

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

<http://go.warwick.ac.uk/wrap/38417>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

**A REVERSE GENETICS APPROACH  
TO STUDY THE PATHOGENESIS OF  
PNEUMONIA VIRUS OF MICE**

Yashar Mohammadzadeh Sadigh, B.Sc., M.Sc.

**A thesis submitted for the degree of Ph.D**

School of Life Sciences, University of Warwick, UK

February, 2011

## Table of content

CHAPTER 1 : INTRODUCTION.....	1
1.1 PVM and RSV: An overview and virus taxonomy .....	2
1.1.1 General features of the virion.....	5
1.2 Molecular biology and replication cycle of pneumoviruses .....	6
1.2.1 Attachment and entry.....	7
1.2.1.1 G glycoprotein.....	8
1.2.1.1.1 Second open reading frame in the PVM G gene .....	12
1.2.1.1.2 Proteolytic processing of G glycoprotein .....	13
1.2.1.2 Fusion (F) glycoprotein .....	14
1.2.1.3 Uncoating .....	15
1.2.2 Genome structure and organisation .....	16
1.2.3 RNA dependent RNA polymerase complex .....	18
1.2.3.1 Nucleocapsid (N) protein of pneumoviruses.....	19
1.2.3.2 Phosphoprotein (P) protein of pneumoviruses .....	20
1.2.3.3 Large polymerase (L) protein of pneumoviruses.....	21
1.2.3.4 M2-1 (22K) protein.....	22
1.2.4 Transcription and gene expression .....	23
1.2.4.1 Gene expression gradient .....	24
1.2.4.2 mRNA capping .....	25
1.2.4.3 Transcription termination and polyadenylation.....	26
1.2.5 Genome replication.....	26
1.2.5.1 M2-2 protein.....	29

1.2.6	Virus assembly and release .....	30
1.2.6.1	Matrix (M) protein .....	30
1.2.6.2	Virus release .....	31
1.3	Pneumovirus pathogenicity.....	31
1.3.1	Virus factors affecting pathogenicity.....	31
1.3.1.1	Non-structural (NS) protein 1 and 2 .....	31
1.3.1.2	Effect of the G glycoprotein in pneumoviruses pathogenesis.....	32
1.3.1.3	Effect of the F glycoprotein in pneumoviruses pathogenesis.....	34
1.3.2	Host factors affecting pathogenicity.....	36
1.3.2.1	The effect of tachykinins on respiratory disease .....	37
1.3.3	Long lasting HRSV infection.....	38
1.3.4	Using PVM to study the pathogenesis of HRSV.....	39
1.3.5	Aims.....	40
CHAPTER 2 : MATERIALS AND METHODS .....		41
2.1	Solutions .....	42
2.1.1	AEC peroxidase substrate .....	42
2.1.2	ABTS Substrate .....	42
2.1.3	Carbonate coating buffer.....	42
2.1.4	Luria-Bertani broth and agar media.....	42
2.1.5	Ketamine and xylazine cocktail.....	42
2.1.6	Phenol:chloroform mixture .....	42
2.1.7	SDS-PAGE sample buffer.....	43
2.1.8	SDS-PAGE electrophoresis buffer .....	43
2.1.9	TAE Buffer.....	43
2.1.10	TB solution.....	43
2.1.11	TBE Buffer .....	43
2.1.12	TBS and TBST buffers .....	43

2.1.13 X-Gal and IPTG solutions.....	43
2.2 Cell culture and classic virology techniques.....	44
2.2.1 Viruses and cell lines .....	44
2.2.2 Mammalian cell culture .....	44
2.2.3 Freezing of mammalian cells .....	44
2.2.4 Transfection of mammalian cell lines using Lipofectamine 2000 .....	45
2.2.5 Transfection of mammalian cell lines using Turbofect .....	45
2.2.6 Virus cultivation .....	45
2.2.7 Micro-plaque assay .....	46
2.3 Molecular biology techniques .....	46
2.3.1 Plasmid DNA extraction .....	46
2.3.2 DNA Digestion.....	47
2.3.3 Gel extraction .....	47
2.3.4 DNA ligation and low melting point agarose dependent ligation .....	47
2.3.5 Preparation of competent bacterial cells.....	48
2.3.6 Heat-shock based transformation of <i>E. coli</i> .....	48
2.3.7 RNA isolation.....	49
2.3.8 pBS-T vectors preparation .....	49
2.3.9 Blue and white colony screening.....	49
2.3.10 Reverse transcription polymerase chain reaction (RT-PCR) .....	49
2.3.11 Polymerase Chain Reactions (PCR) .....	50
2.3.12 Quick change site-directed mutagenesis .....	51
2.3.13 Primer design for Quick change site-directed mutagenesis .....	51
2.3.14 “Round the horn” site directed mutagenesis PCR .....	51
2.3.15 DNA sequencing.....	52
2.3.16 Nucleotide similarity searches.....	52
2.3.17 SDS - polyacrylamide gel electrophoresis (SDS-PAGE) .....	53

2.3.18 Western blot .....	54
2.3.19 Fluorescent detection of PVM P protein.....	55
2.4 Reporter genes assay .....	56
2.4.1 GFP ELISA .....	56
2.4.1.1 Rabbit anti-GFP polyclonal antibody .....	57
2.4.2 Luciferase activity assay .....	57
2.5 Animal studies.....	57
<b>CHAPTER 3 : PATHOGENESIS OF PVM .....</b>	<b>59</b>
3.1 Introduction.....	60
3.2 Evaluation of the pathogenicity of a working stock of PVM strain J3666.....	60
3.3 Effect of consecutive passages of PVM strain J3666 on the pathogenicity of the virus.....	63
3.4 Analysing the effect of consecutive passages on virus genome nucleotide sequence .....	66
3.4.1 The fragment 2 region.....	67
3.4.2 The fragment 3 region.....	68
3.4.3 The fragment 4 region.....	71
3.4.4 The fragment 5 region.....	71
3.5 Analysing the SH gene end and G ORF start sequence consistency in PVM strain J3666 population .....	71
3.6 Conclusion and discussion.....	76
<b>CHAPTER 4 : A REVERSE GENETICS APPROACH TO STUDY THE EFFECT OF THE G GENE ORGANISATION IN PVM PATHOGENICITY .....</b>	<b>81</b>
4.1 Introduction.....	82
4.2 Plasmid p15FL-2G .....	82
4.3 Complete sequencing of p15FL-2G .....	86
4.4 Amending plasmid p15FL-2G .....	87
4.5 Insertion of fragment 3 into p15FL2G.....	90

4.6	Description of the full length genomes.....	90
4.7	Rescue of plasmid p15FL-2G .....	91
4.7.1	Growth of recombinant viruses in cell culture .....	95
4.8	Testing the pathogenesis of recombinant viruses.....	98
4.9	rPVM-G(15).....	98
4.10	rPVM-G(15 $\Delta$ ORF1) .....	99
4.11	rPVM-G(J3666).....	100
4.12	rPVM-G(J3666 $\Delta$ ORF1).....	102
4.13	Discussion .....	103
<b>CHAPTER 5 : EFFECT OF THE FIRST ORF ON PVM G PROTEIN EXPRESSION</b>		
	.....	105
5.1	Introduction .....	106
5.2	Construction of the PVM minigenome carrying G gene of PVM.....	106
5.3	Quick change mutagenesis to make pWt/NcoI .....	108
5.4	The Luciferase gene region.....	109
5.5	The G-GFP region .....	110
5.5.1	Amplifying the G gene from PVM strains; general features .....	110
5.5.1.1	Amplifying G gene from PVM strain 15 template .....	111
5.5.1.2	Amplifying the G gene from strain J3666 template .....	111
5.5.2	Amplification of the GFP gene.....	112
5.5.3	The overlap PCR to join G and GFP together.....	112
5.6	Ligation of G-GFP fragments into pNcoI/PstI-Luc .....	115
5.7	Deletion of the first ORF in both constructs.....	115
5.8	Correction of the first ORF stop codon in pG-GFPJ3666 .....	116
5.9	Cloning the G-GFP fragments under T7 promoter .....	116
5.10	Expression of the G-GFP fusion proteins in Vaccinia Virus T7 infected cells	118

5.11 Measuring the G-GFP fusion proteins expression in BSRT-7 cells transfected with pG-GFP plasmids .....	120
5.12 Discussion .....	123
CHAPTER 6 : DISCUSSION .....	126



## List of Figures

Figure 1.1 Pneumovirus structure.....	5
Figure 1.2 Comparison between RSV and PVM vRNA.....	7
Figure 1.3 Schematic structure of the G glycoprotein of HRSV.....	11
Figure 1.4 Comparison of the nucleotide sequence of the G gene of PVM strain 15 (Warwick) and PVM strain J3666.....	12
Figure 1.5 Comparison of the G gene organisation in PVM strain 15 Warwick and PVM strain J3666.....	13
Figure 1.6 The HRSV N protein structure.....	20
Figure 1.7 The process involving cap synthesis in eukaryotes and in VSV.....	25
Figure 1.8 HRSV genome replication.....	28
Figure 1.9 Genomic sequence of the leader domain of Pneumoviruses.....	28
Figure 1.10 Comparison between leader sequence in the genome and antigenome of HRSV.....	29
Figure 2.1 Round the horn mutagenesis PCR.....	52
Figure 3.1. Weight loss of BALB/c mice infected with different amounts of PVM strain J3666 passaged 4 times in BSC-1 cells.....	61
Figure 3.2. Clinical score of BALB/c mice infected with different amounts of PVM strain J3666 passaged 4 times in BSC-1 cells.....	63
Figure 3.3. Weight loss and clinical score of mice infected with 5000, 500 and 250 pfu of PVM passages 5, 6 and 7 in tissue culture cells.....	65
Figure 3.4 The PVM genome and the location of primers on the genome.....	67
Figure 3.5 Alignment of the SH protein of PVM strain J3666 and PVM strain J3666 passages 5, 6 and 7.....	70

Figure 3.6 Amino acid comparison between G proteins of PVM strain J3666, G protein of strain 15 (ATCC), G protein of strain J3666 passages 5, 6 and 7.....	70
Figure 3.7 Describing the G gene organisation in PVM strain 15 (ATCC) and PVM strain J3666.....	72
Figure 3.8 RT-PCR amplification, amplifying the SH-G fragment from cDNA of PVM strain J3666.....	73
Figure 3.9 EcoRI digestion of pGEM-SHG plasmids.....	73
Figure 4.1 Plasmid p15FL-2G (17944 bp) .....	83
Figure 4.2 The syntheses of p15FL2G-GJ3666, pFL2G-GJ3666ΔORF1, pFL2G-G15 and pFL2G-G15ΔORF1 .....	84
Figure 4.3 Digestion of the 2718 bp fragment generated by PCR amplification from pF3-BSHT1-BGL2 plasmid using J5 and F14 primers.....	89
Figure 4.4 The position of the 1stORFOut-15-R and 1stORFOut-15-R primers to each other.....	91
Figure 4.5 The rescue process to obtain infectious virus from cells .....	93
Figure 4.6 Digestion of the M-SH PCR fragment from the wildtype and recombinant viruses containing a <i>SalI</i> restriction site.....	94
Figure 4.7 Detection of cells infected with the recombinant viruses using the 26/C3/B5 monoclonal antibodies detecting PVM P protein .....	94
Figure 4.8 Growth of stock or recombinant PVM isolates in BS-C-1 cell culture system at two temperatures, 31°C and 37°C .....	96
Figure 4.9 Change in body wight in groups of mice infected with rPVM-G(15) .....	98
Figure 4.10 Clinical score for mice infected with the rPVM-G(15) .....	99

Figure 4.11 Change in body weight in groups of mice infected with rPVM-G(15ΔORF1) .....	100
Figure 4.12 Clinical score for mice infected with rPVM-G(15ΔORF1) .....	100
Figure 4.13 Change in body weight in groups of mice infected with rPVM-G(J3666) .....	101
Figure 4.14 Clinical score for mice infected with rPVM-G(J3666) .....	101
Figure 4.15 Change in body weight in groups of mice infected with rPVM-G(J3666ΔORF1).....	102
Figure 4.16 Clinical score for mice infected with rPVM-G(J3666ΔORF1) .....	102
Figure 5.1 The synthesis of pG-GFP minigenome.....	107
Figure 5.2 Schematic structure of pWt/NdeI. Naturally occurring restriction sites are shown in black and the recombinant restriction site introduced upstream of the CAT gene is shown in red. ....	108
Figure 5.3 Quick Change mutagenesis reaction to introduce an NcoI restriction site before the start codon of CAT gene.....	109
Figure 5.4 Restriction enzyme digestion of pWt/NcoI and pNcoI/PstI with PstI and NcoI restriction enzymes.....	110
Figure 5.5 PCR amplification of the PVM strain 15 (Warwick) G gene using primers PstIFor and PstIRev.....	111
Figure 5.6 Amplification of the GFP gene from pWT/NdeI DNA.....	112
Figure 5.7 The overlap PCR method used to join the G and GFP fragments to synthesis the G-GFP fragments for the strains 15 and J3666.....	113
Figure 5.8 The overlap PCR reaction. The fragment G (blue) and the fragment GFP (green) are shown in the overlap PCR reaction.....	114

Figure 5.9 The overlap PCR to join G and GFP fragments together.....	114
Figure 5.10 Ligation of the G-GFP fragment into pNcoI/PstI-Luc.....	115
Figure 5.11 Deletion of the first ORF of the G gene in pG-GFP15 and pG-GFPJ3666 .....	116
Figure 5.12 Strategy used for constructing pBS-G-GFP plasmids.....	117
Figure 5.13 PstI digestion of the pBS-GGFP constructs.....	118
Figure 5.14 Detection of G-GFP fusion protein using SDS-PAGE and western blot .....	119
Figure 5.15 The ratio between the GFP expression and luciferase activity.....	122

## List of Tables

Table 1.1 Classification of family Paramyxoviridae based on the reclassification of International Committee on the Taxonomy of Viruses (ICTV) in 2009 (ICTV, 2009). .....	4
Table 1.2 The relative molecular mass of G glycoproteins of PVM strain 15 (Warwick) adapted from (Ling & Pringle, 1989b).....	10
Table 1.3 Comparison between HRSV, BRSV and PVM genome length.....	16
Table 1.4 Alignment of the sequences of gene starts in PVM.....	18
Table 1.5 Comparison of gene end sequence in HRSV strain A2.....	18
Table 1.6 Percentage molar ratio of HRSV mRNA indicating the genome expression gradient (Barik, 1992).....	24
Table 2.1 Ingredients to prepare a 8%, 10% and 12% SDS-PAGE electrophoresis gel .....	54
Table 2.2 Clinical scoring system from (Cook et. al. 1998).....	58
Table 3.1 Nucleotide differences between PVM strain J3666 passages 5, 6 and 7 used in this study and PVM strain J3666.....	69
Table 3.2 Sequence differences among the SH-G clones and PVM strain J3666 sequence.....	75
Table 3.3 Sequence differences among the SH-G clones and PVM strain J3666 sequence.....	76
Table 4.1 Comparison of sequence of PVM strain 15 (Warwick) and p15FL-2G.....	86
Table 4.2 DNA concentrations used to rescue PVM from cDNA.....	92
Table 5.1 The concentrations of the minigenomes and helper plasmids, coding for the RNP complex of PVM, used to transfect BSRT-7/5 cells .....	120
Table 5.2 The luciferase activities achieved from two independent transfections of BSRT-7/5 cells .....	121
Table 5.3 The concentrations of GFP proteins achieved from two independent transfections .....	121

## **Acknowledgement**

I would like to acknowledge the advice and guidance of Professor Andrew Easton, my supervisor during my PhD studies. His endless encouragement and guidance from the initial to the final level enabled me to develop an understanding of the subject. The completion of this work without his support would not have been possible. I would like to thank my parents, family members and Miss Sabrina Powell for their support and encouragement. I would like to thank Dr. Roger Ling, Dr. Anthony Marriott, Dr. Phillip Gould, Dr. Bo Meng, Dr. Magali Droniou and Dr. Nicole Edworthy for their guidance and help during this work. I also thank the members of the animal house, genomics and preparation laboratories for facilities provided by them. I would like to thank the Ministry of Health and Medical Education of Iran for providing financial support.

## **Declaration**

I hereby declare that all the results presented in this thesis were obtained by myself under the supervision of Prof. A.J. Easton, unless stated otherwise. This thesis has not been submitted for a degree in any other institution.

## Summary

Human and bovine respiratory syncytial viruses (HRSV and BRSV), along with pneumonia virus of mice (PVM) are the members of the genus *Pneumovirus* in the subfamily *Pneumovirinae* of the family *Paramyxoviridae*. Although both HRSV and BRSV have been associated with important diseases in human and livestock, there is no clearcut description of the molecular aspects of their pathogenesis. For HRSV, the lack of a suitable study model is one of the main reasons hampering the study of aspects of pathogenesis of the virus. HRSV infects a wide range of animal models, however most of the common laboratory animal models are not sufficiently permissive to study the infectivity of the virus.

PVM naturally infects mice and causes a disease indistinguishable from that of HRSV in humans. Two strains of PVM have been described: strain 15 (Warwick) which is not pathogenic and strain J3666 which is highly pathogenic. The main difference between these two strains lies in the organisation of the gene encoding the attachment (G) glycoprotein. The G gene in strain J3666 has two ORFs. The larger second ORF codes for the G glycoprotein and is located downstream of the first ORF which has no known function. The strain 15 G gene also contains two ORFs but in this case both the first and main ORFs overlap each other.

The aim of the project was to investigate the molecular basis for pathogenesis of PVM as a model for pneumoviruses. As a first step, the pathogenesis of PVM strain J3666 was reevaluated and the effect of consecutive tissue culture passages on the pathogenicity of the virus was examined. It was shown that consecutive passages of PVM strain J3666 caused attenuation of the virus. To investigate the possible mutations causing the attenuation the majority of the virus genome was sequenced from three passage stocks where the transition from pathogenic to non-pathogenic occurred. No differences in the genome sequences for the three passage stocks were found. However, sequence analysis of individual clones of the SH and G genes of the viruses showed evidence that the stocks contained a mixed population of sequences.

A robust reverse genetics system was established to rescue recombinant PVM from cDNA using co-transfection of plasmids coding for the ribonucleoprotein complex of the virus (N, L, M2-1 and P proteins) and a cDNA copy of the virus genome cloned under the control of the bacteriophage T7 RNA polymerase. Using this system, four viruses differing in their G gene organisation were generated and used to infect mice to study the effect of mutations on pathogenicity. It was shown that the viruses with the G gene of strain 15 (Warwick) lacking the first ORF manifest a modest increase in their pathogenicity compared to the non-pathogenic PVM strain 15(Warwick) parent. The recombinant viruses containing the G gene organisation of strain J3666 showed the highest level of pathogenicity.

The reverse genetics system was used to study the role of the first ORF in G glycoprotein expression. Using a dicistronic minigenome construct, the effect of the presence or absence of the first ORF in both the strain 15 and strain J3666 G gene organisation was studied. It was shown that the presence of the first ORF of the G gene in the strain 15 (Warwick) suppressed the expression of the G protein, while the first ORF in the strain J3666 did not have any significant effect on G protein expression.



## Abbreviations

A	adenine
ATP	adenosine 5' triphosphate
BAL	bronchoalveolar lavage
BRSV	bovine respiratory syncytial virus
BSA	bovine serum albumin
bp	base pairs
C	cytosine
CAT	chloramphenicol acetyl transferase
cDNA	complementary DNA
CPE	cytopathic effect
cRNA	complementary RNA
C-terminus	carboxy-terminus
dATP	2'deoxyadenosine 5' triphosphate
dCTP	2'deoxycytidine 5' triphosphate
dGTP	2'deoxyguanosine 5' triphosphate
DI	defective interfering
DIG	digoxigenin
DNA	deoxyribonucleic acid
dNTPs	mixture of four deoxynucleoside triphosphate
DTT	dithiothreitol
dTTP	2'deoxythymidine 5' triphosphate
EDTA	ethylenediaminetetra-acetic acid
ELISA	enzyme-linked immunosorbent assay
EM	electron microscopy
FCS	foetal calf serum
G	guanine
GAG	glycosaminoglycan
GMEM	Glasgow modified Eagle's medium
GTP	guanosine 5' triphosphate
HA	haemagglutinin
HAE cells	human airway epithelial cells
HEPES	N-2-hydroxyethylpiperazine-N-2-ethanesulphonic acid

hours	hr
HRSV	human respiratory syncytial virus
HRP	horse radish peroxidase
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
IRES	internal ribosome entry site
kb	kilobase pairs
L	litre
mg	milligram
ml	millilitre
minutes	min
moi	multiplicity of infection
MOPS	3-[N-morpholino] propanesulfonic acid
$M_r$	relative molecular mass
mRNA	messenger RNA
NA	neuraminidase
NK cells	Natural killer cell
NDV	Newcastle disease virus
NKA	neurokinin A
NKB	neurokinin B
NKK	neuropeptide K
NKY	neuropeptide- $\Upsilon$
N-terminus	amino-terminus
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pep27	polypeptide of 27 amino acids
pfu	plaque forming unit
pg	picogram
PVDF	polyvinylidene fluoride
PVM	pneumonia virus of mice
RNP	ribonucleoprotein
RSV	respiratory syncytial virus
RT-PCR	reverse transcription polymerase chain reaction

SDS	sodium dodecyl sulphate
SeV	Sendai virus
SP	substance P
SV5	simian virus 5
T	thymine
U	uracil
UTP	uridine 5' triphosphate
vRNA	viral RNA
VSV	vesicular stomatitis virus
$\mu$ l	microlitre

# **CHAPTER 1**

## **INTRODUCTION**

## 1.1 PVM and RSV: An overview and virus taxonomy

Pneumonia virus of mice (PVM) was first purified from apparently healthy mice in 1938 during an attempt to isolate viruses from mice with “various acute non-influenzal diseases of the respiratory tract” (Horsfall & Hahn, 1940). The lung extract of healthy mice, which were used as the control system, caused transmissible pulmonary consolidations in mice intranasally infected with the extract (Horsfall & Hahn, 1940). A wide range of mammalian hosts including humans, primates, rodents and rabbits were reported being seropositive against PVM (Horsfall & Curnen, 1946; Pringle & Eglin, 1986). A recent study showed the presence of a viral agent closely related to PVM capable of producing acute pneumonia in dogs. The canine virus was isolated from nasal and pharyngeal swab specimens and tested using a pool of human respiratory syncytial virus (HRSV) monoclonal antibodies consisting of four monoclonal antibodies against P, N, M2 and F proteins. Testing the canine virus against the antibodies, individually, indicated that the agent reacts with the P and M2 monoclonal antibodies, but not with N and F antibodies. Sequencing of PCR fragments amplified from the N gene of the canine virus indicated 96% - 97% alignment to the N gene of PVM, and a fragment amplified from the L gene was found to be 96% identical to the L gene of PVM (Renshaw *et al.*, 2010).

About 2 decades after the description of PVM, HRSV was isolated in 1956 from a symptomatic chimpanzee with coryza during an outbreak of illness (Chanock *et al.*, 1957). This virus, chimpanzee coryza agent (CCA), was subsequently shown to be a serious human pathogen (Chanock *et al.*, 1957). HRSV is now recognised as the most important viral pathogen of serious respiratory tract infection in paediatrics, and is responsible for both upper and lower respiratory tract infections (Collins *et al.*, 2001). Annually, a large range of patients including infants, neonates, immunocompromised patients, and elderly people are infected with the virus (Collins *et al.*, 2001).

The similarity between the cytopathology of HRSV and Newcastle disease virus (NDV) was the first evidence suggesting relatedness of the two viruses (Kisch *et al.*, 1962). Later, despite the lack of neuraminidase (NA) and hemagglutination (HA) activities in HRSV, the possibility of the classification of HRSV into the “Parainfluenza group” was suggested (Waterson & Hobson, 1962). Based on the virus morphology and the lack of neuraminidase and hemagglutination activities and to distinguish it from

other members of the family *Paramyxoviridae*, HRSV was classified separately in the genus pneumovirus, of the family *Paramyxoviridae* (Venkatesan *et al.*, 1983).

The family *Paramyxoviridae*, one of the largest families of viruses, includes major contagious and infectious viruses causing a wide range of disease in humans and animals. The family *Paramyxoviridae* contains two subfamilies: *Paramyxovirinae* and *Pneumovirinae*. The subfamily *Pneumovirinae* is distinguished from *Paramyxovirinae* based on their morphologically different nucleocapsid. However, the major differences between them are the number of encoded proteins and a different form of attachment glycoprotein (Lamb & Kolakofsky, 2001). In turn, the family *Paramyxoviridae* is related to other families of virus having single component negative sense RNA genomes as part of the order *Mononegavirales*.

The sub-family *Paramyxovirinae* includes the genera avulavirus, henipavirus, respirovirus, rubulavirus, and morbillivirus, and the subfamily *Pneumovirinae* contains the genera pneumovirus and metapneumovirus (Table 1.1). There are three members described in the genus pneumovirus: human respiratory syncytial virus (HRSV), bovine respiratory syncytial virus (BRSV) and murine pneumonia virus (MPV). HRSV is the type member of the genus and has two known subtypes: genotype A and B. MPV, traditionally known as pneumonia virus of mice (PVM), is the murine counterpart of HRSV. Most information known is about HRSV as the type member of the genus. In the description of pneumoviruses below, differences between PVM and HRSV, and also difference between the two HRSV genotypes are indicated where it is required. A summary of the classification of the family *Paramyxoviridae* is given in Table 1.1.

Family	Subfamily	Genus	Type species
Paramyxoviridae	Paramyxovirinae	avulavirus	Newcastle disease virus (NDV)
		Henipavirus	Hendra virus (HeV)
		morbillivirus	measles virus (MeV)
		rubulavirus	mumps virus (MuV)
		respirovirus	Sendai virus (SeV)
	Pneumovirinae	pneumovirus	human respiratory syncytial virus (HRSV)
		metapneumovirus	avian metapneumovirus (AMPV)

Table 1.1 Classification of family *Paramyxoviridae* based on the reclassification of International Committee on the Taxonomy of Viruses (ICTV) in 2009 (ICTV, 2009). The type member of each genus is provided separately under type species column.

HRSV was classified into two distinct subtypes A and B by using a cross protection experiment with related glycoproteins (Mufson *et al.*, 1985). Mufson and colleagues (1985) used a collection of monoclonal antibodies against the G, F, M, N and P proteins of RSV strain Long and examined seven different viruses isolated at different times. The main difference was observed in the G glycoproteins in which the two subtypes shared one of the six epitopes that was tested. In the same study differences in the epitopes of the F, M and N proteins were reported: one of the two epitopes examined in the F glycoprotein, two of six epitopes in the M protein, and one of six in the N protein were different between subtypes A and B (Mufson *et al.*, 1985). In another study 26 different HRSV strains were classified into three groups which were defined as common, more common and rare strains (Anderson *et al.*, 1985).

Soon after isolation of bovine respiratory syncytial virus (BRSV) the serologic relatedness between HRSV and BRSV was shown which suggested a taxonomical relationship between the two viruses (Doggett *et al.*, 1968; Paccaud & Jacquier, 1970). Subsequently, PVM was classified as another member of the genus mainly because of the morphological similarities with human and bovine RS viruses (Berthiaume *et al.*, 1974).

### 1.1.1 General features of the virion

Both PVM and HRSV are enveloped viruses surrounded by the lipid bilayer which is derived from the host cell plasma membrane. The fusion (F) and attachment (G) glycoproteins project from the surface of the viral envelope. A small hydrophobic (SH) protein is the third viral surface protein. However, due to its hydrophobic nature and small size the SH protein is not readily detectable on the virus surface. The matrix (M) protein forms a protein layer between the envelope and the nucleocapsid. The nucleocapsid consists of the virus genomic RNA (vRNA) tightly packaged into the nucleocapsid protein and surrounded by the RNP complex. The virus structure is depicted in Figure 1.1 and different parts of the virus are annotated.

The virion of HRSV and PVM show a pleomorphic appearance with sizes ranging from 100 to 350 nm. The filamentous forms range from 60 to 200 nm in diameter and a length of up to 10  $\mu\text{m}$  (Bachi & Howe, 1973; Lamb & Kolakofsky, 2001; Norrby *et al.*, 1970). PVM particles are 100-120 nm in diameter with a length of 2-3  $\mu\text{m}$  in the filamentous forms, and 80-120 nm in diameter in the spherical morphology (Compans *et al.*, 1967; Gallaspy *et al.*, 1978).

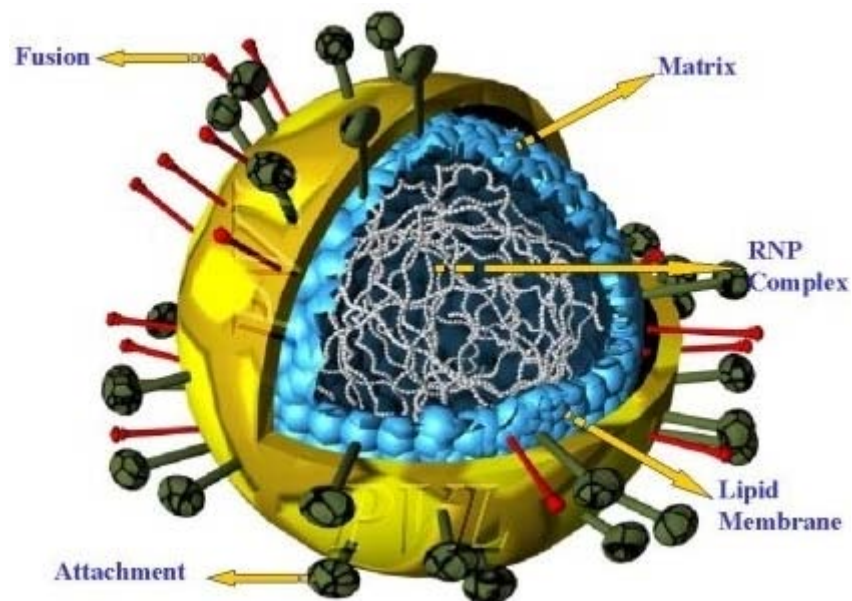


Figure 1.1 pneumovirus structure. The attachment (G) and fusion (F) glycoproteins, lipid membrane, the matrix (M) protein, and ribonucleoprotein complex (RNP complex) are indicated.

As described above, no HA or NA activity has been reported for HRSV (Richman *et al.*, 1971). However, this characteristic is not a general feature of



pneumoviruses; Ling and Pringle (1989b) using a hemagglutination inhibition experiment showed that G glycoprotein of PVM has intrinsic hemagglutination effect on murine red blood cells. Ling and Pringle (1989b) also reported that a monoclonal antibody directed against the G glycoprotein of PVM strain 15 (Warwick) strongly inhibited PVM from agglutinating murine erythrocytes.

The lack of a suitable animal model for HRSV has hampered the study of different aspects in the pathogenesis of the virus. Occasionally BRSV, the bovine counterpart of HRSV, has been used to study the pathogenesis of pneumoviruses. PVM, has also been proposed as a good candidate to investigate unknown aspects of pneumovirus-host interaction.

## **1.2 Molecular biology and replication cycle of pneumoviruses**

The replication cycle of pneumoviruses starts by binding of the virus at the cellular surface components, which is mediated by the G and/or F glycoproteins (section 1.2.1). Virus entry to the cell happens after the attachment and is mediated by the F glycoprotein which initiates fusion between the viral envelope and the cellular membrane. Following the fusion process, the nucleocapsid complex containing the viral genome is released into the cytoplasm (section 1.2.1.3). The viral RNA (vRNA) is used as a template in the RNA synthesis process involving transcription and replication. The transcription process involves synthesis of messenger RNA (mRNA) and the replication cycle involves the synthesis of a complimentary copy of vRNA (cRNA) which acts as the antigenomic RNA. The cRNA is then transcribed to the vRNA by the viral RNP complex (Section 1.2.4). In contrast to Sendai virus, PVM has been shown to be capable of growth in enucleated BS-C-1 cells indicating that nuclear functions provided by the cell are not necessary for the virus replication (Cash *et al.*, 1979). The HRSV and PVM genomes comprise 10 genes separated by intragenic sequences.

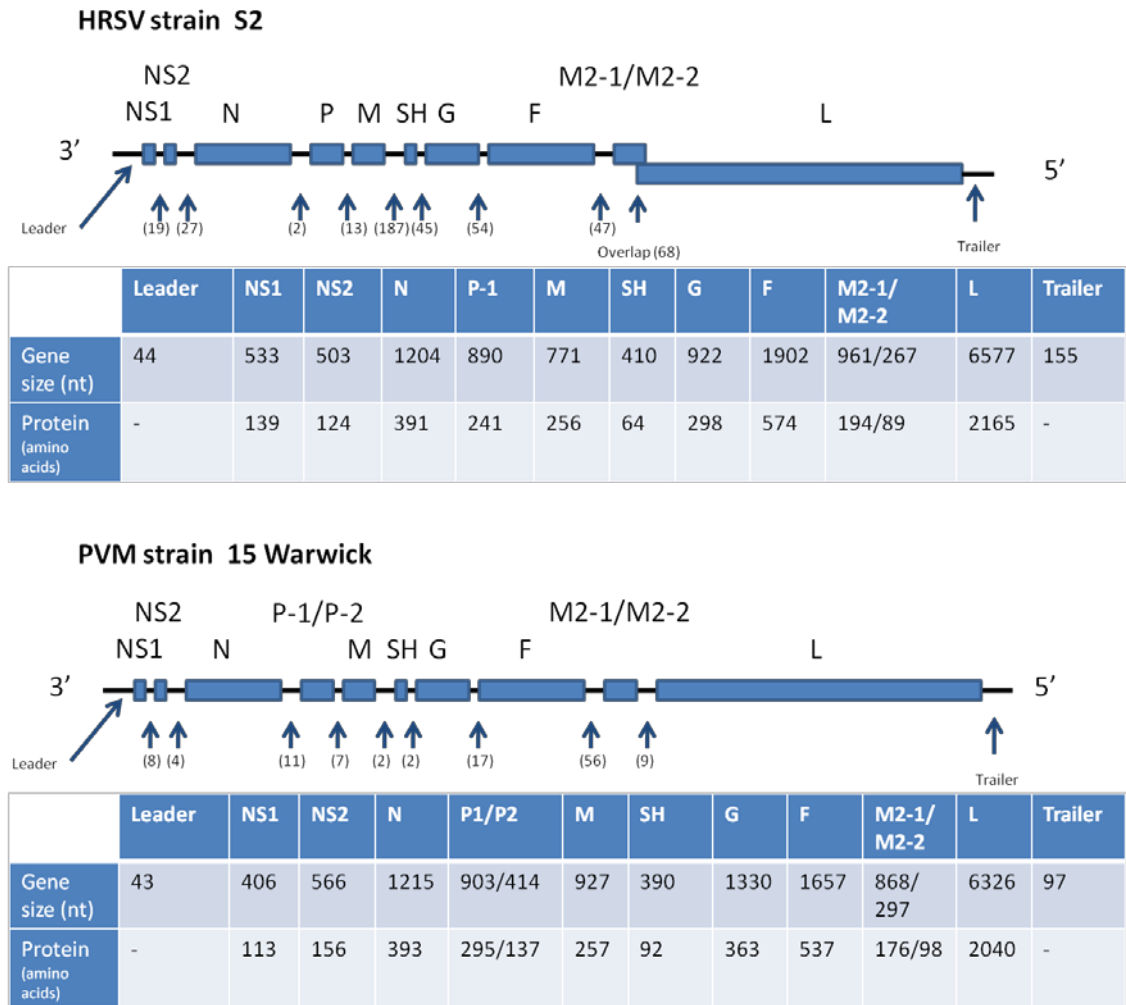


Figure 1.2 Comparison between RSV and PVM vRNA. There are 10 genes across the genome responsible for coding 11 or 12 known proteins respectively. The name of each gene is indicated. Numbers in brackets refer to the number of nucleotides in the intergenic regions. The size of P2 and M2-2 genes and proteins are provided and separated from the number representing the size of P1 and M2-1 by “/”.

During virus replication, at least 11 proteins are synthesised and processed: these are the viral structural proteins (M, SH, F, and G), nucleocapsid complex proteins (N, L, P and M2-1 proteins), and non-structural proteins (NS-1, NS-2, and M2-2). The replication cycle of pneumoviruses is completed by the assembly and release of the virus (Section 1.2.5.1). A schematic structure of the vRNA of PVM strain 15 Warwick and HRSV strain S2 is depicted in Figure 1.2.

### 1.2.1 Attachment and entry

HRSV uses glycosaminoglycan (GAG) proteins for attachment to the host cells with HRSV entry estimated to have a half life of 30 min (Techarpornkul *et al.*, 2002).

There is no report on the entry rates of other members of the *Paramyxoviridae*. The process of attachment of HRSV to host cells is not clearly understood. Two major glycoproteins of HRSV, glycoprotein G (the large glycoprotein) and glycoprotein F, have been shown to be responsible for the attachment and virus docking on the host cells.

The G glycoprotein of HRSV was shown to be responsible for virus attachment (Levine *et al.*, 1987). Subsequently, it was shown that monoclonal antibodies directed against the HRSV G glycoprotein are capable of preventing virus entry (Martinez & Melero, 1998). The observation that viruses lacking the gene encoding the G protein are able to infect *in vivo* and *in vitro* has raised question about the role of the G protein in virus attachment (Karron *et al.*, 1997; Martinez & Melero, 1998). It has been proposed that the F protein may be able to attach to the cell surface through an unknown mechanism.

#### **1.2.1.1 G glycoprotein**

The G glycoprotein is a type II glycoprotein with a hydrophobic region near its N terminus. The G glycoprotein is smaller than, its counterpart, the HN glycoprotein, among the members of *Paramyxovirinae*, and its length varies from 282 to 319 amino acids among different strains of the HRSV (Lamb & Kolakofsky, 2001). No cellular receptor has been identified for the G glycoprotein, though there is evidence that cell surface GAG are important in the virus binding and infection (Feldman *et al.*, 1999). The G glycoprotein appears not to be an essential glycoprotein for virus growth. HRSV lacking the G glycoprotein can grow efficiently in Vero cells but their growth was impaired in HEp-2 cell lines (Teng *et al.*, 2001).

Maturation of the G glycoprotein occurs in the endoplasmic reticulum and Golgi compartments of the cell. The G glycoprotein produces homo-oligomers in the endoplasmic reticulum and *O*-glycosylation follows after the oligomerisation in the trans-Golgi compartment (Collins & Mottet, 1992).

The G protein of HRSV has an  $M_r$  of 36000 in its non-glycosylated state. Its glycosylated form is reported to have an  $M_r$  of 84000-90000 under reducing SDS-PAGE electrophoresis conditions. Recently a new form of the G glycoprotein of HRSV grown in human airway epithelium (HAE) reported, and described as a dimer of the  $M_r$  90000 form or a highly glycosylated form of the G protein with extensive *O*- linked glycosylation (Figure 1.3.A) (Kwilas *et al.*, 2009). About 58% of the mass of the mature

G protein is occupied by *O*- and *N*-linked carbohydrates (55% *O*-linked and 3% *N*-linked). The external portion of the G protein contains between 77 and 91 potential sites for the attachment of *O*-linked oligosaccharides depending on the strain. There are four sites reported for the attachment of *N*-linked glycoproteins (Lambert, 1988; Teng & Collins, 2002; Wertz *et al.*, 1985). The number of potential *N*-glycosylation sites varies between different strains: four *N*-glycosylation sites were reported for strain A2 whereas there are eight *N*-glycosylation sites reported for the Long strain (Teng & Collins, 2002). The *O*- and *N*-glycosylation sites make two distinct domains in the structure of G glycoprotein separated by a short central domain which is highly conserved among different HRSV strains.

The central region of the HRSV G protein overlaps a stretch of four cysteine moieties, clustered in a 13-residue stretch (from amino acids 164 to 176) positioned close to each other which makes disulfide bonds between each other in a 1:4 and 2:3 organisation (Teng & Collins, 2002). This generates a cysteine noose in the G glycoprotein structure. The fact that this domain is highly conserved among different strains of HRSV has led to the suggestion that it is a potential receptor binding domain of the virus. However, a recombinant HRS virus lacking the cysteine noose in the G protein was capable of efficient growth *in vivo*. The poor growth of the virus lacking G gene *in vivo* suggests the presence of other regions on the G glycoprotein that may play an important function in virus infectivity (Teng & Collins, 2002). The conserved region contains a CX3C chemokine motif at amino acid positions 182–186 (Harcourt *et al.*, 2006; Tripp *et al.*, 2001).

Figure 1.3 shows the schematic structure of the G protein (Figure 1.3.A), and compares G protein of different strains of HRSV with PVM (Figure 1.3.B). As it is depicted in the Figure 1.3.A, the hydrophobic region of the HRSV G protein lies between residues 38 to 66. In the G protein of HRSV strain A2 amino acids 38-63, and in the G protein of PVM strain J3666 amino acids 37-59 serve as both signal sequence and the transmembrane anchor Figure 1.3.B.

A temperature sensitive form of HRSV derived from HRSV strain B1 termed strain *cp52*, lacks both SH and G genes and instead contains a fusion between SH (Section 0) and G genes with 91 nucleotide in length and has five amino acid coding changes in F and L. The HRSV strain *cp52* has been extensively used in HRSV studies to analyse the SH and/or G gene functions in the virus pathogenesis (Karron *et al.*, 1997).

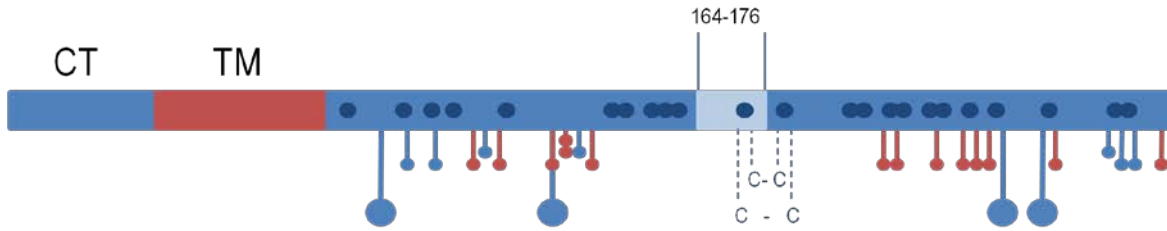
The presence of abundant amounts of proline residues in the primary sequence gives another unique feature to the pneumovirus G glycoproteins. Approximately 30 proline residues (approximately 10% of the total residues) are gathered at the carboxy terminus of the G glycoprotein. The presence of high proportions of serine, threonine, proline residues and extensive *O*-linked carbohydrates are features that the G glycoprotein shares with various mucinous glycoproteins (Wertz *et al.*, 1985).

G1 and G2 glycoproteins were reported as two different forms of the G glycoprotein for PVM strain 15 (Warwick), with G2 derived from G1 (Ling & Pringle, 1989b). The estimated relative molecular mass for G1 and G2 glycoproteins of PVM strain 15 (Warwick) was calculated as 76400 and 62000, respectively. The molecular mass after removal of both *N*- and *O*- linked oligosaccharides were very close to each other. The relative molecular mass for G glycoproteins reported by Ling and Pringle (1989b) is listed in Table 1.2.

	Glycosylated form	N-linked sugars removed	O-linked sugars removed	Removal of both N- and O-linked sugars
G1	76400	58400	57600	39600
G2	62000	48200	51000-58600	37200-44800

Table 1.2 The relative molecular mass of G glycoproteins of PVM strain 15 (Warwick) adapted from (Ling & Pringle, 1989b).

A



B

PVMJ3666	MGRNFEVSGSIT-NLNFERTQHPDTFRTGVKVNQMCKLIAGVLTSAAVAVCVGVIMYSVF
PVM15 W	-----MCKLIAGVLTSAAVAVCVGVIMYSVF
PVM15 ATCC	MGRNLEVSGSIT-NLNFERTQHPDTFRTVVKVNQMCKLIAGVLTSAAVAVCVGVIMYSVF
HRSVA2	MSKNKDQRTAKTLERTWDTLNHLLFISSCLYKLNLSVAQITLSILAMIISTSLIIAATI
HRSVB1	MSKHKNQRTARTLEKTWDTLNHLIVISSCLYRLNLSIAQIALSVLAMIISTSLIIAATI
	: .: .*: *: :...*: ::
PVMJ3666	---TSNHKANST---QNATTRNSTSTP---PQPTAGLPTTEQG-----TIPRFTKP
PVM15 W	---TSNHKANST---QNATTRNSTSTP---PQPTAGLPTTEQG-----TIPRFTKP
PVM15 ATCC	---TSNHKANST---QNATTRNSTSTP---PQPTAGLPTTEQG-----TIPRFTKP
HRSVA2	FIASANHKVTPPTAIIQDATSQIKNTTPTYLTQNPQLGISPSNPSEITSQITTTILASTTP
HRSVB1	FIISANHKVTLTTVTVQTIKNHTEKNITTYLTQVPPERVSSSKQPTTTSPIHNSATTSP
	::***. . * * . . . . . * :. . . . . * * *
PVMJ3666	PTKTATHHEITEPVKMATPSEDPYQCSSNGYLDRPDLPENFKLVLDVICKPPGPEHHNTS
PVM15 W	PTKTATHHEITEPVKMATPSEDPYQCSSNGYLDRPDLPENFKLVLDVICKPPGPEHHNTS
PVM15 ATCC	PTKTATHHEITEPVKMATPSEDPYQCSSNGYLDRPDLPENFKLVLDVICKPPGPEHHNTS
HRSVA2	GVKSTLQSTTVKTKNTTTTQTQPSKPTTKQRQNKPP-----SKPNDFH---
HRSVB1	NTKSETHHTTAQTKGRTTTTSTQTNKPSTKPRLKNPP-----KKPKDDYH---
	.*: : .: . :*. . . : : : . . * . * :.*
PVMJ3666	CYEKREINPGSVCPDLVTMKANMGLNNGGGEDAAPYIEVTTLSTYSNKRAMCVHNGCDQG
PVM15 W	CYEKREINPGSVCPDLVTMKANMGLNNGGGEDAAPYIEVTTLSTYSNKRAMCVHNGCDQG
PVM15 ATCC	CYEKREINPGSVCPDLVTMKANMGLNNGGGEDAAPYIEVTTLSTYSNKRAMCVHNGCDQG
HRSVA2	-FEVFNFPVCSI C SNNPT C WAI C KRI PNKKPGKTTTKPTKKPT-----
HRSVB1	-FEVFNFPVCSI C GNNQL C KSI C KTI PSNKPKKPTIKPTNKPT-----
	:* : : * * : * : . : . . . : * . *
PVMJ3666	FCFFLSGLSTDQERAVLELGGQQAIMELHYDSYWKHYWSNSNCVVPRTNCNLTDQTEILF
PVM15 W	FCFFLSGLSTDQERAVLELGGQQAIMELHYDSYWKHYWSNSNCVVPRTNCNLTDQTEILF
PVM15 ATCC	FCFFLSGLSTDQERAVLELGGQQAIMELHYDSYWKHYWSNSNCVVPRTNCNLTDQTEILF
HRSVA2	-----LKT-T-KKD-----PKPQTTSKEV----
HRSVB1	-----TKT-TNKR-----PKTPAKTTKE-----
	. * : . . . . . * . . . . . *
PVMJ3666	PRFNNKNQSQCTTCADSAGLDNKFYLTCDGLLRTLPLVGLPSLSPQAYKAVPTQTGTGTTT
PVM15 W	PRFNNKNQSECTTCADSAGLDNKFYLTCDGLLRTLPLVGLPSLSPQAYKAVPTQTGTGTTT
PVM15 ATCC	PRFNNKNQSQCTTCADSAGLDNKFYLTCDGLLRTLPLVGLPSLSPQAYKAVPTQTGTGTTT
HRSVA2	PTTKPTEEPTINTTKTNIITLLTSNTTGNPELTSQMETFHSTSSSEG-NPSPSQVSTTSE
HRSVB1	TTTNPTKKPTLTTTERDSTTSQSTVLDTTTTLEHTIQQSLHSTTPEN-TPNSTQPTASE
	. : . . . . * . . . . . * : * : : . . . . * . . . .
PVMJ3666	APTSESRHPTPAPRRSKPLSRKKRALCGVDSGREPKPTMPYWCPLQLFPRRSNS
PVM15 W	APTSETRHPTPAPRRSKPLSRKKRALCGVDSGREPKPTMPYWCPLQLFPRRSNS
PVM15 ATCC	APTSETRHPTPAPRRSKPLSRKKRALCGVDSGREPKPTMPYWCPLQLFPRRSNS
HRSVA2	YPSQPSSPPN-TPRQ-----
HRSVB1	-PSTSNSTQN-TQSHA-----
	*: . . . : .

Figure 1.3.A Schematic structure of the G glycoprotein of HRSV. The structure of the G glycoprotein of HRSV, based on the description by Collins and colleagues (2001). The cytoplasmic tail (CT) and transmembrane domain (TM) are shown with blue and red colours, respectively. The presence of proline residues is shown by small dark blue circles. The O-glycosylation (blue spikes) and N-glycosylation (red spikes) sites are indicated. The conserved area is indicated between amino acid residues 164 and 176. The cysteine noose is shown by dotted lines and the di-sulphide binds are shown between cysteine molecules. B. Comparison of G protein sequences among pneumoviruses. HRSV strain A2 and B1 (as indicated) are compared with PVM strains (J3666, Warwick –shown as PVM 15 W, and 15 – shown as PVM 15 ATCC). The predicted transmembrane domain across the molecule is indicated in red. The conserved area in the G protein of HRSV is indicated in cyan. The cysteine residues responsible for the cysteine noose are indicated in yellow. The presence of a potential furin cleavage site (Arg-X-(Lys/Arg)-Arg) in the G protein of PVM strains is indicated with green background. The “\*” character indicates positions which have a single, fully conserved residue. The “:” character represents strongly and “.” characters indicates weakly conserved residues (Thompson *et al.*, 2002). The prediction for the transmembrane domains was made using the HMMTOP server (Tusnady & Simon, 2001). The schematic structure of the G glycoprotein (A) and the sequences (B) are presented from N terminus to C terminus.

```

PVM_15_W      3' UCCUAUUCAUGAUAGGAUAACCUUGGUUUGCUCUGGACAUCUCGUCGAGUGUGUUCUCU
PVM_J3666    3' UCCUAUUCAUGAUAGGAUAACCUUAGUUUACUCUGGACAUCUCGUCGAGUAUGUUCUCU
*****

PVM_15_W      GGUGUUCGACUGAAGUGGAUCAUACCCUCCUUGAAUCUUCACUCACCGUCGUA AUGGU
PVM_J3666    GAUGAUC AACUGAAGUGAAUCAUACCCUCCUUGAAUCUUCACUCACCGUCGUA AUGGU
* * * * *

PVM_15_W      AAACUUGAAACUCUCUUGAGUCGUAGGACUGUGUAAAUCCUGACAACAUUUUU CACUUGG
PVM_J3666    AAACUUGAAACUCUCUUGAGUCGUAGGACUGUGUAAAUCCUGACCACA-UUUU CACUUGG
*****

PVM_15_W      UUUACACAUU CGAAUAACGUCCACACGAGUGUUCACGACGACACCGUCA AACACACCCCC 5'
PVM_J3666    UUUACACAUUCGAAUAACGUCCACACGAGUGUUCACGACGACACCGUCA AACACACCCCC 5'
*****

```

Figure 1.4 Comparison of the nucleotide sequence (genomic sense) of the G gene of PVM strain 15 (Warwick) and PVM strain J3666. A uridine insertion in the sequence of PVM strain 15 (Warwick) changes the open reading frame for the PVM strain 15 and causes a premature stop in protein synthesis. In the strain J3666, the open reading frame starts from the same place of the strain 15 (Warwick) and protein synthesis continues to the main stop codon of the ORF at position 1273 of the gene (not shown). PVM strain J3666 and PVM strain 15 (Warwick) are shown as PVM\_J3666 and PVM\_15\_W respectively. Start codons are shown in green and stop codons are indicated in red. The mutation site is shown in yellow. The “\*” character indicates positions which have a single, fully conserved residue.

### 1.2.1.1.1 Second open reading frame in the PVM G gene

Analysis of the sequence of the G gene of the non-pathogenic strain 15 (Warwick) and the pathogenic strain J3666 showed that the G protein was encoded by the second ORF in the mRNA. This was also seen with the HRSV G gene. In strain J3666 G mRNA a small (37nt) ORF is located upstream of the major G protein coding

ORF. The strain J3666 G protein contains a cytoplasmic domain followed by a putative transmembrane domain. In contrast, the strain 15 (Warwick) G mRNA contains a first ORF which overlaps the G protein – coding major ORF. The strain 15 (Warwick) G protein does not contain a significant intracellular domain preceding the transmembrane region (Randhawa *et al.*, 1995). The difference in gene organisation in the G protein coding ORFs of the two viruses is the result of the presence of an additional nucleotide in a string of U residues (4 in strain J3666 and 5 in strain 15 G mRNA) which results in a frame shift (see Figure 1.4). The organisation of the G gene of both viruses is shown diagrammatically in Figure 1.5.

A short open reading frame (ORF) of 15 codons was described overlapping the start of major ORF of the HRSV G gene (Wertz *et al.*, 1985). The function of the ORF or the function of its possible product has not been clearly described.

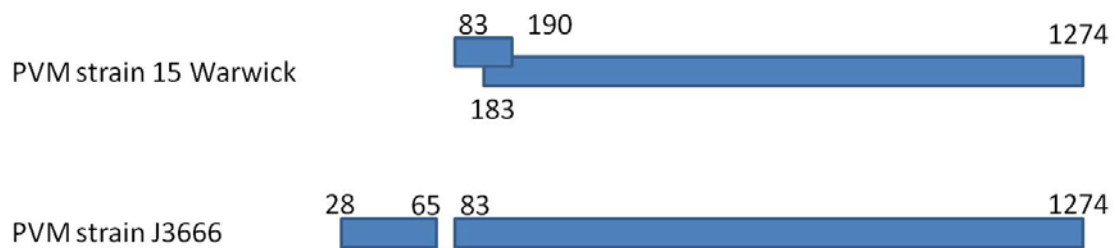


Figure 1.5 Comparison of the G mRNA organisation in PVM strain 15 Warwick and PVM strain J3666. There is a second open reading frame in PVM strain 15 Warwick overlapping the main open reading frame. In PVM strain J3666, first ORF is upstream and separate from the second ORF (Randhawa *et al.*, 1995). The numbers refer to the nucleotide position in the G gene mRNA. The G gene is presented from 5' to 3'.

#### 1.2.1.1.2 Proteolytic processing of G glycoprotein

It was first reported by Spring and Toplin (1983) that the “75K glycoprotein” (G glycoprotein) of HRSV is susceptible to tryptic digestion *in vitro* producing two fragments with  $M_r$  value of 40,000 and 29,000. However, the authors reported that digestion of a pool of HRS virus with trypsin, chymotrypsin, or elastase does not have any effect on the virus infectivity *in vitro* (Spring & Tolpin, 1983). Recently, it was reported that Vero grown viruses predominantly contain a form of G glycoprotein with  $M_r$  of 55,000 which lacks its C terminus whereas viruses grown in HEp-2 cells possessed the G glycoprotein with a relative molecular mass of 90,000 (Kwilas *et al.*, 2009). In the same report, it was shown that the virus grown in Vero cells is less infectious *in vitro*. The data indicated a vulnerability of the G glycoprotein to be



processed post-translationally, and the virus grown in cells capable of producing truncated G glycoprotein are less infectious and less dependent on GAG. The pathogenicity of viruses with the G glycoprotein lacking its C terminal has remained unclear.

A soluble form of the HRSV G glycoprotein (Gs) is formed by initiation of translation from an alternative start codon in the G mRNA. This produces a protein 47 residues shorter than the full-length G protein and this is processed further by proteolytic enzymes which digest at amino acid residue at positions 66 and/or 75 (Hendricks *et al.*, 1988; Roberts *et al.*, 1994). Roberts and colleagues (1994) reported a further processing possibility generating a third form which is also secreted. The precise role of the Gs form is unclear but it has been implicated in some aspects of pathogenicity.

Ling and Pringle (1989b) reported a shorter form of the G glycoprotein in cells infected with PVM. Krempf and colleagues (2007) reported the presence of the same form of the G glycoprotein in infected cells. In both reports no secreted form of PVM G glycoprotein was detected. Thus, the shorter form of the G glycoprotein of PVM is thought to be not the counter part of the secreted G glycoprotein produced in cells infected with HRSV.

#### **1.2.1.2 Fusion (F) glycoprotein:**

The fusion (F) protein of RSV, similar to the fusion proteins of other members of the family *Paramyxoviridae*, is a type I transmembrane surface protein which mediates membrane and envelope fusion. The HRSV F protein is synthesised as a large precursor polypeptide F0 with a length of 574 amino acids, which is activated by cleavage at two separate sites by a furin-like cellular endoprotease to produce two glycoproteins (F1 and F2) which are linked together by disulfide bonds (Johnson & Collins, 1988).

The first cleavage site in the HRSV F protein is at residue 109 and consists of an arginine-rich sequence (RARR). The second cleavage site is at residue 136 and consists of a lysine and arginine-rich sequence (KKRKRR), with a 27 amino acid gap between the two sites. The second site at position 136 is equivalent to the cleavage site found in the other F proteins of members of the *Paramyxoviridae* family (Zimmer *et al.*, 2001). A polypeptide of 27 amino acids in length (pep27) is released upon the proteolytic cleavage. The exact function for the polypeptide is not known. However, it has been

reported that the BRSV pep27 protein is related by sequence and function to tachykinins (1.3.2.1) (Zimmer *et al.*, 2003). There are three N-glycosylation sites in pep27 and it has been shown that mutation in at least one of the sites increases syncytia formation suggesting that differential glycosylation of pep27 may modulate fusion (Rawling *et al.*, 2008).

The F proteins in *Paramyxoviridae* have three main hydrophobic domains: one appears in the N terminal of the F0 and acts as the signal peptide for translocation to the endoplasmic reticulum, the second is the fusion peptide and is close to the C-terminus of the F1 region, and the third is located at the N-terminus of the F2 region. The HRSV F protein is capable of attaching to cells and producing syncytia independently without the presence of the G glycoproteins. This is in contrast with the situation for the Paramyxoviruses where the presence of attachment HN protein is essential for docking of the virus (Rawling *et al.*, 2008).

Similar to the F protein of RS viruses, the F protein of PVM is a glycosylated protein which is activated by cleavage into F1 and F2 polypeptides. The presence of both F1 and F2 proteins in infected cells was described by Ling and Pringle (1989b). The results identified an F(1,2) glycoprotein with a mobility slower than that of F1 and two different forms of F1 with different motilities ( $M_r$  value of 20000 and 12000). This led to a proposal of two possible structures for the PVM F protein: the first one is the common F structure like the other *Paramyxoviridae* members in which F0 is processed and cleaved to F1 and F2 fragments which happens rapidly for the PVM F protein. The F2 protein was proposed to be very small, with  $M_r$  value of 5000 (Ling & Pringle, 1989b).

### **1.2.1.3 Uncoating**

The uncoating process involves the release of the ribonucleoprotein complex (RNP) from the virus particle into the cytoplasm. This involves removal of the M protein layer from the RNP complex. Asenjo and colleagues (2008) have demonstrated that the liberation of M protein from the RNP needs phosphorylation of the P protein at the serine residue at position 54. Administration of LiCl inhibits the cellular kinase responsible for the phosphorylation and causes accumulation of P protein but not M protein in cytoplasmic inclusion bodies which suggests involvement of M protein with the RNP.

## 1.2.2 Genome structure and organisation

A clear description of HRSV genome, its polarity and the number of mRNA encoded from the genome was reported by Huang and Wertz (1982). Figure 1.2 compares genome structure of HRSV and PVM. HRSV possesses a monopartite single stranded negative sense genome with a length varying from 15140 to 15225 nucleotides (in HRSV strain B1). In PVM the genome length ranges from 14885 to 14887 depending on the strain. Table 1.3 provides the genome length for different members of the pneumovirus genus.

Virus	Strain	Nucleotide length	Gene bank accession number
PVM	J3666	14885	NC_006579
		14885	AY743909
PVM	15 (Warwick)	14887	AY743910
PVM	15 (ATCC)	14886	AY729016
HRSV	A2	15222	M74568
HRSV	Line 19	15191	FJ614813
HRSV	S2	15190	U39662
HRSV	B1	15225	NC_001781
HRSV	<i>cp52</i>	13933	AF013255
BRSV	A51908	15140	NC_001989

Table 1.3 Comparison between HRSV, BRSV and PVM genome length. The nucleotide length and the relevant accession number in the gene bank are provided for each of the virus species.

Unlike the situation with other members of the order *Mononegavirales*, the genome of pneumoviruses do not follow the rule of six which requires the number of nucleotide of the genome to be divisible by number 6 (Samal & Collins, 1996). The vRNA in HRSV contains a complementary structure in both the 3' and 5' ends of the molecule (Mink *et al.*, 1991). The leader region (Le), a 44 nucleotide region at the 3' end of the genome, controls transcription, whereas replication is under the control of the trailer complement region (TrC), a 155 nucleotide region at the 5' end of the genome.

Both trailer and leader sequences are highly conserved among different members of pneumoviruses. The leader and trailer regions are complementary and show 81% homology. However, despite the presence of the complementary structures, it was reported that the HRSV genome has a linear structure (Cowton *et al.*, 2006). Among the conserved nucleotides the first 10 nucleotides are identical and nucleotides 11-26 share 81% of similarity. Complete transcription of the vRNA results in production of complementary RNA (cRNA) which acts as the antigenome to mediate the genome replication (see 1.2.5).

The conserved nucleotides in the leader region are responsible for the efficiency of replication and transcription of the genome, and provide the possibility of balancing between transcription and replication. In the 3' leader region the nucleotide at position 4 has been shown to be particularly important in determining the frequency of replication compared to transcription. The presence of a G residue at this position yielded higher levels of transcription while in contrast the presence of an A residue reduced both transcription and replication to 25% of wildtype levels (Fearn's *et al.*, 2002). This and additional mutational analysis data suggest that the leader region is a critical *cis*-acting element which determines the rates of transcription and replication. Generally, the role(s) of the nucleotides in the leader region can be divided into three groups: 1. residues important for RNA replication but not transcription (1U, 2G, 6U and 7U); 2. residues important for both transcription and replication (3C, 5C, 8U, 9U, 10U and 11U); 3. residues less important for transcription and replication and which tolerate alteration (residues 12-26).

In the genome, each gene is flanked by a gene start and gene end sequence and a highly variable intragenic region separates genes from each other. The sequence of both gene start and gene end are not conserved among viral species, but the gene start is highly conserved and gene end is semi-conserved within each virus genome. Collins and Wertz (1985b) have mapped and reported the gene start sequences in HRSV as 5'-GGGGCAAU-3' and the semi-consensus gene end sequence as 5'-AGU(U/A)A(N)<sub>1-4</sub>-poly(A)-3' in the complementary RNA sense, and 3'-UCA(A/U)U(N)<sub>1-4</sub>-poly(U)-5' in the genomic RNA sense. Table 1.5 compares the gene end sequence of HRSV strain A2. Transcription efficiency (Section 1.2.4.3) is controlled by the nucleotide arrangement of the sequences (Harmon *et al.*, 2001). In PVM the gene start sequence is more variable in comparison with other members of pneumoviruses (Table 1.4) (Dibben & Easton, 2007).

HRSV genes are transcribed to produce 10 mRNA molecules responsible for synthesis of the structural (SH, G, F, M, N, P and L) and non-structural (NS1, NS2, M2-1, and M2-2) proteins.

Gene name	Gene sequence
NS1	AGGACAAGU
NS2	AGGACAAGU
M	AGGACAAAU
N	AGGAUAAAU
P	AGGAUAAAU
SH	AGGAUAAAU
G	AGGAUAAGU
F	AGGACAAAU
M2	AGGAUGAGU
L	AGGAUCAAU
	**** * *

Table 1.4 Alignment of the sequences of gene starts in PVM. Sequences are provided in their genomic order. The consensus nucleotides are indicated by a “\*” sign (Thompson *et al.*, 2002).

Gene name and sequence		Termination efficiency
NS1	UCAAUUAUUAUUUUG	+
NS2	UCAUUAAAUUUUUAA	+
N	UCAAUUAUUUUUUUA	++
P	UCAAUGUUUUUUUC	++
M	UCAAUUAUUUUUUUA	++
SH	UCAAUUAUUUUUUUA	+++
G	UCAAUGAAUUUUUG	++
F	UCAAUAUAUUUUUGU	+
M2	UCAAUAAAUUUUUCC	+
L	UCAAUAAAUUUUUAA	+
	*** * ***	

Table 1.5 Comparison of gene end sequence in HRSV strain A2. Termination efficiency is shown with +, ++, and +++ signs indicating inefficient (15 to 40%), efficient (65 to 80%), and highly efficient (95%) respectively (Harmon *et al.*, 2001). The consensus nucleotides are indicated by a “\*” sign (Thompson *et al.*, 2002).

### 1.2.3 RNA dependent RNA polymerase complex

Transcription and replication in the subfamily *Pneumovirinae* is mediated and controlled by viral proteins (N, P, L and M2-1) and *cis*-acting genomic factors. Although the HRSV N, P, and L proteins are capable of directing transcription and replication functions, the presence of the M2-1 protein is essential to obtain fully processed transcription. Transcription requires interplay of the N, P, L and M2-1 proteins. It has been suggested that replication does not require the activity of the M2-1,

but it requires the M2-2 protein for switching from transcription to replication (Melero, 2006).

### 1.2.3.1 Nucleocapsid (N) protein of pneumoviruses

The nucleocapsid protein encoded by HRSV is 391 amino acids in length. The N protein, similar in function to the N protein of other members of the family, makes stable complexes with the vRNA to form the nucleocapsid structure, and binds to the replication intermediate cRNA to protect them against RNase activity and possibly to prevent formation of secondary RNA structures. The N protein is considered the major component of the virus particle. The N protein interacts with the vRNA in association with the virus P protein, and it appears that the fully phosphorylated form of the P protein is essential for obtaining a stable P:N complex (Castagne *et al.*, 2004). It has been shown that when the N protein is expressed in eucaryotic or procaryotic hosts, it binds to the cellular RNA structures non-specifically and forms a nucleocapsid-like structure which is visible using electron microscopy. Moreover, the N protein is capable of self binding to produce ring-like structures (Murphy *et al.*, 2003).

The N protein in *Paramyxovirinae* has a modular structure (Karlin *et al.*, 2003). Sequence similarity between the N protein of PVM and the N protein of *Paramyxovirinae* suggests structural similarities between N proteins (Barr *et al.*, 1991). Two main domains were defined for the N protein: N<sub>core</sub> and C<sub>core</sub>. The N<sub>core</sub> occupies about two thirds of the amino-terminus of the N protein and contains in HRSV the region necessary for production of N:RNA and N:N structures (Murphy *et al.*, 2003; Murray *et al.*, 2001). The C<sub>core</sub> is essential for interaction with the P protein (Karlin *et al.*, 2003). Figure 1.6 shows domains of the N protein of HRSV which are important in P and N interaction and the RNA binding domains.

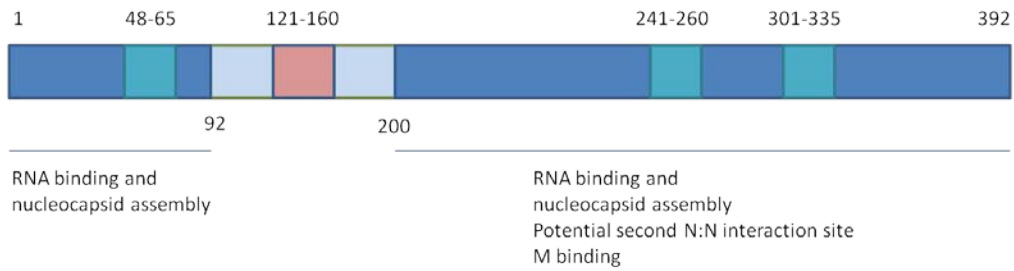


Figure 1.6 The HRSV N protein structure. The potential P binding sites amino acids 48 to 65, 241 to 260 and 301 to 335 are shown with turquoise colour. The region with a potential effect on helical stability is shown in pink colour (amino acids 121 to 160). The predicted functions of N terminus (amino acids 1 to 92) and C terminus ends of the protein (amino acids 200 to 392) are identified. Adapted from Murphy *et al.*, (2003).

There is a high level of similarities (60%) in amino acid sequence between the N protein of PVM and its HRSV counterpart. The C terminus (residues 245 to 315) is highly conserved (96% identity) suggesting a conserved function of this domain (Barr *et al.*, 1991).

During HRSV infection, inclusion bodies are formed in the cytoplasm of infected cells. There are believed to be the result of aggregations of the P and N proteins. The presence of RNA (both genomic and antigenomic polarities) has also been shown in these inclusion bodies (Garcia-Barreno *et al.*, 1996; Garcia *et al.*, 1993)

### 1.2.3.2 Phosphoprotein (P) protein of pneumoviruses

The phosphoprotein of HRSV with 241 amino acids is highly conserved between strains. Using a panel of mutations and a minigenome system, the function of the P protein of BRSV has been analysed thoroughly (Khattar *et al.*, 2001). The P protein is highly structured and regulates the transcription and/or replication process. Amino acids 41 to 60 of the P protein appear to negatively regulate the replication process. It has been shown that the removal of a phosphorylation site on the P protein of BRSV (232S) appears not to have any effect on transcription (Khattar *et al.*, 2001), confirming previous reports (Villanueva *et al.*, 2000). Barik and colleagues (1995) had previously suggested that phosphorylation of P protein at position 232 was essential for viral transcription and virus assembly. More recently, phosphorylation of threonine at position 108 has been shown to be responsible for the interaction of the P and M2-1 proteins (Asenjo *et al.*, 2006).

A putative calcium binding domain (229-DESSDNDLSLEDF-241) has been proposed for the P protein of BRSV. The binding of calcium to this domain results in

conformational changes and exposure of hydrophobic residues allowing interaction with target protein(s) (Ikura, 1996; Khattar *et al.*, 2001). Mutations in the calcium binding domain resulted in deficiency in binding to the N protein. It is noteworthy that the mutations did not affect the affinity of P protein to L protein (Khattar *et al.*, 2001).

Based on the immunoprecipitation studies done by Garcia-Barreno and colleagues (1996), the C terminus of the P protein appears to be essential for interaction with the N protein when both proteins are expressed *in vivo* with the first 6 amino acids from the C terminus being the most important in the interaction with the N protein (Garcia-Barreno *et al.*, 1996). Slack and Easton (1998), using a yeast two hybrid system, identified two distinct sites near the C terminus of the P protein which were important in N and P protein interaction. Confirming the importance of the C terminus of the P protein in its interaction with N protein, two point mutations in the C terminus were identified which were responsible for producing temperature sensitivity for the interaction between the P and N proteins (Lu *et al.*, 2002).

In addition to the full length P protein, it has been shown that a protein (P-2) with  $M_r$  of 19,000 is produced from an internal in-frame start codon from the P protein mRNA of HRSV strain RSN-2 *in vitro* (Caravokyri & Pringle, 1992).

The P protein of PVM (295 amino acids) is longer than that of HRSV. The homology between the two proteins is scattered across the molecule but suggests two conserved sites divided by a non-conserved region in the middle of the molecule (Barr *et al.*, 1994). The precipitation of the P protein of HRSV with monoclonal antibody raised against PVM P protein suggested an antigenic relatedness between these two proteins (Ling & Pringle, 1989a).

A second and internal ORF of the P protein in PVM with a product of 137 amino acids was identified (Barr *et al.*, 1994). The function of this protein is unknown. A reverse genetics approach by Dibben and Easton (2008) indicated that the P-2 protein of PVM was not essential for transcription or replication.

### **1.2.3.3 Large polymerase (L) protein of pneumoviruses**

The large protein of HRSV is 2165 amino acids in length, with  $M_r$  of 250,226 for strain A2 (Collins *et al.*, 1987). The L protein is rich in leucine and isoleucine amino acids and in neutral pH has an estimated positive charge of +75 (Melero, 2006). Aligning five sequences from the order *Mononegavirales*, Poch and colleagues (1990) suggested six conserved regions in the L protein. Based on this comparison the main



catalytic domains of L protein were predicted to lie within the conserved blocks which were flanked by non-conserved sequences with the hinge function. The function of each domain was predicted based on the analogy and the biophysical characteristic of conserved amino acids and their surrounding residues, and it was suggested that region III acts as the active site; lysine residues within regions II and IV acts as ribonucleotide binding domains; a highly conserved region inside domain II acts as the template recognition site; and region VI acts as the polyadenylation or protein kinase site. Region III of the L polymerase contains the GDNQ motif which is present in the all DNA and RNA polymerase molecules (Poch *et al.*, 1990; Poch *et al.*, 1989; Stec *et al.*, 1991). It was suggested that a GDN motif is a modified form of the GDD polymerase motif and was suggested as the common ancestor of polymerase proteins (Kamer & Argos, 1984; Stec *et al.*, 1991).

#### **1.2.3.4 M2-1 (22K) protein**

The HRSV M2-1 protein, previously called the 22K protein and a unique protein to the subfamily of *Pneumovirinae*, is 194 amino acids in length and is one of the most important regulatory proteins of the virus affecting control of gene transcription and genome replication (Dibben *et al.*, 2008). When first identified, it was believed the M2-1 protein was a second matrix protein as it was dissociated from N protein in the same situation similar to that of the M protein (Collins & Wertz, 1985a; Huang *et al.*, 1985). The main function of the M2-1 protein was first discovered by Hardy and Wertz (1998), and described as a transcriptional antitermination factor. Using EM analysis, the M2-1 protein was described as a tetramer with a  $M_r$  of 89,000 and with a diameter of 7.6 nm. Using a panel of mutants, the oligomerisation domain was mapped across amino-acid residues 32-63 which form a potential  $\alpha$ -helix structure (Tran *et al.*, 2009). On a one dimensional SDS-PAGE system two different forms of M2-1 protein have been shown, while on the two dimensional SDS-PAGE system more than two forms have been reported (Routledge *et al.*, 1987). More recently, it has been reported that the difference in the mobility of the two forms shown in the one dimensional SDS-PAGE system is due to differences in the phosphorylation of the M2-1 protein with the slower being phosphorylated and the faster form not being phosphorylated (Hardy & Wertz, 2000). The phosphorylation site has been mapped by two groups individually. Using a mass spectrophotometry technique, Cartee and Wertz (2001) identified the

phosphorylation sites as serine 58 and serine 61. In contrast, site directed mutagenesis confirmed the phosphorylation sites as threonine 56 and serine 58 (Cuesta *et al.*, 2000).

The M2-1 protein has a zinc binding domain known as Cys3–His1 motif (C<sub>7</sub>-X<sub>7</sub>-C<sub>15</sub>-X<sub>5</sub>-C<sub>21</sub>-X<sub>3</sub>-H<sub>25</sub>). The cysteine and histidine residues in this motif have been predicted to be important in maintaining the readthrough ability of the RNP complex in a sub-genomic HRSV based system bearing the M-SH gene junction (Hardy & Wertz, 2000). An M2-1 protein mutated at cysteine residues at positions 7 and 15, and histidine residue at position 25 was not able to act as a potent transcription factor in the RNP complex (Hardy & Wertz, 2000).

Amino acid residues 59 to 80 of the M2-1 protein share a similarity of 95% among HRSV and BRSV strains. The similarity of this region in comparison with the region in PVM is only 45%. However, in the PVM M2-1 protein, residues 70 to 80 show a higher level of similarity (90%) with the same region of the M2-1 protein of other members of the genus. The aromatic and hydrophobic nature of the amino acid residues in this region, and the high level of conservation among other strains and species of the genus suggest that this region may be the RNA binding domain of the molecule (Cuesta *et al.*, 2000).

#### **1.2.4 Transcription and gene expression**

The HRSV genome contains a single promoter located in the leader region which controls RNA synthesis (Dickens *et al.*, 1984). For the RNA synthesis in transcription or replication, two different models are hypothesised. In the first model, transcription and replication start at the same site of the leader sequence. Having been transcribed, the RNA faces three possibilities: 1. the synthesis of an immature RNA that dissociates from the template near the leader region, 2. RNA being packaged into the N protein and forms a stable complex which results in the synthesis of cRNA, and 3. RNA associates with the N protein and makes a stable complex, but RNA elongation ends at the leader region and reinitiates at the gene start region. In the second model, it is proposed that transcription and replication start at distinct but overlapping sites of the leader regions. Nucleotides 3C, 5C, 8U, 9U, 10U, and 11U of the leader sequence are the common elements of the two promoters (see Section 1.2.2 for description on the leader sequence). Nucleotides 1U, 2G, 6U, and 7U of the leader sequence are involved in the recognition of the RNP complex for replication complex and the gene start

signals for the transcription initiation (Fearnis *et al.*, 2002). There are other mechanisms proposed for transcription initiation which are reviewed by Cowton *et al.* (2006).

It has been suggested that the transcriptional complex having been bound to the leader region, starts scanning the genome for the gene start sequence and initiates mRNA synthesis from the NS1 gene (Cowton & Fearnis, 2005). The *cis* factor for termination of gene transcription is the gene end sequence (Section 1.2.2) (Harmon *et al.*, 2001). It has been shown that if the polymerase does not recognise the gene end signal, it will result in the production of a polycistronic RNA containing the intergenic junction (Dickens *et al.*, 1984).

One of the fundamental differences between the genomic structure of PVM and HRSV is the overlap of the M2 and L genes seen in HRSV. The L gene shares its 5' sequence with 3' end of M2 gene. In PVM, however, both M2 and L genes form separate and distinctive genes on the genome (Figure 1.1) (Collins *et al.*, 1987).

#### 1.2.4.1 Gene expression gradient

Early reports indicated that after termination in transcription, the RNP complex dissociates from the cRNA templates and a proportion of the RNP complex restart transcription by relocating to the promoter sequence in the 3' end of vRNA. The remaining proportion continues transcription of the next gene (Dickens *et al.*, 1984). This stop-start process generates a gradient of transcription with mRNAs representing leader proximal genes being more abundant than those from leader-distal genes. The gradient of expression of HRSV genes was confirmed by Barik (1992) in which the concentration of individual viral mRNA molecules were calculated using a slot blot assay and the molar ratio of mRNA was calculated against the NS1 mRNA. The percentage of mRNA for each gene of RSV is shown in Table 1.6.

Gene	NS1	NS2	N	P	M	SH	G	F	M2-1	L
Molar ratio	100	95	90	68	52	32	21	18	15	3

Table 1.6 Percentage molar ratio of HRSV mRNA indicating the genome expression gradient (Barik, 1992).

### 1.2.4.2 mRNA capping

Neither the vRNA nor cRNA contain a 5' cap structure. However, mRNA molecules are capped (m<sup>7</sup>G(5')ppp(5')Gp) and polyadenylated (Barik, 1993). Barik (1993), using radioactive S-adenosyl-methionine, showed that the methyl group, but not the G residues, is added by host enzymes. In the same study it was shown that both the cap structure formation and the cap methylation are coupled to transcription and on-going transcription is required for cap formation and methylation and that the cap itself does not have any effect on transcription.

More recently, it has been reported that the L protein of vesicular stomatitis virus (VSV), a member of family *Rhabdoviridae* within the order *Mononegavirales*, is responsible for capping of pre-mRNA molecules (Li *et al.*, 2006), and specifically that the region V of the L protein is responsible in the pre-mRNA capping process (Li *et al.*, 2008b). Mutational analysis has shown that the capping process happens not in the conventional way of mRNA capping, but in a unique way using histidine residue in the V domain of the L protein (known as the HR motif) (Ogino *et al.*, 2010). It is likely that a similar process occurs in pneumoviruses with the L protein carrying out the capping process. The proposed mechanism for capping mRNA is shown in Figure 1.7 in comparison with the normal cellular capping process.

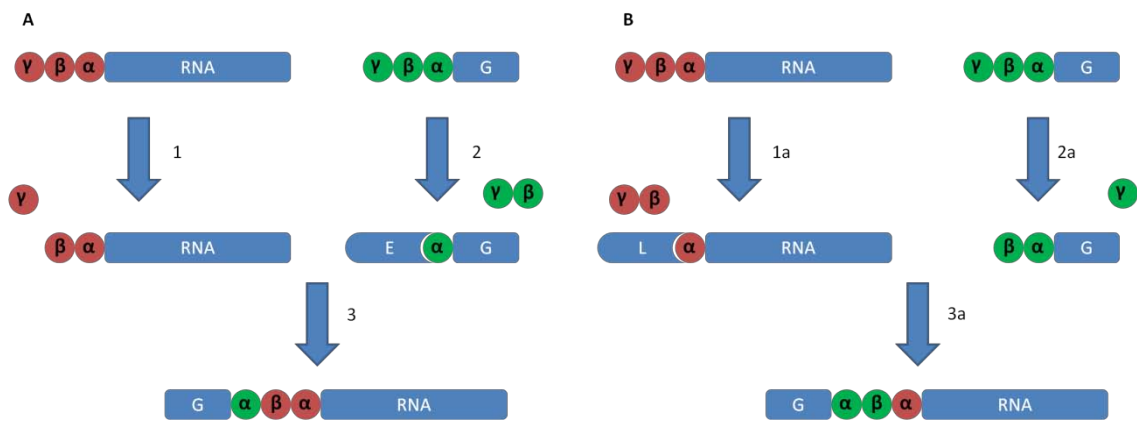


Figure 1.7 The process involving cap synthesis in eukaryotes and in VSV. A. In eukaryotic cells the  $\gamma$  phosphate of pre-mRNA is removed by the action of RNA 5'-triphosphatase (RTPase) (1). In the next steps (2 and 3) a GMP is transferred by a guanylyltransferase (GTase) enzyme action to the pre-mRNA to produce the 5' cap structure. B. In VSV the L protein mediates both reactions and removes two phosphate groups from the tri-phosphate-pre-mRNA structure (1a) and transfers a GDP to the pre-mRNA structure by its RNA:GDP polyribonucleotidyltransferase (PRNTase) activity (2a and 3a). Adapted from Ogino *et al.* (2010).

### 1.2.4.3 Transcription termination and polyadenylation

Harmon and colleagues (2001) have summarised the *cis* factors involving in the transcription termination in three categories: the two conserved regions of the Gene End sequence (the 3'-UCAAU-5' and the U track) and the central region which is located between the conserved regions (Section 1.2.2). Analysis of the gene end sequence of the M gene of HRSV vRNA has shown that nucleotides at positions 2 to 6 are important for transcription termination, whereas the nucleotides at positions 1 and 7 are not important to obtain efficient transcription termination. At position 8 of the gene end sequence, an A or U residue allows termination. The presence of four U residues is necessary for efficient transcription termination, and genes of HRSV with shorter U residues have been shown to fail to produce mRNA efficiently and produce readthrough polycistronic RNA templates (Harmon *et al.*, 2001). Poly-adenylation occurs when the polymerase complex reaches to the gene end sequence. The presence of a U rich sequence in the gene end directs the RNP complex to generate the poly A sequence.

In contrast with the gene start sequence, the gene end sequence is not well conserved among the pneumoviruses. In PVM the gene end sequence consists of uAGUuAnnn(A)<sub>n</sub> and in HRSV it consists of AGU(U/A)Annnn(A)<sub>n</sub> (Chambers *et al.*, 1991; Melero, 2006). The gene end sequence is followed by the non-conserved intergenic sequence (Chambers *et al.*, 1991). It was shown that the length or structure of the intergenic sequence does not affect the termination of transcription (Kuo *et al.*, 1996). This finding was challenged with another report describing the importance of the M/SH gene junction variation in the expression of the SH (Moudy *et al.*, 2004). Moudy and colleagues (2004) reported the P/M gene junction as the most variable gene junction with no detectable variation in the gene expression.

### 1.2.5 Genome replication

It has been proposed that the RNP complex which carries out genome replication is the same complex used in transcription. However, the switch from transcription to replication is not completely understood. The study of Sendai virus, a paramyxovirus, has suggested that the abundance of the N protein is important in controlling the switch between the RNA transcription and replication processes. The presence of the N protein in the cytoplasm of the infected cells may cover the gene end and gene start signals, and as a result the RNP complex does not respond to the gene end signals and instead reads through the cRNA (Vidal & Kolakofsky, 1989).

As for transcription, genome replication is controlled by the leader (Le) sequence of the genomic RNA to synthesis cRNA and the complementary sequence of trailer region of the vRNA (TrC) to synthesis vRNA from the cRNA template (Figure 1.8). Figure 1.9 shows the Le region from several pneumoviruses, and Figure 1.10 compares the Le and the TrC regions of HRSV. Genome replication requires switch of the RNP complex from transcription phase to the replication (Figure 1.8).

Recently, it has been shown that the HRSV RNP complex tends to start RNA synthesis with a purine residue and specifically it prefers adenosine to start vRNA or cRNA synthesis. Deletion of the first nucleotide of the template RNA or introduction of a mutation substituting a U residue in the template urges the RNP complex to start the RNA synthesis from the first available pyrimidine residue in the template (the third nucleotide residue of the template) (Noton *et al.*, 2010). The RNP polymerase binds to the nucleotides 3, 5, 8, 9, 10 and 11, facing its active site towards nucleotides 1 and 2. The RNA polymerase then uses ATP molecule to start RNA synthesis independently of the template to start synthesis of RNA (Noton *et al.*, 2010). It has been suggested that the *de novo* addition of ATP to the nascent RNA in HRSV is mediated by the same secondary ATP binding site which was recognized in the VSV L protein (Massey & Lenard, 1987; Noton *et al.*, 2010).

Initiation of the assembly of the nascent RNA before the RNP complex reaches to the first intergenic region increases the possibility of the read-through RNA synthesis and hence leads to replication rather than transcription (Vidal & Kolakofsky, 1989). Further studies on Sendai virus have confirmed coupled packaging of the nascent RNA with the RNP complex making the replication occur with a higher processivity in comparison to transcription (Gubbay *et al.*, 2001). In a recent study of HRSV, Noton and colleagues (2010) observed that if the RNA synthesis begins at position 3 of the leader sequence it fails to proceed more than 50 nucleotides. They have linked this phenomenon to the possibility of the attachment of the N protein to the first two nucleotides of the leader sequence during coupled RNA synthesis and packaging process, with the binding of the N protein ensuring processivity of RNA synthesis during replication.

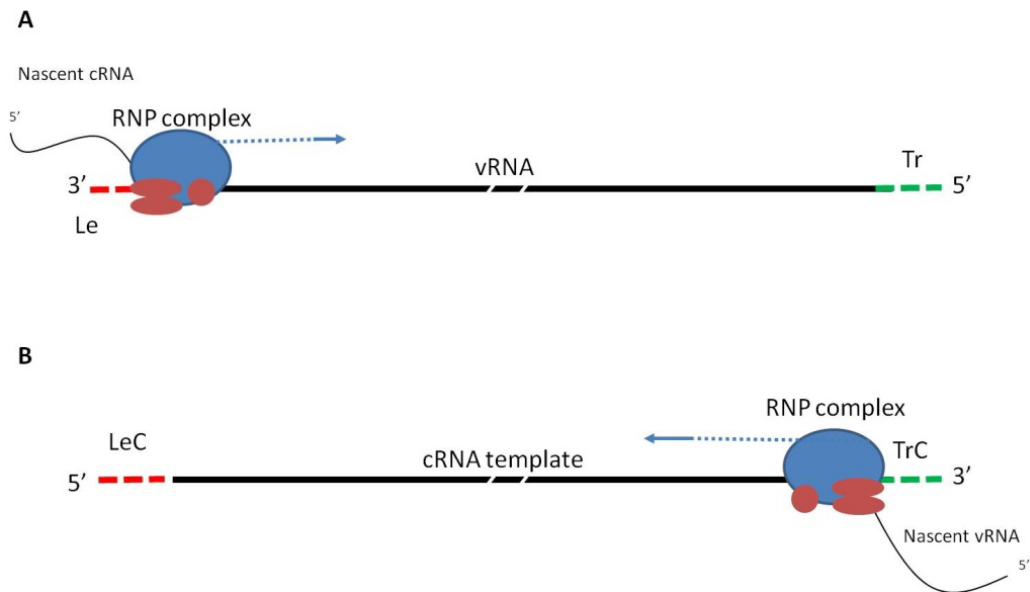


Figure 1.8 HRSV genome replication. The RNP complex (depicted as a compartment in blue and red colors) recognizes the genomic leader (Le) sequence or the trailer complementary sequence (TrC) sequence to dock on the RNA template and starts RNA synthesis. The blue arrows show the direction that the RNP complex moves. The Le and leader complementary sequences (LeC) are shown in red color. The trailer (Tr) and TrC sequences are shown in green color. The red compartments of the RNP complex represent P, N, and M2-1 proteins, and the blue compartment represents the large polymerase protein.

PVM15 (Warwick)	UGCGCUUUUUUACGUAU--UGUUUUGAUAGUUGGACUUUUUUCAAUCCUGUU--
PVM15 (ATCC)	UGCGCUUUUUUACGUAU--UGUUUUGAUAGUUGGACUUUUUUCAAUCCUGUU--
PVMJ3666	UGCGCUUUUUUACGUAU--UGUUUUGAUAGUUGGACUUUUUUCAAUCCUGUUC--
BRSV	UGCGCUUUUUUACGCAUAUUGUUUGGACAUGUAGGUUUUUUCUAGCCCCG----
HRSVB1	UGCGCUUUUUUACGCAUGAUGUUUGAACGUGUAAGCCUUUUUUACCCCCG----
HRSV A2	UGCGCUUUUUUACGCAUGUUGUUUGAACGUAUUUGGUUUUUUUACCCCCG----
HRSV19	----CUUUUUUACGCAUGUUGUUUGAACGCAUUUGGUUUUUUUACCCCCGUUUAU
	***** ** *

Figure 1.9 Genomic sequence of the leader domain of pneumoviruses. The leader sequence of members of genus pneumovirus are compared together. The genomic sequence of the RNA is shown. See Table 1.3 for the accession numbers of the sequences in the genebank. The identical nucleotides are shown in yellow. The “\*” character indicates positions which have a single, fully conserved residue (Thompson *et al.*, 2002).

cRNA_Leader	UGCUCUUUUUUUCACAGUUUUUGAUUAUAGAGCAUUAAAU---
vRNA_Leader	UGCGCUUUUUUACGCA--UGUUGUUUG-AACGUUUUGUUUUU
	*** ***** * ** * *** ** * * *** *

Figure 1.10 Comparison between leader sequence in the genome and antigenome of HRSV. The leader sequence of genome is indicated by vRNA\_Leader and the genome sequence of the antigenome is indicated by cRNA\_Leader. The "\*" character indicates positions which have a single, fully conserved residue (Thompson *et al.*, 2002).

### 1.2.5.1 M2-2 protein

The M2-2 protein (with a  $M_r$  value of 9-10,000) is synthesised from the second ORF of the M2 gene, partially overlapping the M2-1 open reading frame (Ahmadian *et al.*, 1999; Collins *et al.*, 1996a; Collins & Wertz, 1985a). From three tandem AUG codons at the 5' end of the M2-2 ORF, only the first and the second codons can be used to synthesise the M2-2 protein which is functional in the minigenome system (Cheng *et al.*, 2005). Expression of the M2-2 ORF from the M2 mRNA requires termination of the expression of the first ORF and re-initiation of expression of the second ORF (M2-2) in a process called coupled translation (Ahmadian *et al.*, 2000). In the coupled translation process the overlapping sequence between M2-1 and M2-2 directs the ribosome towards the initiation signal of the M2-2 ORF, with an efficiency determined by the sequence upstream of the overlap region (Gould & Easton, 2007).

In a HRSV sub-genomic mini-replicon system, it was shown that the M2-2 protein acts as a negative regulatory protein in RNA synthesis (Collins *et al.*, 1996a; Hardy & Wertz, 1998). It has been suggested that the M2-2 protein maintains the balance between the genome replication and transcription (Bermingham & Collins, 1999). In cells infected with HRSV strain A2 there is a rapid increase in mRNA production until it levels off, whilst in cells infected with a HRSV strain A2 lacking the M2-2 gene an initial delay was observed in the mRNA synthesis followed by a sudden increase in the level of mRNA higher than that of the wildtype which led to an accumulation of mRNA. In contrast, in comparison with the wildtype virus there was a significant reduction in the accumulation of cRNA with the M2-2 deleted virus (Bermingham & Collins, 1999). It has been also reported that over-expression of M2-2 protein in HRSV infected cells completely inhibits virus growth (Cheng *et al.*, 2005).



In HRSV, the M2 mRNA overlaps the beginning of the L gene by 68 nt. Transcription from the L gene start sequence frequently terminates at the M2 gene end sequence which has the effect of down regulating expression of the L gene mRNA and hence the level of L protein (Collins *et al.*, 1987). Comparison between mRNA and protein synthesis of an M2-2 deleted recombinant HRSV and rescued HRSV strain A2 are similar (Jin *et al.*, 2000a). Attenuation of HRSV with deleted M2-2 in BALB/c mice and cotton rats has also been reported (Jin *et al.*, 2000a).

### **1.2.6 Virus assembly and release**

Similar to other members of *Paramyxoviridae*, assembly of HRSV occurs in the cytoplasm of the infected cells in assembly sites close to the plasma membrane. The virus glycoproteins are transferred to the assembly sites of the infected cells using cellular proteins.

Assembly in *Paramyxoviridae* occurs in two steps; assembly of the RNP complex and assembly of the envelope. The assembly of RNP complex to the RNA occurs during replication of the new genomic RNA.

#### **1.2.6.1 Matrix (M) protein**

The matrix (M) protein of pneumoviruses acts as a connecting layer between the viral nucleocapsid and the viral membrane. Its interaction with different viral proteins including N, G, F and M2-1 has been demonstrated. There are two main functions reported for the M protein; switching off viral transcription and directing virus assembly towards budding. It has been shown for members of the order *Mononegavirales* that the M protein is essential for virus budding, suggesting it has intrinsic capabilities of directing virus budding (Mebatsion *et al.*, 1996; Mebatsion *et al.*, 1999; Teng & Collins, 1998). The cytoplasmic domain of the G protein, in particular the first 6 amino acids and especially serine at position 2 and asparagine at position 6, play a major role in the interaction of G and M proteins (Ghildyal *et al.*, 2005). The interaction between the F and M proteins is independent from the interaction between the G and M proteins (Ghildyal *et al.*, 2005). This suggests that there are separate domains on the M protein responsible for interaction with F and G glycoproteins.

It has been shown that the first 110 amino acids from the N-terminus of the N protein are responsible for the interaction between M and N proteins (Li *et al.*, 2008a).

### **1.2.6.2 Virus release**

Sample titration from the apical and basolateral surfaces of polarised human airway epithelial (HAE) cells grown on 0.4 µm semipermeable membranes indicated that virus budding occurs from the apical surface of cells and the virus is released into the luminal periciliary fluid and/or possibly into overlaying mucus layer *in vivo*. This would result in a “vectorial pattern”, forming a circular form of infection with each initial focal point in the centre and radial spread of the progeny HRSV into adjacent cells (Zhang *et al.*, 2002).

## **1.3 Pneumovirus pathogenicity**

### **1.3.1 Virus factors affecting pathogenicity**

Pathogenesis in viruses is frequently multifactorial, and requires interaction between many viral and host proteins. Other factors such as regulatory sequences may also play a role. For example, certain attenuated derivatives of HRSV have a 4C nucleotide substitution at their leader position which has been shown to affect the amount of genome replication (Section 1.2.2). Other viral factors affecting the pathogenicity are described below (Fearn's *et al.*, 2002).

#### **1.3.1.1 Non-structural (NS) protein 1 and 2**

The HRSV non-structural proteins 1 and 2, previously known as 11K and 14K proteins, are found in abundance in the infected cells, but in little amounts associated with virion extracts. Therefore, they are designated as non-structural proteins (Huang *et al.*, 1985). The NS1 and NS2 proteins are unique to pneumoviruses and lack counterparts among other negative strand non-segmented RNA viruses. The NS1 gene of HRSV strain A2 is 552 nucleotides in length and encodes a polypeptide of 139 amino acids. The NS2 gene is 503 nucleotides in length and encodes a polypeptide of 124 amino acids in length (Collins & Wertz, 1985b). The NS1 and NS2 gene positions close to the leader/promoter region makes them the most abundantly transcribed genes in HRSV infected cells (Collins & Wertz, 1985b; Glazier *et al.*, 1977).

In a novel experiment, recombinant rabies viruses carrying either the BRSV NS1 or NS2 genes were generated to assess the ability of these genes to confer reduced sensitivity to interferon. The data showed that both NS1 and NS2 can inhibit the IFN response in the host cell but that both act together synergistically for the greatest effect

(Schlender *et al.*, 2000). It has been proposed that the NS1 and NS2 proteins may also contribute to host specificity by exerting the greatest effect on the IFN systems of the relevant host animal (Bossert & Conzelmann, 2002; Schlender *et al.*, 2000).

The importance of NS1 and NS2 proteins in antagonizing the host innate immune response was also analysed using recombinant BRSV virus deleted in either or both of the genes (Schlender *et al.*, 2000). It was shown that BRS recombinant viruses lacking NS1, NS2 or both proteins grow efficiently in Vero, HEp-2 and BSR-T7/5 cells. However, their growth was reduced by 100-fold in MDBK cells, suggesting further that NS1 and NS2 proteins are IFN antagonists. The evidence from the experiment was confirmed by co-cultivating Vero cells with MDBK cells in a two-chamber culture system in which the chambers were isolated from each other by a permeable filter. The analysis showed that cells responded to soluble factors released from MDBK cells by inhibition of the growth of the viruses. A similar result was observed with macrophages as the effector (Schlender *et al.*, 2000; Valarcher *et al.*, 2003).

The function of NS1 and NS2 proteins in virus replication and plaque morphology of HRSV strain A2 was analysed by Jin and colleagues (2000). It was shown that the presence of the NS1 and NS2 genes in the HRSV genome is not essential for recombinant virus recovery or replication *in vivo* or *in vitro* (Jin *et al.*, 2000b). However, deletion of NS1 and/or NS2 genes in the recombinant viruses resulted in the generation of smaller plaque morphology and 100-fold reduction in virus replication in Vero and HEp-2 cells. Replication of the recombinant viruses in the lower respiratory tract of cotton rats was also reduced (Jin *et al.*, 2000b).

The function of the NS1 and NS2 proteins and their impact in virus replication and interferon antagonism in PVM strain 15 (ATCC) was studied by Buchholz *et al.* (2009). It was shown that the deletion of either the NS1 or NS2 genes did not affect the ability of the recombinant PVM to grow in interferon negative Vero cells. However, recombinant viruses lacking NS1 or NS2 genes did not grow in mouse embryonic fibroblast cells (Buchholz *et al.*, 2009).

### **1.3.1.2 Effect of the G glycoprotein in pneumoviruses pathogenesis**

The G glycoprotein of pneumoviruses has been identified as one of the major determinant factors of pathogenesis. As noted earlier (Section 1.2.1.1) the structural similarity between the CX3C motif of the G glycoprotein and fractalkine chemokine suggests an interaction between the CX3C motif and CXCR1 (fractalkine receptors)

which enables the G glycoprotein to mimic fractalkine function (Tripp *et al.*, 2001). This function of the G protein was attributed to the CX3C motif of the G protein (2006). Mice infected with recombinant HRSV deficient in the CX3C motif or lacking the G gene showed significantly higher levels of pulmonary CX3CR1<sup>+</sup>, CD4<sup>+</sup> and CD8<sup>+</sup> cells than in mice infected with the HRSV strain A2 (Harcourt *et al.*, 2006). These data suggest the importance of the G glycoprotein in controlling the influx of cytotoxic CD8<sup>+</sup> T cells, which have an important role in controlling the infection and virus clearance.

HRSV G protein stimulation of naive spleen cells induces IL4 and CXCL1 production (Harcourt *et al.*, 2006). The mechanism of this function of the G protein has not been identified, but this is attributed to the interaction between the G protein and extracellular heparin molecules, or the interaction between the cysteine noose and fractalkine receptors (Harcourt *et al.*, 2006).

The conserved cysteine rich region of the G glycoprotein acts as an antiinflammatory factor to antagonize the proinflammatory effect of F glycoprotein (Polack *et al.*, 2005). The G glycoprotein has also been shown to change cytokine production of the infected cells by blocking production of inflammatory cytokines, decreasing nuclear translocation of NF- $\kappa$ B. The similarity between the GCRR domain of the G protein (see 1.2.1.1) and the fourth sub-domain of TNFR1 suggests an interaction between these two molecules which may explain the G gene contribution in immunomodulatory function (Langedijk *et al.*, 1998).

More recently it has been shown that the G glycoprotein inhibits TLR 3 and 4 mediated type I interferon response by modulating the TICAM-I pathway (Shingai *et al.*, 2008). The down regulation of IFN- $\beta$  in MLE-15 cells (mouse lung epithelial cells) infected with a recombinant virus lacking the G gene was related to down regulation of interferon stimulated gene-15 (ISG-15) protein which its expression of which rapidly increases in response to viral infections (Moore *et al.*, 2008). In cells infected with recombinant HRSV lacking G gene, the mRNA and expression levels of suppressor of cytokine signalling 3 (SOCS3) were significantly higher in comparison with cells infected with the wildtype HRSV or HRSV lacking NS1 and NS2 genes (Moore *et al.*, 2008), indicating the effect of the G protein in inhibiting the innate immunity cellular signalling pathway.

The fast release of the Gs from the infected cells (after 24 hr of infection) helps the virus to modulate both cellular and humoral immunity by using the Gs as a decoy

for anti-G specific neutralising antibodies and as an imitative protein for CX3C chemokine in controlling monocyte trafficking (Collins & Graham, 2008).

G glycoprotein activates the humoral response. However, the antibody response to the G glycoprotein is sub-group specific (Section 1.1) and cross reactivity between anti-G antibodies rarely happens (Oshansky *et al.*, 2009). In contrast, anti-F antibodies are effectively capable of neutralizing the virus (Kao *et al.*, 1984; Melero *et al.*, 1997; Trudel *et al.*, 1991).

### **1.3.1.3 Effect of the F glycoprotein in pneumoviruses pathogenesis**

Unlike other paramyxoviruses, cleavage of the pneumovirus F protein is not a key determinant of pathogenicity and virus virulence. In contrast with the G glycoprotein, the HRSV F glycoprotein has been reported to interact with toll like receptor 4 (TLR4) and activates the signalling pathway leading to up-regulation of TNF $\alpha$ , IL-6 and IL-12 production (Collins & Graham, 2008; Oshansky *et al.*, 2009).

Antibody response is against two types of F protein: the mature form which is present on the virus particles and the immature form which is the unfolded form of the F protein. The immature form does not contain all of the epitopes necessary for neutralisation (Lopez *et al.*, 1998). Early release of the immature form from lysed infected cells results in an inefficient antibody response (Oshansky *et al.*, 2009). The major neutralising antibodies are against F protein.

It has been reported that the F protein in HRSV strain Line 19 is more mucogenic than its counterparts in HRSV strain A2 or HRSV strain Long (Moore *et al.*, 2009). Moore and colleagues (2009) have shown that BALB/cJ mice infected with a chimeric HRSV virus strain A2 carrying the F gene from the strain Line 19 showed lower IFN- $\alpha$  production, higher lung viral load, higher lung IL-13 level, higher airway mucin production level and higher airway hyper responsiveness than those were infected with recombinant HRSV strains A2 and Long. In comparison with the sequence of the F protein in the strains A2 and Long, the F protein of HRSV strain Line 19 contains 4 unique amino acid substitutes in its sequence (Moore *et al.*, 2009).

The importance of pep27 of BRSV and the furin like cleavage sites of the F protein (Section 1.2.1.2) in BRSV replication in the cell culture has been investigated (Zimmer *et al.*, 2002). Although the pep27 or the furin cleavage sites are not required for efficient virus growth in the cell culture (Zimmer *et al.*, 2002), in a separate study (Valarcher *et al.*, 2006), it was shown that calves infected with recombinant viruses

carrying mutations in the furin cleavage sites of the F protein or lacking the pep27 demonstrate a less gross and microscopic pulmonary pathology in comparison with the calves infected with the recombinant viruses carrying the authentic F glycoprotein.

#### Small hydrophobic (SH) protein

The SH protein is a small (64 amino acids for HRSV strain A2, subgroup A, and 65 amino acid for HRSV subgroup B) and highly conserved transmembrane glycoprotein transcribed and translated from the SH gene. The amino acid sequence of the SH protein is highly conserved among both HRSV subtypes (Carter *et al.*, 2010).

The SH protein was shown to be non-essential in HRSV replication and was categorised as an accessory protein of HRSV (Bukreyev *et al.*, 1997; Whitehead *et al.*, 1999). The SH protein accumulates in different forms (designated SH<sub>0</sub>, SH<sub>p</sub>, SH<sub>t</sub> and SH<sub>g</sub>) inside infected cells. SH<sub>0</sub> is the non-glycosylated and full length form of the SH protein with a relative molecular mass of 7,500, SH<sub>g</sub> contains an N-linked carbohydrate side chain (*Mr* 13,000-15,000), SH<sub>p</sub> has more modification by addition of polylactosaminoglycan oligosaccharides to the N-linked side chain (*Mr* 21,000-65,000) and SH<sub>t</sub> is distinct nonglycosylated form of the SH protein with a shorter N-terminus as its translation starts from the second start codon of the SH gene (Anderson *et al.*, 1992; Collins & Mottet, 1993). The N-linked glycosylation sites and the central hydrophobic region are highly conserved among HRSV and BRSV subtypes (Anderson *et al.*, 1992). Lipid rafts in the cellular secretory pathway such as the endoplasmic reticulum, Golgi compartment and cell surface are the main sites that SH protein is observed in abundance in the infected cells. The appearance of SH protein in viral filamentous structures is minimal (Rixon *et al.*, 2004).

The function of the SH protein is not clearly known. However, an anti-apoptotic function has been reported (Fuentes *et al.*, 2007), and large plaque formation phenomenon seen in SH deleted recombinant HRSV is attributed to the anti-apoptotic function of the SH protein (Bukreyev *et al.*, 1997; Techaarpornkul *et al.*, 2002).

There is evidence that the SH protein is phosphorylated via a MAPK p38-dependent pathway during the virus infection cycle, and the phosphorylated SH protein is linked with a pentamer homo-oligomer structure (Rixon *et al.*, 2005). Observations using electron microscopy techniques suggest the presence of pentamer and hexamer homo-oligomer viroporin channels formed by the SH glycoprotein (Carter *et al.*, 2010; Gan *et al.*, 2008; Kochva *et al.*, 2003). The function of these channels in the viral and

cellular life cycle, their effect on the apoptosis (Lang *et al.*, 2005), and the possibility of their contribution in the virulence remains to be clarified.

Deletion of the SH gene from recombinant HRSV viruses did not have any significant effect on the virus growth *in vitro* (Bukreyev *et al.*, 1997). However, a deleted SH recombinant virus showed 40-fold reduction in replication in the lower respiratory tract of chimpanzees infected with the virus and the chimpanzees showed less rhinorrhea in comparison with the chimpanzees receiving the wild-type HRSV (Whitehead *et al.*, 1999).

Sequence analysis of the SH gene among clinical isolated showed little or no sequence variability (Chen *et al.*, 2000), indicating conservation of SH gene among the clinical isolates and suggesting an evolutionary importance of the SH gene for maintaining viral pathogenicity.

The expression of MIP, MCP-1, and IP-10 was related to the biological function of G and/or SH glycoproteins. The mRNA expression level for MIP, MCP-1 and IP-10 was reported as being higher in the bronchoalveolar lavage (BAL) cells of mice infected with HRSV strain *cp52* than those of mice infected with HRSV strain B1. The difference between the expression levels was related to the lack of biological activity of G and/or SH glycoproteins in the strain *cp52* (Tripp *et al.*, 2000a).

### **1.3.2 Host factors affecting pathogenicity**

There are host factors that may affect the severity of disease associate with HRSV infection. The importance of these host factors in the infection of HRSV was observed by Prince *et al.* (1979). To develop a mice model for HRSV infection, 20 strains of inbred mice were infected with HRSV strain Long (Prince *et al.*, 1979). Among the infected mice strains the CBA/CaCHN strain of mice was reported as the most resistant while the strain DBA/2N was reported the most permissive strain (Prince *et al.*, 1979). In a separate study, Anh *et al.* (2006) tested susceptibility of 6 different strains of mice (129/Sv, BALB/c, C3H/HeN, C57BL/6, DBA/2 and SJL) to PVM infection. the strain SJL showed the most resistance to PVM infection while the rest showed high susceptibility to the infection with a greater degree for the strain 129/Sv which was followed by the strains DBA/2, C3H/HeN, BALB/c and C57BL/6 in order.

The host factors affecting the pathogenicity related with HRSV infection may include age, gender, race, congenital heart disease, bronchopulmonary dysplasia, immunosuppression or immunodeficiency, narrow airways, premature birth, low birth

weight and the level of immunity of individuals during the time of infection (Collins & Graham, 2008; DeVincenzo, 2005; Leader & Kohlhase, 2003; Oshansky *et al.*, 2009; Walsh *et al.*, 1997; Welliver, 2003). The relation of nucleotide polymorphism in certain genes with severity of disease caused by HRSV also has been examined. These include genes involved in immune response including genes encoding cytokines and chemokines (including IL-4, IL-8, IL-10, IL-13, IL-18, RANTES/CCL5), or surface and signal transduction elements (including TLR-4, CD14, IL-4R, CX3CR1, CCR5 and surfactant proteins) (Amanatidou *et al.*, 2006; Awomoyi *et al.*, 2007; Collins & Graham, 2008; Hull *et al.*, 2003; Miyairi & DeVincenzo, 2008; Paulus *et al.*, 2007; Puthothu *et al.*, 2007).

### **1.3.2.1 The effect of tachykinins on respiratory disease**

Mammalian tachykinins are traditionally categorized as neurotransmitters. They all are short polypeptides (about 11 amino acids) processed and post-translationally modified from longer polypeptides (Page, 2005). Despite early reports indicating production of tachykinins solely from neural cells, there are several reports suggesting the production of tachykinins in other tissues and cells including inflammatory sites, macrophages, eosinophils, lymphocytes, and dendritic cells (Bost & Pascual, 1992; Killingsworth *et al.*, 1997; Weinstock *et al.*, 1988). There are also reports showing the presence of tachykinins in the peripheral circulation and even in placental tissue (Page, 2005; Page *et al.*, 2000). The tachykinins in these cells exert their effects in different regulatory and inflammatory mechanisms.

Five tachykinin peptides have been identified in mammals; substance P (SP), neurokinin A (NKA), neuropeptide K (NKK), neuropeptide- $\Upsilon$  (NKY), and neurokinin B (NKB). All of the tachykinins share a similar C-terminal of Phe-X-Gly-Leu-Met-amid where X is an aromatic or hydrophobic amino acid (Hietala *et al.*, 2005). The conserved C terminal region acts as the central activation factor to activate tachykinin receptors (Page, 2005). At least three receptors have been reported to interact with tachykinins: the neurokinin-1, neurokinin-2 and neurokinin-3 receptors (NK-1R, NK-2R, and NK-3R). It has been shown that different tachykinins have different affinities in binding to the receptors, as SP binds to NK-1R more efficiently than the others. NKA and NKB bind to NK-2R and NK-3R with a higher affinity in comparison with the others (O'Connor *et al.*, 2004; Page, 2005).



SP was first discovered in 1931 and was subsequently isolated from bovine hypothalamus and characterized in 1971. In humans and animals, SP increases ventilation (Saaresranta & Polo, 2002). It has been shown that SP has a bronchoconstriction effect *in vitro*, though no such effect is reported *in vivo* (Saaresranta & Polo, 2002). The amount of SP in HPLC purified tracheal and lung extracts has been measured at 13.0 and 5.4 pmol/g tissue, respectively (Groneberg *et al.*, 2006).

It seems likely that human respiratory inflammation causes SP and NKA production from neuronal or non-neuronal cells located in the respiratory system. In guinea pig models, 24 hr after allergen challenge the amount of SP and NKA was increased 3- to 4-fold. In a study on NK-1R depleted animal models it was shown that deletion of the receptor did not have any significant effect on the accumulation of inflammatory factors such as antigen specific IgE and inflammatory cells in bronchoalveolar lavage (BAL) fluids (Tournoy *et al.*, 2003). The effect of tachykinins in signal transduction in respiratory systems and its importance in the inflammation remains to be understood (Groneberg *et al.*, 2006).

It has been shown that SP is produced in association with the G and/or SH glycoproteins in lungs of mice acutely infected with HRSV (Tripp *et al.*, 2000b). In this study the level of SP in the BAL of naïve mice was measured by a competitive ELISA test and compared to those of HRSV strain B1 and HRSV strain *cp52* infected mice (Karron *et al.*, 1997). It was shown that that SP was significantly lower in the group of mice infected by *cp52*. The absence of G and SH glycoproteins in the strain *cp52* suggested that either or both of the glycoproteins contributed to the alteration of the level of SP detected.

### **1.3.3 Long lasting HRSV infection**

The contribution of HRSV to chronic obstructive pulmonary disease (COPD) has been sought for some time. As with acute respiratory disease caused by HRSV the main site of inflammation in COPD is the small airways where obstruction and inflammation happens upon the infection (Wilkinson *et al.*, 2006). The infection of a recombinant HRSV expressing GFP persisted for more than one month in well differentiated human airway epithelium (WD-HAE) culture, with normal cilia movement and no obvious cytopathology or syncytia formation (Zhang *et al.*, 2002). This suggested that long term infection may be possible.

Two reports have described long-lasting pneumovirus infection *in vivo*. To investigate the mechanisms involved in latency, persistency or immune evasion, no virus was detected in broncho alveolar fluid isolated from BALB/c mice infected with HRSV, while the virus was detectable in lung homogenates of the mice using PCR after 100 days of post infection (Schwarze *et al.*, 2004). The mice did not demonstrate any signs of infection during the period. In a similar study BRSV was isolated directly from tracheobronchial and mediastinal lymph nodes of calves infected with BRSV in up to 71 days post infection (Valarcher *et al.*, 2001).

#### **1.3.4 Using PVM to study the pathogenesis of HRSV**

Generally speaking, mouse models are not fully permissive to HRSV infection. Mice least sensitive to HRSV are CBA/CaHN and the most sensitive are DBA/2N (Byrd & Prince, 1997). In comparison with mice, cotton rats are more permissive to HRSV infection replicating to a 100 times greater titre than that in DBA/2N mice (Byrd & Prince, 1997; Prince *et al.*, 1978). However, the lack of reagents in characterising the details of the immunity against HRSV infection is one of the major limiting factors in HRSV studies in cotton rats. The availability of congenic, transgenic and knock-out strains of cotton rats is another limiting factor in HRSV studies (Byrd & Prince, 1997). Another limitation in using mice models of HRSV is the difference between human and mice eosinophils in their response to disease (Rosenberg *et al.*, 2009). Considering that eosinophils play an important role in the HRSV infection, using mouse models in studying HRSV pathogenesis may not reflect the aspects of the eosinophil's contribution in the virus pathogenicity.

Another alternative way to study the pathogenesis of HRSV is studying the pathogenesis of PVM in mice and extrapolating the findings to HRSV. PVM infection in mice reflects the same pathology as HRSV infection in human patients. It has been reported that intranasal inoculation of mice with 250 pfu of PVM in a volume of 50  $\mu$ l causes fatal disease in mice (Cook *et al.*, 1998). In contrast, a virus load of  $10^4 - 10^6$  pfu is necessary to infect mice (Cannon *et al.*, 1987; Easton *et al.*, 2004; Taylor *et al.*, 1984). Moreover, the pathogenic PVM develops clear symptoms including weight loss, abnormal gait, and tremor during its replication in mice which makes monitoring the progress of the disease simpler without the need for obtaining histopathology samples.

As in HRSV infection in human tissues, PVM infection in mice is accompanied by an influx of granulocytes into the lower respiratory system. Production of eosinophil

attractants like MIP-1 $\alpha$ , eotaxin and RANTES is a common feature between the infection caused by PVM in mice and the infection of HRSV in human tissues, and using a gene microarray expression of MIP-2 (mouse orthologue for IL-8) and MCP-2 has been shown (Bonville *et al.*, 2006; Domachowske *et al.*, 2000a; Domachowske *et al.*, 2000b).

### **1.3.5 Aims**

The main challenge in the study of pathogenesis of HRSV is the lack of a suitable animal model. Mouse models are partially permissive for HRSV infection but the study of the pathogenicity can be achieved only through histopathological analysis. PVM is genetically close to HRSV and displays a pathology in mice which resembles that of HRSV in humans. It therefore has been proposed that the PVM infection in mice may be suitable model for the study of pneumoviruses in their host.

The aims of this project were:

- To establish a reverse genetics system for PVM which could be used to study the effects of specific mutations on pathogenicity.
- To investigate the pathogenicity in mice of a series of stocks of PVM which had been progressively passaged in tissue culture and to determine the nucleotide sequence of the genes of pathogenic and non-pathogenic viruses. This data would be compared with the published sequences of pathogenic and non-pathogenic PVM strains.
- To introduce specific mutations into the G gene of recombinant PVM viruses and determine the effect on pathogenicity in mice.
- To investigate possible consequences of specific mutations associated with pathogenicity on molecular processes such as the level of gene expression.

## **CHAPTER 2**

# **MATERIALS AND METHODS**

## **2.1 Solutions**

### **2.1.1 AEC peroxidase substrate**

AEC (3-amino-9-ethylcarbazole) was dissolved in DMSO at a concentration of 3.3 mg/ml and kept frozen in small aliquots at -20°C until the time of use. 20 mM of CH<sub>3</sub>COONa (pH 5.6) and 0.6ml AEC solution were added to water in order, and 16 µl of H<sub>2</sub>O<sub>2</sub> was added to the mixture prior to use to have a final volume of 10 ml of the substrate.

### **2.1.2 ABTS Substrate**

ABTS (2,2'-azino-bis-[3-ethylbenzothiazoline-6-sulfonic acid]) substrate was prepared by dissolving 100 mg of ABTS (SIGMA) in 100 ml of buffer. To prepare the buffer 1.67 g of the buffer powder (Roche; consists of sodium perborate, citric acid, and disodium hydrogen phosphate) was dissolved in 100 ml of distilled water.

### **2.1.3 Carbonate coating buffer**

Coating buffer contained 0.34% (w/v) Na<sub>2</sub>CO<sub>3</sub> (anhydrous), 0.57% (w/v) NaHCO<sub>3</sub>, pH 9.6. The pH was adjusted using NaHCO<sub>3</sub>.

### **2.1.4 Luria-Bertani broth and agar media**

10 g of tryptone, 5 g of yeast extract and 10 g of NaCl were dissolved in one litre of water. Bacteriological agar (10 g/l) was added to prepare LB agar. After being autoclaved the solution was aliquoted in appropriate tubes or dishes.

### **2.1.5 Ketamine and xylazine cocktail**

To obtain the anesthetic condition in mice a ketamine/xylazine cocktail was used. The mixture contained 10 mg/ml of the anesthetic ketamine and 1.8 mg/ml of the analgesic xylazine dissolved in sterile phosphate buffered saline (PBS). Adult BALB/c mice were anaesthetized by intraperitoneal injection with 150 µl of the mixture.

### **2.1.6 Phenol:chloroform mixture**

A mixture of phenol (50% v/v) and chloroform (50% v/v) was used to purify DNA from protein contaminants.

### **2.1.7 SDS-PAGE sample buffer**

For preparing 10 ml of 4x sample buffer, 3 ml of 20% (w/v) SDS solution, 3 ml of glycerol, 1.6 ml of  $\beta$ -mercaptoethanol and 0.006 g bromophenol blue were mixed together. The solution was made prior to use and kept at 4°C for short term storage.

### **2.1.8 SDS-PAGE electrophoresis buffer**

SDS-PAGE electrophoresis buffer (1x) was made by dissolving 25 mM Tris, 250 mM glycine (pH 8.3) and 0.1% (w/v) SDS in distilled water.

### **2.1.9 TAE Buffer**

Tris-acetate (TAE) buffer was specifically used to run low melting agarose gels. To make 50x of Tris acetate buffer 242 g of Tris base, 57.1 ml glacial acetic acid and 100 ml of 0.5 M EDTA (pH 8.0) were mixed together and dissolved into one litre of distilled water.

### **2.1.10 TB solution**

TB solution composed 10 mM PIPES, 55 mM  $MnCl_2$  and 15 mM  $CaCl_2$  and 250 mM KCl. After dissolving the PIPES  $CaCl_2$  and KCl in distilled water and adjusting the pH to 6.7,  $MnCl_2$  was added to the solution. The solution was prepared freshly prior to use and filter sterilized.

### **2.1.11 TBE Buffer**

Tris-borate (TBE) buffer was prepared in a 10x concentrated format. The concentrated solution was diluted to 1x to prepare the working concentration. To make the 10x concentrated solution 54 g of Tris base, 27.5 g of boric acid and 20 ml of 0.5 M EDTA (pH 8) were dissolved in one litre of distilled water.

### **2.1.12 TBS and TBST buffers**

50 mM Tris and 150 mM NaCl were mixed together and the final volume was adjusted to 1L. For making 1X TBST 1 ml of Tween-20 was added to 1X TBS.

### **2.1.13 X-Gal and IPTG solutions**

A 20 mg/ml stock of X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside) in DMF (dimethylformamide) was prepared. The solution was kept at -20°C

and stored in the dark. A 100 mM stock solution of IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) in distilled water was prepared filter sterilized and kept at -20°C and in the dark until the time of usage.

## **2.2 Cell culture and classic virology techniques**

### **2.2.1 Viruses and cell lines**

BS-C-1 (African green monkey kidney cell lines) and BSR-T7/5 (BHK21 cells expressing bacteriophage T7 RNA polymerase) cells were obtained from Dr. R. Ling, pneumovirus Laboratory, University of Warwick (Buchholz *et al.*, 1999). All cell lines were grown at 37°C in Glasgow minimum essential medium (GMEM) supplemented with 10% fetal calf serum (FCS), 2 mM glutamine, 100 units/ml penicillin and 100  $\mu$ g/ml streptomycin. A maintenance medium with the same amount of antibiotics and L-glutamine, but with 2% FCS was prepared and used to keep cells.

PVM strain J3666 and PVM strain 15 (Warwick) were obtained from Prof. A. Easton, pneumovirus Laboratory, University of Warwick.

### **2.2.2 Mammalian cell culture**

BS-C-1 cells were grown to confluence in Glasgow modified Eagles' medium (GMEM) supplemented with 10% fetal bovine serum (FBS), Penicillin (100 units/ml), streptomycin (50 mg/ml) and L-glutamine (2 mM final concentration). After obtaining confluence, cells were detached from the surface by washing with trypsin 1:250 prior to dispersal in growth medium to obtain a single cell suspension. The cell suspension was divided between tissue culture dishes to have a final suspension of cells containing  $5 \times 10^5$  cells/ml. The tissue culture dishes were incubated at 37°C in 5% (v/v) CO<sub>2</sub> and air.

For the BSR-T7/5 cell line, the same procedure was repeated. GMEM medium supplemented with Geneticin (Invitrogen) to a final concentration of 1 mg/ml was used to cultivate the cells.

### **2.2.3 Freezing of mammalian cells**

Mammalian cells were frozen in a mixture of 90% serum plus 10% DMSO. Freshly prepared confluent cells were detached from the growth surface and pelleted at low speed. The cells were resuspended in freezing mix which had been freshly prepared

and filter sterilized prior to use. The number of cells and the volume of freezing mixture were adjusted to approximately  $5 \times 10^5$  per ml. The suspension was placed into cryo-tubes in 1 ml aliquots. The cryo-tubes were moved to a  $1^\circ\text{C}$  freezing container (Mr. Frosty™, Nalgene Labware) to control the reduction of the temperature to a rate of  $1^\circ\text{C}$  for every minute. The freezing container containing the samples was moved to a  $-70^\circ\text{C}$  freezer. The frozen samples were moved to the gas phase of liquid nitrogen for longer storage.

#### **2.2.4 Transfection of mammalian cell lines using Lipofectamine 2000**

Lipofectamine 2000 (Invitrogen) was used to transfect BSR-T7/5 cells. Briefly, sub-confluent BSR-T7/5 cells were freshly prepared in  $10\text{ cm}^2$  tissue culture plates the day before transfection. Prior to transfection, the medium was replaced with GMEM supplemented with 2% FBS. The DNA mixture (see Sections 4.7 and 5.11 for the quantities) and Lipofectamine 2000 were mixed separately in  $250\ \mu\text{l}$  of OPTIMEM® (Invitrogen). For each  $1\ \mu\text{g}$  DNA,  $3\ \mu\text{l}$  of Lipofectamine 2000 was used. After a minimum of 5 min of incubation at room temperature, both the diluted DNA and the transfection agent were mixed together. The mixture was incubated for a minimum of 30 min, and then added drop-wise to the prepared cell culture.

#### **2.2.5 Transfection of mammalian cell lines using Turbofect**

The medium bathing the cells was replaced with GMEM containing 2% (v/v) FBS immediately prior to the transfection. The DNA to be transfected was resuspended in  $400\ \mu\text{l}$  of OPTIMEM® (Invitrogen). For every  $1\ \mu\text{g}$  DNA  $1.5\ \mu\text{l}$  of Turbofect was used. The desired amount of Turbofect reagent was added to the DNA suspension and mixed thoroughly. After 15 to 20 min incubation at the room temperature the DNA:Turbofect mixture was added to the cell culture medium.

#### **2.2.6 Virus cultivation**

Roller bottles containing sub-confluent BS-C-1 cells were infected with PVM virus in such a way to obtain a final moi of 0.05 pfu/cell. Both PVM strain J3666 and strain 15 (Warwick) were incubated at  $31^\circ\text{C}$  until 80-90% cell destruction was achieved. At this point the roller bottles were frozen and thawed. The freeze and thaw cycle was followed by scraping the attached cells into the medium using sterile glass beads. The suspension of virus and cell lysates were kept at  $-70^\circ\text{C}$  in small aliquots.



### **2.2.7 Micro-plaque assay**

Serial dilutions of virus were prepared in the maintenance GMEM medium containing 2% (v/v) FBS immediately before the titration and kept on ice during the assay. Freshly prepared sub-confluent BS-C-1 cells in a 96-well plate were used to quantify the titre of the viruses (section 2.2.1). The medium bathing cells was replaced with 100  $\mu$ l of the virus dilutions in triplicates. Cells were incubated for 1 hr at 31°C to complete adsorption. The adsorption was followed by washing cells to remove unbound viruses, and the cells were overlaid with maintenance medium. The titre of all of viruses was measured at 31°C after 24 hr incubation. Cells were fixed with 1:1 (v/v) mixture of chilled methanol and acetone for 30 min at room temperature. The fixed cells were washed with PBS 3 times, and incubated with 1/40 dilution of the monoclonal antibody (26/C3/B5) which is specific for the P protein of PVM (Ling & Pringle, 1989a). The mAb was diluted in PBS containing 1% gelatin (Ling & Pringle, 1989a). The cells were incubated with the antibody for 1 hr at 37°C. After removing antibody, the cells were washed 3 times with PBS. Goat anti-mouse antibody (BioRad) was used as the secondary antibody and cells were incubated with 1:2000 concentration of this for 1 hr. The incubation was followed by washing the cells 3 times with PBS. The washing process was followed by addition of AEC peroxidase substrate (section 2.1.1) and plates were observed until a brown color was developed. The number of plaques represented by brown cells was counted and the titre of virus stock was calculated.

## **2.3 Molecular biology techniques**

### **2.3.1 Plasmid DNA extraction**

QIAGEN<sup>®</sup> plasmid preparation columns were used to prepare stocks of DNA. Briefly, for DNA preparation in mini scale, 5 ml of an overnight bacterial culture grown in LB broth supplemented with the appropriate antibiotic (mainly Ampicillin at final concentration of 100  $\mu$ g/ml) was used. Bacterial cells were pelleted by centrifugation at 2500 g for 5 min and resuspended in 250  $\mu$ l of buffer P1 (provided in the kit). The cells were lysed by the addition of 250  $\mu$ l of buffer P2 (provided in the kit). The tube containing the lysate was inverted 4-6 times to mix the solution thoroughly. The mixture was neutralized using 350  $\mu$ l of buffer N3 (provided in the kit), and centrifuged at 18000 g for 15min. The supernatant was transferred to a QIAprep spin column and centrifuged down for 1 min to remove the solution from the column leaving the DNA in

the column. The column was washed using 0.75µl of buffer TE and centrifuged for 1min. DNA was eluted using 50µl of the elution buffer provided in kit (EB buffer, provided in the kit).

For larger scale DNA preparation (maxi-preparation), 500 ml of freshly prepared bacterial culture was used. A cell lysate was obtained by adding 10 ml of P2 buffer to 10 ml of resuspended cells in P1 buffer. The lysate was neutralized using 10 ml of chilled P3 buffer. A QIAGEN-tip 500 column was equilibrated by applying 10 ml of QBT buffer (provided in the kit). The supernatant was transferred to the column and washed twice by adding 30 ml of QC buffer (provided in the kit). The buffer was passed through the column by gravity. DNA was eluted in 15 ml of buffer QF (provided in the kit) and precipitated using 10.5 ml of isopropanol. The DNA pellet was washed in 5 ml of 70% ethanol, air dried and reconstituted in an appropriate volume of buffer EB provided in the kit.

### **2.3.2 DNA Digestion**

To carry out a single or double digestion with a restriction endonuclease, the appropriate amount of DNA was diluted into the appropriate buffer for the enzyme according to the manufacturers' instructions. BSA was added to the DNA mixture if required. At the final step the relevant enzyme(s) were added to the mixture and the reaction was incubated according to the enzyme manufacturers' instructions.

### **2.3.3 Gel extraction**

To prepare DNA fragments from agarose gels, the DNA was separated by electrophoresis in a 1% agarose gel in 1x TBE buffer. The desired fragment was excised from the agarose gel using a new razor blade or scalpel blade. Gel extraction was carried out using a QIAGEN<sup>®</sup> gel extraction kit following the manufacturer's instruction. Briefly, the excised gel was dissolved at 65°C in the QG buffer (provided in the kit) at pH 7.5. The solution containing the DNA fragment was passed through the silica membrane column provided. The DNA trapped in the column was eluted with the buffer provided.

### **2.3.4 DNA ligation and low melting point agarose dependent ligation**

T4 DNA ligase (Fermentas) was used to ligate DNA fragments together. To avoid repeated freeze and thaw the 10x ligation buffer provided with the enzyme was

aliquoted in 2  $\mu$ l volumes and used only once. DNA fragments ready to be ligated (insert and vector) were mixed with water, buffer and enzyme to achieve 20  $\mu$ l of reaction in total. The reaction mix was incubated at 16°C for 2 hr. In ligations of two fragments with cohesive ends, the molar ratio of insert and vector was adjusted in such a way that insert was 3 times more than the vector. For ligation of fragments with blunt ends 3-fold excess of vector was used. Aliquots of the ligation product were used in transformation of *Escherichia coli* (*E. coli*)(section 2.3.6).

In situations where it was not possible to get an appropriate concentration of a DNA fragment after gel purification, low melting point (LMP) Agar (TopVision® Genetic Quality LMP, Fermentas) was used. LMP agarose gels were prepared in 1x TAE buffer. The electrophoresis was conducted at 4°C. The desired fragment was excised from the gel and the agarose containing the DNA was melted at 60°C. The molten agarose was added directly to ligation reaction. The concentration of LMP agarose was calculated not to exceed 5% of total volume (20  $\mu$ l) of the ligation mix. Finally, the ligation tube was incubated at appropriate temperature, either at 16°C for 3 hr, or at 4°C for overnight ligation.

### **2.3.5 Preparation of competent bacterial cells**

Competent bacterial cells were produced using the method of Inoue, *et. al.* (1990). A freshly prepared *E. coli* liquid culture originating from one single colony, was grown at 18-20°C in a 250 ml volume of LB medium until the optimal turbidity was achieved ( $A_{600} = 0.6$ ). The cultures were centrifuged at 2500 g for 10 min at 4°C in 50 ml volumes, and pellets were resuspended in 20 ml of ice cold TB solution, individually. The suspensions were incubated in an ice bath for 10 min and bacteria were pelleted as before. The pellets were gently resuspended in 5 ml of TB solution individually. DMSO was added with gentle swirling to a final concentration of 7% (v/v). Following incubation in an ice bath for 20 min, the cell suspension was dispensed into 100 or 200  $\mu$ l aliquots. The aliquots were immediately frozen using liquid nitrogen and kept at -70°C until required.

### **2.3.6 Heat-shock based transformation of *E. coli***

The competent cell aliquots were thawed in an ice bath. The DNA suspension was added in a volume less than 10% of the volume of competent cell solution and mixed gently prior to incubation on ice for 45 min. The incubation was followed by a

heat shock at 42°C for 1 min. Sterile LB medium was added to make 1 ml of bacterial suspension and this was transferred into a sterile polypropylene tube and incubated at 37°C with vigorous shaking. The transformed bacteria were grown on LB agar plates containing the appropriate antibiotic.

### **2.3.7 RNA isolation**

RNA was isolated using Trizol-LS<sup>®</sup> (Invitrogen) reagent and manufacturer's instructions were followed thoroughly. To minimize the effect of cell culture passages on the sequence of viruses, the RNA extraction was performed directly using virus stocks with known titre. The number of the passages of the virus is specified where required. The yield of RNA extracted from 0.25 ml of virus suspension was enough for RT-PCR experiments (2.3.10).

### **2.3.8 pBS-T vectors preparation**

*EcoRV* digested pBlueScript II vectors (2961 bp) were purified from a 1% agarose gel (Section 2.3.3). T residues were added to the 3' ends of the fragment in a polymerase reaction containing 5 units of *Taq* DNA polymerase, 100 µM of dTTP and *Taq* DNA polymerase buffer (Fermentas) containing 1.5 mM Mg<sup>2+</sup>. The reaction was conducted at 72°C for 2 hr. The reaction product was phenol:chloroform purified (Section 2.1.6) and self ligated to remove the digested DNA lacking the 3'T residues. The linear fragment of pBS-T, with an estimated molecular weight of 3000 bp, was visualized and purified on a 0.7% agarose gel.

### **2.3.9 Blue and white colony screening**

From IPTG and X-Gal stocks (Section 2.1.13), 1 ml of each was added to 1 litre of molten LBA (Section 2.1.4) ready to be dispensed into 10 cm diameter bacterial culture dishes. The appropriate antibiotic for selection was added, mixed evenly and poured into the dishes.

### **2.3.10 Reverse transcription polymerase chain reaction (RT-PCR)**

For producing cDNA from viral RNA templates RevertAid<sup>™</sup> H minus reverse transcriptase (Fermentas) was employed. The reaction was set following the manufacturer's guidance. Briefly, RNA samples (0.1 – 5 ng), gene specific (10 µM) or random primers (0.2 µg or 100 µM) and RNase free water were mixed together to

obtain 12.5  $\mu$ l of the mixture. At this point, the sample was heated to 65°C for 5 min and chilled on ice immediately. The reaction buffer containing DTT, RiboLock™ RNase inhibitor (Fermentas), dNTPs (100  $\mu$ M of each dATP, dCTP, dGTP and dTTP) and 200 units of the reverse transcriptase were added to the mixture to end up having a total volume of 20  $\mu$ l. The sample was incubated at 25°C if random primers were used and was followed with incubation of samples at 42°C for 60 min. Otherwise the sample was incubated for 43°C for 60 min. To clear the cDNA product from RNA templates, incubation of the sample with RNase H (Fermentas) for 30 min was carried out, and was followed by stopping the RNase reaction at 65°C for 30 min. The cDNA was phenol:chloroform (section 2.1.6) purified and ethanol precipitated, and was used appropriately in PCR reactions.

### **2.3.11 Polymerase Chain Reactions (PCR)**

The polymerase chain reaction technique was employed for three major purposes:

1. Amplification of specific DNA fragments from DNA or cDNA templates using standard PCR procedure for detection, cloning or sequencing purposes.
2. Introducing site-specific mutations into a target plasmid DNA.
3. Ligase independent adhesion of DNA fragments; which was achieved using overlap PCR technique.

The use of PCR technique in site specific mutation and overlap PCR is explained separately in sections 2.3.12, 2.3.13 and 2.3.14. In the standard PCR for sequencing and detection, *Taq* DNA polymerase (Fermentas) enzyme was employed unless otherwise specified. For PCR to produce DNA for cloning purposes, *Pfu* DNA polymerase (Fermentas) was used. To minimize the error rate in DNA amplification, the proofreading KOD DNA polymerase (Novagen) was employed to produce DNA fragments from single-stranded cDNA templates for sequencing purposes. It is specified in the text if a sequence was obtained from DNA fragments amplified with KOD DNA polymerase enzyme.

For the standard PCR, the reaction mixture was prepared following the manufacturers' guidance. Each PCR mixture contained 100  $\mu$ M of appropriate pair of primers, the appropriate PCR buffer, 2 mM of Mg<sup>2+</sup> salt, and dNTP mix (100  $\mu$ M of the

each of dATP, dCTP, dGTP, and dTTP). Primers were designed using Lasergene PrimerSelect software. In the primer design, extra care was taken to avoid secondary structure formation, intra molecular interaction and GC content less than 50%. Most of the primers were designed ending at their 3' end with C or G rich content to prevent “breathing of 3' ends” unless the situation was inevitable. The PCR condition was adjusted according to enzyme manufacturers' instructions.

### **2.3.12 Quick change site-directed mutagenesis**

Quick change site-directed mutagenesis was carried out using a suitable proofreading polymerase enzyme. For fragments less than 5kb, *Pfu* polymerase (Fermentas) was used, and for fragments with longer length, Phusion polymerase (NEB Biolabs) was used. The thermocycler machine was programmed to run 18 cycles to amplify DNA based on the enzyme manufacturers' instructions. The extension time was calculated for each enzyme specifically (2 min per 1 kb for *Pfu* polymerase, and for 30 seconds for Phusion polymerase). Thereafter, the reaction mixture was digested with *DpnI* enzyme to remove the parental DNA. Aliquots of the *DpnI* digested mixture were used to transform competent *E. coli*.

### **2.3.13 Primer design for Quick change site-directed mutagenesis**

Primers designed for the mutagenesis reaction had three main characteristics: they were complementary and designed to anneal to the same sequence on opposite strands of the plasmid; they had the mutation site embedded in the middle of the primer; and they were not shorter than 25 base pairs with a  $T_m$  of not less than 78°C. The  $T_m$  was calculated using the equation:  $T_m = 81.5 + 0.41(\%GC) - 675/N - \%mismatch$ .

The “%mismatch” was defined as zero for calculating  $T_m$  for primers intended to introduce insertions or deletions, and for mutagenesis was calculated by dividing the number of nucleic acids to the total length of the primer and multiplied to 100. An annealing temperature of 55°C was used uniformly.

### **2.3.14 “Round the horn” site directed mutagenesis PCR**

This technique was designed based on the standard PCR. The primers used in this mutagenesis technique were designed in a way that they abut at the 5' ends (Figure 2.1), and ligation of the fragments amplified with these primers would produce the plasmid carrying the desired mutation or lacking a specific part of the template

plasmid to induce a deletion. Before running the PCR reaction the primers were phosphorylated at their 5' ends using T4 polynucleotide kinase (Fermentas). The PCR reaction was carried out as described in section 2.3.11. The resulting PCR product was digested with *DpnI* to remove the parental, methylated, plasmid template. The PCR product was separated on an agarose gel and the desired fragment was excised from the gel. The purified DNA was ligated to itself using T4 DNA ligase (Fermentas). The ligation mixture was used to transform competent *E. coli*, and the recombinant bacteria were grown in the presence of the appropriate antibiotic. Single colonies were isolated from the culture plates and plasmid DNA was isolated for confirmation of insertion of the described mutation.

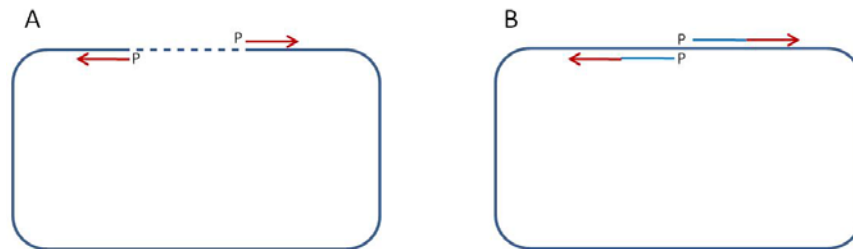


Figure 2.1 “Round the horn” site directed mutagenesis PCR. A. The pair of primers was designed at both sides of the deletion site. The deletion site is depicted in dotted lines. B. Primers carrying the mutagenesis sites were designed abut each other. The tail of the primers carrying the mutated sequence is depicted in light blue color. After completion of the DNA amplification, pre-phosphorylated primers were ligated together to obtain the full plasmid.

### 2.3.15 DNA sequencing

All sequencing samples were prepared individually in 0.5 ml tubes or in 96 well sequencing grade plates. It was ensured the samples were containing the advised amounts of template DNA and primer. The samples were sent to the genomics facilities of university of Warwick, department of life sciences, and they were sequenced using Sanger sequencing method. The sequencing was performed using ABI PRISM 3130xl Genetic Analyser systems.

### 2.3.16 Nucleotide similarity searches

MUSCLE on-line software (available on EBI server: <http://www.ebi.ac.uk/Tools/sequence.html>) was used to find sequence similarities between DNA fragments. Seqman from the DNA star laser gene suit (Lasergene) was used to compile sequencing results.

### **2.3.17 SDS - polyacrylamide gel electrophoresis (SDS-PAGE)**

To prepare a mini SDS-PAGE gel (BioRad Protean III gel), the components (Table 2.1.A) were mixed and the gel was cast into the sealed mould. About two third of the cast was filled with the resolving gel (Table 2.1.A) and the edge of the gel was levelled and covered with isopropanol to remove air bubbles from the gel liquid and to prevent the gel from oxidation. The ready stacking gel mix (Table 2.1.B) was added to the mould containing polymerised and washed resolving gel. At this point the teflon comb was inserted between the two glass sheets and into the liquid gel to prepare wells. The gel was left until complete polymerisation was achieved in the stacking gel. Then combs were removed and the wells were washed with distilled water to remove any non-polymerised acrylamide. The electrophoresis unit was assembled based on the manufacturer's instruction and filled with the SDS-PAGE electrophoresis buffer (section 2.1.8). Samples were prepared in the electrophoresis sample buffer (section 2.1.7), and boiled for 5 min or incubated at 37°C for 30 min, and the samples were applied into the prepared wells. A voltage of 8 volt/cm was applied to the unit until the bromophenol blue dye reached to the resolving gel. A voltage of 15 volt/cm was applied until the bromophenol blue dye reached to the bottom of the gel.



**A**

Ingredients	Amount required for 10 ml of 8% resolving SDS-PAGE gel	Amount required for 10 ml of 10% resolving SDS-PAGE gel	Amount required for 10 ml of 12% resolving SDS-PAGE gel
H <sub>2</sub> O	4.6	4	3.3
30% Acrylamide / 0.8% bisacrylamide mix	2.7	3.3	4
1.5 M Tris (pH 8.8)	2.5	2.5	2.5
10% SDS	0.1	0.1	0.1
10% ammonium persulfate	0.1	0.1	0.1
TEMED	0.006	0.004	0.004

**B**

Ingredients	Amount required for 10 ml of 5% stacking SDS-PAGE gel
H <sub>2</sub> O	1.4
30% Acrylamide mix / 0.8% bisacrylamide	0.33
1.5 M Tris (pH 6.8)	0.25
10% SDS	0.02
10% ammonium persulfate	0.02
TEMED	0.002

Table 2.1 Ingredients to prepare a 8%, 10% and 12% SDS-PAGE electrophoresis gel. Ingredients were mixed in order and poured into the prepared and sealed cast as it is explained in the text. A. Ingredients for preparing the resolving gel. B. Ingredients for preparing the stacking gel.

### 2.3.18 Western blot

A Trans-Blot<sup>®</sup> SD semi-dry electrophoretic transfer cell was used. Prior to the transfer, the SDS-PAGE gel was trimmed and equilibrated with transfer buffer (24 mM Tris base, 192 mM glycine, and 20% methanol) for 5 min. Whatman papers and polyvinylidene fluoride (PVDF) membranes were cut to the dimensions of the gel. The PVDF membrane was soaked in methanol for 2 min and thereafter was soaked in the

transfer buffer. A stack of Whatman papers was made and soaked with the transfer buffer. These were placed on top of the anode electrode. Any air bubbles trapped between the soaked membranes were removed. The PVDF membrane and the gel were placed on the stack of the Whatman papers, leaving the gel on top of the PVDF membrane. More Whatman papers soaked with transfer buffer were then placed on the gel. Air bubbles between the layers were removed and the cathode plate was placed carefully on the stack leaving the stack in the space provided between the cathode and the anode. Constant voltage of 10 volt was used to perform the transfer for 20 min and the PVDF membrane was then removed for further processing.

A blocking procedure was performed on the membrane with 5% bovine serum albumin (BSA) mixture in TBS (Section 2.1.12) in a rocking platform, at room temperature for 1 hr. The blocking was followed by washing the membrane in TBST (Section 2.1.12) 3 times each for 15 min. The membrane was then incubated with rabbit anti-GFP polyclonal antibody (Abcam) diluted 1:5000 in TBS containing 5% BSA for 1 hr. The membrane was then rinsed as before and incubated with horse radish peroxidase (HRP)-conjugated goat anti-rabbit antibody at 37°C and for 1 hr. The membrane was rinsed, and incubated with Lumi-Light western blotting substrate (Roche) for 5 min according to the manufacturer's guidance. The membrane was blot-dried and transferred to a radiography cassette. X-ray medical film (Fuji) was exposed to the membrane for different periods of time. The exposed films were developed using a Curix 60 Agfa film developing machine.

### **2.3.19 Fluorescent detection of PVM P protein**

A monolayer of the BS-C-1 cells were prepared on a sterile glass coverslip and subsequently infected as required for the experiment. The cells were fixed at room temperature for 30 min using an ice cold methanol:acetone (50:50) mixture. Subsequently, the cells were washed 3 times with PBS for 15 min each time. The cells were then incubated in 5% BSA in PBS for 1 hr. The incubation was followed by 3 washes, each for 15 min as before. Monoclonal antibody against the PVM P protein, designated 26/3/B5, (Ling & Pringle, 1989a) was used at a 1/40 dilution in PBS. The cells were incubated with the antibody for 1 hr. The incubation was followed by 3 washes with PBS as before. The coverslips were then incubated with anti mouse Alexa<sup>®</sup> Fluor 488 goat anti-mouse IgG (Invitrogen), (1:3000) diluted in PBS containing 1% (w/v) BSA. From this point, the experiment was carried out in the dark. The coverslips were washed in

PBS 3 times, each for 15 min and incubated with 0.2  $\mu\text{g}$  of DAPI (Sigma) in PBS for 5 min, and were washed 3 times as before and, air-dried, and mounted onto glass slides using Vectorshield (Vector laboratories) and sealed and fixed using fast dry nail varnish.

Fluorescence was analysed using a Leica SP2 confocal microscope linked to a DM RE7 upright microscope. AlexaFluor488 was excited at 488 nm and the emitted light between 505 and 550 nm was recorded. DAPI was excited at 405 nm and the emitted wavelength between 410 and 550 nm was collected. Images were analyzed with Leica software.

## 2.4 Reporter genes assay

The eGFP and Luciferase genes were employed as the reporter genes in the experiments (Chapter 3). To quantify the expression level of GFP and the activity of luciferase, GFP and luciferase assays were used.

### 2.4.1 GFP ELISA

The required numbers of wells (ELISA Immulon 2HB plates) were coated with a 1:3000 dilution of 1 mg/ml stock of goat anti-GFP polyclonal IgG antibody (Rockland) in coating buffer (Section 2.1.3). The binding process was carried out at 4°C overnight. After binding, the wells were washed with PBS/Tween (0.1% Tween 20 in 1x PBS) 3 times. Irrelevant non-specific interactions were blocked using 0.2 ml of blocking buffer (1% BSA in PBS) for 2 hr at room temperature. The ELISA plate was dried and after being wrapped in cling film, and stored at -20°C until required.

Prior to the usage of coated ELISA plates, the antibody coated wells were rehydrated with 0.2 ml/well of PBS/Tween. Standard GFP (Clontech) was prepared in the blocking buffer from 300 to 0 pg concentrations, and used to plot a standard curve to quantify the amount of GFP in each sample. Thereafter, desired samples and the previously prepared GFP standards were added to wells. After 1 hr of incubation at 37°C a rabbit anti-GFP polyclonal antibody (Abcam) was diluted to 1:32000 in PBS/Tween (0.1% Tween 20 in 1x PBS) and was added to the wells. After 1 hr incubation at 37°C the wells were washed with PBS/Tween. Anti-rabbit goat HRP conjugated antibodies (Bio Rad) were diluted to 1:1000 in PBS/Tween (0.1% Tween 20 in 1x PBS) and added to the wells. The ELISA plate was incubated for 1 hr at 37°C, followed by 5 washes with PBS Tween. The ABTS substrate was added until a visible

green color appeared in wells treated with standard amounts of GFP. The absorbance was measured using a Labsystems Multiskan RC plate reader machine at a wave length of 405 nm. The OD<sub>405</sub> of samples were converted to concentrations using the standard curve using the Labsystems Multiskan software.

#### **2.4.1.1 Rabbit anti-GFP polyclonal antibody**

Rabbit anti-GFP polyclonal IgG known as Ab290 (Abcam) was diluted in water to obtain a 1:80 dilution of the antibody. The antibody-water mixture was diluted in glycerol to obtain a 1:1 mixture of antibody and glycerol with a final concentration of 1:160 of the antibody. The solution was kept at -20°C in small aliquots until the time of usage.

#### **2.4.2 Luciferase activity assay**

Prior to the measurement of luciferase activity, the wells containing transfected cells were washed with PBS and lysed in Luciferase Cell Culture Lysis Buffer (CCLR) reagent (25 mM Tris-Phosphate, 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid, 10% glycerol, and 1% Triton® X-100). Cells were scraped from the surface of the wells into the CCLR buffer. This material was centrifuged for 1 min at 18000 g. Finally, 10 µl of cell lysate was transferred to each opaque well of white luminescence micro-wells (LumiNunc, Nunc), and 50 µl of luciferase assay reagent (Promega) was added to the wells containing the cell lysate. The light intensity was measured immediately for a period of 10 seconds using a Luminoskan® Ascent microplate luminometer (Thermo Scientific) immediately after addition of the luciferase assay reagent.

### **2.5 Animal studies**

In this research 4-8 weeks old BALB/c male or female mice were obtained from in-house facilities and Charles river laboratories. The mice obtained from Charles river laboratories were specific pathogen free (SPF) and were used to breed mice in the in-house facilities. To study the pathogenesis of PVM, BALB/c mice were anesthetized with intraperitoneal injection of ketamine and xylazine cocktail (Section 2.1.5), and after obtaining full anesthetic condition were infected intranasally with 50 µl of the known titre of virus suspension. A group of control mice were inoculated intranasally with PBS. The weights of mice in each group were monitored on a daily basis. Clinical

examination was performed every day and signs of infection were recorded using the scoring system (Table 2.2) developed by Cook, *et. al.* (1998). Mice were euthanized by cervical dislocation if weight loss exceeded 25%. Experiments were conducted for a minimum of 14 days, and at the end of the experiment, all of the remaining mice were euthanized by cervical dislocation. Subsequently, the lungs were removed for virus isolation. Removed tissues were kept at -70°C until the day required for analysis.

Clinical score	Clinical signs
1	Healthy with no signs of illness
2	Consistently ruffled fur, especially on neck, Piloerection, breathing may be deeper and mice less alert
3	Laboured breathing. Frequently showing tremors and lethargy
4	Abnormal gait and reduced mobility. Laboured breathing. Frequently emaciated. May show cyanosis of tail and ears
5	Death

Table 2.2 Clinical scoring system from (Cook *et. al.* 1998). Five categories were defined: Categories 1 and 6 were defined as healthy and dead respectively. Categories 2 was defined as sick. Category 3 was defined as sicker. Category 4 was defined as very sick with the signs of cyanosis and abnormal gait in the infected mice.

## **CHAPTER 3**

# **PATHOGENESIS OF PVM**

### 3.1 Introduction

As discussed in chapter 1, PVM has at least two distinguishable strains. PVM strain 15 (Warwick) has been shown to be non-pathogenic and PVM strain J3666 and PVM strain 15 deposited with the ATCC were reported as being pathogenic (Anh *et al.*, 2006; Cook *et al.*, 1998; Domachowske *et al.*, 2004; Krempl & Collins, 2004; Rosenberg *et al.*, 2005). Sequence analyses identified several differences between the two pathogenic and the non-pathogenic strains raising the possibility of the contribution of the sequence differences to the pathogenesis that were observed (Krempl *et al.*, 2005; Thorpe & Easton, 2005). Among these, the differences in the G gene, encoding the attachment glycoprotein of the virus, have been linked with the pathogenesis (Krempl *et al.*, 2005; Krempl *et al.*, 2007; Randhawa *et al.*, 1995).

In many cases the passage of viruses in tissue culture leads to a loss of pathogenicity (Claassen *et al.*, 2005; Cohen *et al.*, 1989; Dardiri, 1969; Hearn *et al.*, 1966; Taylor *et al.*, 1993). PVM strain 15 (ATCC) has been continuously passaged in mice since its deposition, whereas PVM strain 15 (Warwick) has been extensively passaged in BS-C-1 cells in tissue culture. It is possible that the passage of the virus in tissue culture has led to the loss of pathogenicity in mice (Cook *et al.*, 1998; Krempl & Collins, 2004). To investigate this, virus stocks were prepared following sequential passage of the pathogenic strain J3666 in tissue culture and the pathogenicity was investigated in mice. Stocks of the pathogenic PVM strain J3666 which had previously been passaged exclusively in mice were prepared by sequential passage in BS-C-1 cells. Initial analysis showed that five passages in tissue culture did not eliminate the ability of the virus to induce a fatal pneumonia in mice (Prof. A. J. Easton, personal communication).

### 3.2 Evaluation of the pathogenicity of a working stock of PVM strain J3666

It was necessary to prepare a reference stock of PVM strain J3666 for the experiments in this study. A working stock of PVM strain J3666 was prepared by taking as a seed stock virus which had been passaged 3 times in BS-C-1 cells and which retained pathogenicity. This was then used to prepare a large volume of a single stock by growth in BS-C-1 cells as described in section 2.2.6. The titre of the virus was calculated as  $1.6 \times 10^6$  pfu per millilitre. On the basis of previous data, it was

anticipated that 500 pfu of PVM was likely to be a lethal dose (A. Easton, personal communication). This was taken as the initial starting point and two-fold dilutions of virus were prepared for infection of mice.

Each group contained typically five mice. Following infection with the various dilutions of PVM strain J3666, as described in section 2.5, mice were checked daily. The body weight and the clinical score for each group in the experiment were recorded. The daily difference of the body weight throughout the experiment compared to that of the first day of the experiment and expressed as a percentage was calculated for each group. The data are shown in Figure 3.1. The clinical score for each mouse was monitored daily based on the scoring system described in section 2.5. The mean clinical score was calculated for each group and the data are shown in Figure 3.2.

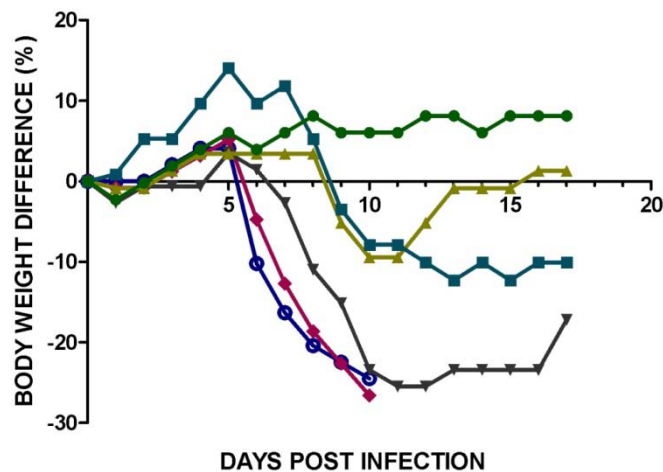


Figure 3.1. Weight loss of BALB/c mice infected with different amounts of PVM strain J3666 passaged 4 times in BS-C-1 cells. After inoculation with different doses of virus, mice were examined daily and the percentage of their body weight compared to day one was determined for each group and expressed as a percentage. Weight loss is presented until the first animal in any group died or was humanely killed. The data shown is representative of two independent experiments. The amount of virus (pfu) in each 50 µl inoculum: 500 pfu ●, 250 pfu ◆, 125 pfu ▲, 62.5 pfu ▼, 31.2 pfu ■, and 15.5 pfu ●.

For the groups that were inoculated with 500 and 250 pfu of PVM, weight loss began on day 6, and progressed quickly. At day 11 post-infection, one of the mice in the group infected with 250 pfu died due to the severe symptoms of the infection. The severity of disease for mice inoculated with either 500 pfu or 250 pfu was sufficiently extensive on day 11 that all animals were humanely killed to prevent further suffering (Figure 3.2), as required by the Home Office license under which this work was done.



The group which received 125 pfu of the virus first showed weight loss on day 6 post infection, as for the groups receiving 500 and 250 pfu. However, the rate of weight loss subsequently proceeded at a slower pace. The rate of weight loss in this group of mice reached 25.47% on day 11 post infection day and began to recede on day 13 as the mice recovered. The mice did not achieve the starting weight by the end of the experiment, and one of the mice did not recover from the infection (Figure 3.1).

Mice in the groups inoculated with 62.5 and 31.2 pfu of virus were the last to show onset of weight loss which first occurred on day 9. The weight loss was transient and mild, with a maximum of 9.48% for the mice treated with 62.5 pfu. This group of mice recovered quickly from the infection. The group of mice infected with 31.25 pfu of the virus, however, did not fully recover to their original weight by the end of the experiment (Figure 3.1). Mice treated with 15.6 pfu of PVM did not show any weight loss. These data demonstrated that an inoculum containing 250 pfu of this virus stock was a lethal dose for BALB/c mice.

In mice infected with 500 pfu and 250 pfu inocula, the signs of disease started from days 5 and 6 post infection and progressed rapidly until day 9 when the mice were culled for humane reasons. The onset of clinical signs of disease in the group of mice infected with 125 pfu was on day 7 post infection. In the group infected with 125 pfu, the signs of disease progressed until day 10 when they reached a plateau and the mice started recovering from day 12 post infection. In the groups infected with 62.5 and 31.2 pfu the signs of disease started on day 10. The group infected with 62.5 pfu fully recovered while the group infected with 31.2 pfu showed the lowest clinical score until the end of the experiment. Clinical signs of disease were not observed in the group infected with 15.5 pfu (Figure 3.2).

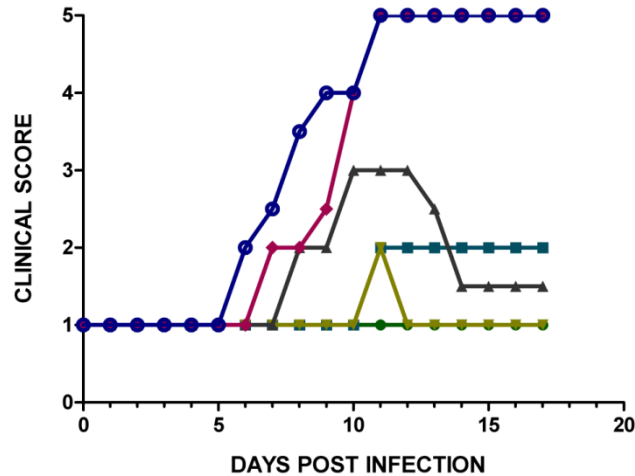


Figure 3.2. Clinical score of BALB/c mice infected with different amounts of PVM strain J3666 passaged 4 times in BS-C-1 cells. The scoring system described in section 2.5 was used to show the severity of infection in mice infected with PVM strain J3666. Mean clinical score for each group of mice was calculated and shown in the graph. The data shown is representative of two independent experiments. Clinical score is presented until the first animal in any group died or was humanely killed. The amount of virus (pfu) in each 50µl inoculum: 500 pfu —●—, 250 pfu —◆—, 125 pfu —▲—, 62.5 pfu —★—, 31.2 pfu —■—, and 15.5 pfu —●—.

### 3.3 Effect of consecutive passages of PVM strain J3666 on the pathogenicity of the virus

As described above, passage in tissue culture has been shown to reduce pathogenicity of many viruses and it was of interest to investigate this for PVM. If it was possible to generate stocks of non-pathogenic virus from a pathogenic virus it may be possible to investigate the changes responsible for the altered characteristics. The virus that was tested by Cook and colleagues (Cook *et al.*, 1998) had been subjected to 5 continuous passages in cell culture after being isolated from mice lungs and retained its pathogenicity (Prof. Andrew Easton, personal communication). In the work presented here, the ability of the virus from the same number of passages and other virus stocks from further consecutive passages (from passage 6 to 10) to generate disease in mice was tested.

For this purpose successive passages from the stock virus (passage 4) described in section 3.2, were made on BS-C-1 cells, and the number of passages was extended to 10. The viruses were kept at -70°C until the day of challenge. The titre of the stock prepared from each passage was calculated and BALB/c mice were infected with 5000, 500 and 250 pfu of virus in an inoculum volume of 50 µl.

As can be seen in Figure 3.3, virus from passages 5 and 6 were capable of producing clinical signs of disease and weight loss when 5000 pfu was administered. The progress of weight loss and clinical signs were more rapid in the passage 5 compared to the passage 6. In mice infected with passage 5 virus, as for passage 6 virus, weight loss and clinical signs of disease started in day 6. The weight loss alongside the clinical syndrome proceeded quickly and led to very severe disease in passage 6, but in passage 6 from day 6 onwards clinical signs and weight loss progressed until day 10. At day 10, both infection and clinical symptoms reached a plateau, and started to increase from day 10 onwards resulting a second phase of the infection with more severe signs. Similar to passages 5 and 6, the onset of clinical signs of disease and weight loss in the group infected with passage 7 virus was day 6, and the symptoms reached to a plateau in the same period of time as in passage 6, but in contrast with passage 6 virus, mice started to recover quickly from day 10 onwards indicating the virus after 7 successive passages has lost its virulence (Figure 3.3.A and B).

Mice inoculated with passage 4 virus (Section 3.2) demonstrated clinical signs and weight loss from day 5 when they were inoculated with 500 pfu of the virus (Figures 3.1 and 3.2). Comparing the results with those obtained from mice inoculated with the same amount of virus from passage 5 indicates that the the onset of clinical manifestations is on the same day and the virulence of the virus from both passages are similar. However, comparison between the clinical signs demonstrated by mice inoculated with 250 pfu of the passages 4 and 5 indicate that the virus in passage 5 is less virulent than the virus in passage 4, as the mice inoculated with passage 5 started recovering from the infection while mice inoculated with passage 4 did not recover.

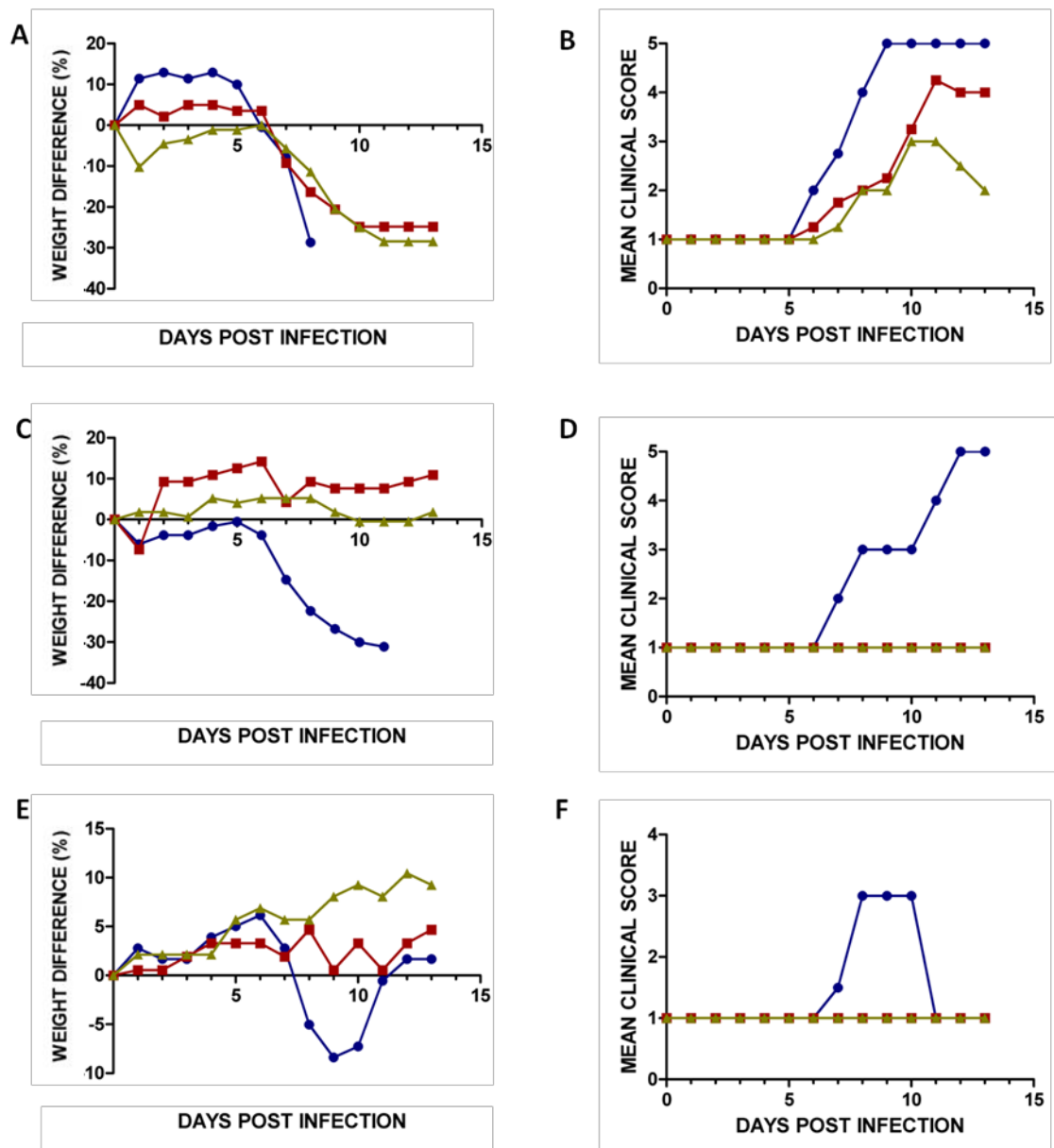


Figure 3.3. Weight loss and clinical score of mice infected with 5000, 500 and 250 pfu of PVM passages 5 (A & B), 6 (C & D) and 7 (E & F). The body weight differences (A, C & E) and clinical scores (B, D & F) were monitored on a daily basis. Weight loss is presented until the first animal died or was humanely killed because of the severe symptoms. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, and 250 pfu ▲.

In the groups of mice infected with 500 pfu of the passage 5 virus stock the clinical score did not begin to rise until day five post infection (Figure 3.3.C). Clinical signs and weight loss developed quickly among the mice within the group (Figure 3.3.C and D). In contrast with the results achieved from passage 5, no sign of disease

appeared in the groups in which mice were infected with 500 pfu of passages 6 and 7 virus stocks, indicating that infection with passage 5 is more severe than the infection with passage 6, and the infection with passage 6 is more severe than the infection with passage 7.

Data obtained from groups of mice infected with 250 pfu of passages 5, 6 and 7 confirmed the above conclusions. Among the mice infected with 250 pfu, only the passage 5 virus stock was able to cause clinical signs of disease. As can be seen in Figure 3.3.E and 3.3.F, mice started to recover from the clinical signs of infection with this stock from day 11 onwards. However, the rate of recovery was slow and the mice in this group did not fully recover by the end of the experiment.

### **3.4 Analysing the effect of consecutive passages on virus genome nucleotide sequence**

Spontaneous mutagenesis is a characteristic of RNA viruses. It has been calculated that the mutation rate for RNA viruses is between  $10^{-3}$  to  $10^{-6}$  per site per replication (Schrag *et al.*, 1999). As a result, RNA virus populations evolve into genetically related but diverse subpopulations known as quasispecies. Here, it was decided to analyse the sequence of the PVM strain J3666 stocks used in the pathogenicity study described in Section 3.3.

While it was desirable to determine the entire genome sequence of the stocks, the limited amount of the stock of PVM strain J3666 passages 5, 6 and 7 made this impractical. Therefore, an alternative technique was chosen to amplify large fragments of the genome and sequence the areas that are more prone to spontaneous mutagenesis. The complete sequence of PVM strains J3666, 15 (Warwick) and 15 (ATCC) have been determined (Krempf *et al.*, 2005; Thorpe & Easton, 2005). Comparison between the sequences suggests little variation among the NS1, NS2 and L genes of PVM. Therefore, to concentrate on the parts of the genome that are more prone to mutagenesis, the rest of the genome except the NS1, NS2, and L genes were sequenced.

The approach employed involved amplifying four fragments (F2, F3, F4 and F5; Figure 3.4). This would amplify from the N gene to the first third of the L gene. For this purpose, total RNA from PVM strain J3666 passages 5, 6 and 7 stocks (Section 3.3) was prepared. The synthesis of the cDNA for fragment 2 used primer pair P15FLN3.2 and P15FLM2.2 (Appendix I), for fragment 3 primer pair P15FLSH3.2 and P15FLF2.2

(Appendix I), for the Fragment 4 primer pair P15FLM23 and P15FLM25 (Appendix I), and for the Fragment 5 primer pair P15FLL5 and P15FLL7 (Appendix I). The fragments were then sequenced in their entireties and the data assembled. A comparison with the sequence of PVM strain J3666 (Accession number: NC\_006579) was made.

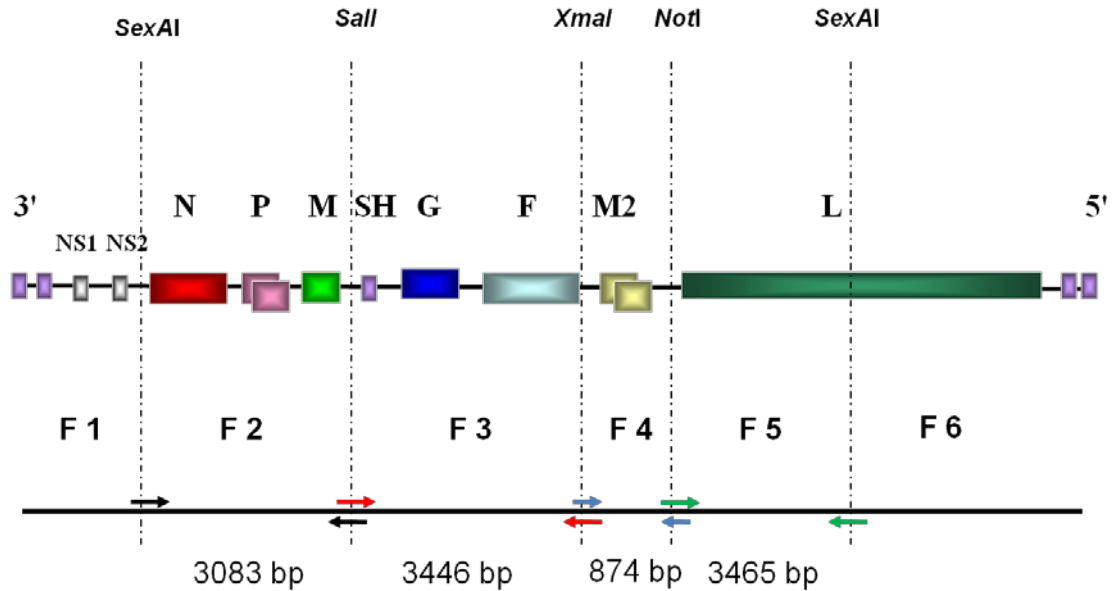


Figure 3.4 The PVM genome and the location of primers on the genome. Primer pair P15FLN3.2 and P15FLM2.2 amplifying fragment 2 (F2) with 3083 bp in length are shown in black (F2), primer pair P15FLSH3.2 and P15FLF2.2 amplifying fragment 3 (F3) with 3446 bp length are shown in red, primer pair P15FLM23 and P15FLM25 amplifying fragment 4 (F 4) with 874 bp in length is shown in cyan, and primer pair P15FLL5 and P15FLL7 amplifying fragment 5 (F 5) with 3465 bp in length is shown in green.

### 3.4.1 The fragment 2 region

Primer pair P15FLN3.2 and P15FLM2.2 was used to amplify the fragment 2 from cDNA templates. DNA amplification was carried out using KOD DNA polymerase (see 2.3.11) to reduce the possibility of causing random mutations during DNA amplification. As indicated in Figure 3.4, fragment 2 contains the N, P and M genes. The sequence of the fragment 2 was obtained using primers NS2B, N1, N5, N3, N4, P.3, P2F, P2R, P1, P.2, and P15FLM2.2 (Appendix I). No sequence conflict reflecting spontaneous mutations was identified.

### 3.4.2 The fragment 3 region

Fragment 3 was amplified from cDNA made from PVM strain J3666 passages 5, 6 and 7 using P15FLF2.2 and P15FLSH3.2 (Figure 3.4). The sequence of fragment 3 was obtained using primers SH2, G7, P15FLSH3, G(4878)R, SH32, SH4360, G9, F14, F7, F9, F22, F11, P15FLSH3.2, F22, and P15FLF2.2 (Appendix I). Unlike the sequence of the fragment 2, several sequence changes were identified in the sequences of fragment 3 of the passages 5, 6 and 7 virus stocks when they were compared against the published sequence of the strain J3666 (Accession number: NC\_006579). All three passages of the virus stock had the same sequences, though they differed from the published sequence of PVM strain J3666 (Accession number: NC\_006579). The identified sequence differences are summarised in Table 3.1.

Compared with the sequence of PVM strain J3666 (accession number: NC\_006579), the SH gene from passage 5, 6 and 7 viruses carry 6 point mutations: a silent mutation at position 70 (C→U), a mutation at position 269 (C→U) resulting in an H<sub>87</sub> → Y<sub>87</sub> change in the amino acid sequence of the SH glycoprotein, a mutation at position 283 (C→U) causing a premature stop codon, two mutations at positions 292-293 (CC→UU) and a mutation at position 299 (C→U). Table 3.1 summarizes the mutations affecting the sequence of SH gene in PVM strain J3666 passages 5, 6 and 7. The amino acid sequence of the SH glycoprotein is compared against the sequence of the SH glycoprotein of PVM strain J3666, 15 Warwick and 15 ATCC in Figure 3.5.

The most striking difference in the G gene of the strain J3666 (Accession number: NC\_006579) and strain J3666 passages 5, 6 and 7 was the mutation at nucleotide 65 of the G gene changing a UGA stop codon to AGA. As a result, the first ORF of the G gene (Sections 1.2.1.1 and 3.5) was placed in frame with the second ORF such that the G glycoprotein in strain J3666 passages 5, 6 and 7 is 18 amino acid longer at the N-terminal than that of PVM strain J3666 and PVM strain 15 (ATCC) (Section 3.5). This feature of the G glycoprotein of the passages 5, 6 and 7 was investigated in more depth and the results are discussed in Section 3.5. The other mutations appearing in the G genes of the strain J3666 passages 5, 6 and 7 were at positions 68 (C→U), 104 (A→G), 118 (C→U), 165 (G→U), 236 (G→A) and 1121 (U→A). The mutation at position 68 is in the noncoding region of the G gene of the strain J3666. The mutations at positions 104, 165, 236, and 1121 resulted in amino acid changes in the G glycoprotein. The amino acid sequence of the G glycoprotein of the strains J3666, 15

(ATCC), 15 (Warwick) are compared with the amino acid sequence of the G protein of PVM strain J3666 Passages 5, 6 and 7 in Figure 3.6.

PVM strain J3666	PASSAGE 5	PASSAGE 6	PASSAGE 7	NUCLEOTIDE CHANGE
C(4381)	U	U	U	SH.269 His → Tyr
C(4395)	U	U	U	SH.283 SL
C(4399)	U	U	U	SH.287 Stop
CC(4404-4405)	UU	UU	UU	SH.292-293 NC*
C(4411)	U	U	U	SH.299 NC (Stop)*
U(4575)	A	A	A	G.65 Amber → Lys
C(4578)	U	U	U	G.68 NC → Leu†
A(4614)	G	G	G	G.104 Ser → Gly
C(4628)	U	U	U	G.118 NS
G(4675)	U	U	U	G.165 Gly → Val
G(4746)	A	A	A	G.236 Gly → Arg
U(5631)	A	A	A	G.1121 Ser → Thr
U(6581)	G	G	G	F.721 Ile → Glu
U(6826)	A	A	A	F.966 SL
U(6852)	C	C	C	F.992 Val → Ala

Table 3.1 Nucleotide differences between PVM strain J3666 passages 5, 6 and 7 used in this study and PVM strain J3666 (Accession number: NC\_006579). The antigenomic sequences of the viruses were compared together. The genomic positions of the nucleotides in strain J3666 genome are provided in parentheses. The nucleotide changes in the passages 5, 6 and 7 are provided. The result of the nucleotide changes in the amino acid sequence of the SH protein is shown. The mutation site in the SH gene is shown in “gene.nucleotide position” format. The occurrence of the mutations in non-coding sequences of each gene is shown by NC and the silent mutations are indicated as SL. \* The effect of mutation is shown after considering the mutation at position 299 of the SH gene. † The effect of mutation is shown after considering the mutation at position 65 of the G gene.



```

*           20           *           40           *
J3666      : MDPNMTSYQITFEINMTSSRIGTYITLALTALLLACAVINTVCALIMACS : 50
15-Warwick : MDPNMTS QIT EINMTSSRIGT TT A TA LLACAVINTVCALIMACS : 50
15-ATCC    : MDPNMTS QIT EINMTSSRIGT TT A TALLLACAVINTVCALIMACS : 50
PVM-P.5    : MDPNMTSYQITFEINMTSSRIGTYITLALTALLLACAVINTVCALIMACS : 50
PVM-P.6    : MDPNMTSYQITFEINMTSSRIGTYITLALTALLLACAVINTVCALIMACS : 50
PVM-P.7    : MDPNMTSYQITFEINMTSSRIGTYITLALTALLLACAVINTVCALIMACS : 50
           MDPNMTS QIT EINMTSSRIGTy T A TALLLACAVINTVCALIMACS

           60           *           80           *           100
J3666      : SRSIATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNLHSRNTTQHHKQQKL : 100
15-Warwick : SRS ATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNL SRNTT*---*--- : 92
15-ATCC    : SRS ATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNL SRNTT*---*--- : 92
PVM-P.5    : SRSIATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNL SRNTT*---*--- : 92
PVM-P.6    : SRSIATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNL SRNTT*---*--- : 92
PVM-P.7    : SRSIATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNL SRNTT*---*--- : 92
           SRS ATSGIVSSQCTVHPNHPPPSYGVNVTGLPGNLySRNTT

*
J3666      : SFNKPQARQLYPAR* : 114
15-Warwick : ----- : -
15-ATCC    : ----- : -
PVM-P.5    : ----- : -
PVM-P.6    : ----- : -
PVM-P.7    : ----- : -

```

Figure 3.5 Alignment of the SH protein of PVM strain J3666 and PVM strain J3666 passages 5, 6 and 7. Differences in aminoacid sequences are indicated in black background. Mutated amino acids of the protein chains are shown in a black background the presence of a termination in the protein sequence is shown by a "\*" sign in each sequence. Numbers on the right represent the length of the protein molecule. Each protein molecule indicated on the left. SH: SH protein of PVM strains J3666, SH P.5, P.6, and P.7: SH proteins of PVM strain J3666 passages 5, 6, ad 7 respectively.

```

*           20           *           40           *
J3666      : -----MGRNFEVSGSITNLFERTQHPDTFRTGVKVN : 32
15-Warwick : ----- : -
15-ATCC    : -----MGRN EVSGSITNLFERTQHPDTFRT VKVN : 32
PVM-P.5    : MRPVEQLIQENYKLTSLSMGRNFEV GSITNLFERTQHPDTFRT VKVN : 50
PVM-P.6    : MRPVEQLIQENYKLTSLSMGRNFEV GSITNLFERTQHPDTFRT VKVN : 50
PVM-P.7    : MRPVEQLIQENYKLTSLSMGRNFEV GSITNLFERTQHPDTFRT VKVN : 50
           mgrn ev gsitnlnfertqhpdtfrt vkvn

           60           *           80           *           100
J3666      : QMCKLIAGVLTSAAVAVCVGIMYSVFSTSNHKANSTQNATTRNSTSTPPQ : 82
15-Warwick : -MCKLIAGVLTSAAVAVCVGIMYSVFSTSNHKANSTQNATTRNSTSTPPQ : 49
15-ATCC    : QMCKLIAGVLTSAAVAVCVGIMYSVFSTSNHKANSTQNATTRNSTSTPPQ : 82
PVM-P.5    : QMCKLIAGVLTSAAVAVCV VIMYSVFSTSNHKANSTQNATTRNSTSTPPQ : 100
PVM-P.6    : QMCKLIAGVLTSAAVAVCV VIMYSVFSTSNHKANSTQNATTRNSTSTPPQ : 100
PVM-P.7    : QMCKLIAGVLTSAAVAVCV VIMYSVFSTSNHKANSTQNATTRNSTSTPPQ : 100
           qMCKLIAGVLTSAAVAVCV VIMYSVFSTSNHKANSTQNATTRNSTSTPPQ

```

Figure 3.6 Amino acid comparison between G proteins of PVM strain J3666 (Accession number: NC\_006579), G protein of strain 15 (ATCC) (Accession number: AY729016), G protein of strain J3666 passages 5, 6 and 7. The differences between the aminoacid sequences are indicated in black. The consensus sequence among all of the viruses is shown in the last line of each row. Numbers on the right indicate the corresponding protein chain length.

### **3.4.3 The fragment 4 region**

Primer pair P15FLM23 and P15FLM25 (Appendix I) was used to synthesise the fragment 4. After obtaining the purified fragment 4, its sequence was obtained using P15FLM23, P15FLM25, and 22K7 primers (Appendix I). No nucleotide changes were identified between the sequenced fragment 4 and the sequence of PVM strain J3666 (Accession number: NC\_006579).

### **3.4.4 The fragment 5 region**

Primer pair P15FLL5 and P15FLL7 (Appendix I) was used to amplify the Fragment 5. The sequence of Fragment 5 was obtained after being purified using PL1, PL23, L1D, L1F, L1H, L1J, L2C and PL29 primers (Appendix I). No changes were observed among the sequence of the fragment 5 of the passages 5, 6, 7 in comparison with PVM strain J3666 (Accession number: NC\_006579).

## **3.5 Analysing the SH gene end and G ORF start sequence consistency in PVM strain J3666 population**

The sequencing data presented in section 3.4 revealed that the sequences obtained from passages 5, 6 and 7 have a difference in the sequence of the G gene in comparison with the sequence of PVM strain J3666 which were reported by Randhawa et. al. (1995). These results were consistent with the sequencing results obtained from cDNA constructs which were amplified from PVM strain J3666 for use in the construction of G-GFP minigenomes (Chapter 5). The sequences studied in Section 3.4 and Chapter 5 did not contain the stop codon for the first ORF of the G gene (Section 1.2.1.1). Instead, a point mutation was present which resulted in a change of the stop codon “UGA” into “AGA” in the mRNA. This point mutation would have changed the structure of the G gene, resulting in the generation of a single ORF through the fusion of ORFs 1 and 2. The G protein arising from this would therefore contain an amino terminal sequence derived from the first ORF. A different type of PVM G gene which is lacking the first ORF at AUG<sub>29</sub> was described by Krempl and Collins (2004). It was reported that the main ORF in PVM 15 (ATCC) contains a single ORF starting from position 83 (Figure 3.7). The data suggested that the sequence obtained from the virus resulted from a mixture of two subpopulations of PVM, one with two ORFs as described in Section 1.2.1.1.1 and reported by Randhawa and colleagues (1995) and the

other one containing one single ORF. A diagram representing these two G gene organisations is shown in Figures 1.5 and 3.7.

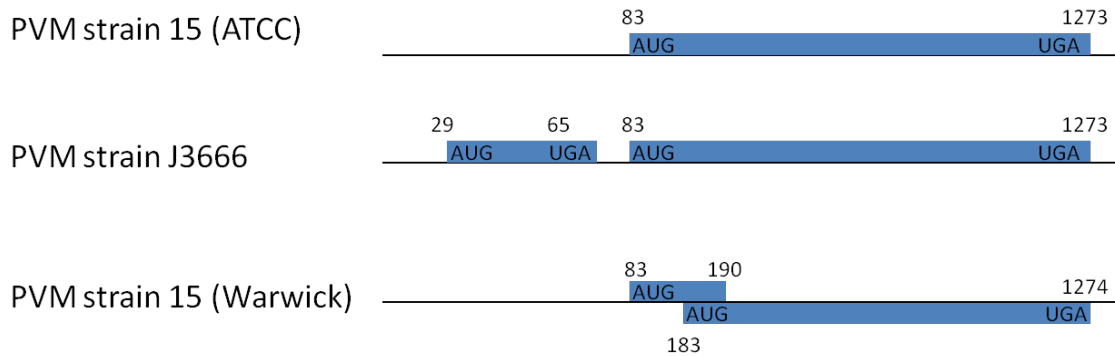


Figure 3.7 Describing the G mRNA organisation in PVM strain 15 (ATCC), PVM strain J3666 and PVM strain 15 (Warwick). In PVM strain 15 (ATCC) the presence of one unique ORF provides a longer N terminal for the G gene. In contrast, in the strain J3666 the stop codon at position 65 of the gene makes two separate ORFs. In the strain 15 (Warwick) a premature stop codon generates an ORF starting from AUG(83). Furthermore, an “A” insertion in the sequence of the ORF causes a frameshift mutation resulting in production of a stop codon at position 190. Subsequently, a start codon at position 190 of the G gene of the strain 15 (Warwick) is formed. For more details refer to the text.

Because of the importance of the organisation of the G gene in our studies, it was decided to explore the G gene organisation in the Warwick PVM strain J3666 working stock of passage 4 virus (Section 3.2) in more depth. For this, the G gene was amplified from total virus RNA extracts using RT-PCR with primers SH4360 and G4878(R) (Appendix I). Primer SH4360 anneals to the SH gene (nucleotides 206 and 227), and primer G4878(R) anneals to the G gene nucleotides 304-324. This made it possible to amplify a fragment from the SH and G genes with 517 nucleotide in length which contained 150 nucleotides from the SH ORF (responsible for encoding 49 amino acids from the C terminal of the SH protein) and 296 nucleotides from the two G ORFs of the strain J3666 (Accession number: NC\_006579) (responsible for encoding the first ORF and 81 amino acids from the N terminal of the second ORF of the G gene of PVM strain J3666). A standard PCR was conducted amplifying the SH-G fragment using *Taq* polymerase. The SH-G fragment (Figure 3.8) was purified from an agarose gel. The “T” overhangs in the pGEM-T (Promega) and pBS-T vectors (Section 2.3.8) and the “A” sequence in the SH-G fragment added by *Taq* DNA polymerase were ligated together to construct the pBS-SHG and pGEM-SHG plasmids prior to transformation.

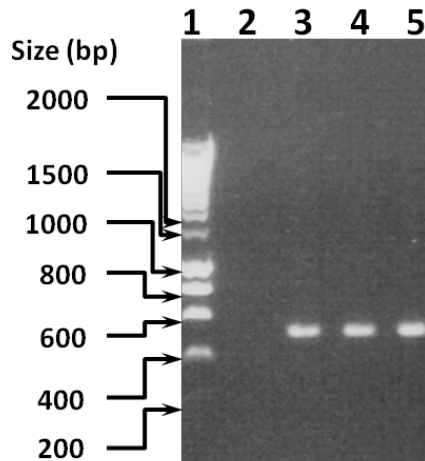


Figure 3.8 RT-PCR amplification, amplifying the SH-G fragment from cDNA of PVM strain J3666. Lane 1 Hyper ladder I, Lane 2 contains negative control, lanes 3, 4 and 5 contain the 518 nt PCR product, designated as SH-G fragment, from the RT-PCR.

The presence of the SH-G fragment in pBS-SHG clones was confirmed by *XhoI* and *EcoRI* restriction enzyme digestion of pBS-SHG (data not shown), and *EcoRI* digestion of pGEM-SHG (Figure 3.9). Prior to the sequencing, the presence of the SH-G fragments was confirmed by PCR with primers SH4360 and G4878(R) (Appendix I).

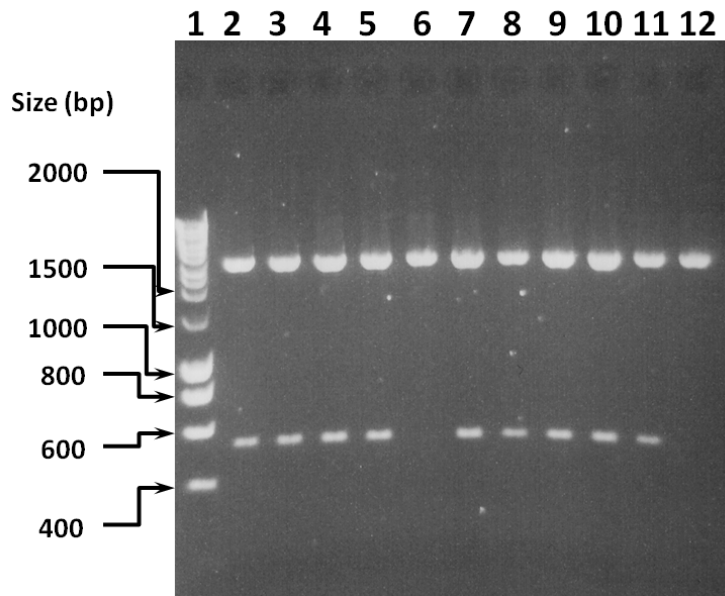


Figure 3.9 *EcoRI* digestion of pGEM-SHG plasmids. Plasmids pGEM-SHG1-11 digested with *EcoRI* restriction enzyme. The pGEM vectors are designed in a way to release the insert when they are digested with *EcoRI* enzyme. The SH-G fragments are visible at 518 bp. Lanes 1 to 12 in order contain ladder and pGEM-SH-G1-11. Plasmids pGEM-SH-G6 and pGEM-SH-G11 do not contain the 518 bp fragment.

The nucleotide sequence of 14 plasmids, 5 from one ligation reaction with pBS-T vector and 9 from an independent RT-PCR/ligation reaction with pGEM-T vector were determined. The sequences were compared against the reference sequence of the strain J3666 in the GenBank (accession number: AY743909). It was found that 12 of the sequences represent a G gene containing only one ORF (Figure 3.7). In contrast, the sequences obtained from pBS-SHG2 and pB-SHG9 clones (15% of the clones) reflected the organisation of the G gene published by Randhawa and colleagues (1995).

The results suggest the presence of at least two subpopulations of PVM strain J3666: a population containing two ORFs in the G gene as previously published (Randhawa *et al.*, 1995) (from now on strain J3666-A), and a population containing a single ORF as described above (Figure 3.7) (from now on strain J3666-B). Two of the 14 (15%) sequenced plasmids (pBS-GSH2 and pBS-GSH9) reflected the organisation of the G gene of the subpopulation J3666-A, and the rest (85% of the sequenced clones) showed the organisation of the G gene for the subpopulation of J3666-B.

Table 3.2 compares the sequences of the pBS-SHG2 and pBS-SHG9 with the sequence of PVM strain J3666 (accession number AY743909). Both pBS-SHG2 and pBS-SHG9 contain the first ORF preceding the main ORF of the G gene. The first ORF contains 12 codons and is terminated by a “UAG” stop codon as indicated. There is a gap of 15 non-coding nucleotides between the first ORF and the second ORF (Figure 3.7).

Gene	Changes in the nucleic acids sequence	Changes in the amino acids sequence
<b>SH</b>	U(262) → C	SL
	U(267) → C	Lys → His
	U(383) → C	NC
	U(387) → C	NC
<b>G</b>	U(21) → C	NC
	C(118) → U	SL
	U(196) → C	SL
	U(229) → C	SL

Table 3.2 Sequence differences among the SH-G clones and PVM strain J3666 sequence. Sequence comparison was made among clones pBS-SHG2, pBS-SHG9 and PVM strain J3666 (accession number: AY743909). The nature of each nucleotide change is indicated. The numbers in the parentheses indicates the nucleotides position in each gene. The effects of the mutation in the aminoacid sequences are shown. The occurrence of the mutations in non-coding sequences of each gene is shown by NC and the silent mutations are indicated as SL.

Table 3.3 compares the sequence of pGEM-SHG-1, 2, 3, 4, 6, 7, 8, 9, and 10 and pBS-SHG-5, 6 and 7, in which the G gene organisation was that of the subpopulation J3666-B, with the sequence of PVM strain J3666 (Warwick).

Gene	Changes in the nucleic acids sequence	Changes in the amino acids sequence	Plasmids
<b>SH</b>	U(250) → A	SL	pGEM-SHG7
	C(269) → U	His → Tyr	pGEM-SHG-1-4,6-10 pBS-SHG-5-7
	C(283) → U	SL	pGEM-SHG-1-4,6-10 pBS-SHG-5-7
	C(287) → U	Gln → Stop codon	pGEM-SHG-1-4,6-10 pBS-SHG-5-7
	A(291) → G	NC *	pGEM-SHG-6
	C(299) → U	NC *	pGEM-SHG-1-4,6-10 pBS-SHG-5-7
<b>G</b>	U(65) → A	Stop codon → Lys	pGEM-SHG-1-4,6-10 pBS-SHG-5-7
	A(66) → U	Lys → Met *	pBS-SHG-7
	A(83) → G	Met → Val	pGEM-SHG-2
	A(104) → G	Ser → Gly	pGEM-SHG-1-4,6-10 pBS-SHG-5-7
	A(113) → U	Ile → Phe	pBS-SHG-7
	G(165) → U	Gly → Val	pGEM-SHG-1-4,6-10 pBS-SHG-6 †
	U(234) → C	Val → Ala	pGEM-SHG-4 †
	G(236) → A	Gly → Arg	pGEM-SHG-1-4,7-10 pBS-SHG-5 ‡

Table 3.3 Sequence differences among the SH-G clones and PVM strain J3666 sequence. Sequence comparison was among clones pGEM-SHG-1, 2, 3, 4, 6, 7, 8, 9, and 10; pBS-SHG-5, 6 and 7; and PVM strain J3666. The nature of each nucleotide change is indicated. The numbers in parentheses indicate the nucleotide position in each gene. The possible effects of the mutation in the amino acid sequences are shown. The occurrence of the mutations in non-coding sequences of each gene is shown by NC and the silent mutations are indicated as SL. (\*) The result of the mutation was determined after applying the mutation affecting the ORF. (†) In pBS-SHG-5 and 7 remained undefined. (‡) In pBS-SHG-6 and 7 remained undefined.

### 3.6 Conclusion and discussion

As has been described for other viruses (Churchill *et al.*, 1969; Cohen *et al.*, 1989; Dardiri, 1969; Hearn *et al.*, 1966; Kwilas *et al.*, 2009; Taylor *et al.*, 1993), successive passage of a pathogenic stock of PVM in BS-C-1 cells generated attenuation in disease (Figure 3.3). This may explain the different pathogenicity seen with PVM strain 15 (Warwick) (which was obtained from ATCC then extensively passaged in tissue culture) and PVM strain 15 (ATCC) which has been passaged only in mice.

The pattern in the onset of the clinical signs of disease and weight loss in the group of mice infected with PVM strain J3666 (Working stock) was consistent with what was reported by Krempl et al. (2007). In results presented here, for the group of mice that received 500 pfu and 250 pfu, the weight loss and signs of disease started from day 5 and extending to day 10 when mice were euthanatized (Figure 3.1). For the group of mice infected with 125 pfu of PVM strain J3666 initially the same pattern was observed, but mice started recovering after day 10. In Krempl et al. (2007), a group of mice infected with 500 pfu showed weight loss one day earlier than observed here, and the experiment was terminated at day 8 post infection. The slight difference between the data reported by Krempl et al. (2007) and the results reported in this chapter could be explained by the 80 µl volume of the inocula used by Krempl et al. (2007). As shown by Cook *et al.* (1998), inoculum volume contributes to pathogenicity, as larger volumes are likely to deliver the virus deeper into the lungs.

Comparison between results reported by Cook et al. (1998) and those here suggest an inconsistency. Cook *et al.* (1998) reported that mice infected with 60 pfu of PVM strain J3666 developed severe disease which resulted in death of 20% of the animals (Cook *et al.*, 1998). However, the data reported here indicate that mice receiving 62.5 pfu developed mild disease. A significant difference in experimental protocol between the experiments here and those of Cook et al. (1998) was in the anaesthetic agent used; Cook et al. (1998) used ether and here mice were anaesthetised with ketamine (Section 2.1.52.1.5). The possible role of anaesthetic agents in mice infected with respiratory viruses cannot be neglected (Knight *et al.*, 1983; Shope, 1934).

In the experiments conducted with Cook et al. (1998) the virus was passaged 3 times in the cell culture (Prof. Andrew Easton, personal communication), while in this research the experiments were conducted using the passage 4 of PVM which was prepared during the time of experiment and was derived directly from the passage number 3 which was used by Cook et al. (1998). The experiments conducted with passages 5, 6 and 7, the virus obtained from the available stocks in the laboratory, all derived directly from the stock used by Cook et al. (1998).

Passage 5 virus was shown to be more pathogenic than passage 6 and 7 virus stocks. The passage 5 virus was able to produce clinical disease in the three administered doses (Figure 3.3). As expected the onset of infection with passage 5 was more rapid and produced more disease when 5000 pfu was administered, with progressively less severe disease seen when the inoculum was reduced to 500 pfu and



250 pfu. Passage 6 virus was capable of producing clinical signs of disease when the inoculum contained 5000 pfu, but not with 500 pfu or 250 pfu, as can be seen in Figure 3.3. When 5000 pfu of passage 7 virus was used, the clinical score peaked on day 8, after which the animals recovered.

In pneumoviruses, production of defective interfering (DI) particles has been reported for HRSV strain Randall during passage of the virus in cell culture and when the moi for the virus was calculated between 1 and 5 (Treuhaft & Beem, 1982). While they are not competent to replicate alone, DI particles require factors provided by a helper virus to be able to replicate. A protective effect for DI particles against heterologous or homologous viruses in animal models has been reported (Doyle & Holland, 1973; Easton *et al.*, 2011; Scott *et al.*, 2011). This is due to interferon induction and interference with the viruses in their replication cycle (Doyle & Holland, 1973; Easton *et al.*, 2011; Scott *et al.*, 2011). Although the presence of DI particles in batches of PVM strain J3666 passages 5, 6 and 7 was not directly examined the possible effect of DI particles in reducing the pathogenesis of the virus can't be ignored. However, virus stocks were prepared using very low moi (typically 0.01-0.05 pfu/cell or lower). This significantly reduces the likelihood of DI virus production. It is likely that among the factors affecting the pathogenicity of PVM strain J3666, adaptation to growth in monkey kidney cells through consecutive passage of PVM in cell culture is one of the main determinants in the attenuation of the virus. The molecular cause of the attenuation remains to be determined.

The genome sequence of PVM passages 5, 6 and 7 was determined and compared to understand its correlation with nucleotide changes and the differences in pathogenesis. As discussed above, the pathogenicity of PVM strain passages 5, 6 and 7 reduced gradually from the highest (in the passage 5) to the lowest (in the passage 7). Surprisingly, no genome sequence changes among PVM strain J3666 passages 5, 6 and 7 were identified. As mentioned in section 3.4, it was designed to sequence a part of genome divided into 4 fragments. The sequence of the NS1 and NS2 genes and the majority of the L gene remained undetermined. Thus, it is possible that the attenuation is due to mutations in the areas of the genome whose sequence remained unknown. Also, it wouldn't be possible to identify the presence of mutations when they are in low frequencies. To identify such mutations different clonal isolates of the population of the virus need to be sequenced, although the cloning experiment increases the risk of nucleotide change.

Comparing the sequences of passages 5, 6 and 7 with PVM strain J3666 (accession number: NC\_006579) indicated mutations across the fragment 3. Table 3.1 summarizes the mutations in the fragment 3. Briefly, the coding mutations seen in the SH gene were C(4381)→U changing a histidine residue to a tyrosine residue and C(4399)→U mutating glutamine residue to a stop codon resulting in the synthesis of a SH protein with a truncated C terminal; the coding mutations in the G gene are U(4575)→A changing an amber stop codon to a lysine residue resulting in the production of a G glycoprotein with a larger N-terminal, C(4578)→U changing the noncoding sequence between the two ORFs to a leucine codon, A(4614)→G changing a serine residue to a glycine residue, G(4675)→U changing a glycine residue to a valine residue, G(4746)→A changing a glycine residue to an arginine, U(5631)→A changing a serine residue to a threonine residue; in the F gene: U(6581)→G changing an isoleucine residue to a glutamine residue and U(6852)→C changing a valine residue to an alanine residue.

In an attempt to investigate the possible presence of subpopulation of viruses in the working stock of PVM strain J3666, 14 independent clones containing parts of the SH and G genes derived from two separate RT-PCR amplifications from virus genomic RNA or antigenomic RNA were sequenced. The analysis showed that at least two distinct subpopulations of viruses were present. The minor subpopulation, (J3666-A) represented by 2 clones, had a G gene with two ORFs organised as reported by Randhawa *et al.* (1995). The remaining clones represented a major subpopulation generated a single, large ORF which would encode a G protein with an amino terminal extension (J3666-B). These data indicate that PVM stocks contain a population of viruses with some sequence variation. Using these stocks it is not possible to be confident about the role(s) of specific gene sequences in pathogenicity. It is therefore more desirable to generate homogenous populations using reverse genetics.

A further possibility is that factors like post-translational processing of proteins including changes in the glycosylation patterns of glycoproteins and proteolytic cleavage may affect the pathogenesis. Recently, comparison of HRSV grown in HEp-2 cells with HRSV grown in Vero cell lines showed that the HEp-2 grown virus contained a smaller G glycoprotein lacking its C terminal (Kwilas *et al.*, 2009), and the progeny viruses recovered from Vero cells were inefficient in using cellular glycosaminoglycan (GAG) molecules as the receptor. However, these viruses were able to initiate an infection in well-differentiated human air-way cells that do not produce GAG molecules

at their apical sites. It is therefore possible that the attenuation seen in the passages 6 and 7 was due to the accumulation of post-translational processing in the virus population.

Sequence analysis of the SH gene (Section 3.5) revealed the presence of two subpopulations. The major subpopulation (J3666-B) expressed a truncated SH protein and a minor subpopulation (J3666-A) expressed a full length SH protein. The mutations discussed in Section 3.5 could affect the structure and function of SH protein, causing a defective virus particle. Considering that the SH protein of HRSV is a transmembrane protein and forms a transmembrane channel across the membrane (Carter *et al.*, 2010) the hydrophobicity of the SH protein was determined to analyse the effect of the mutations on the hydrophobicity (Kyte & Doolittle, 1982). To predict the possible effects on the hydrophobicity of the SH protein, the HMTOP server was used (Kyte & Doolittle, 1982; Tusnady & Simon, 2001). Analysing the results (data not shown) suggested that the SH protein of PVM strain 15 (Warwick) may lack a transmembrane domain, while in the SH protein of PVM strain J3666, the amino acids 24 to 46 produce transmembrane helices.

Unlike in PVM, the SH gene of HRSV strains has been reported to vary very little in coding capacity (Chen *et al.*, 2000). In contrast, the SH gene of human metapneumovirus is very vulnerable to mutations when it is grown *in vitro* (Biacchesi *et al.*, 2007). It is possible that the SH gene of PVM studied in this research has been mutated during passages in tissue culture. HRSV lacking the SH gene was only partially attenuated in chimpanzees' lower respiratory tract (Whitehead *et al.*, 1999). Studying the differences in the SH proteins of PVM strains could explain the importance of the SH gene in the virulence of the virus.

## **CHAPTER 4**

# **A REVERSE GENETICS APPROACH TO STUDY THE EFFECT OF THE G GENE ORGANISATION IN PVM PATHOGENICITY**

## 4.1 Introduction

A reverse genetics system for HRSV was first described in 1991 by rescuing synthetic HRSV genome homologue in which a CAT gene replaced virus genes in the genome (Collins *et al.*, 1991). Soon after this achievement, Collins *et al.* (1995) reported production of infectious virus from cDNA of HRSV strain A2. Since then, the technique has been used to investigate a wide range molecular aspect of HRSV *in vitro* and *in vivo*. However, the main obstacle in studying the pathogenesis of HRSV, lack of a suitable animal model, has hampered the study of pathogenesis. As discussed in chapter one, PVM is a versatile substitute to study molecular aspects of pneumoviruses, and employing a reverse genetics system for PVM would be beneficial in extending its ability as a model system.

In chapter 3, the possibility of the contribution of genetic factors to the pathogenicity of pneumoviruses was analysed. A reverse genetics system for PVM was developed by Dibben (Dibben, 2006). However, the early attempts to obtain infectious virus from cDNA failed (Dibben, 2006). Another reverse genetics system for PVM from which virus could be generated was developed by Krempl *et al.* (2007). The work presented in this chapter, describes the establishment of a reverse genetics system for PVM and its use in studying the pathogenicity of the virus. Using this reverse genetics system, infectious viruses were rescued carrying either of the two organisations of the G gene (Section 1.2.1.1). To study the effect of the organisation of the G gene, the first ORF of the G gene was deleted and infectious viruses were rescued. The infectious viruses were used in a BALB/c mouse model to analyse the pathogenicity of the virus. The clinical scoring system (Section 2.5) was used to evaluate the pathogenicity. The differences between the results obtained from the reverse genetics system described in this section and the reverse genetics system developed by Krempl and colleagues (2007) are discussed below.

## 4.2 Plasmid p15FL-2G

The plasmid carrying a cDNA copy of the PVM strain 15 (Warwick) genome, p15FL-2G, was constructed by Dr. O. Dibben and was available in the pneumovirus Laboratory, University of Warwick (Dibben, 2006). This was constructed from 6 different fragments with four (fragments 2 to 5) defined by restriction enzyme sites inserted during the cloning process (*SexAI* at positions 1006, *SalI* at position 4089,

*Xma*I at position 7261 and *Not*I at position 8134), and one natural site (*Sex*AI at position 11600) (Figure 4-1). The cDNA was cloned between a bacteriophage T7 transcription promoter and the hepatitis delta ribozyme in such a way that T7 transcription would result in the production of an antigenomic copy of the PVM genome shown to be required for efficient recovery of infectious virus for other negative sense RNA viruses. The construct has 2G residues at the 5' end of the antigenome RNA to ensure efficient transcription by the T7 RNA polymerase. These are presumed to be eliminated during virus replication as seen for all other recombinant negative sense RNA viruses.

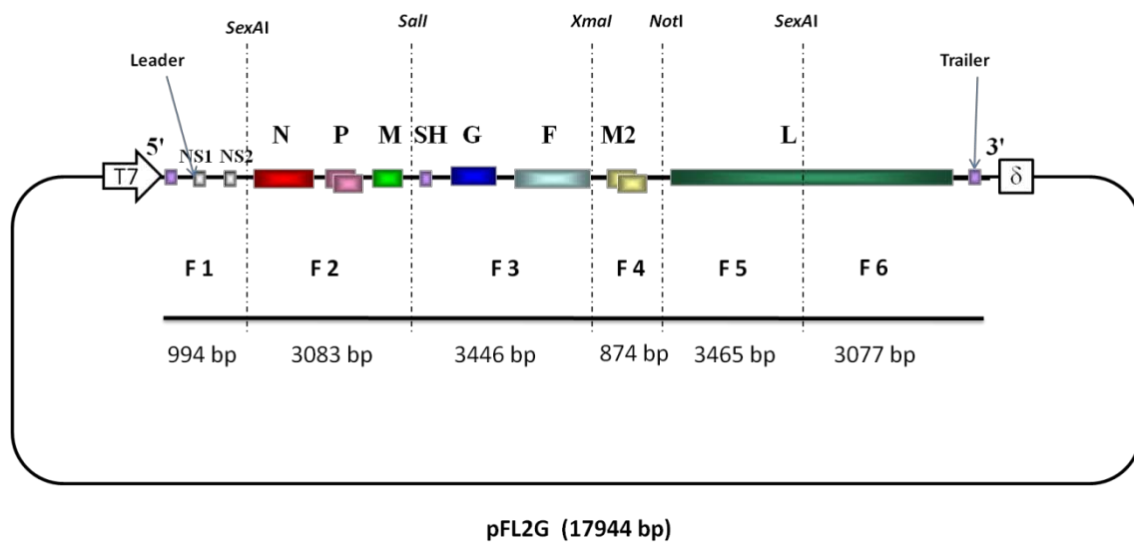


Figure 4.1 Plasmid p15FL-2G (17944 bp). The genes, restriction sites, ribozyme and T7 polymerase transcription promoter are shown. P2 and M2-2 overlapping genes are shown.

Plasmid p15FL-2G was subjected to further modifications to synthesise pFL2G-G15, pFL2G-G15ΔORF1, pFL2G-GJ3666 and pFL2G-GJ3666ΔORF1. The cloning procedure is explained in the text and summarized in Figure 4.2.

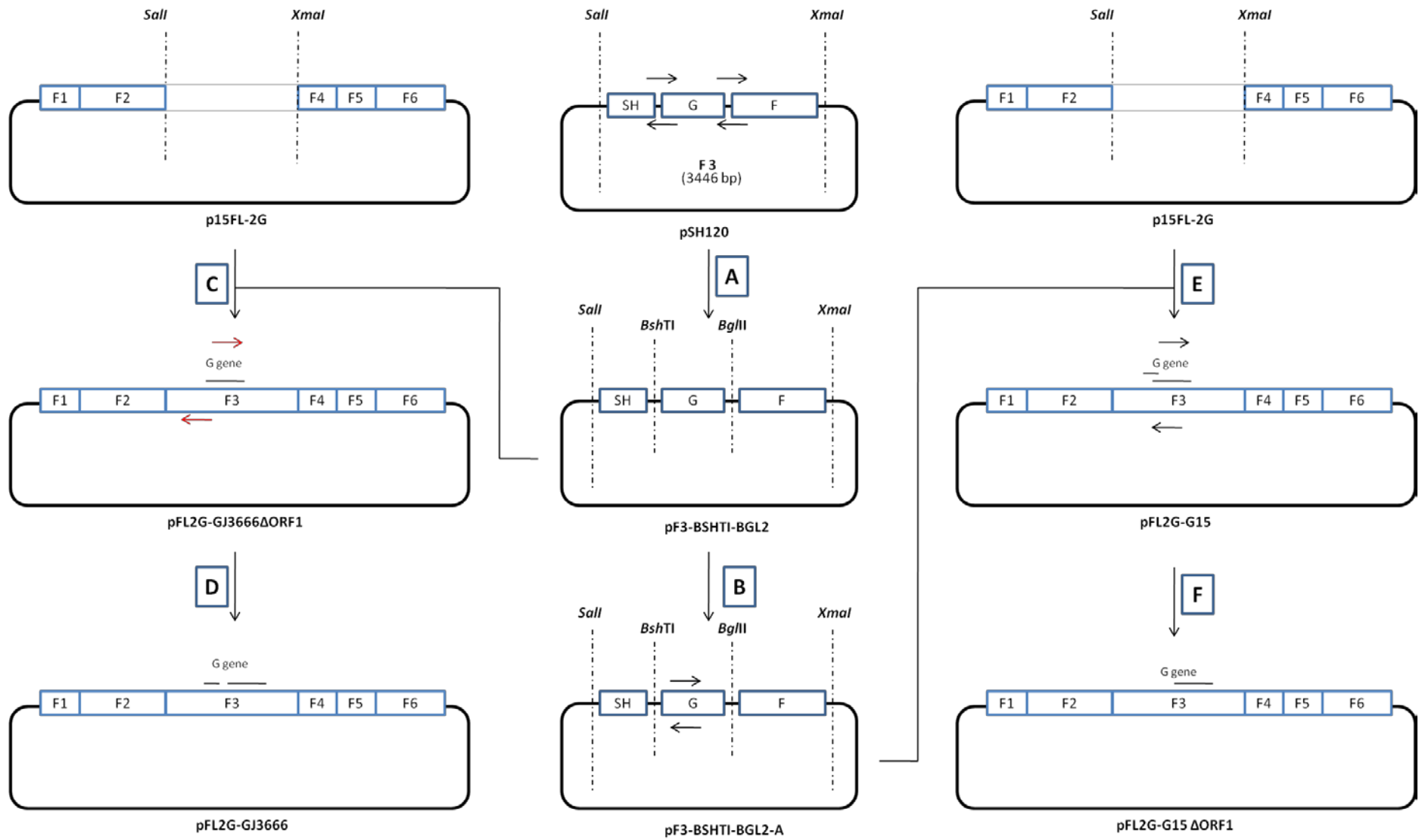


Figure 4.2 – The syntheses of p15FL2G-GJ3666, pFL2G-GJ3666ΔORF1, pFL2G-G15 and pFL2G-G15ΔORF1. To synthesise plasmid pF3-BSHTI-BGLII, the plasmid pSH120 was subjected to two Quick Change mutagenesis reactions using primer pair QCF3BGL2F and QCF3BGL2R (arrows shown in black) as indicated in step A. Plasmid pF3-BSHTI-BGL2 was subjected to a Quick Change PCR reaction using primer pair QCFL2GA4728F and QCFL2GA4728R (arrows shown in black) to synthesise the pF3-BSHTI-BGL2-A plasmid (step B). The fragment 3 from pF3-BSHTI-BGLII was excised using *SalI* and *XmaI* restriction enzymes and ligated into p15FL-2G double digested with *SalI* and *XmaI* restriction enzymes, leading to production of pFL2G-GJ3666ΔORF1 (step C). Primer pair G\_Round(F) and G\_Round(R), arrows shown in red, were designed to introduce the nucleotide differences between PVM strain 15 and strain J3666 into the first ORF of the G gene using the round the horn site directed mutagenesis PCR reaction (step D) to produce pFL2G-GJ3666. To produce pFL2G-G15, the fragment 3 from pF3-BSHTI-BGLII-A (obtained in step B) was excised using *SalI* and *XmaI* restriction enzymes and ligated into p15FL-2G double digested with *SalI* and *XmaI* (step E). Plasmid pFL2G-G15 was then mutated using primer pair 1stORFOut-15-F and 1stORFOut-15-R (arrows shown in black) and Quick Change mutagenesis reaction to synthesis pFL2G-G15ΔORF1 (step F). The black lines indicated with “G gene” define the organisation of ORFs in the various G gene constructs.



### 4.3 Complete sequencing of p15FL-2G

Before starting the rescue process, the complete nucleotide sequence of the full length genome plasmid was determined. For this purpose the plasmid p15FL-2G was sequenced using G9, J1, J4E, J5, M1, M2, N1, N3, N5, NS1B, NS2A, NS2B, P2A, P3, P4, SH2, 22K7, F7, F9, F11, F14, F17, G7, J7, L1F, L1H, L1J, L2C, L2F, L10, PL23, PL29, L2G, L2H, L3C, L3D, L3E, L3F, L3G, PL48, and L50 primers (Appendix I). Sequencing data were analysed and 30 changes between the sequenced plasmid and the published sequence of PVM strain 15 (Warwick) (GenBank; accession number AY743910.1) were identified. These are shown in Table 4.1, and some, as indicated, confirmed the location of restriction endonuclease sites inserted during the cloning process of p15FL-2G. Two of those were introduced in the coding region of the second ORF of the M gene which caused sequence changes in a hypothetical polypeptide synthesised from a short second ORF [A(4048) → G (4089) resulting in the change of a histidine residue to a valine residue, A(4051) → C resulting in the change of a histidine residue to an asparagine residue, A(4063) → C resulting in the change of a threonine residue to a proline residue]. The rest of the restriction sites introduced by Dibben (2006) were placed in the noncoding regions (NTR) of the M, SH, F and M2 genes.

Table 4.1 describes the changes found across the full length genome of PVM strain 15 (Warwick). Briefly, in the M gene, an A(3690)→C mutation changes a tyrosine residue to a serine residue, G(3705)→A changes a glycine residue to asparagine residue; in the G gene G(5465)→C changes a glutamic acid residue to a glutamine residue; in F gene A(5954)→T changes of a tyrosine residue to a phenylalanine residue and C(7278)→G changes a phenylalanine residue to a leucine residue. These mutations were observed and reported by Dibben (2006) during constructing the p15FL-2G plasmid.

Sequence of PVM 15 (Warwick)	Sequence of p15FL-2G	Position	Comments
970 CA	1012 GT	NTR of NS1	Introduction of SexAI restriction site
3690 A	3731 C	M gene, 1 <sup>st</sup> ORF	Tyr → Ser
3705 G	3746 A	M gene, 1 <sup>st</sup> ORF	Gly → Asp
3796 A	3837 G	M gene, 1 <sup>st</sup> ORF	SL
4048 A	4089 G	M gene, 2 <sup>nd</sup> ORF	His → Val

4051 A	4092 C	M gene, 2 <sup>nd</sup> ORF	His → Asp Introduction of <i>SaI</i> restriction site
4059 A	4100 C	M gene, 2 <sup>nd</sup> ORF	SL
4063 A	4104 C	M gene, 2 <sup>nd</sup> ORF	Thr → Pro
4065 A	4106 C	M gene, 2 <sup>nd</sup> ORF	SL
4218 C	4259 T	SH gene	SL
5465 G	5506 C	G gene	Glu → Gln
5548 A	5589 G	G gene	SL
5954 A	5995 T	F gene	Tyr → Phe
5958U	5999 C	F gene	SL
6127U	6168 C	F gene	SL
6531 C	6572 C	F gene	SL
6720U	6761 C	F gene	SL
7278 C	7319 G	F gene	Phe → Leu
7296-7499 AAAC	7537-7540 CGGG	NTR of F gene.	Non-coding region between F and G genes. Introduction of <i>XmaI</i>
8370-8374 AGUAA	8409-8413 GGCCG	NTR of M2 gene.	Non-coding region between M <sub>2-2</sub> and L gene. Introducing <i>NotI</i> restriction site
8902 U	8943 C	L gene	SL
12247 A	12288 G	L gene	SL
13030 A	13071 G	L gene	SL
13516 A	13557 G	L gene	SL

Table 4.1 Comparison of sequence of PVM strain 15 (Warwick) and p15FL-2G. Silent mutations are indicated as SL.

#### 4.4 Amending plasmid p15FL-2G

The structure of the PVM G gene for the non-pathogenic strain 15 and the pathogenic strain J3666 was described in chapter one. To understand the effect of these different organisations of the G gene on the pathogenesis of PVM, it was necessary to amend the plasmid carrying the full-length PVM strain 15 viral genome in such a way to make it feasible to swap the G gene with G gene from other viruses such as PVM strain J3666. For this purpose, two further restriction sites were inserted, one before and one after the G gene. The presence of restriction sites would make it possible to remove the G gene entirely from the full length genome and replace it with the desired gene.

Two restriction enzymes, *Bsh*TI and *Bg*II were inserted into the plasmid pSH120 (Figure 4.2), the plasmid carrying fragment 3 used to create the full length PVM genome cDNA using the quick change mutagenesis technique. The use of plasmid

pSH120 as the template provided advantages including a wider choice from the restriction enzymes which are frequently present on the full length genome, use of *pfu* proof reading enzyme which potentially is not able to amplify very large constructs such as the full length genome, and the ease of confirming the sequence in the smaller construct.

To insert the *Bsh*TI and *Bgl*II restriction sites into pSH120, two different Quick Change mutagenesis PCR reactions were performed. First, the plasmid pSH120 was used as the template and the quick change mutagenesis PCR primer QCF3BSHT1F and QCBSHT1R pair (Appendix I) designed to insert the *Bsh*TI restriction site used. Having been digested with restriction enzyme *Dpn*I to remove the template, the PCR product was used to transform competent *E. coli*. Liquid cultures were grown from individual colonies and the presence of mutated pSH120 was confirmed by restriction digestion of the isolated DNA from the individual bacterial cultures by *Bsh*TI restriction enzyme. The mutated plasmid was named pF3BshT1. Plasmid DNA from four colonies was individually sequenced with QCF3BSHT1F, QCBSHT1R, SH95+, SH2, G4E, G7, G7, F7, F14, F9, and F11 primers (Appendix I) to confirm that no additional mutations had been inserted. The plasmid DNA mixture was then used as the template to insert the second restriction site (*Bgl*II) into the plasmid.

Quick change mutagenesis PCR was performed with the primer pair QCF3BGL2R and QCF3BGL2F (Appendix I) designed to insert a *Bgl*II restriction site. Having been digested with *Dpn*I to remove the template, the PCR product was used to transform competent *E. coli*. Plasmid DNA from individual colonies was prepared and the presence of the *Bgl*II site confirmed by digestion followed by nucleotide sequencing. The mutated plasmid was named pF3-BSHT1-BGL2 plasmid. Using J5 and F14 primers, a PCR fragment was amplified from the pF3-BSHTI-BGL2. Digestion of the DNA fragment with *Bsh*TI and *Bgl*II is shown in Figure 4.3.

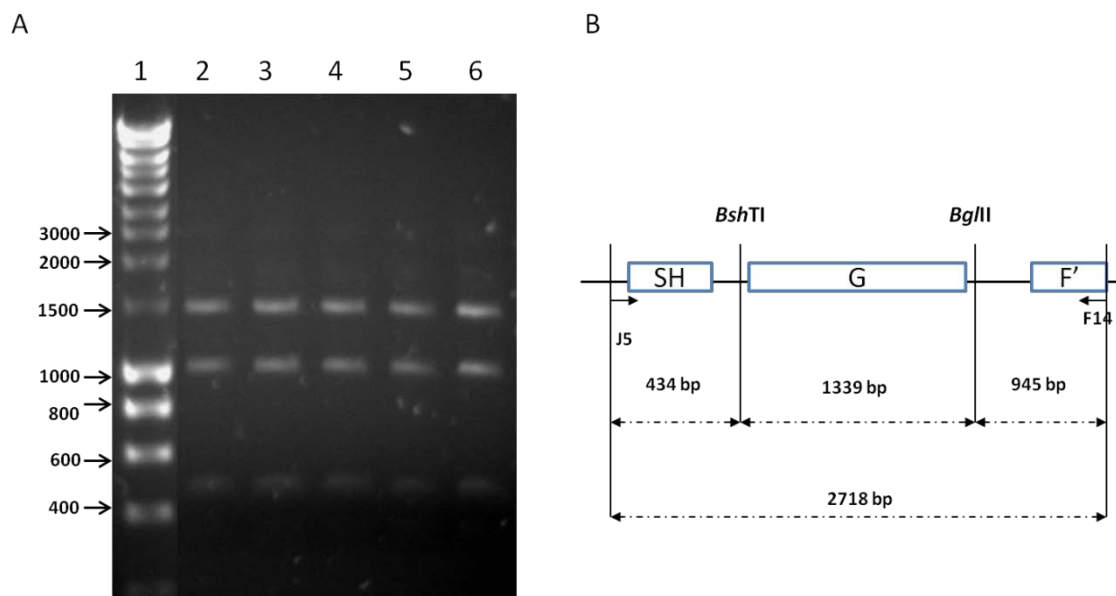


Figure 4.3 Digestion of the 2718 bp fragment generated by PCR amplification from pF3-BSHT1-BGL2 plasmid using J5 and F14 primers. Following digestion with *Bsh*TI and *Bgl*II restriction enzymes, three fragments were generated of 434 bp, 1339 bp and 945 bp (A); lane 1 Hyperladder I DNA size markers, Lanes 2 – 6 the digested PCR fragment. The strategy was used to amplify the 2718 bp fragment and the position of the restriction endonucleases are shown in B.

As discussed in Chapter 1, a significant difference in nucleotide sequence between PVM strain J3666 and PVM strain 15 (Warwick) is in their G gene organisation which in the strain 15, contains 2 overlapping ORFs, the second and largest of which encodes the G protein. In contrast, the strain J3666 G gene contains a short ORF upstream of the main ORF which encodes the G protein (Figure 1.5). Originally, plasmid p15FL-2G was amplified from PVM strain 15 (Warwick), so it was expected to have a genome structure similar to that published. However, sequencing data (Dibben, 2006) revealed that the major ORF of the G gene in the plasmid p15FL-2G had the same structure of the major ORF of the G gene in PVM strain J3666 *i.e.* it lacked the additional U residue. However, only one ORF was present. To return the sequence to that of PVM 15 (Warwick) an A residue was inserted into the G gene in pF3-BSHT1-BGL2 (from now on pF3BB). The presence of the A residue at position 169 of the G gene in the mutated plasmid (pF3BBA) would generate a G gene with a structure similar to that found in PVM strain 15 (Warwick) with two overlapping ORFs. This would generate two plasmids similar to each other with a difference only in the

organisation of the G gene, making it possible to compare the effect of the mutation on pathogenesis.

To construct the mutation, pF3BB was used as the template and a primer pair (QCFL2GA4728F and QCFL2GA4728R) designed to insert the A residue at position 4725 of p15FL-2G. After *DpnI* digestion, the PCR product was used to transform competent *E. coli*. Four colonies (designated as pF3BBA-C1, pF3BBA-C2, pF3BBA-C3 and pF3BBA-C4) were selected and their sequence obtained using T7, J4, SH2, SH95+, G4E, G7, G9, F17, F7, F14, F9 and F11 primers (Appendix I). Following confirmation of the sequence, the plasmid pF3BBA-C3 was chosen to use in the experiments.

#### 4.5 Insertion of fragment 3 into p15FL2G

In order to clone the fragment 3 into p15FL2G, the fragment 3 from the plasmid pF3-BB was digested with the restriction enzymes *SalI* and *XmaI* and was replaced with fragment 3 from either pF3-BB or pF3-BBA. The presence of *BglII* restriction site in the fragment 3 of the full length genome was confirmed by restriction double digestion of the full length plasmid with *BglII* and *EcoRI*, and *NdeI* and *BglII* (not shown).

#### 4.6 Description of the full length genomes

The adenosine deletion in pF3-BB plasmid caused an extension in the open reading frame responsible for coding the G glycoprotein to achieve the same sequence of the G glycoprotein described for PVM strain J3666. However, because the rest of the full length genome is similar to PVM strain 15 (Warwick), the G gene lacked the first ORF found in the G gene of PVM strain J3666 (Randhawa *et al.*, 1995). Therefore, the p15FL-2G construct carrying the F3-BB fragment was called pFL2G-GJ3666 $\Delta$ ORF1. On the same basis, the full length plasmid carrying the fragment 3 with the composition similar to that was found in the PVM strain 15 (Warwick) was called pFL2G-G15.

For the purpose of this research it was necessary to synthesise two more full length genome plasmids based on the sequence of pFL2G-GJ3666 $\Delta$ ORF1 and pFL2G-G15 plasmids which are different only in the presence or absence of the first ORF of the G gene in comparison with their parental plasmids. “Round the horn” site directed mutagenesis PCR (Section 2.3.14) was used to insert the nucleotide changes in the G gene of plasmid pFL2G-GJ3666 $\Delta$ ORF1. For this purpose the G\_Round(R) and G\_Round(F) primer pair (Appendix I) were designed to contain the sequence of the G

gene strain J3666. The position of the primers were chosen in such a way that annealing at their 5' ends would result in generation of a fragment spanning the beginning region of the G gene of PVM (Figure 4.4). The mutagenesis was carried out as explained in Section 2.3.14. The successfully mutated plasmid, designated as pFL2G-GJ3666, was identified by sequencing of the G gene. To make pFL2G-G15 $\Delta$ ORF1 the Quick Change mutagenesis PCR method (Section 2.3.12) was carried out to delete the first ORF in pFL2G-G15 using the primer pair 1stORFOut-15-F and 1stORFOut-15-R (Appendix I) which were designed to mutate the AUG codon of the first ORF to GCG. The successfully mutated plasmids were selected after analyzing the sequence of G gene.

```

          CCTATTTCATGATAGGATAACCTTAGTTTACTCTGGACATCTCGTCGAGT
1  ACAAAAAAATTAGGATAAGTACTATCCTATTGGAACCAAACGAGACCTGTAGAGCAGCTCA 60
   TGTTTTTTAAATCCTATTTCATGATAGGATAACCTTGTTTCTCTGGACATCTCGTCGAGT

          ATGTTCTCTTGATGA-P
61  TACAAGAGAACCCTAGCTGACTTCACCTAGTATGGGAAGGAACCTCGAAGTGAGTGGCA 120
   ATGTTCTCTTGTTGATCGACTGAAGTGGATCATACCCTTCCTTGAACTTCACACCCGT
          P-AGTTGACTTCACTTAGTATGGGAAGGAACCTTGAAGTGAGTGGCA

121 GCATTACCAATTTGAACTTTGAGAGAACTCAGCATCCTGACACATTTAGGACTGGTGTA 180
   CGTAATGGTTAAACTTGAACTCTCTTGAGTTCGTAGGACTGTGTAATCCTGACCACATT
          GCATTACCAATTTGAACTTTGAGAGAACTCAGCATCC

```

Figure 4.4 The position of the G\_Round(F) and G\_Round(R) primers to each other. The phosphorylated 5' ends of the primers are indicated (Section 2.3.14). The G\_Round(R) primer is shown in blue and G\_Round(F) is shown in red.

#### 4.7 Rescue of plasmid p15FL-2G

The sequence of p15FL-2G obtained in Section 4.3 showed that there were no significant changes between sequences of the the p15FL-2G clone and PVM strain 15 (Warwick) sequence submitted to the GenBank (Accession number: AY743910) that might hamper the process of rescuing infectious virus. The genome of negative stranded RNA viruses is not able to start infection *de novo*, and require the viral polymerase complex to synthesis antigenomic RNA to act as the intermediary RNA from the genomic RNA. Hence, it is necessary to provide the viral polymerase complex to support replication of the virus. In order to provide the polymerase complex, four plasmids encoding the PVM N, P, M2-1 and L proteins were co-transfected with the p15FL-2G plasmid. As discussed in chapter one, the stoichiometry of individual proteins in the polymerase complex is highly important to achieve successful replication and/or transcription. It was therefore necessary to find the right combination of concentrations for the polymerase complex plasmid to obtain a successful rescue.

Successful rescue systems for other members of pneumoviruses were reviewed and 5 methods which had been used previously to rescue HRSV, BRSV and APV were considered. Methods 1 and 2 were used to rescue BRSV from cDNA (Buchholz *et al.*, 1999; Yunus *et al.*, 2001), and method 3 and 4 (Collins *et al.*, 1995; Jin *et al.*, 1998) were used to rescue HRSV from cDNA. Method 5 was employed by Dr. Roger Ling to rescue APV (Dr. R. Ling, unpublished data). Table 4.2 describes the concentrations of the helper plasmids which were used in these methods.

The various amounts of DNA shown in Table 4.2 were mixed and BSR-T7/5 cells were transfected as described in Section 2.2.4. The transfected cells were kept at 31°C and checked regularly for the presence of CPE. To passage the material the transfected BSR-T7/5 cells were subjected to a freeze and thaw cycle before being scraped into the medium and one third of the culture was transferred to a fresh BS-C-1 cell culture. This point was considered as the first passage. Passaging of viruses was continued until extensive cell death was observed. Among the methods of rescue mentioned in Table 4.2, methods 1 and 5 resulted in successful production of recombinant viruses. The method 5 was used throughout as the method to rescue recombinant viruses and further experiments were planned based on the viruses obtained from this method. The process for passaging viruses following transfection is summarised in Figure 4.5. The recombinant viruses were defined as rPVM. Rescue of pFL2G-G15, pFL2G-G15ΔORF1, pFL2G-GJ3666, pFL2G-GJ3666ΔORF1 and p15FL-2G resulted in production of rPVM-G(15), rPVM-G(15ΔORF1), rPVM-G(J3666), rPVM-G(J3666ΔORF1) and rPVM, respectively. The recombinant viruses were subjected to continuous passages until a titre greater than  $1 \times 10^5$  pfu/ml for each was achieved. rPVM-G(J3666ΔORF1) and rPVM-G(15) were subjected to six continuous passages in BS-C-1 cells to obtain this titre. rPVM-G(J3666) and rPVM-G(15ΔORF1) were subjected to two passages in BS-C-1 cells to obtain the desired titre.

Method used to rescue virus	Full length	pN	pP	pM2-1	pL	References
1	10 µg	4 µg	4 µg	2 µg	2 µg	(Buchholz <i>et al.</i> , 1999)
2	0.4 µg	0.4 µg	0.2 µg	0.4 µg	0.4 µg	(Jin <i>et al.</i> , 1998)
3	0.4 µg	0.4 µg	0.4 µg	0.1 µg	0.1 µg	(Collins <i>et al.</i> , 1995)
4	1 µg	0.4 µg	0.3 µg	0.1 µg	0.2 µg	(Yunus <i>et al.</i> , 2001)
5	3 µg	0.3 µg	0.3 µg	0.24 µg	0.15 µg	Dr. R. Ling, unpublished data.

Table 4.2 DNA concentrations used to rescue PVM from cDNA.

The presence of virus was confirmed by RT-PCR specific for the viral genome. The presence of *XbaI* and *SalI* restriction enzyme sites which are specific for the recombinant virus (obtained from using the method 2 mentioned in Table 4.2) was confirmed by amplifying the relevant part of the viral genome. For this, the 751 bp “M-SH” fragment was amplified from nucleotide 3571 to 4322 of the PVM genome using M2 and SH2 primers and the cDNA of the virus as the template (Appendix I). The DNA fragment was purified and then digested with *SalI* restriction enzyme which would generate two fragments of 273 bp and 477 bp (Figure 4.6). This confirmed that recombinant PVM had been generated from the plasmid.

In addition, to show successful infection of the cell by the PVM monoclonal antibody 26/C3/B5 was used to detect the PVM P protein in cell culture. Figure 4.7 shows cells infected with recombinant PVM surrounded by non-infected cells and expressing the P protein. The full sequence of plasmids pFL2G-GJ3666 and pFL2G-G15ΔORF1 which were generated using site directed mutagenesis were determined using the G9, J1, J4E, J5, M1, M2, N1, N3, N5, NS1B, NS2A, NS2B, P2A, P3, P4, SH2, 22K7, F7, F9, F11, F14, F17, G7, J7, L1F, L1H, L1J, L2C, L2F, L10, PL23, PL29, L2G, L2H, L3C, L3D, L3E, L3F, L3G, PL48, and L50 primers (Appendix I), to confirm that no mutations had arisen during the cloning procedure. A sequence conflict between their sequence and sequences of p15FL-2G, and pF3BBA was observed resulting in a LS<sub>423</sub>K → LN<sub>423</sub>K change in the F protein synthesised from the pFL2G-GJ3666 and pFL2G-G15ΔORF1 plasmids.

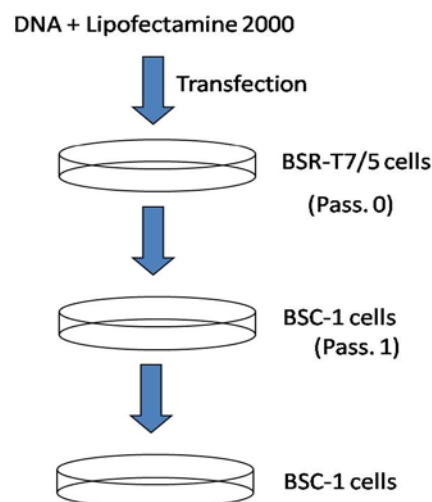


Figure 4.5 The rescue process to obtain infectious virus from cells. Each step of rescue of the recombinant viruses is indicated as a passage (pass.) number. Transfected BSR-T7/5 cells (pass. 0) were lysed and cell lysate obtained from BSR-T7/5 cultures was used to infect BSC-1 cells (pass. 1).



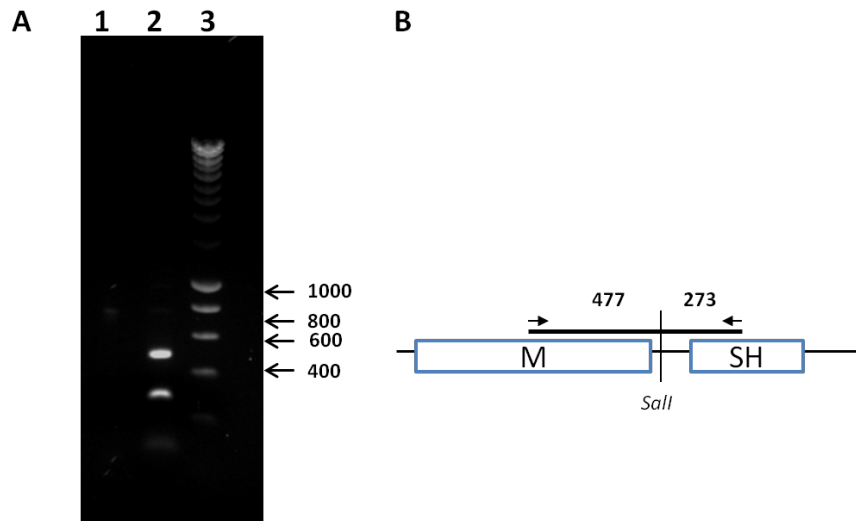


Figure 4.6 A. Digestion of the M-SH PCR fragment from the wildtype and recombinant viruses containing a *SalI* restriction site. The PCR fragments (751 bp) were amplified from the cDNA of the wildtype recombinant viruses (B). Following purification, the PCR fragments were digested with *SalI* restriction enzyme generating 477 bp and 273 bp fragments in the DNA fragment amplified from the recombinant viruses. The DNA amplified from the wildtype virus remained undigested. Lane 1, M-SH DNA fragment from wildtype virus, after being digested with *SalI* restriction enzyme; lane 2, the digested M-SH fragment; lane 3, DNA Hyperladder I size markers. The size of DNA markers are shown on the right of the gel.

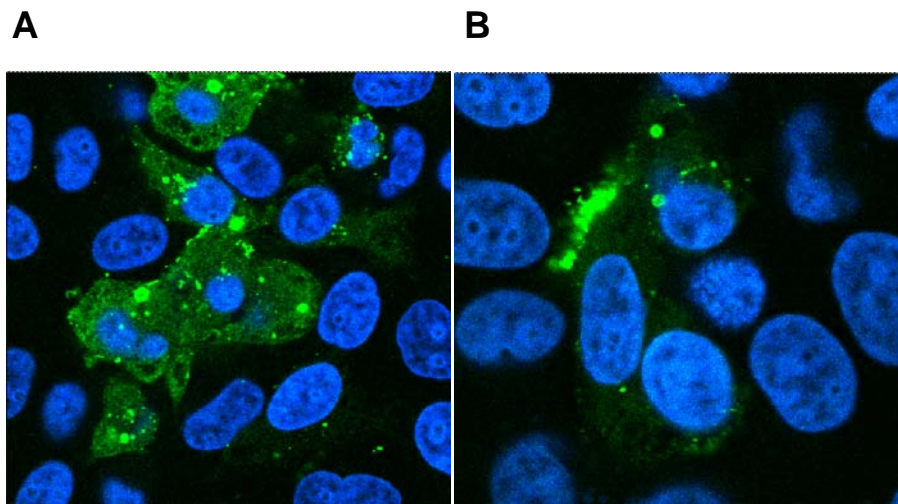


Figure 4.7 Detection of cells infected with the recombinant viruses using the 26/C3/B5 monoclonal antibodies detecting PVM P protein. The infected cells are shown in green surrounded by the uninfected cells. The nuclei of cells were visualised using DAPI dye. A. Cells infected with rPVM-G(15) and B. Cells infected with rPVM-G(J3666).

#### 4.7.1 Growth of recombinant viruses in cell culture

The presence of the inserted mutations in the viral genome [including *SalI*, *XmaI*, *NotI*, *BshTI*, *BglII* and *NotI* restriction sites and deletion of the first ORF of the G genes in strains J3666 and 15 (Warwick)] raised the possibility that the mutations may reduce the growth rate of the viruses *in vitro*. An experiment was designed to test the growth of the recombinant viruses in cell culture. To do this BS-C-1 cells were cultured in 24 well plates in GMEM medium supplemented with 2% FCS. The cells were infected with either rPVM-G(15), rPVM-G(15 $\Delta$ ORF1), rPVM-G(J3666), rPVM-G(J3666 $\Delta$ ORF1), rPVM passage 7, rPVM passage 8, or PVM strain J3666 viruses (Section 4.7) with an moi of 0.05. The cultures were incubated at 31 and 37°C.

Samples were harvested by scraping the cells into the medium at 0, 24, 48, 72, 96, and 120 hr post infection. Three individual samples were harvested from three different infected wells for each time point. All of the samples were kept at -70°C until titration using the micro-plaque assay technique (Section 2.2.7) to detect virus yield. Each titration was conducted in triplicate. An average was calculated for each time point. Results obtained from the titrations are depicted in Figure 4.8.

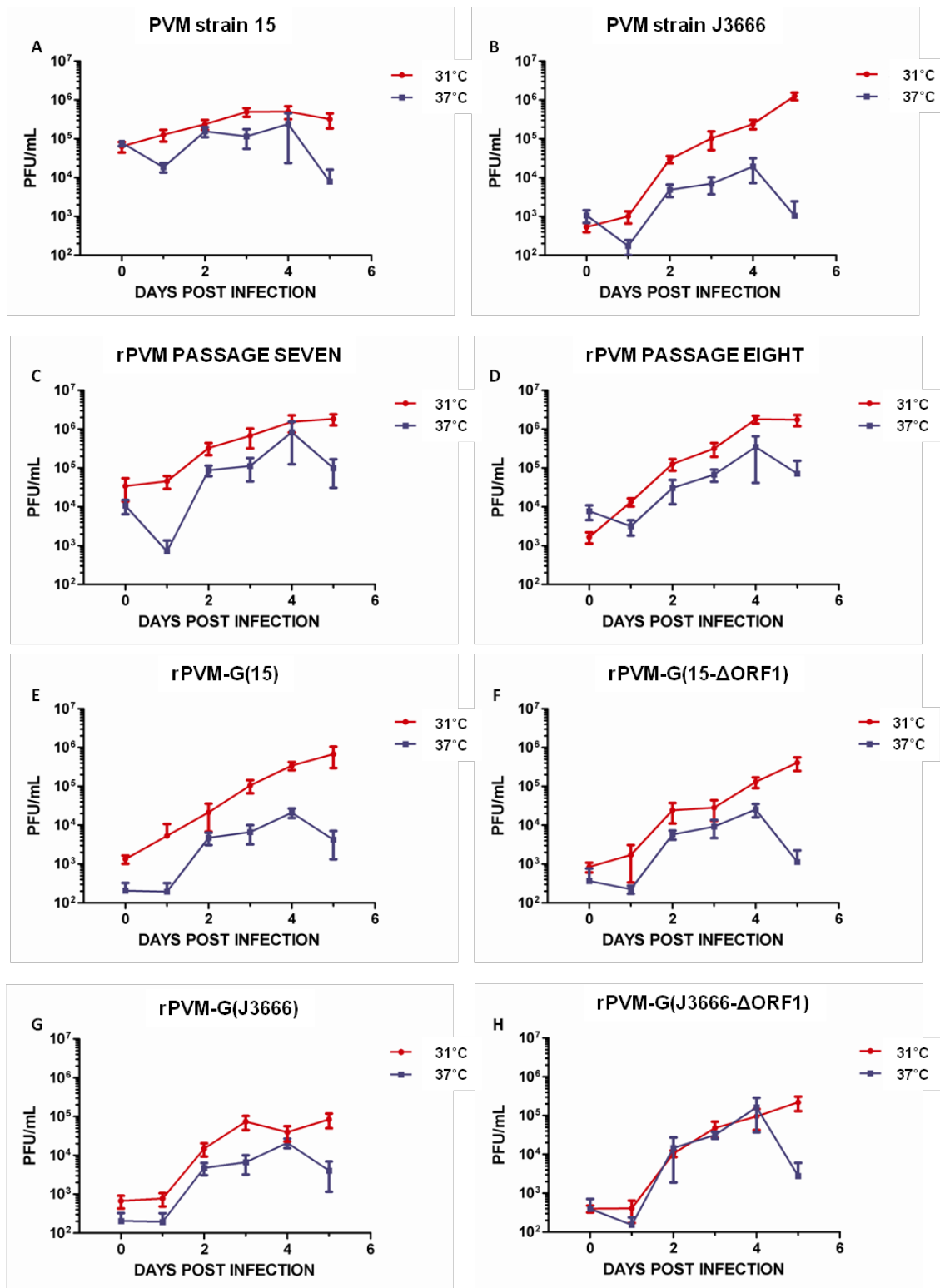


Figure 4.8 Growth of stock or recombinant PVM isolates in BS-C-1 cell culture system at two temperatures, 31°C and 37°C. The growth of the wildtype PVM strain 15, strain J3666, rPVM passage 7, rPVM passage 8, rPVM-G(15), rPVM-G(15ΔORF1), rPVM-G(J3666) and rPVM-G(J3666ΔORF1) are shown in graphs A to H. The curve shown in blue indicates the growth of viruses at 37°C and the curve shown in red indicates the growth of the viruses at 31°C.

The yield of wildtype PVM strain 15 at both 31°C and 37°C was higher on day 0 than for the other viruses tested. The reasons for this are unclear, but the final yield of viruses did not differ greatly from the others. As expected, wildtype PVM strain J3666 showed a significant reduction in yield when grown at 37°C. This virus has previously been shown to be temperature sensitive (Prof. A. Easton, personal communication). In general the yield of all of the viruses was less when grown at 37°C.

The growth rates of two consecutive passages of recombinant virus (rPVM) generated from plasmid p15FL-2G were very similar. For both virus stocks there was a small reduction in yield at 24 hr post infection followed by a rise thereafter when grown at 37°C. The initial drop in yield was not seen when the viruses were grown at 31°C. By day 4 post infection the viruses had reached maximum yields at both growth temperatures. These viruses contained a G gene encoding a protein equivalent to that for PVM strain J3666. This is similar to rPVM-G(J3666ΔORF1).

Comparison of the titres of rPVM passage 7 and rPVM passage 8 virus with rPVM-G(J3666ΔORF1) at 31°C showed that the latter achieved a final yield approximately 10-fold lower than the other two viruses. The only differences in genome sequence between these viruses is that rPVM-G(J3666ΔORF1) contains unique restriction enzyme sites for *Bgl*III and *Bsh*TI introduced to aid cloning. The differences in yield may indicate that the sequence alteration adversely affects the growth of rPVM-G(J3666ΔORF1).

The recombinant viruses carrying the PVM strain 15 G gene showed a similar growth curve to that of rPVM. rPVM-G(15) showed lower yields of virus when grown at 37°C than at 31°C but was propagated at 37°C without difficulty. No substantial differences were seen in the growth characteristics *in vitro* of rPVM-G(J3666) and rPVM-G(J3666ΔORF1).

Overall, the recombinant viruses all grew to similar levels *in vitro* with slightly higher yields achieved when grown at 31°C. The recombinant viruses with a G gene originating from PVM strain J3666 reached slightly lower yields than those with G genes derived from PVM strain 15 (Warwick).

#### 4.8 Testing the pathogenesis of recombinant viruses

The infectivity of the recombinant viruses derived from p15FL-2G clones and rescued from BS-C-1 cells were tested in 4-8 weeks old BALB/c mice. Each virus will be considered in turn.

#### 4.9 rPVM-G(15)

The virus was titrated in BS-C-1 cells and appropriate dilutions to make 250, 500 and 5000 pfu in 50 µl of inocula were prepared. The inoculum was administered intranasally to the animals (Section 2.5). In the control groups mice were inoculated with 50 µl of sterile PBS. Mice were checked and monitored daily. Figure 4.9 and Figure 4.9 show the change in the body weight and clinical signs, respectively for the mice. No weight loss was detected during the time course, and the clinical score of the animals was recorded (Figure 4.10) indicating that the virus was entirely non-pathogenic.

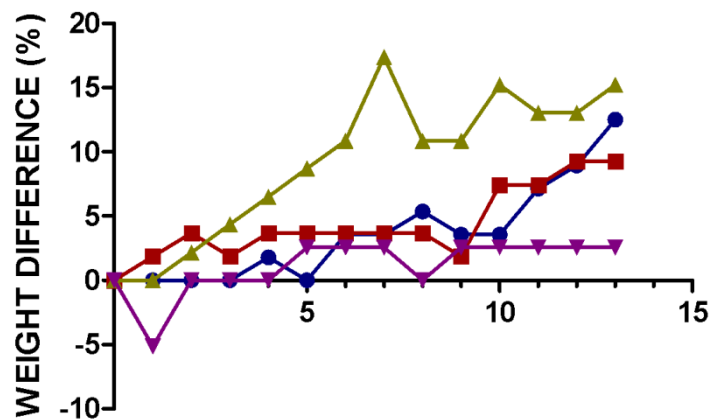


Figure 4.9 Change in body weight in groups of mice infected with rPVM-G(15). Each mouse was infected with 50 µl of the three levels of virus (5000 pfu, 500 pfu and 250 pfu). Body weight was monitored daily and the percentage of change in body weight normalised to the first day is shown. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.

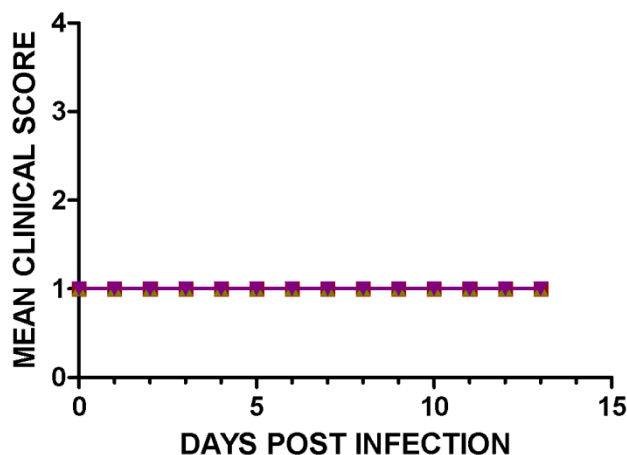


Figure 4.10 Clinical score for mice infected with the rPVM-G(15). Each mouse was infected with 50  $\mu$ l of the three concentrations (5000 pfu, 500 pfu and 250 pfu). Clinical score was monitored on a daily basis and expressed as an average for each experimental group. The data shown is representative of two independent experiments. The inocula were 5000 pfu  $\bullet$ , 500 pfu  $\blacksquare$ , 250 pfu  $\blacktriangle$ , PBS  $\blacktriangledown$ .

#### 4.10 rPVM-G(15 $\Delta$ ORF1)

rPVM-G(15 $\Delta$ ORF1) was used to infect mice and differences in the weight of the infected mice and in the clinical score were monitored (Figure 4.11). Slight weight loss was observed only in the group of mice that received 5000 pfu of the virus (Figure 4.11). Starting from day 6, weight loss reached a peak at day 9. The mice started recovering completely after the day 9. However, no changes in clinical signs of disease were seen Figure 4.12. Although no clinical signs of disease were observed during the infection, the weight loss suggests that deletion of the first ORF of the G gene of strain 15 did alter the virulence of the virus.

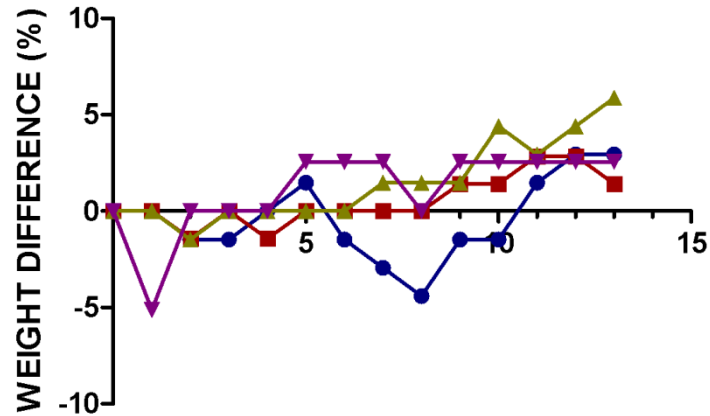


Figure 4.11 Change in body weight in groups of mice infected with rPVM-G(15ΔORF1). Each mouse was infected with 50 μl of the three levels (5000 pfu, 500 pfu and 250 pfu) of rPVM-G(15ΔORF1). Body weight was monitored daily and the percentage of change in body weight normalised to the first day is shown. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.

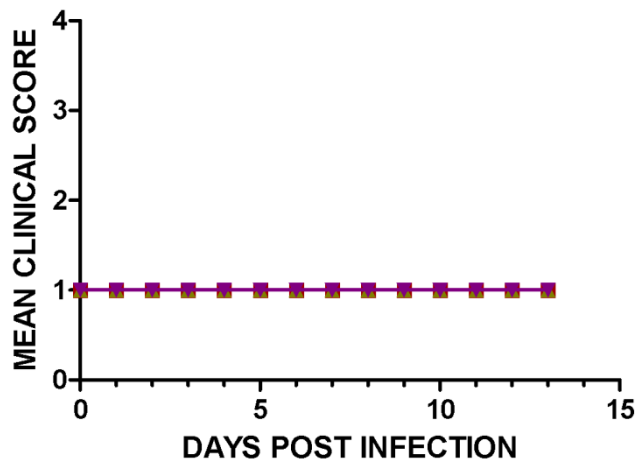


Figure 4.12 Clinical score for mice infected with rPVM-G(15ΔORF1). Each mouse was infected with 50 μl of the three levels (5000 pfu, 500 pfu and 250 pfu). Clinical score was monitored on a daily basis and expressed as an average for each experimental group. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.

#### 4.11 rPVM-G(J3666)

Mice infected with 5000 pfu of rPVM-G(J3666) showed very mild signs of disease and weight loss, but this was not seen with other doses. The level of weight loss was slightly greater at its peak (9.3 %) than seen with rPVM-G(15ΔORF1) (Figure 4.13). The weight loss in these animals started on day 6 and reached its peak at day 9.

The mice recovered completely after day 9. The mice recovered completely after day 9. While the virus was not capable of producing lethal disease in the animals (Figure 4.14), a low level of clinical signs were seen in the mice infected with 5000 pfu. As is normal, the appearance of clinical signs on day 8 followed the first sign of weight loss. Animals infected with 500 and 250 pfu showed no changes in weight or clinical signs of disease. These data indicated that rPVM-G(J3666) was more virulent than rPVM-G(15ΔORF1).

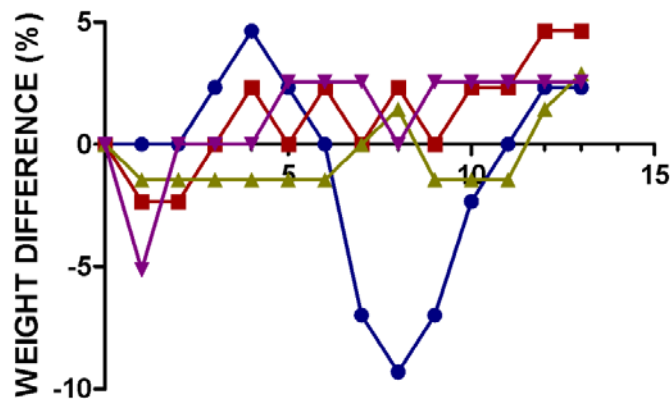


Figure 4.13 Change in body weight in groups of mice infected with rPVM-G(J3666) Each mouse was infected with 50  $\mu$ l of the three concentrations (5000 pfu, 500 pfu and 250 pfu) of rPVM-G(J3666). Body weight was monitored daily and the percentage of change in body weight normalised to the first day is shown. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.

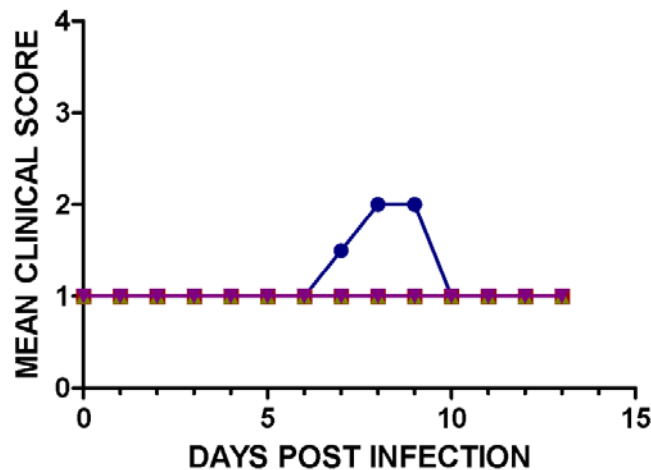


Figure 4.14 Clinical score for mice infected with rPVM-G(J3666). Each mouse was infected with 50  $\mu$ l of the three concentrations (5000 pfu, 500 pfu and 250 pfu) of rPVM-G(J3666). Clinical score was monitored on a daily basis and expressed as an average for each experimental group. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.



#### 4.12 rPVM-G(J3666ΔORF1)

In the group of mice infected with rPVM-G(J3666ΔORF1) no clinical signs of disease or weight loss were observed (Figure 4.15 and Figure 4.16) with any dose of virus.

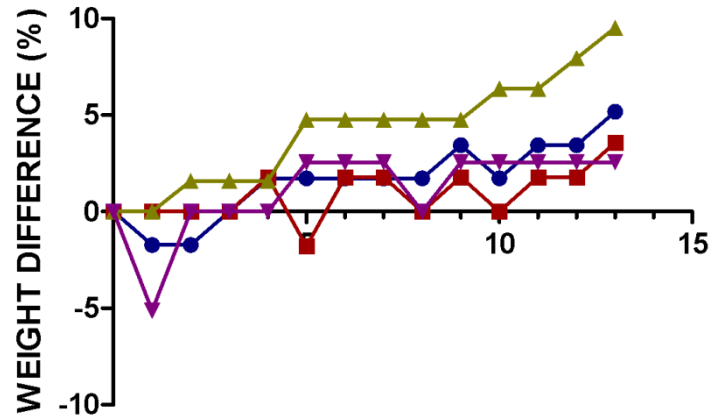


Figure 4.15 Change in body weight in groups of mice infected with rPVM-G(J3666ΔORF1). Each mouse was infected with 50  $\mu$ l of the three concentrations (5000 pfu, 500 pfu and 250 pfu) of rPVM-G(J3666ΔORF1). Body weight was monitored daily and the percentage of change in body weight normalised to the first day is shown. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.

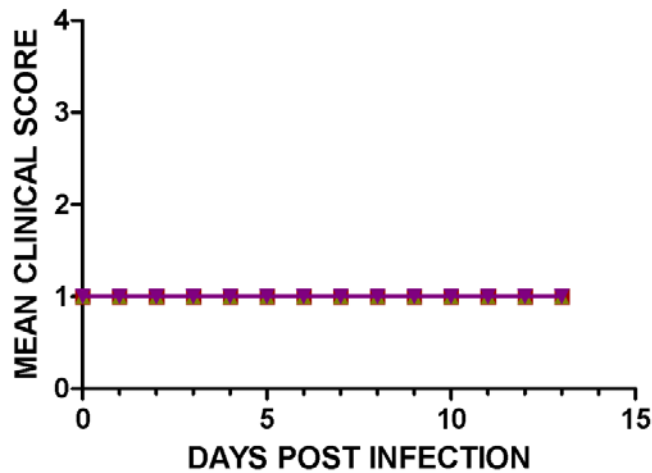


Figure 4.16 Clinical score for mice infected with rPVM-G(J3666ΔORF1). Each mouse was infected with 50  $\mu$ l of the three concentrations (5000  $\mu$  pfu, 500 pfu and 250 pfu) of rPVM-G(J3666ΔORF1). Clinical score was monitored on a daily basis and expressed as an average for each experimental group. The data shown is representative of two independent experiments. The inocula were 5000 pfu ●, 500 pfu ■, 250 pfu ▲, PBS ▼.

#### 4.13 Discussion

In the study presented in this chapter, the nucleotide sequence of the plasmid p15FL-2G was confirmed. The plasmid contained a full length cDNA copy of the PVM strain 15 (Warwick) genome. However, a single nucleotide alteration resulted in the presence of a G gene which expressed a protein similar to that described for PVM strain J3666 but which lacked a short upstream ORF found in the strain J3666 G gene. This plasmid was used to generate infectious virus. Using this genome as a base, four more plasmids were synthesized: pFL2G-G15 carrying the G gene of PVM strain 15 (Warwick), pFL2G-G15 $\Delta$ ORF1 carrying the G gene of PVM strain 15 (Warwick) with its first ORF deleted, pFL2G-GJ3666 carrying the G gene of PVM strain J3666 and pFL2G-GJ3666 $\Delta$ ORF1 carrying the G gene of PVM strain J3666 with its first ORF deleted. The five cDNA named above were used in cell culture to recover infectious viruses, and respectively recombinant viruses rPVM-G(15), rPVM-G(15 $\Delta$ ORF1), rPVM-G(J3666) and rPVM-G(J3666 $\Delta$ ORF1) were recovered.

Recombinant and wildtype viruses demonstrated a better growth rate at 31°C than at 37°C in tissue culture, but all could be propagated at 37°C. The recombinant viruses rPVM-G(J3666) and rPVM-G(J3666 $\Delta$ ORF1) generated 10 fold less yield than the other recombinant viruses. The virulence of the recombinant viruses in mice was examined and showed that rPVM-G(15), which has the same G gene organisation as the non-pathogenic PVM strain 15 (Warwick) was not pathogenic when mice were infected with 250, 500 or 5000 pfu. However, when mice were infected with 5000 pfu of rPVM-G(15 $\Delta$ ORF1) a small (4%) but significant weight loss was seen compared to the weight of the animals at the time of challenge. Control, uninfected, mice and animals infected with 250 pfu or 500 pfu of rPVM(15 $\Delta$ ORF1) showed a gain in weight, and no clinical signs of disease. Removal of the first ORF in PVM strain 15 (Warwick), which overlaps with the ORF encoding the G glycoprotein, results in a small but measurable increase in pathogenicity.

The G gene of the pathogenic strain J3666 contains a short ORF upstream, but not overlapping, the ORF which encodes the G glycoprotein. The PVM strain J3666 G glycoprotein contains additional amino acids at the amino terminus compared to the G protein of strain 15 (Warwick). When mice were infected with rPVM-G(J3666), those receiving 5000 pfu showed weight loss which reached a maximum of 9% of the starting

weight on day 9 post infection and mild clinical signs of disease which followed the weight loss, beginning on day 8.

Animals infected with 250 pfu or 500 pfu of rPVM-G(J3666) showed little change in weight and no clinical signs of disease. All animals challenged with any of the 3 selected doses of rPVM-G(J3666 $\Delta$ ORF1) remained clinically normal and showed no significant difference in weight.

During this research, it was assumed that the low number of passages may reduce the possibility of generating DI particles. Taken together, the data suggest that the presence of the first ORF in the G gene of PVM may play an important role in virus virulence. A possible explanation may be that the upstream ORF affects the level of expression of the G protein from the downstream ORF. It is well known that the presence of an upstream ORF in cellular mRNAs frequently leads to a reduction in the level of protein produced by the downstream ORF (Child *et al.*, 1999b; Kozak, 1984). Thus, pathogenicity of PVM may be affected by the nature and amount of G protein produced. The effect of the first ORF in the G gene of PVM was examined in more detail in Chapter 5.

During construction of the plasmid p15FL-2G the area between nucleotides 900 and 1300 of the F gene was reported to containing unclear sequences (Dibben, 2006). This conclusion was made after sequencing the plasmid with two individual primers (F9 and F20) on several occasions (Dibben, 2006). In the work presented here, the sequences were obtained using the same primer (primer F9, Appendix I) which was used by Dibben (2006). Interestingly, the same ambiguous sequencing results with trace containing more than one peak at each nucleotide were achieved. The reason for the diversity in the nucleotide sequence of the F gene in this region remains unclear. Due to time constraints, the mutation in the F gene of plasmids pFL2G-G15 $\Delta$ ORF1 and pFL2G-GJ3666 was left unchanged.

## **CHAPTER 5**

# **EFFECT OF THE FIRST ORF ON PVM G PROTEIN EXPRESSION**

## 5.1 Introduction

As discussed in Section 1.2.1.1.1, the G gene of PVM strain 15 (Warwick) contains two overlapping open reading frames. In the G gene of PVM strain J3666, however, a shorter ORF precedes the main ORF of the gene, and the main ORF is longer than that of strain 15 (Warwick) (Section 1.2.1.1.1). In other systems it has been shown that the presence or absence of an upstream ORF can affect the translational level of the second ORF (Meijer *et al.*, 2000; Peabody & Berg, 1986; Poyry *et al.*, 2004; Wang & Sachs, 1997). The consistent presence of an upstream ORF in the G gene in PVM strains raises the question of the contribution of the first ORF in controlling the expression level of G glycoprotein from the larger second ORF.

To analyse the contribution of the first ORF of the G gene in translation of the G glycoprotein, it was necessary to compare different levels of G protein in the absence and presence of the first ORF. For this purpose, a reverse genetics system using a synthetic minigenome was employed. A dicistronic minigenome system expressing two reporter genes was constructed for PVM. The first gene transcribed from the synthetic genome expressed the luciferase (Luc) protein. The second gene expressed the G glycoprotein from either PVM strain 15 (Warwick) or strain J3666. To measure the quantity of G protein expressed, the G gene was tagged at its 3' end with green fluorescent protein (GFP) and the amount of GFP was measured using an ELISA, based on the assumption that the quantity of GFP would reflect the quantity of G protein expressed. The activity of luciferase protein coded from the sequence in the first cistron of the minigenome was used to normalise the data. Using these basic constructions it was possible to assess the effect of the presence and absence of the upstream ORF in the G gene. The normalised results were compared to each other to study the effect of first ORF in the translational control of the G gene in both strains of PVM.

## 5.2 Construction of the PVM minigenome carrying G gene of PVM

The plasmid was made based on the available dicistronic minigenome plasmid (pWt/NdeI) by using the presence of the naturally occurring restriction sites, and introducing a novel restriction site upstream of the CAT gene in the minigenome. The process used in the approach is summarised in Figure 5.1.

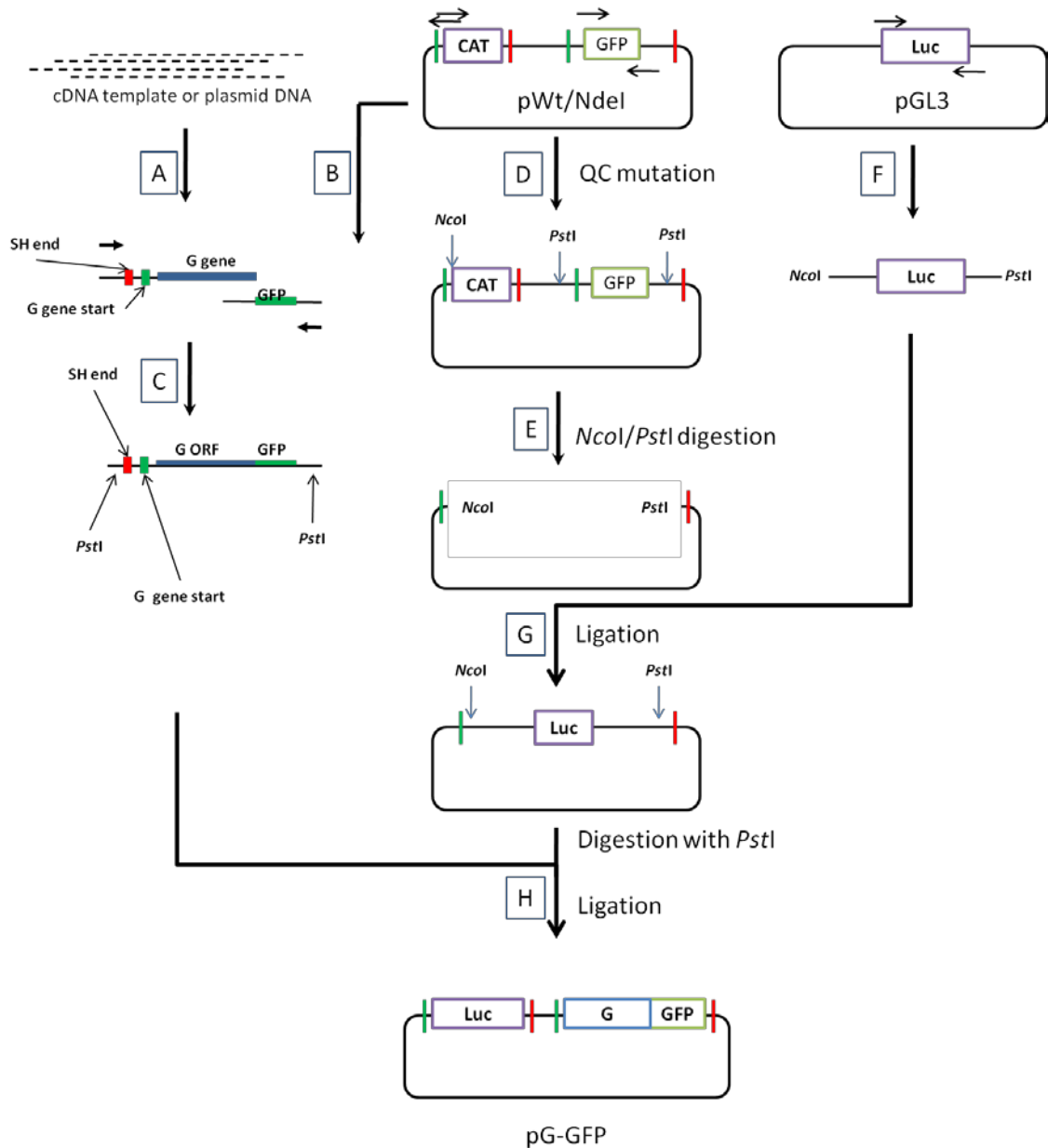


Figure 5.1 The synthesis of pG-GFP minigenome. The construction of pG-GFP was performed in 8 steps indicated in the diagram as steps A to H. In step A, the G gene was amplified using cDNA or plasmid DNA templates; in step B, GFP was amplified using pWt/NdeI as the template. For step C, the G and GFP genes were joined together using an overlap PCR technique. Step D, an *NcoI* restriction site was inserted in pWt/NdeI using quick change mutagenesis technique. Subsequently, in step E, the mutated pWt/NdeI was digested using *NcoI* and *PstI* restriction enzymes. Subsequently, in step F, the luciferase gene was amplified from pGL3 template DNA, and ligated (step F) into the digested pWt/NdeI DNA. Finally in step H, the products of steps G and C were ligated together to make the pG-GFP (see text for further detail).

The plasmid pWt/NdeI (Figure 5.2) was constructed and optimised by O. Dibben and L. Thorpe to analyse the effect of gene leaders and gene trailers on the protein synthesis of PVM (Dibben & Easton, 2007; Dibben *et al.*, 2008). It consists of

two genes expressing chloramphenicol acetyl transferase (CAT) and green fluorescent protein (GFP) each of them flanked by PVM gene start (GS) and gene end (GE) signals making them two individual genes controlled by PVM transcription regulation sequences. The genomic leader and trailer regions of the PVM genome RNA flank both genes, and the whole area was bounded by T7 transcription start and stop sequences. T7 directed transcription produces an RNA molecule representing a synthetic (mini-) genome of PVM which can be replicated and transcribed by the relevant virus protein complexes.

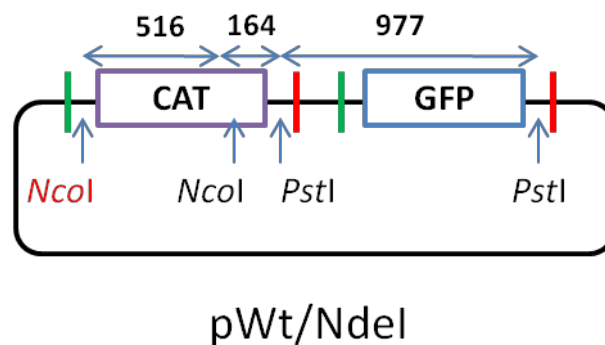


Figure 5.2 Schematic structure of pWt/NdeI. Naturally occurring restriction sites are shown in black and the recombinant restriction site introduced upstream of the CAT gene is shown in red. Gene start sequences are shown with vertical lines and in green, likewise gene end sequences are shown in red. The sizes (bp) of DNA between the restriction sites are shown.

### 5.3 Quick change mutagenesis to make pWt/NcoI

Due to lack of a suitable quantitative detection assay for CAT as the product of the first gene in the minigenome this was removed from the first cistron in the pWt/NdeI minigenome and replaced with the firefly luciferase gene which is readily detectable (Section 2.4.2). For this, two point mutations were inserted upstream of the translation initiation of the CAT gene in pWt/NdeI, to make an *NcoI* restriction site, using the quick change mutagenesis technique (Section 2.3.12). NCOIFOR and NCOIREV primers (Appendix I) were designed for this purpose to mutate TA to CC at position 55 and 56 of the pWt/NdeI (Figure 5.3). The mutated plasmid was designated pNcoI/PstI. Digestion of pNcoI/PstI with *NcoI* and *PstI* restriction enzymes was used to identify the successfully mutated plasmid. As indicated in Figure 5.2 digestion of pNcoI/PstI with *NcoI* and *PstI* restriction enzymes resulted in production of 516, 164, 977 and 3263 bp fragments (Figure 5.4).

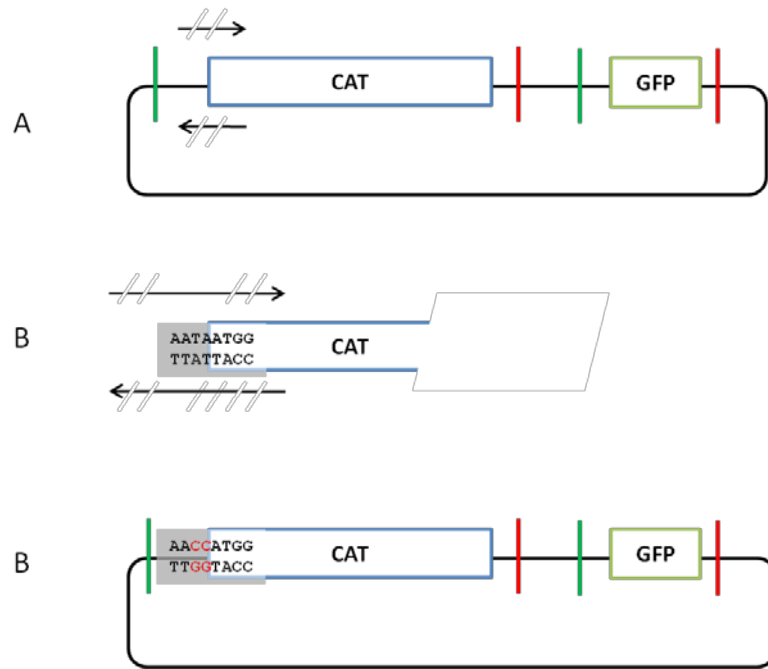


Figure 5.3 Quick Change mutagenesis reaction to introduce an *NcoI* restriction site before the start codon of CAT gene. A. The approximate position of the mutagenesis primers to mutate the nucleotides 55 and 56 of the plasmid pWt/NdeI. B. The mutagenesis site on the plasmid pWt/NdeI. C. Both nucleotides at residues 55 and 66 of the plasmid pWt/NdeI were mutated to cytosine. The primer pair used in the reaction is shown as two black arrows named *NcoI*Rev and *NcoI*For.

#### 5.4 The Luciferase gene region

Amplification of the luciferase gene was conducted using plasmid pGL3 (Promega) as the template and primers *Pst*LucR and *Nco*LucF (Appendix I). The primer *Nco*LucF was designed in such a way as to amplify *NcoI* restriction site from the pGL3 plasmid and the primer *Pst*LucR was designed to add the *PstI* restriction site to the 3' end of the 1653 base pair fragment. Gel purification of the PCR product was followed by digestion of the amplified luciferase gene with *NcoI* and *PstI*. The digested luciferase gene was then ligated into pNco/*PstI* which had also been digested with *NcoI* and *PstI*. Six colonies were randomly selected and the colonies containing pNco/*PstI*-Luc were identified after DNA purification and digestion of the purified DNA using *PstI* and *NcoI* restriction enzymes. Figure 5.4 shows a typical digestion result of three of the colonies; the 1653 bp fragment of the luciferase gene was released upon digestion with *PstI* and *NcoI* restriction enzymes.



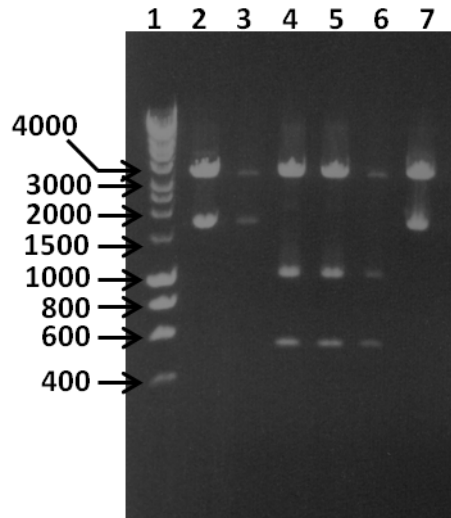


Figure 5.4 Restriction enzyme digestion of pWt/NcoI and pNcoI/PstI with *PstI* and *NcoI* restriction enzymes. Lanes 1 to 7 contain DNA size markers, pWt/NcoI-Luc colony number 1, pWt/NcoI-Luc colony number 2, pNcoI/PstI colony number 1, pNcoI/PstI colony number 2, pNcoI/PstI colony number 3, and pWt/NcoI-Luc colony number 3, respectively.

## 5.5 The G-GFP region

To measure the quantity of expressed G protein, a GFP tag was inserted at the C terminus. To join the two genes together an overlap PCR technique was used.

### 5.5.1 Amplifying the G gene from PVM strains; general features

In the G gene of the PVM strains 15 (Warwick) and J3666, there are 14 nucleotide differences including 13 point mutations and one single nucleotide deletion (Randhawa *et al.*, 1995). For this analysis it was essential to include all of the gene differences in the minigenome. Therefore, it was decided to make two minigenomes each containing the G gene from one PVM strain. To achieve this, the gene was amplified from each strain individually. The plasmid pF3-BBA (Section 4.2) carrying a cDNA copy of the relevant region of strain 15 (Warwick), was used as the template to amplify the G gene. A preparation of cDNA was generated from RNA of PVM strain J3666 (Section 2.3.10) and used as the template for amplification of the G gene for PVM strain J3666.

The presence of a PVM gene start and gene end sequence was required to direct transcription of the G genes in the minigenome construct. These were present on the minigenome pWt/NdeI. Digestion of pWt/NdeI with *PstI* and *NcoI* retained the gene start and stop sequence flanking the Luc gene. The authentic gene start of the G gene,

occurring naturally in the genome of the PVM, was retained in the amplified sequence of the G gene. A reverse primer (PstIRev: Appendix I) was designed based on the 3' end of G gene, which is conserved between both strains. Three features were considered for the primer; first, it was necessary to remove the stop codon from the G gene to fuse together the reading frame of the G and GFP genes; second, for the overlap PCR reaction it was necessary to add the initial nucleotides of the GFP gene to the 3' end of G gene. To achieve this, the reverse primer amplifying the G gene was designed to carry 19 nucleotides of the complementary sequence from the beginning of GFP gene as a tag, and similarly 20 nucleotides from 3' end of G gene was added to the forward primer to amplify GFP gene.

#### 5.5.1.1 Amplifying G gene from PVM strain 15 template

Amplification of the G gene from pF3-BBA template (Section 4.2) with primers PstIFor and PstIRev (Appendix I) generated the expected product with a length of 1393 nucleotides and a non-specific product with about 800 base pairs (Figure 5.5). The 1393 base pair PCR product was excised and purified from the gel.

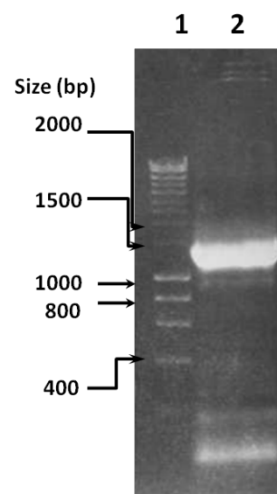


Figure 5.5 PCR amplification of the PVM strain 15 (Warwick) G gene using primers PstIFor and PstIRev. Lane 1 shows the Hyperladder I size marker. The sizes of the markers are shown on the left. Lane 2 shows the 1393 nt product of amplification of the PVM strain 15 (Warwick) G gene.

#### 5.5.1.2 Amplifying the G gene from strain J3666 template

To make a minigenome of the G gene with the organisation found in PVM strain J3666, the G gene was amplified from cDNA from PVM strain J3666 RNA. For this purpose, RNA extracted directly from 250 µl of cell culture harvested virus, and cDNA

was made from both sense and antisense RNA of the virus using random primers. The G gene was amplified using the PstIRev and the PstIGJ3666 primer pair (Appendix I) and used in an overlap PCR as described earlier.

### 5.5.2 Amplification of the GFP gene

The GFP gene was amplified from pWt/NdeI using GFP-For and GFP-Rev primer pair (Appendix I). The primer pair was designed to leave a tag complementary to the 3' end of the genomic sense of the G gene at the 5' end of the PCR product, as before. This would give the opportunity to use it in overlap PCR to join the G and GFP genes together. A *PstI* restriction site was placed at the 3' end of the GFP gene for cloning purposes. The PCR products were separated by agarose gel electrophoresis (Figure 5.6) and the 779 nucleotide fragment was excised from the gel.

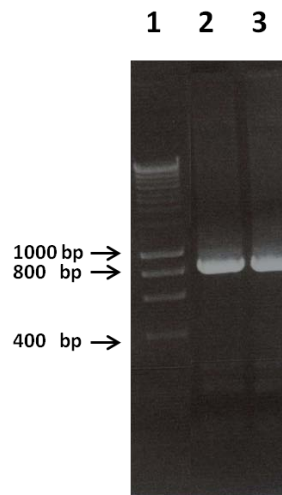


Figure 5.6 Amplification of the GFP gene from pWT/NdeI DNA. Lane 1 shows the, Hyperladder I size markers. The sizes of the markers are shown on the left. Lanes 2 and 3 shows the PCR product with the expected 779 base pair size.

### 5.5.3 The overlap PCR to join G and GFP together

To conduct a successful overlap PCR it was necessary to design overlapping primers for both G and GFP genes. For this, the primer GFPFor was designed carrying 23 nucleotides from the 3' end of the G gene. As a result, the amplified fragment was expected to carry the 23 nucleotide of the G gene at the 5' end of the fragment (Section 5.5.2). The second primer, PstIRev, was designed to carry the first 19 nucleotides from the sense strand of the GFP fragment. This would result in amplification of a fragment from the G gene containing 19 nucleotides from the GFP

gene. It was assumed that in a PCR reaction the overlapped region of the fragments would anneal to each other and act as primers and synthesise the other fragment leading to the completion of the G-GFP fragment synthesis. Using the PstIFor primer, for the strain 15 (Warwick), or PstIGJ3666 primer, for the strain J3666, as the forward primers individually and GFPRev as the reverse primer would give the opportunity to further amplify the G-GFP fragment. The principle of the overlap PCR which was used here is explained in Figure 5.7. The overlap sequence in the sense strand of the G fragment annealing to the overlap sequence of the sense strand of the GFP fragment is shown in Figure 5.8.

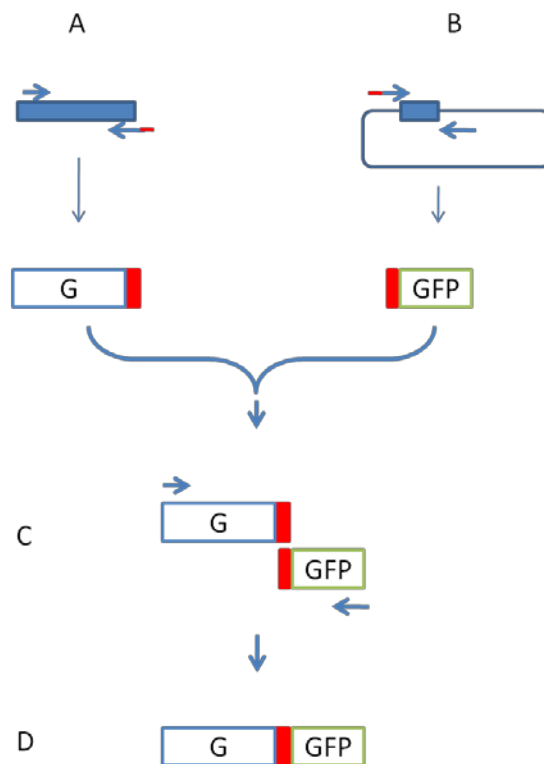


Figure 5.7 The overlap PCR method used to join the G and GFP fragments to synthesis the G-GFP fragments for the strains 15 and J3666. A. Amplification of the G fragment using PstFor, for the strains 15 (Warwick), or PstIGJ3666, for the strain J3666, as the forward primers and PstRev as the reverse primer. The pF3-BBA was used as the template to synthesis G(15) for PVM fragment. The RT-PCR was used to amplify the G gene from genomic RNA isolated from strain J3666. B. Amplification of the GFP fragment using the GFPFor and GFPRev primer pair. C. The overlap PCR reaction. D. the fragment G-GFP. The PstRev and GFPFor primers are shown in blue carrying the overlapping region (shown in red). The G fragment is shown in blue with the overlap sequence at its 3' end (red). The GFP fragment is shown in green with the overlap sequence at its 5' end (red).

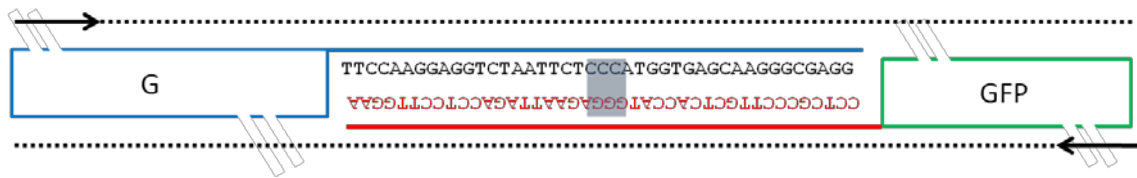


Figure 5.8 The overlap PCR reaction. The fragment G (blue) and the fragment GFP (green) are shown in the overlap PCR reaction. The overlapped regions between the sense G fragment and antisense GFP fragments are annealed together. The stop codon in G fragment was replaced with “CCC” and shown inside a blue box.

To optimise the overlap PCR reaction for the G gene of PVM strain 15, an overlap PCR was conducted in buffers with either 2, 3 or 4 mM of  $Mg^{2+}$ . The annealing temperature was set at 58°C and kept constant during optimisation of  $Mg^{2+}$  optimisation, based on the assumption that higher annealing temperature would eliminate the possibility of production of nonspecific inter- and intra-molecular secondary structures. The process was used to achieve correct annealing between 5' and 3' ends of the GFP and G genes. The highest yield was achieved with concentrations of 3 and 4 mM  $Mg^{2+}$  (Figure 5.9). The PCR product was gel purified and kept at -20°C until the time to be used. Equal concentrations of the G (Section 5.5.1) and the GFP fragments (Section 5.5.2) were used as the template. The GFPFor and PstIRev primers (Appendix I) were designed to overlap at their 5' end. Therefore, it was expected that the fragments G and GFP contain overlapping ends at their 3' (for the G fragment) and 5' (for the GFP fragment) ends. The primers PstIFor and GFPRev were used in the overlap PCR .

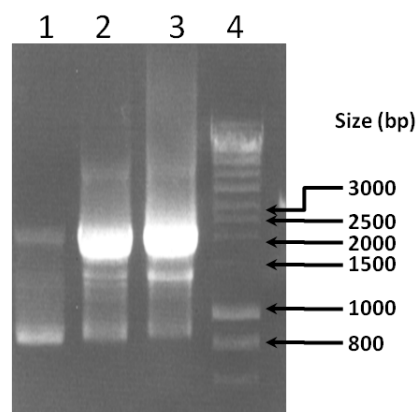


Figure 5.9 The overlap PCR to join G and GFP fragments together. Gel purified G and GFP genes from previous PCR amplifications were used with different  $Mg^{2+}$  concentrations. Lanes 1, 2 and 3 show the products of reaction in the presence of 2, 3 and 4 mM  $Mg^{2+}$  , respectively. Lane 4 shows the Hyperladder I size markers with the sizes shown on the right.

## 5.6 Ligation of G-GFP fragments into pNcoI/PstI-Luc

G-GFP fragments obtained from the process explained above were digested with *PstI* restriction enzyme and ligated into *PstI* digested pNcoI/*PstI*-Luc vector (Figure 5.10). Following transformation of *E. coli*, plasmid DNA prepared from bacterial colonies was digested with *PstI* and the plasmids carrying the G-GFP fragment were selected and sequenced using primer G4878(R) (Appendix I) to confirm the orientation of the G-GFP fragment in the vector.

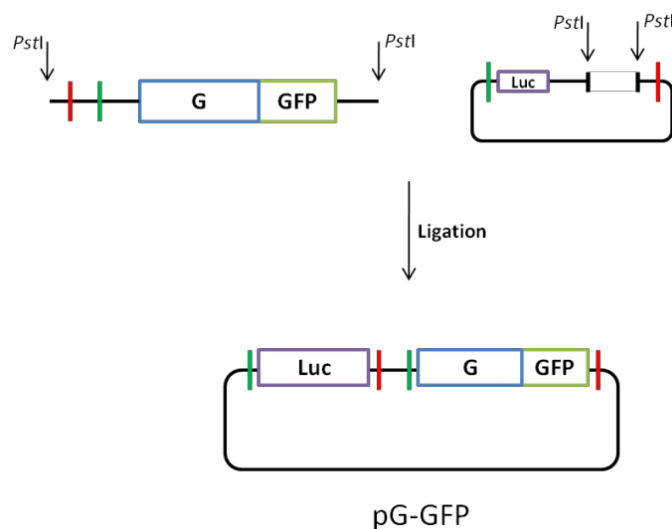


Figure 5.10 Ligation of the G-GFP fragment into pNcoI/*PstI*-Luc. The *PstI* digested G-GFP was ligated into the *PstI* digested pNcoI/*PstI*-Luc vector. The PVM gene start and stop sequences are shown in green and red, respectively.

## 5.7 Deletion of the first ORF in both constructs

Both constructed G-GFP minigenome plasmids (pG-GFP-15 and pG-GFP-J3666) were used in a single step quick change mutagenesis PCR reaction to mutate the start codon of the first ORF to GCG using primers 1stORFOut-J6-F and 1stORFOut-J6-R (Appendix I) for pG-GFPJ3666 and primers 1stORFOut-15-R and 1stORFOut-15-F for pG-GFP-15. This ensured that the first ORF in both constructs could no longer be translated and the transcribed mRNA contained only one single ORF (Section 1.2.1.1). The mutated plasmids were referred to as pG-GFP-15 $\Delta$ ORF-1 and pG-GFPJ3666 $\Delta$ ORF-1 (Figure 5.11).

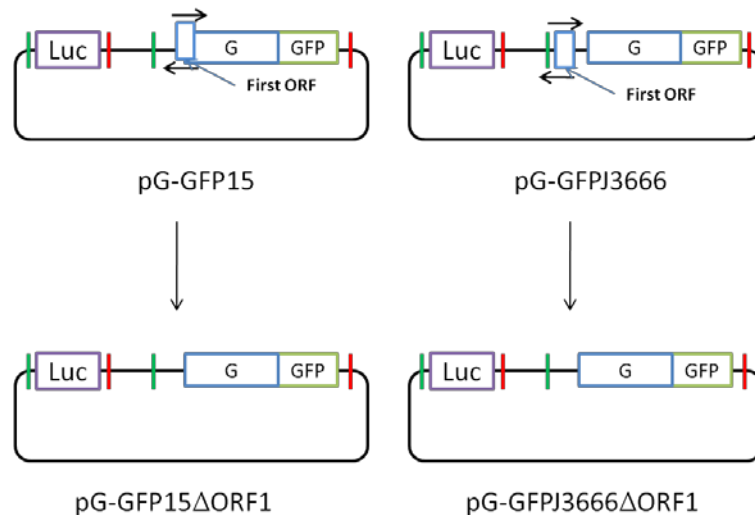


Figure 5.11 Deletion of the first ORF of the G gene in pG-GFP15 and pG-GFPJ3666. The primer pair 1stORFOut-15-F and 1stORFOut-15-R, for the strain 15, and the primer pair 1stORFOut-J6-F and 1stORFOut-J6-R as indicated with black arrows, for the strain J3666, were used to mutate the start codon of the first ORF of the G gene to GCG. The Luc gene is indicated. The PVM gene start and stop sequences are shown with green and red vertical lines. The first ORF in each construct is indicated.

## 5.8 Correction of the first ORF stop codon in pG-GFPJ3666

As discussed in Chapter 4, the sequence of the G gene in strain J3666 did not contain the stop codon for the first ORF of the G gene. As a result, and because the aim of the experiment was to analyse the effect of the first ORF in expression of the main ORF of the G gene, a Quick Change mutagenesis PCR primer pair [MG\_STOP(F) and MG\_STOP(R)] was designed to restore the stop codon for the first ORF. Following the mutagenesis PCR, the *DpnI* digested PCR products were used to transform competent *E. coli*. The sequence of the clones obtained and one of the clones containing the corrected stop codon was randomly selected.

## 5.9 Cloning the G-GFP fragments under T7 promoter

Before transfection of the pG-GFP plasmids into the BSR-T7/5 cells, it was decided to test the expression of the G-GFP ORF independent from the PVM transcription system. For this purpose, the G-GFP fragments were cloned into pBlueScript II plasmids (Figure 5.12). The presence of the T7 polymerase promoter upstream of the G-GFP ORF would provide the opportunity to control the ORF transcription in the presence of the T7 polymerase. BS-C-1 cells infected with a recombinant vaccinia virus expressing bacteriophage T7 RNA polymerase was chosen for this purpose.

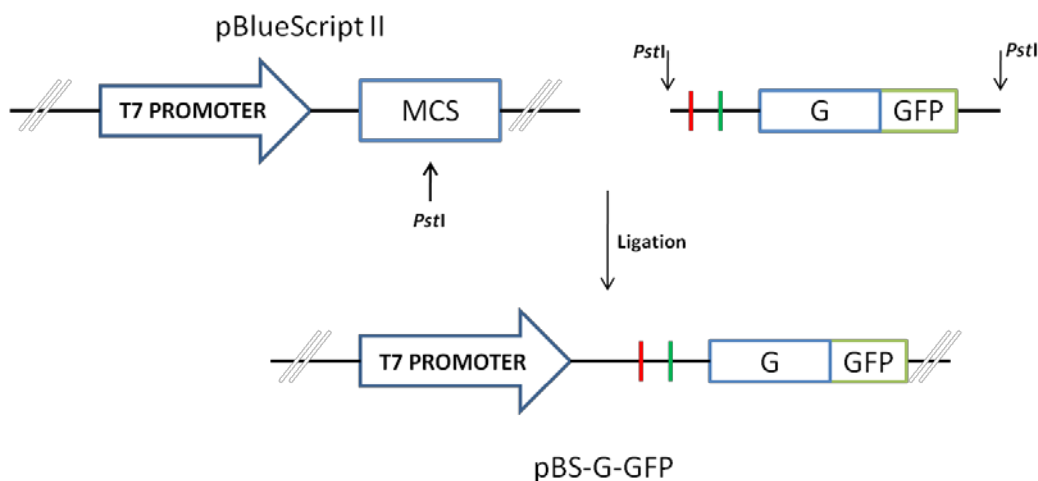


Figure 5.12 Strategy used for constructing pBS-G-GFP plasmids. The multiple cloning site and the T7 promoter of pBlueScript II are depicted. The *PstI* restriction sites are indicated in pBlueScript II vector and G-GFP fragment. The digested G-GFP fragments were inserted into the digested pBlueScript II vector. The PVM gene end and start sequences are shown in red and green, respectively. The gene end and start sequences were inserted into the pBS-G-GFP plasmids during cloning procedure but was not in use during the experiment.

The G-GFP fragment was digested from pG-GFP15, pG-GFP15 $\Delta$ ORF1, pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1 plasmids individually by digestion with *PstI*, and the fragments were purified. These fragments were inserted into pBluescriptII, digested with *PstI*. The ligation mixture was transformed into competent *E. coli* (Sections 2.3.5 and 2.3.6), and the transformed bacteria were cultured on LB agar supplemented with IPTG and X-Gal (Section 2.1.13 and Section 2.3.9). On the following day, 3 white colonies for each construct (pBS-G-GFP15, pBS-G-GFP15 $\Delta$ ORF1, pBS-G-GFPJ3666 and pBSG-GFPJ3666 $\Delta$ ORF1) were selected and grown in liquid cultures to provide DNA. The presence of the G-GFP fragment in pBS-G-GFP15, pBS-G-GFP15 $\Delta$ ORF1, pBS-G-GFPJ3666 and pBSG-GFPJ3666 $\Delta$ ORF1 constructs was confirmed by digestion with *PstI* restriction enzyme. The orientation of the start codon of the G-GFP fragments in 3 plasmid stocks prepared from each of the pBS-G-GFP15, pBS-G-GFP15 $\Delta$ ORF1, pBS-G-GFPJ3666 and pBSG-GFPJ3666 $\Delta$ ORF1 plasmids was confirmed by sequencing the constructs using universal T7 and G4878(R) primers (Appendix I).



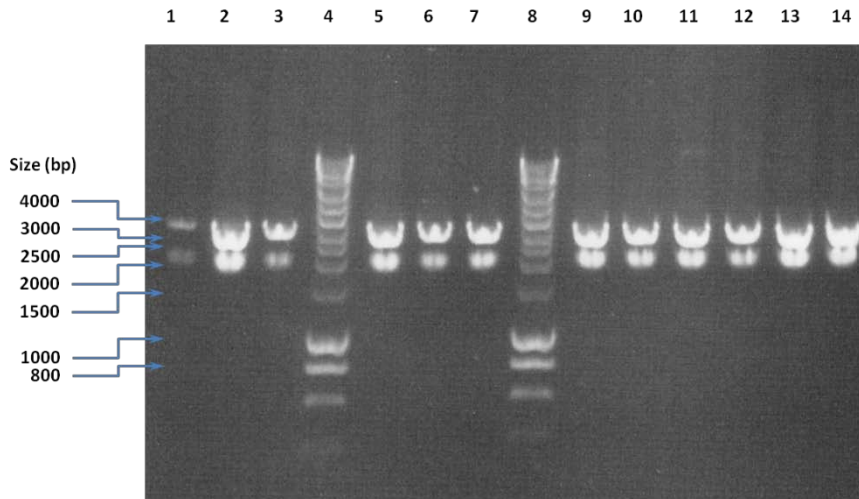


Figure 5.13 PstI digestion of the pBS-GGFP constructs. Lanes 4 and 8 contain DNA size markers. Lanes 1 – 3 contain pstI digested pBS-G-GFP15, lanes 5 – 7 contain pBS-G-GFP15 $\Delta$ ORF1, and lanes 9 -14 contain pBS-G-GFPJ3666 and pBSG-GFPJ3666 $\Delta$ ORF1. The DNA size markers are indicated on the left.

### 5.10 Expression of the G-GFP fusion proteins in Vaccinia Virus T7 infected cells

The pT7-G-GFP constructs were transfected into BS-C-1 cells infected with recombinant vaccinia virus expressing bacteriophage T7 RNA polymerase (moi= 1). The lysate was used in a GFP ELISA to standardise loading on a polyacrylamide gel (Section 2.4.1).

The expression of the G-GFP fusion protein was confirmed by western blotting (Section 2.4.1). Protein samples were incubated at 37°C for 30 min prior to electrophoresis on an 8% polyacrylamide gel (Puffer *et al.*, 2000). The results of the western blot analysis are shown in Figure 5.14.

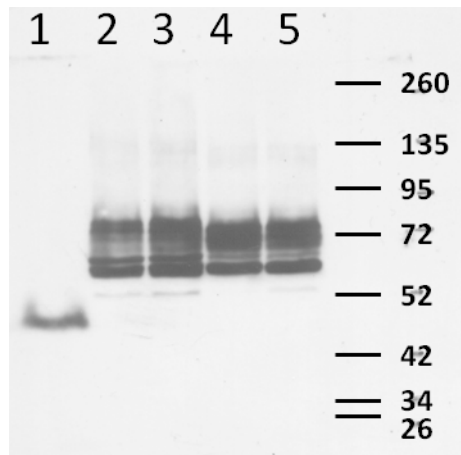


Figure 5.14 Detection of G-GFP fusion protein using SDS-PAGE and western blot. GFP was detected as described in Section 2.3.18. Numbers on the left indicate the position of the protein weight markers (Spectra™, Fermentas). Numbers on the top indicate the lane number. Lanes 1 to 5 contain recombinant GFP, BS-C-1 cell lysates infected with T7 expressing vaccinia virus and transfected with pBS-G-GFPJ3666ΔORF1, pBS-G-GFPJ3666, pBS-G-GFP15ΔORF1, and pBS-G-GFP15, respectively. All samples incubated for 30 min at 37°C as explained in the text.

The relative molecular mass ( $M_r$ ) of GFP is 26900. GFP readily forms a dimer which can be seen in lane 1 in Figure 5.14. The non-covalently formed dimer is stable when incubated at 37°C prior to electrophoresis. This was necessary as no GFP was detected on western blots when the protein samples were heated to 100°C (data not shown). The molecular weight of non-glycosylated PVM strain J3666 G protein is predicted to be 43600 and the nonglycosylated PVM strain 15 G protein is predicted to be 39800 (Randhawa et al., 1995). Following glycosylation the G protein increases in size by approximately 40000. In Figure 5.14 the G-GFP fusion protein from both strains 15 and J3666 G gene produced two bands one of approximately 58000 and a larger band of approximately 73000 with several bands between these two sizes. It is unlikely that the G-GFP fusion protein will be glycosylated and the 73000 protein is likely to be the full length fusion of the G and the GFP proteins. The smaller proteins are most likely to be cleavage products which have been degraded at the amino terminus. Degradation of fusion proteins is commonly observed. However, it is also possible that the PVM G protein contain a furin - like cleavage site which is responsible for the production of the smaller products discussed below (Section 5.11). The data clearly show that the G gene constructs are capable of directing the synthesis of the expected fusion proteins.

### 5.11 Measuring the G-GFP fusion proteins expression in BSRT-7 cells transfected with pG-GFP plasmids

In Section 5.9, the expression of G-GFP fragments cloned in pBlueScript II vectors were shown. To achieve the expression of the G-GFP in the minigenome constructs, BSRT/7-5 cells, which constitutively express bacteriophage T7 RNA polymerase, were prepared in 6 well plates. Each of the pG-GFP15, pG-GFP15 $\Delta$ ORF1, pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1 plasmids was transfected into BSR-T7/5 cells along with the plasmids coding for the RNP complex, P, N, L and M2-1 proteins which are necessary for controlling RNA transcription of PVM (Section 1.2.3). Table 5.1 shows the optimised amounts of the plasmid DNA used in the transfection (Dibben & Easton, 2007; Dibben *et al.*, 2008). In each experiment, three wells of the BSR-T7/5 cells were not transfected with the plasmids and were designated as the negative controls. The negative controls were treated with the same amount of the transfection reagent. For each of the pG-GFP plasmids the transfection was performed in triplicate. The experiment was repeated on two different occasions.

	pN	pM2-1	pL	pP	Minigenome
Concentration ( $\mu$ g)	0.4	0.1	0.2	0.2	0.4

Table 5.1 The concentrations of the minigenomes and helper plasmids, coding for the RNP complex of PVM, used to transfect BSR-T7/5 cells (Dibben & Easton, 2007; Dibben *et al.*, 2008). pP, pL, pM2-1 and pN indicate the name of plasmids coding for P, L, M2-1, and N proteins of PVM.

The transfected BSR-T7/5 cells were incubated for 48 hr at 37°C. Thereafter, the transfected BSR-T7/5 cells and the negative control cells were lysed in CCLR 1X lysing buffer (Section 2.4.1). The samples were used directly in the GFP ELISA to measure the level of the G-GFP fusion protein (Section 2.4.1). The GFP ELISA was performed in triplicate. The luciferase activity in the samples expressed from the Luc gene in the first cistron of the G-GFP minigenomes was determined as described in Section 2.4.2. Table 5.2 shows the mean of the luciferase activities in the sample, and Table 5.3 shows the means of the GFP concentrations in 100  $\mu$ l of each of the samples.

**A**

Minigenome	Luc activity	Luc activity	Luc activity
pG-GFP15	127.6	145.4	95.46
pG-GFP15 $\Delta$ ORF1	97.17	99.54	93.58
pG-GFPJ3666	28.41	27.79	36.21
pG-GFPJ3666 $\Delta$ ORF1	51.08	34.6	40.4

**B**

Minigenome	Luc activity	Luc activity	Luc activity
pG-GFP15	193.7	76.66	190.9
pG-GFP15 $\Delta$ ORF1	110.8	129.9	126.6
pG-GFPJ3666	77.94	56.43	73.77
pG-GFPJ3666 $\Delta$ ORF1	39.87	30.91	38.79

Table 5.2 The luciferase activities achieved from two independent transfections (A and B) of BSR-T7/5 cells with the PVM support plasmids coding for the RNP complex and the minigenomes pG-GFP15, pG-GFP15 $\Delta$ ORF1, pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1. The luciferase assay results were obtained from 3 separate wells derived from each of the two independent experiences.

**A**

Minigenome	GFP Concentration (pg)	GFP Concentration (pg)	GFP Concentration (pg)
pG-GFP15	4431.5	5971	2415
pG-GFP15 $\Delta$ ORF1	5479.9	5000	4701.2
pG-GFPJ3666	1099.8	916.47	1262.8
pG-GFPJ3666 $\Delta$ ORF1	2142.9	1254.3	1360.5

**B**

Minigenome	GFP Concentration (pg)	GFP Concentration (pg)	GFP Concentration (pg)
pG-GFP15	4686.6	1369	4970
pG-GFP15 $\Delta$ ORF1	6674.1	7500	9062.5
pG-GFPJ3666	3389	1887.8	4329.4
pG-GFPJ3666 $\Delta$ ORF1	1819.7	1238	507

Table 5.3 The mean concentrations of GFP proteins achieved from two independent transfections (A and B) of BSR-T7/5 cells with the PVM support plasmids coding for the RNP complex and the minigenomes pG-GFP15, pG-GFP15 $\Delta$ ORF1, pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1. Each transfection experiment generated 3 separate cell sample. Each sample was tested three times to determine the level of GFP present to generate the mean values shown.

The ratio between the GFP concentration in each minigenome and its corresponding luciferase activity was calculated to normalise for transfection efficiencies. The ratios among the BSR-T7/5 cells are compared in Figure 5.15.

The data in Figure 5.15 show that in the context of a PVM minigenome removal of the first, upstream, ORF from the G gene of PVM strain 15 results in a significant increase in expression of the G-GFP fusion protein ( $p= 0.03$  with a paired t-test). In contrast, removal of the upstream ORF from the PVM strain J3666 G gene did not significantly affect the level of protein expression ( $p = 0.54$ ).

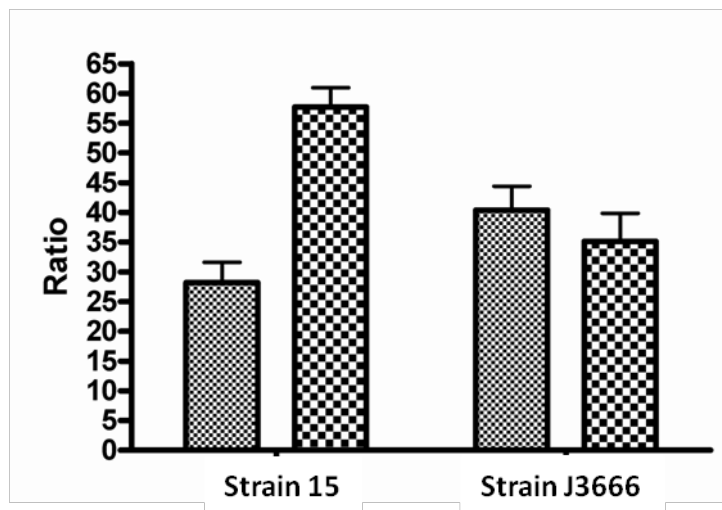




Figure 5.15 The ratio between the GFP expression and luciferase activity. The ratios calculated from BSR-T7/5 cells transfected with the minigenomes containing the sequence of PVM strain 15 are shown as “Strain 15”. Likewise, “Strain J3666” refers to those transfected with the minigenomes of the PVM J3666. The cells transfected with the non-modified minigenomes (pG-GFP15 and pG-GFPJ3666) are shown as , and the cells transfected with the minigenomes with their first ORF deleted (pG-GFP15 $\Delta$ ORF1 and pG-GFPJ3666 $\Delta$ ORF1) are shown as . Error bars show the standard deviation from the mean and were calculated after the mean of GFP concentrations (Table 5.3) normalised against the luciferase activity (Table 5.2). The test was repeated at least three times. Data were found to be reproducible.

## 5.12 Discussion

In eukaryotic systems, the 43 S ribosome subunit binds to the cap structure of the mRNA and scans until the first available AUG codon is found by the 43 S ribosomal subunit, and thereafter, following the association with specific initiation factors and the 60 S ribosomal subunit the translation starts. This model, known as the scanning model explains the effect of the upstream ORF in suppression of a main and downstream ORF (Kozak, 1978). The efficiency of initiation of translation at the first AUG is affected by the nucleotide sequence in which it sits. The presence of a so-called strong translation initiation codon upstream of an ORF with a known and functional product may change the translation efficiency of a downstream main ORF (Kozak, 1989). When the upstream start codon is in a better context than the main ORF, translation of the main ORF is often suppressed. This effect has been shown for many eukaryotic and viral genes (Child *et al.*, 1999b; Kozak, 1984; Meijer *et al.*, 2000; Peabody & Berg, 1986; Wang & Sachs, 1997).

In the work presented in this chapter, the effect of the first ORF in controlling the expression of the G protein of PVM in both strains 15 (Warwick) and J3666 was analysed. In Section 1.2.1.1 it was explained that the G gene of the strain 15 (Warwick) contains two overlapping ORFs (Thorpe & Easton, 2005). In contrast, in the strain J3666 G gene a small ORF consisting of 12 codons was shown to be present upstream of the major ORF (Randhawa *et al.*, 1995). Based on the model explained above, it was expected that the presence of these two ORFs in both of the strains might affect the translational efficiency of the main ORF. Considering the important role of the G gene in the inflammation resulting from the virus infection, it was hypothesised that the presence or absence of the first ORF in either of the strains may change the level of the expression of the G protein from the main ORF resulting in a change in the inflammatory response.

To investigate the level of the G protein expression *in vitro* and the contribution of the first ORF to the expression level of the second ORF of the G gene, a minigenome system was developed. To control for the possibility that transfection efficiency may affect the results, a dicistronic system was developed in which the first cistron expressed the luciferase gene and the second cistron expressed the G gene. The G gene contained the second ORF coding the G glycoprotein fused with GFP at the C terminal

of the G protein (referred to as G-GFP). For each of the strains two variants were synthesised differing in the presence or absence of the first ORF of the G gene. Subsequently, the minigenomes pG-GFP15, pG-GFP15 $\Delta$ ORF1, pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1 were co-transfected with plasmids responsible for the synthesis of the RNP complex of PVM into BSR-T7/5 cells. Following the transfection, the concentration of the GFP and the activity of the luciferase protein were determined and the ratio between them was calculated. The data suggested a significant difference in the ratio of G-GFP and luciferase gene expression between the pG-GFP15 and pG-GFP15 $\Delta$ ORF1 constructs, while no significant difference between the ratios of the pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1 constructs was observed.

The relative level of G-GFP expression in cells transfected with pG-GFP15 $\Delta$ ORF1 was twice that of the cells transfected with pG-GFP15. The difference between the G-GFP level was shown to be statistically significant. In cells transfected with the constructs of the strain J3666 (pG-GFPJ3666 and pG-GFPJ3666 $\Delta$ ORF1), the level of the G-GFP expression was not significantly altered.

The sequence context of the strain J3666 first ORF initiation codon is very poor (CAAAGA), which suggests that ribosomes will not initiate strongly. This may suggest that ribosomes prefer to begin translation on the initiation codon of the second major ORF which is in a better context for translation (AGUAUGG). Removal of the first ORF may therefore have little effect. For the PVM strain 15 (Warwick) the initiation codon of the first Orf is also in a poor context (CAAAGA) as is the initiation codon of the second ORF (AUAAUGU). The dramatic effect of removal of the first ORF may therefore be due simply to the major ORF becoming the first available site for translation initiation in the mRNA. Other possible reasons which must be considered could include the position of the first ORF, the length of the first ORF and the context of the main ORF versus the context of the first ORF. Alternatively, though less likely, it is possible that the effect of the first ORF in the expression of the G ORF is influenced by the cell type (Child *et al.*, 1999a). To investigate this possibility, it would be necessary to analyse the consistency of the data in other cell lines that may support the minigenome expression.

Expression of the G-GFP fusion protein was confirmed by western blot analysis. Protein from cells transfected with plasmids encoding the four G-gene constructs under the control of a bacteriophage T7 promoter was analysed by Western blot. The data

showed that all constructs generated a protein with a molecular mass expected for the G-GFP fusion proteins. The blots also indicated that there was a level of degradation of the fusion protein.

Analysis of the amino acid sequence of the PVM G protein identified the presence of a potential furin-like cleavage site (RKKR) in the G protein of PVM between amino acids 362 and 365 in the strain J3666, and amino acids 329 and 332 in the strain 15 (Warwick). This raises the possibility of the cleavage of the G-GFP protein is due to this cleavage site.

Recently, it was shown that the G protein of HRSV grown in Vero cells is truncated, possibly due to cleavage at a C-terminal arginine residue (Kwilas *et al.*, 2009). It is possible that the putative furin cleavage site in the G glycoprotein of PVM is utilised to cleave the G glycoprotein in a similar way. The effect of this furin-like cleavage site in the maturation of the G glycoprotein remains to be determined.



# **CHAPTER 6**

## **DISCUSSION**

HRSV, a member of pneumoviruses, is one of the most common causes of lower respiratory tract disease of infants, elderly and immunocompromised patients. After more than five decades since its discovery, many aspects of the molecular biology of the virus remain unclear. Most importantly, there is no safe vaccine available to prevent the disease. The only unsuccessful attempt in immunizing children with formalin inactivated viruses caused lungs sensitisation in vaccinated patients and led to a catastrophic situation of enhancement of disease when the patients encountered with the wildtype virus later in their life. Currently, passive immunization with human immunoglobulins or a recombinant humanized monoclonal antibody (palivizumab) is the only way of protecting high risk patients against the virus.

A major barrier in HRSV research is the lack of a suitable model to study the interplay between the host and virus. Chimpanzees (*Pan troglodytes*) are the only highly permissive animal models in which to study HRSV *in vivo* (Belshe *et al.*, 1977). There are several limitations in using chimpanzee as a research model including high cost of maintenance, their endangered status in the wild and ethical issues due to their genetic similarity to humans (Knight, 2008).

Inbred mice have been extensively used to study the pathogenesis of HRSV. However, the semipermissive nature of mice to HRSV infection does not fully reflect all features of HRSV infection in humans. In particular, mice are not very permissive to primary isolates of HRSV, and HRSV infection in mice results in little or no inflammatory response. Another limitation of the HRSV – mouse model is the requirement of large inocula to infect mice to produce detectable signs of clinical disease (Rosenberg *et al.*, 2005).

Studying BRSV or PVM in their natural hosts have been considered as alternatives in studying the pathogenesis of HRSV. The genome of BRSV strain A51908 is 15149 nucleotides in length and shares 73% of identity with HRSV and the genomic organisation of both HRSV and BRSV is similar (Collins *et al.*, 1996b). Extensive cross reactivity between HRSV and BRSV proteins was reported. The genetic similarity between HRSV and BRSV makes BRSV an attractive candidate for the study of the molecular biology and pathogenesis of pneumoviruses. However, genetic diversity among cattle, the cost of maintenance and heterogenicity of the response of outbred cattle to BRSV infection are considered major limitations in using BRSV in cattle as a general model for pneumovirus infection *in vivo* models (Collins *et al.*, 1996b).

Using PVM in its natural murine host is considered as an alternative for studying the pathogenesis of pneumoviruses. PVM shares 52% nucleotide sequence similarity with HRSV, and they share a similar genetic organisation except for the overlap of the M2 and L genes in HRSV which is absent in PVM. The pattern of disease of PVM in mice resembles the disease of HRSV in humans. Efficient replication of PVM in the mouse respiratory system is accompanied by an inflammatory response, mucous production and airway obstruction (Easton *et al.*, 2004). The replication of PVM in lung epithelium leads to inflammatory cytokine production and granulocyte influx which is similar to the situation with HRSV. This similarity between the pathogenesis of HRSV and PVM in their hosts suggests that PVM is a potential powerful tool to study the pathogenesis of pneumoviruses in a natural host.

In the research presented in this thesis, the pathogenesis of PVM was studied, and as the first step, the effect of consecutive tissue culture passages on the pathogenicity and genotype of PVM was analysed (CHAPTER 3).

Using the pathogenic strain J3666 it was shown that passage in tissue culture led to a loss of pathogenicity in mice. The results indicated that passages 6 and 7 were attenuated, but passage 5 was able to cause disease in mice. To investigate the molecular basis of the difference, the sequence of the virus genomes was determined. The NS1 and NS2 genes and most of the L gene were excluded from the study. The L, NS1 and NS2 genes are highly conserved among the strains of PVM and no genetic diversity between them has been shown. In particular, the NS1 and NS2 genes of the pathogenic strain J3666 and the non-pathogenic strain 15 were identical (Thorpe & Easton, 2005). Therefore, it was thought that the genetic diversity between the passages that are closely related to each other would be minimal.

Surprisingly, no differences in the genome sequences were found between passages 5, 6 and 7. However, the sequence of passages 5, 6 and 7 were different from the published sequence of PVM strain J3666 (Accession number: NC\_006579). The sequence differences between the published data and that obtained here were mainly located in the SH and G genes (Table 3.1). Briefly, in passages 5, 6 and 7, a C → U point mutation at nucleotide residue 269 caused a His → Tyr substitution and C → U mutation at SH gene position 283 introduced a stop codon resulting in shortening of the open reading frame. A comparison between the amino acid sequence of the SH protein encoded by the viruses from passages 5, 6 and 7 with the amino acid sequence of the SH protein of PVM strains J3666 and 15 (Warwick) revealed that the sequence of the

passages 5, 6 and 7 was similar to the sequence of the strain J3666 while its length was same as that of PVM strains 15 (Warwick) and 15 (ATCC). In the G gene, a U to A mutation at position 65 changed the stop codon of the first ORF of the G gene (Section 1.2.1.1.1) to a lysine residue resulting in an extension of the ORF and as a result production of a G glycoprotein with a larger N-terminal. The other changes observed in the G gene were: C(68)→U in the noncoding sequence between the two ORFs, A(104)→G changing a serine residue to a glycine residue, G(165)→U changing a glycine residue to a valine residue, G(236)→A changing a glycine residue to an arginine residue, and U(1121)→A changing a serine residue to a threonine residue. Two changes were observed in the nucleic acid sequence of the F gene U(721)→G changing an isoleucine residue to a glutamine residue and U(992)→C changing a valine residue to an alanine residue.

To investigate further the sequences present in the virus stocks, 14 independent clones from two separate RT-PCRs were prepared. These clones contained sequences from the SH and G genes of the PVM strain J3666 virus stock prepared after 5 passages in tissue culture. The sequence analysis showed that only 2 of the 14 clones contained a sequence of the G gene identical to that described by Randhawa et al. (1995). The remaining 12 clones contained a G gene in which the mutation in the stop codon of the first ORF made a single ORF extending from the first ORF to the second ORF. These data strongly suggest that the virus stocks used contain a mixed population of virus sequences. It is therefore not possible to determine which sequence(s) are responsible for the pathogenic phenotype. However, a major difference in the sequences lay in the organisation of the G gene encoding the G glycoprotein. The differences in organisation may play a role in G gene expression and this was investigated further.

Collins *et al.* (1995) described a reverse genetics system for HRSV to study the replication requirements of HRSV *in vitro*. Soon after its first description, the reverse genetics approach was used to study the effect of gene deletion or gene modification in the pathogenesis of HRSV. It was identified that a point mutation in the gene start of the M2 gene is responsible for attenuation and temperature sensitive phenotype in vaccine candidate *cpts248/404* (Whitehead *et al.*, 1998a; Whitehead *et al.*, 1998b). Using the reverse genetics approach, it was shown that the SH deleted recombinant viruses were attenuated in mice (Bukreyev *et al.*, 1997), showing a moderate reduction in growth in the upper and lower respiratory tract of chimpanzees. In chimpanzees the  $\Delta$ SH produced less disease than the wildtype virus (Whitehead *et al.*, 1999). NS1 and NS2

deleted HRSV and also HRSV  $\Delta$ NS2 are highly attenuated in the lower respiratory tract of chimpanzees (Whitehead *et al.*, 1999). The presence of other mutations including a missense mutation changing the Tyrosine→Alanine at amino acid 1321 of L gene caused attenuation in the virus (Whitehead *et al.*, 1999; Whitehead *et al.*, 1998b). Teng and Collins (1999) showed that the NS2 gene is not essential but is important for the efficient replication of HRSV, and the recombinant viruses with tandem stop codons in their NS2 gene produce revertant mutants to restore the NS2 gene function (Teng & Collins, 1999).

Following the advances achieved in the study of HRSV, a reverse genetics system was employed to study the virulence of BRSV (Buchholz *et al.*, 1999; Schlender *et al.*, 2000). Using a recombinant vaccinia virus, Schlender *et al.* (2000) demonstrated that both NS1 and NS2 act together to suppress host interferon response. The function of NS1 and NS2 genes of PVM in antagonising the interferon  $\alpha$  and  $\beta$  and interferon  $\lambda$  (IL-28) have also shown (Heinze *et al.*, 2011). This suggests that both NS1 and NS2 genes in pneumoviruses may act as interferon antagonists to modulate the innate immune response (Heinze *et al.*, 2011; Schlender *et al.*, 2000; Whitehead *et al.*, 1999).

More recently, a reverse genetics system was used to investigate the effect of the G gene of PVM in pathogenesis. It was shown that deletion of the G gene, or deletion of the cytoplasmic tail of the G protein resulted in virus attenuation in mice (Krempl *et al.*, 2007).

The common feature in all of these studies was that the genetic backbone of the recombinant viruses was derived from a pathogenic virus. The studies therefore, employed a principle of establishing which mutations resulted in a loss of pathogenicity. While useful, this approach suffers from the problem that the role(s) of specific genes in reducing pathogenicity is not directly addressed. In the work described here the genetic backbone for the reverse genetics studies was that of PVM strain 15 (Warwick) which is not pathogenic in mice even when large doses are administered (Cook *et al.*, 1998). Thus any mutation introduced into the virus genome are assessed by the appearance of pathogenicity and resulted in the gain of function. Having established a robust reverse genetics system based on work by Dibben (2006) and Thorpe and Easton (2005), it was used to specifically investigate the role of the G protein of PVM and its expression in pathogenesis. The structures of the G gene and the G glycoprotein of HRSV and PVM were reviewed briefly in Section 1.2.1.1.1. While the absence of the G protein *in vitro* does not have any significant effect on HRSV replication, its role in the virus

immunogenicity and pathogenicity *in vivo* has been a major topic in the pneumoviruses molecular biology (Teng *et al.*, 2001). It has been reported that a recombinant HRSV in which the G gene was deleted was highly attenuated in the respiratory tract of mice (Teng *et al.*, 2001). In the same study, it was shown that HRSV containing only the secreted form of the G glycoprotein is less attenuated (Teng *et al.*, 2001). These data suggest an important role for the G glycoprotein in the virus replication *in vivo*.

The G gene in both HRSV and PVM contain two ORFs (Wertz *et al.*, 1985). In PVM strain J3666 the two ORFs are placed close to each other and in PVM strain 15 (Warwick) the ORFs are overlapping each other (Randhawa *et al.*, 1995; Thorpe & Easton, 2005). The presence of an upstream ORF has been shown to modulate the gene expression of a downstream ORF for many eukaryotic genes (Child *et al.*, 1999b; Kozak, 1984; Meijer *et al.*, 2000). The possible effect of the presence or absence of the first ORF of the G gene in PVM strains 15 (Warwick) and J3666 in the expression of the downstream ORF of the G gene was analysed. It was shown that in recombinant viruses with a common genetic backbone derived from the non-pathogenic strain 15 (Warwick), the gene organisation of the G gene in PVM strain J3666 can affect the pathogenicity of the virus. The recombinant viruses in which the G gene was based on that from PVM strain 15 (Warwick) observed no signs of pathogenesis even when 5000 pfu was inoculated into mice. Surprisingly, removal of the first, overlapping ORF of the PVM strain 15 (Warwick) G gene resulted in transient weight loss in mice inoculated with 5000 pfu. The transient weight loss is indicative of mild pathogenesis. The recombinant virus containing the PVM strain J3666 G gene showed clear signs of infection and weight loss in mice infected with the highest dose of virus. When the first ORF in this construct was removed the virus became non-pathogenic. These results indicate that the presence of the first ORF upstream of the main ORF encoding the PVM G glycoprotein has an effect on pathogenesis.

To examine whether amending the organisation of the G gene affects the expression levels of the G protein, a series of minigenome replicons was designed. The dicistronic minigenome replicons, known as pG-GFP minigenomes, were designed to express the luciferase gene from their first gene and the G-GFP fusion protein from their second cistron. Both cistrons were designed to be controlled by the viral transcriptional system, and to minimize the possible effects of the SH-G intergenic region and the G gene start sequence in the expression of the G gene to contain the same intergenic region and G gene start of the G gene in PVM. The analysis of the expression

levels of the G-GFP fusion protein indicated that deletion of the first ORF in the constructs resembling the organisation of the G gene in the strain 15 (Warwick) is associated with a significant increase in the expression of the G protein. This is consistent with the pattern frequently seen in eukaryotic genes. In contrast, deletion of the first ORF in the strain J3666 G gene was associated with little or no effect. The level of expression of the G protein may be affected by the sequences surrounding the AUG initiation codon as discussed in Chapter 5.

The central role of the G protein in HRSV virulence and in modulating the immune system was reviewed in Section 1.3.1.2. The immunomodulatory function of the G glycoprotein has been associated with the presence of a domain which mimics the fractalkine CX3C (1.3.1.2) (Tripp *et al.*, 2001). Fractalkine (also known as CX3CL1) was reported as a new class of chemoattractants found in various non-lymphoid tissues, such as the small intestine, colon, heart, brain, lung, kidney and pancreas, of healthy individuals (Bazan *et al.*, 1997). It consists of a 373 amino acid chain in which the chemokine domain is located at the end of a mucin-like stalk and a transmembrane domain through which it is anchored to the plasma membrane (Bazan *et al.*, 1997). The receptor for fractalkine, CX3CR1, is found in high levels on NK cells and to a lesser extent on monocytes and T cells (Imai *et al.*, 1997; Umehara *et al.*, 2001). The triggering of the receptor results in release of perforin and granzymes from NK cells in a cytolytic response (Umehara *et al.*, 2001; Yoneda *et al.*, 2000). It has been shown that exposure of NK cells to the soluble form of fractalkine (s-fractalkine) increases the cytotoxicity of the NK cell activities (Yoneda *et al.*, 2000) and granular exocytosis of NK cells responds in a dose dependent manner (Yoneda *et al.*, 2000). The chemokine mimicry function of the G glycoprotein of HRSV might contribute in attracting NK cells, adhesion to the cellular membranes of infected epithelial cells expressing the G glycoprotein, and result in the damages of the lungs, and the levels of G protein expressed by the virus may affect the intensity of this response.

The effect of HRSV G and/or SH proteins in reducing the number of pulmonary NK cells and the expression of IFN- $\gamma$  was shown by Tripp *et al.* (1999). The CX3C region of the G glycoprotein was reported to be associated with a reduction in the number of CX3CL1<sup>+</sup>, HRSV-specific and IFN- $\gamma$  expressing lymphocytes, while in the same population the number of IL-4 expressing cells increased (Harcourt *et al.*, 2006). These data emphasize the chemokine mimicry role of the G glycoprotein in controlling and affecting cytokine and chemokine production in the infection sites.

There are numerous examples of using animal models to study the pathogenesis of human viruses. Such models include those in which the virus replicates in its natural host. The models have been extensively used to explore different features in molecular biology of pathogenesis and the mechanisms involved in host defence. Woodchuck hepatitis virus is one such system which has been employed to elucidate the molecular biology and pathogenesis of hepatitis B virus (HBV) (Shafritz & Lieberman, 1984; Snyder *et al.*, 1982; Summers *et al.*, 1978). Studying the interaction between PVM and the mouse and extrapolating and evaluating the data to a larger model (BRSV and bovine) may facilitate the discovery of the factors influencing the pathogenesis of pneumoviruses. Availability of a reverse genetics system for PVM could facilitate studying of virus pathogenesis. A clear description of the pathogenesis of PVM could be extrapolated to explain the pathogenesis of HRSV.

The data presented here have shown that the PVM system is available for manipulation to explore the mechanisms of pathogenesis of pneumoviruses. Further studies are likely to identify additional factors which play a role in the disease process.



## **REFERENCES**

**Ahmadian, G., Chambers, P. & Easton, A. J. (1999).** Detection and characterization of proteins encoded by the second ORF of the M2 gene of pneumoviruses. *J Gen Virol* **80** ( Pt 8), 2011-2016.

**Ahmadian, G., Randhawa, J. S. & Easton, A. J. (2000).** Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism. *EMBO J* **19**, 2681-2689.

**Amanatidou, V., Sourvinos, G., Apostolakis, S., Tsilimigaki, A. & Spandidos, D. A. (2006).** T280M variation of the CX3C receptor gene is associated with increased risk for severe respiratory syncytial virus bronchiolitis. *Pediatr Infect Dis J* **25**, 410-414.

**Anderson, K., King, A. M., Lerch, R. A. & Wertz, G. W. (1992).** Polylactosaminoglycan modification of the respiratory syncytial virus small hydrophobic (SH) protein: a conserved feature among human and bovine respiratory syncytial viruses. *Virology* **191**, 417-430.

**Anderson, L. J., Hierholzer, J. C., Tsou, C., Hendry, R. M., Fernie, B. F., Stone, Y. & McIntosh, K. (1985).** Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J Infect Dis* **151**, 626-633.

**Anh, D. B., Faisca, P. & Desmecht, D. J. (2006).** Differential resistance/susceptibility patterns to pneumovirus infection among inbred mouse strains. *Am J Physiol Lung Cell Mol Physiol* **291**, L426-435.

**Asenjo, A., Calvo, E. & Villanueva, N. (2006).** Phosphorylation of human respiratory syncytial virus P protein at threonine 108 controls its interaction with the M2-1 protein in the viral RNA polymerase complex. *J Gen Virol* **87**, 3637-3642.

**Asenjo, A., Gonzalez-Armas, J. C. & Villanueva, N. (2008).** Phosphorylation of human respiratory syncytial virus P protein at serine 54 regulates viral uncoating. *Virology* **380**, 26-33.

**Awomoyi, A. A., Rallabhandi, P., Pollin, T. I., Lorenz, E., Sztejn, M. B., Boukhvalova, M. S., Hemming, V. G., Blanco, J. C. & Vogel, S. N. (2007).** Association of TLR4 polymorphisms with symptomatic respiratory syncytial virus infection in high-risk infants and young children. *J Immunol* **179**, 3171-3177.

**Bachi, T. & Howe, C. (1973).** Morphogenesis and ultrastructure of respiratory syncytial virus. *J Virol* **12**, 1173-1180.

**Barik, S. (1992).** Transcription of human respiratory syncytial virus genome RNA in vitro: requirement of cellular factor(s). *J Virol* **66**, 6813-6818.

**Barik, S. (1993).** The structure of the 5' terminal cap of the respiratory syncytial virus mRNA. *J Gen Virol* **74** ( Pt 3), 485-490.

**Barik, S., McLean, T. & Dupuy, L. C. (1995).** Phosphorylation of Ser232 directly regulates the transcriptional activity of the P protein of human respiratory syncytial virus: phosphorylation of Ser237 may play an accessory role. *Virology* **213**, 405-412.

**Barr, J., Chambers, P., Harriott, P., Pringle, C. R. & Easton, A. J. (1994).** Sequence of the phosphoprotein gene of pneumonia virus of mice: expression of multiple proteins from two overlapping reading frames. *J Virol* **68**, 5330-5334.

**Barr, J., Chambers, P., Pringle, C. R. & Easton, A. J. (1991).** Sequence of the major nucleocapsid protein gene of pneumonia virus of mice: sequence comparisons suggest structural homology between nucleocapsid proteins of pneumoviruses, paramyxoviruses, rhabdoviruses and filoviruses. *J Gen Virol* **72** ( Pt 3), 677-685.

**Bazan, J. F., Bacon, K. B., Hardiman, G., Wang, W., Soo, K., Rossi, D., Greaves, D. R., Zlotnik, A. & Schall, T. J. (1997).** A new class of membrane-bound chemokine with a CX3C motif. *Nature* **385**, 640-644.

**Belshe, R. B., Richardson, L. S., London, W. T., Sly, D. L., Lorfeld, J. H., Camargo, E., Prevar, D. A. & Chanock, R. M. (1977).** Experimental respiratory syncytial virus infection of four species of primates. *Journal of medical virology* **1**, 157-162.

**Bermingham, A. & Collins, P. L. (1999).** The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc Natl Acad Sci U S A* **96**, 11259-11264.

**Berthiaume, L., Joncas, J. & Pavilanis, V. (1974).** Comparative structure, morphogenesis and biological characteristics of the respiratory syncytial (RS) virus and the pneumonia virus of mice (PVM). *Arch Gesamte Virusforsch* **45**, 39-51.

**Biacchesi, S., Murphy, B. R., Collins, P. L. & Buchholz, U. J. (2007).** Frequent frameshift and point mutations in the SH gene of human metapneumovirus passaged in vitro. *J Virol* **81**, 6057-6067.

**Bonville, C. A., Bennett, N. J., Koehnlein, M., Haines, D. M., Ellis, J. A., DelVecchio, A. M., Rosenberg, H. F. & Domachowske, J. B. (2006).** Respiratory dysfunction and proinflammatory chemokines in the pneumonia virus of mice (PVM) model of viral bronchiolitis. *Virology* **349**, 87-95.

**Bossert, B. & Conzelmann, K. K. (2002).** Respiratory syncytial virus (RSV) nonstructural (NS) proteins as host range determinants: a chimeric bovine RSV with NS genes from human RSV is attenuated in interferon-competent bovine cells. *J Virol* **76**, 4287-4293.

**Bost, K. L. & Pascual, D. W. (1992).** Substance P: a late-acting B lymphocyte differentiation cofactor. *Am J Physiol* **262**, C537-545.

**Buchholz, U. J., Finke, S. & Conzelmann, K. K. (1999).** Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. *J Virol* **73**, 251-259.

**Buchholz, U. J., Ward, J. M., Lamirande, E. W., Heinze, B., Krempl, C. D. & Collins, P. L. (2009).** Deletion of nonstructural proteins NS1 and NS2 from pneumonia virus of mice attenuates viral replication and reduces pulmonary cytokine expression and disease. *J Virol* **83**, 1969-1980.

**Bukreyev, A., Whitehead, S. S., Murphy, B. R. & Collins, P. L. (1997).** Recombinant respiratory syncytial virus from which the entire SH gene has been deleted grows efficiently in cell culture and exhibits site-specific attenuation in the respiratory tract of the mouse. *J Virol* **71**, 8973-8982.

**Byrd, L. G. & Prince, G. A. (1997).** Animal models of respiratory syncytial virus infection. *Clin Infect Dis* **25**, 1363-1368.

**Cannon, M. J., Stott, E. J., Taylor, G. & Askonas, B. A. (1987).** Clearance of persistent respiratory syncytial virus infections in immunodeficient mice following transfer of primed T cells. *Immunology* **62**, 133-138.

**Caravokyri, C. & Pringle, C. R. (1992).** Effect of changes in the nucleotide sequence of the P gene of respiratory syncytial virus on the electrophoretic mobility of the P protein. *Virus Genes* **6**, 53-62.

**Cartee, T. L. & Wertz, G. W. (2001).** Respiratory syncytial virus M2-1 protein requires phosphorylation for efficient function and binds viral RNA during infection. *J Virol* **75**, 12188-12197.

**Carter, S. D., Dent, K. C., Atkins, E., Foster, T. L., Verow, M., Gorny, P., Harris, M., Hiscox, J. A., Ranson, N. A., Griffin, S. & Barr, J. N. (2010).** Direct visualization of the small hydrophobic protein of human respiratory syncytial virus reveals the structural basis for membrane permeability. *FEBS Lett* **584**, 2786-2790.

**Cash, P., Preston, C. M. & Pringle, C. R. (1979).** Characterisation of murine pneumonia virus proteins. *Virology* **96**, 442-452.

**Castagne, N., Barbier, A., Bernard, J., Rezaei, H., Huet, J. C., Henry, C., Da Costa, B. & Eleouet, J. F. (2004).** Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. *J Gen Virol* **85**, 1643-1653.

**Chambers, P., Matthews, D. A., Pringle, C. R. & Easton, A. J. (1991).** The nucleotide sequences of intergenic regions between nine genes of pneumonia virus of mice establish the physical order of these genes in the viral genome. *Virus Res* **18**, 263-270.

**Chanock, R., Roizman, B. & Myers, R. (1957).** Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization. *Am J Hyg* **66**, 281-290.

**Chen, M. D., Vazquez, M., Buonocore, L. & Kahn, J. S. (2000).** Conservation of the respiratory syncytial virus SH gene. *J Infect Dis* **182**, 1228-1233.

**Cheng, X., Park, H., Zhou, H. & Jin, H. (2005).** Overexpression of the M2-2 protein of respiratory syncytial virus inhibits viral replication. *J Virol* **79**, 13943-13952.

**Child, S. J., Miller, M. K. & Geballe, A. P. (1999a).** Cell type-dependent and -independent control of HER-2/neu translation. *Int J Biochem Cell Biol* **31**, 201-213.

**Child, S. J., Miller, M. K. & Geballe, A. P. (1999b).** Translational control by an upstream open reading frame in the HER-2/neu transcript. *J Biol Chem* **274**, 24335-24341.

**Churchill, A. E., Chubb, R. C. & Baxendale, W. (1969).** The attenuation, with loss of oncogenicity, of the herpes-type virus of Marek's disease (strain HPRS-16) on passage in cell culture. *J Gen Virol* **4**, 557-564.

**Claassen, E. A., van der Kant, P. A., Rychnavska, Z. S., van Bleek, G. M., Easton, A. J. & van der Most, R. G. (2005).** Activation and inactivation of antiviral CD8 T cell responses during murine pneumovirus infection. *J Immunol* **175**, 6597-6604.

**Cohen, J. I., Rosenblum, B., Feinstone, S. M., Ticehurst, J. & Purcell, R. H. (1989).** Attenuation and cell culture adaptation of hepatitis A virus (HAV): a genetic analysis with HAV cDNA. *J Virol* **63**, 5364-5370.

**Collins, P. L., Chanock, R. M. & Murphy, B. R. (2001).** Respiratory Syncytial Virus. In *Fields virology*, 4th edn, pp. 1443-1485. Edited by B. N. Fields, D. M. Knipe, P. M. Howley & D. E. Griffin. Philadelphia: Lippincott Williams & Wilkins.

**Collins, P. L. & Graham, B. S. (2008).** Viral and host factors in human respiratory syncytial virus pathogenesis. *J Virol* **82**, 2040-2055.

**Collins, P. L., Hill, M. G., Camargo, E., Grosfeld, H., Chanock, R. M. & Murphy, B. R. (1995).** Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development. *Proc Natl Acad Sci U S A* **92**, 11563-11567.

**Collins, P. L., Hill, M. G., Cristina, J. & Grosfeld, H. (1996a).** Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc Natl Acad Sci U S A* **93**, 81-85.

**Collins, P. L., McIntosh, K. & Chanock, R. (1996b).** Respiratory syncytial virus. In *Fields virology*, 3rd edn, pp. 1313-1351. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott-Raven Publishers.

**Collins, P. L., Mink, M. A. & Stec, D. S. (1991).** Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. *Proc Natl Acad Sci U S A* **88**, 9663-9667.

**Collins, P. L. & Mottet, G. (1992).** Oligomerization and post-translational processing of glycoprotein G of human respiratory syncytial virus: altered O-glycosylation in the presence of brefeldin A. *J Gen Virol* **73** ( Pt 4), 849-863.

**Collins, P. L. & Mottet, G. (1993).** Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory syncytial virus. *J Gen Virol* **74** ( Pt 7), 1445-1450.

**Collins, P. L., Olmsted, R. A., Spriggs, M. K., Johnson, P. R. & Buckler-White, A. J. (1987).** Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. *Proc Natl Acad Sci U S A* **84**, 5134-5138.

**Collins, P. L. & Wertz, G. W. (1985a).** The envelope-associated 22K protein of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polytranscript. *J Virol* **54**, 65-71.

**Collins, P. L. & Wertz, G. W. (1985b).** Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus. *Virology* **143**, 442-451.

**Compans, R. W., Harter, D. H. & Choppin, P. W. (1967).** Studies on pneumonia virus of mice (PVM) in cell culture. II. Structure and morphogenesis of the virus particle. *J Exp Med* **126**, 267-276.

**Cook, P. M., Eglin, R. P. & Easton, A. J. (1998).** Pathogenesis of pneumovirus infections in mice: detection of pneumonia virus of mice and human respiratory syncytial virus mRNA in lungs of infected mice by in situ hybridization. *J Gen Virol* **79** ( Pt 10), 2411-2417.

**Cowton, V. M. & Fearn, R. (2005).** Evidence that the respiratory syncytial virus polymerase is recruited to nucleotides 1 to 11 at the 3' end of the nucleocapsid and can scan to access internal signals. *J Virol* **79**, 11311-11322.

**Cowton, V. M., McGivern, D. R. & Fearn, R. (2006).** Unravelling the complexities of respiratory syncytial virus RNA synthesis. *J Gen Virol* **87**, 1805-1821.

**Cuesta, I., Geng, X., Asenjo, A. & Villanueva, N. (2000).** Structural phosphoprotein M2-1 of the human respiratory syncytial virus is an RNA binding protein. *J Virol* **74**, 9858-9867.

**Dardiri, A. H. (1969).** Attenuation of duck plague virus and its propagation in cell culture. *Arch Gesamte Virusforsch* **27**, 55-64.

**DeVincenzo, J. P. (2005).** Factors predicting childhood respiratory syncytial virus severity: what they indicate about pathogenesis. *Pediatr Infect Dis J* **24**, S177-183, discussion S182.

**Dibben, O. & Easton, A. J. (2007).** Mutational analysis of the gene start sequences of pneumonia virus of mice. *Virus Res* **130**, 303-309.

**Dibben, O., Thorpe, L. C. & Easton, A. J. (2008).** Roles of the PVM M2-1, M2-2 and P gene ORF 2 (P-2) proteins in viral replication. *Virus Res* **131**, 47-53.

**Dibben, O. E. (2006).** Development of a reverse genetics system for Pneumonia Virus of Mice. In *Department of Biological Sciences*: University of Warwick. PhD thesis.

**Dickens, L. E., Collins, P. L. & Wertz, G. W. (1984).** Transcriptional mapping of human respiratory syncytial virus. *J Virol* **52**, 364-369.

**Doggett, J. E., Taylor-Robinson, D. & Gallop, R. G. (1968).** A study of an inhibitor in bovine serum active against respiratory syncytial virus. *Arch Gesamte Virusforsch* **23**, 126-137.

**Domachowske, J. B., Bonville, C. A., Dyer, K. D., Easton, A. J. & Rosenberg, H. F. (2000a).** Pulmonary eosinophilia and production of MIP-1alpha are prominent responses to infection with pneumonia virus of mice. *Cell Immunol* **200**, 98-104.

**Domachowske, J. B., Bonville, C. A., Gao, J. L., Murphy, P. M., Easton, A. J. & Rosenberg, H. F. (2000b).** The chemokine macrophage-inflammatory protein-1 alpha and its receptor CCR1 control pulmonary inflammation and antiviral host defense in paramyxovirus infection. *J Immunol* **165**, 2677-2682.

**Domachowske, J. B., Bonville, C. A. & Rosenberg, H. F. (2004).** Animal models for studying respiratory syncytial virus infection and its long term effects on lung function. *Pediatr Infect Dis J* **23**, S228-234.

**Doyle, M. & Holland, J. J. (1973).** Prophylaxis and immunization in mice by use of virus-free defective T particles to protect against intracerebral infection by vesicular stomatitis virus. *Proc Natl Acad Sci U S A* **70**, 2105-2108.

**Easton, A. J., Domachowske, J. B. & Rosenberg, H. F. (2004).** Animal pneumoviruses: molecular genetics and pathogenesis. *Clin Microbiol Rev* **17**, 390-412.

**Easton, A. J., Scott, P. D., Edworthy, N. L., Meng, B., Marriott, A. C. & Dimmock, N. J. (2011).** A novel broad-spectrum treatment for respiratory virus infections: influenza-based defective interfering virus provides protection against pneumovirus infection in vivo. *Vaccine* **29**, 2777-2784.

**Fearns, R., Peeples, M. E. & Collins, P. L. (2002).** Mapping the transcription and replication promoters of respiratory syncytial virus. *J Virol* **76**, 1663-1672.

**Feldman, S. A., Hendry, R. M. & Beeler, J. A. (1999).** Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J Virol* **73**, 6610-6617.

**Fuentes, S., Tran, K. C., Luthra, P., Teng, M. N. & He, B. (2007).** Function of the respiratory syncytial virus small hydrophobic protein. *J Virol* **81**, 8361-8366.

**Gallaspy, S. E., Coward, J. E. & Howe, C. (1978).** Persistent infection of BHK-21 cells with pneumonia virus of mice. *J Virol* **26**, 110-114.



**Gan, S. W., Ng, L., Lin, X., Gong, X. & Torres, J. (2008).** Structure and ion channel activity of the human respiratory syncytial virus (hRSV) small hydrophobic protein transmembrane domain. *Protein Sci* **17**, 813-820.

**Garcia-Barreno, B., Delgado, T. & Melero, J. A. (1996).** Identification of protein regions involved in the interaction of human respiratory syncytial virus phosphoprotein and nucleoprotein: significance for nucleocapsid assembly and formation of cytoplasmic inclusions. *J Virol* **70**, 801-808.

**Garcia, J., Garcia-Barreno, B., Vivo, A. & Melero, J. A. (1993).** Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology* **195**, 243-247.

**Ghildyal, R., Li, D., Peroulis, I., Shields, B., Bardin, P. G., Teng, M. N., Collins, P. L., Meanger, J. & Mills, J. (2005).** Interaction between the respiratory syncytial virus G glycoprotein cytoplasmic domain and the matrix protein. *J Gen Virol* **86**, 1879-1884.

**Glazier, K., Raghov, R. & Kingsbury, D. W. (1977).** Regulation of Sendai virus transcription: evidence for a single promoter in vivo. *J Virol* **21**, 863-871.

**Gould, P. S. & Easton, A. J. (2007).** Coupled translation of the second open reading frame of M2 mRNA is sequence dependent and differs significantly within the subfamily Pneumovirinae. *J Virol* **81**, 8488-8496.

**Groneberg, D. A., Harrison, S., Dinh, Q. T., Geppetti, P. & Fischer, A. (2006).** Tachykinins in the respiratory tract. *Curr Drug Targets* **7**, 1005-1010.

**Gubbay, O., Curran, J. & Kolakofsky, D. (2001).** Sendai virus genome synthesis and assembly are coupled: a possible mechanism to promote viral RNA polymerase processivity. *J Gen Virol* **82**, 2895-2903.

**Harcourt, J., Alvarez, R., Jones, L. P., Henderson, C., Anderson, L. J. & Tripp, R. A. (2006).** Respiratory syncytial virus G protein and G protein CX3C motif adversely affect CX3CR1+ T cell responses. *J Immunol* **176**, 1600-1608.

**Hardy, R. W. & Wertz, G. W. (1998).** The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J Virol* **72**, 520-526.

**Hardy, R. W. & Wertz, G. W. (2000).** The Cys(3)-His(1) motif of the respiratory syncytial virus M2-1 protein is essential for protein function. *J Virol* **74**, 5880-5885.

**Harmon, S. B., Megaw, A. G. & Wertz, G. W. (2001).** RNA sequences involved in transcriptional termination of respiratory syncytial virus. *J Virol* **75**, 36-44.

**Hearn, H. J., Jr., Chappell, W. A., Demchak, P. & Kominik, J. W. (1966).** Attenuation of aerosolized yellow fever virus after passage in cell culture. *Bacteriol Rev* **30**, 615-623.

**Heinze, B., Frey, S., Mordstein, M., Schmitt-Graff, A., Ehl, S., Buchholz, U. J., Collins, P. L., Staheli, P. & Krempl, C. D. (2011).** Both Nonstructural Proteins 1 and 2 of Pneumonia Virus of Mice are Inhibitors of the Interferon Type I and III Response In Vivo. *J Virol* **85**, 4071-4084.

**Hendricks, D. A., McIntosh, K. & Patterson, J. L. (1988).** Further characterization of the soluble form of the G glycoprotein of respiratory syncytial virus. *J Virol* **62**, 2228-2233.

**Hietala, J., Nyman, M. J., Eskola, O., Laakso, A., Gronroos, T., Oikonen, V., Bergman, J., Haaparanta, M., Forsback, S., Marjamaki, P., Lehtikoinen, P., Goldberg, M., Burns, D., Hamill, T., Eng, W. S., Coimbra, A., Hargreaves, R. & Solin, O. (2005).** Visualization and quantification of neurokinin-1 (NK1) receptors in the human brain. *Mol Imaging Biol* **7**, 262-272.

**Horsfall, F. L. & Curnen, E. C. (1946).** Studies on Pneumonia Virus of Mice (Pvm) : Ii. Immunological Evidence of Latent Infection with the Virus in Numerous Mammalian Species. *J Exp Med* **83**, 43-64.

**Horsfall, F. L. & Hahn, R. G. (1940).** A Latent Virus in Normal Mice Capable of Producing Pneumonia in Its Natural Host. *J Exp Med* **71**, 391-408.

**Huang, Y. T., Collins, P. L. & Wertz, G. W. (1985).** Characterization of the 10 proteins of human respiratory syncytial virus: identification of a fourth envelope-associated protein. *Virus Res* **2**, 157-173.

**Huang, Y. T. & Wertz, G. W. (1982).** The genome of respiratory syncytial virus is a negative-stranded RNA that codes for at least seven mRNA species. *J Virol* **43**, 150-157.

**Hull, J., Rowlands, K., Lockhart, E., Moore, C., Sharland, M. & Kwiatkowski, D. (2003).** Variants of the chemokine receptor CCR5 are associated with severe bronchiolitis caused by respiratory syncytial virus. *J Infect Dis* **188**, 904-907.

ICTV.Virus Taxonomy: 2009 Release. Retrieved from: <http://www.ictvonline.org/virusTaxonomy.asp?version=2009>.

**Ikura, M. (1996).** Calcium binding and conformational response in EF-hand proteins. *Trends Biochem Sci* **21**, 14-17.

**Imai, T., Hieshima, K., Haskell, C., Baba, M., Nagira, M., Nishimura, M., Kakizaki, M., Takagi, S., Nomiya, H., Schall, T. J. & Yoshie, O. (1997).** Identification and molecular characterization of fractalkine receptor CX3CR1, which mediates both leukocyte migration and adhesion. *Cell* **91**, 521-530.

**Inoue, H., Nojima, H. & Okayama, H. (1990).** High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-28.

**Jin, H., Cheng, X., Zhou, H. Z., Li, S. & Seddiqui, A. (2000a).** Respiratory syncytial virus that lacks open reading frame 2 of the M2 gene (M2-2) has altered growth characteristics and is attenuated in rodents. *J Virol* **74**, 74-82.

**Jin, H., Clarke, D., Zhou, H. Z., Cheng, X., Coelingh, K., Bryant, M. & Li, S. (1998).** Recombinant human respiratory syncytial virus (RSV) from cDNA and construction of subgroup A and B chimeric RSV. *Virology* **251**, 206-214.

**Jin, H., Zhou, H., Cheng, X., Tang, R., Munoz, M. & Nguyen, N. (2000b).** Recombinant respiratory syncytial viruses with deletions in the NS1, NS2, SH, and M2-2 genes are attenuated in vitro and in vivo. *Virology* **273**, 210-218.

**Johnson, P. R. & Collins, P. L. (1988).** The fusion glycoproteins of human respiratory syncytial virus of subgroups A and B: sequence conservation provides a structural basis for antigenic relatedness. *J Gen Virol* **69** ( Pt 10), 2623-2628.

**Kamer, G. & Argos, P. (1984).** Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res* **12**, 7269-7282.

**Kao, C. L., McIntosh, K., Fernie, B., Talis, A., Pierik, L. & Anderson, L. (1984).** Monoclonal antibodies for the rapid diagnosis of respiratory syncytial virus infection by immunofluorescence. *Diagn Microbiol Infect Dis* **2**, 199-206.

**Karlin, D., Ferron, F., Canard, B. & Longhi, S. (2003).** Structural disorder and modular organization in Paramyxovirinae N and P. *J Gen Virol* **84**, 3239-3252.

**Karron, R. A., Buonagurio, D. A., Georgiu, A. F., Whitehead, S. S., Adamus, J. E., Clements-Mann, M. L., Harris, D. O., Randolph, V. B., Udem, S. A., Murphy, B. R. & Sidhu, M. S. (1997).** Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant. *Proc Natl Acad Sci U S A* **94**, 13961-13966.

**Khattar, S. K., Yunus, A. S., Collins, P. L. & Samal, S. K. (2001).** Deletion and substitution analysis defines regions and residues within the phosphoprotein of bovine respiratory syncytial virus that affect transcription, RNA replication, and interaction with the nucleoprotein. *Virology* **285**, 253-269.

**Killingsworth, C. R., Shore, S. A., Alessandrini, F., Dey, R. D. & Paulauskis, J. D. (1997).** Rat alveolar macrophages express preprotachykinin gene-I mRNA-encoding tachykinins. *Am J Physiol* **273**, L1073-1081.

**Kisch, A. L., Johnson, K. M. & Chanock, R. M. (1962).** Immunofluorescence with respiratory syncytial virus. *Virology* **16**, 177-189.

**Knight, A. (2008).** The beginning of the end for chimpanzee experiments? *Philos Ethics Humanit Med* **3**, 16.

**Knight, P. R., Bedows, E., Nahrwold, M. L., Maassab, H. F., Smitka, C. W. & Busch, M. T. (1983).** Alterations in influenza virus pulmonary pathology induced by diethyl ether, halothane, enflurane, and pentobarbital anesthesia in mice. *Anesthesiology* **58**, 209-215.

**Kochva, U., Leonov, H. & Arkin, I. T. (2003).** Modeling the structure of the respiratory syncytial virus small hydrophobic protein by silent-mutation analysis of global searching molecular dynamics. *Protein Sci* **12**, 2668-2674.

**Kozak, M. (1978).** How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15**, 1109-1123.

**Kozak, M. (1984).** Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. *Nucleic Acids Res* **12**, 3873-3893.

**Kozak, M. (1989).** The scanning model for translation: an update. *J Cell Biol* **108**, 229-241.

**Krempl, C. D. & Collins, P. L. (2004).** Reevaluation of the virulence of prototypic strain 15 of pneumonia virus of mice. *J Virol* **78**, 13362-13365.

**Krempl, C. D., Lamirande, E. W. & Collins, P. L. (2005).** Complete sequence of the RNA genome of pneumonia virus of mice (PVM). *Virus Genes* **30**, 237-249.

**Krempl, C. D., Wnekowicz, A., Lamirande, E. W., Nayebagha, G., Collins, P. L. & Buchholz, U. J. (2007).** Identification of a novel virulence factor in recombinant pneumonia virus of mice. *J Virol* **81**, 9490-9501.

**Kuo, L., Fearn, R. & Collins, P. L. (1996).** The structurally diverse intergenic regions of respiratory syncytial virus do not modulate sequential transcription by a dicistronic minigenome. *J Virol* **70**, 6143-6150.

**Kwilas, S., Liesman, R. M., Zhang, L., Walsh, E., Pickles, R. J. & Peeples, M. E. (2009).** Respiratory syncytial virus grown in Vero cells contains a truncated attachment protein that alters its infectivity and dependence on glycosaminoglycans. *J Virol* **83**, 10710-10718.

**Kyte, J. & Doolittle, R. F. (1982).** A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* **157**, 105-132.

**Lamb, R. A. & Kolakofsky, D. (2001).** *Paramyxoviridae: The Viruses and Their Replication*. In *Fields virology*, 4th edn, pp. 1305-1340. Edited by B. N. Fields, D. M. Knipe, P. M. Howley & D. E. Griffin. Philadelphia: Lippincott Williams & Wilkins.

**Lambert, D. M. (1988).** Role of oligosaccharides in the structure and function of respiratory syncytial virus glycoproteins. *Virology* **164**, 458-466.

**Lang, F., Foller, M., Lang, K. S., Lang, P. A., Ritter, M., Gulbins, E., Vereninov, A. & Huber, S. M. (2005).** Ion channels in cell proliferation and apoptotic cell death. *J Membr Biol* **205**, 147-157.

**Langedijk, J. P., de Groot, B. L., Berendsen, H. J. & van Oirschot, J. T. (1998).** Structural homology of the central conserved region of the attachment protein G of respiratory syncytial virus with the fourth subdomain of 55-kDa tumor necrosis factor receptor. *Virology* **243**, 293-302.

**Leader, S. & Kohlase, K. (2003).** Recent trends in severe respiratory syncytial virus (RSV) among US infants, 1997 to 2000. *J Pediatr* **143**, S127-132.

**Levine, S., Klaiber-Franco, R. & Paradiso, P. R. (1987).** Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J Gen Virol* **68** (Pt 9), 2521-2524.

**Li, D., Jans, D. A., Bardin, P. G., Meanger, J., Mills, J. & Ghildyal, R. (2008a).** Association of respiratory syncytial virus M protein with viral nucleocapsids is mediated by the M2-1 protein. *J Virol* **82**, 8863-8870.

**Li, J., Rahmeh, A., Morelli, M. & Whelan, S. P. (2008b).** A conserved motif in region v of the large polymerase proteins of nonsegmented negative-sense RNA viruses that is essential for mRNA capping. *J Virol* **82**, 775-784.

**Li, J., Wang, J. T. & Whelan, S. P. (2006).** A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. *Proc Natl Acad Sci U S A* **103**, 8493-8498.

**Ling, R. & Pringle, C. R. (1989a).** Polypeptides of pneumonia virus of mice. I. Immunological cross-reactions and post-translational modifications. *J Gen Virol* **70** ( Pt 6), 1427-1440.

**Ling, R. & Pringle, C. R. (1989b).** Polypeptides of pneumonia virus of mice. II. Characterization of the glycoproteins. *J Gen Virol* **70** ( Pt 6), 1441-1452.

**Lopez, J. A., Bustos, R., Orvell, C., Berois, M., Arbiza, J., Garcia-Barreno, B. & Melero, J. A. (1998).** Antigenic structure of human respiratory syncytial virus fusion glycoprotein. *J Virol* **72**, 6922-6928.

**Lu, B., Brazas, R., Ma, C. H., Kristoff, T., Cheng, X. & Jin, H. (2002).** Identification of temperature-sensitive mutations in the phosphoprotein of respiratory syncytial virus that are likely involved in its interaction with the nucleoprotein. *J Virol* **76**, 2871-2880.

**Martinez, I. & Melero, J. A. (1998).** Enhanced neutralization of human respiratory syncytial virus by mixtures of monoclonal antibodies to the attachment (G) glycoprotein. *J Gen Virol* **79** ( Pt 9), 2215-2220.

**Massey, D. M. & Lenard, J. (1987).** Inactivation of the RNA polymerase of vesicular stomatitis virus by N-ethylmaleimide and protection by nucleoside triphosphates. Evidence for a second ATP binding site on L protein. *J Biol Chem* **262**, 8734-8737.

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

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

**Meijer, H. A., Dictus, W. J., Keuning, E. D. & Thomas, A. A. (2000).** Translational control of the *Xenopus laevis* connexin-41 5'-untranslated region by three upstream open reading frames. *J Biol Chem* **275**, 30787-30793.

**Melero, J. A. (2006).** Molecular Biology of Human Respiratory Syncytial Virus. In *Perspectives in Medical Virology*, pp. 1-42. Edited by C. Patricia: Elsevier.

**Melero, J. A., Garcia-Barreno, B., Martinez, I., Pringle, C. R. & Cane, P. A. (1997).** Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol* **78** ( Pt 10), 2411-2418.

**Mink, M. A., Stec, D. S. & Collins, P. L. (1991).** Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* **185**, 615-624.

**Miyairi, I. & DeVincenzo, J. P. (2008).** Human genetic factors and respiratory syncytial virus disease severity. *Clin Microbiol Rev* **21**, 686-703.

**Moore, E. C., Barber, J. & Tripp, R. A. (2008).** Respiratory syncytial virus (RSV) attachment and nonstructural proteins modify the type I interferon response associated with suppressor of cytokine signaling (SOCS) proteins and IFN-stimulated gene-15 (ISG15). *Virol J* **5**, 116.

**Moore, M. L., Chi, M. H., Luongo, C., Lukacs, N. W., Polosukhin, V. V., Huckabee, M. M., Newcomb, D. C., Buchholz, U. J., Crowe, J. E., Jr., Goleniewska, K., Williams, J. V., Collins, P. L. & Peebles, R. S., Jr. (2009).** A Chimeric A2 Strain of Respiratory Syncytial Virus (RSV) with the Fusion Protein of RSV Strain Line 19 Exhibits Enhanced Viral Load, Mucus, and Airway Dysfunction. *J Virol* **83**, 4185-4194.

**Moudy, R. M., Sullender, W. M. & Wertz, G. W. (2004).** Variations in intergenic region sequences of Human respiratory syncytial virus clinical isolates: analysis of effects on transcriptional regulation. *Virology* **327**, 121-133.

**Mufson, M. A., Orvell, C., Rafnar, B. & Norrby, E. (1985).** Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol* **66** ( Pt 10), 2111-2124.

**Murphy, L. B., Loney, C., Murray, J., Bhella, D., Ashton, P. & Yeo, R. P. (2003).** Investigations into the amino-terminal domain of the respiratory syncytial virus nucleocapsid protein reveal elements important for nucleocapsid formation and interaction with the phosphoprotein. *Virology* **307**, 143-153.

**Murray, J., Loney, C., Murphy, L. B., Graham, S. & Yeo, R. P. (2001).** Characterization of monoclonal antibodies raised against recombinant respiratory syncytial virus nucleocapsid (N) protein: identification of a region in the carboxy terminus of N involved in the interaction with P protein. *Virology* **289**, 252-261.

**Norrby, E., Marusyk, H. & Orvell, C. (1970).** Morphogenesis of respiratory syncytial virus in a green monkey kidney cell line (Vero). *J Virol* **6**, 237-242.

**Noton, S. L., Cowton, V. M., Zack, C. R., McGivern, D. R. & Fearn, R. (2010).** Evidence that the polymerase of respiratory syncytial virus initiates RNA replication in a nontemplated fashion. *Proc Natl Acad Sci U S A* **107**, 10226-10231.

**O'Connor, T. M., O'Connell, J., O'Brien, D. I., Goode, T., Bredin, C. P. & Shanahan, F. (2004).** The role of substance P in inflammatory disease. *J Cell Physiol* **201**, 167-180.

**Ogino, T., Yadav, S. P. & Banerjee, A. K. (2010).** Histidine-mediated RNA transfer to GDP for unique mRNA capping by vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci U S A* **107**, 3463-3468.

**Oshansky, C. M., Zhang, W., Moore, E. & Tripp, R. A. (2009).** The host response and molecular pathogenesis associated with respiratory syncytial virus infection. *Future Microbiol* **4**, 279-297.

**Paccaud, M. F. & Jacquier, C. (1970).** A respiratory syncytial virus of bovine origin. *Arch Gesamte Virusforsch* **30**, 327-342.

**Page, N. M. (2005).** New challenges in the study of the mammalian tachykinins. *Peptides* **26**, 1356-1368.

**Page, N. M., Woods, R. J., Gardiner, S. M., Lomthaisong, K., Gladwell, R. T., Butlin, D. J., Manyonda, I. T. & Lowry, P. J. (2000).** Excessive placental secretion of neurokinin B during the third trimester causes pre-eclampsia. *Nature* **405**, 797-800.

**Paulus, S. C., Hirschfeld, A. F., Victor, R. E., Brunstein, J., Thomas, E. & Turvey, S. E. (2007).** Common human Toll-like receptor 4 polymorphisms--role in susceptibility to respiratory syncytial virus infection and functional immunological relevance. *Clin Immunol* **123**, 252-257.

**Peabody, D. S. & Berg, P. (1986).** Termination-reinitiation occurs in the translation of mammalian cell mRNAs. *Mol Cell Biol* **6**, 2695-2703.

**Poch, O., Blumberg, B. M., Bougueleret, L. & Tordo, N. (1990).** Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J Gen Virol* **71** ( Pt 5), 1153-1162.



**Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989).** Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J* **8**, 3867-3874.

**Polack, F. P., Irusta, P. M., Hoffman, S. J., Schiatti, M. P., Melendi, G. A., Delgado, M. F., Laham, F. R., Thumar, B., Hendry, R. M., Melero, J. A., Karron, R. A., Collins, P. L. & Kleiberger, S. R. (2005).** The cysteine-rich region of respiratory syncytial virus attachment protein inhibits innate immunity elicited by the virus and endotoxin. *Proc Natl Acad Sci U S A* **102**, 8996-9001.

**Poyry, T. A., Kaminski, A. & Jackson, R. J. (2004).** What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev* **18**, 62-75.

**Prince, G. A., Horswood, R. L., Berndt, J., Suffin, S. C. & Chanock, R. M. (1979).** Respiratory syncytial virus infection in inbred mice. *Infect Immun* **26**, 764-766.

**Prince, G. A., Jenson, A. B., Horswood, R. L., Camargo, E. & Chanock, R. M. (1978).** The pathogenesis of respiratory syncytial virus infection in cotton rats. *Am J Pathol* **93**, 771-791.

**Pringle, C. R. & Eglin, R. P. (1986).** Murine pneumonia virus: seroepidemiological evidence of widespread human infection. *J Gen Virol* **67 ( Pt 6)**, 975-982.

**Puffer, B. A., Sharron, M., Coughlan, C. M., Baribaud, F., McManus, C. M., Lee, B., David, J., Price, K., Horuk, R., Tsang, M. & Doms, R. W. (2000).** Expression and coreceptor function of APJ for primate immunodeficiency viruses. *Virology* **276**, 435-444.

**Puthothu, B., Krueger, M., Forster, J., Heinze, J., Weckmann, M. & Heinzmann, A. (2007).** Interleukin (IL)-18 polymorphism 133C/G is associated with severe respiratory syncytial virus infection. *Pediatr Infect Dis J* **26**, 1094-1098.

**Randhawa, J. S., Chambers, P., Pringle, C. R. & Easton, A. J. (1995).** Nucleotide sequences of the genes encoding the putative attachment glycoprotein (G) of mouse and tissue culture-passaged strains of pneumonia virus of mice. *Virology* **207**, 240-245.

**Rawling, J., Garcia-Barreno, B. & Melero, J. A. (2008).** Insertion of the two cleavage sites of the respiratory syncytial virus fusion protein in Sendai virus fusion protein leads to enhanced cell-cell fusion and a decreased dependency on the HN attachment protein for activity. *J Virol* **82**, 5986-5998.

**Renshaw, R. W., Zylich, N. C., Laverack, M. A., Glaser, A. L. & Dubovi, E. J. (2010).** Pneumovirus in dogs with acute respiratory disease. *Emerg Infect Dis* **16**, 993-995.

**Richman, A. V., Pedreira, F. A. & Tauraso, N. M. (1971).** Attempts to demonstrate hemagglutination and hemadsorption by respiratory syncytial virus. *Appl Microbiol* **21**, 1099-1100.

**Rixon, H. W., Brown, G., Aitken, J., McDonald, T., Graham, S. & Sugrue, R. J. (2004).** The small hydrophobic (SH) protein accumulates within lipid-raft structures of the Golgi complex during respiratory syncytial virus infection. *J Gen Virol* **85**, 1153-1165.

**Rixon, H. W., Brown, G., Murray, J. T. & Sugrue, R. J. (2005).** The respiratory syncytial virus small hydrophobic protein is phosphorylated via a mitogen-activated protein kinase p38-dependent tyrosine kinase activity during virus infection. *J Gen Virol* **86**, 375-384.

**Roberts, S. R., Lichtenstein, D., Ball, L. A. & Wertz, G. W. (1994).** The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. *J Virol* **68**, 4538-4546.

**Rosenberg, H. F., Bonville, C. A., Easton, A. J. & Domachowske, J. B. (2005).** The pneumonia virus of mice infection model for severe respiratory syncytial virus infection: identifying novel targets for therapeutic intervention. *Pharmacol Ther* **105**, 1-6.

**Rosenberg, H. F., Dyer, K. D. & Domachowske, J. B. (2009).** Respiratory viruses and eosinophils: exploring the connections. *Antiviral Res* **83**, 1-9.

**Routledge, E. G., Willcocks, M. M., Morgan, L., Samson, A. C., Scott, R. & Toms, G. L. (1987).** Heterogeneity of the respiratory syncytial virus 22K protein revealed by Western blotting with monoclonal antibodies. *J Gen Virol* **68 (Pt 4)**, 1209-1215.

**Saaresranta, T. & Polo, O. (2002).** Hormones and breathing. *Chest* **122**, 2165-2182.

**Samal, S. K. & Collins, P. L. (1996).** RNA replication by a respiratory syncytial virus RNA analog does not obey the rule of six and retains a nonviral trinucleotide extension at the leader end. *J Virol* **70**, 5075-5082.

**Schlender, J., Bossert, B., Buchholz, U. & Conzelmann, K. K. (2000).** Bovine respiratory syncytial virus nonstructural proteins NS1 and NS2 cooperatively antagonize alpha/beta interferon-induced antiviral response. *J Virol* **74**, 8234-8242.

**Schrag, S. J., Rota, P. A. & Bellini, W. J. (1999).** Spontaneous mutation rate of measles virus: direct estimation based on mutations conferring monoclonal antibody resistance. *J Virol* **73**, 51-54.

**Schwarze, J., O'Donnell, D. R., Rohwedder, A. & Openshaw, P. J. (2004).** Latency and persistence of respiratory syncytial virus despite T cell immunity. *Am J Respir Crit Care Med* **169**, 801-805.

**Scott, P. D., Meng, B., Marriott, A. C., Easton, A. J. & Dimmock, N. J. (2011).** DI influenza A virus protects in vivo against disease caused by a heterologous influenza B virus. *J Gen Virol*.

**Shafritz, D. A. & Lieberman, H. M. (1984).** The molecular biology of hepatitis B virus. *Annu Rev Med* **35**, 219-232.

**Shingai, M., Azuma, M., Ebihara, T., Sasai, M., Funami, K., Ayata, M., Ogura, H., Tsutsumi, H., Matsumoto, M. & Seya, T. (2008).** Soluble G protein of respiratory syncytial virus inhibits Toll-like receptor 3/4-mediated IFN-beta induction. *Int Immunol* **20**, 1169-1180.

**Shope, R. E. (1934).** The Infection of Ferrets with Swine Influenza Virus. *J Exp Med* **60**, 49-61.

**Slack, M. S. & Easton, A. J. (1998).** Characterization of the interaction of the human respiratory syncytial virus phosphoprotein and nucleocapsid protein using the two-hybrid system. *Virus Res* **55**, 167-176.

**Snyder, R. L., Tyler, G. & Summers, J. (1982).** Chronic hepatitis and hepatocellular carcinoma associated with woodchuck hepatitis virus. *Am J Pathol* **107**, 422-425.

**Spring, S. B. & Tolpin, M. D. (1983).** Enzymatic cleavage of a glycoprotein of respiratory syncytial virus. *Arch Virol* **76**, 359-363.

**Stec, D. S., Hill, M. G., 3rd & Collins, P. L. (1991).** Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology* **183**, 273-287.

**Summers, J., Smolec, J. M. & Snyder, R. (1978).** A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. *Proc Natl Acad Sci U S A* **75**, 4533-4537.

**Taylor, G., Stott, E. J., Hughes, M. & Collins, A. P. (1984).** Respiratory syncytial virus infection in mice. *Infect Immun* **43**, 649-655.

**Taylor, K. L., Murphy, P. C., Asher, L. V., LeDuc, J. W. & Lemon, S. M. (1993).** Attenuation phenotype of a cell culture-adapted variant of hepatitis A virus (HM175/p16) in susceptible New World owl monkeys. *J Infect Dis* **168**, 592-601.

**Techarpornkul, S., Collins, P. L. & Peeples, M. E. (2002).** Respiratory syncytial virus with the fusion protein as its only viral glycoprotein is less dependent on cellular glycosaminoglycans for attachment than complete virus. *Virology* **294**, 296-304.

**Teng, M. N. & Collins, P. L. (1998).** Identification of the respiratory syncytial virus proteins required for formation and passage of helper-dependent infectious particles. *J Virol* **72**, 5707-5716.

**Teng, M. N. & Collins, P. L. (1999).** Altered growth characteristics of recombinant respiratory syncytial viruses which do not produce NS2 protein. *J Virol* **73**, 466-473.

**Teng, M. N. & Collins, P. L. (2002).** The central conserved cystine noose of the attachment G protein of human respiratory syncytial virus is not required for efficient viral infection in vitro or in vivo. *J Virol* **76**, 6164-6171.

**Teng, M. N., Whitehead, S. S. & Collins, P. L. (2001).** Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication in vitro and in vivo. *Virology* **289**, 283-296.

**Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2002).** Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**, Unit 2 3.

**Thorpe, L. C. & Easton, A. J. (2005).** Genome sequence of the non-pathogenic strain 15 of pneumonia virus of mice and comparison with the genome of the pathogenic strain J3666. *J Gen Virol* **86**, 159-169.

**Tournoy, K. G., De Swert, K. O., Leclere, P. G., Lefebvre, R. A., Pauwels, R. A. & Joos, G. F. (2003).** Modulatory role of tachykinin NK1 receptor in cholinergic contraction of mouse trachea. *Eur Respir J* **21**, 3-10.

**Tran, T. L., Castagne, N., Dubosclard, V., Noinville, S., Koch, E., Moudjou, M., Henry, C., Bernard, J., Yeo, R. P. & Eleouet, J. F. (2009).** The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *J Virol* **83**, 6363-6374.

**Treuhaff, M. W. & Beem, M. O. (1982).** Defective interfering particles of respiratory syncytial virus. *Infect Immun* **37**, 439-444.

**Tripp, R. A., Jones, L. & Anderson, L. J. (2000a).** Respiratory syncytial virus G and/or SH glycoproteins modify CC and CXC chemokine mRNA expression in the BALB/c mouse. *J Virol* **74**, 6227-6229.

**Tripp, R. A., Jones, L. P., Haynes, L. M., Zheng, H., Murphy, P. M. & Anderson, L. J. (2001).** CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. *Nat Immunol* **2**, 732-738.

**Tripp, R. A., Moore, D., Jones, L., Sullender, W., Winter, J. & Anderson, L. J. (1999).** Respiratory syncytial virus G and/or SH protein alters Th1 cytokines, natural killer cells, and neutrophils responding to pulmonary infection in BALB/c mice. *J Virol* **73**, 7099-7107.

**Tripp, R. A., Moore, D., Winter, J. & Anderson, L. J. (2000b).** Respiratory syncytial virus infection and G and/or SH protein expression contribute to substance P, which mediates inflammation and enhanced pulmonary disease in BALB/c mice. *J Virol* **74**, 1614-1622.

**Trudel, M., Nadon, F., Seguin, C. & Binz, H. (1991).** Protection of BALB/c mice from respiratory syncytial virus infection by immunization with a synthetic peptide derived from the G glycoprotein. *Virology* **185**, 749-757.

**Tusnady, G. E. & Simon, I. (2001).** The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849-850.

**Umehara, H., Bloom, E., Okazaki, T., Domae, N. & Imai, T. (2001).** Fractalkine and vascular injury. *Trends Immunol* **22**, 602-607.

**Valarcher, J. F., Bourhy, H., Lavenu, A., Bourges-Abella, N., Roth, M., Andreoletti, O., Ave, P. & Schelcher, F. (2001).** Persistent infection of B lymphocytes by bovine respiratory syncytial virus. *Virology* **291**, 55-67.

**Valarcher, J. F., Furze, J., Wyld, S., Cook, R., Conzelmann, K. K. & Taylor, G. (2003).** Role of alpha/beta interferons in the attenuation and immunogenicity of recombinant bovine respiratory syncytial viruses lacking NS proteins. *J Virol* **77**, 8426-8439.

**Valarcher, J. F., Furze, J., Wyld, S. G., Cook, R., Zimmer, G., Herrler, G. & Taylor, G. (2006).** Bovine respiratory syncytial virus lacking the virokinin or with a mutation in furin cleavage site RA(R/K)R109 induces less pulmonary inflammation

without impeding the induction of protective immunity in calves. *J Gen Virol* **87**, 1659-1667.

**Venkatesan, S., Elango, N. & Chanock, R. M. (1983).** Construction and characterization of cDNA clones for four respiratory syncytial viral genes. *Proc Natl Acad Sci U S A* **80**, 1280-1284.

**Vidal, S. & Kolakofsky, D. (1989).** Modified model for the switch from Sendai virus transcription to replication. *J Virol* **63**, 1951-1958.

**Villanueva, N., Hardy, R., Asenjo, A., Yu, Q. & Wertz, G. (2000).** The bulk of the phosphorylation of human respiratory syncytial virus phosphoprotein is not essential but modulates viral RNA transcription and replication. *J Gen Virol* **81**, 129-133.

**Walsh, E. E., McConnochie, K. M., Long, C. E. & Hall, C. B. (1997).** Severity of respiratory syncytial virus infection is related to virus strain. *J Infect Dis* **175**, 814-820.

**Wang, Z. & Sachs, M. S. (1997).** Ribosome stalling is responsible for arginine-specific translational attenuation in *Neurospora crassa*. *Mol Cell Biol* **17**, 4904-4913.

**Waterson, A. P. & Hobson, D. (1962).** Relationship between respiratory syncytial virus and Newcastle disease-parainfluenza group. *Br Med J* **2**, 1166-1167.

**Weinstock, J. V., Blum, A., Walder, J. & Walder, R. (1988).** Eosinophils from granulomas in murine schistosomiasis *mansoni* produce substance P. *J Immunol* **141**, 961-966.

**Welliver, R. C. (2003).** Review of epidemiology and clinical risk factors for severe respiratory syncytial virus (RSV) infection. *J Pediatr* **143**, S112-117.

**Wertz, G. W., Collins, P. L., Huang, Y., Gruber, C., Levine, S. & Ball, L. A. (1985).** Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc Natl Acad Sci U S A* **82**, 4075-4079.

**Whitehead, S. S., Bukreyev, A., Teng, M. N., Firestone, C. Y., St Claire, M., Elkins, W. R., Collins, P. L. & Murphy, B. R. (1999).** Recombinant respiratory syncytial virus bearing a deletion of either the NS2 or SH gene is attenuated in chimpanzees. *J Virol* **73**, 3438-3442.

**Whitehead, S. S., Firestone, C. Y., Collins, P. L. & Murphy, B. R. (1998a).** A single nucleotide substitution in the transcription start signal of the M2 gene of respiratory

syncytial virus vaccine candidate cpts248/404 is the major determinant of the temperature-sensitive and attenuation phenotypes. *Virology* **247**, 232-239.

**Whitehead, S. S., Juhasz, K., Firestone, C. Y., Collins, P. L. & Murphy, B. R. (1998b).** Recombinant respiratory syncytial virus (RSV) bearing a set of mutations from cold-passaged RSV is attenuated in chimpanzees. *J Virol* **72**, 4467-4471.

**Wilkinson, T. M., Donaldson, G. C., Johnston, S. L., Openshaw, P. J. & Wedzicha, J. A. (2006).** Respiratory syncytial virus, airway inflammation, and FEV1 decline in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* **173**, 871-876.

**Yoneda, O., Imai, T., Goda, S., Inoue, H., Yamauchi, A., Okazaki, T., Imai, H., Yoshie, O., Bloom, E. T., Domae, N. & Umehara, H. (2000).** Fractalkine-mediated endothelial cell injury by NK cells. *J Immunol* **164**, 4055-4062.

**Yunus, A. S., Khattar, S. K., Collins, P. L. & Samal, S. K. (2001).** Rescue of bovine respiratory syncytial virus from cloned cDNA: entire genome sequence of BRSV strain A51908. *Virus Genes* **23**, 157-164.

**Zhang, L., Peeples, M. E., Boucher, R. C., Collins, P. L. & Pickles, R. J. (2002).** Respiratory syncytial virus infection of human airway epithelial cells is polarized, specific to ciliated cells, and without obvious cytopathology. *J Virol* **76**, 5654-5666.

**Zimmer, G., Budz, L. & Herrler, G. (2001).** Proteolytic activation of respiratory syncytial virus fusion protein. Cleavage at two furin consensus sequences. *J Biol Chem* **276**, 31642-31650.

**Zimmer, G., Conzelmann, K. K. & Herrler, G. (2002).** Cleavage at the furin consensus sequence RAR/KR(109) and presence of the intervening peptide of the respiratory syncytial virus fusion protein are dispensable for virus replication in cell culture. *J Virol* **76**, 9218-9224.

**Zimmer, G., Rohn, M., McGregor, G. P., Schemann, M., Conzelmann, K. K. & Herrler, G. (2003).** Virokinin, a bioactive peptide of the tachykinin family, is released from the fusion protein of bovine respiratory syncytial virus. *J Biol Chem* **278**, 46854-46861.

## **APPENDIX I: Primers sequence**

The sequences are listed 5' to 3'.

1stORFOut-15-F

CCACAAGCTGACTTCACCTAGTGCGGGAAGAACTTAGAAGTGAGTGGCAG

1stORFOut-15-R

CTGCCACTCACTTCTAAGTTCCTTCCCGCACTAGGTGAAGTCAGCTTGTGG

1stORFOut-J6-F

AGTACTATCCTATTGGAATCAAGCGAGACC TGTAGAGCAGCTCATAC

1stORFOut-J6-R

GTATGAGCTGCTCTACAGGTCTCGCTTGAT TCCAATAGGATAGTACT

22K7

CGATGGGTCTGGACCATCC

F11

ACTGTTTACTATCTTAGCAA

F14

ACAGTCTATAGAACTTCTTA

F17

TCCACTGCACTACTATAGAT

F20

GTCAACGCTGACACACTGGTTTAT

F7

TATTGAGTCATGCAAGAGCA

F9

CATTGTAGGCGGCATGGCTG

G(4878)R

GGAGGGGTGGATGTGCTGTTC

G\_Round(F)

AGTTGACTTCACTTAGTATGGGAAGGAACCTTTGAAGTGAGTGGCAGCATT  
CCAATTTG AACTTTGAGAGAACTCAGCATCC

G\_Round(R)

AGTAGTTCTCTTGATGAGCTGCTCTACAGGTCTCATTGATTCCAATAGGA  
TAGTACT TATCC



G4E  
GGCCAGGATAAGTACTATCCTATTGG

G9  
CCCGAATTCATTA ACTACTGATAAGGTT

GFFor  
TTCCAAGGAGGTCTAATTCTCCCATGGTGA GCAAGGGCGAGG

GFPrev  
AATTTACTGCAGTTATGATCTAGAGTC

J1  
ACGCGAAAAAATGCATAACAAAAC

J4  
CTGCTGGTCATGTTGATCTCGA

J5  
GATTTACACAGCACAACAACC

J7  
AAGTACACCTGGCCTCACAGTT

L1D  
GGTCATCTGCTCAGTAAGTTGT

L1F  
CTAGAGTTTGCCTTACACCAAC

L1H  
CATCCTGCAATTCTTTCCGG

L1J  
CTCTTTGAGGCTTATATCAGCTG

L2C  
GAAGGACACCAGATTTCTTGAC

L2F  
CAGCTCTGTTGCAAACATAGAC

L2G  
ACTGAGAGGTTTGGAGACGA

L2H  
CTGAGTAGTAAACAGTTCCACTC

L3C  
CTCACTACAGGAAGTTCAGTG

L3D  
CGTTCTCGATGACGGACATGT

L3E  
TCATAGGTGAAGGTGCTGGA

L3F  
TAGTGGTGTTAAGCACTGGTTG

L3G  
TACAGCGAGGTATTGCACTC

L3H  
GCCCTGTTTACCATAACAAATGG

L50  
CTTACAACTCTTCCTCTCTAGC

M1  
CAGCCAACATATCACTAACTGTG

M2  
CACTGTACCCAGCATTACTCCCA

MG\_Stop(F)  
CCTGTAGAGCAGCTCATACAAGAGA ACTACTAGTTGACTTCACTTAGTATGG  
GAAGG

MG\_Stop(R)  
CCTTCCCATACTAAGTGAAGTCAACTAGTAGTTCTCTTGTATGAGCTGCTCT  
ACAGG

N1  
GTAAAATGTGTGGACACACAG

N3  
GGATTCTATCACATCAGAAAT

N5  
TGAAAGCCGAGAAAGCCAGGT

N4  
GCGAAGAAGTGGCAGAAATAG

NcoIFor  
CCTGAAAAAAGTTAGGATAAATAACCATGG AGAAAAAAATCACTGG

NcoIRev  
CCAGTGATTTTTTTCTCCATGGTTATTTAT CCTAACTTTTTTCAGG

NcoLucF

AGCCACCATGGAAGACGCC

NS1B

CTTGCCCTGTAGAACTAAACACG

NS2A

CCGAGCCTACAAAACATCACTAG

NS2B

GAGCTGACTCAGATCACTCCAA

P15FLSH3.2

ACCAACGTCGACATGACAACCACACATACACACCCAC

P15FLF2.2

AATTATGCCCGGGTACAATTGTCATGATAAAACTGTGAGG

P15FLL5

TTTAATATTTGCGGCCGCTCAGTTATTAACCCAAAATTGTTAATTATGTAG

P15FLL7

GGAGGAGACCTGGTTTCATTATTCTTGTCAGC

P15FLL7.2

GTGTCTGCACCTGGTTTCATTATTCTTGTCAGC

P15FLSH3

GGAGGAGTCGACATGACAACCACACATACACAC

P15FLSH3.2

ACCAACGTCGACATGACAACCACACATACACACCCAC

P15FLM2

GGAGGGGTCGACGTTGGTGGCGGTGGTTGT

P15FLM2.2

GTGTGGTTGTCATGTCGACGTTGGTGGCGGTGGTTGTTGTG

P15FLM23

GGAGGGCCCGGGCATAATTGAGTTAGTTAA

P15M23.2

GACAATTGTACCCGGGCATAATTGAGTTAGTTAATTA AAAACTTAGG

P15FLM25

GGGAAAGGAGGCGGCCGCAAAAAAAGGAATACAG

P15FLN3.2  
CTCAGATCACTCCAACCAGGTGAGAGATCCACCACCTATAGC

P2A  
CTTTGTGGAACCCGAGGAG

P2F  
CGTCGAATTCACCATGTCACCAAATATAACATGCCCCC

P2R  
CGTCTCTAGACTACAGATGAAGAACCCGGCTCTTGG

P3  
GAGTCTGATGTTGACATTGAGAC

P4  
CAAGTCTGACATCTTGACTGTC

PL1  
CACATTA ACTTCTTGTT CATCA

PL23  
GGAAAATTATACAAAGAAGTGTTG

PL29  
GTCAATATGGTGTCTGGAAC

PL48  
CTAGACATGTGAGAAGGTCCCA

PstIFor  
GGGGGGCTGCAGTCCTTCAATAAACCCCAG GC

PstIGJ3666  
GGGGGCTGCAGCAGCCTTTGCAGTATGTCGTCTA

PstIRev  
CCTCGCCCTTGCTCACCATGGGAGAATTAG ACCTCCTTGGA

PstLucR  
CCCCCTGCAGTTACACGGCGATCTTTCCG C

QCF3BSHT1F  
CCTTCAATAAACCCCAGGCCAGACCGGTTTACCCTGCTAGACG

QCF3BSHT1R  
CGTCTAGCAGGGTAAACCGGTCTGGCCTGGGGTTTATTGAAGG

QCF3BGL2F  
AGGAGGTCTAATTCTTAAGAGATCTATTCCTGAATTA ACTTCAG

QCF3BGL2R

CTGAAGTTAATTCAGGAATAGATCTCTTAAGAATTAGACCTCCT

QCFL2GA4728F

CCTGACACATTTAGGACTGTTGTAAAAAGTGAACCAAATGTGTAAGC

QCFL2GA4728R

GCTTACACATTTGGTTCACCTTTTACAACAGTCCTAAATGTGTCAGG

SH2

AACTGTGCATTGGCTGCTGAC

SH4360

GTTTCATCCCAATCACCTCCAC

T7

TAATACGACTCACTATAGG