovirus. *J.Virol.* **66**, 3683-3689.

Montefiori, D.C., Cornell, R.J., Zhou, J.Y., Zhou, J.T., Hirsch, V.M., and Johnson, P.R. (1994). Complement control proteins, CD46, CD55, and CD59, as common surface constituents of human and simian immunodeficiency viruses and possible targets for vaccine protection. *Virology* **205**, 82-92.

Morgan, R.A., Nussbaum, O., Muenchau, D.D., Shu, L., Couture, L., and Anderson, W.F. (1993). Analysis of the functional and host range-determining regions of the murine ecotropic and amphotropic retovirus envelope proteins. *J.Virol.* **67**, 4712-4721.

Morimoto, K., Ni, Y.-J., and Kawai, A. (1992). Syncytium formation is induced in the murine neuroblastoma cell cultures which produce pathogenic type G proteins of the rabies virus. *Virology* **189**, 203-216.

Munoz-Barroso, I., Durell, S., Sakaguchi, K., Appella, E., and Blumenthal, R. (1998). Dilation of the human immunodeficiency virus-1 envelope glycoprotein fusion pore revealed by the inhibitory action of a synthetic peptide from gp41. *J.Cell Biol.* **140**, 315-323.

Murphy, F.A. and Harrison, A.K. (1980). Electron microscopy of the rhabdoviruses of animals. *Rhabdoviruses Vol 1* (Bishop, D.H.L., Ed.) CRC Press Inc., Florida. 65-106.

Myers, G. (1997). Retroviral sequences. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 709-755.

Neda, H., Wu, C.H., and Wu, G.Y. (1991). Chemical modification of an ecotropic murine leukemia virus results in redirection of its target cell specificity. *J.Biol.Chem.* **266**, 14143-14146.

Neil, J.C., Fulton, R., Rigby, M., and Stewart, M. (1991). Feline leukaemia virus: Generation of pathogenic and oncogenic variants. *Curr.Top.Microbiol.Immunol.* **171**, 67-93.

Nilson, B.H.K., Morling, F.J., Cosset, F.-L., and Russell, S.J. (1996). Targeting of retroviral vectors through protease-substrate interactions. *Gene Ther.* **3**, 280-286.

Notter, M.F., Leary, J.F., and Balduzzi, P.C. (1982). Adsorption of Rous sarcoma virus to genetically susceptible and resistant chicken cells studied by laser flow cytometry. *J.Virol.* **41**, 958-964.

Nowak, R. (1994). Xenotransplants set to resume. *Science* **266**, 1148-1151.

O'Hara, B., Johann, S.V., Klinger, H.P., Blair, D.G., Rubinson, H., Dunn, K.J., Sass, P., Vitek, S.M., and Robins, T. (1990). Characterization of a human gene conferring sensitivity to infection by gibbon ape leukemia virus. *Cell Growth Differ.* **1**, 119-127.

O'Neill, R.R., Khan, A.S., Hoggan, M.D., Hartley, J.W., Martin, M.A., and Repaske, R. (1986). Specific hybridization probes demonstrate fewer xenotropic than mink cell focus-forming murine leukemia virus *env*-related sequences in DNAs from inbred laboratory mice. *J.Virol.* **58**, 359-366.

Olah, Z., Lehel, C., Anderson, W.B., Eiden, M.V., and Wilson, C.A. (1994). The cellular receptor for gibbon ape leukemia virus is a novel high affinity sodium-dependent phosphate transporter. *J.Biol.Chem.* **269**, 25426-25431.

Onions, D., Cooper, D.K.C., Alexander, T.J.L., Brown, C., Claasen, E., Foweraker, J., Harris, D.L., Mahy, B., Osterhaus, A.D.M.E., Pastoret, P.-P., Yamanouchi, K., and Minor, P. (1999). An assessment of the risk of xenozoonotic disease in pig-to-human xenotransplantation. in preparation.

Onions, D., Hart, D., Mahoney, C., Galbraith, D., and Smith, K. (1998). Endogenous retroviruses and the safety of porcine xenotransplantation. *Trends Microbiol.* **6**, 430-431.

Oprandy, J.J. and Chang, P.W. (1983). 5-bromodeoxyuridine induction of hematopoietic neoplasia and retrovirus activation in the soft-shell clam, *Mya arenaria*. *J.Invertebr.Pathol.* **42**, 196-206.

Ott, D., Friedrich, R., and Rein, A. (1990). Sequence analysis of amphotropic and 10A1 murine leukemia viruses: Close relationship to mink cell focus-inducing viruses. *J.Virol.* **64**, 757-766.

Overell, R.W., Weisser, K.E., and Cosman, D. (1988). Stably transmitted triple-promoter retroviral vectors and their use in transformation of primary mammalian cells. *Mol.Cell Biol.* **8** , 1803-1808.

Owens, R.J., Dubay, J.W., Hunter, E., and Compans, R.W. (1991). Human immunodeficiency virus envelope protein determines the site of virus release in polarized epithelial cells. *Proc.Natl.Acad.Sci.USA* **88**, 3987-3991.

Owens, R.J. and Rose, J.K. (1993). Cytoplasmic domain requirement for incorporation of a foreign envelope protein into vesicular stomatitis virus. *J.Virol.* **67**, 360-365.

Panganiban, A.T. and Fiore, D. (1988). Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* **241**, 1064-1069.

Paradis, K., Langford, G., Long, Z., Heneine, W., Sandstrom, P., Switzer, W.M., Chapman, L.E., Lockey, C., Onions, D., The XEN 111 Study Group, and Otto, E. (1999). Search for cross-species transmission of porcine endogenous retrovirus in patients treated with living pig tissue. *Science* **285**, 1236-1241.

Patience, C., Patton, G.S., Takeuchi, Y., Weiss, R.A., McClure, M.O., Rydberg, L., and Breimer, M.E. (1998). No evidence of pig DNA or retroviral infection in patients with short-term extracorporeal connection to pig kidneys. *Lancet* **352**, 699-701.

Patience, C., Takeuchi, Y., and Weiss, R.A. (1997a). Infection of human cells by an endogenous retrovirus of pigs. *Nature Med.* **3**, 282-286.

Patience, C., Wilkinson, D.A., and Weiss, R.A. (1997b). Our retroviral heritage. *Trends Genet.* **13**, 116-120.

Pattnaik, A.K. and Wertz, G.W. (1991). Cells that express all five proteins of vesicular stomatitis virus from cloned cDNAs support replication, assembly, and budding of defective interfering particles. *Proc.Natl.Acad.Sci.USA* **88**, 1379-1383.

Pepinsky, R.B. and Vogt, V.M. (1979). Identification of retrovirus matrix proteins by lipid-protein cross-linking. *J.Mol.Biol.* **131**, 819-839.

Perez, L.G., Davis, G.L., and Hunter, E. (1987). Mutants of the Rous sarcoma virus envelope glycoprotein that lack the transmembrane anchor and cytoplasmic domains: Analysis of intracellular transport and assembly into virions. *J.Virol.* **61**, 2981-2988.

Pincus, T., Hartley, J.W., and Rowe, W.P. (1975). A major genetic locus affecting resistance to infection with murine leukemia viruses. IV. Dose-response relationships in *Fv-1* sensitive and resistant cell cultures. *Virology* **65**, 333-342.

Pino-Chavez, G., Cary, N., Young, V., Cozzi, E., Horsley, J., Langford, G., and White, D. (1996). The effect of h-DAF expression on endothelial cell activation of porcine xenografts. *Transplant.Proc.* **28**, 590-591.

Porter, C.D., Lukacs, K.V., Box, G., Takeuchi, Y., and Collins, M.K.L. (1998). Cationic liposomes enhance the rate of transduction by a recombinant retroviral vector in vitro and in vivo. *J.Virol.* **72**, 4832-4840.

Prehaud, C., Coulon, P., Lafay, F., Thiers, C., and Flamand, A. (1988). Antigenic site II of the rabies virus glycoprotein: Structure and role in viral virulence. *J.Virol.* **62**, 1-7.

Pulvertaft, R.J.V. (1964). Cytology of Burkitt's tumour (African lymphoma). *Lancet* **1**, 238-240.

Q-One Biotech Ltd. and Immutran Ltd. (1997). International Patent Application WO97/40167

Rabson, A.B. and Graves, B.J. (1997). Synthesis and processing of viral RNA. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratroy Press, New York. 205-261.

Rademacher, T.W., Parekh, R.B., and Dwek, R.A. (1988). Glycobiology. *Annu.Rev.Biochem.* **57**, 785-838.

Radsak, K., Fuhrmann, R., Franke, R.P., Schneider, D., Kollert, A., Brucher, K.H., and Drenckhahn, D. (1989). Induction by sodium butyrate of cytomegalovirus replication in human endothelial cells. *Arch.Virol.* **107**, 151-158.

Ragheb, J.A. and Anderson, W.F. (1994). Uncoupled expression of Moloney murine leukemia virus envelope polypeptides SU and TM: a functional analysis of the role of TM domains in viral entry. *J.Virol.* **68**, 3207-3219.

Ramsdale, E.E., Kingsman, S.M., and Kingsman, A.J. (1996). The "putative" leucine zipper region of murine leukemia virus transmembrane protein (P15e) is essential for viral infectivity. *Virology* **220**, 100-108.

Rasheed, S. and Gardner, M.B. (1980). Characterisation of cat cell cultures for expression of retrovirus, FOCMA and endogenous sarc genes. *Proceedings of the Third International Feline Leukaemia Virus Meeting* (Essex, M. and McClelland, A.J., Eds.) Elsevier, North Holland. 393-400.

Reagan, K.J. and Wunner, W.H. (1985). Rabies virus interaction with various cell lines is independent of the acetylcholine receptor. *Arch.Virol.* **84**, 277-282.

Rein, A., Mirro, J., Haynes, J.G., Ernst, S.M., and Nagashima, K. (1994). Function of the cytoplasmic domain of a retroviral transmembrane protein: p15E-p2E cleavage activates the membrane fusion capability of the murine leukemia virus Env protein. *J.Virol.* **68**, 1773-1781.

Rein, A. and Schultz, A. (1984). Different recombinant murine leukemia viruses use different cell surface receptors. *Virology* **136**, 144-152.

Rimsky, L.T., Shugars, D.C., and Matthews, T.J. (1998). Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. *J.Virol.* **72**, 986-993.

Rolls, M.M., Webster, P., Balba, N.H., and Rose, J.K. (1994). Novel infectious particles generated by expression of the vesicular stomatitis virus glycoprotein from a self-replicating RNA. *Cell* **79**, 497-506.

Roman, L.M., Scharm, A., and Howell, K.E. (1988). Immunoselection of stably transformed MDCK cells expressing the vesicular stomatitis virus G-protein at the basolateral surface. *Exp.Cell Res.* **175**, 376-387.

Rong, L., Edinger, A., and Bates, P. (1997). Role of basic residues in the subgroup-determining region of the subgroup A avian sarcoma and leukosis virus envelope in receptor binding and infection. *J.Virol.* **71**, 3458-3465.

Rose, J.K. and Bergmann, J.E. (1983). Altered cytoplasmic domains affect intracellular transport of the vesicular stomatitis virus glycoprotein. *Cell* **34**, 513-524.

Rose, J.K., Doolittle, R.F., Anilionis, A., Curtis, P.J., and Wunner, W.H. (1982). Homology between the glycoproteins of vesicular stomatitis virus and rabies virus. *J.Virol.* **43**, 361-364.

Rosenberg, N. and Jolicoeur, P. (1997). Retroviral pathogenesis. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 475-585.

Rosengard, A.M., Gary, N.R.B., Langford, G.A., Tucker, A.W., Wallwork, J., and White, D.J.G. (1995). Tissue expression of human complement inhibitor, decay accelerating factor, in transgenic pigs. *Transplantation* **59**, 1325-1333.

Roth, M.G., Srinivas, R.V., and Compans, R.W. (1983). Basolateral maturation of retroviruses in polarized epithelial cells. *J.Virol.* **45**, 1065-1073.

Rous, P. (1911). Transmission of a malignant new growth by means of a cell-free filtrate. *J.Am.Med.Assoc.* **56**, 198-198.

Roux, P., Jeanteur, P., and Piechaczyk, M. (1989). A versatile and potentially general approach to the targeting of specific cell types by retroviruses: Application to the infection of human cells by means of major histocompatibility complex class I and class II antigens by mouse ecotropic murine leukemia virus-derived viruses. *Proc.Natl.Acad.Sci.USA* **86**, 9079-9083.

Rowe, W.P. and Kozak, C.A. (1980). Germ-line insertions of AKR murine leukemia virus genomes in AKV-1 congenic mice. *Proc.Natl.Acad.Sci.USA* **77**, 4871-4874.

Russell, S.J., Hawkins, R.E., and Winter, G. (1993). Retroviral vectors displaying functional antibody fragments. *Nuc.Acids Res.* **21**, 1081-1085.

Rustici, M., Bracci, L., Lozzi, L., Neri, P., Santucci, A., Soldani, P., Spreafico, A., and Niccolai, N. (1993). A model of the rabies virus glycoprotein active site. *Biopolymers* **33**, 961-969.

Ruta, M., Murray, M.J., Webb, M.C., and Kabat, D. (1979). A murine leukemia virus mutant with a temperature sensitive defect in membrane glycoprotein synthesis. *Cell* **16**, 77-88.

Sachs, D.H. (1994). The pig as a potential xenograft donor. *Vet.Immunol.Immunopathol.* **43**, 185-191.

Saifuddin, M., Parker, C.J., Peeples, M.E., Gorny, M.K., Zolla-Pazner, S., Ghassemi, M., Rooney, I.A., Atkinson, J.P., and Spear, G.T. (1995). Role of virion-associated glycosylphosphatidylinositol-linked proteins CD55 and CD59 in complement resistance of cell line-derived and primary isolates of HIV-1. *J.Exp.Med.* **182**, 501-509.

Samuelson, L.C., Wiebauer, K., Snow, C.M., and Meisler, M.H. (1990). Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. *Mol.Cell Biol.* **10**, 2513-2520.

Sandrin, M.S., Fodor, W.L., Mouhtouris, E., Osman, N., Cohney, S., Rollins, S.A., Guilmette, E.R., Setter, E., Squinto, S.P., and McKenzie, I.F. (1995). Enzymatic remodelling of the carbohydrate surface of a xenogenic cell substantially reduces human antibody binding and complement-mediated cytolysis. *Nature Med.* **1**, 1261-1267.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc.Natl.Acad.Sci.USA* **74**, 5463-5467.

Sarma, P.S. and Log,T. (1973). Subgroup classification of feline leukemia and sarcoma viruses by viral interference and neutralization tests. *Virology* **54**, 160-169.

Sarma, P.S., Log, T., Jain, D., Hill, P.R., and Huebner, R.J. (1975). Differential host range of viruses of feline leukemia-sarcoma complex. *Virology* **64**, 438-446.

Sattentau, Q.J., Clapham, P.R., Weiss, R.A., Beverley, P.C., and Montagnier, L. (1988). The human and simian immunodeficiency viruses HIV-1, HIV-2 and SIV interact with similar epitopes on their cellular receptor, the CD4 molecule. *AIDS* **2**, 101-105.

Sawadogo, M. and Sentenac, A. (1990). RNA polymerase B (II) and general transcription factors. *Annu.Rev.Biochem.* **59**, 711-754.

Schlegel, R., Tralka, T.S., Willingham, M.C., and Pastan, I. (1983). Inhibition of VSV binding and infectivity by phosphatidylserine: Is phosphatidylserine a VSV-binding site? *Cell* **32** , 639-646.

Schnell, M.J., Buonocore, L., Boritz, E., Ghosh, H.P., Chernish, R., and Rose, J.K. (1998). Requirement for a non-specific glycoprotein cytoplasmic domain sequence to drive efficient budding of vesicular stomatitis virus. *EMBO J.* **17**, 1289-1296.

Schnell, M.J., Buonocore, L., Kretzschmar, E., Johnson, E., and Rose, J.K. (1996). Foreign glycoproteins expressed from recombinant vesicular stomatitis viruses are incorporated efficiently into virus particles. *Proc.Natl.Acad.Sci.USA* **93**, 11359-11365.

Seganti, L., Superti, F., Bianchi, S., Orsi, N., Divizia, M., and Panà, A. (1990). Susceptibility of mammalian, avian, fish, and mosquito cell lines to rabies virus infection. *Acta virol.* **34**, 155-163.

Seif, I., Coulon, P., Rollin, P.E., and Flamand, A. (1985). Rabies virulence: Effect on pathogenicity and sequence characterization of rabies virus mutations affecting antigenic site III of the glycoprotein. *J.Virol.* **53**, 926-934.

Shakin-Eshleman, S.H., Remaley, A.T., Eshleman, J.R., Wunner, W.H., and Spitalnik, S.L. (1992). *N*-linked glycosylation of rabies virus glycoprotein. *J.Biol.Chem.* **267**, 10690-10698.

Shinnick, T.M., Lerner, R.A., and Sutcliffe, J.G. (1981). Nucleotide sequence of Moloney murine leukaemia virus. *Nature* **293**, 543-548.

Shokralla, S., He, Y., Wanas, E., and Ghosh, H.P. (1998). Mutations in a carboxy-terminal region of vesicular stomatitis virus glycoprotein G that affect membrane fusion activity. *Virology* **242**, 39-50.

Simons, K. and Garoff, H. (1980). The budding mechanisms of enveloped animal viruses. *J.gen.Virol.* **50**, 1-21.

Skehel, J.J. and Wiley, D.C. (1998). Coiled coils in both intracellular vesicle and viral membrane fusion. *Cell* **95**, 871-874.

Smith, A.J., Cho, M.-I., Hammarskjöld, M.-L., and Rekosh, D. (1990). Human immunodeficiency virus type 1 $Pr55^{gag}$ and $Pr160^{gag\text{-}pol}$ expressed from a simian virus 40 late replacement vector are efficiently processed and assembled into viruslike particles. *J.Virol.* **64**, 2743-2750.

Smith, D.M. (1993). Endogenous retroviruses in xenografts. *N.Engl.J.Med.* **328**, 142-143.

Smith, S.D., Shatsky, M., Cohen, P.S., Warnke, R., Link, M.P., and Glader, B.E. (1984). Monoclonal antibody and enzymatic profiles of human malignant T-lymphoid cells and derived cell lines. *Cancer Res.* **44**, 5657-5660.

Somia, N.V., Zoppe, M., and Verma, I.M. (1995). Generation of targeted retroviral vectors by using single-chain variable fragment: An approach to in vivo gene delivery. *Proc.Natl.Acad.Sci.USA* **92**, 7570-7574.

Sommerfelt, M.A. and Weiss, R.A. (1990). Receptor interference groups of 20 retroviruses plating on human cells. *Virology* **176**, 58-69.

Soneoka, Y., Cannon, P.M., Ramsdale, E.E., Griffiths, J.C., Romano, G., Kingsman, S.M., and Kingsman, A.J. (1995). A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nuc.Acids Res.* **23**, 628-633.

Song, S.U., Gerasimova, T., Kurkulos, M., Boeke, J.D., and Corcas, V.G. (1994). An Env-like protein encoded by a *Drosophila* retroelement: Evidence that *gypsy* is an infectious retrovirus. *Genes Dev.* **8**, 2046-2057.

Spiegel, M., Bitzer, M., Schenk, A., Rossmann, H., Neubert, W.J., Seidler, U., Gregor, M., and Lauer, U. (1998). Pseudotype formation of Moloney murine leukemia virus with Sendai virus glycoprotein F. *J.Virol.* **72**, 5296-5302.

Starzl, T.E., Fung, J., Tzakis, A., Todo, S., Demetris, A.J., Marino, I.R., Doyle, H., Zeevi, A., Warty, V., and Michaels, M. (1993). Baboon-to-human liver transplantation. *Lancet* **341**, 65-71.

Stewart, M.A., Warnock, M., Wheeler, A., Wilkie, N., Mullins, J.I., Onions, D.E., and Neil, J.C. (1986). Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. *J.Virol.* **58**, 825-834.

Stoker, M. and Macpherson, I. (1964). Syrian hamster fibroblast cell line *BHK*21 and its derivatives. *Nature* **203**, 1355-1357.

Stoye, J. (1998). No clear answers on safety of pigs as tissue donor source. *Lancet* **352**, 666-667.

Stoye, J.P. and Coffin, J.M. (1987). The four classes of endogenous murine leukemia virus: Structural relationships and potential for recombination. *J.Virol.* **61**, 2659-2669.

Stoye, J.P. and Coffin, J.M. (1995). The dangers of xenotransplantation. *Nature Med.* **1**, 1100

Suomalainen, M. and Garoff, H. (1994). Incorporation of homologous and heterologous proteins into the envelope of Moloney murine leukemia virus. *J.Virol.* **68**, 4879-4889.

Superti, F., Derer, M., and Tsiang, H. (1984a). Mechanism of rabies virus entry into CER cells. *J.gen.Virol.* **65**, 781-789.

Superti, F., Hauttecoeur, B., Morelec, M.-J., Goldoni, P., Bizzini, B., and Tsiang, H. (1986). Involvement of gangliosides in rabies virus infection. *J.gen.Virol.* **67**, 47-56.

Superti, F., Seganti, L., Tsiang, H., and Orsi, N. (1984b). Role of phospholipids in rhabdovirus attachment to CER cells. *Arch.Virol.* **81**, 321-328.

Suzuka, I., Shimizu, N., Sekiguchi, K., Hoshino, H., Kodama, M., and Shimotohno, K. (1986). Molecular cloning of unintegrated closed circular DNA of porcine retrovirus. *FEBS Letters* **198**, 339-343.

Swanstrom, R. and Wills, J.W. (1997). Synthesis, assembly, and processing of viral proteins. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 263-334.

Tailor, C.S. and Kabat, D. (1997). Variable regions A and B in the envelope glycoproteins of feline leukemia virus subgroup B and amphotropic murine leukemia virus interact with discrete receptor domains. *J.Virol.* **71**, 9383-9391.

Tailor, C.S., Takeuchi, Y., O'Hara, B., Johann, S.V., Weiss, R.A., and Collins, M.K.L. (1993). Mutation of amino acids within the gibbon ape leukemia virus (GALV) receptor differentially affects feline leukemia virus subgroup B, simian sarcoma-associated virus, and GALV infections. *J.Virol.* **67**, 6737-6741.

Takeuchi, Y., Patience, C., Magre, S., Weiss, R.A., Banerjee, P., Le Tissier, P., and Stoye, J.P. (1998). Host range and interference studies of three classes of pig endogenous retrovirus. *J.Virol.* **72**, 9986-9991.

Takeuchi, Y., Porter, C.D., Strahan, K.M., Preece, A.F., Gustafsson, K., Cosset, F.-L., Weiss, R.A., and Collins, M.K.L. (1996). Sensitization of cells and retroviruses to human serum by (α1-3) galactosyltransferase. *Nature* **379**, 85-88.

Takeuchi, Y., Simpson, G., Vile, R.G., Weiss, R.A., and Collins, M.K.L. (1992a). Retroviral pseudotypes produced by rescue of a Moloney murine leukemia virus vector by C-type, but not D-type, retroviruses. *Virology* **186**, 792-794.

Takeuchi, Y., Vile, R.G., Simpson, G., O'Hara, B., Collins, M.K.L., and Weiss, R.A. (1992b). Feline leukemia virus subgroup B uses the same cell surface receptor as gibbon ape leukemia virus. *J.Virol.* **66**, 1219-1222.

Tanaka, J., Sadanari, H., Sato, H., and Fukuda, S. (1991). Sodium butyrate-inducible replication of human cytomegalovirus in a human epithelial cell line. *Virology* **185**, 271-280.

Taniguchi, S. and Cooper, D.K.C. (1997). Clinical xenotransplantation – a brief review of world history. *Xenotransplantation* (Cooper, D.K.C., Kemp, E., Platt, J.L. and White, D.J.G., Eds.) Springer-Verlag, Germany.

Telesnitsky, A. and Goff, S.P. (1997). Reverse transcriptase and the generation of retroviral DNA. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 121-160.

Temin, H.M. (1964a). Homology between RNA from Rous sarcoma virus and DNA from Rous sarcoma virus-infected cells. *Proc.Natl.Acad.Sci.USA* **52**, 323-329.

Temin, H.M. (1964b). Nature of the provirus of Rous sarcoma virus. *Natl.Cancer Inst.Monogr.* **17**, 557-570.

Temin, H.M. (1976). The DNA provirus hypothesis. *Science* **192**, 1075-1080.

Temin, H.M. (1980). Origin of retroviruses from cellular moveable genetic elements. *Cell* **21**, 599-600.

Temin, H.M. (1985). Reverse transcription in the eukaryotic genome: Retroviruses, retrotransposons and retrotranscripts. *Mol.Biol.Evol.* **6**, 455-468.

Temin, H.M. (1993). Origin and general nature of retroviruses. *The retroviridae Vol 1* (Levy, J.A., Ed.) Plenum Press, New York. 1-18.

Temin, H.M. and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226**, 1211-1213.

The General Hospital Corporation (1997). International Patent Application WO 97/21836

Thoulouze, M.-I., Lafage, M., Montano-Hirose, J.A., and Lafon, M. (1997). Rabies virus infects mouse and human lymphocytes and induces apoptosis. *J.Virol.* **71**, 7372-7380.

Thoulouze, M.-I., Lafage, M., Schachner, M., Hartmann, U., Cremer, H., and Lafon, M. (1998). The neural cell adhesion molecule is a receptor for rabies virus. *J.Virol.* **72**, 7181-7190.

Ting, C.N., Rosenberg, M.P., Snow, C.M., Samuelson, L.C., and Meisler, M.H. (1992). Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev.* **6**, 1457-1465.

Ting, Y.T., Wilson, C.A., Farrell, K.B., Chaudry, G.J., and Eiden, M.V. (1998). Simian sarcoma-associated virus fails to infect chinese hamster cells despite the presence of functional gibbon ape leukemia virus receptors. *J.Virol.* 72, 9453-9458.

Todaro, G.J., Benveniste, R.E., Lieber, M.M., and Sherr, C.J. (1974). Characterization of a type C virus released from the porcine cell line PK(15). *Virology* **58**, 65-74.

Toyishama, K. and Vogt, P.K. (1969). Enhancement and inhibition of avian sarcoma viruses by polycations and polyanions. *Virology* **38**, 414-426.

Tralka, T.S., Yee, C.L., Rabson, A.B., Wivel, N.A., Stromberg, K.J., Rabson, A.S., and Costa, J.C. (1983). Murine type C retroviruses and intracisternal A-particles in human tumors serially passaged in nude mice. *J.Natl.Cancer Inst.* **71**, 591-599.

Tucker, S.P., Srinivas, R.V., and Compans, R.W. (1991). Molecular domains involved in oligomerization of the Friend murine leukemia virus envelope glycoprotein. *Virology* **185**, 710-720.

Tuffereau, C., Benejean, J., Alfonso, A.-M.R., Flamand, A., and Fishman, M.C. (1998). Neuronal cell surface molecules mediate specific binding to rabies virus glycoprotein expressed by a recombinant baculovirus on the surfaces of lepidopteran cells. *J.Virol.* **72**, 1085-1091.

Van Beveren, C., van Straaten, F., Galleshaw, J.A., and Verma, I.M. (1981). Nucleotide sequence of the genome of a murine sarcoma virus. *Cell* 27, 97-108.

van Zeijl, M., Johann, S.V., Closs, E., Cunningham, J., Eddy, R., Shows, T.B., and O'Hara, B. (1994). A human amphotropic retrovirus receptor is a second member of the gibbon ape leukemia virus receptor family. *Proc.Natl.Acad.Sci.USA* **91**, 1168-1172.

Vanin, E.F., Kaloss, M., Broscius, C., and Nienhuis, A.W. (1994). Characterization of replication-competent retroviruses from nonhuman primates with virus-induced T-cell lymphomas and observations regarding the mechanisms of oncogenesis. *J.Virol.* **68**, 4241-4250.

Varmus, H. and Brown, P. (1989). Retroviruses. *Mobile DNA* (Berg, D.E. and Howe, M.M., Eds.) American Society for Microbiology, Washington,D.C. 53-108.

Vaughan, H.A., Loveland, B.E., and Sandrin, M.S. (1994). Gal alpha(1,3)Gal is the major xenoepitope expressed on pig endothelial cells recognized by naturally occurring cytotoxic human antibodies. *Transplantation* **58**, 879-882.

Venables, P.J.W., Brookes, S.M., Griffiths, D., Weiss, R.A., and Boyd, M.T. (1995). Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function. *Virology* **211**, 589-592.

Villareal, L.P. (1997). On viruses, sex and motherhood. *J.Virol.* **71**, 859-865.

Vogt, P.K. (1997). Historical introduction to the general properties of retroviruses. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 1-25.

Vogt, V.M. (1997). Retroviral virions and genomes. *Retroviruses* (Coffin, J.M., Hughes, S.H., and Varmus, H.E., Eds.) Cold Spring Harbor Laboratory Press, New York. 27-69.

Wagner, R.R. (1990). Rhabdoviridae and their replication. *Virology* (Fields, B.N., Knipe, D.M., *et al.*, Eds.) Second edition. Raven Press, New York. 867-881.

Walker, P.J. and Kongsuwan, K. (1999). Deduced structural model for animal rhabdovirus glycoproteins. *J.gen.Virol.* **80**, 1211-1220.

Wang, H., Dechant, E., Kavanaugh, M.P., North, N.A., and Kabat, D. (1992). Effects of ecotropic murine retroviruses on the dual-function cell surface receptor/basic amino acid transporter. *J.Biol.Chem.* **267**, 23617-23624.

Wang, H., Kavanaugh, M.P., North, N.A., and Kabat, D. (1991). Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. *Nature* **352**, 729-731.

Weclewicz, K., Ekström, M., Kristensson, K., and Garoff, H. (1998). Specific interactions between retrovirus env and gag proteins in rat neurons. *J.Virol.* **72**, 2832-2845.

Weiss, R.A. (1980). Rhabdovirus pseudotypes. *Rhabdoviruses Vol 3* (Bishop, D.H.L., Ed.) CRC Press Inc., Florida. 51-65.

Weiss, R.A. (1993). Cellular receptors and viral glycoproteins involved in retrovirus entry. *The retroviridae Vol 2* (Levy, J.A., Ed.) Plenum Press, New York. 1-108.

Weiss, R.A. (1998). Transgenic pigs and virus adaptation. *Nature* **391**, 327-328.

Weiss, R.A. and Bennett, P.L.P. (1980). Asssembly of membrane glycoproteins studied by phenotypic mixing between mutants of vesicular stomatitis virus and retroviruses. *Virology* **100**, 252-274.

Weissenhorn, W., Carfí, A., Lee, K.-H., Skehel, J.J., and Wiley, D.C. (1998). Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol.Cell* **2**, 605-616.

Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426-430.

White, D. (1996). Alteration of complement activity: a strategy for xenotransplantation. *Trends Biotechnol.* **14**, 3-5.

White, J.M. (1992). Membrane fusion. *Science* **258**, 917-924.

Whitt, M.A., Buonocore, L., Prehaud, C., and Rose, J.K. (1991). Membrane fusion activity, oligomerization, and assembly of the rabies virus glycoprotein. *Virology* **185**, 681-688.

Whitt, M.A., Chong, L., and Rose, J.K. (1989). Glycoprotein cytoplasmic domain sequences required for rescue of a vesicular stomatitis virus glycoprotein mutant. *J.Virol.* **63**, 3569-3578.

WHO (1978). Ebola haemorrhagic fever in Zaire, 1976. *Bull.World Health Organ.* **56**, 271-293.

Wilson, C., Reitz, M.S., Kayama, H.O., and Eiden, M.V. (1989). Formation of infectious hybrid virions with gibbon ape leukemia virus and human T-cell leukemia virus retroviral envelope glycoproteins and the *gag* and *pol* proteins of Moloney murine leukemia virus. *J.Virol.* **63**, 2374-2378.

Wilson, C.A., Wong, S., Muller, J., Davidson, C.E., Rose, T.M., and Burd, P. (1998). Type C retrovirus released from porcine primary peripheral blood mononuclear cells infects human cells. *J.Virol.* **72**, 3082-3087.

Witkor, T.J., Gyorgy, E., Schlumberger, H.D., Sokol, F., and Koprowski, H. (1973). Antigenic properties of rabies virus components. *J.Immunol.* **110**, 269-276.

Witte, O.N. and Baltimore, D. (1977). Mechanism of formation of pseudotypes between vesicular stomatitis virus and murine leukemia virus. *Cell* **11**, 505-511.

Wool-Lewis, R.J. and Bates, P. (1998). Characterization of Ebola virus entry by using pseudotyped viruses: Identification of receptor-deficient cell lines. *J.Virol.* **72**, 3155-3160.

Wu, B.W., Cannon, P.M., Gordon, E.M., Hall, F.L., and Anderson, W.F. (1998). Characterization of the proline-rich region of murine leukemia virus envelope protein. *J.Virol.* **72**, 5383-5391.

Wunner, W.H., Dietzschold, B., Smith, C.L., Lafon, M., and Golub, E. (1985). Antigenic variants of CVS rabies virus with altered glycosylation sites. *Virology* **140**, 1-12.

Wunner, W.H., Reagan, K.J., and Koprowski, H. (1984). Characterization of saturable binding sites for rabies virus. *J.Virol.* **50**, 691-697.

Wunsch, M., Schultz, A.S., Koch, W., Friedrich, R., and Hunsmann, G. (1983). Sequence analysis of Gardner-Arnstein feline leukaemia virus envelope gene reveals common structural properties of mammalian retroviral envelope genes. *EMBO J.* **2**, 2239-2246.

Ye, Y., Neethling, F.A., Niekrasz, M., Koren, E., Richards, S.V., Martin, M., Kosanke, S., Oriol, R., and Cooper, D.K. (1994). Evidence that intravenously administered alpha-galactosyl carbohydrates reduce baboon serum cytotoxicity to pig kidney cells (PK15) and transplanted pig hearts. *Transplantation* **58**, 330-337.

Yee, J.-K., Friedmann, T., and Burns, J.C. (1994). Generation of high-titer pseudotyped retroviral vectors with very broad host range. *Meth.Cell Biol.* **43**, 99-112.

Yoshinaka, Y., Katoh, I., Copeland, T.D., and Oroszlan, S. (1985a). Murine leukemia virus protease is encoded by the *gag-pol* gene and is synthesized through suppression of an amber termination codon. *Proc.Natl.Acad.Sci.USA* **82**, 1618-1622.

Yoshinaka, Y., Katoh, I., Copeland, T.D., and Oroszlan, S. (1985b). Translational readthrough of an amber termination codon during synthesis of feline leukemia virus protease. *J.Virol.* **55**, 870-873.

Young, J.A.T., Bates, P., and Varmus, H.E. (1993). Isolation of a chicken gene that confers susceptibility to infection by subgroup A avian leukosis and sarcoma viruses. *J.Virol.* **67**, 1811-1816.

Young, J.A.T., Bates, P., Willert, K., and Varmus, H.E. (1990). Efficient incorporation of human CD4 protein into avian leukosis virus particles. *Science* **250**, 1421-1423.

Young, J.A.T., Varmus, H.E., and Bates, P.F. (1994). A protein related to the LDL receptor is a cellular receptor specific for subgroup A avian leukosis and sarcoma viruses. *Cellular receptors for animal viruses* (Wimmer, E., Ed.) Cold Spring Harbor Laboratory Press, New York. 61-73.

Young, R.A. (1991). RNA polymerase II. *Annu.Rev.Biochem.* **60**, 689-715.

Yu, X., Yuan, X., Matsuda, Z., Lee, T.-H., and Essex, M. (1992). The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J.Virol.* **66**, 4966-4971.

Yu, X., Yuan, X., McLane, M.F., Lee, T.-H., and Essex, M. (1993). Mutations in the cytoplasmic domain of human immunodeficiency virus type 1 transmembrane protein impair the incorporation of Env proteins into mature virions. *J.Virol.* **67**, 213-221.

Zaidi, A., Schmoeckel, M., Bhatti, F., Waterworth, P., Tolan, M., Cozzi, E., Chavez, G., Langford, G., Thiru, S., Wallwork, J., White, D., and Friend, P. (1998). Life-supporting pig-to-primate renal xenotransplantation using genetically modified donors. *Transplantation* **65**, 1584-1590.

Závada, J. (1972). Pseudotypes of vesicular stomatitis virus with the coat of murine leukaemia and of avian myeloblastosis viruses. *J.gen.Virol.* **15**, 183-191.

Závada, J. (1982). The pseudotypic paradox. *J.gen.Virol.* **63**, 15-24.

Zhang, L. and Ghosh, H.P. (1994). Characterization of the putative fusogenic domain in vesicular stomatitis virus glycoprotein G. *J.Virol.* **68**, 2186-2193.

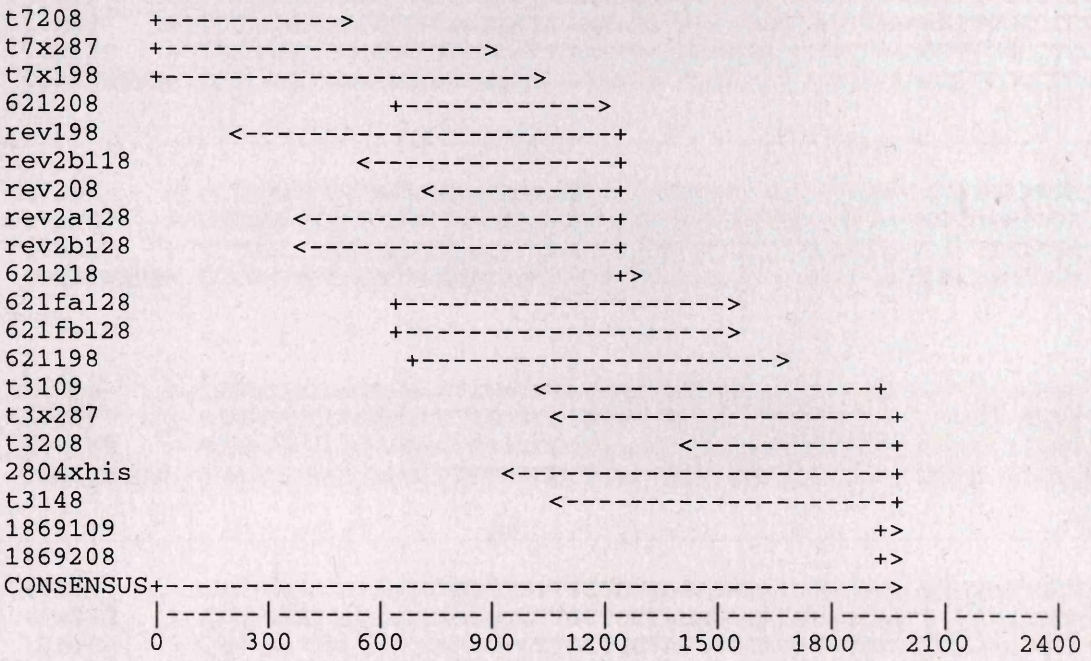
Zhao, Y., Zhu, L., Benedict, C.A., Chen, D., Anderson, W.F., and Cannon, P.M. (1998). Functional domains in the retroviral transmembrane protein. *J.Virol.* **72**, 5392-5398.

APPENDICES

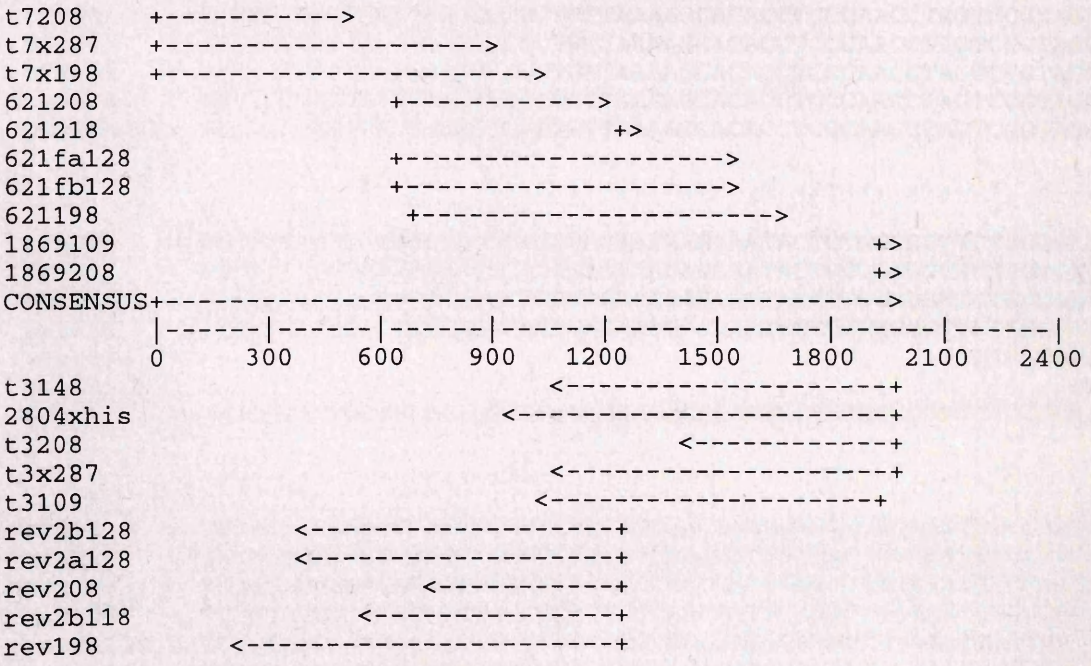
Appendices

A.1 Arrangement of assembled fragments obtained during the sequencing of PoERV-B *env*

All fragments



Fragments sorted by strand



A.2 Alignment of fragments obtained during the sequencing of PoERV-B *env*

```
t7208     > TAGAGTCGaCCCGGGCGGCCGCACCTGGCTCCATGCATCCCACGTTAAGCTGGCGCCACC 60
t7x287    >                GcGGCCGCRCCTGGCTCCATGCAT.CCACGTTAAGCTGGCGCCACC 46
t7x198    >             CCGGGCgGCCGCACCTGGCTCCATGCATCCCACGTTAAGCTGGCGCCACC 50
CONSENSUS> TAGAGTCGACCCGGGCGGCCGCACCTGGCTCCATGCATCCCACGTTAAGCTGGCGCCACC 60

                .........+.........+.........+.........+.........+.........+

t7208     > TCCCGACTCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCCTCCATCG 120
t7x287    > TCCCGACTCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCCTCCATCG 106
t7x198    > TCCCGACTCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCCTCCATCG 110
CONSENSUS> TCCCGACTCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCCTCCATCG 120

                .........+.........+.........+.........+.........+.........+

t7208     > CCTGGTTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCT 180
t7x287    > CCTGGtTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCT 166
t7x198    > CCTGGtTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCT 170
CONSENSUS> CCTGGTTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCT 180

                .........+.........+.........+.........+.........+.........+

t7208     > CGAACCCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCA 240
t7x287    > CGAACCCCCATAGACCTTTATCCCTTACCTGGCTGATTaTTGACCCTGATACGGGTGTCA 226
t7x198    > CGAACCCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCA 230
CONSENSUS> CGAACCCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCA 240

                .........+.........+.........+.........+.........+.........+

t7208     > CTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCT 300
t7x287    > CTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCT 286
t7x198    > CTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCT 290
rev198    <  TGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGCCCTGAACTGCATtTCT 59
CONSENSUS> CTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCT 300

                .........+.........+.........+.........+.........+.........+

t7208     > GCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGTAGTTATG 360
t7x287    > GCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGTAGTTATG 346
t7x198    > GCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGTAGTTATG 350
rev198    < GCCTCCGANTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGTAGTTATG 119
CONSENSUS> GCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGTAGTTATG 360

                .........+.........+.........+.........+.........+.........+

t7208     > GGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCT 420
t7x287    > GGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCT 406
t7x198    > GGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCT 410
rev198    < GGTNCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCT 179
rev2a128  <                                                CTGgGgAATCCT 12
rev2b128  <                                                      AATCCT 6
CONSENSUS> GGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCT 420

                .........+.........+.........+.........+.........+.........+

t7208     > TCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTC 480
t7x287    > TCTGTAGGAGATGGAGCtGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTC 466
t7x198    > TCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTC 470
rev198    < TCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTC 239
rev2a128  < TCTgTAgGAGATggAgCTGCGTCACCTCCAACGATGGAGACTGGAAATggCCgATCTCTC 72
rev2b128  < TCTGTAGNANATGGANCTGCGTCACCTCCAACNATGGANACTGGAAATgNCCNATCTCTC 66
CONSENSUS> TCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTC 480
```

```
t7208     > TCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAA                526
t7x287    > TCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCGAGTACAAAATGATGA 526
t7x198    > TCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAAATGATGA 530
rev198    < TCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAAATGATGA 299
rev2a128  < TCCAggACCGTgTAAAATTCTCCTTTGTCAATTCCGgCCCGGGCAAGTACAAAATGATGA 132
rev2b128  < TCCANGACCGTGTAAAATTCTCCTTTgTCAATTCCGGCCCGGGCAAGTACAAAATGATGA 126
CONSENSUS> TCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAAATGATGA 540

           .........+.........+.........+.........+.........+.........+

t7x287    > AACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATATAAAGATAAGTTTCA 586
t7x198    > AACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTTCA 590
rev198    < AACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTTCA 359
rev2b118  <                                CAGACTTAGATTATCTAAAGATAA.TTTCA 30
rev2a128  < AACTATATAAAGATAAGAgCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTTCA 192
rev2b128  < AACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAA.TTTCA 186
CONSENSUS> AACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTTCA 600

           .........+.........+.........+.........+.........+.........+

t7x287    > CTGAAAAAGGAAAAcAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAG 646
t7x198    > CTGAAAAAGGAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAG 650
rev198    < CTGAAAAAGGAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAG 419
rev2b118  < CTGAAAAAGGAAAACAGGAAAATATTCAAAANTgGATAAATGNTATGANCTgGAGAATAT 90
rev2a128  < CTGAAAAAGGAAAACAGGAAAATATTCAAAAgTGGATAAATGgTATGAGCTggtGAATAG 252
rev2b128  < CTGAAAAAGGAAAACAGGAAAATATTCAAAAGTGGATAAATggTATGAGCTGGgGAATAg 246
CONSENSUS> CTGAAAAAGGAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAG 660

           .........+.........+.........+.........+.........+.........+

t7x287    > TTTTTTATAATTATGGCGGGGGAGCAGGGTAGACTTTAACCATTCGACTTAGGATAGAGA 706
t7x198    > TTTTTTATAAATATGGCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATAGAGA 710
621208    >                          GGAGCAGGGTC.ACTTTAACCaKKCGCCTTAGGATAGAcA 40
rev198    < TTTTTTATAAATATGGCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATAGAGA 479
rev2b118  < TTTTTTATAAATATGNCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATAGAGA 150
rev2a128  < TTTTTTATAAATATAgCGGgGGAgCAGggTCCACTTTAACCATTCgCCTTAGGATAGAGA 312
rev2b128  < TtTTTTATAAATATAgCGGGgGAGCAGg.TCCACTTTAACCATTCGCCTTAGGATAGAGA 306
621fa128  >                          GGaG.AGGGTCCACTTTRACCATcCGCSTTAGGATAGAGA 40
621fb128  >                          GGaGcAGGGTCCACTTTaACCATcCGCSTTAGGATAGAGA 40
621198    >                                aCTTTAA.CATTCGCCTTACGATACACA 28
CONSENSUS> TTTTTTATAAATATGGCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATAGAGA 720

           .........+.........+.........+.........+.........+.........+

t7x287    > CGGGGAgAGAGCCCCCTGTSGCGATGGGACGCGATAGAGTACTGGCTG.AGGGGGGAGCC 766
t7x198    > CGGGGACAGAACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCC 770
621208    > CGGGGACAGAA.CCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAHGGCCCCC 100
rev198    < CGGGGACAGAACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGCCCCC 539
rev2b118  < CGGGGACAGAACCCCCTGTRGCAATGGGACCCGATAAAGTACTgGCTGAACAGGGCCCCC 210
rev208    <                                      AAAGTACTGGCTGAACAGGGCCCCC 25
rev2a128  < CGGGGACAGAACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGCCCCC 372
rev2b128  < CGGGGACAGAACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGCCCCC 366
621fa128  > CGGGGACAGARCCCCCTGTGGCAATGGGACCCGATAAAGTACTGGcTGAaCARGGCCCCC 100
621fb128  > CGGGGACAGARCCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAaCARGGCCCCC 100
621198    > CGGGGaCACAA.CCCCTGTCGCAATGGGA.CCGATAAACTACTGGCTGAACAcSG.CCCC 88
CONSENSUS> CGGGGACAGAACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGCCCCC 780

           .........+.........+.........+.........+.........+.........+
```

```
t7x287    > GGGCTGGAGCCA.GGCATAACTTGaCGGTGCGCCAGTT.AtCTCGCTGCGG.STGACATA 826
t7x198    > GGCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATA 830
621208    > GGCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATA 160
rev198    < GGCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATA 599
rev2b118  < GGCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATA 270
rev208    < GGCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATA 85
rev2a128  < KGCCTGGAGCCACCGCATAACTTGCCG.TGCCCCAATTAACCTCGCTGCGGCCTGACATA 432
rev2b128  < GGCCTGGAGCCACCGCATAACTTGCCG.TGCCCCAATTAACCTCGCTGCGGCCTGACATA 426
621fa128  > GGcCTGGaGCCACCGCATAAcTTGCCGGTGCCCCAATTaACCTCGCtGCGGCCTGASATA 160
621fb128  > GGCCTGGaGCCACCGCATAAcTTGCCGGTGCCCCAATTaACCTCGCtGCGGCCTGAcATA 160
621198    > GGCCTGGaGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCtGCGGCCTGACATA 148
CONSENSUS> GGCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATA 840

           .........+.........+.........+.........+.........+.........+

t7x287    > AcACAGNCGCCTAGC.ACAGTA.CACT.GATTGATTCTTAc..ACACGCCTAG.AGCTCC 886
t7x198    > ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 890
621208    > ACACAGCCGCCTAGCAACAGTACCACTRGATTGATTCCTACCAACACGCCTAGAAACTCC 220
rev198    < ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 659
rev2b118  < ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 330
rev208    < ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 145
rev2a128  < ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 492
rev2b128  < ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 486
621fa128  > ACACAGCCGCCTAGCAACAGTACCACTGGaTTNAKTCCTACCAACACGCCTAGAaACTCC 220
621fb128  > ACACAGCCGCCTAGCAACAGTACCACTGGaTTNAKTCCTACCAACACGCCTAGAaACTCC 220
621198    > ACACAGCCGCCTAGCAACACTACCACTGGAtTCATTCCTACCAACACGCCTAGAAACTCC 208
CONSENSUS> ACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCC 900

           .........+.........+.........+.........+.........+.........+

t7x287    > C.A.GTG.TGCT                                                 898
t7x198    > C.AGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGgGAGCTTTCCAA 950
621208    > CCAGGTGTTCCTGTTAAGACAGGACAGAtACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 280
rev198    < CCAGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 719
rev2b118  < CCAGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 390
rev208    < CCAGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 205
rev2a128  < CCAGGTGTTCCTGTTAAGACAGgACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 552
rev2b128  < CCAGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 546
621fa128  > CCAGGTGTTCCTGTKAAGASAGGACAGAGACTCTTCAGTcTCATCCAGGGAGCTTTCCAA 280
621fb128  > CCAGGTGTTCCTGTtAAGASAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 280
621198    > CCAGGTGTTCCTGTTAAGACAGGACAGAYACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 268
CONSENSUS> CCAGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAA 960

           .........+.........+.........+.........+.........+.........+

t7x198    > GCCATCAACTCCACCGACCCTGATGC.ACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 1010
621208    > GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 340
rev198    < GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 779
rev2b118  < GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 450
rev208    < GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 265
rev2a128  < GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 612
rev2b128  < GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 606
621fa128  > GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 340
621fb128  > GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 340
621198    > GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGG 328
2804xhis  <   CATCA.CTC.ACN.ACC.T.AT.cCACT.CT.CTN.TTG.CTTT..CTATC.TCAGN. 58
CONSENSUS> GCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTgTTGGCTTTGTCTATCCTCAGGG 1020

           .........+.........+.........+.........+.........+.........+
```

```
t7x198     > C.TCCTTATTATGAGGGGATGGCTAAA                                        1037
621208     > CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 400
rev198     < CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 839
rev2b118   < CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 510
rev208     < CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 325
rev2a128   < CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 672
rev2b128   < CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 666
621fa128   > CCTCCTTATTATGAGGGGATGGCTAAAGAaAGAaAATTCAATGTGACCAAAGAGCATAGA 400
621fb128   > CCTCCTTATTATGAGGGGATGGCTAAAGARAGAaAATTCAATGTGACCAAAGAGCATAGA 400
621198     > CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 388
t3109       <                    GATGGCTAAAGAAAGAgAATTCAATGTGACCAAAGAGCATAGA 43
2804xhis   < .CTCTATAT.A.....GGATg.CTAA.....gAAAAATCA.TNTcACCAAAgANCATAgA 118
CONSENSUS> CCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGA 1080

           .........+.........+.........+.........+.........+.........+

621208     > AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 460
rev198     < AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 899
rev2b118   < AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 570
rev208     < AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 385
rev2a128   < AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 732
rev2b128   < AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 726
621fa128   > AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 460
621fb128   > AATCAATGTACATGGGGGTCCCGAAATAaGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 460
621198     > AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 448
t3109       < AATCAATGTACATGGGG.TCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAA.GGG 103
t3x287      <                      CCAGAAATAAGCTTACCCTCNCTGAAGGTTCCGGGaAGgGN 41
2804xhis   < AATCAATNTACATgcCC.TCCA.AAATAANCTTACCNTCACTaAANTTTCcGGAAA.GAg 178
t3148       <           CA..GGGGTCCCGAA.TA.GCT.ACCcTCACTGAGGTTT.CGGGA.GGGG 50
CONSENSUS> AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGG 1140

           .........+.........+.........+.........+.........+.........+

621208     > ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 520
rev198     < ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 959
rev2b118   < ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 630
rev208     < ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 445
rev2a128   < ACATGCATAGGAAAA.CTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 792
rev2b128   < ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 786
621fa128   > ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 520
621fb128   > ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 520
621198     > ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 508
t3109       < ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 163
t3x287      < ACATgCATAGGAAAAGCTCCCCCATaCCACCANCACTTTtGCTNTGGTNCTGTGNTTTAT 101
2804xhis   < ACAT.CATAgAAAAggCTCCCCCATCCCACCAACACCTT.gCTATAgTACTGTGgTaTAT 238
t3148       < ACATGCATAGGAAA.GCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTGTAT 110
CONSENSUS> ACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTAT 1200

           .........+.........+.........+.........+.........+.........+

621208     > GAGCAGGCCTCAGAAAATCA                                                540
rev198     < GAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 1019
rev2b118   < GAGCAGGCCTCARAAAATCAGTATTTAGTACCTGGTTATAACAGGTGgTGGNCATgCAAT 690
rev208     < GAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 505
rev2a128   < GAGCAG.CCTCAGAAAATCAGTATTTAGTACCTNGTTATAACAG.TG.TGg.CATGCAAT 852
rev2b128   < GAGCAGGCCTCAgAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 846
621fa128   > GAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 580
621fb128   > GAGCAGGCCTCAGAaAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 580
621198     > GAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 568
t3109       < GAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 223
t3x287      < GANCAGACCTCAGAAAATCAGTATTTAGTACCTGGGTKTRNCAGGTGGTGgGcATGCAAT 161
2804xhis   < gAgCAGcCCTCAGAAAATCAGTATTTAgTACCTGgTTATAACAGgTGgTgCgCATNCAAT 298
t3148       < GAGCAGSCCTCAGAAAATCAGTATTTAGTACCTGGTTATaACAGGTGG.GGGCATGCAAT 170
CONSENSUS> GAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAAT 1260

           .........+.........+.........+.........+.........+.........+
```

```
rev198    < AcNCGGTCAA                                                                  1029
rev2b118  < ACT.GGTTARCC                                                                702
rev208    < ACTGG..TAACCC                                                               518
rev2a128  < AcT.GGTTAaCCC                                                               865
rev2b128  < AcT.GGTTARCCCG                                                              860
621218    >                    GTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC              44
621fa128  > ACTGGGTTAACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC              640
621fb128  > ACTGGGTTAACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC              640
621198    > ACTGGGTTAACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC              628
t3109     < ACTGGGTTAACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC              283
t3x287    < ACTGGNTTANCCCCCTGTGTTTCCNCNTCAGTCTTCAACcAATCCATAGATTTCTGTGTC              221
2804xhis  < ACTNcCTTAACCCCCTGTGTTTCCACCTCATTCTTCAACCAATCCAAAGATTTCTgTcTC              358
t3148     < ACTGGGTTAACCCCCTGTGTTCCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC              230
CONSENSUS> ACTGGGTTAACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTC             1320

                .........+.........+.........+.........+.........+.........+

621218    > ATGGTCCA                                                                    52
621fa128  > ATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAgTGGTCCTTGATgAATAT              700
621fb128  > ATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATgAATAT              700
621198    > ATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATAT              688
t3109     < ATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATAT              343
t3x287    < ATGGTCCAAATCGTCCCCCGAGTGTACTACCaTCCTGAGGAAGTGGTCCTTGATGAATAT              281
2804xhis  < ATgNTCCAAATCNTCCCCCGAgTNTACTACCATCCTGAGGAAGTGgTCCTTGATGAATAT              418
t3148     < ATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATAT              290
CONSENSUS> ATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATAT             1380

                .........+.........+.........+.........+.........+.........+

621fa128  > GACTATCgGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTC              760
621fb128  > GACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTC              760
621198    > GACTATCGGTATAACCGACCAAAAACAgAACCCGTATCCCTTACCCTAgCTGTAATGCTC              748
t3109     < GACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTC              403
t3x287    < GACTATCGGTATAACaGACcAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTC              341
t3208     <                                                                GCTC       4
2804xhis  < AACTATCGgTATAACCgACCAAAAAGAgAACCCTTATCCCTTACCCTAGCTgTAATgCTC              478
t3148     < GACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTC              350
CONSENSUS> GACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTC             1440

                .........+.........+.........+.........+.........+.........+

621fa128  > gGATTAGGGACGGCCGTTggCGTANgAACAGGGACAgCTgCCCTgATCACAgGACCACAN              820
621fb128  > GGATTAGGGACGGCCGTTgGCGTAGGAACAGGGACAGCTgCCCTGATCACAgGACCACAg              820
621198    > GGATTAGGGACGGCCGTTGGCGTAgGAACAGGGACAGCTGCCCTGATCACAgGACCACAG              808
t3109     < GGATTAGGGACGGC.GTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAG              463
t3x287    < GGATTAGGGACGCC.GTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAG              401
t3208     < GGATTAGGGACGGC.GTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAG               64
2804xhis  < GGATTAGggACGCCCgTTGgCgTAGgAACAGggACAgCTgCCCTgATCACAAGACCACAg              538
t3148     < GGATT.GGGACGGc.GTTGGCGTAGGAAC.GGGACAGCTGCCCTGATCACAGGACCACAG              410
CONSENSUS> GGATTAGGGACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAG             1500

                .........+.........+.........+.........+.........+.........+
621fa128  > CANCTANAGAAAggACTTNGTgAgCTACATgCGNCCATgACAgAAtATCT                       870
621fb128  > CAGCTANAGAAAgGACTTgGTgAGCTACATgCGgCCATgACAgAAtATCTCCGAgCCTTA           880
621198    > CAGCTAGAGAAAGGACTTGGTGAGCTACATGCgGNCATGACAgAAGATCTCCGAGCCTTA           868
t3109     < CAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTA           523
t3x287    < CaGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTA           461
t3208     < CAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTA           124
2804xhis  < CAGCTAGAGAAAGGACTTGgTGAgCTACATGCGCCCATAACAGAAgATCTCCGAgCCTTA           598
t3148     < CAGCKRGAGAARGGACTtGGTGAGCTACAtGCGGCCATGACAGAAGATCTCCgAGCCTTA           470
CONSENSUS> CAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTA           1560

                .........+.........+.........+.........+.........+.........+
```

```
621fb128 > AANGA                                                                885
621198   > AAgGAGTCTGTTAGCA.CCTAgA.GAcTCCCTGACTTCTTTgTCTGAAGTgGTTCTACAg 928
t3109    < AAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAG 583
t3x287   < AAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAG 521
t3208    < AAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAG 184
2804xhis < AAGGAgTCTgTTAgCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGgTTCTACAG 658
t3148    < AAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCtGAAGtGGTTCTACRG 530
CONSENSUS> AAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAG 1620

          .........+.........+.........+.........+.........+.........+

621198   > NACCgGAGgGNANTAGATCTGCTGTcTCTA.gAgA.gGTGNGT.ATGTGCAGC.TTAAA. 988
t3109    < AACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCCtTaAAa 643
t3x287   < AACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCCTTAAAA 581
t3208    < AACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCCTTAAAA 244
2804xhis < AACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGtTTATGTgCAGCCTTAAAA 718
t3148    < AaCCGGAGGGGATKAGATCTGCTGTTTCTAgGAGAAGGTGGGTTATGTGCAGCCTTAAAR 590
CONSENSUS> AACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCCTTAAAA 1680

          .........+.........+.........+.........+.........+.........+

621198   > GA.GNATGT.GCTTCTATGTAGATCACTCA                                  1018
t3109    < GAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTT 703
t3x287   < GAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTT 641
t3208    < GAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTT 304
2804xhis < GAAGAATgTTGCTTCTATGTAgATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTT 778
t3148    < GAAGAATGTGGCTTCTATGTAGATCACTCRGGAGCCATCAGAGACTCCATGAACARGCTt 650
CONSENSUS> GAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTT 1740

          .........+.........+.........+.........+.........+.........+

t3109    < AGAAAAAAGTTAGAGAGGCGTCGAAGGGAaAGAGAGGCTGACCAGGGGTGGTTTGAAGGA 763
t3x287   < AGAAAAAAGTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGA 701
t3208    < AGAAAAAAGTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGA 364
2804xhis < AGAAAAAAGTTAGAGAGgCgTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGA 838
t3148    < RGAAaaaaGTKRgAGAGGCGTSgAgGGGARRgAGAgGCTGACCRGGGGTGGTGGGAgGGA 710
CONSENSUS> AGAAAAAAGTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGA 1800

          .........+.........+.........+.........+.........+.........+

t3109    < TGGTTCAACAGGTCtCCTTGGATGACCACCCYGCTTTCTGCTCTGACGGGGCCCCTAGTA 823
t3x287   < TGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCTAGTA 761
t3208    < TGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCTAGTA 424
2804xhis < TGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGg.CCCCTAgTA 898
t3148    < gGGtTCAACaGGTCKCCtgGGATGACCACCcWGCTTTCtGVTCgGASGGGGCCCCtAGTA 770
CONSENSUS> TGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCTAGTA 1860

          .........+.........+.........+.........+.........+.........+
t3109    < GTCCTGCTCCTGTYACTtACAGTTGGGCCTtGCtTAATTAATAGGTTTGTTGCCTTTGTT 883
t3x287   < GTCCTGCTCCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTTGCCTTTGTT 821
t3208    < GTCCTGCtCCTGTWACTWACAGTTGGGCCTWGATTAAaTAATAGGTtTGTTGCCTTTGTT 484
2804xhis < gTCCTgCTCCTGTTACTTACAgTTGg.CCTTNCTTAATTAATAGATTTNTT.CCTTTNTT 958
t3148    < GTCCGGCRCCTGTRACTRACAGTGGGGCCgAGCTDAATTAAKRGgTgGGTGGCCKgGGTg 830
CONSENSUS> GTCCTGCTCCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTTGCCTTTGTT 1920

          .........+.........+.........+.........+.........+.........+

t3109    < AGAGAACGAGTGAGTGCAGTCCAGATCATGGTACTTAGGCAAC                     926
t3x287   < AGAGAACGAGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAG           873
t3208    < AGAGAACGAGTGAGWGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCATCTG 544
2804xhis < AgAGAACGAGTGAGTGCANTCCANATCATN.TACTTANCaGaCaG.ASCAAN.CCT.CT. 1018
t3148    < AGAGAAgGAGTGAGAGCAGaCCAGATCATGGtACTTaGGCAACAGTACCAAG.CCA.NTG 890
1869109  >                                    AGGCAACAGTACCAAGGCCTTCTG 24
1869208  >          AGTGAATASMGT.CAGATCATGGTACTTAGGCAACAGTACCAAGGCCTTCTG 52
CONSENSUS> AGAGAACGAGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTTCTG 1980
```

```
t3208     < AGCc                                                    548
2804xhis  < AgCC                                                    1022
t3148     < AGaCCAA                                                 897
1869109   > AGCCAAGGAGAAAcTGAcCTCTAGGCGGCCGC                        56
1869208   > AgCCAAGGMGAaAcTGAcCTCTAGGCGGCCGC                        84
CONSENSUS> AGCCAAGGAGAAACTGACCTCTAGGCGGCCGC                        2012

          .........+.........+.........+.........+.........+.........+
```