
THE DETECTION AND ROLE OF HUMAN ENDOGENOUS RETROVIRUSES IN MULTIPLE SCLEROSIS AND OTHER NEUROLOGICAL DISEASES

RICHARD G. SMITH

A thesis submitted in partial fulfilment of the requirements of the
requirements of the University of Wolverhampton for the degree of
Master of Philosophy

SEPTEMBER 2009

This work or any part thereof has not been previously presented in any form to the University or to any other body whether for the purposes of assessment, publication or for any other purpose (unless previously indicated). Save for any express acknowledgements, references and / or bibliographies cited in the work, I confirm that the intellectual content is the work is the result of my own efforts and of no other person.

The right of Richard G. Smith to be identified as author of this work is asserted in accordance with ss.77 and 78 of the Copyright, Designs and Patents Act 1988. At this date copyright is owned by the author.

Signature :

Date : 23 January 2010

Abstract

Human endogenous retroviruses (HERVs) are estimated to form approximately 5% of the human genome. While the majority of sequences are defective, containing premature stop mutations and frameshift mutations, a number encode fully functional proteins. HERVs have been proposed as aetiological agents for a variety of autoimmune diseases, including rheumatoid arthritis (RA), systemic lupus erythematosus, and multiple sclerosis (MS), and have been detected in a variety of tumours. The study aims to develop tools to detect and investigate human endogenous retroviruses in order to establish their roles in MS and anaplastic astrocytomas. A method of detecting and quantifying levels of HERV-W *env* messenger ribonucleic acid (mRNA) and MSRV *gag* by reverse transcriptase polymerase chain (RT-PCR) reaction in a variety of cell lines was developed, with PCR products detected in all cell lines tested, and in particular, high levels of transcription occurring in the BeWo choriocarcinoma cell line. In the astrocytoma cell lines, those with P53 mutation had higher levels of HERV-W *env*. MSRV *gag* variants were also detected in these cell lines, but stimulation with interferon- γ , a proinflammatory cytokine, did not alter expression significantly. An antibody against an epitope of MSRV *gag* has been successfully developed, purified and tested to determine the expression of a predicted linear epitope. This epitope was recognised in all cell lines tested, but unusually for a HERV showed nuclear expression. Further analysis is needed to confirm the identity of the protein detected. Finally a number of retroviral peptides with homology to putative antigens were predicted using a novel bioinformatics approach, of which two, HERV-W *env*₄₁₂ and MSAV *gag*₂₇₄, were tested in an enzyme-linked immunosorbent assay of plasma samples from MS patients, patients with other neurological diseases and normal healthy donors. No significant differences in antibody titres were found between the sample groups for either peptide.

Abbreviation List

ALV	avian leucosis virus
ANN	artificial neural network
ANOVA	analysis of variance
AP	alkaline phosphatase
APC	antigen presenting cell
ATP	adenosine 5'-triphosphate
BLAST	basic local alignment search tool
BLOSUM	blocks substitution matrix
BSA	bovine serum albumen
CASP	critical assessment of techniques for protein structure prediction
CDK	cyclin-dependent kinase
cDNA	complementary DNA
CNS	central nervous system
DAPI	4'-6'-diamidino-2-phenylindole
DCM	dichloromethane
DIPEA	diisopropylethylamine
DMEM	Dulbecco's modified Eagle's medium
DMF	N,N-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphates
EAE	experimental autoimmune encephalomyelitis
ECACC	European Collection of Cell Cultures
EDSS	Expanded Disability Status Scale
EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbent assay
ENV	envelope
ERV	endogenous retrovirus
FCA	Freund's complete adjuvant
FCS	foetal calf serum
FIA	Freund's incomplete adjuvant
Fmoc	fluorenylmethoxycarbonyl
GAG	group-specific antigen
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
HBSS	Hanks' buffered salt solution
HCTU	2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylammonium-hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HERV	human endogenous retrovirus
HHV	human herpesvirus
HIV	human immunodeficiency virus
hnRNP	heterogenous nuclear riboprotein
HOBT	1-hydroxybenzotriazole
HRES	HTLV-related endogenous retrovirus sequence
HRP	horse radish peroxidase
HSRV	human spumaretrovirus
HSV	herpes simplex virus
HTLV	human T cell lymphoma virus
IDDM	insulin-dependent diabetes mellitus
IFN	interferon
IL	interleukin
LTR	long terminal repeat
LINE	long interspersed element

MAG	myelin associated glycoprotein
MALDI-TOF	matrix-assisted laser desorption ionisation time of flight mass spectrometry
MBP	myelin basic protein
MEM	minimum essential medium
MHC	major histocompatibility complex
MOG	myelin oligodendrocyte glycoprotein (MOG)
MoMuLV	Moloney murine leukaemia virus
mRNA	messenger RNA
MS	multiple sclerosis
MSRV	multiple sclerosis-associated retrovirus
MTT	3-(4,-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
NEAA	non-essential amino acids
NHD	normal healthy donor
OASIS	old astrocyte specifically induced substance
OPD	O-phenylenediamine
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PDB	protein database
PERV	porcine endogenous retrovirus
PLP	proteolipid protein
POL	polymerase
PPD	purified protein derivative
PROT	protease
PSI-BLAST	position-specific iterative basic local alignment search tool
PTEN	phosphatase and tensin homologue
RA	rheumatoid arthritis
RNA	ribonucleic acid
RPMI	Roswell Park Memorial Institute
SCID	severe combined immunodeficiency
SINE	short interspersed element
SZRV	schizophrenia-associated retrovirus
Tert-BoC	tert-butyloxycarbonyl
TGF	transforming growth factor
TIS	triisopropylsilane
TM	transmembrane
TMB	3'3',5'5' tetramethylbenzidine
TMEV	Theiler's murine encephalitis virus
TNF	tumour necrosis factor
tRNA	transfer RNA
VZV	varicella-zoster virus

Contents

Chapter 1	Introduction	7
1.1	Human endogenous retroviruses	7
1.2	Multiple sclerosis	10
1.2.1	Symptoms	10
1.2.2	Immunopathology	11
1.2.3	Epidemiology	14
1.3	The retroviral link to MS	14
1.4	HERVs in other neurological diseases	17
1.5	Astrocytic neoplasms	18
1.6	How can HERVs contribute to the disease process	20
1.7	Aims and objectives of the project	25
Chapter 2	Analysis of B cell epitopes	27
2.1	Introduction	29
2.2	Methods and theory	29
2.2.1	Primary structure prediction	34
2.2.2	Secondary structure prediction	35
2.2.3	Alignment of epitopes	36
2.2.4	Three dimensional structure prediction using homology modelling	36
2.3	Results	46
2.4	Discussion	46
Chapter 3	Determination of antibody reactivity against HERV-W epitopes in MS patient plasma by ELISA	49
3.1	Introduction	49
3.2	Methods	51
3.2.1	Peptide synthesis	51
3.2.2	Patient information	54
3.2.3	ELISA of patient samples	55
3.3	Results	56
3.4	Discussion	60

Chapter 4	Detection of HERV-W <i>env</i> and MSRV <i>gag</i> transcripts in glioma cell lines and modulation of MSRV <i>gag</i> expression levels with IFN-γ	63
4.1	Introduction	63
4.2	Methods	65
4.2.1	Culture of cells grown as monolayers	65
4.2.2	Reverse transcriptase PCR	65
4.2.3	RNA extraction and integrity check	69
4.2.4	Primer design	71
4.2.5	cDNA synthesis	71
4.2.6	Polymerase chain reaction	74
4.2.7	Processing of PCR products for sequencing	75
4.2.8	Treatment of CHME3 cells with IFN- γ	75
4.2.9	MTT assay for cell viability of CHME3 cell lines treated with IFN- γ	75
4.3	Results	77
4.3.1	Magnesium chloride optimisation	77
4.3.2	Sensitivity testing of HERV-W <i>env</i> multiplex system in the BeWo cell line	77
4.3.3	Comparison of HERV-W <i>env</i> transcripts in cancer cell lines	77
4.3.4	Detection of MSRV <i>gag</i> transcripts in cancer cell lines	82
5.3.5	Modulation of HERV-W <i>gag</i> expression in the CHME3 microglial cell line	88
4.4	Discussion	90
Chapter 5	Detection of MSRV <i>gag</i>₃₁₁ epitope in astrocytoma cell lines	93
5.1	Introduction	93
5.2	Methods	94
5.2.1	Immunisation of rabbits	94
5.2.2	ELISA for testing antibody reactivity in rabbit serum and purified terminal bleed	95
5.2.3	Affinity purification of rabbit bleeds	95
5.2.4	Culture of cells grown as monolayers	98
6.2.3	Indirect immunofluorescence of cell lines	98
5.3	Results	100
5.3.1	ELISA reactivity of MSRV <i>gag</i> ₃₁₁ antibodies in rabbit bleeds	100
5.3.2	Purification of antibodies and response of purified fraction	101
5.3.3	Detection of MSRV <i>gag</i> ₃₁₁ epitope in astrocytoma cell lines	101
5.4	Discussion	115

Chapter 6	Summary and conclusions	120
6.1	Prediction and screening of epitopes in MS patient samples	120
6.2	Detection of HERV-W RNA transcripts in glioma cell lines	121
6.3	Detection of HERV-W protein products in glioma cell lines	124
6.4	Further applications	125
6.5	Summary	127
References		128
Acknowledgements		152
Appendix A - Physicochemical parameters of amino acids		153
Appendix B - Epitopes identified		154
Appendix C - Ethical approval		157

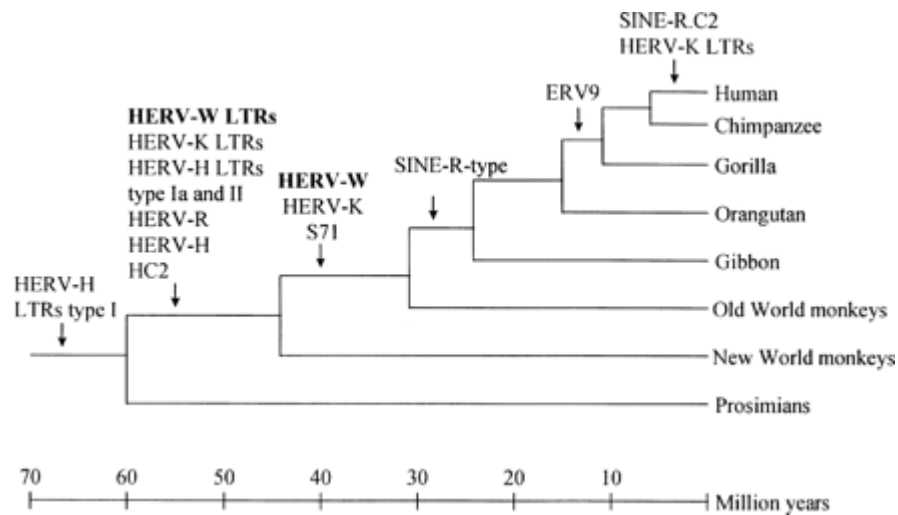
1 Introduction

1.1 Human endogenous retroviruses

Human endogenous retroviruses (HERVs) are estimated to form approximately 5% of the human genome. In contrast to exogenous retroviruses such as human immunodeficiency virus (HIV) and human T cell lymphoma virus (HTLV), HERVs are not transmitted from person to person, but are transmitted vertically (inherited) through the germ line. HERVs represent footprints of ancient retroviral infection, and as such have been termed “fossil viruses” (Sverdlov, 2000). Phylogenetic analysis has indicated that integration into the genome occurred over 60 million years ago, and that a particular family, HERV-W is present in humans and old world monkeys (Kim et al., 1999) but not in new world primates indicating that HERV-W integrated into primate chromosomes at some point in time between the divergence of old world and new world apes and the emergence of man, approximately 33 million years ago (Figure 1.1).

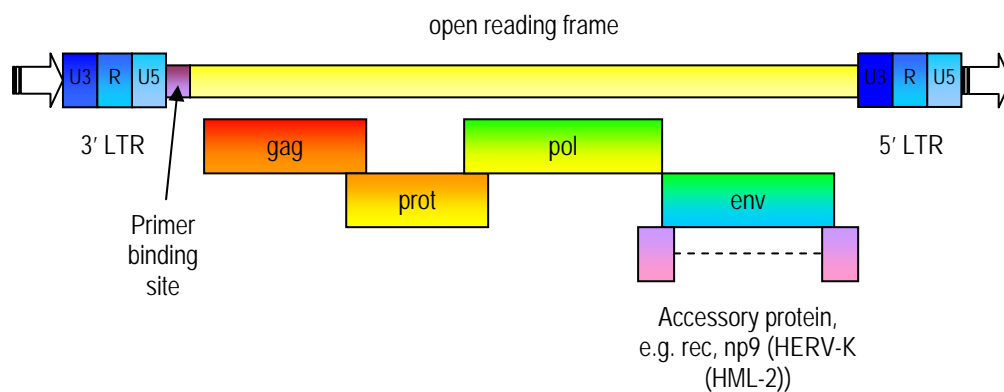
Structurally HERVs closely resemble their endogenous counterparts, comprising of *POL*, *ENV*, and *GAG* genes flanked by long terminal repeats (LTRs) which are important in the integration of retroviruses into the host genome (Figure 1.2). Downstream from the 3' LTR is a short 18 nucleotide sequence known as the primer binding site, where host transfer RNA (tRNA) hybridises to the genome and initiates reverse strand DNA synthesis. The open reading frame consists of a number of genes important in the retroviral life cycle. The *GAG* gene encodes the group-associated antigen, a capsid protein, which in the deltaretroviruses is in the same reading frame as the *PROT* (protease) and *POL* (RNA polymerase gene), and translated as a single precursor protein before being processed and cleaved in the cytoplasm. In other retroviral families *PROT* and *POL* may be present in different reading frames. *ENV* encodes an envelope protein which is always expressed as a single entity, and has two distinct regions, a transmembrane (TM) and a receptor-binding surface (SU) domain. A number of retroviruses also encode accessory proteins through alternative splicing, such as *REC* and *NP9* in the HERV-K (HML-2) family.

Figure 1.1 Phylogenetic tree of integration of HERV proviruses into the primate germline (from Kim et al., 1999)



HERV-W proviruses integrated into the primate genome ≥ 33 million years ago, before the divergence of old and new world primates. HERV-W LTRs are present in both old and new world monkeys. By comparison ERV-9 integrated more recently (~ 12 million years ago), and is only found in hominids.

Figure 1.2 Structure of HERVs
(adapted from Goff, 1996)



The HERV open reading from consists of *GAG*, *PROT*, *POL* and *ENV* genes surrounded by long terminal repeats at the 3' and 5' end. Some HERVs encode accessory proteins using alternative splicing, such as the *REC* and *NP9* gene products.

Previously retroviruses were classified according to morphological and biological criteria, particularly with regard to particle formation observed under electron microscopy. Type A retroviruses produce unenveloped particles, which accumulate in the cytoplasm of cells, while retroviruses of types B, C, and D produce enveloped particles which display different morphology, for example type C retroviruses assemble at the plasma membrane and have a spherical inner core, whilst type D retroviruses assemble in the cytoplasm, and upon budding exhibit a cylindrical core. After formalisation of the genera by the International Committee on Taxonomy of Viruses (Goff, 1996), retroviruses are classified according to a number of criteria, including the morphology, genome structure, receptor utilisation, and the tRNA used as a primer.

HERVs are currently classified according to their degree of sequence similarity to their exogenous counterparts (Table 1.1). Class I HERVs, which include HERV-H, HERV-W, ERV-3, and ERV-9, show similarities to the exogenous gammaretroviruses, such as Moloney Murine Leukaemia Virus (MoMuLV). The class II HERVs are all related to exogenous alpha-, beta-, and gammaretroviruses, which use tRNA^{lys} as a primer binding site. The final class of HERVs, class III, are related to the foamy viruses (spumaretroviruses).

Table 1.1 Classification of endogenous retroviruses
(adapted from Nelson et al., 2003)

Class I	<i>related to gammaretroviruses (e.g. MoMuLV)</i>	
Group 1	HERV-HF	HERV-H, HERV-F
Group 2	HERV-RW	HERV-W, MSRV, HERV-R (ERV 9), HERV-P
Group 3	HERV-ERI	HERV-E, HERV-R (ERV 3), RRHERV-I
Group 4	HERV-T	
Group 5	HERV-IP	HERV-I, HERV-IP-T47D
Group 6	HERV-FRD	
Class II	<i>related to alpha-, beta- and deltaretroviruses (e.g. Mouse Mammary Tumour Virus, Avian Leucosis Virus)</i>	
Group 1	HERV-K (HML-1)	
Group 2	HERV-K (HML-2)	HERV-K10, HERV-K18, HERV-K-HTDV
Group 3	HERV-K (HML-3)	
Group 4	HERV-K (HML-4)	HERV-K-T47D
Group 5	HERV-K (HML-5)	
Group 6	HERV-K (HML-6)	
Group 7	HERV-K (HML-7)	
Group 8	HERV-K (HML-8)	
Group 9	HERV-K (HML-9)	
Group 10	HERV-K (HML-10)	
Class III	<i>related to spumaretroviruses (foamy viruses)</i>	
HERV-L		

Multiple copies of HERVs are found in the human haploid genome, with at least 70 *GAG*, 100 *PRO* and 30 *ENV* sequences present in the case of HERV-W (Voisset et al., 2000). The vast majority of these sequences are defective, containing frameshift mutations and premature stop codons (Kim and Lee, 2001). Few HERVs encode full-length proteins (De Parseval et al., 2003), though a number do. HERV-W *ENV* encodes syncytin, an envelope protein which mediates cell fusion in the placenta (Mi et al., 2000), and may have a role in embryogenesis. HERV-W env can also induce syncytia formation in primate and pig cells (Blond et al., 2000).

While it is clear that endogenous retroviruses play a beneficial role, co-evolving with their hosts, HERVs have been linked with pathogenic roles. HERVs have been linked with breast and germ-line cancers (Weiss, 2001), psychiatric conditions such as schizophrenia (Deb-Rinker et al., 1999; Karlsson et al., 2001), and a variety of autoimmune disorders such as RA, systemic lupus erythematosus (SLE), insulin-dependent diabetes mellitus (IDDM), Sjögren's syndrome, and multiple sclerosis (MS) (Obermayer-Straub and Manns, 2001). With pathogenic roles for HERVs yet to be clearly established, concern has been expressed with xenotransplantation and the possibility of porcine endogenous retroviruses (PERVs) causing human disease (Michie, 2001).

1.2 Multiple sclerosis

1.2.1 Symptoms

Multiple sclerosis (MS) is a chronic disease that tends to begin in late childhood or early in adult life, involving demyelination of axons of neuronal cells of the central nervous system (CNS). It is the most common of the demyelinating diseases, with a lifetime risk of diagnosis of 5.5 per 1,000 women in the United Kingdom population, and 3.1 per 1,000 for men (Alonso et al., 2007).

The symptoms of MS manifest in a number of ways, including disturbances in visual acuity occasionally resulting in double vision, affected motor control of the limbs, lack of coordination, bowel and bladder incontinence, spasticity, sensory impairment (e.g. loss of pain, touch, and temperature) (Charil et al., 2003). MS manifests differently in different people, and there are four main forms – benign MS (where symptoms are mild and recovery occurs after a few attacks); relapsing-remitting MS (the majority of cases, where symptoms flare up and relapse for a while); secondary progressive MS (where the disease progressively worsens after attacks); and primary progressive MS (where there is no distinct relapse or remission period and symptoms steadily worsen) (Lublin and Reingold, 1996).

1.2.2 Immunopathology

Pathologically, MS manifests as demyelinated plaques known as lesions, with large numbers of infiltrating lymphocytes and macrophages present. These lesions can occur at any location in the CNS white matter, but are commonly found in optic nerves, periventricular white matter, brain stem, cerebellum and the white matter in the spinal cord (Charil et al., 2003).

The events that initiate MS are unclear (possibly due to viral or bacterial infection causing upregulation of endothelial adhesion molecules and proteases), but immune cells cross the blood-brain barrier and cause damage (Figure 1.3). The brain is an immunologically privileged site, and while T cells are present in the brains of healthy individuals no autoimmune response occurs, although T cells reactive to myelin components are found in the blood of healthy individuals (Steinman, 1996). Initial activation of these cells is likely to occur outside the central nervous system, e.g. as a result of molecular mimicry with other antigens.

Disruption of the blood-brain barrier is caused by matrix metalloproteases (MMPs) enabling an increase in lymphocyte infiltration into the CNS. A number of these show increased expression in MS, including MMP-7, MMP-9 (Cossins et al., 1997; Lindberg et al., 2001), and a disintegrin and metalloproteinase domain 17 (ADAM-17) (Plumb et al., 2006). Additionally, ADAM-17 cleaves membrane-bound pro TNF- α (TNF- α) into soluble TNF- α (Black et al., 1997), a potent pro-inflammatory cytokine.

Activated T cells then encounter antigens in the CNS, and encounter antigen presenting cells, such as microglia, in the brain triggering clonal expansion of T cells, cytokine production and the further recruitment of infiltrating lymphocytes, including B cells and cytotoxic T cells. Activated T_H2 cells produce interleukins (including IL-4, IL-5, IL-6, IL-10 and IL-13), inducing B cells to secrete immunoglobulins which bind to components of the myelin sheath. In some lesions, oligodendrocyte loss is also seen (Kornek & Lassman, 2003), possibly as a result of attack by invading cytotoxic T cells recognising MHC-I (major histocompatibility complex) bound antigens. Table 1.2 details the role of different types of T cell involved in MS pathogenesis, as well as other cell types such as B cells, macrophages/microglia and endothelial cells.

Figure 1.3: Immunopathology of MS

Following migration through the blood-brain barrier, autoreactive T cells encounter antigen presenting cells in the CNS, leading to T cell activation, and recruitment of B cells, macrophages, and cytotoxic T cells resulting in tissue damage via a number of routes

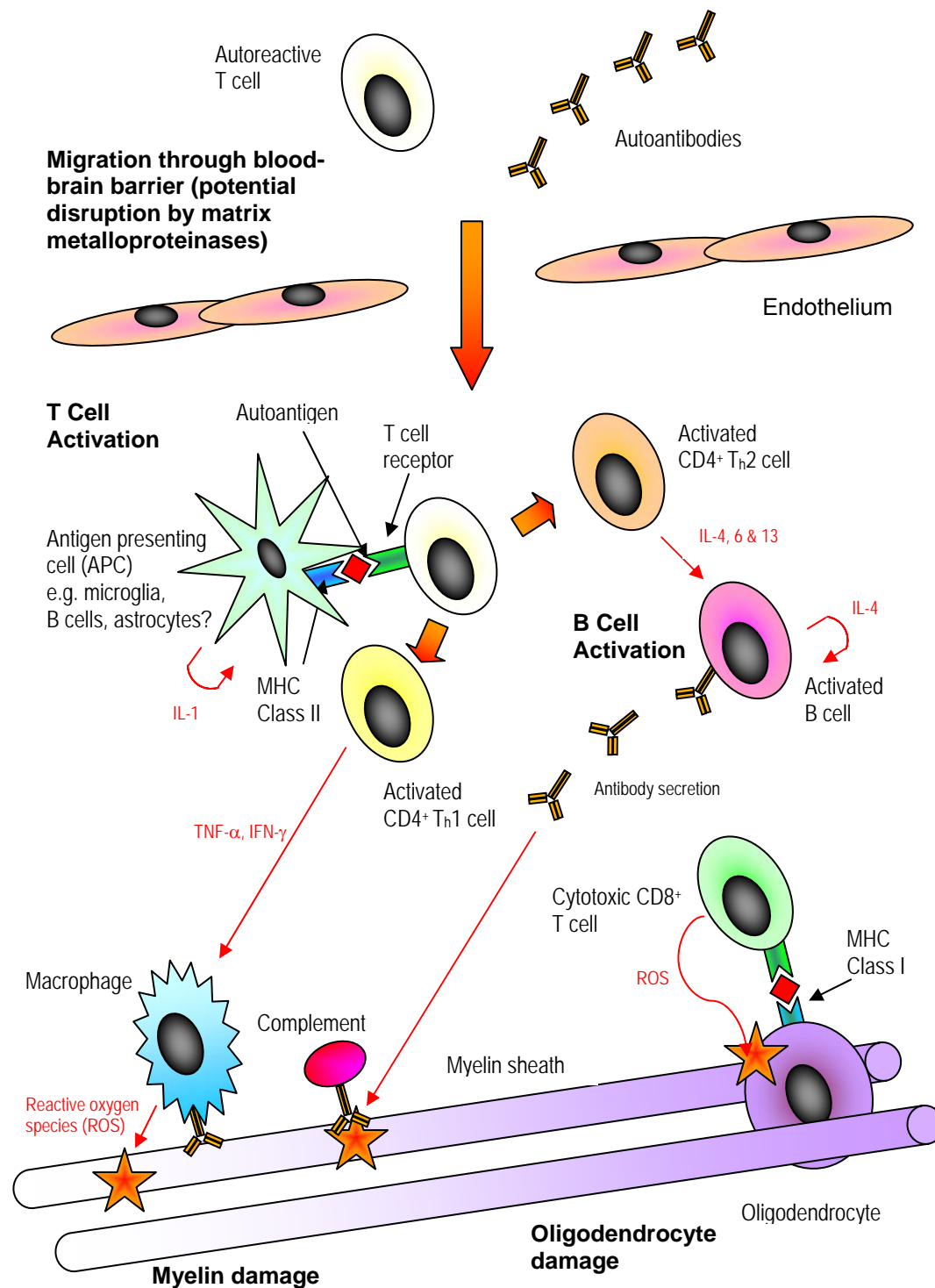


Table 1.2 Cell types involved in MS

Cell type	Subtype	Role	Stimulus	Products
T cell	CD4 ⁺ T _H 1	Macrophage activation Cytotoxic activity	MHC-II	TNF- α IFN- γ
	CD4 ⁺ T _H 2	B cell activation	MHC-II	IL-4 IL-6 IL-13
	CD8 ⁺ cytotoxic $\gamma\delta$	MHC-I mediated cytotoxicity	MHC-I	ROS
		MHC independent cytotoxicity	heat shock proteins	
B cell		Antibody secretion	IL-4	IL-4
		Clonal expansion	IL-6 IL-13	
Microglia		Antigen presentation	IL-1 IFN- γ	IL-1 TNF- α
Macrophage		Antigen presentation	IL-1	IL-1
		Inflammation	IFN- γ	IFN- γ
		Effector cell in demyelination		TNF- α
Endothelial cell		Disruption of BBB (by MMPs) enables lymphocyte infiltration		

While MS is considered to be a T cell-mediated disease, largely due to the presence of T cells present in MS lesions, it is clear that B cells have an important role in the pathology of the disease. Firstly, B cells produce antibodies which bind to components of the myelin sheath (such as myelin basic protein (MBP), myelin associated glycoprotein (MAG), myelin oligodendrocyte glycoprotein (MOG), and proteolipid protein (PLP) or other potential autoantigens in the CNS (e.g. heat shock proteins and arrestins (see Table 2.1 for a more extensive list). The presence of oligoclonal bands are found in approximately 95% of MS patients (Link & Huang, 2006), and is an important diagnostic marker of the disease. IgM in the CSF, produced by CD5⁺ B cells is of a low affinity and targets lipid rich epitopes (Boes, 2000) such as components of the myelin sheath. These immunoglobulins opsonise myelin, leading indirectly via chemokine induction, resulting in destruction by macrophages, which themselves are recruited by pro-inflammatory cytokines such as tumour necrosis factor alpha (TNF- α) and interferon gamma (IFN- γ). Antibody binding to myelin components may also bind to complement, triggering a cascade resulting in further tissue damage.

Only a small proportion (approximately 7%) of antibodies are of the IgM subclass (Esiri, 1977), suggesting a role for memory B cells (which appear to be the main effector cells) and secreted IgG. There is evidence of clonal expansion of B cells within MS lesions (Baranzini et al., 1999), and in the CSF (Colombo et al., 2000). Antibodies present may be produced within the CNS by B cells, or penetrate the blood-brain barrier.

The second role of B cells in MS are as antigen presenting cells (APCs). A number of cell types are capable of acting as APCs in the CNS, including infiltrating B cells, microglia (Hayes et al., 1987), and astrocytes (Zeinstra et al., 2000). Upregulation of MHC-II in these cells leads to T_H1 and T_H2 cell activation.

1.2.3 Epidemiology

MS is geographically distributed worldwide, but temperate areas, predominantly populated by those of European descent, have a higher incidence of the disease than elsewhere in the world. In North America the disease has a North-South gradient with between 150-200 cases per 100,000 population in the Northern United States and Canada, compared with a risk of 30 per 100,000 in the southern states (Cook, 1996).

Migration studies have shown that those who emigrate from a high risk country to a lower risk country after the age of 15 retain the prevalence of their country of origin (Kurtzke, 1997), suggesting that MS is a disease of geography rather than genetics, and that an event (infection or other environmental factors are a possibility) that happens during mid teenage years influences the development of MS. Certainly MS tends to be diagnosed most frequently during young adulthood. Hawkes (2002) suggested that MS may, at least in some cases, be a sexually transmitted disease, largely on epidemiological data where the incidence of MS increased in the Faroe Islands, the Orkneys and Shetland, and Iceland during and after World War II, where significant Allied troops were stationed. These clusters may indicate that MS (or a trigger for the disease) may be an infectious agent. Other arguments for a sexually transmitted agent include the similar ages of onset, high incidences where there is a liberal sexual attitude (and conversely, low rates in stricter countries), though others have argued that Hawkes' paper is speculative and that genetic susceptibility may explain much of the epidemiology, and that there is no evidence for vertical transmission of MS, in contrast to exogenous retroviral infection, such as HIV and HTLV (Stewart, 2002).

1.3 The retroviral link to MS

A number of pathogens have been implicated as potential causes of MS, but no link has been definitively established (Table 1.3). A number of retroviruses have been considered as candidates as a number have links with similar neurodegenerative disorders, such as HTLV-1, which causes tropical spastic paraparesis, and the Visna-Maedi virus which causes an MS-like disease in sheep. Other exogenous retroviruses such as human spumaretrovirus (HSRV) (Lycke et al., 1994) have been investigated as potential causes of MS, but no association has been found.

Table 1.3: Viruses associated with MS
(adapted from Cermelli & Jacobsen, 2000)

Group	Family	Viruses
RNA Viruses	Coronaviridae	Coronavirus
	Flaviviridae	Tick borne encephalitis viruses
	Retroviridae	HTLV-I
		HERVs / retrovirus-like particles
	Paramyxoviridae	Measles virus
		Mumps virus
		Parainfluenza virus 1
		Simian virus 5
	Rhabdoviridae	Rabies virus
	Togaviridae	Togavirus
DNA Viruses	Herpesviridae	Herpes Simplex type 1 & 2
		Varicella zoster virus
		Epstein-Barr virus
		Cytomegalovirus
		Human herpesvirus 6
Unconventional agents		Scrapie agent
		Bone marrow agent

Retroviral activity was discovered in a leptomeningeal cell line derived from an MS patient by Perron et al. (1989), leading to the possibility that an unknown retrovirus could be responsible for the disease. Phylogenetic analysis showed that MS associated retrovirus (MSRV) is “related to, but distinct” (Perron et al., 1997) from ERV-9, an endogenous retrovirus. ERV-9 appears to be widely distributed in various tissues in healthy individuals, and it has been argued that MSRV and ERV-9 are virtually indistinguishable, and therefore MSRV is present in various tissues of the whole population, and not just MS patients. MSRV, along with other similar sequences isolated from MS patients have been classified as part of the HERV-W family of endogenous retroviruses (Blond et al., 1999), using a primer binding site with similarities to that of avian retroviruses which use tRNA^{trp}.

MSRV *pol* transcripts have been detected in the CSF of nearly 50% (13 of 27) of MS patients tested, but not in controls with other neurological disorders. In the same study, transcripts were also detected in the plasma and B cell cultures derived from MS patients, but not in controls (Perron et al., 1997). Similarly HERV-W mRNA has been detected in the serum of MS patients (Garson et al., 1998). MSRV *pol* sequences are found in the serum and CSF of healthy individuals, though at a much lower frequency than that of MS patients (Nowak et al., 2003). In a study of Sardinian MS patients, MSRV was detected in the serum in 80% of cases, compared to a rate of 13% in healthy blood donors (Dolei et al., 2002). In a 6 year follow-up study (Sotgiu et al., 2006), patients who were positive for MSRV *pol* sequences had a poorer prognosis than those who were negative, with a higher relapse rate, disability accumulation and a higher chance of progression (with 2 of 10 MSRV *pol* positive cases entering the progressive phase, compared to none in the MSRV *pol* negative group).

HERV-W *env* transcripts have been determined, using a real-time PCR system, to be more highly expressed in pathological sections of MS patient brain tissue (Antony et al. 2007). HERV-W *env* transcripts appear to have particularly high expression in some areas of the brain, particularly in the cerebellum and pons (Kim et al., 2008).

Another study used random phage peptides, some of which had amino acid similarity to MSRV/HERV-W envelope protein, and detected increased antibody titres in MS patients (Jolivet-Reynaud et al., 1999). Fluorescent *in situ* hybridisation (FISH) also shows an increased copy number of pol sequences on chromatin fibres of MS patients compared with healthy controls (Zawada et al., 2003). Antony et al. (2006) indicated that HERV-W *env* expression is significantly higher in MS brain tissue than that of controls but this finding is not reflected in patient serum. Another study by Mameli et al. (2007a) confirmed that HERV-W *env* expression is upregulated in MS patients when compared with normal and pathological controls, though translation into protein is limited to active MS lesions. Unlike the Antony study, expression of RNA was found to be increased in peripheral blood mononuclear cells (PBMCs) (Mameli et al., 2007), though a number of sequences (approximately 30%) isolated from PBMCs show evidence of recombinations from proviruses located at a number of loci - 5q11.2, 6q21, 7q21.2 (ERVWE1), 14q21.3, 15q21.3, 17q12 and Xq22.3 in both MS patients and controls (Laufer et al., 2009).

Immunohistochemistry has indicated that HERV-W products are expressed in the brain of both normal and MS patients, though the expression pattern differs (Perron et al., 2005). In MS patients, *env* antigen was overexpressed only in recent demyelinating regions, where gag accumulation was higher in demyelinated white matter. In controls *env* was expressed widely within white matter, whilst gag was expressed widely in neurons.

Other human retroviral sequences have been linked to MS, including HERV-H/RGH-2, a type C-related endogenous retrovirus (Christensen et al., 1997). HERV-H/RGH-2 has been detected in the plasma of 24 out of 33 MS patients (Christensen et al., 2000), and over 50% of MS patients had antibodies directed against RGH-2 synthetic peptides (Christensen et al., 2003). In the latter study a variant pro-*env* splice variant was expressed in 40% of MS PBMCs, compared to 10% in controls. The HTLV-related endogenous retrovirus sequence (HRES-1) has been found to be widely expressed in peripheral blood mononuclear cells of both MS patients and healthy controls (Rasmussen et al., 1995).

HERV-W has been linked to other neurological diseases such as schizophrenia, a psychiatric disorder that affects approximately 1% of the population. The sequence of schizophrenia associated retrovirus (SZRV) shows a high degree of homology to that of MSRV (Deb-Rinker et al., 1999). SZRV RNA has been found in 29% of individuals with recent-onset schizophrenia/schizoaffective disorder (Karlsson et al., 2001). HERV-W LTRs have also been

localised at the chromosomal location Xq21.3, which has been linked to psychosis (Kim & Crow, 2002). Raised levels of HERV-W *gag* from the putative gene *PTD015*, located at 11q13.5 has been described in mononuclear cells of patients with schizophrenia (Yao et al., 2008), but expression in astrocytes appears to be reduced (Weis et al., 2006) in the condition (as well as in subjects with major depression and bipolar disorder).

1.4 Detection of HERVs in other autoimmune diseases

HERVs, including sequences related to ERV-9, HERV-K and HERV-H have been detected in synovial fluid cells of patients with RA compared with those of controls (Nakagawa et al., 1997). Using a multiplex PCR approach, Ejtehadi et al., (2006) found a higher expression of HERV-K10 *gag* transcripts in PBMCs of RA patients, when compared to patients with non-inflammatory osteoarthritis and normal healthy donors (NHDs). Similarly, quantitative PCR investigating levels of HERV-K18 superantigen transcripts is increased in juvenile RA patients when compared to controls, but is not higher in SLE patients (Sicat et al., 2005). Other HERV-K transcripts have been described in RA, such as *rec9* of HERV-K (HML-2) (Ehlhardt et al., 2006), though rather than being upregulated in the disease state, it was found that expression was decreased in cultured synovial tissue from RA patients when compared to osteoarthritis controls. Studies looking for MSRV *pol* transcripts in RA serum, by contrast have found no significant differences in expression between patients and controls (Gaudin et al., 2000).

There is also evidence that HERV-E (clone 4-1) is detectable in SLE patients, and *gag* sequences have been detected in PBMCs of lupus patients but not controls, with inactivated stop codons present in part of the *gag* region (Ogasawara et al., 2000). In addition to mRNA, antibodies reactive against a recombinant *gag* protein are found in 48.3% of SLE patients and 35.0% of Sjögren's syndrome patients, but not in healthy controls (Hishikawa et al., 1997).

By way of contrast, although Conrad et al., 1997 identified an autoreactive T cell superantigen, reportedly encoded by HERV-K18 in IDDM patients, other studies have failed to detect significant differences in HERV-K18 mRNA expression, with *env* transcripts detected in 100% of patients and controls, in PBMCs (Jaeckel et al., 1999) and plasma samples (Badenhoop et al., 1999). Additionally, Badenhoop et al. (1999) found that there was no significant difference in antibody reactivity in the plasma samples between IDDM patients and healthy controls.

1.5 Astrocytic neoplasms

Intracranial neoplasms can occur in all areas within the cranial vault, such as the brain, meninges and pituitary gland, and can affect all types of cell present in the CNS. Cell types affected include glial cells (astrocytes, oligodendrocytes, microglia, ependymal cells, and radial glia), meninges, blood vessels, lymphatic cells, endocrine glands, as well as the skull. Tumours of the glial cells are referred to as gliomas. The incidence of primary brain tumours is roughly 4,300 cases per annum in the United Kingdom (McKinney, 2004), equating to 7 cases per 100,000 in the population. Typically, the symptoms of gliomas can manifest in a variety of ways, including headaches (sometimes with nausea and vomiting), movement or sensation changes affecting specific areas of the body, seizures, confusion, memory loss and personality changes (Wen and Kasari, 2008).

The World Health Organisation currently classifies astrocytic neoplasms into four malignancy grades based on histological/cytological and immunocytochemical characteristics. Lower grade tumours closely resemble normal astrocytes while higher grade tumours show evidence of nuclear pleomorphism, mitotic activity, and widespread necrosis and angiogenesis at the tumour site. Grade I tumours, or pilocytic astrocytomas most commonly occur in children, and tend to be non-aggressive with progression rarely occurring. The higher grade tumours (grades II, III and IV) tend to occur in adults, with the incidence peaking between 25-50 years in grade II astrocytomas, and between the ages of 45 and 70 years in grade IV (glioblastoma multiforme, GBM). The prognosis differs according to grade with glioblastoma patients having a mean survival time of 9-11 months compared to about 7 years in patients with grade II astrocytomas (Collins, 2004).

The molecular genetic basis of glioblastomas falls into two groups (for a detailed review see Rasheed et al., 1999). *De novo* glioblastomas show a high frequency of loss of 10q chromosomal heterozygosity, amplification and rearrangement of epidermal growth factor receptor (EGFR), alterations in phosphatase and tensin homologue (*PTEN* gene), deletions of cyclin-dependent kinase (CDK) inhibitor 2a, but a low frequency of P53 mutation. In progressive glioblastomas P53 mutation is frequent with amplification of platelet derived-growth factor receptor alpha polypeptide. Amplification of epidermal growth factor receptor, alterations in phosphatase and tensin homologue, and deletions of CDK inhibitor 2a (*CDKN2A*) occur with lower frequency in progressive tumours. Other changes present in gliomas include loss of heterozygosity for 9p, 13q, 17p, 19q and 22q, with a gain of 7q, but have also lost 10q (Inda et al., 2003).

p53 mutations are infrequent in *de novo* cases of glioblastoma, but commonplace in secondary cases which have progressed from lower grade tumours (Nozaki et al., 1999). Loss of p53 function results in breakdown of the regulation of cell growth, leading to genomic

instability and malignant transformation. The vast majority of mutations that deactivate p53, tend to occur in the DNA binding domain (Walker et al., 1999), affecting its role as a transcriptional regulator.

p53 plays an important role in cell cycle arrest via the upregulation of the *P21* gene (*WAF-21*, *Cip-1*), which encodes for a protein which binds to and inhibits a number of cyclin dependent kinases, including CDK 4 as well as binding to the C-terminal domain of proliferating cell nuclear antigen (PCNA), inhibiting DNA synthesis at S phase (although not DNA repair) (for review see Levine, 1997). The second role that p53 plays is in triggering apoptosis of cells that are damaged beyond repair. p53 induces the transcription of Bax, a pro-apoptotic protein while also simultaneously downregulating *Bcl2*, which encodes a negative regulator of apoptosis which competes with Bax (Basu & Haldar, 1998). Finally, p53 has 3'→5' exonuclease activity, indicating a possible role in DNA repair (Albrechtsen et al., 1999).

Loss of heterozygosity results in allelic loss of tumour suppressor genes, unmasking defective copies of tumour suppressor genes which function abnormally. One such tumour suppressor gene is *PTEN*, located at 10q23, and is a common loss in glioblastomas (a loss of heterozygosity is seen in approximately 74% of cases, of which *PTEN* is defective in 60% of those cases) (Wang et al., 1997). *PTEN* is a lipid phosphatase, activating phosphokinase B/Akt which signals cells to stop dividing and undergo apoptosis (Knobbe et al., 2002). Mutation of this gene causes an escape from these processes, and cellular proliferation. Cyclin dependant kinases are involved in progression of cell cycle. *CDKN2A* is one such inhibitor, and is associated with a loss of heterozygosity at 9p (Ueki et al., 1996). It inhibits CDK 4, which is almost exclusively expressed in higher grade tumours (Lam et al., 2000), and acts by phosphorylating the retinoblastoma (*Rb*) oncogene product, expressed at G1 phase of cell cycle (Harbor et al., 1999). Dephosphorylation of *Rb* inhibits the cell cycle, allowing the cell cycle to progress beyond the G1 phase, preventing cell cycle arrest.

Conversely, gains result in an increase in expression of oncogenes. A commonly seen gain in glioblastoma multiforme patients is at 7p (Misra et al., 2004), usually in tandem with 10q loss (Inda et al., 2003). Of note, *EGFR* is located at 7p12 (Wang et al., 1993), and is overexpressed in up to 50% of glioblastomas (Ekstrand et al., 1991). *EGFR* is a transmembrane protein that binds a number of ligands, including epidermal growth factor and transforming growth factor β . Binding of these ligands results in activation of a number of signal transduction pathways and cellular proliferation. Interestingly a number of glioblastomas also express EGFRvIII, a characteristic mutant of *EGFR* with a 267 base pair deletion coding for amino acids in the extracellular domain (Wong et al., 1992), resulting in an autocrine loop (Ramnarain et al., 2006), without the need for an external stimulus. Regardless of whether wild-type *EGFR* is overexpressed due to chromosome gains, or the presence of mutant EGFRvIII, cellular proliferation is markedly increased as a result.

1.6 How can HERVs contribute towards the disease process?

A number of factors can influence HERV expression, including exogenous factors, such as exogenous viral infection and ultraviolet (UV) light, and endogenous factors such as cytokines and chemokines (Figure 1.4). Normally most HERVs are transcriptionally silent and are heavily methylated. The use of demethylating agents such as 5-aza-deoxycytidine increases the transcription of a number of silent genes, such as HERV-E4.1 (Ogasawara et al., 2003). UV light can induce expression of a number of retroviruses, e.g. HIV (Stanley et al., 1989), HERV-K-T47D, and ERV9 (Hohenadl et al., 1999). Patients with SLE are often photosensitive, and exposure to UV light can exacerbate systemic symptoms (Kind et al., 1993). As MS incidence is a disease associated with latitude, it may be possible that UV light may be an aetiological factor.

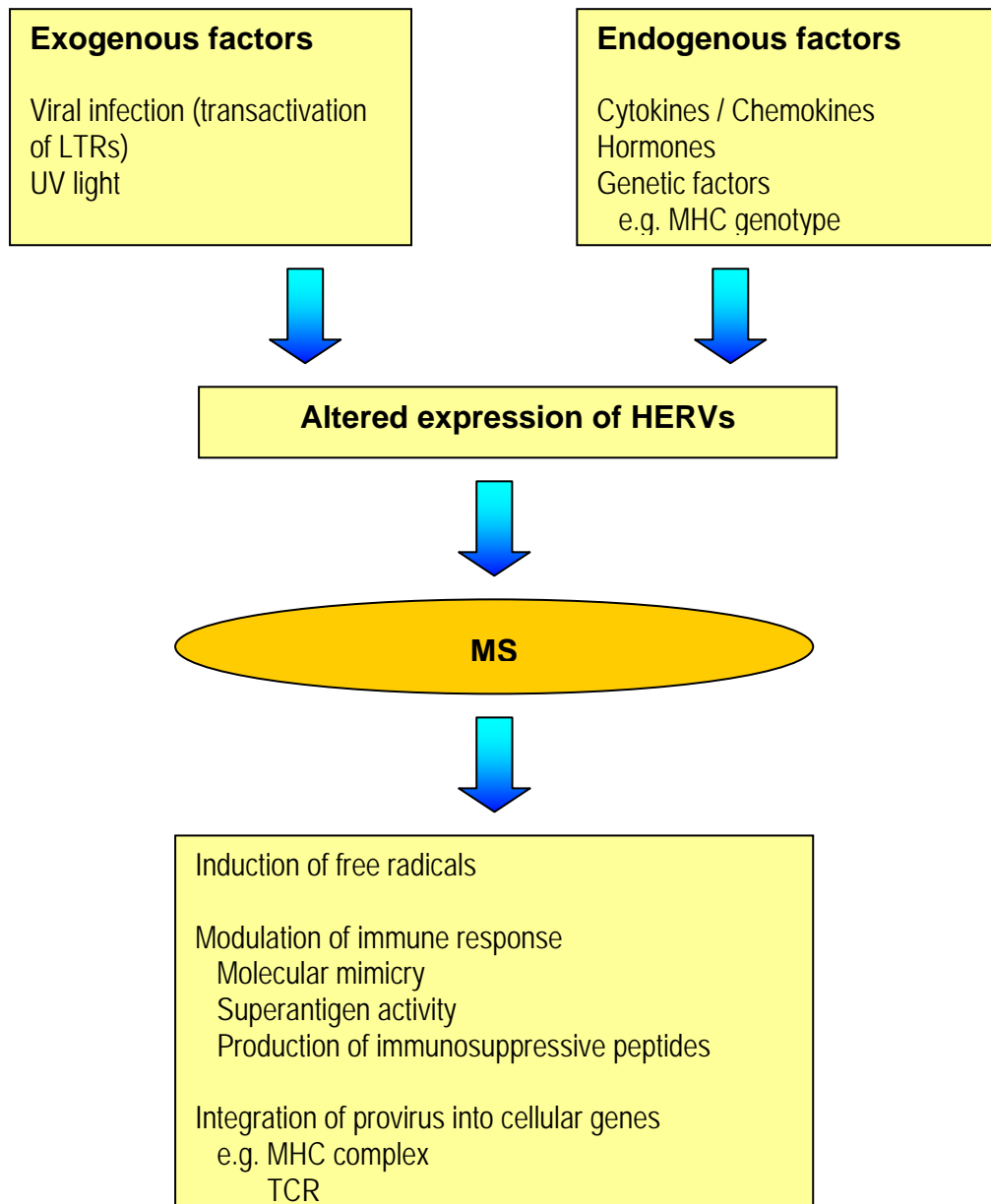
HERV expression can also be transactivated by external viruses, notably members of the herpesvirus family, including herpes simplex type 1 (HSV-1), varicella-zoster virus (VZV), Epstein-Barr virus (EBV) and human herpesvirus 6 (HHV-6) (Brudek et al. 2004). A synergistic inflammatory response occurs in peripheral blood taken from both MS patients and controls when HERV antigens and HSV-1 and HHV-6 antigens. Immediate early protein 1 has been shown to transactivate HERV-W LTRs *in vitro* (Lee et al., 2003). HERV-W *env*, though not *gag* transcripts are upregulated *in vitro* in the CCF-STTG1 astrocytoma cell line following HSV-1 infection. HERV-W *env* and *gag* transcripts are also upregulated in CCF-STTG1, as well as in the histiocytic cell lymphoma cell line, U937 and a human kidney cell line, 293F. The results for influenza A transactivation showed a similar pattern to that of serum deprivation (with the exception of 293F), indicating that cellular stress may partly be responsible for the effects seen (Nellåker et al., 2006). Other HERV families can also be transactivated by herpesviruses, such as HERV-K by HSV-1 (Kwun et al., 2002) and EBV (Sutkowski et al., 2001). Cytokine expression also can affect HERV expression. IL-4, IL-6, IFN- α , and TNF- α increase the expression of MSRV *pol*, whilst IFN- β (which is used in MS therapy) has the converse effect (Serra et al., 2003). Similarly the promoter for *ERVWE1* (HERV-W *env*) is activated *in vivo* upon stimulation of the U87-MG glioblastoma cell lines with the proinflammatory cytokines IFN- γ , TNF- α , IL-1 and IL-6, and inhibited by IFN- β , which dampens the immune response in the CNS (Mameli et al., 2007b).

It is possible that alterations in HERV expression mediated by such factors could trigger MS, either by directly targeting the central nervous system, or by altering the immune response towards it. A number of mechanisms have been proposed for retroviral involvement in MS, such as free radical induction, modulation of the immune response (molecular mimicry, superantigen activity, production of immunosuppressive peptides), and insertion into cellular genes, such as T cell receptors and the MHC complex (Adelman & Marchonis, 2002).

Figure 1.4 How HERVs may contribute towards MS

Adapted from Adelman & Marchalonis (2002)

A number of factors can influence the expression of HERVs, such as viral infection and cytokines. Altered expression of HERV products could lead to MS through a variety of routes.



The aberrant expression of HERV-W gene products may directly damage the brain, causing plaque formation. A 17kDa glycosylated protein secreted by macrophages/monocytes, has been detected in the cerebrospinal fluid of relapsing-remitting MS patients containing reverse transcriptase activity (Ménard et al., 1997b). The presence of MSRV RNA was also detectable in the supernatants of these macrophages/monocytes in cell culture (Ménard et al., 1997a). It is likely that macrophages secrete a number of similar neurotoxins during the course of severe neurological disease associated with HIV infection. Syncytin may induce reactive oxygen species (such as nitric oxide), as well as IFN 1β which damages oligodendrocytes, preventing demyelination (Antony et al., 2004). Such syncytin-mediated direct damage may be via the induction of old astrocyte specifically induced substance (OASIS), an endoplasmic reticulum stress sensor and the subsequent production of nitric oxide synthase (Antony et al., 2007). Overexpression of OASIS results in the suppression of ER stress-induced cell death where there is an accumulation of unfolded proteins in the ER lumen (Kondo et al., 2005).

Molecular mimicry occurs when the immune system fails to discriminate self from non-self, and antibody responses are directed against self antigens with similar epitopes to foreign antigens encountered, such as proteins present in bacteria, viruses, fungi, and protozoa. The best characterised example is rheumatic fever where antibodies directed against cell membrane antigens present on certain strains of group A streptococci which are cross-reactive with human heart tissue antigens, resulting in damage to heart muscle and valves (Kaplan, 1963).

Myelin components, such as myelin basic protein (MBP) and myelin proteolipid protein (PLP) show similarities in (small parts) of the structure to proteins produced by a number of bacteria and viruses (Shaw et al., 1986; Wuchterpfennig & Stromiger, 1995). Animal models have provided a useful tool in trying to understanding the pathology of the disease. Experimental autoimmune encephalomyelitis (EAE) can be induced in mice by injecting them with myelin proteins, such as MBP and PLP in complete Freund's adjuvant (Martin and McFarland, 1997). Another common animal model used to model is Theiler's Murine Encephalitis Virus (TMEV), with pathologic and immunological similarities to MS. In TMEV-infected mice it appears that myelin antigens are presented to naïve T cells by antigen presenting cells in the CNS, known as microglia. It has also been shown that an amino acid substitution at the T cell receptor primary contact site prevents the pathology of the disease (Miller et al., 2000). Non-pathogenic variants of TMEV can be engineered to cause early onset disease by inserting mimic peptides as well as myelin components such as PLP (Olson et al., 2001).

It has been reported that a member of the HERV-K family encodes a superantigen that may contribute to IDDM by stimulating a particular subset of T cells ($v\beta 7$) (Conrad et al., 1994; Conrad et al., 1997). It may be possible that another HERV may act in a similar way by stimulating T cells which are responsible for the destruction of MBP. The MSRV envelope protein induces a superantigen-type response in T lymphocytes encoding $V\beta 16$ chain as a surface receptor (Perron et al., 2001). In addition, immunising MSRV particles into a humanised severe combined immunodeficiency (SCID) mouse resulted in a T cell mediated pathogenesis, leading to brain damage, death and the overexpression of the proinflammatory cytokines $IFN-\gamma$ and $TNF-\alpha$ (Firouzi et al., 2003).

In addition to influencing T cell responses directly, MSRV may promote an inflammatory response by activating monocytes to produce such cytokines, driving the immune system towards a T_h1 response. The SU subunit of MSRV has been shown *in vitro* to directly stimulate monocytes to produce $IFN-\gamma$ and $TNF-\alpha$ in a CD14 and TLR4-dependant manner (Rolland et al., 2006).

Both HERV-H (De Parseval et al., 2001; Manengety et al., 2001) and HERV-W encode immunosuppressive peptides in their envelope genes. These proteins have a high degree of homology to a peptide, CKS-17 derived from MoMuLV (Figure 1.5), which has been shown to downregulate immune effector functions (Cianciolo et al., 1985), causing an imbalance between the T_h1 and T_h2 response. CKS-17 has been shown to downregulate T_h1 cytokines *in vitro*, including IL-2, IL-12 and $IFN-\gamma$ while increasing IL-10 production (Haraguchi et al., 1995). In the placenta the immunosuppressive domain may protect the developing foetus from immune response (in addition to its role in placental differentiation), but if expressed elsewhere at abnormal levels, the delicate balance of the immune response may alter. It is worth noting that T_h1 cytokines drive the inflammatory response, whilst the T_h2 cytokines dampen the T_h1 response.

Figure 1.5 Comparison of immunosuppressive domains of HERVs

Moloney MuLV (CKS-17)	LQNRRGLDLLFLKEGCL
HERV-H (RGH-1)	CRGLDLLFLKEGGE
HERV-H (p62)	LQNRRGLDLLTAEKGCL
HERV-W (7q)	LQNRRALDLLTAERGST

The CKS-17 p15E immunosuppressive domain is highly conserved amongst human endogenous retroviruses and retroviral elements. Differences from the CKS-17 sequence, such as insertions and substitutions are highlighted in red.

HERVs may contribute towards pathogenesis of tumours in a number of ways, by retrotransposition (causing chromosomal instability), promoter insertion, immunomodulation, disruption of normal HERV-related functions, recombination, or by the production of fusion proteins (for review see Larsson & Andersson, 1998). Retroviruses and retrotransposons are mobile elements which can integrate into the genome and disrupt gene function. Integration of retroelements into the coding region of tumour suppressor genes may act in a similar manner to a point mutation, disrupting protein function. Small interspersed elements (SINEs) related to HERV-K have also been found integrated into the DNA repair gene, *XRCC1* (Misra et al., 2001). Similarly, the insertion of Long interspersed element 1 (LINE-1) into the tumour suppressor gene, *APC* which is mutated in familial polyposis coli and sporadic colorectal tumours (Miki et al., 1992).

Insertion of retroelements adjacent to proto-oncogenes, may place the gene under the control of a viral promoter, altering the level of expression. In B cell lymphomas of chickens, the avian leucosis virus (ALV) integrates adjacent to the proto-oncogene, *C-MYC*, and removes it from cellular control, placing the gene under the influence of a viral promoter (Hayward et al., 1981). In rats, the *c-Ha-Ras* oncogene is overexpressed tenfold due to the insertion of a defective retrovirus in N-methyl-N-nitrosourea induced rat mammary cancer adjacent to the promoter (Bera et al., 1998). Hypomethylation of the MMTV promoter induces the expression of *c-neu*, in transgenic *c-neu* transgenic mice, indicating that the promoter is usually heavily methylated (Zhou et al., 2001). Another example is that of HERV-E (PTN) which inserts adjacent to the growth factor pleiotrophin, acting as an additional promoter in choriocarcinoma (Schulte et al., 1998). HERVs may also be attributable as a consequence of cell transformation, where disruption in normal cell function results in hypomethylation of HERV LTRs and retrotransposons, e.g. HERV-W retrotransposons in malignant ovarian cancers (Menendez et al., 2004).

1.7 Aims and objectives of the project

The research project aims to identify potential HERV epitopes, and to develop primers and antibodies against HERVs in order to provide a set of tools to enable detection of HERVs at serological, RNA and protein level, and to investigate potential mechanisms of how HERVs may play a role in the pathogenesis of multiple sclerosis and cancer. The overall approach, summarised in Figure 1.6, involved the following:

- **Antibody titre determination**

A novel bioinformatic approach was utilised to identify HERV epitopes *in silico*, as potential molecular mimics, with sequence homology with autoantigens implicated in the pathogenesis of multiple sclerosis, (Chapter 2). Peptides of predicted epitopes were synthesised, and used in an enzyme-linked immunosorbent assay to analyse antibody titres in multiple sclerosis patients (Chapter 3).

- **Analysis of mRNA expression**

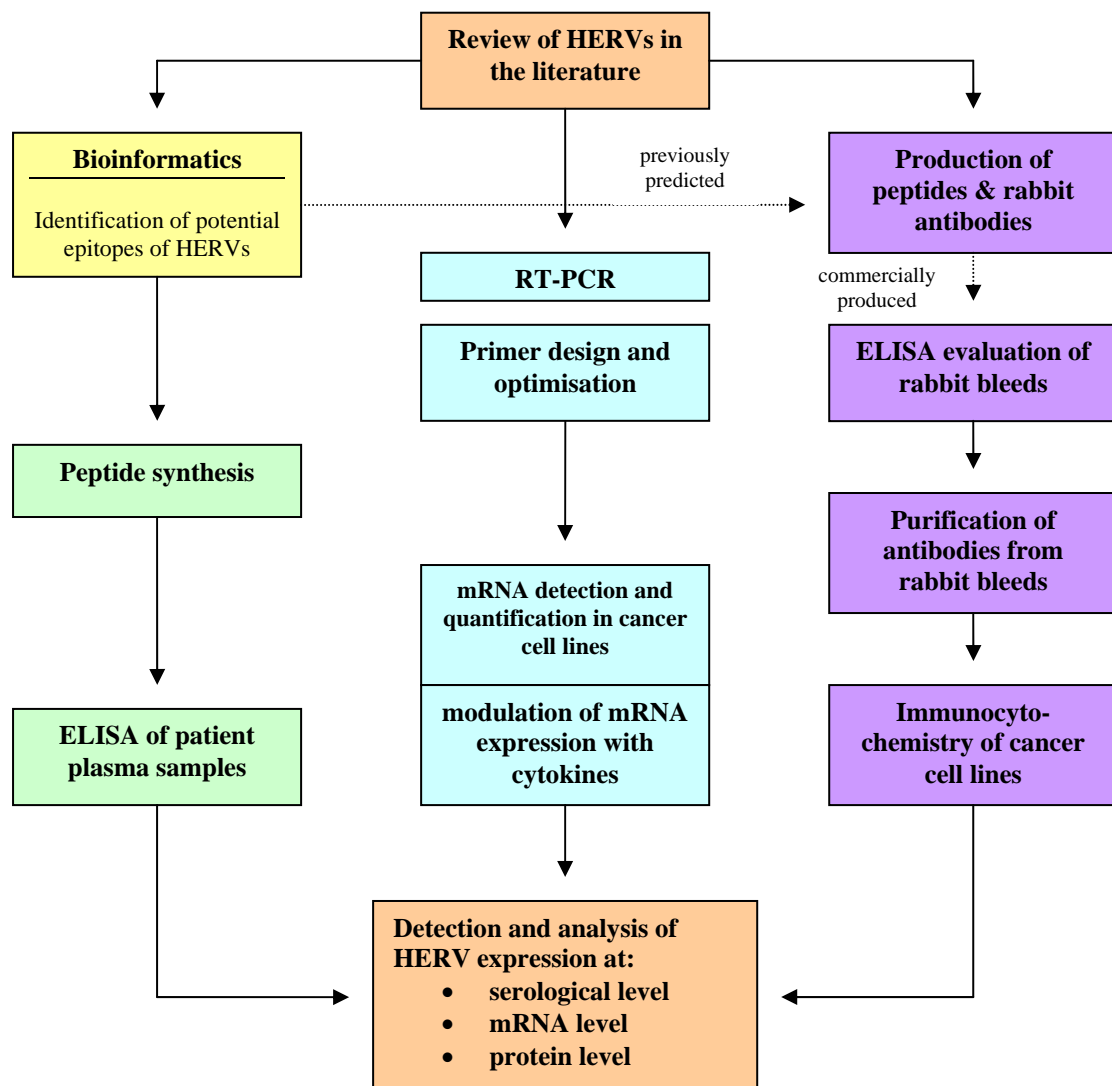
A reverse transcriptase polymerase chain reaction was developed and used to detect and quantify levels of HERV-W *env* and MSRV *gag* expression *in vitro* in cancer cell lines, with a particular focus on astrocytomas/glioblastomas as HERV-W transcripts are known to be widely expressed in the brain, including cancer cell lines derived from astrocytomas (Yi et al., 2002; Nakamura et al., 2003; Stauffer et al., 2004). In order to investigate whether HERV expression is upregulated in the presence of proinflammatory cytokines involved in MS pathology, the reverse transcriptase PCR (RT-PCR) system was used to investigate MSRV *gag* expression in a microglial cell line exposed to interferon- γ (Chapter 4).

- **Analysis of protein expression**

Rabbit serum containing polyclonal antibodies against a peptide previously predicted by a colleague (who used a similar approach to that taken in Chapter 2) were raised commercially. These antibodies were evaluated and purified in order to apply in an indirect immunofluorescence-based protocol to determine the subcellular localisation of MSRV *gag* protein, or more precisely, the epitope, is detectable in astrocytoma cell lines and whether mRNA expression can correlate with protein expression. (Chapter 5)

Figure 1.6 Overview of research project methodology

The research project utilised a number of techniques in order to detect and evaluate HERV expression at serological (green), mRNA (turquoise), and protein (lavender level). A novel bioinformatics approach (yellow) was utilised to predict epitopes to construct as peptides for serological detection and to and raise antibodies for protein detection.



2 Analysis of B cell epitopes

2.1 Introduction

While MS is thought to be a T-cell mediated disease, largely due to the presence of infiltrating T cells at MS lesions and through studies of autoimmune encephalitis there is increasing evidence that B cells and antibodies have an important role in the disease. Over 90% of MS patients show the presence of oligoclonal antibodies in cerebrospinal fluid, which is a diagnostic marker of the disease (Bernard et al., 1981). IgG antibodies and complement, associated with macrophages, are found in actively demyelinating lesions with patients with multiple sclerosis (Breij et al., 2008). It is interesting to note that antibody-producing plasma cells appear to be more prevalent in late lesions with oligodendrocyte loss and extensive destruction of the myelin sheath (Ozawa et al. 1994).

A number of putative autoantigens have been proposed for MS, and those which epitope mapping was carried out on are summarised in Table 2.1. These include components of myelin, as well as a number of other proteins ranging from heat shock proteins to proteins involved in signal transduction. A large number of potential autoantigens have been investigated, including a variety of other molecules such as neurofascin (Mathey et al., 2007), a cell adhesion molecule found on axons; transketolase (Lovato et al., 2008), an enzyme involved the pentose phosphate pathway; and neuron-specific enolase (enolase 2) (Forooghian et al., 2007), a metalloenzyme involved in glycolysis.

The shape (or tertiary structure) that a protein adopts is largely dictated by its amino acid sequence (primary structure). Amino acids in the polypeptide chain interact with each other, and it is these interactions which hold a protein in shape, and thus influence its biochemical properties. Antibodies bind to the native conformation of the protein, and these regions which are recognised by antibodies are known as epitopes (or antigenic determinants).

Using bioinformatics, or *in silico* analysis, the aim was to identify continuous (linear) epitopes present in retroviruses implicated in the aetiology of MS, and to investigate the possibility of molecular mimicry with a number of putative autoantigens. Continuous epitopes are comprised of single segments of polypeptide chain, where discontinuous (conformational) epitopes are formed by different segments which have been brought together due to protein folding. Identification of linear epitopes enables us to produce short peptides for use in ELISA screening, and allows the production of polyclonal antibody for use in screening tissues by immunohistochemistry and western blotting.

Table 2.1 Putative autoantigens in MS

Autoantigen	Function	B cell	T cell	References
<u>Myelin-associated</u>				
Myelin Basic Protein	Component of myelin sheath	✓	✓	Wajgt & Gorny, 1991 Burns et al., 1983
Myelin Associated Glycoprotein	Component of myelin sheath	✓	✓	Wajgt & Gorny, 1991 Johnson et al., 1986
Myelin Oligodendrocyte Glycoprotein	Component of myelin sheath	✓	✓	Sun et al., 1991b
Proteolipid Protein	Component of myelin sheath	✓	✓	Sun et al., 1991a
NoGo A	Neurite outgrowth inhibitor	✓		Reindl et al., 2003
Oligodendrocyte specific protein (Claudin 11)	Component of myelin sheath	✓	✓	Bronstein et al., 1999 Vu et al., 2001
<u>Other proteins</u>				
Alpha B Crystallin	Heat shock protein	✓	✓	Agius et al. (1999) Chou et al. (2004)
Arrestin	Retinal protein. Signal transduction.	✓		Ohguro et al., 1993
Beta Arrestin	CNS protein. Signal transduction.	✓		Ohguro et al., 1993
2'3'-cyclic nucleotide 3' phosphodiesterase	Signal transduction. Synaptic transmission.	✓		Walsh & Murray, 1998
Heterogeneous ribonucleoprotein (hnRNP)	Involved in mRNA expression and transport.	✓		Sueoka et al., 2004
HSP 60	Heat shock protein	✓	✓	Prabhakar et al., 1994 Ruiz-Vazquez & De Castro, 2003
HSC 70	Heat shock protein. Co-localises with MBP?		✓	Mycko et al., 2004
HSP 70	Heat shock protein. Co-localises with MBP?		✓	Mycko et al., 2004
Transaldolase	Involved in pentose phosphate pathway (generation of NADPH)	✓	✓	Colombo et al., 1997

A number of different proteins have been proposed as candidate autoantigens in MS (for B and T cells). These include a number associated with the myelin sheath, and a wide array of other proteins such as signalling molecules and heat shock proteins.

A number of criteria were considered to be important in identifying epitopes, including primary structure analysis (hydrophilicity, polarity, flexibility and accessibility), and secondary structure prediction (identification of alpha helices and beta turns). The protocol followed is summarised in Figure 2.1.

2.2 Methods and theory

2.2.1 Primary structure prediction

One of the major factors in determining protein structure is hydrophilicity. In a solvent environment, hydrophobic residues remain in the core of the protein, while the hydrophilic residues (which are largely polar or charged), such as arginine, aspartic acid, and glutamic acid are exposed on the protein surface and accessible to antibody. Mutations of hydrophilic/polar residues can affect protein folding and thus structure, such as the R1137P mutation of fibrillin in Marfan Syndrome (Thomas et al., 1995).

Another important consideration is that antibody-antigen interactions are dependant on non-covalent forces. These include electrostatic forces, hydrogen bonds, Van der Waals forces, and hydrophobic interactions (where molecules pack together to exclude water).

Hopp and Wood (1981) devised a method based on hydrophilicity to determine antigenic determinants in a number of proteins. Amino acids were assigned a value adapted from those previously assigned by Levitt (1976). The authors used a computer program to work along the polypeptide chain to produce a plot of values averaged over a 6 amino acid stretch. Known epitopes were found to be located at the peaks of these plots.

Amino acid sequences obtained from GenBank (<http://www.ncbi.nlm.nih.gov/Entrez/>) were inputted into the ProtScale tools found on the ExPaSy server (<http://ca.expasy.org>). Values greater than 1.2 and 10 for hydrophilicity and polarity respectively were taken to be significant. The ProtScale tools work in a similar way to the Hopp and Wood study, utilising a sliding window technique where the mean value over a given window size (9 is the size used) is assigned to the central amino acid (Figure 2.2). The tools provide both graphical (Figure 2.3) and numerical output. The default window size of 9 amino acids was selected, as this is considered to be an optimal value for detecting structural features of globular proteins, in contrast to a 19 amino acid window more suitable for the analysis of transmembrane domains (Claverie and Notredame, 2003).

Figure 2.1 Overview of methodology

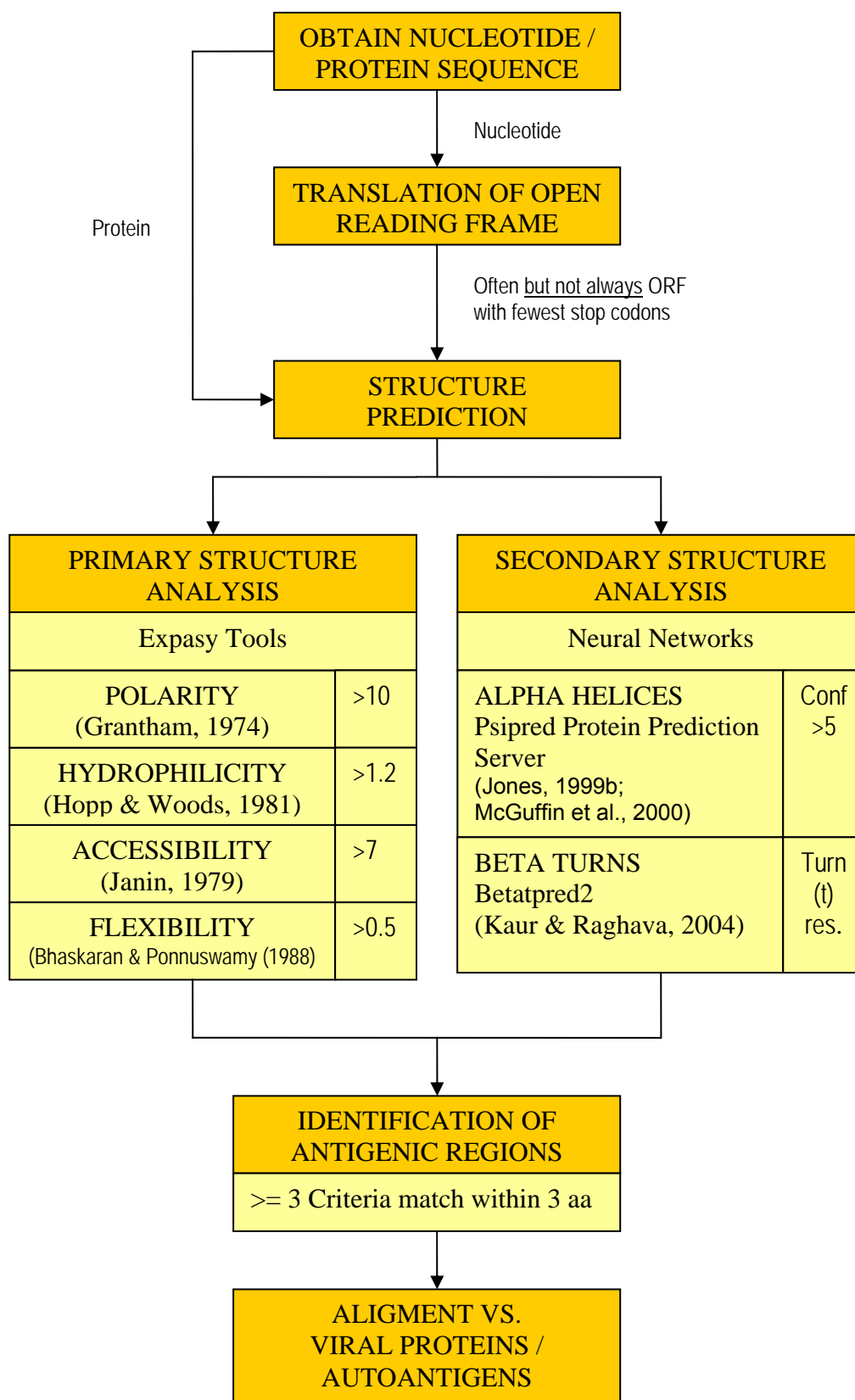
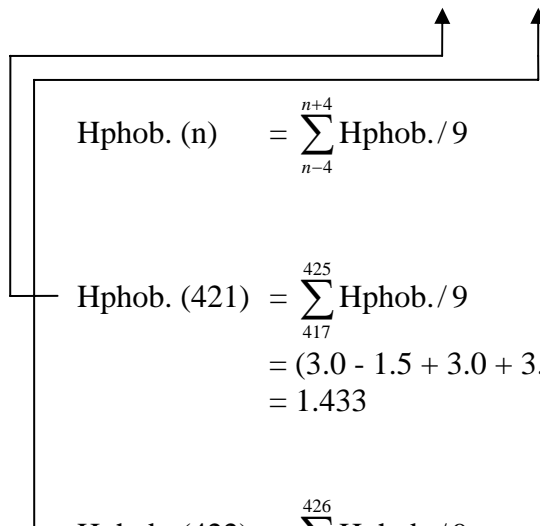


Figure 2.2 Overview of the 'sliding window technique'

In the sliding window technique, parameter values (listed in Appendix A) (e.g. for hydrophilicity) are obtained for each individual amino acid. The mean is taken over a 9 amino acid stretch and assigned to the central amino acid. This gives an indication of how hydrophilic the region is, rather than just the individual amino acid. For hydrophilicity, values of 1.2 or greater are taken to be significant.

417	418	419	420	421	422	423	424	425	426
K	V	K	E	I	R	D	R	I	Q
3.0	-1.5	3.0	3.0	-1.8	3.0	3.0	3.0	-1.8	0.2



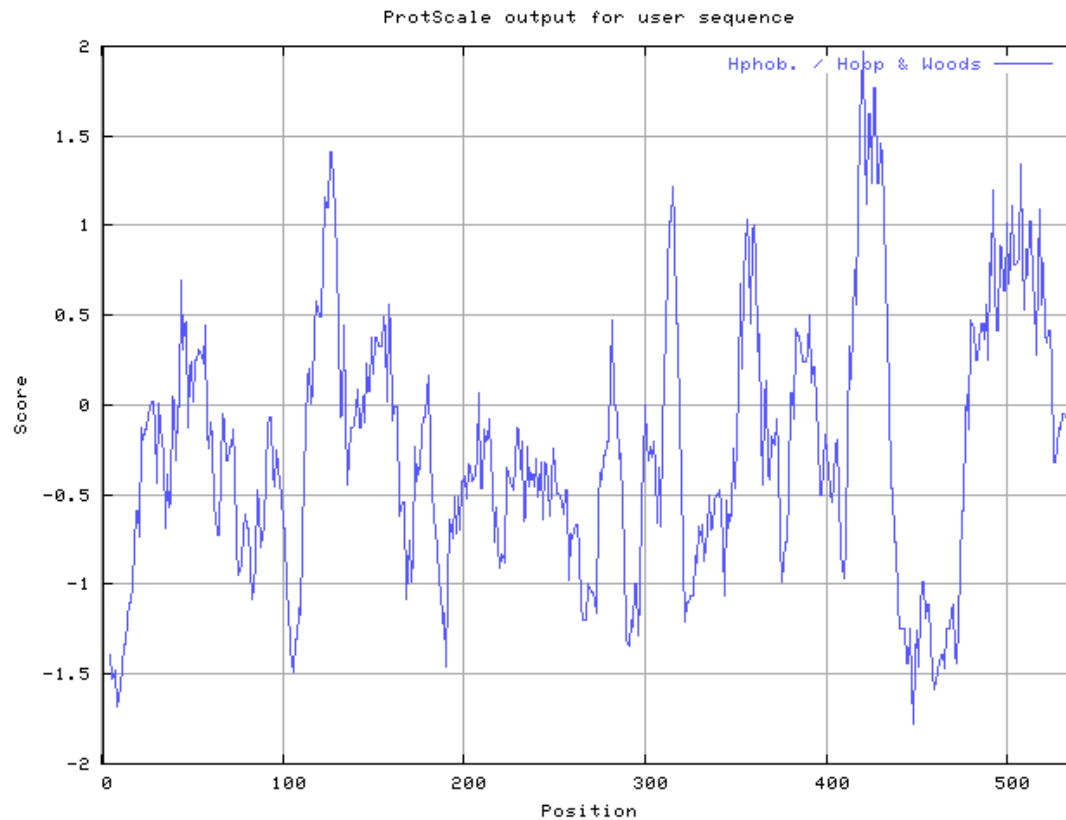
$$\text{Hphob. (n)} = \sum_{n-4}^{n+4} \text{Hphob.} / 9$$

$$\begin{aligned} \text{Hphob. (421)} &= \sum_{417}^{425} \text{Hphob.} / 9 \\ &= (3.0 - 1.5 + 3.0 + 3.0 - 1.8 + 3.0 + 3.0 + 3.0 - 1.8) / 9 \\ &= 1.433 \end{aligned}$$

$$\begin{aligned} \text{Hphob. (422)} &= \sum_{418}^{426} \text{Hphob.} / 9 \\ &= (-1.5 + 3.0 + 3.0 - 1.8 + 3.0 + 3.0 + 3.0 - 1.8 + 0.2) / 9 \\ &= 1.122 \end{aligned}$$

Figure 2.3 ProtScale output for HERV-W env (syncytin) hydrophilicity

An example of a graphical ProtScale output for syncytin using the Hopp and Woods scale for hydrophilicity. Note that there is a strong hydrophilic peak at approx. 420 aa, and a number of smaller peaks. Such peaks often correspond to exposed regions on a protein's surface.

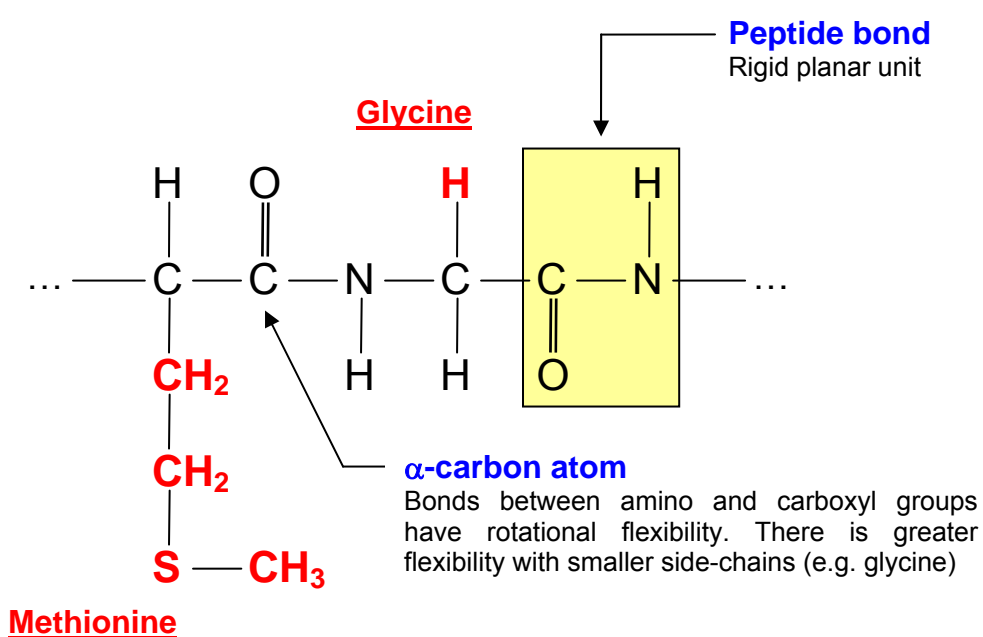


While the peptide bonds in a polypeptide chain are rigid, planar structures as the C-N bond which has partial double-bond characteristics, flexibility in the chain is present in the bonds surrounding the α -carbon atoms (Figure 2.4). Amino acids with larger side chains, such as methionine collide with each other and hinder flexibility. The distribution of amino acids with more flexible side-chains are important in the function of the protein (Vihinen et al., 1987), and therefore would be more likely to be present near the surface of the protein. The flexibility scores were based on values calculated by Bhaskaran & Ponnuswamy (1988). Amino acids were scored on a scale from 0 (minimum flexibility) to 1 (maximum), and values of greater than 0.5 were taken to be significant. The authors showed that flexibility correlated with both hydrophilicity and polarity.

The final criterion used was that of accessibility. For obvious reasons, residues that are inaccessible are unlikely to be interacting with other proteins, such as antibody, if they are buried away. The scale of surface accessibility to solvents was obtained from Janin (1979).

Figure 2.4 Illustration of protein flexibility

In proteins the peptide bond is a rigid planar unit. Flexibility in a protein occurs around the α -carbon atom. Larger side chains tend to have less rotational flexibility as these are more likely to collide with other residues.



2.2.2 Secondary structure prediction

Polypeptide chains can also fold into regular structures, such as alpha helices and beta turns. The arrangement of these structures is called the secondary structure. Alpha helices are rod-like structures, formed by a coiled polypeptide chain where the each CO group is hydrogen-bonded to the NH₂ group of the residue 4 further up in the chain. Alpha helices can coil round each other, forming 'coiled coils' giving rise to strength and rigidity in a protein, particularly in the case of structural proteins, such as keratin in hair. The ends of alpha helices are often polar, and as such are likely to be found on the surface of proteins.

Beta turns are hairpin structures where the direction of the polypeptide chain reverses on itself, and the CO group is hydrogen-bonded to the NH₂ group 3 further up in the chain. In contrast to alpha helices, the structures are stabilised by hydrogen bonding between residues in different strands.

Earlier methods of secondary structure prediction involved looking at the properties of amino acids within the protein sequence. The method of Chou and Fasman (1974) analysed the amino acid of 15 proteins with known three-dimensional structures, and calculated the probability that each particular amino acid would be found within an alpha helix or beta strand. Each amino acid was then classified into their propensity to form or break secondary structures. By summing up these propensities along segments of the sequence, it is possible to determine whether the segment as a whole is part of an alpha helix, a beta strand or a coil region. A similar method, the GOR method developed by Garnier et al. (1978) involved the use of a sliding window technique averaging across a 17 amino acid window size to increase the accuracy. Such methods have been shown to be 56-60% accurate, in comparison to the 70-80% accuracies claimed (Jones, 2003).

As such methods are only moderately accurate and quite cumbersome, modern methods of secondary structure prediction rely on studying multiple alignments carried out by Artificial Neural Networks (ANNs). Such neural networks are collections of mathematical models which aim to mimic adaptive biological learning. These networks are comprised of a number of highly interconnected processing elements (analogous to synapses) linked to weighted connections (analogous to neurons). Such networks are "trained" by identifying particular output patterns (e.g. secondary structure) and relating them to the input pattern (e.g. amino acid sequence), enabling the neural network to learn relationships between the patterns. In addition, neural networks have the advantage of being reproducible, rapid by carrying out large numbers of tasks in parallel, and capable of dealing with complex data.

PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred>) was used to predict alpha helices in protein sequences. PSIPRED was selected as it is ranked as the most accurate prediction program in the Critical Assessment of Techniques for Protein Structure Prediction (CASP) 4 meeting (<http://predictioncenter.org/casp4/Casp4.html>). A neural network utilises PSI-BLAST (position-specific iterative basic local alignment search tool) position scoring matrices rather than sequence alignments used in a number of methods. PSI-BLAST operates by carrying out multiple BLAST searches and creating a multiple alignment, producing a profile which can predict homologies between distantly related sequences (Altschul et al., 1997). PSI-PRED acts by utilising these intermediates, producing a large amount of data which is filtered by the 2 neural network layers, giving a prediction of secondary structure (helix, strand or coil, with a confidence level of 0 (poor) to 9 (high)). Helix predictions with a confidence level of 5 or greater, i.e. those with a greater chance of adopting a structure based on the PSI-PRED prediction algorithm, were taken as a threshold value.

PSIPRED does not predict beta-turns, so another web-based server, BetaTPred2 (<http://www.imtech.res.in/raghava/betatpred2/>) was used. This program takes a similar approach to PSIPRED as it utilises a neural network trained on PSI-BLAST position scoring matrices, but to recognise beta-turns (Kaur & Raghava, 2004). In addition to PSI-BLAST position scoring matrices, PSIPRED prediction results are also used as an input. The authors claim an overall accuracy of 75.5% (Kaur & Raghava, 2003).

2.2.3 Alignment of epitopes

Once the epitopes for both retroviruses and the putative autoantigens were identified (Appendix B), they were aligned to identify sequence homology, and potential molecular mimicry. Lalign (http://www.ch.embnet.org/software/LALIGN_form.html) was used in preference to BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) as BLAST is capable of only identifying the “best” match, this is not necessarily the one that is of the most interest. LAlign passes the query through a substitution matrix (e.g. blocks substitution matrix - BLOSUM 50, the default used), which aligns the peptide sequences in a manner which the conservation score is the highest. Each residue is given a score based on how well it is conserved, and these are summed.

Mutations of the gene affect protein sequence differently. Certain residues are less likely to be conserved (Valdar and Jones, 2004), such as tryptophan (which is coded for by only one RNA triplet, UGG), and the scores given for conservation are higher. Conversely, if gene mutation is less likely to alter the residue, as is the case with leucine (coded for by 9 different RNA triplets), scores assigned are lower. If the mutation changes the residue to a residue with different physicochemical properties, the substitution matrix applies a penalty.

2.2.4 Three-dimensional structure prediction using homology modelling

In order to obtain an accurate three-dimensional model where there is no known structure, it is important to carry out fold recognition in order to obtain an accurate model (Shortle, 1997). In order to evaluate where the predicted epitopes of selected proteins were in three dimensions, the sequences were inputted into an online fold recognition program, GenTHREADER (<http://bioinf.cs.ucl.ac.uk/psipred>). GenTHREADER works by carrying out multiple sequence alignments and comparing these to fold library sequences in the Brookhaven Protein Data Bank (Jones, 1999b). Like the PSIPRED protein structure prediction server, GenThreader utilises a neural network to give a measure of confidence in the evaluation. Three-dimensional structures of homologous proteins are highly conserved, with proteins sharing 50% identity of residues over a domain exhibiting almost identical structures (Claverie & Notredame, 2003), and even down to 20% identity a good model can be predicted.

Where a homologous protein was identified, with a p-value assigned by GenThreader of <0.001 (or 'high'), the PDB sequence was retrieved from the European Bioinformatics Institute's PDBsum database (<http://www.ebi.ac.uk/pdbsum/>). Downloaded PDB sequences were then visualised using Cn3D version 4.1 (National Center for Biotechnology Information), available (<http://www.ncbi.nlm.nih.gov/structure/CN3D/cn3d.shtml>).

2.3 Results

Comparison of the retroviral epitopes identified with those of putative autoantigens yielded a number of interesting results (Table 2.2 – see *Appendix B for a list of epitopes*). With the exception of alpha beta crystallin and claudin 11 (oligodendrocyte specific protein), all of the autoantigens showed potential for molecular mimicry with the retroviruses identified, in particular the Visna virus (largely due to the large number of epitopes identified, as Visna encodes large, polar proteins).

Of particular interest is the number of retroviral sequences with similarity to NoGo A, a myelin protein produced by oligodendrocytes which is involved in the inhibition of neurite growth (Chen et al., 2001). Serum IgG and CSF IgM antibodies against NoGo A are detectable in MS patients (Reindl et al., 2003). It may be more likely that these antibodies indicate levels of tissue damage, rather than acting as an autoantibody, though this is likely to be true for the vast majority of autoantigens.

A large number of epitopes were identified for NoGo A, as it is a large, highly hydrophilic (and polar) protein. Epitopes from both HERV-H and HERV-W (including the MSRV variant) env proteins show a high level of similarity to NoGo A epitopes (Table 2.3a). HERV-H env shares an identical 8 amino acid stretch with NoGoA, and the probability of this arising by chance is extremely low ($E(10,000): 0.0023$). The similarity between HRES-1 and NoGo A is due to a stretch of glutamic acid residues.

An identified MSRV pol epitope shows similarity to both isoforms of myelin associated glycoprotein precursor (Table 2.3b). The glycosylated form of MAG is a lectin-binding protein, which has N-linked oligosaccharide sidechains attached, is responsible for mediating myelin-neuron cell-cell interactions, MAG is located in the periaxonal space of the myelin sheath, and although less abundant, it is more accessible than MBP which lines the interior of the myelin sheath where it interacts with acidic lipids (Quarles, 2007).

MS patients commonly suffer from visual disturbances, and consequently there is frequently damage of the optic nerve. A number of retinal antigens, such as the arrestins have been investigated as putative autoantigens. Ohguro et al. (1993) carried out a small scale study finding 8 out of 14 MS patients were seropositive for antibodies to S-arrestin/beta arrestin, but not in any of the 50 controls tested. The authors showed that the antibodies recognised the major epitope of S-arrestin, located between residues 290-320. While bioinformatics analysis identified an epitope between 296-311, no cross-reactivity with any of the viral epitopes was found by alignment. However, potential mimicry between HERV-W/MSRV envelope and gag proteins was identified at another predicted S arrestin epitope located between residues 229-243.

Epitope mapping has previously been used within our research group to identify potential molecular mimics, and homology between MSRV gag protein and MBP has been found (Table 2.3).

Table 2.2 Overview of alignments of retroviral epitopes against autoantigen epitopes

Potential molecular mimicry was found with a number of retroviruses with all of the autoantigens with the exception of alpha beta crystallin and claudin 11. A large number of matches were found with NoGo A, compared with the other proteins investigated.

	#	HERV-H		HERV-W		HRES-1		HTLV-1			MSRV			Visna		
		env	gag	env	gag	p15	p25	gpp	gp46	tax	env	gag	pol	env	gag	pol
		2	1	2	2	2	3	9	1	2	3	2	4	12	7	12
α B Crystallin	1															
β Arrestin 1A	5															
β Arrestin 1B	7															
Claudin 11	1															
CNP	10															
hnRNP	8															
HSP 60	12															
HSP 70 1A	14															
HSP 70 1B	14															
HSP 70 8-1	14															
HSP 70 8-2	9															
MAG A	4															
MAG B	3															
MBP	3															
NoGo A	18															
PLP	1															
S Arrestin	10															
T'aldolase 1	4															

Key

■ – 4 aa alignment ; ■ – 5 aa alignment ; ■ – 6 or greater aa alignment

Abbreviations:

CNP = 2', 3'-Cyclic Nucleotide 3' Phosphodiesterase; HERV = HERV; HRES = HTLV-related endogenous sequence; hnRNP = Heterogenous nuclear riboprotein; HSP = Heat shock protein; HTLV – Human T cell lymphoma virus; MAG = Myelin associated glycoprotein; MBP = Myelin basic protein; PLP = Proteolipid protein

Table 2.3 Comparison of human endogenous retroviral epitopes identified with autoantigens.

(a) NoGo A

HERV-H env (230)	SPCSSDSPPRPSSRL
	: : : : : :
NoGo A (8)	PLVSSDSPPRPQPA
HERV-H gag (190)	QFTLKKVAGAKGIVK
	. : : : . : :
NoGo A (665)	VSLKKVSGIKEEIKEPENIN
HERV-W env (117)	GGVQDQAREKHVKEVIS
	: : :
NoGo A (352)	KLVKEDEVVSSEKAKDSFNEKRVAV
HERV-W env (412)	GIVTEKVKEIRDRIQRRAEELRNTGP
	. : : : :
NoGo A (503)	QIVTEKNTSTKTSNPFL
HRES-1 p15 (68)	RALHPPSGRDREEEEMGY
	: . . . : : :
NoGo A (23)	FKYQFVREPEDEEEEEEEEEDEDEDLEELEVLER
MSRV env (120)	IQGQAREKQVKEAISQ
	: : :
NoGo A (352)	KLVKEDEVVSSEKAKDSFNEKRVAV
MSRV env (412)	RIVTEKVKEIRDRIQC
	. : : : :
NoGo A (503)	QIVTEKNTSTKTSNPFL
MSRV pol (580)	TFLPDNEEKIEHNCQQV
	: . : : . :
NoGo A (321)	EEIIVKNKDEEEKLVSNIL

(b) Myelin Associated Glycoprotein

MSRV pol (731)	LHCQGHQEEEEEREIEGNRQADIEAKKAARQDSPLEML
	: : : :
MAG A/B (184)	AVLGRLREDEGTWVQ

Table 2.3 Comparison of human endogenous retroviral epitopes identified with autoantigens (*continued*).

(c) S Arrestin

HERV-W env (134)	GGVQDQAREKHVKEVIS
	:: ::..
S Arrestin (229)	TVTNTEKTVKKIKA
MSRV env (120)	IQGQAREKQVKEAISQ
	:: ::.
S Arrestin (229)	TVTNTEKTVKKIKA
MSRV gag (274)	TCVLEGLRKTRKKPMN
	: .. .:: ::
S Arrestin (229)	TVTNTEKTVKKIKA

Table 2.4 Comparison of human endogenous retroviral epitopes (*previously predicted in the research group by Denise Roden*) with myelin basic protein.

MSRV gag (311)	LREALRKHTSLSP
	:::
MBP (41)	LSP

These tables show LAlign outputs comparing predicted retroviral epitopes against those of putative autoantigens. (:) represents identical amino acids, whilst (.) represents chemical similarity (where the Point Accepted Matrix (PAM) (Dayhoff et al., 1978) substitution score is less than or equal to 0).




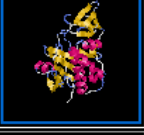

HERV-W env and MSRV gag were entered into the GENThreader program in order to determine where HERV-W env 412 and MSAV gag 274 for homology modelling, in order to determine where are likely to be located in three dimensions.

GENThreader identified a number of alignments to protein data bank structures, with HERV-W env and MSRV env showing at least a high degree of homology to three alignments and one alignment respectively (Figures 2.5 and 2.6). Proteins predicted to have domains with similar structures to HERV-W env included domains envelope proteins of HERV-FRD_6p24.1, MoMuLV and a strain of Zaire ebolavirus, whilst MSRV gag had a high homology to the gag polyprotein of AKV murine leukaemia virus (Table 2.5). The HERV-W env₄₁₂ epitope was not present in the homologous domains of the envelope protein structures, but the MSAV gag₂₇₄ epitope was present towards the C-terminus of the AKV murine leukaemia virus gag polyprotein (Figure 2.7). One further point of note is that the alignment indicates that the complete sequence of the AKV murine leukaemia virus shows homology to MSAV gag₂₇₄.

Table 2.5 Details of proteins identified by GENThreader as having high homology matches, or better, with HERV-W env and MSAV gag

Query protein	PDB ID	UniProt ID	Protein name	Homologous amino acids on retrovirus	HERV-W env ₄₁₂ / MSAVgag ₂₇₄ epitope present?
HERV-W env	1y4m	P60508	HERV-FRD_6p24.1 provirus ancestral envelope polyprotein	357-410	No
	1mof	P03385	MoMuLV envelope glycoprotein p15	296-309	No
	2ebo	Q05320	Zaire ebolavirus (strain Mayinga-76) (ZEBOV) envelope glycoprotein	353-406	No
MSAV gag	1u7k	P03336	AKV murine leukemia virus (AKR (endogenous) murine leukemia virus) gag polyprotein	143-285	MSAV gag ₂₇₄

Figure 2.5 GENThreader alignment summary for HERV-W env

Conf.	Net Score	p-value	PairE	SolveE	Aln Score	Aln Len	Str Len	Seq Len	Alignment	SCOP Codes	Structure
CERT	56.740	9e-05	-75.0	-1.6	291.6	53	53	538	1y4mA0	h.3.2.1	
HIGH	51.232	0.0003	-85.2	5.8	270.0	53	53	538	1mofA0	h.3.2.1	
HIGH	49.331	0.0005	-99.1	0.3	236.0	73	74	538	2eboA0	h.3.2.1	
LOW	33.541	0.021	-259.2	-9.2	46.0	178	235	538	2zgwA0	-	
LOW	31.880	0.031	-224.2	-12.1	31.0	205	290	538	3eixA0	-	

One homologous PDB structure alignment, with a 'certain' homology match ($p < 0.0001$) and 2 with a 'high' match ($p < 0.001$) to HERV-W env were identified by GENThreader. 82 alignments with a 'low' ($p < 0.1$) match were also identified (*data not shown*).


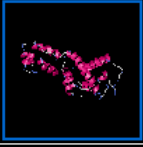



Key

Conf	description of confidence level
Score	raw score from support vector machine
p-val	probability of false positive
Epair	pairwise energy for model
Esolv	solvation energy for model
AlnSc	sequence alignment score
Alen	Length of alignment
Dlen	Length of PDB entry
Tlen	Length of target sequence
PDB_ID	PDB identifier (+ chain code + domain code in CATH format)

Confidence levels

CERT	p-value < 0.0001
HIGH	p-value < 0.001
MEDIUM	p-value < 0.01
LOW	p-value < 0.1
GUESS	p-value ≥ 0.1

Figure 2.6 GENThreader alignment summary for MSAVgag

Conf.	Net Score	p-value	PairE	SolvE	Aln Score	Aln Len	Str Len	Seq Len	Alignment	SCOP Codes	Structure
CERT	66.481	1e-05	-117.8	-5.3	324.0	129	131	352	1u7kA0	a.73.1.1	
LOW	35.649	0.013	-107.8	-8.7	102.0	170	207	352	1eiaA0	a.28.3.1 a.73.1.1	
LOW	33.103	0.023	-88.8	-3.9	108.0	80	95	352	1mn8A0	a.61.1.6	
LOW	31.281	0.035	-75.4	-7.8	82.0	155	199	352	1qrjB0	a.28.3.1	
LOW	30.156	0.046	-137.8	-10.8	59.0	145	189	352	1zr5A0	c.50.1.2	

One homologous PDB structure alignment, with a 'certain' homology match ($p < 0.0001$) was identified by GENThreader. 9 alignments with a 'low' ($p < 0.1$) match were also identified (*data not shown*).

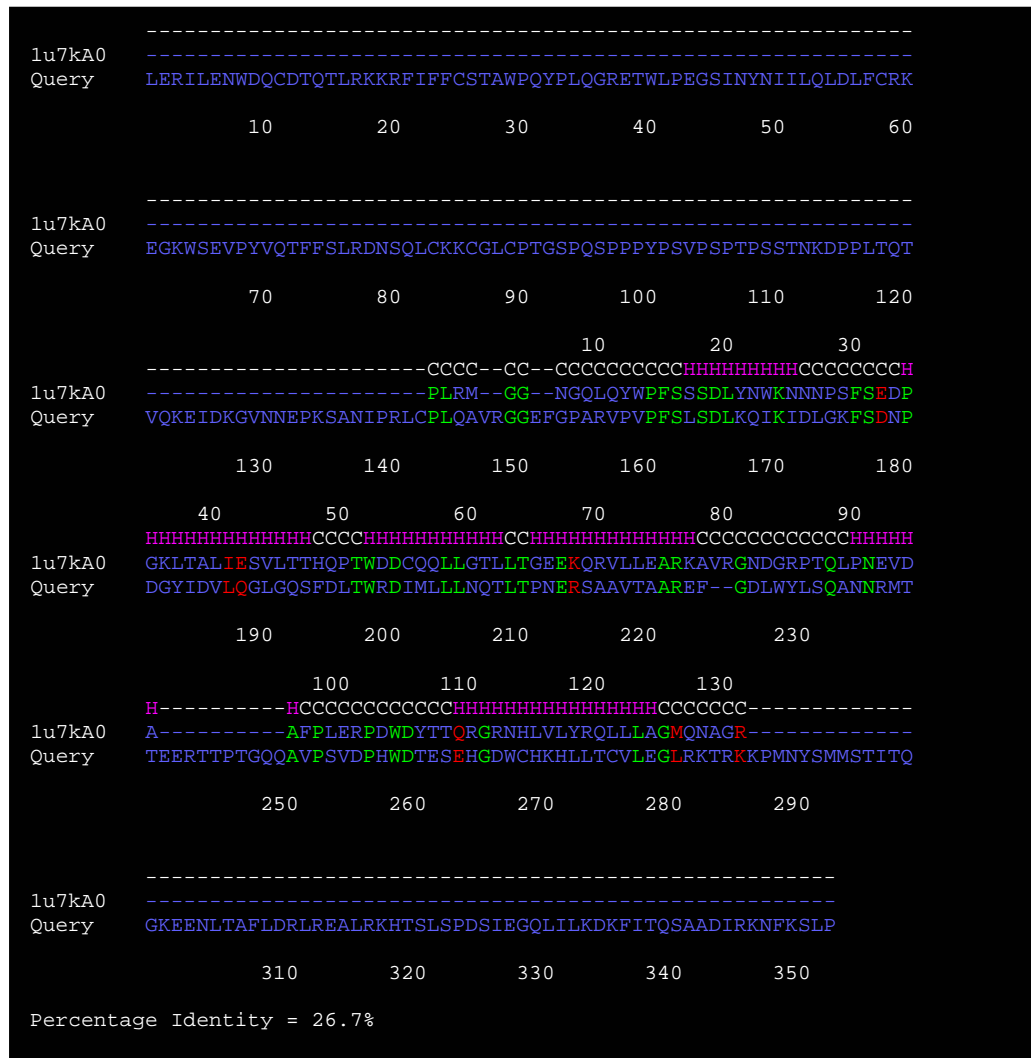
Key

Conf	description of confidence level
Score	raw score from support vector machine
p-val	probability of false positive
Epair	pairwise energy for model
Esolv	solvation energy for model
AlnSc	sequence alignment score
Alen	Length of alignment
Dlen	Length of PDB entry
Tlen	Length of target sequence
PDB_ID	PDB identifier (+ chain code + domain code in CATH format)

Confidence levels

CERT	p-value < 0.0001
HIGH	p-value < 0.001
MEDIUM	p-value < 0.01
LOW	p-value < 0.1
GUESS	p-value \geq 0.1

Figure 2.7 GENthreader alignment of AKV murine leukemia virus (AKR (endogenous) murine leukemia virus) gag polyprotein against MSAVgag



The PDB structure alignment of AKV murine leukemia virus (AKR (endogenous) murine leukemia virus) gag polyprotein (PDB ID: 1u7k) (residues 1-131) shows homology with amino acids 143-285 of MSAV gag (the query sequence). Amino acids highlighted in green are identical, whilst those highlighted in red are physicochemically different. The MSRV gag₂₇₄ epitope is homologous to the C-terminal of the query.

Figure 2.8 Three dimensional structure of AKV murine leukemia virus (AKR (endogenous) murine leukemia virus) gag polyprotein highlighting area homologous to the predicted MSAV gag₂₇₄ epitope



The figure illustrates the three dimensional structure of AKV murine leukaemia virus gag polyprotein as visualised using Cn3d 4.1 (National Institute of Biotechnology Information). The PDB sequence, 1u7k, was obtained by X-ray crystallography by Mortuza et al., 2004.

Areas highlighted in green represent residues forming alpha helices, and those in gold, beta pleated sheets. The area of the chain highlighted in yellow indicates the position amino acids 120-131 of AKV murine leukaemia virus, which is homologous to the MSRV gag₂₇₄ epitope,

Imaging of the AKV murine leukaemia virus (Figure 2.8) in three dimensions illustrates an area homologous to the MSRV gag₂₇₄ epitope is located at the C-terminus towards the end of an alpha helix. It is quite likely, therefore that the MSRV gag₂₇₄ is exposed on the outside of the protein, but this is not conclusive as the structure of the MSRV gag protein has only been estimated for part of the protein, between amino acids 143-285.

2.4 Discussion

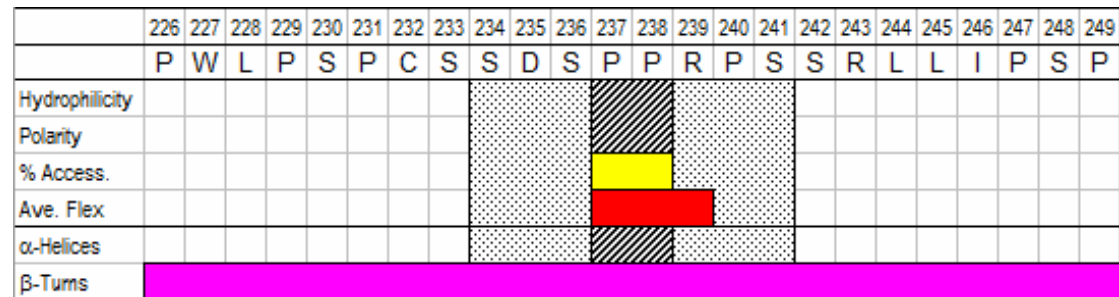
The method applied in this study uses a number of different criteria to predict epitopes, including physicochemical properties in the case of the primary structure prediction; as well as structural predictions based on homology modelling to determine whether the epitopes are likely to be presented on the surface of the protein, i.e. between alpha-helices and beta turns, in the case of secondary structure prediction. The primary prediction methods on their own have a low accuracy, approximately 50% in the case of using hydrophilicity profiles (Hopp and Woods, 1981), but taken together with secondary structure analysis, which has a higher level of accuracy, between 76.5%-78.3% in the case of PSIPRED (Jones, 1999b), the system as a whole can be made more robust.

One problem with this method is that only linear epitopes can be identified. Another is that the thresholds for the primary structure prediction rely on artificial thresholds, albeit ones that have been used within the research group for a period of time. This has yielded problems in that certain hydrophobic proteins, such as PLP had no epitopes predicted for it, even though antibodies to this protein have been detected time after time in MS patients. Conversely, with large, polar proteins many epitopes have been predicted, and a large number of these may be irrelevant. An alternative approach may have been to look for the most hydrophilic areas of the proteins, for example by setting the threshold for each protein individually, basing each threshold on the maximum and minimum values as a percentage. Some epitopes may be missed, but the amount of irrelevant data would be minimised, and epitopes for hydrophobic proteins can be determined.

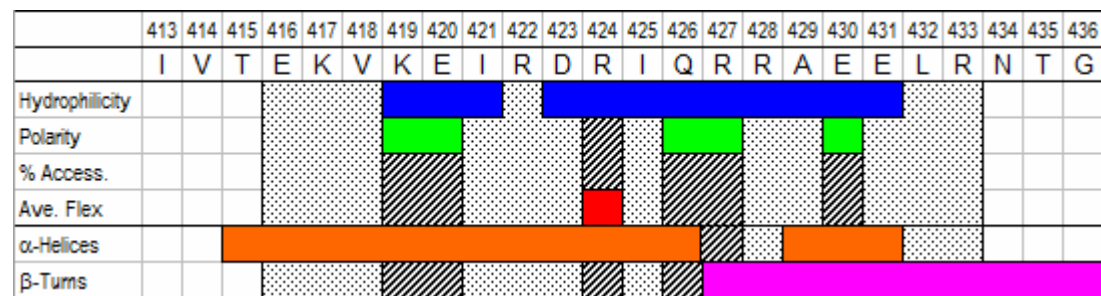
The use of bioinformatics or *in silico* analysis does not prove that molecular mimicry is taking place, but rather a powerful tool to provide a large amount of data that can be used alongside experimental work. In order to test for potential molecular mimicry, two of the predicted epitopes, HERV-W env₄₁₂ and MSRV gag₂₇₄ were selected to be synthesised as short peptides, and tested in a pilot study to test for seropositivity in MS patients and controls in an ELISA based protocol.

Figure 2.9 Comparison of HERV-H env₂₃₀, HERV-W env₄₁₂ and MSAV gag₂₇₄ epitope properties

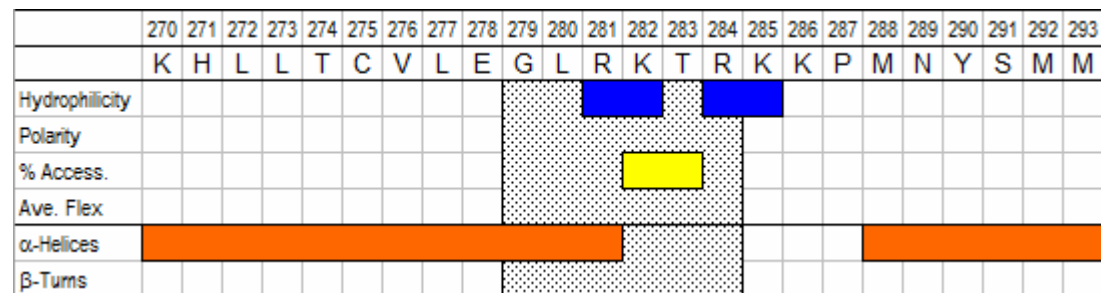
HERV-H env₂₃₀



HERV-W env₄₁₂



MSAV gag₂₇₄



Legend

Primary structure

■ hydrophilicity >1.2 ■ polarity >10 ■ accessibility >7 ■ flexibility >0.5

Secondary structure

■ alpha helix (conf >5) ■ beta turn

■ >3 thresholds on amino acid

■ >3 thresholds within 3 residues

In order to further evaluate whether the epitopes of the peptides for production were likely to be located on the outside of the protein, homology modelling on the basis of known structural data was performed using a protein threading method. The results indicated that the MSRV gag₂₇₄ epitope is likely to be exposed on the protein surface. *Ab initio* analysis may yield further data, particularly on the HERV-W env sequence, where no suitable model was determined.

Ab initio prediction methods, such as Robetta (Kim et al., 2004) and Folding@home (Marianayagam et al., 2005) can be used to determine the three dimensional structure where homology modelling cannot. These methods rely on a large amount of computing power, and in the case of Folding@home, a distributed computing network utilising unused central processing unit power on computers that have the client installed is used (Shirts and Pande, 2000). *Ab initio* methods still have limitations, and are currently more suited for shorter proteins, with Robetta optimised for short, single domain proteins of less than 120 amino acids (Kim et al., 2004).

There were a number of reasons for selecting HERV-W env₄₁₂ and MRSV gag₂₇₄ for further investigation. Both HERV-W env₄₁₂ and MSRV gag₂₇₄ are highly hydrophilic regions (Figure 2.9), and are predicted to lie between structural features (alpha helices and beta-pleated sheets), and therefore are more likely to be on the outside of the protein. By contrast, the epitope predicted for HERV-H env₁₉₀ is not hydrophilic and is predicted to sit in the middle of a beta sheet. Even though this epitope shows an extremely high homology (8 amino acids) to a predicted epitope of NoGo A (Table 2.3a), the sequence has few characteristics of a good epitope. A further consideration in the choice of peptides was that both HERV-W env and MSRV gag proteins have been detected in demyelinating region of MS lesions, with a high accumulation of gag in demyelinated white matter (Perron et al., 2005) whilst random phage peptides with amino acid similarity to HERV-W env were found to be present in raised titres in MS patients (Jolivet-Reynaud et al., 1999).

Ultimately, the peptides predicted can be used to produce rabbit polyclonal antibodies in order to provide a standard curve in order to accurately estimate titres, and to detect the epitopes by immunocytochemistry in cancer cell lines (in the manner of the antibodies used in Chapter 5, as HERVs are known to be aberrantly expressed in tumours and are transcriptionally active in the brain (see Chapter 4 for a review of the literature), as a precursor to investigating pathological sections of MS tissues and controls.

3 Determination of antibody reactivity against retroviral epitopes in MS patient plasma by ELISA

3.1 Introduction

A number of potential epitopes were identified *in silico* using bioinformatic analysis, as outlined in Chapter 2. Short (~15aa) sequences from two of these predicted epitopes, HERV-W env₄₁₂ and MSRV gag₂₇₄ (highlighted in Table 3.1) were selected to be synthesised as peptides, and screened in small pilot study involving an ELISA-based assay to test for potential reactivity of antibodies present in plasma samples of a number of well-defined MS patients and controls (including patients in remission, people with other neurological disorders and healthy volunteers). Assaying the peptides was necessary to determine immune responses in patients to particular predicted molecular mimics as a precursor to production of polyclonal antibodies. The production of polyclonal antibodies would provide a positive control for the ELISA, enabling titres to be estimated (by comparing to a calibration curve) and also a tool to investigate protein expression (e.g. immunohistochemistry in pathological sections, western blotting of protein extracts from peripheral blood, cerebrospinal fluid or pathological tissue).

Table 3.1 HERV-W env and MSRV gag peptides used in the ELISA system and their homology to MS autoantigens

HERV-W env (412)	GIVTEKVKEIRDRIQRRAEELRNTGP
	. : : : :
NoGo A (503)	QIVTEKNTSTKTSNPFL

MSRV gag (274)	T ^C VLEGLRKTRKKPMN
	: . . . : : :
S Arrestin (229)	TVTNNTEKTVKKIKA

HERV-W env₄₁₂ showed homology to a predicted epitope of NoGoA, with a stretch of 5 identical amino acids. NoGoA is a myelin protein produced by oligodendrocytes which is involved in the inhibition of neurite growth in undamaged CNS tissues (Chen et al., 2000). Serum IgG and CSF IgM antibodies against the terminal domain of NoGo A are detectable in relapsing-remitting MS patients, but the levels did not differ significantly from patients with other inflammatory neurological diseases (Reindl et al., 2003). There was no response in SLE patients, suggesting that anti-Nogo A antibodies are detectable in serum as a response to CNS injury, though not specific to MS.

MSRV gag₂₇₄ was identified as an immunodominant region of MSRV gag, showing a short stretch of amino acids homologous to S arrestin. S arrestin has been long been identified as an important autoantigen in uveitis, with a strong T cell response to the protein in uveitis patients. Immunisation of a particular epitope of bovine S arrestin (also known as S antigen), peptide M (DTNLASSTIIKEGIDKTV) causes experimental autoimmune uveitis in Lewis Rats (Donoso et al., 1987). Peptides with homology to peptide M, including a gag-pol polyprotein from a baboon endogenous retrovirus and hepatitis B virus DNA polymerase, act as molecular mimics *in vitro* (Singh et al., 1990).

Autoantibodies to S-arrestin or similar antigen (such as β -arrestin) were present in 8 out of 14 MS patients, with higher titres of antibody during active disease than in remission in 2 of these patients (Ohguro et al., 1993). The authors of this study showed that the major recognition site of the autoantibodies was at the immunogenic site identified in uveitis patients. As neurological symptoms, such as optic neuritis are a common complaint in MS patients, retinal proteins may be considered potential candidate autoantigens in the disease.

Samples of each of the two peptides were coated onto 96 well plates and ELISAs were carried out on these in order to screen for reactivity of antibodies present in plasma samples of well-defined relapsing-remitting MS patients during remission, patients with other neurological diseases, and normal healthy donors.

3.2 Methods

3.2.1 Peptide synthesis

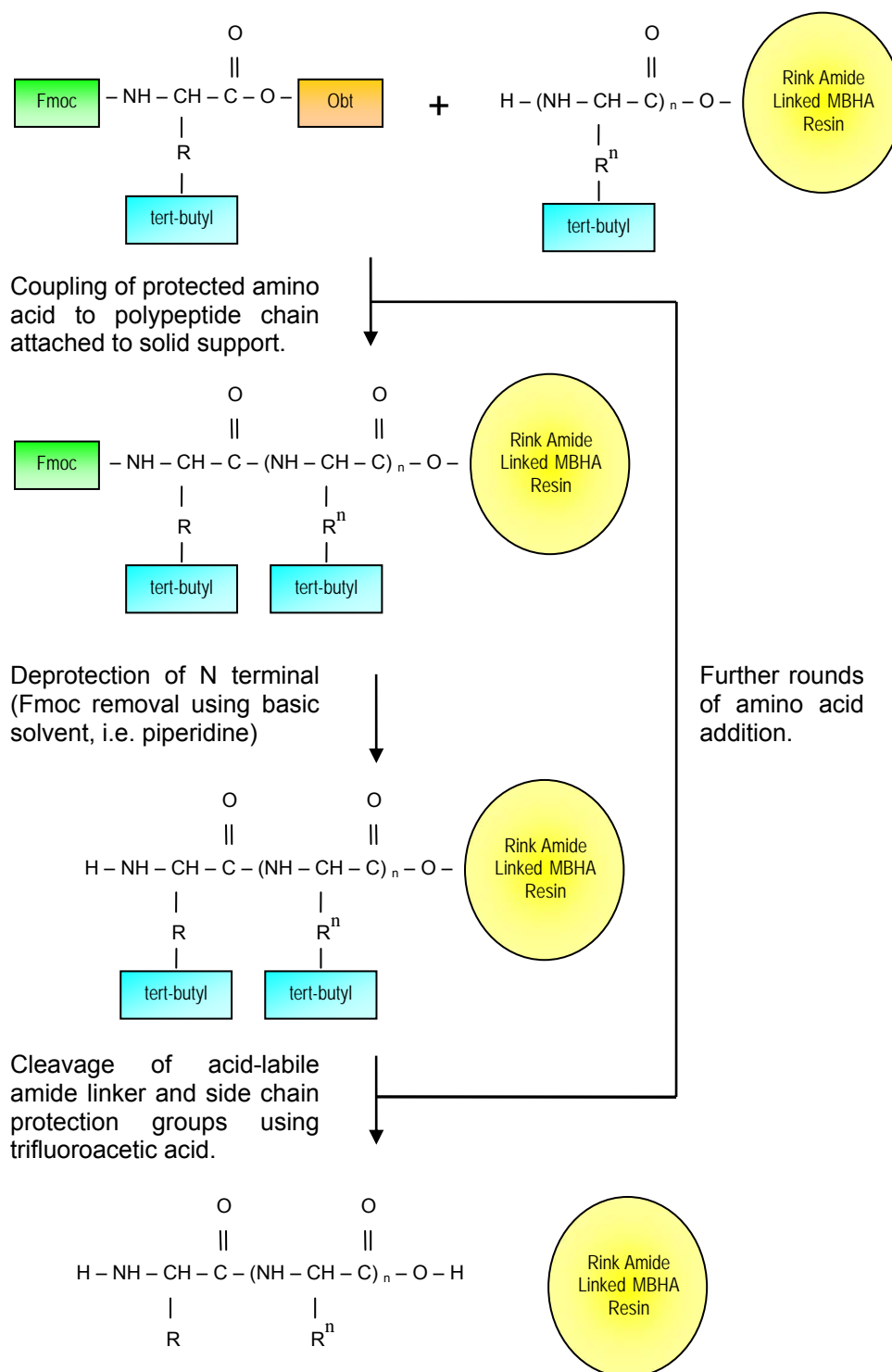
Solid phase peptide synthesis is a method of constructing peptides which are anchored to a resin (the solid phase) by an acid-cleavable linker group. This enables peptides to be produced with each synthesis step occurring in the same reaction vessel, where the reagents can be removed by filtration while the beads remain. An overview of the method is detailed in Figure 3.1.

The growing polypeptide chain is made by adding a series of amino acids to the C-terminal, in contrast to *in vivo* protein synthesis where ribosomes add amino acids from the N-terminal. To prevent polymerisation of the amino acids, a protective group such as Fluorenylmethoxycarbonyl (Fmoc) is employed. In order to remove the protective group and facilitate further rounds of peptide synthesis, a deprotection step is necessary. Incubation in basic conditions, e.g. in 20% piperidine removes Fmoc protection, whilst side chains with acid-cleavable protection (e.g. tert-Butyloxycarbonyl (tert-Boc)) remain.

The amino acids are added in excess (typically a 2-fold amount is used) in order to keep yields high. If each coupling step of a 15 aa peptide was 99% efficient, the final yield would be 86.0% (0.99^{15}), whereas if the coupling efficiency decreased to 95%, the final yield would be much lower at 46.3% (0.95^{15}). It is therefore vital to check whether coupling and deprotection steps have been successful by carrying out a Kaiser test. The Kaiser test relies on the detection of free amine groups, where ninhydrin reacts with unprotected amino acids, giving rise to blue beads. Where the amino acid has coupled, the Fmoc from the coupled amino acid prevents the reaction taking place, and no colour change occurs.

For an amino acid to couple, it is necessary to activate the terminal carboxyl acid group on the growing polypeptide chain. This is done with the addition of Diisopropylethylamine (DIPEA) and aromatic oximes (e.g. 1-hydroxybenzotriazole (HObt)) or uronium salts (e.g. 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminiumhexafluorophosphate (HCTU)). Aromatic oximes and uronium salts also act to protect against racemisation, where amino acids spontaneously switch between D- and L-forms.

Figure 3.1 Basic overview of solid phase peptide synthesis methodology.



Peptide synthesis involves repeated steps of addition of amino acids to a peptide chain coupled to a resin. Once the desired residues have been added, the peptide is cleaved from the resin and protective side chains.

When the chain is complete, the polypeptide is cleaved from the resin in acidic conditions (e.g. trifluoroacetic acid). This process also removes the acid-cleavable side chain protection groups. Triisopropylsilane (TIS) is added to act as a scavenger, preventing reactions with the newly exposed side chains present on the polypeptide chain. The peptide is then precipitated from solution in diethyl ether, washed and vacuum-dried. As a final check, the mass is confirmed with matrix-assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF).

Laboratory protocol

100 μ mole rink amide labelled MBHA resin (Novabiochem) was added to a glass reaction vessel and allowed to swell in 4 ml dichloromethane (DCM) for 1 hour. Following the incubation the DCM was removed under vacuum.

To remove the Fmoc protective group, the resin was treated by the addition of 4 ml 20% (v/v) piperidine in N,N-dimethylformamide (DMF) for 5 minutes, drained under vacuum and reincubated in 4 ml 20% (v/v) piperidine for an additional 15 minutes. Following removal of 20% (v/v) piperidine, the resin was washed 3 times for one minute in DMF.

In order to establish whether the Fmoc was successfully removed, a Kaiser test was conducted to establish whether free amines were present on the resin. A small sample of resin was removed, washed in ethanol, and centrifuged briefly in order to be able to remove the ethanol. One drop of each of solution A (5% (w/v) ninhydrin in ethanol), solution B (80% (w/v) phenol in ethanol) was added to the beads, and solution C (20 μ M potassium cyanide in pyridine), and heated at 120 °C for 5 minutes. The presence of blue beads indicates that there are free amines present, and that the Fmoc removal was successful. If the beads were clear (i.e. Fmoc groups still attached), the resin was treated with piperidine and reincubated with 20% (v/v) piperidine as before.

200 μ mole (2 eq) of the relevant Fmoc-protected amino acid (Novabiochem) were added to 2 eq HCTU/HOBt in DMF, and activated with 4 eq of DIPEA. This was then added to the beads in the reaction vessel, and incubated for 1 hour on an orbital rocker. The resin was then washed 3 times for one minute in DMF.

To establish whether coupling was successful, a Kaiser test was performed on a small sample of beads. The presence of clear beads indicates that there are no free amines present, and that coupling of the amino acid was successful. If the beads were not clear, the amino acid was recoupled as before. If the coupling was successful, further rounds of peptide synthesis (removal of the Fmoc protective group and coupling of the next amino acid) were carried out until the peptide chain was complete.

Once the polypeptide chain is completed, the Fmoc protective group was removed by treating the resin with the addition of 4 ml 20% (v/v) piperidine in N,N-dimethylformamide (DMF) for 5 minutes, drained under vacuum and reincubated in 4 ml 20% (v/v) piperidine for an additional 15 minutes. The resin was washed 3 times with DMF for 1 minute, 3 times with DCM for 1 minute, and finally 3 times with methanol for 1 minute. The resin was then freeze dried overnight.

The polypeptide chain was cleaved from the resin by the addition of 95% TFA : 2.5% TIS : 2.5 % H₂O, followed by filtration through glass wool. Excess TFA was removed in nitrogen gas, and the peptide was then precipitated in a 10 fold volume of diethyl ether at 4 °C, and centrifuged at 2,500 g for 2 minutes. The supernatant was discarded, and the pellet re-centrifuged as above a further 3 times. The predicted mass of the peptide was confirmed with matrix-assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF) (Kratos Kompact Probe).

3.2.2 Patient information

Peripheral blood of MS patients and ONDs were taken by venepuncture at St. Thomas's Hospital outpatient Neurology Clinic by Dr M. K. Sharief (consultant Neurologist) and those of NHDs at the University of Greenwich. Patient plasma samples used, provided by Dr. Laurence Harbige, University of Greenwich, were from relapsing-remitting MS patients, aged between 27 and 55 years (n=24); patients with other neurological diseases (mean age=49 years), (including stroke (n=2), epilepsy (n=2) and Parkinson's disease (n=1)) (n=5); and normal healthy donors (n=10, mean age=39 years). Ethical approval for the project was granted by the University of Wolverhampton School of Applied Sciences Ethics Committee on 26th January 2005 (Appendix C).

3.2.3 ELISA of MS patient plasma

Performed at the Medway School of Sciences, University of Greenwich, Chatham Maritime, Kent

ELISA provides a way of determining the levels of specific antibody (direct ELISA) or antigen (sandwich ELISA) in a particular sample. In a direct ELISA, a multiwell plate is coated with the antigen of interest, and a sample is applied. Specific antibody binds to the antigen on the plate, and this can be detected using an enzyme-labelled anti-species antibody. The enzyme, usually Horse Radish Peroxidase (HRP) or Alkaline Phosphatase (AP), catalyses a reaction in a substrate that is then added to the wells, producing a colour change that can be measured spectrophotometrically.

Laboratory protocol

50 μ l of 400 ng HERV-W env₄₁₂ or MSRV gag₂₇₄ peptide in phosphate buffered saline (PBS), pH 7.2 stock was added to the wells of a Nunclon Δ 96 well plate (Nunc), and was left to coat overnight at 37 °C. The plates were washed thoroughly four times in PBS (Sigma) with 0.05% Tween-20 using an automated plate washer (WellWash Mk4), and blotted dry.

The plates were then blocked with 150 μ l of blocking buffer (PBS with 1% bovine serum albumen (BSA); 0.05% Tween-20) and incubated at room temperature for 2½ hours, before being washed four times in PBS (Sigma) with 0.05% Tween-20 and blotted by hand. This was done to reduce potential unwanted interactions of antibody with the plasticware. Plasma samples and controls were diluted 1:50 in blocking buffer. 50 μ l of sample dilution was added in triplicate to the wells of the plate, alongside appropriate blocking buffer alone to act as a negative control. This was then incubated at room temperature for 2½ hours before further washing and blotting by hand, as before.

Secondary antibody, F(ab')₂ goat anti-human IgG:HRP (Serotec), was diluted to 1:2000, according to the manufacturer's instructions, and 50 μ l of diluted secondary antibody was added to each well of the plate, and incubated at room temperature for 1½ hours. The plate was then washed and blotted by hand as before.

OPD (Sigma) tablets were diluted to 0.4 mg/ml in phosphate citrate buffer with sodium perborate (Sigma) to give active substrate buffer. 50 µl of substrate was added to each well of the plate, which was incubated for 40 minutes at 37 °C in the dark. The plate was read at 450 nm (OPD) in a Multiskan MS (Labsystems) microplate reader. Following colour development, the reaction was stopped by the addition of 50 µl 2N hydrochloric acid. The stopped product was also read at 492 nm (OPD).

3.3 Results

Peptides for HERV-W env₄₁₂ and MSAV gag₂₇₄ were successfully manufactured, and the molecular weights were determined to be accurate by MALDI-TOF mass spectroscopy (Figure 3.2). IgG responses to the HERV-W env₄₁₂ peptide was low (virtually indistinguishable from peptide negative controls, a measure of non-specific binding to the plate by primary antibodies in the serum, which were subtracted from the signal) in all of the types of plasma sample (NHD, OND and MS) tested, with no significant differences between the groups (Figure 3.3) (One way analysis of variance (ANOVA); $F=0.02$; $p=0.976$), indicating that antibody levels recognising the epitope present were non-existent.

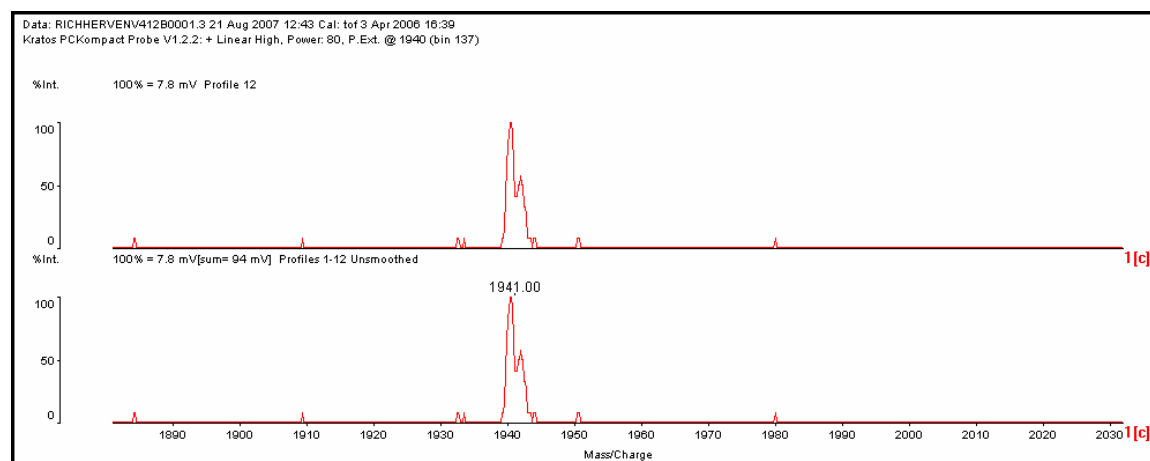
A stronger IgG response against the MSRV gag₂₇₄ peptide was present amongst NHD and MS patients (One way T-test, $P=0.004$ and $P<0.0005$ respectively), but not OND patients (One way T-test, $P=0.137$). As with the HERV-W env₄₁₂ peptide, no significant differences were found between each group (Figure 3.4) (One way ANOVA; $F=0.30$; $p=0.740$) with marginally lower antibody reactivity present in OND controls.

The sensitivity of the assay was poor for both peptides (and thus a high signal to noise ratio), with levels of antibody response uncorrected for background having mean absorbances of 0.160 (Std Dev=0.029) and 0.174 (Std Dev=0.039) for HERV-W env₄₁₂ and MSRV gag₂₇₄ respectively. This compared to background levels for peptide negative controls of 0.130 (Std Dev=0.024) and 0.142 (Std Dev=0.026), serum-free controls of 0.094 (Std Dev=0.018) and 0.098 (Std Dev=0.020) and secondary antibody negative controls of 0.095 (Std Dev=0.026) and 0.091 (Std Dev=0.019) for the same peptides.

Figure 3.2 MALDI-TOF analysis of HERV-W₄₁₂ and MSAV gag₂₇₄ peptides

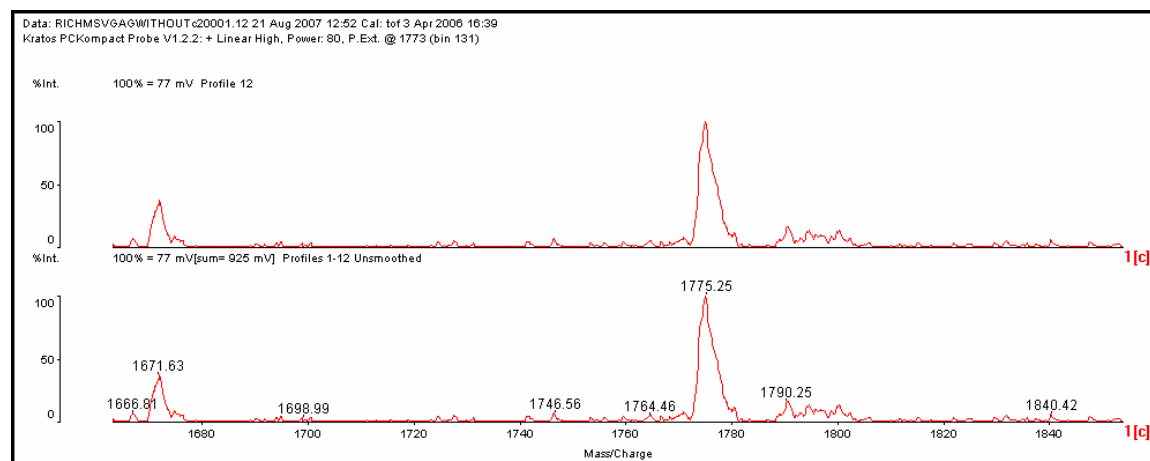
(a) HERV-W₄₁₂

H-GIVTEKVKEIRDRIQR-NH₂ (predicted MW=1940.31)



(b) MSAV gag₂₇₄

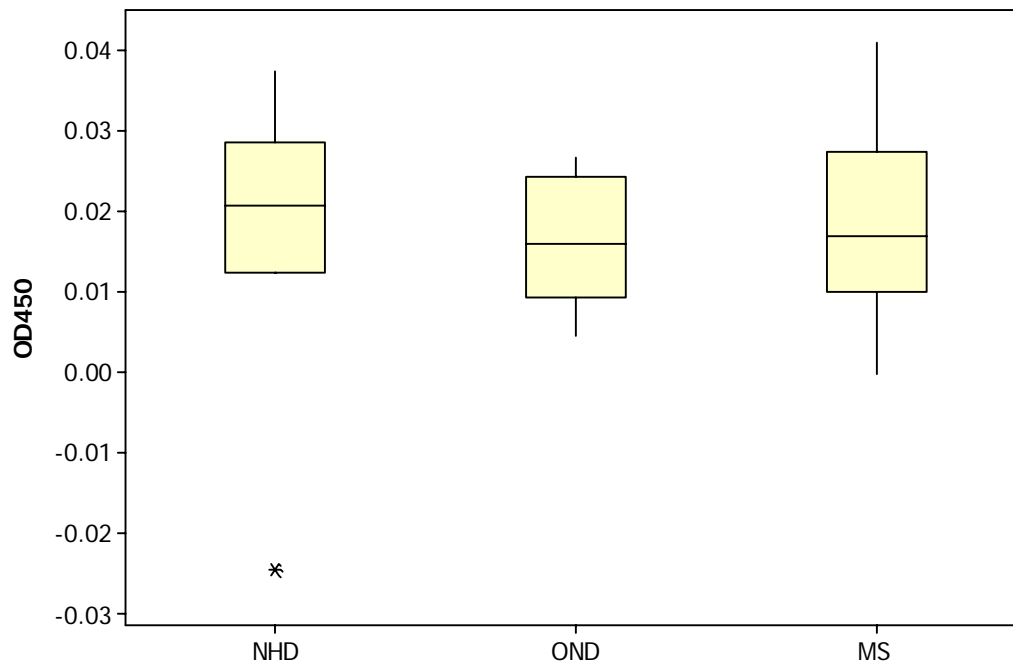
H-CVLEGLRKTRKKPMN-NH₂ (predicted MW=1773.23)



In Figure 3.2 (a), a peak at approximately 1941 Da is present, indicating the presence of HERV-W₄₁₂ in the sample (predicted MW=1940.31). As the peptide is not purified (e.g. by high performance liquid chromatography), a number of smaller peaks are present. This is illustrated in Figure 3.2 (b), where in addition to MSAV gag₂₇₄ (peak at 1775.25, predicted MW=1773.23), there is a peak at approximately 1671.73 Da corresponding to MSAV gag₂₇₄ without the cysteine residue at the C-terminal due to incomplete coupling (MW=1670.08).

MALDI-TOF performed by Dr. Sarah Jones

Figure 3.3 Boxplot comparison of reactivity against HERV-W env₄₁₂ peptide (absorbance at OD₄₅₀ adjusted for peptide negative background) between normal healthy donors (n=10), patients with other neurological diseases (n=5) and active MS patients (n=24).

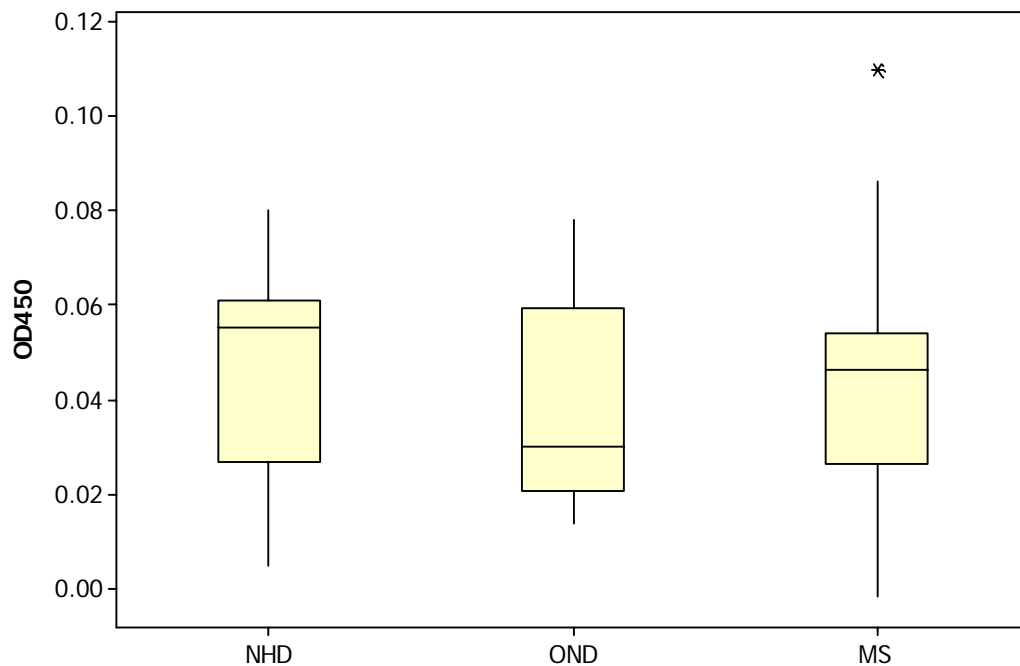


NHDs (median=0.02058, μ = 0.01803, StdDev=0.01717) have a higher level of antibody reactivity (though not significant) against HERV-W env₂₇₄ than patients with OND (median=0.01600, μ =0.01663, StdDev=0.00843) and MS (median=0.01692, μ =0.01790, StdDev=0.01090) (One way ANOVA; F=0.02; p=0.976).

* represents an outlier

Number of replicates = 6; 3 plates run with duplicate wells on each plate

Figure 3.4 Boxplot comparison of reactivity against MSRV gag₂₇₄ peptide (absorbance at OD₄₅₀ adjusted for peptide negative background) between normal healthy donors (n=10), patients with other neurological diseases (n=5) and active MS patients (n=24).



NHDs (median=0.05542, μ =0.04792, StdDev=0.0221) and MS patients (median=0.04658, μ =0.04580, StdDev=0.02387) have a higher level of antibody reactivity (though not significant) against MSRV gag₂₇₄ than patients with OND (median=0.0302, μ =0.0381, StdDev=0.0244) (One way ANOVA; F=0.30; p=0.740).

* represents an outlier

Number of replicates = 6; 3 plates run with duplicate wells on each plate

3.4 Discussion

There was no difference in the level of response to either the HERV-W env₄₁₂ or MSRV gag₂₇₄ amongst patients or controls. The HERV-W env₄₁₂ peptide showed homology to a predicted immunodominant domain of NoGo A, and as with the study by Reindl et al., 2003, no statistical significance of IgG reactivity was found between MS patients, patients with OND, and NHD controls. That particular study looked at both IgM and IgG responses to the large N-inhibitory domain located between residues 1-979 of the sequence, which includes the epitope predicted and tested in this study. In contrast to IgG, Reindl (2003) showed that IgM responses were significantly raised in MS patients and OND controls, suggesting a response to brain injury. A more recent study by Ruprecht et al. (2008) has indicated extremely low levels of immune response to recombinant HERV env and MSAV gag proteins. This study used a larger sample group, with 50 multiple MS patient and 59 controls, with only one MS patient having detectable antibody reactivity against HERV-W env, and none for MSAV gag. No response was found in controls.

HERV-W gag₄₁₂ was found to have homology to an epitope of NoGoA predicted by the bioinformatics analysis (NoGoA₅₀₂). The functional domain of NoGoA (NoGo-66) is not located within this particular epitope, but towards the C-terminus of the molecule (GrandPre et al., 2000). The functional sequence is located extracellularly on oligodendrocytes (Oertle et al., 2003) and on the inner and outer sequences of the myelin sheath (Huber et al., 2002). As NoGoA is thought to play a role in regeneration of the myelin sheath as a result of CNS injury, it might be expected that it is upregulated in conditions such as MS. Immunohistochemical analysis has recently shown that NoGoA expression is upregulated in the oligodendrocytes surrounding active MS lesions (Satoh et al., 2005) but not in other cell types, such as astrocytes. Consequently it is surprising that no significant titres were found either in MS patients or OND controls. As has been discussed in Chapter 2, a large number of epitopes (18) were predicted for NoGo A, and it is more than likely that the responses seen in other studies were reactive to a different epitope.

The MSRV gag₂₇₄ was homologous to S-arrestin, a retinal antigen known to be important in the pathogenesis of autoimmune uveitis (Banga et al., 1987). As many MS patients suffer from ophthalmic conditions, research into possible retinal antigens has been carried out by a number of research groups (Ohguro et al., 1993; Gorczyca et al., 2004). In the Ohguro study, IgG antibody levels against arrestins were present in 8 out of 14 MS patients, of which 2 of these had a higher titre during active disease than during remission. Gorczyca found a number of different retinal antigens were involved in MS including 46 kDa and 41 kDa forms of arrestin and S-arrestin, as well as a number of others. The 46 kDa and 41 kDa forms were present in patients with reduced visual acuity. No significant evidence was found to support a role of molecular mimicry between MSRV gag₂₇₄ and S-arrestin, though it may be worth testing in a small subset of patients.

It is worth noting that antibody levels were measured in plasma samples rather than patient serum. Plasma is considered a more accurate representation of levels in the blood *in vivo* (Sakon et al., 1999), but the presence of clotting factors, such as platelets may potentially activate pathways, for example in assays where platelets scavenge vascular endothelial growth factor (VEGF) (George et al., 2000). As the sensitivity of the assay was poor, it was difficult to effectively discriminate the samples. This could be for a variety of different reasons. The peptide might not have bound efficiently to the plate, resulting in a low signal. To remedy this, higher peptide concentrations could be increased until saturation, or higher affinity plates such as Nunc Immuno- Plates or Corning Co-Star 3590 . High affinity plates are designed to bind proteins via hydrogen-bonding rather than weaker, short-range Van der Waals forces, e.g. Nunc Polysorp plates have a large number of hydrophobic binding sites available for binding hydrophilic molecules. Other studies have utilised biotinylated peptides and bound them to streptavidin coated plates as a method of high-affinity binding (Jayasena et al., 2001).

Antibody binding sites can also be disrupted or obscured when binding to polystyrene ELISA plates, and short peptides may not adopt the same structure independently of the protein as other amino acids elsewhere may influence the structure, holding the epitope in place. Such factors are difficult to investigate without an established positive control. Altering other factors such as blocking buffer components, incubation times and temperatures, concentrations of primary and secondary antibodies used, and choice of substrate can also change the sensitivity of the assay. Another consideration is that antibodies present may have only been weakly cross-reactive to the peptide. Non-specific binding is unlikely, as the controls without peptide have considerably lower activity than the samples.

In summary, no suitable candidate for producing an antibody was found in this study. It is important to note that the two peptides produced were amongst a number identified for possible production. Reactivity was found against the peptides, but at low levels, and without significant differences between patient groups.

4 Detection of HERV-W *env* and MSRV *gag* transcripts in glioma cell lines and modulation of MSRV *gag* expression levels in a microglial cell line with IFN- γ

4.1 Introduction

A number of primers were developed in order to detect and investigate levels of HERV-W expression in astrocytoma cell lines. HERVs are known to be expressed in a number of cancer cell lines, including brain tumours (Yi et al., 2002). As astrocytes (and microglia) are cell types where HERV-W expression is increased, astrocytoma cell lines were investigated to assess the transcription of HERV-W gene products. A multiplex RT-PCR system was optimised to give semi-quantitation of HERV-W *env* transcripts in a number of cell lines, while another system was developed to detect and evaluate MSRV *gag* transcripts. BeWo, a cell line derived from a placental choriocarcinoma was selected as a control cell line, as it is commonly used as a model of syncytin (HERV-W *env*) modulated syncytialisation (Kudo & Boyd, 2002). Glioma cell lines investigated comprised of astrocytoma short term cultures (IN859, IN1472, IN1528, and IN1612) and established glioblastoma cell lines (U251-MG and U373-MG). Other cell lines investigated in the HERV-W *env* system included 2 cell lines where other HERVs, notably HERV-K10 and HERV-K (HDTV) are known to be expressed, namely MCF-7 (a breast cancer cell line) and Tera 1 (a teratocarcinoma cell line).

After design of the primers it is necessary to optimise the RT-PCR system in order to get an efficient reaction resulting in strong, clean bands. A number of factors are important such as the annealing temperature, the number of cycles, and concentrations of reagents (notably that of primers, template, *Taq* DNA polymerase, and magnesium chloride concentration). Sequencing of the products is also necessary to indicate whether the desired amplicon has been targeted successfully. Development of a RT-PCR allows the expression of transcription to be determined (either between different cell types or treatment conditions, e.g. by cytokines, hormones, or transactivation by exogenous agents), and in conjunction with a known housekeeping gene, semi-quantification of their level.

HERV-W gene products have been detected in a variety of different tumours, though in particular ovarian cancers. LTRs are hypomethylated in malignant ovarian tissue in comparison to non-malignant tissues (Menendez et al., 2004), which would be expected to result in an increase in expression of HERV-W gene products. However a more recent report by Hu et al. (2006) indicated that HERV-W RNA sequences are downregulated in ovarian cancer. In endometrial carcinomas, syncytin is increased at both mRNA and protein levels in

endometrial carcinomas compared with control tissues (Strick et al., 2007). By silencing syncytin expression with siRNAs, fusion of the EnCa cell lines was disrupted, whilst treatment of the cells with transforming growth factor β (TGF- β) increases syncytin expression and cell fusion.

In contrast to this, studies looking at HERV-W gene products have indicated that expression takes place in a wide variety of healthy tissues (including the brain, placenta, testis, uterus, kidney and liver) and a number of diverse cancer cell lines. HERV-W *env* and *pol* transcripts were expressed in every sample tested, malignant or otherwise, while HERV-W *gag* was expressed in normal brain, placenta, testes and spleen, but not in any of the cancer cell lines tested). Sequence analysis indicated that *env* and *gag* displayed a high degree of sequence homology (>90%) to the HERV-W sequence at 7q21-22, *pol* transcripts had a large number of point and frameshift mutations (Yi et al., 2002). While HERV-W *env* sequences are often intact, *pol* sequences show little potential for translation (Yi et al., 2004).

It has been shown that a number of HERVs retain transcriptional activity in the brain by expressed sequence tag (EST) analysis (Nakamura et al., 2003), and by digital expression profiling (Stauffer et al., 2004), including HERV-W and members of the HERV-K family. It has also been shown that HERV-K short interspersed elements are altered in a primary human glioma. These elements have homology with a number of genes including the tumour suppressor gene *BRCA2* and the DNA repair gene *XRCC1* (Misra et al., 2001). HERVs may contribute towards tumorigenesis in a number of ways, by retrotransposition (causing chromosomal instability), promoter insertion, immunomodulation, disruption of normal HERV-related functions, recombination, or by the production of fusion proteins (for review see Larsson & Andersson, 1998).

HERV-W was initially identified in MS patients, and at a molecular level MSRV *pol* transcripts have been detected in the CSF, plasma, B cell cultures of affected individuals at a higher level than healthy donors and patients with other neurological diseases (Perron et al., 1997; Nowak et al., 1993). Additionally, there is an increased copy number of *pol* sequences on chromatin fibres of MS patients (Zawada et al., 2003). For a more detailed discussion of HERV-W and its link to MS, please see Sections 1.3 and 1.6.

A number of cytokines are important in the pathogenesis of MS, including IL-4, IL-5, IL-6, IL-10, IL-13, TNF- α , and IFN- γ . MSRV *pol* has been shown to be upregulated by IL-4, IL-6, IFN- γ and TNF- α , while IFN- β (which is used in MS therapy) has the converse effect (Serra et al., 2003). In order to investigate the effects of IFN- γ treatment on MSAV *gag* transcription, the RT-PCR system was applied to a microglial cell line, CHME3. IFN- γ was used at a concentration of 5 u/ml and 50 u/ml used by Serra. MSAV *gag* was investigated for three reasons – firstly, as expression of Gag protein was found to be increased within demyelinated white matter in active MS cases (Perron et al., 2005); secondly, that MSAV *gag* is not expressed in normal brain tissue (Yi et al., 2004); and finally, that the effects of cytokine stimulation of MSRV *gag* have not been shown *in vitro*. As LTRs of retroviruses have IFN- γ responsive elements, such as ERV-9 (Andersson et al, 1998), it may be possible that upregulation of MSRV *gag* occurs if the MSAV *gag* sequence is present in the MHC region. A microglial cell line was selected as microglia are key cells in mediating inflammatory responses in the CNS, and IFN- γ strongly induces the upregulation of MHC-II in microglia (Panek and Benveniste, 1995).

4.2 Methods

4.2.1 Culture of cells grown as monolayers

Cells were grown to $\geq 70\%$ confluency in T75 tissue culture flasks (Sarstedt) under the appropriate conditions shown in Table 4.1 at 37 °C. To passage the cells further, the supplemented media was aspirated, and cells were washed twice with 2ml Hanks' Buffered Salt Solution (HBSS) (Sigma) in order to remove serum from the cells, which contains trypsin inhibitors (Chen et al. 1992). These were then detached by the addition of 2 ml of Trypsin-ethylenediaminetetraacetic acid (EDTA) solution (Sigma), followed by incubation for 2 minutes at 37 °C (for complete detachment of cells). Trypsinisation of the cells was ended by the addition of 8 ml of supplemented media, and an appropriate dilution of cells (between 1:2 and 1:10) was set up in fresh flasks containing 25 ml of fresh supplemented media.

4.2.2 Reverse transcriptase PCR

Reverse transcriptase PCR (RT-PCR) is a method of analysing levels of gene expression. This involves a number of steps, including extraction of RNA from the sample (i.e. cultured cells, clinical samples), reverse-strand synthesis of complementary DNA (cDNA), PCR amplification of the gene of interest, and quantification of the products by densitometry.

The first step of RT-PCR involves RNA isolation. The method used here is a variation of the Chomczynski and Sacchi (1987) method, yielding total RNA. The method is cheap and rapid, giving a typical yield of 8-15 µg of RNA from cultured epithelial cells. In order to minimise the possibility of DNA contamination, all isolated RNA was DNase treated, and analysed spectrophotometrically and on a denaturing gel.

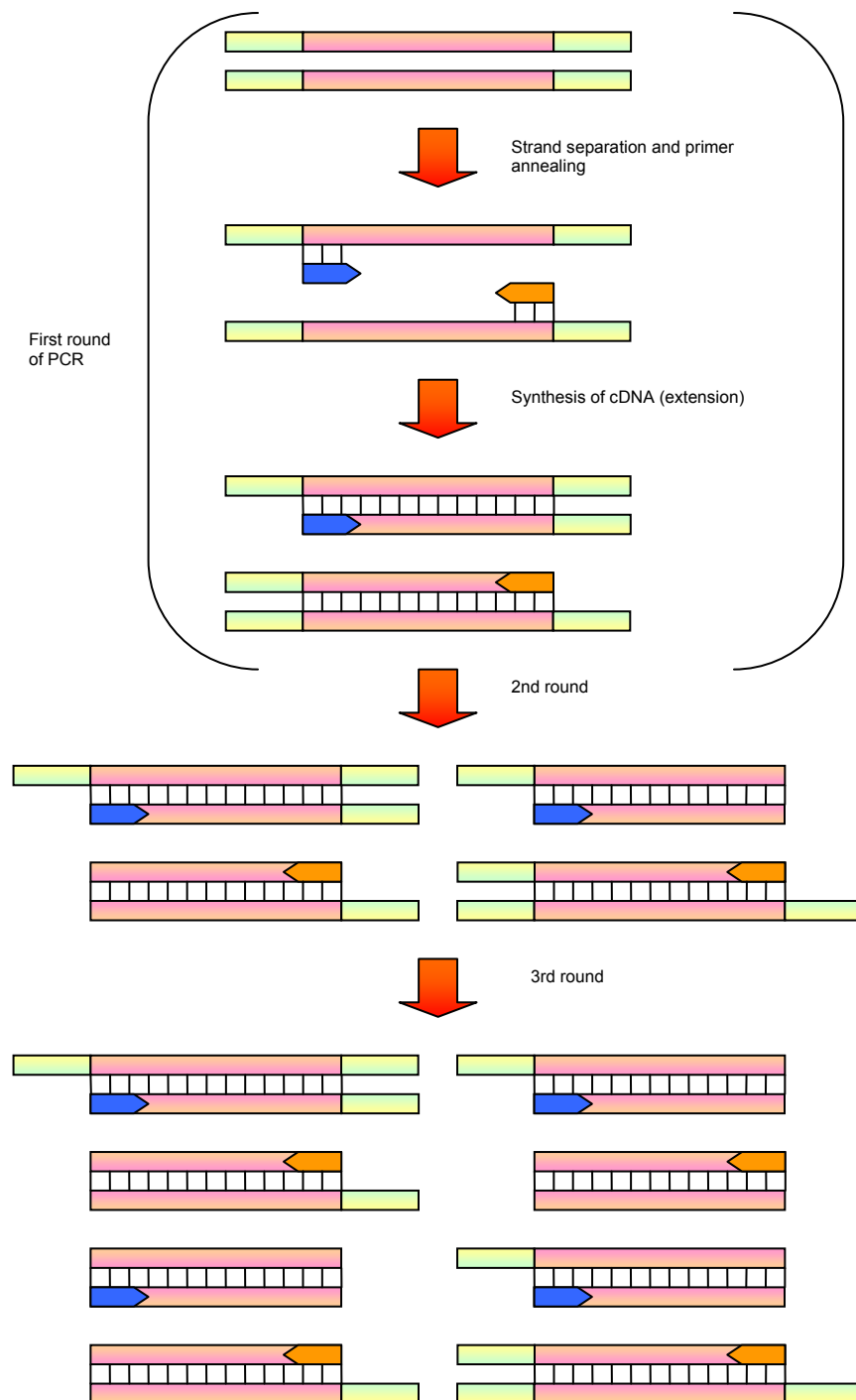
Using reverse transcriptase (RT), derived from avian myeloblastosis virus (AMV) or Moloney murine leukaemia virus (MOMULV), cDNA copies of the isolated RNA are generated. MOMULV RTs considered to be more suitable for longer cDNAs (e.g. >5kb), while AMV RT is more rapid, and stable at higher temperatures (Sambrook et al., 1989). cDNA synthesis requires a number of co-factors for the enzyme, such as magnesium chloride, as well as nucleotides and primers. Three different types of primers for cDNA synthesis are commonly used – oligo(dT) primers (which recognise the poly-A tail present on mRNA transcripts), random hexamers (which prime total RNA), and gene-specific primers (typically those used in the PCR reaction).

PCR amplifies a specific DNA sequence between 2 oligonucleotide primers which are complementary to each DNA strand flanking the region of interest. The DNA strand is pulled apart by heating the PCR mixture to 94 °C, which are then annealed to the primers at a lower temperature when the mixture cools (Figure 4.1). The primers act as a starting point for thermostable DNA polymerases to synthesise complementary DNA in the presence of each of the four deoxyribonucleoside triphosphates (dNTPs) in the appropriate buffer conditions. The newly synthesised strands act as further templates for DNA synthesis, with each cycle doubling the amount of DNA, and largely consisting of the fragment spanned by the primers.

This product can be visualised by separating the DNA fragments by gel electrophoresis, and staining with a DNA intercalating dye, such as ethidium bromide and SYBR green, which can be visualised under UV light. The size of the fragments can be determined by comparing with fragments of a known size (e.g. in a DNA ladder). The product can be verified by sequencing, or by Southern blotting (blotting the DNA onto a membrane and hybridising to a probe complementary to a region of the target),

The level of RNA expression can be determined relative to that of a housekeeping gene, such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH), histidyl tRNA synthetase, or 18s ribosomal RNA. Using densitometric analysis to get an expression ratio, it is possible to obtain relative expression of the gene of interest. It is possible to absolutely quantify by comparing expression against known copy numbers of cloned DNA. Real time PCR, where the levels of DNA are reported during amplification has made this process quicker and more accurate.

Figure 4.1 Overview of PCR amplification



PCR amplification relies on repeated cycles of denaturation (strand separation) of a DNA template, followed by annealing (binding of the primers) and elongation (DNA synthesis) stages to give further DNA template. At every cycle the amount of DNA is doubled, and exponential amplification occurs until the reaction efficiency drops as reagents are consumed and the DNA polymerase loses activity.

Table 4.1 Cell lines used and culture conditions

A list of the cell lines used for RNA extraction, and the culture conditions used. All the cell lines, except the IN series and CHME3 are established cell lines. The former are short term cell cultures derived from patient biopsies, and the latter is a foetal cell line with limited passage.

Cell Line	Description	Conditions	Source
BeWo	Choriocarcinoma	F12 Ham's (Sigma) w/ 10% foetal calf serum (FCS) (PAA) [5% CO ₂]	European Collection of Cell Cultures (ECACC)
CHME 3	Foetal microglia	Dulbecco's modified Eagle's medium (DMEM) (PAA) w/ 10% FCS (PAA) [5% CO ₂]	Prof. Nicola Woodroffe, Sheffield- Hallam University*
IN859	Grade IV astrocytoma	F10 Ham's w/ 4-(2-hydroxyethyl)-1- piperazineethanesulfonic acid (HEPES) (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
IN1472	Grade IV astrocytoma	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
IN1528	Grade IV astrocytoma	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
MCF-7	Breast cancer	Minimum essential medium (MEM) (Sigma) w/ 10% FCS (PAA) 1% non-essential amino acids (NEAA) (Sigma) 1% L-Glutamine (Sigma)[5% CO ₂]	ECACC
Tera 1	Teratocarcinoma	RPMI (Roswell Park Memorial Institute)-1640 (Sigma) w/ 10% FCS (PAA) 1% L-Glutamine (Sigma) [5% CO ₂]	ECACC
SW480	Colon adenocarcinoma	L-15 Leibovitz (Sigma) w/ 10% FCS (PAA) 1% L-Glutamine (Sigma)	ECACC
U251-MG	Grade IV : Glioblastoma Multiforme	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
U373	Grade III Glioblastoma	DMEM (PAA) w/ 10% FCS [5% CO ₂]	ECACC

* With the kind permission of Prof. Marc Tardieu, University of Paris South

4.2.3 RNA extraction and integrity checks

Cell pellets from cultured cells were lysed in 1 ml TRI reagent (Sigma) per 5×10^6 cells, and left to stand for 5 minutes at room temperature to allow disassociation of nucleoprotein complexes. TRI reagent contains a mixture of phenol and guanidine thiocyanate, a chaotropic agent which disrupts protein structure on the cell surface. 200 μ l of chloroform was added to each 1 ml of cell lysate, and vigorously shaken for 15 seconds before centrifugation at 12,000g for 15 minutes at 4 °C. The mixture separates into 3 separate phases: a red organic phase, containing protein; a white interphase, consisting of DNA; and an aqueous phase, containing RNA.

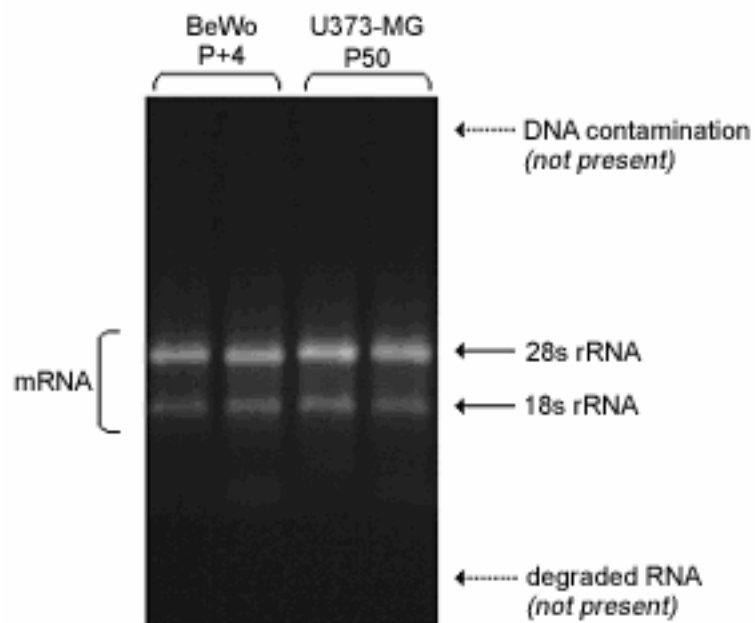
Using a pipette, without disturbing the interphase (in order to minimise DNA contamination) the aqueous phase was removed, and placed in a fresh eppendorf with 500 μ l isopropanol. The eppendorfs were then centrifuged at 12,000g for 10 minutes at 4 °C. The supernatant was decanted, and the pellet resuspended in 1 ml 75% (v/v) ethanol in nuclease-free water. After briefly vortexing, the resuspended pellets were centrifuged at 7,500 g for 5 minutes at 4 °C. The ethanol was then removed and the pellet air-dried for 5 minutes before being dissolved in 100 μ l of (v/v) 85% nuclease free water; 10% RNase-free DNase (Promega); 5% RNasin ribonuclease inhibitor (Promega). In order to remove any trace of contaminating DNA, the dissolved RNA was incubated at 37 °C for 60 minutes, followed by inactivation of the enzymes at 65 °C for 10 minutes. 10 μ l aliquots were removed for RNA quantification and integrity testing.

RNA was quantified spectrophotometrically using a Genova Mk3 analyser (Jenway). 4 μ l samples were diluted 1:20 in nuclease free water in a quartz cuvette, and RNA was quantified and the $A_{260} : A_{280}$ ratio was measured. Samples with $A_{260} : A_{280}$ ratios of between 1.7 and 1.9 were considered suitable for use in RT-PCR.

RNA samples were separated by 1% agarose in Loening Buffer (35.9 mM TRIS, 34 mM sodium phosphate (monobasic), 1mM EDTA, stained with ethidium bromide) gel electrophoresis and visualised under UV light. When visualised under UV light, a mRNA band should be visible, together with bands for 28s and 18s rRNA, with the 28s band of roughly double the intensity of that of the 18s band (Figure 4.2). Samples showing evidence of DNA contamination (a band near the top of the gel) or significant RNA degradation (a smear near the bottom) were considered unsuitable for use in RT-PCR. Samples for RT-PCR were snap-frozen on dry ice and stored at -80 °C until needed.

Figure 4.2 RNA integrity gel

RNA integrity gels are run in order to establish that RNA is of an adequate quality for use in RT-PCR. DNA contamination would mean that PCR products would be obtained from genomic DNA in addition to RNA, affecting reliability of results.



4.2.4 Primer Design

HERV-W *env* and MSRV *gag* primers (Figures 4.3a & b respectively) were designed using DNASTar (Lasergene) computer software. 5 different primers for HERV-W *env* and a single primer pair for MSRV *gag* were selected on the basis of predicted annealing temperature (ideally 59 °C), similar melting temperatures (within 2 °C), GC % content (ideally 40-60 %), and in order to reduce the possibility of hairpin structures (where the primer can fold on itself) and primer-dimers (where the primers bind to each other). Such structures can be minimised by choosing primers where the sequences are not complementary to each other, e.g. avoiding palindromic sequences AGTCTTGACT) and long repeats (e.g. CCCCC and GGGGG). To check that sequences were unique to HERV-W *env* and MSRV *gag*, and not to other sequences (such as other retroviruses) were entered into BLAST.

Previously published *histidyl tRNA synthetase* primers (Figure 4.3c) (Ejtehad et al., 2005) were used as a housekeeping gene. This housekeeping gene was selected as it had been successfully multiplexed with other HERV primers, such as for HERV-K10, and had previously been shown to be stable (Ejtehad et al., 2005; Ejtehad et al., 2006). In the genomic DNA, the forward primer is located in exon 1, and the reverse primer is located in exon 4, giving a genomic product length of 11,440 base pairs. A 319 bp product can only be derived from mRNA, whilst conventional PCR is unlikely to produce an amplicon of 11.4 kbp. Housekeeping genes should be stably expressed amongst different cell types tested, and be unaffected by cell treatments used. This is particularly important where real-time PCR is used.

4.2.5 cDNA synthesis

1 µg of total extracted RNA, as well as a nuclease free water (Promega) control were denatured at 70 °C prior to cDNA synthesis. Reverse transcription was carried out in 20 µl reaction mixtures containing 1 x RT buffer (Promega), 5 mM magnesium chloride (Promega), 1 mM dNTP mixture (Promega), 20 units/ml RNasin® Plus ribonuclease inhibitor (Promega), 15 units/ml AMV reverse transcriptase (Promega) and 0.5 µg random hexamer primers (Promega) per µg RNA. Reverse transcriptase negative controls were also included. Incubation at 42 °C for 1 hour in a PTC-100 thermocycler (MJ Research Inc.) was followed by heating the cDNA to 99 °C to inactivate the reverse transcriptase

Figure 4.3 Primers used to detect HERV-W *env*, MSRV *gag* and *histidyl-tRNA synthetase*

(a) **NM_014590** Homo sapiens endogenous retroviral family W, env (C7), member 1 (syncytin) (ERVWE1), mRNA

HERV-W env 2-1 GTATGTCTAATGGGGGTGGAGTTC

%GC = 50.0, Tm = 60.62

pL = 419

HERV-W env 2-2 GATGCATTGGGAGTTGGTTGTGTA

%GC = 45.8, Tm = 63.78

```

1321 atcctagttg tcctggagga cttggagtca ctgtctgttg gacttacttc acccaaactg
1381 gtatgtctga tgggggtgga gttcaagatc aggcaagaga aaaacatgta aaagaagtaa
1441 tctcccaact caccgggta catggcacct ctagcccta caaaggacta gatctctcaa
1501 aactacatga aaccctccgt acccatactc gcctggtaag cctatttaat accaccctca
1561 ctgggctcca tgaggtctcg gcccacaaacc ctactaactg ttggatatgc cccccctga
1621 acttcaggcc atatgtttca atccctgtac ctgaacaatg gaacaacttc agcacagaaa
1681 taaacaccac ttccgtttta gtaggacctc ttgtttccaa tctggaaata acccatacct
1741 caaacctcac ctgtgtaaaa ttagcaata ctaca tacac aaccaactcc caatgcatca
1801 ggtgggtaac tcctcccaca caaatagtct gcctaccctc aggaatatatt tttgtctgtg

```

Primers were designed (against reference sequences where possible, accession numbers shown) using DNASTar software (Lasergene), and checked against the NCBI GenBank database using BLAST analysis to check that the sequences were unique to the sequence of interest. Shaded areas of the sequences indicate where the primers bind.

(b) **AF123881** MS associated retrovirus element gag
polyprotein (gag) mRNA, partial cds

MSRV gag 1-1 AAGGAGGGCAAATGGAGTGAAGTG
%GC = 50.0, Tm = 65.39
pL = 571
MSRV gag 1-2 CCTGCTGGCCTGTGGGAGTTGTT
%GC = 60.9, Tm = 69.32

61 atttatattc ttctgcagta ccgcctggcc acaatatacct cttcaaggga gagaaacctg
121 gcttcctgag ggaagtataa attataacat catcttacag ctagacctct tctgtagaaa
181 ggagggcaaa tggagtgaag tgccatatgt gcaaactttc ttttcattaa gagacaactc
241 acaattatgt aaaaagtgtg gtttatgccc tacaggaagc cctcagagtc cacctcccta
301 ccccgagctc ccctccccga ctcccttcctc aactaataag gacccccctt taacccaaac
361 ggtccaaaag gagatagaca aaggggtaaa caatgaacca aagagtgccg atattccccg
421 attatgcccc ctccaagcag tgagaggagg agaattcggc ccagccagag tgccctgtacc
481 tttttctctc tcagacttaa agcaaattaa aatagacctc ggtaaattct cagataaccc
541 tgacggctat attgatgttt tacaagggtt aggacaatcc tttgatctga catggagaga
601 tataatgtta ctactaaatc agacactaac cccaaatgag agaagtgccg ctgtaactgc
661 agcccagagag tttggcgatc tttggtatct cagtcaggcc aacaatagga tgacaacaga
721 ggaaagaaca actcccacag gccagcaggc agttcccagt gtagaccctc attgggacac
781 agaatcagaa catggagatt ggtgccacaa acatttgcta acttgcgctg tagaaggact

(c) **NM_002109** Homo sapiens histidyl-tRNA synthetase (HARS), mRNA

Histidyl s CTTTCAGGGAGAGCGCGTGCG
%GC = 65.0, Tm = 69.48
pL = 319
Histidyl as CCTTCAGGACCCAGCTGTGC
%GC = 60.0, Tm = 63.20

1 tcgatagccg gaagtcaccc ttgctgaggc tggggcaacc accgcaggtc gagacagcag
61 gcggctcaag tggacagccg ggatggcaga gcgtgcggcg ctggaggagc tggtgaaact
121 tcaggggagag cgcgtgcgag gcctcaagca gcagaaggcc agcgccgagc tgatcgagga
181 ggaggtggcg aaactcctga aactgaaggc acagctgggt cctgatgaaa gcaaacagaa
241 atttgtgctc aaaaccccc agggcacaa agactatagt ccccggcaga tggcagttcg
301 cgagaaggctg tttgacgtaa tcatccgttg cttcaagcgc cacggtgcag aagtcattga
361 tacacctgta tttgaactaa aggaaacact gatgggaaag tatggggaag actccaagct
421 tatctatgac ctgaaggacc agggcgggga gctcctgtcc cttcgctatg acctcactgt
481 tccttttgcg cggtatttgg caatgaataa actgaccaac attaaacgct accacatagc

4.2.6 Polymerase chain reaction

In the optimised PCR reaction, 50 µl reaction vessels containing 1 x RT buffer (Promega), 1.5 mM magnesium chloride (Promega), 200 mM dNTP mixture (Promega), 1.25 units Taq DNA polymerase (Promega), and appropriate combinations of magnesium chloride (Promega) and primers (Invitrogen), as outlined in Table 4.2, for the reaction and 10 µl cDNA product were prepared. During the process of magnesium chloride optimization, a range of between 1.5 mM and 3.0 mM (in 0.5 mM increments) was used.

Table 4.2 Quantities of magnesium chloride and primers used in the optimised PCR reaction for HERV-W *env* and MSRV *gag*

	HERV-W <i>env</i>		MSRV <i>gag</i>	
	HERV-W <i>env</i> (sequencing)	Multiplex (quantification)	MSRV <i>gag</i>	<i>Histidyl tRNA synthetase</i>
MgCl ₂ /mM	1.5 mM	1.5 mM	2.5 mM	2.5 mM
HERV-W <i>env</i> 2.1 primer	2.5 pmol	2.5 pmol)	-	-
HERV-W <i>env</i> 2.2 primer	2.5 pmol	2.5 pmol	-	-
MSRV <i>gag</i> 1.1 primer	-	-	2.5 pmol	-
MSRV <i>gag</i> 1.2 primer	-	-	2.5 pmol	-
<i>Histidyl tRNA synthetase</i> sense primer	-	5 pmol	-	5 pmol
<i>Histidyl tRNA synthetase</i> antisense primer	-	5 pmol	-	5 pmol

Following initial denaturation at 94 °C for 3 minutes, 30 cycles of PCR were carried out with denaturation for 1 minute at 94 °C, annealing for 1 minute at 59 °C and extension for 2 minutes at 72 °C, followed by a final extension step at 72 °C for 10 minutes. Products were separated by 2% agarose in Tris-buffered EDTA (Invitrogen) (stained with ethidium bromide) gel electrophoresis and visualised under UV light. Gels were imaged using Grabber software (BioGene), and PCR products were quantified using Scion Imager software (Scion Corp.) by analysing the ratio of HERV-W *env* product to *histidyl tRNA synthetase*.

4.2.7 Processing of PCR products for sequencing

In order to sequence PCR products, it is necessary to remove excess primers and dNTPs as these can interfere with the sequencing reaction. This can be done using chromatographic methods (e.g. Miniprep columns) or by enzymatic methods (such as shrimp alkaline phosphatase – exonuclease digests).

Laboratory protocol

PCR products were purified by the addition of 2 µl of ExoSAP-IT (0.5u shrimp alkaline phosphatase; 5u exonuclease I) (Amersham) per 5 µl PCR product. The samples were heated to 37 °C for 30 minutes followed by enzyme inactivation for at 80 °C for 15 minutes. This method enzymatically digests single stranded RNA and excess dNTPs that can interfere with sequencing reactions (Werle et al., 1994). The advantage of using this method is that it is quick, and avoids sample handling and DNA loss associated with chromatography-based techniques. Samples were diluted to 1-2 µl into 10 µl of nuclease free water, and sent to the Sequencing Service at the University of Dundee for sequencing.

4.2.8 Treatment of CHME3 cells with IFN-γ

CHME3 cells were seeded at 1.15×10^6 cells per T75 flask in DMEM containing 10% FCS and incubated at 37 °C overnight. The media was removed and replaced with 23 ml of fresh media containing 5 or 50 ng/ml recombinant IFN-γ (Sigma), or a cytokine-free control. Flasks were incubated at 37 °C for 48 hours prior to RNA extraction.

4.2.9 MTT assay for cell viability of CHME3 cell lines treated with IFN-γ

In order to investigate the effects of a particular treatment on a cell line, it is necessary to carry out a cell viability assay. The assay is dependant on 3-(4,-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT), a yellow substance in solution, being converted to MTT formazan by mitochondrial dehydrogenases which cleave the tetrazolium ring. The reaction yields a purple product, insoluble in water, which can be measured spectrophotometrically after being dissolved in solvents, e.g. dimethylsulfoxide (DMSO). Alternate cell viability assays include those which report adenosine 5'-triphosphate (ATP) activity, such as CellTiter Glo (Promega).

Laboratory protocol

CHME3 cells were cultured in the wells of a 96 well plate (Sarstedt), at a seeding density of 1×10^5 cells per well (e.g. 200 μ l of 5×10^5 cells/ml per well), with the exception of well A1 which acts as a control. The plate was then incubated at 37 °C for 24 hours. After incubation, the medium was then removed by aspiration and replaced with 200 μ l fresh medium per well containing the 5 or 50 ng/ml recombinant IFN- γ (Sigma) and cytokine-free controls at the relevant concentrations (in replicates of 6). Following a further incubation of the plate at 37 °C for 48 hours, the medium was aspirated and replaced with 300 μ l media containing 0.5 mg/ml MTT stain (Sigma), including well A1 (acting as a background control). The plate was then incubated at 37 °C for 3 hours, the media was then aspirated and 200 μ l DMSO (Sigma) was added to solubilise the product for 10 minutes. Absorbances were then read at 540 nm on a Multiskan MS (Labsystems) microplate reader. The effect of the particular cytokine treatment on % cell viability can be calculated by dividing the absorbance of treated cells (adjusted for background, A1) with those of untreated cells and multiplying by 100.

4.3 Results

4.3.1 Magnesium chloride optimisation

A magnesium chloride concentration gives a higher yield of PCR products at 2.0 mM and 1.5 mM respectively for *histidyl tRNA synthetase* and HERV-W *env* respectively (Figures 4.4a and 4.4b). At higher magnesium chloride concentrations with HERV-W *env*, the fidelity of the reaction is reduced, and mispriming seems to occur, and the band becomes smeared (Figure 4.4b). Conversely, the band intensity of MSR-V *gag* (Figure 4.4c) decreases with increasing MgCl_2 concentration, but produces a clear PCR product in all cases.

4.3.2 Sensitivity testing of HERV-W *env* multiplex system in the BeWo cell line

With a higher RNA input for cDNA synthesis, the yield of HERV-W *env* PCR product is higher (Figure 4.5). Products are detectable at a low RNA input (below 5.0 ng/ μl). A true idea of the sensitivity of the reaction can be achieved by cloning the sequence into a plasmid, and determining the lowest copy number which gives a signal.

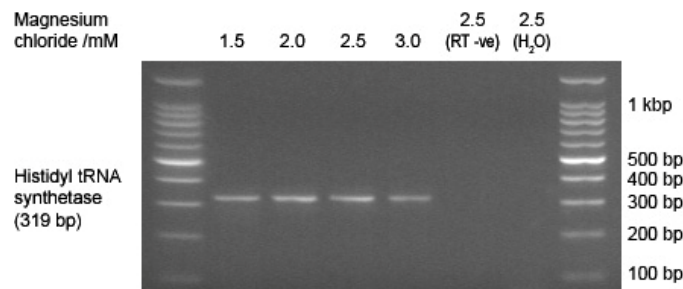
4.3.3 Comparison of HERV-W *env* transcripts in cancer cell lines

HERV-W *env* is expressed at high levels in the BeWo cell line (over 11 times greater than that of U251-MG cell line, $F=18.94$, $p<0.0005$, One-way ANOVA). HERV-W expression is moderate in the established P53 mutant cell lines U251-MG and U373 glioblastoma cell lines, the IN1528 P53 mutant short term culture MCF-7 breast adenocarcinoma cell line, and teratocarcinoma cell line (Tera-1), and low in the short term P53 wild type astrocytoma cultures IN859 and IN1472, and the SW480 colon adenocarcinoma cell line (Figures 4.6 and 4.7).

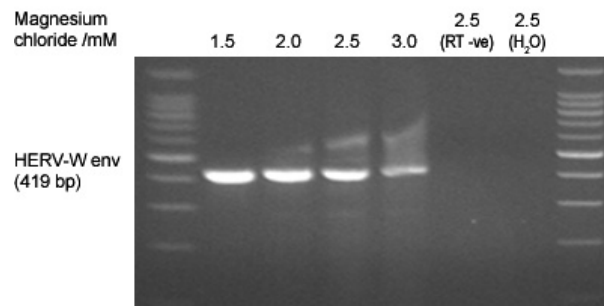
Expression of HERV-W *env* is significantly higher in P53 mutant astrocytoma cell lines ($\mu=0.468$, $sd=0.198$) compared to those with wild type P53 ($\mu=0.1806$, $sd=0.013$) (2 sample T test, $T=5.93$, $p<0.0005$) (Figure 4.8). In all runs used for quantification, RT negative controls (to rule out genomic contamination of the RNA sample) and dH_2O controls (to rule out contamination of the reagents) were blank.

Figure 4.4 Magnesium optimisation gels for histidyl tRNA-synthetase, HERV-W *env*, and MSRV *gag* in BeWo cells

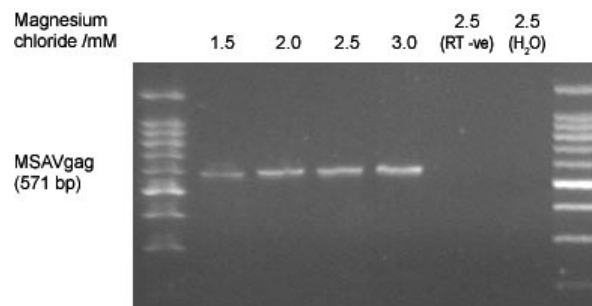
(a) *Histidyl tRNA synthetase*



(b) HERV-W *env*



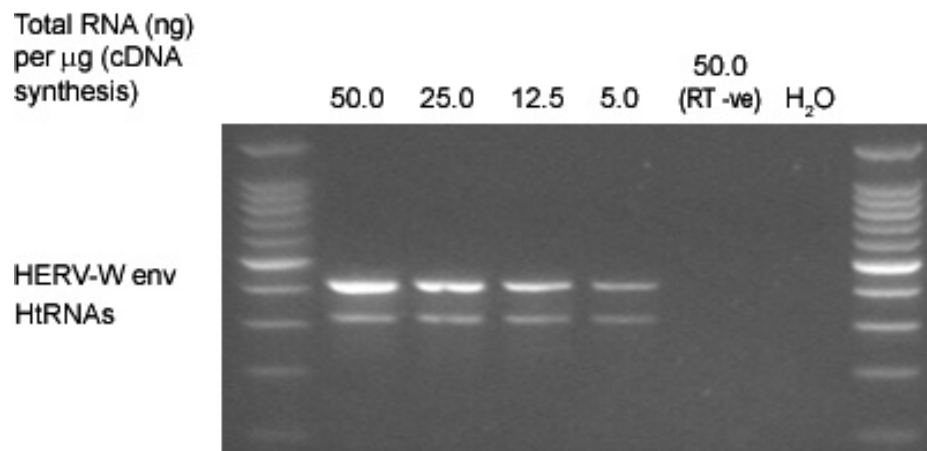
(c) MSRV *gag*



PCR products for all PCR products are detectable between 1.5-3.0 mM, but not in the negative controls. At higher concentrations of MgCl₂ for HERV-W *env* the band becomes less discrete, indicating mispriming. The histidyl tRNA synthetase and MSRVgag primers yield less product at lower magnesium chloride concentrations, and do not misprime at the higher concentrations.

Number of replicates = 3; 1 RNA batch used for cDNA synthesis, PCR runs performed in triplicate

Figure 4.5 Band intensities for histidyl tRNA synthetase and HERV-W *env* at different RNA input concentrations.

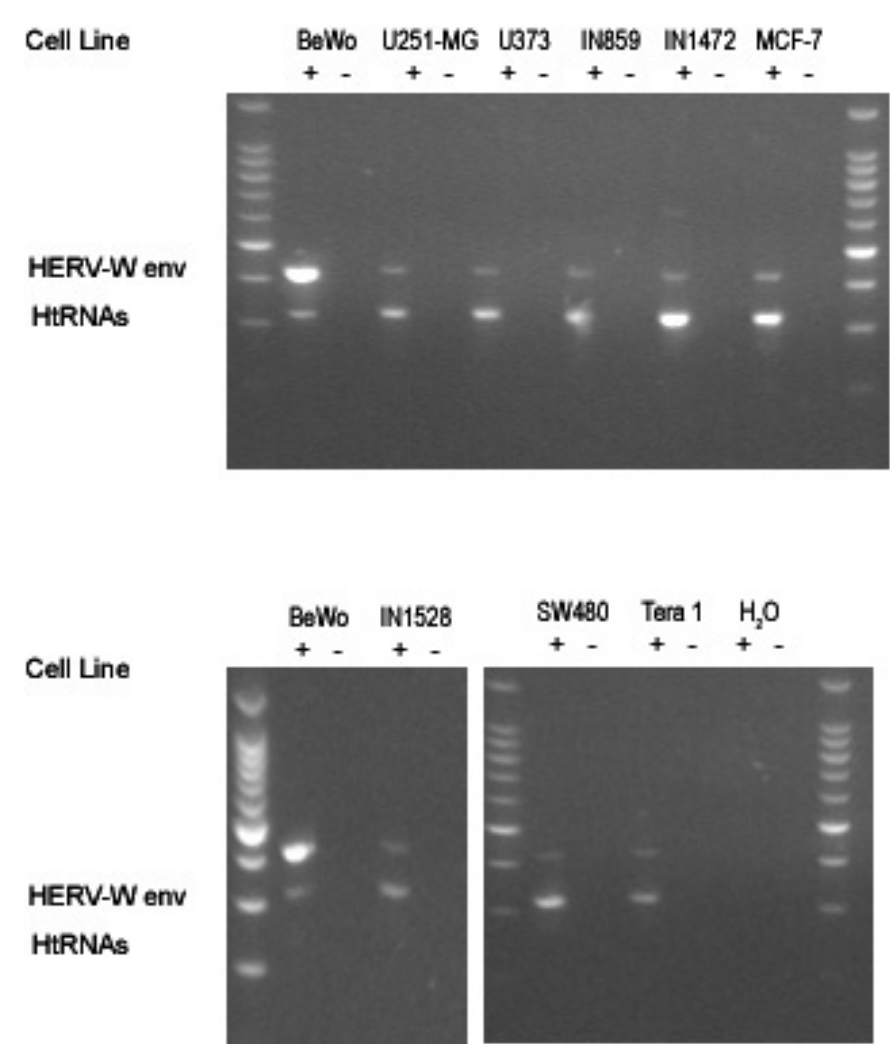


While the intensity of the bands weakens with decreased RNA input, PCR products are readily detectable, even at 5ng per μ l cDNA synthesis reaction. No PCR products are present in the reverse transcriptase negative and water controls.

Number of replicates = 3; 1 RNA batch used for cDNA synthesis, PCR runs performed in triplicate

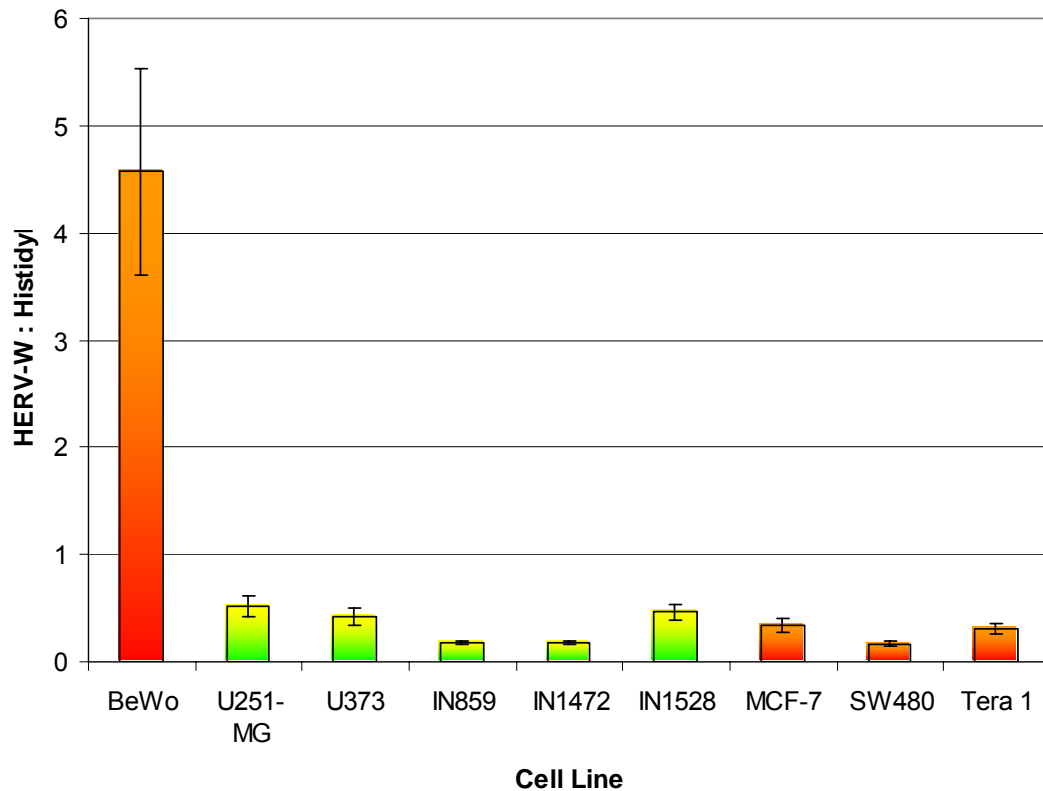
Figure 4.6 Gels illustrating HERV-W *env* RNA levels in different human cell lines in a multiplex PCR reaction.

The ratio of HERV-W *env* (pL=419) to histidyl tRNA synthetase (pL=319) band intensity is stronger in the BeWo cell line compared to the other cell lines tested. The PCR products come from cDNA, and not genomic DNA or contaminant RNA as the RT negative controls and water controls respectively are blank.



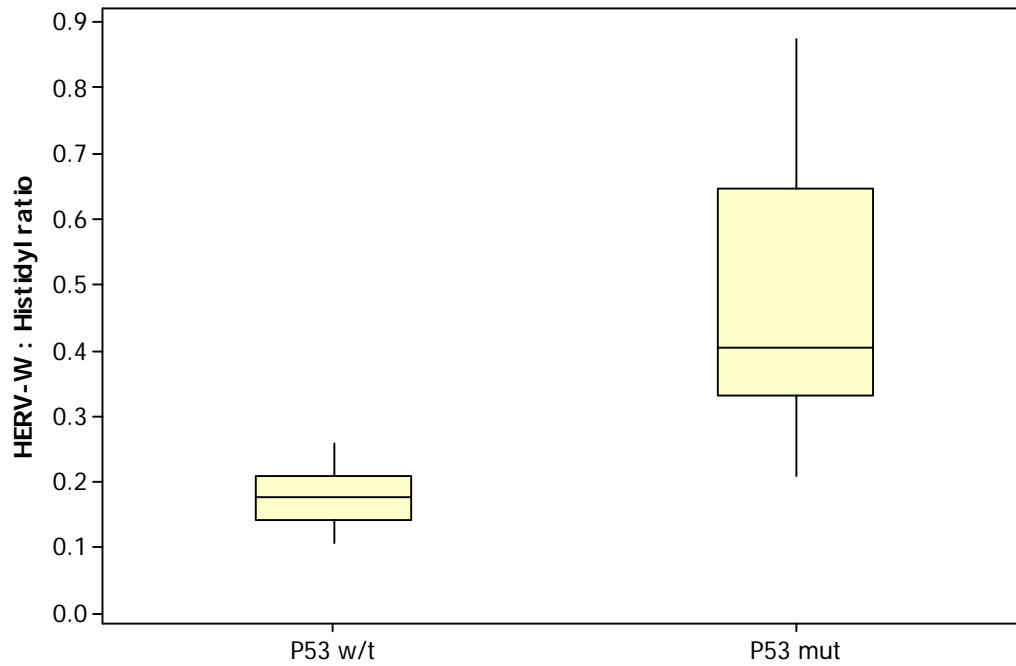
Number of replicates = 6; 3 RNA batches used for cDNA synthesis, PCR runs performed in duplicate

Figure 4.7 Comparison of HERV-W *env* RNA levels in human cancer cell lines



HERV-W *env* RNA is highly expressed in BeWo ($\mu=4.571$, $sd=2.365$) placental choriocarcinoma cell lines ($F=18.94$, $p<0.0005$, One-way ANOVA) compared with the other cell lines tested: U251-MG ($\mu=0.522$, $sd=0.235$); U373 ($\mu=0.417$, $sd=0.198$); IN859 ($\mu=0.184$, $sd=0.041$); IN1472 ($\mu=0.177$, $sd=0.055$); IN1528 ($\mu=0.464$, $sd=0.177$); MCF-7 ($\mu=0.336$, $sd=0.162$); SW480 ($\mu=0.169$, $sd=0.049$); Tera 1 ($\mu=0.312$, $sd=0.117$). Moderate expression is found in the *P53* mutant cell lines, U251-MG, U373 and IN1528, compared with IN859 and IN1528, which are *P53* wild type ($F=5.98$, $p=0.002$, One-way ANOVA).

Figure 4.8 Boxplot illustration comparing HERV-W *env* expression in *P53* wild type and *P53* mutant astrocytoma cell lines



HERV-W *env* transcripts are expressed at higher levels in *P53* mutant astrocytoma cell lines, U251-MG, U373 and IN1528 (pooled $\mu=0.468$, $sd=0.198$, median=0.4040) compared to those with wild type *P53*, IN859 and IN1472 (pooled $\mu=0.1787$, $sd=0.198$, median=0.1787).

4.3.4 Detection of MSRV *gag* transcripts in cancer cell lines

MSRV *gag*, which putatively encodes structural elements of the retrovirus, such as the capsid and nucleocapsid (Komurian-Pradel et al., 1999) was detected in all cell lines tested, including BeWo (a placental choriocarcinoma), the astrocytoma short-term cell cultures IN859, IN1472, IN1582 and IN1528, the Grade IV glioblastoma multiforme cell line U251-MG, the Grade III glioblastoma cell line U373-MG (Figure 4.9), and the foetal microglia cell line (Figure 4.15). No detection of MSRV *gag* or *histidyl tRNA synthetase* was observed in RT negative and water controls, or expression of genomic *histidyl tRNA synthetase* was detected.

In contrast to HERV-W *env*, MSRV *gag* PCR products were less homologous to the sequences that the primers were designed against, namely the reference sequence for syncytin/HERV-W *env* (GenBank accession number NM_014590) and MS associated retrovirus element *gag* polyprotein (*gag*) (GenBank accession number AF123881) respectively. ClustalW alignment of the transcripts against MSRV *gag* (AF123881) using Megalign (DNASTar) with a number of gaps and mutations present in U251-MG and BeWo transcripts, as well as the CHME 3 transcript (Figure 4.11). No significant similarity was found to any sequences present in the GenBank database, with similarities to MSAV *gag* of 68% and 71% for BeWo and U251-MG isolates respectively, and 76% for CHME-3. Consequently, it is likely that a variant of MSRV *gag* has been detected in these cell lines rather than the MSRV *gag* isolates described previously. Phylogenetic analysis using MegAlign and ClustalW alignment indicated that CHME3 resembles MSRV *gag* more closely than either the U251-MG or BeWo sequences, which are more closely related to each other than the CHME3 or MSRV *gag* sequence (Figure 4.12).

Figure 4.9 Alignment of sequence of PCR products with HERV-W *env* (C7)/syncytin

```

1411 AGGCAAGAGA-AAAACATGTAAAAGAAGTAATCTCCCAACTCACC CGGT HERV-W env (C7) (#NM_014590)
      1 -----AGCGTGATCGCTTGTATAGAAATTA-TCTCTCA-CTC ACTCGGGT BeWo isolate
      1 -----CGT-TC AACATGTAAAG--AGTA-TCTCCCAACTCACC CGGT U251-MG isolate

1460 ACATGGCACCT-CTAGCCCCTACAAAGGACTAGATCTCTCAA AACTACAT HERV-W env (C7) (#NM_014590)
      44 AC-TGGCACCTTCTAGCTTCTACAAAGGACTAGATTTCTCTAAA AACTACAT BeWo isolate
      40 ACATGGCACCT-CTAGCCCCTACAAAGGACTAGATCTCTCAA AACTACAT U251-MG isolate

1510 GAAACCCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACC ACCCT HERV-W env (C7) (#NM_014590)
      93 GAAACTTTCCGTATT-ATACTCGCCTGGTAAGCCTATTTAATACC ACCCT BeWo isolate
      89 GAAACCCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACC ACCCT U251-MG isolate

1560 CACTGGGCTCCATGAGGTCTCGGCCCAAACCCCTACTA ACTGTTGGATAT HERV-W env (C7) (#NM_014590)
      142 CACTGGGCTCCATGAGGTCTCGGCTCAA AACCCTACTAACCGTTGGATAT BeWo isolate
      139 CACTGGGCTCCATGAGGTCTCGGCCCAAACCCCTACTA ACTGTTGGATAT U251-MG isolate

1609 GCCTCCCCCTGAACTT-CAGGCCATATGTTTCAATCCCTGTACCTGA ACA HERV-W env (C7) (#NM_014590)
      192 GCCTCTCTCTGAACTTTCAGGCCATATGTTTCAATCCCGTACCTGA ACA BeWo isolate
      189 GCCTCCCCCTGAACTT-CAGGCCATATGTTTCAATCCCTGTACCTGA ACA U251-MG isolate

1658 ATGGAACAACCTTCAGCACAGAAATAAACACCACTT-CCGTTT TAGTAGGAA HERV-W env (C7) (#NM_014590)
      242 ATGGAACAACCTTCAGCACAGAAATAAACACCACTTTC CGTTTAGTAGGA BeWo isolate
      238 ATGGAACAACCTTCAGCACAGAAATAAACACCACTT-CCGTTT TAGTAGGA U251-MG isolate

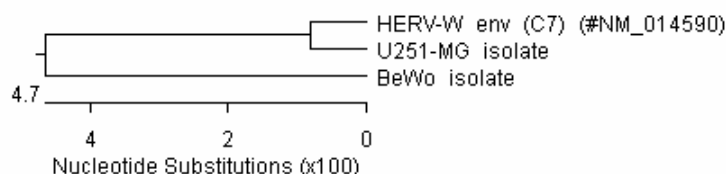
1707 CCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAA ACCTCACCTGTGT HERV-W env (C7) (#NM_014590)
      292 CCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAA ACCTCACCTGTGT BeWo isolate
      287 CCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAA ACCTCACCTGTGT U251-MG isolate

1757 AAAATTTAGCAATACTACATACACAACCAACTCCCAA -TGCATCAGGTGG HERV-W env (C7) (#NM_014590)
      342 AAAATTTAGCAATACTACATACACAACCA-CCCAAAGGGGCAACA BeWo isolate
      337 AAAATTTAGCAATACTACATACACAACCAACTCCAAATGCATCA U251-MG isolate

```

PCR products of BeWo and U251-MG isolates were aligned with HERV-W *env* using the ClustalW algorithm. Identical bases are highlighted in blue, gaps are indicated with -, insertions with ‘.

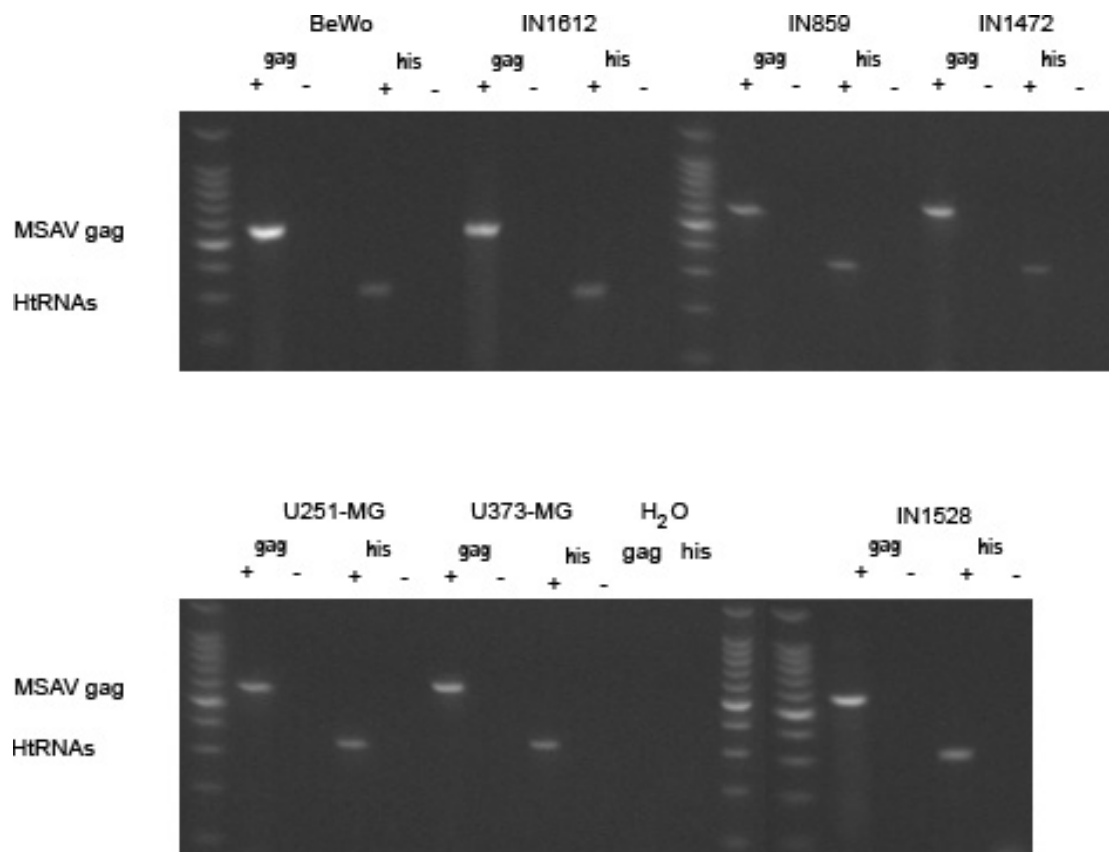
Figure 4.10 Phylogenetic analysis of PCR product sequences with HERV-W *env* (C7)/syncytin



Phylogenetic analysis of PCR products aligned with HERV-W *env* (C7)/syncytin. Both BeWo and U251-MG isolates are highly similar to the reference sequences. Shorter branches, i.e. less frequent nucleotide substitutions indicate a closer relationship between the sequences.

Figure 4.11 Gels illustrating MSRV *gag* and *histidyl tRNA synthetase* PCR products in human cancer cell lines

MSRV *gag* (pL=571) transcripts were detected in all cell lines tested, with prominent expression, particularly in the BeWo choriocarcinoma cell line. *Histidyl tRNA synthetase* (pL=319) bands were expressed uniformly amongst different cell types.



Number of replicates = 6; 3 RNA batches used for cDNA synthesis, PCR runs performed in duplicate

Figure 4.12 Alignment of sequence of PCR products with MSRV *gag*

```

201 GAAGTGCCATATGT ' ' GCAAACTTTCTTTTCATTAAAGAGACAACCTCACAA MSRVgag (#AF123881)
1 -----ACAATGTCAGGTTTCTT-CTCTTACGAGA-CACTCCAT- BeWo Isolate
1 -----CCTACCGATGCAGGGCACTAGATGATTAGGAGA-GAGTCAGCA CHME3 Isolate
1 -----CCTTGGGGAAGGTGTCTT-TAATGA-GAGA-TCCTCGAT- U251-MG Isolate

251 TTATGTAAAAAGTGTGGTTTATGCCCCTACAGGAAGC ' ' CCTCAGAGTCCA MSRVgag (#AF123881)
37 TTATGTAAAAAGGGTG-ATTATATGTCCACAGGAGGACCTCCAAG 'TCCA BeWo Isolate
43 TTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCC '-TCAGAGTCTA CHME3 Isolate
38 TTATTTAGAAATGGTG-TTTATGTCTCC---GAGGACCTCCAAGT--C' U251-MG Isolate

301 CCTCCCTACCCAGCGTCCCCTCCCAGCTCCTTCTCACTAAT-AAGG MSRVgag (#AF123881)
86 CCCCTTACCCCGGGTCCCC---CCCGCTTCCCCTCTGATAATTAAGG BeWo Isolate
92 CCTCCCTACCCAGCGTCCCC---CCGACTCCTTCCCCTCACTAAT-AAGG CHME3 Isolate
82 CTCCCCCGTCTTGGGTCCCC---CCGACTTCCCCTCTAAGAAATAAGG U251-MG Isolate

350 ACCCCCTT ' ' TAACCCAAACGGTCCAAAAGGAGATAGACAAAGGGGTAA MSRVgag (#AF123881)
133 ACTCTCCTTTGG--CCCTAACGGCCCAAAAAG 'GAACACACCCGAAGATA BeWo Isolate
138 ACCCCCTTTT--CCCAAA-GGTCCA 'AAAAAGAAAACAAAGGGGTGA CHME3 Isolate
129 ACCCCCTTTGG--CCCTGAGGGCCCAAAAAGGAAC 'CACACCCGGAGATA U251-MG Isolate

400 ACAATGAACCAAAGAGTGCCAATATTTCCCCG 'ATTATGCCCCCTCCAAG MSRVgag (#AF123881)
181 GTCAGTCCCAAAAGGCC--CTTGGCCCTCTATTGTGCCCCCTCCAG BeWo Isolate
185 AAAAGAACCCAATGGGCC--CATCTCCCTAATATGCCCCCCCAAT CHME3 Isolate
177 CTCTGGTACCCAAAAGGCC--AATGGCCCTCTATTGTGCCCCCTCCAAG U251-MG Isolate

450 CAGTGAG---AGGAGGAGAATTTCGGCCAGCCAGAGTGCCGTGTACCTTTT MSRVgag (#AF123881)
230 CCGTGGGAGAGG 'AGGAGACTTTGTTCTGGTGTGAATA-CATGTTCCCT BeWo Isolate
233 CGGTGGG--AG 'AAAAAATTGTGCCCTTCTCTAGCGC--ATACTCTT CHME3 Isolate
226 CCGTGGGAGAGGGA 'GACTTTGGCCTGGTGTGAGT--CATGTTTCT U251-MG Isolate

497 TCTCTC ' ' TCAGACTTAAAGCAAATTAAAAATAGACCTAGGTAAATT-CTC MSRVgag (#AF123881)
279 TTTCCTCTCA-ACTTAAAAATTAACTCAACTGAACGGATAAAAAATT 'CTC BeWo Isolate
279 TTTCCTCTCACACTTAGGGCAAAATTAAATAAAACCCCGGATT 'TCTT CHME3 Isolate
274 TTTCCTCTCAGACTTAAACAAATCAACTGAACGGAGAAAAATTCTCA U251-MG Isolate

546 AGATAACCTGACGCTATATTGATGTTT 'TACAAGGGTTAGGACAATC MSRVgag (#AF123881)
328 AGATTAACTGATGG--ATGAATATATGTCTTACTAGGATTAGGACATTC BeWo Isolate
329 ATAAACCTGAGGG--GTATTAGGAAATTTTAAAGGGTTAGAGAACC CHME3 Isolate
324 'GATAATCGGGATGG--ATGCATATTGTCTTACTAGGATTAGGACTTTC U251-MG Isolate

596 CTTTGATCTGACATGGAGAGATATAATGTT-ACTACTAAATCAGACACTA MSRVgag (#AF123881)
376 ATTTATTTTGGAAGAAAAATATTATGT 'TGTGTTAAGACAAACCTTA BeWo Isolate
377 TTTTGATTGTATGTGGAAAAATATAATG 'TACTGCTAATCAAAAAGGA CHME3 Isolate
372 TTTTAAATTTGGAAGAAAAATGTTATGTTT 'GT-----CAGACACTA U251-MG Isolate

645 AC ' ' CCCAAATGAGAGAAGTGCCGCTGTAAGTGCAGCCCGAG-AGTTTG MSRVgag (#AF123881)
426 --ACTCCCGACAAAACCGGCCACCTTACTTGCGACCCCTTG-AGTTGGT BeWo Isolate
427 --CCCCATAAGAGAAAGCGGCAGCACTAAGTGGCTCCCGAATAGTTTG CHME3 Isolate
417 --CCCCAAATGAAGAAGTGTGCCCCTAAGTGCATCCCGAACAGTTTG U251-MG Isolate

694 CGATCTTTGGTA-TCTCAGTCAGGCC ' ' AACAAATAGGATGACAACAGAGG MSRVgag (#AF123881)
473 CAATGTTTGGTAC 'CTAAGCCAGG--TAAAGATAAATGACAACA AAAA BeWo Isolate
475 TGATCTCTGGT 'ATCTCGATCAGG--TCGATGATAGGATGACAACAGAGA CHME3 Isolate
465 TGATCTCTGGTATCTCACTCA 'GG--TCAGTGATAGGATGACAACGAGG U251-MG Isolate

743 AAA--GAACAACTCCACAGGCCAGCAGGCAGTTCCTCAGTGTAGACCTC MSRVgag (#AF123881)
521 AAGAGGACCCTCCCGCCG ' BeWo Isolate
523 AAAAGAACAACCTCCACAGGGCC ' CHME3 Isolate
513 AAAGAGAACAACCTCCACCCGACGAGGAAG ' U251-MG Isolate

```

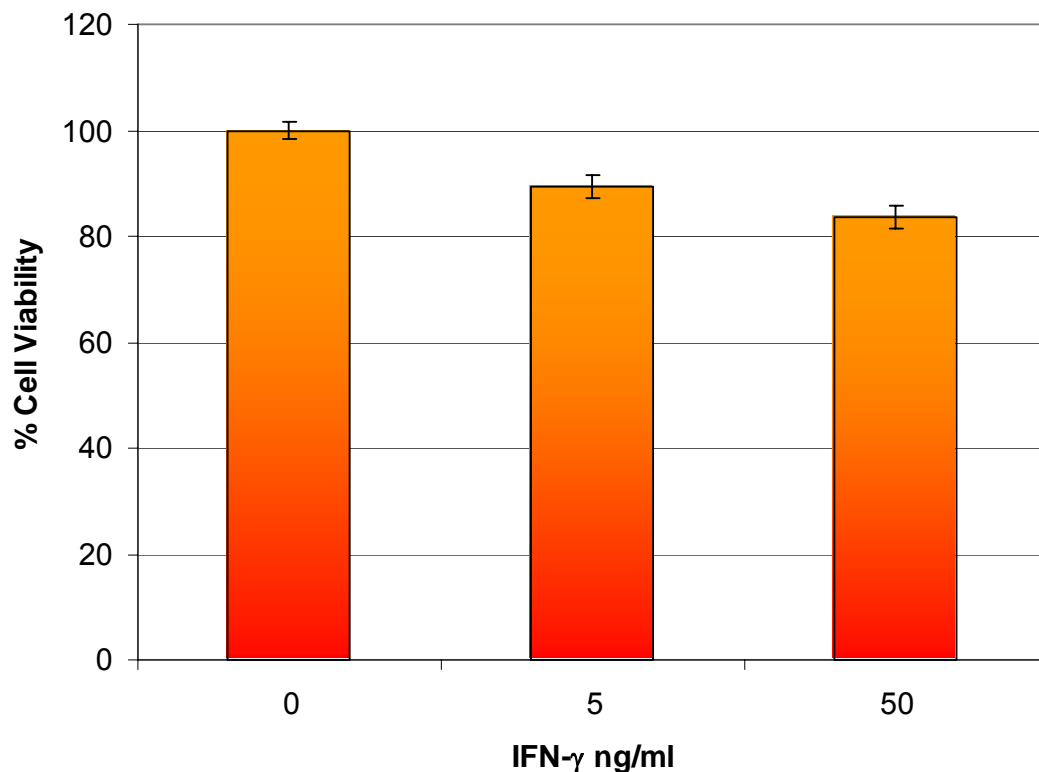
PCR products of BeWo, CHME3 and U251-MG were aligned with MSRV *gag* using the ClustalW algorithm. Identical bases are highlighted in blue, gaps are indicated with -, insertions with '.

Figure 4.13 Phylogenetic analysis of PCR product sequences with MSRV *gag*



Phylogenetic analysis of PCR products aligned with MSRV *gag*. CHME 3 is closely related to the MSRV *gag* prototype sequence, in comparison to isolates sequenced from BeWo and U251-MG cell lines.

Figure 4.14 Cell viability of CHME3 cells treated with IFN gamma

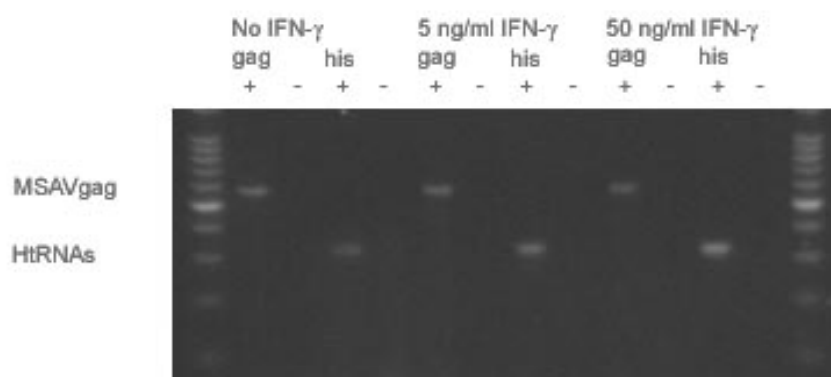


Cell viability of the CHME3 cell line is significantly reduced by IFN- γ ($F=20.52$, $p=0<0.0005$, One-way ANOVA) to 89% and 83% at concentrations of 5 ng/ml and 50 ng/ml respectively.

4.3.5 Modulation of HERV-W gag expression in the CHME3 foetal microglia cell line by IFN- γ

Treatment of cells with 5 ng/ml and 50 ng/ml IFN- γ substantially reduced cell viability to 83-89% of that of untreated cells ($F=20.52$, $p<0.0005$, One-way ANOVA) (Figure 4.14), indicated that IFN- γ had biological activity in the cells, through toxic and anti-proliferative effects. MSRV *gag* expression was decreased in CHME3 cells with increased IFN- γ concentration (Figures 5.15 and 5.16), though this is not significant ($F=2.89$, $p=0.092$, One-way ANOVA). No bands coded for the MSRV *gag* sequence in negative controls, and no genomic *histidyl tRNA synthetase* was detected. However, expression of *histidyl tRNA synthetase* increased as a consequence of cytokine stimulation, indicating poor reliability as a housekeeping gene when cells have been IFN- γ stimulated.

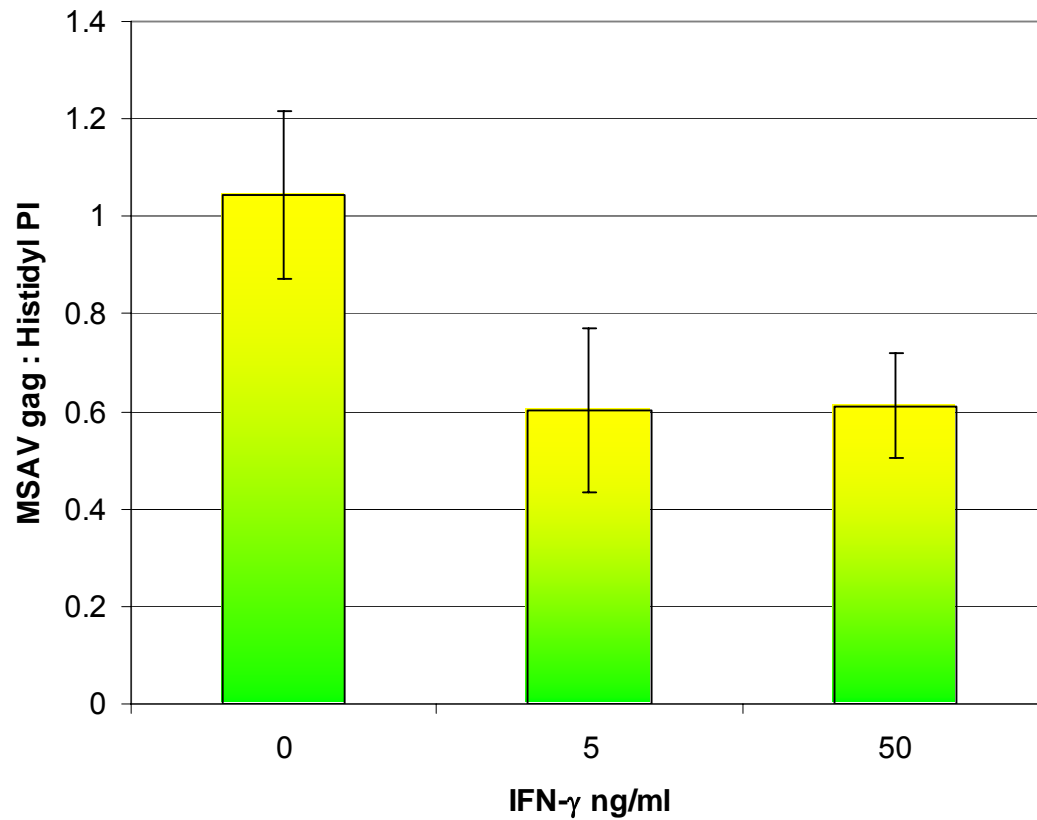
Figure 4.15 Gel illustrating MSRV *gag* RNA levels in CHME3 cell lines treated with IFN- γ



MSRV *gag* (pL=571) transcripts were detectable in both treated and untreated CHME 3 cell lines. *Histidyl tRNA synthetase* (pL=319) expression increased with increased IFN- γ concentration. No transcripts were detected in the reverse transcriptase negative controls.

Number of replicates = 6; 2 RNA batches used for cDNA synthesis, PCR runs performed in triplicate

Figure 4.16 Comparison of HERV-W *gag* mRNA expression in the CHME3 cell line after treatment with IFN- γ for 48 hours



Expression of MSR_V *gag* : *histidyl tRNA synthetase* is decreased with increasing IFN- γ concentration to levels 60% of that of untreated cells at concentrations of 5 and 50 ng/ml. The association is not statistically significant ($F=2.89$, $p=0.092$, One-way ANOVA).

4.4 Discussion

HERV-W *env* mRNA expression is extremely high in the BeWo placental choriocarcinoma cell lines, but is present at lower levels in other tissues. There is strong evidence that although HERV-W *env* appears to be present in all the cell lines tested expression levels are specific to the tissue that the tumour originates from. HERV-W *env* is thought to have a role in mediating trophoblast cell fusion in the placenta (Mi et al., 2000), and the high expression levels in the BeWo cell line are expected. Real time PCR analysis of a panel of HERVs with homology to the gammaretroviruses indicates significantly higher levels of HERV-W expression in the placenta, and slightly higher levels in the brain, testis and uterus (Forsman et al., 2005). Interestingly, the same study shows significant expression of HERV-H and HERV-I in the brain.

Genomic contamination can be ruled out, as the samples have been RNase treated to eliminate genomic DNA. As primers designed for the *histidyl tRNA synthetase* housekeeping gene are intron spanning, the band at 319 bp can only correspond to mRNA. As HERV-W *env* and MSRV *gag* are encoded by a single exon, primers could not be designed in this way, and consequently reverse transcriptase negative controls were required to eliminate the possibility of genomic DNA as a source of the PCR product.

A study by Yi et al. (2004) has indicated that HERV-W *env* is expressed in both normal brain tissue, and in a brain tumour cell line, PFSK-1, though the *gag* gene product is only expressed in normal brain tissue. Sequencing of the products by the authors showed that the clone from normal brain tissue was identical to syncytin, whilst a number of the cancer cells tested showed sequences with greater similarity HERV-W isolates on chromosomes 12,14, 17 and X. Sequence analysis of the PCR product for U251-MG indicated a 100% identity with that of syncytin. The high degree of homology of the transcripts to syncytin indicates that capacity for encoding functional protein is high.

The results show that HERV-W *env* levels differ between the established glioma cell lines and the astrocytoma short term cultures. Comparison of astrocytoma cell lines with mutant P53 (U251-MG, U373-MG, IN1528) against those with wild type P53 (IN859 and IN1472) indicated that expression of HERV-W *env* is considerably increased in the former.

A recent study by Chang et al. (2007) has indicated that HERV-I LTRs, promoters consisting of repeating sequences of nucleotides that flank retroviruses, are negatively regulated by wild type P53 *in vitro*, and that a conformation mutation (V143A) may stimulate transcriptional activity. Wild type P53 has multiple roles in DNA repair, cell cycle arrest and the initiation of apoptosis, and mutation results in breakdown of cell control. Initiation of HERV-W *env* transcription and the presence of such sequences may represent a breakdown of such control. Whether HERV-W *env* is a cause of gliomas or a consequence is not yet clear. Further studies may elucidate whether HERV-W *env* expression levels could be a marker for glioma cell progression. It is of note that HERV-K10 antibodies are raised in patients with seminomas compared to that of healthy controls (Sauter et al., 1995).

MSRV *gag* variants were also detectable in all of the cell lines tested, but as is the case with *pol* (Yi et al., 2004), a large number of frameshifts and point mutations are present, indicating that if the protein is translated it is not likely to be functional, as is the case for other MSRV *gag* isolates, which have premature stop codons and frameshift mutations, such as the CL2 clone which has a stop codon in the capsid region (Komurian-Pradel et al, 1999). In addition to HERV-W, a number of other retroviruses, such as HERV-FRD and ERV-3 do not have functional *gag* regions (Rote et al., 2004). The presence of MSRV *gag* transcripts is striking, as Yi et al. (2002) reported that HERV-W *gag* has a limited distribution, both in cancerous and healthy cells, although the brain tumour line tested was positive for HERV-W *gag*.

Sequence analysis of the variants indicated that there was a large diversity in sequence homology, with CHME3 having the closest homology to that of the GenBank sequence (AF123881) for MSAV *gag*. It would have been worthwhile sequencing the remainder of the PCR products, and testing a number of other tissues (cancerous and non-cancerous) to investigate whether the sequences present in cancer cell lines are more distinct from foetal or benign tissues.

As MSRV *pol* and the presence of particles has shown to be increased in cell lines derived from MS patients by cytokine stimulation, in particular IFN- γ and TNF- α CHME3 cells were stimulated by IFN- γ for 48 hours, and the effects on MSRV *gag* expression. The CHME3 cell line was chosen due to the central role of macrophages and microglia in MS., and that IFN- γ is produced in the CNS by microglia (Kawanokucki et al., 2006) as well as Th1 cells. IFN- γ has a number of effects on macrophages, including anti-microbial and anti-tumour mechanisms (for review, see Schroder et al., 2004). The presence of IFN- γ decreased the viability of CHME3 cells considerably. This was expected as cell proliferation is reduced in IFN- γ exposed cells, which enter S-phase arrest (Matsuoka et al., 1999), and possibly also due to apoptotic effects.

A number of retroviral elements, including a number of retrotransposons (Long interspersed elements (LINEs) and Short interspersed elements (SINES)), HERV-DRB7 (a member of the HERV-K family), and ERV-9 (Andersson et al., 1998) are present in MHC regions. Another effect of IFN- γ is to upregulate both MHC class I and class II expression and antigen presentation. Interestingly, ERV-9 is a retrovirus with a high degree of homology with HERV-W, with approximately 75-90% homology in the *pol* gene (Perron et al., 1997), and has IFN- γ responsive elements in its LTRs (Andersson et al., 1998). If copies of HERV-W *gag* were present in the MHC region, it would be expected that upregulation of MSRV *gag* mRNA expression would occur.

In contrast to *pol*, where expression is upregulated by IFN- γ (Serra et al., 2003), our results indicate that IFN- γ has no significant effect on the expression of MSRV *gag*. It is also worth noting that the expression of the *histidyl tRNA synthetase* housekeeping gene was increased with increasing IFN- γ concentration, affecting the reliability of the results for the IFN- γ quantitation. The choice of housekeeping genes is important in the reliability of RT-PCR, and should be stably expressed in the conditions that are used. The transcription of β actin and glyceraldehyde-3-phosphate dehydrogenase, for example, are significantly increased by exposure to serum, while beta-2-microglobulin and 18s ribosomal RNA are not (Schmittgen et al., 2000). For this reason, applications where enhanced sensitivity is required, such as real time PCR, it is recommended that a panel of housekeeping genes are used.

In summary, primers for HERV-W *env* have been successfully developed, optimised and used in a multiplex PCR system to quantify RNA expression in a variety of cell lines. High levels of expression were present in the placental cell line, in agreement with other studies. Lower levels were detected in astrocytoma cell lines, where the expression of HERV-W *env* showed a correlation with P53 status. No sequences with a high identity to HERV-W *gag* were found, though variants were found in all cell lines tested. Expression of these variants was not altered by exposure to IFN- γ .

The primers developed would be useful in analysing levels of RNA expression in pathological sections of MS and glioma patients in order to see if the expression correlates with the severity of the disease and symptoms. The sequencing data enables the development of probes to investigate by *in situ* hybridisation. The PCR systems, and in particular the multiplex system developed for HERV-W *env* allows investigation *in vitro* to determine the effects of cells cultured in 'disease conditions', e.g. by exposure to other cytokines, steroid hormones, transactivation by exogenous viruses, irradiation, or demethylating agents.

5 Detection of MSRV gag₃₁₁ epitope in astrocytoma cell lines

5.1 Introduction

Bioinformatic analysis identified a potential epitope of HERV-W gag with homology to MBP. This was then immunised into a New Zealand White Rabbit in order to produce polyclonal antibodies to this peptide. After verifying that the rabbit had an appropriate immune response to the peptide, it was necessary to purify the polyclonal IgG using affinity chromatography in order to remove other components of the serum that might interfere with any assays carried out. These purified fractions were then evaluated in order to ensure that the antibody was not denatured during the purification process. Purified antibody was then used in immunocytochemical analysis of a number of astrocytoma cell lines, as well as BeWo (a placental choriocarcinoma cell line), where a large number of retroviruses are expressed, including HERV-E, HERV-R, HERV-H (Forsman et al., 2005) and HERV-W. The aim was to identify whether expression of the MSRV gag₃₁₁ epitope is present in these cell lines, the subcellular localisation, and how the expression differs between cell lines.

PCR analysis indicated that sequences related to HERV-W gag mRNA were detected in astrocytoma cell lines, but this does not necessarily mean that this is translated into a functional product. HERV-W retroviral particles have been demonstrated in cell lines derived from MS patients (Perron et al., 1989; Komurian-Pradel et al., 1999), while retroviral particles (notably from the HERV-K family) are commonly found in germline tumours and testicular tumours (Lower et al., 2001). The vast majority of HERVs do not produce functional products (Kim et al., 2001), with a few notable exceptions, such as syncytin (HERV-W/7q env) (Mi et al., 2000).

Functional gag proteins assemble in the cytoplasm and are transported to the plasma membrane (types A, B, and D retroviruses) or are assembled at the plasma membrane itself (type C retroviruses), and are important in viral assembly. The vast majority of human endogenous virus products which form particles form into immature structures, such as HERV-K-HTDV (human teratocarcinoma derived virus) which forms non-infective particles resembling Type C retroviruses in teratocarcinomas (Boller et al., 1983; Lower et al., 1984). MSRV particles are observable at the surface of the cytoplasmic membrane or clustered in vacuoles of EBV-immortalised lymphoblast cell lines derived from MS patients during peak periods of RT activity (Perron et al., 2001).

In order to detect whether the MSAV gag transcripts detected in astrocytoma cell lines in the previous chapter were translated into protein, a direct immunofluorescence-based protocol was used, involving propagation of cultured cells in culture slides and fixation in paraformaldehyde. Dilutions of purified anti-MSRV gag₃₁₁ antibody were applied to the fixed cells along with anti-alpha tubulin antibody to enable visualisation of cytoskeletal detail upon incubation with fluorophore-conjugated anti-species antibodies. Nuclear detail was revealed by staining with DAPI.

5.2 Methods

5.2.1 Immunisation of rabbits

Antibodies specific for the MSRV gag₃₁₁ peptide were raised commercially by Severn Biotech, Kidderminster, Worcestershire (home office license number: 40\2154). New Zealand White Rabbits were immunised with MSRV gag peptide (LREALRKHTSLSP) (Severn Biotech) bound to purified protein derivative according to the schedule outlined in Table 5.1. Carrier molecules, such as purified protein derivative (PPD) are necessary in order to elicit a sufficient immune response. Test bleeds were taken, and analysed by ELISA at regular intervals.

Table 5.1 Schedule of immunisation and test bleeds

Day	
0	Immunisation with PPD-conjugated peptide + Freund's complete adjuvant (FCA)
28	Booster immunisation with PPD-conjugated peptide + Freund's incomplete adjuvant (FIA)
38	1st test bleed taken
42	2nd booster immunisation with PPD-conjugated peptide + FIA
52	2nd test bleed taken
56	3rd booster immunisation with PPD-conjugated peptide + FIA
66	3rd test bleed taken
70	4th booster immunisation with PPD-conjugated peptide + FIA
80	Terminal bleed taken

5.2.1 ELISA for testing antibody reactivity in rabbit sera and purified terminal bleeds

50 µl of 50 ng MSRV gag₃₁₁ peptide in phosphate buffered saline (PBS), pH 7.2 stock was added to the wells of a Nunclon Δ 96 well plate (Nunc), and was left to coat overnight at 37 °C. The plates were washed thoroughly four times in PBS (Sigma) with 0.05% Tween-20 using an automated plate washer (WellWash Mk4), and blotted dry.

The plates were then blocked with 150 µl of blocking buffer (PBS with 1% bovine serum albumen (BSA); 0.05% Tween-20) and incubated at room temperature for 1 hour, before being washed four times in PBS (Sigma) with 0.05% Tween-20 using the automated plate washer, as before. This was done to reduce potential unwanted interactions of antibody with the plasticware. For the rabbit sera (Severn Biotech), serial dilutions (either a 2-fold series from 1:200 to 1:1600, or a 10-fold series from 1:100 to 1:10,000) of each test bleed were added in duplicate. With the purified antibody, fractions were added in triplicate to the wells at a concentration of 1:200. Negative controls were included, including peptide-free, primary antibody negative, and irrelevant peptide controls (e.g. predicted epitopes for other HERVs such as HERV-K10 gag).

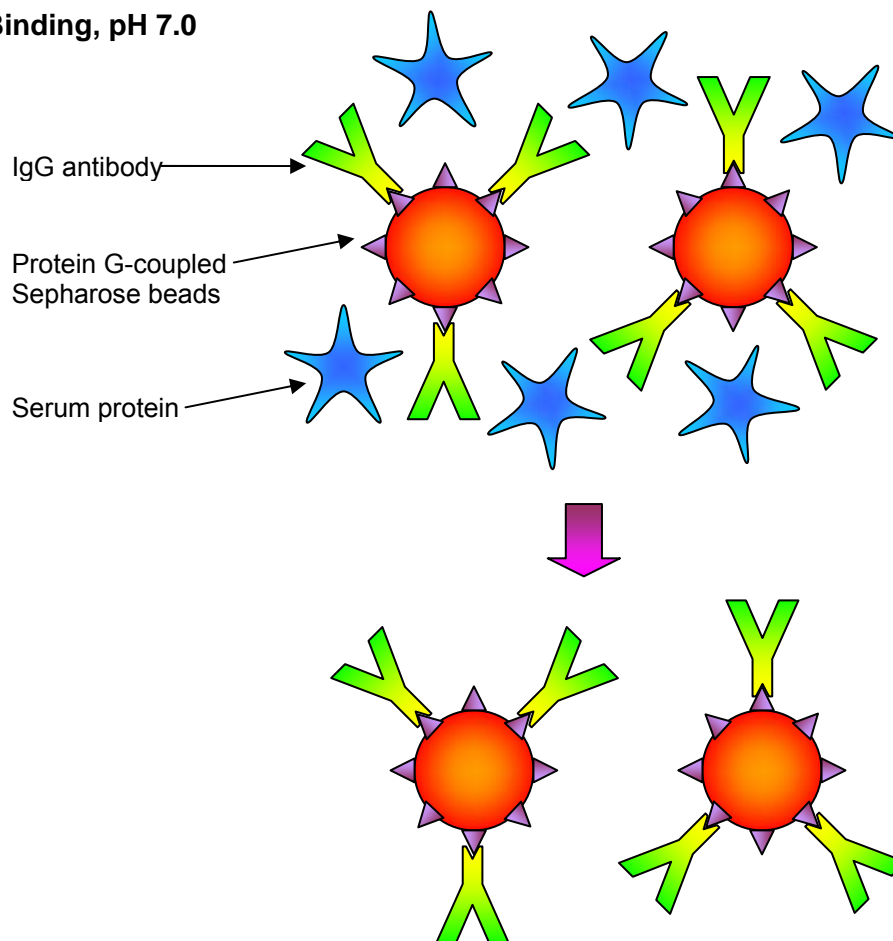
After incubation for 1 hour at room temperature the plates were washed 4 times with PBS and blotted dry. 50 µl of HRP conjugated polyclonal swine anti-rabbit immunoglobulins (Dako) diluted 1:800 in blocking buffer was added to each well in the plate, and incubated at 1 hour at room temperature. After washing and blotting 50 µl of TMB solution (Sigma) (for rabbit sera) or 0.4 mg/ml OPD (Sigma) (for purified terminal bleed antibody) in phosphate-citrate buffer with sodium perborate (Sigma) was added to each well on the plate, and incubation took place at 37 °C in the dark for 10 or 30 minutes (for TMB and OPD respectively). Readings were also taken at 620 nm and 450 nm after stopping the reaction by the addition of 50 µl 2N hydrochloric acid for TMB, and at 450 nm and 492 nm (stopped) for OPD.

5.2.3 Affinity purification of rabbit bleeds

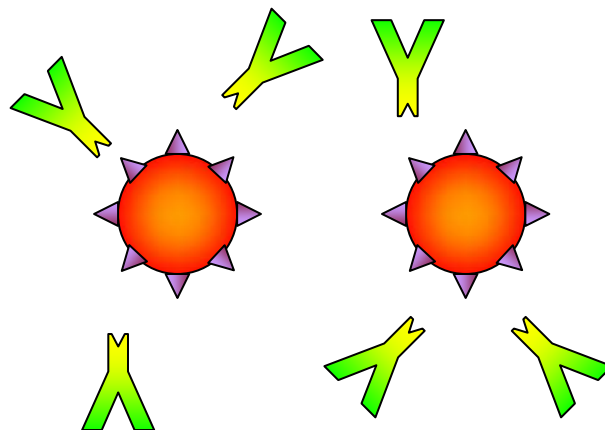
Purification of antibody from serum is often necessary to remove unwanted components, such as albumin, transferrin, and α2-macroglobulin, which may interfere with assays that the antibodies are used in, e.g. by blocking antibody sites in ELISA, or sticking to cells analysed in immunocytochemistry or immunofluorescence.

Figure 5.1 *Schematic representation of the principle of antibody purification using a Protein G column*

(a) Binding, pH 7.0



(b) Elution, pH 2.7



(a) Antibody binds with high affinity to protein G, and remains covalently bound to the matrix while other serum proteins pass through the column. (b) By changing the pH from 7.0 to 2.7, the interaction is disrupted and the purified antibody is eluted.

Affinity chromatography relies on the separation of biological molecules based on reversible interactions with specific ligands bound to a matrix, such as sepharose. The ligands bind reversibly to the antibody of interest at pH 7, whilst other, undesired components of the sample are not retained by the column (Figure 5.1). Antibody purification can be carried out using chromatography with sepharose bound to protein A (derived from *Staphylococcus aureus*) or protein G (derived from Group G *Streptococci*), or by ion exchange chromatography.

Separation using protein A or protein G ligands are rapid, and require little optimisation. Both protein A and protein G are suitable for purifying polyclonal rabbit antibody, though protein G binds to a wider range of IgGs. To disrupt the binding of the antibody to the ligand, the conditions are made more acidic, and the antibody is eluted. Antibodies, as with many proteins denature at low pH, so it is important to neutralise the eluent rapidly.

Laboratory protocol

1.25 ml of rabbit serum terminal bleed (Severn Biotech) was diluted 1:1 in 1.25 ml binding buffer (20 mM sodium phosphate) (Amersham) and loaded into a syringe. The AKTAprime system (GE Healthcare) was washed and prepared as per the User Manual, and a HiTrap Protein G HP column (Amersham) and 500 µl sample loop was connected to the system. The inlet tubing for buffer A was placed in binding buffer, and the tubing for buffer B placed in elution buffer (0.1 M glycine-HCl, pH 2.7) (Amersham), and the waste outlets placed in an empty conical flask. Collection tubes were filled with 100 µl neutralisation buffer (Amersham) (1M Tris, pH 9.0). The neutralisation buffer is necessary to reduce potential protein denaturation in the acidic elution conditions used.

The sample loop was filled by adding 2.5 ml of 1:1 rabbit serum in binding buffer and the recorder was connected to the system, powered on, and pens connected. The chart speed and voltages were set to appropriate values (e.g. 10 mm/min, 0.5 V). The MAbTrap step elution program was selected on the AKTAprime system, and the purification started. Fractions collected were analysed for activity using ELISA, and stored at -20 °C.

5.2.4 Culture of cells grown as monolayers

Cells were grown to $\geq 70\%$ confluency in T75 tissue culture flasks (Sarstedt) under the appropriate conditions shown in Table 5.2 at 37 °C. To passage the cells further, the supplemented media was aspirated, and cells were washed twice with 2ml HBSS (Sigma) and detached by the addition of 2 ml of Trypsin-EDTA solution (Sigma), followed by incubation for 2 minutes at 37 °C (for complete detachment of cells). Trypsinisation of the cells was ended by the addition of 8 ml of supplemented media, and an appropriate dilution of cells (between 1:2 and 1:10) was set up in fresh flasks containing 25 ml of fresh supplemented media.

Table 5.2 Cell lines used and culture conditions

A list of the cell lines used for immunofluorescence, and the culture conditions used. All the cell lines, except the IN series are established cell lines, which are short term cell cultures derived from patient biopsies.

Cell Line	Description	Conditions	Source
BeWo	Choriocarcinoma	F12 Ham's (Sigma) w/ 10% FCS (PAA) [5% CO ₂]	ECACC
IN859	Grade IV astrocytoma	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
IN1472	Grade IV astrocytoma	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
IN1528	Grade IV astrocytoma	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
IN1612	Grade IV astrocytoma	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
U87-MG	Grade III Glioblastoma	MEM (Sigma) w/ 10% FCS (PAA) 1% NEAA (Sigma) 1% L-Glutamine (Sigma) 1% Sodium pyruvate (Sigma)	Prof. Nicola Woodroffe, Sheffield- Hallam University
U251-MG	Grade IV : Glioblastoma Multiforme	F10 Ham's w/ HEPES (Sigma) w/ 10% FCS (PAA)	Institute of Neurology
U373	Grade III Glioblastoma	DMEM (PAA) w/ 10% FCS [5% CO ₂]	ECACC

5.2.5 Indirect immunofluorescence of cell lines

Immunofluorescence is a procedure that uses fluorescently tagged antibodies in order to directly detect antigens within a cell or tissue. Cells are first attached to a slide, using a fixative such as methanol or paraformaldehyde. Paraformaldehyde fixation cross-links proteins and DNA (hence is also used as a fixative in *in-situ* hybridisation), and causes minimal loss of the native structures in the cell. If paraformaldehyde is used as a fixative, it is also necessary to permeabilise the cell membrane to allow entry of antibodies, and other large molecules such as blocking agents.

As with other protocols involving the use of antibodies, a blocking step is employed in order to block unwanted antigenic sites, and decrease background fluorescence. BSA is a commonly used blocking buffer, but powdered milk and normal serum (from the species that the secondary antibody is raised in) are other alternates. The primary antibody is applied, and these bind to the epitopes present in the target protein. It is also useful to have a positive control antibody, which is known to be expressed in the cells in order to avoid false negatives, such as anti-tubulin or against a specific marker (e.g. glial fibrillary acidic protein for astrocytes, or MHC-II for macrophages). Negative controls, either from non-immune serum from the same species as the primary antibody, or antibody raised against a protein not expressed in the cell type under investigation are required to eliminate false positives.

In direct immunofluorescence, the primary antibody is already attached to the fluorophore, and this can be detected under a fluorescent or a confocal microscope. Where indirect immunofluorescence is used, it is necessary to bind a labelled secondary antibody directed against the species in which the primary antibody is raised. While indirect immunofluorescence requires this step, which makes the procedure slower, it is more sensitive as multiple secondary antibodies can bind to the primary antibody, acting as an amplification step.

It is also necessary to counterstain the slides, so that all cells can be identified. Commonly used counterstains in immunofluorescence include 4'-6'-diamidino-2-phenylindole (DAPI) which binds to double stranded DNA (highlighting cell nuclei) and fluorophore-conjugated phalloidin, an F-actin-binding toxin from the death cap mushroom (*Amanita phalloides*) which reveals cytoskeletal detail.

Laboratory protocol

Cells were cultured at a seeding density of 320 cells/well in a BD Falcon CultureSlide (BD Biosciences), and incubated at 37 °C for 24 hours. The media was removed, and adherent cells were washed twice in 300 µl of PBS (Sigma) per well, followed by fixation in 200 µl 4% paraformaldehyde (Sigma) solution (in PBS) per well for 15 minutes. The fixation reaction was quenched by removing the paraformaldehyde solution, and adding 200 µl 0.1 M glycine. The fixed cells were then permeabilised using 0.1% Triton X-100 (Sigma) in PBS for 90 seconds, followed by washing twice in 300 µl of PBS.

The CultureSlide chambers were blocked with 200 µl 4% BSA in PBS at 37 °C for 1 hours, followed by washing twice in 300 µl of PBS. Purified antibody (diluted at 1:25 or 1:50) or control (pre-immune bleeds or antibody negative) were prepared in 1:800 mouse anti-tubulin antibody (Sigma) in 1% BSA/0.1% Tween-20 in PBS. 150 µl amounts of the antibody mixture were added to the appropriate chambers of the CultureSlide and incubated at 4 °C overnight. After incubation, the CultureSlide was washed 3 times in 300 µl of PBS, and secondary antibodies (Alexa-Fluor ® 594 donkey anti-rabbit (Invitrogen) and Alexa Fluor ® 488 goat anti-mouse (Invitrogen)), diluted 1: 400 in 1% BSA/0.1% Tween-20 in PBS) added to the CultureSlide chambers, and incubated at 37 °C for 1 hour in the dark. The CultureSlides were then washed twice in 300 µl of PBS, before being counterstained with 50 µl of 1 µg/ml DAPI in ethanol for 90 seconds. These were then washed twice in 300 µl of PBS, left to dry for 5 minutes, and the chambers were removed. The slides were then mounted in fluorescent mounting medium (Dako). The slides were then visualised and scanned at a 40x resolution under a Olympus BX61 fluorescence microscope and images captured using CytoVision software (Applied Imaging).

5.3 Results

5.3.1 ELISA reactivity of MSR_V gag₃₁₁ antibodies in rabbit test bleeds

The MSR_V gag₃₁₁ peptide elicited a strong immune response in immunized rabbits (Figures 5.2 (2 fold series) and 5.3 (10 fold series)), with titres detectable at dilutions as low as 1:10,000, whereas minimal cross-reactivity was found with peptides derived from the gag epitope predicted in HERV-K10 (RIGKELKQAGRKGNI). The strongest bleed was the second test bleed, though all immune bleeds were deemed suitable for use for further work. The terminal bleed was selected for purification as larger quantities were available.

5.3.2 Purification of antibodies and response of purified fractions

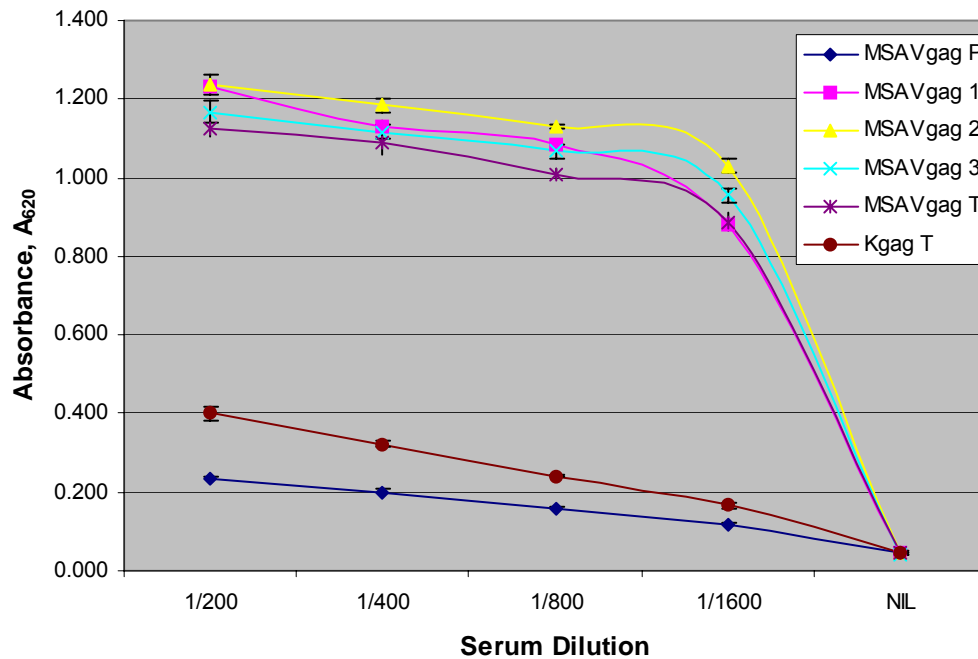
Antibody purification produced a significant antibody peak during elution between approximately 20 and 30 minutes after commencing the run (Figure 5.4), with fractions were collected at 60 second intervals. Testing of the fractions eluted in ELISA indicated that fraction 3 had the highest absorbance (A_{492}) value (Figure 5.5), and therefore the greatest amount of antibody reactivity. The ELISA results also confirmed that the purified antibody still had affinity for the MSRV gag₃₁₁ peptide, and was considered suitable for use for further experimentation.

5.3.3 Detection of MSRV gag₃₁₁ epitope in astrocytoma cell lines

All of the cell lines tested had positive staining (in red) for the MSRV gag₃₁₁ epitope (Figures 5.6-5.13, summarised in Table 5.3). Prominent nuclear staining in red, with prominent foci were present in all the cell lines tested at a 1:25 antibody dilution (of fraction 3), and were prominent at a 1:50 dilution in BeWo, IN1528, IN1612, and U251-MG cell lines. This staining pattern was only observed where MSRV gag₃₁₁ antibody was used, and not in controls using a dilution of serum from the same rabbit prior to immunisation, or where no primary antibody was applied. Whilst it would have been more appropriate to use purified pre-immune bleeds as negative controls, this was not practical as each purification run (repeated 3 times) requires 1.25 ml of sample, whereas only approximately 4.5 ml was provided by the manufacturer for each bleed, with the exception of the terminal bleed (approximately 20 ml). No reactivity against the epitope was identified in the cytoplasm or at the plasma membrane in any test or control, where it might be expected that gag or gag-pol polyproteins accumulate in the cells.

The staining was strongest in the BeWo cell line and the *P53* mutant grade IV astrocytoma cultures, namely the IN1528 and IN1612 short-term cultures, and the established grade IV glioblastoma multiforme cell line U251-MG. The grade III glioblastoma cell lines, U373-MG and U87-MG (with mutant and wild type *P53* respectively) showed less reactivity against the anti-MSRV gag₃₁₁ antibody, as did the *P53* wild type short-term cultures IN859 and IN1472. This data suggests that while not directly linked with *P53* status, expression of the epitope appears to be stronger where the cells are derived from higher grade tumours, and the gross abnormalities that can result from *P53* mutation, such as aneuploidy and loss of heterozygosity in a number of chromosomes.

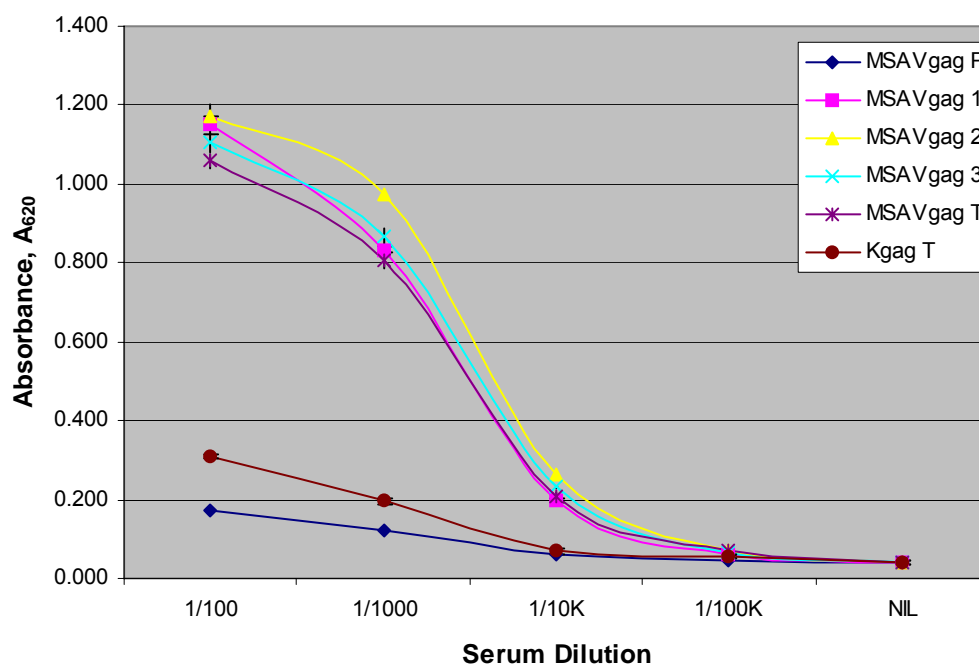
Figure 5.2 2-fold serial dilution for MSRV gag₃₁₁ protein and rabbit test sera



In a 2-fold serial dilution series, antibody titres remain high, even at 1:1600 (ANOVA: $p < 0.001$ for all dilutions; $p = 0.729$ for nil). The 2nd test bleed is the strongest bleed, though all post-immunisation bleeds would be suitable for antibody purification. The secondary antibody used, HRP conjugated polyclonal swine anti-rabbit immunoglobulins (Dako) is specific against all immunoglobulin subtypes, including both IgM and IgG, so no inferences can be made whether the response is IgM or IgG-predominant.

Number of replicates = 6; 3 plates run with duplicate wells on each plate

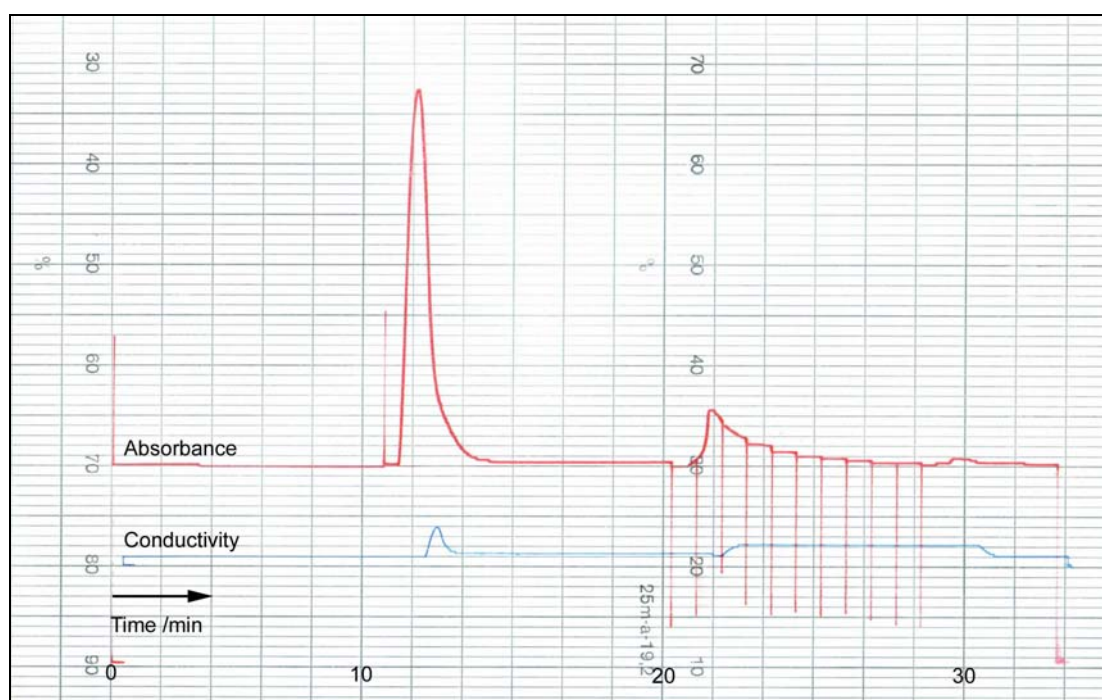
Figure 5.3 10-fold serial dilution for MSR_{SV} gag₃₁₁ protein and rabbit test sera



In a 10-fold serial dilution series, antibody titres are diluted out at >1:100,000 (ANOVA: $p < 0.001$ for all dilutions; $p = 0.781$ for nil). As with the 2-fold series, the 2nd bleed (14 days) gives the highest antibody titre.

Number of replicates = 6; 3 plates run with duplicate wells on each plate

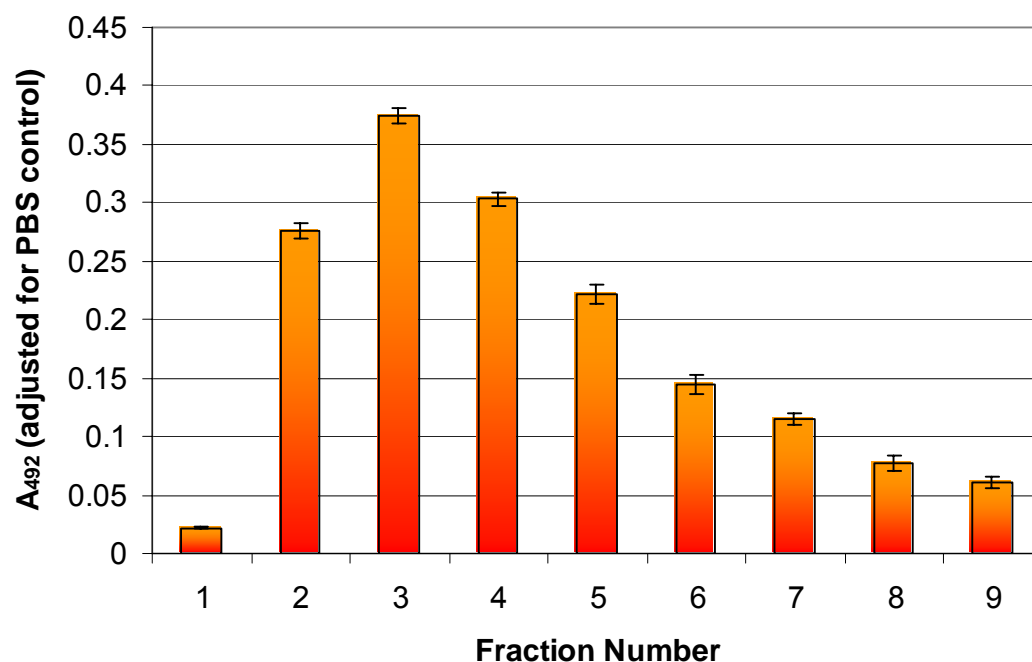
Figure 5.4 Output of recorder for antibody purification for SEV 3 (anti-MSRV gag₃₁₁) terminal bleed (run 2)*



Peaks in absorbance, or protein detected (red trace) were seen at 12 minutes and 22 minutes. The first peak is the flow-through the column consisting of unwanted components of the serum, such as albumen, while the second peak represents the eluted antibody. The blue trace measures the conductivity in the system, and rises during the flow-through peak (due to ions present in the serum) and during the elution step (due to the acidic buffer used to break antibody-antigen interactions).

* Run was performed 3 times

Figure 5.5 Reactivity of purified SEV 3 (anti-MSRV gag₃₁₁) terminal bleed (run 2) antibody fractions



All fractions of purified antibody retain activity against the MSRV gag₃₁₁ peptide. A peak of antibody reactivity (measured at 492 nm) is present in the 3rd fraction (22-23 minutes after commencing the run).

Number of replicates = 6; 2 plates run with triplicate wells on each plate

Table 5.3 Summary of anti-MSRV gag₃₁₁ nuclear staining in cell lines tested

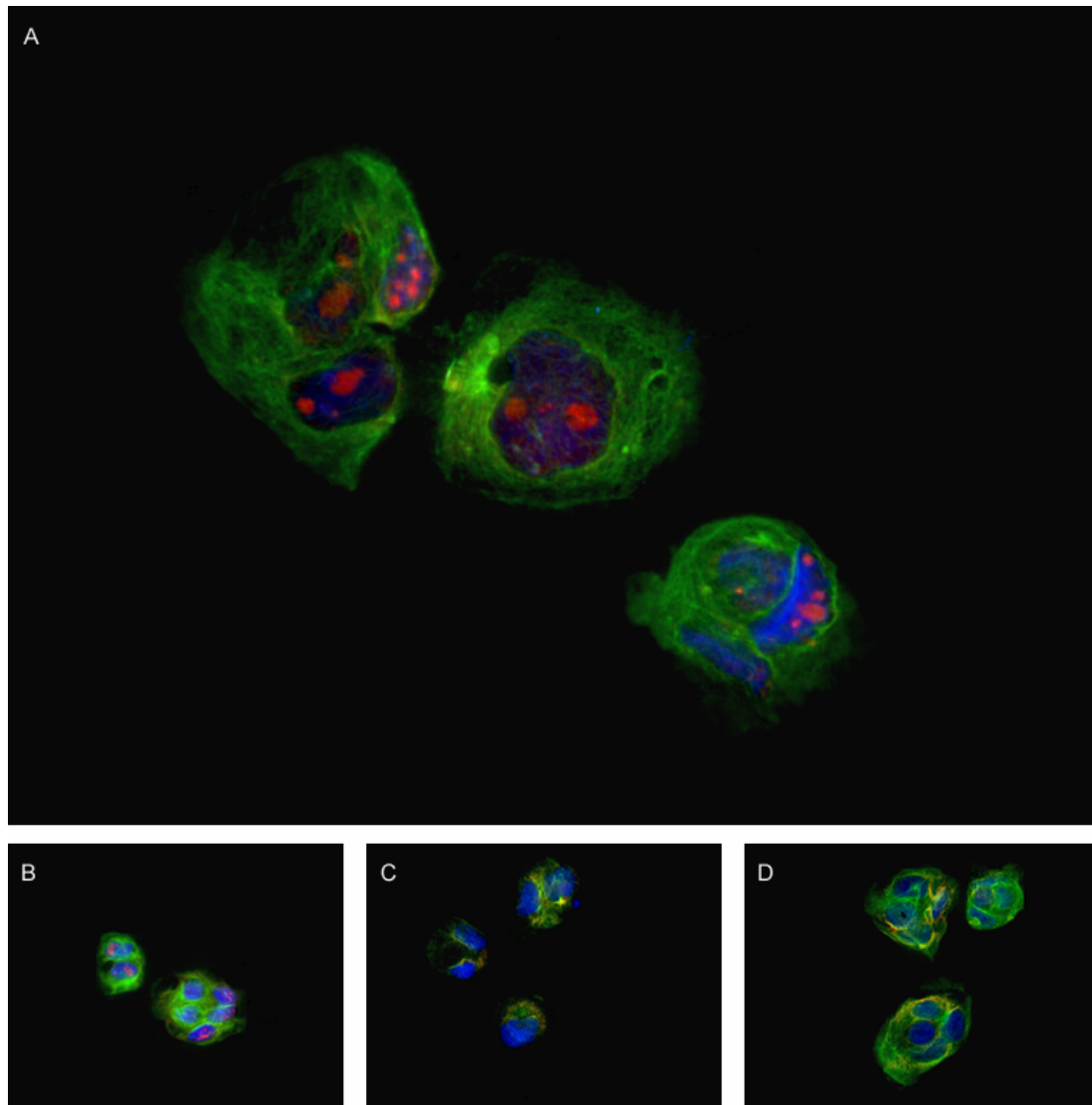
Cell Line	Type	MSRV gag ₃₁₁ antibody dilution				Figure
		1:25	1:50	preimmune serum	PBS	
BeWo	Placental choriocarcinoma	+++	+++	-	-	Fig. 5.6
IN859	Grade IV astrocytoma	++	+	-	-	Fig. 5.7
IN1472	Grade IV astrocytoma	++	+	-	-	Fig. 5.8
IN1528	Grade IV astrocytoma	+++	+++	-	-	Fig. 5.9
IN1612	Grade IV astrocytoma	+++	++	-	-	Fig. 5.10
U87-MG	Grade III glioblastoma	++	+	-	-	Fig. 5.11
U251-MG	Grade IV glioblastoma multiforme	++	++	-	-	Fig. 5.12
U373-MG	Grade III glioblastoma	++	+/-	-	-	Fig. 5.13

Legend

+++	strong staining, punctate distribution
++	moderate staining, punctate distribution
+	weak staining, punctate distribution
+/-	very weak staining, diffuse distribution
-	no staining detected

Number of replicates = 2; fixing and staining performed on 2 different passages of cell lines

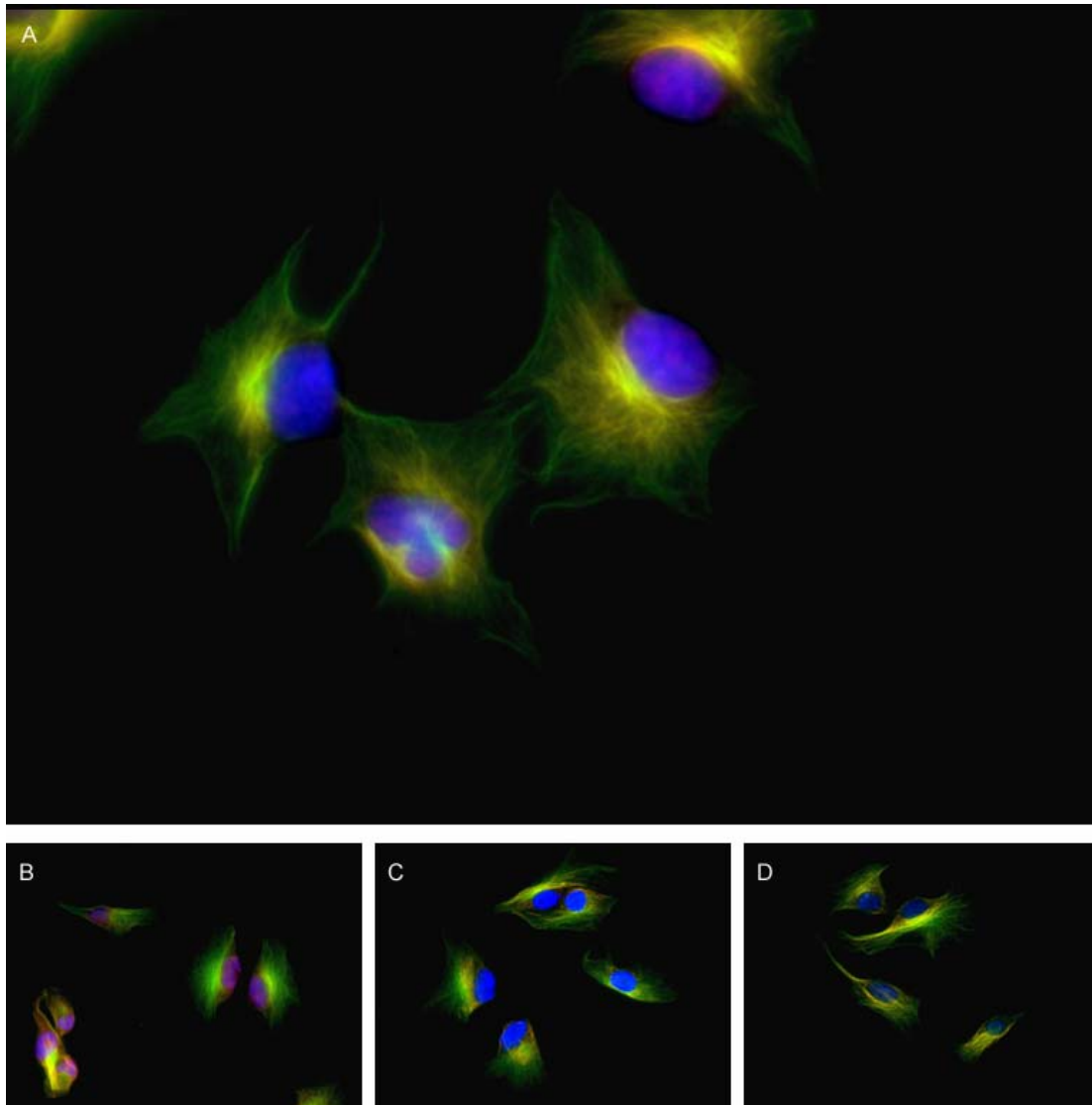
Figure 5.6 Immunofluorescent microscopy of BeWo cells stained for MSRV gag₃₁₁



BeWo cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAP in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:50 and 1:100 dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

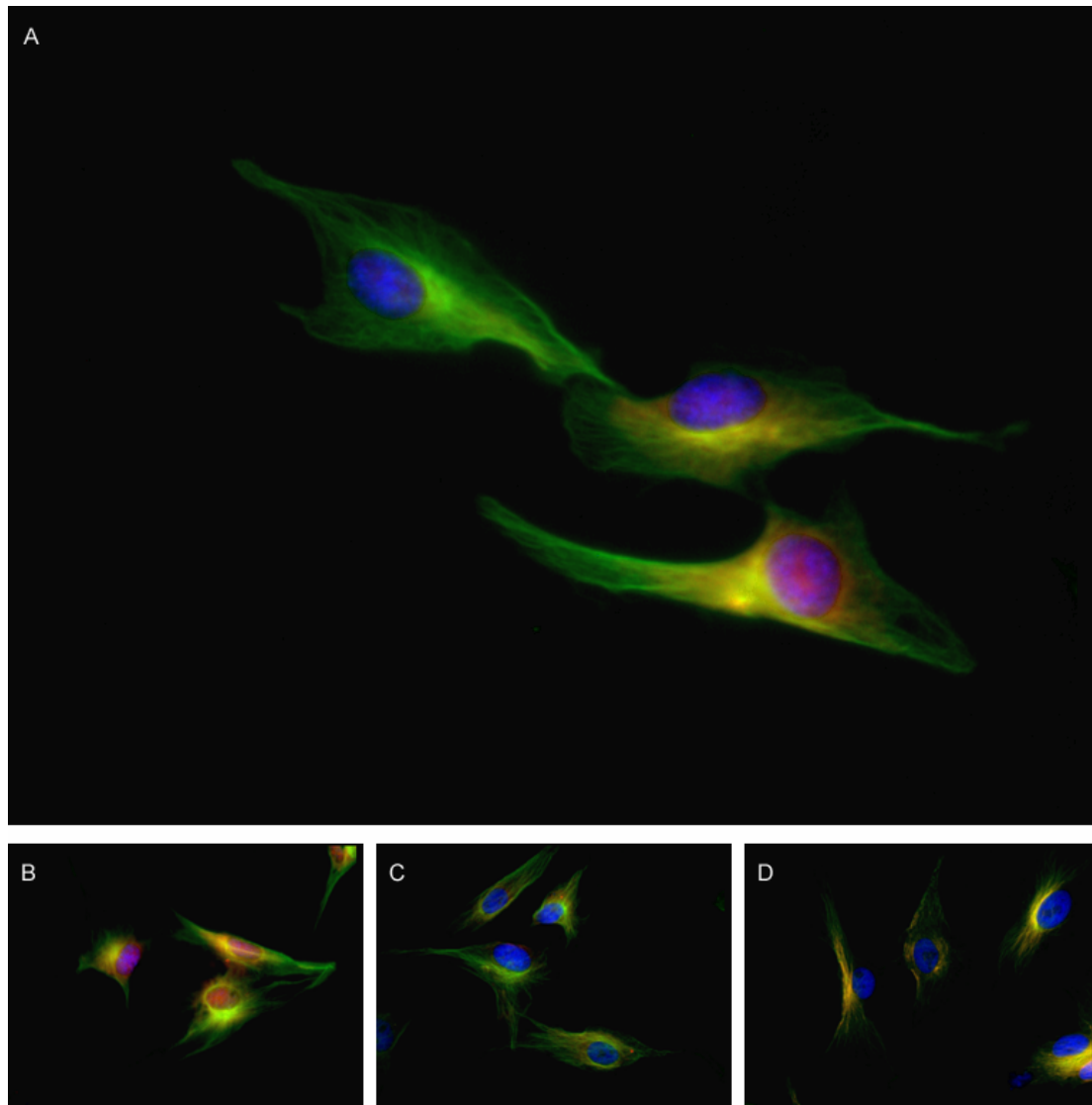
Figure 5.7 Immunofluorescent microscopy of IN859 cells stained for MSRV gag₃₁₁



IN859 cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAP in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:25 and 1:50 (weak staining) dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

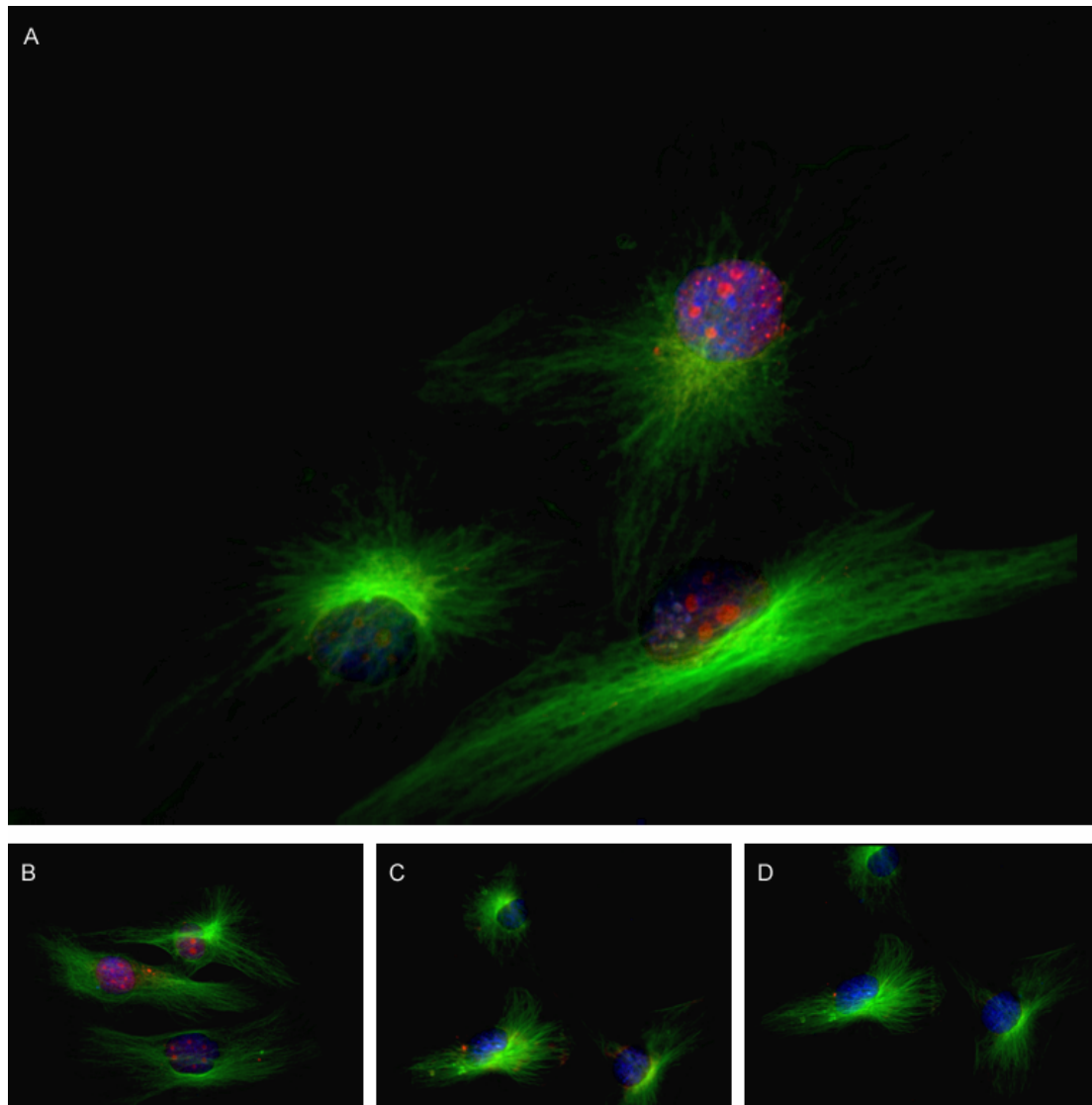
Figure 5.8 Immunofluorescent microscopy of IN1472 cells stained for MSRV gag₃₁₁



IN1472 cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAPI in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:25 and 1:50 (weak staining) dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

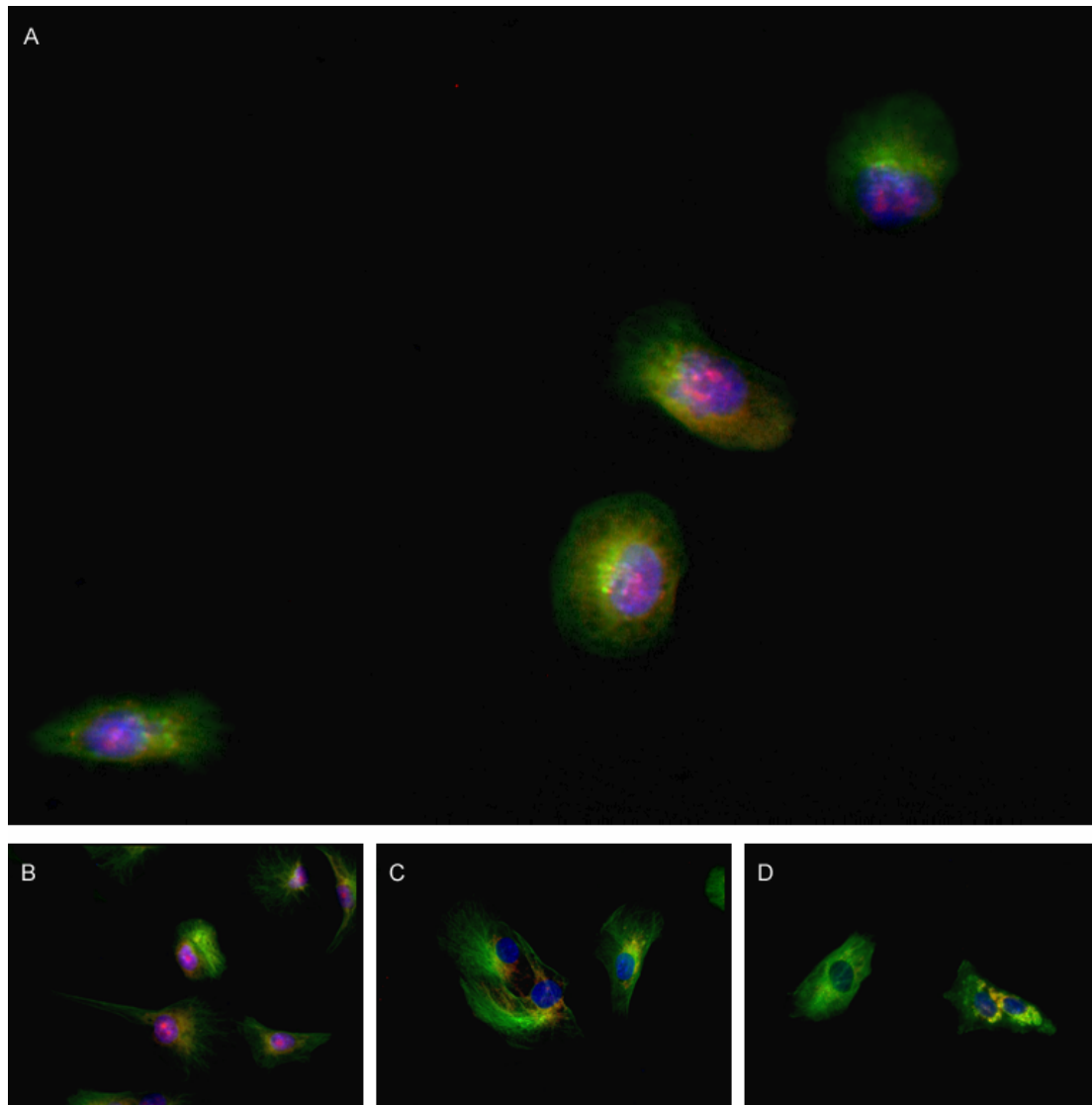
Figure 5.9 Immunofluorescent microscopy of IN1528 cells stained for MSRV gag₃₁₁



IN1528 cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAPI in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Strong punctate staining (red) within the nucleus can be seen in both the 1:25 and 1:50 dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

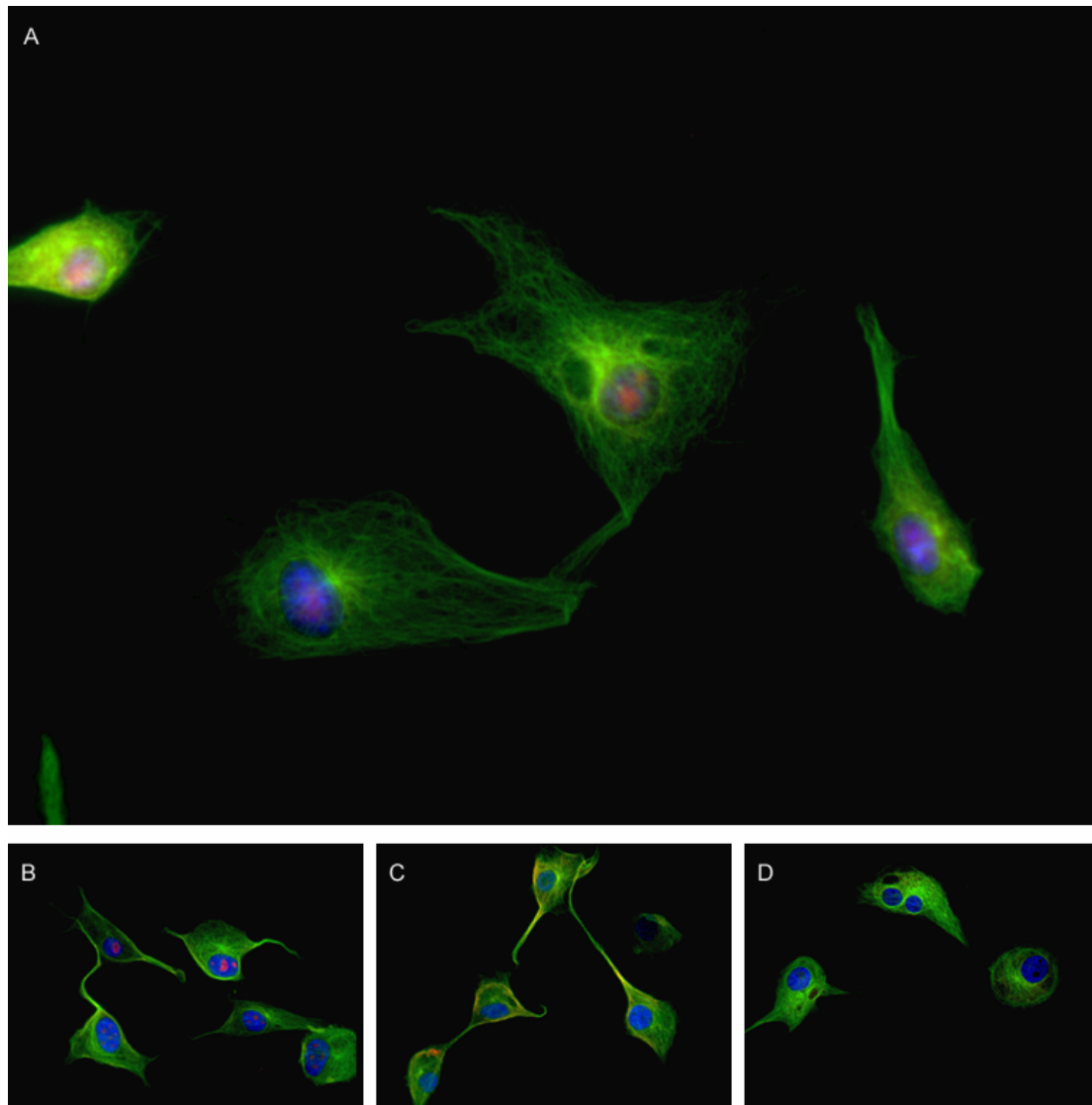
Figure 5.10 Immunofluorescent microscopy of IN1612 cells stained for MSRV gag₃₁₁



IN1612 cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAP in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:50 and 1:100 dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

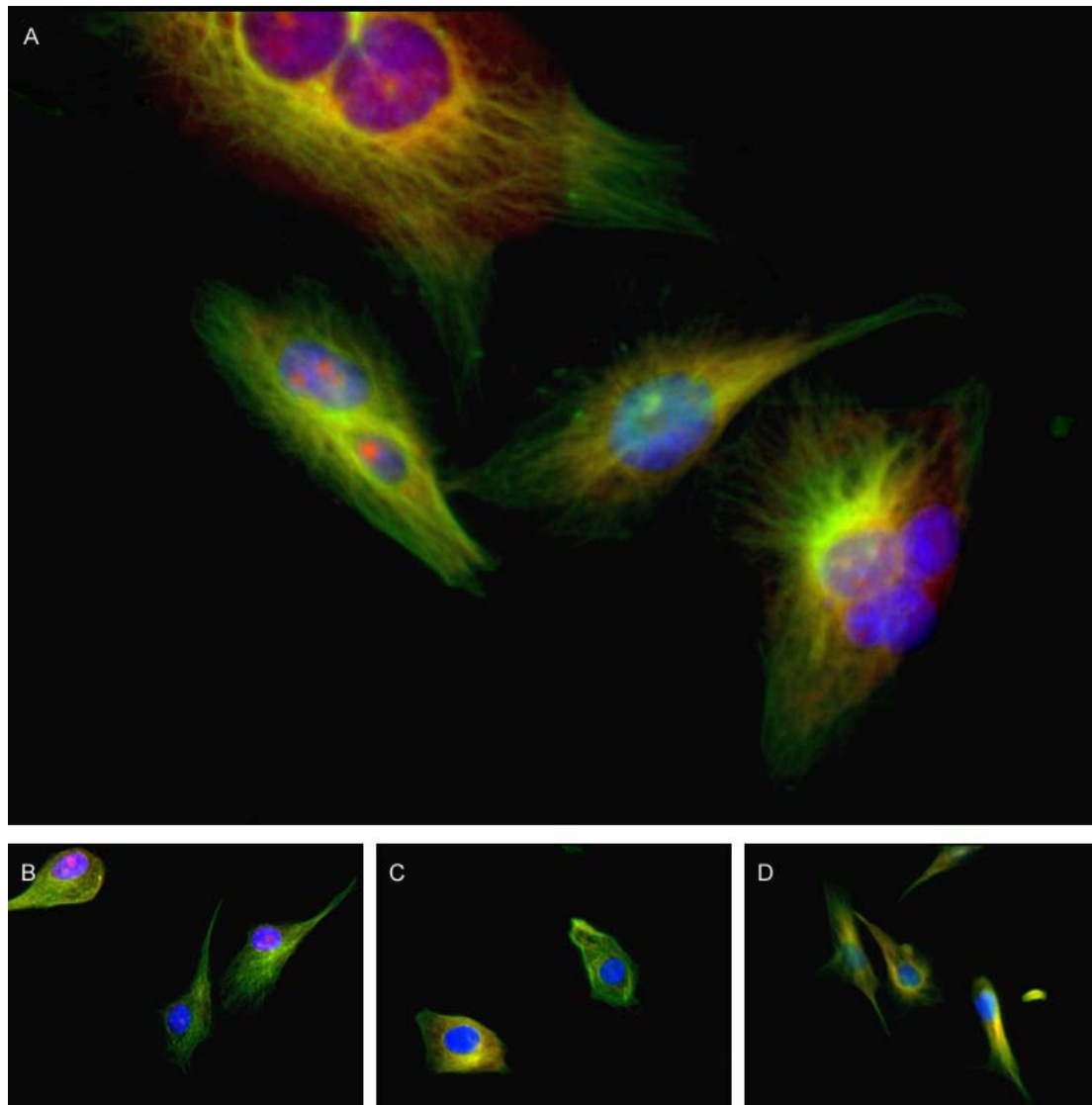
Figure 5.11 Immunofluorescent microscopy of U87-MG cells stained for MSRV gag₃₁₁



U87-MG cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAP in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:50 and 1:100 dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

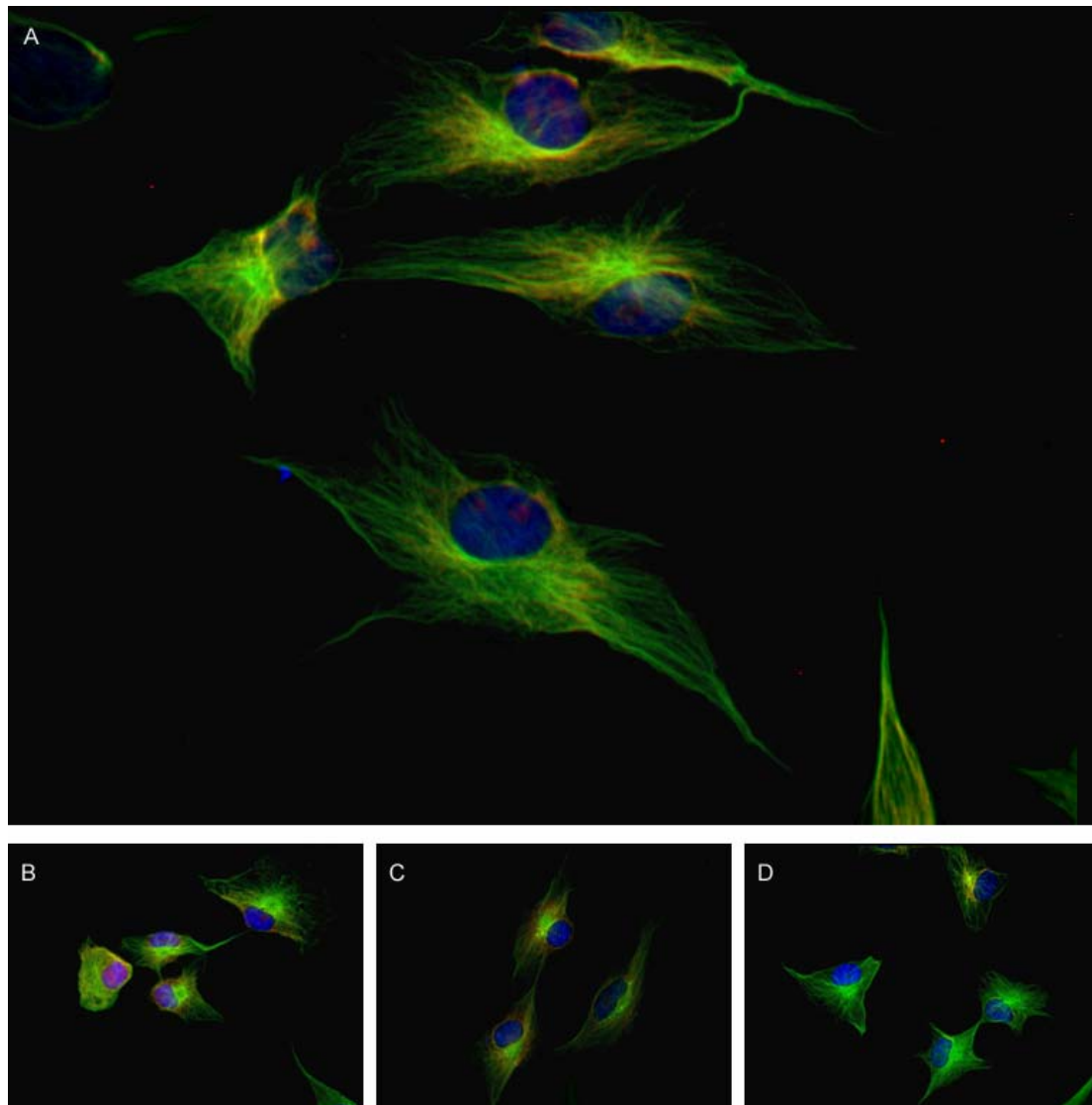
Figure 5.12 Immunofluorescent microscopy of U251-MG cells stained for MSRV gag₃₁₁



U251-MG cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAPI in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:50 and 1:100 dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

Figure 5.13 Immunofluorescent microscopy of U373-MG cells stained for MSRV gag₃₁₁



U373-MG cells were fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton X-100 in PBS and incubated with 1:400 monoclonal mouse anti-tubulin antibody with (A) 1:50 polyclonal rabbit anti-MSRV gag₃₁₁ antibody; (B) 1:25 polyclonal rabbit anti-MSRV gag₃₁₁; (C) 1:200 pre-immune rabbit serum; and (D) PBS. Secondary incubation was performed using 1:400 Alexa-Fluor® 594 donkey anti-rabbit and Alexa-Fluor® 488 goat anti-mouse antibodies and counterstained with 1 µg/ml DAP in ethanol. Slides were scanned at 40x magnification using Cytovision software.

Punctate staining (red) within the nucleus can be seen in both the 1:25 and 1:50 (weak staining) dilutions of anti-MSRV gag₃₁₁, which is absent in both the pre-immune and PBS negative controls.

5.4 Discussion

The MSRV gag₃₁₁ epitope was detected in all of the cell lines tested, but not in any of the negative controls. If retroviral packaging was taking place, we would expect that the gag would be detected in the cytoplasm of the cell, or accumulating at the cell membrane. Retroviruses with a type B or type D morphology assemble in the cytoplasm where they are transported to the plasma membrane for release (Rhee et al., 1990), whilst type C retroviruses assemble at the plasma membrane (Henderson, 1983). Type A particles, such as IAPs associated murine endogenous retroviruses and a number of HERVs such as HERV-K (HTDV) bud into the endoplasmic reticulum. Instead, the epitope was detected in the nucleus of the cells, often exhibiting a punctate distribution. Intranuclear localisation has been demonstrated in a number of exogenous retroviruses including MoMuLV (Nash et al., 1993), foamy viruses (Chang et al., 1995), and HIV Tat and Rev (Truant & Cullen, 1999), as well as the endogenous retrovirus (Rec) product of HERV-K (HTDV) (Magin-Lachman et al., 2001), but not HERV-W.

Analysis of the protein sequence of MSRV gag has identified a potential bipartite nuclear localisation sequence, located between 269-286 aas (Figure 5.14), with the MSRV gag₃₁₁ epitope lying further upstream. The consensus sequence for such nuclear localisation signals is BBX₁₀BXBXB, consisting of two basic residues, followed by an alpha helix of 10 amino acids, and with at least 3 of the following 5 bases being basic residues (Figure 5.15). Such sequences have been identified in MoMuLV (Nash et al., 1993), as well as established nuclear proteins such as nucleoplasmin (a molecular chaperone) (Robbins et al., 1991) and P53. Wild type P53 is predominately located in the cytoplasm of cells, but can shuttle to and from the nucleus due to the presence of nuclear localisation signals and nuclear export sequences (Canman et al., 1997). At least two nuclear localisation sequences are encoded by P53, a monopartite sequence with homology to SV40 (Dang et al., 1989), and a BBX₁₀BXBXB bipartite sequence (Liang & Clarke, 1999) which associates with importin α , and it docks to the nuclear membrane in association with importin β as part of a heterodimeric nuclear import complex. Mutations in these sequences cause aberrant distribution of P53 in the cell, affecting its normal function as a tumour suppressor gene.

Figure 5.14 Protein sequence of MSRV gag (AAD48375) indicating locations of potential nuclear localisation sequence and linear epitope of antibody.

```

1 lerilenwdq cdtqtlrkkf fifcstawp qyplqgretw lpegsinyini ilqldlferk
61 egkwsevpvy qtffslrdns qlckkcglcp tgsppspypy psypsptpss tnkdplltqt
121 vqkeidkgvn nepksanipr lcplqavrgg efgparvpvp fslsdlkqik idlgkfsdnp
181 dgyidvlqgl gqsfdltwr d imllnqtl t pnersaavta arefgdlwyl sqannrmtte
241 erttptgqqa vpsvdphwdt esehgdwch hlltcvlegl rktrkkpmny smmstitqgk
                                     |
                               Potential nuclear localisation sequence

301 eenltafldr lrealrkhts lspdsiegql ilkdkfitqs aadirknfks lp
               |
        Epitope recognised by antibody

```

Figure 5.15 Comparison of nuclear localisation sequences of SV40, MoMuLV gag, MSRV gag, nucleoplasmin and P53 (*adapted from Nash et al., 1993*)

SV40 consensus	PKKKRKV	<u>Key</u>
P53 (NLS1)	PQPKKKD	B – basic amino acid e.g. Lysine (K), Arginine (R), Histidine (H)
Bipartite consensus	BB X ₁₀ BXXB	X – any amino acid
MuLV (alt. 1)	RR RTEDEQKEKE RDRRR	
MuLV (alt. 2)	R RR TEDEQKEKER DRRR	
MSRV gag (alt. 1)	HK HLLTCVLEGL RKTRK K	
MSRV gag (alt. 2)	H KH LLTCVLEGLR KTRKK	
Nucleoplasmin	KR PAATKKAGQA KKKKL	
P53	KR ALPNNTSSSP QPKKK D	

Nuclear localisation sequences fall into 2 categories, with either homology to SV40, or having a bipartite consensus (BBX₁₀BXXB). The predicted nuclear localisation sequence for MSRV gag falls into the latter category, as do MuLV, nucleoplasmin, and the bipartite P53 sequence.

In order to determine whether the putative MSRV gag nuclear localisation sequence is functional, it may be worthwhile constructing a biotin-labelled peptide (and a number of mutants with disruption at key residues) of the putative localisation sequences, and loading these onto a carrier protein such as Transportan-10 and carrying out fluorescence microscopy to determine the localisation of the construct in the cells, or by assaying nuclear extracts for the peptide and mutant constructs. Transportan-10 is a cell penetrating peptide, a chimeric molecule derived from the N-terminal of galanin and the complete sequence of the wasp venom, mastoparan, and enters the cells by a mechanism involving the transcellular pathway rather than endocytosis (Lindgren et al., 2004). Once internalised, the disulphide bridge that binds the cargo is disrupted in the cellular environment.

Alternatively it may be possible that the antibodies are not recognising HERV-W gag, and are recognising an alternate, unidentified epitope. It is not clear from immunocytochemical analysis what protein is being expressed, merely that an antibody recognises a protein target. Had time allowed, it would have been useful to perform western blotting on either whole cell or nuclear extracts in order to determine whether the protein detected is MSRVgag, and if so whether the sequence is truncated as an approximation of molecular weight can be determined. If the protein can be isolated, it is possible to determine the amino acid sequence by the sequencing utilising the Edman degradation reaction or by cleaving into smaller peptides and performing mass spectroscopy (Steen and Mann, 2004). The expected size of a HERV-W gag product is highly variable due to large numbers of premature stop codons present in the open reading frames of such sequences, though truncated forms of 45 kDa have been detected in the placenta (Voisset et al., 2000). Gammaretroviruses, such as HERV-W contain *PROT* and *POL* genes in the same reading frame and polyproteins are produced by translational readthrough of a stop codon at the end of *gag* (Goff, 1996). Consequently a gag-prot-pol polyprotein is also a possibility.

It is worth noting that the PCR data presented in Chapter 4 do not detect *MSRVgag* mRNA transcripts, but a variant with homologies between ~75-85%, and the putative localisation sequence is likely to be heavily mutated (and as consequence functionality is likely to be altered due to altered protein sequence). The antibodies were designed to have cross-homology with MBP (see Chapter 2), though the homology was weak (3 amino acids), but this is expressed by oligodendrocytes rather than astrocytes (Baumann and Pham-Dinh, 2001), and is exported out of the cell forming an integral component of the myelin sheath. Therefore, astrocytoma cell lines were selected rather than oligodendrioma cell lines, as we could be certain that as protein detected would have not been MBP, although it would be interesting to test whether MSRV gag₃₁₁ antibodies developed are reactive against MBP.

Detection of the MSRV gag₃₁₁ epitope varied amongst the cell lines tested, with stronger expression present in the BeWo cell, the P53 mutant short-term grade IV astrocytoma cultures IN1582 and IN1612, and the grade IV glioblastoma multiforme cell line U251-MG. Higher expression of MSRV gag₃₁₁ would be expected in the BeWo cell line, derived from a placental choriocarcinoma, as a wide variety of retroviruses are expressed in the placenta, including HERV-W env (7q)/syncytin and a truncated HERV-W gag protein of 363 aas in length (Voisset et al., 2000), although no evidence suggests a replication-competent HERV-W provirus.

All cell lines tested were positive for *MSRV gag* transcripts (Chapter 4), and this was also the case for MSRV gag₃₁₁ expression. As previously stated, the astrocytoma cell lines with the most marked expression were grade IV tumours with mutant *P53*. Such high grade tumours have a number of other genetic alterations, including loss of heterozygosity at 9p, 10q, 13q, 17p and 22q, and a gain of chromosome 7 (Rasheed et al., 1999). These changes become more frequent with higher grade of tumour, in particular the loss of chromosome 10 loci, whilst the frequency of *P53* mutation is constant in tumours categories from grade II upwards.

A variety of genes are amplified on chromosome 7, including CDK 6 (CDK 6) (Costello et al., 1997), which is located at 7q21 (Johnson et al., 1995). CDK6 acts by phosphorylating the retinoblastoma (Rb) proto-oncogene, and is expressed in the early G1 phase of the cell cycle (Harbor et al., 1999). Interestingly, chromosome 7q21-22 is the location of the HERV-W provirus encoding syncytin, and present in tandem with a HERV-H (RGH-2) provirus (Perron et al., 2000). Costello et al. (1997) suggested that the amplification was specific to CDK6, as the expressed sequence tags (ESTs) nearby were not amplified.

Clearly the molecular pathogenesis of high grade tumours is complex, and a wide variety of genes have their expression upregulated or downregulated, and thus their translation into functional gene products. Other copies of HERV-W proviruses exist on a number of chromosomes, with at least 70 different *HERV-W gag* sequences present in the human chromosome (Voisset et al., 2000). It is more than likely that some of these sequences are active in normal tissue, and that regulation of transcription alters in cancerous tissues. Exactly which sequences and their importance to the pathogenesis of tumours (and indeed MS) is not fully understood, and whether expression is just a consequence of tumourigenesis.

6 Summary and conclusions

6.1 Prediction and screening of epitopes in MS patient plasma

A novel approach has been applied in order to predict potential epitopes where molecular mimicry between HERV proteins and candidate autoantigens in MS. A number of studies have investigated whether antibodies in MS are targeted against molecular mimics. Myelin components such as PLP have short stretches of sequence homology to proteins produced by a number of viruses (Shaw et al., 1986; Wuchterpfennig & Strominger, 1995). In the course of this project, a number of putative autoantigens were shown to have sequence homology to a variety of HERV products.

The first step of the bioinformatics analysis was to identify immunodominant domains in both the autoantigens and the HERVs as potential epitopes. By looking at the primary and secondary structures of the proteins, a large number of candidate epitopes were predicted. Primary sequence analysis looked for groups of residues that are likely to be found on the surface of a protein, based on hydrophilicity (hydrophobic residues form the core of most globular proteins), polarity (polar residues are important in interactions with other molecules), and accessibility and flexibility. Secondary structure analysis utilised web servers 'trained' to compare protein structures to that of known sequences to look for alpha helices and beta turns. These "neural networks" learn the relationships between input and output patterns and evolve. The process can only determine linear epitopes, as no analysis of tertiary structure was carried out.

The next stage of the process involved investigating whether there was any homology between identified HERV epitopes and their autoantigen counterparts, by aligning the large number of sequences obtained using the BLOSUM substitution matrix. Sequences with identity with at least 4 amino acids out of a stretch of 5 were taken to be significant.

A large number of epitopes from the candidate viruses were identified. Two were selected to be produced as peptides, one a HERV-W env peptide with sequence homology to NoGo A (a neurite growth inhibitor) (Chen et al., 2001), and the other a MSRV gag peptide with sequence homology to S-arrestin, a retinal antigen important in the pathogenesis of uveitis, but also detected in MS patients (Ohguro et al., 1993), and screened in a small scale pilot study utilising an ELISA system in order to establish suitability for antibody production and/or a full scale study of MS patients and controls. It is important to note that evidence of *in silico* homology is not evidence that molecular mimicry is taking place.

Screening of the samples in the ELISA system indicated that there was no difference in IgG response to either peptide, HERV-W env₄₁₂ or MSRV gag₂₇₄, between MS patients, OND controls and NHDs, and additionally the levels of response was low. Other studies have indicated that NoGo A is produced as a response to CNS injury, and not just MS, with significant IgM responses, but not IgG (Reindl et al., 2003). NoGo A has been reported to be upregulated in oligodendrocytes surrounding active MS lesions (Sato et al., 2005). Antibody titres to S-arrestin had been shown to be raised in MS patients in a number of studies, particularly where patients had reduced visual acuity (Ohguro et al., 1993; Gorczyca et al., 2004). It may be worth testing the MSRV gag₂₇₄ peptide in a particular subset of patients with ocular symptoms.

The results do not prove that antibodies against HERV epitopes do not cross-react with those of putative autoantigens, as a large number of potential matches were found, but it is extremely unlikely that either of the two peptides produced do. The sensitivity of the assay was poor, and optimisation of the assay may have yielded higher titres of antibody, making discrimination of the samples more difficult due to the low signal to noise ratio. Lack of reactivity could be due to the peptide binding poorly to the plate, or a genuine lack of antibody in patient plasma.

Consequently no suitable candidate was identified for antibody production in this study. Polyclonal antibodies produced could be used to provide a calibration for the ELISA system, enabling effective troubleshooting of the system. They would also provide a tool to investigate patterns of expression in different cell types in pathological specimens of MS tissue, and matched to patients and controls, and whether the expression is more marked in active lesions. It may have been worthwhile producing recombinant proteins of HERV-W to test for reactivity in patient plasma (or serum). This method has been successfully used to detect HERV K10 Gag polyprotein in the serum of patients with testicular tumours and a number of teratocarcinoma-derived cell lines (Sauter et al., 1995). The production of a recombinant protein would also provide a useful tool to screen whether the epitopes that were predicted were reflected in the protein, or whether they are obscured.

6.2 Detection of HERV-W RNA transcripts in glioma cell lines

The main objective of this study was to produce tools to enable detection of HERV products in the central nervous system, both at mRNA level and at protein level. A number of primers were developed to enable the detection and quantification of mRNA in a number of cancer cell lines.

As HERV-W *env* expression is ubiquitous in both cancerous and healthy tissues (Yi et al., 2004), a multiplex PCR system was utilised to investigate the amount of HERV-W *env* in glioma cell lines in relation to *histidyl tRNA synthetase*, which was used as a housekeeping gene. HERV-W gene products have been detected in a variety of tumour types including ovarian tumours (Menendez et al., 2004) and endometrial carcinomas (Strick et al., 2007). A number of studies have shown that HERV-W retains transcriptional activity in the normal brain by a variety of methods including EST analysis (Nakamura et al., 2003), digital expression profiling (Stauffer et al., 2004) and real-time PCR (Forsman et al., 2005). A number of studies have also detected HERV-W expression in a small sample of brain tumour cell lines (Yi et al., 2002; Yi et al., 2004).

A number of glioblastoma and astrocytoma cells were tested in this study to investigate the expression of HERV-W *env* and MSRV *gag* at RNA level, and were largely representative of higher grade tumours (WHO classification III and IV). HERV-W *env* RNA was detected in every cell line tested, including a placental control line and cell lines derived from teratocarcinomas, breast and bowel cancer. Semi-quantitation of HERV-W *env* expression was determined, with levels in the BeWo line significantly higher than that of other cell lines, as in the placenta HERV-W *env* is known to play a key role in syncytialisation and trophoblast development (Mi et al., 2000). PCR products had a high sequence homology (>94%) to the reference HERV-W *env* sequence that encodes syncytin. The high degree of sequence homology to syncytin indicates that translation of the transcript is likely to produce a functional protein.

In the astrocytoma cell lines, expression was significantly higher in the cell lines with mutant P53 than those with wild type protein. Wild type P53 has multiple roles in DNA repair, cell cycle arrest and the initiation of apoptosis, and mutation results in breakdown of cell control. Initiation of HERV-W *env* transcription and the presence of such sequences may represent a breakdown of such control. P53 negatively regulates LTRs of another gammaretrovirus, HERV-I (Chang et al., 2007), and conformational mutation of the protein stimulates transcriptional activity. LTRs are responsible for mediating integration of the exogenous retrovirus into the chromosome.

Whether HERV-W *env* is a cause or a consequence of the genomic instability associated with glioblastomas has yet to be established. Further studies may elucidate whether HERV-W *env* expression is correlated with glioma cell progression. By looking at expression levels in lower grade tumours (anaplastic astrocytomas) and normal astrocytes, a clearer picture of the involvement of HERV-W *env* in the disease may be ascertained. *In situ* PCR or *in situ* hybridisation (using probes developed from the sequence) of pathological specimens of astrocytomas (and also MS) will further indicate the role of MSRV *env* in the disease by giving information about the localisation of the sequences in the central nervous system.

MSRV *gag* variants were also detected in all of the cell lines tested, but as in the case of *pol* (Yi et al., 2004), a large number of frameshifts and point mutations are present, indicating a probable lack of any functional protein. The presence of these transcripts is interesting, as HERV-W *gag* appears to have a limited distribution in cancer cells and healthy tissues (Yi et al., 2002), in contrast to HERV-W *env* which is widely expressed. Sequences of the products had only moderate homology to the prototype MSRV *gag* and to each other, with the sample from the CHME3 foetal microglia cell showing the highest degree of similarity. Analysis of the sequences of the PCR products in a wide variety of tissues and further cell lines could be utilised in order to determine whether particular isolates are associated with particular tissues and tumours.

The effects of IFN- γ on MSRV *gag* transcription levels in the foetal microglial cell line, CHME3, was also determined *in vitro*, as IFN- γ is one of the central cytokines involved in MS, and is involved in a variety of functions including anti-microbial and anti-tumour mechanisms, and in the upregulation of MHC class I and II. A number of other retroviruses and retrotransposons are located in MHC regions, and it might be expected that these sequences are upregulated along with MHC expression. IFN- γ and TNF- α have been shown to upregulate *pol* expression (Serra et al., 2003). No significant evidence of MSRV *gag* upregulation was found, indicating that any increase in expression present in MS is due to a different mechanism. Such mechanisms may include other cytokines, steroid hormones, exogenous viruses (such as members of the herpes family), UV irradiation or demethylating agents. The PCR developed in this thesis provides a useful tool for investigating these possibilities.

In addition to HERV-W and MSRV, a number of other HERVs have been detected in the CNS, and are transcriptionally active, including including HERV-H and HERV-I (Forsman et al., 2005). Notably, the RGH-2 subtype of HERV-H has also been linked with MS (Christensen et al., 1997), and an RGH variant is present in tandem with the HERV-W provirus located at 7q (Perron et al., 2000). It would be interesting to see if HERV-H has a similar expression pattern to that of HERV-W by synthesising primers to use in a similar PCR system to the one detailed here.

6.3 Detection of HERV-W protein products in glioma cell lines

Evidence of mRNA expression does not necessarily mean that functional protein is present. The vast majority of HERVs encode defective proteins, with a large amount of frameshift mutations and premature stop codons present (De Parseval et al., 2003). HERVs are estimated to form 5% of the human genome, and a large amount is untranscribed, and in the HERV-W family alone at least 70 *GAG*, 100 *PRO* and 30 *ENV* sequences are present.

A polyclonal antibody was made against an epitope of MSRV gag, predicted within our research group by Denise Roden, by immunising a short peptide sequence into a rabbit which was then purified using affinity chromatography. The purified antibody was applied to astrocytoma cell lines including those investigated in the MSRV gag PCR. The MSRV gag₃₁₁ epitope was identified in all cell lines and was localised as small punctate foci in the nucleus of the cells. Strong expression of the epitope was detected in the BeWo choriocarcinoma cell line along with three grade IV p53 mutant astrocytoma/glioblastoma cell lines. Expression was less marked in the grade III glioblastoma cell lines and the two p53 wild type astrocytoma cell lines. Expression is not likely to be purely due to the P53 status of the cell lines, but more complex involving a number of other factors. Higher grade tumours have a number of genetic alterations, including loss of heterozygosity at a number of chromosomal loci, and gains at others.

A number of genes are amplified on chromosome 7, including cyclin-dependant kinase 6 (CDK 6) (Costello et al., 1997), which is located at 7q21, the location of the HERV-W provirus which encodes syncytin, and present in tandem with a HERV-H provirus (Perron et al., 2000). CDK 6 is expressed in the G1 phase of the cell cycle and is involved in the phosphorylation of the retinoblastoma (Rb) oncogene.

It is not clear why MSRV gag should localise in the nucleus, as functional gag is important in retroviral assembly, and is usually detected at the cytoplasm or at the cytoplasmic membrane of the cell. Intranuclear localisation of a number of retroviruses has been described – in MoMuLV (Nash et al., 1993), foamy viruses (Chang et al., 1995), HIV Tat and Rev (Truant & Cullen, 1999), and HERV-K (HTDV) (Magin-Lachman et al., 2001), but not as yet in HERV-W.

Analysis of the MSRV gag protein sequence identified a putative bipartite nuclear localisation sequence similar to that described in MoMuLV (Nash et al., 1993). Similar amino acids exist in a number of proteins which are imported into the nucleus including nucleoplasmin and P53. To establish that the nuclear localisation sequence is active *in vitro* it is necessary to determine interactions with nuclear import molecules such as importin α , and to determine whether the sequence is localised into the nucleus.

Exact identification of the MSRV gag₃₁₁ epitope is important as it is not clear exactly what the product is, whether it is MSRV gag (a variant), MSRV gag polyprotein or a different molecule entirely. Western blotting would assist in narrowing down the range of possibilities (and is able to distinguish partial, complete or a polyprotein), but gives only a molecular weight.

Immunocytochemistry is not quantitative. In order to quantify the data other studies would need to be carried out, such as dot-blotting or western blotting nuclear extracts of the cells, and relating the quantity against that present in the cytoplasm or the total cell. This would also be useful for investigating how HERV protein expression is altered in response to exogenous factors such as cytokines.

6.4 Further applications

The primers and antibodies developed can be used in a wide variety of applications where HERV-W involvement is expected, such as MS. The system developed in the process of this project can provide a rapid way of determining levels of other retroviral products, namely HERV-W *env* and MSRV *gag* in these tissues. It is also possible to utilise the designed primers for use with *in situ* PCR to determine localisation of sequences (DNA or RNA) within pathological sections of MS tissue. Sequencing of the products allows the development of probes suitable for use with *in situ* hybridisation studies as an alternate method.

The system can also be used to look at the modulation of HERV-W/MSRV transcription by other factors involved in MS pathogenesis, e.g. cytokines, exogenous agents. The HERV-W env (syncytin) promoter has been shown to be regulated by proinflammatory cytokines (TNF- α , IFN- γ , IL-6, IL-1) and suppressed by IFN- β (Mameli et al, 2007). It would be worthwhile investigating the promoter activity of the syncytin promoter in the different cell lines tested, and establishing whether promoter activity is suppressed in those cell lines with wild type P53.

The epitopes predicted can be used to develop peptides suitable for use in MS patient studies, once the target of antibody has been identified, such as an ELISA system to detect autoantibodies present in patient plasma or serum, and CSF. Other studies have shown cross-reactivity of random phage peptides with cross-homology against MSRV (Jolivet-Reynaud et al., 1999), but no demonstration of molecular mimicry between HERVs and putative autoantigens has been identified. There is very little evidence produced to support the presence of anti HERV-W antibody in the plasma or serum of MS patients, though responses against synthetic peptides of HERV-H (RGH-2) have been characterised (Christensen et al., 2000).

The presence of HERV-W/MSRV gag products has been demonstrated by Perron et al., (2005) in brains of MS patients and controls. In MS patients env antigen was overexpressed only in recent demyelinating regions, where gag accumulation was higher in demyelinated white matter. In controls env was expressed widely within white matter, whilst gag was expressed widely in neurons. It is worth detecting using the antibody developed against MSRV gag displays the predicted epitope in a similar expression pattern in MS patients. Development of antibodies against HERV epitopes predicted *in silico* may also prove a useful tool in such investigations.

The study also raises questions about the role of HERV-W gene and protein products in patients with astrocytic tumours, particularly the link with P53 mutation and HERV-W env expression. Further study at protein level will establish whether HERV-W env is a fully functional product, though the extremely close homology of the HERV-W env isolates to syncytin suggests that it is more than likely. The other major question to be resolved is how p53 may affect the expression of the provirus. Genomic instability associated with high grade tumours results in a number of alterations in the cell, such as loss and gain at a number of chromosomal loci, a number of which contain HERV elements. Investigation of the activity of the HERV-W ENV promoter in cell lines with wild type and mutant P53 protein could explain how this situation arises. This would help to establish whether HERV-W env products are a reliable marker of the disease.

6.5 Summary

Tools for the detection of HERVs at mRNA and protein level were successfully developed and applied in the investigation of the detection of HERV-W/MSRV gag products a number of cell lines, predominately astrocytomas. HERV-W *env* RNA was shown to be present in all lines, but levels were particularly high in the BeWo placental choriocarcinoma cell line, while in astrocytoma cell lines a relationship was found with P53 status. Exactly how P53 status alters HERV-W activity is unclear, but a number of HERV LTRs are active in the genome, including that of HERV-W. MSRV *gag* variants are also widely expressed in astrocytomas, but appear to be heavily mutated and unlikely to code for functional protein. IFN- γ had no significant effect on the levels of MSRV *gag* expressed in the CHME3 microglial cell line.

Immunocytochemical analysis indicated the presence of a protein displaying the MSRV gag₃₁₁ epitope, though this was located in the nucleus rather than in the cytoplasm or the cytoplasmic membrane as expected. Expression of this protein was marked in a number of astrocytoma cell lines and short-term cultures derived from high grade tumours.

Attempts to show molecular mimicry between HERV epitopes and those of autoantigens yielded little, with no significant differences in antibody reactivity between MS patients, OND controls, and NHDs in HERV-W *env* and MSRV gag peptides constructed on the basis of bioinformatics analysis and similarities with epitopes of known autoantigens.

References

- Adelman, M. K. and J. J. Marchalonis (2002). Endogenous retroviruses in systemic lupus erythematosus: candidate lupus viruses. Clin Immunol **102**(2): 107-16.
- Agius, M. A., Kirvan, C. A., Schafer A. L., Gudipati E. and S. Zhu (1999). High prevalence of anti-alpha-crystallin antibodies in multiple sclerosis: correlation with severity and activity of disease. Acta Neurol Scand **100**(3): 139-47.
- Albrechtsen, N., Dornreiter, I., Grosse, F., Kim, E., Wiesmüller, L. and W. Deppert (1999). Maintenance of genomic integrity by p53: complementary roles for activated and non-activated p53. Oncogene **18**(53): 7706-17.
- Alonso A., Jick S. S., Olek M.J. and M. A. (2007). Incidence of multiple sclerosis in the United Kingdom : findings from a population-based cohort. J Neurol **254**(12): 1736-41.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang Z., W. Miller, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**(17): 3389-402.
- Andersson, G., Svensson, A. C., Setterblad, N. and L. Rask (1998). Retroelements in the human MHC class II region. Trends Genet **14**(3): 109-14.
- Antony, J. M., Ellestad, K. K., Hammond, R., Imaizumi, K., Mallet, F., K. G. Warren, et al. (2007). The human endogenous retrovirus envelope glycoprotein, syncytin-1, regulates neuroinflammation and its receptor expression in multiple sclerosis: a role for endoplasmic reticulum chaperones in astrocytes. J Immunol **179**(2): 1210-24.
- Antony, J. M., Izad, M., Bar-Or, A., Warren, K.G., Vojdani, M., F. Mallet, et al. (2006). Quantitative Analysis of Human Endogenous Retrovirus-W env in Neuroinflammatory Diseases. AIDS Res Hum Retroviruses **22**(12): 1253-9.
- Antony, J. M., van Marle, G., Opii, W., Butterfield, D.A., Mallet, F., V. W. Yong, et al. (2004). Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination. Nat Neurosci **7**(10): 1088-95.

Antony J. M., Zhu Y., Izad M., Warren K. G., Vodjgani M., Mallet F. and C. Power (2007). Comparative expression of human endogenous retrovirus-W genes in multiple sclerosis. AIDS Res Hum Retroviruses **23**(10):1251-6

Banga J. P., Kasp E., Ellis B. A., Brown E., Suleyman S. and D. C. Dumonde (1987). Antigenicity and uveitogenicity of partially purified peptides of a retinal autoantigen, S-antigen. Immunology **61**(3): 357-62.

Baranzini, S. E., Jeong M.C, Butunoi, C., Murray, R. S., Bernard, C. C. and J. R. Oksenberg (1999). B cell repertoire diversity and clonal expansion in multiple sclerosis brain lesions. J Immunol **163**(9): 5133-44.

Basu, A. and S. Haldar (1998). The relationship between Bcl2, Bax and p53: consequences for cell cycle progression and cell death. Mol Hum Reprod **4**(12): 1099-109.

Baumann N. and Pham-Dinh (2001). Biology of oligodendrocyte and myelin in the mammalian central nervous system. Physiol Rev **81**: 871-927.

Bera, T. K., Tsukamoto, T., Panda, D. K., Huang, T., Guzman, R. C., S. I. Hwang, et al. (1998). Defective retrovirus insertion activates c-Ha-ras protooncogene in an MNU-induced rat mammary carcinoma. Biochem Biophys Res Commun **248**(3): 835-40.

Bernard, C. C., Randell, V. B., Horvath, L. B., Carnegie, P. R. and Mackay, I. R (1981). Antibody to myelin basic protein in extracts of multiple sclerosis brain. Immunology **43**(3): 447-57.

Bhaskaran, R. and P. K. Ponnuswamy (1984). Dynamics of amino acid residues in globular proteins. Int J Pept Protein Res **24**(2): 180-91.

Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., M. F. Wolfson, et al. (1997). A metalloproteinase disintegrin that releases tumour-necrosis factor- α from cells. Nature **385**(6618): 729-33.

Blond, J. L., Besème, F., Duret, L., Bouton, O., Bedin, F., H. Perron, et al. (1999). Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. J Virol **73**(2): 1175-85.

Blond, J. L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., S. Chapel-Fernandes, et al. (2000). An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J Virol **74**(7): 3321-9.

Boes, M. (2000). Role of natural and immune IgM antibodies in immune responses. Mol Immunol **37**(18): 1141-9.

Boller, K., Frank, H., Löwer, J., Löwer, R. and R. Kurth (1983). Structural organization of unique retrovirus-like particles budding from human teratocarcinoma cell lines. J Gen Virol **64** (Pt 12): 2549-59.

Breij E C, Brink B P, Veerhuis R, van den Berg C, Vloet R, R. Yan et al. (2008). Homogeneity of active demyelinating lesions in established multiple sclerosis. Ann Neurol **63**(1):16-25.

Bronstein, J. M., Lallone, R. L., Seitz, R. S., Ellison, G. W. and L. W. Myers (1999). A humoral response to oligodendrocyte-specific protein in MS: a potential molecular mimic. Neurology **53**(1): 154-61.

Brudek, T., Christensen, T., Hansen, H. J., Bobecka, J. and A. Møller-Larsen (2004). Simultaneous presence of endogenous retrovirus and herpes virus antigens has profound effect on cell-mediated immune responses: implications for multiple sclerosis. AIDS Res Hum Retroviruses **20**(4): 415-23.

Brudek, T., Lühdorf, P, Christensen, T., Hansen, H. J. and A. Møller-Larsen (2007). Activation of endogenous retrovirus reverse transcriptase in multiple sclerosis patient lymphocytes by inactivated HSV-1, HHV-6 and VZV. J Neuroimmunol **187**(1-2): 147-55.

Burns, J., Rosenweig, A., Zweiman, B. and R. P. Lisak (1983). Isolation of myelin basic protein-reactive T-cell lines from normal human blood. Cell Immunol **81**: 435-40.

- Canman, C. E. and M. B. Kastan (1997). Role of p53 in apoptosis. Adv Pharmacol **41**: 429-60.
- Cermelli, C. and S. Jacobson (2000). Viruses and multiple sclerosis. Viral Immunol **13**(3): 255-67.
- Chang, J., Lee, K. J., Jang, K. L., Lee, E. K., Baek, G. H. and Y. C. Sung (1995). Human foamy virus Bel1 transactivator contains a bipartite nuclear localization determinant which is sensitive to protein context and triple multimerization domains. J Virol **69**(2): 801-8.
- Chang, N. T., Yang, W. K., Huang, H. C., Yeh, K. W. and C. W. Wu (2007). The transcriptional activity of HERV-I LTR is negatively regulated by its cis-elements and wild type p53 tumor suppressor protein. J Biomed Sci **14**(2): 211-222.
- Charil A., Zijdenbos A. P., Taylor J., Boelman C., Worsley K. J., A. C. Evans et al. (2003). Statistical mapping analysis of lesion location and neurological disability in multiple sclerosis: application to 452 patient data sets. Neuroimage **19**(3): 532-44.
- Chen, L., Mao, S. J. and W. J. Larsen (1992). Identification of a factor in fetal bovine serum that stabilizes the cumulus extracellular matrix. A role for a member of the inter-alpha-trypsin inhibitor family. J Biol Chem **267**(17): 12380-6.
- Chen, M. S., Huber, A. B., van der Haar, M. E., Frank, M., Schnell, L., Spillman, A. A., et al. (2000). Nogo-A is a myelin-associated neurite outgrowth inhibitor and an antigen for monoclonal antibody IN-1. Nature **403**(6768): 434-9.
- Chou, P. Y. and G. D. Fasman (1974). Prediction of protein conformation. Biochemistry **13**(2): 222-45.
- Chou, Y. K., Burrows, G. G., LaTocha, D., Wang, C., Subramanian, S., D. N. Bourdette, et al. (2004). CD4 T-cell epitopes of human alpha B-crystallin. J Neurosci Res **75**(4): 516-23.
- Christensen, T., Dissing Sørensen, P., Riemann, H., Hansen H. J., Munch, M., S Haahr, et al. (2000). Molecular characterization of HERV-H variants associated with multiple sclerosis. Acta Neurol Scand **101**(4): 229-38.

Christensen, T., Jensen, A. W., Munch, M., Haahr, S., Sørensen, P. D., H. Riemann, et al. (1997). Characterization of retroviruses from patients with multiple sclerosis. Acta Neurol Scand Suppl **169**: 49-58.

Christensen, T., Sørensen, P. D., Hansen, H. J. and A. Møller-Larsen (2003). Antibodies against a human endogenous retrovirus and the preponderance of env splice variants in multiple sclerosis patients. Mult Scler **9**(1): 6-15.

Cianciolo, G. J., Copeland, T. D., Oroszlan, S. and R. Snyderman (1985). Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. Science **230**(4724): 453-5.

Claverlie, J and C. Notredame (2003). Bioinformatics for Dummies. New York: Wiley.

Collins, V. P. (2004). Brain tumours: classification and genes. J Neurol Neurosurg Psychiatry **75** Suppl 2: ii2-11.

Colombo, E., Banki, K., Tatum, A. H., Daucher, J., Ferrante, P., R. S. Murray, et al. (1997). Comparative analysis of antibody and cell-mediated autoimmunity to transaldolase and myelin basic protein in patients with multiple sclerosis. J Clin Invest **99**(6): 1238-50.

Colombo, M., Dono, M., Gazzola, P., Roncella, S., Valetto, A., N. Chiorazzi, et al. (2000). Accumulation of clonally related B lymphocytes in the cerebrospinal fluid of multiple sclerosis patients. J Immunol **164**(5): 2782-9.

Conrad, B., Weidmann, E., Trucco, G., Rudert, W. A., Behboo, R., C. Ricordi, et al. (1994). Evidence for superantigen involvement in insulin-dependent diabetes mellitus aetiology. Nature **371**(6495): 351-5.

Conrad, B., Weissmahr, R. N., Böni, J., Arcari, R., Schüpbach, J. and B. Mach (1997). A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. Cell **90**(2): 303-13.

Cook, S. D. (1996). The epidemiology of multiple sclerosis: Clues to the etiology of a mysterious disease. Neuroscientist **2**(3): 172-180.

Cossins, J. A., Clements, J. M., Ford, J., Miller, K. M., Pigott, R., W. Vos, et al. (1997). Enhanced expression of MMP-7 and MMP-9 in demyelinating multiple sclerosis lesions. Acta Neuropathol (Berl) **94**(6): 590-8.

Costello, J. F., Plass, C., Arap, W., Chapman, V. M., Held, W. A., M. S. Berger, et al. (1997). Cyclin-dependent kinase 6 (CDK6) amplification in human gliomas identified using two-dimensional separation of genomic DNA. Cancer Res **57**(7): 1250-4.

Dang, C. V. and W. M. Lee (1989). Nuclear and nucleolar targeting sequences of c-erb-A, c-myc, N-myc, p53, HSP70, and HIV tat proteins. J Biol Chem **264**(30): 18019-23.

Dayhoff, M. O., Schwartz, R. M. and B. C. Orcutt. (1978). A model of evolutionary change in proteins. pp. 345-352. In: Atlas of Protein Sequence and Structure Vol. 5, suppl. 3. Ed. M. O. Dayhoff. Washington DC : National Biomedical Research Foundation

de Parseval, N., Casella, J., Gressin, L. and T. Heidmann (2001). Characterization of the three HERV-H proviruses with an open envelope reading frame encompassing the immunosuppressive domain and evolutionary history in primates. Virology **279**(2): 558-69.

de Parseval, N., Lazar, V., Casella, J. F., Benit, L. and T. Heidmann (2003). Survey of human genes of retroviral origin: identification and transcriptome of the genes with coding capacity for complete envelope proteins. J Virol **77**(19): 10414-22.

Deb-Rinker, P., Klempan, T. A., O'Reilly, R. L., Torrey, E. F. and S. M. Singh (1999). Molecular characterization of a MSRV-like sequence identified by RDA from monozygotic twin pairs discordant for schizophrenia. Genomics **61**(2): 133-44.

Dolei, A., Serra, C., Mameli, G., Pugliatti, M., Sechi, G., M. C. Ciroto, et al. (2002). Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. Neurology **58**(3): 471-3.

Donoso, L. A., Merryman, C. F., Shinohara, T., Sery, T. W. and A. Smith (1987). S-antigen. Experimental autoimmune uveitis following immunization with a small synthetic peptide. Arch Ophthalmol **105**(6): 838-40.

Ehlhardt S., Seifert M., Schneider J., Ojak A., Zang K. D., and Y. Mehraein (2006). Human endogenous retrovirus HERV-K(HML-2) Rec expression and transcriptional activities in normal and rheumatoid arthritis synovia. J Rheumatol **33**(1):16-23.

Ejtehadi H. D., Freimanis G. L., Ali H. A., Bowman S., Alavi A., Axford J., R. Callaghan et al. (2006). The potential role of human endogenous retrovirus K10 in the pathogenesis of rheumatoid arthritis: a preliminary study. Ann Rheum Dis **65**(5): 612-6.

Ejtehadi H. D., Martin J. H., Junying, J., Roden, D. A., Lahiri, M., Warren, P, et al. (2005). A novel multiplex PCR system detects human endogenous retrovirus-K in breast cancer. Arch Virol **150**(1): 177-184.

Ekstrand, A. J., James, C. D., Cavenee, W. K., Seliger, B., Pettersson, R. F. and V. P. Collins (1991). Genes for epidermal growth factor receptor, transforming growth factor alpha, and epidermal growth factor and their expression in human gliomas in vivo. Cancer Res **51**(8): 2164-72.

Esiri, M. M. (1977). Immunoglobulin-containing cells in multiple-sclerosis plaques. Lancet **2**(8036): 478.

Firouzi, R., Rolland, A., Michel, M., Jouvin-Marche, E., Hauw, J. J., C. Malcus-Vocanson, et al. (2003). Multiple sclerosis-associated retrovirus particles cause T lymphocyte-dependent death with brain hemorrhage in humanized SCID mice model. J Neurovirol **9**(1): 79-93.

Forooghian F., Cheung R. K., Smith W. C., O'Connor P. and H. M. Dosch (2007). Enolase and arrestin are novel nonmyelin autoantigens in multiple sclerosis. J Clin Immunol. **27**(4):388-96.

Forsman, A., Yun, Z., Hu, L., Uzhameckis, D., Jern, P. and J. Blomberg (2005). Development of broadly targeted human endogenous gammaretroviral pol-based real time PCRs Quantitation of RNA expression in human tissues. J Virol Methods **129**(1): 16-30.

Garnier, J., Osguthorpe, D. J. and B. Robson (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol **120**(1): 97-120.

Garson, J. A., Tuke, P. W., Giraud, P., Paranhos-Baccala, G. and H. Perron (1998). Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. Lancet **351**(9095): 33.

Gaudin P., Ijaz S., Tuke P. W., Marcel F., Paraz A., J. M. Seigneurin et al. (2000). Infrequency of detection of particle-associated MSRV/HERV-W RNA in the synovial fluid of patients with rheumatoid arthritis. Rheumatology **39**(9):950-4.

George, M. L., Eccles, S. A., Tutton, M. G., Abulafi, A. M. and R. L. Swift (2000). Correlation of plasma and serum vascular endothelial growth factor levels with platelet count in colorectal cancer: clinical evidence of platelet scavenging? Clin Cancer Res **6** (8): 3147-3152.

Goff, S. P. (1996) Chapter 7: Retroviruses. pp. 1871-1939. In: Fields Virology (3rd edition). Eds in Chief: Fields B.N., Knipe D.M. David M. Knipe. Philadelphia : Lippincott-Raven

Gorczyca, W. A., Ejma, M., Witkowska, D., Misiuk-Hojlo, M., Kuropatwa, M., M. Mulak, et al. (2004). Retinal antigens are recognized by antibodies present in sera of patients with multiple sclerosis. Ophthalmic Res **36**(2): 120-3.

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. Science **185**(4154): 862-4.

GrandPré, T., Nakamura, F., Vartanian, T. and S. M. Strittmatter (2000). Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein. Nature **403**(6768): 439-44.

Haraguchi, S., Good, R. A., James-Yarish, M., Cianciolo, G. J. and N. K. Day (1995). Induction of intracellular cAMP by a synthetic retroviral envelope peptide: a possible mechanism of immunopathogenesis in retroviral infections. Proc Natl Acad Sci U S A **92**(12): 5568-71.

Harbor, J. W., Luo, R. X., Dei Santi, A., Postigo, A. A. and D. C. Dean (1999). Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1. Cell **98**(6): 859-69.

Hawkes, C. H. (2002). Is multiple sclerosis a sexually transmitted infection? J Neurol Neurosurg Psychiatry **73**(4): 439-43.

Hayes, G. M., Woodroffe, M. N. and M. L. Cuzner (1987). Microglia are the major cell type expressing MHC class II in human white matter. J Neurol Sci **80**(1): 25-37.

Hayward, W. S., Neel, B. G. and S. M. Astrin (1981). Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature **290**(5806): 475-80.

Henderson, L. E., Krutzsch, H. C. and S. Oroszlan (1983). Myristyl amino-terminal acylation of murine retrovirus proteins: an unusual post-translational proteins modification. Proc Natl Acad Sci U S A **80**(2): 339-43.

Hohenadl, C., Germaier, H., Walchner, M., Hagenhofer, M., Herrmann, M., M. Stürzl et al. (1999). Transcriptional activation of endogenous retroviral sequences in human epidermal keratinocytes by UVB irradiation. J Invest Dermatol **113**(4): 587-94.

Hopp, T. P. and K. R. Woods (1981). Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A **78**(6): 3824-8.

Hu, L., Hornung, D., Kurek, R., Ostman, H., Blomberg, J. and A. Berggvist (2006). Expression of human endogenous gammaretroviral sequences in endometriosis and ovarian cancer. AIDS Res Hum Retroviruses **22**(6): 551-7.

Huber, A. B., Weinmann, O., Brösamle, C., Oertle, T. and M. E. Schwab (2002). Patterns of Nogo mRNA and protein expression in the developing and adult rat and after CNS lesions. J Neurosci **22**(9): 3553-67.

Inda, M. M., Fan, X., Muñoz, J., Perot, C., Fauvet, D., G. Danglot et al. (2003). Chromosomal abnormalities in human glioblastomas: gain in chromosome 7p correlating with loss in chromosome 10q. Mol Carcinog **36**(1): 6-14.

Jaeckel E., Heringlake S., Berger D., Brabant G., Hunsmann G and M. P. Manns (1999), No evidence for association between IDDMK(1,2)22, a novel isolated retrovirus, and IDDM. Diabetes **48**(1): 209-14.

- Janin, J. (1979). Surface and inside volumes in globular proteins. Nature **277**(5696): 491-2.
- Jayasena, U. L., Gribble, S. K., McKenzie, A., Beyreuther, K., Masters, C. L. and J. R. Underwood (2001). Identification of structural variations in the carboxyl terminus of Alzheimer's disease-associated beta A4[1-42] amyloid using a monoclonal antibody. *Clin Exp Immunol* **124**(2): 297-305.
- Johnson, D., Hafler, D. A., Fallis, R. J., Lees, M. B., Brady, R. O., R. H. Quarles et al. (1986). Cell-mediated immunity to myelin-associated glycoprotein, proteolipid protein, and myelin basic protein in multiple sclerosis. J Neuroimmunol **13**(1): 99-108.
- Jolivet-Reynaud, C., Perron, H., Ferrante, P., Becquart, L., Dalbon, P. and B. Mandrand (1999). Specificities of multiple sclerosis cerebrospinal fluid and serum antibodies against mimotopes. Clin Immunol **93**(3): 283-93.
- Jones D. T. (1999a). GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. J Mol Biol **287**(4): 797-815.
- Jones, D. T. (1999b). Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol **292**(2): 195-202.
- Jones, D. T. (2003). Chapter 9: Protein structure prediction, pp. 135-150. In: Bioinformatics: Genes, Proteins and Computers. Eds: Orengo C., Jones D., Thornton J. Oxford : BIOS Scientific.
- Kaplan, M. H. (1963). Immunologic relation of streptococcal and tissue antigens: properties of an antigen in certain strains of group A streptococci exhibiting an immunologic cross reaction with human heart tissue. J Immunol **90**: 595-606.
- Karlsson, H., Bachmann, S., Schröder, J., McArthur, J., Torrey, E. F. and R. H. Yolken (2001). Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. Proc Natl Acad Sci U S A **98**(8): 4634-9.
- Kaur, H. and G. P. Raghava (2003). Prediction of beta-turns in proteins from multiple alignment using neural network. Protein Science **12**: 627-634.

Kaur, H. and G. P. Raghava (2004). A neural network method for prediction of beta-turn types in proteins using evolutionary information. Bioinformatics **20**(16): 2751-8.

Kawanokuchi J., Mizuno T., Takeuchi H., Kato H., Wang J., N. Mitsuma et al. (2006). Production of interferon-gamma by microglia. Mult Scler. **12**(5): 558-64.

Kim D. E., Chivian D. and D. Baker (2004). Protein structure prediction and analysis using the Robetta server. Nucleic Acids Research **32**: W526–W531.

Kim, H. S. and T. J. Crow (2002). Identification and phylogeny of the human endogenous retrovirus HERV-W LTR family in human brain cDNA library and Xq21.3 region. Journal of Microbiology and Biotechnology **12**(3): 508-513.

Kim, H. S. and W. H. Lee (2001). Human endogenous retrovirus HERV-W family: chromosomal localization, identification, and phylogeny. AIDS Res Hum Retroviruses **17**(7): 643-8.

Kim H. S., Ahn K. and D. S. Kim (2008). Quantitative expression of the HERV-W env gene in human tissues. Arch Virol 2008;153(8):1587-91.

Kim, H. S., Takenaka, O. and T. J. Crow (1999). Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates. J Gen Virol **80**(Pt 10): 2613-9.

Kind, P., Lehmann, P. and G. Plewig (1993). Phototesting in lupus erythematosus. J Invest Dermatol **100**(1): 53S-57S.

Knobbe, C. B., Merlo, A. and G. Reifenberger (2002). Pten signaling in gliomas. Neuro Oncol **4**(3): 196-211.

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F., Ounanian-Paraz, A., Sodoyer, M., C. Ott, et al. (1999). Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles. Virology **260**(1): 1-9.

Kornek, B. and H. Lassmann (2003). Neuropathology of multiple sclerosis-new concepts. Brain Res Bull **61**(3): 321-6.

Kondo, S., Murakami, T., Tatsumi, K., Ogata, M., Kanemoto, S., K. Otori, et al. (2005). OASIS, a CREB/ATF-family member, modulates UPR signalling in astrocytes. Nat Cell Biol **7**(2): 186-94.

Kudo, Y. and C. A. Boyd (2002). Changes in expression and function of syncytin and its receptor, amino acid transport system B(0) (ASCT2), in human placental choriocarcinoma BeWo cells during syncytialization. Placenta **23**(7): 536-41.

Kurtzke, J.F. (1997). Chapter 7: The epidemiology of multiple sclerosis, pp. 91-139. In: Multiple sclerosis: Clinical and pathogenetic basis. Eds: Raine C.S., McFarland H.F., Tourtellotte W.W. London: Chapman & Hall.

Kwun, H. J., Han, H. J., Lee, W. J., Kim, H. S. and K. L. Jang (2002). Transactivation of the human endogenous retrovirus K long terminal repeat by herpes simplex virus type 1 immediate early protein 0. Virus Res **86**(1-2): 93-100.

Lam, P. Y., Di Tomaso, E., Ng, H. K., Pang, J. C., Roussel, M. F. and N. M. Hjelm (2000). Expression of p19INK4d, CDK4, CDK6 in glioblastoma multiforme. Br J Neurosurg **14**(1): 28-32.

Larsson, E. and G. Andersson (1998). Beneficial role of human endogenous retroviruses: facts and hypotheses. Scand J Immunol **48**(4): 329-38.

Laufer G, Mayer J, Mueller BF, Mueller-Lantzsch N, and Ruprecht K (2009). Analysis of transcribed human endogenous retrovirus W env loci clarifies the origin of multiple sclerosis-associated retrovirus env sequences. Retrovirology **15**;6:37.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. Cell **88**(3): 323-31.

Levitt, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. J Mol Biol **104**(1): 59-107.

- Liang, S. H. and M. F. Clarke (1999). A bipartite nuclear localization signal is required for p53 nuclear import regulated by a carboxyl-terminal domain. J Biol Chem **274**(46): 32699-703.
- Lindberg, R. L., De Groot, C. J., Montagne, L., Freitag, P., van der Valk, P., L. Kappos et al. (2001). The expression profile of matrix metalloproteinases (MMPs) and their inhibitors (TIMPs) in lesions and normal appearing white matter of multiple sclerosis. Brain **124**(Pt 9): 1743-53.
- Lindgren, M. E., Hällbrink, M. M., Elmquist, A. M. and U. Langel (2004). Passage of cell-penetrating peptides across a human epithelial cell layer in vitro. Biochem J **377**(Pt 1): 69-76.
- Link, H. and Y. M. Huang (2006). Oligoclonal bands in multiple sclerosis cerebrospinal fluid: an update on methodology and clinical usefulness. J Neuroimmunol **180**(1-2): 17-28.
- Lovato L., Cianti R., Gini B., Marconi S., Bianchi L., A. Armini et al. (2008). Transketolase and 2',3'-cyclic-nucleotide 3'-phosphodiesterase type I isoforms are specifically recognized by IgG autoantibodies in multiple sclerosis patients. Mol Cell Proteomics **7**(12): 2337-49.
- Löwer, R., Löwer, J., Frank, H., Harzmann, R. and R. Kurth (1984). Human teratocarcinomas cultured in vitro produce unique retrovirus-like viruses. J Gen Virol **65** (Pt 5): 887-98.
- Löwer, R., Tönjes, R. R., Korbmacher, C., Kurth, R. and J. Löwer (1995). Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K. J Virol **69**(1): 141-9.
- Lublin F.D. and S. C. Reingold (1996). Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Neurology **46** (4): 907-11.
- Lycke, J., Svennerholm, B., Svenningsson, A., Muranyi, W., Flügel, R. M. and O. Andersen (1994). Human spumaretrovirus antibody reactivity in multiple sclerosis. J Neurol **241**(4): 204-9.

Magin-Lachmann, C., Hahn, S., Strobel, H., Held, U., Löwer, J. and R. Löwer (2001). Rec (formerly Corf) function requires interaction with a complex, folded RNA structure within its responsive element rather than binding to a discrete specific binding site. J Virol **75**(21): 10359-71.

Mameli, G., Astone, V., Arru, G., Marconi, S., Lovato, L., C. Serra, et al. (2007a). Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. J Gen Virol **88**(Pt 1): 264-74.

Mameli, G., Astone, V., Khalili, K., Serra, C., Sawaya, B. E. and A. Dolei (2007b). Regulation of the syncytin-1 promoter in human astrocytes by multiple sclerosis-related cytokines. Virology **362**(1): 120-30.

Mangeney, M., de Parseval, N., Thomas, G. and T. Heidmann (2001). The full-length envelope of an HERV-H human endogenous retrovirus has immunosuppressive properties. J Gen Virol **82**(Pt 10): 2515-8.

Marianayagam N.J., Fawzi N.L., and T. Head-Gordon (2005). Protein folding by distributed computing and the denatured state ensemble. Proc Natl Acad Sci U S A **102**(46): 16684-9

Martin R. and H, F. McFarland (1997). Chapter 14: Immunology of multiple sclerosis and experimental allergic encephalomyelitis. In: Multiple Sclerosis: Clinical and pathogenic basis. Eds. C. S. Raine, H. F. McFarland and W. W. Tourtellotte. London : Chapman & Hall.

Mathey E. K., Derfuss T., Storch M. K., Williams K. R., Hales K., D. R. Woolley et al. (2007). Neurofascin as a novel target for autoantibody-mediated axonal injury. J Exp Med. **204**(10):2363-72.

Matsuoka, M., Nishimoto, I. and S. Asano (1999). Interferon- γ impairs physiologic downregulation of cyclin-dependent kinase inhibitor, p27^{Kip1}, during G1 phase progression in macrophages. Exp. Hematol. **27**(2): 203-9.

McGuffin, L. J., Bryson, K. and D. T. Jones (2000). The PSIPRED protein structure prediction server. Bioinformatics **16**(4): 404-5.

Ménard, A., Amouri, R., Michel, M., Marcel, F., Brouillet, A., J. Belliveau, et al. (1997). Gliotoxicity, reverse transcriptase activity and retroviral RNA in monocyte/macrophage culture supernatants from patients with multiple sclerosis. FEBS Lett **413**(3): 477-85.

Menard, A., Paranhos-Baccala, G., Pelletier, J., Mandrand, B., Seigneurin, J. M., H. Perron et al. (1997). A cytotoxic factor for glial cells: a new avenue of research for multiple sclerosis? Cell Mol Biol (Noisy-le-grand) **43**(6): 889-901.

Menendez, L., Benigno, B. B. and J. F. McDonald (2004). L1 and HERV-W retrotransposons are hypomethylated in human ovarian carcinomas. Mol Cancer **3**(1): 12.

Mi, S., Lee, X., Li, X., Veldmann, G. M., Finnerty, H., Racie, L., E. LaVallie. et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature **403**(6771): 785-9.

Michie, C. (2001). Xenotransplantation, endogenous pig retroviruses and the precautionary principle. Trends Mol Med **7**(2): 62-3.

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., K. W. Kinzler, et al. (1992). Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res **52**(3): 643-5.

Miller, S. D., Olson, J. K. and J. L. Croxford (2001). Multiple pathways to induction of virus-induced autoimmune demyelination: lessons from Theiler's virus infection. J Autoimmun **16**(3): 219-27.

Misra, A., Chosdol, K., Sarkar, C., Mahapatra, A.K. and S. Sinha (2001). Alteration of a sequence with homology to human endogenous retrovirus (HERV-K) in primary human glioma: implications for viral repeat mediated rearrangement. Mutat Res **484**(1-2): 53-9.

Misra, A., Pellarin, M., Shapiro, J.R. and B. G. Feuerstein (2004). A complex rearrangement of chromosome 7 in human astrocytoma. Cancer Genet Cytogenet **151**(2): 162-70.

Muir A., Ruan Q. G., Marron M. P. and J. X. She (1999). The IDDMK(1,2)22 retrovirus is not detectable in either mRNA or genomic DNA from patients with type 1 diabetes. Diabetes **48**(1): 219-22.

Mycko, M. P., Cwiklinska, H., Szymanski, J., Szymanska, B., Kudla G., L. Kilianek, et al. (2004). Inducible heat shock protein 70 promotes myelin autoantigen presentation by the HLA class II. J Immunol **172**(1): 202-13.

Nakagawa K., Brusic V., McColl G., and L. C. Harrison (1997). Direct evidence for the expression of multiple endogenous retroviruses in the synovial compartment in rheumatoid arthritis. Arthritis Rheum **40**(4): 627-38.

Nakamura, A., Okazaki, Y., Sugimoto, J., Oda, T. and Y. Jinno (2003). Human endogenous retroviruses with transcriptional potential in the brain. J Hum Genet **48**(11): 575-81. Epub 2003 Oct 15.

Nash, M. A., Meyer, M. K., Decker, G. L. and R. B. Arlinghaus (1993). A subset of Pr65gag is nucleus associated in murine leukemia virus-infected cells. J Virol **67**(3): 1350-6.

Nellaker, C., Yao, Y., Jones-Brando, L., Mallet, F., Yolken, R. H. and H. Karlsson (2006). Transactivation of elements in the human endogenous retrovirus W family by viral infection. Retrovirology **3**: 44.

Nelson, P. N., Carnegie, P. R., Martin, J., Davari Ejtehad, H., Hooley, P., D. Roden, et al. (2003). Demystified. Human endogenous retroviruses. Mol Pathol **56**(1): 11-8.

Nowak, J., Januszkiewicz, D., Pernak, M., Liweń, I., Zawada, M., J. Rembowska, et al. (2003). Multiple sclerosis-associated virus-related pol sequences found both in multiple sclerosis and healthy donors are more frequently expressed in multiple sclerosis patients. J Neurovirol **9**(1): 112-7.

Nozaki, M., Tada, M., Kobayashi, H., Zhang, C L., Sawamura, Y., H. Abe, et al. (1999). Roles of the functional loss of p53 and other genes in astrocytoma tumorigenesis and progression. Neuro Oncol **1**(2): 124-37.

Obermayer-Straub, P. and M. P. Manns (2001). Hepatitis C and D, retroviruses and autoimmune manifestations. J Autoimmun **16**(3): 275-85.

Oertle, T., van der Haar, M. E., Bandtlow, C. E., Robeva, A., Burfeind, P., Buss, A., et al. (2003). Nogo-A inhibits neurite outgrowth and cell spreading with three discrete regions. J Neurosci **23**(13): 5393-406.

Ogasawara H., Hishikawa T., Sekigawa I., Hashimoto H., Yamamoto N. and N. Maruyama (2000). Sequence analysis of human endogenous retrovirus clone 4-1 in systemic lupus erythematosus. Autoimmunity **33**(1):15-21.

Ogasawara, H., Okada, M., Kaneko, H., Hishikawa, T., Sekigawa, I. and H. Hashimoto (2003). Possible role of DNA hypomethylation in the induction of SLE: relationship to the transcription of human endogenous retroviruses. Clin Exp Rheumatol **21**(6): 733-8.

Ohguro, H., Chiba, S., Igarashi, Y., Matsumoto, H., Akino, T. and K. Palczewski (1993). Beta-arrestin and arrestin are recognized by autoantibodies in sera from multiple sclerosis patients. Proc Natl Acad Sci U S A **90**(8): 3241-5.

Olson, J. K., Croxford, J. L., Calenoff, M. A., Dal Canto, M. C. and S. D. Miller (2001). A virus-induced molecular mimicry model of multiple sclerosis. J Clin Invest **108**(2): 311-8.

Panek R. B. and E. N. Benveniste (1995). Class II MHC gene expression in microglia. Regulation by the cytokines IFN-gamma, TNF-alpha, and TGF-beta. J Immunol **154**(6): 2846-54.

Perron, H., Garson, J. A., Bedin, F., Beseme, F., Paranhos-Baccala, G., F. Komurian-Pradel, et al. (1997). Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc Natl Acad Sci U S A **94**(14): 7583-8.

Perron, H., Geny, C., Laurent, A., Mouriquand, C., Pellat, J., J. Perret, et al. (1989). Leptomeningeal cell line from multiple sclerosis with reverse transcriptase activity and viral particles. Res Virol **140**(6): 551-61.

Perron, H., Jouvin-Marche, E., Michel, M., Ounanian-Paraz, A., Camelo, S., A. C. Dumon, et al. (2001). Multiple sclerosis retrovirus particles and recombinant envelope trigger an abnormal immune response in vitro, by inducing polyclonal V β 16 T-lymphocyte activation. Virology **287**(2): 321-32.

Perron, H., Lazarini, F., Ruprecht, K., P  choux-Login, C., Seilhean, D., V. Sazdovitch, et al. (2005). Human endogenous retrovirus (HERV)-W ENV and GAG proteins: physiological expression in human brain and pathophysiological modulation in multiple sclerosis lesions. J Neurovirol **11**(1): 23-33.

Perron, H., Perin, J. P., Rieger, F. and P. M. Alliel (2000). Particle-associated retroviral RNA and tandem RGH/HERV-W copies on human chromosome 7q: possible components of a 'chain-reaction' triggered by infectious agents in multiple sclerosis? J Neurovirol **6**(Suppl 2): S67-75.

Plumb, J., McQuaid, S., Cross, A. K., Surr, J., Haddock, G., R. A. Bunning, et al. (2006). Upregulation of ADAM-17 expression in active lesions in multiple sclerosis. Mult Scler **12**(4): 375-85.

Prabhakar, S., Kurien, E., Gupta, R. S., Zielinski, S. and M. S. Freedman (1994). Heat shock protein immunoreactivity in CSF: correlation with oligoclonal banding and demyelinating disease. Neurology **44**(9): 1644-8.

Quarles RH (2007). Myelin-associated glycoprotein (MAG): past, present and beyond. J Neurochem **100**(6):1431-48.

Ramnarain, D. B., Park, S., Lee, D. Y., Hatanpaa, K. J., Scoggin, S. O., H. Otu, et al. (2006). Differential gene expression analysis reveals generation of an autocrine loop by a mutant epidermal growth factor receptor in glioma cells. Cancer Res **66**(2): 867-74.

Rasheed, B. K., Wiltshire, R. N., Bigner, S. H. and D. D. Bigner (1999). Molecular pathogenesis of malignant gliomas. Curr Opin Oncol **11**(3): 162-7.

Rasmussen, H. B., Geny, C., Deforges, L., Perron, H., Tourtelotte, W., A. Heltberg, et al. (1995). Expression of endogenous retroviruses in blood mononuclear cells and brain tissue from multiple sclerosis patients. Mult Scler **1**(2): 82-7.

Reindl, M., Khantane, S., Ehling, R., Schanda, K., Lutterotti, A., C. Brinkhoff, et al. (2003). Serum and cerebrospinal fluid antibodies to Nogo-A in patients with multiple sclerosis and acute neurological disorders. J Neuroimmunol **145**(1-2): 139-47.

Rhee, S. S., Hui, H. X. and E. Hunter (1990). Preassembled capsids of type D retroviruses contain a signal sufficient for targeting specifically to the plasma membrane. J Virol **64**(8): 3844-52.

Robbins, J., Dilworth, S. M., Laskey, R. A. and C. Dingwall (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell **64**(3): 615-23.

Rolland, A., Jouvin-Marche, E., Saresella, M., Ferrante, P., Cavaretta, R., A. Créange, et al. (2005). Correlation between disease severity and in vitro cytokine production mediated by MSRV (multiple sclerosis associated retroviral element) envelope protein in patients with multiple sclerosis. J Neuroimmunol **160**(1-2): 195-203.

Rote, N. S., Chakrabarti, S. and B. P. Stetzer (2004). The role of human endogenous retroviruses in trophoblast differentiation and placental development. Placenta **25**: 673-83.

Ruprecht K., Gronen F., Sauter M., Best B., Rieckmann P. and N. Mueller-Lantzsch (2008). Lack of immune responses against multiple sclerosis-associated retrovirus/human endogenous retrovirus W in patients with multiple sclerosis. J Neurovirol **14**(2):143-51.

Ruiz-Vazquez, E. and P. de Castro (2003). 2-6-11 motif in heat shock protein 60 and central nervous system antigens: a preliminary study in multiple sclerosis patients. J Physiol Biochem **59**(1): 1-9.

Sakon M., Kita, Y., Takeda, Y., Higaki, N., Ohzato, H., K. Umeshita, et al. (1999). Measurement of hepatocyte growth factor in serum and plasma. Int J Clin Lab Res **29**(3): 110-113.

Sambrook J., Fritsch E.F. and T. Maniatis (1989). Molecular cloning : a laboratory manual. New York : Cold Spring Harbor Laboratory.

Satoh, J., Onoue, H., Arima, K. and T. Yamamura (2005). Nogo-A and nogo receptor expression in demyelinating lesions of multiple sclerosis. J Neuropathol Exp Neurol **64**(2): 129-38.

Sauter, M., Schommer, S., Kremmer, E., Remberger, K., Dölken, G, I. Lemm, et al. (1995). Human endogenous retrovirus K10: expression of gag protein and detection of antibodies in patients with seminomas. J Virol **69**(1): 414-21.

Schmittgen, T. D. and B. A. Zakrajsek (2000). Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. J Biochem Biophys Methods **46**(1-2): 69-81.

Schroder, K., Hertzog, P. J., Ravasi, T. and D. A. Hume (2004). Interferon-gamma: an overview of signals, mechanisms and functions. J Leukoc Biol **75**(2): 163-89.

Schulte, A. M. and A. Wellstein (1998). Structure and phylogenetic analysis of an endogenous retrovirus inserted into the human growth factor gene pleiotrophin. J Virol **72**(7): 6065-72.

Serra, C., Mameli, G., Arru, G., Sotgiu, S., Rosati, G. and A. Dolei (2003). In vitro modulation of the multiple sclerosis (MS)-associated retrovirus by cytokines: implications for MS pathogenesis. J Neurovirol **9**(6): 637-43.

Shaw, S. Y., Laursen, R. A. and M. B. Lees (1986). Analogous amino acid sequences in myelin proteolipid and viral proteins. FEBS Lett **207**(2): 266-70.

Shirts M. and V. S. Pande (2000). COMPUTING: Screen Savers of the World Unite! Science **290**(5498): 1903-4.

Shortle D (1997). Structure prediction: Folding proteins by pattern recognition. Folding proteins by pattern recognition. Curr. Biol **7**: R151–R154

Sicat J., Sutkowski N. and B. T. Huber (2005). Expression of human endogenous retrovirus HERV-K18 superantigen is elevated in juvenile rheumatoid arthritis. J. Rheumatol **32**(9): 1821-31.

Singh, V. K., Kalra, H. K., Yamaki, K., Abe, T., Donoso, L. A. and T. Shinohara (1990). Molecular mimicry between a uveitopathogenic site of S-antigen and viral peptides. Induction of experimental autoimmune uveitis in Lewis rats. J Immunol **144**(4): 1282-7.

Sotgiu, S., Arru, G., Mameli, G., Serra, C., Pugliatti, M, G. Rosati, et al. (2006). Multiple sclerosis-associated retrovirus in early multiple sclerosis: a six-year follow-up of a Sardinian cohort. Mult Scler **12**(6): 698-703.

Stanley, S. K., Folks, T. M. and A. S. Fauci (1989). Induction of expression of human immunodeficiency virus in a chronically infected promonocytic cell line by ultraviolet irradiation. AIDS Res Hum Retroviruses **5**(4): 375-84.

Stauffer, Y., Theiler, G., Sperisen, P., Lebedev, Y. and C. V. Jongeneel (2004). Digital expression profiles of human endogenous retroviral families in normal and cancerous tissues. Cancer Immun **4**: 2.

Steen H. and M. Mann (2004). The abc's (and xyz's) of peptide sequencing. Nat Rev Mol Cell Biol **5**(9): 699-711.

Steinman, L. (1996). Multiple sclerosis: a coordinated immunological attack against myelin in the central nervous system. Cell **85**(3): 299-302.

Stewart, G. J. (2002). Infection and multiple sclerosis – a new hypothesis? J Neurol Neurosurg Psychiatry **73**(4): 358-9.

Strick, R., Ackermann, C., Langbein, M., Swiatek, J., Schubert S. W., S. Hashemolhosseini et al. (2007). Proliferation and cell-cell fusion of endometrial carcinoma are induced by the human endogenous retroviral Syncytin-1 and regulated by TGF-beta. J Mol Med **85**(1): 23-38.

Sueoka, E., Yukitake, M., Iwanaga, K., Sueoka, N., Aihara, T. and Y. Kuroda (2004). Autoantibodies against heterogeneous nuclear ribonucleoprotein B1 in CSF of MS patients. Ann Neurol **56**(6): 778-86.

Sun, J. B., Olsson, T., Wang, W. Z., Xiao, B. G., Kostulas, V., S. Fredrikson, et al. (1991). Autoreactive T and B cells responding to myelin proteolipid protein in multiple sclerosis and controls. Eur J Immunol **21**(6): 1461-8.

Sun, J., Link, H., Olsson, T., Xiao, B. G., Andersson, B. G., H. P Ekre, et al. (1991). T and B cell responses to myelin-oligodendrocyte glycoprotein in multiple sclerosis. J Immunol **146**(5): 1490-5.

Sutkowski, N., Conrad, B., Thorley-Lawson, D. A. and B. T. Huber (2001). Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. Immunity **15**(4): 579-89.

Sverdlov, E. D. (2000). Retroviruses and primate evolution. Bioessays **22**(2): 161-71.

Thomas, P. J., Qu, B. H. and P. L. Pedersen (1995). Defective protein folding as a basis of human disease. Trends Biochem Sci **20**(11): 456-9.

Truant, R. and B. R. Cullen (1999). The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. Mol Cell Biol **19**(2): 1210-7.

Ueki, K., Ono, Y., Henson, J. W., Efird, J. W., von Deimling, A. and D. N. Louis (1996). CDKN2/p16 or RB alterations occur in the majority of glioblastomas and are inversely correlated. Cancer Res **56**(1): 150-3.

Valdar, S. J. and Jones, D. T. (2004). Chapter 4: Amino acid residue conservation, pp. 49-101. In: Bioinformatics: Genes, Proteins and Computers. Eds: Orengo C., Jones D., Thornton J. Oxford : BIOS Scientific.

Vihinen, M. (1987). Relationship of protein flexibility to thermostability. Protein Eng **1**(6): 477-80.

Voisset, C., Bouton, O., Bedin, F., Duret, L., Mandrand, B., F. Mallet, et al. (2000). Chromosomal distribution and coding capacity of the human endogenous retrovirus HERV-W family. AIDS Res Hum Retroviruses **16**(8): 731-40.

Vu, T., Myers, L. W., Ellison, G. W., Mendoza, F. and J. M. Bronstein (2001). T-cell responses to oligodendrocyte-specific protein in multiple sclerosis. J Neurosci Res **66**(3): 506-9.

Wajgt, A. and M. Gorny (1983). CSF antibodies to myelin basic protein and to myelin-associated glycoprotein in multiple sclerosis. Evidence of the intrathecal production of antibodies. Acta Neurol Scand **68**(5): 337-43.

Walker, D. R., Bond, J. P., Tarone, R. E., Harris, C. C., Makalowski, W., M. S. Boguski, et al. (1999). Evolutionary conservation and somatic mutation hotspot maps of p53: correlation with p53 protein structural and functional features. Oncogene **18**(1): 211-8.

Walsh, M. J. and J. M. Murray (1998). Dual implication of 2',3'-cyclic nucleotide 3' phosphodiesterase as major autoantigen and C3 complement-binding protein in the pathogenesis of multiple sclerosis. J Clin Invest **101**(9): 1923-31.

Wang, Y., Minoshima, S. and N. Shimizu (1993). Precise mapping of the EGF receptor gene on the human chromosome 7p12 using an improved fish technique. Jpn J Hum Genet **38**(4): 399-406.

Wang, S. I., Puc, J., Li, J., Bruce, J. N., Cairns, P, D. Sidransky, et al. (1997). Somatic mutations of PTEN in glioblastoma multiforme. Cancer Res **57**(19): 4183-6.

Weis, S., Llenos, I. C., Sabunciyan, S., Dulay, J. R., Isler, L., R. Yolken, et al. (2007). Reduced expression of human endogenous retrovirus (HERV)-W GAG protein in the cingulate gyrus and hippocampus in schizophrenia, bipolar disorder, and depression. J Neural Transm **114**(5): 645-55.

Weiss, R. A. (2001). Retroviruses and cancer. Current Science **81**(5): 528-534.

Wen P. Y. and S. Kesari (2008). Malignant glioma in adults. N Engl J Med **359**(5): 492-507.

Werle, E., Scheider C., Renner, M., Völker, M. and W. Fiehn (1994). Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res **22**(20): 4534-5.

Wong, A. J., Ruppert J.M., Bigner, S. H., Grzeschik, C. H., Humphrey, C. H., P. A. Humphrey, et al. (1992). Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc Natl Acad Sci U S A **89**(7): 2965-9.

Wucherpfennig, K. W. and J. L. Strominger (1995). Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell **80**(5): 695-705.

Yao, Y., Schröder, J., Nellåker, C., Bottmer, C., Bachmann, S., R. H. Yolken, et al. (2008). Elevated levels of human endogenous retrovirus-W transcripts in blood cells from patients with first episode schizophrenia. Genes Brain Behav. **7**(1): 103-12.

Yi, J. M., Kim, H. M., Lee, W. H. and H. S. Kim (2002). Molecular cloning and phylogenetic analysis of new human endogenous retrovirus HERV-W family in cancer cells. Curr Microbiol **44**(3): 216-20.

Yi, J. M., Kim, H. M. and H. S. Kim (2004). Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells. J Gen Virol **85**(Pt 5): 1203-10.

Zawada, M., Liwień, I., Pernak, M., Januszkiewicz-Lewandowska, D., Nowicka-Kujawska, K., J. Rembowska, et al. (2003). MSRV pol sequence copy number as a potential marker of multiple sclerosis. Pol J Pharmacol **55**(5): 869-75.

Zeinstra, E., Wilczak, N., Streefland, C. and J. De Keyser (2000). Astrocytes in chronic active multiple sclerosis plaques express MHC class II molecules. Neuroreport **11**(1): 89-91.

Zhou, H., Chen, W. D., Qin, X., Lee, K., Liu, L., S. D. Markowitz, et al. (2001). MMTV promoter hypomethylation is linked to spontaneous and MNU associated c-neu expression and mammary carcinogenesis in MMTV c-neu transgenic mice. Oncogene **20**(42): 6009-17.

Acknowledgements

I wish to thank my supervisory team of Dr. Gill Conde and Dr. Paul Nelson for their advice and support through the course of my study, and to Dr. Angie Williams, Keith Holding and and the rest of the technical staff for their assistance.

Thanks also go to the following people for their technical help or sample provision in the following procedures:

Chapter 2:	<u>Bioinformatics</u>	Graham Freimanis Denise Roden
Chapter 3:	<u>ELISA of MS patient samples</u>	Dr. Lawrence Harbige, University of Greenwich (provision of patient plasma samples) Mingyan Xiang (assistance at Greenwich) Dr. Sarah Jones and Prof. John Howl (help with peptide synthesis)
Chapter 4:	<u>PCR development</u>	Dr. Hora Ejtehad Graham Freimanis Dr. Iain Nicholl
Chapters 3-4:	<u>Provision of cell lines</u>	Rachael Baker (MCF-7; SW-480) Dr. Sarah Brown (IN1528, IN1612) Sonia Guidi (IN859, IN1472, U251-MG) Dr. Sarah Jones (U373) Phil Warren (Tera-1) Prof. Nicola Woodroffe, Sheffield Hallam University (CHME3, U87-MG)

A big thankyou to my colleagues, friends, and especially my family in helping me through the bad times and sharing in the better ones over the last few years.

Appendix A Physicochemical parameters of amino acids

Amino acid			Polarity (Grantham, 1974)	Hydrophilicity (Hopp & Woods, 1981)	Accessibility (Janin, 1979)	Flexibility (Bhaskaran & Ponnuswamy (1998))
Alanine	Ala	A	8.100	-0.500	6.600	0.360
Arginine	Arg	R	10.500	3.000	4.500	0.530
Asparagine	Asn	N	11.600	0.200	6.700	0.460
Aspartic acid	Asp	D	13.000	3.000	7.700	0.510
Cysteine	Cys	C	5.500	-1.000	0.900	0.350
Glutamine	Gln	Q	10.500	0.200	5.200	0.490
Glutamic acid	Glu	E	12.300	3.000	5.700	0.500
Glycine	Gly	G	9.000	0.000	6.700	0.540
Histidine	His	H	10.400	-0.500	2.500	0.320
Isoleucine	Ile	I	5.200	-1.800	2.800	0.460
Leucine	Leu	L	4.900	-1.800	4.800	0.370
Lysine	Lys	K	11.300	3.000	10.300	0.470
Methionine	Met	M	5.700	-1.300	1.000	0.300
Phenylalanine	Phe	F	5.200	-2.500	2.400	0.310
Proline	Pro	P	8.000	0.000	4.800	0.510
Serine	Ser	S	9.200	0.300	9.400	0.510
Threonine	Thr	T	8.600	-0.400	7.000	0.440
Tryptophan	Trp	W	5.400	-3.400	1.400	0.310
Tyrosine	Tyr	Y	6.200	-2.300	5.100	0.420
Valine	Val	V	5.900	-1.500	4.500	0.390

Appendix B Epitopes identified

B.1 Retroviral epitopes

HERV-H (RGH-1) env

97-111 LDRSSKTSPDISH
230-244 SPCSSDSPRPSSRLL

HERV-H gag

190-204 QFTLKKVAGAKGIVK

HERV-W env (Syncytin)

117-134 GGVQDQAREKHVKEVIS
412-437 GIVTEKVKEIRDRIQRRAEEL
RNTGP

HERV-W gag

134-148 KEIDKGVNNEPKSAN
307-322 ITQGKEENPTAFLDRL

HRES-1 p15

68-86 RALHPPSGRDREEEEMGY
96-118 ACARGGRGGAREDFGARRKHV
RG

HRES-1 p25

37-53 PRLRPRHRHPQDPRSPG
56-71 PRHRRPPRPDPRAPPA
107-131 LGGEGPGAGDRRREGPDSPR
QPPV

HTLV-1 gag-pro-pol

105-120 SSTDPPDSDPQIPPP
227-242 LRVQANNPQQGLRRE
301-317 SLAYSANKECQKLLQ
330-344 RACQAWTPKDKTKVL
397-413 KPTIPEPEPEDALLD
654-669 IDLSSSSPGPDLSSL
854-868 CALQRHTDPRDQIYL
1252-1270 GAISATQKRKETSSEAIS
1455-1472 RPVGGPADPKEKDHQHHG

HTLV1 gp46

87-102 HWIKKPNRNGGGYY

HTLV1 tax

318-337 SLLFNEKEADDNDHEPQISP
339-353 GLEPLSEKHFRETEV

MSRV env

120-135 IQGQAREKQVKEAISQ
412-428 RIVTEKVKEIRDRIQC
424-439 RIQCRAEELQNTERWG

MSRV gag

123-138 KEIDKGVNNEPKSANI
274-289 TCVLEGLRKTRKKPMN

MSRV pol

1-15 SSSRTEGARGKCQPM
345-359 CGYKVSQPKARLCSQ
580-596 TFLPDNEEKIEHNCQQV
731-765 LHCQGHQEEEEIEGNRQAD
IEAKKAARQDSPLEML

Visna env

1-17 MASKEKPSRTTTRGME
29-48 QELVKRQQQEEEEQQGLVSG
57-73 DLLGTEGKDIKKVNIWE
191-218 RQWMKENEKEYKERTNKTED
IDDLVAG
301-315 LRCQDEGKSPGGCVQ
328-342 EAMKYLRGKKSRYGG
338-356 SRYGGIKDKNGELKLPLS
636-651 SYMEAQGENKRSRNL
776-791 QQWEEIEQHEGNLS
876-896 EAPVELEEKQKRNGDTNGCA
897-911 SLEHERRTSHRSFIQ
950-975 NGWNGENQHKKKKERVDCQDR
EQMPT

Visna gag

1-18 MAKQGSKEKKGYPELKEV
71-86 KGLTPEETSKREFASL
112-134 ISMKEGLHENKEAKGEKVEQL
YP
212-226 LIQKGLNNEEAERWVR
284-299 MLVKQKNTESYEDFIA
347-362 TVEEKMQACRDVGSEG
414-436 HMQKDCRQKKQQGNRRGPRV
VP

Visna pol

14-29 LPAEETAGKQQEEGAT
173-188 KEIVDRLEKEGKVGRA
210-226 MLIDFRELNKQTEDLA
350-367 GFMLPEDKRQEGYPAKW
453-468 GNYYDEEKDIYGQLDQ
571-568 TYYTDGGKKNGRGS LGY
596-610 RIHEEGTNQQLELRA
668-683 PGHKGIPQNEEIDRYI
695-709 ILQKRAEDAGYDLIC
799-816 EPWGETRKTERGEQFGS
868-883 MPSTLRGSNKRGIHW
1028-1045 QRIQQQSKSKQEKIRFCY

B.2 MS autoantigen epitopes

Alpha B Crystallin

100-116 VH GKHEERQDEHGFISR

Beta Arrestin 1 Isoform A

149-166 FCAENLEEKIHKRNSVRL
219-233 HVTNNTNKT VKKIKI
352-367 MHPKPKEEPPHREVPE
374-389 TNLIELDTND DDIVFE
392-417 ARQLKGMKDDKEEEDGTGS
PQLNN

Beta Arrestin 1 Isoform B

90-114 APEDKKPLTRLQERL
149-166 FCAENLEEKIHKRNSVRL
219-233 HVTNNTNKT VKKIKI
286-301 GLALDGK LKHEDTNLA
344-359 MHPKPKEEPPHREVPE
366-381 TNLIELDTND DDIVFE
384-409 ARQLKGMKDDKEEEDGTGS
PQLNN

Claudin 11 (Oligodendrocyte Specific Protein)

39-53 PTCRKLDELGSKGLW

2', 3'-Cyclic Nucleotide 3' Phosphodiesterase

1-15 MSSSGAKDKPELQFP
33-48 FILRGLPGSGKSTLAR
48-62 RVIVDKYRDGTMVVS
74-89 GAFSEYKRLDEDLAA
98-116 ILVLDDTNHERERLEQLFE
146-166 QWQLSADDLKKLKPGLKDFL
173-187 FLTKKSSETLRKAGQ
204-221 RQFVPGDEPREKMDLVTY
330-348 EILRQEKGGSRGEEVGELS
382-398 KPVPTQGSRKGGALQS

Heterogenous Nuclear Riboprotein A2/B1 Isoform B

6-25 ETVPLERKKREKEQFRKLFI
51-66 VVMRDPASKRSRGFGF
114-133 LFVGGIKEDTEEHHLRDYFE
142-157 EIITDRQSGKKRGFGF
158-172 VTFDDHDPVDKIVLQ
195-209 EVQSSRSRGGRGFGF
226-240 NFRGSDGYGSGRGF
334-353 GNYGPGSGSGSGGYGGRSRY

Heat Shock Protein 60 (Chaperonin)

78-94 VTVAKSIDLKDKYKNIGA
122-140 SIAKEGFEEKISKGANPVE
172-187 ATISANGDKEIGNIIS
186-200 ISDAMKKVGRKGVIT
199-213 ITVKDGKTLNDELEI
245-260 LLSEKKISSIQSIVPA
303-319 PGFGDNRKNQLKDMAI
357-373 LLKKGDKAQIEKRIQE
378-397 LDVTTSEYEKEKLNERLAKL
406-428 VGGTSDVEVNEKKDRVTDAL
NAT
477-491 MTIAKNAGVEGSLIV
544-562 VVVTEIPKEEKDPGMGAMG

Heat Shock 70kDa Protein 1A

67-82 VFDAKRLIGRKFGDPV
93-107 QVINDGDKPKVQVSY
148-163 AYFNDSQRQATKDAGV
183-197 YGLDRTGKGERNVLI
239-262 NHFVEEFKRKHKKDISQNKRA
VRR
319-333 KALRDAKLDKAQIHD
358-373 DLNKSINPDEAVAYGA
379-393 ILMGDKSENVQDLLL
441-457 QVYEGERAMTKDNNLLG
488-503 VTATDKSTGKANKITI
502-517 TITNDKGRLSKEEIER
517-540 RMVQEAKEYKAEDVQRERVS
AKN
552-578 AVEDEGLKGKISEADKKKVLD
KCQEVI
579-598 GGFGAGGPKGGSGSGPTIEE

Heat Shock 70kDa Protein 1B

93-107 QVINDGDKPKVQVSY
148-163 AYFNDSQRQATKDAGV
183-197 YGLDRTGKGERNVLI
239-262 NHFVEEFKRKHKKDISQNKRA
VRR
319-333 KALRDAKLDKAQIHD
358-373 DLNKSINPDEAVAYGA
379-393 ILMGDKSENVQDLLL
441-457 QVYEGERAMTKDNNLLG
489-503 TATDKSTGKASKITI
502-517 TITNDKGRLSKEEIER
517-540 RMVQEAKEYKAEDVQRERV
SAKN
552-578 AVEDEGLKGKISEADKKKVLD
KCQEVI
579-605 SWLDANTLAEKDEFEHKRKEL
EQVCN
620-639 GGFGAQGPKGGSGSGPTIEE

Heat Shock 70kDa Protein 8 Isoform 1

67-82 VFDAKRLIGRRFDDA
148-163 AYFNDSQRQATKDAGT
182-197 AYGLDKKVGAEARNVLI
240-263 HFIAEFKRKHKKDISQNKRA
VRR
318-333 EKALRDAKLDKSQIHD
358-373 ELNKSINPDEAVAYGA
378-393 AILSGDKSENVQDLLL
441-457 QVYEGERAMTKDNNLLG
488-504 VSAVDKSTGKENKITIT
502-517 TITNDKGRLSKEDIER
517-544 RMVQEAKEYKAEDKQRDKVS
SKNSLES
550-576 KATVEDEKLQGGKINDEKQKI
LDKCNE
578-604 INWLDKNQTAEKEEFHQKKE
LEKVCN
627-643 GGGAPPSGGASSGPTIEE

Heat Shock 70kDa Protein 8 Isoform 2

67-82 VFDAKRLIGRRFDDAV
148-163 AYFNDSQRQATKDAGT
182-196 AYGLDKKVGAEARNVLI
240-263 HFIAEFKRKHKKDISENKRAV
RRL
318-333 EKALRDAKLDKSQIHD
358-373 ELNKSINPDEAVAYGA
378-393 AILSGDKSENVQDLLL
441-458 QVYEGERAMTKDNNLLG
474-490 GGGAPPSGGASSGPTIE

229-243 RL
TVTNNTEKTVKKIKA
296-311 GIALDGKIKHEDTNLA
359-379 LMHPQPEDPAKESYQDANLVF
384-405 RHNLDAGEAEEGKRDKNDAD
E

Transaldolase 1

70-84 RKLGGSQEDQIKNAI
126-141 IELYKEAGISKDRILI
198-212 VANTDKKSYEPLEDP
271-286 VPVLSAKAAQASDLEK

Myelin Associated Glycoprotein**Isoform A Precursor**

184-199 AVLGRLEDEGTWVQ
533-548 ITQTRKKNVTESPSF
564-585 ISGAPEKYESERRLGSERRLL
G
598-612 SHSDLGKRPTKDSYT

Myelin Associated Glycoprotein**Isoform B Precursor**

184-199 AVLGRLEDEGTWVQ
533-548 ITQTRKKNVTESPSF
564-580 ISGAPEKYESKEVSTLE

Myelin Basic Protein

44-61 RFFGGDRGAPKRGSGKVP
99-103 LPQKSHGRTQDENPV
168-185 KIFKLGRDSRSGSPMAR

NoGo A (Reticulon 4 Isoform A)

8-22 PLVSSSDSPPRQPA
23-57 FKYQFVREPEDEEEEEEEEE
DEDEDLEELEVLER
126-146 AAVSPSKLPEDEPPARPPPP
173-190 PAAPKRRGSSGSVDETLF
263-285 GTLQENVSEASKEVSEKAKTL
LI
321-342 EEIIVKNKDEEEKLVSNIL
352-376 KLVKEDEVVSSEKAKDSFNEK
RVAV
391-409 ERVWEVKDSKEDSDMLAAG
412-426 IESNLESKVDKKCFA
427-450 DSLEQTNHEKDSESSNDDTSF
PST
481-506 PLLGDPTSENKTDEKKIEKK
AQIVT
503-519 QIVTEKNTSTKTSNPFL
665-684 VSLKKVSGIKEIKEPENIN
732-747 SELVEDSSPDSEPVDL
750-765 DDSIPDVPQKQDETVM
778-792 SMIEYENKEKLSALP
926-947 PKVEEKISFSDDFSKNGSATS
971-997 KVLVKEAEKKLPDTEKEDRS
PSAIFS

Proteolipid Protein

115-129 ATVTGGQKGRGSRGQ

S Arrestin

2-17 AASGKTSKSEPNHVIF
14-29 HVIFKKISRDKSVTIY
49-63 VLVDPLVKGKKVYV
98-112 VGAASTPTKLQESLL
108-122 QESLLKKLGSNTYPF
154-176 KAFATDSTDAEEDKIPKKSSV

Appendix C Ethical approval

RES20A Form submitted to School of Applied Sciences Ethics Committee

Submitted: 14/10/2003

Approved: 26/01/2005



SCHOOL OF APPLIED SCIENCES

RES 20A

Ethics and Safety Committee: submission of project for approval

This form should be word processed – no handwritten forms can be considered

CATEGORY A PROJECTS:

There is no severe or significant interference with the subjects' physical or psychological wellbeing. The subjects are not considered vulnerable to the procedures or topic of the project proposed. With the exclusion of activities in SHS as listed below:-

Taking samples of urine, saliva, skin scrapes for sweat, capillary blood and venous blood taken by a trained phlebotomist and administration of drugs in quantities constituting no abnormal hazard to health.

Projects involving access to confidential records may be considered Category A provided that the investigator's access to these is part of his/her normal professional duties.

Category A projects will be approved at divisional level on an individual basis and monitored by the School Ethics and Safety Committee. The School Ethics and Safety Committee will not normally expect to see individual Category A projects but would expect to see, on a semester basis, a record of projects that have been approved at divisional level.

Schools with a large number of Category A projects at u/g level should put appropriate generic procedures in place, and these will be approved by the School Ethics and Safety Committee.

Date approved at School Ethics Committee:	
Covering letter of approval from School:	

Investigator	Project Title	Level (u/g, p/g)	Supervisor
P.Nelson, G.Conde, D.Roden, R.Smith, + Res Assistants/Techn icians, PhD/MSc/BSc students	Development of antibody reagents to endogenous viral antigenic regions implicated in multiple sclerosis	Tissue sections/cell lines: probable ng/pg levels of antigen	Dr. P.Nelson Dr. G.Conde

Signed: _____
(Dean of School/Head of Division/Chair of
SRC/Chair of Divisional Ethics Committee)

Date: _____

Signed: _____
(Chair of School Ethics Committee)

Date: _____

Protocol/Explanation of Method please see attached sheet	
Full details of the methodology are attached. In brief, antibodies will be used as primary reagents on tissue sections and cell cytospin preparations. Antibodies will be revealed using a secondary antibody conjugated to enzyme or fluorescent moiety. Signal intensity will then be visualised using light/immunofluorescent microscope and photographed. In ELISA, antibodies will be tested against peptide and revealed using a coloured/chemiluminescent substrate.	
Information Sheet for Subjects and Informed Consent Form Brain tissue donated with prior consent to the MS Tissue Bank (letter attached).	NA
Details of how information will be held Data of HERV expression will be held on work computer hard drive/CD/Floppy disk.	*
Details of if/how results will be fed back to participants Brain tissue has been donated following death of MS and other patients. Research data appertaining to HERV expression will be provided to the MS Tissue Bank and MS Society (Annual Report). Levels of HERV expression in cell lines will be submitted to Prof. Woodroffe	*
Letter of consent from any Collaborating Institutes Details enclosed	
Letter of consent from Head Teacher, if participants under the age of 16 years of age	NA
Is any other External Ethical Approval required? NO If so, which Committee? N/A	

Signed: Dr. PN Nelson

Date: 10th November 2003