THE ROLE OF THE VIRUS-RECEPTOR INTERACTION IN THE DEVELOPMENT OF FeLV-RELATED ERYTHROID HYPOPLASIA (PURE RED CELL APLASIA)

Karen Willemina Adema

This thesis is submitted for the degree of PhD

Retrovirus Research Laboratory,
Department of Veterinary Pathology,
University of Glasgow,

Abstract

Feline leukaemia virus is one of the most common causes of viral disease in the domestic cat. Virus infection causes a number of diseases, including lymphoma, immunodeficiency, reproductive problems and a form of anaemia known as pure red cell aplasia (PRCA). There are three known subgroups of the virus (FeLV-A, FeLV-B and FeLV-C), as determined previously by interference assays and each subgroup displays a distinct *in vitro* cell tropism. FeLV-A is present in all natural isolates, whereas FeLV-B and FeLV-C are more seldom isolated (FeLV-B is present in approximately 40% of all natural isolates, while FeLV-C is present in only 1% of all natural isolates). The subgroups are determined by the amino acid sequence of the envelope glycoprotein with each subgroup of virus utilising a distinct cellular receptor. It is the specificity of the FeLV-C receptor interaction that is thought to underlie the development of PRCA with the selective destruction of erythroid progenitors by this subgroup of virus.

Subgroup C *envs* were cloned from a panel of primary isolates (the subgroup of these isolates had been established previously by superinfection interference). The *envs* were then characterised in order to determine the amino acids responsible for the characteristic *in vitro* tropism of subgroup C viruses. These analyses revealed that the subgroup C-component was more abundant in some primary isolates than in others. Furthermore, in some of the primary isolates more than one C-component was identified, suggesting that each isolate may consist of multiple viral variants. In agreement with previous studies, the *env* sequences were found to be relatively conserved with the majority of changes located within discrete variable regions, and in particular, the first variable region, the "VRA" region. However, a small number of characteristic mutations (unique to the isolate) were distributed throughout the entire envelope glycoprotein.

Although these mutations may simple reflect drift in the Env sequence due to the low fidelity of the viral reverse transcriptase, an alternative explanation is that they compensate for mutations within the VRA and ensure that the protein adopts a viable conformation enabling receptor recognition.

The *env* clones that displayed an *in vitro* tropism similar to the prototypic strain FeLV-C/Sarma shared two common motifs in the VRA region; the loss of a six amino acid stretch (^{61}TNVKHG^{66}) and a valine to tryptophan substitution on position 63 ($V_{63}W$) (tryptophan residues have been implicated in the formation of a hydrophobic pocket that is thought to be critical for interaction between the virus and its receptor).

In contrast with previous findings, the ability to infect porcine cells was found not to be a unique property of subgroup B viruses, both subgroup A and subgroup C Env clones were isolated that were capable of mediating infection of the porcine cell line ST Iowa (for example, an A-component of primary isolate FA27 and a C-component of primary isolate FY981). The presence of an aspartic acid to asparagine mutation on position 51 ($D_{51}N$) in the A-component of the isolate appeared to correlate with the expanded cell tropism. However, the C-components of primary isolate FA27, which also displayed the $D_{51}N$ mutation, were incapable of infecting ST Iowa cells. This may be due to additional mutations in the *env* resulting in an altered conformational structure of the C-component Env compared with the A-component that was capable of infecting this cell line. These findings suggest that a single mutation may not be sufficient to alter the *in vitro* tropism of a clone; additional mutations may ensure the correct folding of the protein and thus the accessibility of the receptor binding site. Although previous studies suggested that the amino acid sequence of the envelope glycoprotein of FeLV-A is highly conserved between isolates, it was found that minor differences in sequence were distributed throughout the entire *env*. These variations in the envelope glycoprotein may account for the differences in *in vitro* tropism displayed by the novel FeLV-A isolates.

To understand fully the extent of expression of the feline FeLV-C receptor (feFLVCR) and its possible role in the onset of PRCA, a panel of both haematopoietic and non-haematopoietic tissues was screened for feFLVCR mRNA. In contrast with the human FeLV-C receptor (thought to be expressed preferentially in haematopoietic tissues), the feFLVCR is expressed in both haematopoietic and non-haematopoietic tissues and there was no single tissue where the receptor was consistently expressed to higher or lower levels. Next, the extent of receptor expression in cells that are involved in the erythroid maturation was examined. Bone marrow samples were collected and subsequently enriched for or depleted of cells that are involved in erythroid maturation using a monoclonal antibody directed against cells of this lineage. Although only a small number of bone marrow samples were tested, the feFLVCR appeared to be expressed at high levels in the bone marrow populations that were enriched for cells of the erythroid lineage. These findings support the hypothesis that the normal function of this gene is disrupted upon infection and subsequently leads to the onset of PRCA.

This thesis is dedicated to my grandparents,
Opa en Oma Adema and
Oma and Opa Brinkhuis

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Brian J. Willett, for all his help, support and understanding throughout the years that I have worked with him and I am grateful for all the opportunities he has given me. Furthermore, I would like to thank everybody in the Retrovirus Research Laboratory and the Department of Veterinary Pathology, in particular Dr Linda Scobie, for all their help and technical support over the years. I am also grateful to the Wellcome trust for funding this research.

I would like to say a big thank you to Donna for all the nights out, yoga sessions and her friendship over the years and to Dot for all the munrobagging trips up north, crazy runs along the West Highland Way and plenty bottles of wine. I've kept my sanity because of you guys!! Furthermore, I would like to thank everybody that have made my stay in Scotland a memorable one, especially Donna, Dot, Sam, Mark, Linda and Annemarie, cheers guys, hope to see you again in The Netherlands!!! Last, but certainly not least, I am extremely grateful to my family for all their support and their never ceasing confidence in me! Without it, I wouldn't have persevered!! Furthermore, I'm extremely grateful to Marco and dad for all their help and patience with the printing!

TABLE OF CONTENTS

CHAPTER ONE - Introduction

CHAPTER TWO - Methods

CHAPTER THREE - Cloning and sequencing of novel isolates of FeLV-C

CHAPTER FOUR - Enrichment of FeLV-C isolates by biological selection

CHAPTER FIVE – Expression of the FeLV-C receptor feFLVCR *in vivo*.

CHAPTER SIX – Identification and characterisation of a putative FeLV-A receptor.

CHAPTER SEVEN - Discussion

Appendices

LIST OF FIGURES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

LIST OF TABLES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER ONE

INTRODUCTION

1.1 General introduction

Feline leukaemia virus (FeLV) is a naturally occurring retrovirus and a widespread pathogen of the domestic cat. Infection may result in the development of a number of diseases, including immunodeficiency, haematopoietic tumours and anaemia. FeLV is a member of the retroviridae and can be classified into four main subgroups, A, B, C and T, and the three subgroups have distinct cell tropism. The main determinant of the cell tropism of virus is the interaction between the envelope glycoprotein of the virus and a cellular reception. FeLV-C is the subgroup that causes a fatal pure red cell aplasia (PRCA) in the cat and it has been shown that mutations in a single region of the envelope glycoprotein of FeLV-A reproduce the anaemogenic phenotype of FeLV-C, suggesting that the interaction between the virus and the receptor underlies the development of PRCA.

1.2 Structure and classification of the *retroviridae*

Retroviruses are a large family of viruses that cause many different diseases, including immunodeficiency, neurological disorders and tumours. Infection with the virus can also lead to a viraemia with no obvious signs of disease, an "inapparent" infection. The name is derived from the opposite way of working of the viruses: all pathogens translate their DNA into RNA whereas the retroviruses have an RNA genome that gets translated by the enzyme "Reverse Transcriptase" (RT). The virion structure, genome organisation and replication strategies are similar for all of the *retroviridae*, despite the diversity of host species and host-virus interactions. Retrovirus virions are enveloped particles that are approximately 100 nm in diameter. Contained within each particle is the viral genome consisting of two linear single stranded (SS) RNA molecules encoding for the *gag*, *pol* and *env* genes, arranged from the 5' capped end to the 3' polyadenylated end of the genome. The envelope gene (*env*) encodes for the surface glycoprotein of the virus. The internal structural proteins, the capsid, or core, and nucleocapsid are encoded for by the group-specific antigen (*gag*) gene. The polymerase (*pol*) gene encodes for the proteins involved in viral replication.

A number of retroviruses (such as the spuma- and lentiviruses) contain additional genes that are involved in the regulation of virus replication, these viruses are referred to as complex retroviruses; viruses lacking these additional genes are referred to as simple retroviruses, such as the oncovirinae.

Until recently, the *retroviridae* were classified in accordance with their morphology in electron microscopy images, however this classification has been modified as more is known about their structure and the diseases they cause. The oncovirinae have been regrouped into 5 new genera, depending on virus structure and outcome of infection in the host (see table 1). Most of the retroviruses that were previously classified as oncovirinae, cause cancer in their host species, however there are some oncovirinae that are non-pathogenic. The lentivirinae have long incubation periods and cause lifelong, persistent infections regardless of an extensive immune response. They induce two different types of diseases, they either induce a degenerative and inflammatory disease of specific connective tissues (for example the Maedi Visna viruses) or they cause pathological changes in the immune system, which result in severe immunodeficiency syndromes (the most well known example being Human Immunodeficiency Virus (HIV)). These differences in disease are due to the differences in main target cells: the first group mainly infects macrophages whereas the second group are able to infect both macrophages and lymphocytes (Jindrak L. and Grubhoffer L., 1999). In contrast, the spumavirinae cause non-pathogenic lifelong infections in naturally or accidentally infected hosts although they do cause vacuolation of cells *in vitro*.

Genus	Morphology	Examples
α-retrovirus	C type	RSV, ASLV.
β-retrovirus	B and D type	JSRV, MMTV.
γ-retrovirus	C type	FeLV, MuLV, PERV, RD114, GALV.
δ-retrovirus		BLV, HTLV-1, STLV-1.
ε-retrovirus		WDSD.
Lentivirus		HIV, SIV, FIV
Spumavirus		HFV, SFV

Table 1.1: *Classification of the retroviruses.*

RSV, Rous Sarcoma Virus; ASLV, avian sarcoma/leukosis virus; JSRV, Jaagsiekte Sheep Retrovirus; MMTV, mouse mammary tumour virus; FeLV, Feline Leukaemia Virus; PERV, Porcine Endogenous RetroVirus; GALV, Gibbon Ape Leukaemia Virus; MuLV, Murine Leukaemia Virus; BLV, Bovine Leukaemia virus; HTLV, Human T-cell Leukaemia Virus; STLV, Simian T-cell Leukaemia Virus; WDSD, Walleye Dermal Sarcoma Virus; HIV, Human Immunodeficiency Virus; FIV, Feline Immunodeficiency Virus; SIV, Simian Immunodeficiency Virus; HFV, Human Foamy Virus; SFV, Simian Foamy Virus

1.3 γ-Type retroviridae

The γ-type retroviridae were previously classified as C-type mammalian retroviruses and contain the following retroviruses: feline leukaemia virus (FeLV), murine leukaemia virus (MuLV), porcine endogenous retrovirus (PERV), Gibbon ape leukaemia virus (GALV) and the endogenous feline retrovirus RD-114. Of this group, MuLV is perhaps the most well studied retrovirus, indeed it is commonly used to study leukaemogenesis and carcinogenesis in general. Furth et al (Furth J. et al., 1933) originally identified the virus by breeding high-leukaemia-incidence mouse strains in 1933. Independent isolates of the agent were successfully identified by a number of people, such as Gross, Friend, Moloney and Rauscher (Fan H., 1999). Each isolate induces distinct leukaemias, for instance Moloney MuLV induces T cell lymphoma, whereas Friend MuLV induces an erythroleukemia. The predominant infection sites of MuLV are the haematopoietic tissues, such as bone marrow and lymph nodes. The latency of these viruses in mice ranges from 2 months to 18 months, dependent on virus strain and mouse strain. The MuLV have been classified using host range studies and the principle of interference (discussed in more detail in section 1.4.3), determining the differences in the interaction of the viral surface glycoprotein with the cell surface receptor, this has led to the following five subtypes. Ecotropic MuLV (e-MuLV) is an exogenous virus that is capable only of infecting murine cells. Xenotropic MuLV (X-MuLV) is an endogenous virus and is capable of infecting non-murine cells only. Amphotropic and polytropic MuLV (A-MuLV and P-MuLV) are capable of infecting both murine and non-murine cells. GALV is a retrovirus of the gibbon ape that causes T-cell leukaemia. It is transmitted horizontally and is closely related to the endogenous virus of the Asian mouse *Mus Caroli*, which suggests trans species transmission (Reitz Jr M.S., 1999). The virus most closely resembles a new world monkey virus, simian sarcoma associated virus (SSAV) (Delassus S. et al., 1989). However it also bears resemblance to both MuLV and FeLV. All pigs carry several copies of the RNA genome of porcine endogenous retroviruses (PERV) (Patience C. et al., 1997). Although PERV is not known to cause any disease in pigs, the virus can be activated to produce virions that are infectious for human cells and thus could have severe implications for the use of pig organs for xenotransplantation (Akiyoshi D.E. et al., 1998). There are three subgroups know to date, PERV-A, PERV-B and PERV-C, as determined by host range and interference studies (Takeuchi Y. et al., 1998).

RD-114 is an endogenous, xenotropic retrovirus with multiple copies of the complete viral genome present in all domestic cats (Hardy W.G.Jr., 1993). RD-114 is not known to cause any disease and does not recombine with either FIV or FeLV to cause recombinant, infectious retroviruses.

1.4 Feline leukaemia virus

1.4.1 Discovery and origins of feline leukaemia virus

FeLV was first discovered in 1964 by Jarrett et al in Scotland (Jarrett W.F.H. et al., 1964). An increased incidence of lymphosarcomas in a multiple cat household had been reported, suggesting that a contagious agent caused these. In order to determine if a contagious agent was the cause of these lymphosarcomas, four kittens were injected with material from a lymphosarcoma (Mackey L. et al., 1975). All kittens developed symptoms and died or were euthanised at different times. Analysis of tissue from a kitten euthanised after 18 months showed virus-like particles in intracellular vesicles. The virus-like particles resembled particles observed in mice and chickens infected with leukaemia-causing virus. FeLV is thought to originate from an ancestral rodent virus and shows striking similarities both in sequence and in genetic structure with murine leukaemia virus, a leukaemogenic γ-type retrovirus of the laboratory mouse (Benviste R.E. et al., 1975; Neil J.C. and Onions D.E., 1985).

1.4.2 Structure of the virus

FeLV is a γ-type retrovirus with two linear single stranded (SS) RNA molecules surrounded by a protein core, an inner protein coat and a viral envelope. The genome is approximately 8 kilobases (kb) in length and consists of the genes *gag*, *pol* and *env*, arranged from the 5' capped end to the 3' polyadenylated end of the genome (Hardy W.G.Jr, 1983a), (Hardy W.G.Jr, 1983b).

The *gag* gene (group-specific antigen) encodes the structural proteins p10 (10 kDa), p12 (12 kDa), p15 (15 kDa) and p27 (27 kDa), which are similar in all FeLV isolates. P10 is the nucleocapsid (NC) and is associated with the virion RNA, p15 is the matrix protein (MA) that forms the outer layer around the core and p27 is the internal capsid protein (CA) that is the major component of the core. The precise function of p12 (an unnamed protein) is unknown. P27 is the protein that is recognised by all the available diagnostic FeLV tests. The *pol* gene (polymerase) encodes the replication enzymes protease (PR), integrase (IN) and reverse transcriptase (RT). RT is involved in the transcription of RNA into double stranded DNA. This DNA integrates into the cellular chromosome, forming the provirus. The virus genome then gets translated alongside the cellular genome. The *env* gene (envelope) encodes the viral precursor protein gp85, which the enzyme protease subsequently cleaves into the glycosylated extracellular surface glycoprotein gp70 (70 kDa), the non-glycosylated transmembrane protein p15E (15 kDa) and a signal peptide. The viral envelope is a phospholipid membrane with spikes formed by the gp70 protein, these spikes are held in place by the hydrophobic p15E protein. The receptor on the cell surface recognises these spikes and this results in cell entry. A cartoon of the structure of the virus and the location of its antigens is depicted in figures 1.1 and 1.2.

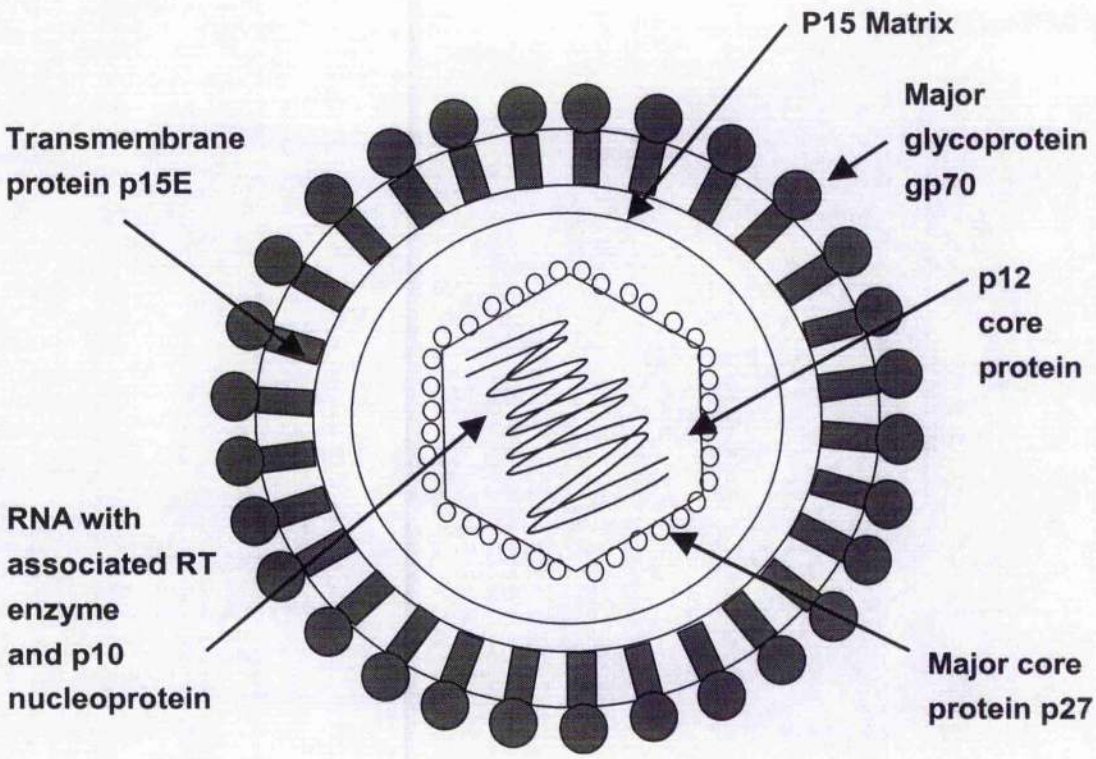
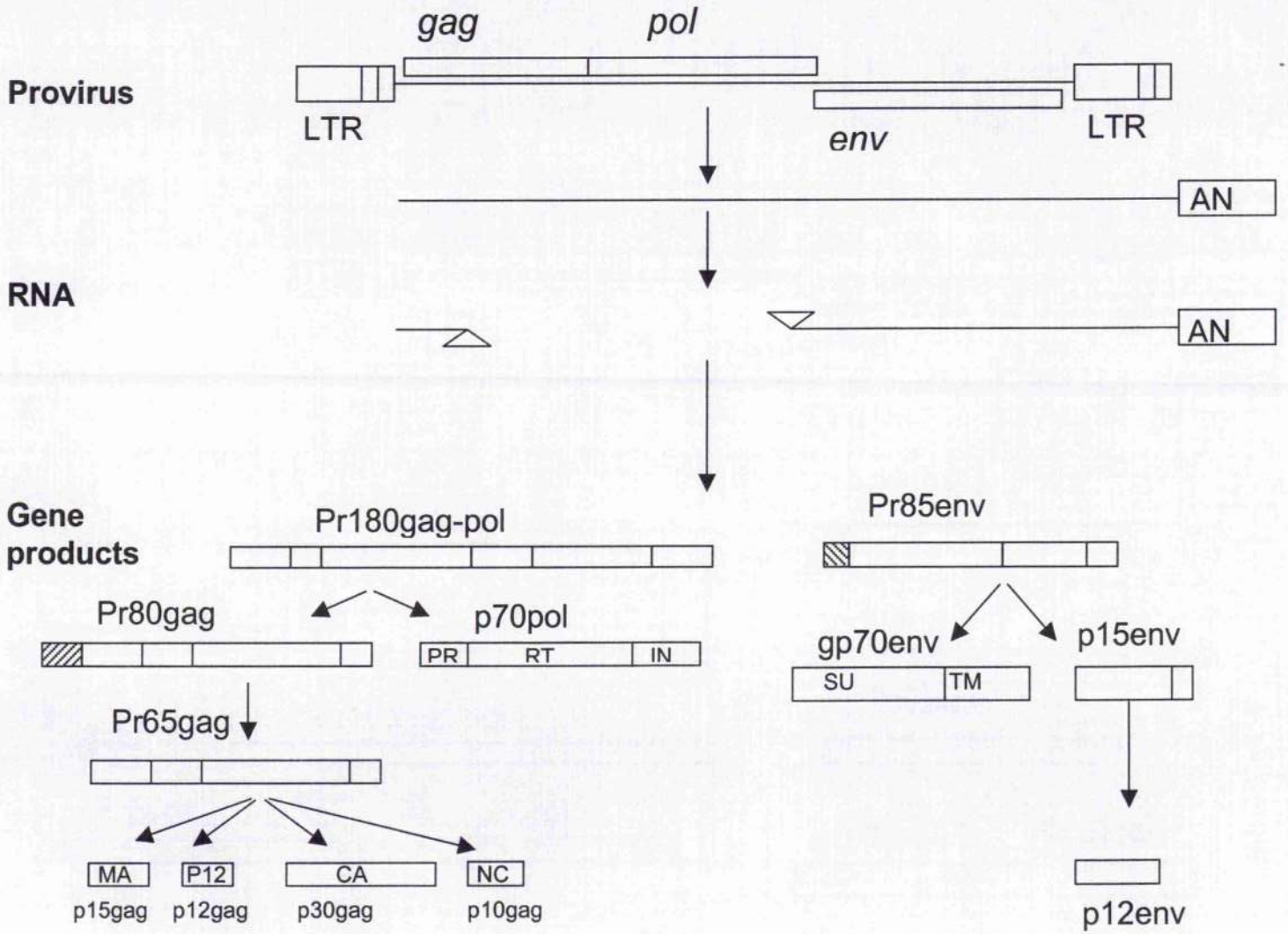


Figure 1.1: *FeLV particle and its antigens.*

Figure 1.2: *The proviral structure, RNA transcripts and gene products of FeLV. Adapted from Neil and Onions (Neil J.C. and Onions D.E., 1985). The provirus is flanked by two long terminal repeats (LTR) and contains two open reading frames encoding for gag-pol and env. Two polyadenylated RNA transcripts (An), one RNA transcript covers the entire genome whereas the other RNA transcript is spliced to encode the env coding sequence. When the RNA is translated, the longer RNA transcript is translated to produce the internal virion protein precursor Pr65gag. The precursor is cleaved by the viral protease and the products include the p15 matrix protein (MA), the p27 capsid protein (CA) and the p10 nucleocapsid (NC). The envelope protein precursor Pr85env is cleaved and its products include the glycosylated mature gp70 virion surface glycoprotein (SM) and the non-glycosylated transmembrane anchor protein p15E.*

(Opposite page)


gag
pol
Provirus
LTR
env
LTR
AN
RNA
AN
Gene products
Pr180gag-pol
Pr85env
Pr80gag
p70pol
PR
RT
IN
gp70env
p15env
SU
TM
Pr65gag
MA
P12
CA
NC
p15gag
p12gag
p30gag
p10gag
p12env

1.4.3 Subgroups of FeLV

There are four known subgroups of the virus, FeLV-A, FeLV-B, FeLV-C and FeLV-T; FeLV-A, FeLV-B and FeLV-C were identified first using viral interference and neutralisation assays (Sarma P.S. and Log T., 1973). Subsequently, FeLV-T was described as an immunodeficiency causing FeLV isolate by Overbaugh et al (Overbaugh J. et al., 1988a) and Moser et al (Moser M. et al., 1998) showed the interference properties of two closely related FeLV variants. Interference was first demonstrated in avian type C viruses and has since then been used to identify subgroups of many viruses. Reciprocal receptor interference is the process whereby cells chronically infected with one virus remain resistant to superinfection by viruses that use the same cell surface receptor. These infected cells remain susceptible to infection by viruses that use different receptors. Non-reciprocal interference (see figure 1.3) has been observed in murine and avian retroviruses and after identification of the receptors, it was shown that some viruses could use more than one receptor and therefore were able to sustain multiple infection (Sommerfelt M.A., 1999). Viral subgroups are defined by differences in the amino acid sequence of the envelope glycoprotein and therefore their host ranges and cell surface receptor usage.

1.4.3.1 FeLV-A

FeLV-A is the only subgroup that can be transmitted between cats and therefore is present in all isolates (Jarrett O. et al., 1984). FeLV-A can infect only feline cells (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). It is minimally pathogenic in the absence of other subgroups (Roy-Burman P., 1996). Recently, Moser et al has shown that their prototypic FeLV-A isolate, 61E, is capable of infecting not only feline cells, but also some canine and human cell lines (Moser M. et al., 1998). None of the FeLV subgroups are able to infect either mouse or rat cell lines.

1.4.3.2 FeLV-B

FeLV-B is the result of recombination between FeLV-A and endogenous FeLV (Elder J.H. and Mullins J.I., 1983; Stewart M.A. et al., 1986; Overbaugh J. et al., 1988a; Overbaugh J. et al., 1988b). The FeLV-B *env* sequences differ in five regions of the gp70 (Neil J.C. et al., 1987). Probes derived from these divergent regions of the envelope glycoprotein were used to show that the FeLV-B *env* gene is derived from endogenous proviral DNA (Stewart M.A. et al., 1986). There are approximately 15 copies of the enFeLV in the genome of the domestic cat (Benviste R.E. et al., 1975; Roy-Burman P., 1996). However, these endogenous FeLV sequences do not generate infectious virus and do not cause disease (Roy-Burman P., 1996). FeLV-B is present in approximately 40% of all natural isolates and causes mild, non-fatal anaemia and is also linked to the occurrence of lymphosarcomas and other lympho- or myeloproliferative diseases (Jarrett O. et al., 1984). It has a broader *in vitro* host range than FeLV-A as it is capable of infecting both feline and non-feline cells (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993).

1.4.3.3 FeLV-C

FeLV-C is the result of point mutations in the envelope glycoprotein of FeLV-A and shows a close similarity with FeLV-A. The majority of the differences are located in the variable regions of the envelope glycoprotein. The determinants of the FeLV-C subgroup and disease phenotype have been located to the 3' *pol* and 5' *env* region of the virus, which includes the VRA region of the envelope glycoprotein (Riedel N. et al., 1988). In a study by Rigby et al (Rigby M.A. et al., 1992), the region of the envelope glycoprotein that is involved in receptor recognition and pathogenesis has been narrowed down to a single cysteine bounded domain encompassing the VR1 region. The FeLV-C strains known to date differ in this region, suggesting that each isolate arose independently from FeLV-A. FeLV-C is the most rare subgroup; it is present in only 1% of all natural isolates (Jarrett O. et al., 1984). This subgroup has the broadest in vitro host range; it can infect both feline and non-feline cell lines and it is the only subgroup capable of infecting guinea pig cells (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993).

FeLV-C causes aplastic anaemia, a disease that is characterised by the fatal erythroid hypoplasia or pure red cell aplasia (PRCA). The main haematological symptoms of PRCA are a normal number of early erythroid precursor cells (BFU-E), but a decreased number of late erythroid precursor cells (CFU-E). These symptoms suggest a defect in the maturation from BFU-E to CFU-E (Abkowitz J.L et al., 1985). Clinical signs are not obvious until the anaemia is severe, they include weight loss, fever and listlessness (Hoover E.A. et al., 1974; Onions D.E. et al., 1982; Abkowitz J.L et al., 1985; Abkowitz J.L et al., 1998). It has been shown that new-born specific pathogen-free (SPF) kittens that are infected with FeLV-C develop pure red cell aplasia within eight weeks, however 8-week-old kittens that are infected with FELV-C do not become infected, and showing that the infection is age related (Hoover E.A. et al., 1974; Riedel N.O. et al., 1986). Previous research has shown that inserting the point mutations, which distinguish FeLV-C from FeLV-A, into the envelope gene of FeLV-A generates a virus with an expanded cell tropism, including non-feline cell lines. These results suggest that the interaction between the virus and its cell surface receptor underlie the development of PRCA (Onions D.E. et al., 1982; Abkowitz J.L et al., 1987; Rigby M.A. et al., 1992; Brojatsch J. et al., 1992; Abkowitz J.L et al., 1998).

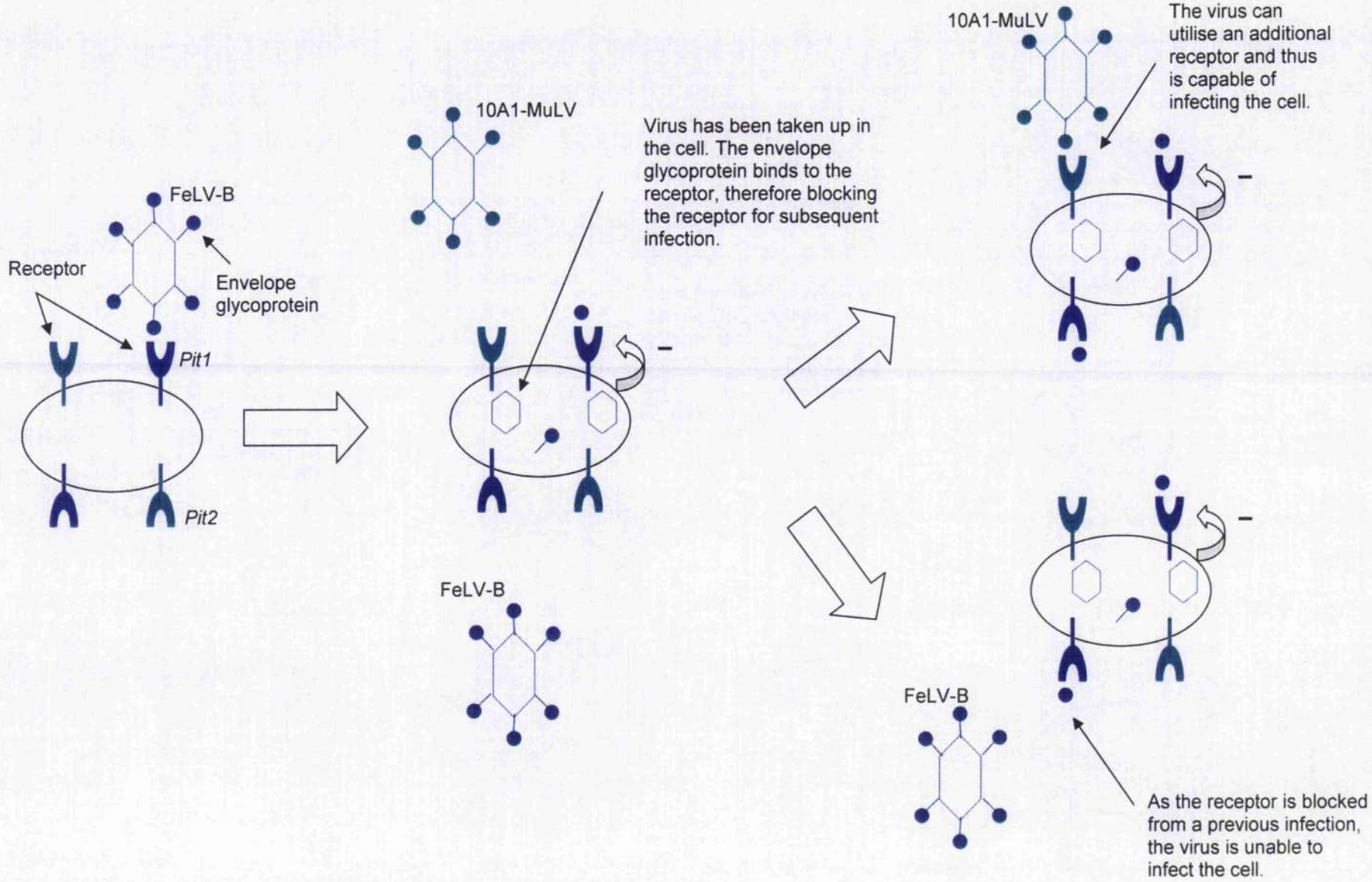
1.4.3.4 FeLV-T

In 1988, Overbaugh et al (Overbaugh J. et al., 1988a) described a naturally occurring FeLV isolate that induces a fatal immunodeficiency similar to AIDS in humans. In this study, it was shown that the isolate could not induce a disease without the help of a replication-competent virus. Later studies showed that there is a 98% sequence homology between their two isolates, suggesting they are closely related isolates that were cloned from a single virus (Overbaugh J. et al., 1992). It was shown that there are two regions with differences in the amino acid sequence essential for the conversion into the cytopathic virus, namely a region near the N-terminus and a region of 7 amino acids near the C-terminus. More recently, Rohn et al (Rohn J.L. et al., 1998) showed that the immunodeficiency-inducing isolate is unable to establish superinfection interference against a homologous challenge, suggesting that the cytopathic properties of this isolate might be due to the lack of interference against superinfection.

When the cell tropism of both the immunodeficiency-inducing isolate and the non-cytopathic isolate was studied, Moser et al (Moser M. et al., 1998) found no interference occurs between the two clones, suggesting they use distinct receptors for viral entry and therefore these isolates are distinct subgroups.

Figure 1.3: *Principle of non-reciprocal interference.*

(Opposite page)


FeLV-B
Receptor
Envelope glycoprotein
Pit1
Pit2
10A1-MuLV
Virus has been taken up in the cell. The envelope glycoprotein binds to the receptor, therefore blocking the receptor for subsequent infection.
FeLV-B
10A1-MuLV
The virus can utilise an additional receptor and thus is capable of infecting the cell.
FeLV-B
As the receptor is blocked from a previous infection, the virus is unable to infect the cell.

1.4.4 Determinants of host range and pathogenicity

1.4.4.1 γ-Type retroviruses envelope glycoproteins

The SU domains of FeLV, MuLV and GALV are composed of three distinct, highly variable domains; two variable regions VRA and VRB and a proline rich region PRR. The remainder of the SU envelope glycoprotein is highly conserved between species (figure 1.4a and 1.4b). The virus subtype variations are mainly found in VRA and VRB. VRA begins approximately 50 amino acids from the amino terminal end of the SU envelope glycoprotein, is 35 to 80 amino acids long and is flanked by two conserved cysteine residues. A conserved region consisting of 38 amino acids separates VRA and VRB; this region contains a cysteine residue. The second variable region, VRB, is 6 to 30 amino acids is size and contains a conserved cysteine residue (Battini J-L. et al., 1992).

1.4.4.2 MuLV envelope glycoprotein studies

In order to investigate the domains of the envelope gene that are involved in host range and pathogenicity, Battini et al (Battini J-L. et al., 1995) and Gray et al (Gray K.D. and Roth M.J., 1993) studied the envelope glycoprotein of MuLV. Both studies concluded that the VRA region is the main determinant in receptor recognition, although VRA alone is not sufficient for efficient entry into the cell. The PRR is necessary for efficient virus entry, but mutations in this region do not alter receptor recognition. Therefore, it is thought that the PRR is involved in maintaining the correct conformational shape of the receptor binding domain and therefore allowing efficient binding of the virus.

1.4.4.3 FeLV envelope glycoproteins studies

Riedel et al (Riedel N. et al., 1988) investigated the main determinants involved in the induction of pure red cell aplasia (PRCA) as a result of FeLV infection and found an 886 base pair region that encodes 72 amino acids of the 3' end of the *pol* gene and 241 amino acids of the N-terminal region of the envelope glycoprotein.

After infection of SPF cats with the chimaeras made, these cats developed anaemia, confirming that this region is involved in the onset of PRCA. The chimaeras also had a wider host range, proving that this region is also involved in the determination of host range of the virus. A few years later Brojatsch et al (Brojatsch J. et al., 1992) and Rigby et al (Rigby M.A. et al., 1992) constructed chimaeras in order to narrow down the domains involved. Brojatsch et al looked into 3 anaemia-inducing FeLV-C isolates, FA27, FS246 and FZ215. Sequence analysis showed that FA27 and FS246 isolates were more closely related to FeLV-A, however there was a single amino acid difference between FZ215 and FeLV-C, suggesting that this was not an independent isolate. Chimaeras were made by inserting the previously determined region into a FeLV-A backbone. The chimaeras that included the SU domain of the envelope glycoprotein of FeLV-C all had a broader host range, including guinea pig cells, proving that the host range determinant is situated in this region of the envelope. After examining the sequences of the different isolates, it was observed that there was only one consistent amino acid change in the FeLV-C isolates; this is a $K_{98}L$ change in the VR1 region, affecting the secondary structure. However, this structural change on its own was insufficient to alter the host range specificity. Rigby et al narrowed down the subgroup specificity; host range and pathogenicity determinants to a cysteine bounded region of gp70, including the VR1 region. The cysteine residues are not only conserved amongst all FeLV isolates, but also amongst other γ-retroviruses, such as MuLV. The isolates studied varied in sequence, showing that they had arisen independently from FeLV-A. The chimaeras that were constructed induced anaemia in neonates, although this anaemia was less virulent and progressed more slowly than the PRCA induced by FeLV-C.

Figure 1.4a: *Schematic structure of the RNA genome of FeLV. The gp70 surface glycoprotein (SU) and the location of the variable regions VRA, VRB and PRR are shown in more detail.*

(Opposite page)

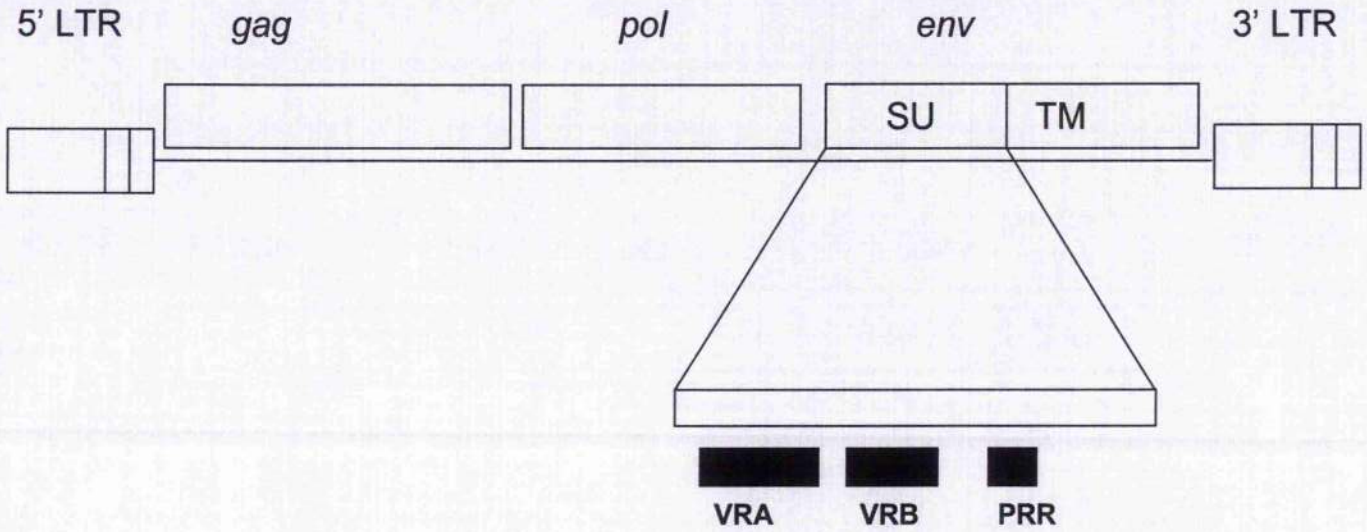

5' LTR
gag
pol
env
3' LTR
SU
TM
VRA
VRB
PRR

Figure 1.4b: *Amino acid sequence alignment from the SU proteins of the γ type retroviruses upstream of the PRR. Arrows indicate the start of the SU protein and the PRR. An asterisk indicates a conserved cysteine residue, whereas a dot indicates a gap in alignment. The standard one letter code for amino acids is used and the amino acid position is indicated on the right.*

⇨ SU protein

```
e-mulv marstlsk.............plknkvnprgpliplillmlrgvstaspgssphqvynitwevtng.dretvwatsgnhplwtwwpdltp
galv   mvllpgsmlltsnlhhlrhqmspgswkrliillscvfggggt.....slqnknphqpmtltwqvlsq.tgdvvwdtkavqppwtwwptlkp
felv-a mespthpk.............pskdktlswnlaflvgilftidigmanp..sphqiynvtwvitnvqtntqanatsmlgtltdayptlhv
felv-b mespthpk.............pskdktlswnlvflvgilftidigmanp..sphqvynvtwtitnlvtgtkanatsmlgtltdafptmyf
felv-c mespthpk.............pskdktfpwnlvflvgilfqidmgmanp..sphqvynvtwvitnvqtnsranatsmlgtltdayptlyv
```

VRA

```
e-mulv dlcmlah hgpsywgleyqspfssppgppccsggsspgcsrdceepltsltprcntawnrlkldqtthksnegfyvcpgphrpresks..
galv   dvcalaa .sleswdipgtdvssskrvrppdsdytaaykqitwgai......gcsyprartrmasst......fyvcprdgrtlsearr.
felv-a dlcdlvg ...dtwepivlnptnvkhgarysssky.................gckttd.rkkqqqtyp.....fyvcpghapslgpkgth
felv-b dlcdiig ...ntwnpsdqepfpgy...........................gcdqpm.rrwqqrntp.....fyvcpghanrkq.....
felv-c dlcdlvg ...dtwepiapdprswaryssssth....................gckttd.rkkqqqtyp.....fyvcpghapsmgpkgty
```

VRB

```
e-mulv cggpdsfycaywgcettgraywkpssswdfitvnnnlt sdqavqv.......ckdn.............  kwcnplvirftdagrrvts.
galv   cggleslyckewdcettgtgywlsksskdlitvkwdqn sewtqkfqq.....chqt.............  gwcnplkidftdkgk..lsk
felv-a cggaqdgfcaawgcettgeawwkptsswdyitvkrgss qdns..........ce...............  gkcnplvlqftqkgrq.as.
felv-b cggpqdgfcavwgcettgetywrptsswdyitvkkgvt qgiyqcsgggwcgpcydkavhsstt.gaseg  grcnplilqftqkgrq.ts.
felv-c cggaqdgfcaawgcettgeawwkptsswdyitvkrgsn qdns..........ck...............  gkcnplvlqftqkgrq.as.
```

⇨ PRR

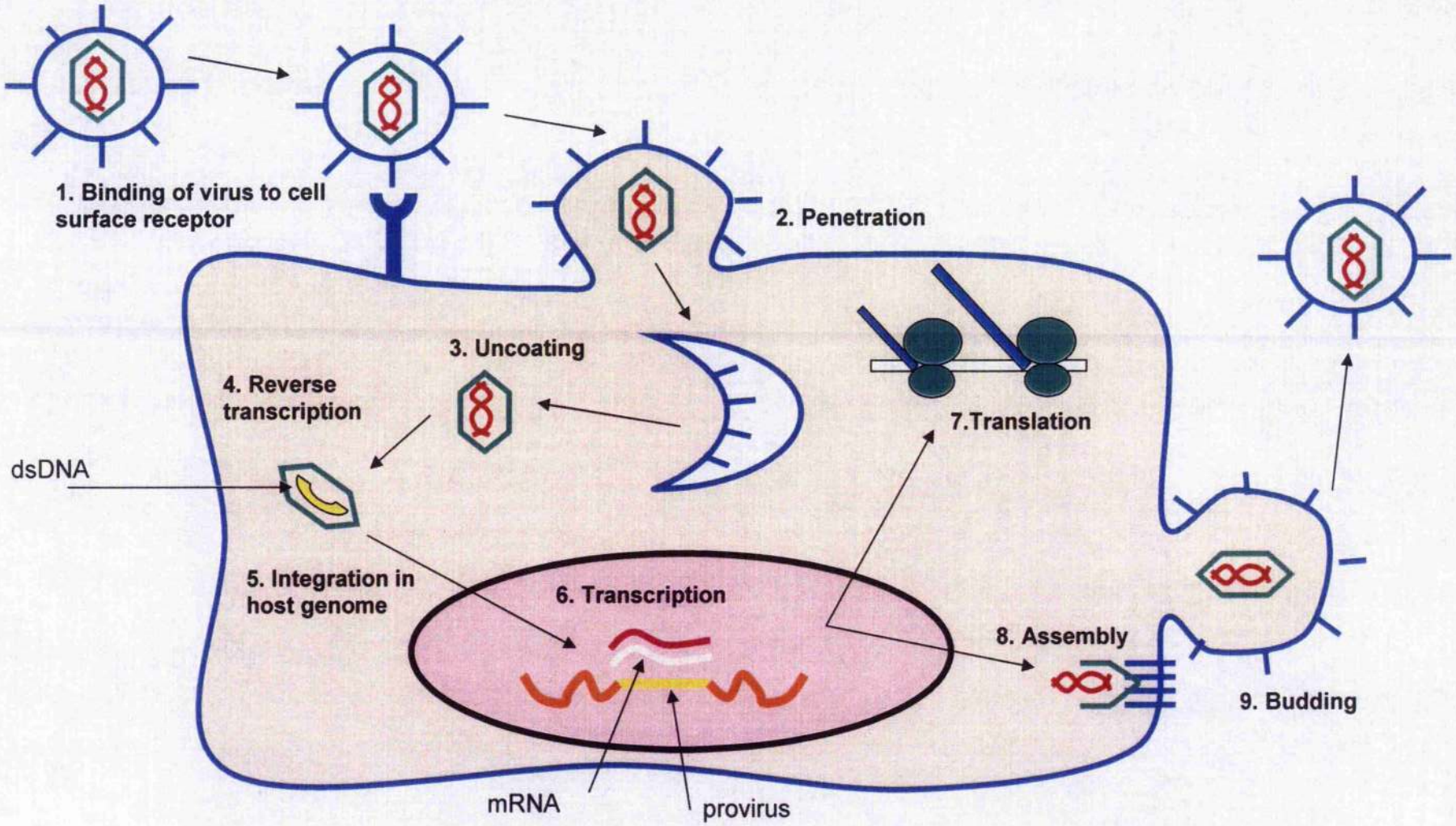
```
e-mulv .wttghywglrlyvsgq.dpgltfgirlryqnlgprvp
galv   dwitgktwglrfyvsg..hpgvqftirlki.tnmpava
felv-a .wdgpkmwglrlyrtgy.dpialftvsrqvstitppqa
felv-b .wdgpkswglrlyrsgy.dpialfsvsrqvmtitppqa
felv-c .wdrpkmwglrlyrsgy.dpialfsvsrqvmtitppqa
```

1.4.5 Replication and life cycle

FeLV is thought to be transmitted predominantly by biting and licking (the virus is present in saliva). Following exposure to the virus, the epithelial surfaces of the eye, mouth and nose are penetrated and the virus enters the local lymph nodes, where it replicates in lymphocytes. At this stage, most cats are able to fight off infection and become immune to the virus. After replication in lymphocytes, the virus enters the bone marrow where it is able to replicate to high titres. The virus binds to specific cell surface receptors on the target cell, after which it enters the cell and uncoats to reveal the RNA genome. The RNA genome is transcribed into cDNA by the RT enzyme of the virus. This cDNA is flanked by two long terminal repeats (LTR) and integrates into the host cellular DNA, this process is mediated by the viral protein IN. The LTR regions contain enhancers that drive the transcription of the viral RNA and promoters that have processing signals for the cleavage and poly-adenylation of RNA transcripts. The viral mRNA (both spliced and unspliced mRNA) is then transported to the cytoplasm and translated into viral proteins. These initial products are polyproteins, which contain all the protein domains for the genes. At this stage, the virus particles are assembled at the cell surface from where they derive a host cell derived outer membrane and virus encoded spikes. Cleavage of the polyproteins by the viral PR results in mature viral particles, which are subsequently released into the bloodstream without damage to the infected cell (see figure 1.5). At this point the structural protein p27 is detectable in the blood. The virus spreads to the salivary glands and the infection cycle can start again from the beginning. This cycle takes approximately 4 weeks (Hardy W.G.Jr, 1983a; Hardy W.G.Jr, 1983b; Jarrett O., 1994; Jarrett O., 1999). Transmission through the placenta is rare, as most queens that are infected with the virus are infertile (Caney S., 2000), however, infection can occur from an infected queen to her new-born kitten via the milk (Hardy W.G.Jr et al., 1976; Jarrett O., 1999).

Figure 1.5: *Life cycle of a retroviral particle.*

(Opposite page)


1. Binding of virus to cell surface receptor
2. Penetration
3. Uncoating
4. Reverse transcription
dsDNA
5. Integration in host genome
6. Transcription
7.Translation
8. Assembly
9. Budding
mRNA
provirus

1.4.6 Epidemiology and immune response

Infection with FeLV is a significant cause of death amongst young adult cats and most infections occur between one and five years of age. In the United Kingdom, the prevalence of FeLV is 18% in sick cats and 5% in healthy cats, compared to 19% in sick cats and 6% in healthy cats for FIV (Hosie M.J. et al., 1989). Similar results were found in the United States, where the prevalence of FeLV is 13% in sick cats and 4% in healthy cats, compared to 24% in sick cats and 4% in healthy cats for FIV (Shelton G.H. et al., 1989). There is a distinct difference between the outcome of infection in multiple cat households and in free-roaming cats, due to the close contact cats in multiple cat households have with each other and their shared usage of bowls and litter pans. Of these cats, approximately 30 to 40% become viraemic, whereas approximately 70% of FeLV negative cats have virus neutralising (VN) antibodies, indicating that they have been in contact with the virus (Jarrett O., 1994). Of free-roaming cats that are exposed to the virus, 28% become persistently infected, 42% become transiently infected and 30% remain uninfected (Hoover E.A. and Mullins J.I., 1991; Hardy W.G.Jr., 1993). Persistently infected cats have high levels of viral capsid protein p27 in their blood and this can be used to diagnose their FeLV infection. Transiently infected cats are able to fight off the infection and become immune to subsequent FeLV infection. Within four to eight weeks after infection, virus expression is brought to a halt, although the virus is not completely eliminated and unexpressed non-replicating FeLV is present in the bone marrow. "Latent" FeLV in bone marrow is distinct from endogenous FeLV (Hardy W.G.Jr., 1993). The latent FeLV can be reactivated *in vitro*, although this is uncommon (Rojko J.L. et al., 1982). Pacitti et al (Pacitti A.M. and Jarrett O., 1985) looked at the long-term effects of latent infection. In a previous study, 4 viraemic cats were placed in a household consisting of 19 susceptible 4-month-old kittens. The virus neutralising antibody titre and the anti-feline oncornavirus-associated cell membrane antigen (FOCMA) antibody titre was determined, as well as virus isolation from the blood and bone marrow.

It was shown that the period of latency is short; within months of infection the bulk of cats is able to clear the infection from the bone marrow. However, a small portion of cats remained latently infected for a longer period. At 138 weeks, when the study was concluded, only 14% of the cats remained latently infected.

Similar results were found by Hoover et al (Hoover E.A. et al., 1976) when comparable experiments were carried out. It is not known why the latent stage of the virus comes to an end, however there are 2 possible reasons for this:

1. The virus infects a small portion of the immature cells in the bone marrow. Therefore, after a period of time, these cells gradually differentiate, leaving the bone marrow free of virus.
2. The latently infected cells are eliminated by the immune response of the cat.

The cats remain negative for the presence of p27 in serum; they develop persistent VN antibodies and FOCMA titres. FOCMA is a cell membrane antigen that is produced by cells infected with either FeLV or feline sarcoma virus FeSV (these viruses are antigenically indistinguishable). Studies have shown that cats which develop FeLV malignancies do not have high FOCMA antibody titres and that inadequate anti-FOCMA response is a cause, rather than an effect, of ensuing FeLV lymphosarcoma (Essex M. et al., 1975; Essex M. et al., 1979). Some cats produce antibodies against FOCMA, but these antibodies are not capable of neutralising the virus (Essex M. et al., 1971; Hardy W.G.Jr., 1993). However, the true nature and function of FOCMA remains illusive and needs to be studied in more detail to ascertain its role in FeLV infection. Whether the FeLV infected cat is able to fight off infection appears to be dependent on several factors, including the age of the cat, status of the immune system at the time of exposure and the amount of virus at the time of infection. Susceptibility to infection decreases with age; cats that are infected with FeLV at an early age are more likely to develop persistent viraemia and FeLV-related disease than cats infected later in life. Similar results were observed in FeSV (Hoover E.A. et al., 1976). The outlook for persistently infected cats is very poor, most viraemic cats die within 3 years from infection, mostly from non-neoplastic disease (Hardy W.G.Jr et al., 1976).

1.4.7 Clinical signs

Infection with the virus can cause both neoplastic and non-neoplastic diseases such as immunodeficiency, reproductive failure, haematopoietic tumours and anaemia (see table 1.2) (Jarrett O. et al., 1984). Studies by, amongst others, Mullins et al (Mullins J.I. et al., 1984) and Neil et al (Neil J.C. et al., 1987) have shown that the different types of cancer caused by FeLV infection is due to the activation of cellular oncogenes, resulting in the unlimited proliferation of cells characteristic of leukaemia. Similarly, it has been shown previously that the haematological abnormalities characteristic for pure red cell aplasia are induced by the specific FeLV-C viral strains (Linenberger M.L. and Abkowitz J.L, 1995). Interestingly, infection with FeLV-A alone does not appear to cause disease, leading to the conclusion that FeLV-A evolves in order to generate pathogenic variants of the disease.

Malignant haemapoietic diseases	Tissues affected
Lymphosarcoma	*thymic lymphosarcoma, multicentric lymphosarcoma, alimentary lymphosarcoma.*
Leukaemias	*Lymphoid leukaemia, myeloid leukaemia, erythroleukaemia.*
Non-malignant haemapoietic diseases	**Symptoms**
Anaemia	*Haemolytic anaemia, pure red cell aplasia.*
Immunodeficiency	*Comparable to AIDS in humans, thus susceptible to a range of opportunistic infections.*
Marrow aplasia	*Hypoplasia of the bone marrow cells*
Non-malignant non-haemapoietic diseases	**Symptoms**
Reproductive failure	*Spontaneous abortion or death soon after birth*
Enteritis	*Degeneration of the epithelial cells located in villous crypts*

Table 1.2: *Diseases associated with FeLV infection.*

1.4.7.1 Pure red cell aplasia

FeLV-C causes aplastic anaemia, a disease that is characterised by the fatal erythroid hypoplasia or pure red cell aplasia (PRCA). The main haematological symptoms of PRCA are similar to the PRCA in humans which are a normal number of early erythroid precursor cells (BFU-E), but a decreased number of late erythroid precursor cells (CFU-E), suggesting a defect in the maturation from BFU-E to CFU-E as can be seen in figure 1.6 (Abkowitz J.L et al., 1985). However, two separate studies by Testa et al (Testa N. et al., 1983) and Boyce et al (Boyce J.T. et al., 1981) have reported a significant drop in BFU-E. This could be the result of different techniques or different amounts of virus. Other symptoms associated with PRCA are severe anaemia with haematocrit levels of 4 to 15%, absence of reticulocytes in blood, no haemoglobinised cells in the bone marrow, high levels of erythropoietin and a severely diminished erythropoiesis (Linenberger M.L. and Abkowitz J.L, 1995). However, the red cell survival is normal. Clinical signs are not obvious until the anaemia is severe; they include weight loss, fever and listlessness (Hoover E.A. et al., 1974; Onions D.E. et al., 1982; Abkowitz J.L et al., 1987; Abkowitz J.L et al., 1998). Feline PRCA does not respond well to immunosuppressive treatments, such as prednisone or cyclosporine (Linenberger M.L. and Abkowitz J.L, 1995). As the onset of anaemia is more acute in neonatal and immunosuppressed cats then in adult cats, the PRCA results from viral replication rather then as a consequence of the immune response of the host (Hoover E.A. et al., 1974; Riedel N.O. et al., 1986; Linenberger M.L. and Abkowitz J.L, 1995).

Several groups have tried to identify the viral components that are involved in the development of PRCA, however to date the regions that are essential in the onset of PRCA still remain to be identified (Riedel N. et al., 1988; Rigby M.A. et al., 1992; Brojatsch J. et al., 1992). Rigby et al (Rigby M.A. et al., 1992) have shown with their chimaeras that the VRA region of the envelope is the major determinant of tropism of the virus as a chimaera of the VRA region of FeLV-C in FeLV-A backbone renders guinea pig cells susceptible. They have also shown that this region is important in the onset of acute anaemia, suggesting that the cell surface receptor is necessary for normal erythroid maturation.

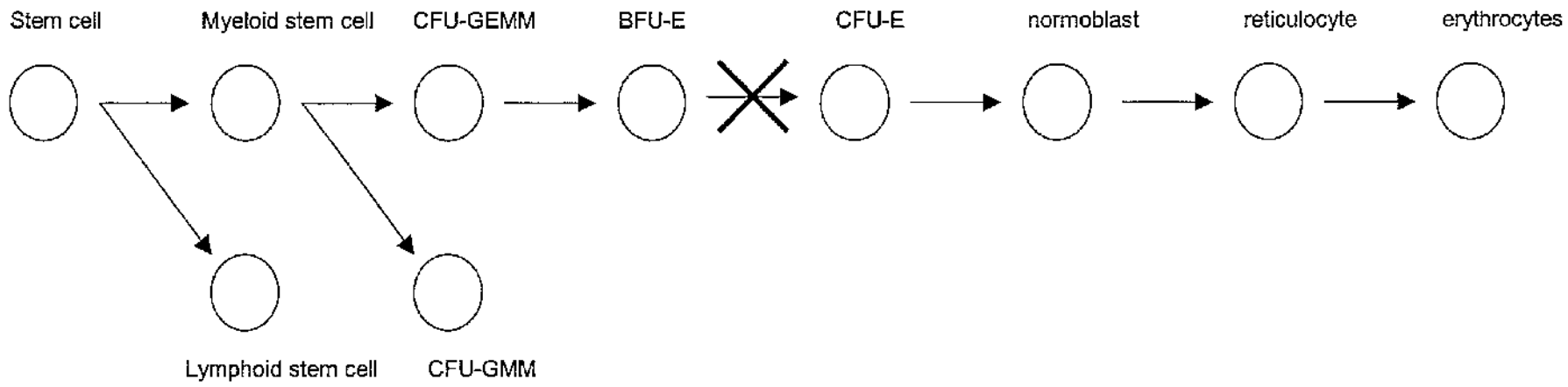


Figure 1.6: *Normal haematopoiesis and the defect in erythroid maturation in the onset of pure red cell aplasia (PRCA).*

1.5 Viral receptors

The identification of viral receptors is important for two reasons. Firstly, as the receptor plays an important role in virus entry, the expression of the receptor on tissues and cells is a major determinant in the pathogenesis and cell tropism of the virus. Secondly, once receptors have been identified, novel drugs that inhibit virus-cell interactions can be designed. Receptors are components on the cell surface to which a virus binds specifically and this interaction result in virus entry. This is in contrast to a virus-binding site, which is a component on the cell surface that permits virus attachment, but binding of the virus does not result in virus entry. It was previously thought that the binding of a ligand to its receptor was a one step process, dependent on temperature, pH and ionic strength. With the discovery of large numbers of receptors and their subsequent characterisation, it is now believed that virus binding is a multi-step process. Each step involves conformational changes in either host or receptor or viral envelope glycoprotein, allowing subsequent steps to take place. The secondary interactions between virus and target cell strengthens adhesion, enabling virus uptake by either of the two following pathways:

1. Fusion of the viral envelope with the plasma membrane.
2. Cellular uptake of virions by endocytosis.

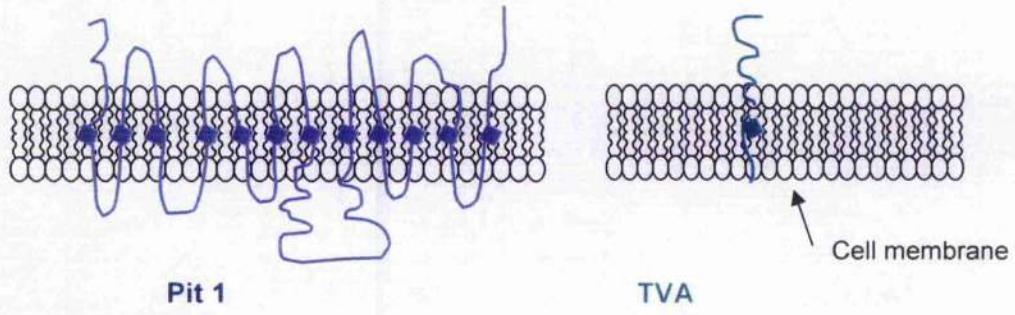
1.5.1 Identification and characterisation of viral receptors

There are several methods to identify and characterise viral receptors; the most common approach is to transfer the gene of a potential receptor into non-permissive cell line and demonstrate the ability of this gene to render the cell line permissive for infection. An alternative method is to screen a panel of monoclonal antibodies against cell surface antigens to determine whether any of the antibodies protect the cells against viral infection (Haywood A.M., 1994; Jindrak L. and Grubhoffer L., 1999). A range of cell surface molecules has been identified as viral receptors, ranging from immunoglobulin-like molecules to multi transmembrane transporter proteins. In certain cases (for example the retroviral receptors) binding of the viral envelope glycoprotein severely disrupts normal cellular function of the molecule.

1.5.2 Retroviral receptors

Over the years, there has been considerable interest in identifying and characterising retroviral receptors. Retroviral infection leads to serious disease in both humans and animals. Similarly, retroviruses can be used for the delivery of DNA in gene therapy. Over the years, a number of retroviral receptors have been identified and characterised. There are many types of cell surface proteins that can act as retroviral receptors; however, it is thought that members of the same genera use similar cell surface molecules as their receptor. As described previously, retroviral envelope glycoproteins are oligomers of SU anchored to the viral membrane by the TM. The SU bears epitopes for specific receptor recognition. Binding of the virus to the cellular receptor leads to conformational changes in the SU, exposing a hydrophobic peptide in the TM domain. Fusion between the virus and the cell is catalysed by the exposed peptide. It is thought that fusion is a multi-step process, however the exact mechanism is not yet fully understood. The conformational changes to expose the peptide are pH-independent. The level of receptor expression does not always correspond with the level of infection, suggesting there are other cellular factors involved in virus entry (Sommerfelt M.A., 1999).

The structure of retrovirus receptors can be divided into 5 general groups (figure 1.7). Protein receptors that have been identified for γ-type retroviruses are generally multiple transmembrane spanning proteins (multiple TM receptor). The number of times a receptor spans the membrane is variable in the different receptors. Examples of this group are *Pit1* (FeLV-B), FLVCR (FeLV-C) and mCAT-1 (E-MuLV) (figure 1.7a). Several retroviruses (for instance HIV) require an additional co-receptor for efficient entry in the cell. For example, CD4 is the main receptor for HIV-1 with the chemokine CCR5 acting as a co-receptor or fusion receptor. After initial binding of the HIV SU to CD4, conformational changes occur in the viral SU, uncovering the binding site for CCR5 and allowing fusion to occur between virus and cell. To date, several chemokines, that can act as co-receptor for HIV, have been identified, such as CCR3, CCR5, CXCR4 and SDF-1 (figure 1.7b). FeLV-T appears to be the only γ-type retrovirus that requires an additional factor for successful entry; it uses *Pit1* as a primary receptor and the secreted protein FeLIX as an additional factor (figure 1.7c).

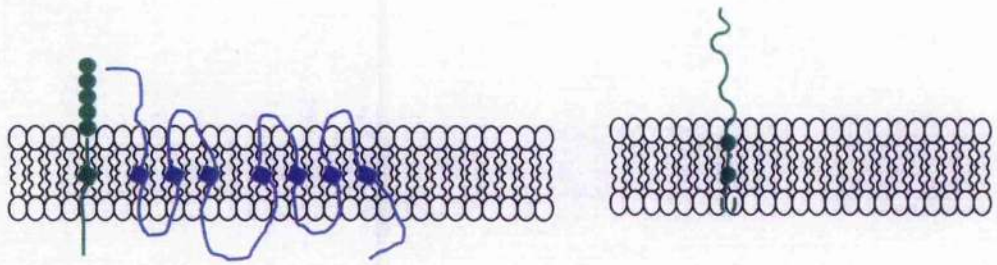


Pit 1

TVA

a. Multiple TM

d. Single TM

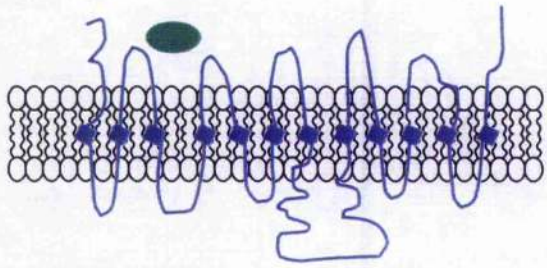


CD4 CCR5

HYAL2?

b. Single TM with multiple TM

e. GPI anchor



FeLIX?

Pit1?

c. Multiple TM with secreted factor

Figure 1.7: *Different classes of proteins utilized as a receptor by the γ-type retroviridae.*

Most α-retroviruses use single transmembrane spanning proteins (single TM) as a receptor. To date, the receptors for the avian sarcoma/leukosis viruses (ASLV) subgroups A, B, D and E have been identified. They are all member of the tumour necrosis factor (TNF) family (figure 1.7d). To date the receptor for Jaagsiekte sheep retrovirus (JSRV), HYAL-2 is the only glycosylphosphatidylinositol-anchored (GPI-anchored) protein identified for retroviruses. These receptors have an amino terminal endoplasmic reticulum signal sequence and a hydrophobic carboxy terminal end. The hydrophobic end is removed during anchor addition, allowing the tail part of the receptor to anchor the protein to the cell membrane (figure 1.7e) (Overbaugh J. et al., 2001). Eleven receptor groups of retroviruses have been identified on human cells on the basis of receptor interference studies. These receptors have distinct structures and functions in the cell (Weiss R.A. and Tailor C.S., 1995).

1.5.3 Receptors of the γ-type retroviruses

1.5.3.1 MuLV receptors

The first γ-type retrovirus receptor that was identified was the receptor for e-MuLV, which is a cationic amino acid transporter, mCAT-1 (Albritton L.M. et al., 1989). It has been shown previously that the gene encoding mCAT-1 plays a critical role in the haematopoiesis and growth in the development of mice, resulting in death when this gene is non functional (Perkins C.P. et al., 1997). The A-MuLV receptor has been identified as *Pit2*, a sodium dependent inorganic phosphate transporter with 10 transmembrane spanning domains (Miller D.G. et al., 1994). This molecule is closely related to the FeLV-B receptor *Pit1* (discussed in more detail in section 1.5.3); there is 62% amino acid homology and both molecules are expected to have 10 transmembrane spanning domains and 5 extracellular loops (Fan H., 1999). A putative human receptor for X-MuLV has been identified and characterised; X-receptor is a protein with eight transmembrane spanning domains and is widely expressed in both haematopoietic and non- haematopoietic tissues (Battini J-L. et al., 1999; Tailor C.S. et al., 1999a).

Infection studies on Chinese Hamster Ovary cells (CHO cells), a cell line that is not susceptible to infection with both X-MuLV and P-MuLV, showed that when these cells were transfected with the X-receptor, these cells were not only susceptible to X-MuLV infection, but also susceptible to infection with P-MuLV. These findings lead to the conclusion that X-MuLV and P-MuLV are indeed a single family of retroviruses, despite the differences in infectivity of mice. The viruses have co-evolved due to differences in the receptor genes of the host (Tailor C.S. et al., 1999a).

1.5.3.2 RD114 receptors

RD114, human endogenous retrovirus type W, all strains of simian immunosuppressive type D retroviruses and the avian reticuloendotheliosis group (including spleen necrosis virus and baboon endogenous) virus utilise the same receptor, ASCT2 (Rasko J.E.J. et al., 1999; Tailor C.S. et al., 1999b; Marin M. et al., 2000). This receptor is a sodium-dependent neutral-amino-acid transporter that has 10 transmembrane spanning domains (Tailor C.S. et al., 1999b). The majority of these viruses were also capable of utilising the related transporter ASCT1, possibly due to the high degree of conserved sequence in the receptor binding domain (Marin M. et al., 2000; Lavillette D. et al., 2002).

1.5.3.3 PERV receptors

Recently, Ericsson et al (Ericsson T.A. et al., 2003) identified the receptors utilised by PERV-A; PhuR-A1 and PhuR-A2. In keeping with the γ-type retrovirus receptors identified previously, the proteins are thought to have 10 to 11 transmembrane spanning domains, however the exact number of TM domains and the function of the protein remains to be confirmed.

γ-type retrovirus	Receptor	Transmembrane spanning domains	Function of the protein
FeLV-C	FLVCR	12	Organic anion transporter
FeLV-B, GALV, WMSV	*Pit1*	12	Na-dependent inorganic phosphate transporter
FeLV-T	FeLIX and *Pit1*	Soluble (FeLIX)	Env-like protein. Na-dependent inorganic phosphate transporter.
X-MuLV, P-MuLV	X-receptor	8	G-protein-coupled signalling? Inorganic phosphate transporter?
A-MuLV	*Pit2*	10	Na-dependent inorganic phosphate transporter
E-MuLV	mCAT-1	14	Na-independent basic amino acid transporter
RD114	ASCT-2	10	Na-independent neutral amino acid transporter
PERV-A	HuPAR1/ HuPAR2	10-11?	?
FeLV-A	?	?	?

Table 1.3: *γ-type retroviruses and their receptors*

1.5.4 Identification of receptors used by FeLV

A putative receptor on feline cells for FeLV-A has been identified using immuno-precipitation, revealing a 70 kDa molecule (Ghosh S.K. et al., 1992). To date, however, this receptor has not been characterised. However, the receptors for FeLV-B, FeLV-C and FeLV-T have been identified and characterised.

1.5.4.1 FeLV-B receptor

FeLV-B shares its receptor with gibbon ape leukaemia virus (GALV) (Takeuchi Y. et al., 1992). Their receptor, *pit1*, has 10 transmembrane domains and 5 extracellular loops; it is a Na^+-dependent inorganic phosphate transporter. The function of *Pit1* is inhibited when Na^+ is replaced with either K^+ or Li^+. Chimaeric studies with the functional human *pit1* and the non-functional mouse *pit1* have identified the region involved in receptor binding (Johann S.V. et al., 1993). These differences are also found in the characterised *Pit1* from different species. Amino acid sequence differences of this molecule account for the differences in permissivity of human, rat and murine *Pit1* (Weiss R.A. and Tailor C.S., 1995). Amphotropic murine leukaemia virus (A-MuLV) uses the related transporter *Pit2* as its receptor. There is approximately 60% homology between *Pit1* and *Pit2* and both are widely expressed in tissues, although the level of expression varies between tissues. Interestingly, the envelope glycoproteins of FeLV-B and GALV show few similarities, even though they use the same receptor. In contrast, FeLV-B and A-MuLV share a high homology in their envelope glycoproteins (van Zeijl M. et al., 1994). Recent studies by Sugai et al (Sugai J. et al., 2001) tried to clarify the usage of *Pit1* and *Pit2* by the different viruses. Studies on MuLV have shown that the N-terminal domain of the SU is a major determinant for receptor specificity. In this region, there are 2 variable regions (VRA and VRB) and a proline rich region (PRR). The VRA region of FeLV-B is sufficient to be able to use *Pit1* as a receptor, whereas VRB and PRR are secondary determinants for *Pit2* usage. It was shown that an arginine in the VRA region is sufficient to change the usage of *Pit2* by a virus. This amino acid is found in all viruses that are able to use *Pit2* as a receptor.

As there are amino acid sequence differences between the receptor binding sites of the different species, the receptor-binding site is not conserved between species and these differences contribute to control in the tropism of the virus. Both *pit1* and *pit2* are widely expressed *in vivo* and the viruses that use them mainly cause leukaemias, it is thought that other post-entry factors are involved in causing pathogenic effects (Sommerfelt M.A., 1999).

1.5.4.2 FeLV-C receptors

Both the human and the feline homologue of the FeLV-C receptor have been identified and characterised (Tailor C.S. et al., 1999c), (Quigley J.G. et al., 2000). The feline homologue, feFLVCR, is a protein with a molecular mass of 60 kDa; the cDNA encoding the protein has a size of 1.8kb. The protein has twelve membrane spanning domains and shows a high similarity to a hypothetical protein in *Caenorhabditis Elegans* that encodes a putative D-glucarate transporter. It is suggested that the FeLV-C envelope SU protein can act as a dominant negative protein, which inhibits the function of the FLVCR, resulting in PRCA. Consequently, it is thought that FLVCR plays a crucial role in erythopoiesis (Quigley J.G. et al., 2000). The human homologue, huFLVCR, has been identified by Tailor et al (Tailor C.S. et al., 1999c). The cDNA has a predicted size of 2.0 kb. In accordance with feFLVCR, huFLVCR has 12 membrane spanning domains and shows a high similarity to a hypothetical protein in *Caenorhabditis Elegans* that encodes a putative D-glucarate transporter. Northern blot analysis showed that the protein is widely expressed in haematopoietic tissues, most abundantly in foetal liver. Relatively little is expressed in non-haematopoietic tissues. This is consistent with evidence that FeLV-C infects different lineages of haematopoietic cells *in vivo* (Dean G.A. et al., 1992). There is 83% amino acid similarity and 89% protein similarity between the human and feline homologue of FLVCR. The majority of differences are found in the N-terminal domain. The putative D-glucarate transporter is a member of the major facilitator superfamiiy of transporters (MFS), suggesting that the receptor for FeLV-C is a member of the same family. There is further evidence that the FeLV-C receptor is a member of this superfamily as the predicted structure is similar to the structure of members of this family. As FeLV-A is closely related to FeLV-C, it is proposed that FeLV-A use another member of the MFS family as a receptor (Tailor C.S. et al., 1999c).

The MFS is an ancient superfamily with at least 17 distinct branches. There are only slight similarities between the different branches since they have separated shortly after the origin of the superfamily (Pao S.S. et al., 1998). The prediction of the FeLV-A receptor also being related to the FeLV-C receptor is strengthened by the findings described by Bupp et al (Bupp K. and Roth M.J., 2002). Bupp et al describe the screening of a library containing Env variants with random amino acids in the host range determining region of the envelope glycoprotein of FeLV-A, resulting in the identification of a functional variant (EF) with an altered *in vitro* tropism. The host range was distinct from both FeLV-A and FeLV-C and interference assays on the canine cell line D17 indicated that EF utilised a distinct receptor from FeLV-A. It was shown that FeLV-C interfered with both FeLV-A and EF infection on D17 cells, but neither FeLV-A nor EF was capable of interfering with FeLV-C infection (thus this is non-reciprocal interference). These findings indicate that there are two receptors present on D17 cells; FeLV-C is capable of using both receptors, whereas FeLV-A and EF are capable only of utilising one of the molecules present on the cell surface.

1.5.4.3 FeLV-T receptors

FeLV-T appears to be the only γ-type retrovirus that requires an additional factor for successful entry. It uses *Pit1* as primary receptor and the secreted protein FeLIX as a co-receptor (Anderson M.M. et al., 2000). FeLIX is a soluble FeLV-B related glycoprotein that is either expressed by feline T-lymphocytes or by endogenous FeLV glycoproteins (Anderson M.M. et al., 2000; Lauring A.S. et al., 2002).

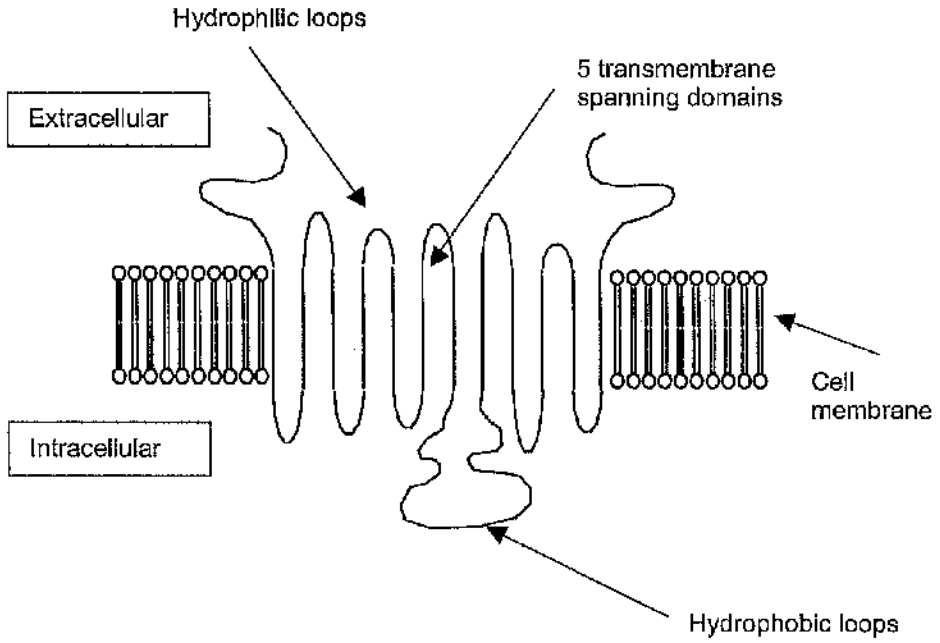


Figure 1.8a: *FeLV-B receptor, Pit1.*

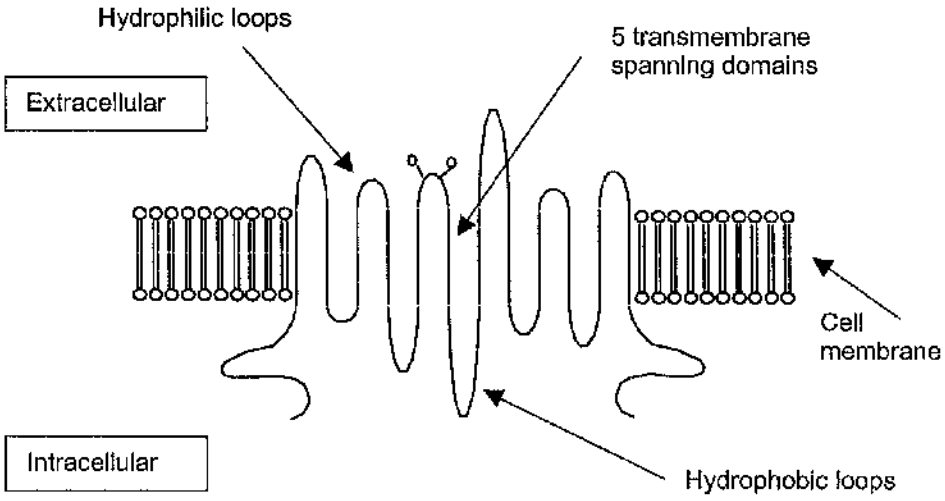


Figure 1.8b: *FeLV-C receptor, FLVCR.*

1.6 Main aims

The aims of the project described in this thesis were threefold; they were:

- To amplify *env* gene products representative of a mix of subgroups within a panel of primary isolates.
- To examine the expression of the feline FeLV-C receptor in a panel of both haematopoietic and non-haematopoietic tissues
- To identify and characterise the FeLV-A cellular receptor.

1.6.1 Amplification of representative *env* gene products

In order to identify amino acid sequence motifs that define the subgroup C phenotype and thus the *in vitro* cell tropism of the *env* genes studied, *env* gene products representative of a mix of subgroups within a panel of primary isolates were isolated and characterised. Previous studies indicated that FeLV-A had the most restricted cell tropism (it was capable only of infecting feline cells), whereas both FeLV-B and FeLV-C were capable of infecting both feline and non-feline cells (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). Thus, an expanded cell tropism would indicate the presence of subgroup C viruses.

1.6.2 Expression of the FeLV-C receptor feFLVCR

The expression of the feline FeLV-C receptor feFLVCR was studied in a panel of both haematopoietic and non-haematopoietic tissues to establish if the detrimental effect FeLV-C infection has is a result of preferential expression of the receptor on cells that are involved in erythroid maturation or if binding of the virus results in inhibition of the normal function of the molecule in the production of erythrocytes.

1.6.3 Identification of the FeLV-A receptor

To date, the FeLV-A receptor has not been identified and characterised. Therefore, attempts were made to identify this receptor using methods that have been used successfully for the identification of, amongst others, the feline FeLV-C receptor FLVCR (Tailor C.S. et al., 1999c).

CHAPTER TWO

METHODS

Methods used throughout the thesis are described in this chapter, whereas methods specific to one area are described in the appropriate chapters. Many of the methods are based on standard techniques and can be found in laboratory manuals such as Current Protocols in molecular biology and Sambrook (Ausubel F.A. et al., 1992) (Sambrook J. et al., 1989).

2.1 GROWTH AND MANIPULATION OF MAMMALIAN CELLS

2.1.1. Basic techniques

All procedures involving manipulation of mammalian cells were performed using standard aseptic procedures and, where possible, were performed in a microbiological safety cabinet, class II.

2.1.1.1 Cryopreservation of cells

In order to preserve stocks of cell lines for long term use, cells were stored in the vapour phase of liquid nitrogen. Cells to be frozen were grown to mid-log phase (as described below) and removed into a sterile 50ml centrifuge tube (using trypsin-EDTA where necessary). The cells were centrifuged at 400xg for five minutes, the supernatant discarded and the cells resuspended in freezing medium to a concentration of approximately 2 x 10^6 cells/ml. The cell suspension was transferred in 1ml aliquots, to labelled cryovials and brought to -130°C in a controlled rate cell freezer (Kryo 10 - Planer Products Ltd., Sunbury on Thames, U.K.). The vials were then transferred to a liquid nitrogen freezer. Cell stocks were revived by rapid thawing in a 37°C waterbath, resuspended in 10% DMEM growth medium to remove the DMSO and subsequently cultured using standard techniques (described below).

2.1.1.2 Cell counting

Cells were counted using a haemocytometer (Weber Scientific International) as follows. Cells were resuspended in the appropriate growth medium and introduced to the haemocytometer chamber and cells were counted under an inverted microscope with 4X or 10X objective. Cells lying on the top and right hand perimeter of each large (one millimetre) square were included; those on the bottom or left hand were excluded. Cell concentration (cells/ml) was calculated by multiplying the mean number of cells per large square by 10^4 and correcting for the dilution factor.

2.1.1.3 Cell growth and maintenance

Cells were maintained in 10% DMEM growth medium and grown at 37°C with 5% CO_2 unless otherwise stated. The growth medium contained 10% foetal calf serum, 2mM glutamine; the antibiotics penicillin and streptomycin, which were added to a final concentration of 100IU/ml and 100µg/ml respectively to prevent contamination. When cells reached 80-90% confluency, which was typically every 3-4 days, cells were subcultured by adding a 1:10 dilution of trypsin-EDTA, spinning the cells at 1000rpm and resuspending the cell pellets in growth medium.

2.1.1.4 Routinely used cell lines

** 293T cell line (HEK293T)*

This is a human embryonic kidney line transformed with adenovirus E1a and carrying a temperature sensitive T antigen to enhance gene expression (Graham F.L. et al., 1977). The T antigen allows for episomal replication of plasmids that contain the SV40 origin and early promotor region. The cells are maintained in 10% DMEM growth medium in the presence of 400µg/ml G418. Prior to setting up cells for transfection assays, the G418 is omitted from the growth medium and cells are detached from the flask by gentle tapping.

** 3T3 cell line (NIH 3T3)*

This is a cell line of Swiss albino mouse fibroblasts that grow as a monolayer with contact inhibition. The cell line was established in 1962 by disaggregating Swiss albino mouse embryos and is described by Todaro et al (Todaro G.J. et al., 1965).

** 104C1 cell line*

104C1 are guinea pig foetal cells that grow in an epithelial like monolayer and were a kind gift from Prof. D.E. Onions (Q1 Biotech, Glasgow, U.K.). Evans et al (Evans C.H. et al., 1975) first described this cell line in 1975.

** HO6T1 cell line*

This cell line is a derivative of the Crandell feline kidney (CrFK) cell line, which was first described by Crandell et al in 1973 (Crandell R.A. et al., 1973) (Crandell R.A. et al., 1973). The CrFK cell line is an epithelial-like cell line derived from the kidney of a female cat. HO6T1 cells were generated by transfecting CrFK cells with the plasmid pHO6T1 carrying an activated Ha-ras oncogene (Spandidos D.A. and Wilkie N.M., 1984).

** HeLa cell line*

The HeLa cell line is a cell line derived from a cervical tumour and cells grow in an epithelial like monolayer. This cell line was established by Gey et al and first described in 1952 (Gey G.O. et al., 1952).

** ST Iowa cell line*

This cell line was originally derived from porcine testes from an eighty to ninety day old foetus and was first described in 1966 by McClurkin et al (McClurkin A.W. and Norman J.O., 1966).

** Mv1Lu cell line*

This cell line was derived from the lungs of several near full term foetuses of the mink, *Mustella vison*, and grows in an epithelial like monolayer. This cell line was first described in 1974 by Henderson et al (Henderson I.C. et al., 1974).

** MDCK cell line*

This cell line is also known as Madin-Darby canine kidney cells and was derived from the kidneys of a normal, healthy female cocker spaniel. The cells grow in a heteroploid epithelial monolayer and are maintained in MEM supplemented with non essential amino acids and 5% foetal calf serum. The cells were first described in 1958 by Madin et al (Madin S.H. and Darbey N.B.Jr., 1958) and characterised in more detail in 1966 by Gaush et al, (Gaush C.R. et al., 1966).

2.1.2 Transfection of mammalian cells

2.1.2.1 Calcium Phosphate transfection method

Cells were transfected using the $CaPO_4$ method as described previously by Gorman (Gorman C.M. et al., 1985) using 24μg of each of the plasmids. The plasmids used for transfections were the FeLV *env* of interest in the retroviral vector VR1012, pHit60 and mfg LacZ. PHit60 contains the *gag* and *pol* of murine leukaemia virus, therefore enabling the production of infectious particles and therefore infection of the target cells. The marker gene mfg LacZ encodes for β-galactosidase. When staining the transfected cells with x-gal, the x-gal will be converted into an insoluble blue compound by the enzyme β-galactosidase. A subsequent plate was transfected with vector only, Hit60 and mfg LacZ to determine negative background. This negative control plate was treated similarly to the plates with the gene of interest.

On day 1, $6x10^6$ 293T cells were plated in 150mm dishes in 16mls 10% DMEM. The following day, chloroquine was added to the cells to a final concentration of 25μM chloroquine to enhance DNA uptake. The DNA to be transfected was made up to 1752μl with distilled water and 248μl 2M $CaCl_2$ was added. 2000μl 2xHBS was added slowly whilst blowing air in the mixture and this DNA mixture was added to the cells. After an 8 to 9hr incubation, the transfection mixture was replaced with fresh 10% DMEM growth medium and cells were incubated overnight at 37°C. On day 3, the cells were washed twice with PBS and fresh growth medium was added. The following day, the supernatant of the transfected 293T cells was removed and filtered using a 0.45micron filter. The supernatants were titrated to determine the titre and subsequently stored at -70°C until further use. The 150mm 293T dishes were stained for transfection efficiency by fixing the cells for 10 minutes in PBS containing 0.5% glutaraldehyde, washed three times with PBS and β-galactosidase activity was determined by incubation with x-gal buffer (5mM potassiumferrocyanide, 5mM potassiumferricyanide, 2mM $MgCl_2$) containing 1mg/ml x-gal (5-bromo-4-chloro-3-indolyl β-D galactopyranoside).

2.1.2.2 Assay to determine virus titre using β-galactosidase

The day prior to collection of supernatants, HO6T1, FEA and 104C1 cells were seeded in 24 well plates at a final concentration of $2*10^4$ cells per well and incubated overnight at 37°C in the presence of 5% CO_2. Five fold dilutions of the filtered supernatants were prepared in 10%DMEM, polybrene was added to a final concentration of 4μg/ml and 900μl of the dilutions was added to a well each of HO6T1, FEA and 104C1. After 72hr incubation, the 24 well plates were fixed for 10 minutes in PBS containing 0.5% glutaraldehyde, washed three times with PBS and β-galactosidase activity was determined by incubation with x-gal buffer (5mM potassiumferrocyanide, 5mM potassiumferricyanide, 2mM $MgCl_2$) containing 1mg/ml x-gal (5-bromo-4-chloro-3-indolyl β-D galactopyranoside) in dimethylformamide. The titre of the supernatants was determined using the following formula:

$$\frac{\text{Number of colonies} * \text{dilution plated}}{\text{ml plated per well}}$$

2.1.2.3 Antibody staining

To establish the efficient uptake of the FeLV envelope glycoprotein constructs, the transfected 293T cells were stained with the anti-gp70 monoclonal antibody 6-15 (Weijer K. et al., 1986). After removal of the supernatant, cells were washed with PBS and subsequently fixed with ice cold methanol for approximately 10 minutes. Cells were scraped off the wells using a cell scraper and transferred to tubes, pelleted for 5 minutes at 1000rpm and washed twice in PBS containing 1% (w/v) BSA and 0.1% (w/v) azide. Subsequently, cell pellets were resuspended in 40µl KW1 (1:100 dilution in PBS/1% BSA/ 0.1% Azide) and incubated for 30 minutes at 4°C. Again, cell pellets were washed twice in PBS containing 1% (w/v) BSA and 0.1% (w/v) azide and subsequently resuspended in 40µl FITC α-mouse (1:50 dilution in PBS/1% BSA/ 0.1% Azide), after which the cells were incubated for 30 minutes at 4°C. After a further two washes in PBS containing 1% (w/v) BSA and 0.1% (w/v) azide, cell pellets were resuspended in a small volume of PBS. Simultaneously, an additional sample of the cells was treated and stained similarly with an antibody against FIV (AE11) to establish the background. Stained cells were examined under a fluorescent microscope (Das Mikroskop, Leica , Solms, Germany) using 40x magnification and ultraviolet light of a wavelength of approximately 250nm. To ensure the presence of cells in the field examined, the cells were also examined using bright field microscopy.

2.2 RECOMBINANT DNA TECHNIQUES

2.2.1 Storage and growth of bacteria

Plasmids were maintained in *E.coli* strain INVαF'. To enable storage of these *E.coli* host strains and of transformants obtained during this work, glycerol stocks were prepared. The desired bacterial culture was streaked onto a 1.5% agar plate (1.5% agar in LB medium). In cases where the bacterial stock contained a plasmid conferring ampicillin resistance the medium was supplemented with 50μg/ml ampicillin and in cases where the bacterial stock contained a plasmid conferring kanamycin resistance the medium was supplemented with 20 μg/ml kanamycin. The plate was incubated overnight at 37°C and the following day single colonies were picked using a sterile toothpick, into 2.5ml LB medium (supplemented as appropriate with kanamycin or ampicillin) in sterile bijoux. The cultures were incubated at 37°C overnight in an orbital incubator. Confirmation that the overnight culture was derived from bacteria containing the desired recombinant plasmid was achieved by DNA extraction and restriction digest (section 2.2.2.4). Glycerol stocks were prepared by the addition of 200μl of glycerol to one millilitre of culture and were stored at -70°C. Bacterial stocks were revived for subsequent work by using a sterile toothpick to scratch the surface of the stock, following which it was grown in the appropriate medium as described above.

2.2.2 Extraction and purification of DNA

Plasmid DNA was isolated using a modification of the alkaline lysis technique described by Birnboim and Doly (Birnboim H.C. and Doly J., 1979).

2.2.2.1 Small Scale Plasmid Preparations

Small-scale plasmid preparations were performed using QIAprep® Spin Miniprep. One and a half millilitres of an overnight culture of the desired transformant was removed to an eppendorf tube and spun at 13.000rpm in a microcentrifuge for 5 minutes to pellet the bacteria. The supernatant was discarded and the bacterial pellet was resuspended in 250µl buffer P1, lysed in 250µl buffer P2 and neutralised in 350µl buffer N3. The lysate was cleared from denatured cellular components by centrifugation at 13.000rpm in a microcentrifuge for 10 minutes. The cleared lysate is applied to the QIAprep spin column and the solution is drawn through the columns by applying a vacuum. The QIAprep spin columns contain sillca-gel membrane to adsorb plasmid DNA in a high salt buffer. The column is washed once with buffer PE and spun once at 13.000rpm in a microcentrifuge for 1 minute to remove all residual buffer PE. To elute the DNA 30µl of the low salt buffer EB is added to the column, incubated for 1 minute and spun at 13.000rpm in a microcentrifuge for 1 minute.

2.2.2.2 Large Scale plasmid Preparations using a Caesium Chloride Gradient

A 200ml overnight culture of *E.coli* containing plasmid DNA was grown in the presence of 20µg/ml kanamycin and incubated overnight at 37°C with shaking. Bacteria were spun down in a Beckman J2-21 centrifuge (JA10 rotor) at 8671xg for 10 minutes and supernatant was removed. The bacterial pellet was resuspended in 5ml 50mM TRIS containing 25% sucrose. 1ml of 20mg/ml lysozyme in 0.25 mM TRIS pH8.0, mixed by inversion and incubated on ice for 10 minutes. 8ml TRITON mix was added quickly, mixed by inversion and incubated on ice for 15 minutes, after which the bacterial debris was pelleted by spinning in a Beckman J2-21 centrifuge (JA20 rotor) at 48384xg for 30 minutes.

The supernatant was removed, caesium chloride was added to a final concentration of 1 g/ml and 100µl ethidium bromide (10mg/ml) was added. The tubes were spun for 18 hours at 60,000rpm in an ultracentrifuge (Ti70 rotor). The plasmid band was removed using a 24-gauge needle and a second gradient spin was loaded. The samples were spun in a benchtop ultracentrifuge for 4 hours at 100,000rpm. Again, the plasmid band was removed using a 24-gauge needle and the ethidium bromide was removed through several extractions with butanol. DNA was then ethanol precipitated by adding 2 volumes of 1xTE (10mM TRIS.Cl, 1mM EDTA pH7.5), 1/10th volume of 3M NaOAc and 2 volumes of ethanol at 4°C for 30 minutes. The nucleic acid concentration was determined as described in section 2.2.3.3.

2.2.2.3 Large Scale plasmid Preparations using QIAGEN's endofree™ plasmid maxi kit

A 100ml overnight culture of *E.coli* containing plasmid DNA was grown in the presence of 20µg/ml kanamycin and incubated overnight at 37°C with shaking. Bacteria were spun down in a Beckman J2-21 centrifuge (JA14 rotor) at 7518xg for 10 minutes and supernatant was removed. The bacterial pellet was resuspended in 10ml resuspension buffer P1 and 10ml lysis buffer P2 was added, mixed gently by inversion and incubated at RT for 5 minutes. 10ml Ice-cold neutralisation buffer P3 was added, mixed gently by inversion and this was added to the resuspended bacteria. The lysate was applied to a QIAfilter cartridge to remove bacterial debris, proteins, genomic DNA and detergent and incubated for 10 minutes after which the cell lysate was filtered into a 50ml tube. To remove endotoxins, 2.5ml buffer ER was added and incubated on ice for 30 minutes. In this time, the QIAGEN-tip 500 was equilibrated by addition of 10ml buffer QBT. After the endotoxin removal, the filtered lysate was applied to the QIAGEN-tip and allowed to enter the resin by gravity flow. The QIAGEN-tip was washed twice with buffer QC after which the DNA was eluted from the QIAGEN-tip with 15ml buffer QN and precipitated with 10.5ml isopropanol. The DNA was precipitated in a Beckman J2-21 centrifuge (JA20 rotor) at 48384xg for 30 minutes and supernatant was removed.

After a wash with endotoxin-free 70%EtOH, the DNA pellet was air-dried briefly and resuspended in 200µl endotoxin-free buffer TE. The nucleic acid concentration was determined as described in section 2.2.3.3.

2.2.3 Determination of nucleic acid concentration by spectrophotometry

The sample was diluted 1:200 by addition of 2.5µl of nucleic acid to 497.5µl of dH_2O. The optical density was measured at 260nm and 280nm, in comparison to a blank containing dH_2O. An OD reading of 1.0 at 260 nm corresponds to an approximate nucleic acid concentration of 50µg/ml for double stranded DNA or 40µg/ml for RNA. The ratio of the OD readings at 260 nm and 280 nm (OD_{260}/OD_{280}) was used to estimate the purity of the nucleic acid. Pure preparations of DNA and RNA have an OD_{260}/OD_{280} of 1.8 and 2.0, respectively; a lower value suggesting possible protein contamination.

2.2.4 Restriction endonuclease digestion

Typically, 0.5-1µg of DNA was digested in a 10µl reaction mix containing 10 units of the desired restriction enzyme and the appropriate buffer. The reactions were incubated at 37°C for a minimum of one hour and separated on an agarose gel as described in section 2.2.3.5. When the isolation of restriction fragments was required, 5-10µg of DNA was digested in a 50µl reaction mix containing 10 units of the desired restriction enzyme and the appropriate buffer. The reactions were then incubated at 37°C for a minimum of one hour and separated on an agarose gel as described in section 2.2.3.5.

2.2.5 Electrophoresis of DNA

DNA fragments of 1.0 - 10 kb were separated and identified by agarose gel electrophoresis using a Bio-rad wide mini sub® cell GT gel kit (Hybaid). All analysis was carried out on agarose gels, which varied from 0.8% to 1.5% depending on the size of the fragment to be visualised. Gels were cast in 1xTBE and 0.5µg/ml EtBr was added to the gel. DNA samples were prepared by adding the appropriate volume of 10X DNA Gel Loading Buffer and loaded into the wells. Gels were run at 90volts in 1xTBE for 60 minutes to 90 minutes. The separated DNA was visualised using a short wave UV transilluminator and photographed using a gel print 2000i (MWG biotech AG, Ebersberg, Germany).

2.2.6 Purification of restriction enzyme fragments

The DNA fragment of interest was purified from an agarose gel using the QIAquick® Gel Extraction Kit that uses a silica-gel membrane. The adsorption of the DNA only takes place in the presence of high salt buffers. The DNA is eluted from the column by the addition of low salt buffer. The DNA was digested with the desired restriction enzyme and separated by electrophoresis on an agarose gel as described previously (sections 2.2.3.4 and 2.2.3.5). The DNA fragment of interest was excised from the gel using a sterile scalpel and placed in an eppendorf tube. Three volumes of binding and solubilization buffer QC was added and the mixture was incubated at 50°C for 10 minutes or until the fragment was dissolved completely. The mixture was then applied to a QIAquick column and washed once with buffer PE to remove impurities such as agarose, restriction enzymes or primers. An additional centrifugation step was included to remove any residual buffer PE as this may interfere with subsequent enzymatic reactions. The purified fragment was eluted with 30µl buffer EB, incubated for 1 minute and spun down at maximum speed for 1 minute. To establish the efficiency of the gel purification, a small aliquot of the sample was run on an agarose gel as described in section 2.2.3.5.

2.2.7 Ligation of vector and target DNA

Equal quantities of vector and insert were mixed with an appropriate volume of ligation buffer and 4 units of DNA ligase in a 10µl reaction. The ligation reaction was incubated at 14°C overnight and stored thereafter at -20°C if not used immediately. A negative control, omitting insert DNA, was generally set up in parallel to check for background when performing the subsequent bacterial transformations.

2.2.8 Transformation of INVαF' cells with plasmid DNA

Cells were thawed on wet ice, mixed gently and 25µl of cells were used per transformation. One microlitre of ligation reaction (containing either the recombinant plasmid or the negative control as described previously) was added to the cells whilst mixing carefully. Cells were incubated on ice for 30 minutes and subsequently heat shocked in a 42°C waterbath for 30 seconds. After cells were placed on ice for 2 minutes, 250µl SOC medium was added and the cells were incubated at 37°C for 1 hour with shaking at 225rpm. Typically, 50µl and 200µl cells were plated on LB plates containing 20µg/ml kanamycin and incubated overnight at 37°C.

2.2.9 Preparation of a [α-^{32}P]-dCTP probe

Invitrogen's RadPrime DNA Labelling System was used to prepare [α-^{32}P]-dCTP labelled DNA probes. Approximately 50ng of the gel-purified fragment was denatured at 100°C for 10 minutes and immediately chilled on ice. Random primers were annealed to the denatured DNA template, extended by the *klenow* fragment in the presence of 50 µCi [α-^{32}P]-dCTP and incubated at 37°C for 10 minutes. The probe was applied to a NICK ™ column (Amersham Pharmacia Biotech) to remove unincorporated [α-^{32}P]-dCTP. The column was washed once with 350µl 1xTE pH8.0. This flow through was discarded. The second flow through of 400µl was collected.

The radioactivity was determined by counting 2µl of the prepared probe in liquid scintillation counting and the total counts/ml was determined by multiplying the counts by 200 to take the dilution into account. The [α-^{32}P]-dCTP probe was stored at -20°C until required. Before use, the probe was denatured at 100°C for 10 minutes and immediately chilled on ice.

2.2.10 Automated sequencing using the LI-COR 4000

The LI-COR model 4000 automated sequencer (LI-COR inc., Lincoln, USA) utilises an infrared detection system, whereby DNA fragments are detected following labelling with IRD41 labelled primers, as they run through the denaturing polyacrylamide gel. In addition, sequencing reactions were performed using a thermostable DNA polymerase (SequiTherm DNA polymerase), allowing reactions to be performed in a thermal cycler. This variation on the original method of chain termination sequencing, known as 'cycle sequencing', was first described by Innis et al (Innis M.A. et al., 1988). This method allows direct sequencing of dsDNA without alkaline denaturation, requires less template and is more efficient at sequencing templates that are G/C rich or have high secondary structure. Custom made IRD41 labelled primers were designed and ordered from MWG Biotech, see figure 2.1 for the position of the primers in the FeLV-A/Glasgow-1 envelope glycoprotein (Accession number P08359) (Stewart M.A. et al., 1986).

Primer name	Primer binding site
FeLV env F69	5'-GTC CAA CGC ACC CAA AAC C-3'
FeLV env F795	5'-ATC AAA AAC CCC CAT CCC-3'
FeLV env F974	5'-ACC CCA ACA AAA CTA AAG AC-3'
FeLV env R1019	5'-GTC GAG AAA CCA GGC AGA GC-3'
FeLV env R1482	5'-GCT TTA GTC CCT GTT CCG ACC-3'
FeLV env F1737	5'-TAA AAC AGC GGC AAC AAC-3'
FeLV env R1858	5'-TAG GAG TAG GAT TAG TAA GGG-3'

Table 2.1: *Primers used for sequencing the envelope glycoprotein.*

1 → *FeLV env F69* 110

ctgcaggacc aaccaccaat caagacctct cggacagccc cagctcagac gatccatcaa gatggaaagt ccaacgcacc caaaaccctc taaagataag actctctcgt

111 220

ggaacttagc gtttctggtg gggatcttat ttacaataga cataggaatg gccaatccta gtccacacca aatatataat gtaacttggg taataaccaa tgtacaaact

221 330

aacacccaag ctaacgccac ctctatgtta ggaaccttaa ccgatgccta ccctacccta catgttgact tatgtgacct agtgggagac acctgggaac ctatagtcct

331 440

aaacccaacc aatgtaaaac acggggcacg ttactcctcc tcaaaatatg gatgtaaaac tacagataga aaaaaacagc aacagacata ccccttttac gtctgccccg

441 550

gacatgcccc ctcgttgggg ccaaagggaa cacattgtgg aggggcacaa gatgggtttt gtgccgcatg gggatgtgag accaccggag aagcttggtg gaagcccacc

551 660

tcctcatggg actatatcac agtaaaaaga gggagtagtc aggacaatag ctgtgaggga aaatgcaacc ccctggtttt gcagttcacc cagaagggaa gacaagcctc

661 770

ttgggacgga cctaagatgt ggggattgcg actataccgt acaggatatg accctatcgc tttattcacg gtgtcccggc aggtatcaac cattacgccg cctcaggcaa

771 → *FeLV env F795* 880

tgggaccaaa cctagtctta cctgatcaaa aacccccatc ccgacaatct caaacagggt ccaaagtggc gacccagagg ccccaaacga atgaaagcgc cccaaggtct

881 → *FeLV env F974* 990

gttgccccca ccaccatggg tcccaaacgg attgggaccg gagataggtt aataaattta gtacaaggga catacctagc cttaaatgcc accgacccca acaaaactaa

991 *FeLV env R 1019* ← 1100

agactgttgg ctctgcctgg tttctcgacc accctattac gaagggattg caatcttagg taactacagc aaccaaacaa acccccccc atcctgccta tctactccgc

1101 1210
aacacaaact aactatatct gaagtatcag ggcaaggaat gtgcataggg actgttccta aaacccacca ggctttgtgc aataagacac aacagggaca tacaggggcg

1211 1320
cactatctag ccgcccccaa cggcacctat tgggcctgta acactggact caccccatgc atttccatgg cggtgctcaa ttggacctct gatttttgtg tcttaatcga

1321 1430
attatggccc agagtgactt accatcaacc cgaatatgtg tacacacatt ttgccaaagc tgtcaggttc cgaagagaac caatatcact aacggttgcc cttatgttgg

1431 *FeLV env R1482 ←* 1540
gaggacttac tgtagggggc atagccgcgg **gggtcggaac agggactaaa gc**cctccttg aaacagccca gttcagacaa ctacaaatgg ccatgcacac agacatccag

1541 1650
gccctagaag aatcaattag tgccttagaa aagtccctga cctccctttc tgaagtagtc ttacaaaaca gacggggcct agatattcta ttcttacaag agggagggct

1651 *→ FeLV env F1737* 1760
ctgtgccgca ttgaaagaag aatgttgctt ctatgcggat cacaccggac tcgtccgaga caatatggcc aaattaagag aaagac**taaa acagcggcaa caac**tgtttg

1761 *FeLV env R1858 ←* 1870
actcccaaca gggatggttt gaaggatggt tcaacaagtc cccctggttt acaaccctaa tttcctccat tatgggc**ccc ttactaatcc tactccta**at tctcctcttc

1871 1980
ggcccatgca tccttaaccg attagtacaa ttcgtaaaag acagaatatc tgtggtacag gctttaattt taacccaaca gtaccaacag ataaagcaat acgatccgga

1981 2090
ccgaccatga tttccaatta aatgtatgat tccatttagt ccccagaaaa agggggaat gaaagacccc ctaccccaaa atttagccag ctactgcagt ggtgtcattt

2091 2200
cacaaggcat ggaaaattac tcaagtatgt tcccatgaga tataaggaag ttagaggcta aaacaggata tctgtggtta agcacctggg ccccggcttg aggccaagaa

2201 2310

cagttaaacc ccggatatag ctgaaacagc agaagtttca aggccgctgc cagcagtctc caggctcccc agttgaccag agttcgacct tccgcctcat ttaaactaac

2311 2420

caatccccac gcctctcgct tctgtgcgcg cgctttctgc tataaaacga gccatcagcc cccaacgggc gcgcaagtct ttgctgagac ttgaccgccc cgggtacccg

2421 2520

tgtacgaata aacctcttgc tgattgcatc tgactcgtgg tctcggtgtt ccgtgggcac ggggtctcat cgccgaggaa gacctagttc gggggtctttca

Figure 2.1: *Position of the sequencing primers relative to the envelope glycoprotein sequence of FeLV-A/Glasgow-1. Arrows indicate the direction of the primers and recognition sites of the primers are underlined. Nucleotide position 1 corresponds with a position of 5920 in the complete published FeLV-A/Glasgow-1 nucleotide sequence.*

2.2.10.1 Sequencing reactions

Protocols for cycle sequencing were provided by LI-COR. The dsDNA was diluted by adding 1μg to 40μl in 1mM TRIS-HCl pH 8.5. The following were combined in a 250μl micro-amp reaction tube: 1μl 5' IRD 800 labelled primer, 2μl of A, T, G, or C long read termination mix (45 μM each of dATP, dCTP, dTTP and 7-deaza-dGTP and either 0.03 mM ddGTP, 0.3 mM ddCTP, 0.45 mM ddATP or 0.45 mM ddTTP as appropriate) and 5μl of diluted dsDNA stock. The tubes were placed in a thermal cycler (Perkin Elmer) and subjected to an initial denaturation step of 95°C for two minutes followed by 30 cycles of 95°C for 30 seconds; 60°C for 30 seconds; 70°C for 30 seconds; terminating in a 4°C soak. The annealing temperature used was dependent on the melting temperature of the primer used; generally the annealing temperature was 2°C below the melting temperature of the primers used. After cycling was complete, four microlitres of stop solution (95% (w/v) formamide, 10 mM EDTA (pH 7.6), 0.1% xylene cyanol and 0.1% bromophenol blue) was added to each of the reaction mixes. Reactions were stored at - 20°C for up to one week prior to gel electrophoresis; whilst exposure to light was avoided.

2.2.10.2 Gel electrophoresis

Sequencing reactions were run on 4% denaturing polyacrylamide gels, as detailed in the manufacturer's instructions. Glass plates (66cm) were used with a 32 well rectangular toothed comb. The plates were cleaned thoroughly before use and 340μl bind-silane was added to the notched glass plate at the area to be in contact with the comb (to aid the formation of wells) and at the bottom of the plate (to aid the formation of a plug). Bind-silane contains 10μl 10% acetic acid and 330μl silane solution (40μl γ-methacryloxy-propyltrimethoxysilane in 10 ml 100% ethanol). After this had dried, the plates were assembled and placed in a casting stand. A gel mix was prepared using 27.7gr urea, 5.3ml millilitres Long Ranger™ polyacrylamide gel mix (AT Biochem.), 7.9ml 10x TBE and dH_2O was added to a final concentration of 65ml. To the gel mix was added 440μl 10% APS. To prepare a plug, 3ml of the gel mix were removed and 40μl 10% APS plus 4μl TEMED was added to the mix.

A pasteur pipette was used to inject the mix between the bottom of the glass plates to form a plug. After polymerisation of the plug, 40µl TEMED was added to the remainder of the gel solution, which was mixed and the gel was poured using a 50ml syringe. The comb was inserted and the gel was left to polymerise for approximately 2hrs at an angle. The gel was then transferred to the electrophoresis apparatus, both buffer tanks were filled with 1x TBE, the comb was removed and wells were flushed. The gel was pre-electrophoresed for 30 - 45 minutes, during which time the scanning microscope was focussed and the gain controls were adjusted. Sequencing reactions were denatured at 95°C for three minutes prior to loading; generally 1.5 µl was loaded per well. The gel was run at 45.0 W, with the scanner programmed to collect 32 frames (approximately 1000 bp). Sequence data was read automatically; base ambiguities were checked by visual inspection of the gel image.

2.2.11 Sequencing using the ABI PRISM 3100 genetic analyser

During the later stages of this project an ABI PRISM 3100 genetic analyser (Applied biosystems, Warrington, U.K.) became available, which considerably increased the ease and throughput of sequencing. The kit used (ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit version 2.0) to generate the PCR products required for the ABI PRISM 3100 genetic analyser contains AmpliTaq ® DNA Polymerase BigDye terminators and all required components for the sequencing reactions. The BigDye terminators are labelled with high-density dyes, allowing the laser in the ABI PRISM 3100 genetic analyser to determine the correct sequence. Primers were designed and ordered as previously, see figure 2.1 for primer binding sites of the primers used.

2.2.11.1 Sequencing reactions

dsDNA to be sequenced was prepared as described previously (section 2.2.3.2.1) and the dsDNA was eluted in 50µl elution buffer. The PCR reactions contained 4µl dsDNA, 4µl ABI Prism® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit, version 2.0, 2µl 5x sequencing buffer, 3.2pmol primer and dH_2O was added to a final volume of 20µl. The tubes were placed in a thermal cycler (Perkin Elmer) and subjected to 25 cycles of 96°C for 10 seconds; 50°C for 5 seconds; 60°C for 4 minutes; terminating in a 4°C soak. The PCR reactions were purified using Performa DTR gel filtration system columns (EDGE biosystems, U.K.) as described in the product protocol. Spinning at 750*g for 2 minutes rehydrated the columns and the PCR reactions were added to the middle of the column and subsequently spun at 750*g for 2 minutes. The purified reactions were dried in a vacuum drier for approximately an hour and resuspended in 25µl HiDi formamide (Applied biosystems, Warrington, U.K.), a 96 well plate (ABgene, U.K.) was loaded and stored at 4°C until required.

2.2.11.2 Sample collection and sequence analysis

The analysis of the sequencing reactions was carried out as described by the manufacturers. A septa (Applied biosystems, Warrington, U.K.) was placed on the 96 well plate and this was inserted in the box. Conditions were chosen depending on the polymer used and the length of the capillary used and were as described by the manufacturers. To obtain approximately 800bp of reliable sequence, 80cm capillaries were used with the polymer POP6, again as described by the manufacturers.

2.2.12 Analysis of sequences obtained from either the LI-COR 4000 or the ABI PRISM 3100 genetic analyser

The sequences obtained were analysed using the NCBI blast nucleotide search on their website (http://www.ncbi.nlm.nih.gov/BLAST/) to ensure the isolates were indeed FeLV-related. Ambiguities were checked and corrected where possible and the sequences were further analysed using the following software: Wisconsin Package Version 10.0, Genetics Computer Group (GCG) (Madison, Wisc.). The sequences were translated into protein sequences and a pileup analysis was done to compare the sequences to published sequences.

2.2.13 Envelope glycoprotein structure predictions

FeLV *env* structures were generated using "Swiss model" (www.expasy.org, hosted by the Swiss Institute of Bioinformatics, Basel, Switzerland). Envelope glycoprotein structure predictions were prepared using PDB-viewer (www.expasy.org, as above). The FeLV *env* structures were based on the Friend MuLV gp70 (Fass D. et al., 1997), (Accession number AAA46478, accession code in the Protein Data Bank (PDB): 1AOL).

CHAPTER THREE

CLONING AND SEQUENCING OF NOVEL ISOLATES OF FELV-C

3.1 Introduction

FeLV-A is present in all naturally occurring isolates of FeLV, FeLV-B is present in approximately 40% of naturally occurring isolates and FeLV-C is the subgroup that is isolated least frequently and it is present in approximately 1% of isolates (Jarrett O. et al., 1984). FeLV-C is the subgroup that causes a fatal pure red cell aplasia (PRCA) in the cat and it has been shown previously (Hoover E.A. et al., 1974) that mutations in a single region of the envelope glycoprotein of FeLV-A reproduce the anaemogenic phenotype of FeLV-C, suggesting that the interaction between the virus and the receptor underlies the development of PRCA.

A panel of primary isolates of FeLV, some of which have been described previously by Brojatsch et al (Brojatsch J. et al., 1992) and Rigby et al (Rigby M.A. et al., 1992), were examined in this study. These primary isolates of FeLV (see table 3.1 for a description and origin) were classified previously into subgroups A, B or C by interference assays [described by Sarma et al (Sarma P.S. and Log T., 1973)] and contained either an A/C mixture or and A/B/C mixture. Three of these isolates (FA27, FS246 and FZ215) have been partially characterised previously. However discrepancies were found between the sequences published in both studies. Furthermore, the envelope glycoprotein genes were only sequenced partially in the studies described previously.

Therefore, the aim of this study was to identify and characterise *env* gene products representative of the mixture of subgroups present in the primary isolates in order to identify amino acid sequence motifs that define the subgroup C phenotype and thus the *in vitro* cell tropism of the *env* genes studied. Previous studies indicated that FeLV-A had the most restricted cell tropism of the FeLV subgroups and that is was capable of infecting feline cells only (Jarrett O. et al., 1973). In contrast, FeLV-B and FeLV-C were found to have a broader tropism and were capable of infecting both feline and non-feline cells (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). Thus, an expanded cell tropism would indicate the presence of subgroup C viruses.

Isolate	**Description of isolate**
FA27	*Isolated by limiting dilution from an A/C complex obtained from an anaemic cat in Great Britain. Passaged in FEA cells and anaemogenic in neonatal cats. (Onions D.E. et al., 1982)*
FY981	*Isolated by limiting dilution from an A/B/C complex obtained from an anaemic cat in Great Britain in 1980. Passaged in FEA cells and anaemogenic in neonatal cats.*
FS246	*Isolated by limiting dilution from an A/C complex obtained from an anaemic cat in Scotland in 1977. Passaged in FEA cells and anaemogenic in neonatal cats. (Onions D.E. et al., 1982)*
FA621	*Isolated by limiting dilution from an A/C complex obtained from an anaemic cat in Scotland and passaged in FEA cells.*
FX343	*Isolated by limiting dilution from an A/C complex obtained from an anaemic cat in Scotland and passaged in FEA cells.*
FZ215	*Isolated by limiting dilution from an A/B/C complex obtained from an anaemic cat in Great Britain in 1979. Passaged in FEA cells and anaemogenic in neonatal cats.*
L3728F	*Isolated by limiting dilution from an A/C complex obtained from an anaemic cat in Scotland in 1994 and passaged in FEA cells.*
L3950F	*Isolated by limiting dilution from an A/C complex obtained from an anaemic cat in Scotland in 1994 and passaged in FEA cells.*

Table 3.1: *Source and description of primary FeLV isolates.*

3.2 Methods

3.2.1 Expansion of primary isolates

Isolates that were classified previously as subgroups A/C or A/B/C were used to infect the non-restrictive cell line FEA (FEA cells do not select for the A, B or C subgroups). The day prior to infection, FEA cells were seeded in 12 ½ cm^2 flasks at a concentration of $1x10^5$ cells per flask. The following day, supernatants were collected, clarified by centrifugation and 2ml of SN was added to each flask of FEA cells and incubated overnight at 37°C. The remainder of the supernatant was stored at -70°C until required. The following day, FEA cells were re-infected with the frozen supernatant and incubated for a further 4 hours, after which they were washed once with PBS and fresh 10% DMEM was added. Supernatant samples of the FEA-infected cells were collected and stored at -70°C at regular intervals. These supernatants were assayed for FeLV p27 by ELISA (FeLV p27 ELISA IDEXX Laboratories Limited, Buckinghamshire, U.K). If cultures were positive by ELISA, cells were expanded to prepare DNA (as described previously in section 3.2.2) and to enable storage in the vapour phase of liquid nitrogen (as described previously in section 2.2.1.2.1).

3.2.2 Preparation of High Molecular Weight DNA from cell pellets

DNA was prepared from p27-positive cell cultures by the method of Boom et al [(Boom R. et al., 1990), (QIAamp® DNA blood mini kit, QIAGEN)]. To prepare DNA, 10^7 cells were resuspended in 200µl PBS and added to a sterile 1.5ml microcentrifuge tube containing 20µl protease. The cells were lysed by adding 200µl buffer AL and vortexed to ensure efficient mixing and thus lysing of the cells. Buffer AL contained 25µg/ml RNAse A to ensure RNA-free DNA. The mixture was incubated at 56°C for 10 minutes before spinning briefly to remove drops from the inside of the lid. An equal volume of 100% ethanol was added before the lysate was applied to a spin column and was allowed to enter the silica gel resin.

After binding of the DNA, the column was washed with 750µl buffer AW1 and 750µl buffer AW2 subsequently to remove residual debris. The column was then centrifuged at 13000 rpm to remove all traces of wash buffer and DNA was eluted by adding 200µl buffer AE to the column. The columns were incubated for 5 minutes at room temperature and subsequent centrifugation at 13000 rpm to release the DNA from the columns.

3.2.3 Amplification of envelope glycoprotein gene by PCR

Envelope glycoprotein genes were amplified from total cellular genomic DNA by using the polymerase chain reaction (PCR). Oligonucleotide primers were synthesised corresponding to the 5' start codon and 3' stop codon and incorporating *SalI* and *NotI* restriction sites, the oligonucleotide primers are shown in figure 3.1. A Kozak-sequence was included in the sequence of the 5' primer to improve expression levels as described previously by Kozak (Kozak M., 1987). In each PCR reaction, the final concentrations of the components were the following: 1x buffer 2 (Roche applied science, East Sussex, U.K.), 1 mM dNTP (Promega, U.K.), 37.5 pmol of each primer (MWG, U.K.) and 50U of enzyme 1 (Roche applied science, East Sussex, U.K.). The amplification was performed under the following conditions: denaturation at 94°C for 3 minutes, followed by 30 cycles of denaturation at 94°C for 30secs, annealing at 50°C for 1 minute and extension at 72°C for 3 minutes after which there was an additional 10 minutes extension step at 72°C. The PCR gave a 1.8kb product, as predicted from the FeLV-A sequence and the samples were visualised by agarose gel electrophoresis as described previously (section 2.2.3.5).

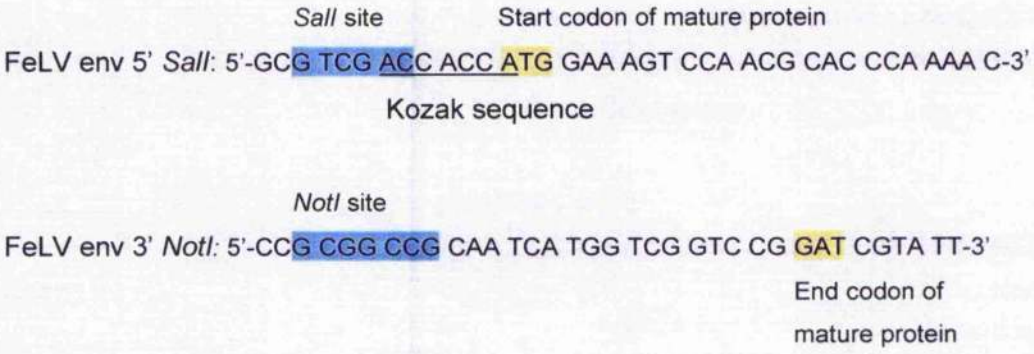


Figure 3.1: *Structure of primers used to amplify the envelope glycoprotein of FeLV.*

3.2.4 Molecular cloning of FeLV *env* gene products

The PCR products were ligated into the pCR1 vector (Invitrogen NV, Leek, The Netherlands) using the TA cloning system as described previously (section 2.2.3.7) and transformed into *E.coli* INVαF' competent cells (section 2.2.3.8) as described previously. Colonies were picked, inoculated in 2.5ml L-broth containing 50μg/ml ampicillin and grown overnight at 37°C. DNA was prepared using the "small-scale plasmid preparation" method (section 2.2.3.2.1) and the plasmid DNA was digested using 10 units of the restriction enzyme *EcoRI* to establish the presence of an insert as described previously (section 2.2.3.5).

3.2.5 Molecular cloning into the vector VR1012

Prior to ligation into the eukaryotic expression vector VR1012, the positive "miniprep" DNA's were digested in a 50μl reaction mix containing 25 units of the restriction enzyme *SalI* and 25 units of the restriction enzyme *NotI*. The reactions were incubated at 37°C for a minimum of one hour and subsequently separated by agarose gel electrophoresis as described previously (section 2.2.3.6). Again, the fragment of interest was excised from the agarose gel and purified as described previously.

Simultaneously, 5μg of the eukaryotic expression vector VR1012 was digested and purified. Insert and vector were then ligated and transformed, after which small-scale plasmid preparations were performed to identify positive colonies.

3.2.6 Nucleic acid sequencing of the novel *env* genes

The cloned envelope gene products were sequenced using either a LICOR model 4000 automated sequencer (section 2.2.3.10) or an ABI PRISM 3100 genetic analyser (section 2.2.3.11) using a panel of primers based on the published sequence of FeLV-A/Glasgow-1 (accession number P08359, (Stewart M.A. et al., 1986)). Primer design utilised the Omiga 2.0 software package (Oxford Molecular Ltd.), the primer binding sites are described in table 2.1 and figure 2.1. The sequences obtained were analysed using Wisconsin Package Version 10.0, Genetics Computer Group (Madison, Wisconsin) as described previously in section 3.2.5. Each nucleic acid sequence was translated using the "Translate" component of the GCG software package and the predicted amino acid sequence examined for integrity of the *env* open reading frame.

The novel *env* sequences were compared to the published sequences of FeLV-A/Glasgow-1 (Accession number P08359, (Stewart M.A. et al., 1986)), FeLV-B Gardner-Arnstein (Accession number P03391, (Wunsch M. et al., 1983)) or FeLV-C-Sarma (Accession number M14331, (Riedel N.O. et al., 1986)) using the "Bestfit" component of the GCG software package according to the local homology algorithm of Smith et al (Smith T.F. et al., 1985). Both nucleotide and amino acid alignments were performed using the "Pileup" component of the GCG software package according to the progressive alignment method described by Feng and Doolittle (Feng D.F. and Doolittle R.F., 1987).

3.2.7 Preparation of FeLV/MuLV pseudotypes in 293T cells by transient transfection

Cells were transfected using the $CaPO_4$ method as described previously by Gorman (Gorman C.M. et al., 1985). On day 1, $3x10^6$ 293T cells were plated in 100mm dishes in 10mls 10% DMEM. On day 2, chloroquine was added to the cells to a final concentration of 25μM (the pH is increased by this endosomalytic agent, which in turn inhibits lysosomal hydrolytic enzymes and thus decreases the risk of degradation of endocytosed material (Luthman H. and Magnusson G., 1983)), 9μg of the FeLV *env* of interest, 9μg pHit60 (which contains MuLV *gag-pol* to enable formation of FeLV *env*/MuLV *gag-pol* pseudotypes (Soneoka Y. et al., 1995)) and 9μg mfg LacZ, (carrying the β-galactosidase marker gene (Dranoff G. et al., 1993)) were combined and the total volume made up to 1314μl with distilled water. 168μl of 2M $CaCl_2$ was then added without mixing. 1500μl 2xHBS was added slowly and the DNA mixture was added to the cells. After a 9 hour incubation, the transfection mixture was replaced with fresh 10%DMEM and the cells were then incubated overnight at 37°C. On day 3, the cells were washed twice with PBS and fresh growth medium was added. The following day, the supernatant of the transfected 293T cells was removed and filtered using a 0.4micron filter. The filtered viral pseudotypes were titred on HO6T1 cells and then used to infect a range of cell lines in order to establish the tropism of the novel envelope glycoprotein. In parallel with the viral pseudotype studies, 293T cells transfected with each VR1012-*env* construct were fixed using ice-cold methanol and stained with the 6-15 monoclonal antibody (anti-gp70) first described in 1986 by Weijer et al (Weijer K. et al., 1986). Expression of mature envelope glycoprotein was then evaluated by immunofluorescence as described previously (section 2.2.1.2.2.).

3.3 Results

3.3.1 Isolation, PCR amplification and characterisation of novel envelope glycoprotein genes (*envs*)

Partial nucleic acid sequences of the envelope glycoproteins of FA27, FZ215 and FS246 have been described previously by Brojatsch et al (Brojatsch J. et al., 1992) and Rigby et al (Rigby M.A. et al., 1992). However, the sequences of the envelope glycoprotein genes of the three isolates did not cover the entire SU region and discrepancies were found between the sequences published by Brojatsch et al and Rigby et al. In order to identify a representative sequence of each primary isolate covering the entire SU and TM regions and to clarify the discrepancies described previously, a representative *env* clone of each primary isolate was amplified, cloned and sequenced. The *env* clones were then translated and the predicted amino acid sequence was examined for integrity of the *env* open reading frame. Amino acid alignments were performed on these sequences and are shown in figure 3.2.

FA27

Differences were apparent when comparing the three envelope glycoprotein sequences of primary isolate FA27. Three changes clustered in the VRA region and altered the polarity of the envelope glycoprotein. These substitutions were $R_{69}W$, $S_{72}R$ and $A_{84}T$. Outwith the VRA region, the sequence published by Rigby et al contains an additional aspartate residue and a glycine residue has been deleted (D_{114} and G_{122} respectively). When compared with both the sequence published by Brojatsch et al and the sequence described in this study, the alterations in the sequence described by Rigby et al result in an isolate that is slightly more polar than the isolates described by Brojatsch et al in this study. However, as only part of the envelope glycoprotein was sequenced by Rigby et al and Brojatsch et al, it was conceivable that mutations further on in the envelope glycoprotein may counteract the slight change in polarity and subsequently the infectivity of this isolate.

FS246

The amino acid sequences described previously for FS246 and the amino acid sequence characteristic of the FS246 isolate in this study were very similar. The overall polarity of this region of the envelope glycoprotein differed between the sequence published by Rigby et al and the sequence published by Brojatsch et al and the sequence described in this study as there was a $K_{87}E$ substitution resulting in a change from a positively charged residue to a negatively charged residue and a $H_{101}Y$ substitution which resulted in a change from a positively charged residue to a neutral residue. Therefore, the overall polarity predicted for the *env* described by Rigby et al was slightly lower than for the two other *env* sequences. As described for isolate FA27, it is conceivable that mutations further on in the envelope glycoprotein counteract this change in polarity in order that the isolate remains capable of infection. The addition and deletion of two amino acids in the FA27 sequence published by Rigby et al can also be found in the FS246 sequence described in the same paper, these are an additional aspartate residue and a deleted glycine residue, as is described for isolate FA27 (D_{114} and G_{122} respectively).

FZ215

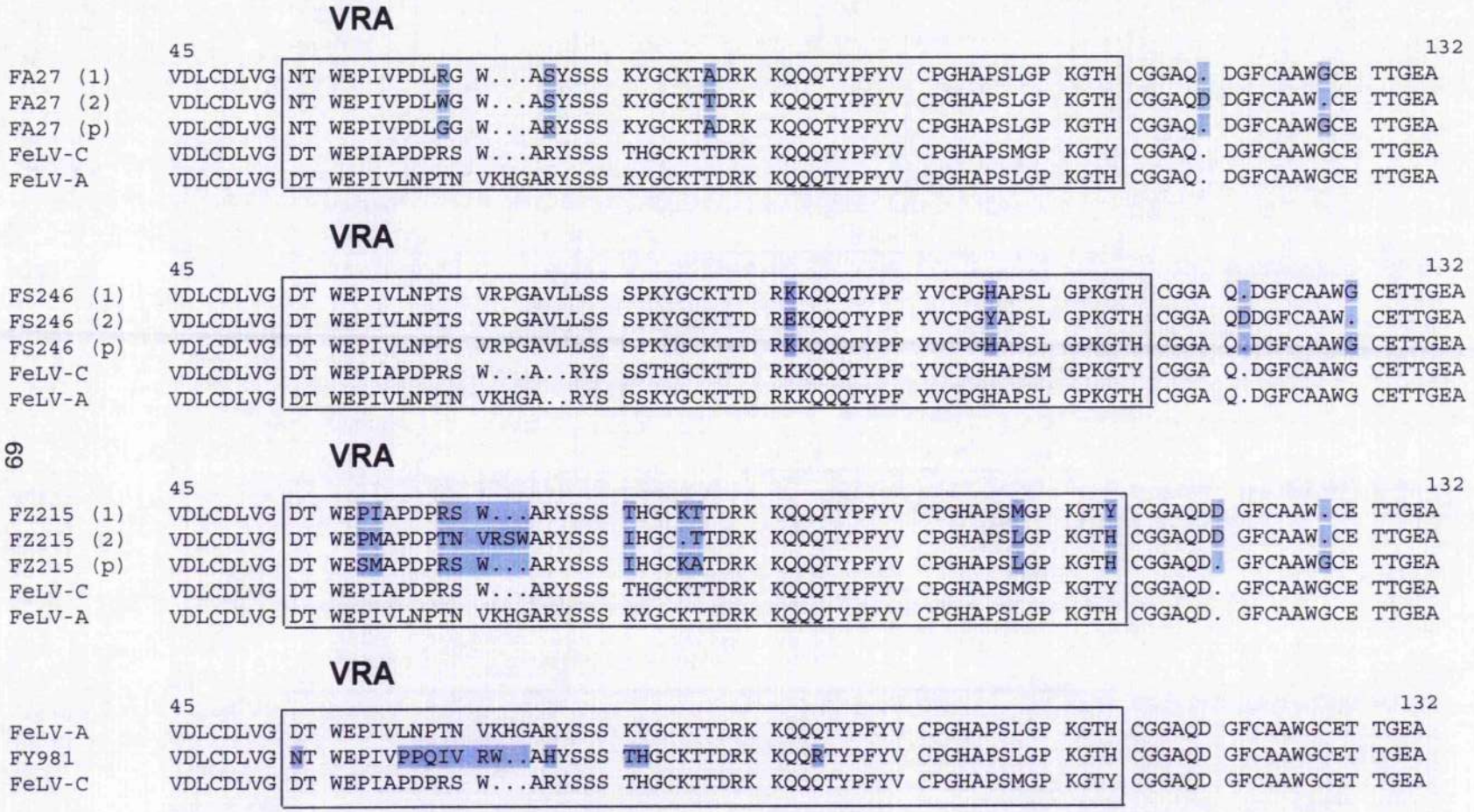
The amino acid sequence of the FZ215 envelope glycoprotein described by Brojatsch et al is virtually identical to the published FeLV-C/Sarma sequence, bar one amino acid difference ($L_{298}S$) outwith the region shown in figure 3.2. Therefore, Brojatsch concluded that this isolate was not an independent isolate, but rather both this isolate and the FeLV-C/Sarma were derived from the same virus isolate. The *env* clone isolated in this study for the FZ215 isolate was very similar to the FZ215 isolate described by Brojatsch et al, however when compared with the published amino acid sequence of FeLV-C/Sarma, there were more differences in the sequence described in this thesis, therefore it was conceivable that this *env* clone was indeed an independent isolate and not derived from the same isolate as FeLV-C/Sarma. The sequence published by Rigby et al is significantly different from the two other FZ215 isolates, having gained a stretch of three amino acids (^{66}RSW^{68}) at the start of the VRA region, this stretch is very similar to the sequence found in FeLV-A/Glasgow-1.

FY981

In order to compare the unique envelope glycoproteins identified in this study, an *env* clone of the primary isolate FY981 was amplified, cloned and subsequently sequenced. This *env* clone was considered characteristic of the FY981 isolate. An amino acid alignment of this representative FY981 clone, FeLV-A/Glasgow-1 and FeLV-C/Sarma is shown in figure 3.2. There were few differences in the VRA region of this FY981 *env* clone, when compared with FeLV-A/Glasgow-1, however the VRA region of this representative *env* clone was shorter than the VRA region of FeLV-A/Glasgow-1, indicating it was a C component of this isolate.

Figure 3.2: *Amino acid alignments of the primary isolates characterised by Brojatsch et al and Rigby et al. The amino acid sequences numbered (1) have been published previously by Brojatsch et al (Brojatsch J. et al., 1992), the amino acid sequences numbered (2) have been published previously by Rigby et al (Rigby M.A. et al., 1992) and the amino acid sequences numbered (p) are the amino acids that were considered characteristic for the three primary isolates compared in this amino acid alignment. The amino acid sequences of both FeLV-A/Glasgow-1 and FeLV-C/Sarma were included in the amino acids alignment and have been published previously.*

(Opposite page)


VRA
45
132
FA27 (1) VDLCDLVG NT WEPIVPDLRG W...ASYSSS KYGCKTADRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQ. DGFCAAWGCE TTGEA
FA27 (2) VDLCDLVG NT WEPIVPDLWG W...ASYSSS KYGCKTTDRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQD DGFCAAW.CE TTGEA
FA27 (p) VDLCDLVG NT WEPIVPDLGG W...ARYSSS KYGCKTADRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQ. DGFCAAWGCE TTGEA
FeLV-C VDLCDLVG DT WEPIAPDPRS W...ARYSSS THGCKTTDRK KQQQTYPFYV CPGHAPSMGP KGTY CGGAQ. DGFCAAWGCE TTGEA
FeLV-A VDLCDLVG DT WEPIVLNPTN VKHGARYSSS KYGCKTTDRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQ. DGFCAAWGCE TTGEA
VRA
45
132
FS246 (1) VDLCDLVG DT WEPIVLNPTS VRPGAVLLSS SPKYGCKTTD RKKQQQTYPF YVCPGHAPSL GPKGTH CGGA Q.DGFCAAWG CETTGEA
FS246 (2) VDLCDLVG DT WEPIVLNPTS VRPGAVLLSS SPKYGCKTTD REKQQQTYPF YVCPGYAPSL GPKGTH CGGA QDDGFCAAW. CETTGEA
FS246 (p) VDLCDLVG DT WEPIVLNPTS VRPGAVLLSS SPKYGCKTTD RKKQQQTYPF YVCPGHAPSL GPKGTH CGGA Q.DGFCAAWG CETTGEA
FeLV-C VDLCDLVG DT WEPIAPDPRS W...A..RYS SSTHGCKTTD RKKQQQTYPF YVCPGHAPSM GPKGTY CGGA Q.DGFCAAWG CETTGEA
FeLV-A VDLCDLVG DT WEPIVLNPTN VKHGA..RYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL GPKGTH CGGA Q.DGFCAAWG CETTGEA
VRA
45
132
FZ215 (1) VDLCDLVG DT WEPIAPDPRS W...ARYSSS THGCKTTDRK KQQQTYPFYV CPGHAPSMGP KGTY CGGAQDD GFCAAW.CE TTGEA
FZ215 (2) VDLCDLVG DT WEPMAPDPTN VRSWARYSSS IHGC.TTDRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQDD GFCAAW.CE TTGEA
FZ215 (p) VDLCDLVG DT WESMAPDPRS W...ARYSSS IHGCKATDRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQD. GFCAAWGCE TTGEA
FeLV-C VDLCDLVG DT WEPIAPDPRS W...ARYSSS THGCKTTDRK KQQQTYPFYV CPGHAPSMGP KGTY CGGAQD. GFCAAWGCE TTGEA
FeLV-A VDLCDLVG DT WEPIVLNPTN VKHGARYSSS KYGCKTTDRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQD. GFCAAWGCE TTGEA
VRA
45
132
FeLV-A VDLCDLVG DT WEPIVLNPTN VKHGARYSSS KYGCKTTDRK KQQQTYPFYV CPGHAPSLGP KGTH CGGAQD GFCAAWGCET TGEA
FY981 VDLCDLVG NT WEPIVPPQIV RW..AHYSSS THGCKTTDRK KQQRTYPFYV CPGHAPSLGP KGTH CGGAQD GFCAAWGCET TGEA
FeLV-C VDLCDLVG DT WEPIAPDPRS W...ARYSSS THGCKTTDRK KQQQTYPFYV CPGHAPSMGP KGTY CGGAQD GFCAAWGCET TGEA

At present, three sequences that represent FeLV-A have been published; they are FeLV-A/Glasgow-1 (accession number P08359, (Stewart M.A. et al., 1986)), 61E (accession number M18247, (Donahue P.R. et al., 1988)) and FeLV strain Rickard, subgroup A (accession number AF052723, (Chen H. et al., 1998)) The published FeLV-A sequences show a high degree of homology in the envelope glycoprotein and have a VRA region of a similar length. Therefore, two criteria were used to define a unique *env* clone as the C-component of the primary isolate. The criteria were, firstly, a high degree of sequence divergence in the VRA region of the *env* clone when compared with FeLV-A/Glasgow-1 and secondly, the presence of length polymorphisms in the VRA region of the envelope glycoprotein. Conversely, sequence conservation and the absence of length polymorphisms in the VRA region of the *env* clones when compared with FeLV-A/Glasgow-1 indicated that the clone constituted the "A component" of the primary isolate. As FeLV-B/Gardner Arnstein is highly divergent from both FeLV-A/Glasgow-1 and FeLV-C/Sarma, subgroup B components were identified based on similarity to FeLV-B/Gardner Arnstein.

After small scale plasmid screening, thirty four small scale plasmid preps contained an insert of the expected size and thus were sequenced using the FeLV *env* primer F69, the primer binding site of which is described in figure 2.1 and table 2.1. This primer binds prior to the start of the mature protein and therefore provides the sequence of the VRA region, enabling the identification of novel *env* clones. In order to classify each *env* clone as either the A-component or the C-component of the primary isolate, the predicted amino acid sequence was examined. Similarly, the integrity of the *env* open reading frame was confirmed in order to determine whether the *env* gene had the potential to encode a viable Env. The sequences were compared with both FeLV-A/Glasgow-1 and FeLV-C/Sarma and changes in the amino acid sequence relative to the published FeLV-A/Glasgow-1 sequence were identified. All of the unique envelope glycoproteins isolated had intact open reading frames. Table 3.2 lists the similarities found when comparing the *env* clones with both FeLV-A/Glasgow-1 and FeLV-C/Sarma and figure 3.3 shows the amino acid alignment. Amino acid 1 in the amino acid sequence compares to amino acid 2026 in the published FeLV-A genome.

Env clone #	% Similarity to VRA region of FeLV-A	% Similarity to VRA region of FeLV-C	Subgroup predicted by sequencing	Further characterisation
FA27-17	97.4%	94.7%	C	Yes
FA27-19	85.5%	79.9%	A	No
FA27-53	99.0%	94.4%	A	Yes
FA27-55	98.7%	94.7%	C	Yes
FA621-15	77.6%	72.2%	B (98.1%)	No
FA621-17	98.2%	92.4%	A	No
FA621-42	77.8%	79.6%	B (96.3%)	No
FA621-43	98.8%	90.4%	A	No
FA621-44	67.1%	97.8%	B (89.0%)	No
FS246-1	99.4%	92.5%	A	No
FS246-4	98.7%	93.7%	C	Yes
FS246-37	98.9%	93.9%	C	Yes
FS246-38	89.6%	82.9%	A	No
FS246-39	87.1%	82.1%	A	No
FS246-40	99.3%	94.2%	A	Yes
FX343-26	99.0%	93.9%	A	Yes
FX343-28	98.2%	93.6%	A	Yes
FY981-9	87.4%	80.7%	C (variant I)	No
FY981-10	97.4%	93.2%	C (variant II)	Yes
FY981-12	94.1%	88.9%	C (variant I)	No
FY981-13	95.9%	90.7%	C (variant I)	No
FY981-14	97.9%	93.7%	C (variant I)	Yes
FY981-22	94.8%	90.4%	C (variant II)	No
FY981-24	95.9%	89.8%	C (variant I)	No
FZ215-9	98.2%	93.4%	A	Yes
FZ215-10	39.4%	55.4%	B (61.5%)	No
FZ215-11	71.9%	71.9%	B (92.3%)	No
FZ215-12	77.0%	89.1%	B (98.9%)	Yes
FZ215-46	66.8%	65.3%	B (91.5%)	No
L3128F-49	96.8%	91.1%	A	No
L3128F-50	97.8%	92.6%	A	No
L3950F-32	82.3%	76.7%	A	No
L3950F-52	85.7%	80.8%	C	No
L3950F-66	99.6%	92.6%	A	No

As shown in figure 3.3, the *env* genes that were classified as the A-components of the primary isolates display a high degree of similarity to the published FeLV-A/Glasgow-1 sequence. With the exception of the *env* clones derived from primary isolate FS246, in which the VRA region is three amino acids longer than the VRA region of FeLV-A/Glasgow-1, the *env* clones that appear to be C-components of the primary isolates all contain a VRA region that is shorter than the FeLV-A/Glasgow-1 VRA region. This shows that the criteria used to classify the unique envelope glycoproteins are valid.

FA27

Isolate FA27, an A/C mixture in the original isolate, yielded four *env* gene clones. Both FA27-17 and FA27-55 appeared to be the C components of the mixture, whereas *env* clone FA27-53 was thought to be the A component. The fourth *env* clone, FA27-19, was identical to FA27-53 and therefore was not studied further.

FS246

Isolate FS246 was an A/C mixture of which six *env* genes were identified; three of the *env* clones identified were the A-component (FS246-1, FS246-38 and FS246-40), three of the *env* clones identified appeared to be the C component (FS246-4, FS246-37 and FS246-39). The three *env* clones thought to be the A component were identical, therefore only *env* clone FS246-40 was characterised further. Of the C component *env* clones, FS246-39 was identical to FS246-4 and therefore not characterised further.

FY981

Although the original FY981 isolate contained an A/B/C mixture, neither the A component nor the B component were isolated following expansion on FEA cells, suggesting that the C component of this isolate was relatively abundant. In total, seven envelope clones were isolated, five of which (FY981-9, FY981-12, FY981-13, FY981-14 and FY981-24 were identical to the representative FY981 sequence and were termed "variant I". Clones FY981-10 and FY981-22 had two amino acid differences ($P_{74}S$ and $P_{76}S$ respectively when compared with the representative FY981 amino acid sequence and were termed "variant II". Thus, only two FY981 *env* clones were characterised further, they were FY981-10 and FY981-14, representative of variant II and variant I respectively.

FX343

Isolate FX343 again was an A/C mixture of which two *env* clones (FX343-26 and FX343-28) were identified. Both of these clones appeared to be the A component, however, amino acid differences were observed in both sequences of these *env* clones, therefore the two sequences were characterised further.

FZ215

Isolate FZ215 was an A/B/C mixture, of which five *env* gene clones were identified, four of which (FZ215-10, FZ215-11, FZ215-12 and FZ215-40) were the B-component while the fifth clone, FZ215-9, was the A-component. These results suggested that the B-component was relatively abundant in this isolate.

FA621

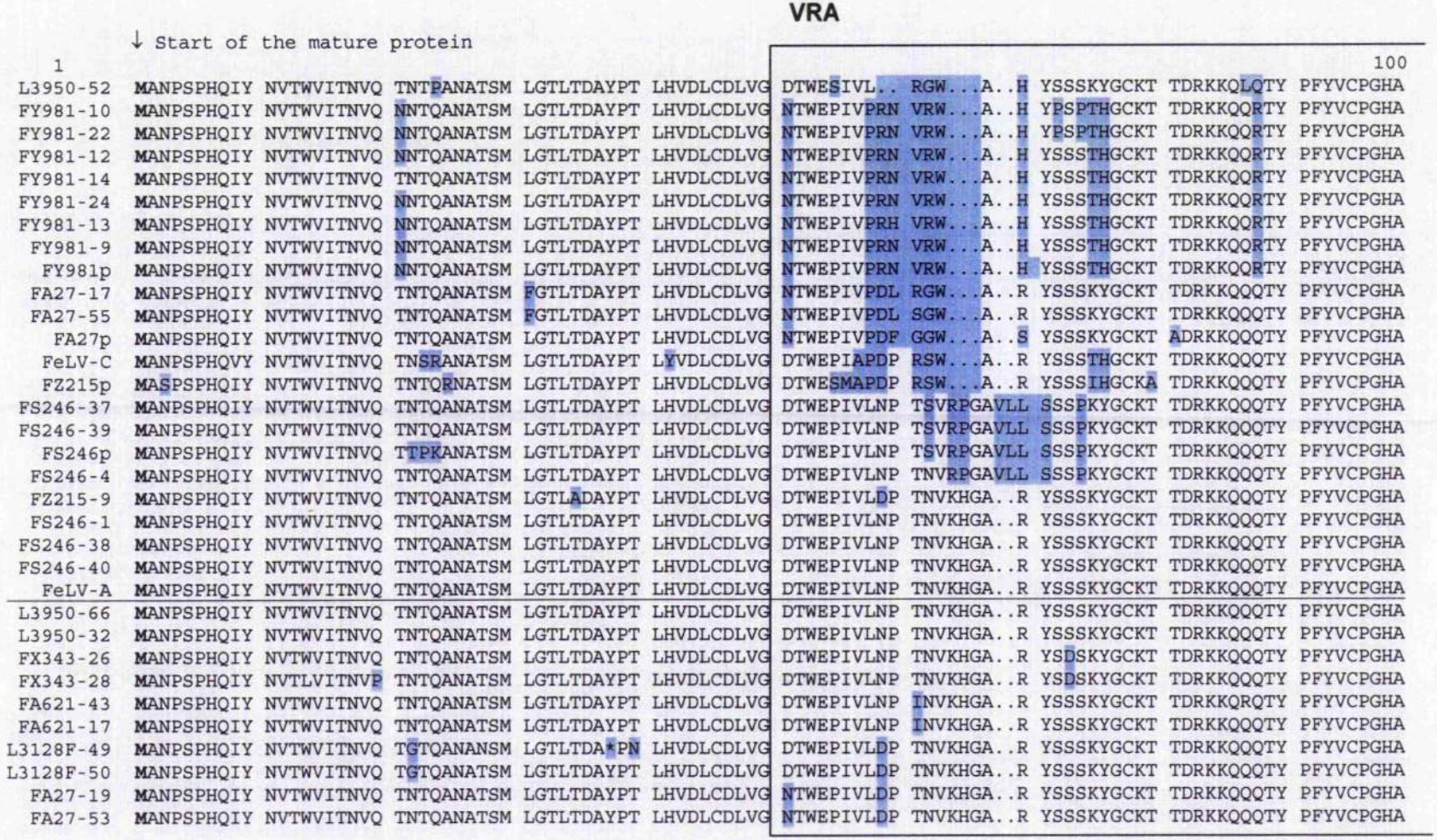
Primary isolate FA621 was an A/C mixture, however when sequencing the five envelope genes obtained, three were similar to FeLV-B/Gardner Arnstein (FA621-15, FA621-42 and FA621-44) and two appeared to be the A-component (FA621-17 and FA621-43). These *env* clones were not characterised further, as the regions of the *env* clones sequenced were identical to the sequences published of FeLV-A/Glasgow-1 and FeLV-B/Gardner Arnstein. One *env* clone, FA621-44 contained two stop codons prior to the VRA region, suggesting this protein would be non-functional.

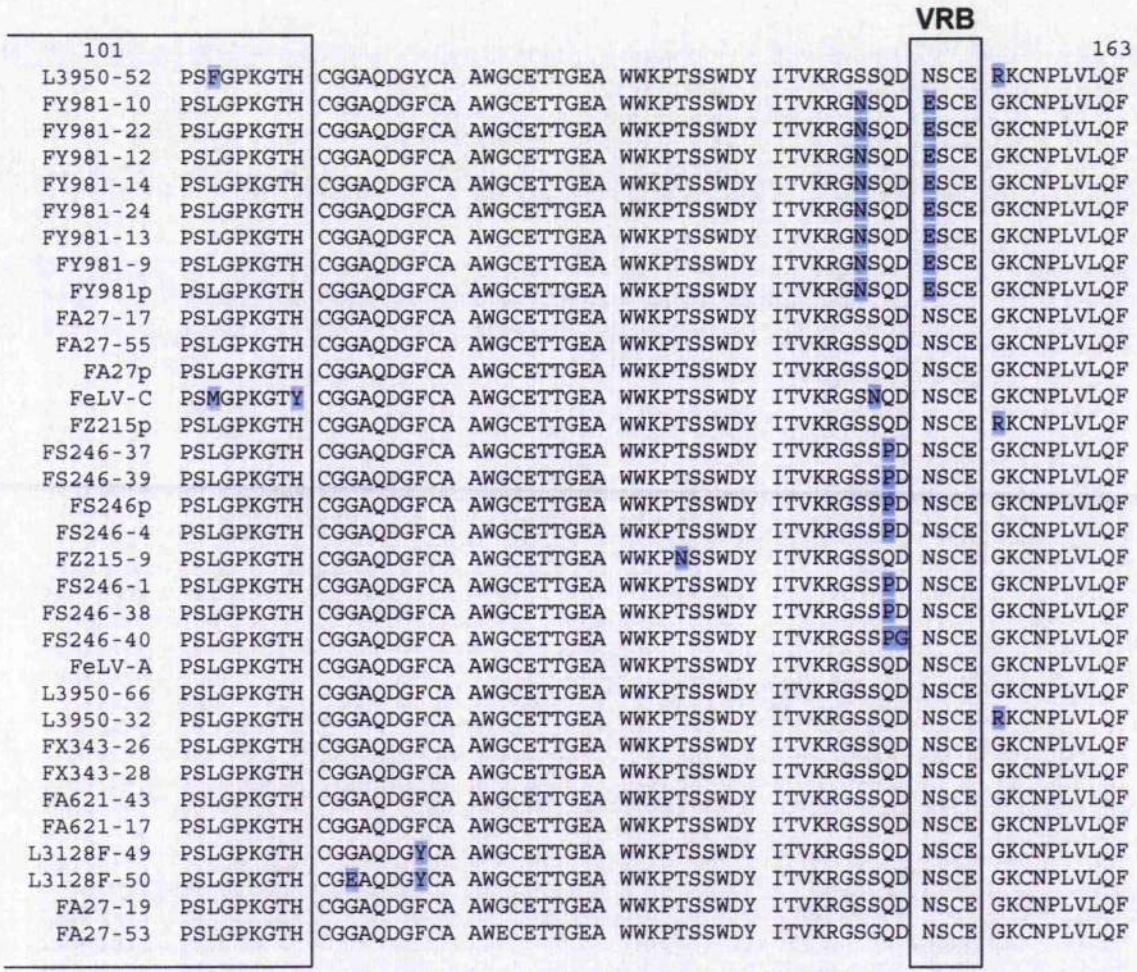
L3128F

When the two *env* genes obtained from primary isolate L3128F were sequenced, the *env* genes (L3128F-49 and L3128F-50) appeared to be the A-component of the primary isolate. *Env* clone L3128F-50 had a stop codon prior to the VRA region, suggesting this clone would be non-functional. These clones were not characterised further.

L3950F

Similarly, primary isolate L3950F yielded three envelope genes, L3950F-32, L3950F-52 and L3950F-66. One *env* clone, L3950F-52 appeared to be the C-component and two *env* clones (L3950F-32 and L3950F-66) were the A-components of his primary isolate. However, these clones were not characterised further as they appeared to be identical to the published FeLV-A/Glasgow-1 sequence.


VRA
↓ Start of the mature protein
1
100
L3950-52 MANPSPHQIY NVTWVITNVQ TNTPANATSM LGTLTDAYPT LHVDLCDLVG DTWESIVL.. RGW...A..H YSSSKYGCKT TDRKKQLQTY PFYVCPGHA
FY981-10 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YPSPTHGCKT TDRKKQQRTY PFYVCPGHA
FY981-22 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YPSPTHGCKT TDRKKQQRTY PFYVCPGHA
FY981-12 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YSSSTHGCKT TDRKKQQRTY PFYVCPGHA
FY981-14 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YSSSTHGCKT TDRKKQQRTY PFYVCPGHA
FY981-24 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YSSSTHGCKT TDRKKQQRTY PFYVCPGHA
FY981-13 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRH VRW...A..H YSSSTHGCKT TDRKKQQRTY PFYVCPGHA
FY981-9 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YSSSTHGCKT TDRKKQQRTY PFYVCPGHA
FY981p MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...A..H YSSSTHGCKT TDRKKQQRTY PFYVCPGHA
FA27-17 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...A..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FA27-55 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL SGW...A..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FA27p MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPDF GGW...A..S YSSSKYGCKT ADRKKQQQTY PFYVCPGHA
FeLV-C MANPSPHQVY NVTWVITNVQ TNSRANATSM LGTLTDAYPT LYVDLCDLVG DTWEPIAPDP RSW...A..R YSSSTHGCKT TDRKKQQQTY PFYVCPGHA
FZ215p MASPSPHQIY NVTWVITNVQ TNTQRNATSM LGTLTDAYPT LHVDLCDLVG DTWESMAPDP RSW...A..R YSSSIHGCKA TDRKKQQQTY PFYVCPGHA
FS246-37 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TSVRPGAVLL SSSPKYGCKT TDRKKQQQTY PFYVCPGHA
FS246-39 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TSVRPGAVLL SSSPKYGCKT TDRKKQQQTY PFYVCPGHA
FS246p MANPSPHQIY NVTWVITNVQ TTPKANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TSVRPGAVLL SSSPKYGCKT TDRKKQQQTY PFYVCPGHA
FS246-4 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVRPGAVLL SSSPKYGCKT TDRKKQQQTY PFYVCPGHA
FZ215-9 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLADAYPT LHVDLCDLVG DTWEPIVLDP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FS246-1 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FS246-38 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FS246-40 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FeLV-A MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
L3950-66 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
L3950-32 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FX343-26 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSDSKYGCKT TDRKKQQQTY PFYVCPGHA
FX343-28 MANPSPHQIY NVTLVITNVP TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGA..R YSDSKYGCKT TDRKKQQQTY PFYVCPGHA
FA621-43 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP INVKHGA..R YSSSKYGCKT TDRKKQRQTY PFYVCPGHA
FA621-17 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP INVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
L3128F-49 MANPSPHQIY NVTWVITNVQ TGTQANANSM LGTLTDA*PN LHVDLCDLVG DTWEPIVLDP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
L3128F-50 MANPSPHQIY NVTWVITNVQ TGTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLDP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FA27-19 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVLDP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA
FA27-53 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVLDP TNVKHGA..R YSSSKYGCKT TDRKKQQQTY PFYVCPGHA


VRB
101
163
L3950-52 PSFGPKGTH CGGAQDGYCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE RKCNPLVLQF
FY981-10 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981-22 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981-12 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981-14 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981-24 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981-13 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981-9 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FY981p PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPLVLQF
FA27-17 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FA27-55 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FA27p PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FeLV-C PSMGPKGTY CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSNQD NSCE GKCNPLVLQF
FZ215p PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE RKCNPLVLQF
FS246-37 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPLVLQF
FS246-39 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPLVLQF
FS246p PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPLVLQF
FS246-4 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPLVLQF
FZ215-9 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPNSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FS246-1 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPLVLQF
FS246-38 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPLVLQF
FS246-40 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPG NSCE GKCNPLVLQF
FeLV-A PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
L3950-66 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
L3950-32 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE RKCNPLVLQF
FX343-26 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FX343-28 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FA621-43 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FA621-17 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
L3128F-49 PSLGPKGTH CGGAQDGYCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
L3128F-50 PSLGPKGTH CGEAQDGYCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FA27-19 PSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPLVLQF
FA27-53 PSLGPKGTH CGGAQDGFCA AWECETTGEA WWKPTSSWDY ITVKRGSGQD NSCE GKCNPLVLQF

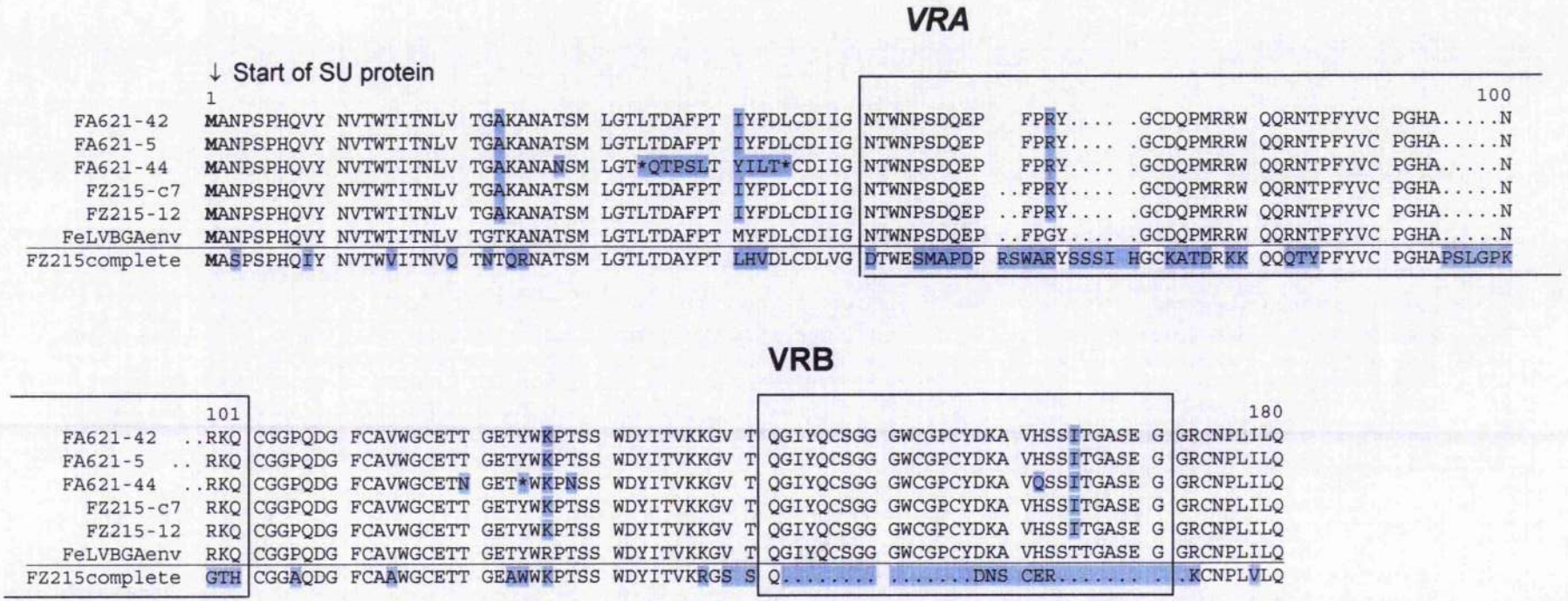


Figure 3.3b: *Amino acid alignment of a section of the SU proteins, which includes both variable regions VRA and VRB, of the unique envelope glycoproteins most closely resembling FeLV-B/Gardner Arnstein. The arrow indicates the start of the mature SU protein, whereas a dot indicates a gap in the alignment. The standard one letter code for the amino acids is used and the amino acid position is indicated above the alignment, this position is relative to the published FeLV-B/Gardner Arnstein genome.*

As summarised in table 3.3, when the nucleic acid sequences of the novel isolates were compared with FeLV-A, five nucleotide sequences were similar to FeLV-A/Glasgow-1, (FA27-53; 95.5%, FS246-40; 98.7%, FX343-26; 98.3%, FX343-28; 97.9% and FZ215-9; 97.5% similarity to FeLV-A respectively), while six nucleotide sequences were similar to FeLV-C/Sarma (FA27-17; 90.2%, FA27-55; 89.2%, FY981-10; 88.9%, FY981-14; 89.5% FS246-4; 92.3% and FS246-37; 90.8% similarity to FeLV-C/Sarma respectively). One nucleotide sequence was similar to FeLV-B/Gardner Arnstein (FZ215-12; 98.4% similarity). A multiple nucleotide sequence alignment was performed on these sequences, with the exception of the isolate that was related to FeLV-B/Gardner Arnstein. The alignment is shown in figure 3.4 and position 1 in this nucleotide sequence correlates with nucleotide 6079 in the published FeLV-A genome. The majority of the sequence variation clusters in the VRA region of the envelope glycoprotein. It is striking that on examining the nucleotide alignment, the majority of changes in the nucleotide sequence are found in similar positions in the VRA region of the *env* gene.

A separate nucleotide alignment was performed on the isolate that was related to FeLV-B/Gardner Arnstein, this alignment is shown in figure 3.5. Position 1 in this nucleotide sequence correlates with nucleotide 211 in the published FeLV-B genome. At present, two sequences that represent FeLV-B have been published; they are FeLV-B/Gardner Arnstein (Accession number P03391, (Wunsch M. et al., 1983)) and FeLV-B/Snyder Theilen (Accession number K01208, (Nunberg J.H. et al., 1984)). In order to maintain the nucleotide alignment, an adenine on position 380 in the published FeLV-B/Snyder Theilen sequence was deleted. The nucleotide alignment was lost from nucleotide 1510 in the published sequence of FeLV-B/Snyder Theilen and thus this sequence was deleted from that position onwards. When FZ215 was compared with the published sequences of the FeLV-B/Gardner Arnstein and FeLV-B/Snyder Theilen *env* genes, the majority of the differences were localised to the VRA and the VRB regions. A third cluster of changes was present from nucleotide 1115, however these changes did not affect the amino acid sequence (synonymous changes) and thus were discounted.

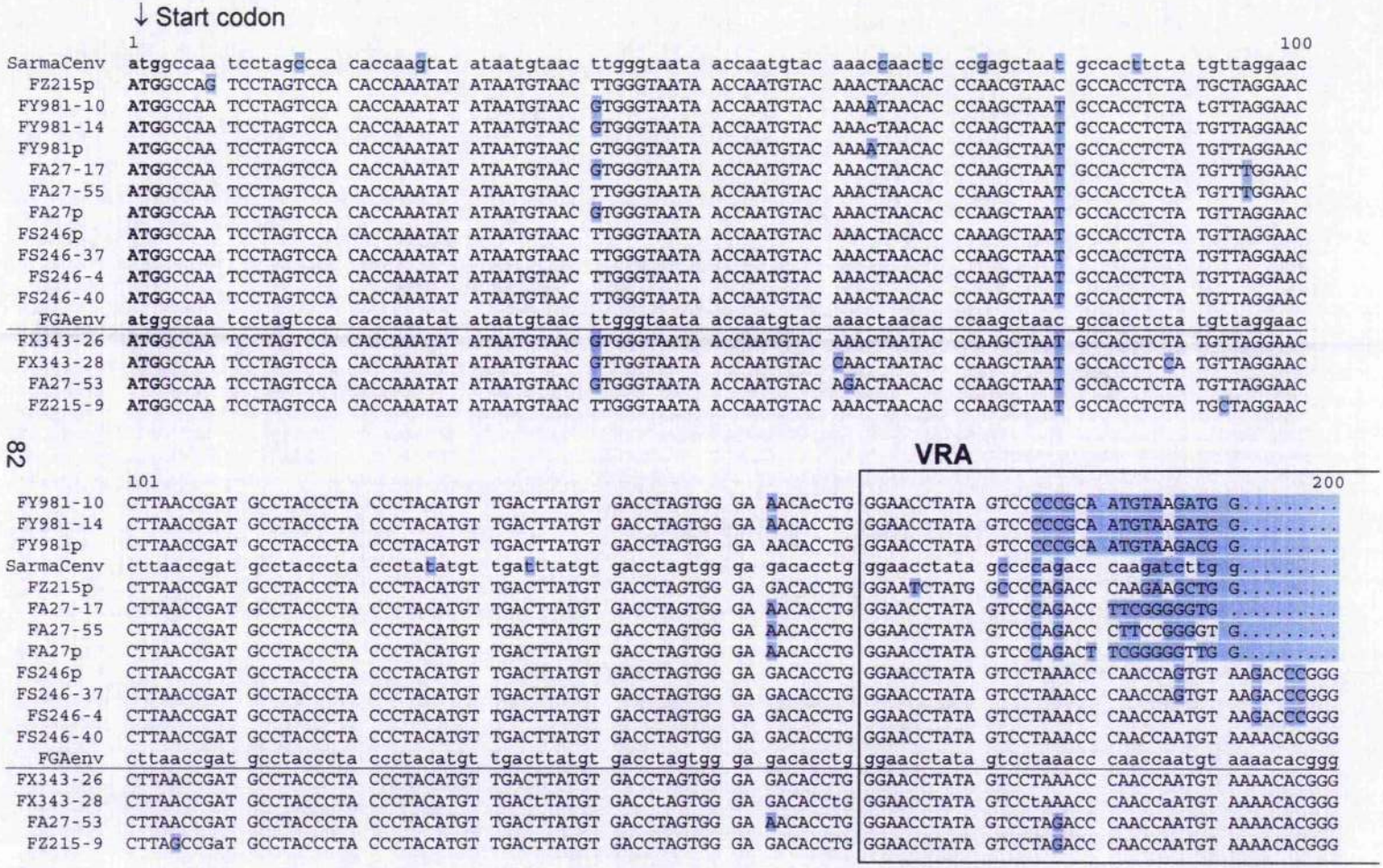
The sequences were then translated into amino acid sequences to establish the integrity of the *env* open reading frame and changes from the published FeLV A sequence. All of the novel *env* gene clones had intact open reading frames and figure 3.6 shows the amino acid alignment using the "Pileup" component on GCG. Position 1 in the amino acid sequence correlates with position 6079 in the published FeLV-A genome. The percentage similarity found for the novel *env* clones compared to FeLV-A/Glasgow-1 and FeLV-C/Sarma were equivalent to those found for the nucleic acids, albeit slightly higher, and are summarised in table 3.3. This higher degree of amino acid similarity suggests that some of the mutations detected in the nucleic acid sequence are "synonymous" (i.e. not affecting the coding sequence). When the sequence comparisons were restricted to the VRA region, the degree of amino acid similarity was considerably lower, consistent with this region being the major site for variation; the results of these comparisons are illustrated in table 3.4.

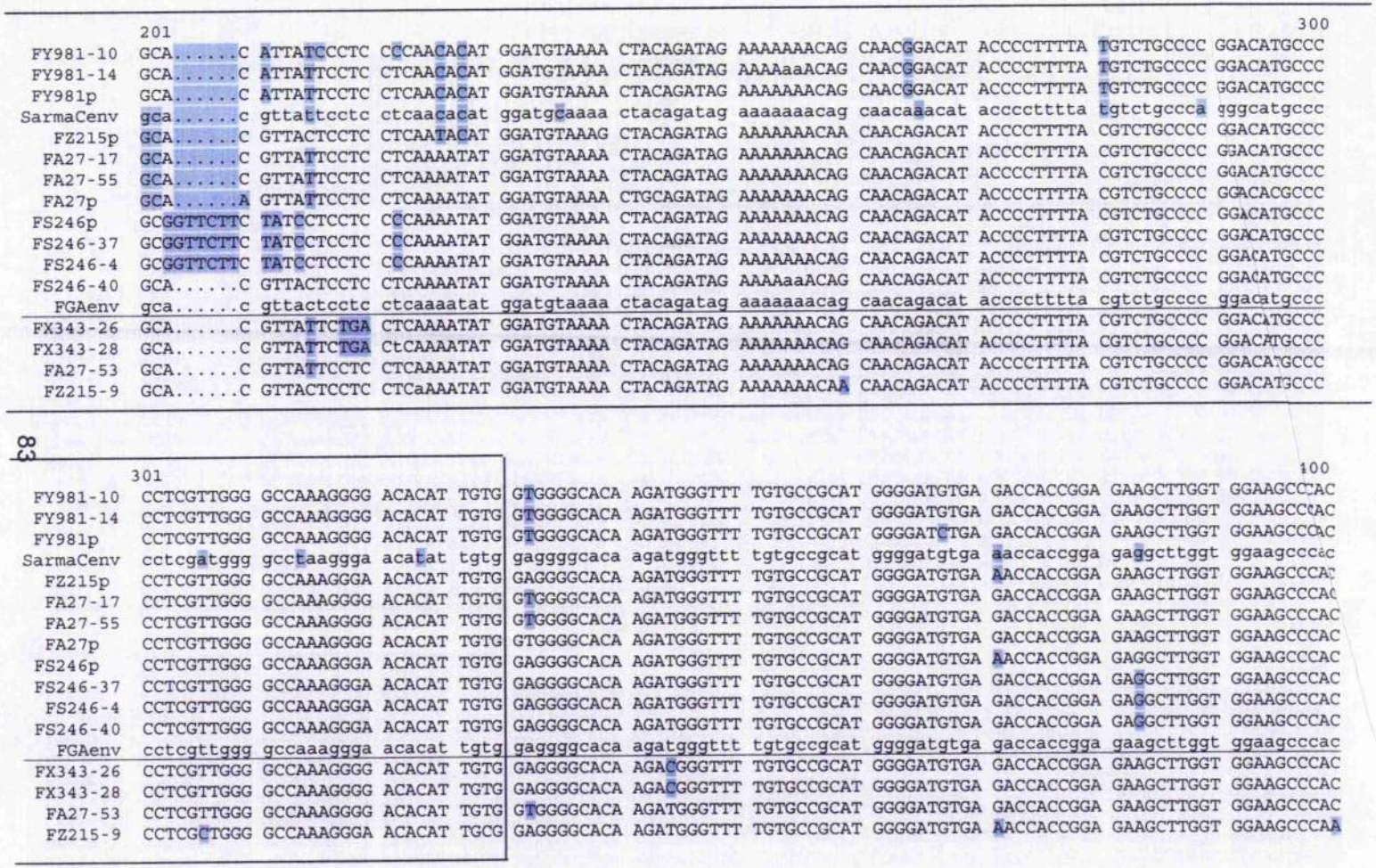
Env clone #	Nucleotide similarity to FeLV-A Glasgow 1	Amino acid similarity to FeLV-A Glasgow 1	Nucleotide similarity to FeLV-C/Sarma	Amino acid similarity to FeLV-C/Sarma	Subgroup predicted by sequencing
FA27-17	94.9%	97.4%	90.2%	94.7%	C
FA27-53	95.5%	99.0%	88.4%	94.4%	A
FA27-55	96.1%	97.4%	89.2%	94.7%	C
FS246-4	98.5%	98.7%	92.3%	93.7%	C
FS246-37	98.1%	98.9%	90.8%	93.9%	C
FS246-40	98.7%	99.3%	90.1%	94.2%	A
FY981-10	95.7%	97.4%	88.9%	93.2%	C
FY981-14	96.5%	97.9%	89.5%	93.7%	C
FX343-26	98.3%	99.0%	90.0%	93.9%	A
FX343-28	97.9%	98.2%	89.9%	93.6%	A
FZ215-9	97.5%	98.2%	89.9%	93.4%	A
FZ215-12	77.4%	77.0%	90.1%	89.1%	B (98.6% and 98.9%)

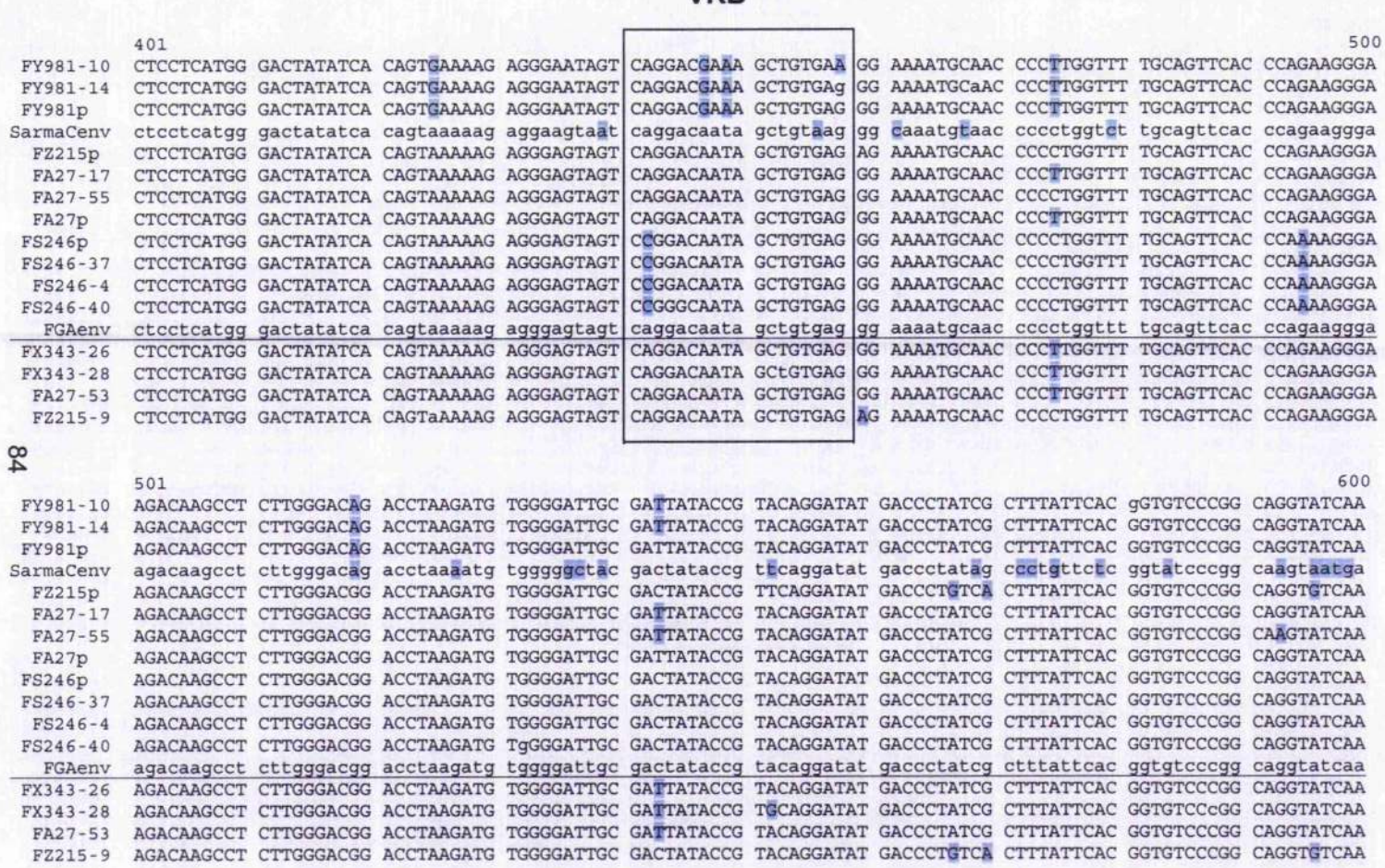
Table 3.3: *Nucleotide and amino acid sequence homology of the complete env genes relative to the published sequences of the envelope glycoproteins of FeLV-A/Glasgow-1 and FeLV-C/Sarma.*

Env clone #	Amino acid similarity to FeLV-A Glasgow 1	Amino acid similarity to FeLV-C/Sarma	Predicted subgroup determined by sequencing
FA27-17	95.4	93.2	C
FA27-53	97.3	92.5	A
FA27-55	93.9	91.9	C
FS246-4	96.6	90.9	C
FS246-37	95.9	90.2	C
FS246-40	98.6	92.4	A
FY981-10	92.5	88.5	C
FY981-14	94.6	90.5	C
FX343-26	99.3	93.1	A
FX343-28	98.0	91.2	A
FZ215-9	97.3	91.8	A
FZ215-12	75.4	75.0	B

Table 3.4: *Amino acid sequence homology of the VRA region of the novel env genes relative to the published sequences of FeLV-A/Glasgow-1 and FeLV-C/Sarma. The VRA region is a stretch of fifty-six amino acids starting from amino acid 50, relative to the published sequence of FeLV-A/Glasgow-1 (^{50}DTWEPI to PKGTH^{110}).*


↓ Start codon
1
100
SarmaCenv atggccaa tcctagccca caccaagtat ataatgtaac ttgggtaata accaatgtac aaaccaactc ccgagctaat gccacttcta tgttaggaac
FZ215p ATGGCCAG TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAACGTAAC GCCACCTCTA TGCTAGGAAC
FY981-10 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAAATAACAC CCAAGCTAAT GCCACCTCTA tGTTAGGAAC
FY981-14 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAAcTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FY981p ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAAATAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FA27-17 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTTGGAAC
FA27-55 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCtA TGTTTGGAAC
FA27p ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FS246p ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTACACC CAAAGCTAAT GCCACCTCTA TGTTAGGAAC
FS246-37 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FS246-4 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FS246-40 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FGAenv atggccaa tcctagtcca caccaaatat ataatgtaac ttgggtaata accaatgtac aaactaacac ccaagctaac gccacctcta tgttaggaac
FX343-26 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAATAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FX343-28 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTTGGTAATA ACCAATGTAC CAACTAATAC CCAAGCTAAT GCCACCTCCA TGTTAGGAAC
FA27-53 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AGACTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FZ215-9 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGCTAGGAAC
82
VRA
101
200
FY981-10 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACAcCTG GGAACCTATA GTCCCCCGCA ATGTAAGATG G.....
FY981-14 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCCCGCA ATGTAAGATG G.....
FY981p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCCCGCA ATGTAAGACG G.....
SarmaCenv cttaaccgat gcctacccta ccctatatgt tgatttatgt gacctagtgg ga gacacctg ggaacctata gcccagacc caagatcttg g.....
FZ215p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAATCTATG GCCCCAGACC CAAGAAGCTG G.....
FA27-17 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCAGACC TTCGGGGGTG .....
FA27-55 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCAGACC CTTCCGGGGT G.....
FA27p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCAGACT TCGGGGGTTG G.....
FS246p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAAACC CAACCAGTGT AAGACCCGGG
FS246-37 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAAACC CAACCAGTGT AAGACCCGGG
FS246-4 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAAACC CAACCAATGT AAGACCCGGG
FS246-40 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAAACC CAACCAATGT AAAACACGGG
FGAenv cttaaccgat gcctacccta ccctacatgt tgacttatgt gacctagtgg ga gacacctg ggaacctata gtcctaaacc caaccaatgt aaaacacggg
FX343-26 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAAACC CAACCAATGT AAAACACGGG
FX343-28 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACtTATGT GACCtAGTGG GA GACACCtG GGAACCTATA GTCCtAAACC CAACCaATGT AAAACACGGG
FA27-53 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCTAGACC CAACCAATGT AAAACACGGG
FZ215-9 CTTAGCCGaT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAGACC CAACCAATGT AAAACACGGG


201
300
FY981-10 GCA.....C ATTATCCCTC CCCAACACAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACGGACAT ACCCCTTTTA TGTCTGCCCC GGACATGCCC
FY981-14 GCA.....C ATTATTCCTC CTCAACACAT GGATGTAAAA CTACAGATAG AAAAaaACAG CAACGGACAT ACCCCTTTTA TGTCTGCCCC GGACATGCCC
FY981p GCA.....C ATTATTCCTC CTCAACACAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACGGACAT ACCCCTTTTA TGTCTGCCCC GGACATGCCC
SarmaCenv gca.....c gttattcctc ctcaacacat ggatgcaaaa ctacagatag aaaaaaacag caacaaacat acccctttta tgtctgccca gggcatgccc
FZ215p GCA.....C GTTACTCCTC CTCAATACAT GGATGTAAAG CTACAGATAG AAAAAAACAA CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FA27-17 GCA.....C GTTATTCCTC CTCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FA27-55 GCA.....C GTTATTCCTC CTCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FA27p GCA.....A GTTATTCCTC CTCAAAATAT GGATGTAAAA CTGCAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACACGCCC
FS246p GCGGTTCTTC TATCCTCCTC CCCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FS246-37 GCGGTTCTTC TATCCTCCTC CCCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FS246-4 GCGGTTCTTC TATCCTCCTC CCCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FS246-40 GCA.....C GTTACTCCTC CTCAAAATAT GGATGTAAAA CTACAGATAG AAAAaaACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FGAenv gca.....c gttactcctc ctcaaaatat ggatgtaaaa ctacagatag aaaaaaacag caacagacat acccctttta cgtctgcccc ggacatgccc
FX343-26 GCA.....C GTTATTCTGA CTCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FX343-28 GCA.....C GTTATTCTGA CtCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FA27-53 GCA.....C GTTATTCCTC CTCAAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAG CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
FZ215-9 GCA.....C GTTACTCCTC CTCaAAATAT GGATGTAAAA CTACAGATAG AAAAAAACAA CAACAGACAT ACCCCTTTTA CGTCTGCCCC GGACATGCCC
301
400
FY981-10 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FY981-14 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FY981p CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATCTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
SarmaCenv cctcgatggg gcctaaggga acatat tgtg gaggggcaca agatgggttt tgtgccgcat ggggatgtga aaccaccgga gaggcttggt ggaagcccac
FZ215p CCTCGTTGGG GCCAAAGGGA ACACAT TGTG GAGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA AACCACCGGA GAAGCTTGGT GGAAGCCCAC
FA27-17 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FA27-55 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FA27p CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FS246p CCTCGTTGGG GCCAAAGGGA ACACAT TGTG GAGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA AACCACCGGA GAGGCTTGGT GGAAGCCCAC
FS246-37 CCTCGTTGGG GCCAAAGGGA ACACAT TGTG GAGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAGGCTTGGT GGAAGCCCAC
FS246-4 CCTCGTTGGG GCCAAAGGGA ACACAT TGTG GAGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAGGCTTGGT GGAAGCCCAC
FS246-40 CCTCGTTGGG GCCAAAGGGA ACACAT TGTG GAGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAGGCTTGGT GGAAGCCCAC
FGAenv cctcgttggg gccaaagggа acacat tgtg gaggggcaca agatgggttt tgtgccgcat ggggatgtga gaccaccgga gaagcttggt ggaagcccac
FX343-26 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GAGGGGCACA AGACGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FX343-28 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GAGGGGCACA AGACGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FA27-53 CCTCGTTGGG GCCAAAGGGG ACACAT TGTG GTGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA GACCACCGGA GAAGCTTGGT GGAAGCCCAC
FZ215-9 CCTCGCTGGG GCCAAAGGGA ACACAT TGCG GAGGGGCACA AGATGGGTTT TGTGCCGCAT GGGGATGTGA AACCACCGGA GAAGCTTGGT GGAAGCCCAA


VRB
401
500
FY981-10 CTCCTCATGG GACTATATCA CAGTGAAAAG AGGGAATAGT CAGGACGAAA GCTGTGAA GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FY981-14 CTCCTCATGG GACTATATCA CAGTGAAAAG AGGGAATAGT CAGGACGAAA GCTGTGAg GG AAAATGCaAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FY981p CTCCTCATGG GACTATATCA CAGTGAAAAG AGGGAATAGT CAGGACGAAA GCTGTGAG GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
SarmaCenv ctcctcatgg gactatatca cagtaaaaag aggaagtaat caggacaata gctgtaag gg caaatgtaac cccctggtct tgcagttcac ccagaaggga
FZ215p CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG AG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAGAAGGGA
FA27-17 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FA27-55 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG GG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAGAAGGGA
FA27p CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FS246p CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CCGGACAATA GCTGTGAG GG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAAAAGGGA
FS246-37 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CCGGACAATA GCTGTGAG GG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAAAAGGGA
FS246-4 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CCGGACAATA GCTGTGAG GG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAAAAGGGA
FS246-40 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CCGGGCAATA GCTGTGAG GG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAAAAGGGA
FGAenv ctcctcatgg gactatatca cagtaaaaag agggagtagt caggacaata gctgtgag gg aaaatgcaac cccctggttt tgcagttcac ccagaaggga
FX343-26 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FX343-28 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCtGTGAG GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FA27-53 CTCCTCATGG GACTATATCA CAGTAAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG GG AAAATGCAAC CCCTTGGTTT TGCAGTTCAC CCAGAAGGGA
FZ215-9 CTCCTCATGG GACTATATCA CAGTaAAAAG AGGGAGTAGT CAGGACAATA GCTGTGAG AG AAAATGCAAC CCCCTGGTTT TGCAGTTCAC CCAGAAGGGA
501
600
FY981-10 AGACAAGCCT CTTGGGACAG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC gGTGTCCCGG CAGGTATCAA
FY981-14 AGACAAGCCT CTTGGGACAG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FY981p AGACAAGCCT CTTGGGACAG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
SarmaCenv agacaagcct cttgggacag acctaaaatg tgggggctac gactataccg ttcaggatat gaccctatag cctgttctc ggtatcccgg caagtaatga
FZ215p AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GACTATACCG TTCAGGATAT GACCCTGTCA CTTTATTCAC GGTGTCCCGG CAGGTGTCAA
FA27-17 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FA27-55 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAAGTATCAA
FA27p AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FS246p AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GACTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FS246-37 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GACTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FS246-4 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GACTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FS246-40 AGACAAGCCT CTTGGGACGG ACCTAAGATG TgGGGATTGC GACTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FGAenv agacaagcct cttgggacgg acctaagatg tggggattgc gactataccg tacaggatat gaccctatcg ctttattcac ggtgtcccgg caggtatcaa
FX343-26 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FX343-28 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GATTATACCG TGCAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCcGG CAGGTATCAA
FA27-53 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GATTATACCG TACAGGATAT GACCCTATCG CTTTATTCAC GGTGTCCCGG CAGGTATCAA
FZ215-9 AGACAAGCCT CTTGGGACGG ACCTAAGATG TGGGGATTGC GACTATACCG TACAGGATAT GACCCTGTCA CTTTATTCAC GGTGTCCCGG CAGGTGTCAA

```
          601                                                                                                  700
FY981-10  CCATTACGCC GCCTCAGGCA ATGGGACCAA AcCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
FY981-14  CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCAA aaACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
FY981p    CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
SarmaCenv ccattacgcc gcctcaggcg atgggaccca acttagtctt acctgatcaa aaacccccat cccgacaatc tcaaacaaag tccaaggtga caacccagag
FZ215p    CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCCT ACCTAATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CGACCCAGAT
FA27-17   CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCGA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CGACCCAGAG
FA27-55   CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCtT ACCTGATCGA AAACCCCCAT CCCGACAATC TCAAaCAGGG TCCAaAGTGG CGACCCAGAG
FA27p     CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCGA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CGACCCAGAG
FS246p    CCATTACACC GCCTCAGGCA ATGGGACCGA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
FS246-37  CCATTACACC GCCTCAGGCA ATGGGACCGA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
FS246-4   CCATTACACC GCCTCAGGCA ATGGGACCGA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
FS246-40  CCATTACACC GCCTCAGGCA ATGGGACCGA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CAACCCAGAG
FGAenv    ccattacgcc gcctcaggca atgggaccaa acctagtctt acctgatcaa aaacccccat cccgacaatc tcaaacaggg tccaaagtgg cgacccagag
FX343-26  CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGAAAATC TCAAACAGGG TCCAAAGTGG CGACCCAGAG
FX343-28  cCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CGACCCAGAG
FA27-53   CCATTACGCC GCCTCAGGCA ATGGGACCAA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG CGACCCAGAG
FZ215-9   CCATTACGCC GCCTCAGGcA ATGGGACCAA ACCTAGTCTT ACCTGATCAA AAACCCCCAT CCCGACAATC TCAAACAGGG TCCAAAGTGG cGACCCAGAG

          701                                                                                                  800
FY981-10  GCCCC.AAAC AAATGAAAGC GCCCCAAGGT CTATTGCCCC CACCACCATG AGTCCCAAAC GAATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FY981-14  GCCCC.AAAC AAATGAAAGC GCCCcAAGGT CTATTGCCCC CACCACCATG AGTCCCAAAC GAATTGGGAC TGGAGATAGG TTAATAAATT TAGTACAAGG
FY981p    GCCCC.AAAC AAATGAAAGC GCCCCAAGGT CTATTGCCCC CACCACCATG AGTCCCAAAC GAATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
SarmaCenv gcccc.aaat aactagcagc accccaaggt ctgtcgcctc cgctaccatg ggtcccaaac ggatagggac cggagataga ttaataaatt tagtgcaagg
FZ215p    GCCCC AAAC GAATGAAAGC GCCCCAAGAT CTGTTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGACAGG TTAATAAATT TAGTACAAGG
FA27-17   GCCCC.AAAC AAATGAAAGC GCCCCAAGGT CTATTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FA27-55   GCCCCAAAAC AAATGAAAGC GCCCCAAGGT CTATTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FA27p     GCCCC AAAC AAATGAAAGC GCCCCAAGGT CTATTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CAGAGATAGG CTAATAAATT TAGTACAAGG
FS246p    GCCCC AAAC GAATGAAAGC GCCCCAAGGT CTGTTGGCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FS246-37  GCCCC.AAAC GAATGAAAGC GCCCCAAGGT CTGTTGGGCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATt TAGTACAAGG
FS246-4   GCCCC.AAAC GAATGAAAGC GCCCCAAGGT CTGTTGGCCC AACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FS246-40  GCCCC.AAAC GAATGAAAGC GCCCCAAGGT CTGTTGGCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FGAenv    gcccc.aaac gaatgaaagc gccccaaggt ctgttgcccc caccaccatg ggtcccaaac ggattgggac cggagatagg ttaataaatt tagtacaagg
FX343-26  GCCCc.AAAC A...GAAAGC GCCCCAAGGT CTGTTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FX343-28  GCCCC.AAAC AACTGAAAGC GCCCCAAGGT CTGTTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FA27-53   GCCCC.AAAC AAATGAAAGC GCCCCAAGGT CTATTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGATAGG TTAATAAATT TAGTACAAGG
FZ215-9   GCCCC.AAAC GAATGAAAGC GCCCCAAGAT CTGTTGCCCC CACCACCATG GGTCCCAAAC GGATTGGGAC CGGAGACAGG TTAATAAATT TAGTACAAGG
```

```
          801                                                                                                 900
SarmaCenv gacataccta gccttaaatg ccaccgaccc caacaaaact aaagactgtt ggctctgcct ggtttctcga ccaccctatt acgaagggat tgcagtctta
  FZ215p  GACATACCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
 FY981-10 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
 FY981-14 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
 FY981p   GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
  FA27-17 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGTCT AGTCTCCCGA CCGCCCTATT ATGAAGGAAT TGCGGTATTG
  FA27-55 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGTCT AGTCTCCCGA CCGCCCTATT ATGAAGGAAT TGCGGTATTG
  FA27p   GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
 FS246p   GACATACCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
 FS246-37 GACATACCTA GCCTTAAATG CCACCGACCC cAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
  FS246-4 GACATACCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTTTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
 FS246-40 GACATACCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
   FGAenv gacataccta gccttaaatg ccaccgaccc caacaaaact aaagactgtt ggctctgcct ggtttctcga ccaccctatt acgaagggat tgcaatctta
 FX343-26 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTCTCTCGA CCACCCTATT ACGAAGGGGT TGCAATCTTA
 FX343-28 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTCTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
  FA27-53 GACATATCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA
  FZ215-9 GACATACCTA GCCTTAAATG CCACCGACCC CAACAAAACT AAAGACTGTT GGCTCTGCCT GGTTTCTCGA CCACCCTATT ACGAAGGGAT TGCAATCTTA

          901                                                                                                1000
 FY981-10 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC AATCTATTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
 FY981-14 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC AATCTATTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
 FY981p   GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC AATCTATTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
SarmaCenv ggtaactaca gcaaccaaac aaaccccccc ccatcctgcc tatctaccgc gcaacataaa ctgactatat cagaagtgtc cgggcaaggt ttgtgcatag
  FZ215p  GGTAACTATA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTATTCC GCAACACAAA CTGACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
  FA27-17 GGCAATTACA GCAACCAAAC CAACCCCCCC CCATCCTGCC TATCCGACCC ACAACATAAA CTGACTATAT CAGAAGTGTC CGGGCAAGGT TTGTGCATAG
  FA27-55 GGCAATTACA GCAACCAAAC CAACCCCCCC CCATCCTGCC TATCCGACCC ACAACATAAA CTGACTATAT CAGAAGTGTC CGGGCAAGGT TTGTGCATAG
  FA27p   GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTATTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
 FS246p   GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTACTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
 FS246-37 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTACTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
  FS246-4 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTACTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
 FS246-40 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTACTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
   FGAenv ggtaactaca gcaaccaaac aaaccccccc ccatcctgcc tatctactcc gcaacacaaa ctaactatat ctgaagtatc agggcaagga atgtgcatag
 FX343-26 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTATTCC GCAACACAAA TTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
 FX343-28 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTATTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
  FA27-53 GGTAACTACA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTATTCC GCAACACAAA CTAACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
  FZ215-9 GGTAACTATA GCAACCAAAC AAACCCCCCC CCATCCTGCC TATCTATTCC GCAACACAAA CTGACCATAT CTGAAGTATC AGGGCAAGGA CTGTGCATAG
```

```
          1001                                                                                                1100
FY981-10  GGACTGTTCC TAAGACCCAC CAGGCTCTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
FY981-14  GGACTGTTCC TAAGACCCAC CAGGCTCTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
FY981p    GGACTGTTCC TAAGACCCAC CAGGCTCTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
SarmaCenv ggactgttcc taagacccac caagctttgt gcaaaaagac acaaaaagga cataaaggga ctcactacct ggcagccccc aacggcacct attgggcctg
FZ215p    GGACTGTTCC TAAGACCCAC CAGGCTTTGT GCAATGAGAC GCAACAGGGA CATACAGGGG CACACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCTTG
FA27-17   GGACTGTTCC TAAGACCCAC CAAGCTTTGT GCAAAAAGAC ACAAAAAGGA CATAAAGGGA CTCACTATCT AGCAGCCCCT AGCGGTACCT ACTGGGCATG
FA27-55   GGACTGTTCC TAAGACCCAC CAAGCTTTGT GCAAAAAGAC ACAAAAAGGA CATAAAGGGA CTCACTATCT AGCAGCCCCT AGTGGTACCT ACTGGGCATG
FA27p     GGACTGTTCC TAAGACCCAC CAGGCTTTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
FS246p    GGACTGTTCC TAAAACCCAC CAGGCTTTGT GCAATAAAAC ACAACAGGGA CATACAGGAG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCATG
FS246-37  GGACTGTTCC TAAAACCCAC CAGGCTTTGT GCAATAAAAC ACAACAGGGA CATACAGGAG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCATG
FS246-4   GGACTGTTCC TAAAACCCAC CAGGCTTTGT GCAATAAAAC ACAACAGGGA CATACAGGAG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCATG
FS246-40  GGACTGTTCC TAAAACCCAC CAGGCTTTGT GCAATAAAAC ACAACAGGGA CATACAGGAG CGCACTATCT AGCCGCCCCC AACGGCCCCT ATTGGGCATG
FGAenv    ggactgttcc taaaacccac caggctttgt gcaataagac acaacaggga catacagggg cgcactatct agccgccccc aacggcacct attgggcctg
FX343-26  GGACTGTTCC TAAGACCCAC CAGGCTTTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
FX343-28  GGACTGTTCC TAAGACCCAC CAGGCTTTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
FA27-53   GGACTGTTCC TAAGACCCAC CAGGCTTTGT GCAATGAGAC ACAACAGGGA CATACAGGGG CGCACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCCTG
FZ215-9   GGACTGTTCC TAAGACCCAC CAGGCTTTGT GCAATGAGAC GCAACAGGGA CATACAGGGG CACACTATCT AGCCGCCCCC AACGGCACCT ATTGGGCTTG

          1101                                                                                                1200
FY981-10  TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FY981-14  TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FY981p    TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
SarmaCenv taacactgga ctcaccccat gcatttccat ggcagtgctc aattggacct ctgatttttg tgtcttaatc gaattatggc ccagagtaac ctaccatcaa
FZ215p    TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCA TTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FA27-17   CAACACCGGG CTAACCCCAT GCATCTCCAT GGCAGTGCTC AACTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FA27-55   CAACACCGGG CTAACCCCAT GCATCTCCAT GGCAGTGCTC AACTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FA27p     TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FS246p    TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FS246-37  TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FS246-4   TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FS246-40  TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FGAenv    taacactgga ctcaccccat gcatttccat ggcggtgctc aattggacct ctgatttttg tgtcttaatc gaattatggc ccagagtgac ttaccatcaa
FX343-26  TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FX343-28  TAACACTGGA CTCGCCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FA27-53   TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCT CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
FZ215-9   TAACACTGGA CTCACCCCAT GCATTTCCAT GGCGGTGCTC AATTGGACCG CTGATTTTTG TGTCTTAATC GAATTATGGC CCAGAGTGAC TTACCATCAA
```

```
          1201                                                                                                1300
FY981-10  CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGGGA ACCGATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
FY981-14  CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGGGA ACCGATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
FY981p    CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGGAA CCGAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
SarmaCenv cccgaatata ttacacaca tttcgacaaa gctgtcaggt tccgaagaga acctatatca ctaaccgttg cccttatgtt gggaggactc accgtagggg
  FZ215p  CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCGGTATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
  FA27-17 CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
  FA27-55 CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
  FA27p   CCCGAATATG TGTACACACA TTTTGCCGAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
  FS246p  CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACGGTTG CCCTTATGTT GGGAGGACTT ACTGTAGGGG
FS246-37  CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACGGTTG CCCTTATGTT GGGAGGACTT ACTGTAGGGG
  FS246-4 CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATACCA CTAACGGTTG CCCTTATGTT GGGAGGACTT ACTGTAGGGG
FS246-40  CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACGGTTG CCCTTATGTT GGGAGGACTT ACTGTAGGGG
   FGAenv cccgaatatg tgtacacaca ttttgccaaa gctgtcaggt tccgaagaga accaatatca ctaacggttg cccttatgtt gggaggactt actgtagggg
FX343-26  CCCGAATATG TGTACACACA TTTTGCCGAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
FX343-28  CCCGAATATG TGTACACACA TTTTGCCGAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
  FA27-53 CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAATATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG
  FZ215-9 CCCGAATATG TGTACACACA TTTTGCCAAA GCTGTCAGGT TCCGAAGAGA ACCAGTATCA CTAACTGTTG CCCTTATGTT GGGAGGACTC ACTGTAGGGG

          1301                                                                                                1400
FY981-10  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
FY981-14  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
FY981p    GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
SarmaCenv gcatagccgc gggggtcgga acagggacta aagccctcct tgaaacagcc cagttcagac aactacaaat agccatgcac acagacatcc aggccctgga
  FZ215p  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACGGACATCC AAGCCCTAGA
  FA27-17 GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
  FA27-55 GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
  FA27p   GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
  FS246p  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
FS246-37  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
  FS246-4 GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
FS246-40  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
   FGAenv gcatagccgc gggggtcgga acagggacta aagccctcct tgaaacagcc cagttcagac aactacaaat ggccatgcac acagacatcc aggccctaga
FX343-26  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATAC AGGCCCTAGA
FX343-28  GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATAC AGGCCCTAGA
  FA27-53 GCATAGCCGC AGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AGGCCCTAGA
  FZ215-9 GCATAGCCGC GGGGGTCGGA ACAGGGACTA AAGCCCTCCT TGAAACAGCC CAGTTCAGAC AACTACAAAT GGCCATGCAC ACAGACATCC AAGCCCTAGA
```

```
          1401                                                                                                1500
FY981-10  AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FY981-14  AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FY981p    AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
SarmaCenv agagtcaatt agtgccttag aaaaatccct gacctccctc tctgaggtag tcctacaaaa taggcggggc ctagatattc tgttcttaca agagggaggg
FZ215p    AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FA27-17   AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FA27-55   AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FA27p     AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FS246p    AGAATCAATT AGTGCCTTAG AAAAGTCCCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FS246-37  AGAATCAATT AGTGCCTTAG AAAAGTCCCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FS246-4   AGAATCAATT AGTGCCTTAG AAAAGTCCCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FS246-40  AGAATCAATT AGTGCCTTAG AAAAGTCCCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FGAenv    agaatcaatt agtgccttag aaaagtccct gacctccctt tctgaagtag tcttacaaaa cagacggggc ctagatattc tattcttaca agagggaggg
FX343-26  AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FX343-28  AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT ACTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FA27-53   AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG
FZ215-9   AGAGTCAATT AGTGCCTTAG AAAAGTCTCT GACCTCCCTT TCTGAAGTAG TCTTACAAAA CAGACGGGGC CTAGATATTC TATTCTTACA AGAGGGAGGG

          1501                                                                                                1600
FY981-10  CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FY981-14  CTCTGTGCCG CATTAAAAGA AGAATGTTGT TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FY981p    CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
SarmaCenv ctctgtgccg cattaaaaga agaatgctgc ttctatgcag atcacaccgg actcgtccga gacaatatgg ctaaattaag agaaagacta aaacagcggc
FZ215p    CTCTGTGCCG CATTGAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CCAAATTAAG AGAAAGACTA AAACAGCGGC
FA27-17   CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAGAGACTA AAACAGCGGC
FA27-55   CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FA27p     CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FS246p    CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTACG AGAAAGACTA AAACAGCGGC
FS246-37  CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTACG AGAAAGACTA AAACAGCGGC
FS246-4   CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTACG AGAAAGACTA AAACAGCGGC
FS246-40  CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTACG AGAAAGACTA AAACAGCGGC
FGAenv    ctctgtgccg cattgaaaga agaatgttgc ttctatgcgg atcacaccgg actcgtccga gacaatatgg ccaaattaag agaaagacta aaacagcggc
FX343-26  CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FX343-28  CTCTGTGTCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FA27-53   CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CTAAATTAAG AGAAAGACTA AAACAGCGGC
FZ215-9   CTCTGTGCCG CATTAAAAGA AGAATGTTGC TTCTATGCGG ATCACACCGG ACTCGTCCGA GACAATATGG CCAAATTAAG AGAAAGACTA AAACAGCGGC
```

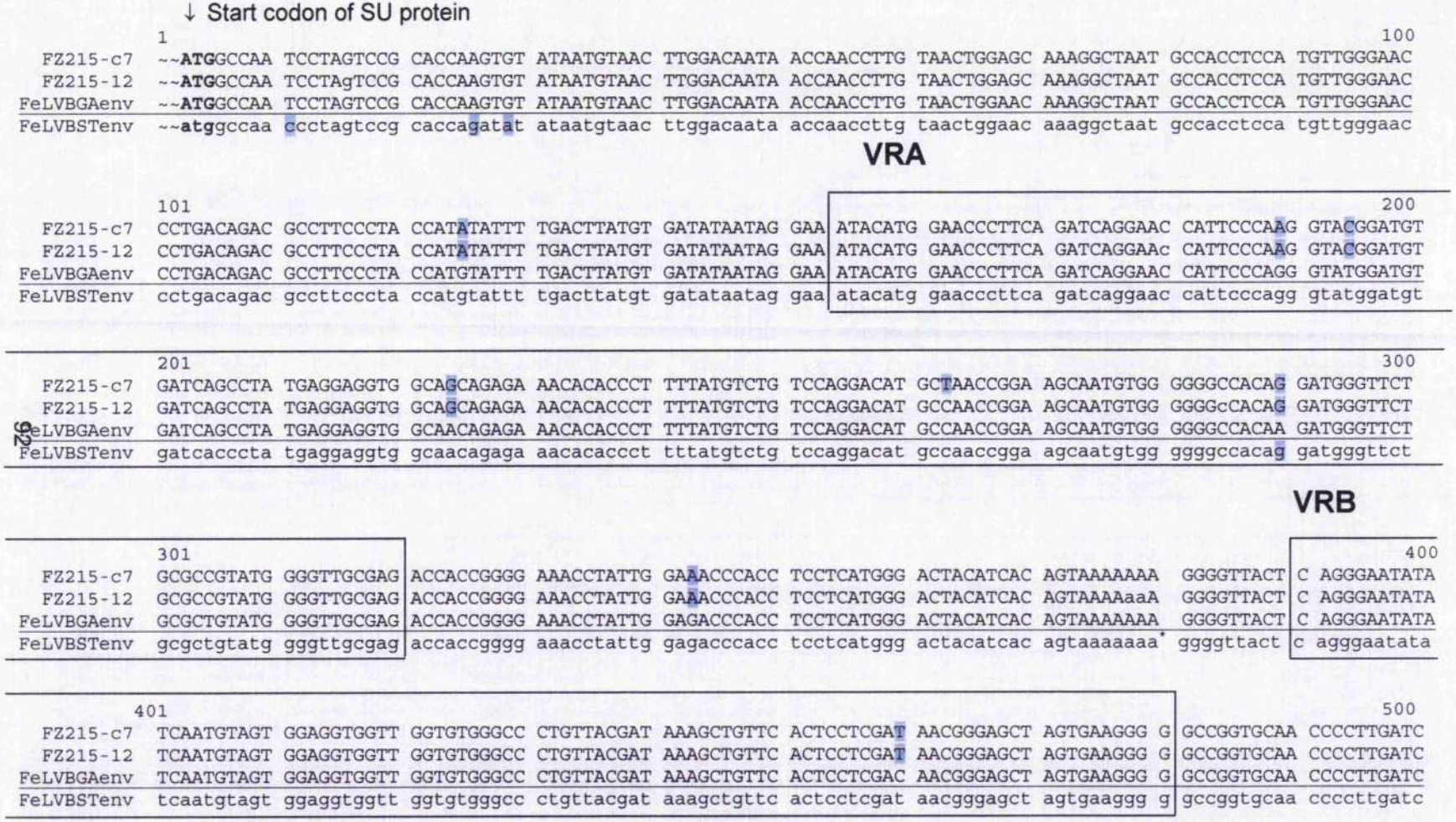
```
          1601                                                                                                1700
FY981-10  AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FY981-14  AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FY981p    AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
SarmaCenv aacaactgtt tgattcccaa cagggatggt ttgaaggatg gttcaacaag tcccccтggt ttacaaccct aatttcctcc atcatgggcc ccttactaat
FZ215p    AACAATTGTT TGACTCCCAA CAGGGAGGGT TTGAAGGATG GTTCAACAAG TCCCCTAGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FA27-17   AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTGTGGGCC CCTTACTAAT
FA27-55   AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FA27p     AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FS246p    AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FS246-37  AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FS246-4   AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FS246-40  AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FGAenv    aacaactgtt tgactcccaa cagggatggt ttgaaggatg gttcaacaag tcccccтggt ttacaaccct aatttcctcc attatgggcc ccttactaat
FX343-26  AACAGCTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FX343-28  AACAGCTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FA27-53   AACAACTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT
FZ215-9   AACAATTGTT TGACTCCCAA CAGGGATGGT TTGAAGGATG GTTCAACAAG TCCCCCTGGT TTACAACCCT AATTTCCTCC ATTATGGGCC CCTTACTAAT

          1701                                                                                                1800
FY981-10  CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FY981-14  CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FY981p    CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
SarmaCenv cctactccta attctcctcc tcggcccatg catccttaac cgattagtgc aattcgtaaa agacagaata tctgtggtac aagccttaat tttaacccaa
FZ215p    CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FA27-17   CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATCCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FA27-55   CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FA27p     CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FS246p    CCTACTCCTA ATTCTCCTCT TCGGCCCGTG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FS246-37  CCTACTCCTA ATTCTCCTCT TCGGCCCGTG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AGGCTTTAAT TTTAACCCAA
FS246-4   CCTACTCCTA ATTCTCCTCT TCGGCCCGTG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AGGCTTTAAT TTTAACCCAA
FS246-40  CCTACTCCTA ATTCTCCTCT TCGGCCCGTG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AGGCTTTAAT TTTAACCCAA
FGAenv    cctactccta attctcctct tcggcccatg catccttaac cgattagtac aattcgtaaa agacagaata tctgtggtac aggctttaat tttaacccaa
FX343-26  CCTACTCCTA ATTCTCCTCT TCGGGCCATG CATCCTTAAC CGATTAGTAC AATTCATAAt AGACAGAATA TCTGTGGTAC AAGCTCTAAT TTTAACCCAA
FX343-28  CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTcATAAG AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FA27-53   CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAATA TCTGTGGTAC AAGCTTTAAT TTTAACCCAA
FZ215-9   CCTACTCCTA ATTCTCCTCT TCGGCCCATG CATCCTTAAC CGATTAGTAC AATTCGTAAA AGACAGAACT ATTGTGGTAC AAGCTTTAAT TTTAACCCAA
```


1801
1843
FY981-10 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FY981-14 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FY981p CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
SarmaCenv cagtaccaac agatacaaca atacgattcg gaccgaccat gat
FZ215p CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FA27-17 CAGTACCAAC AGATAAGGCA ATACGATCCG GACCGACCAT GAT
FA27-55 CAGTACCAAC AGATAAGGCA ATACGATCCG GACCGACCAT GAT
FA27p CAGTACCAAC AGATAAGGCA ATACGATCCG GACCGACCAT GAT
FS246p CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FS246-37 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FS246-4 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FS246-40 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FGAenv cagtaccaac agataaagca atacgatccg gaccgaccat gat
FX343-26 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FX343-28 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
FA27-53 CAGTACCAAC AGATAAGGCA ATACGATCCG GACCGACCAT GAT
FZ215-9 CAGTACCAAC AGATAAAGCA ATACGATCCG GACCGACCAT GAT
↑ Stop codon

Figure 3.5: *Nucleotide alignment from the SU proteins of the isolated unique envelope regions of the unique envelope glycoproteins identified that most closely resembled FeLV-B/Gardner Arnstein. The arrow indicates the start of the SU protein and a dot indicates a gap in alignment. The envelope glycoprotein amino acid sequences of both FeLV-B/Gardner Arnstein and FeLV-B/Snyder-Theilen have been published previously. To maintain the nucleotide alignment, an adenine on position 380 in the published FeLV-B/Snyder-Theilen amino acid sequence has been deleted.*

(Opposite page)


↓ Start codon of SU protein
VRA
VRB
FZ215-c7
FZ215-12
FeLVBGAenv
FeLVBSTenv

```
          501                                                                                                 600
  FZ215-c7  TTGCAATTTA CCCAAAAGGG AAGACAAACA TCTTGGGATG GACCTAAGTC ATGGGGGCTA CGACTATACC GTTCAGGATA TGACCCTATA GCCCTGTTCT
  FZ215-12  TTGCAATTTA CCCAAAAGGG AAGACAAACA TCTTGGGATG GACCTAAGTC ATGGGGGCTA CGACTATACC GTTCAGGATA TGACCCTATA GCCCTGTTCT
FeLVBGAenv  TTGCAATTTA CCCAAAAGGG AAGACAAACA TCTTGGGATG GACCTAAGTC ATGGGGGCTA CGACTATACC GTTCAGGATA TGACCCTATA GCCCTGTTCT
FeLVBSTenv  ttgcaattta cccaaaaggg aagacaaaca tcttgggatg gacctaagtc atgggggcta cgactatacc gttcaggata tgaccctata gccctgttct

          601                                                                                                 700
  FZ215-c7  CGGTATCCCG GCAAGTAATG ACCATTACGC CGCCTCAGGC CATGGGACCA AATCTAGTCC TGCCTGATCA AAAACCCCCA TCCAGGCAAT CTCAAATAGA
  FZ215-12  CGGTATCCCG GCAAGTAATG ACCATTACGC CGCCTCAGGC CATGGGACCA AATCTAGTCC TGCCTGATCA AAAACCCCCA TCCAGGCAAT CTCAAATAGA
FeLVBGAenv  CGGTATCCCG GCAAGTAATG ACCATTACGC CGCCTCAGGC CATGGGACCA AATCTAGTCC TGCCTGATCA AAAACCCCCA TCCAGGCAAT CTCAAATAGA
FeLVBSTenv  cggtatcccg gcaagtaatg accattacgc cgcctcaggc catgggacca aatctagtcc tgcctgatca aaaacccccа tccaggcaat ctcaaataga

          701                                                                                                 800
  FZ215-c7  GTCCCGAGTA ACACCTCACC ATTCCCAAGG CAACGGAGGC ACCCCAGGTA TAACTCTTGT TAATGCCTCC ATTGCCCCTC TAAGTACCCC TGTCCCCCCG
  FZ215-12  GTCCCGAGTA ACACCTCACC ATTCCCAAGG CAACGGAGGC ACCCCAGGTA TAACTCTTGT TAATGCCTCC ATTGCCCCTc TAAGTACCCC TGTCACCCCC
FeLVBGAenv  GTCCCGAGTA ACACCTCACC ATTCCCAAGG CAACGGAGGC ACCCCAGGTA TAACTCTTGT TAATGCCTCC ATTGCCCCTC TAAGTACCCC TGTCACCCCC
FeLVBSTenv  gtcccgagta acacctcacc attcccaagg caacggaggc accccaggta taactcttgt taatgcctcc attgcccctc taagtacccc tgtcacccccc

          801                                                                                                 900
  FZ215-c7  CCAAGTCTCA AACGGATTGG GACCGGAGAC AGGTTAATAA ATTTAGTACA AGGGACATAC CTAGCCTTAA ATGCCACCGA CCCCAACAAA ACTAAAGACT
  FZ215-12  GCAAGTCCCA AACGGATTGG GACCGGAGAC AGGTTAATAA ATTTAGTACA AGGGACATAC CTAGCCTTAA ATGCCACCGA CCCCAACAAA ACTAAAGACT
FeLVBGAenv  GCAAGTCCCA AACGGATTGG GACCGGAGAT AGGTTAATAA ATTTAGTACA AGGGACATAC CTAGCCTTAA ATGCCACCGA CCCCAACAGA ACTAAAGACT
FeLVBSTenv  gcaagtccca aacgtatagg gacaggaaat aggttaataa atttagtgca ggggacatat ctagctttaa atgtcactaa ccccaacaaa actaaagact

          901                                                                                                1000
  FZ215-c7  GTTGGCTCTG CCTGGTTTCT CGACCACCCT ATTACGAAGG GATTGCAATC TTAGGTAACT ATAGCAACCA AACAAACCCC CCCCCATCCT GCCTATCTAT
  FZ215-12  GTTGGCTCTG CCTGGTTTCT CGACCACCCT ATTACGAAGG GATTGCAATC TTAGGTAACT ATAGCAACCA AACAAACCCC CCCCCATCCT GCCTATCTAT
FeLVBGAenv  GTTGGCTCTG CCTGGTTTCT CGACCACCCT ATTACGAAGG GATTGCAATC TTAGGTAACT ACAGCAACCA AACAAACCCC CCCCCATCCT GCCTATCTAT
FeLVBSTenv  gttggctctg tctagtctcc cgaccgccct attatgaagg aattgcggta ttgggcaatt acagcaacca aaccaacccc cccccatcct gcctatccga

          1001                                                                                               1100
  FZ215-c7  TCCGCAACAC AAACTGACCA TATCTGAAGT ATCAGGGCAA GGACTGTGCA TAGGGACTGT TCCTAAGACC CACCAGGCTT TGTGCAATGA GACGCAACAG
  FZ215-12  TCCGCAACAC AAACTGACCA TATCTGAAGT ATCAGGGCAA GGACTGTGCA TAGGGACTGT TCCTAAGACC CACCAGGCTT TGTGCAATGA GACGCAACAG
FeLVBGAenv  TCCGCAACAC AAACTAACCA TATCTGAAGT ATCAGGGCAA GGACTGTGCA TAGGGACTGT TCCTAAGACC CACCAGGCTT TGTGCAATGA GACACAACAG
FeLVBSTenv  cccacaacat aaactgacta tatcagaagt gtccgggcaa ggtttgtgca tagggactgt tcctaagacc caccaagctt tgtgcaaaaa gacacaaaaa
```

```
            1101                                                                                                1200
  FZ215-c7  GGACATACAG GGGCACACTA TCTAGCCGCC CCCAACGGCA CCTATTGGGC TTGTAACACT GGACTCACCC CATGCATTTC CATGGCGGTG CTCAATTGGA
  FZ215-12  GGACATACAG GGGCACACTA TCTAGCCGCC CCCAACGGCA CCTATTGGGC TTGTAACACT GGACTCACCC CATGCATTTC CATGGCGGTG CTCAATTGGA
FeLVBGAenv  GGACATACAG GGGCGCACTA TCTAGCCGCC CCCAATGGCA CCTATTGGGC CTGTAACACT GGACTCACCC CATGTATTTC CATGGCGGTG CTCAATTGGA
FeLVBSTenv  ggacataaag ggactcacta tctagcagcc cctagcggta cctactgggc atgcaacacc gggctaatcc catgcatttc catggcagtg ctcaattgga

            1201                                                                                                1300
  FZ215-c7  CCGCTGATTT TTGTGTCTTA ATCGAATTAT GGCCCAGAGT GACTTACCAT CAACCCGAAT ATGTGTACAC ACATTTTGCC AAAGCTGTCA GGTTCCGAAG
  FZ215-12  CCGCTGATTT TTGTGTCTTA ATCGAATTAT GGCCCAGAGT GACTTACCAT CAACCCGAAT ATGTGTACAC ACATTTTGCC AAAGCTGTCA GGTTCCGAAG
FeLVBGAenv  CCTCTGATTT TTGTGTCTTA ATCGAATTAT GGCCCAGAGT GACTTACCAT CAACCCGAAT ATGTGTACAC ACATTTTGCC AAAGCTGCCA GGTTCCGAAG
FeLVBSTenv  cctctgattt ttgtgtccta atcgaattgt ggcccagagt gacctaccat caacccgaat atgttacac acatttcgac aaaactgtca ggctccggag

            1301                                                                                                1400
  FZ215-c7  AGAACCAGTA TCACTAACTG TTGCCCTTAT GTTGGGAGGA CTCACTGTAG GGGGCATAGC CGCGGGGGTC GGAACAGGGA CTAAAGCTCT CCTTGAAACA
  FZ215-12  AGAACCAGTA TCACTAACTG TTGCCCTTAT GTTGGGAGGA CTCACTGTAG GGGGCATAGC CGCGGGGGTC GGAACAGGGA CTAAAGCCCT CCTTGAAACA
FeLVBGAenv  AGAACCAATA TCACTAACTG TTGCCCTCAT GTTGGGAGGA CTCACTGTAG GGGGCATAGC CGCGGGGGTC GGAACAGGGA CTAAAGCCCT CATTGAAACA
FeLVBSTenv  agaaccaata tcactaaccg ttgcccttat gttaggagga ctcactgtag ggggcatagc cgcaggggtc ggaacgggga ctaaagccct cctcgaaaca

            1401                                                                                                1500
  FZ215-c7  GCCCAGTTCA GACAACTACA AATGGCCATG CACACAGACA TCCAAGCCCT AGAAGAGTCA ATTAGTGCCT TAGAAAAGTC TCTGACCTCC CTTTCTGAAG
  FZ215-12  GCCCAGTTCA GACAACTACA AATGGCCATG CACACAGACA TCCAAGCCCT AGAAGAGTCA ATTAGTGCCT TAGAAAAGTC TCTGACCTCC CTTTCTGAAG
FeLVBGAenv  GCCCAGTTCA GACAACTACA AATGGCCATG CACACAGACA TCCAGGCCCT AGAAGAGTCA ATTAGTGCCT TAGAAAAGTC CCTGACCTCC CTTTCTGAAG
FeLVBSTenv  gcccagttca gacaactaca aatggccatg cacacagaca tccaggccct ggaagagtca attagtgcct tagaaaaatc cctgacctcc ctctctgagg

            1501                                                                                                1600
  FZ215-c7  TAGTCTTACA AAACAGACGG GGCCTAGATA TTCTATTCTT ACAAGAGGGA GGGCTCTGTG CCGCATTAAA AGAAGAATGT TGCTTCTATG CGGATCACAC
  FZ215-12  TAGTCTTACA AAACAGACGG GGCCTAGATA TTCTATTCTT ACAAGAGGGA GGGCTCTGTG CCGCATTAAA AGAAGAATGT TGCTTCTATG CGGATCACAC
FeLVBGAenv  TAGTCTTACA AAACAGACGG GGCCTGGATA TTCTATTCTT ACAAGAGGGA GGGCTCTGTG CCGCATTAAA AGAAGAATGT TGCTTCTATG CGGATCACAC
FeLVBSTenv  tagtcttact gaaagacccc cccccacccc gaaacttagt cagccagcta ttgcagtaat accatttccc cgaaacttag tcagccagct attgcagaaa

            1601                                                                                                1700
  FZ215-c7  CGGACTCGTC CGAGACAATA TGGCCAAATT AAGAGAAAGA CTAAAACAGC GGCAACATTT GTTTGACTCC CAACAGGGAT GGTTTGAAGG ATGGTTCAAC
  FZ215-12  CGGACTCGTC CGAGACAATA TGGCCAAATT AAGAGAAAGA CTAAAACAGC GGCAACAATT GTTTGACTCC CAACAGGGAT GGTTTGAAGG ATGGTTCAAC
FeLVBGAenv  CGGACTTGTC CGAGACAATA TGGCTAAATT AAGAGAAAGA CTAAAACAGC GGCAACAACT GTTTGACTCC CAACAGGGAT GGTTTGAAGG ATGGTTCAAC
FeLVBSTenv  gacccccccc caccccgaaa cttagccagc tattgcagta ataccatttc cccgaaactt agtcagccag ctattgcagt aataccattt ccccgaaact
```

```
          1701                                                                                                1800
  FZ215-c7 AAGTCCCCCT GGTTTACAAC CCTAATTTCC TCCATTATGG GCCCCTTACT AATCCTACTC CTAATTCTCC TCTTCGGCCC ATGCATCCTT AACCGATTAG
  FZ215-12 AAGTCCCCCT GGTTTACAAC CCTAATCTCC TCCATTATGG GCCCCTTACT AATCCTACTC CTAATTCTCC TCTTCGGCCC ATGCATCCTT AACCGATTAG
FeLVBGAenv AAGTCCCCCT GGTTTACAAC CCTAATTTCC TCCATTATGG GCCCCTTACT AATCCTACTC CTAATTCTCC TCTTCGGCCC ATGCATCCTT AACAGATTAG
FeLVBSTenv tagtcagcca gctattgcag taataccatt tcacaaggca tggaaaatta cccaagcatg ttcccataag atataaggaa gttagaatt

          1801                                                                                          1897
  FZ215-c7 TACAATTCGT AAAAGACAGA ATATCTGTGG TACAAGCTTT AATTTTAACC CAACAGTACC AACAGATAAA GCAATACGAT CCGGACCGAC CATGAT
  FZ215-12 TACAATTCGT AAAAGACAGA ATATCTGTGG TACAAGCTTT AATTTTAACC CAACAGTACC AACAGATAAA GCAATACGAT CCGGACCGAC CATGAT
FeLVBGAenv TACAATTCGT AAAAGACAGA ATATCTGTGG TACAAGCCTT AATTTTAACC CAACAGTACC AACAGATAAA GCAATACGAT CCGGACCGAC CATGAT
FeLVBSTenv
```

FA27

Clone FA27-53 appeared to be the A-component of FA27, however FA27-53 differed from FeLV-A/Glasgow-1 at six amino acids. As these amino acids were also present in the representative FA27C amino acid sequence, the data suggested that these amino acids were indeed representative of the FA27 isolate. Clones FA27-17 and FA27-55 appeared to be C components of the FA27 isolate. When compared with the A-component of FA27 (FA27-53), there were three amino acid differences outwith the VRA region that were not present in clone FA27-53. A $Q_{218}R$ substitution was present in both unique *env* clones and in the FA27 sequence published previously, therefore this substitution can be considered representative. There were also two amino acid substitutions present in both clones that were not present in the sequence published previously; these were $I_{301}V$ and $I_{318}D$. The latter was a change from a hydrophobic residue without polarity (isoleucine) to an acidic negatively charged residue (aspartate). There were two changes in the VRA region compared with the published FA27 sequence, these were $F_{62}L$ and $S_{67}R$. The first substitution yields an amino acid with an aromatic side chain while the latter substitution resulted in a change in polarity of the residue from an uncharged polar residue (serine) to a positively charged polar residue (arginine).

FS246

When FS246-40, the A-component of the primary isolate FS246, was compared to the published FeLV-A/Glasgow-1 sequence, three amino acid changes were found. Two of these amino acids ($Q_{140}P$ and $A_{248}G$) were present in the representative FS246 sequence and thus they were considered to be consistent with FS246, whereas one substitution, $D_{150}G$, was unique to FS246-40. The two *env* clones thought to be C components of primary isolate FS246, FS246-4 and FS246-37, were very similar to one another, however there were differences between these two *env* clones. Clone FS246-37 was identical to the representative FS246 sequence, with the exception of a three amino acid stretch at the start of the mature peptide; the published sequence has a $T_{24}P_{25}K_{26}$, whereas the clone identified has a similar sequence to FeLV-A, namely $N_{24}T_{25}Q_{26}$. This amino acid stretch appeared to be a loop that links β-sheets and also resulted in an increase in the overall polarity; therefore the change in this stretch of sequence could alter the structure and subsequent function of the protein.

Clone FS246-4, which also appears to be a C component, was very similar to the published FS246 sequence, with the exception of a $S_{64}N$ substitution in the centre of the VRA region (similar to the sequence found in the published FeLV-A/Glasgow-1 sequence) and a $S_{290}F$ substitution, neither of which resulted in a change in polarity of the overall amino acid sequence.

FX343

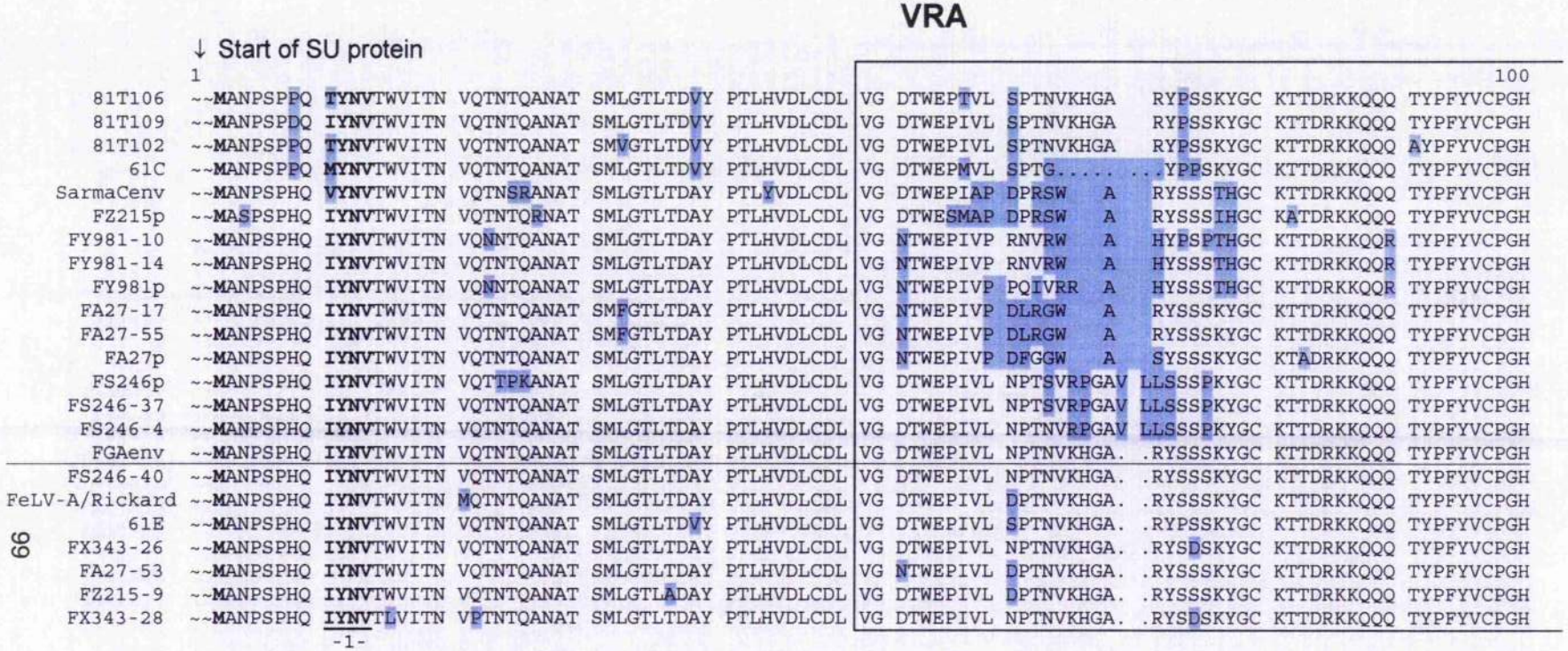
Both clones characterised for primary isolate FX343 were similar to FeLV-A/Glasgow-1, suggesting that they were the A component of this primary isolate. As the parent FX343 sequence was not known, it was difficult to identify residues in the amino acid sequence that conferred the tropism of the clones characterised. There were four amino acid changes between the A-components and FeLV-A/Glasgow-1, suggesting these residues were unique to the primary isolate. However, there were an additional six changes present in FX343-28 when FX343-26 and FX343-28 were compared to one another, these were; $Q_{22}P$, $T_{186}A$, $N_{240}T$, $T_{380}A$, $S_{469}T$ and $A_{511}V$, none of these amino acid substitutions altered the polarity of the sequence. However, the amino acid substitution $N_{240}T$ resulted in the loss of a glycosylation site. Interestingly, there was also a change in the FX343-26 sequence compared to FeLV-A, namely $Q_{224}K$, which was not found in the FX343-28 amino acid sequence, altering the polarity from uncharged to positively charged.

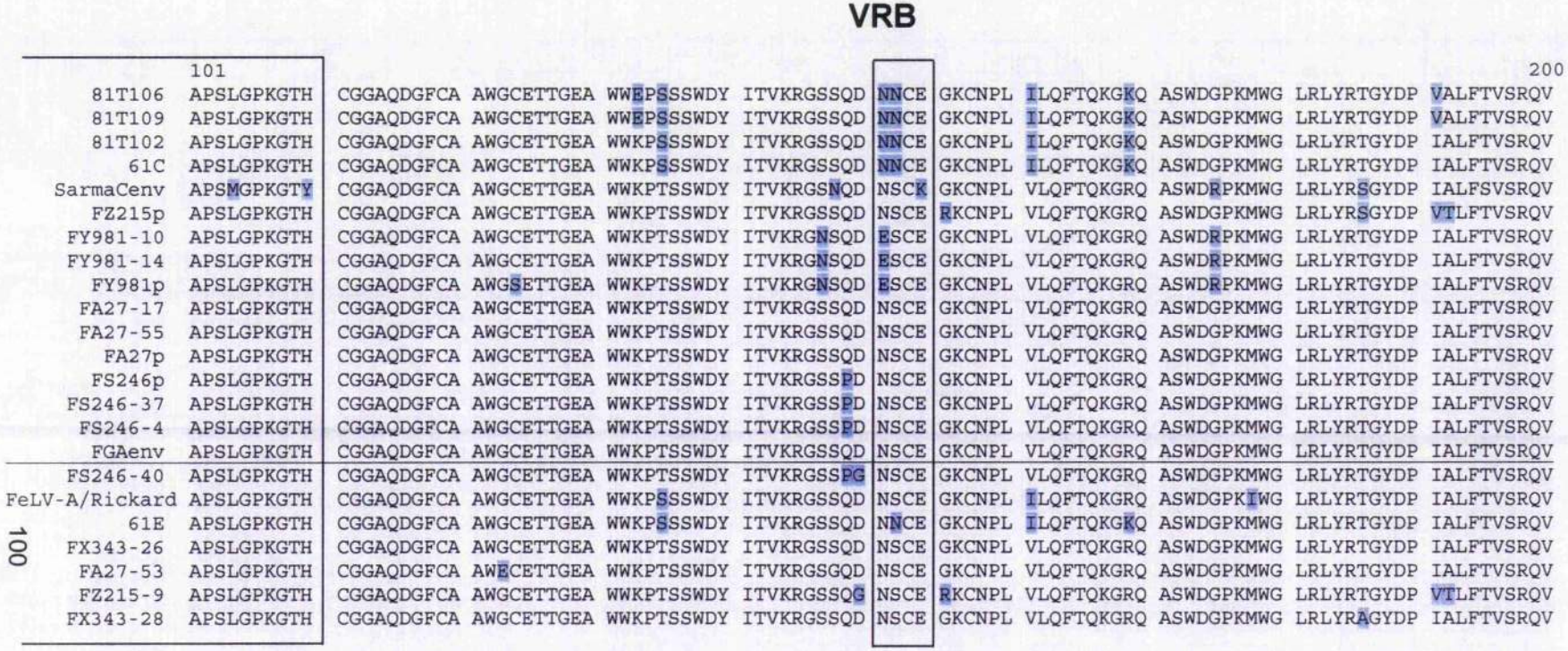
FY981

The *env* clones identified from isolate FY981 were very similar to the FY981p amino acid sequence. Indeed, when compared to the FY981p amino acid sequence, there was only one amino acid difference in *env* clone FY981-14 (termed "variant I"). The threonine residue present at position 23 was also present in the FeLV-A/Glasgow-1 amino acid sequence. Clone FY981-10 (termed "variant II") differed from FY981-14 and FY981p at a stretch of three amino acids; $P_{74}S_{75}P_{76}$ in contrast to $S_{74}S_{75}S_{76}$. The polarity of the $P_{74}S_{75}P_{76}$ and $S_{74}S_{75}S_{76}$ is the same. However, although this variation in the amino acid sequence may not alter the overall polarity of this region of the protein, proline residues are predominantly found in the folds of a protein chain, therefore the presence of these proline residues may significantly alter the structure of the protein.

FZ215

Isolate FZ215 consisted of two unique *env* clones, FZ215-9 and FZ215-12. The A component, FZ215-9, was virtually identical to the previously described FZ215 sequence. However, there were four amino acids that were not present in the representative FZ215 sequence, these were $D_{150}G$, $I_{392}A$, $I_{598}T$ and $S_{599}I$, these amino acid changes should not affect the overall polarity of the protein. These amino acids were similar to those found in the FeLV-A/Glasgow-1 Env, suggesting that these amino acids were characteristic of the A component of this isolate. A separate nucleotide sequence alignment and amino acid sequence alignment was performed for the isolates that resembled FeLV-B Gardner-Arnstein; these are shown in figures 3.5 and 3.7. The majority of unique amino acids were also found in the representative FZ215 clone, with the exception of two changes namely $L_{41}I$ just prior to the VRA region and $T_{165}I$ (compared to the FeLV-B/Gardner Arnstein sequence) in the VRB region. $T_{165}I$ results in a change of polarity from uncharged polar to hydrophobic non-polar, a change that might result in a change in structure.


VRA
↓ Start of SU protein
1
100
81T106 ~~MANPSPPQ TYNVTWVITN VQTNTQANAT SMLGTLTDVY PTLHVDLCDL VG DTWEPTVL SPTNVKHGA RYPSSKYGC KTTDRKKQQQ TYPFYVCPGH
81T109 ~~MANPSPDQ IYNVTWVITN VQTNTQANAT SMLGTLTDVY PTLHVDLCDL VG DTWEPIVL SPTNVKHGA RYPSSKYGC KTTDRKKQQQ TYPFYVCPGH
81T102 ~~MANPSPPQ TYNVTWVITN VQTNTQANAT SMVGTLTDVY PTLHVDLCDL VG DTWEPIVL SPTNVKHGA RYPSSKYGC KTTDRKKQQQ AYPFYVCPGH
61C ~~MANPSPPQ MYNVTWVITN VQTNTQANAT SMLGTLTDVY PTLHVDLCDL VG DTWEPMVL SPTG..... ...YPPSKYGC KTTDRKKQQQ TYPFYVCPGH
SarmaCenv ~~MANPSPHQ VYNVTWVITN VQTNSRANAT SMLGTLTDAY PTLYVDLCDL VG DTWEPIAP DPRSW A RYSSSTHGC KTTDRKKQQQ TYPFYVCPGH
FZ215p ~~MASPSPHQ IYNVTWVITN VQTNTQRNAT SMLGTLTDAY PTLHVDLCDL VG DTWESMAP DPRSW A RYSSSIHGC KATDRKKQQQ TYPFYVCPGH
FY981-10 ~~MANPSPHQ IYNVTWVITN VQNNTQANAT SMLGTLTDAY PTLHVDLCDL VG NTWEPIVP RNVRW A HYPSPTHGC KTTDRKKQQR TYPFYVCPGH
FY981-14 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG NTWEPIVP RNVRW A HYSSSTHGC KTTDRKKQQR TYPFYVCPGH
FY981p ~~MANPSPHQ IYNVTWVITN VQNNTQANAT SMLGTLTDAY PTLHVDLCDL VG NTWEPIVP PQIVRR A HYSSSTHGC KTTDRKKQQR TYPFYVCPGH
FA27-17 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMFGTLTDAY PTLHVDLCDL VG NTWEPIVP DLRGW A RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
FA27-55 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMFGTLTDAY PTLHVDLCDL VG NTWEPIVP DLRGW A RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
FA27p ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG NTWEPIVP DFGGW A SYSSSKYGC KTADRKKQQQ TYPFYVCPGH
FS246p ~~MANPSPHQ IYNVTWVITN VQTTPKANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTSVRPGAV LLSSSPKYGC KTTDRKKQQQ TYPFYVCPGH
FS246-37 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTSVRPGAV LLSSSPKYGC KTTDRKKQQQ TYPFYVCPGH
FS246-4 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTNVRPGAV LLSSSPKYGC KTTDRKKQQQ TYPFYVCPGH
FGAenv ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTNVKHGA. .RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
FS246-40 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTNVKHGA. .RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
FeLV-A/Rickard ~~MANPSPHQ IYNVTWVITN MQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL DPTNVKHGA. .RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
61E ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDVY PTLHVDLCDL VG DTWEPIVL SPTNVKHGA. .RYPSSKYGC KTTDRKKQQQ TYPFYVCPGH
FX343-26 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTNVKHGA. .RYSDSKYGC KTTDRKKQQQ TYPFYVCPGH
FA27-53 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VG NTWEPIVL DPTNVKHGA. .RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
FZ215-9 ~~MANPSPHQ IYNVTWVITN VQTNTQANAT SMLGTLADAY PTLHVDLCDL VG DTWEPIVL DPTNVKHGA. .RYSSSKYGC KTTDRKKQQQ TYPFYVCPGH
FX343-28 ~~MANPSPHQ IYNVTLVITN VPTNTQANAT SMLGTLTDAY PTLHVDLCDL VG DTWEPIVL NPTNVKHGA. .RYSDSKYGC KTTDRKKQQQ TYPFYVCPGH
-1-
66


VRB
101
200
100
81T106 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWEPSSSWDY ITVKRGSSQD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP VALFTVSRQV
81T109 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWEPSSSWDY ITVKRGSSQD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP VALFTVSRQV
81T102 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSSQD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
61C APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSSQD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
SarmaCenv APSMGPKGTY CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSNQD NSCK GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRSGYDP IALFSVSRQV
FZ215p APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE RKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRSGYDP VTLFTVSRQV
FY981-10 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRTGYDP IALFTVSRQV
FY981-14 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRTGYDP IALFTVSRQV
FY981p APSLGPKGTH CGGAQDGFCA AWGSETTGEA WWKPTSSWDY ITVKRGNSQD ESCE GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRTGYDP IALFTVSRQV
FA27-17 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA27-55 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA27p APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FS246p APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FS246-37 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FS246-4 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FGAenv APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FS246-40 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSPG NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FeLV-A/Rickard APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSSQD NSCE GKCNPL ILQFTQKGRQ ASWDGPKIWG LRLYRTGYDP IALFTVSRQV
61E APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSSQD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FX343-26 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA27-53 APSLGPKGTH CGGAQDGFCA AWECETTGEA WWKPTSSWDY ITVKRGSGQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FZ215-9 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQG NSCE RKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP VTLFTVSRQV
FX343-28 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSSQD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRAGYDP IALFTVSRQV

```
               201                                                                                                   300
     81T106    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN KSAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     81T109    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN KSAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     81T102    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN KSAPRSVAPA TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
        61C    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
  SarmaCenv    MTITPPQAMG PNLVLPDQKP PSRQSQTKSK VTTQ.RPQIT SSTPRSVASA TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FZ215p    STITPPQAMG PNLVLPNQKP PSRQSQTGSK VATQMRPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
   FY981-10    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMSPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
   FY981-14    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMSPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FY981p    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMSPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
    FA27-17    STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
    FA27-55    STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
      FA27p    STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FS246p    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVGPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
  FS246-37     STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVGPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
   FS246-4     STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVGPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV FRPPYYEGIA
     FGAenv    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
-----------------------------------------------------------------------------------------------------------------------------
  FS246P-40    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVGPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
FeLV-A/Rickard SAITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RLQTT ESAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
        61E    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
   FX343-26    STITPPQAMG PNLVLPDQKP PSRKSQTGSK VATQ.RPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGVA
    FA27-53    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
    FZ215-9    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
   FX343-28    STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTT ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
```

-2-

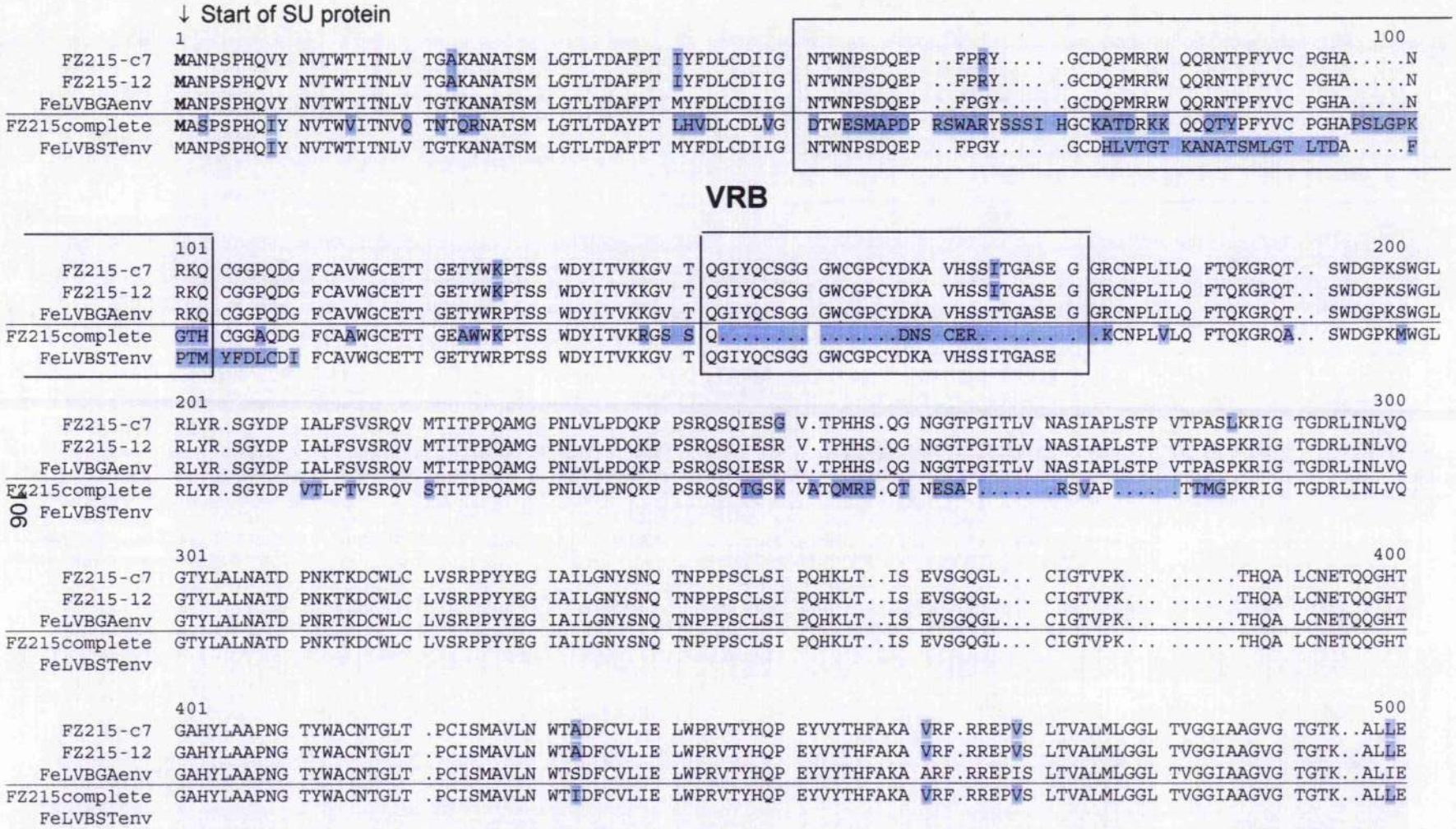
```
               301                                                                                                        400
        81T106 ILGNYSNHTN PPLSCLSIPR HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGAP..G ESQYLAAPNG TYWACNTGLT PCISMAVLNL TSDFCVLIEL
        81T109 ILGNYSNHTN PPLSCLSIPR HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGAP..G ESQYLAAPNG TYWACNTGLT PCISMAVLNL TSDFCVLIEL
        81T102 IVGNYSNHTY PPLSCLSIPR HKVTISEVSW QGLCIGTVPK THQALCNKTQ QGHTGAP..G ESLYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
           61C ILGNYSNQTN PPPSCISIPP HKLTISEVSG QGLCIGTVPK THQALCNKTH QGHTGADYLT APRYLAAPNG TYWACNTGLT PCISMAVLNL TSDFCVLIEL
     SarmaCenv VLGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
        FZ215p ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TIDFCVLIEL
      FY981-10 ILGNYSNQTN PPPSCQSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
      FY981-14 ILGNYSNQTN PPPSCQSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
        FY981p ILGNYSNQTN PPPSCQSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
       FA27-17 VLGNYSNQTN PPPSCLSDPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPSG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
       FA27-55 VLGNYSNQTN PPPSCLSDPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPSG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
         FA27p ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
        FS246p ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
      FS246-37 ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
       FS246-4 ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
        FGAenv ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGMCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
-----------------------------------------------------------------------------------------------------------------------------
      FS246-40 ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG PYWACNTGLT PCISMAVLNW TSDFCVLIEL
FeLV-A/Rickard ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG AYWACNTGLT PCISMAVLNW TSDFCVLIEL
           61E ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
      FX343-26 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
       FA27-53 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
       FZ215-9 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TADFCVLIEL
      FX343-28 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLA PCISMAVLNW TSDFCVLIEL
```

↓ Start of p15E protein

```
                401                                                                                                 500
      81T106    WPRVTYHQPE YVYTHFAKAG RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LERSLTSLSE VVLQNRRGLD
      81T109    WPRVTYHQPE YVYTHFAKAG RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LERSLTSLSE VVLQNRRGLD
      81T102    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQL*MAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
        61C     WPRVTYHQPE YVYTHFAKAG RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
   SarmaCenv    WPRVTYHQPE YIYTHFDKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQIAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FZ215p    WPRVTYHQPE YVYTHFAKAV RFRREPVSLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FY981-10    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FY981-14    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FY981p    WPRVTYHQPE YVYTHFAKAV RFRRERISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     FA27-17    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     FA27-55    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA27p     WPRVTYHQPE YVYTHFAEAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     FS246p     WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FS246-37    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     FS246-4    WPRVTYHQPE YVYTHFAKAV RFRREPIPLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FGAenv    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FS246-40    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
FeLV-A/Rickard  WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
        61E     WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FX343-26    WPRVTYHQPE YVYTHFAEAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     FA27-53    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     FZ215-9    WPRVTYHQPE YVYTHFAKAV RFRREPVSLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FX343-28    WPRVTYHQPE YVYTHFAEAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLTE VVLQNRRGLD
```

```
                501                                                                                                 600
      81T106    ILFLQEGGLC ATLKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFQGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      81T109    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFQGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      81T102    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
        61C     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
   SarmaCenv    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LLGPCILNRL VQFVKDRISV
      FZ215p    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG GFEGWFNKSP RFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
    FY981-10    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
    FY981-14    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FY981p     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA27-17    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIV GPLLILLLIL LFGPCILNRL VQSVKDRISV
     FA27-55    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      FA27p     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FS246p     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
    FS246-37    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FS246-4    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FGAenv     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
    FS246-40    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
FeLV-A/Rickard  ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRELLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
        61E     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
    FX343-26    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFIIDRISV
     FA27-53    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FZ215-9    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRTIV
    FX343-28    ILFLQEGGLC VALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFIRDRISV
```

```
                601                 622
      81T106    VQALILTQQY QQIKQYDPDR PD
      81T109    VQALILTQQY QQIKQYDPDR P~
      81T102    VQALILTQQY QQIKQYDPDR P*
        61C     VQALILTQQY QQIKQYDPDR P~
   SarmaCenv    VQALILTQQY QQIQQYDSDR P*
      FZ215p    VQALILTQQY QQIKQYDPDR P*
    FY981-10    VQALILTQQY QQIKQYDPDR P*
    FY981-14    VQALILTQQY QQIKQYDPDR P*
      FY981p    VQALILTQQY QQIKQYDPDR P*
     FA27-17    VQALILTQQY QQIRQYDPDR P*
     FA27-55    VQALILTQQY QQIRQYDPDR P*
       FA27p    VQALILTQQY QQIRQYDPDR P*
      FS246p    VQALILTQQY QQIKQYDPDR P*
    FS246-37    VQALILTQQY QQIKQYDPDR P*
     FS246-4    VQALILTQQY QQIKQYDPDR P*
      FGAenv    VQALILTQQY QQIKQYDPDR P*
    FS246-40    VQALILTQQY QQIKQYDPDR P*
FeLV-A/Rickard  VQALILTQQY QQIKQYDPDR P*
        61E     VQALILTQQY QQIKQYDPDR P~
    FX343-26    VQALILTQQY QQIKQYDPDR P~
     FA27-53    VQALILTQQY QQIRQYDPDR P*
     FZ215-9    VQALILTQQY QQIKQYDPDR P*
    FX343-28    VQALILTQQY QQIKQYDPDR P*
```


↓ Start of SU protein
VRB
FZ215-c7
FZ215-12
FeLVBGAenv
FZ215complete
FeLVBSTenv

```
                501                                                                                                 600
     FZ215-c7   TAQFRQLQMA MHTDIQALEE SISALEKSLT SLSE...VVL QNRRGLDILF LQEGGLCAAL KEECCFYADH TGLVRDNMAK LRERLKQRQH LFDSQQGWFE
     FZ215-12   TAQFRQLQMA MHTDIQALEE SISALEKSLT SLSE...VVL QNRRGLDILF LQEGGLCAAL KEECCFYADH TGLVRDNMAK LRERLKQRQQ LFDSQQGWFE
   FeLVBGAenv   TAQFRQLQMA MHTDIQALEE SISALEKSLT SLSE...VVL QNRRGLDILF LQEGGLCAAL KEECCFYADH TGLVRDNMAK LRERLKQRQQ LFDSQQGWFE
FZ215complete  TAQFRQLQMA MHTDIQALEE SISALEKSLT SLSE...VVL QNRRGLDILF LQEGGLCAAL KEECCFYADH TGLVRDNMAK LRERLKQRQQ LFDSQQGGFE
   FeLVBSTenv

                601                                                                                         670
     FZ215-c7   GWFNKSPWFT TLISSIMGPL LILLLILLFG PCILNRLVQF VKDRISVVQA LILTQQYQQI KQYDPDRP*
     FZ215-12   GWFNKSPWFT TLISSIMGPL LILLLILLFG PCILNRLVQF VKDRISVVQA LILTQQYQQI KQYDPDRP*
   FeLVBGAenv   GWFNKSPWFT TLISSIMGPL LILLLILLFG PCILNRLVQF VKDRISVVQA LILTQQYQQI KQYDPDRP*
FZ215complete  GWFNKSPRFT TLISSIMGPL LILLLILLFG PCILNRLVQF VKDRISVVQA LILTQQYQQI KQYDPDRP*
   FeLVBSTenv
```

3.3.2 Assaying the production of mature envelope glycoproteins by immunofluorescence

In order to assess whether the novel *env* genes produced polypeptides that were folded correctly, 293T cells were transfected with VR1012-env constructs, fixed with ice-cold methanol and stained using 6-15, a monoclonal antibody recognising the FeLV envelope glycoprotein. FeLV-A/Glasgow-1, FeLV-B/Gardner Arnstein and FeLV-C/Sarma were included as controls. The results of the antibody staining are shown in figure 3.8. From the twelve novel *env* gene products assessed, nine produced envelope glycoproteins that could be detected with the 6-15 antibody, namely FA27-17, FA27-53, FA27-55, FS246-4, FS246-37, FS246-40, FY981-14, FX343-28 and FZ215-9. The remaining three did not produce envelope glycoproteins that were recognised by the 6-15 antibody, namely FZ215-12, FX343-26 and FY981-10. Cells that were successfully transfected were clearly distinguishable as the fluorescence localised on the membrane of these cells, indicating the production and export of the envelope glycoprotein from the cells. The three *env* gene products that were unable to produce a protein that was recognised by the antibody were indistinguishable from the background cells, as can be seen in figure 3.8. To ensure the presence of cells in the field examined, a second image was recorded using bright field microscopy, the results of which can be seen in figure 3.8. As there were cells present in the field photographed, the negative antibody staining was due to the *env* constructs producing a protein that was no longer recognised by the antibody. Two possible explanations for the production of a protein that is not recognised by the antibody are firstly, mutations in the amino acid sequence disrupt the conformational structure of the protein and therefore the protein is no longer recognised by the antibody. Secondly, the mutations in the amino acid sequence are located in the antibody-binding site, therefore the antibody is no longer able to recognise the protein.

FA27-17

FA27-53

FA27-55

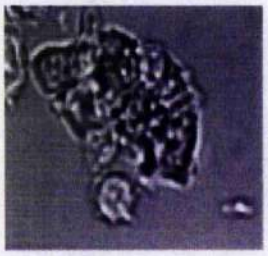
FS246-4

FS246-37

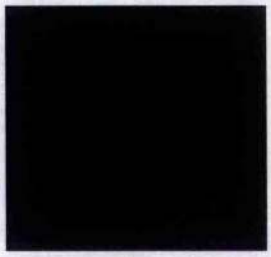
FS246-40

FZ215-9

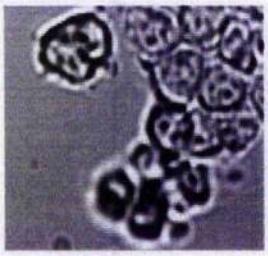
FZ215-12



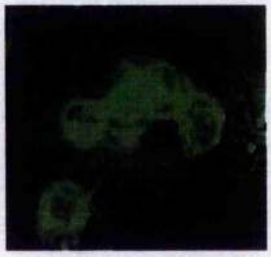
FZ215-12



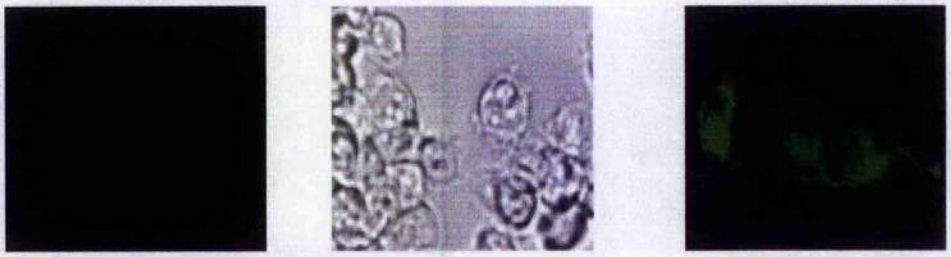
FX343-26



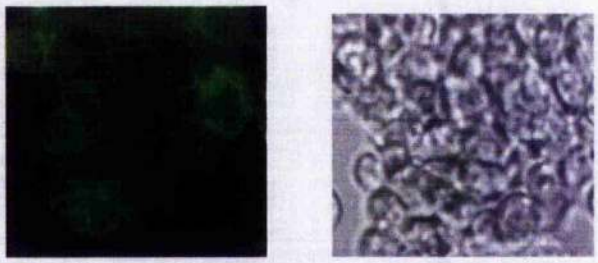
FX343-26



FX343-28



FY981-10 FY981-10 FY981-14



FeLV-A/Glasgow-1 control

FY981 control FeLV-C/Sarma control

FeLV-B/Gardner Arnstein control

Figure 3.8: *Evaluation of the expression of the mature envelope glycoprotein. The 293T cells that were transiently transfected with the FeLV env constructs, a plasmid containing MuLV gag-pol and a marker gene were stained with the anti-gp70 monoclonal antibody 6-15 and examined by immunofluorescence.*

3.3.3 Cell tropism of the FeLV *env*/MuLV *gag-pol* pseudotypes

FeLV *env*/MuLV *gag-pol* pseudotypes were prepared by transfection of 293T cells with the VR1012 *env* constructs. The supernatants containing the pseudotypes were filtered through a 0.45μM filter and dilutions were prepared and subsequently plated onto a range of cell lines to determine the titres and *in vitro* cell tropism, the results of which are summarised in table 3.5. Pseudotypes bearing the FeLV-A, FeLV-B and FeLV-C *env* proteins were capable of infecting the cell lines predicted by previous studies, with FeLV-A having the most restricted *in vitro* cell tropism and FeLV-C the broadest *in vitro* cell tropism (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). FeLV-A infected the feline cell line HO6T1 and Mink cells to titres of 10^3 CFU/ml. FeLV-B infected the feline cell line HO6T1 ($1.1x10^3$ CFU/ml) and in agreement with previous studies, the human cell line HeLa ($6.8x10^3$ CFU/ml), the canine cell line MDCK ($1.2x10^4$ CFU/ml) and the porcine cell line ST ($1.5x10^2$ CFU/ml) (previous studies have shown that porcine cells are permissive for FeLV-B infection but not FeLV-A or FeLV-C infection (Hardy W.G.Jr., 1993)). FeLV-B/Gardner Arnstein infected Mink cells poorly (26 CFU/ml). Similarly, FeLV-C infected the feline cell line HO6T1 ($4.6x10^2$ CFU/ml) and in agreement with previous studies, the human cell line HeLa ($5.7x10^2$ CFU/ml) and the guinea pig cell line 104C1 ($1.5x10^3$ CFU/ml) (guinea pig cells are known to be permissive to FeLV-C infection alone (Hardy W.G.Jr., 1993)). The titres of the FeLV-C/Sarma control were lower than the titres for the other controls included in the studies, on average the titres were five to ten fold lower than other titres (for example HO6T1; FeLV-A $7.4x10^3$ CFU/ml, FeLV-B $1.1x10^3$ CFU/ml and FeLV-C $4.6x10^2$ CFU/ml). In agreement with previous studies, no FeLV subgroups were capable of infecting the murine cell line 3T3 (Jarrett O. et al., 1973).

The constructs, which were not recognised by the monoclonal antibody 6-15, namely FZ215-12, FX343-26 and FY981-10, did not produce pseudotypes capable of infecting the panel of cell lines, suggesting that the proteins formed were non-functional. An additional three constructs, FS246-4, FS246-40 and FX343-28, also did not produce infectious pseudotypes, suggesting that, although the envelope glycoprotein was synthesised, there were further changes in the envelope glycoprotein that rendered these proteins non functional.

	H06T1	104C1	Mink	HeLa CB	MDCK	ST Iowa	3T3
FS246-4	0	0	0	0	0	0	0
FS246-37	$4.3x10^3$	$5.6x10^3$	$9.3x10^3$	$2.5x10^3$	0	0	0
FS246-40	0	0	0	0	0	0	0
FA27-17	9	0	$4.6x10^2$	0	0	0	0
FA27-53	$9.6x10^3$	0	$5.6x10^3$	0	0	$1.5x10^2$	0
FA27-55	$9.3x10^3$	$1.0x10^4$	$1.5x10^4$	$7.4x10^3$	0	0	0
FY981-10	0	0	0	0	0	0	0
FY981-14	$7.1x10^3$	$1.7x10^4$	$1.6x10^4$	$2.5x10^3$	$5.7x10^2$	$4.6x10^2$	0
FZ215-9	$9.6x10^3$	0	$6.2x10^3$	0	0	0	0
FZ215-12	0	0	0	0	0	0	0
FX343-26	0	0	0	0	0	0	0
FX343-28	0	0	0	0	0	0	0
CL29A	$7.4x10^3$	0	$4.0x10^3$	0	0	0	0
FeLV-B/GA	$1.1x10^3$	0	26	$6.8x10^3$	$1.2x10^4$	$1.5x10^2$	0
Sarma C	$4.6x10^2$	$1.5x10^3$	$1.9x10^3$	$5.7x10^2$	0	0	0
FY981	$8.6x10^3$	$1.9x10^4$	$2.0x10^4$	$7.1x10^3$	0	$4.1x10^2$	0

Table 3.5: *Titres (CFU/ml) of the unique env constructs established by infection of a panel of cell lines. The cell lines tested are: the feline kidney cell line HO6T1, the guinea pig cell line 1O4C1, the mink lung cell line Mv1Lu (Mink), the human cervical carcinoma cell line Hela CB, the canine kidney cell line MDCK, the porcine testes cell line ST IOWA and the murine cell line NIH 3T3.*

The *env* constructs most closely resembling FeLV-A and which were shown to produce functional *env* protein (FA27-53 and FZ215-9) were able to infect the same cell lines as the control FeLV-A construct to comparable titres ($9.6x10^3$ CFU/ml and $9.6x10^3$ CFU/ml respectively for HO6T1 infection and $5.6x10^3$ CFU/ml and $6.2x10^3$ CFU/ml respectively for Mink infection). Interestingly, FA27-53 was capable also of infecting the porcine cell line ST Iowa, a cell line thought to be permissive to FeLV-B, but not FeLV-A or FeLV-C infection, to a comparable titre as FeLV-B/Gardner Arnstein, $1.5x10^2$ CFU/ml. When the amino acid alignment (see figure 3.6) was examined, one amino acid change ($D_{53}N$) was present at the start of the VRA region in *env* clone FA27-53 and similarly present in FeLV-B/Gardner Arnstein. Therefore it is possible that this single amino acid substitution is responsible for the extended *in vitro* cell tropism of this clone. There are two possible explanations for this altered tropism of clone FA27-53. Firstly, the FeLV-A receptor present on ST cells is in an alternative conformational structure than the FeLV-A receptor present on HO6T1 cells. Secondly, similar to the proposition by Bupp et al (Bupp K. and Roth M.J., 2002) that FeLV-C is capable of utilising two receptors, there are two FeLV-A receptors and ST cells express a second FeLV-A receptor used by FA27-53. The amino acid substitution described above ($D_{53}N$) is present also in the other FA27 constructs (FA27-17, FA27-55 and FA27p), however these constructs are not capable of infecting porcine cells. This could be due to the fact that there are numerous other changes in the VRA region when compared to FeLV-A/Glasgow-1 that may affect the conformational structure of the envelope glycoprotein and subsequently prevent infection of porcine cells.

The *env* constructs that were thought to be C components and were capable of producing functional *env* protein (FA27-55, FY981-14 and FS246-37) showed a similar host range to FeLV-C/Sarma. The titres of the unique *env* clones were slightly higher than the titres of the control construct, but comparable to titres found for the other constructs. *Env* clone FA27-17, which was also thought to be the C component of primary isolate FA27, was not able to infect the guinea pig cell line 104C1 (a cell line permissive for infection with FeLV-C), however the overall titres of pseudotypes produced from this construct were at least 10 fold lower than the titres achieved with the other constructs ($1.5x10^4$ CFU/ml, $1.6x10^4$ CFU/ml and $9.3x10^3$ CFU/ml respectively compared to $4.6x10^2$ CFU/ml when comparing the titres of mink

infection). Therefore it is conceivable that the failure of pseudotypes derived from FA27-17 to infect 104C1 cells was due to the low titres generated and not due to the *env* being unable to support infection of this cell line. Interestingly, construct FY981-14 had an extended *in vitro* cell tropism and was capable of infecting not only the porcine cell line ST, but also the canine cell line MDCK ($5.7x10^2$ CFU/ml). When pseudotypes derived from the representative FY981 clone (FY981p) were plated on the same panel of cells, it was capable also of infecting the porcine cell line ST ($4.1x10^2$ CFU/ml), suggesting that this a genuine trait of this isolate. However, FY981p was not capable of infecting the canine cell line MDCK. This was considered to be a valid observation as the titres achieved when infecting other cell lines were comparable, if not slightly higher for FY981p when compared to the other FY981 clones. Therefore, an additional amino acid change present in FY981-14 and possibly FeLV-B/Gardner Arnstein, but not present in FY981p, could be responsible for the altered tropism of FY981-14. There was one amino acid difference ($N_{23}T$) present in FY981-14, relative to FY981p, but this threonine residue was present also in the published FeLV-A/Glasgow-1 amino acid sequence (FeLV-A is incapable of infecting MDCK cells). When comparing the FY981 and the FA27 constructs, the $D_{53}N$ substitution discussed for construct FA27-53 was present in both the FY981p and the FY981-14 amino acid sequence, strengthening the involvement of this single amino acid substitution. The deletion in the VRA region of the FY981 constructs was smaller than the deletion in the FA27 constructs that were considered to be C components, suggesting that the unique sequence of the FA27 VRA region is responsible for the lack of infection of ST cells.

Rather surprisingly, only few isolates (FS246-37, FA27-55 and FY981-14) were capable of infecting the human cell line HeLa. With the exception of the FeLV-B control, only isolates that were capable of infecting 104C1 cells were capable of infecting HeLa cells, leading to the conclusion that only pseudotypes that were the C component of a primary isolate were capable of infecting the human cell line HeLa. Given that previous studies have shown that FeLV-A infects the human cell line 293T, it is likely that these findings reflect differential receptor expression on 293T cells and HeLa cells.

Virus isolate	Original mixture of subgroups	*Env* clone number	Subgroup predicted by sequencing	Antibody staining using 6-15 monoclonal antibody	Subgroup established by virus infection
FA27	*A/C*	17	C	+	A or C*
		53	A	+	A?
		55	C	+	C
FS246	*A/C*	4	C	+	-
		37	C	+	C
		40	A	+	-
FX343	*A/C*	26	A	-	-
		28	A	+	-
FZ215	*A/B/C*	9	A	+	A
		12	B	-	-
FY981	*A/B/C*	10	C	-	-
		14	C	+	C

Table 3.6: *Subgroups determined of the unique envelope glycoproteins through both sequencing and infection on a panel of different cell lines.*

** Titres of pseudotypes derived from FA27-17 were too low to confirm that FA27-17 had the cell tropism of an A or C component.*

3.4 Discussion

FeLV can be classified into three main subgroups, A, B and C as determined by interference assays (Sarma P.S. and Log T., 1973). The three subgroups have a distinct cell tropism; subgroup A has the most restricted cell tropism and is capable of infecting feline cells only, whereas subgroups B and C have a broader tropism and are capable of infecting both feline and non feline cells (Hardy W.G.Jr., 1993; Jarrett O. et al., 1973; Sarma P.S. et al., 1975). Thus, an expanded cell tropism could be a useful tool in identifying subgroup C viruses. The main determinants of cell tropism and receptor binding of the virus are located in the N-terminal domain of the envelope glycoprotein, as was found for the murine retroviruses (Gray K.D. and Roth M.J., 1993; Battini J-L. et al., 1995; Bae Y. et al., 1997; Han J-Y. et al., 1997). Previous studies by Brojatsch et al and Rigby et al examined a number of FeLV-C isolates in order to establish the amino acids in the envelope glycoprotein that were responsible for the altered *in vivo* pathogenicity and *in vitro* cell tropism of FeLV-C (Brojatsch J. et al., 1992; Rigby M.A. et al., 1992). In both studies, the envelope glycoproteins of FeLV-C/Sarma, FeLV-C/FA27, FeLV-C/FS246 and FeLV-C/FZ215 were examined, however, the envelope glycoprotein was not sequenced fully in either study and there were discrepancies when the sequences were compared between the two studies (see figure 3.2 for the amino acid alignment of the published sequences and a representative clone of each isolate). As described in section 3.3.1, the differences found between the sequences altered the predicted polarity of the protein only slightly and were not expected to alter the overall structure of the protein. Furthermore, the specific amino acids or the stretch of amino acids involved in determining the *in vivo* pathogenicity and *in vitro* cell tropism were not identified. In order to locate the determinants of the *in vitro* cell tropism, we aimed to clone a representative of the C-phenotype of each of the isolates and subsequently sequence the entire envelope glycoprotein gene to establish the determinants for the C-phenotype. Secondly, we aimed to investigate whether the envelope glycoproteins of subgroup A FeLVs were as conserved as has been suggested previously by screening a selection of A-components of the primary isolates studied.

In the course of these studies, thirty-four novel FeLV *env* clones were identified and were classified into either the A component or the C component of the isolate using the criteria described in section 3.3.1. Accordingly, fifteen clones appeared to be the A-component of the primary isolate and twelve clones appeared to be the C-component of that particular primary isolate. As the envelope glycoprotein of FeLV-B/Gardner Arnstein is highly divergent from both FeLV-A and FeLV-C, it was relatively easy to establish that the remaining seven clones were the B-component of those specific primary isolates.

3.4.1. The sequence of the novel FeLV-C envelope glycoproteins

3.4.1.1. Envelope glycoproteins described in this study are divergent from published envelope glycoprotein sequences

When the envelope glycoprotein sequences published by Brojatsch et al and Rigby et al were compared, contradictions were found in the two versions of the *env* sequences (Brojatsch J. et al., 1992; Rigby M.A. et al., 1992). Therefore, a clone of each of the isolates was sequenced and an amino acid alignment was performed to identify a representative envelope glycoprotein sequence. This alignment (see figure 3.2) showed that there were some discrepancies between the studies; however, the overall polarity of each of the sequences was altered only slightly by the inconsistencies between the studies. Therefore, it is conceivable that the discrepancies found previously did not have a significant effect on the overall structure of the envelope glycoprotein. However, as only fragments of the envelope glycoprotein were sequenced, it was possible that additional changes were present in regions outwith the VRA and VRB domains described previously (see section 1.4.4) which may have affected the structure in such a way that the isolates became non-functional, as described by Rigby et al. Natural isolate FZ215 was described as being FeLV-C/Sarma by Brojatsch et al, as there was only one amino acid difference in the region sequenced.

However, both the FZ215 clone described by Rigby et al and the representative clone sequenced in this thesis were significantly different from FeLV-C/Sarma; therefore the clones used in both the study described by Rigby et al and the study described in this thesis were indeed, novel clones in their own right. When the novel envelope glycoproteins characterised in this thesis were compared with the published FeLV envelope glycoproteins, it became clear that the C component of each primary isolate was represented by more than one sequence.

3.4.1.2. Each primary isolate is represented by multiple envelope glycoprotein sequences

When all the FeLV *env* A and C sequences identified in this study were compared, the majority of the *env* sequence was conserved. However, for each independent isolate, there were characteristic differences throughout the sequence that were unique to that isolate. The majority of these differences were clustered in the variable regions of the envelope glycoprotein, though additional sequence polymorphisms were scattered throughout the entire envelope glycoprotein. For example, the primary isolate FY981 yielded seven *env* clones, five clones of which were termed "variant I" whereas the remaining two *env* clones were termed "variant II", due to a substitution from a stretch of $S_{72}S_{73}S_{74}$ in "variant I" to a stretch of $P_{72}S_{73}P_{74}$ in "variant II". The FY981 "variant II" clones produced an *env* protein that was not recognised by the 6-15 antibody, suggesting that "variant II" was defective. When comparing clone FY981-10 (variant II) and clone FY981-14 (variant I), no additional changes were found in the amino acid sequence, thus it is likely that the $P_{72}S_{73}P_{74}$ substitution in clone FY981-10 is responsible for this loss of infection. Similarly, primary isolate FS246 yielded two variant C components, namely FS246-4 and FS246-37. When the amino acid sequences of FS246-4 and FS246-37 were compared, there were only three amino acid differences present in the entire envelope glycoprotein, nevertheless clone FS246-4 was no longer recognised by monoclonal antibody 6-15 and was incapable of infecting the panel of cell lines tested. Of the three amino acid differences, the first difference ($N_{62}S$) was situated at the start of the VRA region, whereas the remaining two differences were outwith the VRA region ($S_{291}F$ and $S_{428}P$).

Thus, minor differences resulted in clone FS246-4 no longer being recognised by the monoclonal antibody 6-15 and incapable of infecting the panel of cell lines tested. When the amino acid sequences of the C components of primary isolate FA27 were compared, there was a single amino acid difference when comparing FA27-17 and FA27-55. Clone FA27-55 contained a serine residue at position 61, whereas clone FA27-17 contained an arginine residue at this position, consistent with the sequence found in FeLV-C/Sarma. However, the altered amino acid sequence of clone FA27-17 resulted in a diminished titre when used to infect a panel of cell lines. Thus, the results described in this thesis lead to the conclusion that indeed there are multiple envelope glycoprotein sequences present in the viral pool of each primary isolate.

3.4.1.3. The abundance of the C-component within each primary isolate pool

When the number of subgroup C *envs* identified for each primary isolate was compared, it was clear that the C phenotype was more abundant in some primary isolates than in other primary isolates. For example, with isolate FY981 seven out of seven *env* clones were the C component, suggesting a relatively abundant C phenotype within this viral pool. Although this isolate was classified previously as an A/B/C mixture, neither the A component nor the B component were isolated. The primary isolate FZ215, which was originally classified as an A/B/C mixture, yielded five unique *env* clones. Of these five clones, one appeared to be the A component, four appeared to be the B component and no C component was identified, suggesting that the B component was abundant in this isolate while the C component was scarce. Although both primary isolates FX343 and L3728F were classified as A/C mixtures, only the A components were isolated, suggesting that the C component was rare in these viral pools. Both isolate FA27 and isolate FS246 yielded two out of four clones and two out of six clones respectively that were the C component, suggesting that the C component was relatively abundant in these pools. Interestingly, although isolate FA621 was previously classified as an A/C mixture through interference assays, the majority of clones (three out of five clones) appeared to be subgroup B, suggesting that either this original classification was incorrect, or that the C component was rare.

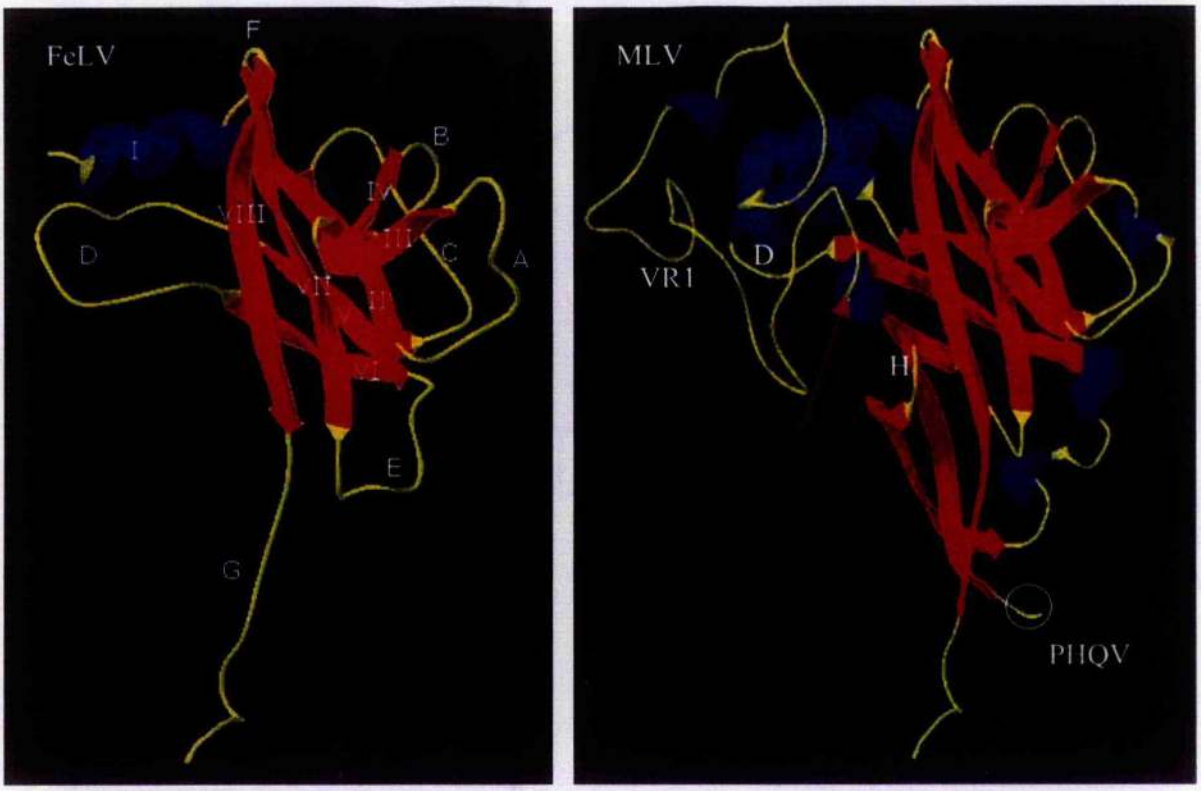
3.4.2. Characteristics of the C phenotype of FeLV

3.4.2.1. The common features of the C-component

To date, no single amino acid or amino acid motif has been identified as a consistent feature of viruses with a subgroup C phenotype. Brojatsch et al proposed that a single lysine to arginine substitution at position 64 in the published FeLV-A sequence was sufficient to alter the *in vitro* host range (Brojatsch J. et al., 1992). However, this $K_{64}R$ substitution was not present consistently in the clones described in this thesis. Indeed, with the exception of the clones characterised from isolate FS246, this substitution was absent. The lack of this single substitution in the isolates described in this study could be due to differences in the method used to perform an amino acid alignment. The amino acid alignment described in this thesis (see figure 3.6) centred on a conserved alanine residue on position 66 in the published FeLV-A/Glasgow-1 sequence, which was present in all *env* clones identified in this study. Consequently, the lysine to arginine substitution described by Brojatsch et al was absent from the amino acid alignment. Therefore, it was highly likely that this amino acid was not the sole determining factor of the altered phenotype of the clones representing the C component.

When Rigby et al studied the receptor specificity and pathogenicity of the same natural FeLV-C isolates described by Brojatsch et al, It was found that the VRA region (referred to as Vr1) was sufficient to alter the receptor specificity from an A phenotype to a C phenotype (Rigby M.A. et al., 1992). The majority of changes found in the envelope glycoproteins described by Rigby et al were in the N-terminus of the envelope glycoprotein; in agreement with other studies (Bae Y. et al., 1997; Riedel N. et al., 1988; Riedel N.O. et al., 1986). Indeed as expected, the majority of changes described in this thesis can be found in the first variable region, the VRA region. The VRA region is predicted to be located on the surface of the envelope glycoprotein, as deduced from the MuLV *env* structure proposed by Fass et al (Fass D. et al., 1997). Therefore changes in the VRA region may alter the surface of the *env* and subsequently receptor recognition by the envelope glycoprotein.

Figure 3.9 shows a possible structure of the FeLV *env* receptor-binding domain as deduced from the structure of the corresponding region of Friend MuLV according to Fass et al (Fass D. et al., 1997).




```
           1                                                    in MLV 50
FS246new   MESPTDPNPS KDKTPSWNLV FLLGILITID IGMANPSPHQ IYNVTWVITN
   FeLVA   MESPTHPKPS KDKTLSWNLA FLVGILFTID IGMANPSPHQ IYNVTWVITN
FZ215new   MESPTHPKPS KDKTLSWDLV FLVGILFTID IGMASPSPHQ IYNVTWVITN
  SarmaC   MESPTHPKPS KDKTFPWNLV FLVGILFQID MGMANPSPHQ VYNVTWVITN
FY981new   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~MANPSPHQ IYNVTWVITN

           51 H                    in MLV               VR1       100
FS246new   VQTTPKANAT SMLGTLTDAY PTLHVDLCDL VGDTWEPIVL NPTSVRPGAV
FeLVA      VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VGDTWEPIVL NPTNV   KHG
FA27C      VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VGNTWEPIVP DL      RGW
FZ215new   VQTNTQRNAT SMLGTLTDAY PTLHVDLCDL VGDTWESMAP DP      RSW
SarmaC     VQTNSRANAT SMLGTLTDAY PTLYVDLCDL VGDTWEPIAP DP      RSW
L39505     VQTNTPANAT SMLGTLTDAY PTLHVDLCDL VGDTWESIVL         RGW
FY981new   VQNNTQANAT SMLGTLTDAY PTLHVDLCDL VGNTWEPIVP PQI      VRR

                          I          II            A     III
           101                                                   150
FS246new   LLSSSPKYGC KTTDRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
FeLVA      ARYSSSKYGC KTTDRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
FA27C      ASYSSSKYGC KTADRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
FZ215new   ARYSSSTHGC KATDRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
SarmaC     ARYSSSTHGC KTTDRKKQQQ TYPFYVCPGH APSMGPKGTY CGGAQDGFCA
L39505     AHYSSSKYGC KTTDRKKQLQ TYPFYVCPGH APSFGPKGTH CGGAQDGYCA
FY981new   ARYSSSTHGC KTTDRKKQQR TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
                         MLV VRA
```

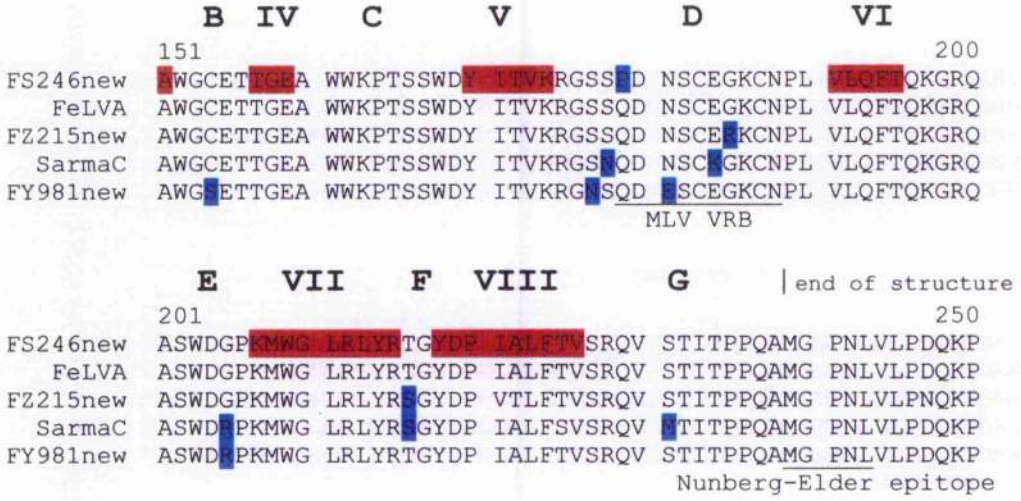


Figure 3.1. Comparison of the predicted structures of the envelope glycoproteins of the feline and murine leukaemia viruses.

When examining both the nucleotide and the amino acid alignment, a pattern emerges in the sequence of the C components of the primary isolates. Whereas the A components are very similar, indeed there are few differences throughout the entire envelope glycoprotein (these will be discussed in more detail in section 3.4.3), the C components are highly divergent, as can be seen in figures 3.3 and 3.6. Deletions appear to cluster in similar areas of the envelope glycoprotein and there appear to be two common motifs found only in the C components, they are the lack of a stretch of six amino acids (^{61}TNVKHG^{66}) and a $V_{63}W$ substitution.

3.4.2.1.1. Deletions in the VRA region of the envelope glycoprotein

When the nucleotide sequences of subgroup C *envs* were aligned, it was apparent that the mutations in the VRA region occurred at the same nucleotide position and consisted of common deletions and mutations (see figure 3.4). The deletions in the nucleotide sequences present in all subgroup C *envs* occurred from position 173 and consisted of:

a) Seventeen mutated nucleotides
b) Nine deleted nucleotides
c) Three nucleotides that were fairly conserved to account for the conserved alanine residue
d) Three deleted nucleotides in all *envs* except FS246 *env* clones to account for the two extra amino acids in FS246

The retroviral RNA genome is transcribed into a proviral DNA copy in a complex process called "reverse transcription" driven by the enzyme "reverse transcriptase" (which is encoded for by the retroviral *pol* gene). Reverse transcription in retroviruses is subject to a relatively high occurrence of errors due to the complexity of the process and the lack of an error correcting function of the enzyme (see also (Varela-Echavarria A. et al., 1992), (Temin H.M., 1993) and (Mansky L.M., 2002)). Due to the presence of an additional enzyme, RNAse H, only a limited stretch of RNA is "unzipped" at any given time in the process. The majority of mistakes during this process occur due to the enzyme dislocating itself from the RNA and rejoining the

structure at a different location. Two possible locations for the occurrence of such a mistake can be seen in figure 3.10, where in position A, the adjacent RNA strands are relatively close or position B where the RNA structure is stressed by adjacent areas of intrastrand hairpin formation. Thus it is highly likely that the clustering of mutations in the nucleotide sequence alignment of the novel FeLV envelope glycoproteins may be the result of mistakes during the reverse transcription process.

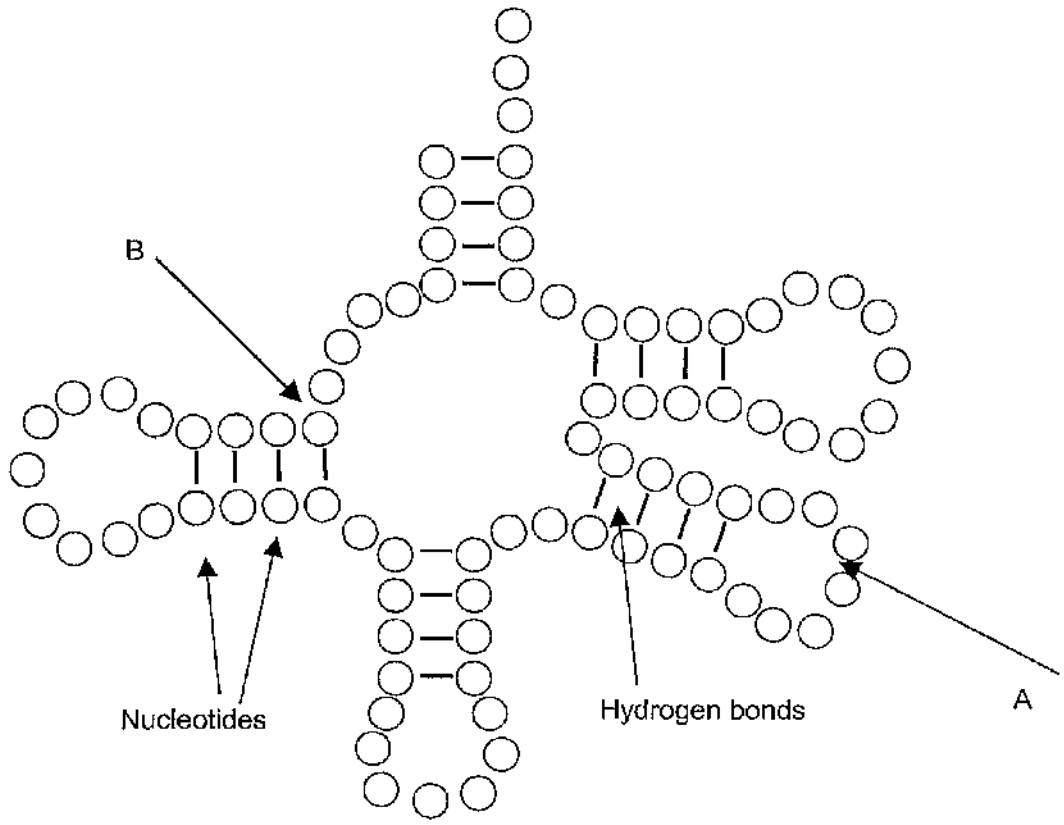


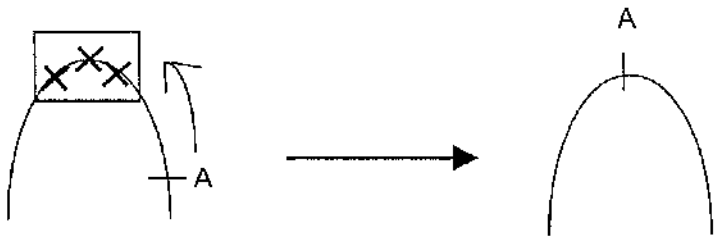
Figure 3.10: *Secondary structure of RNA and possible positions where errors in the transcription could occur.*

3.4.2.1.2. Motifs in the envelope glycoprotein

When the *env* sequences of the C components are compared, there are two main features present in the envelope glycoprotein that are consistent with a C phenotype. Within the VRA region of the C components, a series of amino acids consisting of ^{61}TNVKHG^{66} starting at position 61 in the FeLV-A/Glasgow-1 envelope glycoprotein amino acid sequence is mutated resulting in the disruption of the ^{61}TNVKHG^{66}. A second feature present in all C components is the substitution $V_{63}W$. This series of amino acids may be a critical determinant in the A phenotype and usage of the A receptor as clones that contained "^{61}TNVKHG^{66}" had an A tropism, whereas clones that lacked "^{61}TNVKHG^{66}" had a tropism consistent with the C phenotype. As discussed previously (see figure 3.10 and section 3.4.2.1), this stretch of amino acids is predicted to lie on the exposed surface of the protein and is situated in the putative receptor binding domain, therefore changes in this region of the envelope glycoprotein are likely to alter the region of the protein that determines the ability of the virus to bind its receptor. The valine and the tryptophan residues have hydrophobic side chains and thus may form a hydrophobic pocket on the surface of the envelope glycoprotein; a tryptophan residue at position 142 has been linked previously to the formation of a hydrophobic pocket and subsequent binding of the viral receptor by the envelope glycoprotein of ecotropic MuLV (Zavorotinskaya T. and Albritton L.M., 1999). Thus, even though the tryptophan residue described in this thesis is not the tryptophan residue described by Zavorotinskaya and Albritton, it is conceivable that the tryptophan residue present in the C components has a similar role in the receptor recognition and thus is involved in the altered receptor binding of the C components.

As is clear in figure 3.11, the addition or deletion of even a small number of amino acids can have a major impact on the exposed receptor recognition site on the surface of the envelope glycoprotein. When three amino acids are deleted (as is the case in the C components of the primary isolates), the amino acid sequence after this shifts to the left and thus exposes a novel receptor recognition site (see figure 3.11a). Similarly, the addition of three amino acids (such as in the C components of primary isolate FS246) results in the amino acid sequence shifting to the right and thus the exposure of an additional receptor recognition site (see figure 3.11b).

A: Deletion of three amino acids



B: Addition of three amino acids

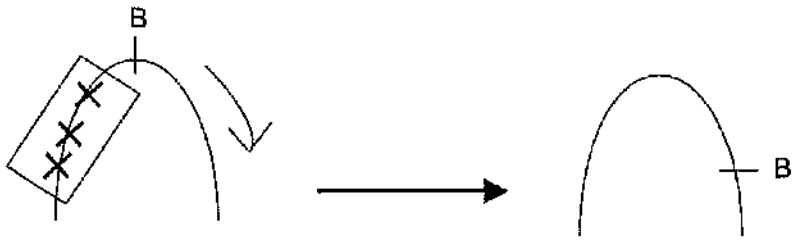


***Figure 3.11**: The effects a deletion or an addition of a stretch of three amino acids can have on amino acids involved in for example receptor binding.*

3.4.2.2. The occurrence of clustered mutations within the envelope glycoprotein

When the nucleotide and amino acid sequences were compared, differences between the sequences were clustered within a number of regions of the *env*. The majority of the sequence changes were present in the first twenty-five amino acids of the VRA region (stretching from D_{53} to G_{79}) while fewer changes were clustered in the VRB region. There were additional amino acid changes throughout the entire envelope glycoprotein and it is possible that these changes occurred to counteract amino acid changes in other regions of the envelope glycoprotein and therefore were there to maintain the structure of the protein, as also discussed briefly by Fass et al (Fass D. et al., 1997).

The clustering of amino acid changes in the envelope glycoprotein sequence rather than the random occurrence of these changes throughout the entire envelope glycoprotein indicates that the accumulation of mutations in the sequence is not random. While the mechanism of acquisition of the mutations (e.g. errors in reverse transcription) is largely random in nature, the process of selection for mutant envelope glycoproteins is likely to be driven by:

a) Pressure to escape the humoral immune response against exposed epitopes
b) Selective pressure for mutations that are tolerated by the envelope glycoprotein structure
c) Selection for mutations that give rise to a variant virus with a replicative advantage, for example a broader cell tropism.

3.4.3. The envelope glycoprotein sequences of the A components

Of the thirty-two envelope glycoproteins identified, fifteen *env* clones were classified as the A component of the primary isolate, using the criteria described previously (lack of both variability and length polymorphisms in the VRA region of the envelope glycoprotein). Comparisons of the FeLV-A envelope glycoprotein sequences that have been published to date (FeLV-A/Glasgow-1 and 61E) revealed few differences between the envelope glycoprotein sequences. However, when the *env* clones identified in this study were compared, there appeared to be a small number of differences between the envelope glycoprotein amino acid sequences. These amino acids were present in other clones of the same primary isolate, for example a $D_{51}N$ substitution and a $N_{59}D$ substitution (relative to the FeLV-A/Glasgow-1 sequence) in the VRA region of the clones derived from primary isolate FA27. As these amino acids were present in all clones sequenced from this particular primary isolate, it was likely that these amino acids were a conserved feature of the envelope glycoprotein. Further, these changes were also present in the published amino acid sequences of FA27 by both Brojatsch et al and Rigby et al (Brojatsch J. et al., 1992), (Rigby M.A. et al., 1992).

These data suggested that although FeLV-A amino acid sequences were very similar, there might be amino acids that were specific for *env* clones derived from a particular isolate, rendering these changes unique to a specific primary isolate. Moreover, minor variations in the VRA region of independent isolates of FeLV-A might affect the cell tropism and immunogenicity of the virus, perhaps influencing the likelihood of producing a subgroup C variant *in vivo*.

3.4.4. Effects of the *env* sequence on antibody staining and biological function

In order to establish the function and the host range of twelve of the unique envelope glycoproteins isolated, FeLV/MuLV pseudotypes were generated and their cell tropism investigated *in vitro* (see table 3.5). Further, recognition of the viral *env* protein by the anti-gp70 antibody 6-15 was examined by immunofluorescence (see figure 3.8).

3.4.4.1. The production of mature envelope glycoprotein *in vitro*

Of the twelve envelope glycoproteins that were fully characterised, three *env* clones (FX343-26, FY981-10 and FZ215-12) failed to react with the anti-gp70 monoclonal antibody 6-15. While the binding site for the monoclonal antibody 6-15 has not yet been characterised, it is known that the antibody recognises the three subgroups of the virus, suggesting that the antibody-binding site is located in a highly conserved region of the envelope glycoprotein. Two possible explanations for the production of a protein that is no longer recognised by this antibody are:

a) the mutations in the amino acid sequence disrupt the conformational structure of the protein and therefore the antibody epitope is no longer available to the antibody.

b) the mutations are located in the antibody-binding site, destroying the epitope and preventing antibody binding.

In each of the three *env* clones that failed to bind the monoclonal antibody, there were a number of changes in the amino acid sequence that were not present in *env* clones identified from the same primary isolate that were recognised by the monoclonal antibody. The mutations for each *env* clone were as follows:

a) FX343-26: $Q_{224}K$
b) FY981-10: $S_{74}P$ and $S_{76}P$
c) FZ215-12: $T_{23}A$ and $T_{165}I$

When the three *env* clones were compared, the mutations appeared to fall in distinct regions of the envelope glycoprotein. Based on the structure of the MLV envelope glycoprotein (see figure 3.12), the predicted locations for these amino acids in the FeLV envelope glycoprotein lie in distinct regions of the protein, it is therefore unlikely that these amino acids form the 6-15 epitope on the envelope glycoprotein. Rather, it is likely that the mutations in the envelope glycoprotein alter the conformational structure of the protein and as a result, the 6-15 epitope is destroyed, leading to a protein no longer recognised by the monoclonal antibody.

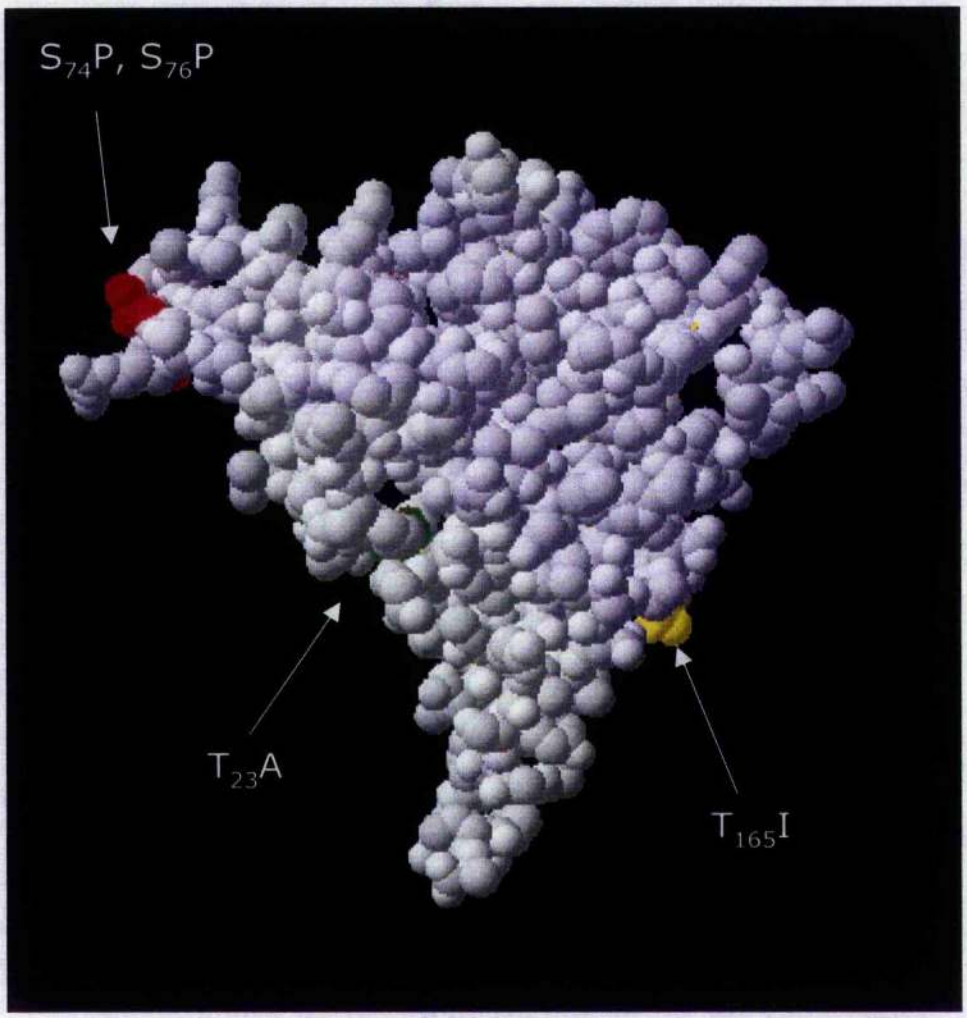


Figure 3.12: *Predicted locations of the $S_{74}P$,$S_{76}P$, $T_{23}A$ and $T_{165}I$ mutations on the FeLV A envelope glycoprotein (structure is modelled on the published coordinates of MLV envelope glycoprotein, pdb number 1AOL).*

In a study by Burns et al (Burns C.C. et al., 1995) into the mutations responsible for an altered phenotype and the delayed onset of immunodeficiency, it was found that two amino acid changes, $E_{536}K$ $E_{537}K$, within a putative cysteine loop of the TM region were sufficient to interfere with envelope glycoprotein processing. In *env* clones that did not produce detectable protein in this study, similar changes were not observed. These data suggested that distinct amino acids are responsible for interfering with the production of viable Env. As the epitope recognised by the 6-15 monoclonal antibody has not yet been characterised, it is not possible to determine the precise residues responsible for the improper folding of these proteins. A further two Env clones (FS246-40 and FX343-28) produced a protein that was recognised by the anti gp70 monoclonal antibody 6-15 but did not produce pseudotypes capable of infecting of a panel of cell lines, suggesting that although correctly folded the mutations disrupted a region that affects the function of the Env, possibly by interfering with the receptor binding site (discussed in more detail in section 3.4.2.2.). Accordingly, these data may suggest that the antibody binding site and the receptor-binding site are distinct regions in the envelope glycoprotein.

3.4.4.2. Altered cell tropism of novel FeLV *env* clones

To determine whether the novel envelope glycoproteins were functional, the cell tropism of FeLV/MuLV pseudotypes prepared from each clone was investigated. Previous studies have examined the *in vitro* cell tropism of FeLV-A, FeLV-B and FeLV-C and have shown that FeLV-A has the most restricted *in vitro* cell tropism and is capable of infecting feline cells alone (Jarrett O. et al., 1973), (Sarma P.S. et al., 1975), (Hardy W.G.Jr., 1993). In contrast, FeLV-B and FeLV-C have a broader *in vitro* cell tropism and both viruses are capable of infecting a range of feline and non-feline cells. FeLV-C is the only subgroup capable of infecting guinea pig cells, whereas porcine cells can be infected by FeLV-B alone (Jarrett O. et al., 1973), (Hardy W.G.Jr., 1993).

The three *env* clones that failed to produce protein recognised by the 6-15 antibody (FX343-26, FY981-10 and FZ215-12) did not produce pseudotypes capable of infecting the cell lines examined, suggesting that the proteins produced were indeed non-functional. Furthermore, three viable *env* clones (FS246-4, FS246-40 and FX343-28) did not produce pseudotypes capable of infecting the cell lines tested, suggesting that although the protein was folded properly, it was non-functional and indicating that mutations outwith the antibody-binding site rendered the protein non-functional. The receptor-binding domain has been identified previously for amphotropic MuLV by amongst others Bae et al and consists of the first two hundred and thirty amino acids of the envelope glycoprotein (Gray K.D. and Roth M.J., 1993; Battini J-L. et al., 1995; Bae Y. et al., 1997; Han J-Y. et al., 1997). Thus, amino acid substitutions within this region may affect the function of the protein. When clone FS246-40 was compared with FeLV-A/Glasgow-1, five amino acid differences were identified; these were $Q_{149}P$, $D_{150}G$, $A_{248}G$, $M_{333}L$ and $T_{371}P$. As the $Q_{149}P$ and $A_{248}G$ mutations were present in the viable FS246 clones, it was unlikely that these amino acids were responsible for the loss of function of the protein. Furthermore, although a $M_{333}L$ substitution was present in FS246-40, the methionine residue at position 333 is present only in FeLV-A/Glasgow-1 and not in any other FeLV envelope glycoprotein. When FeLV-A/Glasgow-1 was sequenced, this methionine residue was present in the primary FeLV-A/Glasgow-1 clone, thus the residue was considered to be a unique feature of FeLV-A/Glasgow-1. In all other clones, this residue is leucine, thus this residue can be excluded from the possible mutations affecting the protein structure. Thus, two substitutions remain that may be involved in the loss of functional protein; these are $D_{150}G$ and $T_{371}P$. $D_{150}G$ is located in the VRB region of the envelope glycoprotein and this substitution results in a change from a negatively charged residue (D) to an uncharged residue (G). Considering the location of this residue, this substitution could affect the interaction between the virus and its receptor. However, this substitution was present also in *env* clone FZ215-9, a clone that was capable of producing a protein that was recognised by the anti-gp70 monoclonal antibody 6-15. Thus it is unlikely that this residue is responsible for the loss of biological function although it is possible that FZ215-9 has mutations in adjacent regions of the structure that compensate for $D_{150}G$ while FS246-40 does not.

In contrast, $T_{371}P$ is situated in a highly conserved region of the envelope glycoprotein and results in a change from a neutral, hydrophilic residue to a residue that is found predominantly in bends in the conformational structure of a protein. Furthermore, this mutation is situated ten amino acids removed from a stretch of amino acids that has been described previously to be important in the onset of immunodeficiency (Quackenbusch S.L. et al., 1990). Thus it is likely that amino acid mutations in this region of the envelope glycoprotein can indeed alter the protein structure sufficiently to account for the loss of a functional protein. When the sequence of clone FX343-28 was compared with FeLV-A/Glasgow-1, there were twelve amino acid differences between the clones. Of these twelve amino acids, five were also present in clone FX343-26 (a clone that was not recognised by the 6-15 monoclonal antibody); these were $S_{75}D$, $T_{318}I$, $M_{333}L$, $K_{348}E$ and $V_{594}I$. A further seven amino acid differences were present in clone FX343-28, but not in clone FX343-26, these were $Q_{22}P$, $T_{186}A$, $N_{240}T$, $T_{380}A$, $S_{389}T$, $A_{611}V$ and $K_{595}I$. Four of the substitutions were relatively minor substitutions ($T_{186}A$, $T_{380}A$, $S_{389}T$ and $A_{611}V$) and were not likely to have an effect on the overall structure of the protein. However, the remaining three substitutions ($Q_{22}P$, $N_{240}T$ and $K_{595}I$) could have an effect on the overall structure of the protein and subsequent loss of a functional protein. Proline is a residue that is predominantly found in the bend of folded protein chains, thus the presence of this residue is likely to have a major effect on the protein structure. Similarly, the occurrence of changes in a highly conserved area of the envelope glycoprotein ($K_{595}I$ and $N_{240}T$) may affect the conformational protein structure and result in the loss of the production of a functional protein. The mutation $K_{595}I$ is situated in the p15E protein, and might affect the interaction between p15E and gp70 and thus result in a protein that is not capable of producing pseudotypes that can infect the panel of cell lines tested.

Pseudotypes bearing the *Envs* most closely resembling FeLV-A (FA27-53 and FZ215-9) were able to infect the feline cell line HO6T1 and mink cells to comparable titres as those derived from the control FeLV-A ($7.4x10^3$ CFU/ml for the control FeLV-A compared to $9.6x10^3$ CFU/ml for both FA27-53 and FZ215-9 when infecting HO6T1 cells). FA27-53, FZ215-9 nor the control FeLV-A derived pseudotypes were capable of infecting any other cell line tested, including the human cell line HeLa and the canine cell line MDCK.

The inability of pseudotypes derived from FeLV-A/Glasgow-1 and the novel FeLV-As to infect human cells, contradicts studies by Moser et al and Bupp et al, in which the host range of FeLV-A was studied (Moser M. et al., 1998), (Bupp K. and Roth M.J., 2002). Both Moser et al and Bupp et al found that FeLV-A does indeed infect human cells. However, both of these studies used the human cell line 293T for infection assays while HeLa cells were used in this thesis. Moreover, the FeLV-A isolates used were distinct; 61E in contrast with FeLV-A/Glasgow-1 used in the infection assays described in this thesis. Indeed, 293T cells supported infection with the FeLV-A/Glasgow-1 isolate used in this study (data not shown). It is possible that distinct isolates of FeLV-A are capable of infecting different cell types and secondly although it is generally stated that FeLV-A is incapable of infecting human cells, the discrepancies when comparing the findings in this thesis and the studies described by both Moser et al and Bupp et al may be due to differences in the human cell line used for the infection assays.

Interestingly, pseudotypes derived from clone FA27-53 were capable of infecting the porcine cell line ST Iowa, in contrast to previous reports in which porcine cells were shown to support FeLV-B infection alone. When the sequence of FA27-53 was examined, there were relatively few differences between FA27-53 and FeLV-A/Glasgow-1. Relative to FeLV-A/Glasgow-1, they are $D_{51}N$, $N_{59}D$, $V_{247}I$, $T_{318}I$, $M_{333}L$, $K_{348}E$ and $K_{614}R$. With the exception of $D_{51}N$ and $N_{59}D$, the amino acid differences lay outwith the region in the envelope glycoprotein associated with cell tropism (determined amongst others by Bae et al (Bae Y. et al., 1997)). When the FA27-53 amino acid sequence was compared with the amino acid sequence of other constructs that were capable of infecting ST Iowa cells (FY981-14, FY981 and FeLV-B/Gardner Arnstein), the $D_{51}N$ substitution alone was present in all constructs which pseudotypes were capable of infecting ST Iowa cells. Thus it is likely that this single amino acid substitution may explain the tropism of FA27-53 *env*-bearing pseudotypes for the porcine cell line ST Iowa. The amino acid substitution discussed ($D_{51}N$) was present also in the other FA27 constructs (FA27-17, FA27-55 and FA27p), however pseudotypes derived from these constructs were not capable of infecting porcine cells.

This could be due to the fact that there were numerous other changes in the VRA region when compared to FeLV-A/Glasgow-1 and therefore altering the conformational structure of the proteins and subsequently the amino acids thought to be involved in infection of porcine cells were no longer available. Thus, the data presented in this thesis suggest that although it was thought that FeLV-A isolates had the same *in vitro* cell tropism, FeLV-As derived from diverse primary isolates did have a unique *in vitro* cell tropism.

The *env*-construct that most closely resembled FeLV-B/Gardner Arnstein, FZ215-12, did not produce protein recognised by the monoclonal antibody 6-15, nor was it capable of producing functional protein as was shown by the lack of infection by FZ215-12 *env*/MuLV *gag-pol* pseudotypes. This suggested that the small number of changes compared to the published FeLV-B/Gardner Arnstein present in this envelope glycoprotein rendered this construct defective. There were two additional changes in the amino acid sequence of FZ215-12 that were found neither in the published FeLV-B/Gardner Arnstein sequence nor in the parent FZ215 sequence. These were a $T_{23}A$ substitution and a $T_{165}I$ substitution; both substitutions resulted in a change from a residue with an uncharged polar side chain to a residue with a non-polar side chain. The side chain of threonine is hydrophilic; whereas both alanine and isoleucine have hydrophobic side chains, thus leading to the conclusion that these minor changes could indeed lead to the production of a non-functional protein.

Pseudotypes derived from the *env* clones thought to be C components (FA27-55, FS246-37 and FY981-14) were capable of infecting both the feline cell line HO6T1 and the guinea pig cell line 104C1 to higher titres than the control FeLV-C/Sarma ($4.6x10^2$ CFU/ml for the control FeLV-C compared to $9.3x10^3$ CFU/ml for FA27-55, $4.3x10^3$ CFU/ml for FS246-37 and $7.1x10^3$ CFU/ml for FY981-14 when infecting HO6T1 cells). The ten fold difference in titres was most likely due to the poorer titres for the FeLV-C/Sarma construct, as the titres found for FA27-55, FS246-37 and FY981-14 were comparable with the titres found for the control FeLV-A/Glasgow-1 construct. Furthermore, *env* clone FY981-14 had a broader tropism than the other C components and the control FeLV-C/Sarma. However, the tropism of FY981-14 was comparable to the tropism of clone FY981p.

Indeed, it was capable of Infecting all cell lines tested, including the porcine cell line ST Iowa thought to be restrictive to FeLV-B infection only. As discussed for FA27-53, the $D_{51}N$ substitution that may be involved in altering the tropism to include infection of the porcine cell line ST Iowa is present also in the amino acid sequence of FY981-14.

CHAPTER FOUR

ENRICHMENT OF FeLV-C ISOLATES BY BIOLOGICAL SELECTION

4.1 Introduction

The FeLV isolates described earlier (section 3.1) were classified into subgroups A, B or C by interference assay. Our initial aim was to amplify *env* gene products representative of the original mix of subgroups within the primary isolate. With some isolates, no C component was detected following the PCR reaction, therefore, in order to enrich for the C component, the primary isolate was plated onto the 104C1 cell line, a guinea pig cell line which is known to be permissive for infection with subgroup C viruses, but restrictive to infection with the A and B subgroups. Thus, it was hoped that through biological selection, the C component would replicate preferentially and thus it would be represented in the pool of viral *envs* isolated.

Biological selection is the process whereby populations of organisms adjust to their environment through the accumulation of beneficial mutations. Through natural selection, these mutations will spread and ultimately the mutations that are the most advantageous will predominate in the viral population (Gerrish P., 2001). As the enzyme reverse transcriptase (encoded by the *pol* gene in the retroviral genome) does not have a proof-reading activity, these random mutations occur frequently and result in a population with a novel phenotype (Miralles R. et al., 2000), (Overbaugh J. and Bangham C.R., 2001).

4.2 Methods

4.2.1 Isolation of FeLV – C subgroups through biological selection

Isolates that were classified previously as subgroups A/C or A/B/C were expanded on the FEA cell line (non-restrictive). The supernatants were then used to infect 104C1 cells. The day prior to infection, 104C1 cells were seeded in 12.5 cm^2 flasks at a concentration of $1x10^5$ cells per flask. The following day, supernatant was collected from FEA stocks, clarified by centrifugation and 2ml of SN was added to each flask of 104C1 cells and incubated overnight at 37°C. The remainder of the supernatant was stored at -70°C until required. The following day, 104C1 cells were re-infected with the frozen FEA supernatant and incubated for a further 4 hours, after which they were washed once with PBS and fresh 10% DMEM was added. Supernatant samples of the 104C1-infected cells were collected at regular intervals and stored at -70°C. The supernatants were then assayed for FeLV p27 by ELISA (FeLV p27 antigen ELISA kit, IDEXX Laboratories Limited, Buckinghamshire, U.K). If the cultures were positive by ELISA, then the cells were expanded to prepare DNA (as described previously in section 3.2.2) and to enable storage in the vapour phase of liquid nitrogen (as described previously in section 2.2.1.2.1).

4.2.2 Amplification of the envelope glycoprotein by PCR

Total cellular DNA from the infected 104C1 cells was subjected to PCR using primers to introduce *SalI* and *NotI* restriction sites to enable direct cloning of the amplified product into the eukaryotic expression vector VR1012. The ligation reactions were subsequently transformed into *E.coli* INVαF' competent cells as described previously in section 3.2.4.

4.2.3 Sequencing of new envelope genes obtained through biological selection

The cloned envelope gene products were sequenced using either a LICOR model 4000 automated sequencer (section 2.2.3.10) or an ABI PRISM 3100 genetic analyser (section 2.2.3.11) as described previously using a panel of primers based on the FeLV-A *env* sequence (Stewart M.A. et al., 1986). The sequences and locations of the primers are described in table 2.1 and figure 2.1. The *env* sequences were analysed using Wisconsin Package Version 10.0, Genetics Computer Group (Madison, Wisconsin) as described in section 3.2.5. Each nucleic acid sequence was translated and the predicted amino acid sequence analysed for integrity of the *env* open reading frame.

4.2.4 Production of FeLV pseudotypes in 293T cells by transient transfection

As described in section 3.2.7, FeLV *env*/MuLV *gag-pol* pseudotypes were generated by transient transfection of 293T cells with FeLV *env*/ MuLV *gag-pol* and MFG-LacZ plasmids. Viral pseudotypes were titred on FEA cells and then used to infect a range of cell lines in order to establish the tropism of the novel envelope glycoprotein. In parallel with the viral pseudotype studies, 293T cells transfected with each VR1012-*env* construct were fixed using ice-cold methanol and stained with the 6-15 monoclonal antibody (anti-gp70). Expression of mature envelope glycoprotein was then evaluated by immunofluorescence as described previously in section 3.2.7.

4.3 Results

4.3.1 Isolation, PCR amplification and characterisation of novel envelope glycoproteins

Following infection of 104C1 cells with the primary isolates FS246 and L3128F, the cultures remained negative for FeLV p27; therefore no DNA was prepared from these cultures. The remaining cultures tested positive for p27 and thus DNAs were prepared for PCR amplification of *env*. Following cloning of the amplified products and small scale plasmid screening, twenty three plasmid DNAs containing inserts were identified and were sequenced using the FeLV *env* primer F69, the primer binding site of which is described previously in figure 2.1 and table 2.1. This primer binds prior to the start codon of the mature gp70 protein and therefore yields the sequence of the VRA region, enabling the identification of novel *env* clones. Table 4.1 and figure 4.1 summarise the results. The sequence characteristics used to classify an *env* clone to be either subgroup A or C were described previously in section 3.3.1.

FA27

Ten novel *env* clones were identified from 104C1-derived FA27, nine of which (FA27-C1, FA27-C2, FA27-C3, FA27-C5, FA27-C7, FA27-C9, FA27-C10, FA27-C11 and FA27-C12) appeared to be subgroup C, whereas one *env* clone appeared to be the A-component (FA27-C6). Of the ten *env* clones, only four were characterised further, they were FA27-C5, FA27-C6, FA27-C7 and FA27-C12; these were unique amongst the clones identified.

FA621

Biological selection of primary isolate FA621 on 104C1 cells yielded six unique *env* clones, one clone (FA621-C4) was the A-component, whereas the remaining five clones (FA621-C1, FA621-C2, FA621-C3, FA621-C6 and FA621-C7) were very similar and appeared to be C components. As FA621-C1 and FA621-C3 were not identical, these two clones were characterised further.

FX343

When isolate FX343 was plated onto the selective 104C1 cell line, one unique *env* clone was isolated, FX343-C8, this was thought to be subgroup C.

FY981

Biological selection of isolate FY981 yielded one unique *env* clone, FY981-C11, which appeared to be a C component.

FZ215

Selection of isolate FZ215 yielded three unique *env* clones; one of the clones was the A-component (FZ215-C6), whereas two *env* clones resembled FeLV-B/Gardner Arnstein (FZ215-C7 and FZ215-C11). Attempts to obtain the complete sequence of FZ215-C11 failed and this isolate was not studied further.

L3950F

When primary isolate L3950F was biologically selected on 104C1 cells, one *env* clone was identified. L3950F-C18 was the A-component of the isolate and was not characterised further.

Of the *env* gene clones that were characterised further, three nucleic acid sequences were similar to FeLV-A/Glasgow-1 (FA27-C6; 98.5%, FA621-C4; 97.9% and FZ215-C6; 98.0% similarity to FeLV-A/Glasgow-1 respectively). Seven nucleic acid sequences satisfied the criteria for subgroup C (FA27-C5; 89.5%, FA27-C7; 90.2%, FA27-C12; 89.5%, FA621-C1; 90.8%, FA621-C3; 89.2%, FA621-C4; 90.9%, FY981-C11; 89.6% and FX343-C8; 91.0% similarity to FeLV-C/Sarma respectively). One nucleic acid sequence was similar to FeLV-B/Gardner Arnstein (FZ215-C7; 98.0% similarity).

A multiple nucleic acid sequence alignment was performed on the sequences, with the exception of the isolate that was similar to FeLV-B Gardner-Arnstein using the "Pileup" component on the software package GCG according to the progressive alignment method of Feng and Doolittle (Feng D.F. and Doolittle R.F., 1987). This alignment is shown in figure 4.2. Position 1 in this sequence correlates with position 6079 in the published FeLV-A genome.

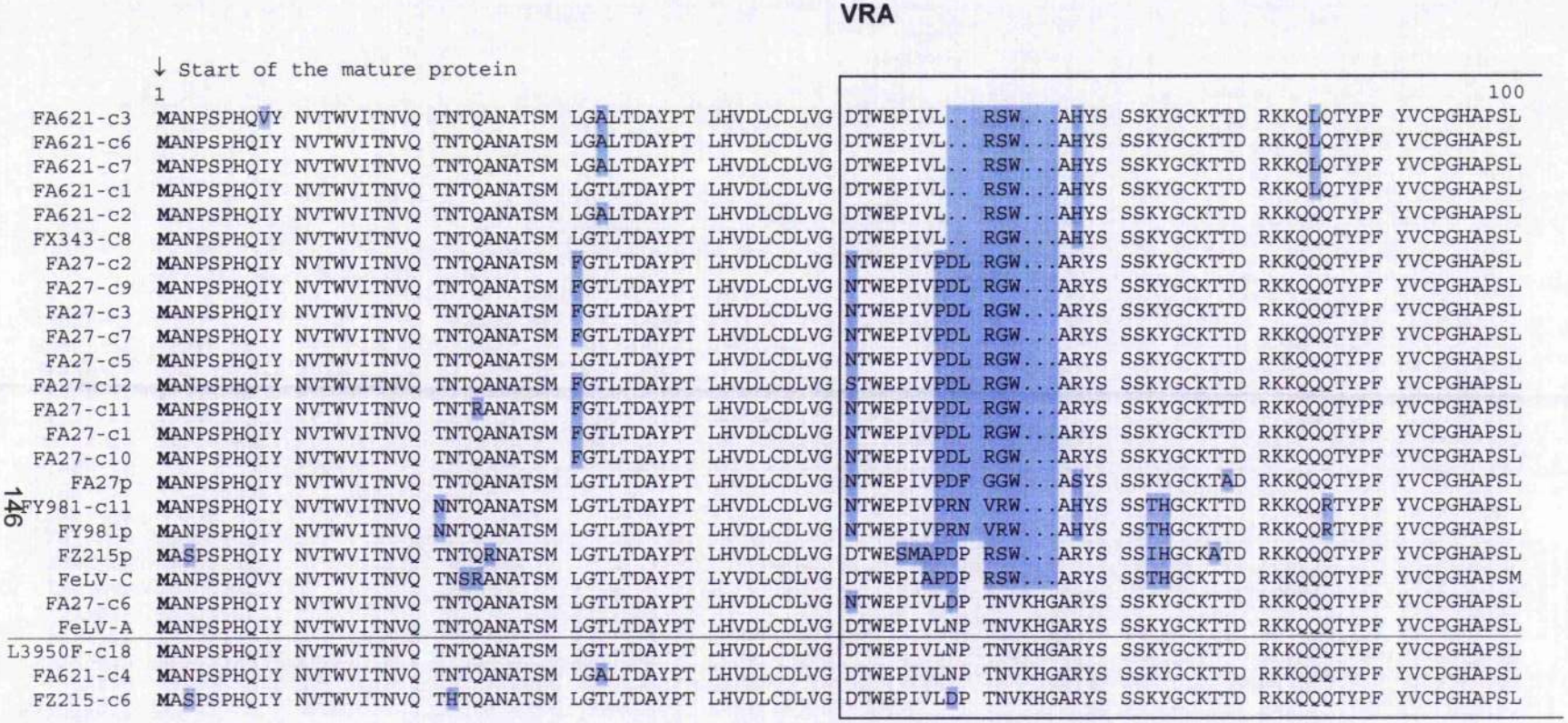
The multiple nucleic acid sequence of the isolates that were similar to FeLV-B Gardner-Arnstein, including isolate FZ215-C7, is shown in figure 3.5.

The sequences were then translated into amino acid sequences to establish the integrity of the *env* open reading frame and changes from the published FeLV A sequence. All of the novel *env* gene clones had intact open reading frames and figure 4.3 shows the amino acid alignment using the "Pileup" component on GCG. Position 1 in the amino acid sequence correlates with position 5981 in the published FeLV-A genome.

The percentage similarity found for the novel *env* clones compared to FeLV-A/Glasgow-1 and FeLV-C/Sarma were equivalent to those found for the nucleic acids, albeit slightly higher, and are summarised in table 4.2. This higher degree of amino acid similarity suggests that some of the mutations detected in the nucleic acid sequence are "synonymous" (i.e. not affecting the coding sequence). When the sequence comparisons were restricted to the VRA region, the degree of amino acid similarity was considerably lower, consistent with this region being the major site for variation; the results of these comparisons are illustrated in table 4.3.

Env clone #	% Similarity to VRA region of FeLV-A	% Similarity to VRA region of FeLV-C	Subgroup predicted by sequencing	Further characterisation
FA27-C1	93.7%	88.9%	C	No
FA27-C2	95.0%	90.4%	C	No
FA27-C3	97.2%	92.6%	C	No
FA27-C5	97.7%	95.1%	C	Yes
FA27-C6	99.3%	94.6%	A	Yes
FA27-C7	97.5%	94.9%	C	Yes
FA27-C9	96.8%	91.6%	C	No
FA27-C10	91.8%	86.8%	C	No
FA27-C11	92.2%	88.1%	C	No
FA27-C12	97.0%	97.0%	C	Yes
FA621-C1	98.0%	94.2%	C	Yes
FA621-C2	95.5%	91.8%	C	No
FA621-C3	98.0%	94.4%	C	Yes
FA621-C4	98.5%	94.0%	A	Yes
FA621-C6	96.9%	93.0%	C	No
FA621-C7	96.3%	92.2%	C	No
FY981-C11	97.7%	93.9%	C	Yes
FX343-C8	98.3%	94.1%	C	Yes
FZ215-C6	98.7%	93.9%	A	Yes
FZ215-C7	89.7%	88.4%	B (98.4%)	Yes
FZ215-C11	74.4%	59.1%	B (87.2%)	No
L3950F-18	99.6%	91.0%	A	No

Table 4.1: *Amino acid sequence similarity of the novel env genes relative to the published sequences of the VRA regions of FeLV-A/Glasgow-1 and FeLV-C/Sarma. The similarities of the env clones, which resemble FeLV-B/Gardner Arnstein, relative to the published sequence of FeLV-B/Gardner Arnstein, have been included in brackets in the table.*


VRA
↓ Start of the mature protein
1
100
FA621-c3 MANPSPHQVY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVL.. RSW...AHYS SSKYGCKTTD RKKQLQTYPF YVCPGHAPSL
FA621-c6 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVL.. RSW...AHYS SSKYGCKTTD RKKQLQTYPF YVCPGHAPSL
FA621-c7 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVL.. RSW...AHYS SSKYGCKTTD RKKQLQTYPF YVCPGHAPSL
FA621-c1 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL.. RSW...AHYS SSKYGCKTTD RKKQLQTYPF YVCPGHAPSL
FA621-c2 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVL.. RSW...AHYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FX343-C8 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL.. RGW...AHYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c2 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c9 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c3 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c7 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c5 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c12 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG STWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c11 MANPSPHQIY NVTWVITNVQ TNTRANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c1 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27-c10 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVPDL RGW...ARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA27p MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPDF GGW...ASYS SSKYGCKTAD RKKQQQTYPF YVCPGHAPSL
FY981-c11 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...AHYS SSTHGCKTTD RKKQQRTYPF YVCPGHAPSL
FY981p MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVPRN VRW...AHYS SSTHGCKTTD RKKQQRTYPF YVCPGHAPSL
FZ215p MASPSPHQIY NVTWVITNVQ TNTQRNATSM LGTLTDAYPT LHVDLCDLVG DTWESMAPDP RSW...ARYS SSIHGCKATD RKKQQQTYPF YVCPGHAPSL
FeLV-C MANPSPHQVY NVTWVITNVQ TNSRANATSM LGTLTDAYPT LYVDLCDLVG DTWEPIAPDP RSW...ARYS SSTHGCKTTD RKKQQQTYPF YVCPGHAPSM
FA27-c6 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVLDP TNVKHGARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FeLV-A MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
L3950F-c18 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FA621-c4 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVLNP TNVKHGARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL
FZ215-c6 MASPSPHQIY NVTWVITNVQ THTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVLDP TNVKHGARYS SSKYGCKTTD RKKQQQTYPF YVCPGHAPSL

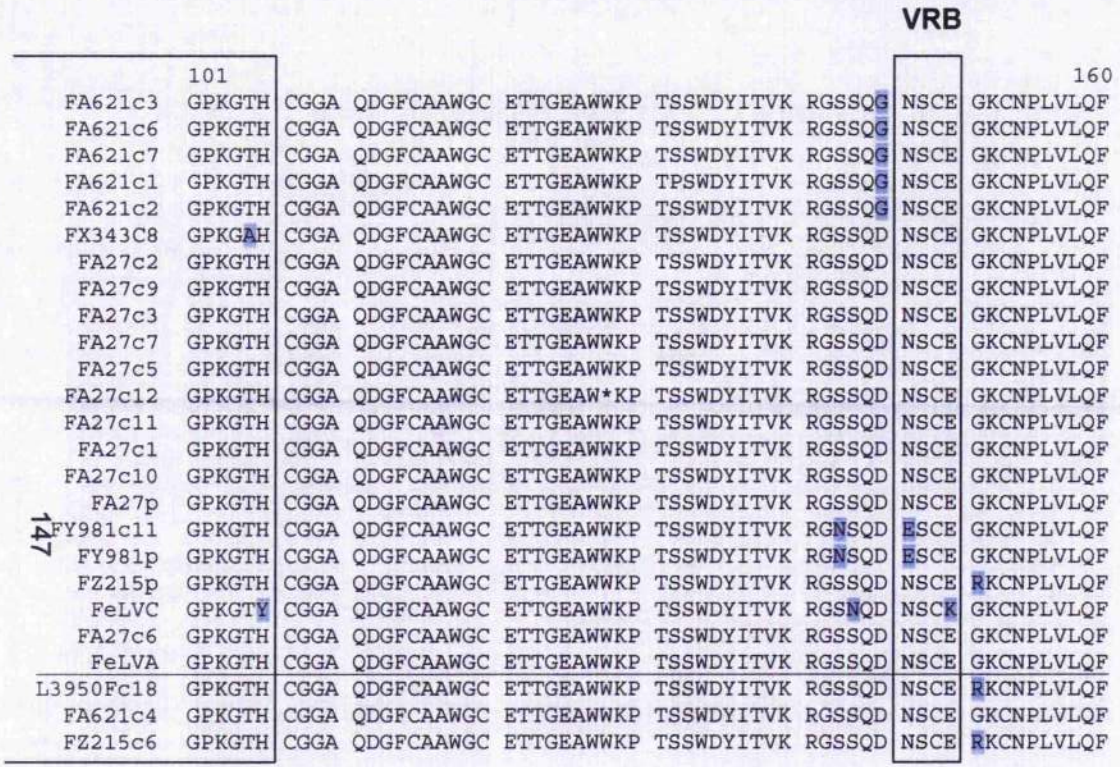
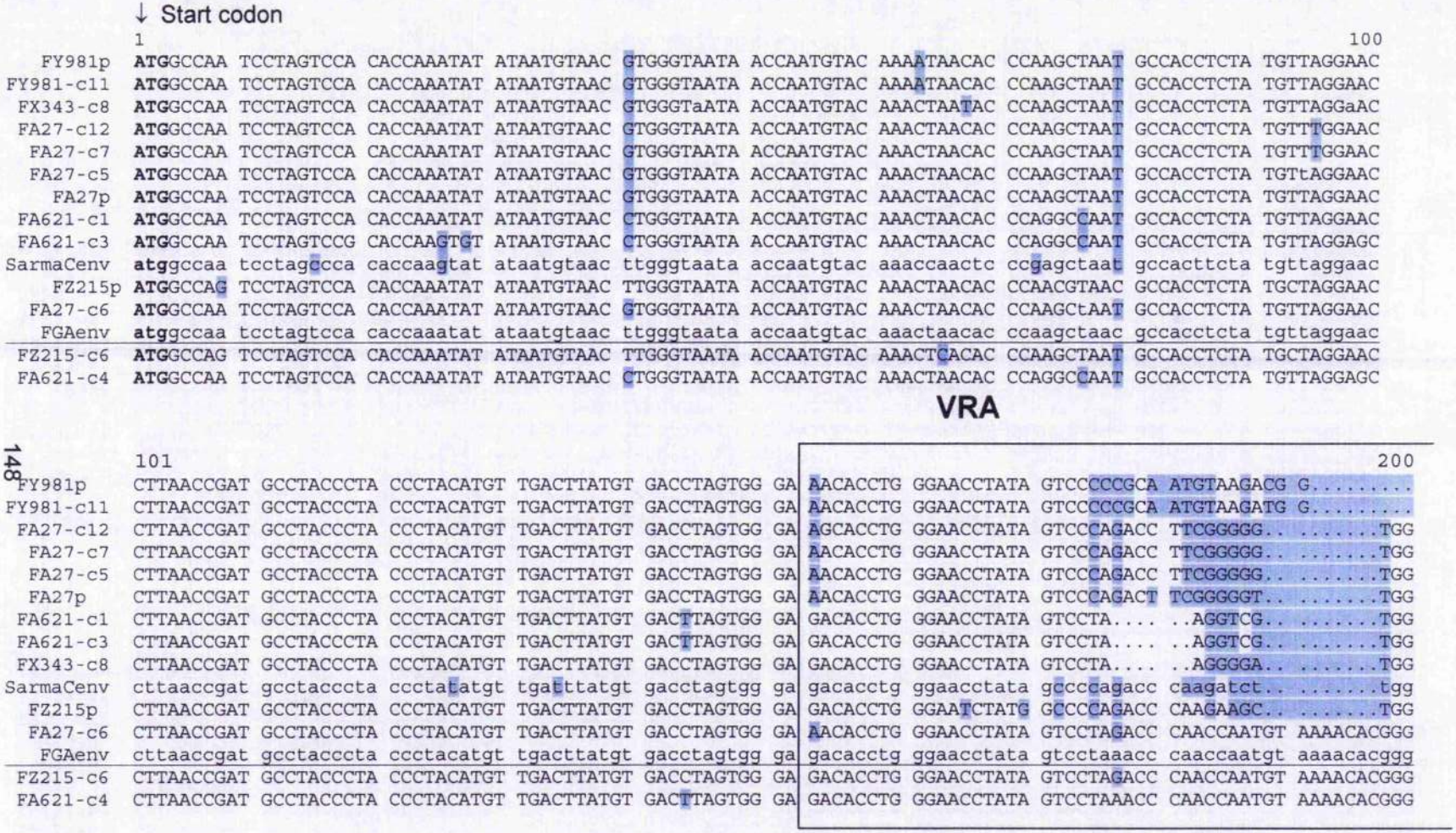
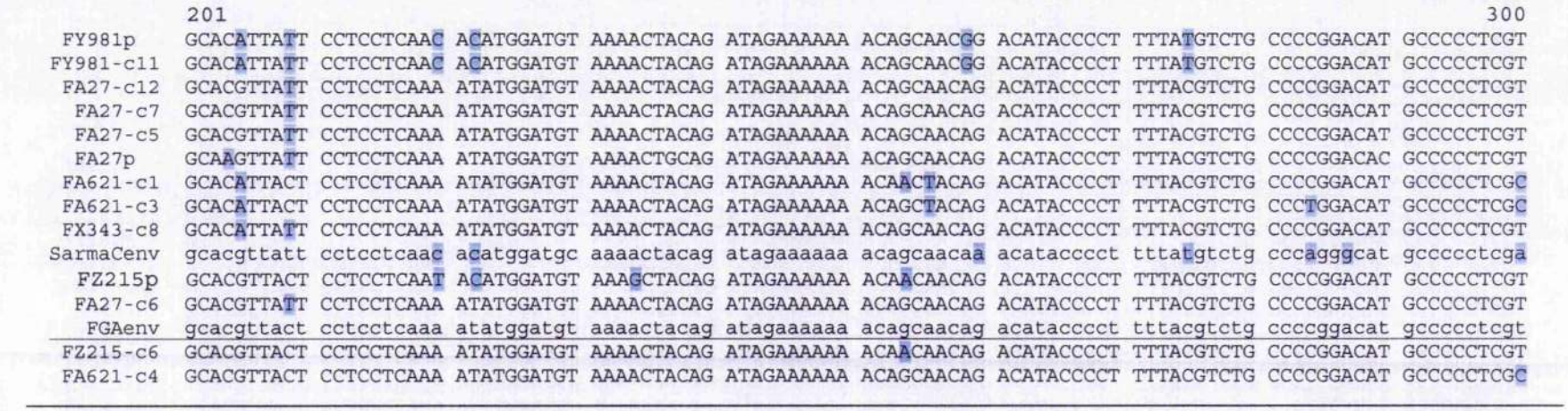
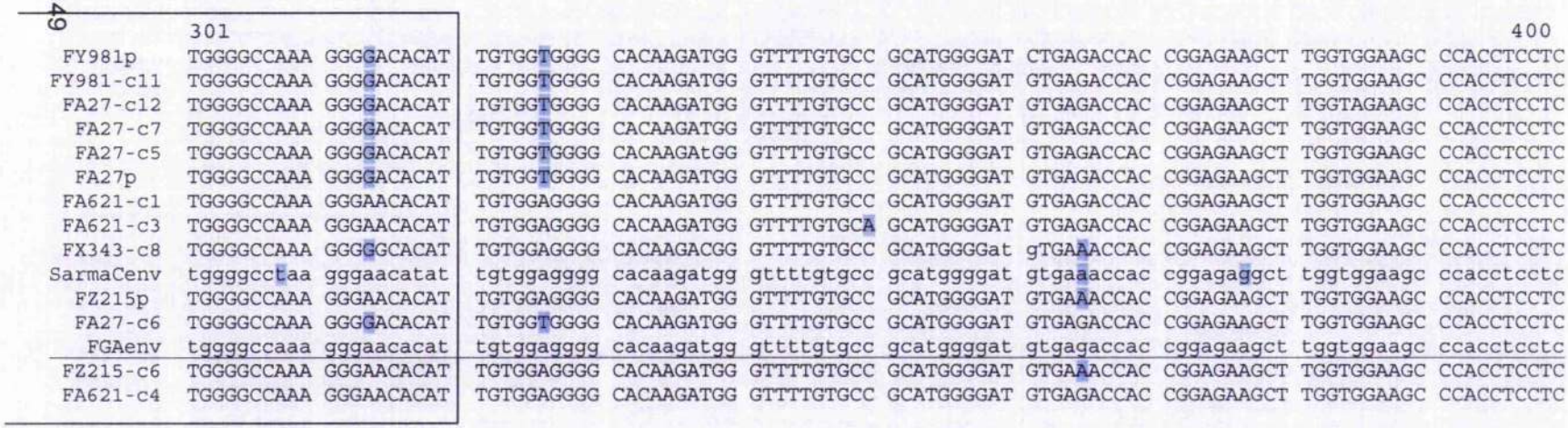

VRB
101
160
FA621c3 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQG NSCE GKCNPLVLQF
FA621c6 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQG NSCE GKCNPLVLQF
FA621c7 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQG NSCE GKCNPLVLQF
FA621c1 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TPSWDYITVK RGSSQG NSCE GKCNPLVLQF
FA621c2 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQG NSCE GKCNPLVLQF
FX343C8 GPKGAH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c2 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c9 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c3 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c7 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c5 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c12 GPKGTH CGGA QDGFCAAWGC ETTGEAW*KP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c11 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c1 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27c10 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FA27p GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FY981c11 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGNSQD ESCE GKCNPLVLQF
FY981p GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGNSQD ESCE GKCNPLVLQF
FZ215p GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE RKCNPLVLQF
FeLVC GPKGTY CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSNQD NSCK GKCNPLVLQF
FA27c6 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FeLVA GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
L3950Fc18 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE RKCNPLVLQF
FA621c4 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE GKCNPLVLQF
FZ215c6 GPKGTH CGGA QDGFCAAWGC ETTGEAWWKP TSSWDYITVK RGSSQD NSCE RKCNPLVLQF
147

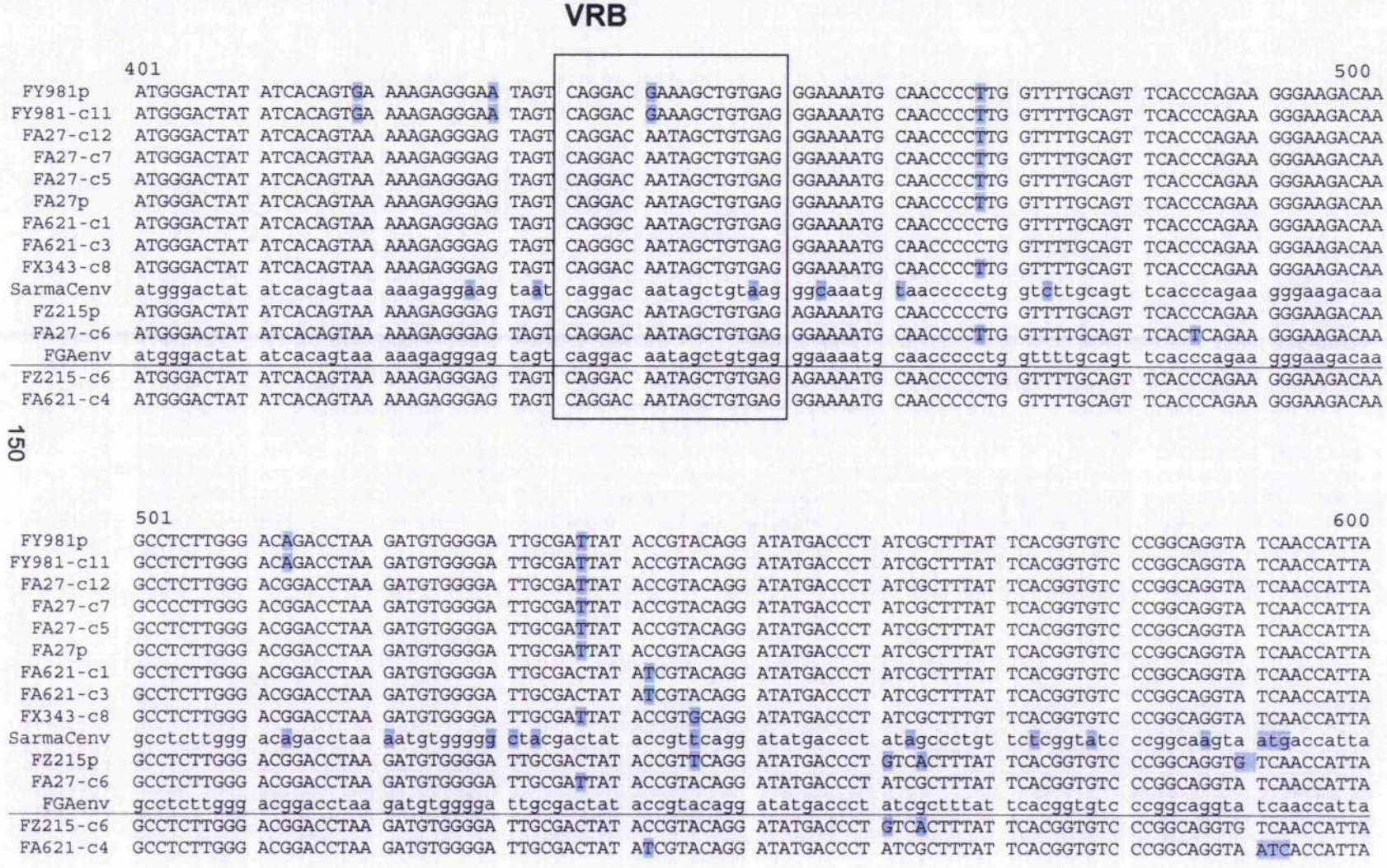
Figure 4.2: *Nucleotide alignment from the SU proteins of the isolated unique envelope regions of the unique envelope glycoproteins identified. The arrow indicates the start of the SU protein and a dot indicates a gap in alignment. A position of 1 in this figure corresponds with position 18237in the published FeLV A nucleic acid genome.*

(Opposite page)


↓ Start codon
1
100
FY981p ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAAATAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FY981-c11 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAAATAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FX343-c8 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTaATA ACCAATGTAC AAACTAATAC CCAAGCTAAT GCCACCTCTA TGTTAGGaAC
FA27-c12 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTTGGAAC
FA27-c7 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTTGGAAC
FA27-c5 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTtAGGAAC
FA27p ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCTAAT GCCACCTCTA TGTTAGGAAC
FA621-c1 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC CTGGGTAATA ACCAATGTAC AAACTAACAC CCAGGCCAAT GCCACCTCTA TGTTAGGAAC
FA621-c3 ATGGCCAA TCCTAGTCCG CACCAAGTGT ATAATGTAAC CTGGGTAATA ACCAATGTAC AAACTAACAC CCAGGCCAAT GCCACCTCTA TGTTAGGAGC
SarmaCenv atggccaa tcctagccca caccaagtat ataatgtaac ttgggtaata accaatgtac aaaccaactc ccgagctaat gccacttcta tgttaggaac
FZ215p ATGGCCAG TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTAACAC CCAACGTAAC GCCACCTCTA TGCTAGGAAC
FA27-c6 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC GTGGGTAATA ACCAATGTAC AAACTAACAC CCAAGCtAAT GCCACCTCTA TGTTAGGAAC
FGAenv atggccaa tcctagtcca caccaaatat ataatgtaac ttgggtaata accaatgtac aaactaacac ccaagctaac gccacctcta tgttaggaac
FZ215-c6 ATGGCCAG TCCTAGTCCA CACCAAATAT ATAATGTAAC TTGGGTAATA ACCAATGTAC AAACTCACAC CCAAGCTAAT GCCACCTCTA TGCTAGGAAC
FA621-c4 ATGGCCAA TCCTAGTCCA CACCAAATAT ATAATGTAAC CTGGGTAATA ACCAATGTAC AAACTAACAC CCAGGCCAAT GCCACCTCTA TGTTAGGAGC
VRA
101
200
FY981p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCCCGCA ATGTAAGACG G.....
FY981-c11 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCCCGCA ATGTAAGATG G.....
FA27-c12 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AGCACCTG GGAACCTATA GTCCCAGACC TTCGGGGG.. .....TGG
FA27-c7 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCAGACC TTCGGGGG.. .....TGG
FA27-c5 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCAGACC TTCGGGGG.. .....TGG
FA27p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCCAGACT TCGGGGGT.. .....TGG
FA621-c1 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACTTAGTGG GA GACACCTG GGAACCTATA GTCCTA.... ..AGGTCG.. .....TGG
FA621-c3 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACTTAGTGG GA GACACCTG GGAACCTATA GTCCTA.... ..AGGTCG.. .....TGG
FX343-c8 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTA.... ..AGGGGA.. .....TGG
SarmaCenv cttaaccgat gcctacccta ccctatatgt tgatttatgt gacctagtgg ga gacacctg ggaacctata gcccagacc caagatct.. .....tgg
FZ215p CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAATCTATG GCCCCAGACC CAAGAAGC.. .....TGG
FA27-c6 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA AACACCTG GGAACCTATA GTCCTAGACC CAACCAATGT AAAACACGGG
FGAenv cttaaccgat gcctacccta ccctacatgt tgacttatgt gacctagtgg ga gacacctg ggaacctata gtcctaaacc caaccaatgt aaaacacggg
FZ215-c6 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACCTAGTGG GA GACACCTG GGAACCTATA GTCCTAGACC CAACCAATGT AAAACACGGG
FA621-c4 CTTAACCGAT GCCTACCCTA CCCTACATGT TGACTTATGT GACTTAGTGG GA GACACCTG GGAACCTATA GTCCTAAACC CAACCAATGT AAAACACGGG


201
300
FY981p GCACATTATT CCTCCTCAAC ACATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACGG ACATACCCCT TTTATGTCTG CCCCGGACAT GCCCCCTCGT
FY981-c11 GCACATTATT CCTCCTCAAC ACATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACGG ACATACCCCT TTTATGTCTG CCCCGGACAT GCCCCCTCGT
FA27-c12 GCACGTTATT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
FA27-c7 GCACGTTATT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
FA27-c5 GCACGTTATT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
FA27p GCAAGTTATT CCTCCTCAAA ATATGGATGT AAAACTGCAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAC GCCCCCTCGT
FA621-c1 GCACATTACT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAACTACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGC
FA621-c3 GCACATTACT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCTACAG ACATACCCCT TTTACGTCTG CCCTGGACAT GCCCCCTCGC
FX343-c8 GCACATTATT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
SarmaCenv gcacgttatt cctcctcaac acatggatgc aaaactacag atagaaaaaa acagcaacaa acataccccct tttatgtctg cccagggcat gccccctcga
FZ215p GCACGTTACT CCTCCTCAAT ACATGGATGT AAAGCTACAG ATAGAAAAAA ACAACAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
FA27-c6 GCACGTTATT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
FGAenv gcacgttact cctcctcaaa atatggatgt aaaactacag atagaaaaaa acagcaacag acatacccct tttacgtctg ccccggacat gcccccctcgt
FZ215-c6 GCACGTTACT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAACAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGT
FA621-c4 GCACGTTACT CCTCCTCAAA ATATGGATGT AAAACTACAG ATAGAAAAAA ACAGCAACAG ACATACCCCT TTTACGTCTG CCCCGGACAT GCCCCCTCGC


149
301
400
FY981p TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT CTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FY981-c11 TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FA27-c12 TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTAGAAGC CCACCTCCTC
FA27-c7 TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FA27-c5 TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGAtGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FA27p TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FA621-c1 TGGGGCCAAA GGGAACACAT TGTGGAGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCCCCTC
FA621-c3 TGGGGCCAAA GGGAACACAT TGTGGAGGGG CACAAGATGG GTTTTGTGCA GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FX343-c8 TGGGGCCAAA GGGGGCACAT TGTGGAGGGG CACAAGACGG GTTTTGTGCC GCATGGGGat gTGAAACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
SarmaCenv tggggcctaa gggaacatat tgtggagggg cacaagatgg gttttgtgcc gcatggggat gtgaaaccac cggagaggct tggtggaagc ccacctcctc
FZ215p TGGGGCCAAA GGGAACACAT TGTGGAGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAAACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FA27-c6 TGGGGCCAAA GGGGACACAT TGTGGTGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FGAenv tggggccaaa gggaacacat tgtggagggg cacaagatgg gttttgtgcc gcatggggat gtgagaccac cggagaagct tggtggaagc ccacctcctc
FZ215-c6 TGGGGCCAAA GGGAACACAT TGTGGAGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAAACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC
FA621-c4 TGGGGCCAAA GGGAACACAT TGTGGAGGGG CACAAGATGG GTTTTGTGCC GCATGGGGAT GTGAGACCAC CGGAGAAGCT TGGTGGAAGC CCACCTCCTC


VRB
401
500
FY981p ATGGGACTAT ATCACAGTGA AAAGAGGGAA TAGT CAGGAC GAAAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FY981-c11 ATGGGACTAT ATCACAGTGA AAAGAGGGAA TAGT CAGGAC GAAAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA27-c12 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA27-c7 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA27-c5 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA27p ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA621-c1 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGGC AATAGCTGTGAG GGAAAATG CAACCCCCTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA621-c3 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGGC AATAGCTGTGAG GGAAAATG CAACCCCCTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FX343-c8 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
SarmaCenv atgggactat atcacagtaa aaagaggaag taat caggac aatagctgtaag ggcaaatg taacccccctg gtcttgcagt tcacccagaa gggaagacaa
FZ215p ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG AGAAAATG CAACCCCCTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA27-c6 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCTTG GTTTTGCAGT TCACTCAGAA GGGAAGACAA
FGAenv atgggactat atcacagtaa aaagagggag tagt caggac aatagctgtgag ggaaaatg caacccccctg gttttgcagt tcacccagaa gggaagacaa
FZ215-c6 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG AGAAAATG CAACCCCCTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
FA621-c4 ATGGGACTAT ATCACAGTAA AAAGAGGGAG TAGT CAGGAC AATAGCTGTGAG GGAAAATG CAACCCCCTG GTTTTGCAGT TCACCCAGAA GGGAAGACAA
150
501
600
FY981p GCCTCTTGGG ACAGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FY981-c11 GCCTCTTGGG ACAGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FA27-c12 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FA27-c7 GCCCCTTGGG ACGGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FA27-c5 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FA27p GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FA621-c1 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGACTAT ATCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FA621-c3 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGACTAT ATCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FX343-c8 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGATTAT ACCGTGCAGG ATATGACCCT ATCGCTTTGT TCACGGTGTC CCGGCAGGTA TCAACCATTA
SarmaCenv gcctcttggg acagacctaa aatgtggggg ctacgactat accgttcagg atatgaccct atagccctgt tctcggtatc ccggcaagta atgaccatta
FZ215p GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGACTAT ACCGTTCAGG ATATGACCCT GTCACTTTAT TCACGGTGTC CCGGCAGGTG TCAACCATTA
FA27-c6 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGATTAT ACCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA TCAACCATTA
FGAenv gcctcttggg acggacctaa gatgtgggga ttgcgactat accgtacagg atatgaccct atcgctttat tcacggtgtc ccggcaggta tcaaccatta
FZ215-c6 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGACTAT ACCGTACAGG ATATGACCCT GTCACTTTAT TCACGGTGTC CCGGCAGGTG TCAACCATTA
FA621-c4 GCCTCTTGGG ACGGACCTAA GATGTGGGGA TTGCGACTAT ATCGTACAGG ATATGACCCT ATCGCTTTAT TCACGGTGTC CCGGCAGGTA ATCACCATTA

```
                601                                                                                                  700
   FY981p CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCAACCC AGAGGCCCCA
FY981-c11 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCAACCC AGAGGCCCCA
 FA27-c12 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCGAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
  FA27-c7 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCGAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
  FA27-c5 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCGAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
    FA27p CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCGAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
 FA621-c1 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAGAAACCC CCATCCCGAC AATCTCAAAC AAGGTCCAAA GTGGCAACCC AGAGGCCCCA
 FA621-c3 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAGAAACCC CCATCCCGAC AATCTCAAAC AAGGTCCAAA GTGGCAACCC AGAGGCCCAA
 FX343-c8 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
SarmaCenv cgccgcctca ggcgatggga cccaacttag tcttacctga tcaaaaaccc ccatcccgac aatctcaaac aaagtccaag gtgacaaccc agaggcccca
   FZ215p CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCCTACCTAA TCAAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGATGCCCCA
  FA27-c6 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
    FGAenv cgccgcctca ggcaatggga ccaaacctag tcttacctga tcaaaaaccc ccatcccgac aatctcaaac agggtccaaa gtggcgaccc agaggcccca
 FZ215-c6 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAAAAACCC CCATCCCGAC AATCTCAAAC AGGGTCCAAA GTGGCGACCC AGAGGCCCCA
 FA621-c4 CGCCGCCTCA GGCAATGGGA CCAAACCTAG TCTTACCTGA TCAGAAACCC CCATCCCGAC AATCTCAAAC AAGGTCCAAA GTGGCAACCC AGAGGGCCCA
```

151

```
                701                                                                                                  800
   FY981p AAC AAATGA AAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGAGTCCC AAACGAATTG GGACCGGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
FY981-c11 AACAAAATGA AAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGAGTCCC AAACGAATTG GGACCGGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
 FA27-c12 AAC.AAATGA AAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCGGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
  FA27-c7 AAC.AAATGA AAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCGGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
  FA27-c5 AAC.AAATGA GAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCAGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
    FA27p AAC.AAATGA AAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCAGAGA TAGGCTAATA AATTTAGTAC AAGGGACATA
 FA621-c1 AACGAATGTA AGCGCCCCAA GGTCCGTTGC CCCCCACCGC CATGGGTCCC AAACGGATTG GGACCAGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
 FA621-c3 ACGAATTGTC AAGCGCCCAA GGTCCGTTGC CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCAGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
 FX343-c8 AAC.AAATGA AAGCGCCCCA AGGTCTGTTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCAGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
SarmaCenv aat.aactag cagcacccca aggtctgtcg cctccgctac catgggtccc aaacggatag ggaccggaga tagattaata aatttagtgc aagggacata
   FZ215p AAC.GAATGA AAGCGCCCCA AGATCTGTTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCGGAGA CAGGTTAATA AATTTAGTAC AAGGGACATA
  FA27-c6 AAC.AAATGA AAGCGCCCCA AGGTCTATTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCGGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
    FGAenv aac.gaatga aagcgcccca aggtctgttg cccccaccac catgggtccc aaacggattg ggaccggaga taggttaata aatttagtac aagggacata
 FZ215-c6 AAC.GAATGA AAGCGCCCCA AGATCTGTTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCGGAGA CAGGTTAATA AATTTAGTAC AAGGGACATA
 FA621-c4 AAC.GAATGT AAGCGCCCCA AGGTCCGTTG CCCCCACCAC CATGGGTCCC AAACGGATTG GGACCAGAGA TAGGTTAATA AATTTAGTAC AAGGGACATA
```

```
            801                                                                                                 900
FY981p      TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FY981-c11   TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FA27-c12    TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GTCTAGTCTC CCGACCGCCC TATTATGAAG GAATTGCGGT ATTGGGCAAT
FA27-c7     TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GTCTAGTCTC CCGACCGCCC TATTATGAAG GAATTGCGGT ATTGGGCAAT
FA27-c5     TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GTCTAGTCTC CCGACCGCCC TATTATGAAG GAATTGCGGT ATTGGGCAAT
FA27p       CCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FA621-c1    CCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FA621-c3    CCTAgCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FX343-c8    TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAaAGAC TGTTGGCTCT GCcTGGTCTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
SarmaCenv   cctagcctta aatgccaccg accccaacaa aactaaagac tgttggctct gcctggtttc tcgaccacct tattacgaag ggattgcagt cttaggtaac
FZ215p      CCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FA27-c6     TCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FGAenv      cctagcctta aatgccaccg accccaacaa aactaaagac tgttggctct gcctggtttc tcgaccaccc tattacgaag ggattgcaat cttaggtaac
FZ215-c6    CCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTTTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
FA621-c4    CCTAGCCTTA AATGCCACCG ACCCCAACAA AACTAAAGAC TGTTGGCTCT GCCTGGTCTC TCGACCACCC TATTACGAAG GGATTGCAAT CTTAGGTAAC
```



```
            901                                                                                                1000
FY981p      TACAGCAACC AAACAAACCC CCCCCCATCC TGCCAATCTA TTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FY981-c11   TACAGCAACC AAACAAACCC CCCCCCATCC TGCCAATCTA TTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FA27-c12    TACAGCAACC AAACCAACCC CCCCCCATCC TGCCTATCCG ACCCACAACA TAAACTGACT ATATCAGAAG TGTCCGGGCA AGGTTTGTGC ATAGGGACTG
FA27-c7     TACAGCAACC AAACCAACCC CCCCCCATCC TGCCTATCCG ACCCACAACA TAAACTGACT ATATCAGAAG TGTCCGGGCA AGGTTTGTGC ATAGGGACTG
FA27-c5     TACAGCAACC AAACCAACCC CCCCCCATCC TGCCTATCCG ACCCACAACA TAAACTGACT ATATCAGAAG TGTCCGGGCA AGGTTTGTGC ATAGGGACTG
FA27p       TACAGCAACC AAACAAACCC CCCCCCATCC TGCCTATCTA TTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FA621-c1    TACAGCAACC AAACCAACCC CCCCCCATCC TGCCTGTCTA CTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FA621-c3    TACAGCAACC AAACCAACCC CCCCCCATCC TGCCTGTCTA CTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FX343-c8    TACAGCAACC AAACAAACCC CCCCCCATCC TGCCTATCTA TTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACCGTGC ATAGGGACTG
SarmaCenv   tacagcaacc aaacaaaccc cccccccatcc tgcctatcta ccccgcaaca taaactgact atatcagaag tgtccgggca aggtttgtgc atagggactg
FZ215p      TATAGCAACC AAACAAACCC CCCCCCATCC TGCCTATCTA TTCCGCAACA CAAACTGACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FA27-c6     TACAGCAACC AAACAAACCC CCCCCCATCC TGCCTATCTA TTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FGAenv      tacagcaacc aaacaaaccc cccccccatcc tgcctatcta ctccgcaaca caaactaact atatctgaag tatcagggca aggaatgtgc atagggactg
FZ215-c6    TATAGCAACC AAACAAACCC CCCCCCATCC TGCCTATCTA TTCCGCAACA CAAACTGACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
FA621-c4    TACAGCAACC ACACCAACCC CCCCCCATCC TGCCTGTCTA CTCCGCAACA CAAACTAACC ATATCTGAAG TATCAGGGCA AGGACTGTGC ATAGGGACTG
```

```
            1001                                                                                                1100
FY981p      TTCCTAAGAC CCACCAGGCT CTGTGCAATG AGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
FY981-c11   TTCCTAAGAC CCACCAGGCT CTGTGCAATG AGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
FA27-c12    TTCCTAAGAC CCACCAAGCT TTGTGCAAAA AGACACAAAA AGGACATAAA GGGACTCACT ATCTAGCAGC CCCTAGCGGT ACCTACTGGG CATGCAACAC
FA27-c7     TTCCTAAGAC CCACCAAGCT TTGTGCAAAA AGACACAAAA AGGACATAAA GGGACTCACT ATCTAGCAGC CCCTAGCGGT ACCTACTGGG CATGCAACAC
FA27-c5     TTCCTAAGAC CCACCAAGCT TTGTGCAAAA AGACACAAAA AGGACATAAA GGGACTCACT ATCTAGCAGC CCCTAGCGGT ACCTACTGGG CCTGTAACAC
FA27p       TTCCTAAGAC CCACCAGGCT TTGTGCAATG AGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
FA621-c1    TTCCTAAGAC CCACCAGGCT TTGTGCAATG CGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
FA621-c3    TTCCTAAGAC CCACCAGGCT TTGTGCAATG CGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
FX343-c8    TTCCTAAGAC CCACCAGGCT TTGTGCAATA AGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
SarmaCenv   ttcctaagac ccaccaagct ttgtgcaaaa agacacaaaa aggacataaa gggactcact acctggcagc ccccaacggc acctattggg cctgtaacac
FZ215p      TTCCTAAGAC CCACCAGGCT TTGTGCAATG AGACGCAACA GGGACATACA GGGGCACACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CTTGTAACAC
FA27-c6     TTCCTAAGAC CCACCAGGCT TTGTGCAATG AGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
FGAenv      ttcctaaaac ccaccaggct ttgtgcaata agacacaaca gggacataca ggggcgcact atctagccgc ccccaacggc acctattggg cctgtaacac
FZ215-c6    TTCCTAAGAC CCACCAGGCT TTGTGCAATG AGACGCAACA GGGACATACA GGGGCACACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CTTGTAACAC
FA621-c4    TTCCTAAGAC CCACCAGGCT TTGTGCAATG CGACACAACA GGGACATACA GGGGCGCACT ATCTAGCCGC CCCCAACGGC ACCTATTGGG CCTGTAACAC
```



```
            1101                                                                                                1200
FY981p      TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FY981-c11   TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA27-c12    CGGGCTAACC CCATGCATCT CCATGGCAGT GCTCAACTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA27-c7     CGGGCTAACC CCATGCATCT CCATGGCAGT GCTCAACTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA27-c5     TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA27p       TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA621-c1    TGGACTCACC CCATGCATCT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA621-c3    TGGACTCACC CCATGCATCT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FX343-c8    TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
SarmaCenv   tggactcacc ccatgcattt ccatggcagt gctcaattgg acctctgatt tttgtgtctt aatcgaatta tggcccagag taacctacca tcaacccgaa
FZ215p      TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCATTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA27-c6     TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FGAenv      tggactcacc ccatgcattt ccatggcggt gctcaattgg acctctgatt tttgtgtctt aatcgaatta tggcccagag tgacttacca tcaacccgaa
FZ215-c6    TGGACTCACC CCATGCATTT CCATGGCGGT GCTCAATTGG ACCGCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
FA621-c4    TGGACTCACC CCATGCATCT CCATGGCGGT GCTCAATTGG ACCTCTGATT TTTGTGTCTT AATCGAATTA TGGCCCAGAG TGACTTACCA TCAACCCGAA
```

```
            1201                                                                                                1300
FY981p      TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GGAACCGAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FY981-c11   TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GGGAACCGAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA27-c12    TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA27-c7     TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA27-c5     TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA27p       TATGTGTACA CACATTTTGC CGAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA621-c1    TATGTGTACT CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCGCTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA621-c3    TATGTGTACT CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCGCTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FX343-c8    TATGTGTACA CACATTTTGC CGAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
SarmaCenv   tatatttaca cacatttcga caaagctgtc aggttccgaa gagaacctat atcactaacc gttgccctta tgttgggagg actcaccgta gggggcatag
FZ215p      TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCGGT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA27-c6     TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FGAenv      tatgtgtaca cacattttgc caaagctgtc aggttccgaa gagaaccaat atcactaacg gttgccctta tgttgggagg acttactgta gggggcatag
FZ215-c6    TATGTGTACA CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAGT ATCACTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
FA621-c4    TATGTGTACT CACATTTTGC CAAAGCTGTC AGGTTCCGAA GAGAACCAAT ATCGCTAACT GTTGCCCTTA TGTTGGGAGG ACTCACTGTA GGGGGCATAG
```



```
            1301                                                                                                1400
FY981p      CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FY981-c11   CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FA27-c12    CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FA27-c7     CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FA27-c5     CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FA27p       CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FA621-c1    CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAGGAGTC
FA621-c3    CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCTTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAGGAGTC
FX343-c8    CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATACAGGCCC TAGAAGAGTC
SarmaCenv   ccgcgggggt cggaacaggg actaaagccc tccttgaaac agcccagttc agacaactac aaatagccat gcacacagac atccaggccc tggaagagtc
FZ215p      CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACGGAC ATCCAAGCCC TAGAAGAGTC
FA27-c6     CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAAGAGTC
FGAenv      ccgcgggggt cggaacaggg actaaagccc tccttgaaac agcccagttc agacaactac aaatggccat gcacacagac atccaggccc tagaagaatc
FZ215-c6    CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAAGCCC TAGAAGAGTC
FA621-c4    CCGCGGGGGT CGGAACAGGG ACTAAAGCCC TCCTTGAAAC AGCCCAGTTC AGACAACTAC AAATGGCCAT GCACACAGAC ATCCAGGCCC TAGAGGAGTC
```

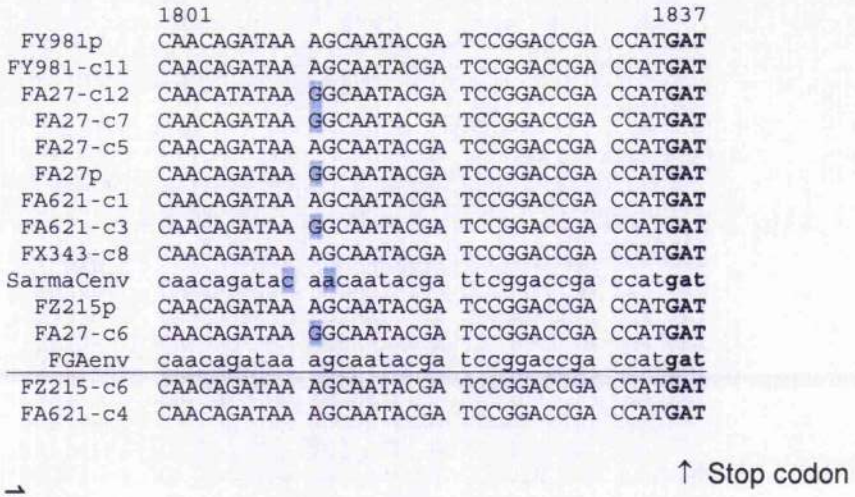

1401
1500
FY981p AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FY981-c11 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA27-c12 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA27-c7 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA27-c5 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA27p AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA621-c1 AATTAGTGCC TTAGAAAAGT CCCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGACTCTGT
FA621-c3 AATTAGTGCC TTAGAAAAGT CCCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGACTCTGT
FX343-c8 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
SarmaCenv aattagtgcc ttagaaaaat ccctgacctc cctctctgag gtagtcctac aaaatAggcg gggcctagat attctgttct tacaagaggg agggctctgt
FZ215p AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA27-c6 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FGAenv aattagtgcc ttagaaaagt ccctgacctc cctttctgaa gtagtcttac aaaacagacg gggcctagat attctattct tacaagaggg agggctctgt
FZ215-c6 AATTAGTGCC TTAGAAAAGT CTCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGGCTCTGT
FA621-c4 AATTAGTGCC TTAGAAAAGT CCCTGACCTC CCTTTCTGAA GTAGTCTTAC AAAACAGACG GGGCCTAGAT ATTCTATTCT TACAAGAGGG AGGACTCTGT

1501
1600
FY981p GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FY981-c11 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA27-c12 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCGCA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA27-c7 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA27-c5 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA27p GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA621-c1 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTTGT CCGTGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA621-c3 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTTGT CCGTGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FX343-c8 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAGC
SarmaCenv gccgcattaa aagaagaatg ctgcttctat gcagatcaca ccggactcgt ccgagacaat atggctaaat taagagaaag actaaaacag cggcaacaac
FZ215p GCCGCATTGA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCCAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FA27-c6 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC
FGAenv gccgcattga aagaagaatg ttgcttctat gcggatcaca ccggactcgt ccgagacaat atggccaaat taagagaaag actaaaacag cggcaacaac
FZ215-c6 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTCGT CCGAGACAAT ATGGCCAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAT
FA621-c4 GCCGCATTAA AAGAAGAATG TTGCTTCTAT GCGGATCACA CCGGACTTGT CCGTGACAAT ATGGCTAAAT TAAGAGAAAG ACTAAAACAG CGGCAACAAC

```
          1601                                                                                                1700
FY981p    TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
FY981-c11 TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
FA27-c12  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC GGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
 FA27-c7  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
 FA27-c5  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
 FA27p    TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
FA621-c1  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATCTC CTCCATTATG GGCCCCTTAC TAATCCTACT
FA621-c3  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATCTC CTCCATTATG GGCCCCTTAC TAATCCTACT
FX343-c8  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
SarmaCenv tgtttgattc ccaacaggga tggtttgaag gatggttcaa caagtccccc tggtttacaa ccctaatttc ctccatcatg ggccccttac taatcctact
 FZ215p   TGTTTGACTC CCAACAGGGA GGGTTTGAAG GATGGTTCAA CAAGTCCCCT AGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
 FA27-c6  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
  FGAenv  tgtttgactc ccaacaggga tggtttgaag gatggttcaa caagtccccc tggtttacaa ccctaatttc ctccattatg ggccccttac taatcctact
FZ215-c6  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
FA621-c4  TGTTTGACTC CCAACAGGGA TGGTTTGAAG GATGGTTCAA CAAGTCCCCC TGGTTTACAA CCCTAATTTC CTCCATTATG GGCCCCTTAC TAATCCTACT
```



```
          1701                                                                                                1800
FY981p    CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
FY981-c11 CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
FA27-c12  CCTAATTCTC CTCTTCGGCC CATgCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
 FA27-c7  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
 FA27-c5  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
 FA27p    CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
FA621-c1  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
FA621-c3  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
FX343-c8  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCA TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
SarmaCenv cctaattctc ctcctcggcc catgcatcct taaccgatta gtgcaattcg taaaagacag aatatctgtg gtacaagcct taattttaac ccaacagtac
 FZ215p   CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
 FA27-c6  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
  FGAenv  cctaattctc ctcttcggcc catgcatcct taaccgatta gtacaattcg taaaagacag aatatctgtg gtacaggctt taattttaac ccaacagtac
FZ215-c6  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
FA621-c4  CCTAATTCTC CTCTTCGGCC CATGCATCCT TAACCGATTA GTACAATTCG TAAAAGACAG AATATCTGTG GTACAAGCTT TAATTTTAAC CCAACAGTAC
```


1801
1837
FY981p CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
FY981-c11 CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
FA27-c12 CAACATATAA GGCAATACGA TCCGGACCGA CCATGAT
FA27-c7 CAACAGATAA GGCAATACGA TCCGGACCGA CCATGAT
FA27-c5 CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
FA27p CAACAGATAA GGCAATACGA TCCGGACCGA CCATGAT
FA621-c1 CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
FA621-c3 CAACAGATAA GGCAATACGA TCCGGACCGA CCATGAT
FX343-c8 CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
SarmaCenv caacagatac aacaatacga ttcggaccga ccatgat
FZ215p CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
FA27-c6 CAACAGATAA GGCAATACGA TCCGGACCGA CCATGAT
FGAenv caacagataa agcaatacga tccggaccga ccatgat
FZ215-c6 CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
FA621-c4 CAACAGATAA AGCAATACGA TCCGGACCGA CCATGAT
↑ Stop codon

FA27

When primary isolate FA27 was plated on the selective cell line 104C1, it yielded four unique *env* clones. The A-component, FA27-C6, differed at seven amino acids from the published FeLV-A/Glasgow-1 sequence ($D_{51}N$, $N_{59}D$, $V_{247}I$, $T_{318}I$, $M_{333}L$, $K_{348}E$ and $K_{614}R$) as can be seen in figure 4.1. These seven amino acids were also present in the FA27p amino acid sequence, suggesting that these amino acids are representative of FA27 and that this *env* clone is the true A component of FA27. The C components (FA27-C5, FA27-C7 and FA27-C12) displayed a high degree of similarity to each other and to the FA27p *env* sequence. The three *env* clones were slightly different from one another, but there were four amino acid changes present in all three *env* clones that were different from both FeLV-A/Glasgow-1 and the FA27p amino acid sequence. The first difference can be found in the VRA region, namely a $H_{67}L$ substitution when compared to FeLV-A/Glasgow-1 amino acid or a $F_{67}L$ substitution when compared to the FA27p amino acid sequence. The $H_{67}L$ substitution could result in the loss of a positive polar side chain, as leucine has a non-polar hydrophobic side chain whereas histidine contains a positively charged polar side chain. When compared to FA27, there is no difference in the nature of the side chain as both amino acids contain non-polar hydrophobic side chains. At position 301, there is an I to V substitution, a mutation that is also present in the FeLV-C/Sarma amino acid sequence. I to V does not result in a change in polarity. Further along the sequence, there is a $T_{318}D$ mutation relative to FeLV-A/Glasgow-1 or an $I_{318}D$ substitution relative to FA27. Both threonine (T) and isoleucine (I) contain two centres of asymmetry and both amino acids do not contain a polar side chain, however threonine is hydrophilic whereas isoleucine is hydrophobic. As aspartate (D) has a negatively charged side chain, the overall polarity is altered. Lastly, a six amino acid stretch (^{351}KGHKGT^{356}) was found in all three *env* clones that was not present in the FA27p amino acid sequence. In the FA27p amino acid sequence, this stretch was identical to the stretch found in FeLV-A/Glasgow-1, namely ^{351}QGHTGA^{356}, however in the FA27 clones and indeed in FeLV-C/Sarma this stretch reads ^{351}KGHKGT^{356}. Both glutamine (Q) and threonine (T) contain a negatively charged side chain, whereas lysine (K) contains a positively charged side chain, resulting in an altered polarity. In addition, the $A_{356}T$ substitution results in the substitution from a hydrophobic residue to a hydrophilic residue.

The hydrophobic side chain in alanine is a side chain that will avoid contact with water and thus this residue will be on the inside of a protein structure, whereas the hydrophilic side chain in threonine is a side chain that encourages contact with water and thus this residue will favour a location on the outside of a protein structure, or will co-localise with other hydrophilic residues within a structure. Therefore, the substitution of a hydrophobic residue with a hydrophilic residue may alter the protein structure. When compared with the FA27p amino acid sequence and the other C components identified, clone FA27-C5 is the sole *env* clone that has maintained a lysine on position 614, which have substituted this lysine to arginine, although this substitution does not alter the overall polarity. With the exception of two amino acid differences ($R_{260}G$ and $R_{614}K$), clone FA27-C7 was identical to clone FA27-C5. Clone FA27-C12 had several substitutions that were not present in the other C components of FA27; these were $N_{53}S$, $H_{523}R$, $W_{561}G$ and $Q_{612}H$ when compared to FA27-C5 and FA27-C7. Only one of these changes resulted in a change in the polarity of the side chain, this being $Q_{612}H$, changing the side chain from an uncharged, polar residue to a positively charged, hydrophilic residue. Interestingly, the $K_{348}E$ substitution that is present in both clone FA27-C6 (the A-component of this primary isolate) and the FA27p clone, is not present in the other three clones identified as C components (FA27-C5, FA27-C7 and FA27-C12). This substitution alters the polarity of the side chain of these amino acids from positively charged for lysine to negatively charged for glutamic acid.

FY981

Env clone FY981-C11 was identical to clone FY981p. The substitution present in *env* clone FY981-14 (characterised in chapter three), $T_{23}N$ relative to FY981p amino acid sequence, is not present in the FY981-C11 clone described in this part of the thesis.

FX343

The *env* clone isolated after biological selection of primary isolate FX343, FX343-C8, appeared to be the C component of this primary isolate. However, as this was the only clone identified for this isolate, the changes found in the amino acid sequence relative to FeLV-A Glasgow-1 cannot be confirmed as representative of the parent sequence.

As discussed previously in chapter three, an isolate was classified as subgroup C when the VRA region of the isolate contained length polymorphisms relative to the VRA region of FeLV-A/Glasgow-1 or displayed substantial divergence from FeLV-A/Glasgow-1. With fifty-one amino acids, the VRA region of clone FX343-C8 was shorter than the VRA region of FeLV-A/Glasgow-1, which contains fifty-six amino acids. Furthermore, there were six additional amino acid differences in the envelope glycoprotein sequence when compared with FeLV-A/Glasgow-1. The first substitution is situated at the end of the VRA region, $T_{99}A$, substituting a hydrophilic residue for a hydrophobic residue. As discussed previously, the substitution of a residue with a hydrophobic side chain with a residue with a hydrophilic side chain may affect the protein structure. Similarly, the converse substitution of a hydrophilic residue with a hydrophobic residue may also affect the protein structure. An identical substitution was found after the VRB region ($T_{186}A$), again affecting the charge. As a $T_{318}I$ substitution was also present in FX343-26 and FX343-28, the $T_{318}I$ mutation would appear to be representative of FX343. $T_{318}I$ results in a change of polarity of the side chain (threonine has a hydrophilic side chain while isoleucine has a hydrophobic side chain); therefore the structure may be altered in this region of the envelope glycoprotein. $M_{333}P$ may affect the structure, as proline is a residue that contributes to bends in a polypeptide sequence and can be considered hydrophilic, whereas methionine (M) has a hydrophobic side chain. $G_{260}R$ results in a change of polarity of the side chain, as glycine has a neutral side chain and arginine contains a basic side chain. The remaining two amino acid differences between FX343-C8 and FeLV-A/Glasgow-1 may not alter the overall polarity of the side chains ($M_{333}P$ and $V_{594}I$).

FA621

Biological selection of primary isolate FA621 on the selective cell line 104C1 yielded three unique *env* clones, one of which (FA621-C4) was the A-component and two clones (FA621-C1 and FA621-C3) that were C components. As with primary isolate FX343, no representative amino acid sequence was known, therefore it was not possible to distinguish between determinants of the subgroup C phenotype and changes in the sequence that were unrelated to structure and function. There are five amino acid differences between the FA621 clones and the published FeLV-A/Glasgow-1 sequence that are present in all three *env* clones, therefore suggesting that these differences are unique to the primary isolate.

There is one subsequent change that is present in FA621-C4 (the A-component) and FA621-C3 (a C component), but not present in the other C component (FA621-C1), this is a $T_{35}A$ substitution prior to the VRA region, which is a change from a hydrophilic residue to a hydrophobic residue. With the exception of the substitutions listed, there is one additional change in the FA621-C4 sequence when compared to FeLV-A/Glasgow-1; $S_{147}R$ two amino acids upstream of the VRB region. The change from an uncharged hydrophilic residue to a positively charged hydrophilic residue could be significant, taking its location into consideration. Both C components of FA621 were very similar, however there were some differences between the amino acid sequences. Clone FA621-C3 has a valine residue rather than an isoleucine at the start of the mature protein (a similar residue is found in the published sequence of FeLV-C/Sarma). There is one additional change that can only be found in clone FA621-C3, this is $T_{241}A$, as discussed previously resulting in a change from a hydrophilic residue to a hydrophobic residue.

FZ215

Biological selection of primary isolate FZ215 yielded two unique *env* clones, FZ215-C6 and FZ215-C7, with clone FZ215-C6 resembling FeLV-A/Glasgow-1 and clone FZ215-C7 resembling FeLV-B/Gardner Arnstein. The A-component of the primary isolate, clone FZ215-C6, contained seven amino acid differences relative to FeLV-A/Glasgow-1, these were also present in clone FZ215p and thus these changes were thought to be representative of the parent isolate. There were two additional changes that were found neither in the published FeLV-A/Glasgow-1 amino acid sequence nor in the FZ215p amino acid sequence. These substitutions were a $N_{24}H$ substitution (which results in a change from an uncharged side chain to a positively charged polar side chain) and a $S_{382}A$ substitution. Serine is the hydroxylated version of alanine, with this difference that alanine is hydrophobic and serine is hydrophilic, therefore this could alter the structure of the protein.

A separate amino acid alignment was performed on the two isolates that closely resemble FeLV-B Gardner-Arnstein (FZ215-C7 and FZ215-12, described in chapter three, see figure 3.7; the isolates were compared to the published sequences of FeLV-B/Gardner Arnstein (Accession number P03391), FeLV-B/Snijder-Theilen (Accession number K01208) and the parent isolate FZ215. The two *env* genes isolated are very similar to one another, indeed there is only one amino acid difference between the two isolates, namely FZ215-C7 has a histidine on position 590 rather than the predicted glutamine in both FeLV-B/Gardner Arnstein and FZ215-12. To sum up, there are two changes in these *Env* amino acid sequences that cannot be found in either the FeLV-B/Gardner Arnstein amino acid sequence or the parent FZ215 amino acid sequence. These changes are $L_{41}I$ just prior to the VRA region and $T_{165}I$ (compared to the FeLV-B/Gardner Arnstein sequence) in the VRB region. As discussed previously in chapter three, the latter substitution results in a change of polarity of the side residues from uncharged polar to hydrophobic, possibly resulting in a change of structure of the envelope glycoprotein.

Env clone #	Nucleotide similarity to FeLV-A Glasgow 1	Amino acid similarity to FeLV-A Glasgow 1	Nucleotide similarity to FeLV-C/Sarma	Amino acid similarity to FeLV-C/Sarma	Subgroup determined by sequencing
FA27-C5	95.8%	97.7%	89.5%	95.1%	C
FA27-C6	98.5%	99.3%	90.7%	94.6%	A
FA27-C7	94.9%	97.5%	90.2%	94.9%	C
FA27-C12	95.8%	97.0%	89.5%	97.0%	C
FA621-C1	96.9%	98.0%	90.8%	94.2%	C
FA621-C3	95.1%	98.0%	89.2%	94.4%	C
FA621-C4	97.9%	98.5%	90.9%	94.0%	A
FY981-C11	96.4%	97.7%	89.6%	93.9%	C
FZ215-C6	98.0%	98.7%	89.8%	93.9%	A
FZ215-C7	76.6%	89.7%	77.6%	88.4%	B
FX343-C8	97.5%	98.3%	91.0%	94.1%	C

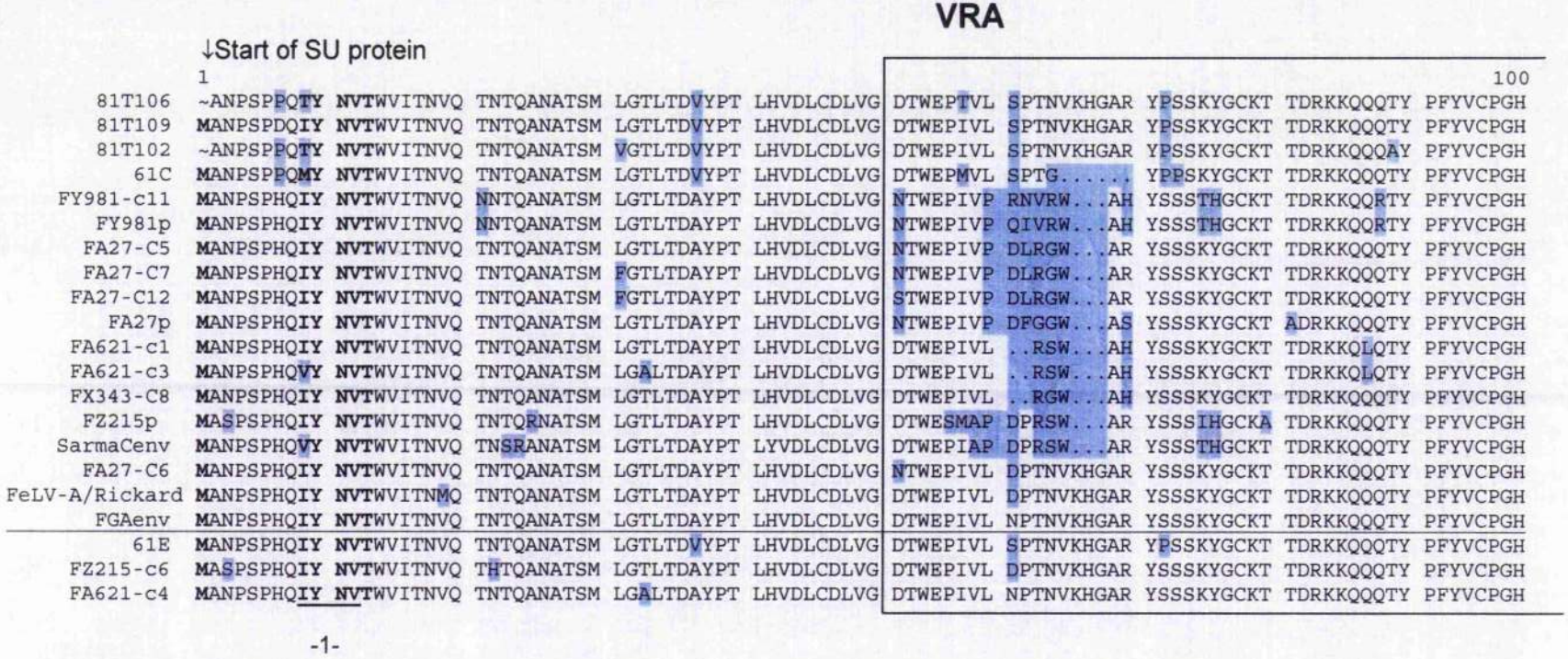
Table 4.2: *Nucleotide and amino acid sequence homology of the env genes isolated following selection in 104C1 cells. Sequences were compared with FeLV-A/Glasgow-1 and FeLV-C/Sarma.*

Env clone #	Amino acid similarity to FeLV-A Glasgow 1	Amino acid similarity to FeLV-C/Sarma	Predicted subgroup determined by sequencing
FA27-C5	95.9%	93.9%	C
FA27-C6	98.6%	93.8%	A
FA27-C7	95.2%	93.2%	C
FA27-C12	94.6%	92.6%	C
FA621-C1	95.6%	92.6%	C
FA621-C3	95.9%	92.6%	C
FA621-C4	98.7%	91.8%	C
FY981-C11	93.9%	89.9%	C
FZ215-C6	97.2%	93.2%	A
FZ215-C7	98.0%	93.1%	B
FX343-C8	75.4%	75.0%	C

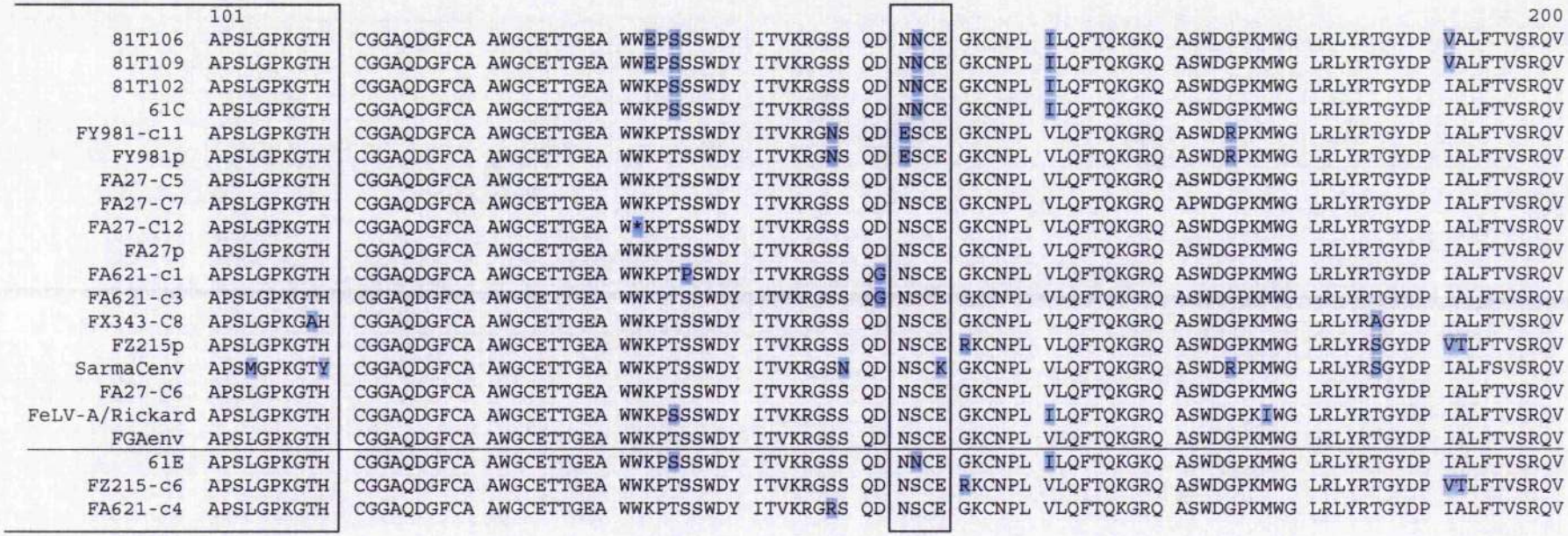
Table 4.3: *Amino acid sequence homology of the VRA region of the novel env genes relative to FeLV-A/Glasgow-1 and FeLV-C/Sarma. The VRA region is a stretch of fifty-six amino acids starting from amino acid fifty, relative to the published sequence of FeLV-A/Glasgow-1 (D_{53} to H_{110}).*

Figure 4.3: *Amino acid alignment from the SU proteins of the isolated unique envelope regions of various isolated subgroups. The arrow indicates the start of the SU protein. A dot indicates a gap in alignment, whereas an asterisk indicates a conserved cysteine residue. The standard one letter code for amino acids is used and the amino acid position is indicated above the figure. A position of 1 in this figure corresponds with position 6079 in the published FeLV A genome. The immunodeficiency defining sequence (region 3 in the figure) has been described previously by Rohn et al (Rohn J.L. et al., 1998). The importance of the histidine residue on position 8 has been discussed previously (Bae Y. et al., 1997). The Nunberg-Elder epitope (region 2 in the figure)has been described previously by Nunberg et al and Elder et al (Nunberg J.H. et al., 1984) and (Elder J.H. et al., 1987).*

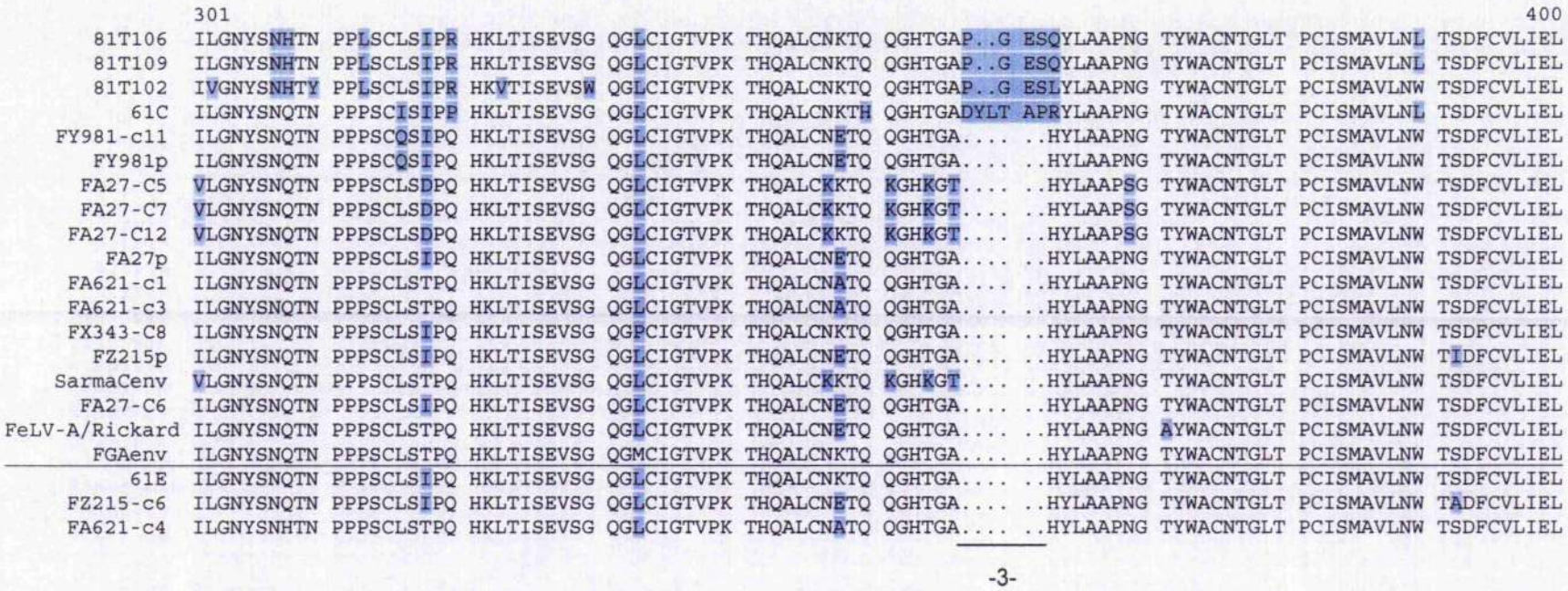
(Opposite page)


VRA
↓Start of SU protein
1
100
81T106 ~ANPSPPQTY NVTWVITNVQ TNTQANATSM LGTLTDVYPT LHVDLCDLVG DTWEPTVL SPTNVKHGAR YPSSKYGCKT TDRKKQQQTY PFYVCPGH
81T109 MANPSPDQIY NVTWVITNVQ TNTQANATSM LGTLTDVYPT LHVDLCDLVG DTWEPIVL SPTNVKHGAR YPSSKYGCKT TDRKKQQQTY PFYVCPGH
81T102 ~ANPSPPQTY NVTWVITNVQ TNTQANATSM VGTLTDVYPT LHVDLCDLVG DTWEPIVL SPTNVKHGAR YPSSKYGCKT TDRKKQQQAY PFYVCPGH
61C MANPSPPQMY NVTWVITNVQ TNTQANATSM LGTLTDVYPT LHVDLCDLVG DTWEPMVL SPTG..... YPPSKYGCKT TDRKKQQQTY PFYVCPGH
FY981-c11 MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVP RNVRW...AH YSSSTHGCKT TDRKKQQRTY PFYVCPGH
FY981p MANPSPHQIY NVTWVITNVQ NNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVP QIVRW...AH YSSSTHGCKT TDRKKQQRTY PFYVCPGH
FA27-C5 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVP DLRGW...AR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FA27-C7 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG NTWEPIVP DLRGW...AR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FA27-C12 MANPSPHQIY NVTWVITNVQ TNTQANATSM FGTLTDAYPT LHVDLCDLVG STWEPIVP DLRGW...AR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FA27p MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVP DFGGW...AS YSSSKYGCKT ADRKKQQQTY PFYVCPGH
FA621-c1 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL ..RSW...AH YSSSKYGCKT TDRKKQLQTY PFYVCPGH
FA621-c3 MANPSPHQVY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVL ..RSW...AH YSSSKYGCKT TDRKKQLQTY PFYVCPGH
FX343-C8 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL ..RGW...AH YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FZ215p MASPSPHQIY NVTWVITNVQ TNTQRNATSM LGTLTDAYPT LHVDLCDLVG DTWESMAP DPRSW...AR YSSSIHGCKA TDRKKQQQTY PFYVCPGH
SarmaCenv MANPSPHQVY NVTWVITNVQ TNSRANATSM LGTLTDAYPT LYVDLCDLVG DTWEPIAP DPRSW...AR YSSSTHGCKT TDRKKQQQTY PFYVCPGH
FA27-C6 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG NTWEPIVL DPTNVKHGAR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FeLV-A/Rickard MANPSPHQIY NVTWVITNMQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL DPTNVKHGAR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FGAenv MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL NPTNVKHGAR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
61E MANPSPHQIY NVTWVITNVQ TNTQANATSM LGTLTDVYPT LHVDLCDLVG DTWEPIVL SPTNVKHGAR YPSSKYGCKT TDRKKQQQTY PFYVCPGH
FZ215-c6 MASPSPHQIY NVTWVITNVQ THTQANATSM LGTLTDAYPT LHVDLCDLVG DTWEPIVL DPTNVKHGAR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
FA621-c4 MANPSPHQIY NVTWVITNVQ TNTQANATSM LGALTDAYPT LHVDLCDLVG DTWEPIVL NPTNVKHGAR YSSSKYGCKT TDRKKQQQTY PFYVCPGH
-1-

VRB


101
200
81T106 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWEPSSSWDY ITVKRGSS QD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP VALFTVSRQV
81T109 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWEPSSSWDY ITVKRGSS QD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP VALFTVSRQV
81T102 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSS QD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
61C APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSS QD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FY981-c11 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNS QD ESCE GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRTGYDP IALFTVSRQV
FY981p APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGNS QD ESCE GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRTGYDP IALFTVSRQV
FA27-C5 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA27-C7 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ APWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA27-C12 APSLGPKGTH CGGAQDGFCA AW*KPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA27p APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA621-c1 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTPSWDY ITVKRGSS QG NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FA621-c3 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QG NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FX343-C8 APSLGPKGAH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRAGYDP IALFTVSRQV
FZ215p APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE RKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRSGYDP VTLFTVSRQV
SarmaCenv APSMGPKGTY CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSN QD NSCK GKCNPL VLQFTQKGRQ ASWDRPKMWG LRLYRSGYDP IALFSVSRQV
FA27-C6 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FeLV-A/Rickard APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSS QD NSCE GKCNPL ILQFTQKGRQ ASWDGPKIWG LRLYRTGYDP IALFTVSRQV
FGAenv APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
61E APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPSSSWDY ITVKRGSS QD NNCE GKCNPL ILQFTQKGKQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV
FZ215-c6 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGSS QD NSCE RKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP VTLFTVSRQV
FA621-c4 APSLGPKGTH CGGAQDGFCA AWGCETTGEA WWKPTSSWDY ITVKRGRS QD NSCE GKCNPL VLQFTQKGRQ ASWDGPKMWG LRLYRTGYDP IALFTVSRQV

```
               201                                                                                                     300
       81T106  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN KSAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
       81T109  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN KSAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
       81T102  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN KSAPRSVAPA TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
          61C  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
    FY981-c11  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMSPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
      FY981p   STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMSPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA27-C5   STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA27-C7   STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA27-C12  STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
       FA27p   STITPPQAMG PNLVLPDRKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA621-c1  STITPPQAMG PNLVLPDQKP PSRQSQTRSK VATQ.RPQTN VSAPRSVAPT AMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA621-c3  STITPPQAMG PNLVLPDQKP PSRQSQTRSK VATQ.RPQTN VSAPRSVAPT TMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FX343-C8  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
      FZ215p   STITPPQAMG PNLVLPNQKP PSRQSQTGSK VATQMRPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
   SarmaCenv   MTITPPQAMG PNLVLPDQKP PSRQSQTKSK VTTQ.RPQIT SSTPRSVASA TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA27-C6   STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSIAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
FeLV-A/Rickard SAITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RLQTT ESAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
       FGAenv  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
---------------------------------------------------------------------------------------------------------------------------
          61E  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TVGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FZ215-c6  STITPPQAMG PNLVLPDQKP PSRQSQTGSK VATQ.RPQTN ESAPRSVAPT TMGPKRIGTG DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
     FA621-c4  STITPPQAMG PNLVLPDQKP PSRQSQTRSK VATQ.RPQTN VSAPRSVAPT TMGPKRIGTR DRLINLVQGT YLALNATDPN KTKDCWLCLV SRPPYYEGIA
                    ---
                    -2-
```


301
400
81T106 ILGNYSNHTN PPLSCLSIPR HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGAP..G ESQYLAAPNG TYWACNTGLT PCISMAVLNL TSDFCVLIEL
81T109 ILGNYSNHTN PPLSCLSIPR HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGAP..G ESQYLAAPNG TYWACNTGLT PCISMAVLNL TSDFCVLIEL
81T102 IVGNYSNHTY PPLSCLSIPR HKVTISEVSW QGLCIGTVPK THQALCNKTQ QGHTGAP..G ESLYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
61C ILGNYSNQTN PPPSCISIPP HKLTISEVSG QGLCIGTVPK THQALCNKTH QGHTGADYLT APRYLAAPNG TYWACNTGLT PCISMAVLNL TSDFCVLIEL
FY981-c11 ILGNYSNQTN PPPSCQSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FY981p ILGNYSNQTN PPPSCQSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA27-C5 VLGNYSNQTN PPPSCLSDPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPSG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA27-C7 VLGNYSNQTN PPPSCLSDPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPSG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA27-C12 VLGNYSNQTN PPPSCLSDPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPSG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA27p ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA621-c1 ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNATQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA621-c3 ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNATQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FX343-C8 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGPCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FZ215p ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TIDFCVLIEL
SarmaCenv VLGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCKKTQ KGHKGT.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FA27-C6 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FeLV-A/Rickard ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG AYWACNTGLT PCISMAVLNW TSDFCVLIEL
FGAenv ILGNYSNQTN PPPSCLSTPQ HKLTISEVSG QGMCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
61E ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNKTQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL
FZ215-c6 ILGNYSNQTN PPPSCLSIPQ HKLTISEVSG QGLCIGTVPK THQALCNETQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TADFCVLIEL
FA621-c4 ILGNYSNHTN PPPSCLSTPQ HKLTISEVSG QGLCIGTVPK THQALCNATQ QGHTGA.... ..HYLAAPNG TYWACNTGLT PCISMAVLNW TSDFCVLIEL

```
                401                                                                                                 500
      81T106    WPRVTYHQPE YVYTHFAKAG RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LERSLTSLSE VVLQNRRGLD
      81T109    WPRVTYHQPE YVYTHFAKAG RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LERSLTSLSE VVLQNRRGLD
      81T102    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQL*MAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
        61C     WPRVTYHQPE YVYTHFAKAG RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
    FY981-c11   WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FY981p    WPRVTYHQPE YVYTHFAKAV RFRRNRISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA27-C5   WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA27-C7   WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA27-C12  WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
       FA27p    WPRVTYHQPE YVYTHFAEAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA621-c1  WPRVTYHQPE YVYSHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA621-c3  WPRVTYHQPE YVYSHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FX343-C8  WPRVTYHQPE YVYTHFAEAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
       FZ215p   WPRVTYHQPE YVYTHFAKAV RFRREPVSLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
     SarmaCenv  WPRVTYHQPE YIYTHFDKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQIAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA27-C6   WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
  FeLV-A/Rickard WPRVTYHQPE YVVTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FGAenv    WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      61E       WPRVTYHQPE YVYTHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FZ215-c6  WPRVTYHQPE YVYTHFAKAV RFRREPVSLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
      FA621-c4  WPRVTYHQPE YVYSHFAKAV RFRREPISLT VALMLGGLTV GGIAAGVGTG TKALLETAQF RQLQMAMHTD IQALEESISA LEKSLTSLSE VVLQNRRGLD
```

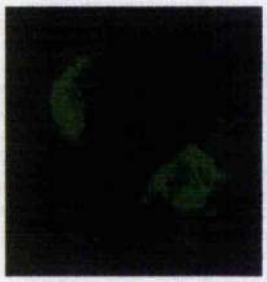
```
                501                                                                                                 600
      81T106    ILFLQEGGLC ATLKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFQGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      81T109    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFQGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      81T102    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
        61C     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
    FY981-c11   ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      FY981p    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA27-C5    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA27-C7    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA27-C12   ILFLQEGGLC AALKEECCFY ADRTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP GFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      FA27p     ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA621-c1   ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA621-c3   ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FX343-C8   ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFIKDRISV
      FZ215p    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG GFEGWFNKSP RFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     SarmaCenv  ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LLGPCILNRL VQFVKDRISV
     FA27-C6    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
  FeLV-A/Rickard ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRELLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      FGAenv    ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
      61E       ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNRSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FZ215-c6   ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
     FA621-c4   ILFLQEGGLC AALKEECCFY ADHTGLVRDN MAKLRERLKQ RQQLFDSQQG WFEGWFNKSP WFTTLISSIM GPLLILLLIL LFGPCILNRL VQFVKDRISV
```

```
                601                 623
      81T106    VQALILTQQY QQIKQYDPDR PD
      81T109    VQALILTQQY QQIKQYDPDR P~
      81T102    VQALILTQQY QQIKQYDPDR P*
        61C     VQALILTQQY QQIKQYDPDR P~
    FY981-c11   VQALILTQQY QQIKQYDPDR P*
      FY981p    VQALILTQQY QQIKQYDPDR P*
     FA27-C5    VQALILTQQY QQIKQYDPDR P*
     FA27-C7    VQALILTQQY QQIRQYDPDR P*
     FA27-C12   VQALILTQQY QHIRQYDPDR P*
      FA27p     VQALILTQQY QQIRQYDPDR P*
     FA621-c1   VQALILTQQY QQIKQYDPDR P*
     FA621-c3   VQALILTQQY QQIKQYDPDR P*
     FX343-C8   VQALILTQQY QQIKQYDPDR P*
      FZ215p    VQALILTQQY QQIKQYDPDR P*
    SarmaCenv   VQALILTQQY QQIQQYDSDR P*
     FA27-C6    VQALILTQQY QQIRQYDPDR P*
  FeLV-A/Rickard VQALILTQQY QQIKQYDPDR P*
      FGAenv    VQALILTQQY QQIKQYDPDR P*
      61E       VQALILTQQY QQIKQYDPDR P~
     FZ215-c6   VQALILTQQY QQIKQYDPDR P*
     FA621-c4   VQALILTQQY QQIKQYDPDR P*
```

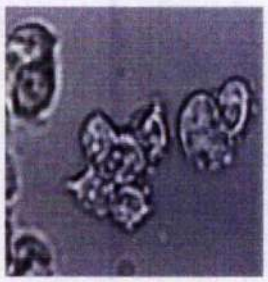
4.3.2 Assaying the production of mature envelope glycoproteins by immunofluorescence

To assess whether the novel *env* gene products were folded correctly, 293T cells were transfected with VR1012 *env* constructs, fixed with ice-cold methanol and stained using 6-15, a monoclonal antibody recognising the FeLV envelope glycoprotein. FeLV-A/Glasgow-1, FeLV-B/Gardner Arnstein, FeLV-C/Sarma and FY981 were included as controls. The results of the antibody staining are shown in figure 4.4. From the eleven novel *env* gene products assessed, nine were capable of producing a protein that was recognised by the monoclonal antibody 6-15 (FA27-C5, FA27-C6, FA27-C7, FA27-C12, FA621-C1, FA621-C3, FA621-C4, FY981-C11 and FZ215-C6). The remaining two constructs (FX343-C8 and FZ215-C7), were incapable of forming protein that was recognised by the antibody.

As described previously in section 3.3.2, cells that were transfected efficiently and were able to produce protein recognised by 6-15 were evident from the localised fluorescence on the surface membrane of the cells, indicating the production and export of the envelope glycoprotein from the cells. The two *env* gene constructs (FZ215-C7 and FX343-C8) that did not produce a protein that was recognised by the antibody were indistinguishable from the background cells, as can be seen in figure 4.4. To ensure the presence of cells in the field examined, a second image was recorded using bright field microscopy, the results of which can be seen in figure 4.4. As there are cells present in the field photographed, the negative antibody staining was due to the *env* constructs producing a protein that was no longer recognised by the antibody. As discussed in chapter three, a likely explanation for this is the disruption of the conformational structure through changes in the amino acid sequence and subsequently, loss of recognition by the antibody.



FA27-C5



FA27-C5

FA27-C6

FA27-C7

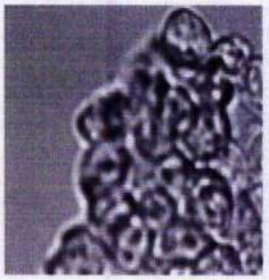
FA27-C12

FA621-C1

FA621-C3

FA621-C4

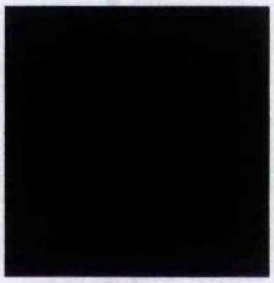
FX343-C8



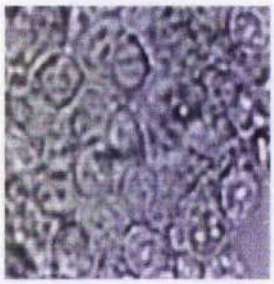
FX343-C8

FY981-C11

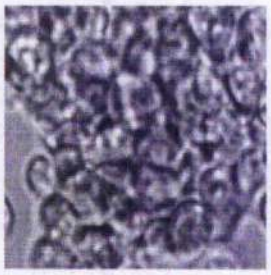
FZ215-C6



FZ215-C7



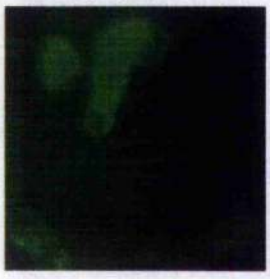
FZ215-C7



FeLV-A/Glasgow-1 control

FY981 control

FeLV-C/Sarma control



FeLV-B/Gardner Arnstein control

Figure 4.4: *Evaluation of the expression of the mature envelope glycoprotein. The 293T cells that were transiently transfected with the FeLV env constructs, a plasmid containing MuLV gag-pol and a marker gene were stained with the anti-gp70 monoclonal antibody 6-15 and examined by immunofluorescence.*

4.3.3 Cell tropism of the FeLV *env*/MuLV *gag pol* pseudotypes

FeLV *env*/MuLV *gag pol* pseudotypes were prepared and plated onto a range of cell lines to determine the tropism. The results of which are shown in table 4.4. As discussed in chapter three, pseudotypes bearing the FeLV-A, FeLV-B and FeLV-C *env* proteins were capable of infecting the cell lines predicted by previous studies, with FeLV-A having the most restrictive *in vitro* tropism and FeLV-C the broadest *in vitro* tropism (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). The titres in this section of the study were generally lower than the titres found in chapter three; this was most likely due to differences in the transfection efficiency of the 293T cells.

FeLV-A infected the feline cell line HO6T1 and the cell line Mink to titres of 10^3 CFU/ml, no other cell lines were susceptible. FeLV-B infected the feline cell line HO6T1 (to a titre of $1.3x10^3$ CFU/ml) and as expected the human cell line HeLa (to a titre of $6.5x10^3$ CFU/ml), the canine cell line MDCK (to a titre of $1.0x10^4$ CFU/ml) and the porcine cell line ST Iowa to a titre of $1.3x10^2$ CFU/ml (porcine cells were shown previously to be permissive to FeLV-B infection but resistant to FeLV-A and FeLV-C infection (Hardy W.G.Jr., 1993)). FeLV-B also infected the cell line Mink, however with a lower titre compared to both FeLV-A and FeLV-C (to a titre of 22 CFU/ml compared to $3.6x10^3$ CFU/ml and $1.7x10^3$ CFU/ml respectively). Similarly, FeLV-C infected the feline cell line HO6T1 (to a titre of $4.0x10^2$ CFU/ml) and as expected the human cell line HeLa (to a titre of $5.4x10^2$ CFU/ml) and the guinea pig cell line 104C1 (to a titre of $1.3x10^3$ CFU/ml), the latter of which is known to be permissive to FeLV-C infection only (Hardy W.G.Jr., 1993). The titres of the FeLV-C control were lower than the titres for the other controls included in the studies, on average the titres were five to ten fold lower than the titres found for other constructs. As expected, none of the constructs produced viruses that were capable of infecting the murine cell line 3T3 (murine cells were shown previously to be resistant to FeLV infection (Jarrett O. et al., 1973)).

The constructs, which were not recognised by the monoclonal antibody 6-15, namely FZ215-C7 and FX343-C8, did not produce pseudotypes capable of infecting the panel of cell lines, suggesting that the proteins formed were non-functional.

	H06T1	104C1	Mink	Hela CB	MDCK	ST Iowa	3T3
FA27-C5	$3.6X10^2$	$6.7X10^2$	$3.1X10^3$	$2.6X10^2$	0	0	0
FA27-C6	$3.6X10^2$	0	$3.2X10^2$	0	0	0	0
FA27-C7	$4.6X10^2$	$1.2\ X10^3$	$3.4X10^3$	$6.2X10^2$	0	0	0
FA27-C12	43	0	$1.6X10^2$	26	0	0	0
FA621-C1	91	60	$6.2X10^2$	17	97	0	0
FA621-C3	$3.6X10^2$	26	$1.2X10^3$	0	0	0	0
FA621-C4	60	0	$1.0X10^2$	0	0	0	0
FY981-C11	$1.5X10^2$	$3.6\ X10^2$	$1.9X10^3$	$1.0X10^2$	0	0	0
FZ215-C6	$1.1X10^2$	0	9	0	0	17	0
FZ215-C7	0	0	0	0	0	0	0
FX343-C8	0	0	0	0	0	0	0
CL29A	$6.8x10^3$	0	$3.6x10^3$	0	0	0	0
FeLV-B/GA	$1.3x10^3$	0	22	$6.5x10^3$	$1.0x10^4$	$1.3x10^2$	0
Sarma C	$4.0x10^2$	$1.3x10^3$	$1.7x10^3$	$5.4x10^2$	0	0	0
FY981	$8.9x10^3$	$2.1x10^4$	$1.8x10^4$	$6.8x10^3$	0	$3.8x10^2$	0

Table 4.4: *Titres (CFU/ml) of the novel unique env clones established by infection of a panel of cell lines. The cell lines tested are: the feline kidney cell line HO6T1, the guinea pig cell line 1O4C1, the mink lung cell line Mv1Lu (Mink), the human cervical carcinoma cell line Hela CB, the canine kidney cell line MDCK, the porcine testes cell line ST IOWA and the murine cell line NIH 3T3.*

The *env* constructs most closely resembling FeLV-A and capable of producing viable protein, FA27-C6, FA621-C4 and FZ215-C6, infected the same panel of cell lines as the FeLV-A control, albeit to lower titres than the control (to titres of $3.6x10^2$ CFU/ml, 60 CFU/ml and 1.1 $x10^2$ CFU/ml respectively for HO6T1 infection and to titres of $3.2x10^2$ CFU/ml, $1.0x10^2$ CFU/ml and 9 CFU/ml respectively for mink infection). Construct FZ215-C6 also infected the porcine cell line ST with a low titre (17 CFU/ml), suggesting that the tropism of subgroup A viruses may not be as restricted as suggested by earlier studies on viral growth *in vitro* (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). The *env* constructs that were classified as C components through sequencing analysis and that were capable of producing viable protein (FA27-C5, FA27-C7, FA27-C12, FA621-C1, FA621-C3 and FY981-C11) infected the same cell lines as the control construct FeLV-C/Sarma (to titres of $3.6x10^2$ CFU/ml, $4.6x10^2$ CFU/ml, 43 CFU/ml, 91 CFU/ml, $3.6x10^2$ CFU/ml and $1.5x10^2$ CFU/ml respectively for HO6T1 infection and to titres of $3.1x10^3$ CFU/ml, $3.4x10^3$ CFU/ml, 26 CFU/ml, 17 CFU/ml, 0 CFU/ml and $1.0x10^2$ CFU/ml respectively for mink infection). With the exception of construct FA27-C12, titres were comparable to the titres obtained with the control construct FeLV-C/Sarma. Construct FA27-C12 had an unusual tropism as it was capable of infecting both HO6T1 and mink cells (to a titre of 43 CFU/ml and 26 CFU/ml respectively), however it was not capable of infecting the guinea pig cell line 104C1, a determining factor for the C-phenotype. This could indicate the presence of an alternative receptor on the cell surface, which is used by FeLV-C/Sarma. Alternatively, the receptor present on 104C1 cells is slightly altered and thus prevents infection by FA27-C12. Interestingly, construct FY981-C11 was incapable of infecting the pig cell line ST Iowa in contrast to the parent control construct FY981 and the previously characterised construct FY981-14. This could be due to the fact that the overall titres of this construct are lower than the titres found for the previously identified construct FY981-14 and therefore infection of ST Iowa cells could not be detected. Construct FA621-C1 had an extended *in vitro* tropism; as it was also capable of infecting the canine cell line MDCK (to a titre of 97 CFU/ml), in contrast to the subgroup C control FeLV-C/Sarma.

As was noted in chapter three, few constructs yielded viruses that were capable of infecting the human cell line HeLa, although as discussed in section 3.4, this could be due to the choice of human cell line tested rather than an inability of the *env* constructs tested in this study (previous studies used 293T cells).

Virus isolate	Original mixture of subgroups	*Env* clone #	Subgroup predicted by sequencing	Antibody staining using 6-15 monoclonal antibody	Subgroup established by virus infection
FA27	A/C	C5	C	+	C
		C6	A	+	A
		C7	C	+	C
		C12	C	+	C?
FA621	A/C	C1	C	+	C
		C3	C	+	C
		C4	A	+	A
FY981	A/B/C	C11	C	+	C
FX343	A/C	C8	C	-	-
FZ215	A/B/C	C6	A	+	B
		C7	B	-	-

Table 4.5: *Subgroups of the novel envelope glycoproteins as determined by sequencing and pseudotype infection on a panel of cell lines from a range of species.*

4.4 Discussion

To date no single amino acid or stretch of amino acids has been identified that can change the subgroup of an isolate of FeLV. Therefore, we aimed to isolate and characterise the C-components of a panel of natural isolates in order to identify the regions involved in defining the phenotype of these isolates. When the primary isolates described previously in chapter three (see table 3.1), were plated onto FEA cells, a non-restrictive cell line, a number of C components could not be isolated, most likely due to the relative scarcity of the C phenotype in the viral pool of a specific primary isolate. Therefore, the panel of primary isolates was plated onto the guinea pig cell line 104C1, which is restrictive to FeLV-C infection alone. Thus, it was anticipated that the remaining C-components of the natural isolates would be enriched, enabling their isolation and characterisation. After biological selection on 104C1 cells, twenty-two novel FeLV *env* clones were identified. Using the criteria described in section 3.3.1 (the presence of length polymorphisms and a high degree of divergence in the VRA region relative to the published FeLV-A/Glasgow-1 envelope glycoprotein sequence), the *env* clones were classified into either the A component or the C component of an isolate. Thus, four clones appeared to be the A component of the primary isolate and sixteen clones appeared to be the C component. As discussed in chapter three, as the envelope glycoprotein of FeLV-B/Gardner Arnstein was highly divergent from both FeLV-A and FeLV-C, it was relatively easy to establish that the remaining two clones were the B-component of that specific primary isolate.

4.4.1. The sequence of the novel FeLV-C envelope glycoproteins

4.4.1.1. Envelope glycoproteins described in this study are divergent from published envelope glycoprotein sequences

Although previous studies by Brojatsch et al (Brojatsch J. et al., 1992) and Rigby et al (Rigby M.A. et al., 1992) into the amino acids determinants of in vitro cell tropism utilised the same isolates of FeLV, the sequences described in the two studies contained significant discrepancies.

Thus, a representative clone of each primary isolate was characterised and an amino acid alignment was prepared (see figure 3.2). When the novel envelope glycoproteins characterised in this thesis were compared with the published FeLV envelope glycoproteins, the data confirmed that the C component of a primary isolate was represented by more than one sequence, and that the sequences described previously may have represented single variants and may not have been representative of the dominant phenotype within the pool.

4.4.1.2. Each primary isolate is represented by multiple envelope glycoprotein sequences

As discussed previously, when the FeLV env A and C sequences identified in this study were compared, the majority of the envelope glycoprotein sequence was conserved and each independent isolate appeared to contain characteristic differences throughout the sequence that were unique to that specific primary isolate. The majority of the differences were present in the variable regions of the envelope glycoprotein, but further differences were scattered throughout the entire envelope glycoprotein. The differences discussed in more detail in chapter three were also present in the env clones described in this section of the thesis. For example, the A component of primary isolate FA27, FA27-C6, was identical to the A component described in chapter three, FA27-53, thus again strengthening the conclusion that the env sequences of both clone FA27-C6 and clone FA27-53 were indeed representative of the A component of primary isolate FA27. Similarly, the C component of primary isolate FY981, FY981-C11, was identical to env clone FY981-14, with the exception of one amino acid (N_{21} found in FY981-C11 compared to T_{21} found in FY981-14), which may not alter the overall structure. When comparing the C components identified for primary isolate FA621 (FA621-C1 and FA621-C3), they were very similar, indeed there were only three amino acid differences present (I_9V, $T_{33}A$ and $A_{261}T$ in clone FA621-C3 relative to clone FA621-C1). Of these differences, the arginine residue on position 251 was present only in clone FA621-C1, compared to a threonine residue in both FA621-C3 and FA621-C4 (the A component of primary isolate FA621).

4.4.1.3. The abundance of the C-component within each primary isolate pool

When the panel of primary isolates was plated onto FEA cells (a non-restrictive cell line), it became apparent that the abundance of the C component in the viral pool varied for each primary isolate. For example, the C component of primary isolate FY981 (previously classified as an A/B/C mixture) was relatively abundant, as all seven env clones identified appeared to be the C component, whereas the C component of primary isolate FZ215 (previously classified as an A/B/C mixture) was not detected. After plating the panel of primary isolates onto 104C1 cells (a cell line restrictive to FeLV-C infection alone) a number of the C components that had not yet been identified, were isolated. For example, after plating primary isolate FA621 onto 104C1 cells, six env clones were isolated and five appeared to be the C component. As discussed in chapter three, after plating this primary isolate onto FEA cells, the C component was not identified. After selection onto 104C1 cells of a number of primary isolates, the number of clones that appeared to be the C component was more abundant than after plating onto FEA cells. For example, primary isolate FA27 yielded ten novel env clones, nine of which appeared to be the C component, thus these results were in contrast with the results discussed in chapter three, where only two out of four clones appeared to be the C component. In conclusion, the C component was isolated from the majority of the primary isolates after plating onto the restricted cell line 104C1. However, to date the C component of two of the panel of primary isolates (FZ215 and L3728F) remains to be isolated, this may be due to the relative scarcity of the C phenotype in the viral pool of the specific primary isolate.

4.4.2. Characteristics of the C phenotype of FeLV

The main features and the clustering of mutations within the envelope glycoprotein described in Chapter 3 were also found in the env clones described in this Chapter. Indeed, the findings in Chapter 4 further strengthen the findings described in Chapter 3. The main features that ascertain a C-phenotype (mutations in a series of amino acids consisting of ^{61}TNVKHG^{66} resulting in the loss of a N-linked glycosylation site and the substitution $V_{63}W$) were present in all isolates that were able to produce pseudotypes capable of infecting 104C1 cells.

These findings strengthen the hypothesis that the while the mechanism by which mutations arise in the viral envelope glycoprotein may be random; the mutations are retained in areas of significance to protein structure (for example the receptor binding site) or biological function (evasion of the host immune response).

4.4.3. The envelope glycoprotein sequences of the A components

Of the twenty-two clones identified, four novel env clones were classified as the A component of the primary isolate, using the criteria described previously (lack of both variation and length polymorphisms in the VRA region of the envelope glycoprotein). When the A components identified in this study were compared, the envelope glycoprotein amino acid sequences were not identical, indeed there were differences present throughout the envelope glycoprotein sequence. For example, when comparing both the A components and the C components identified for primary isolate FA27, a $D_{51}N$ substitution and a $N_{59}D$ substitution (relative to the FeLV-A/Glasgow-1 sequence) in the VRA region were present in all the clones identified to date. Thus, these results strengthen the findings discussed in Chapter 3 that, although FeLV-A amino acid sequences were very similar, there were amino acids present that were specific for a particular primary isolate, rendering these changes unique to that particular primary isolate. Given that these subtle differences in amino acid sequence fall within the predicted receptor binding domain, these observations raise the possibility that variants of FeLV-A may exist that have distinct cell tropisms.

4.4.4. Effects of the *env* sequence on antibody staining and biological function

In order to establish the function and the host range of eleven of the unique envelope glycoproteins isolated, FeLV/MuLV pseudotypes were generated and their cell tropism investigated in vitro (see figure 4.4 and table 4.4). Further, recognition of the viral env protein by the anti-gp70 antibody 6-15 was examined by immunofluorescence (see figure 4.3).

4.4.4.1. The production of mature envelope glycoprotein *in vitro*

Of the eleven *env* constructs characterised further, two constructs (FZ215-C7 and FX343-C8) failed to react with the anti-gp70 monoclonal antibody 6-15, monoclonal antibody 6-15, suggesting that the proteins produced by these env clones were no longer recognised by the antibody. As discussed in chapter three, the binding site of the monoclonal antibody 6-15 has not yet been characterised, however as the antibody recognises all three subgroups, it is likely that the antibody binding site is located in a highly conserved region of the envelope glycoprotein. As described in chapter three in more detail, there were two possible explanations for the production of a protein that was no longer recognised by this antibody (firstly, the conformational structure of the protein was disrupted and thus the antibody epitope was no longer available to the protein and secondly, the mutations were present in the antibody epitope and thus antibody binding was disrupted). As was found for the four env constructs described in Chapter 3 that failed to bind the anti-gp70 antibody 6-15 (FS246-4, FX343-26, FY981-10 and FZ215-12), there were a number of changes in the amino acid sequence that were not present in the env clones identified from the same primary isolate that were recognised by the monoclonal antibody. The mutations for each clone were as follows:

a) FX343-C8: $T_{109}A$ and $G_{260}R$
b) FZ215-C7: $T_{23}A$ and $T_{165}I$

As was discussed previously (see Chapter three), the mutations appeared to fall in distinct regions of the envelope glycoprotein, thus strengthening the findings discussed in chapter three that the mutations in the envelope glycoprotein altered the conformational structure of the protein and therefore disrupted the antibody epitope which lead to the protein failing to react with the anti-gp70 monoclonal antibody 6-15.

4.4.4.2. Altered cell tropism of novel FeLV env clones

In order to establish whether the novel env clones were functional, FeLV/MuLV pseudotypes were generated and the in vitro cell tropism was examined. Previous studies have examined the in vitro cell tropism of FeLV-A, FeLV-B and FeLV-C and have shown that FeLV-A has the most restricted in vitro cell tropism and is capable of infecting feline cells alone (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993). In contrast, FeLV-B and FeLV-C have a broader in vitro cell tropism and are capable of infecting both feline and non-feline cells. Guinea pig cells alone support infection of FeLV-C, whereas porcine cells alone support FeLV-B infection (Jarrett O. et al., 1973; Sarma P.S. et al., 1975; Hardy W.G.Jr., 1993).

Similar to the four env clones that failed to react with the anti gp70 monoclonal antibody 6-15 described in chapter three (FS246-4, FX343-26, FY981-10 and FZ215-12), the two constructs described in this part of the study that failed to produce protein recognised by the 6-15 antibody (FX343-C8 and FZ215-C7) did not produce pseudotypes capable of infecting the cell lines tested, thus suggesting that the proteins produced were indeed non-functional. Pseudotypes derived from the remaining nine viable env constructs (FA27-C5, FA27-C6, FA27-C7, FA27-C12, FA621-C1, FA621-C3, FA621-C4, FY981-C11 and FZ215-C6) were capable of infecting the panel of cell lines tested.

Pseudotypes derived from the three env constructs that appeared to be the A components (FA27-C6, FA621-C4 and FZ215-C6) infected the same range of cells as the control FeLV-A/Glasgow-1, namely the feline cell line HO6T1 and mink cells, although the titres found for the novel env constructs were lower than the titres found for the control FeLV-A/Glasgow-1 construct ($9.8x10^2$ CFU/ml for FeLV-A, compared to $3.6x10^2$ CFU/ml for FA27-C6, 60 CFU/ml for FA621-C4 and $1.1x10^2$ CFU/ml for FZ215-C6). Interestingly, FZ215-C6 derived pseudotypes were capable also of infecting the porcine cell line ST Iowa, in contrast to pseudotypes derived from env clone FZ215-9 (the A component of primary isolate FZ215 described in chapter three), which were incapable of infecting ST Iowa cells. When the amino acid sequences of FZ215-C6 and FZ215-9 were compared, there were two amino acid changes present in FZ215-C6, but not in FZ215-9 (N_3S and $N_{22}H$).

N_3S is present also in the amino acid sequence of FZ215p, however pseudotypes derived from this isolate were not plated onto the porcine cell line ST Iowa. The second substitution, $N_{22}H$, is unique and can be found only in FZ215-C6. The amino acid residue (aspartic acid on position 51) associated with the extended host range of the pseudotypes derived from env constructs FA27-53, FY981-14, FeLV-B/Gardner Arnstein and FY981p (i.e. capable of infecting porcine cells, see 4.4.3) was not present in FZ215-C6. Thus, this aspartic acid residue on position 51 cannot be the sole residue important in the extended *in vitro* cell tropism of the three novel *env* constructs (FA27-53, FY981-14 and FZ215-C6) to include porcine ST Iowa cells.

With the exception of env construct FA27-C12, pseudotypes derived from the *env* constructs that appeared to be the C components (FA27-C5, FA27-C7, FA27-C12, FA621-C1, FA621-C3 and FY981-C11) were capable of infecting the feline cell line HO6T1, the guinea pig cell line 104C1 and mink cells to comparable titres as those derived from the control FeLV-C/Sarma ($4.6x10^2$ CFU/ml for the control FeLV-C compared to $3.6x10^2$ CFU/ml for FA27-C5, $4.6x10^2$ CFU/ml for FA27-C7, 91 CFU/ml for FA621-C1, $3.6x10^2$ CFU/ml for FA621-C3 and $1.5x10^2$ CFU/ml for FY981-C11 when infecting HO6T1 cells). The inability of FA27-C12 to infect 104C1 cells was possibly due to the diminished titres of this construct. In comparison, the control FeLV-C/Sarma pseudotypes were able to infect HO6T1 cells to a titre of 463 CFU/ml, whereas FA27-C12 pseudotypes had a titre of 43 CFU/ml when infecting HO6T1 cells, resulting in a ten-fold reduction of infectivity. With the exception of FA621-C3, all pseudotypes were capable also of infecting HeLa cells, similar to the control FeLV-C/Sarma derived pseudotypes. FA621-C3 derived pseudotypes were incapable of infecting HeLa cells, even though FA621-C3 derived pseudotypes were able to infect HO6T1 cells to comparable titres as those derived from the control FeLV-C (360 CFU/ml in comparison to 463 CFU/ml for the control FeLV-C/Sarma construct). Interestingly, although the titres of FA621-C3 pseudotypes were four fold higher than the titres found for FA621-C1 pseudotypes (which also appeared to be the C component of primary isolate FA621), FA621-C1 derived pseudotypes were capable of infecting HeLa cells. When the amino acid sequences of the two C components were compared, there were 3 amino acid differences between clones C1 and C3, V_9I, $T_{33}A$ and $T_{251}A$.

Two amino acid differences (V_9I and $T_{251}A$) were present in regions of the envelope glycoprotein thought to be associated with cell tropism (determined amongst others by Bae et al (Bae Y. et al., 1997)). As the first substitution (V_9I) was also present in the published FeLV-C/Sarma amino acid sequence and FeLV-C/Sarma derived pseudotypes are incapable of infecting ST Iowa cells, it was unlikely that this residue was involved in altering the in vitro cell tropism of FA621-C3. $T_{33}A$ lies in the region immediately upstream of the VRA and was also found in FA621-C4, an env clone that behaved as subgroup A. The third substitution ($T_{251}A$) is unique to clone C1. The switch from a hydrophilic side chain (threonine) to a hydrophobic side chain (alanine), may be sufficient to alter the conformation of the FA621-C1 env such that it has a broader tropism than FA621-C3, or conversely, may be responsible for the narrower cell tropism of FA621-C1.

The *env* construct most closely resembling FeLV-B/Gardner Arnstein (FZ215-C7) did not produce protein recognised by the monoclonal antibody 6-15, nor was it capable of producing functional protein as was shown by the lack of infection by FZ215-C7 Env-bearing pseudotypes, these findings were in accordance with the novel FZ215 env clone resembling FeLV-B/Gardner Arnstein described in chapter three (FZ215-12). Thus both FZ215-C7 and C12 produced non-functional Env proteins.

CHAPTER FIVE

EXPRESSION OF THE FeLV-C RECEPTOR feFLVCR *IN VIVO*

5.1 Introduction

In 1999, Tailor et al (Tailor C.S. et al., 1999c) described the cloning and characterisation of a cellular receptor for FeLV-C from a cDNA library prepared from a human cell line. Subsequently, the feline homologue of this molecule was cloned and characterised (Quigley J.G. et al., 2000). Both molecules are members of the major facilitator superfamily (MFS) (discussed in detail by Pao (Pao S.S. et al., 1998)) and are multi-transmembrane domain molecules that are thought to function as transporters (see also section 1.5.2). By screening a polyA (+) mRNA northern blot consisting of RNAs from both haematopoietic and non-haematopoietic tissues, Tailor et al (Tailor C.S. et al., 1999c) demonstrated that the huFLVCR transcript was preferentially expressed in haematopoietic tissues, although low levels of the mRNA transcript were also present in non-haematopoietic cells. These findings led to the conclusion that FeLV-C preferentially infected haematopoietic cells *in vivo*, in agreement with previous studies by Dean et al (Dean G.A. et al., 1992) into the preferential haematopoietic target cells of FeLV-C infection. Dean et al (Dean G.A. et al., 1992) described three potential mechanisms by which the FeLV-C Env protein could cause an erythroid cell-specific cytopathic effect and thus pure red cell aplasia PRCA:

1) The FeLV-C receptor is expressed solely on erythroid progenitor cells and thus these cells are targeted selectively by the virus.
2) The binding of FeLV-C to its receptor on the cell surface impairs the normal function of its receptor and cells of the erythroid lineage have an absolute dependence on the function of this molecule for survival.
3) The binding of the virus to its receptor has an effect through a mechanism that, to date has not been characterised.

Dean et al showed that viral Env protein gp70 and capsid protein p27 could be detected in bone marrow populations containing erythroid, but also lymphoid and myeloid cells at similar levels when samples taken from cats infected either with FeLV-A or FeLV-C were compared, suggesting that the FeLV-C receptor was not expressed solely on erythroid progenitor cells.

Further, these findings were strengthened by Tailor et al (Tailor C.S. et al., 1999c) when the preferential expression of huFLVCR on haematopoietic cells was detected, however the huFLVCR expression was not restricted exclusively to erythroid progenitor cells. To date, the impairment of the normal function of the receptor due to the virus binding, has not been investigated. Dean et al concluded that the most plausible explanation for the devastating effect that FeLV-C infection has on erythropoeisis, is inhibition of the normal function of the receptor in the production of erythrocytes. This conclusion is in agreement with the results described by Abkowitz et al (Abkowitz J.L., 1991). Abkowitz et al hypothesised that the FeLV-C/Sarma envelope glycoprotein interfered with the interaction of transferrin and its receptor, suggesting that either the transferrin receptor is also the FeLV-C/Sarma receptor or that gp70 produced intracellularly interfered with the recycling of the transferrin receptor and this subsequently led to a defect in the uptake of transferrin by its receptor.

The main aims of the studies described in this chapter were:

1) To examine the expression of the feline FeLV-C receptor in both haematopoietic and non-haematopoietic tissues.
2) To study the expression of the receptor in bone marrow populations that were either enriched with or depleted of cells involved in the erythroid maturation.

5.2 Methods

The main haematological symptoms of PRCA are a reduced number of late erythroid precursor cells (CFU-E) in spite of a normal number of early erythroid precursor cells (BFU-E), suggesting a defect in the maturation of BFU-E into CFU-E (see figure 1.6 and (Abkowitz J.L et al., 1985)). In order to investigate the role of the FeLV-C receptor in the development of the PRCA associated with FeLV-C infection, a range of both lymphoid (bone marrow, thymus, spleen, mesenteric lymph node MLN, peripheral lymph node PLN) and non-lymphoid (skeletal muscle, saliva gland, heart, brain, liver, kidney) tissues were collected from a number of cats at post mortem and used for RNA preparation and subsequent northern blotting and RT-PCR analysis. In addition, bone marrow cells were fractionated into erythroid and non-erythroid lineages using the monoclonal antibody K1, an antibody thought to recognise erythroid progenitors (Gengozian N., 1998). RNA was prepared from each of the bone marrow fractions and the abundance of the FLVCR RNA transcript quantified by RT PCR.

The tissues were collected from a range of cats at the time of post mortem. The cats were involved in projects underway in the Department of Veterinary Pathology. In total, bone marrow was collected from six cats, four of which were FeLV-negative as assessed by p27 ELISA at the time of post mortem, while two cats were positive for FeLV as assessed by p27 ELISA at the time of post mortem. In addition, a full range of tissues was collected from six cats, three of which were FeLV positive at the time of post mortem while the remaining three were FeLV negative. The tissues that were collected from each cat are summarised in table 5.1.

Cats F4 and F8 were cats from a FIV vaccination trial. These cats were FIV-negative as assessed by p27 ELISA at the time of post mortem and had never been in contact with FeLV. Both bone marrow and a panel of tissues were collected from cat F4, whereas only bone marrow was obtained from cat F8. The tissues were snap frozen in liquid nitrogen and stored at -70°C until required.

	F4	F8	E2	E5	E6	L22	L23	L32	L36
Bone marrow	✓	✓	×	×	×	✓	✓	✓	✓
Kidney	✓	×	✓	✓	✓	×	×	✓	✓
Heart	✓	×	✓	✓	✓	×	×	✓	✓
Saliva gland	✓	×	✓	✓	✓	×	×	✓	✓
Liver	✓	×	✓	✓	✓	×	×	✓	✓
Thymus	✓	×	✓	✓	✓	×	×	✓	✓
Brain	✓	×	×	×	×	×	×	×	×
Spleen	✓	×	✓	✓	✓	×	×	✓	✓
Skeletal muscle	✓	×	✓	✓	✓	×	×	✓	✓
MLN	✓	×	✓	✓	✓	×	×	✓	✓
PLN	×	×	✓	✓	✓	×	×	✓	✓

Table 5.1: *Tissues collected during post mortem examinations from each of the cats. The cat numbers shown in red were positive for FeLV p27 at the time of post mortem, whereas the cats shown in green were negative for FeLV.*

Cats E2, E5 and E6 were cats from an immunopathogenesis study. All three cats were FeLV positive at the time of post mortem, with varying amounts of provirus detected using real time PCR and FeLV p27 ELISA (data not shown). Tissues from this group of cats were collected, snap frozen in liquid nitrogen and stored at -70°C until required.

Cats L22, L23, L32 and L36 were cats from a FeLV vaccination trial. Tissues and bone marrow were collected from two groups of cats; cats L22 and L23 were FeLV-positive after FeLV infection, whereas cats L32 and L36 remained FeLV-negative after FeLV infection (data not shown). As described previously, the tissues were snap frozen in liquid nitrogen and stored at -70°C until required.

5.2.1 Preparation of RNA from tissues

Care was taken to prevent RNA degradation whilst working with tissues and RNA samples. In order to do this, all work surfaces were cleaned with RNase ZAP® (Invitrogen NV, Leek, The Netherlands) to destroy RNAse, all pipettes, filtered tips and reagents were used for RNA work only and isolated from common use equipment to minimise the risk of RNAse contamination. The samples were maintained on ice at all times and long delays were avoided to keep the risk of degradation to a minimum. All centrifugations were carried out in a refrigerated bench top centrifuge at 4°C.

5.2.1.1 Preparation of RNA from tissues using RNAzol™ B

A range of tissues was collected from cats during post mortem. The following tissues were snap-frozen in liquid nitrogen and stored at -70°C until required: thymus, spleen, mesenteric lymph node (MLN), peripheral lymph node (PLN), brain, heart, saliva gland, skeletal muscle, liver and kidney. To prevent RNA degradation, tissues were stored at -70°C until required and maintained on dry ice when removed from storage. The tissues were pulverised using a mortar and pestle filled with liquid nitrogen. The pulverised tissues were then resuspended in 2ml RNAzol™ B (Biogenesis, Dorset, United Kingdom). RNAzol™ B contains phenol and guanidine thiocyanate and is used for the single step method of RNA isolation described in 1987 by Chomczynski and Sacchi (Chomczynski P. and Sacchi N., 1987). In order to separate the phases, 0.2ml chloroform was added per 2ml RNAzol™ B. This was shaken vigorously for a few seconds, incubated on ice for 15 minutes and spun at 3,000rpm for 15 minutes. The upper aqueous layer containing the RNA was transferred to a fresh tube and precipitated with an equal volume isopropanol, incubated on ice for 15 minutes and spun at 3,000rpm for 15 minutes. The RNA pellet was transferred to a 1.5ml tube and washed once with 70% ethanol by centrifugation for 8 minutes at 13,000rpm in a microcentrifuge. The pellet was then dried in a vacuum drier and resuspended in 50µl dH_2O. The RNA concentration was determined using spectrophotometry as described previously (see section 2.2.3.3).

The $A_{260/280}$ ratio was typically 1.9, confirming the purity of the RNA and that there was no protein contamination. The RNA was stored in 20µg aliquots at -70°C until required.

5.2.1.2 Preparation of RNA from tissues using RNA-Bee™

During the course of these studies, the RNAzol™ B reagent was discontinued and thus an alternative product was used. "RNA-Bee™" (Biogenesis, Dorset, United Kingdom) was used to isolate RNA from tissues collected from cat L22 onwards. RNA-Bee™ is similar to RNAzol™ B and therefore follows a similar method for the isolation of RNA and is derived from the single step method described previously by Chomczynski and Sacchi (Chomczynski P. and Sacchi N., 1987). Similar precautions were undertaken when using RNA-Bee™ for the preparation of RNA in order to prevent RNA degradation. Following pulverisation of the tissue, 1ml RNA-Bee™ was added per 50mg pulverised tissue to disrupt the cells and then 0.2ml of chloroform was added per 1ml RNA-Bee™. The mixture was shaken vigorously for approximately 15 seconds and incubated on ice for 5 minutes. The phases were then separated by centrifugation at 3,000rpm for 15 minutes and the upper aqueous layer was transferred to a new centrifuge tube. RNA was precipitated by adding 0.5ml isopropanol and incubating at room temperature for 10 minutes followed by centrifugation at 3,000rpm for 5 minutes. The RNA pellet was washed once with 1ml 70% ethanol per 1ml RNA-Bee™ by centrifugation at 13,000rpm for 5 minutes. The pellet was then air-dried briefly and resuspended in 50µl dH_2O and incubated at 56°C for 10 minutes to aid the solubilisation of the RNA. The final RNA concentration was determined and samples were stored as described in section 2.2.3.3.

5.2.2 Selection of enriched and depleted bone marrow populations using miniMACS columns

In order to investigate whether the FLVCR RNA transcript was differentially expressed in erythroid progenitors, bone marrow was extracted from the femurs of cats at post-mortem and fractionated using the monoclonal antibody K1 (an antibody raised against erythroid progenitors (Gengozian N., 1998)). Gengozian demonstrated that following depletion of bone marrow cells using the K1 antibody, the remaining population showed an 80% reduction in BFU-E progenitor cells. As previous studies have suggested that a defect in the development of CFU-E from BFU-E is thought to underlie the development of PRCA, differential expression of the FLVCR receptor may explain the specific targeting of BFU-E by FeLV-C (Abkowitz J.L et al., 1985). Accordingly, if the expression of the FeLV-C receptor was restricted to cells involved in erythroid maturation, then feFLVCR RNA would be reduced in the K1 depleted population. Conversely, if the expression of feFLVCR was not restricted to cells involved in the erythroid maturation then the feFLVCR RNA would be expressed in both depleted and enriched populations of bone marrow.

5.2.2.1 Extraction of bone marrow from the feline femurs

Femurs were obtained from cats at post-mortem. The ends were removed from the femurs using bone cutters and the bone marrow collected by flushing RPMI medium through the bone with a 20ml syringe and an 18 gauge needle. Once all bone marrow was removed, a single cell solution was obtained by repeatedly pipetting the solution with a 10ml pipette and pelleted by centrifugation at 3000rpm for 10 minutes. Cell pellets were washed twice with RPMI and pelleted by centrifugation at 3000rpm for 10 minutes. The cell pellet was resuspended in 500µl MACS buffer and the cell concentration was determined by counting the cells using a haemocytometer as described previously (section 2.2.1.2.2). $2x10^8$ cells were used per antibody for selection. The cells were pelleted, labelled with 1ml unconjugated primary antibody (as culture supernatant).

The cells were washed once in MACS buffer by centrifugation and labelled with 40µl goat anti-mouse IgG microbeads (Miltenyi Biotec Ltd., Surrey, United Kingdom) to a final 1:5 dilution and incubated at 4°C for 15 minutes. During this incubation period, miniMACS MS columns (Miltenyi Biotec Ltd., Surrey, United Kingdom) were placed in a miniMACS magnetic holder and rehydrated with 500µl MACS buffer. The column matrix consists of ferromagnetic spheres, allowing separation of cells less than 30µM in size. After a subsequent wash step, the cells were resuspended in 500µl MACS buffer and applied to the MS columns. The columns were washed twice with 500µl MACS buffer and all eluates were collected as the "antibody-depleted" population. A further 500µl MACS buffer was added to the columns, after which the columns were removed from the magnetic holder. The K1-positive population was eluted by expelling the MACS buffer from the columns using the plungers supplied. Both populations were counted as previously (section 2.2.1.2.2) to determine the concentration of these populations after which the cells were pelleted, snap frozen and stored at -70°C until required. An unselected bone marrow sample that had been neither enriched nor depleted using the antibody K1 was included in the assay.

5.2.2.2 Preparation of RNA from cells using either RNAzol™ B or RNA-Bee™

As described previously, great care was taken to prevent RNA degradation. To lyse tissue culture cells, 0.2ml RNAzol™ B or RNA-Bee™ was added per 10^6 cells. The RNA was separated and precipitated as described previously (section 5.2.1.1 and section 5.2.1.2 respectively).

5.2.3 RNA gel electrophoresis and transfer onto a nitrocellulose membrane (northern blotting)

The RNA aliquots (containing 20µg RNA) were thawed rapidly in a 60°C water bath and vacuum dried. The samples were then resuspended in 20µl loading buffer and incubated at 65°C for 15 minutes. A 1% agarose gel containing 1x MOPS and 2.1M formaldehyde was prepared and poured in a laminar flow hood.

After incubation, 5µl RNA dye was added and the samples were loaded onto the gel. A 0.24-9.50 Kb RNA ladder was included to determine the size of the RNA bands (Invitrogen ™ Life technologies (Paisley, United Kingdom)). The samples were run for approximately 10 minutes at 100 volts before a circulation pump was attached and the samples were run for a further 3 hours at 100 volts and then stained in dH_2O containing ethidium bromide (1.1 mg/l (w/v) ethidium bromide) for 10 minutes. The gel was destained for 30 minutes in dH_2O twice and then visualised on a UV transilluminator and photographed (section 2.2.3.5). The lanes containing the RNA ladder remained in fresh dH_2O for destaining overnight and were photographed the following morning. The lanes containing the RNA samples were rinsed for 30 minutes in 10x SSC and a northern transfer was set up as described previously (Current Protocols in Molecular Biology, (Ausubel F.A. et al., 1992). The following day, RNA was immobilised using a spectrolinker XL-1500 UV-crosslinker (Spectronics Corporation, New York, USA) using the "optimal crosslink" setting (1200 Watt). The blot was then stored at room temperature until required.

5.2.4 Hybridisation of northern blots

5.2.4.1 Hybridisation of northern blots using "RapidHyb™" hybridisation buffer

The northern blot was hybridised using "RapidHyb™" buffer (Amersham Biosciences Europe GmbH, Freiburg, Germany). The buffer was pre-warmed to 65°C before adding to the blot. The blot was prehybridised at 65°C for a minimum of 15 minutes, during which the [α-^{32}P]-dCTP probe (the preparation of which is described in section 2.2.3.9) and the salmon sperm DNA (ssDNA, to a final concentration of 100 µg/ml) were denatured at 100°C for 10 minutes and subsequently chilled on ice. The ssDNA was added to reduce the background and was added at a 1/1000th volume of the RapidHyb™ buffer. The [α-^{32}P]-dCTP probe was added to a final concentration of $1x10^7$ cpm/ml RapidHyb™ buffer. The blot was allowed to hybridise with shaking for 2 hours after which it was washed with the following stringency washes: 2x SSC rinse, 0.5%SDS/0.5xSSC rinse, 0.5%SDS/0.5xSSC 20 minutes, 0.5%SDS/0.5xSSC 20 minutes.

The final wash for the blot was adjusted, depending on the strength of the probe and the abundance of the gene of interest. The blot probed with the feFLVCR probe was washed using 0.5%SDS/0.5xSSC 20 minutes, whereas the blot probed with the GAPDH probe was washed with 0.5%SDS/0.1xSSC for 20 minutes. The washed blot was sealed in a bag and was then exposed to film (Hyperfilm MP, supplied by Amersham Biosciences UK Limited, United Kingdom) at -70°C using two intensifying screens. The film was developed in an automated processor (Automated Processor: Compact X4 supplied by Xograph, Wiltshire, UK.) following the manufacturer's instructions.

5.2.4.2 Hybridisation of northern blot using ULTRAhyb™ hybridisation buffer

As the feFLVCR gene appeared to be expressed at low levels, a second northern blot was hybridised using ULTRAhyb™ hybridisation buffer (Ambion Europe Ltd, Cambridgeshire, United Kingdom). The buffer was pre-warmed to 68°C before the blot was immersed in sufficient buffer to uniformly wet the blot. The blot was subsequently prehybridised at 42°C for 45 minutes during which time the [α-^{32}P]-dCTP probe (the preparation of which is described in section 2.2.3.9) was denatured at 100°C for 10 minutes and subsequently chilled on ice. The [α-^{32}P]-dCTP probe was added at a final concentration of $1x10^6$ cpm per ml of ULTRAhyb ™ and was allowed to hybridise overnight with shaking at 42°C. The following morning, the blot was washed as follows: 0.5%SDS/2xSSC 2x5 minutes, 0.5%SDS/0.5xSSC 15 minutes, after which the blot was sealed in a polythene bag and was exposed to film at -70°C using two intensifying screens. The film was developed in an automated processor (Automated Processor: Compact X4 supplied by Xograph, Wiltshire, UK.) as described by the manufacturer.

5.2.5 Reverse transcription PCR

5.2.5.1 First strand cDNA synthesis from total RNA

First strand cDNA synthesis consisted of two steps. The first step involved denaturing of the total RNA and an oligo$(dT)_{15}$ primer. The second step was the transcription of cDNA in the presence of the enzyme "reverse transcriptase" ("Expand™ Reverse Transcriptase", Roche Applied Science, East Sussex, United Kingdom). All incubations were performed using a Perkin Elmer 9700 thermal cycler. For the cDNA synthesis, 1μg of total RNA and 0.5μg of an oligo$(dT)_{15}$ primer (Roche Applied Science, East Sussex, United Kingdom) were denatured at 65°C for 10 minutes and cooled immediately to 4°C. To this, a bulk mix containing the following reagents was added: 1x Expand™ Reverse Transcriptase buffer (first strand), 10mM DTT, 1mM each of dATP, dCTP, dGTP and dTTP, 20 units of RNase inhibitor (Roche Applied Science, East Sussex, United Kingdom), 50 units of the enzyme Expand™ Reverse Transcriptase to a total volume of 20μl. This mixture was incubated at 42°C for 1 hour and cooled immediately to 4°C. The cDNA was used immediately for PCR amplification as described below and the remainder was stored at -20°C until required.

5.2.5.2 PCR amplification of cDNA

5μl of the cDNA synthesis product was used in each PCR reaction. Oligonucleotide primers were synthesised corresponding to the 5' start codon and 3' stop codon of the FeLV-C receptor feFLVCR (5'-TCA AAT TGC TGA TTC TGA CTG-3' and 5'-CTC AAC TGC CTG GGT GCC TGG-3') and of the housekeeping gene β-actin (5'-ATC TGG CAC CAC ACC TTC TAC AAT GAG CTG CG-3' and 5'- CGT CAT ACT CCT GCT TGC TGA TCC ACA TCT-3'). In each PCR reaction, the final concentrations of the components were the following: 1x buffer 2 (Roche applied science, East Sussex, United Kingdom), 10mM dNTP (Promega, Southampton, United Kingdom), 37.5pmol of each primer (MWG, United Kingdom) and 2.5 units enzyme 1 (Roche applied science, East Sussex, United Kingdom).

The amplification was performed under the following conditions: denaturation at 94°C for 3 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 50°C for 1 minute and extension at 72°C for 3 minutes after which there was an additional 10 minutes extension step at 72°C. The amplification reaction yielded a 1.6kb product as predicted for the FeLV-C receptor (feFLVCR) and a 1.3kb product as predicted for the housekeeping gene (β-actin). The samples were separated by electrophoresis and visualised by ethidium bromide staining as described previously (section 2.2.3.5).

5.3 Results

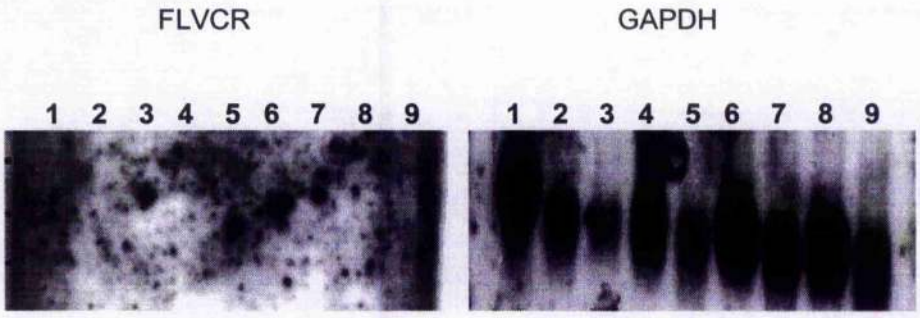
5.3.1. feFLVCR expression in a panel of tissues using northern blot analysis

5.3.1.1 Northern blot hybridisation using RapidHyb™ buffer

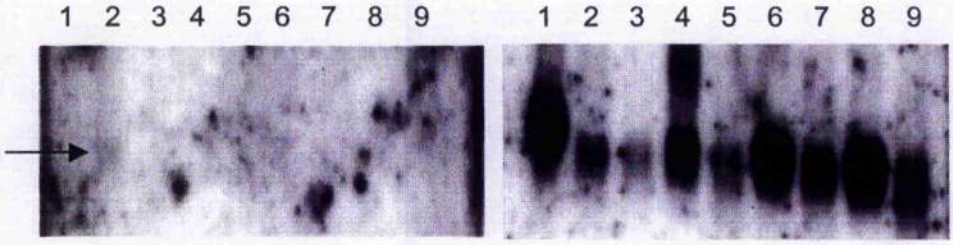
RNA samples from cats F4 and F8 were separated by electrophoresis and northern blotted (section 5.2.3.) in order to analyse the distribution of the FeLV-C receptor mRNA in both haematopoietic and non-haematopoietic tissues. The blot was probed using an α-^{32}P labelled feFLVCR probe containing the entire FLVCR gene and subsequently exposed to the film at -70°C for 3 days before the film was developed. As figure 5.1a shows, no FLVCR bands were detected in the blot probed with the FLVCR probe. However, subsequent screening of the same blot with the control housekeeping gene GAPDH, revealed transcripts of a similar intensity in all lanes, confirming that the lack of message on the blot with the feFLVCR probe was not due to differences in loading of the RNA samples but more likely due to the FLVCR gene being rare in these tissues. The presence of bands rather than smears when probed with the GAPDH probe confirmed that the RNA was intact and had not been degraded. To overcome the problems in analysing the distribution of the FLVCR due to the scarcity of the FLVCR transcript, an alternative hybridisation buffer, ULTRAHYB™, was used to hybridise an additional northern blot. As the manufacturers of ULTRAHYB™ (Ambion Europe Ltd, Cambridgeshire, United Kingdom), described their product as being sensitive enough to pick up rare transcripts, ULTRAHYB™ was used for subsequent hybridisations.

5.3.1.2 FLVCR expression in feline tissues by northern analysis using ULTRAHYB™ hybridisation buffer

RNA samples from cat F4 were separated by electrophoresis and a northern blot was prepared and hybridised as described previously (section 5.2.3.). The blot was exposed to the film at -70°C for 3 days prior to developing the film. As can be seen in figure 5.1b, when the northern blot was probed using the feFLVCR probe, expression of the feFLVCR transcript could not be detected in the majority of the tissues studied. A faint band corresponding to the feFLVCR transcript was visible in the sample lanes bearing mRNA from thymus and salivary gland suggesting that the transcript was present at levels close to the sensitivity of northern blotting. When the same northern blot was probed with the GAPDH probe, transcripts of similar intensity were detected in all RNA samples. This demonstrated that similar amounts of RNA were loaded onto the gel, the RNA was intact and subsequently confirmed that the failure to detect FLVCR transcripts was due to the low level of gene expression. As was shown in the northern blot hybridised using RapidHyb™ buffer, the RNA was not degraded as there were bands rather than smears present in each lane.



A: Northern blot using RapidHyb™ hybridisation buffer.



B: Northern blot using ULTRAhyb™ hybridisation buffer.

Figure 5.1: *Northern blot analysis of total RNA prepared from a panel of tissues collected from cat F4. The northern blots were hybridised using either RapidHyb™ or ULTRAhyb™ with either, ^{32}P-labelled feFLVCR or a housekeeping gene, GAPDH. The order of the lanes is as follows: 1 - MLN, 2 - thymus, 3 - heart, 4 - brain, 5 - spleen, 6 - salivary gland, 7 - liver, 8 - skeletal muscle and 9 - kidney.*

5.3.2 feFLVCR expression in a panel of tissues using RT PCR analysis

5.3.2.1 Expression of the FeLV-C receptor in a panel of tissues collected from cat F4

cDNA was synthesised from RNA collected from cat F4 (a cat from an FIV vaccination trial and thus FeLV negative). The cDNA was subsequently used for PCR amplification with primers encoding either feFLVCR or β-actin. The PCR products were separated by electrophoresis and visualised by ethidium bromide staining (figure 5.2). PCR using the β-actin primers yielded products of a similar intensity, confirming that the cDNA synthesis reactions contained amplifiable cDNA in broadly similar amounts. Therefore, when comparing the intensity of the PCR products following amplification using the feFLVCR primers, the intensity of the PCR products could be correlated broadly with the level of expression of the feFLVCR gene in that particular tissue. FLVCR was relatively abundant in both the liver and the MLN, but expressed at lower levels in kidney, heart and thymus. The saliva gland, spleen and brain showed intermediate levels of FLVCR expression. Thus, in cat F4, there was no evidence for the preferential expression of the FLVCR mRNA in haematopoietic tissues.

5.3.2.2 RT PCR on RNA prepared from tissues collected from cat E2

cDNA was synthesised from RNA collected from cat E2 (a cat from a FeLV immunopathogenesis trial and at the time of post mortem positive for FeLV p27). The cDNA was subsequently used for PCR amplification with primers encoding either feFLVCR or β-actin (figure 5.2). In contrast to cat F4, PCR using the β-actin primers yielded products with varying intensity, in particular the kidney and the skeletal muscle show diminished intensities, suggesting that the cDNA synthesis reactions contained amplifiable cDNA in varying amounts. Thus, the results of feFLVCR RT-PCR from both the kidney and skeletal muscle cDNA should be interpreted with caution.

When amplifying cDNA synthesised from RNA collected from the remaining tissues, PCR using the β-actin primers yielded products of similar intensity, confirming that the cDNA synthesis reactions contained amplifiable cDNA in broadly similar amounts. Thus, variations in the intensity of the feFLVCR PCR products were likely due to differences in the level of feFLVCR mRNA in each tissue. feFLVCR mRNA was not detected in the kidney, heart or thymus. In contrast, the mRNA was relatively abundant in salivary gland, liver and spleen. Lower levels of mRNA were detected in skeletal muscle, mesenteric lymph node and peripheral lymph node. Thus, in cat F4, expression of feFLVCR mRNA did not appear to be restricted to haematopoietic tissues.

5.3.2.3 Expression of the FeLV-C receptor in a panel of tissues collected from cat E5

cDNA was synthesised from RNA prepared from the tissues from cat E5 (a cat from a FeLV immunopathogenesis trial and at the time of post mortem positive for FeLV p27) and subsequently used for PCR amplification using primers specific for either feFLVCR or β-actin. The PCR products were separated by electrophoresis and visualised by ethidium bromide staining (figure 5.2). As was found for cat E2, the levels of β-actin expression varied between the tissues screened, suggesting that the amounts of amplifiable cDNA present varied between the cDNA synthesis reactions, in particular for the heart, liver and skeletal muscle. Thus the results obtained with these tissues following PCR amplification using the feFLVCR primers should be interpreted with caution. The remaining PCR using primers encoding for β-actin yielded products of comparable intensity and thus amplifiable cDNA was present in the cDNA synthesis reactions in broadly comparable amounts. FeFLVCR mRNA was relatively abundant in the kidney, spleen and peripheral lymph node. There were lower levels of mRNA detected in the remaining tissues (salivary gland, thymus, spleen and MLN). Despite the poor quality of the liver cDNA template, a product was detected following feFLVCR PCR, suggesting that the feFLVCR transcript was abundant in this tissue. Thus, in cat E5, feFLVCR was expressed in both haematopoietic and non-haematopoietic tissues.

5.3.2.4 Expression of the FeLV-C receptor in a panel of tissues collected from cat E6

Tissues collected from cat E6 (a cat that was FeLV p27 positive at the time of post mortem) were used for RNA preparation. After cDNA synthesis and PCR amplification, the PCR products were visualised on an agarose gel (figure 5.2). The levels of β-actin expression were broadly similar between the tissues screened, with the exception of the skeletal muscle and mesenteric lymph node samples which showed low intensity products, suggesting that less amplifiable cDNA was present in the cDNA synthesis reactions from these tissues and thus the results for the feFLVCR PCR from the cDNAs derived from these tissues should be interpreted with caution. Interestingly, both the skeletal muscle and MLN show high levels of expression of feFLVCR, suggesting that the FeLV-C receptor was relatively abundant in both tissues. When comparing the remaining tissues, the heart showed relatively abundant expression of feFLVCR, whereas there was very low or no expression of the gene in the salivary gland, thymus and peripheral lymph node. The remaining tissues (kidney, liver and spleen) showed intermediate levels of expression, thus strengthening the findings of the previous studies that the FeLV-C receptor feFLVCR is not preferentially expressed in haematopoietic tissues.

5.3.2.5 Expression of the FeLV-C receptor in a panel of tissues collected from cat L32

The tissues in this part of the experiment were collected from a cat that was part of a FeLV vaccination trial and which remained FeLV p27 negative at the time of post mortem. After cDNA synthesis and PCR amplification, the PCR products were visualised on agarose gels (figure 5.2). Comparable levels of RNA were used for the initial cDNA synthesis, as the intensity of the PCR products in the lanes containing the β-actin products were of equal intensity. Therefore, variations in the intensity of the feFLVCR PCR products were due to differences in the level of expression of feFLVCR mRNA.

Comparing the levels of expression in the lanes containing the feFLVCR PCR products, no feFLVCR mRNA was detected in the spleen while high levels of feFLVCR expression were detected in the heart, salivary gland, skeletal muscle and thymus.

5.3.2.6 Expression of the FeLV-C receptor in a panel of tissues collected from cat L36

The tissues in this RT PCR assay were collected from a cat was part of a FeLV vaccination trial and was FeLV p27 negative at the time of post mortem. The PCR products were visualised on agarose gels and are shown in figure 5.2. The β-actin PCR products from the spleen and kidney were poor in comparison with the remaining tissues, indicating that the results found for feFLVCR expression in both the kidney and the spleen should be interpreted with caution. With the exception of both the MLN and the PLN, feFLVCR was expressed to high levels in all tissues tested from this particular cat. Interestingly, the feFLVCR expression in both the spleen and the kidney was comparable to the expression in the remaining tissues, indicating that, despite the poor quality of the template cDNA, high levels of feFLVCR mRNA could still be detected in these tissues. The distribution of the feFLVCR mRNA in tissue from cat L36 was consistent with expression being widespread in the cat and not restricted to a single lineage of cells.

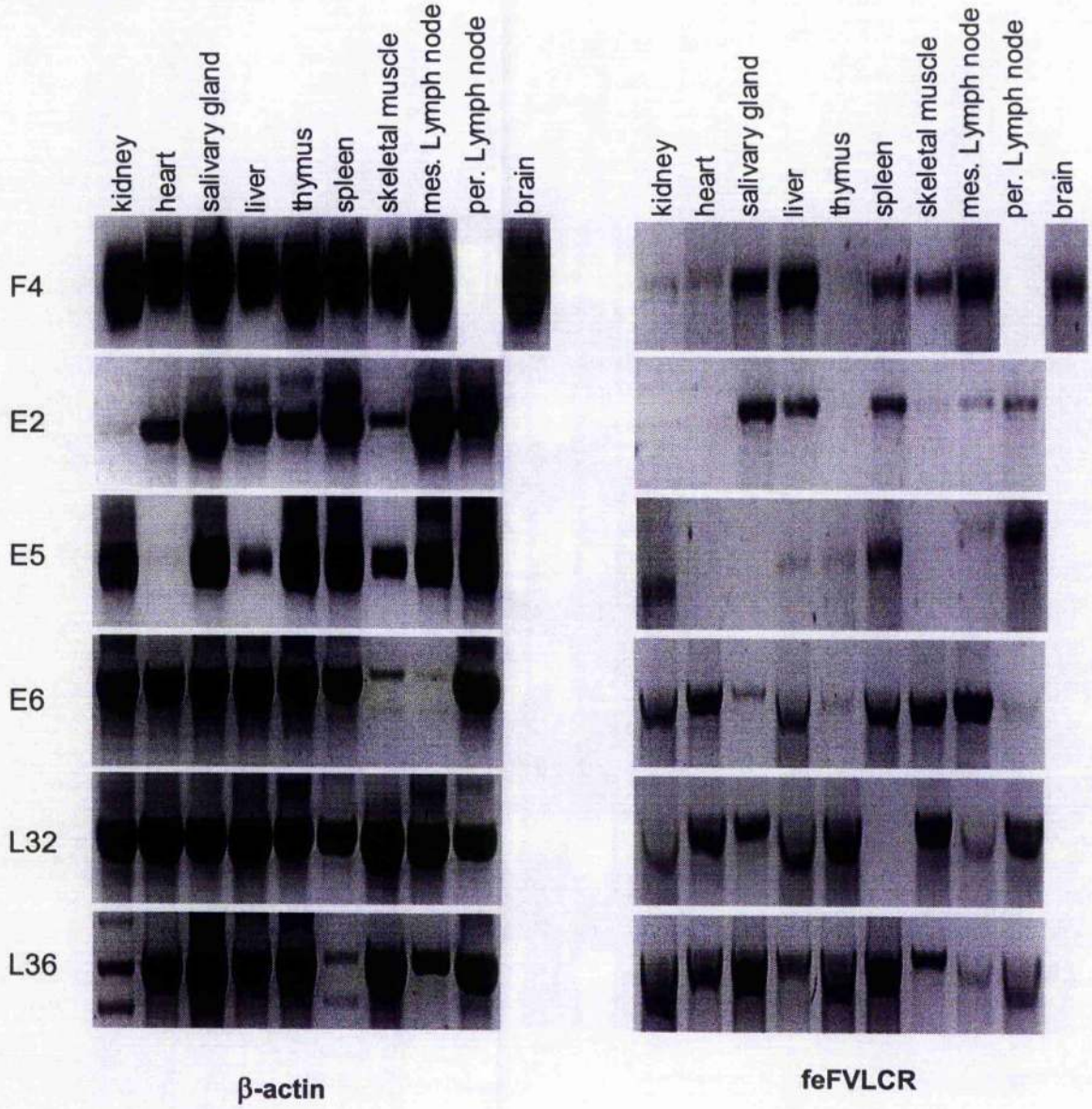


Figure 5.2: *RT PCR on a panel of tissues collected from a range of cats. Oligonucleotide primers were synthesised according to the published sequences feFLVCR or the housekeeping gene β-actin. Amplified products were visualized on agarose gels with ethidium bromide staining.*

	F4		E2		E5		E6		L32		L36	
	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin
Kidney	+/-	+	-	+/-	+	+	+	+	+/-	+	++	+/-
Heart	+/-	+	-	+	-	+/-	+++	+	+	+	+	+
Saliva gland	++	+	++	+	-	+	+/-	+	+	+	++	+
Liver	+++	+	+	+	+/-	+/-	+	+	+	+	+	+
Thymus	-	+	-	+	+/-	+	-	+	++	+	+	+
Brain	++	+	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Spleen	+	+	+	+/-	+	+	+	+	-	+	++	+/-
Skeletal muscle	+	+	-	+	-	+/-	+++	+/-	++	+	+	+
MLN	++	+	+/-	+	+/-	+	+++	+/-	+/-	+	+/-	+
PLN	N.D.	N.D.	+	+	++	+	-	+	+	+	+/-	+

Table 5.2: *Expression of the FeLV-C receptor feFLVCR in a panel of tissues collected from six cats by RT PCR assay. The first column shows the feFLVCR expression, the second column shows the β-actin expression. The level of expression of feFLVCR ranges between from no visible expression (-) to a good level of expression (+++), whereas N.D. indicates that the tissue was not screened for feFLVCR expression. The tissues where the β-actin expression was lower than the remaining tissues from that cat are highlighted in red.*

5.3.3 feFLVCR expression in bone marrow enriched for or depleted of erythroid progenitors

Bone marrow was extracted from the femur of four cats (F4, F8, L22 and L36) and subsequently used to either enrich or deplete cells reactive with the monoclonal antibody K1, an antibody directed against erythroid progenitor cells (Gengozian N., 1998). In a population that has been depleted of cells reacting with this antibody, BFU-E progenitors will have been reduced by approximately 80%. A control bone marrow population that had not undergone this treatment was also included.

5.3.3.1 Selection of bone marrow populations using an antibody directed against erythroid progenitors

After selection of bone marrow populations from cats F4, F8, L22 and L36, the populations were counted, aliquoted and stored at -70°C until required. RNA was prepared as described previously (section 5.2.2.2) and subsequently used for cDNA synthesis and PCR amplification using primers encoding either feFLVCR or β-actin. The PCR products were separated by electrophoresis and visualised by ethidium bromide staining (figure 5.3).

5.3.3.2 RT PCR analysis on mRNA from erythroid progenitors

As can be seen in figure 5.3, the PCR products amplified using primers encoding for β-actin were of equivalent intensity in the samples from F8, L22 and L36, indicating that equivalent amounts of amplifiable cDNA were present in the initial cDNA synthesis. The samples from F4 yielded variable intensity products and thus the results of feFLVCR PCR for this sample should be interpreted with caution. Following RT-PCR for feFLVCR mRNA, a product was detected in control bone marrow samples from the four cats. In samples F4, F8 and L22, depletion of the sample with K1- antibody reduced the levels of feFLVCR mRNA.

Conversely, in samples F4 and L22 there was a strong product for the feFLVCR mRNA in the K1- enriched fraction. Interestingly, F4 had a relatively weak intensity product for β-actin in the K1-enriched sample suggesting that the feFLVCR product must be relatively abundant in the K1- enriched population from this cat. Enrichment for K1+ cells in the L36 bone marrow sample also markedly increased the intensity of the feFLVCR product, although in this sample the K1- depleted sample did not differ in intensity from the control sample. It is notable that the β-actin product was considerably stronger in the K1- depleted sample from this cat and weaker in the K1- enriched sample, suggesting that relative to the β-actin product there had been a marked enrichment for feFLVCR mRNA with K1- selection. No feFLVCR product was detected in bone marrow sample F8 following either enrichment or depletion of the K1+ population. Thus, in three out of four samples, there was evidence for segregation of the feFLVCR transcript with the K1+ erythroid progenitor population.

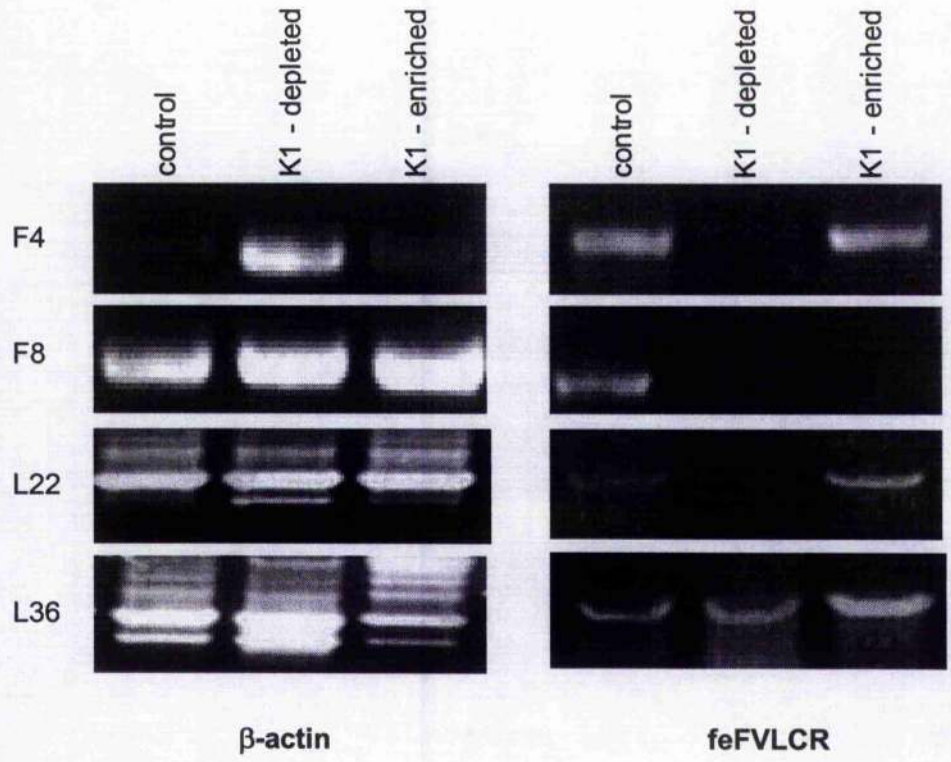


Figure 5.3: *Effect of enrichment of erythroid lineage cells on abundance of feFLVCR mRNA. Bone marrow populations were either enriched for or depleted of cells reactive with the anti-erythroid lineage antibody K1. RNA was isolated from the three cell populations and cDNA prepared. The cDNAs were then assayed for the presence of β-actin or feFLVCR cDNAs by PCR. Amplified products were visualized on agarose gels with ethidium bromide staining.*

	F4		F8		L22		L23		L32		L36	
	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin	feFLVCR	β-actin
Untreated population	+	-	+	+	+	+	-	-	-	-	+	+
K1 depleted population	-	+	-	+	-	+	-	-	-	-	+	++
K1 enriched population	+	+/-	-	+	+	+	-	-	-	-	+	+

Table 5.3: *Expression of the FeLV-C receptor feFLVCR in bone marrow populations collected from six cats by RT PCR assay. The bone marrow populations were either enriched for or depleted from cells that were reactive with the antibody K1. A bone marrow population containing cells that had not undergone this treatment was included as a negative control. The level of expression of feFLVCR ranges between from no visible expression (-) to a good level of expression (+++). The populations where the β-actin expression was lower than the remaining populations from that cat are highlighted in red.*

5.4 Discussion

Both the human (huFLVCR) and the feline (feFLVCR) receptor for FeLV-C have been identified and characterised (Tailor C.S. et al., 1999c, Quigley J.G. et al., 2000). As described previously (see section 1.5.1), both the human and feline receptors are multiple transmembrane domain proteins that are transporters and are members of the major facilitator superfamily (MFS, (Pao S.S. et al., 1998)). Studies on the human FeLV-C receptor, huFLVCR, suggested that the receptor was expressed preferentially on haematopoietic cells, but not solely on erythroid progenitors (Tailor C.S. et al., 1999c). To date, the feline receptor, feFLVCR, has not been characterised in such detail. Dean et al (Dean G.A. et al., 1992) found that FeLV antigens (gp70 and p27) were present in both populations that contained erythroid cells and populations containing myeloid or lymphoid cells and that the proteins were present to comparable percentages in the samples tested. However, Tailor et al (Tailor C.S. et al., 1999c) showed that huFLVCR (the human FeLV-C receptor) was expressed preferentially on haematopoietic cells, but it was not ascertained whether receptor expression was restricted to erythroid progenitor cells. Thus, in an attempt to clarify the possible role of the FeLV-C receptor in the onset of PRCA, a panel of both haematopoietic and non-haematopoietic tissues were screened for expression of feFLVCR. At the same time, bone marrow samples were collected and either depleted or enriched using an antibody reactive with erythroid progenitors and RNA was prepared from these bone marrow populations. The RNA samples were subsequently analysed using either northern blot analysis or RT PCR techniques for expression of the feFLVCR mRNA.

5.4.1 feFLVCR expression in a panel of tissues using northern blot analysis

When the level of expression of the FeLV-C receptor feFLVCR was analysed using northern blotting, it was evident that the feFLVCR mRNA was not expressed at high levels in any of the tissues studied.

The inability to detect the feFLVCR mRNA transcript by using northern blotting is in contrast with the findings of Tailor et al (Tailor C.S. et al., 1999c) where a panel of human tissues was screened using a huFLVCR probe. Tailor et al concluded that the human FeLV-C receptor, huFLVCR, was preferentially expressed in haematopoietic tissues. However, expression of the 2.0kb transcript was also detected in pancreas and kidney, and a faint transcript may have been present in other non-haematopoietic tissues. The northern blot used in the study described by Tailor et al was a custom made blot that contained polyA (+) mRNA, whereas in the study described in this thesis the northern blots contained total RNA. Thus it is conceivable that the lack of feFLVCR expression in the northern blot analysis described here may be due to the type of RNA used initially to prepare the northern blots, as polyA (+) mRNA is more pure that total RNA and thus more of the transcript will be present in polyA (+) mRNA than in total RNA. The amount of transcript present in the type of RNA is also reflected in the amount of RNA loaded onto the initial gel; Tailor et al describe loading 2 μg of polyA (+) mRNA onto the gel, whereas in this study, 20 μg of total RNA was loaded onto the gel.

5.4.2 feFLVCR expression in a panel of tissues using RT PCR analysis

Due to the relative scarcity of the feFLVCR transcript, it was concluded that northern blot analysis was not sensitive enough to detect tissue expression of feFLVCR. Therefore, the RNA prepared from both haematopoietic and non-haematopoietic tissues was screened using RT PCR, a more sensitive technique than northern blotting. RT-PCR revealed that the FeLV-C receptor feFLVCR was expressed widely in both haematopoietic and non-haematopoietic tissues and indicated that no single tissue expressed consistently low or high levels of the feFLVCR transcript when comparing between cats. In agreement with Tailor et al., the majority of lymphoid tissues tested (spleen, thymus, mesenteric and peripheral lymph node) expressed the feFLVCR transcript. However, in contrast with the observations of Tailor et al., the feFLVCR transcript was also detected at similar levels in non-lymphoid tissues such as the salivary gland or the liver.

5.4.3 feFLVCR expression in bone marrow populations following enrichment for cells of the erythroid lineage

The main haematological symptoms of PRCA are decreased levels of CFU-E, but normal levels of BFU-E, suggesting a defect in the maturation from BFU-E to CFU-E (see also figure 1.6). Thus, if the FeLV-C receptor was expressed exclusively on cells involved in erythroid maturation, the screening of an enriched bone marrow population using primers encoding for feFLVCR would result in higher amounts of PCR product, whereas screening depleted bone marrow populations would result in low amounts of PCR product. Following depletion of erythroid progenitors from the bone marrow cell suspensions, levels of the feFLVCR transcript detected by RT-PCR were markedly reduced. Further, in the majority of cases, the feFLVCR transcript appeared to be enriched in the K1+ population (erythroid). These data indicate segregation of feFLVCR transcript expression with cells of the erythroid lineage and do raise the possibility that the selective targeting of erythroid progenitors by FeLV-C is the result of high level expression of the feFLVCR transcript in this lineage of cells and accordingly, an important role for the feFLVCR protein in the development of cells of this lineage. Further samples should be analysed to confirm the validity of these data.

CHAPTER SIX

IDENTIFICATION AND CHARACTERISATION OF A PUTATIVE FeLV-A RECEPTOR

6.1 Introduction

Both the FeLV-B receptor *Pit-1* and the FeLV-C receptor FLVCR have been identified and characterised previously (Takeuchi Y. et al., 1992; Tailor C.S. et al., 1999c; Quigley J.G. et al., 2000). FeLV-B shares its receptor *Pit1* with the gibbon ape leukaemia virus GALV (Takeuchi Y. et al., 1992). The FeLV-B receptor *Pit1* is a Na^+-dependent inorganic phosphate transporter that has 10 transmembrane domains and 5 extracellular loops (see figure 1.8a). Both the human and feline receptors for FeLV-C have been identified (by Tailor et al (the human homologue huFLVCR (Tailor C.S. et al., 1999c)) and Quigley et al (the feline molecule feFLVCR (Quigley J.G. et al., 2000) respectively). Both molecules show a high degree of similarity (83% nucleic acid similarity and 89% amino acid similarity) and both molecules have 12 transmembrane spanning domains and six extracellular loops (see figure 1.8b). To date, the FeLV-A receptor remains to be identified and characterised. Ghosh et al (Ghosh S.K. et al., 1992) attempted to identify a FeLV-A receptor by using purified envelope glycoprotein gp70 of a representative FeLV-A virus 61E (accession number M18247, (Donahue P.R. et al., 1988)) to bind the viral receptor. ^{125}I-labelled gp70 protein bound to FeT cells (3201 cells) was employed to detect membrane proteins that interact with the gp70 following affinity cross-linking. These methods identified a complex of 130 to 140 kDa, leading to the conclusion that the FeLV-A receptor is a cell surface molecule of approximately 70 kDa. To date, this molecule remains to be identified and characterised. This method has been used previously to characterise, amongst others, the interaction of human immunodeficiency virus (HIV) and CD4+ (McDougall J.S. et al., 1986).

As discussed previously (section 1.5.2), a number of receptors of the γ-type retroviruses are members of the major facilitator superfamily, MFS, and it is conceivable that the FeLV-A receptor is also a member of this extensive family of molecules. Furthermore, due to the high degree of homology between FeLV-A and FeLV-C, it is likely that both viruses utilise related molecules.

6.2 Methods

6.2.1 Cloning of a receptor using a retroviral library

Two libraries (a feline T-cell library and a human T-cell library in the pBabe X retroviral vector (Kitamura T. et al., 1995)) were screened to identify the putative FeLV-A receptor (see also figure 6.1). Cells were transfected using the $CaPO_4$ method as described previously by Gorman (Gorman C.M. et al., 1985). On day 1, $8x10^6$ 293T cells were plated in 150mm dishes in 20mls 10% DMEM. On day 2, chloroquine was added to the cells to a final concentration of 25μM (the pH of the medium is increased by this endosomal agent, which in turn inhibits lysosomal hydrolytic enzymes and thus decreases the risk of degradation of endocytosed material (Luthman H. and Magnusson G., 1983)), 24μg of either the feline T-cell library or the human T-cell library, 24μg Hit60 (which contains MuLV *gag-pol* to enable formation of FeLV *env*/MuLV *gag-pol* pseudotypes (Soneoka Y. et al., 1995)) and 24μg mfg LacZ, (carrying the β-galactosidase marker gene (Dranoff G. et al., 1993)) were combined and the total volume made up to 1752μl with distilled water. To this mixture, 248μl of 2M $CaCl_2$ was then added without mixing. 2000μl 2xHBS was added slowly and the DNA mixture was added to the cells. After a 10 hour incubation, the transfection mixture was replaced with fresh 10% DMEM and the cells were then incubated overnight at 37°C. On day 3, the cells were washed twice with PBS and fresh growth medium was added. Target cells 104C1 and 3T3 (both 3T3 and 104C1 are not susceptible to FeLV-A infection, as can be seen in table 3.5) were plated onto six 100mm dishes at a concentration of $5x10^5$ cells per dish and incubated overnight at 37°C. 293T cells were plated in 150mm dishes at a concentration of $8x10^6$ cells per dish and incubated overnight at 37°C. The following day, the supernatant of the transfected 293T cells was removed, filtered using a 0.45μM filter and fresh 10% DMEM was added to a total volume of 100ml.

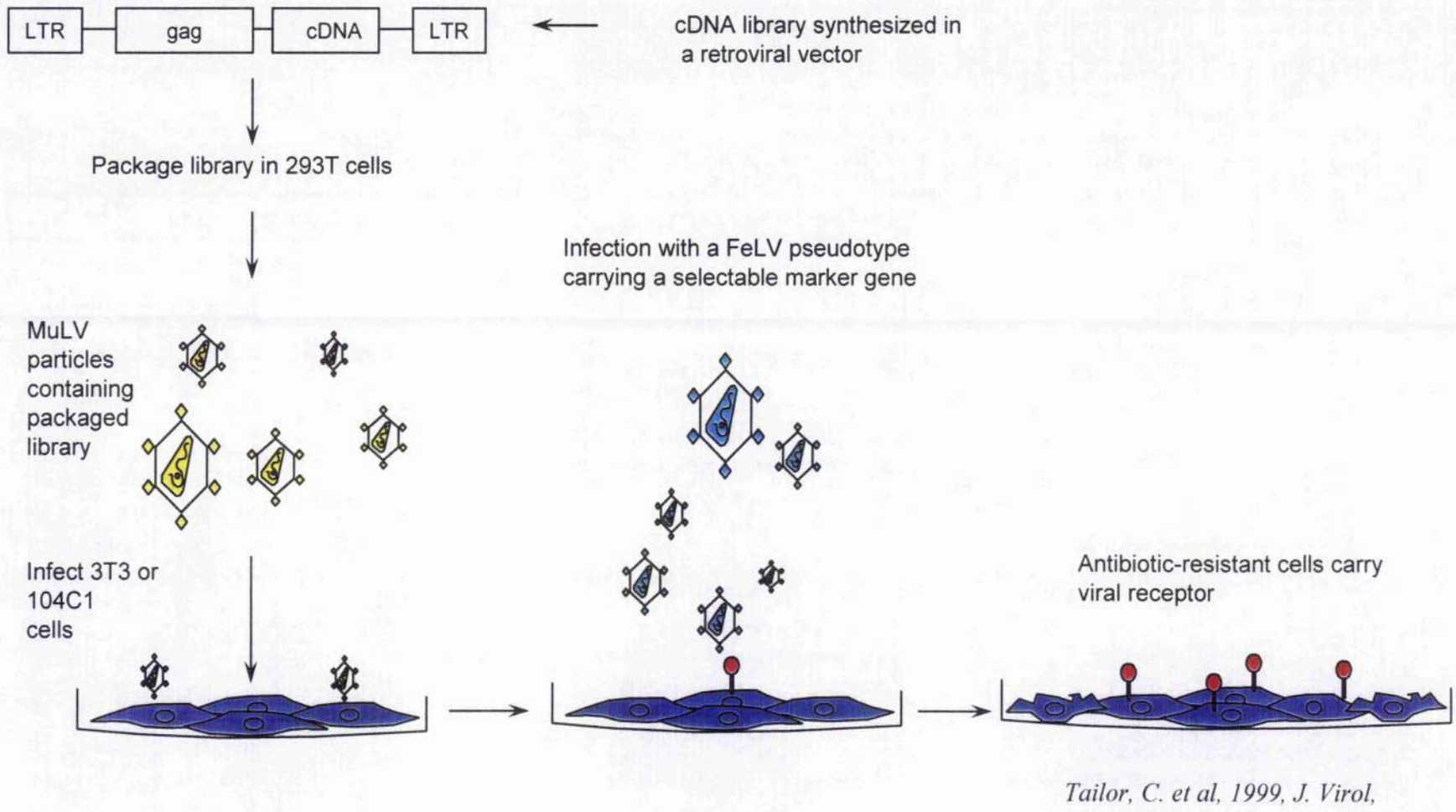


Figure 6.1: Cartoon depicting the method used to identify the FeLV-A receptor.

The transfection efficiency was estimated by staining the cells for LacZ activity. The cells were fixed for 10 minutes in PBS containing 0.5% glutaraldehyde, washed three times with PBS and β-galactosidase activity was determined by incubation with x-gal buffer (5mM potassiumferrocyanide, 5mM potassiumferricyanide, 2mM $MgCl_2$) containing 1mg/ml x-gal (5-bromo-4-chloro-3-indolyl β-D galactopyranoside). Five of the 100mm dishes containing either 3T3 or 104C1 cells were transduced using 10ml of the 293T supernatant in the presence of 4μg/ml (w/v) polybrene (hexadimethrine), to enhance the uptake of retrovirus into the cells (Coelen R.J. et al., 1983)) and incubated overnight at 37°C. The remaining two dishes were a negative control; 10% DMEM containing 4μg/ml (w/v) polybrene was added to the remaining two dishes and these dishes were incubated overnight at 37°C. On day 5, the 100mm dishes containing the transduced 3T3 or 104C1 cells and both control dishes were subcultured into 150mm dishes as described previously in section 2.2.1.2.3. The 150 mm dishes containing 293T cells were transfected using 24μg FeLV-A *env*, 24μg HIT60 and 24μg pBabe Puro (this vector permits the selection of cells that have been transduced successfully by adding growth medium containing puromycin, (Morgenstern J.P. and Land H., 1990)) in the presence of 25μM chloroquine and incubated for 9 hours after which the medium was replaced with fresh growth medium. On day 6, the 293T cells that were transfected were washed twice with PBS, after which 20ml growth medium was added to the cells and cells were incubated for 24 hours at 37°C. After 24 hours incubation the supernatants of the transfected 293T dishes were filtered at 0.45μM and fresh 10% DMEM was added to a total volume of 100ml. Again, 150mm 293T dishes were stained for transfection efficiency by fixing the cells for 10 minutes in PBS containing 0.5% glutaraldehyde, washed three times with PBS and β-galactosidase activity was determined by incubation with x-gal buffer (5mM potassiumferrocyanide, 5mM potassiumferricyanide, 2mM $MgCl_2$) containing 1mg/ml x-gal (5-bromo-4-chloro-3-indolyl β-D galactopyranoside). The 150mm dishes containing the transduced 3T3 and 104C1 cells were infected with the 293T supernatant and were incubated at 37°C for 3 days. At this point, each dish containing 104C1 cells was subcultured into 2 dishes and each dish containing 3T3 cells was subcultured into 3 dishes. After 4 hours incubation to allow the cells to adhere to the dishes, puromycin was added to the dishes at a concentration of 2μg/ml for the 3T3 cells and 0.5μg/ml for the 104C1 cells.

The puromycin-containing medium was refreshed every 2 days and after approximately one and a half weeks, puromycin resistant colonies appeared in the dishes. The puromycin resistant colonies were picked using autoclaved 3MM paper disks soaked in trypsin/versene, placed in a well of a 24 well plate containing 10% DMEM and incubated overnight at 37°C. The following day, the 3MM paper disks were removed, the cells were washed twice with PBS and puromycin was added to the medium as previously. When the cells were confluent, each well was subcultured into 2 wells of a 24 well plate. Once confluent, the cells of one well were infected with FeLV-A *env*/mfg LacZ to determine whether the picked colonies were positive when infected with supernatant containing FeLV-A mfg LacZ particles as described previously (section 3.2.7). The wells corresponding to the cells that were positive when stained for β-galactosidase were used to prepare HMW DNA.

6.2.2 Isolation of receptor cDNA

Cell pellets of colonies that tested positive for FeLV-A *env*/mfg LacZ infection as determine by the β-galactosidase assay described previously (section 3.2.7) were used to prepare HMW DNA (section 3.2.2). The prepared DNA was subjected to PCR amplification using primers flanking the insert in the pBabe X vector (5'-GATCCCAGTGTGCTGGAAAG-3' and 5'-GGTGGGGTCTTTCATTCC-3'). The amplification was performed under the following conditions: denaturation at 94°C for 3 minutes, followed by 30 cycles of denaturation at 94°C for 30secs, annealing at 54°C for 1.5 minutes and extension at 68°C for 7 minutes. The PCR products were expected to be approximately 2.0kb in size.

6.2.3 Molecular cloning of PCR products

The PCR products were cloned into the pCDNA3.1V5HisTOPO vector (Invitrogen NV, Leek, The Netherlands) using the TA cloning system as described previously (section 2.2.3.7) and transformed into *E.coli* INVαF' competent cells (section 2.2.3.8) as described previously.

Colonies were picked, inoculated in 2.5ml L-broth containing 50μg/ml ampicillin and grown overnight at 37°C. DNA was prepared using the "small-scale plasmid preparation" method (section 2.2.3.2.1) and the plasmid DNA was digested using 10 units of the restriction enzyme *BstXI* to establish the presence of an insert (section 2.2.3.5).

6.2.4 Nucleic acid sequencing of a putative FeLV-A receptor

The putative FeLV-A receptor clones were sequenced using an ABI PRISM 3100 genetic analyser (section 2.2.3.11) using either the universal M13 forward primer (5'-GTA AAA CGA CGG CCA GT-3') or the universal M13 reverse primer (5'-CAG GAA ACA GCT ATG AC-3'), these primers can be used with any vector containing the N-terminus of β-galactosidase. The sequences obtained were analysed using the NCBI blast nucleotide search on the website (http://www.ncbi.nlm.nih.gov/BLAST/) to establish possible homology to proteins that function as receptors.

6.3 Results

6.3.1 Screening of the feline T-cell library on murine 3T3 cells or guinea pig 104C1 cells

After puromycin selection of both the cells transduced with the feline T cell library and the negative control cells, colonies were picked and each colony plated into a single well of a 24 well plate. Interestingly, no puromycin resistant colonies were present on the 104C1 control plates, whereas twenty-three puromycin resistant colonies were present on the 3T3 control plates. The plates that contained either 3T3 transduced or 104C1 transduced cells contained one hundred and twenty four and six puromycin resistant colonies respectively. Thus, when comparing the number of puromycin resistant colonies on the 3T3 transduced cells with the number of colonies on the 3T3 control plates, there is a significant increase in the number of colonies appearing after puromycin selection. Once confluent, these colonies were screened for their ability to support FeLV-A *env*/mfg LacZ infection. However, none of the colonies were capable of supporting FeLV-A *env*/mfg LacZ infection. This led to the conclusion that, despite these colonies being puromycin resistant, none of them displayed the putative FeLV-A/Glasgow-1 receptor on the cell surface.

6.3.2 Screening of the human T-cell library on murine 3T3 cells or guinea pig 104C1 cells

Similarly, after puromycin selection of either 104C1 or 3T3 cells that were transduced with the human T-cell library, none of the sixteen or eighty five colonies isolated respectively were able to support FeLV-A *env*/mfg LacZ infection. As was found for the feline T cell library, there were no puromycin resistant colonies present on the 104C1 control plates and twelve puromycin resistant colonies on the 3T3 control plates (these control plates had not been transduced with the supernatant of transfected 293T cells, but otherwise had undergone the same protocol as the transduced cells).

Library screened	Cell line used to obtain puromycin resistant colonies	Number of puromycin resistant colonies isolated	Number of colonies able to support FeLV-A infection
Feline	3T3	124 (23)	0
Human	3T3	85 (12)	0
Feline	104C1	6 (0)	0
Human	104C1	16 (0)	0

***Table 6.1**: Summary of the screening of either the feline or the human T cell library, in order to identify the FeLV-A receptor. In brackets is the number of puromycin resistant colonies present on the control plates.*

6.4 Discussion

To date, the FeLV-A/Glasgow-1 receptor is the sole FeLV receptor that has not been identified and characterised. Ghosh et al (Ghosh S.K. et al., 1992) identified a putative FeLV-A/Glasgow-1 receptor, however, this molecule was not characterised in depth. In order to identify the cellular FeLV-A receptor, the same method, which has been used successfully by Tailor et al (Tailor C.S. et al., 1999c) in the identification of the FeLV-C receptor, was employed. However, when this approach was applied to the identification of the FeLV-A/Glasgow-1 receptor, we found that it was not possible to identify the viral receptor. There are a number of possible explanations for the difficulties encountered when screening either the human or the feline T-cell library, they are:

1. The receptor is being expressed only transiently on the surface of the cell and is removed from the surface due to the molecule being toxic to the cells.

2. mRNA for the FeLV-A receptor was not expressed at sufficiently high levels in the library and thus the probability of isolating the receptor was low.

3. When the cells are taking up the viral particles, this occurs through a non-specific pathway rather than through the receptor being expressed on the cell surface.

While it is unlikely that the uptake of the viral particles is though a non-specific pathway, the more likely the occurrence of false positive puromycin resistant colonies is due to either the libraries being not specific enough or due to the receptor being toxic to the cells.

The occurrence of puromycin resistant colonies on the plates containing control cells that have not been transduced (and thus should not contain the FeLV-A receptor) does indicate that there is "leaking" of the virus into the cells through a non-specific pathway. However, if this were the preferred method of uptake of the viral particles, this would have led to a positive infection when using FeLV-A/mfg LacZ containing supernatants. With respect to the cellular molecule encoding the putative receptor being toxic to the cells, Zhang et al (Zhang C.C. et al., 1999) have described similar problems encountered when attempting to stably express the Estrogen receptor (ER) in the human cell line HeLa. It was found that the molecule encoding for this receptor was highly toxic to these cells and thus was removed from the cell surface immediately, comparable to problems encountered when screening the cDNA library. Furthermore, both Dr C. Tailor and Dr T. Miyazawa attempted to identify the FeLV-A/Glasgow-1 receptor in a similar fashion as described in this thesis and were unable to identify a putative receptor (personal communications), in agreement with the studies described in this thesis. However, it is also possible that the library screened to identify the cellular receptor is not specific enough. A potential resolution to this would be the preparation of a library from a cell line or a tissue that is highly susceptible to FeLV-A infection, for example a feline cell line such as FEA or H06T1. Both these cell lines are readily infected by FeLV-A/Glasgow-1 and thus it is probable that the receptor will be present at high levels in these cell lines.

Another possible way to successfully identify the FeLV-A receptor will be to link the library to a system that is driven by an inducible promoter such as an antibiotic resistance gene. An example of such a system is the tetracycline-regulated system described originally by Gossen et al (Gossen M. and Bujard H., 1992). In brief, this tetracycline-regulated system is based on regulatory elements of the Tn*10* specified tetracycline resistance operon of *Escherichia coli*. The tetracycline repressor TetR negatively mediates transcription of the resistance-mediated genes, meaning that in the presence of tetracycline, TetR does not bind to its operators within the promoter region and thus allows transcription. Similarly, in the absence of tetracycline, TetR is able to bind to its operators, which are located in the promoter region, and thus transcription does not occur.

CHAPTER SEVEN

DISCUSSION

7.1 Introduction

Feline leukaemia virus is a common disease of the domestic cat and infection with the virus invariably leads to death three to five years after infection. Infection causes a number of diseases ranging from lymphoma to erythroid hypoplasia (pure red cell aplasia, PRCA). There are three known subgroups (FeLV-A, FeLV-B and FeLV-C), as determined by interference assays. FeLV-A is the subgroup that is most frequently isolated, it is present in 100% of natural isolates, whereas FeLV-C is most seldom isolated; it is present in only 1% of all natural isolates. FeLV-B is the result of recombination between the FeLV-A envelope glycoprotein and endogenous FeLV-related proviruses that are present in the feline genome. FeLV-A and FeLV-B share approximately 74% amino acid homology when the predicted amino acid sequences of the envelope glycoproteins are compared. FeLV-C arises as the result of point mutations in the FeLV-A envelope glycoprotein and thus these subgroups share approximately 93% homology when comparing the envelope glycoproteins. The three subgroups show distinct in vitro tropisms; FeLV-A is capable of infecting feline cells only (ecotropic), whereas both FeLV-B and FeLV-C display a broader tropism and are capable of infecting both feline and non-feline cells (amphotropic). Infection with FeLV-C leads to PRCA, an anaemia that is characterised by the defective BFU-E to CFU-E maturation (see figure 1.6 for a diagram of erythroid maturation). It is hypothesised that the onset of PRCA is due to the loss of the normal function of the protein encoded by the FeLV-C receptor gene, FLVCR (Dean G.A. et al., 1992). The aim of the studies described in this thesis was to examine the role of the virus-receptor interaction in the development of PRCA.

7.2 Elucidation of the envelope glycoprotein region involved in altered *in vitro* cell tropism of FeLV-C

When the amino acid sequences of the A and C subgroups of FeLV are compared, most of the differences between the two sequences are located in the N-terminal region of the protein, an area referred to as the first variable region, the "VRA" region, (Riedel N. et al., 1988; Rigby M.A., 1989; Brojatsch J. et al., 1992; Bae Y. et al., 1997).

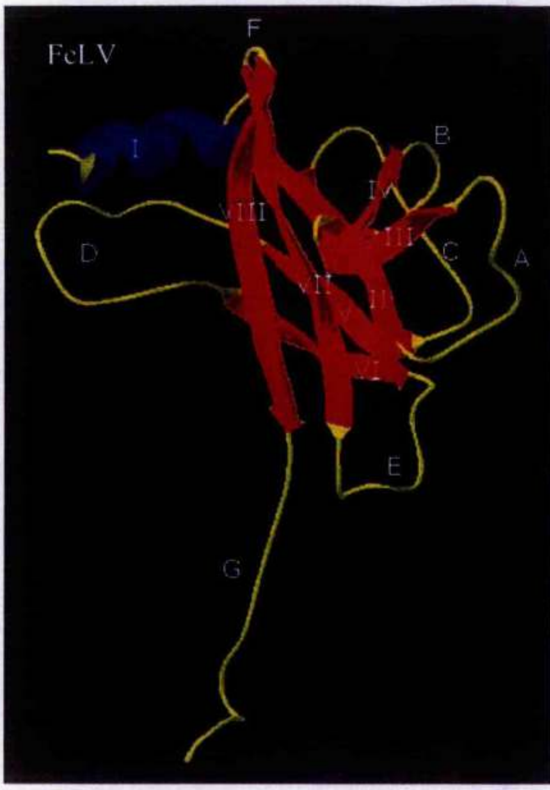
The predicted structure of the FeLV Env protein (based on the Friend MuLV *env* structure (Fass D. et al., 1997)) suggested that the VRA region is located on the surface of the envelope glycoprotein, consistent with this region being a component of the receptor binding domain. When the clones that were characterised during this study were compared, the majority of the mutations observed were located in the VRA region of the envelope glycoprotein, however there were additional clusters of mutations in distinct areas of the entire envelope glycoprotein. To date, no single amino acid or amino acid motif has been identified that confers an altered *in vitro* tropism. Brojatsch et al (Brojatsch J. et al., 1992) proposed that a single lysine to arginine substitution at position 64 in the published FeLV-A sequence would result in an altered *in vitro* host range. With the exception of the clones characterised from primary isolate FS246, this $K_{64}R$ substitution was absent from all clones identified in this thesis. These findings lead to the conclusion that it is not likely that this single lysine to arginine substitution is the sole determinant of an altered *in vitro* tropism and that there are either more amino acids or an entire region of amino acids involved in determining the *in vitro* tropism. Rigby et al (Rigby M.A. et al., 1992) have shown previously that the first variable region of the envelope glycoprotein, the VRA region, is a determinant of *in vitro* tropism. When the amino acid alignments were analysed (see figure 3.6 and 4.3), there were two motifs in the VRA region that were present only in the clones that displayed an *in vitro* tropism consistent with FeLV-C. Firstly, clones that had a tropism consistent with FeLV-A contained a stretch of six amino acids consisting of ^{61}TNVKHG^{66}, whereas clones that lacked this motif showed an *in vitro* tropism similar to FeLV-C, as shown by infection of the guinea pig cell line 104C1. Secondly, all subgroup C components contained a valine to tryptophan substitution at position 63 ($V_{63}W$). The presence of a tryptophan residue (a residue with a hydrophobic side chain) may create a hydrophobic pocket on the surface of the envelope glycoprotein. It has been shown previously by Zavorotinskaya et al (Zavorotinskaya T. and Albritton L.M., 1999) that the presence of a tryptophan residue at position 142 of the ecotropic MuLV envelope glycoprotein is critical for binding of the envelope glycoprotein to the receptor. Similarly, Wang et al (Wang Q-Y. et al., 2002) illustrated the necessity of a conserved tryptophan residue in the viral envelope glycoprotein of the Rous Sarcoma Virus-subgroup A in ensuring correct protein folding and subsequent binding of the receptor.

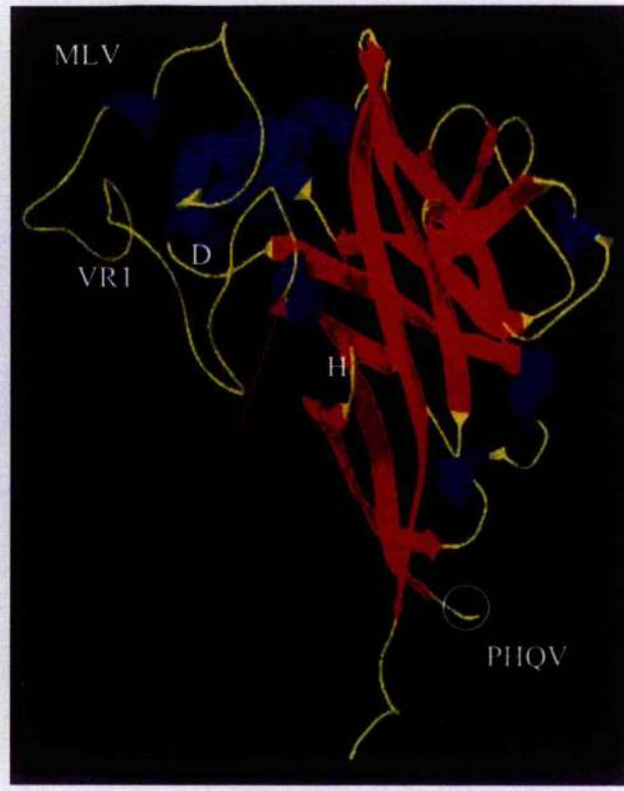
Consequently, the substitution of a valine to a tryptophan residue on position 63 in the clones, that displayed an *in vitro* tropism similar to FeLV-C, may be an important factor in the altered receptor recognition and *in vitro* tropism of these clones. The hypothesis that mutations outwith the VRA region are necessary for correct folding of the protein and in turn an altered *in vitro* tropism of the virus, is comparable to results described previously for human immunodeficiency virus (HIV) (Stamatatos L. and Cheng-Mayer C., 1993; Milich L. et al., 1993; Labrosse B. et al., 2001). The envelope glycoprotein gp120 of HIV contains five variable regions, V1 to V5. The third variable region of gp120 (the V3 loop) contains thirty-five amino acids and displays a high degree of sequence variability between isolates. An altered HIV tropism is linked to changes in the V3 loop, in particular amino acid changes flanking the GPGR-motif in this loop. This motif is predicted to be located on the periphery of the envelope glycoprotein, thus the location of this motif is comparable to the predicted location of the VRA region of the FeLV envelope glycoprotein. The amino acid composition within the V3 loop determines the overall structure and presentation of the loop. In turn, several regions including the V3 loop determine the overall structural conformation of the envelope glycoprotein gp120. The Env conformation defines the extent and type of post binding structural changes that lead to viral entry. Hence, changes in one region of the envelope glycoprotein can affect the function of other regions of the Env and the biological phenotype of an isolate is determined by the overall structure of the envelope glycoprotein and not by a single region (Stamatatos L. and Cheng-Mayer C., 1993; Milich L. et al., 1993; Labrosse B. et al., 2001). Similar findings have been described for the envelope glycoprotein gp120 of feline immunodeficiency virus (FIV). The Env of FIV contains 5 variable regions and the third variable region (V3), which is believed to be a determinant of *in vitro* tropism and receptor recognition, contains 62 amino acids. Verschoor et al (Verschoor E.J. et al., 1995) have shown that a single mutation in the V3 loop of the FIV Env resulted in an altered cell tropism of the virus. Furthermore, mutations outwith the V3 loop, in particular mutations located in the TM of the envelope glycoprotein, alter the *in vitro* tropism by facilitating changes in the conformational structure of the envelope glycoprotein (Vahlenkamp T.W. et al., 1997) Thus, although a region of the protein may be necessary for the function of the protein, other regions of the molecule are involved in maintaining the proper folding of the protein. This may explain the presence of clusters of mutations throughout the FeLV envelope glycoprotein; these mutations are essential to counteract mutations in other regions of the Env and thus result in maintaining the proper folding of the protein.

The importance of amino acid residues in distinct regions of the Env to maintaining structural integrity is underlined by the Envs that were not recognised by the monoclonal antibody (mAb) 6-15. This monoclonal antibody 6-15 is directed against the FeLV envelope glycoprotein and all three subgroups are recognised by this antibody. Two possible explanations for the *env* clones no longer being recognised by mAb 6-15 are 1) the mutations were located within the antibody binding site, resulting in the binding site being destroyed and 2) the mutations were located outwith the putative antibody recognition site and resulted in an altered conformational structure of the envelope glycoprotein and thus disrupted or masked the antibody binding site. As the mAb 6-15 and its antibody binding site remain to be characterised in full, the underlying mechanism behind the inability of the antibody to recognise the *env* clones cannot be confirmed. However, as the mAb 6-15 recognises all three FeLV subgroups, it is likely that the antibody binding site is located in a highly conserved region of the envelope glycoprotein. When the amino acid sequences of the Envs that were no longer recognised by the antibody were compared, it was evident that the mutations that were present in these clones were scattered throughout the envelope glycoprotein and were not restricted to a single region. It is therefore likely that the mutations in the envelope glycoprotein induced an altered conformational structure and thus disrupted the mAb 6-15 binding site.

In order to comprehend more fully the impact that mutations in the VRA region of the envelope glycoprotein could have on the overall structure of the protein, FeLV *env* structures were generated as described in section 2.2.13. As can be seen in figure 7.1, a structure for FeLV gp70 can be modelled successfully based on the MuLV structure. Figure 7.2 illustrates a comparison of the predicted structures of the variable regions of a selection of FeLV-A and FeLV-C Envs. The natural isolates that showed an *in vitro* tropism comparable to FeLV-C/Sarma indeed also showed a similar structure, whereas FeLV-A/Glasgow-1 and isolate FS246 showed structures comparable to one another, but distinct from the structures displayed by FeLV-C/Sarma. As was expected, the structure of FeLV-B/Gardner-Arnstein was highly divergent from both the FeLV-A/Glasgow-1 and the FeLV-C/Sarma envelope glycoprotein structures. Thus, the structural predictions suggest that while several of the FeLV-C may have a similar structure, FS246 provides an exception and thus no single structural element can be identified that confers the subgroup C phenotype.

When a spacefilling structure was prepared of FeLV A and C envelope glycoproteins (figure 7.3), it was evident that the variable regions were in close proximity. Thus mutations in the VRA region may need to be counteracted by mutations elsewhere in the envelope glycoprotein in order to maintain the correct three-dimensional structure and ultimately the function of the protein. Recently, the structure of the Receptor Binding Domain of FeLV-B (FeLV-B RBD) was solved by Barnett et al ((Barnett A.L. et al., 2003), PDB accession code: 1LCS). When the FeLV *env* structure predictions generated in this study were compared with the structure of FeLV-B RBD, it was evident that the structures were very similar, strengthening the validity of the structure predictions based on MuLV and the conclusions that were drawn from these structures and the locations of the variable regions in the envelope glycoprotein.






```
           1                                                in MLV 50
FS246new   MESPTDPNPS KDKTPSWNLV FLLGILITID IGMANPSPHQ IYNVTWVITN
   FeLVA   MESPTHPKPS KDKTLSWNLA FLVGILFTID IGMANPSPHQ IYNVTWVITN
FZ215new   MESPTHPKPS KDKTLSWDLV FLVGILFTID IGMASPSPHQ IYNVTWVITN
  SarmaC   MESPTHPKPS KDKTFPWNLV FLVGILFQID MGMANPSPHQ VYNVTWVITN
FY981new   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~MANPSPHQ IYNVTWVITN

           51 H                    in MLV          VR1       100
FS246new   VQTTPKANAT SMLGTLTDAY PTLHVDLCDL VGDTWEPIVL NPTSVRFGAV
FeLVA      VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VGDTWEPIVL NPTNV   KHG
FA27C      VQTNTQANAT SMLGTLTDAY PTLHVDLCDL VGNTWEPIVP DL      RGW
FZ215new   VQTNTQRNAT SMLGTLTDAY PTLHVDLCDL VGDTWESMAP DP      RSW
SarmaC     VQTNSRANAT SMLGTLTDAY PTLYVDLCDL VGDTWEPIAP DP      RSW
L39505     VQTNTPANAT SMLGTLTDAY PTLHVDLCDL VGDTWESIVL         RGW
FY981new   VQNNTQANAT SMLGTLTDAY PTLHVDLCDL VGNTWEPIVP PQT     VRR

                           I           II          A    III
           101                                                  150
FS246new   LLSSSPKYGC KTTDRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
FeLVA      ARYSSSKYGC KTTDRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
FA27C      ASYSSSKYGC KTADRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
FZ215new   ARYSSSTHGC KATDRKKQQQ TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
SarmaC     ARYSSSTHGC KTTDRKKQQQ TYPFYVCPGH APSMGPKGTY CGGAQDGFCA
L39505     AHYSSSKYGC KTTDRKKQLQ TYPFYVCPGH APSFGPKGTH CGGAQDGYCA
FY981new   AHYSSSTHGC KTTDRKKQQR TYPFYVCPGH APSLGPKGTH CGGAQDGFCA
                     MLV VRA
```

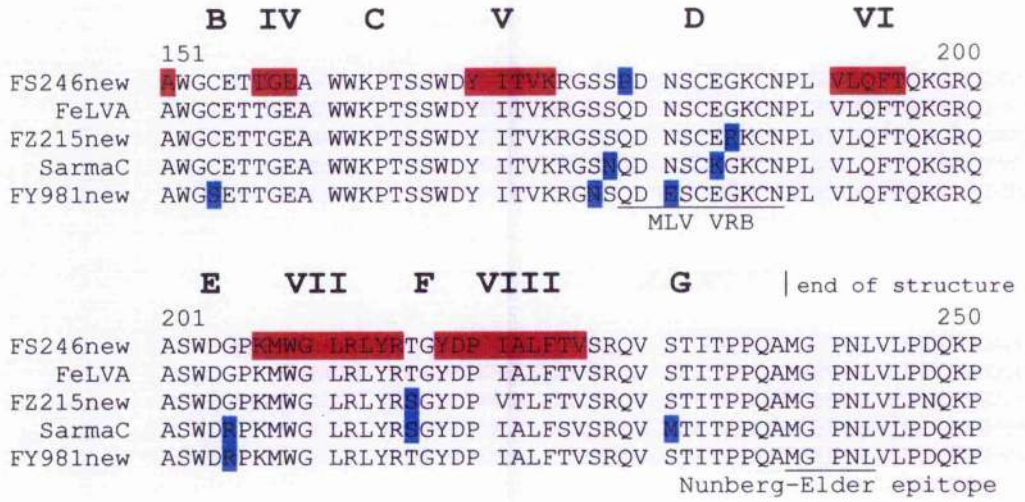


Figure 7.1. Comparison of the predicted structures of the envelope glycoproteins of the feline and murine leukaemia viruses.

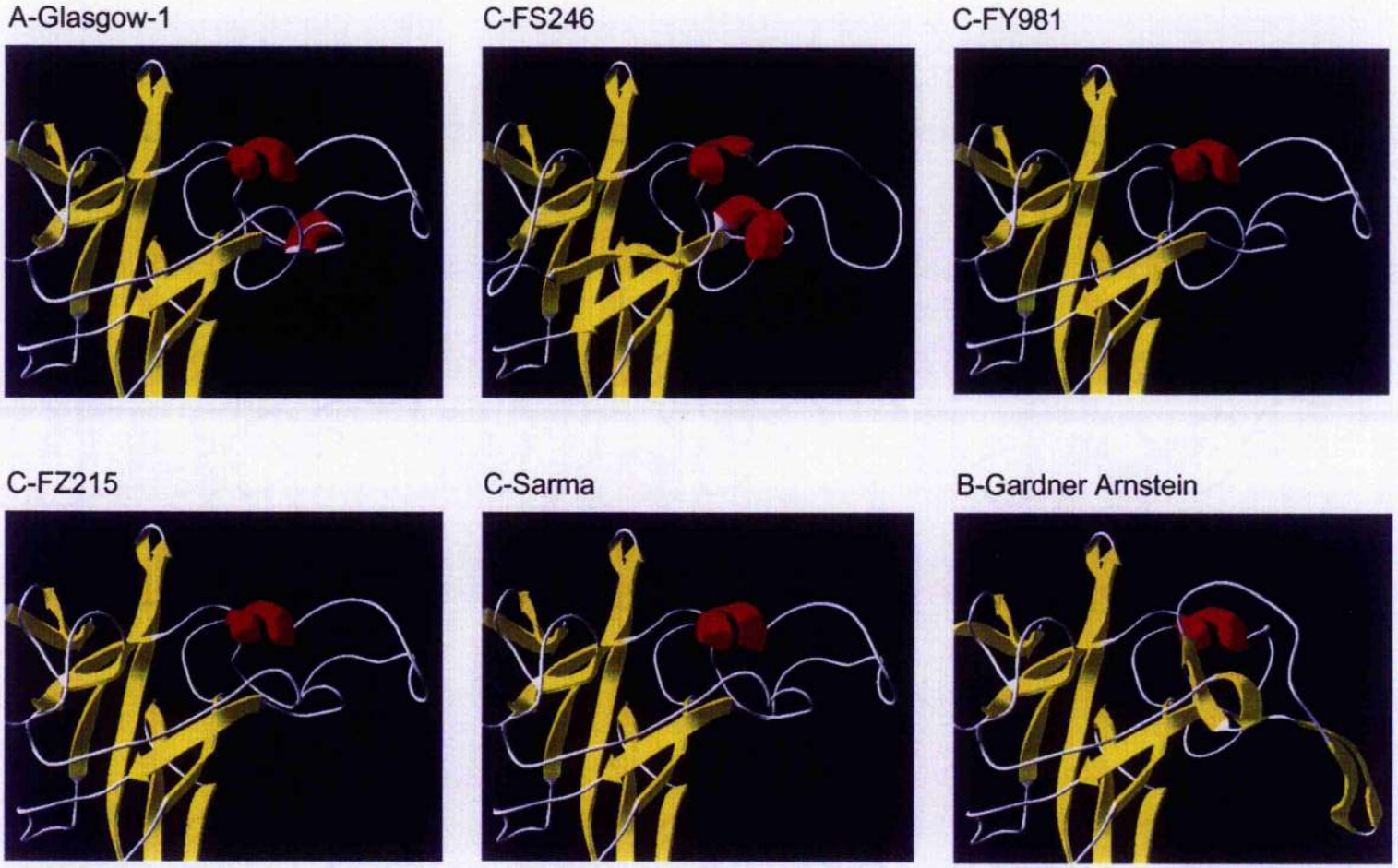


Figure 7.2: *Predicted ribbon structure of the FeLV envelope glycoprotein gp70 of a number of primary isolates. The structure of the envelope glycoprotein is based on the previously published structure of the Friend MuLV gp70 (Fass et al).*

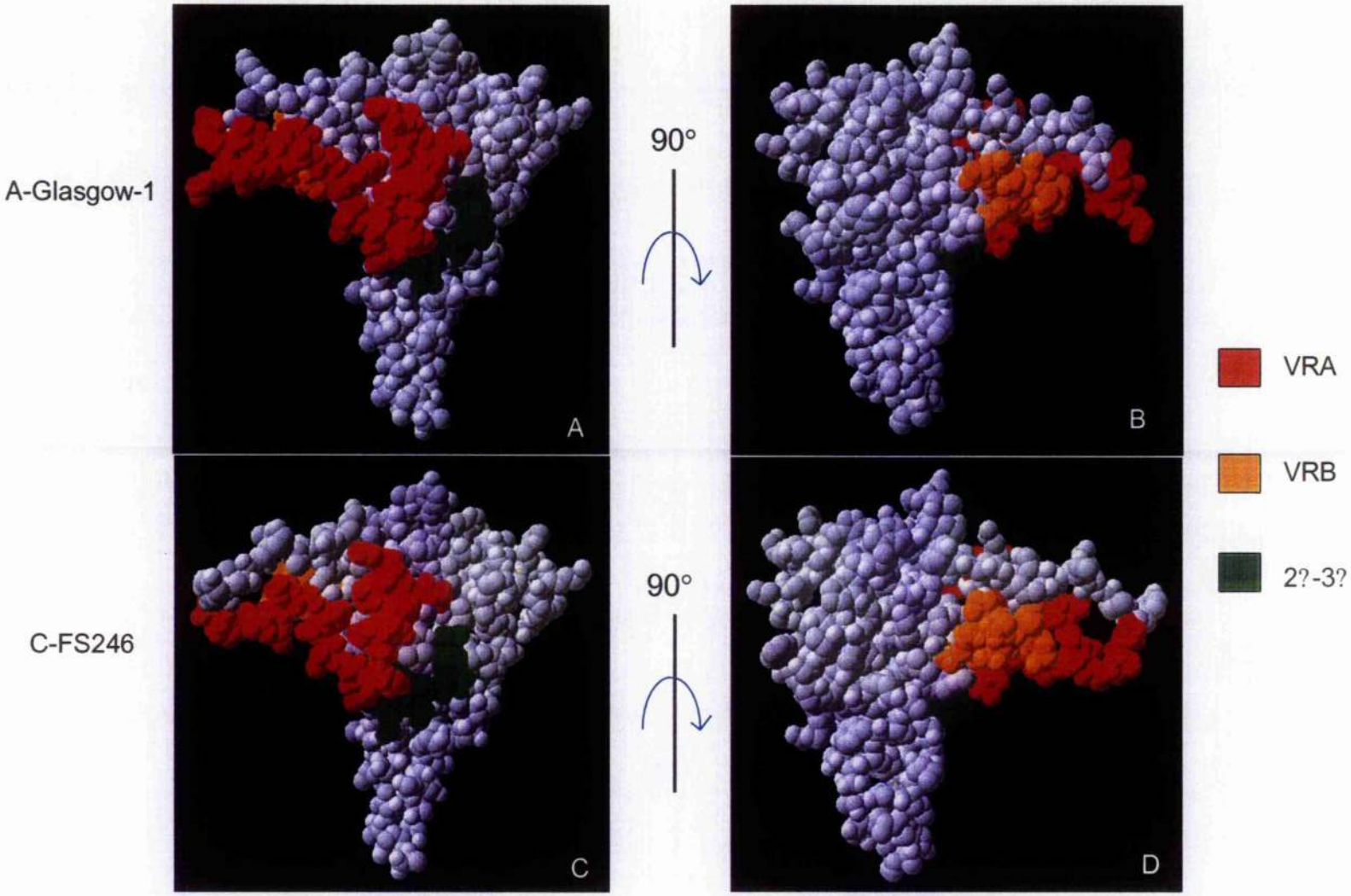


Figure 7.3: *Predicted ball structure of the FeLV envelope glycoprotein gp70 of FeLV-A/Glasgow 1 and primary isolate FS246, which has an in vitro tropism consistent with an FeLV-C phenotype. The structure of the envelope glycoprotein is based on the published structure of the Friend MuLV gp70 (ref Fass et al).*

7.3 Expression of the feline FeLV-C receptor, feFLVCR

In order to characterise in more detail the role of the feline FeLV-C receptor (feFLVCR) in the development of PRCA, the expression of this gene was studied in a panel of both haematopoietic and non-haematopoietic tissues. Due to relatively low levels of mRNA expression in the tissues screened, it was not possible to detect the feline FeLV-C receptor by conventional northern blot analysis. However, the RT-PCR results indicated that the receptor was expressed widely, albeit at low levels. Furthermore, there appeared to be no single tissue where the gene was either consistently expressed to high levels or consistently not expressed. These findings are in contrast with the findings described by Tailor et al (Tailor C.S. et al., 1999) when screening a pre-made poly A(+) northern blot with the human FeLV-C receptor huFLVCR. Tailor et al concluded that huFLVCR was expressed preferentially in haematopoietic tissues, consistent with the hypothesis that the development of PRCA is due to the FeLV-C receptor being expressed at a high level on cells that are involved in erythroid maturation and that this lineage of cells may be particularly sensitive to the cytopathic effect of FeLV-C infection (see also Dean et al, 1992).

It is likely that the anaemia that develops after infection with FeLV is the result of the loss of the normal function of the gene encoding for FLVCR. Perkins et al (Perkins C.P. et al., 1997) have reported similar findings with the ecotropic MuLV receptor mCAT-1. The mCAT-1 gene encodes a basic amino acid transporter and also acts as the receptor for ecotropic MuLV (Albritton L.M. et al., 1989). When a mutation occurs in the mCAT-1 gene, the normal transporter function of the gene is abolished, resulting in mice that develop anaemia and are approximately 25% smaller than their littermates. Yamaguchi et al (Yamaguchi S. et al., 2003) studied the expression of the mCAT-1 gene in a panel of both haematopoietic and non-haematopoietic tissues and found that the gene is expressed widely, with the exception of the heart, liver and muscle tissues. Thus, impairment of FLVCR function may manifest as a lineage-specific defect, despite the receptor being expressed widely in the cat. Similarly, Wang et al (Wang H. et al., 1992) described the effect infection with ecotropic MuLV has on the amino acid transport in the infected cell through the viral receptor.

It was found that when cells were chronically infected with ecotropic MuLV, the uptake and outflow of L-arginine was reduced significantly, whereas when extracellular gp70 was added to cells, the transport of amino acids was not altered. These results indicate that binding of the virus to the cellular receptor does indeed affect the normal function of the molecule.

The main characteristic of PRCA in the cat is the disrupted maturation of BFU-E to CFU-E cells, possibly by impairing the cell surface expression or function of the receptor (Young N.S. et al., 2000). When feline bone marrow populations were either enriched for or depleted of cells that are part of the erythroid lineage and subsequently subjected to RT-PCR analysis, the results indicated that FLVCR receptor expression may be higher in cells of the erythroid lineage in comparison to populations of cells that were depleted of cells of the erythroid lineage. Thus, it is possible that although FLVCR expression is widespread, erythroid cells may be extremely sensitive to perturbations in its the normal cellular function of the molecule. In contrast, in cells of other lineages, the function of FLVCR may be dispensable.

7.4 Identification of the FeLV-A receptor

To date, the FeLV-A receptor remains to be identified and characterised, although a putative binding protein has been identified previously (Ghosh S.K. et al., 1992). However, the 70 kDa molecule, identified by Ghosh et al, remains to be characterised fully. The method utilised in this thesis, has been used successfully previously in the identification of CCR5 as an HIV receptor (Deng H. et al., 1997) and the FeLV-C receptor (Tailor C.S. et al., 1999). Thus, the problems encountered whilst attempting to isolate the FeLV-A receptor are unlikely to be due to a flawed experimental approach, but more likely due to other factors, such as the molecule being toxic to the cells or the libraries screened not being abundant in the receptor cDNA. Comparable problems were encountered when Zhang et al (Zhang C.C. et al., 1999) attempted to stably express the Estrogen Receptor (ER) in the human cell line HeLa. It was found that the receptor was highly toxic to cells and thus the transfected HeLa cells attempted to remove the toxic molecules as soon as possible.

7.5 Future work

7.5.1 Envelope glycoprotein studies

Two features of the FeLV-C envelope glycoprotein, that appeared to be involved in altering the *in vitro* tropism of FeLV, were identified. They were 1) the lack of a ^{61}TNVKHG^{66} motif and 2) a $V_{63}W$ substitution. Furthermore, a $D_{51}N$ mutation appeared to be involved in allowing the infection of porcine cells in clones that were designated the A component of the primary isolate. In order to confirm the role of these features in infection, primers for each motif will be designed and subsequently used to generate clones displaying each feature. Thus, the ^{61}TNVKHG^{66} motif will be introduced into the envelope glycoprotein of FeLV-C/Sarma by PCR-based mutagenesis. The clones obtained will then be used to produce FeLV *env*/MuLV *gag-pol* pseudotypes and their *in vitro* tropism compared with the prototypic A and C strains by plating onto a panel of both feline and non-feline cells lines. Similarly, the $V_{63}W$ substitution will be introduced into the envelope glycoprotein of FeLV-A/Glasgow-1, in order to assess whether it can alter the subgroup of the virus from A to C. Finally, the $D_{51}N$ mutation will be introduced into the envelope glycoprotein of FeLV-A/Glasgow-1, to assess whether the tropism of the virus can be extended to the porcine cell line ST Iowa.

In order to test receptor usage of the novel clones described in this thesis, pseudotypes derived from the novel *env* clones, can be plated onto cells stably expressing the FeLV-C receptor feFLVCR to determine if these novel clones are capable of using feFLVCR efficiently. Similarly, the *env* clones that contained the $D_{51}N$ mutation and were capable of infecting the porcine cell line ST Iowa can be plated onto cells stably expressing either Pit1 or Pit2 to establish the receptor usage of these clones.

7.5.2 Expression of the feline FeLV-C receptor

Preliminary results from the RT-PCR studies on the bone marrow samples collected from a number of cats indicated that the FeLV C-receptor is expressed preferentially in bone marrow populations that have been enriched for cells that are involved in the erythroid maturation. In order to confirm these findings, more bone marrow populations will be collected and subsequently screened using primers encoding feFLVCR, as described in chapter five of this thesis. If similar results are obtained from the additional bone marrow samples, these results will strengthen the hypothesis that PRCA is the result of disruption of the normal function of the receptor gene in cells in which FLVCR is essential for growth.

In addition, the expression of feFLVCR can also be examined using quantitative real time PCR. The main advantage of this technique is that it is a more quantitative method than the RT PCR described in this thesis. Therefore, the amount of feFLVCR present in each tissue can be established and thus give a clearer understanding of the distribution of the receptor in each tissue.

7.5.3 Identification and characterisation of the FeLV-A receptor

Thus far, the FeLV-A receptor remains to be identified. In order to increase the probability of identifying this receptor, there are four possible future directions:

- Firstly, additional cell lines should be screened for suitability for use in library screening. It was found that when attempts were made to prepare murine 3T3 cells that stably expressed the human FeLV-C receptor, these cells were not able to express the gene for any length of time, despite a G418 selection step. In contrast, when the same construct was used to prepare stably expressing CHO cells (a Chinese hamster ovary cell line), it was found that the cells were indeed able to express the gene in the presence of the antibiotic. This would suggest that ectopic expression of the FLVCR molecule might be toxic to some cell lines.
- Secondly, a cDNA library will be constructed from a cell line that is highly susceptible to FeLV-A infection (e.g. the feline cell line FEA). As FEA cells are highly susceptible to FeLV-A infection, the mRNA encoding the FeLV-A receptor is anticipated to be present at high levels in these cells.

Thus increasing the probability of identifying the receptor gene will be increased. The new library will subsequently be screened as described in chapter six.

- Thirdly, it is likely that, although the puromycin resistant cells no longer express the putative FeLV-A receptor, the gene encoding the molecule may still be present in the cellular genomic DNA. Thus, DNA prepared from puromycin resistant cells will be subjected to PCR using primers flanking the insertion sites of the library in order to amplify a putative FeLV-A receptor.
- Finally, the cDNA library will be constructed in a vector in which expression of the cDNA is inducible, for example the "Tet-on" system (BD Biosciences, Oxford, United Kingdom). In such an antibiotic regulated system, the expression of the gene can be either switched on or off in the presence or absence of the antibiotic. Thus, the period during which the gene is being expressed can be limited, enhancing the probability that the receptor is still being expressed on the cell surface and not having been removed due to its toxicity.

When considering the possible directions in which future research may proceed, it is important that the results obtained from the screening of the bone marrow samples are repeated, in order to confirm the cautious conclusions described in chapter five. Similarly, the role of the motifs described in chapters three and four in altering the *in vitro* tropism should be elucidated. If these motifs are indeed crucial in the altered *in vitro* tropism of the novel clones, it may be possible to identify the receptor binding domain on the viral Env protein and thus understand more fully the mechanism by which the virus exerts its devastating effect on erythroid maturation.

APPENDIX A

MATERIALS

Materials in regular use, such as equipment, general reagents and solutions are described in this section. A full list of manufacturers is supplied in appendix B.

A.1 CELL CULTURE MATERIALS

A.1.1. Plasticware

Tissue culture flasks, 100mm cell culture dishes and 150mm cell culture dishes were supplied by Corning incorporated (Corning, NY).
Cryotubes were supplied by Nunc (DK 400, Roskilde, Denmark).
Falcon conical centrifuge tubes were supplied by Becton Dickson U.K. Ltd. (Oxford, United Kingdom).

A.1.2. Solutions, media and supplements

All solutions and media for cell culture were supplied by Invitrogen ™ Life technologies (Paisley, United Kingdom), unless otherwise stated.

A.1.2.1. Media

All media were supplied as sterile solutions and stored at +4°C.

Dulbecco's Modified Eagle's Medium (DMEM).

A.1.2.2. Supplements

Foetal Calf Serum (FCS), (Perbio, Cheshire, United Kingdom): virus screened, mycoplasma screened. FCS was heat inactivated at 56°C for 30 minutes, then stored in 50ml aliquots at -20°C until use.

L-glutamine: supplied as a 100X stock solution containing 200mM L-glutamine. This was stored in 5ml aliquots at -20°C until use.

Penicillin/streptomycin: supplied as a 100X stock solution containing 10.000units penicillin and 10.000μg streptomycin per millilitre. This was stored in 5ml aliquots at -20°C until use.

Trypsin-EDTA: supplied as 10X liquid, stored at -20°C. This was diluted 1:10 in versene prior to use and stored at +4°C.

A.1.3. Cell lines

The cell lines used routinely are described in more detail in section 2.2.1.2.4.

FEA	This cell line is a derivative of the Crandell feline kidney (CrFK) cell line. HO6T1 cells were generated by transfecting CrFK cells with the plasmid pHO6T1 carrying an activated Ha-ras oncogene. H06T1 cells are susceptible to infection with all three FeLV subgroups.
293T (originally referred to as 293tsA1609neo)	This is a human embryonic kidney line transformed with adenovirus E1A and carrying a temperature sensitive T antigen. These cells are susceptible to infection with all three FeLV subgroups.
HeLa cell line	This is a cell line that is derived from a human cervical carcinoma. Its cells grow in an epithelial like monolayer. These cells are susceptible to FeLV-B and FeLV-C infection, but not to FeLV-A infection.
Madin-Darby canine kidney (MDCK) cell line	A cell line derived from the kidney of a normal, adult female cocker spaniel. The cells grow in an epithelial like monolayer.
Mv1Lu cell line	This cell line was derived from the lungs of several near full term foetuses of the mink, *Mustella vison*, and grows in an epithelial like monolayer.

ST Iowa cell line	A cell line derived from the testes of an eighty to ninety days old porcine foetus, this cell line grows in an epithelial like monolayer.
NIH/3T3	A cell line derived from Swiss albino mouse fibroblasts. These cells are, like all murine cell lines, not susceptible to FeLV infection.
104C1	104C1 are a guinea pig cell line and were a kind gift from Prof. D.E. Onions (Q1 biotech, Glasgow, United Kingdom). These cells are susceptible to FeLV-C infection alone.

A.2 RADIOCHEMICALS

[α-^{32}P]-dCTP (specific activity of > 10 mCi/ml at reference date) was supplied by Amersham Life Science (Bucks, United Kingdom) and stored at -20°C until use.

A.3 GENERAL CHEMICALS

Chemicals used were of analytical or ultrapure quality and were supplied by Sigma-Aldrich Company Ltd. (Dorset, United Kingdom), BDH Ltd (Poole, United Kingdom) unless otherwise stated.

Caesium Chloride: Roche applied science (East Sussex, United Kingdom).

Agarose: Life technologies (Paisley, United Kingdom).

Bacterial agar and *tryptone:* Oxoid (Basingstoke, United Kingdom).

A.4 COMPLETE KITS

QIAquick® Gel Extraction Kit supplied by QIAGEN ltd. (Crawley, United Kingdom).
The QIAquick columns comprise of a silica-gel membrane that binds the DNA in the presence of high salt and a pH of ≤7.5. The columns are washed with an ethanol-containing wash buffer, which efficiently removes impurities. The DNA is then eluted with a low salt elution buffer, pH 8.5, or water.

QIAprep® Spin Miniprep Kit supplied by QIAGEN ltd. (Crawley, United Kingdom).
The method used is a modification of the alkaline lysis method first described by Birnboim and Doly in 1979 (Birnboim H.C. and Doly J., 1979). Bacteria are lysed, after which the lysate is neutralised. This lysate is applied to a column that contains a silica-gel membrane. Similar to the QIAquick® Gel Extraction Kit, the silica-gel membrane binds the DNA in the presence of high salt and a pH of ≤7.5, after which the DNA is eluted with a low salt elution buffer, pH 8.5, or water.

Endofree™ Plasmid Maxi Kit supplied by QIAGEN ltd. (Crawley, United Kingdom).
The preparation and purification of DNA using this method consists of four steps, namely lysing of the bacteria, removal of endotoxins, binding of the DNA to a column containing an Anion-Exchange resin, washing to remove impurities and subsequent elution of the purified DNA. Similar to the QIAprep® Spin Miniprep Kit, this is a modification of the alkaline lysis method described by Birnboim and Doly in 1979 (Birnboim H.C. and Doly J., 1979). As with the previously described kits, efficient binding of the DNA is dependent on the correct salt concentration and pH. As an additional step, endotoxins are removed by filtering the lysate through a QIAfilter cartridge and subsequent incubation with an endotoxin removal buffer, which prevents endotoxin molecules to bind to the resin in the columns.

QIAamp® DNA blood mini kit supplied by QIAGEN ltd. (Crawley, United Kingdom).
This kit is used to prepare DNA from a range of tissues, fluids and cultured cells. The sample is lysed using proteinase K and buffer AL. The lysed cells are added to the silica gel membrane-containing column. Similar to QIAGEN's other columns, the pH and salt concentration in the lysate ensure the binding of DNA only to the columns.

The bound DNA is washed twice with buffer AW1 and AW2 subsequently to remove residual contaminants. The DNA is subsequently eluted in buffer AE.

TOPO TA cloning kit supplied by Invitrogen NV (Leek, The Netherlands).
The vector used in this kit is linearized and contains single 3' thymidine (T) overhangs to allow optimal ligation of the PCR products. The PCR products contain 3' deoxyadenosine (A) overhangs due to the use of Taq polymerase which has a non-template-dependent terminal transferase activity and thus adds these single 3' deoxyadenosine (A) overhangs to the PCR products. Due to the overhangs on both vector and PCR product, the ligation takes place at room temperature in 5 minutes. Subsequent transformation of chemically competent cells allow the analysis of the obtained colonies.

ABI Prism® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit, version 2.0 was supplied by Applied Biosystems (Warrington, United Kingdom).

RadPrime DNA Labelling System was supplied by Invitrogen NV (Leek, The Netherlands).
This system is designed for rapid labelling and preparation of high specific activity ^{32}P-labelled probes. Random octamer primers are annealed to the DNA template after which it is extended in the presence of Klenow fragment and a α-^{32}P-dCTP. The prepared DNA probe is highly specific and can be used to detect both DNA and RNA.

FeLV p27 ELISA kit supplied by IDEXX Laboratories Limited (Buckinghamshire, United Kingdom).
This ELISA (Enzyme Linked ImmunoSorbent Assay) kit uses the indirect ELISA principle: a well of the plate is coated with the antigen of interest (in this case FeLV p27) after which the coated plates are incubated with test solutions and the solutions to be tested. Unbound antibodies are washed away after which the coated plates are incubated with an antibody-enzyme conjugate and a substrate is added to the wells and colour change is observed.

Expand reverse transcriptase kit was supplied by Roche applied science (East Sussex, U.K.).

Gene Amp PCR Core Reagents (Roche Molecular Systems Inc, New Jersey, USA) supplied in the U.K. by Applied Biosystems Ltd.

A.5 BACTERIAL STRAINS

E.coli INVαF' cells (Invitrogen NV, Leek, The Netherlands): F' *end*A1 *rec*A1 *hsd*R17 (r_k^-, m_k^+) *sup*E44 *thi*-1 *gyr*A96 *rel*A1 ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *deo*R$^+$ λ$^-$.

The above strain has the ϕ80*lac*ZΔM15 marker, enabling blue-white screening by α-complementation of β-galactosidase encoded by vector DNA (e.g. pCR™ II). The genotypes *end*A1 and *hsd*R17 give improved quality of miniprep DNA; *rec*A1 denotes recombination negative, recommended for stable replication of high copy number plasmids.

A.6 DNA

Plasmid, molecular weight marker and oligonucleotide DNAs were stored at -20°C until required.

A.6.1 Plasmids

* VR1012 (Hartikka J. et al., 1996)
This plasmid DNA expression vector has a modified pUC backbone and contains a kanamycin resistance gene to allow for bacterial growth selection. Furthermore, a CMV early promoter is present to regulate the expression of the inserted gene.

* pHit60 (Soneoka Y. et al., 1995)
The plasmid pHit60 is a CMV-driven MuLV *gag-pol expression plasmid that contains the SV40 ori (SV40 origin of replication). When plasmids containing the SV40 ori are introduced into cells expressing the SV large T antigen (e.g. the 293T*

cell line), the copy number of these plasmids is increased and therefore a higher uptake of plasmid is likely.

* MFG-lacZ (Dranoff G. et al., 1993)
This retroviral vector encodes the E. coli LacZ gene. The lacZ gene product ß-galactosidase catalyses the hydrolysis of the substrate X-gal to produce a blue precipitate, and thus this vector can be used as a marker to ascertain the efficiency of the transfection.

* pBabePuro (Morgenstern J.P. and Land H., 1990)
This retroviral vector contains a Moloney MuLV derived LTR promoter to drive the expression of the genes inserted. A drug resistance marker (puromycin) is present to allow for selection in puromycin containing medium.

A.6.2 Molecular Size Standards

1 kb DNA Ladder (size range 500-10.002 bp) was supplied by New England Biolabs (U.K.) Ltd. (Hitchin, United Kingdom).

0.24-9.5 Kb RNA ladder (size range 240-9.490 bp) was supplied by Invitrogen ™ Life technologies (Paisley, United Kingdom).

A.6.3 Oligonucleotide Primers

Oligonucleotides were supplied by MWG biotech (MWG biotech AG, Ebersberg, Germany).

IRD41 labelled primers for use with the LICOR model 4000 automated sequencer were supplied by Hybaid U.K. Ltd.

The primers used for the ABI PRISM 3100 genetic analyser were supplied by MWG biotech (MWG biotech AG, Ebersberg, Germany).

A.7 ENZYMES

All enzymes and their reaction buffers were stored at -20°C, being removed immediately before use.
Restriction enzymes and their associated reaction buffers were supplied by Gibco BRL. (Invitrogen Lifetechnologies ™, Paisley, United Kingdom).
T4 DNA Ligase was provided by Gibco BRL or Invitrogen (as part of the TA Cloning Kit).

A.8 BUFFERS, SOLUTIONS AND GROWTH MEDIA

A.8.1 Water

Tissue culture grade distilled water was supplied by Gibco BRL. Ultrapure water (for procedures involving recombinant DNA, PCR etc.) was provided by a Millipore Q50 water purification system (Millipore (U.K.) Ltd., Watford, United Kingdom). A Millipore RO10 system was used to supply water for preparation of general solutions and media.

A.8.2 Antibiotics

Ampicillin (Sigma Chemical Company, Dorset, United Kingdom): Solution of 100mg/ml prepared in dH_2O, filter sterilised and stored in aliquots at -20°C until use.

Kanamycin (Sigma Chemical Company, Dorset, United Kingdom) Solution of 20mg/ml was prepared in dH_2O, filter sterilised and stored in aliquots at -20°C until use.

A.8.3 Buffers and solutions

A.8.3.1 Molecular Biology Solutions

1x TE	10 mM Tris.Cl, 1mM EDTA pH 7.5.
1x TBE	90 mM Tris-borate, 2 mM EDTA.
2x HBS	280 mM NaCl, 50 mM HEPES, 1.5 mMNa_2HPO_4, pH to 7.05 with 0.5 M NaOH. Filter sterilise with 0.4 micron filter.
10x MOPS	0.2 M [3-(N-morpholino)propanesulfonic acid], 50 mM NaAc, 10 mM EDTA, pH to 7.0 with 0.5 M NaOH and make up to 1 litre with milliQ water.
20x SSC	3 M NaCl, 0.3 M Tri-Na acetate.
100x Denhardts	2% Ficoll 400; 2% polyvinylpyrrolidine; 2% BSA, Filter sterilise with 0.4 micron filter.
Buffer P1	50mM TRIS-Cl, pH8.0; 10mM EDTA.
Buffer P2	200mM NaOH; 1% SDS. (Prepared immediately prior to use).
Buffer P3	3.0M KOAc, pH5.5.
Buffer QBT	750mM NaCl; 50mM MOPS, pH7.0; 15% isopropanol; 0.15% Triton X-100.
Buffer QC	1.0M NaCl; 50mM MOPS, pH7.0; 15% isopropanol.
Buffer QN	1.6M NaCl; 50mM MOPS, pH7.0; 15% isopropanol.
Buffer TE	10mM TRIS, pH8.0; 1mM EDTA.

RNA loading buffer	500µl formamide; 166µl formaldehyde; 100µl 10x MOPS; 234µl dH_2O.
RNA loading dye	5ml glycerol; 1ml 10x MOPS; 4ml dH_2O; few grains of bromophenolblue. Store at -20°C.
Ethidium Bromide	Made to a working solution of 10mg/ml in dH_2O, stored at room temperature away from light.
10X DNA Gel Loading .Buffer	0.25% Xylene Cyanol FF, 15% glycerol
Sequencing gel (for use with LICOR 4000)	27.7gr urea, 5.3ml Long Ranger Solution, 7.9ml 10X TBE buffer made up to 65ml with dH_2O to which 440µl 10% APS (made immediately prior to use) was added.
10x TBE, for use with the LI-COR 4000	53.9g Tris base, 27.5g boric acid, 3.7g EDTA; to 500ml with dH_2O, stored at 4°C for a maximum of 2 weeks

A.8.3.2 Bacteriological Media

SOC Medium	2% Tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl2, 10mM MgSO4, 20mM glucose.
L-Broth	10g bactotryptone, 5 g yeast extract, 10g NaCl, pH to 7.0 with 0.5 M NaOH, make up to 1 litre. Autoclave.
L-agar	10g bactotryptone, 5 g yeast extract, 10g NaCl, pH to 7.0 with 0.5 M NaOH, make up to 1 litre and add 1.5% Bacto™ Agar. Autoclave.

A.8.3.3 Cell Culture Media

10%DMEM growth media	DMEM supplemented with 10%FCS, 2mM Glutamine, 100IU/ml penicillin and 100μg/ml streptomycin.
Freezing media	Appropriate cell culture media supplemented with an additional 10% FCS and 10% DMSO.
PBS	Phosphate buffered saline: Sodium chloride, 150 mM, and sodium phosphate, 150 mM, pH 7.2 at 25°C.
X-gal buffer	5mM potassiumferrocyanide, 5mM potassiumferricyanide, 2mM $MgCl_2$ in PBS, store at room temperature away from light.
X-gal	5-bromo-4-chloro-3-indolyl ß-D-galactopyranoside to a final concentration of 1mg/ml (w/v) in dimethylformamide, store at -20°C away from light.

A.9 EQUIPMENT

A.9.1. Major equipment

* Benchtop centrifuge: Beckman GS-6 centrifuge (Beckman Coulter, United Kingdom).
* Centrifuge: Beckman J2-21 centrifuge (Beckman Coulter, United Kingdom).
* Ultracentrifuge: Beckman L8-M ultracentrifuge (Beckman Coulter, United Kingdom).
* Microcentrifuge: Biofuge 13 (Heraeus Sepatech, Germany).
* Incubators for tissue culture: Leec Ltd. (Nottingham, United Kingdom).
* Water baths: supplied by Grant Instruments (Cambridge) Ltd. (United Kingdom).
* Spectrophotometer: Model DU640, Beckman (Beckman Coulter, United Kingdom).
* Vacuum dessicator: Freeze drier modulo (Edwards Engineering Corp, NJ, USA)
* Automatic Sequencing Apparatus: Licor Model 4000 sequencer, sequencing plates etc. - Licor Inc., Lincoln, Nebraska.
* ABI PRISM 3100 genetic analyser (Applied biosystems, Warrington, United Kingdom.)
* Automated Processor: Compact X4 supplied by Xograph, Wiltshire, United Kingdom.
* Pipette man (P20, P200, and P1000): supplied by Gilson Medical Electronics (Villiers-le-Bel, France).
* Automatic Sarpette: supplied by Sarstedt.
* Ultraviolet Transilluminator: supplied by UV Products Inc. (San Gabriel, CA).
* Autoradiography (Film) cassettes: with intensifying screens (Cronex), supplied by Amersham Biosciences (Uppsala, Sweden).
* PCR machine: geneamp PCR system 9700 (PE applied biosystems, Warrington, United Kingdom).
* Orbital incubator: Basic orbital incubator supplied by Sanyo Gallenkamp plc, Loughborough, United Kingdom.
* Hybridisation oven: Hybaid mini 10 supplied by Thermo Hybaid – UK, Ashford, United Kingdom
* Liquid scintillator counter: Beckman LS1801 supplied by Beckman Coulter, United Kingdom
* Powerpack: Bio-rad powerpac 3000 supplied by Bio-Rad Laboratories Ltd, Hertfordshire, United Kingdom.

* Gel electrophoresis tank: Bio-rad wide mini sub® cell GT supplied by Bio-Rad Laboratories Ltd, Hertfordshire, United Kingdom.

* Gel visualisation system: gel print 2000i supplied by MWG biotech AG, Ebersberg, Germany.

A.9.2 Consumables

* Screw top 1.5 ml eppendorf tubes, 0.5 ml and 1.5 ml eppendorf tubes were supplied by Treff AG (Degersheim, Switzerland).

* Pipette tips were supplied by Sarstedt.

* Syringes (two, five, 10, 20 and 50 ml) were supplied by Becton Dickinson.

* Round gel loading tips were supplied by Sorenson Bioscience Ltd.

* Filter tip pipette tips (30 µl and 200 µl) were supplied by Rainin Instrument Co. (Woburn, MA); for use in setting up PCR reactions and RNA preparation.

* RNAse free 0.5ml tubes (Ambion (Europe) Ltd. Huntingdon, Cambridgeshire, United Kingdom) were used for all work involving RNA.

* Acrodisc syringe filters (0.22 and 0.4 µm) were supplied by Gelman Sciences (Ann Abor, MI); these were used for sterilising of filtering small volumes of solutions.

* Petri dishes, bijoux and universals were supplied by Greiner, (Stonehouse, Glos. United Kingdom).

* Disposable, sterile scalpels were supplied by Swann-Morton (Sheffield, United Kingdom).

APPENDIX B

LIST OF MANUFACTURERS

1: List of manufacturers of chemicals

Ambion (Europe) Limited
Spitfire Close,
Ermine Business Park,
Huntingdon, PE29 6XY,
United Kingdom.
www.ambion.com

Amersham Biosciences UK Limited
Amersham Place,
Little Chalfont, HP7 9NA,
United Kingdom.

www.amershambiosciences.com

Applied Biosystems
Lingley House,
120 Birchwood Boulevard,
Warrington, WA3 7QH,
United Kingdom.
www.appliedbiosystems.com

BD Biosciences
21 Between Towns Road,
Cowley, OX4 3LY,
Oxford,
United Kingdom.
www.bd.com

BDH Ltd
Merck House,
Poole, BH15 1TD,
United Kingdom.
www.bdh.com

Becton Dickson U.K. Ltd.
Between Towns Road,
Oxford, OX4 3LY,
United Kingdom.
www.bd.com

Corning incorporated
Sullivan Park R&D Center,
Science Center Drive,
Corning, New York 14831-0001,
United States.
www.corning.com

IDEXX Laboratories Limited
Milton Court,
Churchfield Road,
Chalfont St. Peter, SL9 9EW,
United Kingdom.
www.idexx.com

invitrogen BV
De Schelp 26,
9351 NV Leek,
The Netherlands.

www.bioscience.org

Invitrogen ™ Life technologies
3 Fountain Drive,
Inchinnan Business Park,
Paisley,
United Kingdom.
www.lifetechnologies.com

MWG biotech
Anzinger Str. 7,
D-85560 Ebersberg
Germany.

www.mwg-biotech.com

Nunc
Postbox 280,
DK-4000 Roskilde,
Denmark.

www.nalgenunc.com

Perbio Science UK Ltd.
Century House,
High Street,
Tattenhall, CH3 9RJ,
United Kingdom.
www.perbio.com

Roche applied science
Bell Lane,
Lewes, BN7 1LG,
United Kingdom.
www.roche-applied-science.com

New England Biolabs (U.K.) Ltd
73 Knowl Piece,
Wilbury Way,
Hitchin, SG4 0TY,
United Kingdom.
www.NEB.com

Oxoid Limited
Wade Road,
Basingstoke, RG24 8PW,
United Kingdom.

www.oxoid.com

QIAGEN Limited
Boundary Court,
Gatwick Road,
Crawley, RH10 9AX,
United Kingdom.
www.qiagen.com

Sigma-Aldrich Company Ltd.
Dorset,
United Kingdom.

www.sigmaaldrich.com

2: List of manufacturers of equipment

Beckman Coulter (U.K.) Limited
Oakley Court, Kingsmead Business Park,
London Road,
High Wycombe, HP11 1JU,
United Kingdom.
www.beckman.com

Bio-Rad Laboratories Limited
Bio-Rad House,
Maylands Avenue,
Hemel Hempstead, HP2 7TD,
United Kingdom.
www.bio-rad.com

Edwards Engineering Corporation
PO Box 487,
Pompton Plains, NJ 07444-0487
United States.
www.edwards-eng.com

Gelman Sciences inc.
600 South Wagner Road,
Ann Arbor, MI 48103-9019,
United States.
www.medtech1.com

Gilson Medical Electronics
19, avenue des Entrepreneurs,
B.P. 145,
F - 95400 VILLIERS LE BEL,
France.
www.gilson.com

Grant Instruments (Cambridge) Limited
Shepreth, SG8 6GB
United Kingdom.

www.grant.co.uk

Greiner Bio-One Limited
Unit 5, Stroudwater Business Park,
Brunel Way,
Stonehouse, GL10 3SX,
United Kingdom.
www.greiner-lab.com

Millipore (U.K.) Limited
Units 3&5, The Courtyards,
Hatters Lane,
Watford, WD18 8YH,
United Kingdom.
www.millipore.com

LEEC Limited
Private Road No. 7,
Colwick Industrial Estate,
Nottingham, NG4 2AJ,
United Kingdom.
www.leec.co.uk

Leica AG
Oskar-Barnack-Straße 11,
D-35606 Solms,
Germany.

www.leica.com

LI-COR Biosciences UK Limited
St. John's Innovation Centre,
Cowley Road,
Cambridge, CB4 0WS,
United Kingdom.
www.licor.com

Sanyo Gallenkamp plc
Monarch Way,
Belton Park,
Loughborough, LE11 5XG,
United Kingdom.
www.sanyogallenkamp.com

Sorenson Bioscience Limited
6507 South 400 West,
Salt Lake City, Utah,
United States.
www.sorbio.com

Thermo Hybaid – UK
Action Court,
Ashford Road,
Ashford, TW15 1XB,
United Kingdom.
www.hybaid.co.uk

Ultra-Violet Products Limted
Unit 1,
Trinity Hall Farm Estate,
Nuffield Rd,
Cambridge, CB4 1TG,
United Kingdom.
www.uvp.com

Rainin Instrument Limited
Anachem House,
20 Charles Street,
Luton, LU2 0EB,
United Kingdom.
www.rainin.com

Sarstedt Limited
68, Boston Road,
Beaumont Leys,
Leicester LE4 1AW,
United Kingdom.
www.sarstedt.com

Swann-Morton Limited
Owlerton Green,
Sheffield, S6 2BJ,
United Kingdom.
www.swann-morton.com

Treff AG
Taastrasse 16,
CH-9113 Degersheim,
Switzerland.

www.treff-ag.ch

Xograph Imaging Systems Limited
Xograph House,
Hampton Street,
Tetbury, GL8 8LD,
United Kingdom.

www.xograph.com

APPENDIX C

LIST OF ABBREVIATIONS

APS	ammonium persulfate
BSA	Bovine serum albumin
CsCl	Caesium chloride
Ci	Curie
dATP	2' deoxyadenosine 5'-triphosphate
dCTP	2' deoxycytidine 5'-triphosphate
DEPC	diethyl pyrocarbonate
DMEM	Dulbeco's modified Eagles medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxunucleosidetriphosphates
E.coli	*Eschericia coli*
EDTA	EthyleneDiamineTetraAceticacid (sodium salt)
env	Envelope glycoprotein
EtBr	Ethidium bromide
FCS	Foetal calf serum
FeLV	Feline leukaemia virus
FIV	Feline immunodeficiency virus
gag	Group-specific antigen
GALV	Gibbon Ape Leukaemia virus
HBS	HEPES-buffered saline
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
HIV	Human immunodeficiency virus
KAc	Potassium acetate
M	molar
MBq	Megabecquerel
mg	milligram
µg	microgram
$MgCl_2$	magnesium chloride
ml	millilitre
µl	microlitre
mM	millimolar
µM	micromolar
MOPS	3-(N-morpholino)propanesulfonic acid
MuLV	Murine leukaemia virus

NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanograms
°C	degrees centigrade
ORF	Open reading frame
PERV	Porcine endogenous retrovirus
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pol	polymerase
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
mRNA	messenger ribonucleic acid
tRNA	total RNA
rpm	revolutions per minute
RT	reverse transcriptase
RT PCR	Reverse transcription polymerase chain reaction
Temed	N,N,N',N'-tetramethylethylenediamine
U	units
UV	ultraviolet
v/v	volume per volume
V	volts
vol	volume
W	watts
w/v	weight per volume
%	percentage

APPENDIX D

ONE AND THREE LETTER AMINO ACID CODE

Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

APPENDIX E

LIST OF REFERENCES

1. Abkowitz J.L, Holly R.D, and Adamson J.W. (1987). Retrovirus-induced feline pure red cell aplasia: the kinetics of erythroid marrow failure. Journal of cellular physiology 132, 571-577.

2. Abkowitz J.L, Holly R.D, and Grant CK (1985). Retrovirus-induced feline pure red cell aplasia. Journal of clinical investigation 80, 1056-1063.

3. Abkowitz J.L, Taboada M.R., Sabo K.M., and Shelton G.H. (1998). The ex vivo expansion of feline marrow cells leads to increased numbers of BFU-E and CFU-GM but a loss of reconstituting ability. stem cells 16, 288-293.

4. Abkowitz J.L. (1991). Retrovirus-induced feline pure red blood cell aplasia: pathogenesis and response to suramin. Blood 77, 1442-1451.

5. 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. Journal of virology 72, 4503-4507.

6. 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.

7. Anderson M.M., Lauring A.S., Burns C.C., and Overbaugh J. (2000). Identification of a cellular cofactor required for infection by Feline Leukaemia Virus. science 287, 1828-1830.

8. Ausubel F.A., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., and Struhl K. (1992). Current Protocols in Molecular Biology. John Wiley & sons).

9. Bae Y., Kingsman S.M., and Kingsman A.J. (1997). Functional dissection of the moloney murine leukaemia virus envelope glycoprotein gp70. Journal of virology 71, 2092-2099.

10. Barnett A.L., Wensel D.L., Weihua L., Fass D., and Cunningham J.M. (2003). Structure and mechanism of a coreceptor for infection by a pathogenic feline retrovirus. Journal of virology 77, 2717-2729.

11. Battini J-L., Danos O., and Heard J.M. (1995). Receptor-binding domain of murine leukaemia virus envelope glycoproteins. Journal of virology 69, 713-719.

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

13. Battini J-L., Rasko J.E.J., and Miller A.D. (1999). A human cell-surface receptor for xenotropic and polytropic murine leukaemia viruses: possible role in G-protein coupled signal transduction. Proc. Natl. Acad. Sci. USA 96, 1385-1390.

14. Benviste R.E., Sherr C.J., and Todaro G.J. (1975). Evolution of type C viral genes: origin of FeLV. science 190, 886-888.

15. Birnboim H.C. and Doly J. (1979). A rapid alkaline lysis procedure for screening recombinant plasmid DNA. Nucleic Acids Research 7, 1513-1522.

16. Boom R., Sol C.J., Salimans M.M., Jansen C.L., Wertheim-van Dillen P.M., and van der Noordaa J. (1990). Rapid and simple method for purification of nucleic acids. Journal of Clinical Microbiology 28, 495-503.

17. Boyce J.T., Hoover E.A., Kociba G.J., and Olsen R.G. (1981). Feline leukaemia virus-induced erythroid aplasia: in vitro haeomotopoietic culture studies. Experimental Hematology 9, 990-1001.

18. Brojatsch J., Kristal B.S., Viglianti G.A, Khiroya R., Hoover E.A., and Mullins J.I. (1992). Feline leukaemia virus subgroup C phenotype evolves through distinct alterations near the N terminus of the envelope surface glycoprotein. Proc. Natl. Acad. Sci. USA 89, 8457-8461.

19. Bupp K. and Roth M.J. (2002). Altering retroviral tropism using a random-display envelope library. Molecular therapy 5, 329-335.

20. Burns C.C., Poss M.L., Thomas E., and Overbaugh J. (1995). Mutations within a putative cysteine loop of the transmembrane protein of an attenuated immunodeficiency-inducing feline leukemia virus variant inhibit envelope protein processing. Journal of virology 69, 2126-2132.

21. Caney S. (2000). Feline leukaemia virus: an update. In Practice July/August, 397-401.

22. Chen H., Bechtel M.K., Shi Y., Phipps A., Mathes L.E., Hayes K.A., and Roy-Burman P. (1998). Pathogenicity by feline leukaemia virus, Rickard strain, subgroup A plasmid DNA (pFRA). Journal of virology 72, 7048-7056.

23. Chomczynski P. and Sacchi N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical biochemistry 162, 156-159.

24. Coelen R.J., Jose D.G., and May J.T. (1983). The effect of hexadimethrine bromide (polybrene) on the infection of the primate retroviruses SSV 1/SSAV 1 and BaEV. Archives of virology 75, 307-311.

25. Crandell R.A., Fabricant C.G., and Nelson-Rees W.A. (1973). Development, characterization, and viral susceptibility of a feline (Felis catus) renal cell line (CRFK). In Vitro 3, 176-185.

26. Dean G.A., Groshek P.M., Mullins JI, and Hoover EA (1992). Haemotopoietic target cells of anaemogenic subgroup C versus nonanaemogenic subgroup A feline leukaemia virus. Journal of virology 66, 5561-5568.

27. Delassus S., Sonigo P., and Wain-Hobson S. (1989). Genetic organisation of gibbon ape leukaemia virus. Virology 173, 205-213.

28. Deng H., Unutmaz D., Kewalramani V.N., and Littman D.R. (1997). Expression cloning of new receptors used by simian and human immunodeficiency viruses. Nature 388, 296-300.

29. Donahue P.R., Hoover E.A., Beltz G.A., Riedel N., Hirsch V.M., Overbaugh J., and Mullins J.I. (1988). Strong sequence conservation among horizontally transmissible, minimally pathogenic feline leukemia viruses. Journal of virology 62, 722-731.

30. Dranoff G., Jaffee E., Lazenby A., Golumbek P., Levitsky H., Brose K., Jackson V., Hamada H., Pardoll D., and Mulligan R.C. (1993). Vaccination with irradiated tumour cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific and long-lasting anti-tumour immunity. Proc. Natl. Acad. Sci. USA 90, 3539-3543.

31. Elder J.H. and Mullins J.I. (1983). Nucleotide sequence of the envelope gene of Gardner-Arnstein feline leukemia virus B reveals unique sequence homologies with a murine mink cell focus-forming virus. Journal of virology 46, 871-880.

32. Ericsson T.A., Takeuchi Y., Templin C., Quinn G., Farhadian S.F., Wood J.C., Oldmixon B.A., Suling K.M., Ishii J.K., Kitagawa Y., Miyazawa T., Salomon D.R., Weiss R.A., and Patience C. (2003). Identification of receptors for pig endogenous retrovirus. Proc. Natl. Acad. Sci. U. S. A 100, 6759-6764.

33. Essex M., Cotter S.M, Carpenter J.L., Hardy WG Jr, Hess P, Jarrett W.F.H., Schaller JP, and Yohn D.S. (1975). Feline oncornavirus-associated cell membrane antigen. II. Antibody titers in healthy cats from household and laboratory colony enviroments. Journal of the National Cancer Institute 54, 631-635.

34. Essex M., Grant C.K., Cotter S.M., Sliski A.H., and Hardy W.D.,Jr. (1979). Leukaemia specific antigens: FOCMA and immune surveillance. Hamatologie und Bluttransfusion 23, 453-467.

35. Essex M., Klein G., Snyder S.P., and Harrold J.B (1971). Feline sarcoma virus (FSV) induced tumours: correlation between humoral antibody and tumour regression. Nature 233, 195-196.

36. Evans C.H., Cooney A.M., and DiPaolo J.A. (1975). Colony inhibition mediated by nonimmune leukocytes in vitro and skin reactivity in vivo as indices of tumorigenicity of guinea pig cultures transformed by chemical carcinogens. cancer research 35, 1045-1052.

37. Fan H. (1992). Murine leukaemia viruses (retroviridae). In Encyclopedia of virology, Granoff A. and Webster R.G., eds. Academic Press Limited), pp. 995-1001.

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

39. Feng D.F. and Doolittle R.F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. Journal of molecular evolution 25, 351-360.

40. Furth J., Seibold H.R., and Rathbone R.R. (1933). Experimental studies on lymphomatosis of mice. American Journal of Cancer 19, 521-604.

41. Gaush C.R., Hard W.L., and Smith T.F. (1966). Characterisation of an established line of canine kidney cells MDCK. Proceedings of the society for experimental biology and medicine 122, 931-935.

42. Gengozian N. (1998). Development of monoclonal antibodies to erythroid progenitors in feline bone marrow. Veterinary immunology and pathology 64, 299-312.

43. Gerrish P. (2001). The rythm of microbial adaptation. Nature 413, 299-302.

44. Gey G.O., Hoffmann W.D., and Kubicek M.T. (1952). Tissue culture studies of the proliferative capacity of carcinoma and normal epithelium. cancer research 12, 264-265.

45. Ghosh S.K., Hachmann M.H., Hoover EA, and Mullins JI (1992). Identification of a putative receptor for subgroup A feline leukaemia virus on feline T cells. Journal of virology 66, 3707-3714.

46. Gorman C.M., Rigby P.W., and Lane D.P. (1985). Negative regulation of viral enhancers in undifferentiated embryonic stem cells. Cell 42, 519-526.

47. Gossen M. and Bujard H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA 89, 5547-5551.

48. 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. Journal of general virology 36, 59-74.

49. Gray K.D. and Roth M.J. (1993). Mutational analysis of the envelope gene of moloney murine leukaemia virus. Journal of virology 67, 3489-3496.

50. Han J-Y., Cannon P.M., Lai K-M., Zhao Y., Eiden M.V., and French Anderson W. (1997). Identification of envelope protein residues required for the expanded host range of 10A1 murine leukaemia virus. Journal of virology 71, 8103-8108.

51. Hardy W.G.Jr (1983a). Naturally occuring retroviruses (RNA tumour viruses) Part 1, cancer investigation 1, 67-83.

52. Hardy W.G.Jr (1983b). Naturally occuring retroviruses (RNA tumour viruses) part 2. cancer investigation 1, 163-174.

53. Hardy W.G.Jr, Hess P.W., MacEwen E.G., McLelland A.J., Zuckerman E., Essex M., Cotter S.M., and Jarrett O. (1976). Biology of feline leukaemia virus in the natural enviroment. cancer research 36, 582-588.

54. Hardy W.G.Jr (1993). Feline oncoretroviruses. In The retroviridae, volume 2, Levy J.A., ed., pp. 109-168.

55. Hartikka J., Sawdey M., Cornefert-Jensen F., Margalith M., Barnhart K., Nolasco M., Vahlsing H.L., Meek J., Marquet M., Hobart P., Norman J., and Manthorpe M. (1996). An improved plasmid DNA vector for direct injection of skeletal muscle. Human Gene Therapy 7, 1205-1217.

56. Haywood A.M. (1994). Virus receptors: binding, adhesion strengthening and changes in viral structure. Journal of virology 68, 1-5.

57. Henderson L., Lieber M.M., and Todaro G.J. (1974). Mink cell line Mv1Lu (CCL64): virus formation and the generation of "non producer" transformed cell lines with murine and feline sarcoma viruses. Virology 60, 282-287.

58. Hoover E.A., Kociba G.J., Hardy W.G.Jr, and Yohn D.S. (1974). Erythroid hypoplasia in cats inoculated with feline leukaemia virus. Journal of the National Cancer Institute 53, 1271-1276.

59. Hoover E.A. and Mullins J.I. (1991). Feline leukaemia virus infection and diseases. Journal of the American veterinary medicine association 199, 1287-1297.

60. Hoover E.A., Olsen R.G., Hardy W.G.Jr, Schaller J.P., and Mathes L.E. (1976). Feline leukaemia virus infection: age related variation in response of cats to experimental infection. Journal of the National Cancer Institute 57, 365-369.

61. Hosie M.J., Robertson C., and Jarrett O. (1989). Prevalence of feline leukaemia virus and antibodies to feline immunodeficiency virus in cats in the United Kingdom. Veterinary Record 125, 293-297.

62. Innis M.A., Myambo K.B., Gelfand D.H., and Brow M.A. (1988). DNA sequencing with Thermus aquatis DNA polymerase and direct sequencing of polymerase chain reaction amplified DNA. Proc. Natl. Acad. Sci. USA 85, 9436-9440.

63. Jarrett O. (1994). Feline Leukaemia Virus. In Feline medicine and therapeutics. pp. 473-487.

64. Jarrett O. (1999). Strategies of retrovirus survival in the cat. veterinary microbiol 69, 99-107.

65. Jarrett O., Golder M.C., Toth S., Onions D.E., and Stewart M.F. (1984). Interaction between feline leukaemia virus subgroups in the pathogenesis of erythroid hypoplasia. International Journal of Cancer 34, 283-288.

66. Jarrett O., Laird H.M., and Hay D. (1973). Determinants of the host range of feline leukaemia viruses. Journal of general virology 20, 169-175.

67. Jarrett W.F.H., Crawford E.M., Martin W.B, and Davie H. (1964). Leukaemia in the cat. Nature 202, 566-568.

68. Jindrak L., and Grubhoffer L. (1999). Animal virus receptors. Folia Microbiol (Praha) 44, 467-486.

69. Johann S.V., van Zeijl M., Cekleniak J., and O'Hara B. (1993). Definition of a domain of GALVR1 which is necessary for infection by gibbon ape leukaemia virus and which is highly polymorphic between species. Journal of virology 67, 6733-6736.

70. Kitamura T., Onishi M., Kinoshita S., Shibuya A., and Miyajima A. (1995). Efficient screening of retroviral cDNA expression libraries. Proc. Natl. Acad. Sci. USA 92, 9146-9150.

71. Kozak M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Research 15, 8125-8148.

72. Labrosse B., Treboute C., Brelot A., and Alizon M. (2001). Cooperation of the V1/V2 and V3 domains of human immunodeficiency virus type 1 gp120 for interaction with the CXCR4 receptor. Journal of virology 75, 5457-5464.

73. Lauring A.S., Cheng H.H., Eiden M.V., and Overbaugh J. (2002). Genetic and biochemical analyses of receptor and cofactor determinants for T-cell-tropic feline leukaemia virus infection. Journal of virology 76, 8069-8078.

74. Lavillette D., Ruggieri A., Boson B., Maurice M., and Cosset F-L. (2002). Relationships between SU subdomains that regulate the receptor-mediated transition from the native (fusion-inhibited) to the fusion-active conformation of the murine leukaemia virus glycoprotein. Journal of virology 76, 9685-9685.

75. Linenberger M.L. and Abkowitz J.L (1995). Haematological disorders associated with feline retrovirus infections. Bailliere's Clinical Haematology 8, 73-112.

76. Luthman H. and Magnusson G. (1983). High efficiency polyoma DNA transfection of chloroquine treated cells. Nucleic Acids Research 11, 1295-1308.

77. Mackey L.J., Jarrett W., Jarrett O., and Laird H. (1975). Anemia associated with feline leukaemia virus infection in cats. Journal of the National Cancer Institute 54, 207-217.

78. Madin S.H., Darbey N.B.Jr. (1958). Established kidney cell lines of normal adult bovine and ovine origin. Proceedings of the society for experimental biology and medicine 98, 574-576.

79. Mansky L.M. (2002). In vivo analysis of human T-cell leukaemia virus type 1 reverse transcription accuracy. Journal of virology 74, 9525-9531.

80. Marin M., Tailor C.S., Nouri A., and Kabat D. (2000). Sodium-dependent neutral amino acid transporter type 1 is an auxiliary receptor for baboon endogenous retrovirus. Journal of virology 74, 8085-8093.

81. 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. Canadian journal of comparative medical veterinary science 30, 190-198.

82. McDougall J.S., Kennedy M.S., Sligh J.M., Cort S.P., Mawle A., and Nicholson J.K. (1986). Binding of HTLV III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. science 231, 382-385.

83. Milich L., M[illegible] B., and Swanstrom R. (1993). **V3 loop of the human immunod[illegible] virus type 1 env protein: interpreting sequence variabili[illegible]** of virology 67, 5623-5634.

84. Miller D.G., [illegible]wards R.H., and Miller A.D. (1994). **Cloning of the cellular receptor fo[illegible]notropic murine retroviruses reveals homology to that for gibbon ape [illegible]ia virus.** Proc. Natl. Acad. Sci. **USA 91, 78-82.**

85. Miralles [illegible] A., and Elena S.F. (2000). **Diminishing returns of population s[illegible]he rate of RNA adaptation. Journal of virology 74, 3566-3571.**

86. Morgens[illegible] and Land H. (1990). Advanced **mammalian gene transfer: hig[illegible] retroviral vectors with multiple drug selection markers and a co[illegible]ntary helper-free packaging cell line. Nucleic Acids Researc[illegible]**-3596.

87. Moser M., [illegible]thel Burns C., Boomer S., and Overbaugh J. (1998). **The host ran[illegible] interference properties of two closely related Feline Leukemia [illegible]s suggest that they use distinct receptors. Virology 242, 366-377.**

88. Mullins J.I., [illegible] D.S., Binary R.C.Jr., and Cotter S.M (1984). **Viral transduction [illegible]c-myc gene in naturally occuring feline leukaemias. Nature 308, 856-85[illegible]**

89. Neil J.C., [illegible] D., Doggett D.L., and Mullins J.I. (1987). **The role of feline leuk[illegible]virus in naturally occurring leukaemias. Cancer surveys 6.**

90. Neil J.C. a[illegible]ons D.E. (1985). Feline leukaemia viruses: molecular biology and p[illegible]esis. Anticancer research 5, 49-64.

91. Nunberg J.[illegible] Williams M.E., and Innis M.A. (1984). **Nucleotide sequences of the env[illegible]nes of two isolates of feline leukemia virus subgroup B. Journal of V[illegible]49, 629-632.**

92. Onions D.E., Jarrett O., Testa N., Frassoni F., and Toth S. (1982). Selective effects of feline leukaemia virus on early erythroid precursors. Nature 296, 156-158.

93. Overbaugh J. and Bangham C.R. (2001). Selection forces and constraints on retroviral sequence variation. science 292, 1106-1109.

94. Overbaugh J., Donahue P.R., Quackenbusch S.L., Hoover E.A., and Mullins J.I. (1988a). Molecular cloning of a feline leukaemia virus that induces fatal immunodeficiency. science 239, 906-910.

95. Overbaugh J., Hoover E.A., Mullins J.I., Burns D.P.W., Rudensey L., Quackenbush S.L., Stallard V., and Donahue P.R. (1992). Structure and pathogenicity of individual variants within an immunodeficiency disease-inducing isolate of FeLV. Virology 188, 558-569.

96. Overbaugh J., Miller A.D., and Eiden M.V. (2001). Receptors and entry cofactors for retroviruses include single and multiple transmembrane-spanning proteins as well as newly described glycoposphatidylinositol-anchored and secreted proteins. Microbiology and molecular biology reviews 65, 371-389.

97. Overbaugh J., Riedel N., Hoover E.A., and Mullins J.I. (1988b). Transduction of endogenous envelope genes by feline leukaemia virus in vitro. Nature 332, 731-734.

98. Pacitti A.M., Jarrett O. (1985). Duration of the latent state in feline leukaemia virus infection. Veterinary Record 117, 472-474.

99. Pao S.S., Paulsen I.T., and Saier,J.M.H. (1998). Major facilitator superfamily. Microbiology and molecular biology reviews 62, 1-34.

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

101. Perkins C.P. Kar V., Shutter J.R., Del Castillo J., Danilenko D.M., Medlock E.S., Pont... Graham M., Stark K.L., Zuo Y., Cunningham J.M., and Bosselman... (1997a). Anaemia and perinatal death result from loss of the murine ... c retrovirus receptor mCAT-1. Genes and development 11, 914-925.

102. Perkins C.P. Kar V., Shutter J.R., Del Castillo J., Danilenko D.M., Medlock E.S., Pontin... L., Graham M., Stark K.L., Zuo Y., Cunningham J.M., and Bosselman... A. (1997b). Anemia and perinatal death result from loss of the murine ... c retrovirus receptor mCAT-1. Genes development 11, 914-925.

103. Quackenbu... C.L., Donahue,P.R., Dean G.A., Myles M.H., Ackley C.D., ooper M.D., ... JI, and oover EA (1990). Lymphocyte subset alterations and viral determinants of immunodeficiency disease induction by the feline leukaemia virus FeLV-FAIDS. Journal of virology 64, 5465-5474.

104. Quigley J.C. ... ns C.C., Anderson M.M., Lynch E.D., Sabo K.M., Overbaugh ... and Abkowitz J.L. (2000). Cloning of the cellular receptor for feline leuka... virus subgroup C (FeLV-C), a retrovirus that induces red cell ap... ia [... ished erratum appears in Blood 2000 Jul 1;96(1):8]. Blood 95, 10... -10...

105. Rasko J.E.... i J-L, Gottschalk R.J., Mazo I., and Miller A.D (1999). The RD114 ... type D retrovirus receptor is a neutral amino acid transporter. ... Natl. Acad. Sci. USA 96, 2129-2134.

106. Reitz ... M. ... 1999). Gibbon ape leukaemia virus (retroviridae). In Encyclopedia ... of virology, Granoff A. and Webster R.G., eds. Academic Press Limited, ... pp. 617-619.

107. Riedel N., ... E.A., Dornsife R.E., and Mullins J.I. (1988). Pathogenic and ho... determinants of the feline aplastic anemia retrovirus. Proc. Natl. Acad. ... USA 85, 2758-2762.

108. Riedel N.O. Hoover E.A., Gasper P.W., Nicolson M.O., and Mullins J.I. (1986). Mol analysis and pathogenesis of the feline aplastic anaemia retrovirus, fe kaemia virus C-Sarma. Virology 60, 242-250.

109. Rigby M.A. ecular basis of the subgroup determinants of feline leukaemia v 1989. University of Glasgow.

110. Rigby M.A. J.L., Stewart M.A., Kociba G.J., Cheney C.M., Rezanka L.J., Mathe Hartke J.R., Jarrett O., and Neil J.C. (1992). Partial dissociation bgroup C phenotypes and in vivo behaviour in feline leukaemia v s with chimeric envelopes. Journal of general virology 73, 2839-2847.

111. Rohn J.L., M.S., Gwynn S.R., Baldwin D.N., and Overbaugh J. (1998). In olution of a novel, syncytium-inducing and cytopathic feline leu s variant. Journal of virology 72, 2686-2696.

112. Rojko J.L., E.A., Quackenbusch S.L., and Olsen R.G. (1982). Reactivation atent feline leukaemia virus infection. Nature 298, 385-388.

113. Roy-Burman P. (1996). Endogenous env elements: partners in generation of pathogenic feline leukaemia viruses. virus genes 11, 147-161.

114. Sambrook J. itsch E.F., and Maniatis T. (1989). Molecular cloning, a laboratory al. Cold Spring Harbor Laboratory Press).

115. Sarma P.S. and Log T. (1973). Subgroup classification of feline leukaemia and feline a viruses by viral interference and neutralization tests. Virology 54 -179.

116. Sarma P.S. g T., Jain D., Hill P.R., and Huebner R.J. (1975). Differential host range c iruses of feline leukemia-sarcoma complex. Virology 64, 438-446.

117. Shelton G.H., Waltier R.M., Connor S.C., and Grant C.K. (1989). Prevalence of feline immunodeficiency virus and feline leukaemia virus infections in pet cats. Journal of American animal hospital association 25, 7-12.

118. Smith T.F., Waterman M.S., and Burks C. (1985). The statistical distribution of nucleic acid similarities. Nucleic Acids Research 13, 645-656.

119. Sommerfelt M.A. (1999). Retrovirus receptors. Journal of general virology 80, 3049-3064.

120. 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. Nucleic Acids Research 23, 628-633.

121. Spandidos D.A. and Wilkie N.M. (1984). Malignant transformation of early passage rodent cells by a single mutated human oncogene. Nature 310, 469-475.

122. Stamatatos L. and Cheng-Mayer C. (1993). Evidence that the structural conformation of envelope gp120 affects human immunodeficiency virus type 1 infectivity, host range and syncytium-forming ability. Journal of virology 67, 5635-5639.

123. 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 leukaemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. Journal of virology 58, 528-834.

124. Sugai J., Eiden M., Anderson M.M., van Hoeven N., Meiering C.D., and Overbaugh J. (2001). Identification of envelope determinants of feline leukaemia virus subgroup B that permit infection and gene transfer to cells expressing human Pit1 or Pit2. Journal of virology 75, 6841-6849.

125. Tailor C.S., Nouri A, Lee C.G., Kozak C., and Kabat D. (1999a). Cloning and characterisation of a cell surface receptor for xenotropic and polytropic murine leukaemia viruses. Proc. Natl. Acad. Sci. USA 96, 927-932.

126. Tailor C.S., Nouri A, Zhao Y, Takeuchi Y, and Kabat D. (1999b). A sodium-dependent neutral-amino-acid transporter mediates infections of feline and baboon endogenous retroviruses and simian type D retroviruses. Journal of virology 73

127. Tailor C.S., Willett B.J., and Kabat D. (1999c). A putative cell surface receptor for anaemia-inducing feline leukaemia virus subgroup C is a member of a transporter superfamily. Journal of virology 73, 6500-6505.

128. Takeuchi Y., Patience C., Magre S., Weiss R.A., Banerjee P.T., 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.

129. Takeuchi Y., Vile R.G, Simpson G., Collins M.K.L, and Weiss R.A. (1992). Feline leukaemia virus subgroup B uses the same cell surface receptor as gibbon ape leukaemia virus. Journal of virology 66, 1219-1222.

130. Temin H.M. (1993). Retrovirus variation and reverse transcription: Abnormal strand transfers result in retrovirus genetic variation. Proc. Natl. Acad. Sci. USA 90, 6900-6903.

131. Testa N., Onions D.E., Jarrett O., Frassoni F., and Eliason J.F. (1983). Haemopoietic colony formation during the development of pure red cell hypoplasia induced in the cat by feline leukaemia virus. Leukemia research 7, 103-116.

132. Todaro G.J., Habel K., and Green H. (1965). Antigenic and cultural properties of cells doubly transformed by polyoma virus and SV40. Virology 27, 179-185.

133. Vahlenkamp T.W., Verschoor E.J., Schuurman N.N.M.P., van Vliet A.L.W., Horzinek M.C., Egberink H.F., and de Ronde A. (1997). A single amino acid substitution in the transmembrane envelope glycoprotein of feline immunodeficiency virus alters cellular tropism. Journal of virology 71, 7132-7135.

134. van Zeijl M., Johann S.V., Closs E., Cunningham J.M., Eddy R., Shows T.B, and O'Hara B. (1994). A human amphotropic retrovirus is a second member of the gibbon ape leukaemia virus receptor family. Proc. Natl. Acad. Sci. USA 91, 1168-1172.

135. Varela-Echavarria A., Garvey N., Preston B.D., and Dougherty J.P. (1992). Comparison of Moloney murine leukaemia virus mutation rate with the fidelity of its reverse transcriptase in vitro. Journal of biological chemistry 267, 24681-24688.

136. Verschoor E.J., Boven L.A., Blaak H., van Vliet A.L.W., Horzinek M.C., and de Ronde A. (1995). A single mutation within the V3 envelope neutralisation domain of feline immunodeficiency virus determines its tropism for CRFK cells. Journal of virology 69, 4752-4757.

137. Wang H., Dechant E., Kavanaugh M.P., North R.A., and Kabat D. (1992). Effects of ecotropic murine retroviruses on the dual-function cell surface receptor basic amino acid transporter. The journal of biological chemistry 267, 23617-23624.

138. Wang Q-Y., Manicassamy B., Yu X., Dolmer K., Gettins P.G.W., and Rong L. (2002). Characterisation of the LDL-A module mutants of Tva, the subgroup A Rous sarcoma virus receptor and the implications in protein folding. Protein Science 11, 2596-2605.

139. Weijer K., Uytdehaag F.G.C.M., Jarrett O., and Lutz H. (1986). Post-exposure treatment with monoclonal antibodies in a retrovirus system: failure to protect cats against feline leukaemia virus infection with virus neutralising monoclonal antibodies. International Journal of Cancer 38, 81-87.

140. Weiss R.A. and Tailor C.S. (1995). Retrovirus receptors. Cell 82, 531-533.

141. Wunsch M., Schulz 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. The embo journal 2, 2239-2246.

142. Yamaguchi S., Hasegawa M., Suzuki T., Ikeda H., Aizawa S., Hirokawa K., and Kitagawa M. (2003). In vivo distribution of receptor for ecotropic murine leukaemia virus and binding of envelope protein of Friend Murine leukaemia virus. Archives of virology 148, 1175-1184.

143. Young N.S., Abkowitz J.L., and Luzzatto L. (2000). New insights into the pathophysiology of acquired cytopenias. Haematology 18-27.

144. Zavorotinskaya T. and Albritton L.M. (1999). A hydrophobic patch in ecotropic murine leukaemia virus envelope protein is the putative binding site for a critical tyrosine residue on the cellular receptor. Journal of virology 73 10164-10172.

145. Zhang C.C., Krieg S., and Shapiro D.J. (1999). HMG-1 stimulates estrogen response element binding by estrogen receptor from stably transfected HeLa cells. Molecular Endocrinology 13, 632-643.