TITLE NATURAL AND INDUCED VARIATION IN THE FUSION GLYCOPROTEIN GENE OF HUMAN RESPIRATORY SYNCYTIAL VIRUS SUBGROUP A

AUTHOR David John PLOWS

DEGREE Ph.D

AWARDING BODY Warwick University

DATE 1994

THESIS NUMBER DX184327

NATURAL AND INDUCED VARIATION IN THE FUSION GLYCOPROTEIN GENE OF HUMAN RESPIRATORY SYNCYTIAL VIRUS SUBGROUP A

by David John Plows (B.Sc. Warwick)

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Warwick

Research carried out in the

Department of Biological Sciences

Thesis submitted in June 1994

NUMEROUS ORIGINALS IN COLOUR

CONTENTS

Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

ACKNOWLEDGEMENTS

I would like to thank Professor Craig Pringle for his excellent supervision of this project and for his support and encouragement whilst writing this thesis. I would like to take this opportunity to acknowledge the support of my parents, both morally and financially during the course of this PhD and for all my time as a student. I would also like to thank Dr. Andrew Easton for his help during my time in the Pneumovirus Laboratory. I would especially like to thank Dr. Calliope Caravokyri and Dr. Easton for their help in proof-reading this thesis. Finally, I would like to thank all my lab-mates, who shall remain nameless, for their advice and companionship, it has been a pleasure working with you all.

DECLARATION

I hereby declare that all the results presented within this thesis were obtained by myself under the supervision of Professor Craig Pringle, with the exception of those instances where the contribution of others has been acknowledged. These results have not been submitted for a degree at any other institution.

ABBREVIATIONS

ATP	:	Adenosine triphosphate
BSA	:	Bovine serum albumin
CTP	:	Cytidine triphosphate
DMEM	:	Dulbeco's modified Eagles medium
DNA	:	Deoxyribonucleic acid
DTT	:	Dithiothreitol
dATP	:	2'-deoxyadenosine 5'- triphosphate
dCTP	:	2'-deoxycytidine 5'- triphosphate
dGTP	:	2'-deoxyguanosine 5'- triphosphate
dTTP	:	2'-deoxythymidine 5'- triphosphate
ELISA	:	Enzyme linked immunosorbant assay
FCS	:	Bovine foetal calf serum
g	:	Gravity
GMEM	:	Glasgow modified Eagle's medium
GTP	:	Guanosine triphosphate
HEPES	:	N-2-hydroxyethylpiperazine-N-2-ethane sulphonic acid
hr	:	Hour
IPTG	:	Isopropyl-thio-β-D-galactoside
IMP	:	Immunoprecipitation
kD	:	Kilodaltons
LMP	:	Low melting point agarose
M	:	Molar
mA	:	Milliamp
ml	:	Millilitre
mM	:	Millimolar
Mab	:	Monoclonal antibody
MeV	:	Measles virus
min	:	Minute
m.o.i.	:	Multiplicity of infection
M_r	:	Relative molecular weight

mRNA	:	Messanger ribonucleic acid
MuV	:	Mumps virus
NDV	:	Newcastle disease virus
O.D.	:	Optical density
ORF	:	Open reading frame
PBS	:	Phosphate buffered saline
PEG	:	Polyethylene glycol
PIV 3	:	Parainfluenza virus type 3
pfu	:	Plaque forming units
PMSF	:	Phenylmethylsulphonylfloride
PVM	:	Pneumonia virus of mice
RIP	:	Radioimmunoprecipitation
RNA	:	Ribonucleic acid
SDS	:	Sodium dodecyl sulphate
SV	:	Sendai virus
SV 5	:	Simian virus 5
T_c	:	Cytotoxic T-cell
T_h	:	Helper T-cell
TEMED	:	N, N, N', N'-tetra-methylethylenediamine
TRTV	:	Turkey rhinotracheitis virus
ts	:	Temperature-sensitive
TTP	:	Thymidine triphosphate
URT	:	Upper respiratory tract
V	:	Volts
VSV	:	Vesicular stomatitis virus
X-Gal	:	5-Bromo-4-chloro-3-indolyl-β-D-galactoside

SUMMARY

The nucleotide sequences of the fusion (F) protein encoding gene of five isolates of subgroup A respiratory syncytial (RS) virus, representing the five lineages associated with current epidemics of RS virus-associated disease, were determined and compared with F gene sequences from laboratory strains. Overall natural variation in the fusion protein gene among subgroup A isolates is low. Amino acid identities of between 97% to 99.5% were deduced. The relationship of the five lineage isolates, based on the sequence of their fusion protein genes, agrees broadly with evolutionary relationships inferred from comparison of their G, SH and N protein genes.

There is a region of high amino acid variation at the C-terminal end of the F_2-subunit, specifically between residues #101 and #105. None of the apparent amino acid coding changes are located in known epitopes. Using predictive structural models it is suggested that the few amino acid changes observed may alter the fusion protein structure especially in the C-terminal domain of the F_2-subunit. The secondary structure of the F_1-subunit is predicted to remain unaltered. It is hypothesised that amino acid variation in the F_2-subunit may result in antigenic variation, by altering a potential conformational epitope formed by interaction between the N-terminal region of F_1 and the C-terminal region of F_2.

Induced variation in the fusion protein gene of two candidate vaccine strains was investigated. Temperature-sensitive mutant *ts*A1, derived by intensive mutagenesis from the A2 strain, has two distinctive phenotypes; temperature-sensitivity and a retarded fusion protein mobility in non-reducing gels. Previous analysis of the phenotypic characteristics of *ts*A1 indicated that the F gene might be the site of the ts mutation. The sequence data derived in this study suggest that the site of ts lesion is located at residue #66 (Glu → Lys) and that the mobility phenotype is located at residue #102 (Pro → Ser). Mutant *ts*A1 exhibits a complex pattern of reversion with two classes of revertant viruses observed; a fully revertant virus which has wild-type growth characteristics but still retains retarded fusion protein mobility; and a partially revertant virus that possesses near wild-type growth characteristics and wild-type mobility. In fully revertant viruses the correction of the ts phenotype has been identified as the reversion of amino acid #66 to the wild-type residue (Lys → Glu) whilst the coding change at amino acid #102 is retained, resulting in a mobility phenotype similar to that found in mutant *ts*A1. In partially revertant viruses the coding changes at amino acid #66 and amino acid #102 are retained. In partially revertant viruses the coding change correcting for the mobility phenotype, and partially correcting for the ts phenotype, has been tentatively identified as additional coding changes at residues #103 (Thr → Ala) and/or #105 (Asn → Ser). *In vitro* expression of the fusion gene products of mutant, revertant and wild-type viruses in mammalian cells has confirmed that the mobility phenotype is solely a consequence of changes in the fusion protein gene.

Temperature-sensitive mutant *ts*1C is a triple ts mutant derived from the RSS-2 strain. The mutations detected in the fusion gene of mutant *ts*1C are conservative in nature and are not located in known epitopes. Therefore it is unlikely that the coding changes observed in the fusion protein gene account for the ts-phenotype or viral attenuation. It is also thought that the induced mutations have not altered the antigenic properties of the virus.

CHAPTER ONE

CHAPTER ONE

Introduction

1.1 The family *Paramyxoviridae* The family *Paramyxoviridae* comprises pleomorphic enveloped viruses possessing non-segmented genomes of negative-sense, single-stranded RNA. These viruses infect cells by fusion of the viral envelope with the cell surface membrane at neutral pH. The genome is then transcribed from a single promoter into 6 to 10 mRNAs, the majority of which are monocistronic. Independently assembled nucleocapsids are enveloped on the cell surface at sites containing viral envelope proteins, virions being released by budding. Transmission of the virus is airborne, mainly horizontal with no identifiable vectors involved. The effect on infected cells is cytolytic, but temperate and persistent infections are common. Physical features of infection are inclusion bodies, syncytium formation and haemadsorption.

The family *Paramyxoviridae* is subdivided into four genera, *Paramyxovirus, Rubulavirus, Morbillivirus* and *Pneumovirus* (**Table 1**), on the basis of measurement of nucleocapsid diameter and envelope surface projections by electron microscopy, antigenic or biochemical aspects of the large attachment glycoprotein present in the virion envelope, genomic order (**Figure 1**) and strategies used to express different P gene products (**Section 1.4.4**). Biochemically the paramyxoviruses and rubulaviruses are distinguishable from the other two by the ability of their attachment glycoprotein to exhibit both haemagglutinin and neuraminidase activity (H/N) (Kingsbury *et al.*, 1978; Matthews, 1982; Francki *et al.*, 1991). The attachment protein of the morbilliviruses possess haemagglutinin activity (H) only. Within the *Pneumovirus* genus there is some variation, with PVM able to haemagglutinate mouse cells (Berthiaume *et al.*, 1974; Ling and Pringle, 1989b). Analysis of the *Paramyxoviridae* at the molecular level has emphasised the relatedness that exists between the morbilliviruses, paramyxoviruses and rubulaviruses. The paramyxoviruses are closely related to the morbilliviruses in regard to genetic map, arrangement and sizes of ORFs and sequences of intergenic regions. In addition, the

Table 1 The family *Paramyxoviridae*

Subfamily *Paramyxovirinae*

Genus *Paramyxovirus*

- Human parainfluenza virus type 1 (type species)
- Human parainfluenza virus type 3
- Bovine parainfluenza virus type 3
- Mouse parainfluenza virus type 1 (Sendai virus)
- Simian parainfluenza virus type 10

Genus *Morbillivirus*

- Measles virus (type species)
- Dolphin morbillivirus
- Canine distemper virus
- Peste-des-petits-ruminants virus
- Phocine distemper virus
- Rinderpest virus

Genus *Rubulavirus*

- Mumps virus (type species)
- Avian paramyxovirus 1 (Newcastle disease virus)
- Avian paramyxovirus 2 (Yucaipa virus)
- Avian paramyxovirus 3
- Avian paramyxovirus 4
- Avian paramyxovirus 5 (Kunitachi virus)
- Avian paramyxovirus 7
- Avian paramyxovirus 8
- Avian paramyxovirus 9
- Human parainfluenza virus type 2
- Human parainfluenza virus type 4a
- Human parainfluenza virus type 4b
- Porcine rubulavirus (La-Piedad-Michoacan-Mexico virus)
- Simian parainfluenza virus 5
- Simian parainfluenza virus 41

Subfamily *Pneumovirinae*

Genus *Pneumovirus*

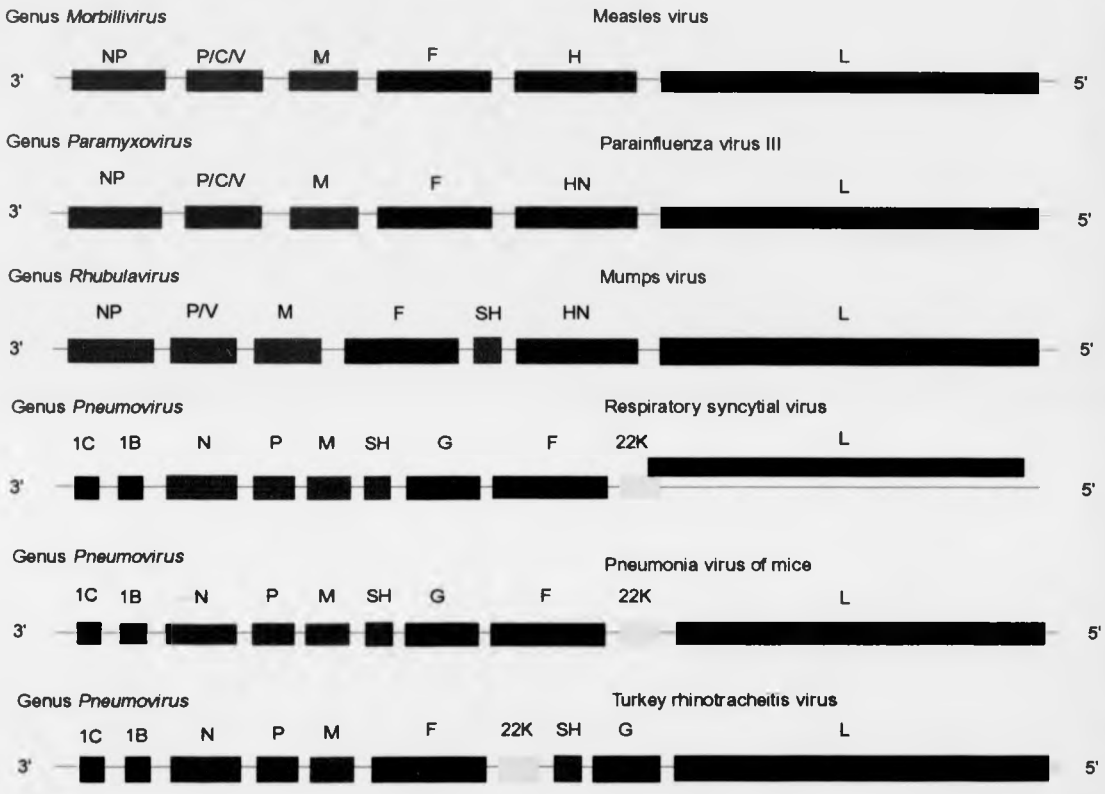
- Human respiratory syncytial virus (type species)
- Bovine respiratory syncytial virus
- Pneumonia virus of mice (murine pneumonia virus)
- Turkey rhinotracheitis virus

Unassigned viruses in the family

- Fer-de-Lance virus
- Nariva virus
- Penguin paramyxoviruses

Table 1 Summary of the classification of individual paramyxoviruses into the four genera, *Paramyxovirus*, *Rubulavirus; Morbillivirus* and *Pneumovirus*, that comprise the *Paramyxoviridae* family. Adapted from the 6th report of the International Committee on the Taxonomy of Viruses (Fauquet *et al.*, 1994).

Figure 1. Genomic organisation of the *PARAMYXOVIRIDAE*



Diagramatical representation of the genomes of members of the family *Paramyxoviridae* , showing the extra genes present in the pneumoviruses. Also showing the change in genomic order between TRTV and the other pneumoviruses. The internal site of initiation (AUG) of the L gene in RS viruses within the 22K gene is also shown.
Adapted from J.Barr,(1994).

morbillivirus NP, P, and L proteins possess sequence relatedness to their paramyxovirus counterparts but less homology to the rubula- and pneumoviruses (Morgan, 1991; data for the L gene sequences is shown in **Table 2**). Nucleotide sequencing of *Paramyxoviridae* L, P and V protein genes also supports this relationship model (Stec *et al.*, 1991).

The genomic composition of pneumoviruses is distinct, possessing four extra genes, one non-glycosylated envelope protein gene (22K), two non-structural protein genes (NS1 and NS2), and a small hydrophobic (SH) protein gene possibly analogous to the SH protein of the rubulaviruses. In addition to their unique genomic structure, the polypeptides encoded by the pneumovirus genome are generally smaller than their homologues in the other genera, and display little amino acid similarity to those of other viruses outside the genus (Spriggs and Collins, 1986a; Barr *et al.*, 1991; Yu *et al.*, 1992a and b; Chambers *et al.*, 1992). The exceptions to this are the L protein (Stec *et al.*, 1991) and the F protein (Spriggs *et al.*, 1986c) which both show a low but significant level of similarity with the corresponding proteins of other members of the family *Paramyxoviridae*.

1.1.1 The genus *Pneumovirus* Molecular studies comparing the known pneumoviruses indicate that they correspond closely by gene-map and by the numbers and types of gene products (Cash *et al.*, 1977; Cavanagh and Barrett, 1988; Ling and Pringle, 1988,1989a and b; Lerch *et al.*, 1989). The genomes of human and bovine RS viruses and PVM are identical in gene-order. There are differences; **1)** the start of the L gene of human and bovine RS viruses lies within the 22K gene but there is no overlap in PVM (A.Easton, personal communication); **2)** there are at least two gene products from the PVM P gene whereas in human RS virus there is only one (J.Barr, 1994). The pneumovirus TRTV has a distinctive gene-order (Ling *et al.*, 1992; Yu *et al.*, 1992b). The SH and G gene pair, and the F-22K gene pair are transposed, giving F-SH-G the same linear relationship as in the paramyxo-, morbilli- and rubulaviruses. However, whilst there is no firm evidence to suggest that TRTV possesses analogues of the RS viral NS1 and NS2 genes, the

Table 2 Conserved amino acids of *Paramyxoviridae* L proteins

	Paramyxovirus	Paramyxovirus	Morbillivirus	Rubulavirus
	PIV3	SV	MeV	NDV
PIV3		72	50	34
SV	61		50	33
MeV	36	37		32
NDV	23	21	20	

Table 2 The total number of conserved amino acids, including identical and conservative replacements, in each alignment pair (figures above the diagonal), and the total number of identical amino acids (figures below the diagonal) are expressed as a percentage of the aligned sequences. The sources of the published sequences are: PIV3 (Galinski *et al.*, 1988); SV (Shioda *et al.*, 1986); MeV (Blumberg *et al.*, 1988); and NDV (Yusoff *et al.*, 1987). Reproduced from E. M. Morgan, Chapter 5, The Paramyxoviruses, Plenum Press, New York, 1991.

possession of genes 22K and SH confirms the inclusion of TRTV in the genus Pneumovirus. There appears to be two subgroups of the ruminant RS viruses, ovine and bovine, that differ antigenically (Alansari and Potgieter, 1993). Significant antigenic relatedness exists between human RS and bovine RS viruses (Stott *et al.*, 1984; Lerch *et al.*, 1989), and between human RS virus and PVM. There is evidence for neutralising activity to PVM or a related virus in 27% to 75% of human adult sera (Horsfall and Hahn, 1940; Horsfall and Curnen, 1946; Pringle and Eglin, 1986). Antigenic cross-reactivity has also been found between human RS virus and PVM N (Gimenez *et al.*, 1984; Ling and Pringle, 1989a) and P proteins (Ling and Pringle, 1989a).

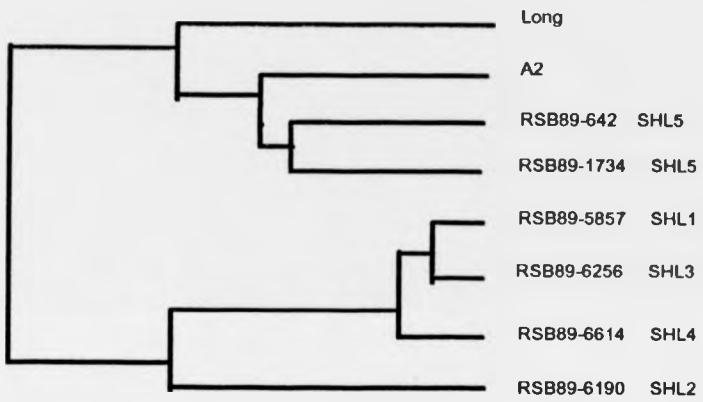
1.1.2 Respiratory syncytial (RS virus) virus Human respiratory syncytial (RS) virus was first isolated in 1957 and shown to be related to Chimpanzee coryza agent isolated in 1956. The virus replicates efficiently in enucleated cells (Follett *et al.*, 1975) and does not shut off host cell metabolism (Levine *et al.*, 1977). More than 90% of RS viral infectivity remains cell-associated (Peeples and Levine, 1980). Human RS virus has been classified into two subgroups on the basis of the reactivity of the G, F, M and N polypeptides with a panel of monoclonal antibodies (Gimenez *et al.*, 1984; Ward *et al.*, 1984; Mufson *et al.*, 1985; Anderson *et al.*, 1985). The A and B subgroups of human RS viruses exhibit differences in the molecular weights of the F_1 and F_2 polypeptides (Nörrby *et al.*, 1986) and the P protein (Cash *et al.*, 1977; Nörrby *et al.*, 1986). Within subgroups the G protein amino acid identity varies from 80 to 99% (Johnson *et al.*, 1987b; Sullender *et al.*, 1990; Cane *et al.*, 1992) whereas between subgroups the amino acid identity is only 53% (Johnson *et al.*, 1987b). The other human RS viral proteins show varying degrees of difference between subgroups (**Table 3**). Intra-group amino acid conservation is higher than inter-group for all proteins which suggests that the two subgroups have been evolving separately for some time (Collins, 1991). Sequence analysis of the SH and G protein genes and restriction analysis of the G and N protein genes of a large number of subgroup A isolates, suggest that there are at least five distinct isolate types that show extensive genetic divergence (**Figure 2**)(Cane and Pringle, 1991, 1992; Cane *et al.*,

Table 3 Inter- and intra-subgroup homology of RS virus polypeptides

Polypeptide	% intra-subgroup homology	% inter-subgroup homology
NS1	98	87
NS2	98	87
N	99	96
P	98	90
M	99	**ND**
SH	84	76
G	80 to 98	53
F	97 to 99.5	91
22K	98	92
L	97	**ND**

Table 3 Inter-subgroup figures are taken from Collins (1991). Intra-subgroup homology figures quoted are from data supplied by; NS2 and NS1 (J.Evans, personal communication); N, P, 22K, SH and L (C.Pringle, K.Tolley, A.Simpson, personal communication), G (Cane *et al.*, 1992), and F (**this Thesis**). **ND**. Data not available.

Figure 2. Relatedness of subgroup A isolates based on their SH and N genes.



Simple dendrogram illustrating the relative relationships between subgroup A isolates based on their N and SH protein genes. Reproduced from Cane and Pringle, 1991.

1992). These lineages appear to be distributed world-wide (Cane *et al.*, 1992) but any difference in virus virulence and immunity at the individual and community level is unknown. Within subgroup B there are at least three distinct isolate types (Akerlind *et al.*, 1988; Anderson *et al.*, 1991; Mufson *et al.*, 1991; Sullender *et al.*, 1991; Nagai *et al.*, 1993). Repeated infections of humans with RS virus has been recognised for many years (Beem, 1966; Henderson *et al.*, 1979). Studies have demonstrated that subgroup A and B isolate-types can co-circulate in the same community and that the relative frequency of their isolation can change from year to year (Mufson *et al.*, 1988; Hendry *et al.*, 1989; Breese-Hall *et al.*, 1990 and Anderson *et al.*, 1991). With the recognition that these individual isolates give rise to varying immune responses (Robinson and Everson, 1992), it has been proposed that isolate variation within a subgroup (and hence subsequent antigenic variation) may contribute towards adult reinfection.

1.2 The RS viral genome Sequence analysis of human RS virus strain A2 vRNA (subgroup A) shows that the genome is a non-polyadenylated strand, 15,222 nucleotides in length and of negative sense (Huang and Wertz, 1982; Collins, 1991; Mink *et al.*, 1991). cDNAs were obtained for 10 viral genes, accounting for 97% of the genome (Collins *et al.*, 1984a; Collins *et al.*, 1986). These genes generate 10 mRNAs, each encoding a single polypeptide whose Mr is consistent with the length of the major translational ORF (Cash *et al.*, 1977; Dubovi, 1982; Venkatesan *et al.*, 1983; Huang and Wertz, 1982; Collins *et al.*, 1984a; Satake *et al.*, 1984; Elango *et al.*, 1985b).

1.2.1 RS virus gene start and end sequences Sequences at the gene-start and -end are shown in **Table 4**. in comparison to another pneumovirus, PVM. Each gene has a conserved gene-start sequence, 3' GGGGCAAAU 5', that encodes the 5' end of the corresponding mRNA. The conserved gene-start signal resembles the SV gene-start signal in length and sequence, being preceded by an A residue that remains un-transcribed. The gene-end is a semi-conserved 12 to 13 nucleotide sequence (3' UC[$^{AA}/_{CU}$] $X_{(n=3-4)}$ $U_{(n=4-7)}$ 5') which is similar to other paramyxoviruses, and thought to function as a signal during

Table 4 Gene start and end sequences of human RS virus and PVM

Gene	Gene start sequence		Gene end sequence	
	human RS virus*	PVM	human RS virus*	PVM
NS1	GGGGCAAAU	AGGACAAGU	UAGUUAAAU	UAGUUAAUU
NS2	GGGGCAAAU	AGGAUAAAU	UAGUAAUUU	UAGUUAUAG
N	GGGGCAAAU	AGGAUAAAU	GAGUUAAU	UAUUUAAUU
P	GGGGCAAAU	AGGACAAAU	UAGUUAC	UAUUUAAUU
M	GGGGCAAAU	AGGAUAAAU	AAGUUAAU	UAGUUAAAU
SH	GGGGCAAAU	AGGAUAAGU	UAGUUAAUU	UAGUUAAC
G	GGGGCAAAU	AGGACAAAU	UAGUUACUU	UAGUUAAUG
F	GGGGCAAAU	AGGAUGAGU	UAGUUAUAU	UAGUUAAUU
22K	GGGGCAAAU**	AGGACAAAU	UAGUUAUU	UAGUUAUAU
L	GGGACAAAA	ND	AUCUUAUAAU	ND
consensus	GGGgCAAAu	AGGAuaAaU	uAgUuAauu	UAgUUAauu

Table 4 Upper case letters denote total conservation of a residue, lower case letters identify the most common residue at positions where conservation was not total. * The gene start and end sequences are for human RS virus strain A2. ** the L gene start is located within the 22K gene. Adapted from (Chambers *et al.*, 1990a). **ND** denotes that the sequence data is unavailable.

sequential transcription (Spriggs and Collins, 1986b). In mRNA, the complement of this sequence constitutes the 3' end immediately preceding the poly(A) tail. In RS virus the gene starts are well conserved but the gene ends show considerable variation, this is in contrast to PVM where the gene start and stop sequences are fairly well conserved. Human RS virus is unique, having the gene-start of the L gene 68 nucleotides upstream of the end of the 22K gene, thus the 22K and the L genes overlap. The L and 22K genes encode separate mRNAs with a common 68 nucleotide sequence, coding in the L gene but non-coding in 22K. The L gene-start sequence is also slightly different from those of the other genes, containing two nucleotide differences (3' CCCUGUUUU). This overlap is not unique to the *Paramyxoviridae*, an analogous situation exists for the *Filovirus* Ebola virus (Sanchez *et al.*, 1987).

During transcription the polymerase initiates gene transcription at the gene start signal (Collins *et al.*, 1986). The gene end signals are believed to cause the viral polymerase to stall and, by reiterative copying of short poly(U) tracts, synthesise a poly(A) tail. It is believed that no other gene sequences other than those identified as gene start and gene end sequences are able to act as structural signals since these are the only regions that display conservation between RS virus subgroups (Collins, 1991).

1.2.2 RS virus intergenic regions There are nine short intergenic regions in the RS virus genome which vary in length from one nucleotide (the N-P junction) to 52 nucleotides (for the G-F junction) and these are considered to be non-specific spacers. RS virus and PVM are notable for the lack of length and sequence conservation in the intergenic regions, a phenomenon which appears to exist in only one other paramyxovirus, SV5 (Paterson *et al.*, 1984; Collins and Wertz, 1985c; Hiebert *et al.*, 1985a; Collins *et al.*, 1985a, 1986). The high level of conservation in the other genera implies that these sequences may possess important transcriptional roles. Lack of homology in the

pneumovirus intergenic regions may indicate a difference in transcriptional function compared with the corresponding regions of many paramyxoviruses.

1.2.3 RS virus leader and trailer regions Sequences at the 3' and 5' ends of the pneumovirus genome (the leader and trailer) are believed to contain signals involved in RNA synthesis. In the RS virus genome, the first gene (NS1) is preceded by a 44-nucleotide leader region encoding a short leader RNA that initiates exactly at the 3' vRNA end. Human RS virus leader is shorter than other paramyxovirus leader sequences and unrelated in sequence, however a feature common to all is their richness in U nucleotides. The leader region is thought to encode at least two distinct nucleotide signals (Mink *et al.*, 1991); **1)** the region responsible for binding the polymerase complex prior to RNA synthesis and **2)** a site responsible for initiating encapsidation of nascent, negative sense genome copies (vRNAs). The importance of the leader region in RNA replication has been exemplified by experiments which demonstrated that a single nucleotide substitution in the 3' leader region can alter the efficiency of transcription (Collins *et al.*, 1991; 1993). Mutational analysis of the RS leader has identified at least two regions involved in RNA replication (nucleotides #1-10 and #21-26; P.Collins, personal communication).

There is a 155-nucleotide trailer region at the 5' end of the RS virus genome, much longer than any other equivalent sequence (54, 40 and 44 in SV, MeV and PIV3), however there is some commonality, with the 3'-terminal nucleotide being a U residue. It is thought that the last 40 nucleotides are also involved in RNA replication (P.Collins, personal communication). In common with other viruses from related groups, the RS virus leader and trailer regions show complementarity extending for 21 (of the first 26) nucleotides. This complementarity is essential for the RNA synthesis since the 5' trailer of the (-) vRNA encodes the 3' leader of the positive sense genome copy (vcRNA) which is thought to be functionally analogous to the 3' leader of the vRNA.

1.3 Replication and transcription of negative strand non-segmented genome RNA viruses

1.3.1 The replication-competent structure Vesicular stomatitis virus has become the model non-segmented negative strand RNA virus due to the knowledge that has been amassed concerning its replicative and transcriptive processes. The viral structure involved in RNA synthesis is the nucleocapsid, containing the polymerase (L), phosphoprotein (NS or P) and the nucleocapsid (N or NP) proteins plus the viral RNA. VSV L protein is considered to encode all RNA transcriptional activities, namely, ribonucleotide polymerase (De and Banerjee, 1985), cap methylase (Hendry *et al.*, 1988b; Hercyk *et al.*, 1988), poly(A) polymerase (Hunt *et al.*, 1984) and possibly protein kinase activity that phosphorylates the NS protein (equivalent of human RS virus P protein) (Barik and Banerjee, 1991; Barik, 1992; Hammond *et al.*, 1992; Beckes and Perrault, 1991). NS protein is required for the synthesis of authentic RNA products, suggesting that it functions as a polymerase accessory factor or transcriptional activator (De and Banerjee, 1985) and also maintains the N protein in a replication competent state. It is thought that this is accomplished by binding to and preventing N protein monomers from forming insoluble aggregates which are unable to support genome encapsidation (Howard and Wertz, 1989) in the form N-NS (Peluso and Moyer, 1988). VSV NS protein is highly phosphorylated with only the fully phosphorylated form (NS_2) being transcriptionally active (Takacs *et al.*, 1992). Also, VSV L is only able to bind to the N-RNA template in the presence of NS protein (Mellon and Emerson, 1978). The phosphorylation status of the NS protein was shown to regulate its ability to bind to the N-RNA template in the presence of L protein which is a crucial pre-requisite for RNA synthesis (Chattopadhyay and Banerjee, 1987a, and b).

1.3.2 RNA synthesis Kinetic studies have shown that the five VSV mRNA species are synthesised in genomic order with quantities decreasing with distance from the transcription start site. The model for VSV RNA synthesis (Banerjee, 1987b) is the stop-start model. This suggests that the polymerase binds to the genome at the 3' polymerase entry site and initially synthesises the leader sequence. The polymerase then re-initiates RNA synthesis at the beginning of the N gene and continues until it meets the termination

signal there adding the poly(A) tail. The polymerase then either dissociates from the template or again re-initiates, this time at the beginning of the NS gene and so the process goes on. The important feature of this model is that once a polymerase molecule becomes removed from the genome, it may only re-attach to the genome at the 3' entry site. In this way, 3' proximal genes are abundantly represented by mRNA while 5' proximal genes are considerably less so. Initiation of both mRNA transcription and genome replication begins when the nucleocapsid complex binds to the 3' end of the genome and begins to synthesise the 47 base long leader sequence. Transcription and replication have fundamental differences in that genome replication requires concomitant protein synthesis whereas transcription does not (Wertz and Levine, 1973). The requirement for ongoing protein synthesis is met entirely by synthesis of the nucleocapsid protein (Patton *et al.*, 1984) and this led to the postulation of a mechanism for the switch between transcription and replication that depended upon the concentration of the N protein (Blumberg *et al.*, 1981). This relies upon the ability of the N protein to encapsidate the nascent RNA molecule as it is synthesised. The polymerase complex begins to synthesise the positive polarity nascent leader RNA to which N protein molecules bind by recognition of an encapsidation signal (Blumberg *et al.*, 1983) believed to reside within the first 10 nucleotides of the leader sequence, although for optimal encapsidation the first 19 nucleotides are needed (Moyer *et al.*, 1990). As the polymerase complex passes down the genome, the growing RNA molecule is encapsidated by associating N protein molecules which somehow suppresses a transcriptional termination signal thought to be located at the leader-N junction. The polymerase complex is thus able to continue RNA synthesis and eventually give rise to a full-length anti-genome molecule fully encapsidated in N protein molecules. If N protein was not present at an appropriate concentration, the leader RNA species would be synthesised but would not be fully encapsidated and consequently the transcription termination signal not suppressed, leading to the generation of the 47 base long leader sequence. The polymerase would then either exit the genome or re

initiate transcription at the beginning of the N gene. This model suggests a self-regulating system where insufficient N protein stimulates increased mRNA synthesis to produce more viral proteins, and excess N protein triggers genome replication which will deplete the pool of viral proteins. The existence of negative sense leader sequences supports this model since it suggests that the polymerase complex initiating on anti-genomes (which are not templates for mRNA synthesis) is able to switch between a mode which makes full-length genomes to a mode in which genome synthesis is prematurely terminated at the postulated termination site, with the subsequent synthesis of a (-) leader sequence. No provision has been made in this model for the role of the VSV matrix protein, which has been shown to inhibit transcription of the VSV genome both *in vitro* and *in vivo* (Kaptur *et al.*, 1991).

The mechanism for paramyxovirus RNA replication is thought to be roughly similar to that found in VSV but the process is harder to define. The ability of the alternative SV P gene products (**Section 1.4.4**) to interact with SV nucleocapsid components (Yamada *et al.*, 1990; Curran *et al.*, 1992) and influence RNA synthetic operations, suggests subtle differences in the mode of RNA synthesis from the VSV model. Whilst having no apparent affect on replication, the C proteins are capable of causing inhibition of transcription (Curran *et al.*, 1992). Thus there is speculation that the C proteins, like N protein, are able to act as transcriptional inhibitors, aiding the switch from transcription to replication. Two other products of the SV P gene, namely the V and W proteins have been found to act as inhibitors of genome replication (Curran *et al.*, 1991a and b) and so perhaps only the P protein is essential for SV RNA synthesis, the negative effectors being accessory factors which are dispensable. The suggestion that the role played by these proteins is non-essential in the rubulaviruses and paramyxoviruses is supported by the finding that members of these genera do not contain a C or V ORF (Lyn *et al.*, 1991; Matsuoka *et al.*, 1991). Perhaps the most important difference between the RNA synthesis mechanisms of VSV and paramyxoviruses concerns the low abundance of

leader sequence in the latter (Kolakofsky *et al.*, 1991). In addition, two groups have identified large amounts of leader-N readthrough RNAs in MuV-infected cells at an abundance 20-fold greater than that found for leader RNA's (Castadenda and Wong, 1989; Chan *et al.*, 1989). These results have been interpreted to suggest that the leader termination signal is either weak or absent in paramyxoviruses, and that the leader RNAs are not an essential operational element for paramyxovirus RNA synthesis (Blumberg *et al.*, 1991). Also, cellular factors appear to influence viral mechanisms (Moyer *et al.*, 1986, 1990; De *et al.*, 1991; Horikami and Moyer, 1991; Leopardi *et al.*, 1993) and because of this, the paramyxovirus model does not suppose that all the processes of RNA synthesis are mediated exclusively by viral components.

1.3.3 RS virus mRNA transcription The 5' end of each mRNA is the complement of the gene-start sequence; the 3' end is the complement of the gene-end signal followed by the poly(A) tail. The 5' ends, by analogy to other nonsegmented negative-strand viruses, are capped and methylated by viral enzymatic activities associated with the nucleocapsid (Emerson, 1985; Kingsbury, 1985; Barik, 1992). The RS viral cap present on mRNA is of type cap 0 (only one methylation event) and has the structure ^{m7}G(5')ppp(5')Gp..., similar to the only other cap identified in a paramyxovirus, NDV (Colonno and Stone, 1975, 1976). Capping, including methylation, appears to be coupled to transcription. Purified RS virus nucleocapsids were shown to be transcriptionally active after the addition of mock-infected cell extract (Barik, 1992) which suggested that a cellular factor was an obligate requirement of transcription, possibly actin (Huang *et al.*, 1993). Polycistronic transcripts have been identified for every gene junction except for the one between the SH and G genes. They are present at 5 to 10% the level of individual mRNAs, although the polytranscript spanning the G-F gene junction is present at very low levels. The reason for this variance is unknown. They are generated by precise transcriptional readthrough across intergenic regions, thought to be created by polymerase error or by a mechanism similar to readthrough that occurs during vRNA replication. This has already been presupposed in RS virus by L gene transcription which initiates within

the 22K gene, so the polymerase must transcribe across the 22K gene-end signal in order to synthesise full-length L mRNA. The most abundant product of L gene transcription is a 68-nucleotide sequence, called L leader RNA, that initiates at the L gene-start and terminates at the 22K gene-end and contains a poly(A) tail. It possesses the first 20 codons of the L ORF and could encode a short protein (Collins, 1991). Full length L mRNA is less abundant and is generated from readthrough of the 22K gene-end signal, thus the presence of the gene-end signal within the L gene attenuates synthesis of full-length L mRNA.

1.3.4 RS virus mRNA translation For most mRNAs, the translational initiation codon is the first methionine codon, a context favoured for efficient translational initiation (Kozak, 1986). However the SH mRNA has an initiation codon in a unfavourable position (Collins and Wertz, 1985b; Johnson and Collins, 1988a) which may allow translation at the second AUG in the ORF, resulting in a truncated form of SH. The second exception is the G mRNA in which both the first and second AUG initiate translation (**Section 1.7.1**) (Wertz *et al.*, 1985; Satake *et al.*, 1985; Johnson and Collins, 1988a, Roberts *et al.*, 1993; Mallipeddi and Samal, 1993). The first AUG initiates a short ORF of 15 to 36 codons, depending on the strain of virus, which overlaps the translational start of the major ORF. Numerous other short ORFs exist in the RS virus mRNAs but are predominantly strain specific and not thought to encode functional polypeptides. However there is a second ORF in the 22K gene that is conserved between RS virus strains. This might encode a small protein significant for RS virus replication. The second ORF of the 22K protein gene overlaps the major ORF and encodes a polypeptide of 90 amino acids with 62% sequence identity between strains. The RS virus P mRNA does not contain an internal ORF for a cysteine-rich sequence that might be analogous to that of the V proteins of the other paramyxoviruses (Cattaneo *et al.*, 1989).

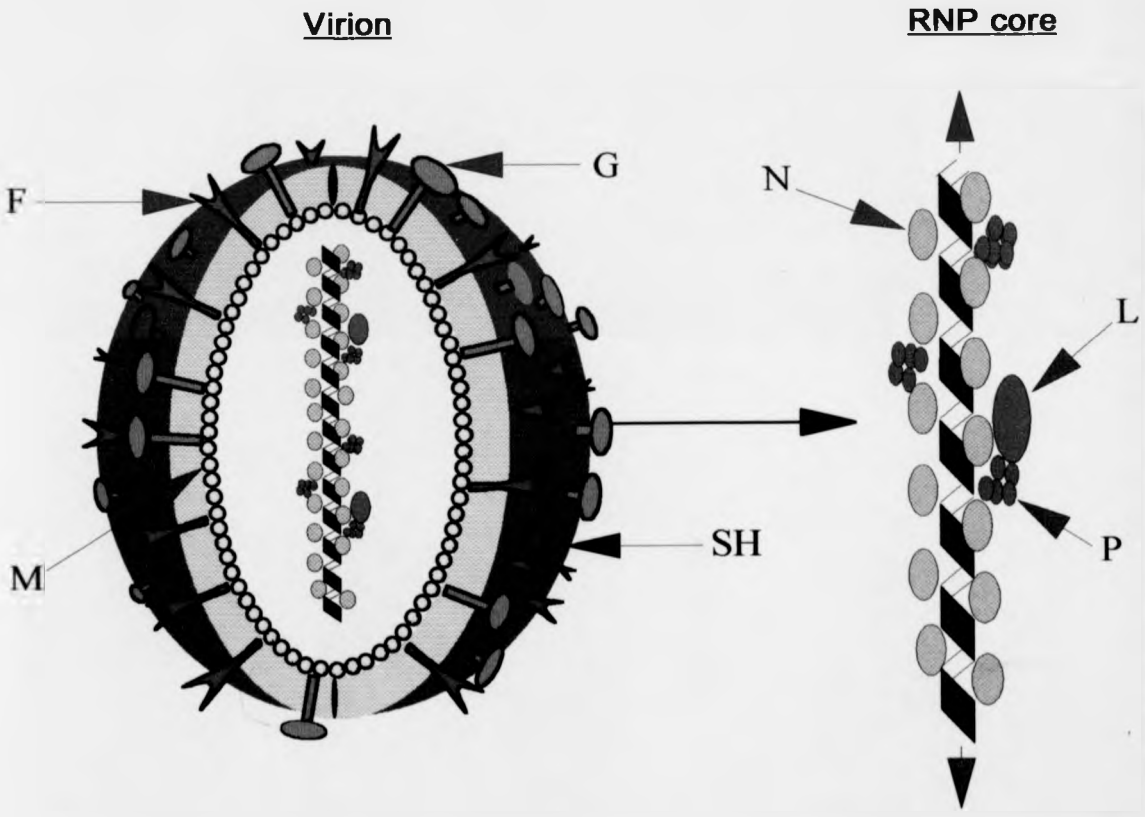
1.4 RS virion structure and functions The virion of RS virus consists of a ribonucleoprotein nucleocapsid contained within a lipoprotein envelope (**Figure 3**).

the 22K gene, so the polymerase must transcribe across the 22K gene-end signal in order to synthesise full-length L mRNA. The most abundant product of L gene transcription is a 68-nucleotide sequence, called L leader RNA, that initiates at the L gene-start and terminates at the 22K gene-end and contains a poly(A) tail. It possesses the first 20 codons of the L ORF and could encode a short protein (Collins, 1991). Full length L mRNA is less abundant and is generated from readthrough of the 22K gene-end signal, thus the presence of the gene-end signal within the L gene attenuates synthesis of full-length L mRNA.

1.3.4 RS virus mRNA translation For most mRNAs, the translational initiation codon is the first methionine codon, a context favoured for efficient translational initiation (Kozak, 1986). However the SH mRNA has an initiation codon in a unfavourable position (Collins and Wertz, 1985b; Johnson and Collins, 1988a) which may allow translation at the second AUG in the ORF, resulting in a truncated form of SH. The second exception is the G mRNA in which both the first and second AUG initiate translation (**Section 1.7.1**) (Wertz *et al.*, 1985; Satake *et al.*, 1985; Johnson and Collins, 1988a, Roberts *et al.*, 1993; Mallipeddi and Samal, 1993). The first AUG initiates a short ORF of 15 to 36 codons, depending on the strain of virus, which overlaps the translational start of the major ORF. Numerous other short ORFs exist in the RS virus mRNAs but are predominantly strain specific and not thought to encode functional polypeptides. However there is a second ORF in the 22K gene that is conserved between RS virus strains. This might encode a small protein significant for RS virus replication. The second ORF of the 22K protein gene overlaps the major ORF and encodes a polypeptide of 90 amino acids with 62% sequence identity between strains. The RS virus P mRNA does not contain an internal ORF for a cysteine-rich sequence that might be analogous to that of the V proteins of the other paramyxoviruses (Cattaneo *et al.*, 1989).

1.4 RS virion structure and functions The virion of RS virus consists of a ribonucleoprotein nucleocapsid contained within a lipoprotein envelope (**Figure 3**).

Figure 3. Diagrammatic representation of the RS virion.



Reproduced by kind permission Dr. J. Randhawa (1993).

Virions assemble at the plasma membrane of the infected cell and mature by budding, during which the intracellular nucleocapsid is packaged within a viral envelope. The envelope consists of a cellular membrane, derived by budding out of the host cell, into which are inserted the fusion (F) and attachment (G) glycoproteins and there is evidence to indicate that the SH protein is also inserted into the membrane (Huang *et al.*, 1985). Positioned under the membrane and associated with it, the surface glycoproteins, and the nucleocapsid, lies the matrix (M) protein which is thought to play a role in virion assembly and budding out of the cell (Bowen and Lyles, 1982). 22K protein appears to be located on the inner face of the envelope (Peeples and Levine, 1979; Huang *et al.*, 1985). The nucleocapsid consists of an RNA-nuclear protein (RNA-N) template and the RNA polymerase, consisting of L and P proteins (Colonno and Stone, 1976; Buetti and Choppin, 1977)(**Section 1.3.1**).

1.4.1 RS virus nucleocapsid proteins

1.4.2 The large (L) protein The L protein is presumed to be the major polymerase subunit responsible for viral transcription and replication (Banerjee, 1987a)(**Section 1.3.1**). Located furthest from the 3' polymerase binding site, mRNA species representing this gene are found in least amounts. The protein is basic and relatively hydrophobic, containing 29% hydrophobic residues, similar characteristics to the L proteins of the other paramyxoviruses. The gene is 6578 bases in size and encodes a protein 2165 residues long, with an estimated Mr of 250 kD (Stec *et al.*, 1991). Human RS and bovine RS viruses have the start of the L gene located within the 22K gene (Collins, 1991; Stec *et al.*, 1991; Zamora and Samal, 1992a)(**Section 1.2.1**), no comparable overlapping L gene ORF is found within the 22K gene of other pneumoviruses (Yu *et al.*, 1992a; Ling *et al.*, 1992; A.Easton, personal communication). The protein has low overall sequence homology to the other paramyxoviruses, with N-terminal and C-terminal regions being poorly related, but there are two conserved regions. The first region (residue #775 to #786) has 58% identity with rabies virus and 92% identity with the comparable regions of

the L proteins of SV, PIV3, MeV and NDV (Collins, 1991). The second region (residues #810 to #815) has 67% identity with rabies virus and 83% identity with SV, PIV3, MeV, NDV and VSV. By analogy with VSV, the pneumovirus L protein would be expected to possess equivalent enzymatic activities of mRNA capping, methylation and polyadenylation and the ability to act as a protein kinase (Hamaguchi *et al.*, 1983; Hunt *et al.*, 1984; Sanchez *et al.*, 1985).

1.4.3 The nucleoprotein (N) The function of the N protein is presumed to be that of a template upon which viral RNA polymerase can act (**Section 1.3.1**). This template is formed when the N protein binds to viral RNA. In VSV it would appear that the specificity of the N protein for viral RNA occurs if the viral NS (the equivalent of *paramyxovirus* genus P protein) and L proteins are present (Pattnaik and Wertz, 1990). Calain and Roux (1993) found that in SV the nucleocapsid needs to contact exactly six nucleotides to form a proper template for efficient replication. RS viral N protein is encoded from a gene 1203 bases in size, giving a protein 391 residues long with an estimated Mr of 43 kD. It differs from the paramyxovirus NP protein and resembles the VSV N protein, being rich in basic amino acids and not found in phosphorylated forms (Cash *et al.*, 1979; Collins *et al.*, 1985a; Lambert *et al.*, 1988). N protein is the most highly conserved RS virus protein between subgroups (96% identity; Johnson and Collins, 1989) however there is antigenic variation in a hydrophilic region, which is exposed on the surface of intact nucleocapsids (Ward *et al.*, 1984). Between two strains (A2 and RSS-2) of the same subgroup there is very high conservation of amino-acids, approximately 99% (C.Pringle, personal communication; Collins, 1991). Comparison between human RS virus N protein and PVM N protein shows 60% amino acid identity. Homology varies within the protein sequence with 38% identity in the amino-terminal third rising to 74% for the remaining amino acids, and in one specific domain (residues #245 to 315) this homology increases to 97% (Barr *et al.*, 1991). The degree of homology possibly accounts for the cross-reactivity observed between human RS virus and PVM N proteins (Gimenez *et al.*, 1984; Ling and Pringle, 1989a). Several structural features common to all

paramyxoviruses have been identified (Barr *et al.*, 1991) and may be relevant for the function of the protein. N protein can be cleaved by trypsin treatment into two fragments, 14 kD and 27 kD. The 27 kD fragment appears to be released or exposed because it can be degraded upon further trypsinisation (Ward *et al.*, 1984). This is similar to the NP protein of SV which has an N-terminal hydrophobic domain that is trypsin-resistant and a C-terminus that is sensitive and hydrophilic.

RS virus N protein accumulates in a cytoplasmic pool followed by association with intermediate filaments of the cytoskeleton (Garcia-Barreno *et al.*, 1988). This might indicate an association of nucleocapsids with the cytoskeleton which is utilised in the replicative cycle as a structural support or transport mechanism (Hamaguchi *et al.*, 1985; Bohn *et al.*, 1986). Lambert and Pons (1983) and Huang *et al.* (1984) observed an association between the P and viral nucleocapsids of human RS viruses, Garcia *et al.* (1993) have also demonstrated the presence of N-P complexes in human RS virus infected cells, and demonstrated that there was interaction between the N, and the P and 22K proteins. Interactions between the nucleoprotein and other viral proteins have been reported for PVM (J.Barr, 1994), SV (Lambert and Pons, 1983; Ryan and Kingsbury, 1988; Homann *et al.*, 1991), and in VSV (Masters and Banerjee, 1988a). N protein is also capable of interacting with itself to form aggregates in measles virus (Spehner *et al.*, 1991; Huber *et al.*, 1991; Fooks *et al.*, 1993) and SV (Buchholz *et al.*, 1993). In SV there appears to be a requirement for NP-P complexes (Horikami *et al.*, 1992) during replication. Investigating defective interfering (DI) viruses, they noted that a reduction or increase in expressed NP or P proteins inhibited replication of the DI's. Their results suggest that an imbalance in the N:P ratio switches polymerase function from replication to transcription.

1.4.4 The phosphoprotein (P) The P protein is acidic, hydrophilic and heavily phosphorylated (Cash *et al.*, 1979; Satake *et al.*, 1984; Lambden, 1985; Lambert, 1988). Navarro *et al.* (1991) mapped the sites of the phosphorylated residues in the RS virus P

protein to the central region of the molecule. The P gene is 914 nucleotides long, giving a polypeptide 241 amino acids long with an estimated Mr of 27 kD. RS virus P protein seems to share little homology with other P proteins but may have low sequence relatedness with a central domain of the P proteins of PIV3 and MeV (Spriggs and Collins, 1986a). Comparison of the P protein between the two subgroups of human RS virus shows that there are two conserved domains (residues #1 to #58 and #86 to #241) with a total of 96% identity, similar to the P proteins of SV and PIV3 which have conserved N- and C-terminal domains. Despite an overall 90% identity between strains of the two subgroups the P protein shows characteristic (apparent) mobility and antigenic differences between subgroups (Mufson *et al.*, 1985; Norrby *et al.*, 1986; Gimenez *et al.*, 1986; Morgan *et al.*, 1987; Akerlind *et al.*,1988; Johnson and Collins, 1990). Pneumovirus P proteins share some sequence homology, comparing human RS to bovine RS viruses (35.6%) and human RS virus to PVM (28%) with more conserved domains (approximately 64% homology) in some regions of the protein. In PVM there is a second, smaller ORF that has the capacity to encode a polypeptide 137 amino acids in length. The P protein genes of RS virus and TRTV do not possess a second open reading frame (Satake *et al.*, 1984; Lambden, 1985; A.Easton, personal communication) but the presence of a second ORF in PVM P means that P genes with multiple ORFs is not an exclusive property of the paramyxo-,rubula- and morbilliviruses.

The P protein has been shown to interact with the nuclear (N or NP) protein in several paramyxoviruses including human RS virus (Lambert and Pons, 1983; Huang *et al.*, 1984; Ryan and Kingsbury, 1988; Baker and Moyer, 1988; Huber *et al.*, 1991; Homman *et al.*, 1991; Garcia *et al.*, 1993). At least one region of the P protein of measles virus and the NS protein of VSV has been identified as being involved in nucleoprotein interaction, this region consists of approximately 40% of the C-terminal region (Gill *et al.*, 1986; Huber *et al.*, 1991; Spehner *et al.*, 1991). The N-terminal region of P is thought to be involved in RNA binding as demonstrated in SV (Curran *et al.*, 1991a), MeV (Huber *et al.*, 1991) and VSV (Gill *et al.*, 1986). The role of the P protein in transcription of human

RS viral genes has been demonstrated using a ts mutant, and implies a role for the P protein in providing specificity for the N-P complex to bind virus specific RNA (Caravokyri *et al.*, 1992), supported by evidence from VSV (Masters and Banerjee, 1988a and b). The pneumoviruses (PVM excepted) encode only one polypeptide from the ORF. The other paramyxoviruses are capable of directing the expression of several polypeptide products, either utilising other reading frames or switching between ORFs by RNA editing. The mechanism of RNA editing has been demonstrated in a number of viruses including SV5 (Thomas *et al.*, 1988), MeV (Catteneo *et al.*, 1989), SV (Vidal *et al.*, 1990), MuV (Paterson and Lamb, 1990; Takeuchi *et al.*, 1990), HPIV-2 (Ohgimoto *et al.*, 1990; Southern *et al.*, 1990), HPIV-4A and 4B (Kondo *et al.*, 1990), BPIV-3 (Pelet *et al.*, 1991), HPIV-3 (Galinski *et al.*, 1992), phocid distemper virus (Blixenkrone-Moller *et al.*, 1992), and NDV (Steward *et al.*, 1993). RNA editing activity results in the insertion of non-templated G nucleotides at specific sites during transcription of the P gene which causes a frame-shift, thus allowing access to alternative reading frames. In addition to the mechanism of RNA editing, many of these viruses contain alternative cistrons which are accessed by independent ribosomal initiation. The P gene of SV has been particularly well studied and has been found to give rise to 8 polypeptides (P, V, C, C', W, X, Y1 and Y2) by using both expression strategies (Giorgi *et al.*, 1983; Shioda *et al.*, 1983; Curran *et al.*, 1991a and b). Hendricks *et al.* (1993) investigating the C protein of SV concluded that differential phosphorylation and the potential for modulation of phosphorylation suggested a regulatory role for the C protein. Three functional domains have been defined. The N-terminal domain I is highly acidic, with domains II and III being basic and thought to be involved in binding to the L protein and the NP-RNA template. The model favoured for P protein activation is the phosphorylation cascade pathway proposed by Barik and Banerjee, (1991). Unphosphorylated P protein (P_0) is converted to a phosphorylated (P_1) species by a casein kinase II -like cellular enzyme at domains I (Barik and Banerjee, 1992) and domain III (Villanueva *et al.*, 1994).

1.5 The membrane-associated proteins

1.5.1 The matrix (M) protein (or M1) The M protein is a nonglycosylated inner component of the viral envelope (Peeples and Levine, 1979; Collins *et al.*, 1984a; Satake and Venkatesan, 1984; Huang *et al*, 1985). The protein is encoded by a gene which is 958 nucleotides long and generates a protein of 256 amino acids and an estimated Mr of 28 kD. The M protein appears to be highly conserved in amino-acid content within subgroup A isolates (C.Pringle, personal communication). The M protein is highly basic and occurs as multiple electrophoretically-distinct species that differ in intramolecular disulphide bonding (Gruber and Levine, 1983). M protein also undergoes phosphorylation involving a subset of intracellular proteins (Lambert *et al.*, 1988). Since the M protein of the pneumoviruses is smaller than that of the morbilli- and paramyxoviruses, the functions of human RS virus M may possibly be split between the M and 22K proteins. The M proteins of the *Morbillivirus* and *Paramyxovirus* genera contain a conserved hydrophobic segment in the C-terminal third that is thought to be involved in membrane interaction (Satake and Venkatesan, 1984; Bellini *et al*, 1986). RS virus contains a similar region (residue # 192 to #217) that is thought to have a similar role (Collins, 1991), however there is little sequence relatedness with the M protein of NDV except for a short region of apparent relatedness (residues # 150 to #220 in RS virus and residues #120 to #190 in NDV) (Chambers *et al*, 1986). The M proteins of many members of the family *Paramyxoviridae* are believed to be involved in functions relating to maturation and assembly of the virion. This theory is supported by studies indicating that the M protein of members of the family *Paramyxoviridae* is able to interact with the nucleocapsid structure and also itself (Peeples, 1991; Homann *et al.*, 1991; Barge *et al.*, 1993). In addition, M protein is able to interact with a variety of cellular components such as membrane lipids and cytoskeletal elements (Caldwell and Lyles, 1986; Faaberg and Peeples, 1988; Tashiro *et al.*, 1993; Sanderson *et al.*, 1993). Interaction between M and F proteins of NDV, and M with F or HN of SV, has also been shown (Peeples and Bratt, 1984; Sanderson *et al.*, 1994). Analysis of ts mutants in the M protein of human RS virus have shown that viral assembly is inhibited at the non-permissive temperature by increasing degradation of the M protein (Caravokyri and Pringle, 1991). Thus it is thought

that the M protein possibly interacts with F protein which protects the matrix protein from degradation by cellular proteases. Additionally, by analogy with VSV and SV, the matrix protein of human RS virus may inhibit host cell and viral transcription by structural hindrance of the polymerase as it attempts to pass along the template (Marx *et al.*, 1974; De *et al.*, 1982; Ogden *et al.*, 1986; Black and Lyles, 1992; Black *et al.*, 1993; Li *et al.*, 1993). The ability of VSV M protein to independently bind to membranes (Li *et al.*, 1993) differs from the situation in SV where the M protein was found to be unable to associate with membranes when expressed on its own (Pattnaik and Wertz, 1991; Sanderson *et al.*, 1993). In either paramyxovirus, rubulavirus or morbillivirus systems, demonstration that a definitive physical relationship exists between M and H/N (or H) proteins has not yet been reported, although the observation that the cellular location of membrane bound M protein appears to be determined by the location of the SV glycoproteins suggests that these protein species do associate (Sanderson *et al.*, 1993).

1.5.2 The 22K protein (or M2) The pneumoviruses are unique in possessing a second non-glycosylated membrane-associated protein, 22K. The 22K protein is 194 residues long with an estimated Mr of 22 kD, encoded by a gene 961 nucleotides long (slightly larger than that found in PVM and TRTV). The amino acid sequence does not contain any region of sufficient hydrophobicity to suggest membrane insertion (Collins and Wertz, 1985c; Elango *et al.*, 1985b) but 22K is the most basic RS viral protein (Dubovi, 1982; Huang *et al.*, 1985; Collins and Wertz, 1985c) and contains three moderately hydrophobic regions which could be involved in membrane interactions (Collins and Wertz, 1985c). In infected cells this protein is resistant to trypsinisation suggesting an internal location (Routledge *et al.*, 1987a). Studies have shown 22K to be associated with internal structures of human RS virus such as the N and P proteins (Routledge *et al.*, 1987a; Garcia *et al.*, 1993; Samal *et al.*, 1993). 22K protein appears to co-localise with the N and P proteins in nucleocapsids whereas M protein appears to co-localise with F and G at the cell membrane. This would suggest a possible role for the 22K protein of indirectly linking

the N and M proteins. The kinetics of 22K protein phosphorylation appear to be similar to that of the P protein and distinct from the M protein, and seem to co-ordinate with virion assembly and release (Lambert *et al.*, 1988). Thus it may be that the 22K protein associates with nucleocapsids at an early point in virion assembly, whereas association with M might be at a later point that immediately precedes budding. The 22K protein exhibits strain-specific variations in apparent mobility and subgroup-specific antigenic differences (Norrby *et al.*, 1986; Morgan *et al.*, 1987; Routledge *et al.*, 1987b). There appear to be conformational variants of 22K possibly due to different disulphide-bond formation (Routledge *et al.*, 1987a, b). Human RS viral 22K protein shows considerable amino acid identity to the 22K proteins of PVM, bovine RS virus and TRTV, and these three polypeptides have similar hydrophobicity profiles which indicates a degree of structural conservation over regions where amino acid identity is not strong (Collins and Wertz., 1985c; Ling *et al.*, 1992; Yu *et al.*, 1992a; Zamora and Samal, 1992b).

1.5.3 The small hydrophobic (SH) protein (or 1A) The SH gene is 410 nucleotides long and encodes a protein of 64 residues. There are four electrophoretically distinct species with an estimated Mr of 4.8, 7.5, 13 to 15, and 21 to 30 kD. (Olmsted and Collins, 1989). The Mr 7.5 kD species appears to be the full-length, unglycosylated, unprocessed form. The Mr 4.8 kD is also unglycosylated and lacks the N-terminal amino acids, due to ribosomes initiating at a second AUG methionine codon (met-23 in the amino acid sequence)**(Section 1.3.4).** The other two major forms of the SH protein are glycosylated (Olmsted and Collins, 1989) and derived from the Mr 7.5 kD species. Most SH accumulates intracellularly as unglycosylated Mr 4.8 kD and Mr 7.5 kD species with the process being independent of other viral proteins (Olmsted and Collins, 1989). Glycosylation of SH is thought to be a late event, occurring in or beyond the trans-Golgi compartment (Collins and Mottet, 1993). All four forms of SH protein appear to be integral membrane proteins with greater stability in membrane association than that found in the M and 22K proteins (Olmsted and Collins, 1989). SH protein is expressed at high

level at the cell surface but does not appear to be secreted (Olmsted and Collins, 1989; Collins and Mottet, 1993). Collins and Mottet (1993) have demonstrated that SH assembles into homo-oligomers that co-sediment with the F protein tetramer. The C-terminus of the Mr 7.5 kD species is trypsin-sensitive and thus extracellular with the N-terminus presumably cytoplasmic (Collins and Mottet, 1993). The Mr 7.5 kD and 30 kD species are probable structural components of the virion (Huang *et al.*, 1985). The SH protein shares 76% homology between the two subgroups, with the two potential glycosylation sites and the second methionine codon at position 23 being conserved. Within subgroups, SH shows 84% homology, the cytoplasmic and transmembrane regions (residues #1 to 41) are highly conserved (89%) but the extracellular domain less so (43%), possibly indicative of response to immunological pressure (Collins, 1991; Collins and Mottet, 1993). Thus the transmembrane and cytoplasmic domains may have functional importance, with the extracellular domain acting as an anchor to fix SH in the membrane. A small hydrophobic protein is also found in other paramyxoviruses such as SV5 and MuV (44 amino acids and 57 amino acids respectively)(Collins and Wertz, 1985b; Hiebert *et al.*, 1985b). There is also a small integral membrane protein in the influenza viruses types A and B (the 97-amino acid M2 protein and the 100-amino acid NB protein)(Williams and Lamb, 1986; Zebedee and Lamb, 1988). These other viral proteins are also inserted in the membrane with the C-terminus orientated extracellularly. One hypothesis of SH function is that it forms ion channels, as does the M2 protein of Influenza A (Pinto *et al.*, 1992). Experiments reported by Hemingway *et al.* (IXth International Congress of Virology, 1993) suggest that SH plays a critical role in the formation of syncytia in an, as of yet, undefined manner.

The nucleotide sequence of the RS virus SH protein gene has revealed the presence of four potential N-linked glycosylation sites, with two located at the amino-terminal end and two located at the carboxyl-terminal end of the protein. The predicted amino acid sequence of the TRTV SH protein has revealed the presence of a potential N-

linked glycosylation site and this site is thought to be used (Ling *et al.*, 1992). An additional species with a molecular weight of 180 kD was also identified and believed to represent an oligomer composed of four or five cross-linked pentameric subunits. Cane and Pringle, (1991) analysed the SH gene sequences of twelve RS virus isolates and described 76% deduced amino acid identity between subgroups A and B, but little variation in deduced amino acid homology within the subgroups (93% to 99%). They then used restriction enzyme analysis of the SH gene sequence of forty-two isolates obtained from a single epidemic season (autumn/winter 1989) and identified at least six different isolate-types of RS virus circulating at the same time in the same geographical region. Five of these isolate-types (termed SH lineage, abbreviated to SHL) were shown to belong to the subgroup A class of RS virus.

1.6 The non-structural proteins

1.6.1 The 1C protein (or NS1) The NS1 protein is 139 amino acids in length and is encoded by a 532 nucleotide gene located at the 3' end of the genome and is the first gene to be transcribed. As a consequence of the 3' leader proximal position of the NS1 gene, NS1 is the most abundant protein in infected cells. However, it has not been detected in the virion. The protein has a net charge of -1 at neutral pH and an approximate Mr of 15.6 kD. However on SDS-PAGE the NS1 protein has an approximate size of Mr 13.6 kD (J.Evans, personal communication). There appears to be the same degree of protein sequence conservation both within and between subgroups as is found in the NS2 sequence. An association with the M protein has been observed, suggesting that NS1 may be involved in maturation (J.Evans, personal communication). PVM NS1 gene has a second open reading frame which is accessed by the third AUG codon from the 5' end of the mRNA, and is capable of encoding a polypeptide 69 amino acids in length. In addition, the first AUG codon of the PVM NS1 gene is followed by a second initiation codon only nine nucleotides downstream. Neither of these AUGs are in good context for initiation (Kozak, 1986) but it is possible that they are both used since two closely migrating forms of the

NS1 polypeptide are seen in PVM-infected BS-C-1 cells (Ling and Pringle, 1989b). Perhaps the poor initiation context for these two in-frame AUG codons allows scanning ribosomes to access of the second ORF.

1.6.2 The 1B protein (or NS2) The NS2 gene lies at the 3' end of the genome downstream of the first gene, NS1. The protein has a net charge of +6 at neutral pH and has an estimated Mr of 14.7 kD, and is 124 amino acid residues in length. It is encoded by a gene of 503 nucleotides. The NS2 protein is of basic nature and is present in high concentrations in the infected cell, and whilst there is no direct evidence showing that it is post-translationally modified, it does appear to have a short half-life and appears to undergo conformational change soon after synthesis (J.Evans, personal communication). The function of NS2 is not known but it appears to be absent from the virion and this, combined with its basic nature, suggests that it could be involved in RNA interactions (Collins and Wertz, 1985d). There is very high sequence conservation (approximately 98%) of the NS2 sequence within subgroup A isolates (J.Evans, personal communication). Between subgroups there is slightly less conservation, with amino acid homology approximately 87% (Collins, 1991).

1.7 The major surface glycoproteins

1.7.1 The attachment (G) protein The G protein is one of two main glycoproteins found on the surface of the virion (Walsh *et al.*, 1984b; Levine *et al.*, 1987). The G protein gene is 923 nucleotides long and encodes a protein of 298 or 292 amino acids depending on subgroup (Wertz *et al.*, 1985; Satake *et al.*, 1985; Johnson *et al.*, 1987b; Sullender *et al.*, 1990, 1991; Sullender and Wertz, 1991) with an estimated Mr of 84 kD to 90 kD. The G protein of PVM is slightly larger (363 amino acids), possibly as a result of the acquisition of the region responsible for haemagglutination. Two forms of human RS virus G have been observed, one is associated with the virion, the other a soluble form (G_s), lacking the first 65 or 74 amino acids which includes the membrane anchor region

(Hendricks *et al.*, 1988). The G_S protein is synthesised from an internal translation initiation site on the same mRNA as the membrane-inserted G protein and appears to be cleaved downstream of the AUG. Cleavage is suggested to remove the membrane-anchor region, thus preventing retention of the G protein in the cellular membrane and therefore allowing secretion to occur (Roberts *et al.*, IXth International Congress of Virology, 1993; Mallipeddi and Samal, 1993).

Full length attachment protein lacks detectable sequence relatedness with its counterparts in the other paramyxoviruses, the HN and the H proteins (Spriggs *et al.*, 1986c),and lacks haemagglutination and neuraminidase activities (Richman *et al.*, 1971). The first initiation codon is followed soon after by a termination codon and this arrangement of initiation codons is likely to down-regulate expression of G protein as scanning ribosomal 40s subunits will probably initiate at the 5' proximal AUG. Interestingly, the first AUG of the TRTV G gene is not followed soon after by a termination codon, and thus expression of TRTV G may not be down-regulated. However down-regulation of TRTV G protein expression relative to other TRTV proteins may occur as a consequence of the position of the G gene on the TRTV genome, since in TRTV the G gene is further from the 3' proximal promoter region (Yu *et al.*, 1992b) and therefore less frequently transcribed. The amino acid sequence of the protein contains a single major hydrophobic domain near the N-terminus (residues #38 to #66). This region has all the structural information required for membrane insertion and anchoring (Hendricks *et al.*, 1987, 1988; Olmsted *et al.*, 1989; Lichtenstein *et al.*, 1991). In common with the other paramyxovirus attachment proteins, RS virus G protein is orientated in the membrane with the N terminus on the cytoplasmic side and the C-terminus extracellular. G protein is heavily modified intracellularly (approximately 60%) by the addition of both N-linked and O-linked carbohydrate, giving a mature form of G with a much larger Mr than is predicted (Collins *et al.*, 1984a; Wertz *et al.*, 1985; Hiebert *et al.*, 1985a; Satake *et al.*, 1985), the presence of O-linked glycosylations being unique to the pneumoviruses in the *Paramyxoviridae* (Collins, 1991). The number and location of N-linked glycosylation sites

in the G protein are not conserved within or between subgroups (Johnson *et al.*, 1987b) and the situation is similar for O-linked glycosylation sites. However these do appear to fall into two clusters, one near to the membrane anchor region and the other near to the C-terminus, with O-linked sugars divided equally between the two clusters (Olmsted *et al.*, 1989). The ectodomain of the G protein, unlike the membrane and intracellular domains, is rich in proline residues which disrupt secondary structure and are thought to contribute to β-turns (Chou and Fasman, 1978). There is also a high degree of serine and threonine residues in the ectodomain, which in other proteins have been shown to contain sites of attachment of O-linked sugars when present in proline-rich regions (Hill *et al.*, 1977; Fiat *et al.*, 1980; Eckhardt *et al.*, 1987). Thus the extracellular structure of G might be similar to that of mucin proteins which share these features with the G protein (Collins, 1991). The high serine and threonine content is one of the few conserved features between the G proteins of the two subgroups. A second feature is the conservation of the four cysteine residues (Alansari and Potgieter, 1993).

Oligomerisation of the G protein can occur as both homo- and hetero-oligomerisation. Lambert (1988) described an Mr 175 kD protein consisting of two disulphide-bonded G proteins being present on SDS-PAGE under nonreducing conditions, and described a purified fraction of G protein which was disulphide bonded to F protein. Close association of the attachment and fusion glycoproteins has also been observed in MeV (Malvoisin and Wild, 1993). Oligomerisation, O-glycosylation and incorporation of G into virions can occur in the absence of N-linked sugars, and even fully unglycosylated G is incorporated into virions (Gruber and Levine, 1985a; Fernie *et al.*, 1985; Satake *et al.*, 1985; Lambert, 1988; Wertz *et al.*, 1989). Also the absence of either type of sugar has a modest effect on maturation. However absence of all glycosylation results in a 10-fold reduction in G protein surface expression (Wertz *et al.*, 1989) thus it would appear that both types of sugar side-chain contribute in some undefined way to G protein maturation.

1.7.2 G protein glycosylation The method of post-translational processing of the G protein remains unclear, but it appears that a glycosylated species (Mr 45 kD) is a major intermediate of the mature G protein (Gruber and Levine, 1985a, b; Fernie *et al.*, 1985; Hendricks *et al.*, 1987, 1988; Wertz *et al*, 1989). This species is predominantly N-linked glycosylated and appears to accumulate prior to migration to the trans-Golgi compartment where most of the O-linked sugars are added (Collins and Mottet, 1992). The intracellular maturation of G is relatively rapid , with the half-time for mature G formation being approximately 15 to 30 minutes (Fernie *et al.*, 1985; Gruber and Levine, 1985b). Human RS viral G protein may be glycosylated to a greater degree than PVM G protein, with carbohydrate increasing the G protein molecular weight by approximately 50 kD compared with 20 kD (Ling *et al.*, 1992). The large molecular weight difference attributed to O-linked sugars may however be an artefact due to reduced SDS binding on account of the O-linked sugars (Collins and Mottet, 1992). Investigations by several groups have determined that the N-linked sugars are added co-translationally whereas the O-linked sugars are added post-translationally (Fernie *et al.*, 1985; Wertz *et al.*, 1989).

1.7.3 G protein variation Akerlind *et al.*(1988) attributed variation in SDS-PAGE mobility of human RS virus G proteins to strain differences in glycosylation. Sullender *et al.* (1991) and Cane *et al.* (1991) observed extensive sequence variation among naturally occurring isolates. Analysis of the carboxy-terminal third of the G protein gene of subgroup A isolates revealed a region prone to polymerase errors resulting in the insertion or deletion of adenosine residues (Cane *et al.*, 1993). Sequence divergence in the G proteins of the two subgroups is localised to two large segments (residues #70 to #163 and residues #177 to 298) that comprise most of the extracellular domain, however the overall hydropathic profile remains similar. In contrast the membrane anchor and cytoplasmic domains are conserved between subgroups (84% homology) but no sequences critical for G protein synthesis and processing have been identified. The ectodomain contains a short segment of exact sequence identity between subgroups

(residues #164 to #176), however this region only shows 50% conservation between bovine and ovine G proteins (Alansari and Potgieter, 1993). This region is proposed as a candidate for attachment to cellular receptors (Johnson *et al.*, 1987a). The fact that it is not conserved between bovine, ovine and human RS virus suggests that it could relate to host specificity (Lerch *et al.*, 1991). The cysteine residues are clustered in an amino acid grouping, 14 residues long, that overlaps the conserved sequence. This is the most hydrophobic region of the ectodomain and has fewer potential glycosylation acceptor sites. The region is also flanked by hydrophilic regions high in potential carbohydrate acceptor sites. Secondary structure predictions define this region as a tight turn between two β-sheets and this sort of structure could be stabilised by intramolecular disulphide bonds which would require the conservation of the cysteine residues (Collins, 1991). The cysteine-rich region of the ectodomain is relatively conserved between subgroups so is thought not to be a factor in the inefficiency of cross-subgroup protection; *i.e.* that it is weakly antigenic. Escape mutants raised against Mab c793 which recognises a conserved epitope shared by all human isolates (Mufson *et al.*, 1985), were found to have lost one of the four conserved cysteine residues (cysteine #182 or #186). Thus losing most of the conserved and subgroup-specific epitopes, whilst retaining strain-variable epitopes (Rueda *et al.*, 1994). The adjacent hydrophilic domains probably would be antigenic but structural heterogeneity could reduce the efficiency of cross-strain neutralisation. This could be brought about by partial utilisation of O-glycosylation sites which could result in an antigenically diverse population of molecules even within a genetically homogeneous virus population (Collins, 1991). Epitope mapping using synthetic peptides has shown that these flanking hydrophilic domains (amino acids #134 to #158 and #174 to #208) are indeed the most reactive peptides in binding to convalescent and monoclonal antibodies (Norrby *et al.*, 1987). Experiments by Palomo *et al.* (1991) ; Garcia-Barreno *et al.* (1989) have demonstrated the crucial role of carbohydrates for evasion of host immune responses. Major changes in the carbohydrate content could occur by frame shift mutations in the G gene. Characterisation of human RS viral escape mutants against a neutralising anti-G monoclonal antibody has produced evidence that most escape mutants do indeed arise

because of frame shift mutations affecting the terminal third of the molecule (Garcia-Barreno *et al.*, 1990; Rueda *et al.*, 1991). Frame-shift mutations have also been observed in isolates of subgroup B (Sullender *et al.*, 1991). Thus frame shift mutations in the G gene could potentially alter the carbohydrate content and lead to changes in antigenicity.

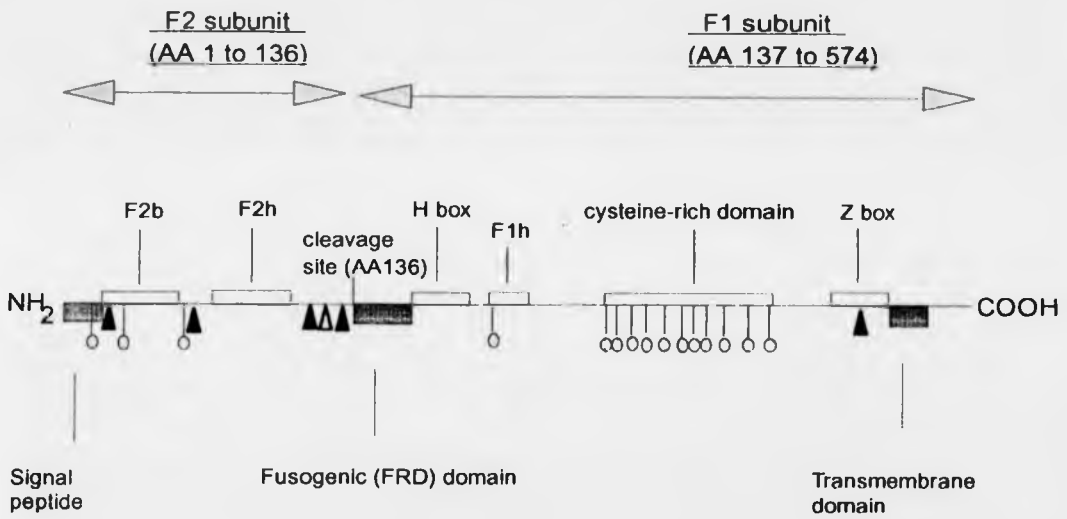
1.7.4 The fusion (F) protein The fusion protein is the second major surface glycoprotein, forming a membrane-associated spike which is involved in fusion of the virion to the host cell, and is also involved in syncytium formation by fusing neighbouring cells together in cell-cell spread of the virus. However the fusion protein when expressed separately does not initiate cellular fusion and syncytium formation efficiently and it appears that other viral proteins are involved. *In vitro* expression experiments suggest that the attachment (G) and SH proteins are required (Heminway *et al.*, 1994) whereas in MuV, MeV and bovine parainfluenza virus 3, the neuraminidase enzyme of the HN glycoprotein is involved in modulating cell fusion cytopathology (Waxham and Wolinsky, 1986; Morrison *et al.*, 1991; Horvath and Lamb, 1992; Tanabayashi *et al.*, 1992). The F gene is 1901 nucleotides long and encodes a protein of 574 amino acids with an estimated Mr 68 kD to 70 kD, from a single major ORF which is 1722 nucleotides long. The F protein of human RS virus is 40% homologous to the fusion proteins of the other pneumoviruses (Chambers *et al.*, 1992). Human RS virus F protein has general structural similarities with the fusion proteins of other paramyxoviruses, which form a more closely-related group (Collins *et al.*, 1984b; Elango *et al.*, 1985a; Spriggs *et al.*, 1986c; Barrett *et al.*, 1987) but only has low sequence homology (<20%). These similarities include overall length, the location of hydrophobic domains and cleavage site and the approximate locations of carbohydrate side chains and cysteine residues.

1.7.5 F protein structure Nucleotide sequence analyses has revealed the presence in the inferred protein product of various secondary structural elements that were well conserved between the F proteins of all morbilliviruses, rubulaviruses and paramyxoviruses (Chambers *et al.*, 1992). Features common to the fusion proteins of all paramyxoviruses

are illustrated in the generalised RS virus F protein (**Figure 4**). The F protein contains an N-terminal hydrophobic signal peptide that directs cotranslational translocation into the rough endoplasmic reticulum, and which is cleaved in the vicinity of amino acids #21 to #25. The most hydrophobic domain is located near the C-terminus (amino acids #525 to #550) and serves as a stop-transfer signal and membrane anchor, leaving a 24-amino acid cytoplasmic domain. The F protein is synthesised as a precursor, F_0, transported to the Golgi apparatus via the endoplasmic reticulum during which the signal region is removed, and then cleaved by a cellular protease to generate two disulphide-linked subunits, NH_2-F_2-ss-F_1-COOH (Fernie and Gerin, 1982; Gruber and Levine, 1983; Huang *et al.*, 1985; Gruber and Levine, 1985b; Fernie *et al.*, 1985). This feature is characteristic of all paramyxoviruses (Choppin and Scheid, 1980). In Sendai virus the disulphide linkage is between the single cysteine residue in the F_2 subunit (residue #70) and residue #199, the most upstream of the cysteine amino acids (Iwata *et al.*, 1994). By analogy with the Sendai virus model, disulphide linking of the two subunits of the RS virus fusion protein is thought to involve either/or both cysteine residues in the C-terminal domain of the F_2 subunit and the most upstream cysteine (residue #212) of the F_1 subunit. Proteolytic cleavage occurs following amino acid #136 in human RS virus (Elango *et al.*, 1985a). Intracellular cleavage occurs before transport to the trans-Golgi region (Gruber and Levine, 1985b; Collins and Mottet, 1991). Like the other paramyxoviruses, cleavage is obligatory for fusion activity and is potentially a factor in tissue tropism. Investigations into the mechanism responsible for cleavage determined that amino acid changes in the F_1 subunit (located in the cysteine-rich region or thereabouts) caused loss of F_0 cleavage, thus alteration of one region of the F protein can influence changes in another that is spatially separate, possibly by affecting protein folding (Anderson *et al.*, 1992). Investigations by Buckland *et al.* (1992) determined that amino acid changes in the C-terminal region of the F_1 subunit (in a putative "leucine-zipper" region located adjacent to the membrane-anchor) could interfere with the fusion function. The cleavage site of RS virus F is highly conserved, with six arginine and lysine residues present, compared to five, four, three and one for SV5, MeV, MuV, PIV3 and SV (Galinski and Wechsler, 1991). In

are illustrated in the generalised RS virus F protein (**Figure 4**). The F protein contains an N-terminal hydrophobic signal peptide that directs cotranslational translocation into the rough endoplasmic reticulum, and which is cleaved in the vicinity of amino acids #21 to #25. The most hydrophobic domain is located near the C-terminus (amino acids #525 to #550) and serves as a stop-transfer signal and membrane anchor, leaving a 24-amino acid cytoplasmic domain. The F protein is synthesised as a precursor, F_0, transported to the Golgi apparatus via the endoplasmic reticulum during which the signal region is removed, and then cleaved by a cellular protease to generate two disulphide-linked subunits, NH_2-F_2-ss-F_1-COOH (Fernie and Gerin, 1982; Gruber and Levine, 1983; Huang *et al.*, 1985; Gruber and Levine, 1985b; Fernie *et al.*, 1985). This feature is characteristic of all paramyxoviruses (Choppin and Scheid, 1980). In Sendai virus the disulphide linkage is between the single cysteine residue in the F_2 subunit (residue #70) and residue #199, the most upstream of the cysteine amino acids (Iwata *et al.*, 1994). By analogy with the Sendai virus model, disulphide linking of the two subunits of the RS virus fusion protein is thought to involve either/or both cysteine residues in the C-terminal domain of the F_2 subunit and the most upstream cysteine (residue #212) of the F_1 subunit. Proteolytic cleavage occurs following amino acid #136 in human RS virus (Elango *et al.*, 1985a). Intracellular cleavage occurs before transport to the trans-Golgi region (Gruber and Levine, 1985b; Collins and Mottet, 1991). Like the other paramyxoviruses, cleavage is obligatory for fusion activity and is potentially a factor in tissue tropism. Investigations into the mechanism responsible for cleavage determined that amino acid changes in the F_1 subunit (located in the cysteine-rich region or thereabouts) caused loss of F_0 cleavage, thus alteration of one region of the F protein can influence changes in another that is spatially separate, possibly by affecting protein folding (Anderson *et al.*, 1992). Investigations by Buckland *et al.* (1992) determined that amino acid changes in the C-terminal region of the F_1 subunit (in a putative "leucine-zipper" region located adjacent to the membrane-anchor) could interfere with the fusion function. The cleavage site of RS virus F is highly conserved, with six arginine and lysine residues present, compared to five, four, three and one for SV5, MeV, MuV, PIV3 and SV (Galinski and Wechsler, 1991). In

Figure 4. General features of the RS virus fusion protein.



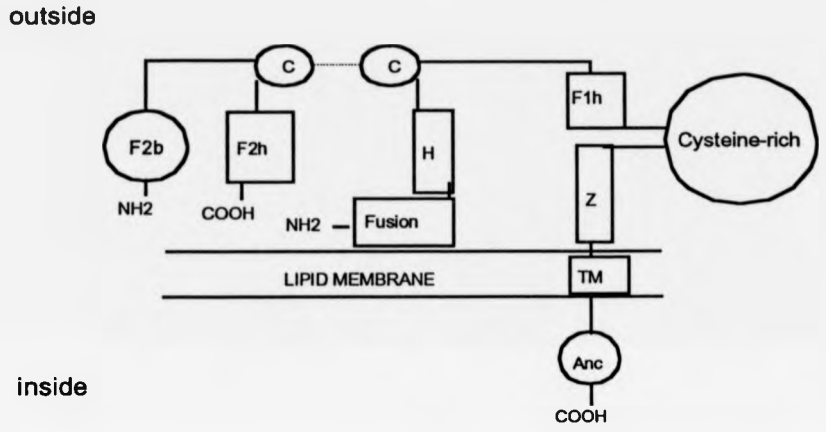
Diagramatical representation of the generalised RS virus fusion protein. Potential N-glycosylation sites are denoted by ▲ and the variable N-glycosylation site denoted by ▲. Cysteine residues are represented by ○. The precise cysteine residues involved in disulphide-bonding between subunits are unknown and so are not shown. The three hydrophobic domains are highlighted in shaded boxes, whilst important structural domains are highlighted in unshaded boxes.

SV F protein, cleavage of the F_o precursor results in a major conformational change (Hsu *et al.*, 1981), as it does also in NDV, where alteration in the cleavage site has an effect on pathogenicity (Toyoda *et al.*, 1987; Glickman *et al.*, 1988). In RS virus F protein, however, uncleaved F_o is antigenically similar to the cleaved form F_2-F_1. Thus perhaps the conformational change is localised to the F_2-F_1 junction and does not alter the major antigenic sites (Collins, 1991). The N-terminus of F_1, created by cleavage, is hydrophobic and thought to initiate fusion by directly interacting with the host cell membrane. The sequence of the F_1 N-terminus is highly conserved among viruses of the *Morbillivirus* and *Paramyxovirus* genera but the corresponding sequence in RS virus is distinct. The first three residues of the F_1 subunit of all paramyxoviruses are conserved: Phe-X-Gly (Collins *et al.*, 1984b; Elango *et al.*, 1985a). However, in some NDV strains the terminal amino acid is Leu not Phe, indicating that even this triplet is not obligatory. Also present in the N-terminus of F_1 is a semi-conserved sequence, Gly-Val-Ala, located just downstream of the N-terminus in RS virus and present slightly upstream in the other paramyxoviruses. However this motif is not exactly conserved within RS virus strains, being Gly-Ile-Ala (strains RSS-2 and 18537; subgroups A and B respectively) and Gly-Thr-Ala [(strain Long; subgroup A)(Baybutt and Pringle, 1987; Johnson and Collins, 1988b; Lopez *et al.*, 1988)]. The lack of sequence identity between RS viruses and the other paramyxoviruses in this region of F_1 may indicate that there is a general requirement only for hydrophobicity and that there is an extended unordered or helical structure at the N-terminus of F_1. There are five or six potential acceptor sites for N-linked carbohydrate within the F protein of RS virus, with estimates of carbohydrate content suggesting that at least five are utilised (Collins *et al.*, 1984b). All but one of these sites occur within the F_2 subunit. This suggests that the F_2 subunit is weakly antigenic due to either shielding or alteration of antigenic sites (Skehel *et al.*, 1984; Elder *et al.*, 1986; Sjoblom *et al.*, 1987) and could explain the observation of differences in immunogenicity of F_2 compared to F_1 (Walsh *et al.*, 1985). The carbohydrate side chains are not required for cleavage of F_o or the incorporation of F protein into virions (Huang *et al.*, 1983; Lambert, 1988). The fusion proteins of the pneumoviruses are unusual in that their F_2 subunits have two or

three cysteine residues depending on strain, whereas those of morbilli- and paramyxoviruses only have one. It is possible that both of these cysteines could be involved in disulphide bond formation with the F_1 subunit. Another common feature of paramyxovirus F proteins is a cysteine-rich region in the F_1 subunit (Spriggs *et al.*, 1986c; Barrett *et al.*, 1987). In the viruses of the *Morbillivirus* and *Paramyxovirus* genera, a 100-amino acid domain, which begins approximately 220 residues from the N-terminus of F_1 and ends approximately 75 residues before the membrane anchor, contains eight cysteine residues that are exactly conserved in number and spacing (Galinski and Wechsler, 1991; Yu *et al.*, 1991). In RS virus there is an analogous domain that is somewhat distinct: it is longer (127 amino acids), begins 176 residues from the N-terminus and ends 85 residues before the membrane anchor. It contains 11 cysteine residues and these differ in sequence position from the other paramyxoviruses thus giving a different pattern of disulphide bonding. However the number and position of cysteine residues is conserved between and within subgroups. In Sendai virus the cysteine-rich domain is thought to form internal disulphide pairings, giving rise to a bunched secondary structure containing at least two tandem cysteine loops (Iwata *et al.*, 1994). The region between the proposed loops in the Sendai virus fusion protein is hydrophilic, suggesting that the bunched structure is exposed to the molecular surface. This is compatible with the supposition that the cysteine-rich domain is recognised by monoclonal antibodies. The situation is thought to be similar for the RS virus fusion protein, and epitope mapping has identified a major fusion-inhibiting, neutralising epitope to this region of the fusion protein (**Section 1.7.8**).

The secondary structure of a generalised fusion protein for all paramyxoviruses has been proposed by Chambers *et al.* (1992) and this is shown diagramatically in **Figure 5**. The N-terminus of the F_2 subunit is separated from the C-terminal helical region (F_2h) by a region predicted to form a loop structure. The N-terminus of the F_1 subunit (the fusion-related domain) is thought to be involved in the fusogenic activity of the protein. There is a sequence of heptad repeats (H) (Chambers *et al.*, 1990b) next to the fusion-related

Figure 5. Secondary structure of a generalised paramyxovirus fusion protein



Suggested arrangement of potential a-helices in paramyxovirus F proteins. Proposed a-helical regions are shown as rectangles with B-sheet and turn-rich regions as circles. The F2-subunit contains one a-helical region (F2h) and one B-sheet region (F2b). The F1-subunit contains several a-helical regions: the fusion related domain (Fusion); the transmembrane domain (TM); an amphipathic alpha-helix region (H); a leucine zipper-like domain (Z); and a final a-helical domain (F1h). There are two B-sheet/turn-rich regions: the cysteine-rich domain; and a intracellular anchor region (Anc). The inferred disulphide bond is denoted by the dotted line between two cysteine residues (c). Reproduced from Chambers *et al.*, (1992).

domain. Heptad repeats are sequences of amino acids arranged with a seven residue periodicity in which positions 1 and 4 are predominantly occupied by small hydrophobic or neutral residues. These heptad repeat regions suggest the presence of extended α-helical conformations, and there is evidence to suggest that at least in the case of NDV, the H region interacts with the F_2-subunit (Wang *et al.*, 1992). Downstream of the H region is another loop-forming sequence which is followed by a second hydrophobic region with a relatively low prediction for α-helix (F_1h). The cysteine-rich domain is predicted to possess β-sheet structures, but no overall structure could be determined for this region because this region is proline-rich. Located C-terminal of the cysteine-rich domain is another hydrophobic heptad repeat sequence region (Z). This is analogous to the leucine zipper region of paramyxovirus proteins identified by Buckland and Wild (1989). The extreme C-terminus of the F_1 subunit contains a final hydrophobic domain thought to be the membrane anchor region. By specific alteration of residues in the hydrophobic terminal domain of paramyxovirus F proteins it has been shown that amino acid changes that increase the helical content of this transmembrane region, increase the cell-fusion activity of mutant F proteins (Horvath and Lamb, 1992).

Human RS virus F protein can be recovered as a non-covalently linked homodimer (2 X F_2-F_1)(Walsh *et al.*, 1985). SV F protein can be isolated as a tetramer of two dimers, and this tetramer is thought to be the membrane spike (Sechoy *et al.*, 1987). Given the general structural similarities between RS virus and SV F proteins, the higher-order structure of RS virus F protein might be similar (Markwell and Fox, 1980; Sechoy *et al.*, 1987; Ng *et al.*, 1989; Collins and Mottet, 1991). On the other hand RS virus F has been isolated as a dimer associated with the G protein monomer (Arumugham *et al.*, 1989). This would suggest the possibility that the RS virus F and G proteins are not segregated into different spikes (Collins, 1991). Fusogenic activity of the SV5 fusion protein has been demonstrated to be dependent on transport of the protein to the cell surface (Horvath and Lamb, 1992). The heptad repeat sequence (the H box region) of the F_1 subunit has been shown to be crucial in fusion protein transport in SV5 (Horvath and Lamb, 1992) and

NDV (Wang *et al.*, 1992). Anderson *et al.* (1992) have suggested that cleavage of human RS virus F protein can be disrupted by amino acid changes distal to the cleavage site, suggesting a conformational change such that the cleavage site is rendered inaccessible to host cell proteases. Intracellular processing of human RS virus fusion protein has been investigated by comparing the maturation and stability of wild-type F, uncleaved mutant F and chimeric F proteins expressed by recombinant vaccinia viruses (Anderson *et al.*, 1992). They found that mutations at residues #301 (Val to Ala) and #447 (Val to Met) resulted in the expression of uncleaved F protein. They also noted that the uncleaved F protein was associated with GRP78-BiP in the ER for a longer time than the wild-type F protein and that protein transport was inefficient. In addition the uncleaved F protein was not recognised by a panel of monoclonal antibodies in ELISA or indirect immunofluorescence assays, suggesting that the protein was miss-folded, and as a result not transported properly or cleaved.

1.7.6 F protein variation The F proteins of human RS virus subgroup A and B strains are highly conserved, exhibiting a high degree of sequence conservation (91%; Collins, 1991). Despite the high degree of sequence conservation there are strain-specific mobility differences and some epitopic variation (Mufson *et al.*, 1985; Norrby *et al.*, 1986; Baybutt and Pringle, 1987; Johnson and Collins, 1988b). Within the human fusion proteins there are two regions of divergence, both occurring in the F_2-subunit. The first is the signal peptide with only 34% homology, indicating that its function is dependent on hydrophobicity rather than amino acid sequence. The second region, amino acids #111 to #129, has only 50% identity. Analysis of the fusion proteins of RS virus subgroup A strains reveals that there is a higher degree of homology, with strains A2 and RSS-2 thought to be 98% identical (Baybutt and Pringle, 1987). Their fusion proteins also reveal the two regions of divergence identified above, the signal peptide is 79% homologous, but the second region, amino acids #111 to #129, has much higher identity of 90% as well as containing an extra N-linked glycosylation site found in strain RSS-2 but not in strain A2. This region is relatively hydrophilic suggesting that it might be exposed in the folded

molecule, contributing to a subgroup specific B- or T-cell epitope (Anderson *et al.*, 1985; Mufson *et al.*, 1985; Pemberton *et al.*, 1987).

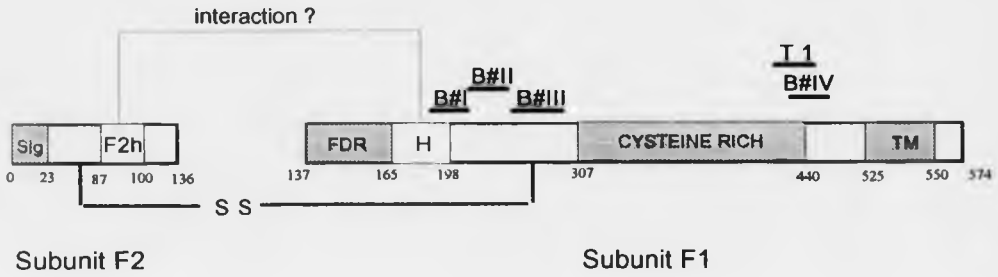
1.7.7 F protein Antigenicity The F protein of RS virus has been shown to be a major antigen that efficiently induced serum neutralising antibodies, specific cytotoxic and helper T lymphocytes, and a high level of resistance to challenge with homologous virus in rodents and monkeys (Olmsted *et al.*, 1986; Bangham and Askonas, 1986; Walsh *et al.*, 1987b; Wertz *et al.*, 1987; Pemberton *et al.*, 1987; Wathen *et al.*, 1989). Studies in cotton rats have suggested that the F protein is the most important RS virus protective antigen, and the protective immunity induced is equally effective against both viral subgroups (Olmsted *et al.*, 1986; Johnson *et al.*, 1987a). Immunising mice or cotton rats with recombinant vaccinia viruses expressing the fusion protein demonstrated protection against challenge and the induction of neutralising antibodies (Olmsted *et al.*, 1986; Walsh *et al.*, 1987b; Wertz *et al.*, 1987). By analogy to PIV-3, the degree of F protein glycosylation is not thought to be relevant in order to elicit a protective response (Ewasyshyn *et al.*, 1993).

1.7.8 F protein B-cell epitopes Three functional types of monoclonal antibody that bind to the F protein have been described using *in vitro* assays (Walsh *et al.*, 1987a). These consist of **1)** non-neutralising, non-fusion-inhibiting antibody; **2)** neutralising, non-fusion-inhibiting antibody; and **3)** neutralising, fusion-inhibiting antibody. However the situation is far from clear cut, with all monoclonal antibodies that are fusion inhibiting able to neutralise virus but not vice-versa, and not all neutralising monoclonal antibodies able to protect mice by passive immunisation (Taylor *et al.*, 1984; Walsh *et al.*, 1989). It is thought that fusion inhibition correlates with protection but protection has been described with monoclonal antibodies that are neither neutralising nor fusion-inhibiting; and monoclonal antibodies that are neutralising but not fusion-inhibiting (Taylor *et al.*, 1992). The F protein has been mapped using sets of functionally defined murine monoclonal antibodies in competitive binding assays to define topological relationships between

antigenic areas. At least three antigenic areas have been defined by this method (A, B and C) all of which have been found to bind neutralising monoclonal antibodies, and one of which includes fusion-inhibiting monoclonal antibodies (antigenic area #IV) (Walsh *et al.*, 1987a; Anderson *et al.*, 1986; Trudel *et al.*, 1987a and b; Beeler and Van Wyke-Coelingh, 1989; Garcia-Barreno *et al*, 1989). Initially it was supposed that three of the four antigenic areas were partially inter-related, leaving one independent fusion-inhibiting antigenic area (Walsh *et al.*, 1987a; Trudel *et al.*, 1987a and b; Garcia-Barreno *et al.*, 1989; Beeler and Van Wyke-Coelingh, 1989; Trudel *et al.*, 1991). It is now thought that the antigenic areas of the neutralising, non-fusion-inhibiting monoclonal antibodies overlap with the fusion-inhibiting site (Walsh *et al.*, 1987b; Beeler and Van Wyke-Coelingh, 1989; Taylor *et al.*, 1992). In terms of epitopes on the fusion protein, the neutralising and fusion-inhibiting antibodies map to two specific regions, the cysteine-rich region and a second region downstream of the fusion-related domain (Martin-Gallardo *et al.*, 1991; Arbiza *et al.*, 1992; Lounsbach *et al.*, 1993).

Positional mapping of the B cell epitopes using peptide analysis has revealed a major binding site of an RS virus neutralising non-fusion-inhibiting monoclonal antibody (antigenic area #I) at positions #205 to #236 (Trudel *et al.*, 1987a and b; Trudel *et al.*, 1991; Bourgeois *et al.*, 1991). Peptides that corresponded to amino acids #205 to #225 were able to induce a neutralising antibody response in rabbits. Polyclonal anti-RS virus rabbit antiserum loses neutralising capacity when reacted with this peptide. However it also lost neutralising capacity when reacted with a peptide corresponding to amino acids #259 to #278 (Bourgeois *et al.*, 1991; Trudel *et al.*, 1991); this region has been defined as antigenic area #II (#255 to #288)(Trudel *et al.*, 1991; Arbiza *et al.*, 1992; Lounsbach *et al.*, 1993). A third antigenic area has been mapped between amino acids #283 to #298 (Martin-Gallardo *et al.*, 1991; Paradiso *et al.*, 1991). The fourth antigenic region has been mapped to amino acids #417 to #450 (Scopes *et al.*, 1990; Arbiza *et al.*, 1992). Epitope maps are summarised in **Table 5**. There is some uncertainty using peptide mapping

Figure 6. Location of epitopes in the RS virus fusion protein.



Map of the B- and T-cell epitopes and antigenic sites. B-cell epitopes are shown as B #I -IV, representing antigenic areas #I to #IV. The T-cell epitope is shown as T1. The potential regions of structural interaction in the folded protein are shown linked by a dotted line.

Table 5 Summary of Fusion protein epitopes

Antigenic areas	Residue #	Antibody Group	Reference
B-Cell Epitopes			
I	205 → 236	Group A	Brideau *et al.*,1991
II	255 → 288	Group A	Lopez *et al.*, 1990
III	283 → 298	Group A	Martin-Gallardo *et al.*, 1991
IV	417 → 450	Group C	Scopes *et al.*, 1990
T-Cell Epitopes			
#1	338→ 355	n.a.	Levely *et al.*, 1991

techniques in defining an antigenic region since the peptide could be just a small linear part of a complex antigenic site. For instance one monoclonal antibody (MAb L4) shows a 100-fold greater binding efficiency to F in dimer form than to reduced F_1. This suggests that its native antigenic site is conformational (Paradiso *et al.*, 1991).

An alternative method to epitope mapping is to raise monoclonal antibody-resistant escape mutant viruses and map the base changes that have occurred. Escape mutants raised against monoclonal antibodies that are in the same competitive binding assay group have shown sequence alterations at amino acids #262, #268 or #272 (Trudel *et al.*, 1987a and b; Garcia-Barreno *et al.*, 1989; Arbiza *et al.*, 1992). However mutant viruses with alterations at #268 and #272 are also resistant to the binding of a non-neutralising monoclonal antibody. Thus the area between amino acids #262 and #272 must be involved in more than one functional epitope. Changes at amino acids #268 and #272 prevent the binding of a monoclonal antibody that has been mapped to amino acids #222 to #232 (antigenic area #I). However escape mutant viruses raised to this monoclonal antibody had two amino acid changes at #190 and #258. This suggests that the two linear regions around antigenic area #I (#222 to #232) and antigenic area #II (#259 to #278) could be involved in a conformational epitope (Arbiza *et al.*, 1992; Lopez *et al.*, 1993). Evidence supporting the existance of a large antigenic site consisting of two or more antigenic areas was reported by Trudel *et al.*, (1991) who found that Mab 7C2 reacted with peptides containing regions #221 to #232 (antigenic area #I) and also containing regions #275 to #288 which overlap antigenic areas #II and #III. Other escape mutant viruses exhibited changes at amino acid #429 (Arbiza *et al.* 1992), confirming the mapping of antigenic area #IV at this region as defined by Scopes *et al.*, (1990). This region could be the independent fusion-inhibiting antigenic site described by Beeler and Van Wyke-Coelingh (1989); Garcia-Barreno *et al.* (1989); Trudel *et al.* (1991).

techniques in defining an antigenic region since the peptide could be just a small linear part of a complex antigenic site. For instance one monoclonal antibody (MAb L4) shows a 100-fold greater binding efficiency to F in dimer form than to reduced F_1. This suggests that its native antigenic site is conformational (Paradiso *et al.*, 1991).

An alternative method to epitope mapping is to raise monoclonal antibody-resistant escape mutant viruses and map the base changes that have occurred. Escape mutants raised against monoclonal antibodies that are in the same competitive binding assay group have shown sequence alterations at amino acids #262, #268 or #272 (Trudel *et al.*, 1987a and b; Garcia-Barreno *et al.*, 1989; Arbiza *et al.*, 1992). However mutant viruses with alterations at #268 and #272 are also resistant to the binding of a non-neutralising monoclonal antibody. Thus the area between amino acids #262 and #272 must be involved in more than one functional epitope. Changes at amino acids #268 and #272 prevent the binding of a monoclonal antibody that has been mapped to amino acids #222 to #232 (antigenic area #I). However escape mutant viruses raised to this monoclonal antibody had two amino acid changes at #190 and #258. This suggests that the two linear regions around antigenic area #I (#222 to #232) and antigenic area #II (#259 to #278) could be involved in a conformational epitope (Arbiza *et al.*, 1992; Lopez *et al.*, 1993). Evidence supporting the existance of a large antigenic site consisting of two or more antigenic areas was reported by Trudel *et al.*, (1991) who found that Mab 7C2 reacted with peptides containing regions #221 to #232 (antigenic area #I) and also containing regions #275 to #288 which overlap antigenic areas #II and #III. Other escape mutant viruses exhibited changes at amino acid #429 (Arbiza *et al.* 1992), confirming the mapping of antigenic area #IV at this region as defined by Scopes *et al.*, (1990). This region could be the independent fusion-inhibiting antigenic site described by Beeler and Van Wyke-Coelingh (1989); Garcia-Barreno *et al.* (1989); Trudel *et al.* (1991).

1.7.9 F protein T-cell epitopes RS virus infection in humans and in experimental animals induces a vigorous cell-mediated response. Target proteins for RS virus-specific cytotoxic T lymphocytes include the fusion protein (Pemberton *et al.*, 1987; Nicholas *et al.*, 1990; Leverly *et al.*, 1991; Cherrie *et al.*, 1992; Connors *et al.*, 1992). Peptide scanning of the fusion protein has identified an immunodominant T-cell epitope spanning residues #338 to #355 on the F_1 subunit (see **Figure 6**). The response appears to be MHC class 2 restricted and mediated by the $CD4^+$ subset of T cells (Leverly *et al.*, 1991). Similar studies carried out in BALB/c mice identified the immunodominant T-cell epitope as spanning residues #51 to #65 on the F_2-subunit. It is therefore possible that recognition of different regions of the fusion protein by T cells from humans and mice results from differences in the nature of exposure to antigen.

1.8 RS virus immunobiology Neutralising antibody is only partially effective in providing protection since repeated infections occur and not until the third infection is there an appreciable reduction in the severity of RS virus disease in children (Hall, 1980). Re-infection may thus be necessary to elicit substantial immunity since there is a pattern of gradually increasing levels of serum antibody with successive infections (Stott and Taylor, 1985). The epidemic pattern may be similar to the antigenic drift observed in the haemagglutinin (HA) and neuraminidase (NA) of Influenza A virus isolates (White and Fenner, 1986), suggesting some immunological pressure (Nagai *et al.*, 1993). Both serum and secretory antibodies are produced upon RS virus infection with higher levels of protection in mice conferred by anti-F rather than anti-G antibodies (Stott *et al.*, 1986; Olmsted *et al.*, 1986; Johnson *et al.*, 1987a; Walsh *et al.*, 1987b).

Both subgroup A and B strains are known to co-circulate within the same community (Mufson *et al.*, 1988; Hendry *et al.*, 1989), thus the immune response may vary depending on the strain or subgroup exposure. Anti-G antibodies have been demonstrated to be 13-fold less effective against heterologous subgroup challenge than against homologous challenge (Johnson *et al.*, 1987a; Stott *et al.*, 1987), with a possible explanation lying in the sequence variation of the G gene (Cane and Pringle, 1992). The role of T-cell immunity is less clear, but cross-reactive RS virus-specific murine cytotoxic T cells have been identified which probably recognise a conserved internal viral antigen which is thought to be the 22K protein (Openshaw *et al.*, 1988; Connors *et al.* 1992).

1.8.1 The humoral response to RS viral infection Studies in mice have demonstrated that antibodies to the F and G proteins neutralise RS virus infectivity, and have also shown that passively transferred monoclonal antibodies to G or F protect animals from lower respiratory tract infection (Taylor *et al.*, 1984; Walsh *et al.*, 1984a). However, in human infants maternally derived antibodies can decrease antibody response to RS viral glycoproteins G and F (Murphy *et al.*, 1986). Murphy *et al.* (1988) have showed that cotton rat RS virus hyperimmune antiserum suppresses the antibody response to RS viral F and G and renders the rats more susceptible to RS viral infection. Also immunosuppressed animals developed F and G antibodies with low neutralising activities. Thus pre-existing serum RS virus antibodies have a quantitative and qualitative effect on the immune response to RS viral infection. This evidence is also supported by work by Popow-Kraupp *et al.* (1989) and Ward *et al.* (1983). The F and G proteins have been demonstrated to give a strong, lasting protective humoral response when maternal antibodies are absent (Olmsted *et al.*, 1986; Stott *et al.*, 1986; Wertz *et al.*, 1987). On the other hand the P, M, SH, NS1 and NS2 proteins are shown to give weak antibody responses (Taylor *et al.*, 1984; Connors *et al.*, 1992), whilst any protection derived from the N and 22K proteins appears to be short-lived (Connors *et al.*, 1992). The degree to which the F and G proteins contribute to protection against RS virus infection indicates that both contribute significantly towards immunity (Brideau *et al.*, 1989). However

this result highlights the difficulties in designing a vaccine to RS virus, because whilst the F protein is well conserved between subgroups (Collins *et al.*, 1984b; Elango *et al.*, 1985a; Baybutt and Pringle, 1987; Johnson and Collins, 1988b) the G protein shows significant differences both within and between subgroups (Anderson *et al.*, 1985; Akerlind *et al.*, 1988; Garcia-Barreno *et al.*, 1989; Sullender *et al.*, 1991; Cane *et al.*, 1991).

1.8.2 Cell-mediated response to RS viral infection Cannon *et al.* (1989) suggest that the level of the T_c response is critical, although the T_c response clears viral infection, an excessive response could play a role in the pathogenesis of infantile bronchiolitis and/or pneumonia. Work by Ananaba and Anderson (1991) supports this view. The T_h response to specific viral proteins in BALB/c mice has been investigated, using vaccinia expressing recombinants, and it has been observed that there is a good response to F and N but a weak response to G (Openshaw *et al.*, 1988). There was no apparent recognition of expressed SH and NS2 proteins. Nicholas *et al.* (1990) further investigated and found that 22K was the major target protein for the CD8+ cytotoxic T lymphocytes. Immunodeficient children are at high risk of complicated, prolonged and fatal RS virus infections probably because of defective cellular immunity (McIntosh and Fishaut, 1980). Welliver *et al.* (1985) noted that children recovering from RS bronchiolitis had reduced CD8+ levels in their peripheral blood. Other indicators of the importance of cellular responses are that the transfer of primed T cells to persistently infected immunodeficient mice will lead to rapid viral clearance in the absence of antibodies (Cannon *et al.*, 1987). However as with the case for the humoral response, the situation is far from clear-cut, murine T_c cells have been shown to recognise different antigenic subtypes of RS virus (Bangham and Askonas, 1986; Openshaw *et al.*, 1988). Cherrie *et al.* (1992) found that whilst the T-cell response in humans to RS viral proteins proceeded in the order N>SH>F>M>22K, the level of recognition of F and 22K was proportional to increased levels of serum IgG, indicative of recent RS virus infection. Connors *et al.* (1992) concluded that resistance to RS virus

infection in mice is mediated by CD8+ T cells recognising the 22K protein, but was also mediated by antibodies recognising F and G.

1.9 Vaccine development RS virus is the major cause of severe lower respiratory tract infections in children during the first year of life (Kim *et al.*, 1973a; Stott *et al.*, 1980). The presence of serum neutralising antibody however, does not prevent reinfection which can occur throughout life at roughly 36-month intervals (Hall *et al.*, 1980). The identification of discrete isolate types within RS virus subgroups (**Section 1.1.2**) has suggested that antigenic variation is sufficient to be a determining factor in reinfection (Hendry *et al.*, 1988a, 1989; Mufson *et al.*, 1988; Breese-Hall *et al.*, 1990; Cane *et al.*, 1991; Anderson *et al.*, 1991).

Difficulties have been encountered in the development of vaccines for children. Inactivated vaccine was considered undesirable due to results where a formalin-inactivated RS virus vaccine administered to children not only failed to protect, but exacerbated the disease response on exposure to virus (Chin *et al.*, 1969; Kim *et al.*, 1969; Kapikian *et al.*, 1969). It is thought that the formalin-inactivation modified some of the epitopes on RS virus F and/or G glycoproteins so that immunisation resulted in an unbalanced immune response, producing non-neutralising antibodies which may have contributed to an Arthus-type reaction. After these results, interest in vaccine development moved to the use of attenuated virus vaccines. Local immunity is thought to play an major role in resistance to illness, consequently a live attenuated virus restricted to replicating in the upper respiratory tract without producing disease was thought to be desirable because lower respiratory tract diseases are the main cause of mortality and severe disease, usually manifesting as bronchiolitis or pneumonia. Temperature-sensitive mutants that are attenuated and have growth restricted to the upper respiratory tract have been derived for many members of the *Paramyxoviridae* (Wright *et al.*, 1973; Portner *et al.*, 1974; Tsipis and Bratt, 1976). Temperature-sensitive mutants have been derived from subgroup A RS viruses, the A2 wild-type strain, after treatment with 5-fluorouridine and 5-fluorouracil

(Gharpure *et al.*, 1969) and from the RSS-2 wild-type strain after treatment with 5-fluorouracil and ICR compounds (McKay *et al.*, 1988), and from the RSN-2 strain of subgroup B (Faulkner *et al.*, 1976). Temperature-sensitive mutants derived from the A2 strain have been classified into 3 complementation groups , A (*ts*A1), B(*ts*A2) and C (*ts*A7) (Gharpure *et al.*, 1969; Richardson *et al.*, 1977; Belshe *et al.*, 1977, 1978; Gimenez and Pringle, 1978; Pringle *et al.*, 1981), all seemingly having a maturation defect defined by reduced surface antigen as detected by immunoflourescence and SDS-PAGE. Two of these ts mutants, *ts*A1 and *ts*A2, have been tested as potential live vaccines. The *ts*A1 mutant did not produce any symptoms when administered intranasally to adult volunteers and was found to be genetically stable, inducing resistance to challenge with the wild-type virus (Wright *et al.*, 1971). However problems occurred when *ts*A1 was inoculated into sero-negative infants because genetically altered revertant viruses were recovered. Also some sero-negative children developed mild symptoms of disease (Kim *et al.*, 1973b). These revertants emerged late in infection and presented a complex pattern of mainly partial (90%) and some full (10%) revertants (Hodes *et al.*, 1974). Attempts to produce more stable mutants were undertaken by isolating second-stage mutant clones of *ts*A1 after exposure to nitrosoguanidine (Richardson *et al.*, 1977), however these were found to be over attenuated (Belshe *et al.*, 1978). Trials with *ts*A2 showed that this mutant was unsuitable as a candidate vaccine due to low infectivity when administered intranasally; however it did appear to be genetically stable (Wright *et al.*, 1982).

Another candidate vaccine has been developed as a third-stage ts mutant (*ts*1C) of the RSS-2 strain, which has been derived entirely in human diploid cells (MRC-5). A high level of protection against wild-type challenge with low disease-inducing capacity was observed in preliminary trials in adult volunteers (McKay *et al.*, 1988; Watt *et al.*, 1990; Pringle *et al.*, 1993). Mutant *ts*1C was sequentially derived from the wild-type RSS-2 strain via the intermediate viruses *ts*1A and *ts*1B, using different mutagenising agents. The three mutants *ts*1A, *ts*1B and *ts*1C have restrictive temperatures of 39°C, 38°C and 37°C respectively (McKay *et al.*, 1988; Pringle *et al.*, 1993). The use of different mutagens to

obtain a triple ts mutant was designed to lessen the likely-hood of loss of temperature-sensitivity by reversion. All three mutant viruses have RNA positive phenotypes and synthesise viral polypeptides intracellularly at their non-permissive temperatures. The single mutant *ts*1A, when administered intranasally in low dose to adult volunteers, reduced disease-inducing potential in comparison to the wild-type virus RSS-2 but produced a poor immune response. The double (*ts*1B) and triple (*ts*1C) mutants, however, gave the greatest reduction in disease-inducing potential with approximately equal levels of immunogenicity (Pringle *et al.*, 1993).

Another approach to vaccine development adopted in recent years is that of administering purified fusion (F) or attachment (G) glycoproteins. The role of RS virus proteins in the immune response has been described previously (**Section 1.8**). Merz *et al.* (1980) suggested that a successful subunit vaccine should induce antibody to inhibit fusion activity but also induce a balanced immune response. Consequently, immunisation with a fragment of the F protein able to induce both neutralising and fusion-inhibiting antibody responses might constitute a safe and effective vaccine. RS virus F and G glycoproteins have been expressed in a variety of systems such as vaccinia virus recombinants (Olmsted *et al.*, 1986; Stott *et al.*, 1986), baculovirus vectors (Wathen *et al.*, 1989), bacterial systems such as expression in *E.coli* (Lounsbach *et al.*, 1993) and mammalian cells (Belshe *et al.*, 1993). Protection against RS virus infection has been achieved by immunisation of Cotton rats with F or G proteins (Stott *et al.*, 1986; Walsh *et al.*, 1987b) or with chimaeric proteins consisting of both F and G sequences (Brideau *et al.*, 1989) but the nature of the protective immune response remains to be clearly defined.

1.10 Objectives of this study The primary question to be examined by this study is whether variation in the fusion gene products of human RS virus subgroup A isolates is greater than that previously reported for subgroup A strains A2, RSS-2, Long and Edinburgh (**Chapter 3**). It is thought that sequence variation in the fusion gene product could provide a basis for antigenic variation within a specific subgroup of human RS

viruses. Therefore providing a possible explanation for reinfection with RS virus in human adults.

Secondary goals are to establish the location and nature of the ts lesions in the candidate vaccine strains *ts*A1 **(Chapter 4)** and candidate vaccine strain *ts*1C **(Chapter 5)**. It is also hoped to provide an explanation for the complex nature of reversion of mutant *ts*A1 to non-ts phenotype.

CHAPTER TWO

CHAPTER TWO

CHAPTER TWO

Materials and methods

2.1 CELLS AND VIRUSES

2.1.1 Mammalian cells MRC-5 and BS-C-1 G23 cells were provided by Professor C .R. Pringle (University of Warwick) and grown in GMEM 5% FCS. The 293 cell line (Graham and Van der Eb, 1973) was provided by Dr. C. Caravokyri (University of Warwick) and grown in DMEM 10% Calf serum. All cells were grown in a humidified incubator at 37°C under 5 % CO_2.

2.1.2 Viruses RS virus strains A2 and RSS-2 (and their ts mutants) were obtained from Professor C. R. Pringle. The A2 wild-type and its mutant *ts*A1 were originally isolated by Dr. R. Chanock (NIH, Maryland, USA). Strain RSS-2 was originally isolated by Professor P. Gardner (University of Newcastle). Mutants *ts*1A, *ts*1B and *ts*1C were originally isolated by McKay *et al.*, (1988). Virus strains RSB89-6190; RSB89-5857; RSB89-6256; RSB89-6614 and RSB89-1734 (subgroup A isolates) were provided by Dr. P. Cane (University of Warwick) and were isolated from nasopharyngeal aspirates (Cane and Pringle., 1991).

Culture media, trypsin, versene, agar, neutral red, amino acids and antibiotic solutions, PBS and double distilled water were provided by the Virus Group media preparation laboratory.

2.1.3 Materials All solutions described in this chapter show compounds along-side their final concentrations.

Crystal violet stain Stock solution of 1.5% (w/v) crystal violet dissolved in ethanol was made, then diluted 1 in 40 with water before use. This was stored at ambient temperature.

Glasgow modification of Eagle's medium (GMEM) Eagle's minimum enabled medium was supplemented with non-essential amino acids as described previously in McKay *et al.*, (1988), plus the following at the final concentrations given; 0.2 mM glutamine, 100 units/ml penicillin, 100 μg/ml streptomycin and 5% or 10% (v/v) bovine FCS. This was sterilised by filtration and stored at 4°C.

Dulbeco's modification of Eagles medium (DMEM) Eagle's minimum enabled medium was supplemented with non-essential amino acids as described previously Graham and Van der Eb (1973), plus the following at the final concentrations given; 0.2 mM glutamine, 100 units/ml penicillin, 100 μg/ml streptomycin and 10 % (v/v) bovine calf serum. This was sterilised by filtration and stored at 4°C.

Overlay agar An equal volume of double strength supplemented GMEM was added to melted Noble agar (1.8%) which had been cooled to 46°C. The overlay was maintained at 46°C until used.

Glutaraldehyde A working solution of 1% (v/v) glutaraldehyde in PBS was prepared and stored at ambient temperature until use.

Neutral red overlay 5 ml of a neutral red stock solution (0.1% w/v) was added to 200 ml overlay agar and maintained at 46°C until needed.

Versene Solutions consisting of 0.02% (w/v) EDTA (disodium salt) and 0.002% (w/v) phenol red in PBS pH 7.4 were prepared and stored at ambient temperature.

TD Solutions consisting of 0.14 M sodium chloride, 5 mM potassium chloride, 0.5 mM di-sodium orthophosphate and 25 mM Tris were prepared and adjusted to pH 7.5 with concentrated hydrochloric acid. Solutions were then autoclaved and stored at ambient temperature.

TDE Solutions were prepared from TD solution by the addition of 0.2 ml 0.25 M EDTA/100 ml pH 8.0, sterilised by autoclaving and kept at ambient temperature until use.

Foetal calf serum from Gibco BRL (Gaithesburg, U.S.A) and tissue culture plastic-ware were from Northumbria Biologicals (Northumberland, U.K.). Crystal violet was from Sigma (St.Louis, U.S.A.) and glutaraldehyde from BDH (Poole, U.K.).

2.1.4 Cell culture BS-C-1 G23 and MRC-5 cells were grown in 175 cm^2 tissue-culture flasks in supplemented GMEM 5% FCS at 37°C and passaged at 3 to 4 day-intervals. Cells were washed with 20 ml versene twice and once with 20 ml trypsin/versene (1 in 5, v/v). A small amount of the latter was left in the flask for further incubation at 37°C, until all the cells became detached from the flask surface. The cell suspension was then diluted with supplemented GMEM containing 5% FCS and one-tenth aliquots were pipetted into new flasks and fresh 50 ml GMEM 5% FCS added.

293 cells were grown in 90 mm^2 tissue culture Petri-dishes in supplemented DMEM containing 10% Calf serum at 37°C and passaged at 3 day-intervals. Cells were washed once with TD and then removed from the monolayer with 2 ml TDE and 0.2 ml trypsin. The cell suspension was added to 0.5 ml Calf serum, pelleted at 10,000 x g for 5 minutes and the pellet resuspended in 10 ml supplemented DMEM. One-fifth aliquots were then pipetted into new dishes.

2.1.5 Preparation of virus stocks Confluent monolayers of either MRC-5 or BS-C-1 G23 cells in 25 cm^2 tissue-culture flasks were inoculated with 200 µl virus (m.o.i. of

approximately 1 pfu/cell), placed at 33°C for one hour, and 5 ml supplemented GMEM containing 5% FCS added. Cells were then incubated at 33°C until 80% CPE was observed. Virus was harvested by immediate freezing of infected cells at -70°C.

2.1.6 Viral titre Confluent MRC-5 or BS-C-1 G23 cells were infected with 200 µl viral stock (m.o.i. of approximately 1 pfu/cell) through a dilution series down to 10^{-6}. The cells were then incubated at 33°C for 1 hr and the supernatant removed. 2.5 ml of overlay agar was then added and the cells incubated at 33°C until syncytia observed. The cells were fixed by adding 1% glutaraldehyde solution onto the agar overnight. The agar was then removed. The cells were then stained with crystal violet for 5 minutes at room temperature and excess stain removed by washing with water. Plaques were then counted under a light microscope.

2.1.7 Selection of ts revertants BS-C-1 G23 cells were infected with 200 µl viral stock (m.o.i. of approximately 1 pfu/cell) through a dilution series down to 10^{-3} as described for viral titres. Cells were then incubated for 7 days at the non-permissive temperature of 39°C. Cells were stained overnight with 2 ml neutral red overlay placed on top of the agar overlay. Viral plaques were identified , removed using a sterile Pasteur pipette and viral stocks grown at 39°C. The revertant viruses were plaque purified a further two times using the above procedure.

2.2 ANALYSIS OF VIRAL PROTEINS

All chemicals mentioned in this and subsequent sections were obtained from BDH (unless otherwise stated) and were of AnalaR grade. Radioisotopes for protein labelling, the fluorographic agent Amplify™, biotinylated anti-human and anti-bovine IgG, biotinylated peroxidase-streptavidin complex, biotinylated fluorescein-streptavidin complex and nitro-cellulose (Hybond-C) were purchased from Amersham International plc. Protein markers were obtained from Sigma. X-ray films were from Fuji Photo Film

approximately 1 pfu/cell), placed at 33°C for one hour, and 5 ml supplemented GMEM containing 5% FCS added. Cells were then incubated at 33°C until 80% CPE was observed. Virus was harvested by immediate freezing of infected cells at -70°C.

2.1.6 Viral titre Confluent MRC-5 or BS-C-1 G23 cells were infected with 200 µl viral stock (m.o.i. of approximately 1 pfu/cell) through a dilution series down to 10^{-6}. The cells were then incubated at 33°C for 1 hr and the supernatant removed. 2.5 ml of overlay agar was then added and the cells incubated at 33°C until syncytia observed. The cells were fixed by adding 1% glutaraldehyde solution onto the agar overnight. The agar was then removed. The cells were then stained with crystal violet for 5 minutes at room temperature and excess stain removed by washing with water. Plaques were then counted under a light microscope.

2.1.7 Selection of ts revertants BS-C-1 G23 cells were infected with 200 µl viral stock (m.o.i. of approximately 1 pfu/cell) through a dilution series down to 10^{-3} as described for viral titres. Cells were then incubated for 7 days at the non-permissive temperature of 39°C. Cells were stained overnight with 2 ml neutral red overlay placed on top of the agar overlay. Viral plaques were identified , removed using a sterile Pasteur pipette and viral stocks grown at 39°C. The revertant viruses were plaque purified a further two times using the above procedure.

2.2 ANALYSIS OF VIRAL PROTEINS

All chemicals mentioned in this and subsequent sections were obtained from BDH (unless otherwise stated) and were of AnalaR grade. Radioisotopes for protein labelling, the fluorographic agent Amplify™, biotinylated anti-human and anti-bovine IgG, biotinylated peroxidase-streptavidin complex, biotinylated fluorescein-streptavidin complex and nitro-cellulose (Hybond-C) were purchased from Amersham International plc. Protein markers were obtained from Sigma. X-ray films were from Fuji Photo Film

Co.(Japan). The monoclonal antibodies used in the subsequent experiments were a kind gift from Drs.E.Nörrby and C.Örvell (Karolinska Institute, Stockholm, Sweden). Polyclonal bovine anti-human anti-RS serum was provided by Drs. Stott and Taylor (IAH, Compton Laboratory).

2.2.1 Materials

SDS-PAGE gel fixing solution Fixing solution was prepared by the addition of 50% (v/v) methanol and 7% (v/v) acetic acid in distilled water and was stored at ambient temperature before use.

Methionine-free GMEM A 1 litre stock solution containing GMEM plus dialysed foetal calf serum (5% final volume) plus 2 mM leucine, 2 mM valine and 4 mM glutamine. 0.01% (v/v) sodium hydrogen carbonate, 100 units/ml of Penicillin and 100 µg/ml of Streptomycin was prepared and sterilised by filtration. This working solution was stored at 4°C.

Phenylmethylsulphonyl fluoride (PMSF) Working stock solutions of 100 mM PMSF (supplied by Sigma) in ethanol were prepared and stored at 4°C. When mentioned in conjunction with buffers, PMSF was added immediately before use to a final concentration of 1 mM.

Phosphate buffered saline (PBS pH 7.4) Stock solutions were prepared, consisting of 2.5 mM potassium chloride, 0.15 M sodium chloride, 1.5 mM potassium dihydrogen orthophosphate and 6.5 mM di-sodium hydrogen orthophosphate, and were sterilised by autoclaving and stored at ambient temperature.

Tris buffered saline (TBS pH 7.6) Stock solutions were prepared, consisting of 150 mM sodium chloride, 5 mM potassium chloride and 25 mM Tris-base adjusted to pH 7.6 and sterilised by autoclave.

PBS-T Stock solutions were prepared prior to use, and consisted of PBS as above, with the addition of 0.1% (v/v) of Tween 20.

Preparation of immunoprecipitin (IMP) Immunoprecipitin (formalin-fixed *S. aureus* A cells, BRL) was pelleted at 10,000 x g for 2 minutes. The supernatant was discarded and the pellet resuspended in the buffer used for immunoprecipitation of viral polypeptides (RIP buffer). IMP solutions were stored at 4°C.

RIP buffer RIP buffer consisting of 0.01 M Tris-HCl (pH 7.4), 0.15 M sodium chloride, 1% (w/v) sodium deoxycholate, 1% (v/v) Triton X-100 and 0.1% (w/v) SDS, was made and sterilised by autoclaving and stored at ambient temperature.

RIP washing buffer Stock solutions were prepared, consisting of 0.5 M lithium chloride and 0.15 M Tris-HCl (pH 8.5). The solution was sterilised by autoclave and stored at ambient temperature.

SDS-PAGE reducing sample buffer (3 X concentrate) This buffer consisted of 0.15 M Tris-HCl (pH 6.7) with 30% (v/v) glycerol, 15% (v/v) β-mercaptoethanol, 6% (w/v) SDS and 0.15 mg/ml bromophenol blue.

SDS-PAGE non-reducing sample buffer (2X concentrate) Stock solutions were prepared, consisting of 0.1 M Tris-HCl (pH 6.7), 20% (v/v) glycerol, 4% (w/v) SDS and 0.1 mg/ml bromophenol blue.

SDS-PAGE running buffer Running buffer consisted of 0.6% (w/v) Tris-base, 0.4% (w/v) glycine and 0.1% (w/v) SDS, and was made up immediately prior to use.

SDS-polyacrylamide single-concentration gels Gel solutions were prepared using a 30% stock solution of acrylamide/bisacrylamide at a w/w ratio of 28.5 to 1.5. All resolving gels were used at a concentration of 10% acrylamide-bis, 0.4 M Tris-HCl (pH 8.9) and 0.1% (w/v) SDS. The gels were polymerised by adding 200 µl 10% ammonium persulphate and 20 µl TEMED.

The stacking gel used in the SDS-PAGE experiments consisted of 3.7% acrylamide-bis, 0.25M Tris-HCl (pH 6.7) and 0.1% (w/v) SDS. The gel was polymerised by adding 150 µl 10% ammonium persulphate and 6 µl TEMED. Acrylamide (electrophoresis grade) was obtained from Fisons (Loughborough, U.K.) and the bisacrylamide from Kodak Chemicals Ltd. (Liverpool, U.K.). Ultrapure ammonium persulphate was from Pharmacia (Uppsala, Sweden) and TEMED from Sigma.

Western transfer buffer Western buffer consisted of 20% (v/v) methanol, 0.3% (w/v) Tris-base and 1.5% (w/v) glycine.

Western blot-blocking solution This solution was prepared immediately prior to use and consisted of 10% (w/v) powdered skimmed milk (Marvel, Cadburys) in 100 ml of PBS.

Western blot-colour development 0.06% (w/v) 4-chloro-1-napthol (supplied by Bio-rad) was dissolved in 20 ml methanol, and added immediately to 0.06% (v/v) hydrogen peroxide in 100 ml of PBS. The solution was mixed immediately before use and developed in the absence of light. The colour reaction was stopped by air drying the filter.

2.2.2 *In vivo* labelling of transfected proteins with [^{35}S]-methionine 293 cells were grown in a 90 mm^2 tissue-culture Petri-dishes and when confluent, were induced to take up foreign DNA. After 40 hours the supernatant was removed and the cells incubated for four hours in 1 ml of prewarmed methionine - free medium containing 100 µCi/ml of [^{35}S]-methionine (specific

activity:>800 Ci/mmol). Cells were then harvested by removal with a sterile policeman and prepared by immunoprecipitation.

2.2.3 Immunoprecipitation of viral proteins Cells were collected as a pellet after a 30 second centrifugation at 10,000 x g. Pellets were resuspended in 1 ml of ice-cold TBS/PMSF. Cells were pelleted once more at 10,000 x g. The final cell pellet was resuspended in 1 ml of TBS/PMSF and the cells lysed by the addition of NP-40 detergent (final concentration of 2% {v/v}), for half an hour on ice. Cellular debris was removed by centrifugation at 10,000 x g for 1 minute. The supernatants were precleared by incubation with 50 µl of IMP for half an hour on ice (to reduce non-specific background). After the removal of the precipitate by centrifugation, supernatants were reacted with 10 µl polyclonal anti-RS virus serum or specific MAb, for more than 4 hours on ice. Samples were then incubated with IMP (at an IMP:Ab volume ratio of 10:1) on ice for 60 minutes. Immune complexes were pelleted at 10,000 x g and washed twice with cold RIP washing buffer and the pellets resuspended in SDS-PAGE sample buffer (reducing or non-reducing) and stored at -20°C. Before electrophoresis, samples were boiled for 3 minutes and clarified from the IMP by a 1 minute centrifugation at 10,000 x g.

2.2.4 Preparation of SDS-PAGE samples Cell monolayers were scraped into suspension with sterile policeman. Cells were collected as a pellet after a 30 second centrifugation at 10,000 x g. Pellets were resuspended in 1 ml of ice-cold PBS/PMSF. Cells were pelleted once more at 10,000 x g. The final cell pellet was resuspended in 100 µl of PBS/PMSF and the cells lysed by the addition of NP-40 detergent (final concentration of 2% {v/v}), for half an hour on ice. Cellular debris was removed by centrifugation for 2 minutes. An equal volume of 2 X SDS-PAGE sample buffer (reducing or non-reducing) was then added to the protein-containing supernatant. Before electrophoresis, samples were boiled for 3 minutes.

2.2.5 Polyacrylamide gel electrophoresis (SDS-PAGE) The discontinuous system of Laemmli. (1970) was used. Resolving gel containing 10% polyacrylamide, were cast directly inside the assembled gel plates (250 x 220 mm with 1.5 mm spacers). Stacking gel was applied on top of the resolving gel once this had set. Gels were run overnight at a constant current of 14 mA until the dye front had reached the bottom of the gel.

2.2.6 Fluorography For radiolabelled-protein samples, gels were fixed with two changes of fixer for 1 hour. Fluorography was carried out by gentle shaking of fixed gels in Amplify™ for 30 minutes. Gels were then dried onto a 3MM Whatman paper under vacuum for 2 hours (BIO-RAD gel drier) and exposed to pre-flashed X-ray film at -70°C.

2.2.7 Western blotting For most experiments viral proteins were transferred directly to nitro-cellulose, the electrophoretic transfer of which was performed according to Towbin *et al.* (1979). Both immunoprecipitated and total cell lysates were used. The electrophoresis gels were equilibrated, after running, in Western transfer buffer for 30 minutes and then laid on pre-wetted 3 MM Whatman paper, resting on a fibre pad which was placed on the bottom half of a gel holder of the Trans-blot™ apparatus (BIO-RAD). A piece of nitro-cellulose (21 x 16 cm), pre-soaked in transfer buffer, was applied tightly onto the gel. Assembly was completed by placing another pre-wetted filter paper on top of the nitro-cellulose and a second fibre pad on top of the paper. The gel holder was inserted into the Trans-blot™ tank, which had been filled with transfer buffer, with the nitro-cellulose towards the anode. Transfer took place at a constant voltage of 70 V for 4 hours. The gel holder was then taken out of the tank, opened and the nitro-cellulose sheets then incubated in 10% blocking solution for 30 minutes at room temperature to allow blocking of non-specific binding sites. Blots were then washed four times in PBS-T and incubated with antibody diluted 1:200 in 1% (w/v) BSA; PBS-T for 3 hours at room temperature. Unbound antibody was removed by washing four times with PBS-T and blots incubated with biotinylated second antibody (anti-mouse, -human or -bovine) diluted 1:400 in 1% (w/v) BSA; PBS-T for one hour at room temperature. After another four

washes in PBS-T, blots were incubated with biotinylated peroxidase - streptavidin complex, diluted 1:3000 in 1% (w/v) BSA ; PBS-T for one hour. Blots were finally washed 3 times in PBS-T and twice in PBS alone (to remove residual detergent) prior to colour development.

2.3 RECOMBINANT DNA TECHNIQUES

2.3.1 Bacterial Cells used in cloning experiments *E. coli* strain TG2 was used in all experiments. The genotype is shown below:

supE hsdΔ5 thi Δ(lac-proAB) Δ(srl-recA)306::Tn10(Tet^r)

2.3.2 Bacterial plasmids and phage used in cloning experiments All DNA subcloning was carried out in the Bluescribe™ (pBSM13) plasmid obtained from Stratagene Cloning Systems, (San Diego, California, U.S.A.). DNA sequencing was performed using the bacteriophage M13 mp18/19 vectors described by Messing and Vieira. (1982).

2.3.3 Materials

Luria-Bertani (LB) medium This medium contained 1% (w/v) tryptone, 0.5% (w/v) sodium chloride and 1% (w/v) yeast extract. Medium was made sterile by autoclaving at 121°C for 20 minutes. LB agar plates were made with the same ingredients but containing 1.5% Bacto-agar and sterilised as above. Indicator plates for selection of recombinant clones also contained 0.5 mM X-Gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside), 1 mM IPTG (isopropyl-thiogalactoside) and ampicillin (100 μg/ml).

H-Top medium H-top medium was used in transformation of M13 phage and contained 0.8% (w/v) Tryptone, 0.5% sodium chloride, 0.5% (w/v) yeast extract and 0.6% (w/v) Bactoagar. Media was sterilised by autoclaving for 20 minutes at 121°C.

TE Stock solutions were prepared, consisting of 10 mM Tris-HCl (pH 8.0) and 0.1 mM EDTA

TNE Stock solutions were prepared, consisting of 10 mM Tris-HCl (pH 8.0), 100 mM sodium chloride and 1 mM EDTA

10 X TBE Stock solutions were prepared, consisting of 0.89 M Tris-HCl (pH 8.0), 0.89 M boric Acid and 20 mM EDTA (pH 8).

6 X Tracking buffer Stock solutions were prepared, consisting of 0.25% bromophenol blue and 30% glycerol.

3 X Boiling mix Stock solutions were prepared, consisting of 15% 2-mercaptoethanol, 6% SDS, 100 mM Tris (pH 6.7) and 30% glycerol. Bromophenol blue was added till a dark blue colour was obtained.

Phenol Distilled phenol was melted at 65°C and equilibrated against several changes of equal volumes of TNE.

Phenol/Chloroform Consisted of a mixture of equal volumes phenol and chloroform.

2 X HEPES Consisted of 0.3 M sodium chloride, 0.01 M potassium chloride, 0.01 M glucose and 0.05 M HEPES in distilled water. The solution was sterilised by filtration and stored at 4°C as a stock solution.

100X Phosphate salts A stock solution was prepared, consisting of 0.08 M di-sodium hydrogen orthophosphate (anhydrous). The solution was sterilised by filtration and stored at 4°C.

TS Stock solutions were prepared from TD stock solution by the addition of 0.5 ml 20 mg/ml Ca^{2+}/Mg^{2+} salts and then sterilisation by filtration. Stocks were kept at 4° C.

Isotonic lysis buffer This solution was made up with 0.15 M sodium chloride; 1.5 mM magnesium chloride; 0.65% NP40 (v/v) and 1 mM Tris. The solution was made up to pH 7.8 and sterilised by autoclaving. The solution was stored at room temperature until use.

Phenol extraction buffer The solution comprised 7 M Urea; 0.4 M sodium chloride; 1 mM EDTA; 1% SDS and 1 mM Tris. The solution was made up to pH 7.8 and sterilised by autoclaving. The solution was stored at room temperature until use.

2.3.4 Small scale plasmid preparation (mini preps) The method used was a slightly modified version of Serghini et al. (1989). 1,5 ml of culture grown overnight at 37°C with minimal shaking was centrifuged for 5 minutes at 15,000 x g and the cell pellet resuspended in 100 µl of TNE buffer and 100 µl of phenol/chloroform, and centrifuged for a further 5 minutes at 15,000 x g. The top aqueous layer was transferred to fresh eppendorf tubes to which 35 µl of a 10 M solution of ammonium acetate and 270 µl of cold 100% ethanol had previously been added. The mixture was incubated at 4°C for 20 minutes after which the nucleic acids were pelleted for 10 minutes at 15,000 x g. The alcohol supernatant layer was removed and traces of ethanol removed by desiccation under vacuum. The pellet of nucleic acids was resuspended in either 30 or 40 µl of distilled water since the yields of plasmid DNA obtained were dependent upon plasmid copy number. Aliquots of 10 µl were used in restriction enzyme analysis in the presence of 5 µg of RNAse A.

2.3.5 Large scale plasmid preparation (Maxi preps) Plasmid DNA was prepared using the alkali lysis method of Birnboim and Doly. (1979). Briefly, cells from a 500 ml overnight culture were pelleted at 2,000 x g for 20 minutes and resuspended in 8 ml of

solution I (50 mM glucose, 10 mM EDTA (pH 8.0) and 25 mM Tris-HCl solution containing lysozyme at a final concentration of 5 mg per ml). After a 10 minute interval, 10 ml of solution II (1% SDS and 0.2 M sodium hydroxide) was added, mixed by inversion and kept on ice for 10 minutes. 12 ml of cold 5M potassium acetate was added , mixed by inversion and iced for another 10 minutes followed by centrifugation for 30 minutes at 30,000 x g, 4°C. The cleared supernatant was subsequently treated with an equal volume of a phenol/chloroform solution and the two immiscible phases separated by centrifugation for 30 minutes at 2,000 x g, 4°C. The aqueous layer was transferred to fresh tubes and 0.6 times the volume of isopropanol was added and the mixture left at room temperature for 15 minutes. DNA was recovered by centrifugation for 30 minutes at 2,000 x g, 4°C. The alcohol layer was removed and the nucleic acid pellet desiccated under vacuum. The pellet was resuspended in 1.5 ml distilled water RNAse treated (10 µg/ml) at 37°C for 30 minutes.

2.3.6 Caesium chloride gradient DNA banding The method used was as described in Maniatis *et al.* (1982). DNA prepared using the large scale plasmid prep described above was diluted to a final volume of 3.8 ml with distilled water to which 4.2 g of caesium chloride and 50 µl of a 10 mg/ml solution of ethidium bromide were added. The mixture was then transferred to Beckman heat seal tubes (5 ml) and filled up with a solution of 3.8 ml water, 4.2 g caesium chloride and 50 µl of 10 mg/ml solution of ethidium bromide. The tubes were sealed and centrifuged for at least 18 hours at 400,000 x g, 20°C. Banded DNA was excised from the caesium gradient and the ethidium bromide removed by extraction with 6 rounds of equal volumes of amyl alcohol. Plasmid DNA was then dialysed overnight against 5 litre of a 0.1% solution of SSC in order to remove the caesium chloride. The DNA was precipitated by addition of 1/10 volume of a 3 M sodium acetate solution and 2.5 times the volume of 100% ethanol, stored at either -70°C for 4 hr or -20°C overnight and then pelleted at 10,000 x g at 4°C for 30 minutes. Traces of alcohol were removed by desiccation under vacuum and the total concentration of plasmid DNA estimated by optical density reading at 260 nm of resuspended DNA.

2.3.7 Preparation of competent *E.coli.* Method used was the same as that described in Maniatis *et al.* (1982). 70 ml of LB medium was inoculated with 1 ml of *E. coli* cells were grown at 37°C with minimal shaking to an optical density (O.D.) reading at 650 nm of ~0.4 and kept on ice for at least 30 minutes. 50 ml of cells were pelleted at 2,000 x g, 4°C for 5 minutes, resuspended in 25 ml of a cold freshly made solution of magnesium sulphate and pelleted once again at 2,000 x g, 4°C for 5 minutes. The pelleted cells were then resuspended in 2.5 ml of a cold freshly made solution of calcium chloride and kept on ice for 1 hour before beginning the transformation procedure.

2.3.8 Transformation of competent *E. coli* with plasmid DNA The method of bacterial transformation was derived from that of Cohen *et al.* (1972) . Plasmid DNA was added to 100 µl aliquot of competent cells and kept on ice for 30 minutes, shocked at 42°C for 2 minutes, replaced on ice for a further 5 minutes and then left at room temperature for 5 minutes. 1 ml of sterile LB medium was added and the cells incubated at 37°C with minimal shaking for 40 minutes. 0.25 ml aliquots were then plated onto LB agar plates containing the appropriate selection antibiotic.

2.3.9 Transformation of competent *E. coli* with phage M13 DNA M13 vectors derived by Messing and Vieira. (1982) were transformed into competent TG2 bacteria using a protocol similar to that described above except that after the 42°C shock, cells were kept on ice for 5 minutes and then added to a mixture of 3 ml of molten H-Top agar containing 75 mM of X-Gal, 150 mM of IPTG and 80 µl of an *E.coli* culture that had been grown overnight, diluted 1/100 with LB medium and incubated at 37°C with minimal shaking for 1 hour. This whole mixture was then poured onto LB plates, the H-top allowed to set before incubating overnight at 37°C.

2.3.10 Preparation of phage M13 single stranded DNA Using the method described in Maniatis *et al.* (1982), 5 ml of LB medium was inoculated with M13 and grown for 6

hours at 37°C with minimal shaking. 1.5 ml of M13 infected *E. coli* cells were pelleted at 10,000 x g for 5 minutes. 1.2 ml of the supernatant was transferred to fresh tubes to which 250 µl of a 20% (v/v) polyethylene glycol 6000/3 M sodium chloride solution had been added. The two were mixed by inversion and kept at room temperature for 30 minutes. The PEG precipitated M13 was pelleted (10,000 x g, 4 minutes) and the supernatant removed. Any remaining supernatant was removed after re-spinning the tubes for 20 seconds at 10,000 x g. The cell pellet was then resuspended in 120 µl of Tris-HCl (pH 8.0) and 50 µl of buffered phenol was the added. The two immiscible phases were the separated by centrifuging at 10,000 x g for 5 minutes. 100 µl of the top aqueous layer was transferred to fresh tubes and 50 µl of chloroform was added and again the two phases separated as before. 80 µl of the aqueous layer was transferred to fresh tubes and 10 µl of a 3 M sodium acetate solution and 250 µl of cold 100% ethanol were added.

2.3.11 Transfection of mammalian cells Plasmid DNA was introduced into mammalian cells using a modified protocol of Chen and Okayama.(1987). Before the start of the experiment, the medium was replaced with supplemented DMEM plus 10 % Foetal calf serum for at least 3 hours. The DNA-calcium phosphate precipitate was then prepared by adding 10-20 µg DNA and 3 µg salmon sperm DNA to 100 µl 1.25 M calcium chloride and making up the volume to 500 µl with sterile distilled water. This solution was then added slowly to a solution of 10 µl 100 X phosphate salts solution and 490 µl 2 X HEPES. The mixture was allowed to precipitate for 0.5 hours at room temperature. The media was then removed from the cells and 1 ml of supplemented DMEM plus 10% FCS added, plus 500 µl of the DNA-phosphate precipitate. The precipitate was allowed to settle on the cells for 5 hours whilst the cells were maintained in a 37°C incubator under 5% CO_2. After 5 hours the cell media was removed and the cell monolayer induced to take up the bound DNA by the addition of 2 ml 80% glycerol in TS. The glycerol shock was stopped after 50 seconds by removal of the glycerol and washing once with 5 ml TS solution. This was immediately removed and 10 ml supplemented DMEM plus 10% calf

serum was added to the cells. The cells were then incubated for 2 days at 37°C under 5% CO_2 before harvesting.

2.3.12 Restriction endonucleases DNA to be analysed by restriction digests was prepared either by the mini-,maxi- or large scale preparation methods described above. All digests were carried out in the manufacturer's recommended buffers in the presence of 1-2 units of enzyme and 50μg/ml bovine serum albumin. All digests were carried out at the recommended temperature for between 2-6 hours. Restriction enzymes were supplied by either Amersham, Gibco-BRL or Pharmacia.

2.3.13 De-phosphorylation of linearised vector DNA 2-5 units of calf intestinal phosphatase (CIAP) were added during the last hour of restriction endonuclease digestion of vector DNA so as to minimise vector re-ligation by removal of the 3' terminal phosphate of linearised DNA.

2.3.14 Blunting 3' overhanging DNA T4 DNA polymerase when compared with DNA polymerase I (Klenow Fragment) possesses an equivalent 5→3' polymerase but more active 3→5' exonuclease activity. In the presence of excess deoxynucleotides, the amount of 3→5' exonuclease activity is reduced considerably when compared to 5→3 polymerase activity. Additionally it's 3→5' exonuclease activity is only functional on ssDNA or on dsDNA with 3' overhangs but not on 5' overhangs (3' recess). Thus, T4 DNA polymerase was chosen in place of DNA polymerase I (Klenow fragment) for blunting of 5' and 3' overhangs of dsDNA following digestion with restriction enzymes. Restriction endonuclease treated DNA was first phenol/chloroform extracted and the ethanol precipitated by the method used in the mini-prep protocol. DNA was then blunted in the presence of 0.25 mM dATP, dTTP, dGTP, dCTP, 33 mM Tris-HCl (pH 7.5), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT, 0.1 μg BSA and 1-2 units of T4 DNA polymerase and the reaction incubated at 37°C for 45 minutes. Blunted DNA

was then either purified from LMP agarose gels or phenol/chloroform extracted and ethanol precipitated as described above.

2.3.15 DNA Ligation Ligation of DNA fragments was carried out overnight at 16°C in the presence of 66 mM Tris-HCl (pH 7.6), 5 mM magnesium chloride, 5 mM DTT, 0.1 µg BSA, 1 mM hexamminecobalt chloride, 1 mM ATP, 0.5 mM spermidine-HCl and 1-2 units T4 DNA ligase. Typically the final reaction volume was between 20-25 µl. DNA fragments purified from LMP agarose gels were heated to 65°C for 10 minutes and the ligation carried out in a final volume of between 60-100 µl, where the final volume was in at least 8 fold excess of the volume of DNA fragment added.

2.3.16 Isolation of RNA from RS virus-infected cells This method is based on that of Kumar and Lindberg.(1972). When infected cells reached 40% syncytium formation the cells were scrapped off the monolayer using a sterile policeman. The cells were pelleted at 10,000 x g for 5 minutes in a MSE™ chillspin. Pellets were then washed once in 10 ml PBS and lysed on ice in 5 ml isotonic lysis buffer for 30 minutes. 5 ml of phenol extraction buffer was then added along with 10 ml phenol/chloroform. The solutions were mixed by inversion and the mixture centrifuged for 10 minutes at 10,000 x g, the clear top-layer being extracted. To this was added a further 10 ml phenol/chloroform and the preceding step repeated a further time. The top-layer was cleared of residual phenol by the addition of 8 ml chloroform, further mixing and spinning for another 10 minutes at 10,000 x g. Total RNA was ethanol-precipitated, centrifugation at 10,000 x g for 20 minutes, and the resulting pellet dried under vacuum. The pellet of total RNA was then resuspended in 50 µls sterile distilled water and stored at -20°C.

2.3.17 cDNA synthesis from purified RNA The 1st strand of cDNA was synthesised from the total RNA from RS virus-infected cells using avian myeloblastosis virus (AMV) reverse transcriptase (Life Sciences, Inc.). Reverse transcription was carried out in the presence of 0.5 mM each of dATP, dCTP, dGTP, dTTP, 50 mM Tris-HCl (pH 8.3), 6

mM magnesium chloride, 40 mM potassium chloride, 3 μg of pd(T) (12-18; Pharmacia) and 10 units AMV reverse transcriptase.. The reaction was incubated at 42°C for 45 minutes, phenol/chloroform extracted and ethanol precipitated. cDNA was resuspended in 30 μl of distilled water. 3 μl aliquots were used to amplify DNA fragments using the polymerase chain reaction.

2.3.18 Amplification of DNA by polymerase chain reaction (PCR) Amplification was based on the method of Saiki *et al.* (1988) and was achieved by 30 cycles of ; DNA denaturation at 93°C for 0.75 minutes, oligonucleotide annealing at 56°C for 0.75 minutes and extension at 74°C for 2 minutes in the presence of 10 mM potassium chloride, 10 mM ammonium sulphate, 20 mM Tris-HCl (pH 8.8), 2 mM magnesium sulphate, 0.1% (v/v) Triton X-100, 100 μg/ml BSA, 0.5 mM each of dATP, dTTP, dCTP, dGTP and 1 unit of Vent™ DNA polymerase (New England Biolabs, MA, USA).

2.3.19 Oligomer primers used in PCR and sequencing experiments Two primers were used for cDNA synthesis and PCR amplification experiments; PCRF1 which is complimentary to the 3' end of the fusion gene; and PCRF2, complimentary to the 5' end. Six internal primers (SEQF1 → SEQF6) and the M13 primer (-40)(**Table 6**) were used to determine the gene sequence.

2.3.20 Agarose gels Restriction endonuclease digested DNA was visualised under UV illumination after running on either a 1% or 2% agarose gel depending upon the size of DNA fragments. Agarose was dissolved in 1 X TBE solution containing 0.05 μg/ml of ethidium bromide and then cast in either a 6 x 8, 8 x 10.5, 11.5 x 14 or 20 x 25 cm horizontal tray. 6 X tracking dye buffer was added to samples to be analysed and the gel run in the presence of 1 X TBE buffer at between 100 and 180 volts.

Table 6 Oligomers used during sequencing, cDNA synthesis and PCR amplification

PRIMER	NUCLEOTIDES	SEQUENCE 5' → 3' (mRNA sense)
PCRF1	30mer	GGATCCCGGGGCAAATAACAATGGAGTTGC
PCRF2	33mer	GGATCCCGGGTTACTAAATGCAATATTATTTAT
-40	17mer	CAGCACTGACCCTTTTG
SEOF1	18mer	CACATTACCTTGTCTACG
SEOF2	16mer	CGACATAGATTCCAGG
SEOF3	16mer	GGTCCCTTAAATCACA
SEOF4	17mer	GTGTGTAGGGGAGATAC
SEOF5	15mer	GAGACGTTGTAACTG
SEOF6	17mer	CAGAGATACATTTTCCA

Primers PCRF1 and PCRF2 were complimentary to the 5' (mRNA sense) ends of the fusion protein gene and were used to amplify the target gene prior to cloning and sequencing. At the immediate 5' ends of these oligomers were included a BamH1 (GGATCC) and Sma1 (CCCGGG) restriction endonuclease DNA sequence. These were used to clone the PCR product into the vector DNA. All other primers were used for sequencing experiments and were designed to anneal to different unique DNA sequences within the fusion gene.

2.3.21 Sequencing reactions All sequencing reactions were carried out using a method derived from that of Sanger *et al.* (1977) and as described in the sequenase™ kit instructions. Radioactive labelling of the target DNA was achieved using α-labelled ^{35}SdATP.

2.3.22 Sequencing Gels Sequencing reactions were run on a 5.7% acrylamide, 0.3% bis-acrylamide, 7 M urea dissolved in 1X TBE gel. To polymerise the gel, 265 µl 10% (w/v) ammonium persulphate and 112 µl TEMED were added. A 31 x 38.5 cm gel was cast horizontally between two glass plates separated by a 0.4→1.2 mm wedge spacer. The gel was run in the presence of 1 X TBE running buffer at a constant 80 watts.

CHAPTER THREE

CHAPTER THREE

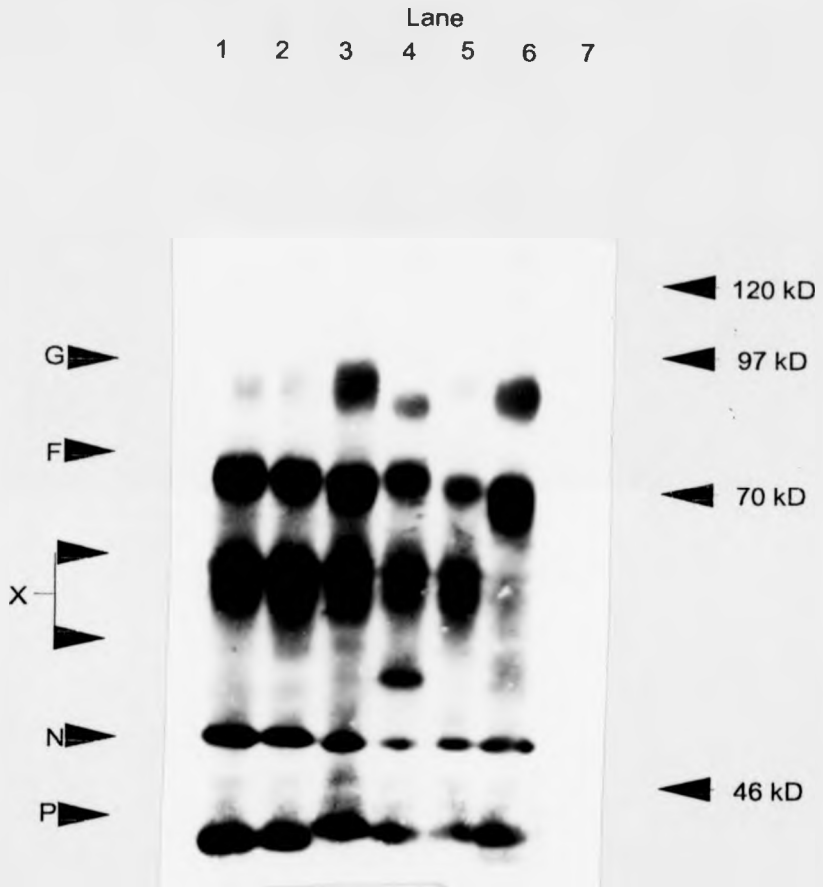
Comparison of fusion protein genes and their products from RS virus subgroup A isolates

3.1 Introduction Investigations of the role of RS virus proteins in the immune response have established the importance of the fusion protein as the major antigenic unit responsible for protection (**Section 1.8.1**). Characterisation by limited nucleotide sequencing and restriction mapping of five discrete subgroup A isolate-types circulating during a winter 1989 epidemic in the Birmingham, U.K. area (Cane and Pringle, 1991; SH lineages SHL1 to SHL5) has led to the hypothesis that strain-variation may lead to changes in antigenicity, accounting for reinfection. Evidence supporting this hypothesis comes from work by Sullender *et al.* (1991) and Cane *et al.* (1991) who observed extensive sequence variation in the G protein gene among naturally occurring RS virus isolates. Changes observed in the attachment (G) protein gene included frame shift mutations leading to major alterations in the carbohydrate content (Garcia-Barreno *et al.*, 1990; Rueda *et al.*, 1991). Therefore the major aim of this project was to determine whether there was similar variation in the fusion protein gene of RS virus subgroup A isolates, and determine the potential effects of this variation on fusion protein structure in an attempt to explain whether amino acid variation could lead to antigenic variation.

3.2 Results All RS viruses analysed belonged to subgroup A and included the prototype strain A2 and five natural isolates, RSB89-5857, RSB89-6190, RSB89-6256, RSB89-6614 and RSB89-1734, representing isolate-types SHL 1 to 5 respectively (described by Cane and Pringle, 1991). All RS viruses were grown in the BS-C-1 G23 cell line.

3.2.1 SDS-PAGE variation Analysis of the fusion proteins of both subgroup A (strains A2 and RSS-2 and isolates RSB89-1734 and RSB89-1202) and subgroup B (isolates RSB89-5987 and RSB89-5832) viruses by SDS-PAGE under non-reducing conditions (**Figure 7**) revealed characteristic differences in electrophoretic mobility of

Figure 7. SDS-PAGE analysis of subgroup A and subgroup B RS viral proteins .



Total protein was harvested from virus-infected BS-C-1 cells and analysed by 10% SDS-PAGE under non-reducing conditions and the proteins transfered via Western-blotting onto a nitrocellulose filter which was developed with bovine polyclonal anti-RS virus antibody as described in Chapter 2. Four RS viral proteins are shown,the attachment protein (G), the fusion protein (F), the nucleoprotein (N) and the phosphoprotein (P). Fusion protein breakdown products (X), identified by anti-F monoclonal antibodies (data not shown), are also seen. In lane 4 the 48 kD product might possibly represent a precursor protein of the mature F. Lanes 1 and 2 are subgroup B virus proteins (RSB89 -5987 and RSB89-5832 respectively), lanes 3, 4, 5 and 6 are subgroup A viral proteins (RSB89-1202, RSB89 -1734, RSS-2 and A2 respectively). Lane 7 is a negative control of BSC-1 cellular protein.

viral proteins between the two RS viral subgroups. The phosphoprotein (P) has accelerated mobility in the subgroup B viruses whereas the nucleoprotein (N) appears conserved between subgroups. The attachment glycoproteins of each virus sample show differences in mobility as well as differences in their reactivity with the bovine polyclonal anti-RS antibody, indicating that there is heterogeneity both within and between RS virus subgroups. **Figure 7** shows also that there are differences in mobility of the fusion proteins both between and within subgroups. This suggested that variation in the fusion proteins of subgroup A RS viruses might be greater than previously reported.

The fusion protein mobility differences could be explained as a consequence of variation in the number and position of cysteine residues leading to alterations in the structure of the protein, however the cysteine residues are absolutely conserved in number and position throughout subgroup A viruses (Collins, 1991). Fusion protein mobility differences could also be a consequence of variation in the number and position of potential N-linked glycosylation sites, and there has been a difference in the number of glycosylation sites reported between strain A2 and other subgroup A viruses (Collins, 1991), with strain A2 possessing five sites and strain RSS-2, Long and Edinburgh possessing six sites. An alternative explanation is that amino acid substitutions could alter the secondary structure of the fusion protein. It is this final suggestion that this study sought to address.

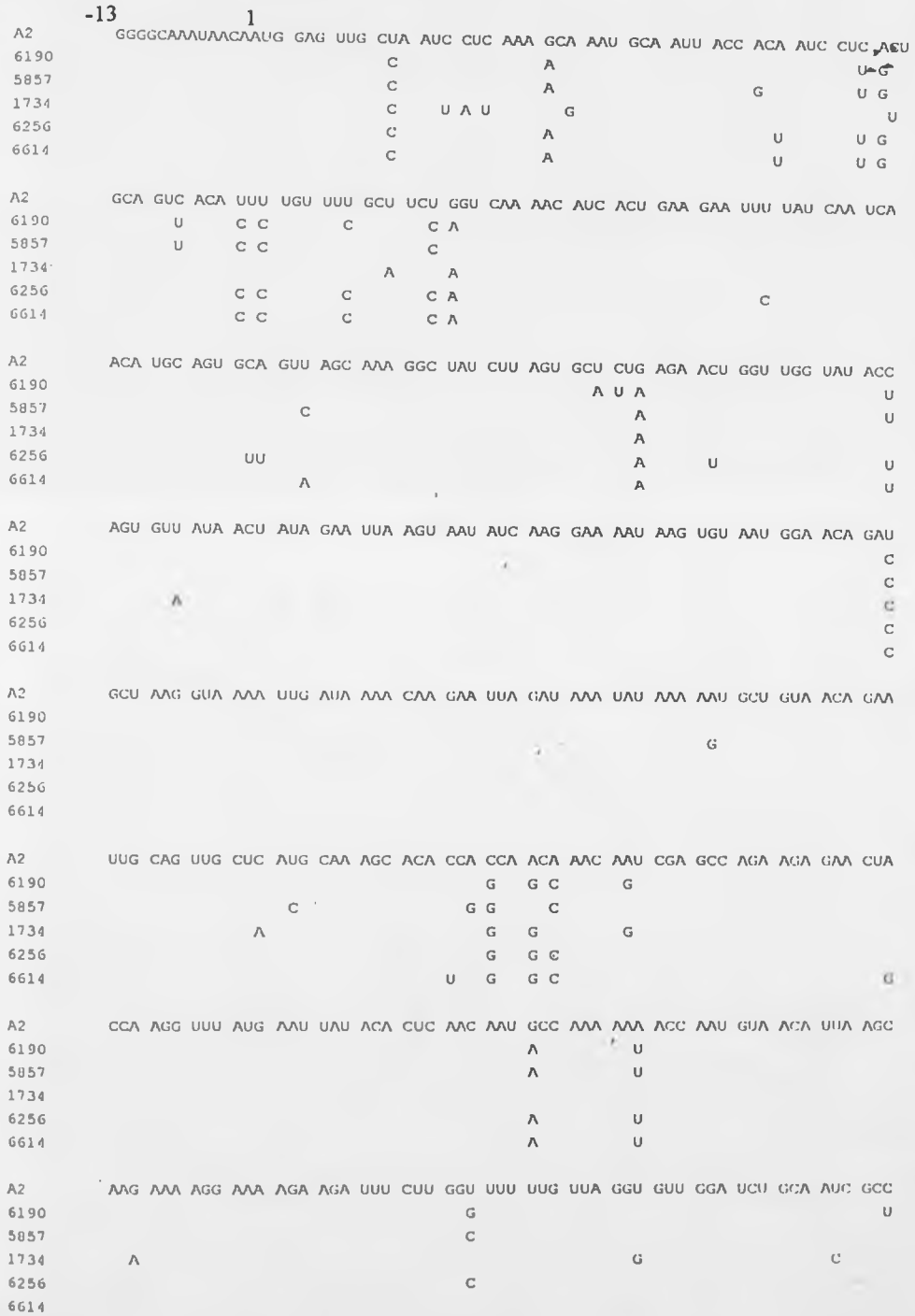
3.2.2 Variation in the fusion gene and its predicted product Determination of the nucleotide sequence of the fusion genes of five recent RS virus subgroup A isolates, representing five lineages defined by Cane and Pringle (1991), was carried out by PCR amplification of the major ORF via a cDNA intermediate generated from total RNA from RS virus-infected cells. At least two plaque-purified colonies of phage M13 containing the fusion genes of each isolate were sequenced by di-deoxy chain termination. Ambiguity in the sequence data was deemed to have possibly arisen from errors in either the cDNA or PCR process. Therefore extra fusion protein genes were sequenced from the product of a

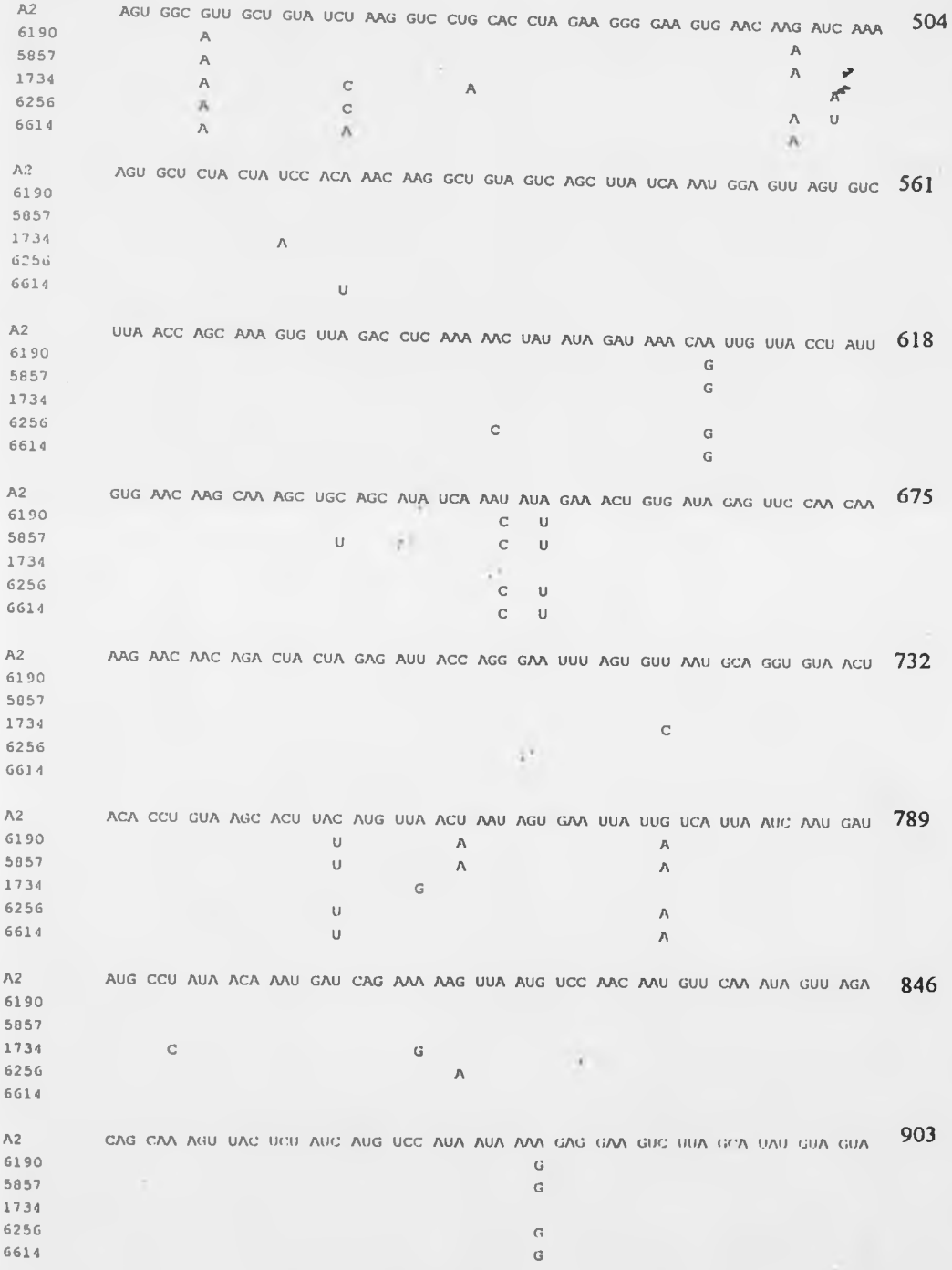
different cDNA/PCR reaction. The full nucleotide and predicted amino acid sequences for the five isolates and laboratory strain A2 are shown in **Figures 8 and 9**.

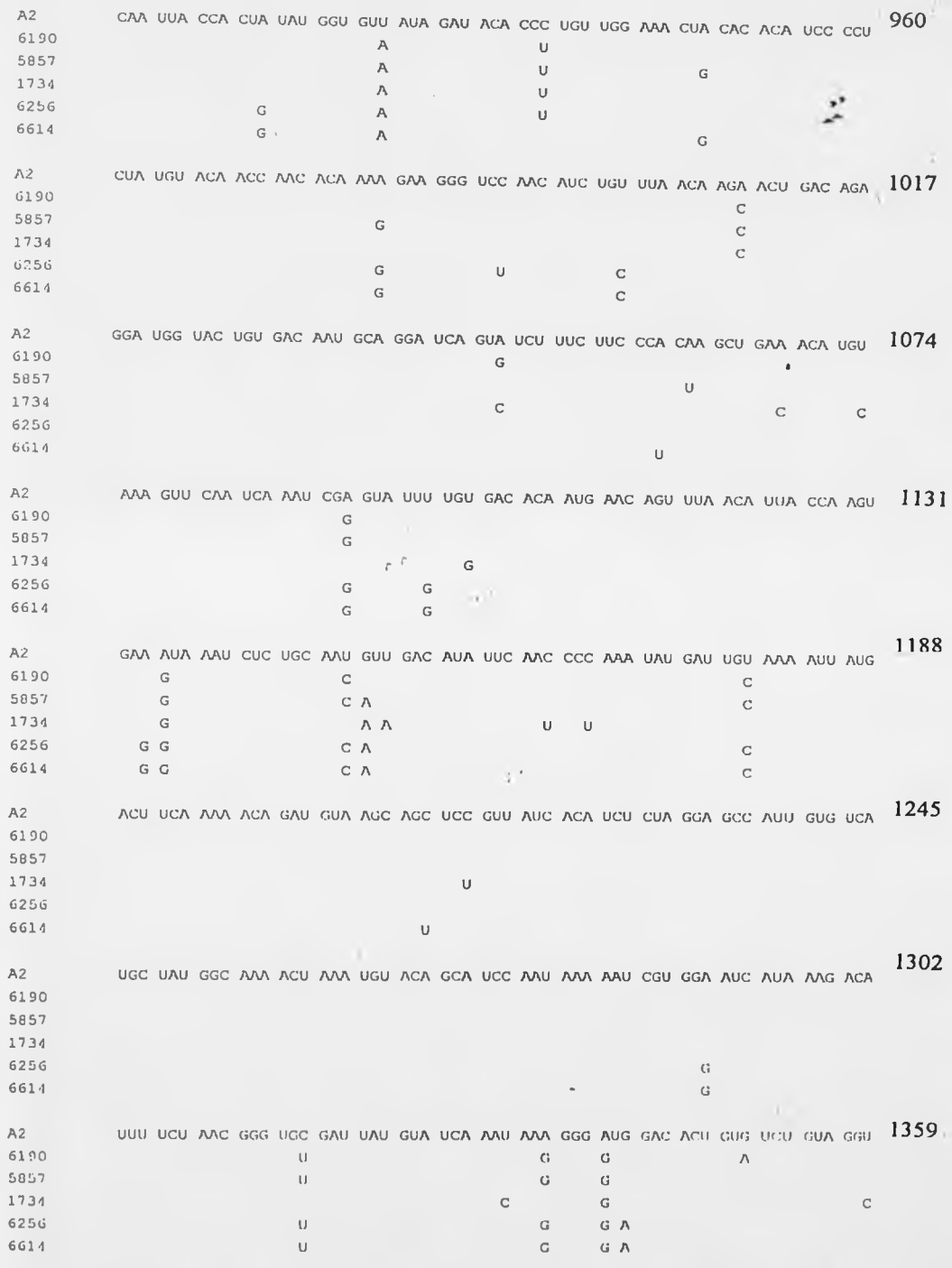
3.2.3 Fusion protein identity in subgroup A viruses The relative relatedness of both nucleotide and predicted amino acid sequences of each subgroup A virus are summarised in **Tables 7 and 8** respectively. **Table 7** summarises the nucleotide data presented in this thesis, and **Table 8** includes an additional three F protein gene sequences described in the literature. Comparison of the fusion gene product from subgroup A isolates shows that amino acid identity between subgroup A viruses is very high, ranging from 97% to 99.5%, whilst at the nucleotide level the pattern is very similar with homologies from 94.7% to 99.4%. The relationship of each isolate type is shown diagramatically in **Figure 10**.

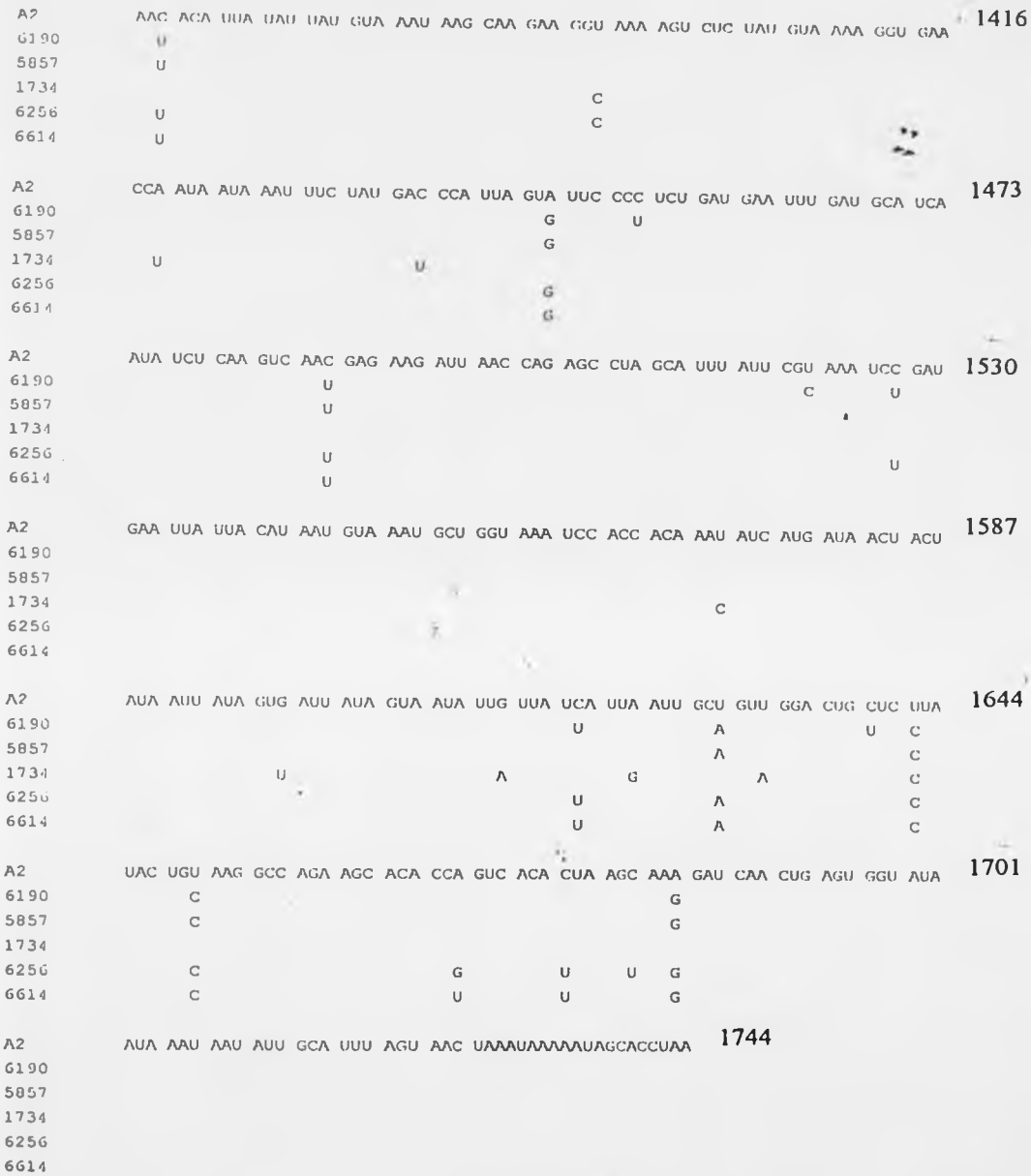
Within the fusion protein gene, amino acid changes are not uniformly distributed. The F_1 subunit is highly conserved (96.2%) among subgroup A isolates, whereas the F_2 subunit is less conserved (86.1%). Within the fusion protein gene the spread of nucleotide substitution is roughly uniform, however there are proportionally more amino acid changes in the F_2 subunit. Thus in the F_2-encoding region of the gene nearly 50% of nucleotide changes cause amino acid changes. Within the F_1-encoding region of the gene, only 16.5% of nucleotide changes lead to coding changes. Overall, most of the nucleotide changes (70%) occur in the third position of codon triplets, 15% occur in the first position and 9.5% in the second position. There are also double nucleotide changes (5.5%), usually involving the first and third positions. These results reflect a pattern of random nucleotide substitutions in the fusion gene. Most of the amino acid coding changes are conservative, maintaining size, charge and hydrophobicity profiles. The location and number of N-linked glycosylation sites and cysteine residues of all subgroup A isolates and strains, are highly conserved, with the exception of the loss of one glycosylation site in the F_2-subunit of strain A2 and isolate RSB89-1734.

Figure 8. The fusion gene sequence from subgroup A isolates compared to wild-type strain A2.



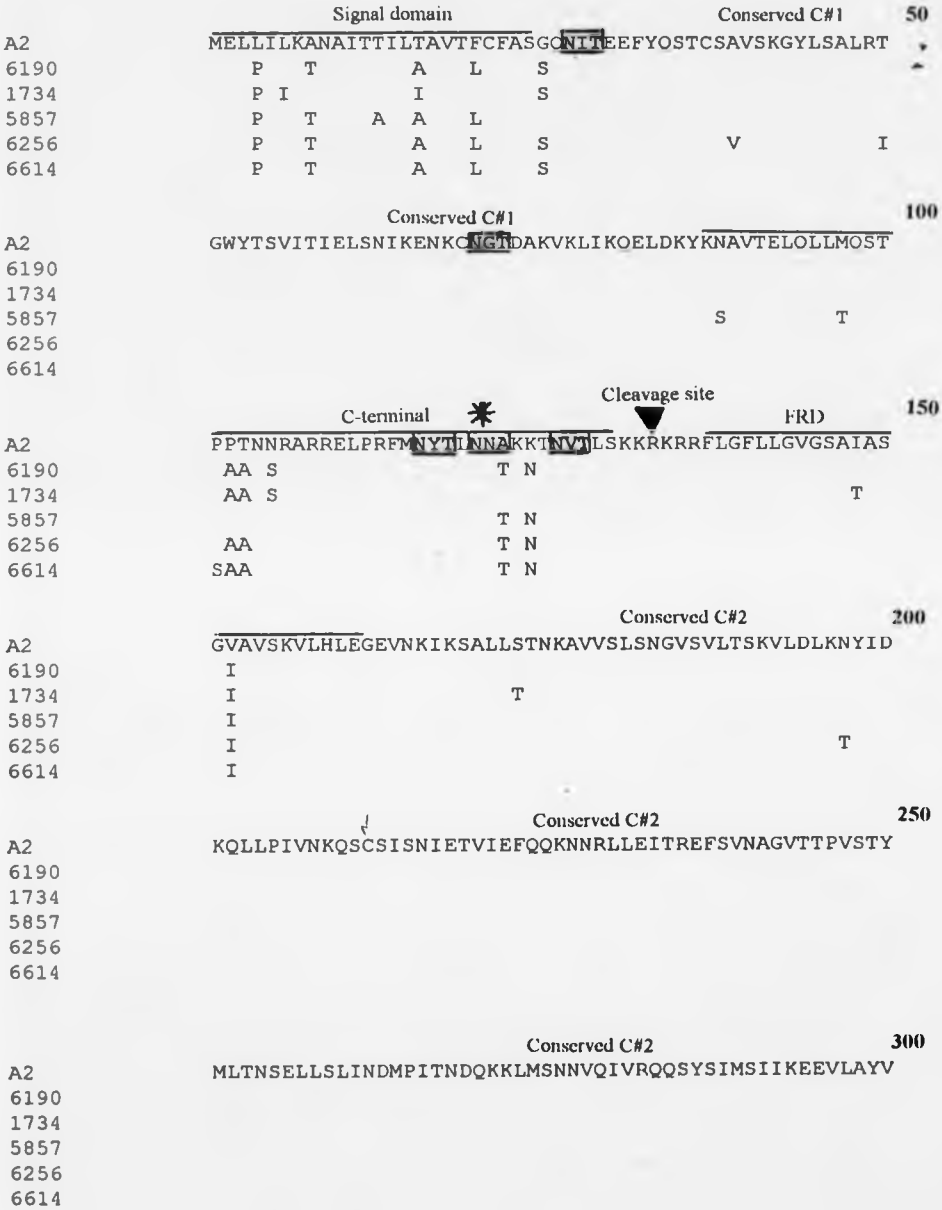

A2 AGU GGC GUU GCU GUA UCU AAG GUC CUG CAC CUA GAA GGG GAA GUG AAC AAG AUC AAA 504
6190 A
5857 A
1734 A C A
6256 A C
6614 A A A U
A2 AGU GCU CUA CUA UCC ACA AAC AAG GCU GUA GUC AGC UUA UCA AAU GGA GUU AGU GUC 561
6190
5857
1734 A
6256
6614 U
A2 UUA ACC AGC AAA GUG UUA GAC CUC AAA AAC UAU AUA GAU AAA CAA UUG UUA CCU AUU 618
6190 G
5857 G
1734
6256 C G
6614 G
A2 GUG AAC AAG CAA AGC UGC AGC AUA UCA AAU AUA GAA ACU GUG AUA GAG UUC CAA CAA 675
6190 C U
5857 U C U
1734
6256 C U
6614 C U
A2 AAG AAC AAC AGA CUA CUA GAG AUU ACC AGG GAA UUU AGU GUU AAU GCA GGU GUA ACU 732
6190
5857
1734 C
6256
6614
A2 ACA CCU GUA AGC ACU UAC AUG UUA ACU AAU AGU GAA UUA UUG UCA UUA AUC AAU GAU 789
6190 U A A
5857 U A A
1734 G
6256 U A
6614 U A
A2 AUG CCU AUA ACA AAU GAU CAG AAA AAG UUA AUG UCC AAC AAU GUU CAA AUA GUU AGA 846
6190
5857
1734 C G
6256 A
6614
A2 CAG CAA AGU UAC UCU AUC AUG UCC AUA AUA AAA GAG GAA GUC UUA GCA UAU GUA GUA 903
6190 G
5857 G
1734
6256 G
6614 G

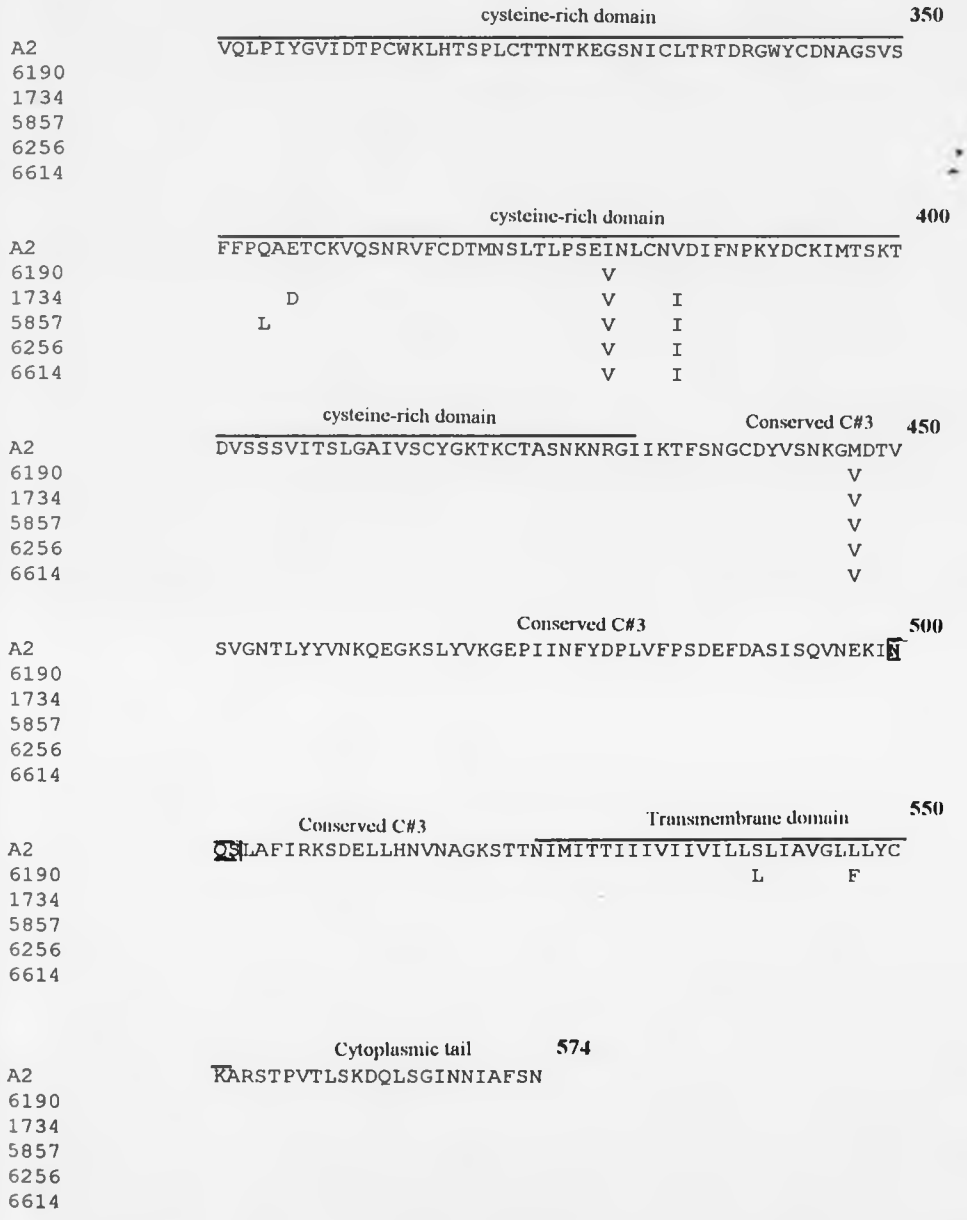

A2 CAA UUA CCA CUA UAU GGU GUU AUA GAU ACA CCC UGU UGG AAA CUA CAC ACA UCC CCU 960
6190 A U
5857 A U G
1734 A U
6256 G A U
6614 G A G
A2 CUA UGU ACA ACC AAC ACA AAA GAA GGG UCC AAC AUC UGU UUA ACA AGA ACU GAC AGA 1017
6190 C
5857 G C
1734 C
6256 G U C
6614 G C
A2 GGA UGG UAC UGU GAC AAU GCA GGA UCA GUA UCU UUC UUC CCA CAA GCU GAA ACA UGU 1074
6190 G
5857 U
1734 C C C
6256
6614 U
A2 AAA GUU CAA UCA AAU CGA GUA UUU UGU GAC ACA AUG AAC AGU UUA ACA UUA CCA AGU 1131
6190 G
5857 G
1734 G
6256 G G
6614 G G
A2 GAA AUA AAU CUC UGC AAU GUU GAC AUA UUC AAC CCC AAA UAU GAU UGU AAA AUU AUG 1188
6190 G C C
5857 G C A C
1734 G A A U U
6256 G G C A C
6614 G G C A C
A2 ACU UCA AAA ACA GAU GUA AGC AGC UCC GUU AUC ACA UCU CUA GGA GCC AUU GUG UCA 1245
6190
5857
1734 U
6256
6614 U
A2 UGC UAU GGC AAA ACU AAA UGU ACA GCA UCC AAU AAA AAU CGU GGA AUC AUA AAG ACA 1302
6190
5857
1734
6256 G
6614 G
A2 UUU UCU AAC GGG UGC GAU UAU GUA UCA AAU AAA GGG AUG GAC ACU GUG UCU GUA GGU 1359
6190 U G G A
5857 U G G
1734 C G C
6256 U G G A
6614 U G G A



Nucleotide sequence data (shown in codon triplets) for the fusion genes of five isolate-types of RS virus subgroup A. All isolates are described in **Section 2.1.2**. The fusion gene sequence for subgroup A isolates representative of each lineage is shown below the fusion gene sequence of the prototype strain A2. **Nucleotide numbers are aligned above the sequence data.**

Figure 9. Fusion protein sequences from RS virus subgroup A isolates compared to wild-type strain A2





Deduced amino acid sequence data for the fusion genes of five isolate-types of RS virus subgroup A. All isolates are described in **Section 2.1.2.** The fusion protein sequence for each isolate is shown below the fusion protein sequence of the prototype strain A2. N-linked glycosylation sites are denoted as hatched boxes. The variable N-linked glycosylation site is marked by an asterix above the box. Above the amino acid sequence are markings indicating areas of conserved residue identity throughout subgroup A viruses (C#1 to C#3). Three hydrophobic regions are also denoted : **1)** the signal peptide; **2)** the putative fusion related domain, FRD; **3)** the transmembrane domain. The cleavage site is also highlighted.

Table 7..Homology of the nucleotide sequences of the fusion genes of five subgroup A isolates and two laboratory strains, A2 and RSS-2

Isolate	A2	RSB89 - 6190	RSS-2	RSB89-1734	RSB89-5857	RSB89-6256	RSB89-6614
A2		96.7	97.3	97.0	97.0	96.2	96.5
RSB89-6190	57		98.3	95.3	98.5	97.8	97.8
RSS-2	46	29		95.8	99.5	98.0	98.0
RSB89-1734	51	81	73		95.3	94.9	94.9
RSB89-5857	52	26	9	81		98.0	98.1
RSB89-6256	66	38	35	88	35		98.4
RSB89-6614	60	38	34	88	32	24	

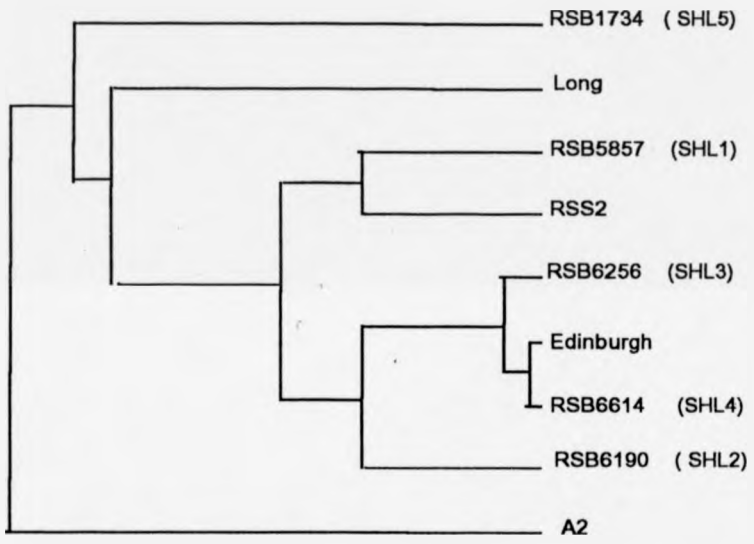
N.B. reading below the diagonal gives the actual number of nucleotide changes. Reading above the diagonal gives the percentage identity between isolates. The published sequence data of subgroup A strains A2 and RSS-2 are included for comparison. Sequence data for strain A2 was determined by Collins *et al.*, 1984b, and strain RSS-2 by Baybutt and Pringle, 1987.

Table 8..Homology of the predicted amino acid sequences in the fusion proteins of five isolate-types and four laboratory strains

Isolate	A2	RSB89 - 6190	RSS-2	Long	Edinb*	RSB89-1734	RSB89-5857	RSB89-6256	RSB89-6614
A2		97.2	97.6	98.1	97.7	97.6	97.6	97.2	97.6
RSB89-6190	16		98.3	97.9	98.8	97.7	97.9	98.6	98.9
RSS-2	14	10		97.9	99.1	97.4	99.4	98.6	98.9
Long	11	12	12		98.4	97.7	97.6	97.9	98.3
Edinb*	13	7	5	9		97.9	98.9	99.1	99.5
RSB89-1734	14	13	73	13	12		97.0	97.7	98.1
RSB89-5857	14	12	2	14	7	17		98.3	98.6
RSB89-6256	16	8	8	12	5	13	10		99.3
RSB89-6614	14	6	6	10	3	11	8	4	

N.B. Reading below the diagonal gives the actual number of amino acid changes, reading above the diagonal gives the percentage identity between isolates. The amino acid data was obtained from the following references, strain A2 (Collins *et al.*, 1984b), strain RSS-2 (Baybutt and Pringle, 1987), strain Edinburgh (Elango *et al.*, 1985a) and strain Long (Lopez *et al.*, 1988). *Edinb denotes the Edinburgh strain.

Figure 10 Relatedness of F gene sequences within subgroup A isolates



Graphical representation of relative relatedness of the fusion gene products from subgroup A strains and isolates. Fusion sequence data for SHL 1 - 5 isolates (This Thesis). Strain A2 sequence (Collins et al.,1984b); RSS-2 (Baybutt and Pringle, 1987); Edinburgh (Elango et al.,1985a) and Long (Lopez et al.,1988). Data analysed by the Clustal computer data handling program (Higgins and Sharp, 1988).

3.2.4 Variation in the F_2-subunit of the fusion protein The analysis of amino acid variation is based on the total number of coding changes for all five isolates. All coding changes are defined as amino acid differences in comparison with the published sequence data for a consensus of fusion gene cDNAs from strain A2 (Collins *et al.*, 1984b). The A2 strain is the only strain of RS virus fully sequenced, and is regarded as the prototype subgroup A RS virus.

The nucleotide sequence encoding the F_2-subunit possesses three defined regions; **1)** a hydrophobic signal domain; **2)** a conserved region (C#1) situated between the signal and C-terminal domains, and **3)** a C-terminal region (**Section 3.2.2 Figure 9**). The aggregate of all coding changes in comparison to the prototype strain A2 and present in the F_2-subunit are shown in **Table 9**.

The signal domain possesses the lowest degree of amino acid sequence identity throughout the fusion protein, with approximately 59.6% residue homology. Within the signal region domain (residues #1 to #22) there are equal numbers of coding and non-coding nucleotide changes. There are seven coding changes in this region of the gene, of which three occur in the first coding position of the triplet (residues #8; #13 and #16), two coding changes that in the second position (residue #4 and an alternative change in residue #16), and two multiple nucleotide substitutions (both in the first and third positions of the triplet), at residue #6 and #20. All but one of the amino acid changes observed are conservative with residue characteristics remaining unchanged. The exception to this is the coding change at residue #4 (Leu → Pro) that is a non-conservative change. Despite the low degree of homology, the overall pattern of hydrophobicity of the encoded protein as defined by Kyte and Doolittle (1982) remains very similar (data not shown). It therefore appears that amino acid changes are tolerated in this region as long as the overall hydrophobic nature is retained.

Table 9 Total number of coding changes in the F_2 subunit-encoding region of the fusion gene

Residue #	Amino acid	nucleotide change	Comments	Domain
#4	Leu → Pro	C**U**A → C**C**A	Non-conservative	Signal
#6	Leu → Ile	**C**U**C** → **A**U**U**	Conservative	Signal
#8	Ala → Thr	**G**CA → **A**CA	Conservative	Signal
#13	Thr → Ala	**A**CA → **G**CA	Conservative	Signal
#16	Thr → Ala	**A**CU → **G**CU	Conservative	Signal
	Thr → Ile	A**C**U → A**U**U	Conservative	Signal
#20	Phe → Leu	**U**U**U** → **C**U**C**	Conservative	Signal
#25	**Gly → Ser**	**G**GU → **A**GU	**Non-conservative**	**Conserved C#1**
#39	Ala → Val	G**C**A → G**U**A	Conservative	Conserved C#1
#50	Thr → Ile	A**C**U → A**U**U	Conservative	Conserved C#1
#88	Asn → Ser	A**A**U → A**G**U	Conservative	C-terminal
#97	Met → Thr	A**U**G → A**C**G	Conservative	C-terminal
#101	**Pro → Ser**	**C**CA → **U**CA	**Non-conservative**	**C-terminal**
#102	**Pro → Ala**	C**C**A → G**C**A	**Non-conservative**	**C-terminal**
#103	Thr → Ala	**A**CA → **G**CA	Conservative	C-terminal
	Thr → Ala	**A**C**A** → **G**C**C**	Conservative	C-terminal
#105	Asn →Ser	A**A**U → A**G**U	Conservative	C-terminal
#122	Ala → Thr	**G**CC → **A**CC	Conservative	C-terminal
#124	**Lys → Asn**	**A**A**A** → **A**A**U**	**Non-conservative**	**C-terminal**

Nucleotide bases that are underlined and in bold character denote the position and nature of change. Those changes considered to be of structural importance are shown in highlighted boxes.

The conserved region (C #1) of subunit F_2 (residues #24 to #87) has 98.5% homology within subgroup A viruses. There are nine non-coding nucleotide changes and three amino acid coding substitutions in this region of the fusion protein, with nucleotide changes in either the first codon position (residue #25) or the second position (residue #39 and #50). This region of the protein appears to be less tolerant of amino acid alterations, with two of the identified coding changes (residues #39 and #50) being conservative substitutions. The coding change at residue #25 (Gly → Ser) is a non-conservative change removing the potential of glycine to disrupt α-helixes. The two potential N-glycosylation sites in this domain (residues #25 to 28 and residues #70 to 72) are conserved in all subgroup A viruses.

The C-terminal region of subunit F_2 contains a relatively variable region (residues #88 to #136) with 83.33% homology. There are four non-coding nucleotide changes and nine coding changes in the F_2-subunit C-terminal region of the gene (**Table 9**). Four of these base substitutions occur in the first codon position (residue #101, #102, #103 and #122) and three nucleotide substitutions in the second codon position (residue #88, #97 and #105). There is also a nucleotide substitution in the third codon position at residue #124 and a multiple nucleotide substitution at residue #103. The pattern of coding changes in this domain of the fusion protein highlights a highly variable region between residues #101 and #105 inclusive. There are two or three N-linked glycosylation sites in this domain depending on strain, with a coding change at residue #122 forming the extra N-linked glycosylation site in all subgroup A viruses except strain A2 and isolate RSB89-1734. Not all amino acid substitutions in this region are conservative in nature, with the changes at residues #101 and #102 retaining charge characteristics but altering the structural characteristics of these residues as defined by Chou and Fasman (1978). The change at residue #124 (Lys → Asn) changes the residue charge from basic to neutral with possible structural effects. The consequences of these residue changes will be discussed later.

The cleavage site (residues #131 to #136) of all subgroup A isolates is located at the C-terminal end of this domain and remains unaltered in subgroup A viruses, with only one non-coding nucleotide change at residue #131 (Lys AAA → AAG), being observed.

3.2.5 Variation in the F_1-subunit of the fusion protein The F_1 subunit is highly conserved in amino acid sequence, with overall amino acid identity of 96.2%. The region of the fusion gene encoding the F_1 subunit can be divided into six distinct regions; **1)** a hydrophobic fusion related domain (FRD), believed to be the functional domain of the fusion protein; **2)** a conserved sequence (C #2) between the FRD and a cysteine-rich domain; **3)** a cysteine-rich domain believed to be an important structural region, and by analogy to the fusion protein gene of SV, possessing intra-molecular disulphide bonds that form at least two loop structures thought to lie at the molecular surface of the protein (Iwata *et al.*, 1994); **4)** a conserved region (C #3) between the cysteine-rich and transmembrane domain; **5)** a hydrophobic transmembrane region; and **6)** a cytoplasmic tail domain believed to interact with the matrix protein (**Section 3.2.2 Figure 9**). The aggregate of all coding changes in comparison to the prototype strain A2 and present in the F_1-subunit are summarised in **Table 10**.

The FRD (residues #137 to #166) is well conserved within subgroup A viruses (93.4% amino acid identity), with only two out of eight coding triplets affected by nucleotide substitutions giving rise to amino acid coding changes. The coding changes are conservative in type and involve changes in either the first position of the triplet (residue #152) or the second position (residue #148). The first three residues of the FRD (Phe-Leu-Gly) are conserved throughout all subgroup A isolates, thus preserving the conserved Phe-X-Gly found in all paramyxoviruses (Collins *et al.*, 1984b; Elango *et al.*, 1985a). The semi-conserved sequence, Gly-Val-Ala, located just downstream of the N-terminus in the fusion related domain at residues #151 to #153 does show variation from the A2 consensus sequence, with residue #152 (Val) being substituted for Ile, and this is in agreement with data published by Baybutt and Pringle.(1987), Johnson and Collins

Table 10 Coding changes in the gene region encoding the F_1-subunit

Residue #	Amino acid	Nucleotide change	Comments	Domain
#148	Ile → Thr	A**U**C → A**C**C	Conservative	FRD
#152	Val → Ile	**G**UU → **A**UU	Conservative	FRD
#173	Ser → Thr	**U**CC → **A**CC	Conservative	Conserved #2
#197	Asn → Thr	A**A**C → A**C**C	Conservative	Conserved #2
#354	Gln → Leu	C**A**A → C**U**A	Conservative	Cysteine-rich
#356	**Glu → Asp**	**GA**A → **GA**C	**Non-conservative**	**Cysteine-rich**
#379	Ile → Val	**A**UA → **G**UA	Conservative	Cysteine-rich
#384	Val → Ile	**G**UU → **A**UU	Conservative	Cysteine-rich
	Val → Ile	**GU**U → **AU**A	Conservative	Cysteine-rich
#447	Met → Val	**A**UG → **G**UG	Conservative	Conserved #3
	Met → Val	**AU**G → **GU**A	Conservative	Conserved #3
#540	**Ser → Leu**	**UC**A → **UU**A	**Non-conservative**	**Membrane anchor**
#547	Leu → Phe	**C**UC → **U**UC	Conservative	Membrane anchor

Nucleotide bases that are underlined and in bold character denote the position and nature of change. Those changes considered to be of structural importance are shown in highlighted boxes.

(1988b) and Lopez *et al.* (1988). Neither change is thought to affect this important region of the fusion protein.

The conserved region (C#2; residue #167 to #306), exhibits 98.56% amino acid identity. This is a relatively large region of the gene and contains eighteen nucleotide changes. However, only two of the eighteen result in coding changes (**Table 10**). The two coding changes are located at residue #173 (Ser → Thr); and residue #197 (Asn → Thr). These changes only occur in one isolate in each case, RSB89-1734 at residue #173 and isolate RSB89-6256 at residue #197, and both changes result in conservative amino acid substitutions that are not expected to affect the structure of this region of the fusion protein.

The cysteine-rich region (residue #307 to #431) is also highly conserved between subgroup A isolates (96.77%) with the number and location of cysteine residues remaining unaltered. There are four coding changes in this domain. These are a mixture of nucleotide substitutions in the first codon position (residue #379 and residue #384); second codon position (residue #354), third codon position (residue #356), or a combination of first and third positions (residue #384). All but one of the four amino acid substitutions are conservative in nature. The change at residue #356 (Glu → Asp) retains the acidic charge but alters the structural characteristics of the amino acid from a strong β-sheet breaker to an indifferent residue (Chou and Fasman, 1978). It is not known whether these changes would disrupt the intra-molecular disulphide bonding, but if so then this would have important consequences to the structure (and possibly antigenicity) of the fusion protein.

The conserved region (C #3) between the cysteine-rich and transmembrane domains (residue #432 to #524) has 97.4% homology between subgroup A isolates, and possesses fourteen non-coding nucleotide substitutions and only one amino acid coding change at residue #447 (Met → Val) which is considered a conservative change. This coding change is found in all subgroup A isolates examined, and will be discussed later in

this chapter. There is one potential N-linked glycosylation site (residues #500 to #502) in the C #3 domain, and this is conserved in all subgroup A viruses examined in this study.

The transmembrane (membrane anchor) domain (residue #525 to #551) possesses eight non-coding nucleotide changes and in all but one isolate-type virus there is 100% amino acid homology. The exception is isolate RSB89-6190 which shows 90% homology (**Table 10**) with two coding nucleotide changes, one in the first position of the codon triplet at residue #547 (Leu → Phe), and in the second codon position at residue #540 (Ser → Leu). The change at residue #547 is conservative whereas the change at residue #540 is non-conservative, but it still retains the hydrophobic characteristics required in this membrane-anchor region.

The intracellular tail region (residues #552 to #574) has six non-coding changes. The lack of variation in the cytoplasmic domain may relate to the function of this region, which may interact with the matrix protein.

3.3 Discussion

3.1.1 The pattern of amino acid changes Conservation in amino acid content between subgroup A viruses is very high (97 to 99.5%, depending on strain). Whilst most amino acid changes observed in the different fusion proteins are present in one or more isolates, there are some changes that appear in all isolates except the prototype strain A2 (**Table 11.**). Collins *et al.*, (1984b) observed nucleotide heterogeneity among different F gene cDNA clones derived from strain A2. The published sequence is therefore a consensus sequence. This heterogeneity may be a consequence of RNA polymerase error and/or errors in reverse transcription during cDNA synthesis. A second explanation is that these amino acid changes may reflect adaptation to different cell culture regimes, but this is unlikely as no changes were observed in the hypervariable region of the G protein after ten passages in different cultured cells (Cane *et al.*, 1993). The third explanation is that the A2

Table 11 Changes found in all isolates except strain A2

Residue #	Amino acid	nucleotide change	Comments	Domain
#4	Leu → Pro	C**U**A → C**C**A	Non-conservative	F_2 Signal
#16	Thr → Ala	**A**CU → **G**CU	Conservative	F_2 Signal
	Thr → Ile	AC**U** → A**U**U	Conservative	F_2 Signal
#152	Val → Ile	**G**UU → **A**UU	Conservative	F_1 FRD
#379	Ile → Val	**A**UA → **G**UA	Conservative	F_1 Cysteine-rich
#447	Met → Val	**A**UG → **G**UG	Conservative	F_1 Conserved #2
	Met → Val	**A**U**G** → **G**U**A**	Conservative	F_1 Conserved #2

Nucleotide bases that are underlined and in bold character denote the position and nature of change.

strain itself is heterogenous. Evidence supporting this theory comes not only from the amino acid variation observed in the fusion protein gene of this strain, but also from the mobility profiles of the phosphoprotein, which runs as a triplet on SDS-PAGE analysis (Caravokyri, 1990; and **Section 4.3 and Figure 16**), and not as a single band as previously published (Lambden, 1985). There is also evidence that the sequence of at least one other gene of the A2 substrain used in this project does not correspond exactly with the published sequence; specifically the NS1 gene (J.Evans, personal communication) where there are two reported coding changes. There may be two substrains of the A2 virus in use in different laboratories.

Anderson *et al.*, (1992) identified the amino acid change at residue #447 in a cleavage mutant derived from the A2 strain, and suggested that it might play a role in that phenotype, possibly by disrupting the correct protein-folding required for transport through the ER and Golgi apparatus where cleavage occurs.

There are other amino acid changes that occur in only one isolate; isolate RSB89-1734 (#6 Leu → Ile; #148 Ile → Thr; and #173 Ser → Thr); isolate RSB89-5857 (#13 Thr → Ala; residue #88 Asn → Ser; #97 Met → Ser; and #354 Gln → Leu); isolate RSB89-6190 (#540 Ser → Leu; and #547 Leu → Phe); isolate RSB89-6256 (#39 Ala → Val; #50 Thr → Ile; and #197 Asn → Thr); and isolate RSB89-6614 (#101 Pro → Ser).. It is not known whether these are an accurate reflection of sequence divergence or errors in the fusion gene introduced during the cDNA and PCR amplification process.

The C-terminal region of the F_2-subunit (residues #101 to #105) is highly variable. The pattern of amino acid changes here is of interest, showing several coding changes from the A2 consensus sequence. Residue #102 is changed in all isolates except RSB89-5857, as is also the case at residue #103. The change at residue #102 is in a position where nucleotide variability has been reported previously (Collins *et al.*, 1984b; Anderson *et al.*, 1992). Residue #105 is changed only in isolates RSB89-6190 and RSB89-1734. The

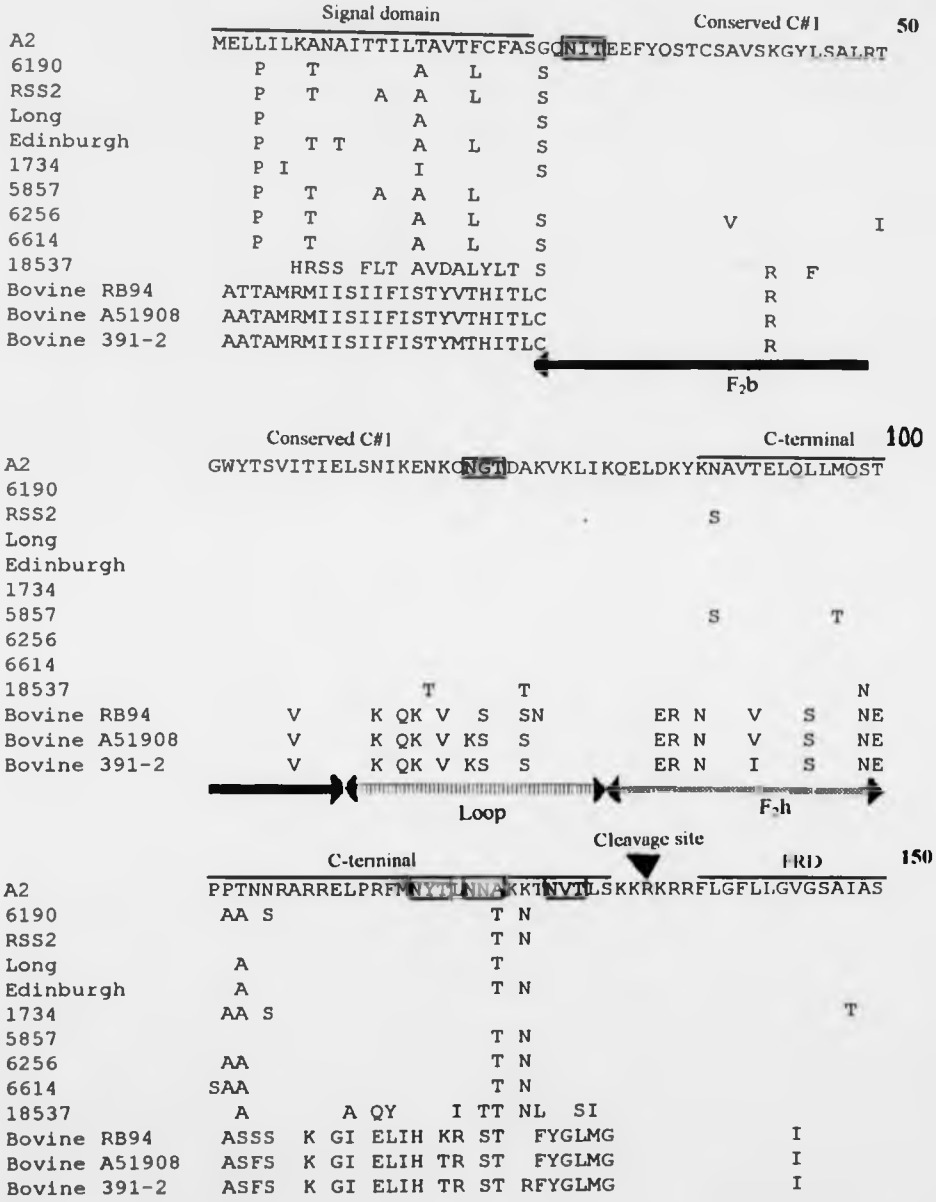
potential consequences of changes in amino acid sequence in this region are discussed in the next section.

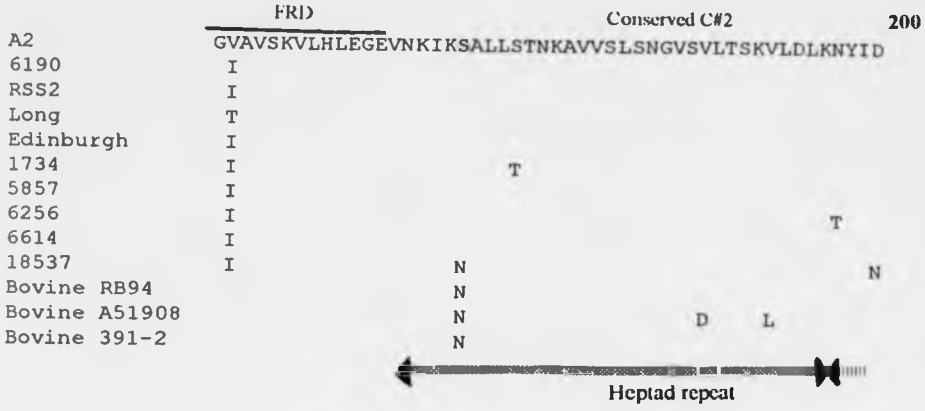
3.3.2 F_2-subunit variation within RS viruses The previously published sequences of the RS virus subgroup A strains RSS-2, Edinburgh and Long (Baybutt and Pringle, 1987; Elango *et al.*, 1985a; and Lopez *et al.*, 1988), the subgroup B strain 18537 (Johnson and Collins, 1988b), and the bovine strains RB94, A51908 and 391-2 (Walravens *et al.*, 1990; Himes and Gershwin, 1992; and Lerch *et al.*, 1991) are considered together in this section (**Figure 11**). They are also considered in relation to the structural features common to all paramyxovirus fusion proteins (Chambers *et al.*, 1992).

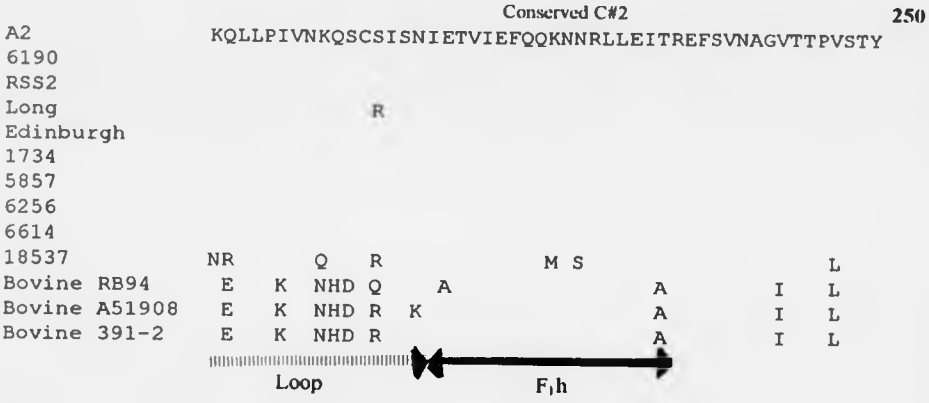
Variation in the signal region appears to leave the hydrophobic nature of this domain unaltered and it is thought that this is the principal requirement for signal function with sequence motif being less important (Collins, 1991). The predicted amino acid sequence for the F_2-subunit signal region of strains Long, RSS-2 and Edinburgh show overall agreement with one change at residue #10 (Ala $\rightarrow$ Thr) in strain Edinburgh that is not found in other strains and isolates. This coding change is conservative in nature. Changes at residues #16 and #20 are also found in the subgroup B strain 18537 virus (Johnson and Collins, 1988b) although the significance of this similarity is unclear. Within subgroup A viruses the signal domain identity is 59.6%, whereas between subgroups there is 45.8% homology (based on the comparison of subgroup A strain A2 and subgroup B strain 18537). Analysis of bovine RS virus signal domain sequences (strains RB94, Walravens *et al.*, 1990; A51908, Himes and Gershwin, 1992; and 391-2, Lerch *et al.*, 1990) revealed that of 24 amino acid residues, only one was shared in common, the initial methionine. However the hydrophobicity profile in this region is similar to that of subgroup A and B viruses (data not shown).

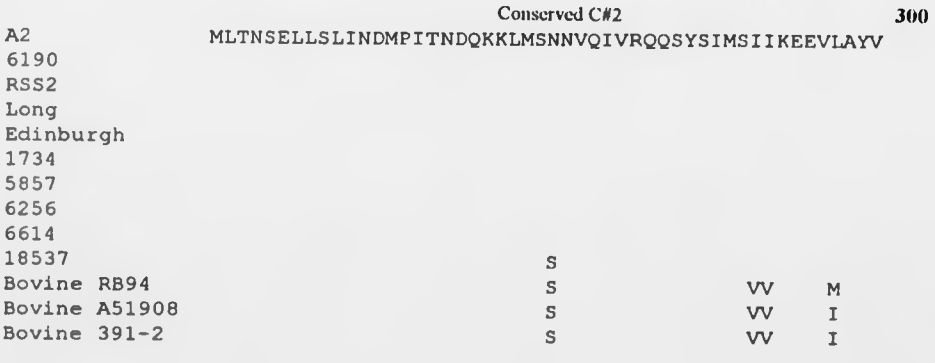
The conserved region (C #1) of subunit F_2 has 98.5% amino acid identity within subgroup A viruses and shows lower homology (93.6%) with subgroup B strains (Johnson and Collins, 1988b) and much lower homology with bovine RS virus strains (77.8%;

Figure 11. Fusion protein sequences from subgroup A isolates and strains, as compared to subgroup B and bovine RS virus strains




FRD
Conserved C#2
200
A2 GVAVSKVLHLEGEVNKIKSALLSTNKAVVSLSNGVSVLTSKVLDLKNYID
6190 I
RSS2 I
Long T
Edinburgh I
1734 I T
5857 I
6256 I T
6614 I
18537 I N N
Bovine RB94 N
Bovine A51908 N D L
Bovine 391-2 N
Heptad repeat


Conserved C#2
250
A2 KQLLPIVNKQSCSISNIETVIEFQQKNNRLLEITREFSVNAGVTTPVSTY
6190
RSS2
Long R
Edinburgh
1734
5857
6256
6614
18537 NR Q R M S L
Bovine RB94 E K NHD Q A A I L
Bovine A51908 E K NHD R K A I L
Bovine 391-2 E K NHD R A I L
Loop
F1h


Conserved C#2
300
A2 MLTNSELLSLINDMPITNDQKKLMSNNVQIVRQQSYSIMSIIKEEVLAYV
6190
RSS2
Long
Edinburgh
1734
5857
6256
6614
18537 S
Bovine RB94 S VV M
Bovine A51908 S VV I
Bovine 391-2 S VV I

cysteine-rich domain 350

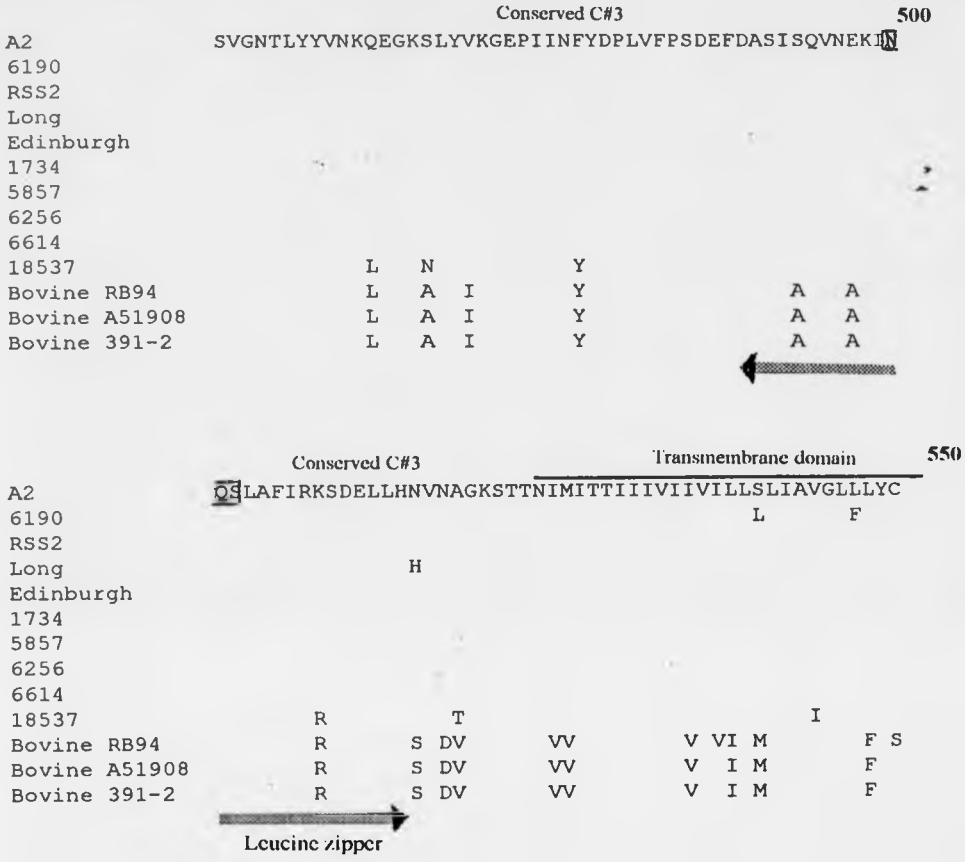
```
A2              VQLPIYGVIDTPCWKLHTSPLCTTNTKEGSNICLTRTDRGWYCDNAGSVS
6190
RSS2
Long
Edinburgh
1734
5857
6256
6614
18537                                I
Bovine RB94                          DN
Bovine A51908                        DN
Bovine 391-2                         DN
```

cysteine-rich domain 400

```
A2              FFPQAETCKVQSNRVFCDTMNSLTLPSEINLCNVDIFNPKYDCKIMTSKT
6190                                 V
RSS2              L                  V    I
Long                                 V
Edinburgh                            V    I
1734                D                V    I
5857              L                  V    I
6256                                 V    I
6614                                 V    I
18537               D                VS   T    S
Bovine RB94                         TDV   T    A
Bovine A51908      T                TDV   T    T
Bovine 391-2                        TDV   T    T
```

cysteine-rich domain Conserved C#3 450

```
A2              DVSSSVITSLGAIVSCYGKTKCTASNKNRGIIKTFSNGCDYVSNKGMDTV
6190                                                        V
RSS2                                                        V
Long                                                  A     V
Edinburgh                                                   V
1734                                                        V
5857                                                        V
6256                                                        V
6614                                                        V
18537           I                                           V
Bovine RB94     I       I                                 R V
Bovine A51908   I       I                                   V
Bovine 391-2    I       I                                   V
```



Summary of the amino acid sequence for the fusion proteins of human and bovine RS viruses including data derived from this thesis, and the published sequence data for subgroup A strains A2 (Collins *et al.*, 1984b); RSS-2 (Baybutt and Pringle, 1987); Edinburgh (Elango *et al.*, 1985a); Long (Lopez *et al.*, 1988). Also include are subgroup B strain 18537 (Johnson and Collins, 1988b), and bovine strains RB94 (Walravens *et al.*, 1990); A51908 (Himes and Gershwin, 1992) and 391-2 (Lerch *et al.*, 1991). Possible sites of N-linked glycosylation are denoted by hatched boxes. Important structural features identified by Chambers *et al.*, (1992) as common to paramyxovirus fusion proteins are shown below the amino acid sequence. Detailed explanation of these structural features can be found in **Section 1.7.5.**

Himes and Gershwin, 1992). Chambers *et al.*, (1992) suggest that this region consists of a β-sheet forming region (F_2b; residues #23 to 61) followed by a loop structure (residues #62 to 78) of variable length, containing both the conserved cysteine residues found in the F_2-subunit. Most variation between human and bovine strains occurs in the predicted loop structure, but how this affects the structure of the F_2-subunit is unclear. There is evidence (**Section 4**; this thesis) that a single coding change in the conserved C#1 region can lead to restricted growth at temperatures above 37°C. This supports the theory that the structure of this region is crucial for correct fusion protein function.

The C-terminal domain of F_2 is variable within subgroup A viruses. The cleavage site of all human and bovine strains and isolates is strictly conserved, with the sequence motif (residue #131 to #136 Lys-Lys-Arg-Lys-Arg-Arg). The extra (fifth) N-linked glycosylation site present in the C-terminal domain is found in all bovine and human RS viruses except two subgroup A viruses. This might explain the observed differences in fusion protein mobility on SDS-PAGE described in **Section 3.2.1**. The degree of amino acid homology within the C-terminal domain of subgroup A viruses is 76% (only the signal domain is more divergent) and between subgroups is approximately 69.7%, and between human and bovine strains considerably lower at 17.3%. Variation in the C-terminal end of subunit F_2, may have an effect on fusion protein function. This region (residues #79 to 100) has been suggested to comprise an α-helical structure (F_2h; Chambers *et al.*, 1992) and alterations in this region may affect the secondary structure of the protein and also account for the mobility differences. The analysis of mutant *ts*A1 and its revertant viruses described in **Chapter 4** of this thesis support this theory, and suggest that these changes can partially correct for alterations in the loop structure encoded by the conserved C#1 region.

3.3.3 F_1-subunit variation within RS viruses The F_1-subunit FRD shows high amino acid identity (96%) amongst all RS viruses, both human and bovine. Isolate RSB89-1734 possesses an exclusive coding change at residue #148 and actually demonstrates lower identity (92%) with strain A2 than that found between bovine strains and strain A2 (96%).

It was originally considered to be a sequencing error, but further sequencing of this region of isolate RSB89-1734 using PCR product from separate cDNA preparations verified the observation. The change from the strain A2 sequence at residue #152 is found in both subgroup A and B viruses, although in subgroup A strain Long the change is to threonine.

The conserved region (C #2) of the F_1-subunit is well conserved within subgroup A (97.8%), between subgroups (92.8%), and between human and bovine strains of RS viruses (87.1%). It possesses three discrete structural regions, two α-helixes (H and F_1h) separated by a loop-forming sequence (Chambers *et al.*, 1992). The α-helical regions are well conserved in both bovine and human strains, but the loop-forming sequence does show variation. This mirrors the change found in the loop-forming sequence present in the F_2-subunit, although the reason for this is unknown. The non-conservative substitution at residue #213 (Ser → Arg) is only present in strain Long and subgroup B strain 18537. This change could disrupt the loop structure and alter the folding in the F_1-subunit and possibly alter the antigenicity of the protein by disrupting epitopes in antigenic area #I which has been mapped to this region of the protein (**Section 1.7.8**). There is an extra coding change present in subgroup B and bovine strains, consisting of a conservative substitution of Asp by Ser at residue #276. This change also occurs in a region defined as a major epitope (antigenic area #II; **Section 1.7.8**). Therefore both these changes could potentially contribute to the antigenic difference of subgroups A and B, as well as between subgroup A and bovine viruses.

The cysteine-rich domain of F_1 does not possess any identifiable structural features, and yet is highly conserved between and within human subgroups (96.0%) and between human and bovine strains (92.7%). The coding changes reported in strain RSS-2 (residue #428 Asn → Asp and residue #434 Thr → Arg) are exceptional and appear to be errors in the sequence published by Baybutt and Pringle (1987) and shown in this thesis, **Section 5.2.2**. It is also thought that the substitution at residue #379 (Ile → Val) which is observed in all human and bovine isolates of RS viruses may be an error in the sequence

published by Collins *et al.*, (1984b). This is supported by the nucleotide sequence of a temperature-sensitive mutant (*ts*A1) derived from the wild-type strain A2, which exhibited the same sequence at this position as all other subgroup A isolates **(Section 4.2.2.)** and thus differed from the published sequence of strain A2. The number and position of the cysteine residues remain unaltered and these are thought to contribute to an extended open structure (possibly consisting of at least two stem loops) formed by intra-molecular disulphide bonding. Epitope mapping **(Section 1.7.8)** has identified a T-cell epitope at residues #338 to 355 and a major fusion-inhibiting, neutralising B-cell epitope (antigenic area #IV) that are present in this domain.

The conserved region (C #3) of F_1 possesses one identified structural feature (the leucine zipper, Z) which is thought to comprise an α-helical region immediately upstream of the transmembrane anchor region. This structural feature is highly conserved in both human and bovine RS viruses, suggesting an essential structural requirement. The C#3 domain has 97.4% homology within subgroup A viruses, and 94.6% homology with subgroup B strain 18537, and 88.1% homology with bovine strains. There is a conservative amino acid substitution at position #442 (Val → Ala) in strain Long that is not found in other strains and isolates. The change in amino acid #447 occurs in both subgroups of human RS viruses and also in bovine RS virus strains and also reflects the variable nature of this residue identified by Collins *et al.*, (1984b). This is supported by the nucleotide sequence of mutant *ts*A1 which exhibited the same residue at this position as that found in all viruses reported in this study **(Section 4.2.2.)**. There is one coding change only found in strain Long at residue #515 (Asn → His). This could imply an error in the published sequence, however there is a change at this position in bovine RS viruses. It is unclear therefore, whether this coding change is truly present in the Long strain. If this change is real it could potentially alter the predicted protein structure in this region by replacement of a α-helix indifferent residue with a strong α-helix forming residue and alter the residue charge from neutral to basic. This could potentially lengthen the series of α-helixes formed by the leucine zipper (Z) region described in **Section 1.7.5**. The consequences of this are unclear, however it might be

a response to the coding change at residue #213 (Ser → Arg) which has the potential to disrupt the loop feature downstream of the H box and FRD.

The transmembrane (membrane anchor) domain of F_1 is nearly exactly conserved within subgroup A viruses (100% homology within all strains except isolate RSB89-6190 which is 90% homologous). Between viruses of both human subgroups there is less homology (96.2%), and even less between human and bovine strains (69.2%). The two coding changes found in isolate RSB89-6190 are unique in viruses of both subgroups. The change at residue #547 (Leu → Phe) is conservative, whilst the change at residue #540 (Ser → Leu) is non-conservative. This therefore, may alter the structure of the membrane anchor, but it is unclear what likely effect this might have on function. However, because this change is present in a naturally occurring isolate, it is presumed unlikely that this is a deleterious change. The possibility that these coding changes were errors in the derived sequence data was considered and so this region of the fusion protein gene was re-sequenced from new molecular clones derived from a separate PCR product obtained from a different cDNA preparation. No nucleotide errors in the derived sequence data were found, thus indicating that the sequence at this region is correct.

The cytoplasmic tail domain is completely conserved within subgroup A viruses, and is also conserved (87.5%) between human RS virus subgroups, but is less homologous (66.7%) between human and bovine strains of RS viruses. The total conservation of amino acid sequence within subgroup A strains and isolates highlights the importance of this region of the protein which is thought to interact with some inner component of the virus, possibly the matrix (M) protein.

3.3.4 Fusion protein (F1 and F2) relatedness within subgroup A isolates The sequence data generated in this study found that the deduced amino acid identities for the fusion proteins of subgroup A isolates varied from 97% to 99.5%. This is a higher degree of homology than that found in the attachment (G) proteins of the same isolates (Cane *et*

al., 1991). It was also found that the relative relatedness of individual isolates based on their fusion proteins (**Figure 10**) varied slightly from that reported for their SH and N protein genes (Cane and Pringle, 1991)(**Section 1.2 Figure 2**). The isolates can be arranged into three main groupings based on the relatedness of their fusion proteins; the first consists of isolates RSB89-6256, RSB89-6614, RSB89-6190 and the Edinburgh strain. The second group consists of isolate RSB89-5857 and strain RSS-2, and the third consisting of isolate RSB89-1734 and strain Long. Strain A2 appears to be more distantly related to each of the other isolates and strains.

The differences in the consensus sequence of strain A2 (Collins *et al.*, 1984b) and its possible causes have been discussed in **Section 3.3.1**. The divergence of the A2 strain is not related to its earlier time of isolation, because the Long strain predates it. The nucleotide and predicted amino acid sequences for RS virus subgroup A isolates reported in this study are presented in comparison to the published A2 sequence. The reason for this is historical and is not meant to indicate that all isolates deviate from strain A2, but rather that strain A2 is somewhat different from all other subgroup A strains and isolates, possibly because of the reasons mentioned above. Overall, however, the range of nucleotide and predicted amino acid homology identified in this study would remain unchanged even using the modified A2 data. Relatedness of subgroup A fusion genes is similar but not identical to the pattern exhibited by their attachment (G) protein genes. However, the fusion protein gene data are less definitive in discriminating strain differences due to the small numbers of amino acid changes present.

3.3.5 Predicted fusion protein secondary structure In order to determine whether changes in amino acid residues could alter the secondary structure of the fusion proteins, predictions of the fusion protein structure of isolates representing each of the five SHL lineages, and the prototype strain A2, were carried out using a model based on the algorithm derived by the Chou and Fasman (1978) method and analysed on the GCG™ computer programme (Devereux *et al.*, 1984). Their predictions may be no more than 50% accurate but in the

absense of definitive structural analysis by X-ray crystallography, are useful for purposes of inter-strain comparison. Structural changes predicted by computer analysis are only intended as an indication of possible disruption of conserved and identifiable structural features, and no conclusion as to the actual structure of the protein can be drawn. The Chou and Fasman predictions were confirmed with an alternative model using the algorithm derived by Garnier-Osguthrope-Robson (1978) and analysed using the Microgenie™ computer programme (Queen and Korn, 1984). These predictions are based solely on the primary amino acid sequence and do not predict the effects of post-translational modification, such as intra-molecular disulphide bonding, on the folding of the fusion protein.

For the purposes of this discussion the fusion protein structure is separated into the component subunits, F_2 and F_1. The Chou and Fasman predictions for six F_2-subunits corresponding to the prototype strain A2 and isolates representative of the five lineages are shown in **Figure 12** and the corresponding Garnier-Osguthrope-Robson prediction shown in **Figure 13**.

The predicted F_2-subunit structure of the lineages falls into five distinct groups. The first of these contains strain A2 and isolate RSB89-5857 which are nearly identical except for the number of N-linked glycosylation sites at the C-terminal end of the subunit which is greater in isolate RSB89-5857. This group is also unique in possessing an orientation opposite to that found in the other four groups. The secondary structure predictions show variation in the number of coil-forming motifs located in the conserved β-sheet region (F_2b) identified by Chambers *et al.*, (1992). Isolate RSB89-6256 possesses fewer random coil-forming residues than those found in other viruses. It is thought that amino acid changes at residues #39 (Ala → Val) and #50 (Thr → Ile) account for these structural modifications. There is also variation between the five groups in the position and size of the α-helix region (F_2h) located in the C-terminal region and identified by Chambers *et al.*, (1992). Variation in this region gives rise to three distinct structural groups with isolates

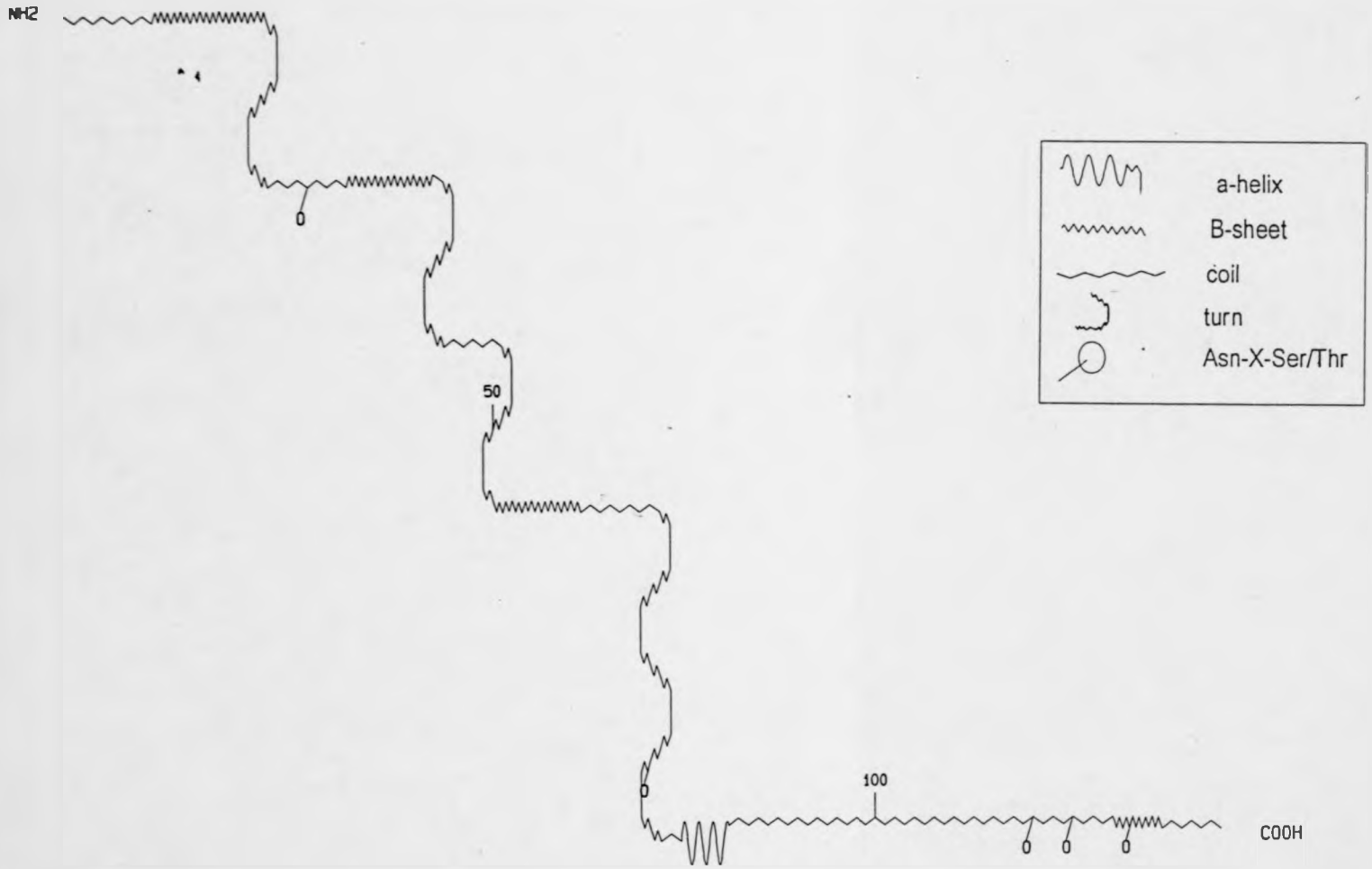
Figure 12 Secondary structure predictions for the F2-subunits of five natural subgroup A isolates and wild-type strain A2; using the algorithm of Chou and Fasman, 1978.

Predicted secondary structure of the F_2-subunit of the fusion proteins representative of each of the five lineages identified by Cane and Pringle, (1991). Secondary structure predictions used the method of Chou and Fasman (1978), which is based solely on the primary amino acid sequence and does not take into account the effects of post-translational modifications of the protein. However the GCG™ computer programme used for graphically displaying the predicted structures, indicates the position of potential sites of N-linked glycosylation (*i.e.* the tripeptide sequence Asn-X-Ser/Thr). Explanations of the symbols used to depict regions of random coil, α-helical or β-sheet structure are given in the box on the left. α-helical and β-sheet regions are also indicated on the structure plot. Five discrete structural groups are predicted, with strain A2 and isolate RSB89-5857 possessing similar structure but differing in the number of N-linked glycosylation sites in the C-terminal region of the subunit. Detailed descriptions of the differences between structural groups are given in the text.

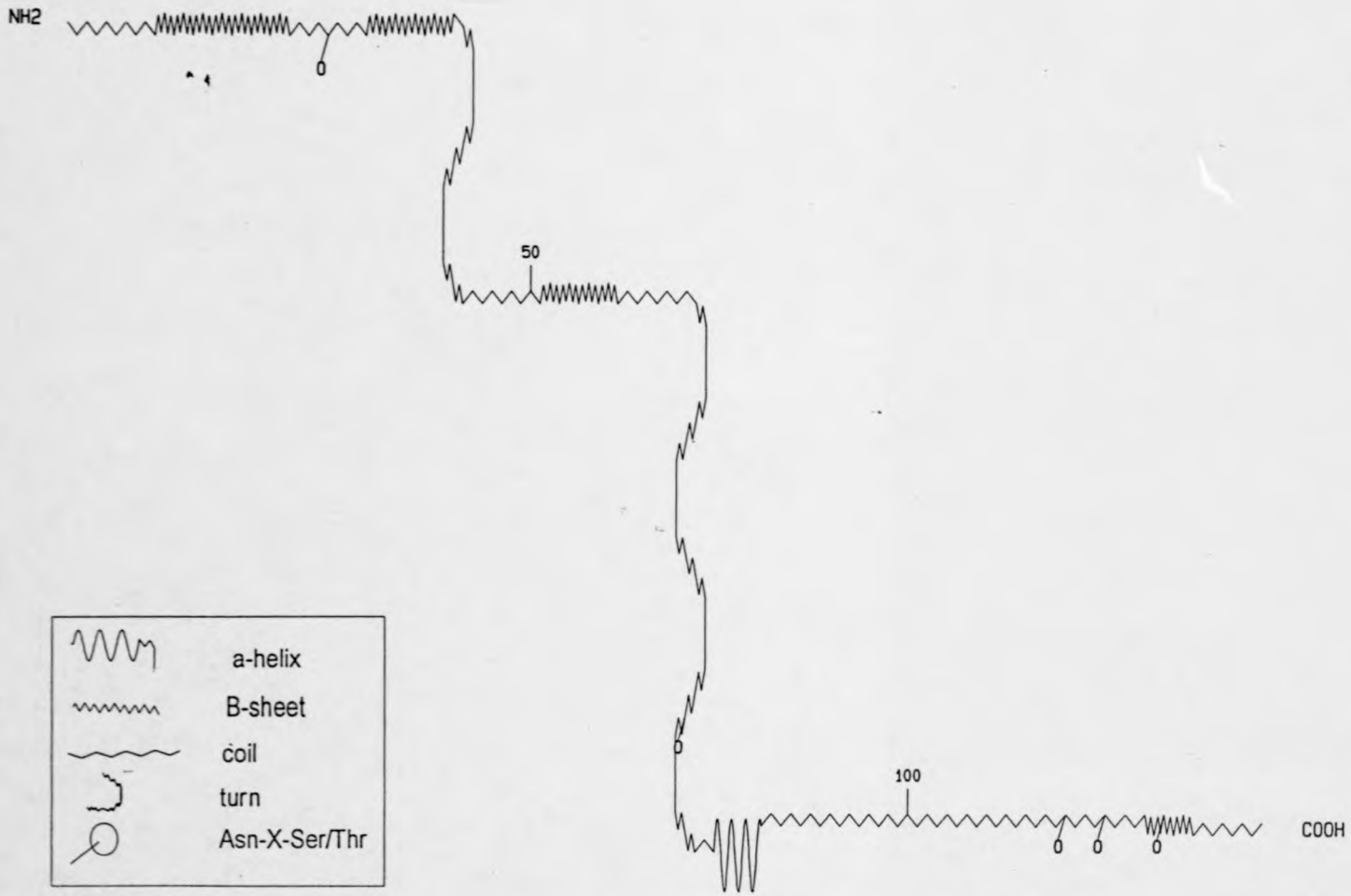
natural subgroup
sman, 1978.

sentative of each
cture predictions
e primary amino
modifications of
ly displaying the
sylation (*i.e.* the
lepict regions of
α-helical and β-
ural groups are
but differing in
ubunit. Detailed
in the text.

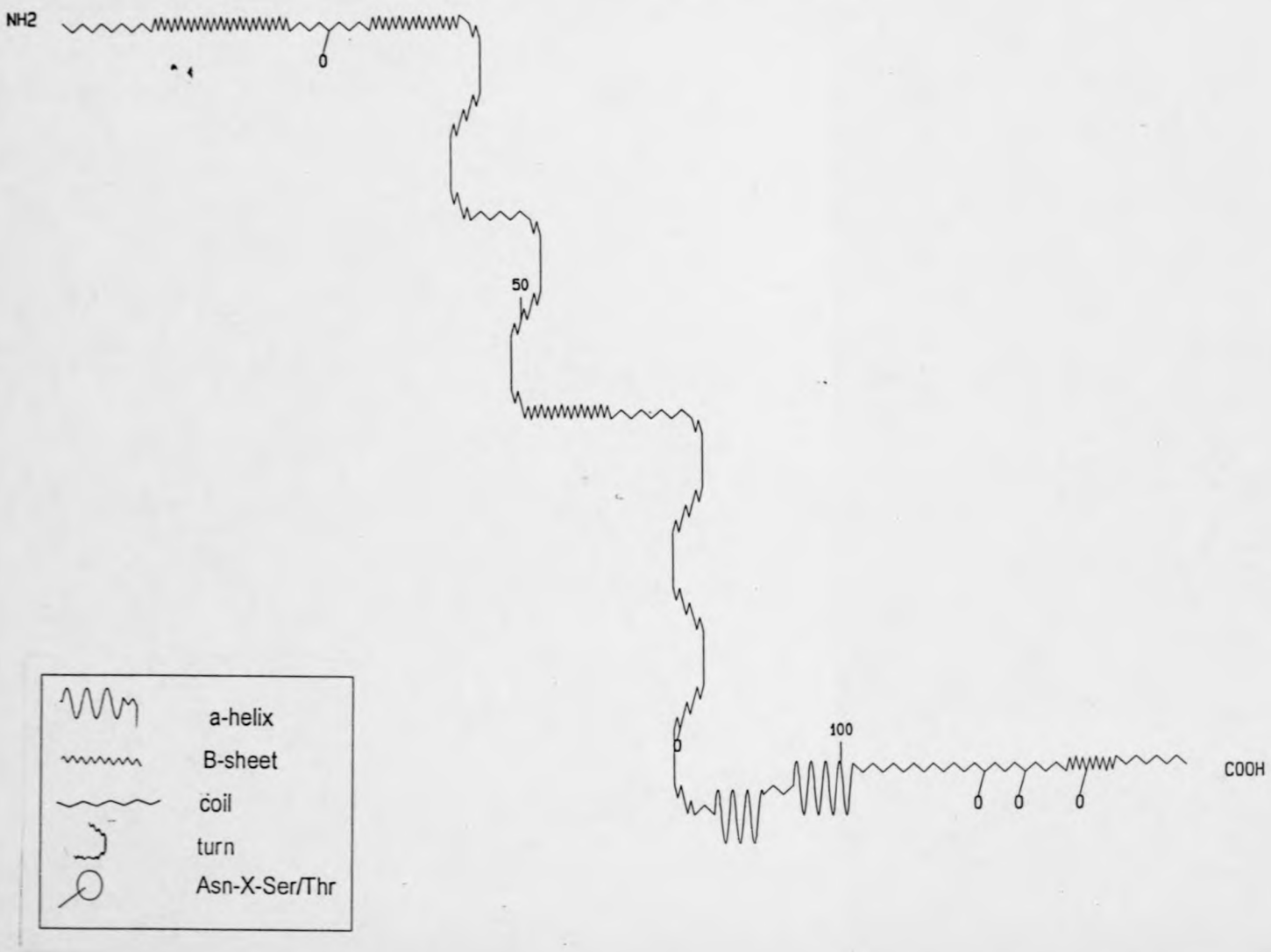
Isolate RSB89-6190



Isolate RSB89-6256



Isolate RSB89-6614



Isolate RSB89-1734

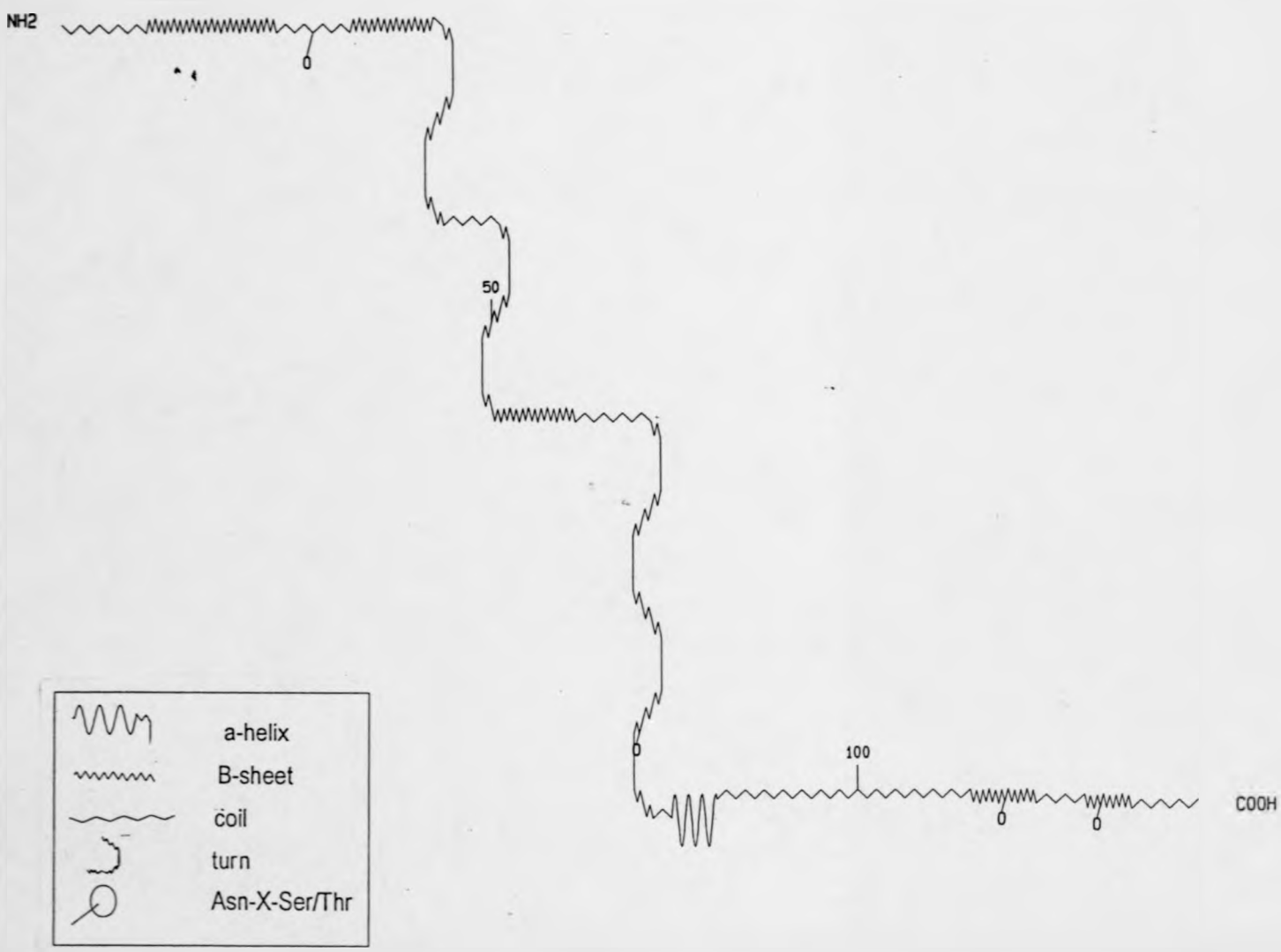
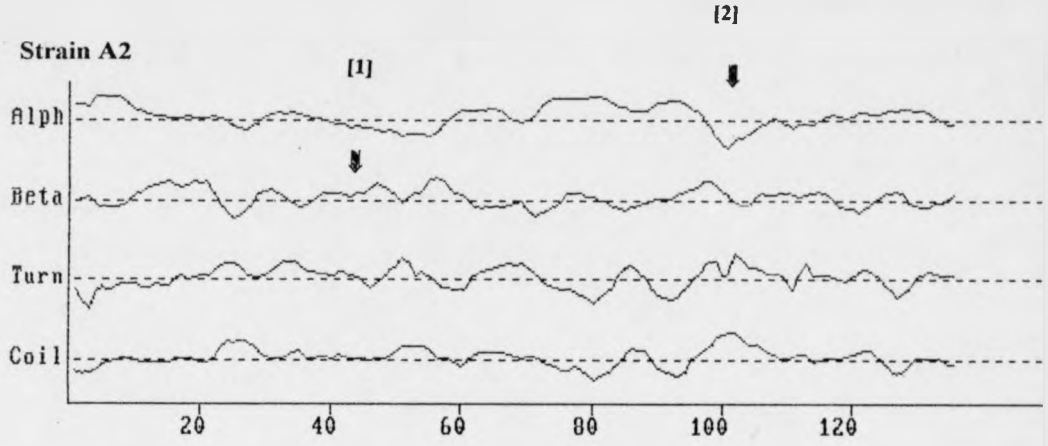


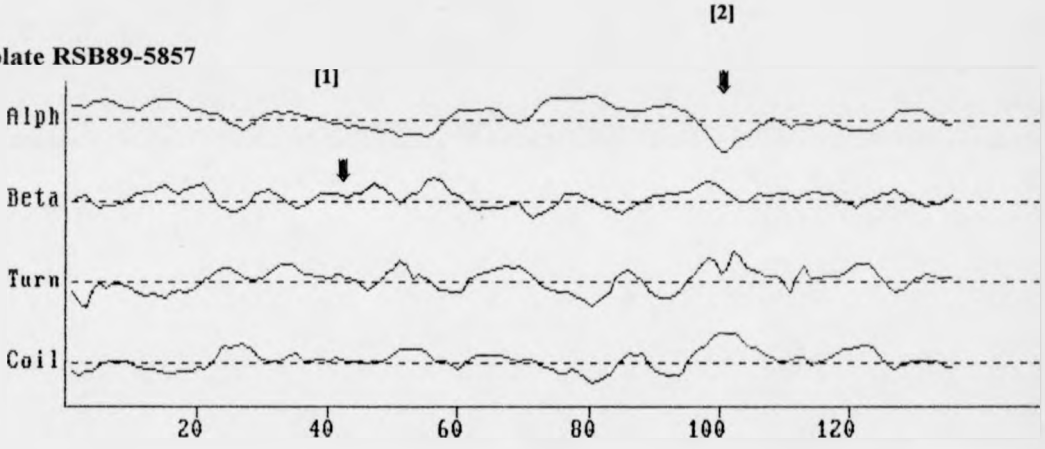
Figure 13 Secondary structure predictions for the F2-subunits of five natural subgroup A isolates and wild-type strain A2; using the algorithm of Garnier *et al.*, 1978.

Predicted secondary structure of the F_2-subunits of the fusion proteins representative of each of the five lineages identified by Cane and Pringle, (1991), according to the algorithm of Garnier *et al.*, (1978). Computation of the data and the graphic output of predicted structures were obtained with the Microgenie™ computer programme. The relative probabilities of the four secondary structures occurring along each protein (α-helix : Alph; β-sheet : Beta; β-turn : Turn; and random coil : Coil) are shown in four separate graphs, one above the other. Values above the horizontal axes are considered as positive. Especially high values in a graph ("peaks") indicate that there is a greater (predicted) probability of that structure being formed. In any given region, the protein is most likely to have the type of secondary structure which corresponds to the graph with the highest positive value at that position. Positions of differing structural predictions are marked as **[1]** and **[2]**.

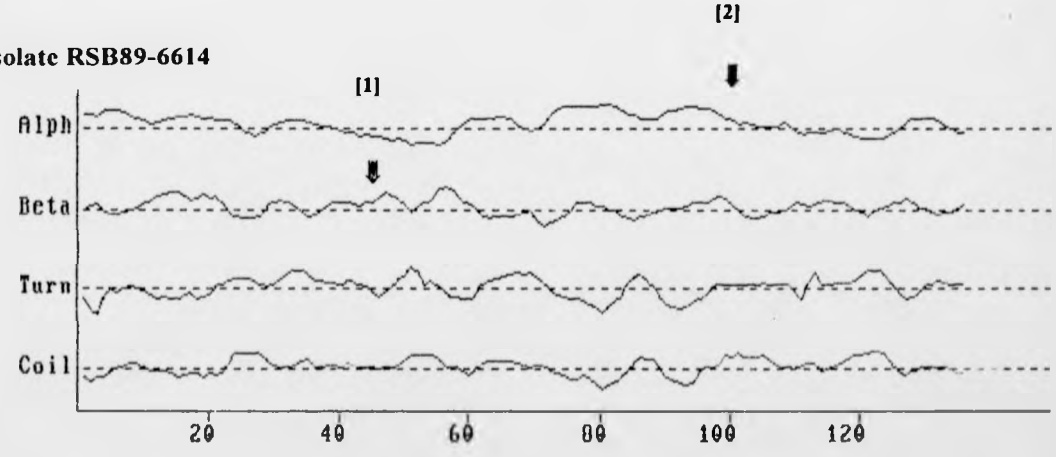
Strain A2



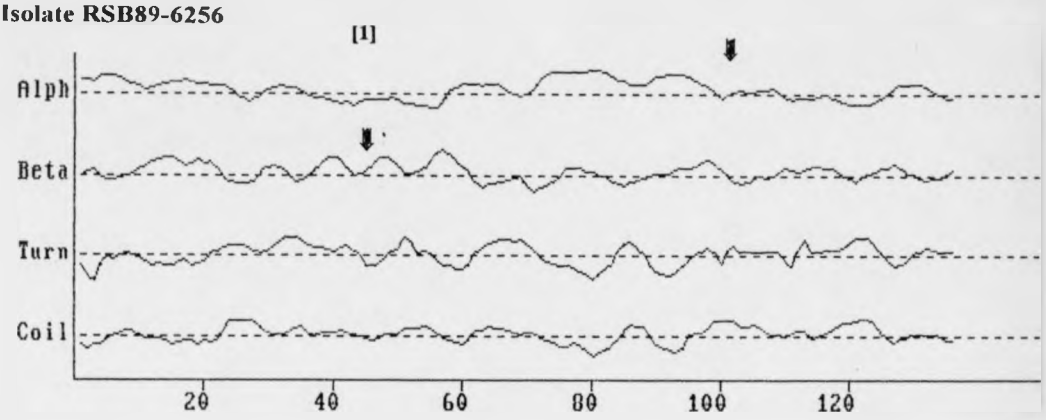
Isolate RSB89-5857



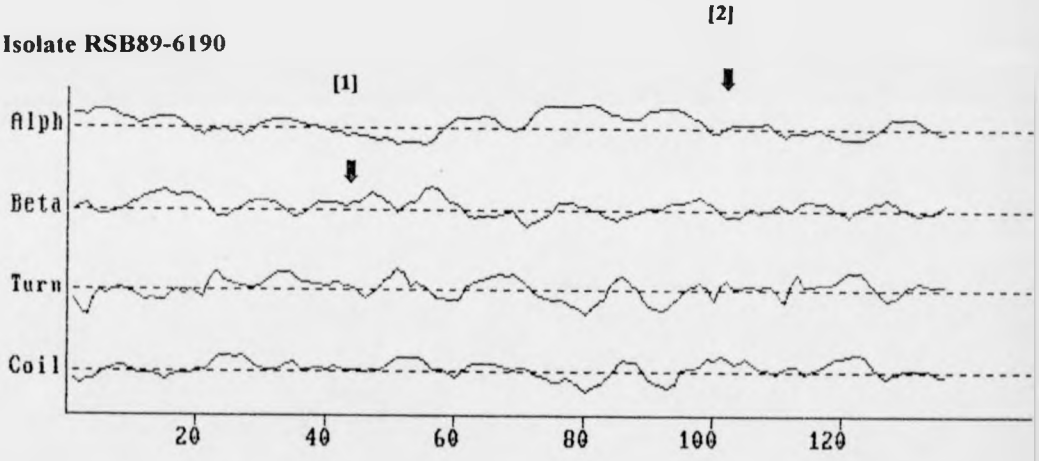
Isolate RSB89-6614



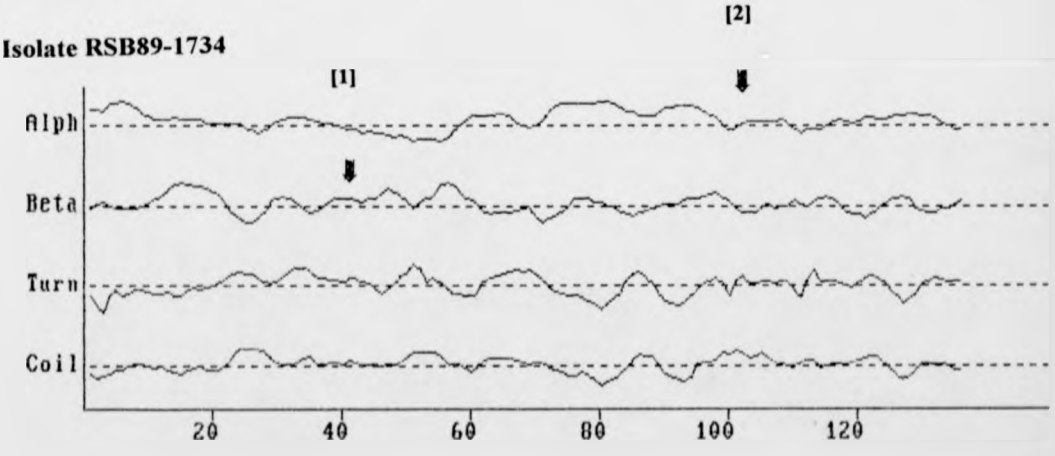
Isolate RSB89-6256



Isolate RSB89-6190



Isolate RSB89-1734



RSB89-6256; RSB89-619 and RSB89-1734 possessing an extended α-helical region around residue #100, isolate RSB89-6614 has this α-helical sequence truncated and positioned further downstream, and isolate RSB89-5857 and strain A2 have an α-helical sequence disrupted by coil-forming residues. Amino acid substitutions at residues #88, #101, and #102 are thought to be responsible for these changes in the length and position of the α-helical region (F_2h).

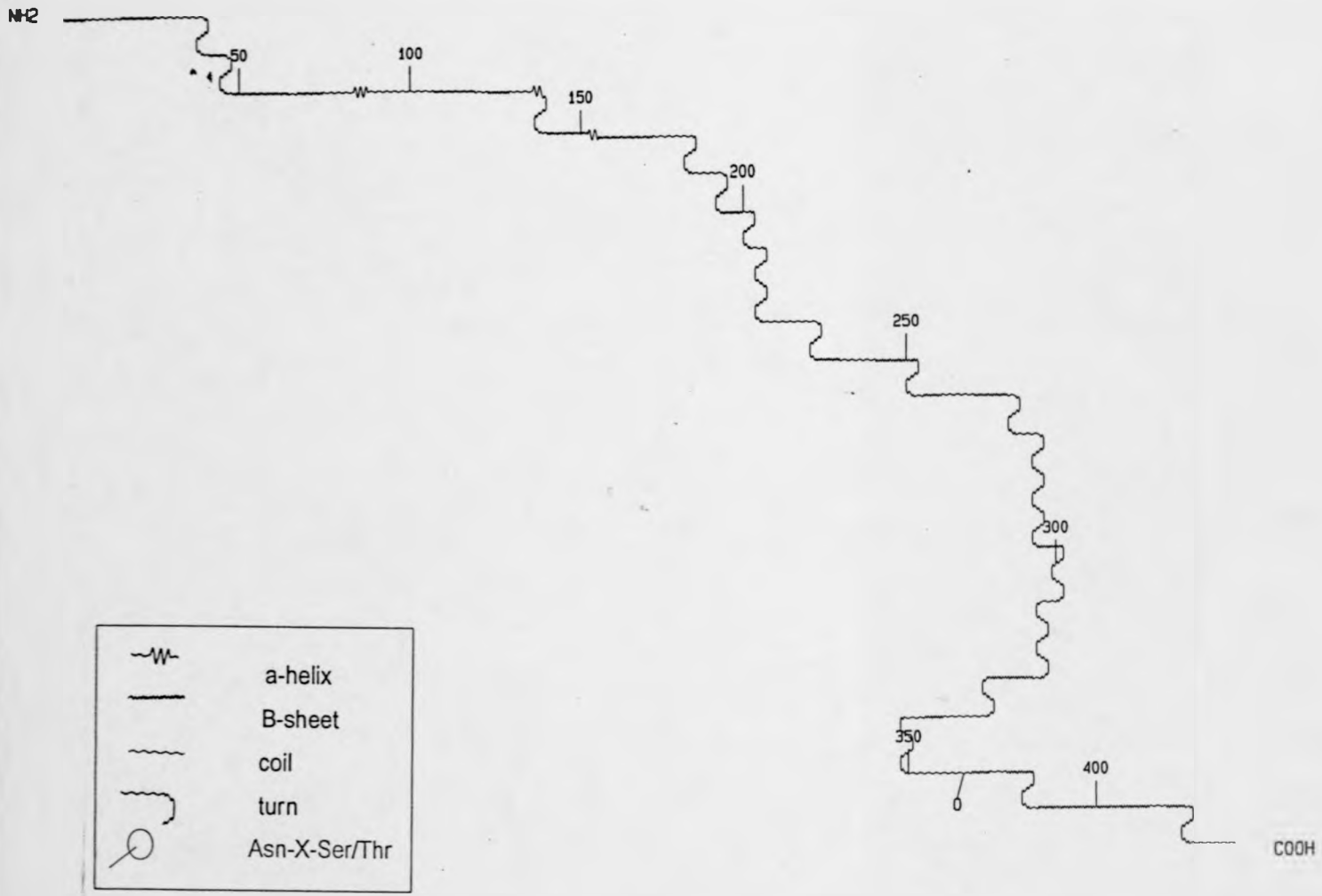
The Chou and Fasman predictions for six F_1-subunits corresponding to the prototype strain A2 and isolates representative of the five lineages are shown in **Figure 14** and the corresponding Garnier-Osguthrope-Robson prediction shown in **Figure 15**. It appears that the predicted structure of the F_1-subunit is highly conserved, with two distinct groups differing only in their orientation. Similar to that found in the predicted structure of the F_2-subunit, strain A2 and isolate RSB89-5857 show identical structural predictions, whilst the remaining isolates have F_1-subunits which are predicted to be identical. The change in orientation of the F_1-subunit is due to changes found in the F_2-subunit, and because this is cleaved from the F_1-subunit during post-translational modification, the altered orientation of the F_1-subunit predicted in this model is probably an artefact. This is supported by the secondary structure predictions using the Garnier *et al.* (1978) algorithm described in **Figure 15**. Therefore no change in the orientation of this subunit is thought to exist. Additionally, the computer predictions do not give an indication of true protein structure merely the position of certain features, so changes in orientation would be irrelevant.

3.3.6 Conclusion From this study it can be concluded that fusion protein gene variation among subgroup A isolates is low (homology 97 to 99.5%). From the predicted structural models it would appear that the few coding changes observed have the potential to change fusion protein structure especially in the C-terminal domain of the F_2-subunit, whereas the secondary structure of the F_1 subunit remains conserved. Comparing the location of coding changes in subgroup A strains to identified epitopes on the fusion protein

Figure 14 Secondary structure predictions for the F1-subunits of five natural subgroup A isolates and wild-type strain A2; using the algorithm of Chou and Fasman, 1978.

Predicted secondary structure of the F_1-subunit of the fusion proteins representative of each of the five lineages identified by Cane and Pringle, (1991). Secondary structure predictions used the method of Chou and Fasman (1978), which is based solely on the primary amino acid sequence and does not take into account the effects of post-translational modifications of the protein. However the GCG™ computer programme used for graphically displaying the predicted structures, indicates the position of potential sites of N-linked glycosylation (*i.e.* the tripeptide sequence Asn-X-Ser/Thr). Explanations of the symbols used to depict regions of random coil, α-helical or β-sheet structure are given in the box on the left. α-helical and β-sheet regions are also indicated on the structure plot. Structural analysis predicts very similar structural features, the two groups differing only in orientation, which is a prediction of protein structure that the computer models are not sufficiently accurate to make. These two groups are shown along with their representative viruses.

Strain A2 and isolate RSB89-5857



Isolates RSB89-6614; RSB89-6256; RSB89-6190; and RSB89-1734

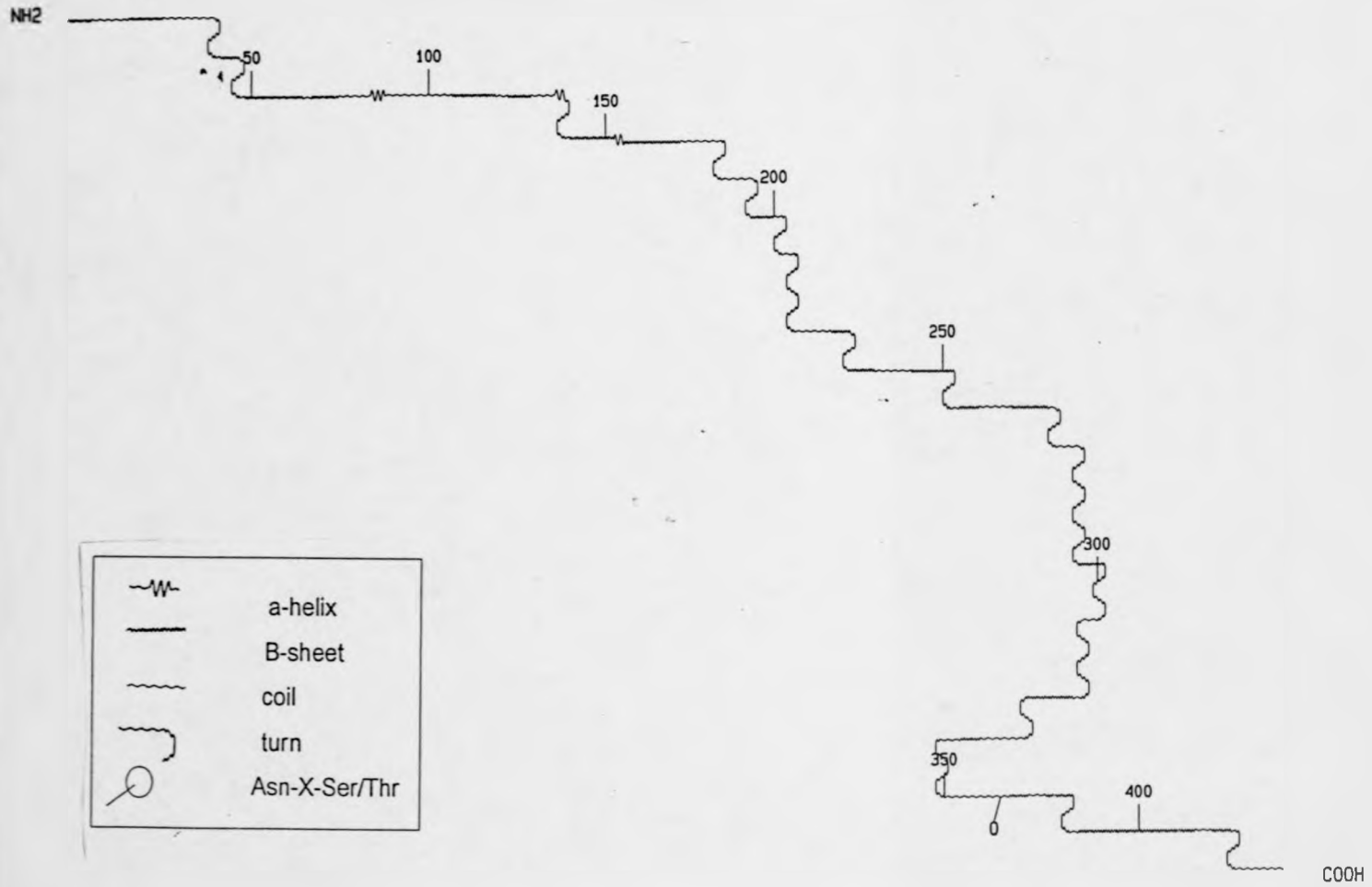
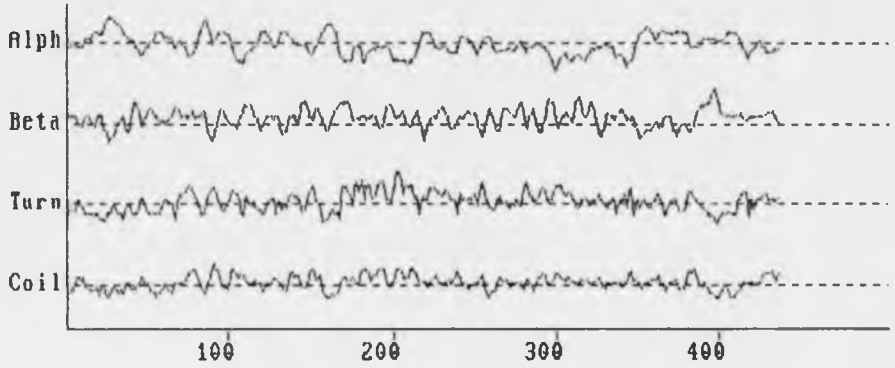


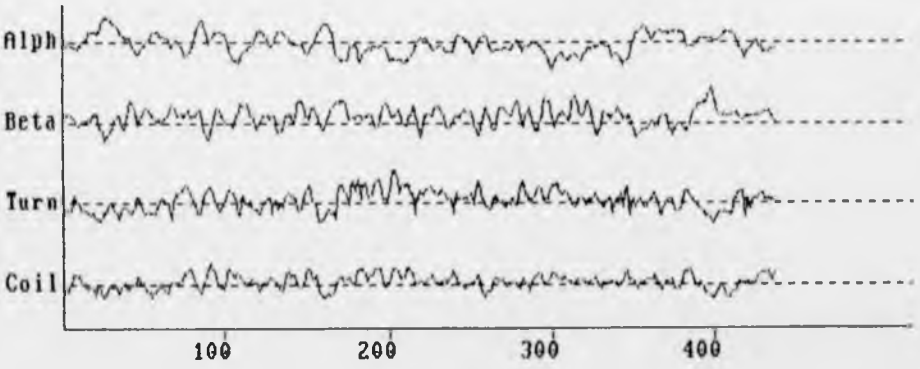
Figure 15 Secondary structure predictions for the F1-subunits of five natural subgroup A isolates and wild-type strain A2 using the algorithm of Garnier *et al.*, 1978.

Predicted secondary structure of the F_1-subunits of the fusion proteins representative of each of the five lineages identified by Cane and Pringle, (1991), according to the algorithm of Garnier *et al.*, (1978). Computation of the data and the graphic output of predicted structures were obtained with the Microgenie™ computer programme. The relative probabilities of the four secondary structures occurring along each protein (α-helix : Alph; β-sheet : Beta; β-turn : Turn; and random coil : Coil) are shown in four separate graphs, one above the other. Values above the horizontal axes are considered as positive. Especially high values in a graph ('peaks') indicate that there is a greater (predicted) probability of that structure being formed. In any given region, the protein is most likely to have the type of secondary structure which corresponds to the graph with the highest positive value at that position. The predictions suggest that all the F_1-subunits possess very similar structural features. Two predictions of secondary structure (strain A2 and isolate RSB89-6190) are therefore presented as representative for all viruses analysed.

Strain A2



Isolate RSB89-6190



(Section 1.7.8) has revealed that changes in structural features occur away from known epitopes which are located in the conserved F_1-subunit. However changes in amino acids distal to epitopes could possibly lead to alterations in the secondary structure of the protein which might sufficiently alter antigenic sites and lead to antigenic variation. The possibility of this has already been suggested by Arbiza *et al.*, (1992) and Lopez *et al.*, (1993) who identified a monoclonal antibody escape mutant at residue #190 which allowed escape from selection at antigenic site #I (residues #205 to #236), an area fifteen residues downstream from the amino acid change. Thus the amino acid variation observed in the F_2-subunit may result in antigenic variation, possibly by altering the conformational epitope suggested as being formed by antigenic areas #I to #III **(Section 1.7.8 Figure 6)** and involving interaction between the H box region of F_1 and the F_2h region of the F_2-subunit.

The coding change found in subgroup B strain 18537 and bovine RS viruses at amino acid #276 (asparagine to a serine) occurs in a region defined as a major epitope **(Section 1.7.8)** in subgroup A viruses. This change could potentially contribute to the antigenic variation between subgroups A and B, by disrupting an antigenic site in subgroup A strains or creating a different antigenic site in subgroup B strains.

CHAPTER FOUR

CHAPTER FOUR

Temperature--sensitive mutant *ts*A1-a putative fusion (F) protein gene mutant

4.1 Introduction Several temperature-sensitive (ts) mutant viruses have been derived from wild-type strains of RS virus in efforts to produce candidate vaccines. The first candidate vaccine was derived from strain A2 by Chanock and colleagues (Gharpure *et al.*, 1969; Wright *et al.*, 1982). This mutant was found to induce a good immune response, with growth restricted to the upper respiratory tract. A secondary phenotype has been described for this mutant. The mobility of the F protein in SDS-PAGE analysis was retarded under non-reducing conditions (Belshe *et al.*, 1978). When this candidate vaccine was used in children, non-ts revertant viruses were isolated from the lower respiratory tract. These revertants emerged late in infection and in most cases represented incomplete reversion to the wild-type phenotype (Hodes *et al.*, 1974) with viral titres lower than that of the wild-type virus at the non-permissive temperature.

Analysis of the phenotype of mutant *ts*A1 suggested that the lesion was present in the fusion gene, possibly in the F_2-subunit encoding region of the gene (Belshe *et* al., 1978; Caravokyri, 1990). The nature of the reversion phenomenon appeared to be complex with two types found. The terminology of Caravokyri, (1990) is used here. **(1)** 90% of clones were fully revertant viruses which had wild-type growth characteristics at the non-permissive temperature. However these revertants still retained the mutant mobility phenotype. **(2)** 10% of clones were partial revertant viruses which had intermediate growth characteristics at the non-permissive temperature but had reverted to the wild-type fusion protein mobility.

It was therefore the purpose of this study to identify the nature of the ts lesion and to examine any nucleotide changes in the fusion protein gene of the two classes of revertant viruses in an effort to explain the nature of reversion.

4.2 Results All viruses were grown in BS-C-1 G23 African green monkey kidney cell culture at the appropriate temperatures. Fusion genes were amplified using PCR from total RNA from RS virus-infected cells via a cDNA product. The genes were then ligated into sequencing vectors M13mp18/19 and sequenced by di-deoxy chain termination. At least two plaque-purified clones containing the fusion protein gene were sequenced for each virus.

4.2.1 Sequence of the *ts*A1 fusion protein gene The deduced nucleotide sequence for the fusion gene of mutant *ts*A1 and its predicted amino acid sequence revealed seven coding changes in comparison to the wild-type sequence of strain A2 (**Table 12**).

What is immediately apparent is that as predicted by Caravokyri, (1990), nearly all coding changes in the mutant occurred in the F_2-subunit of the protein. The number of coding changes appeared higher than expected (as compared to those occurring in mutant *ts*1C, **Section 5.2.2**) but this was probably due to the method used to generate this mutant, where a high concentration of mutagen and low stringency selection were used. Coding changes in the fusion gene product are described below.

There were changes at residue #4 and at residue #8 and these are also found in several isolates and strains of subgroup A viruses (**Section 3.2.4**). The change at residue #8 occurred at a nucleotide that has previously been defined as prone to variability (Anderson *et al.*, 1992). Both these changes in residue are thought to be conservative in nature with charge, hydrophobicity, and secondary structure forming characteristics remaining unchanged. These changes are thought not to be the site of the ts lesion or mobility phenotype because of their conservative nature, the fact that they are present in several subgroup A RS viruses, and the fact that they occur in the signal domain of the F_2-subunit which is removed during transport through the endoplasmic reticulum to the Golgi apparatus (**Section 1.7.5**).

Table 12 Coding changes in mutant *ts*A1 compared to wild-type strain A2

Residue #	Nucleotide change	Amino acid	Comments	Domain
#4	C**U**A → A**U**A	Leu → Ile	Conservative	F_2 signal domain
#8	G**C**A → A**C**A	Ala → Thr	Conservative	F_2 signal domain
#66	G**A**A → A**A**A	Glu → Lys	Non-conservative	F_2 conserved #1 domain
#81	C**A**A → C**U**A	Gln → Leu	Conservative	F_2 conserved #1 domain
#102	C**C**A → U**C**A	Pro → Ser	Non-conservative	F_2 C-terminal domain
#379	A**U**A → G**U**A	Ile → Val	Variable residue	F_1 cysteine-rich domain
#447	A**U**G → G**U**G	Met → Val	Variable residue	F_1 conserved #3 domain

NB Those bases that are underlined and in bold type denote the changed nucleotide and its position in the codon triplet. Coding changes that are thought to be significant are shown highlighted in shaded boxes. Variable residues are those defined as such by Collins *et al.*, (1984b).

The change at residue #66 is a non-conservative change that could potentially have a significan effect on fusion protein structure due to the replacement of an acidic residue with a basic residue (Glu – Lys). This change alters the secondary-structure forming characteristics of this residue from strong α-heli forming/ strong β-sheet breaking residue to one that is a weak α-helix former/ moderate β-sheet breaker as defined by Chou and Fasman, (1978).

The coding change at residue #81 (Gln → Leu) is conservative in nature and was not predicted t have a structural effect. This change was thought unlikely to be the site of the ts lesion due to the patter of coding changes in the revertant viruses (**Sections 4.2.3 and 4.2.4.**).

The coding change at residue #102 (Pro → Ser) is a non-conservative change and may have th potential to alter protein structure by replacing an α-helix breaking residue with an α-helical formin residue (Chou and Fasman, 1978). However, a change from proline to alanine at residue #102 was foun in many subgroup A RS viruses (**Section 3.2.4**), and a change from Pro → Ser was reported in 50% o cDNA clones sequenced by Collins *et al.*, (1984b). This feature will be discussed later in **Section 4.3**.

There were two coding changes located in the F_1-subunit of the fusion protein, at residues #379 (Il → Val) and residue #447 (Met → Val). These are not thought to be actual coding changes specific to thi mutant, but rather to be errors in the original published consensus sequence of the wild-type strain A (Collins *et al.*, 1984b; **Section 3.3.1**) since these changes are found in all naturally occurring isolates o RS virus, both subgroup A and B and also in bovine strains of RS viruses (**Section 3.3.3**). Therefore ther are thought to be two potential ts lesion-causing changes in mutant *ts*A1 (residues #66 and/or #102).

4.2.2 Analysis of the fusion gene sequence of fully and partially revertant viruses Revertant viruses were isolated from mutant *ts*A1, essentially by picking and purifying viral plaques that grew at the non-permissive temperature. This was possible because of the high rate of reversion of the parental mutant in the absence of mutagenisation. As described previously (Caravokyri, 1990), fully revertant viruses had a rate of growth equivalent to that of the wild-type virus and were far more prevalent (approximately 90%)

than partially revertant viruses (10%). The revertant virus isolates were classified on the basis of their efficiency of plating (**Table 13**). A coefficient of efficiency of plating < 0.3 and $> 10^{-2}$ was used to define a partially revertant virus (*i.e.* one that remained partially temperature-sensitive). The revertants were also tentatively characterised in terms of the electrophoretic mobility profiles of their fusion proteins under non-reducing conditions (**Figure 16**). The pattern of reversion is complex, with the partial revertant viruses being present at lower frequency than full revertants. One fully revertant and one partially revertant virus was analysed for their fusion gene sequence.

4.2.3 Fusion gene sequence of a fully revertant virus Sequence analysis of the fusion gene of the fully revertant virus is summarised in **Table 14** and important coding changes shown in **Figures 17 and 18**. Determination of the predicted amino acid sequence of this virus revealed that residues #4, #8 and #81 were restored to the wild-type sequence. Importantly, the change at position #66 (Glu → Lys) had been lost, with the residue reverting to the amino acid (glutamic acid) found in the wild-type strain. The effect of this was to remove one of the two potential ts-causing mutations. The remaining coding changes found in mutant *ts*A1 had been retained, including the two variable residues of strain A2 (residues #379 and #447 of the F_1-subunit). Unexpectedly, one extra coding change had been acquired at residue #509 (Ser → Thr), located in the conserved region (C#3) of the F_1-subunit of the fusion protein. This change is conservative in nature and may not affect fusion protein function or mobility, but since it occurs in the conserved C#3 region of the protein, an effect on structural features cannot be dismissed.

4.2.4 Fusion gene sequence of a partially revertant virus Sequence analysis of the fusion gene of the partial revertant virus is summarised in **Table 15** and important coding changes shown in **Figures 17 and 18**. A more complex pattern of changes was observed in this revertant virus. Most of the coding changes acquired in mutant *ts*A1, including the mutations at residues #66 and #102 potentially associated with the ts lesion, had been retained

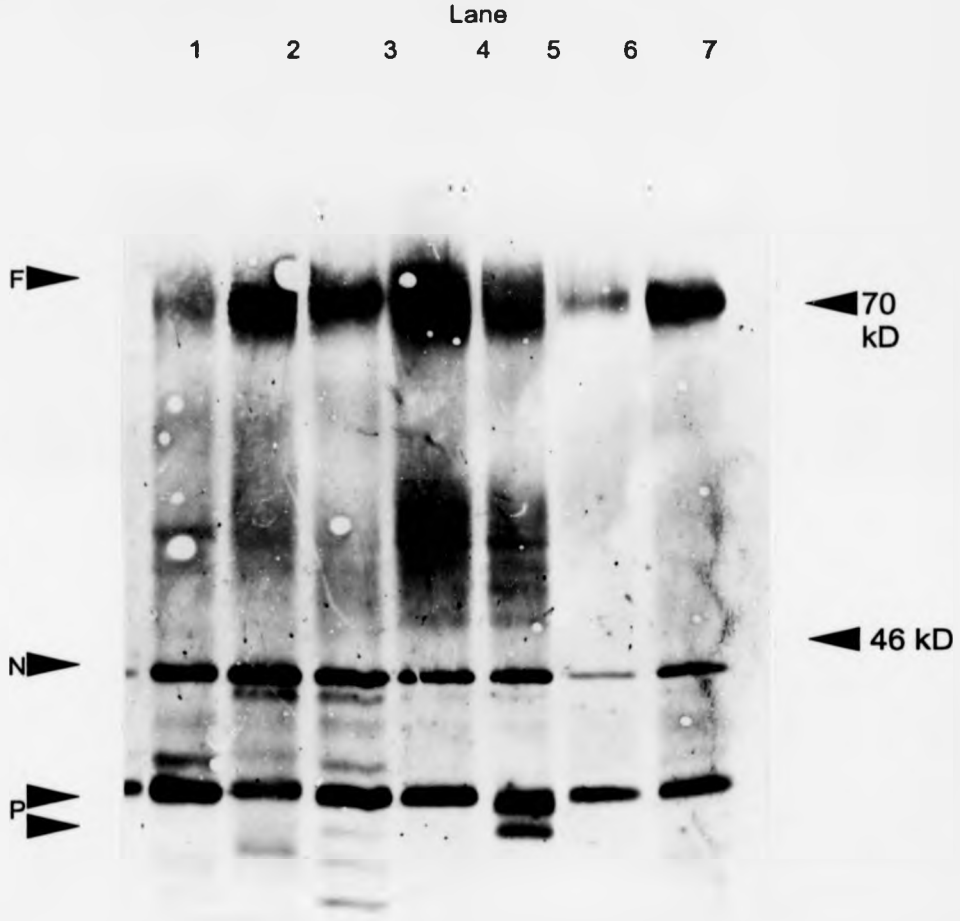
Table 13 Efficiency of plating of mutant *ts*A1 and its ts$^+$ revertants

Virus clone	Viral yield (pfu/ml)		Efficiency of plating [1]
	33° C	39° C	
Wild-type A2	1.2×10^6	0.8×10^6	0.67
Mutant *ts*A1	1.9×10^5	2.3×10^2	1.21×10^{-3}
Full revertant 1.1	1.6×10^5	1.5×10^5	0.94
Full revertant 5.0	6.5×10^5	4.1×10^5	0.63
Full revertant 8.1	3.2×10^4	2.7×10^4	0.84
Partial revertant 3.4	1.9×10^5	2.8×10^4	0.15

[1]....Efficiency of plating = $\frac{\text{Viral yield at 39°C}}{\text{Viral yield at 33°C}}$

The non-permissive temperature for mutant *ts*A1 is >37°C.

Figure 16 . Identification of type of revertant by fusion-protein mobility on SDS-PAGE analysis.



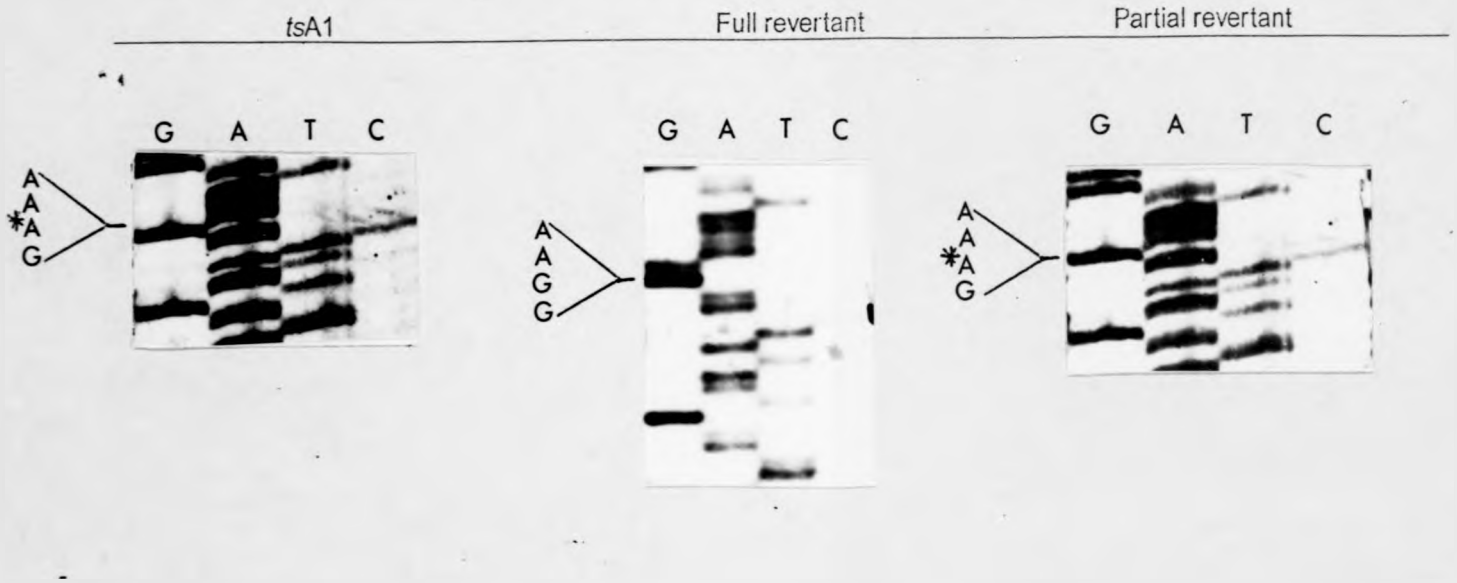
Total protein was harvested from virus-infected BS-C-1cells at the permissive temperature and the proteins immunoprecipitated (to reduce the sample volume) with bovine polyclonal anti-RS virus antibody at a dilution of 1:100 as described in **Section 2.2.3**. Samples were then run on a 10% SDS-PAGE under non-reducing conditions and the proteins transfered via Western-blotting onto a nitrocellulose filter which was developed as described in Chapter 2. Three RS viral proteins are shown; the fusion protein (F), the nucleoprotein (N) and the phosphoprotein (P). Lanes 1, 2, 3 and 7 show characteristic retarded fusion mobility found in the fully revertant viruses. Lane 4 shows wild--type fusion protein mobility characteristic of the partially revertant viruses. Lane 5 is the wild-type virus A2, and lane 6 is the mutant virus *ts*A1.

Table 14 Coding changes in a fully revertant virus in comparison to mutant *ts*A1

Residue #	Nucleotide change	Amino acid	Protein domain	Comments
#4	**A**UA → **C**UA	Ile → Leu	F_2 signal	Restores A2 sequence
#8	**A**CA → **G**CA	Thr → Ala	F_2 signal	Restores A2 sequence
#66	**A**AA → **G**AA	Lys → Glu	F_2 conserved #1	Restores A2 sequence
#81	C**U**A → C**A**A	Leu → Gln	F_2 conserved #1	Restores A2 sequence
#102	**none.**	Ser	F_2 C-terminal	Retains *ts*A1 sequence
#379	**none.**	Val	F_1 cysteine-rich	Retains *ts*A1 sequence
#447	**none.**	Val	F_1 conserved #3	Retains *ts*A1 sequence
#509	**U**CC → **A**CC	Ser → Thr	F_1 conserved #3	Conservative change

NB Those letters underlined and in bold type denote the changed nucleotide and its position in the codon triplet. **None.** is where the predicted amino acid is the same as that of mutant *ts*A1. The coding change that is thought to be significant is shown highlighted.

Figure 17. Coding change at residue # 66 in mutant tsA1 and its revertant viruses, in comparison to the published sequence of the wild-type virus strain A2.



nucleotide sequence	A2 (Collins et al.,1984b)		TTA AGT AAT ATC AAG GAA AAT AAG
	tsA1		TTA AGT AAT ATC AAG AAA AAT AAG
	Full revertant		TTA AGT AAT ATC AAG GAA AAT AAG
	Partial revertant		TTA AGT AAT ATC AAG AAA AAT AAG
			* *
Predicted amino acid sequence:	A2	LSNIKENK	
	tsA1	LSNIKKNK	
	Full revertant	LSNIKENK	
	Partial revertant	LSNIKKNK	
		*	

N.B. Nucleotide changes are denoted by an asterix at the relevant position.

Figure 18. Coding change at residue # 102 , #103 and #105 in mutant *ts*A1 and its revertant viruses , in comparison to the published sequence of the wild-type strain A2

nucleotide sequence:	A2 (Collins et al.,1984b)	A AGC ACA CCA CCA ACA AAC AAT CGA
	tsA1	A AGC ACA CCA TCA ACA AAC AAT CGA
	Full revertant	A AGC ACA CCA TCA ACA AAC AAT CGA
	Partial revertant	A AGC ACA CCA TCA GCC AAC AGT CGA
		* * * *
Predicted amino acid sequence:	A2	STPPTNNR
	tsA1	STPSTNNR
	Full revertant	STPSTNNR
	Partial revertant	STPSANSR
		** *

N.B. Nucleotide changes are denoted by an asterix at the relevant position.

Table 15 Coding changes in a partially revertant virus in comparison to mutant *ts*A1

Residue #	Nucleotide change	Amino acid	Protein domain	Comments
#4	**A**UA → **C**UA	Ile → Leu	F_2 signal	Restores A2 sequence
#8	**A**CA → **G**CA	Thr → Ala	F_2 signal	Restores A2 sequence
#66	**none**	Lys	F_2 conserved #1	Retains *ts*A1 sequence
#81	**none**	Leu	F_2 conserved #1	Retains *ts*A1 sequence
#102	**none**	Ser	F_2 C-terminal	Retains *ts*A1 sequence
#103	**A**C**A** → **G**C**C**	Thr → Ala	F_2 C-terminal	Non-conservative change
#105	A**A**U → A**G**U	Asn → Ser	F_2 C-terminal	Conservative change
#379	**none**	Val	F_1 cysteine-rich	Retains *ts*A1 sequence
#447	**none**	Val	F_1 conserved #3	Retains *ts*A1 sequence
#509	**U**CC → **A**CC	Ser → Thr	F_1 conserved #3	Conservative change

NB Those letters that are underlined and in bold type denote the changed nucleotide and its position in the codon triplet. **None.** is where the predicted amino acid is the same as that of mutant *ts*A1. The coding change that is thought to be significant is highlighted in a shaded box.

in the partial revertant, whereas residues #4 and #8 had changed to restore the A2 sequence at these positions. However, there were a series of additional coding changes acquired in the F_2- subunit encoding region of the gene. Immediately downstream of residue #102 there were two coding changes (residue #103 Thr → Ala and residue #105 Asn → Ser). The change at residue #105 is conservative in nature and the change at residue #103 is non-conservative, replacing an α-helix indifferent residue with a strong α-helix forming residue (Chou and Fasman, 1978). These changes are considered to be potentially important. The coding changes observed in the F_1-subunit of mutant *ts*A1, at residues #379 and #447 are retained in the partially revertant viruses as they are in the fully revertant viruses, and the partially revertant virus had also acquired the extra coding change at residue #509 (Ser → Thr) found in the fully revertant viruses.

A summary of coding changes in mutant *ts*A1, and both classes of revertant viruses is given in **Table 16**.

4.3 Discussion There are two amino acid changes that are likely to be associated with the ts lesion and electrophoretic mobility phenotype, residue #66 (Glu → Lys) and residue #102 (Pro → Ser). The presence of a serine residue at position #102 in two out of four cDNA clones analysed by Collins *et al.*, (1984b), does not exclude this change as a true difference in coding sequence between the wild-type virus and its mutant. There is evidence that the original isolate of strain A2 was heterogenous (reviewed in **Section 3.3.1**), and it is feasible that the plaque purification process used during the generation of mutant *ts*A1 could have isolated a variant of the A2 virus that had the proline residue at this position. It is also equally possible that during the time between the original isolation of strain A2 and subsequent derivation of its ts mutant, and the time of the elucidation of the fusion gene sequence, the virus could have acquired mutations, possibly due to adaptation to growth in cultured cells.

The change at residue #66 involved the replacement of an acidic amino acid with one that is basic, thus altering the predicted secondary-structure forming characteristics from

Table 16 Summary of amino acid changes in mutant *ts*A1 and its revertant viruses in comparison to wild-type A2.

Virus	Fusion mobility	F_2-subunit							F_1-subunit		
		Signal		Conserved C#1		C-terminal			Cysteine--rich	Conserved C#3	
		#4	#8	#66	#81	#102	#103	#105	#379*	#447*	#509
A2 w.t.	wild-type	Leu	Ala	Glu	Gln	Pro	Thr	Asn	Ile	Met	Ser
*Ts*A1	retarded	Ile	Thr	Lys	Leu	Ser	Thr	Asn	Val	Val	Ser
Partial revertant	wild-type	Leu	Ala	Lys	Leu	Ser	Ala	Ser	Val	Val	Thr
Full revertant	retarded	Leu	Ala	Glu	Gln	Ser	Thr	Asn	Val	Val	Thr

N.B. The changes denoted by * are considered to represent errors in the published sequence of the wild-type strain A2 (Collins *et al.*, 1984b).

strong α-helix forming/ strong β-sheet breaking residue to one that is a weak α-helix former/ moderate β-sheet breaker. Therefore this amino acid change has the potential to alter the loop structure (residues #62 to 78) that separates the N-terminal region from the C-terminal (F_2h containing) region of the F_2-subunit. This loop-forming region is common to all paramyxovirus fusion proteins (Chambers *et al.*, 1992) and is therefore thought to be an important structural feature of the protein. The acquisition of the coding change at residue #66 could possibly destabilise the loop and thereby account for the restricted growth of mutant *ts*A1 at high temperatures. Loss of this change could possibly re-stabilise the loop structure and allow for growth at the higher temperature. The change at residue #102 alters the amino acid characteristic from a strong α-helix breaker to an α-helix indifferent residue. This could remove an α-helix breaking residue and potentially lengthen the F_2h sequence of α-helixes (residue #79 to #100) described by Chambers *et al.*, (1992) as common to all paramyxoviruses, and hence structurally important. The coding change at residue #509 (Ser → Thr) whilst considered conservative, occurs in the conserved C#3 region of the F_1-subunit and may have a structural effect. However, the pattern of acquisition and loss of coding changes in the F_2-subunit conforms to acquisition or loss of phenotypes, suggesting that it is changes in the F_2-subunit that account for the two phenotypes. The other coding changes found in the fusion protein of mutant *ts*A1 are thought to be irrelevant.

The acquisition of wild-type growth characteristics at the non-permissive temperature by fully revertant viruses accompanied the loss of the coding change at residue #66. Therefore this mutation may be the site of the ts lesion, possibly accomplishing this in the way described above.

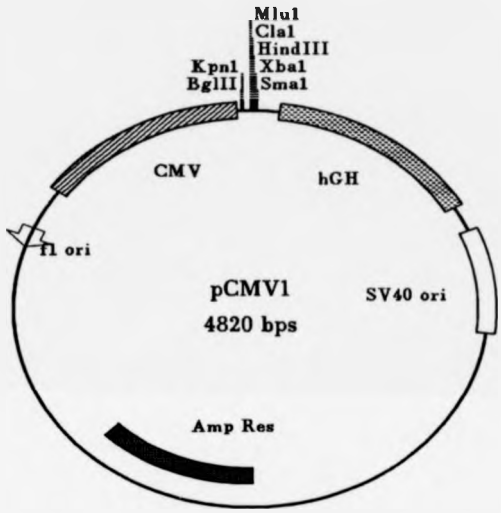
Analysis of the additional coding changes observed in the partially revertant viruses (specifically residues #103 and #105), and comparison with the known acquisition of wild-type mobility in this class of revertant, suggests that the mutation causing the mobility shift is the change at residue #102. It is proposed that the mobility phenotype is caused by disruption in the C-terminal secondary structure (specifically F_2h) of the F_2 subunit by extending the α-

helix structure. Acquisition of the extra coding change at residue #103 may modify the disrupted C-terminal secondary structure sufficiently to counteract the change at residue #102 which appeared first in mutant *ts*A1. With the partial revertant virus still retaining the presumptive ts lesion at residue #66 it is proposed that as well as correcting for mobility, the change at residue #103 also partially reverses the ts lesion, possibly by modification of the local structure of this subunit and stabilising the disrupted loop structure.

The genetic complexity of mutant *ts*A1 is probably a consequence of the conditions of mutagenesis which were designed to maximise the yield of mutants, rather than induce single site mutations.

4.4 Expression of the fusion proteins of mutant *ts*A1 and its revertant viruses The above conclusions are that sequence changes in the fusion gene lead to structural changes in the fusion gene product that accounted for both the ts and F protein mobility phenotypes. However the mobility phenotype might be a function of changes in the interaction of the fusion gene with other viral proteins. In order to confirm the role of the F protein it was necessary to express the fusion gene product in isolation from the other viral proteins. The method chosen was to insert the fusion genes from the ts mutant, both types of revertant viruses, and the wild-type strain A2, seperately into an expression vector and transiently transfect mammalian cells in tissue-culture and examine the expressed viral fusion protein. The expression vector chosen for this work was vector pCMV1 (**Figure 19**). This utilises the human cytomegalovirus immediate early promoter which is of high efficiency. The fusion genes previously sequenced were ligated into this vector to determine whether the electrophoretic mobility phenotype was retained when the F gene were expressed independently of the rest of the genome. Expression levels were enhanced by the use of type 293 mammalian cells which had been adapted from HeLa cells by the insertion of the adenovirus E1A gene (Graham and Van der Ebb, 1973). The E1A product is a powerful transactivator of the CMV promoter and has also been demonstrated as being a DNA binding protein that can

Figure 19. Eukaryotic expression vector pCMV1



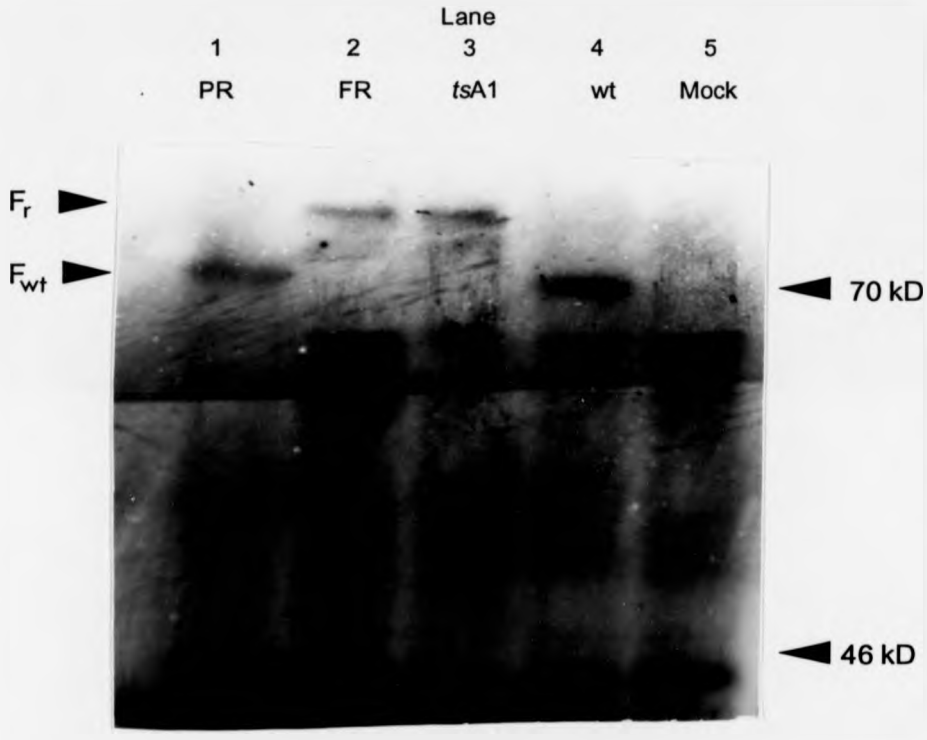
Plasmid details.

The pCMV plasmids are a series of mammalian expression vectors developed in David Russell's laboratory (University of Texas, Dallas, USA.). The vector backbone is pTZ18R, containing a bacteriophage f1 origin of replication for production of single stranded DNA, and an ampicillin-resistance gene. The CMV region consists of nucleotides -760 to +3 of the strong promoter-regulatory region of human cytomegalovirus (Towne strain) major immediate early gene. The human growth hormone fragment (hGH) contains transcription termination and poly-adenylation signals representing sequences 1533 to 2157 of this gene. The SV40 origin of replication and early region promoter-enhancer are also included in the construct. The promoter in this fragment is oriented such that transcription proceeds away from the CMV/hGH cassette.

stabilise transfected DNA and prevent early degradation by cellular enzymes (Akusjarvi, 1993). After transfection the cellular proteins were labelled with [^{35}S] methionine and the fusion protein was removed from the protein pool by immunoprecipitation with a polyclonal bovine anti-RS virus serum. After analysis on SDS-PAGE under non-reducing conditions, the mobility phenotype was determined by autoradiography (**Figure 20**). Normal wild-type mobility was observed in strain A2 and the partially revertant virus, and retarded mobility observed in both the *ts*A1 mutant and the fully revertant virus. This confirmed that the coding changes in the F gene were responsible for the mobility differences observed, and that interaction with other viral proteins was not involved in this phenotype.

4.5 Conclusion There are two distinctive phenotypes observed in the mutant virus *ts*A1, temperature-sensitivity and a retarded fusion protein mobility in non-reducing gels. It is suggested that the site of the ts lesion is located at residue #66 (Glu → Lys) and that of the mobility phenotype at residue #102 (Pro → Ser). The nature of reversion is complex with two classes of revertant viruses observed; a fully revertant virus which has wild-type growth characteristics but still retains retarded fusion protein mobility; and a partially revertant virus that possesses near wild-type growth characteristics and wild-type mobility. The coding change correcting for the mobility phenotype, and partially correcting for the ts phenotype in the partial revertant viruses, has been tentatively identified as additional coding changes at residues #103 and #105. For the fully revertant viruses it has been deduced that changing residue #66 to the wild-type amino acid fully corrects for the ts phenotype. *In vitro* expression of the fusion gene products of mutant, revertant and wild-type viruses in mammalian cells has confirmed that the mobility phenotype is solely a consequence of changes in the fusion protein gene.

Figure 20. Expression of the fusion proteins of mutant tsA1, Full and Partial revertant viruses.



Viral fusion protein genes were inserted into the pCMV1 expression vector and transfected into monolayers of 293 cells. After 36 hours the expressed proteins were labeled for 4 hours with radioactive methionine. Viral proteins were then immunoprecipitated with murine monoclonal anti-F antibody (Mab F8) at a dilution of 1:100. Samples were then run on a 10% SDS-PAGE under non-reducing conditions overnight. This over-run was to ensure that the expressed fusion protein mobility differences were emphasised. The gel was then dried and exposed as described in Chapter 2 in order to visualise the expressed fusion proteins which have either wild-type mobility (Fwt) or retarded mobility (Fr). Lane 1 shows the fusion protein from the partially revertant virus. Lane 2 shows the fusion protein from the fully revertant virus. Lane 3 shows the fusion protein isolated from the mutant tsA1 and lane 4 shows the fusion protein from the wild-type virus A2. Lane 5 is the negative control of the expression vector containing no insert.

CHAPTER FIVE

CHAPTER FIVE

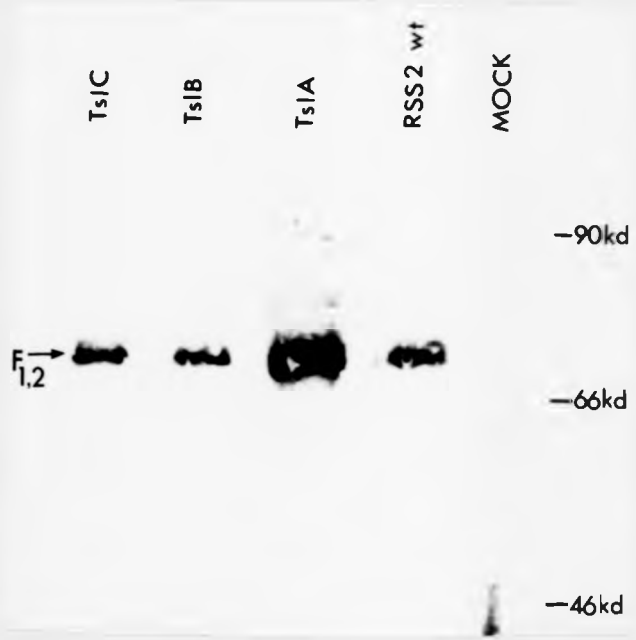
Genetic mutations in the fusion gene of vaccine-candidate mutant *ts*1C

5.1 Introduction A candidate ts-mutant vaccine has been derived from the RSS-2 strain of RS virus (**Section 1.10**). The conditions under which mutant *ts*1C was derived were chosen to optimise the recovery of single-site ts lesions, with the generation of the mutant *ts*1C via intermediate viruses (*ts*1A and *ts*1B). Different mutagenising agents were employed at each stage of the process. Since the ts lesion in the *ts*A1 candidate vaccine derived by Chanock and colleagues (Gharpure *et al.*, 1969; Wright *et al.*, 1982) appeared to be located in the fusion protein gene (**Section 4.3**), it was appropriate to investigate the F protein gene of mutant *ts*1C. A formalin-inactivated vaccine derived by Chanock and colleagues in the 1960s, induced an unbalanced immune response in children which was responsible for exacerbated disease on subsequent natural infection of the children with RS viruses. This incident is now considered to be due to the destruction of F and/or G protein epitopes by the inactivation treatment. It was important therefore to ascertain whether in the derivation of mutant *ts*1C the fusion protein gene had been altered and whether any observed nucleotide changes could potentially affect known epitopes on the fusion protein.

5.2 Results The isolation of the wild-type strain RSS-2 and all subsequent characterisation of the original isolate, the derivation of the ts mutants, and preparation of stock virus were carried out in human diploid (MRC-5) cells (McKay *et al.*, 1988). Therefore all viral manipulations described here were performed in this cell line.

5.2.1 SDS-PAGE analysis of fusion proteins SDS-PAGE analyses of the fusion gene products from the three mutants (*ts*1A, *ts*1B and *ts*1C) synthesised at the permissive temperature have confirmed that the fusion proteins from mutants (*ts*1A, *ts*1B and *ts*1C) cannot be differentiated from wild-type RSS-2 strain by electrophoretic mobility (**Figure 21**).

Figure 21. Comparison of the fusion proteins of *ts* mutants *ts*1A, *ts*1B and *ts*1C by SDS-PAGE analysis.



Total protein was harvested from virus-infected MRC-5 cells at the permissive temperature and the fusion proteins immunoprecipitated with mouse monoclonal anti-RS virus fusion protein antibody (Mab F8) at a dilution of 1:100. Samples were then run on a 12% SDS-PAGE under non-reducing conditions and the proteins transfered via Western-blotting onto a nitrocellulose filter which was developed as described in **Section 2**.

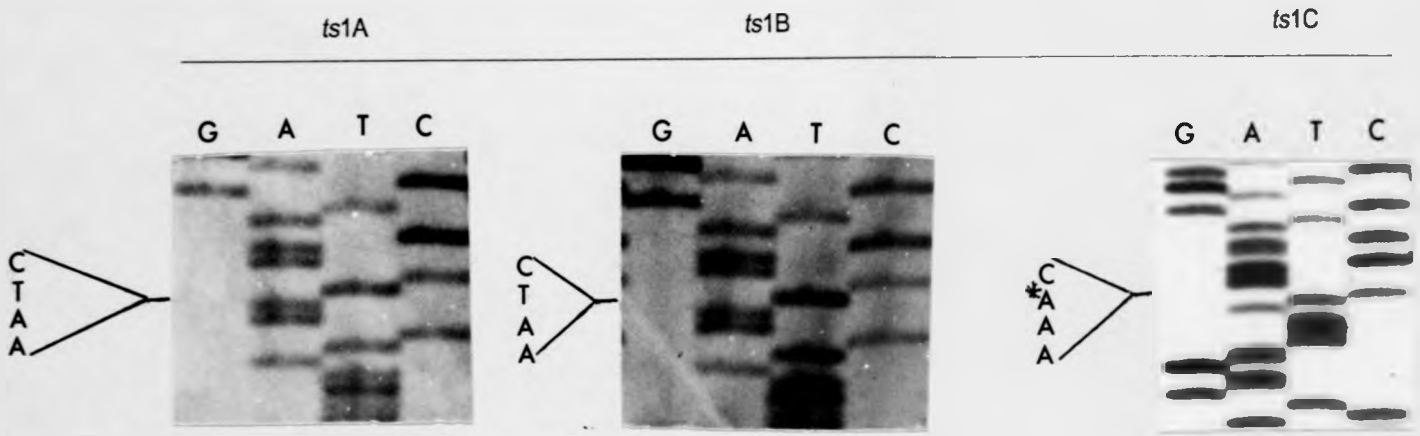
The fusion gene products from all three mutants and the wild-type virus showed the same mobility for $F_{1,2}$ on non-reducing and for F_1 and F_2 subunits on reducing gel SDS-PAGE (C.R.Pringle, personal communication). Quantitative differences in samples are probably due to inequalities in the amounts loaded in the different tracks.

5.2.2 The fusion protein gene sequence of mutant *ts*1C Determination of the nucleotide sequence of the fusion protein gene of *ts*1C and its intermediates was carried out by PCR amplification of the full length gene via a cDNA intermediate generated from total RNA extracted from virus-infected cell monolayers. The primers, PCRF1 and PCRF2 were designed to be complementary to the 5' and 3' ends of the gene respectively, both incorporated a Sma 1 restriction endonuclease site which was used to insert the full-length gene directly into vectors M13mp18/19. Two plaque-purified M13 cultures of each mutant were sequenced as described **(Section 2.3.21)**.

There are eight nucleotide differences between mutant *ts*1C and the wild-type RSS-2, potentially coding for four amino acid changes. Nucleotide changes occur in all three codon positions and as expected, mutations in the third codon position (residues #73 GA**U** → GA**C**; #166 AA**G** → AA**A**; #316 CU**A** → CU**G** and #447 GU**G** → GU**U**) do not change the encoded amino acid residue. However mutations in both the first codon (#434 **U**CA → **A**CA and #523 **G**AU → **A**AU) and the second codon positions (#1301 A**G**A → A**C**A and #1580 A**C**A → A**U**A) do give rise to potential amino acid changes. The four amino acid changes appear to identify three mutational sites. The nucleotide changes for each of the three mutational sites #1, #2 and #3 are shown in detail for each ts mutant **(Figures 22, 23 and 24)** and the data for *ts*1C are summarised in **Table 17**.

Mutation site #1 (residue #35 Ser → Thr) is located in the F_2 subunit of the fusion protein, in a putative α-helical region F_2h **(Section 1.7.5 Figure 5)**. The serine to threonine exchange represents a conservative change with relative size, charge and hydrophobicity profiles unaltered. Mutation site #2 lies within the F_1 subunit in the cysteine-rich domain

Figure 22. Coding changes at mutation site #1 between the published sequence of the wild-type virus RSS-2 and the sequence of mutant *ts*1C and its intermediate viruses, ts1A and ts1B

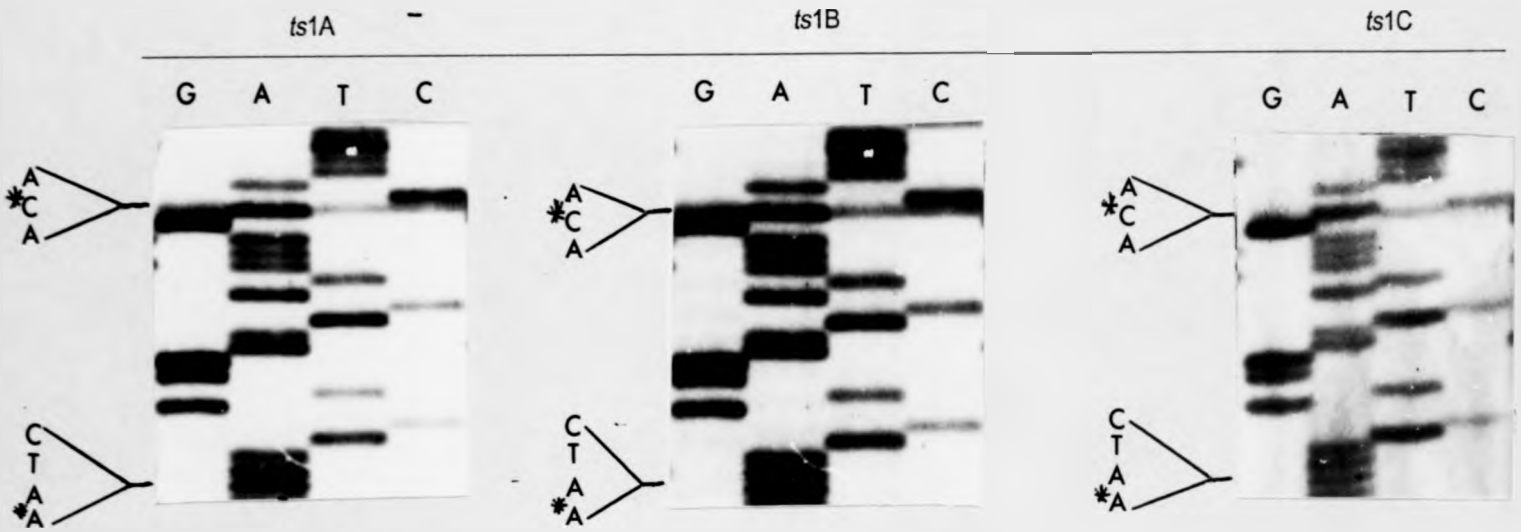


Nucleotide:	RSS-2 (Baybutt and Pringle,1987)	TTT TAT CAA TCA A
sequence	*ts*1A	TTT TAT CAA TCA A
	*ts*1B	TTT TAT CAA TCA A
	*ts*1C	TTT TAT CAA ACA A
		*

Predicted amino acid sequence:	RSS-2	FYQS
	*ts*1A	FYQS
	*ts*1B	FYQS
	*ts*1C	FYQT
		*

N.B. Nucleotide changes are denoted by an asterix at the relevant position.

Figure 23. Coding changes at mutation site #2 between the published sequence of the wild-type virus RSS-2 and the sequence of mutant ts1C and its intermediate viruses, ts1A and ts1B



Nucleotide:	RSS-2 (Baybutt and Pringle,1987)	AAT AAA GAT CGT GGA ATC ATA AAG AGA TTT T
sequence	*ts*1A	AAT AAA AAT CGT GGA ATC ATA AAG ACA TTT T
	*ts*1B	AAT AAA AAT CGT GGA ATC ATA AAG ACA TTT T
	*ts*1C	AAT AAA AAT CGT GGA ATC ATA AAG ACA TTT T
		* *

Predicted amino acid sequence:	RSS-2	NKDRGIIKRF
	*ts*1A	NKNRGIIKTF
	*ts*1B	NKNRGIIKTF
	*ts*1C	NKNRGIIKTF
		* *

N.B. Nucleotide changes are denoted by an asterix at the relevant position.

Figure 24. Coding changes at mutation site #3 between the published sequence of the wild-type virus RSS-2 and the sequence of mutant ts1C and its intermediate viruses, ts1A and ts1B

Nucleotide:	RSS-2 (Baybutt and Pringle,1987)	AAA TCC ACC ACA AA
sequence	*ts*1A	AAA TCC ACC ATA AA
	*ts*1B	AAA TCC ACC ATA AA
	*ts*1C	AAA TCC ACC ATA AA
		*

Predicted amino acid sequence:	RSS-2	KSTT
	*ts*1A	KSTI
	*ts*1B	KSTI
	*ts*1C	KSTI
		*

N.B. Nucleotide changes are denoted by an asterix at the revelant position

Table 17 Summary of the coding and non-coding changes in the fusion gene of *ts*1C compared to the wild-type virus RSS-2

Nucleotide changes		Coding changes	
Nucleotide #	Codon Triplet	Amino acid #	Amino acid
103	**UCA → ACA**	**35**	**Ser → Thr**
219	GAU → GAC	73	No change
498	AAG → AAA	166	No change
948	CUA → CUG	316	No change
1282	**GAU → AAU**	**428**	**Asp → Asn**
1301	**AGA → ACA**	**434**	**Arg → Thr**
1341	GUG → GUU	447	No change
1580	**ACA → AUA**	**523**	**Thr → Ile**

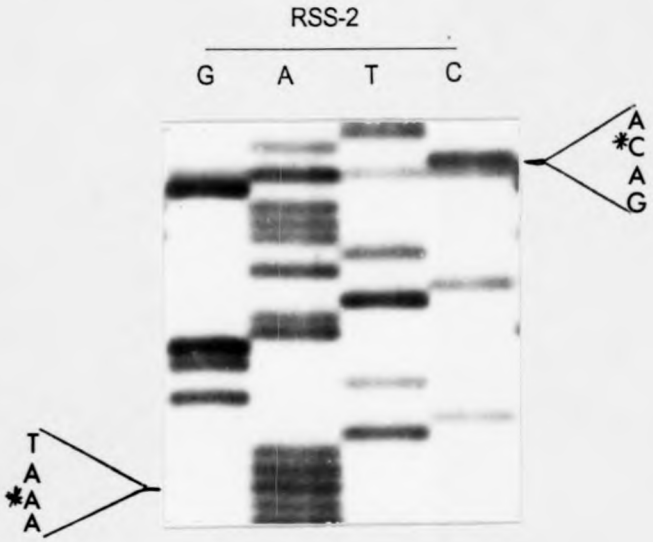
N.B. Letters in bold and underlined denote a base within the codon which was mutated in *ts*1C. Coding changes are highlighted in shaded boxes.

(residues #312 to #440; **Section 1.7.4 Figure 4**). There are two changes in amino acids (residue #428 Asp → Asn and residue #434 Arg → Thr) which have altered bases in their first and second codons respectively (#428 GAU → AAU and #434 AGA → ACA). These changes are less conserved in that the aspartic acid to asparagine change removes an acidic amino acid and replaces it with a neutrally charged residue, but the hydrophilic nature of this residue is retained. The change of arginine to threonine also alters the charge of the substituted molecule from basic to neutral but the hydrophilic nature once again is retained. However, there is a size change from large to small. The retention of hydrophilic characteristics implies that these amino acids lie on the surface of the folded protein. The substituted amino acid sequence present at this site is identical to that found in the wild-type strain A2 (Collins *et al.*, 1984b).

The residues present in *ts*1C at position #428 and #434 are found in all strains of RS viruses that have been sequenced, other than the RSS-2 wild-type virus (**Section 3.2.5**). This raises two possibilities, firstly the original published sequence data for strain RSS-2 could be in error at these two positions; secondly, these two sites may indeed be valid mutations that appear to be linked suggesting a possible structural role. To investigate the first possibility, the F gene from wild-type strain RSS-2 was PCR amplified, cloned into M13mp19 and sequenced across the three identified mutational sites. The nucleotide sequence at sites #1 and #3 were found to be as published, however the sequence at site #2 was found to agree with that found in mutant *ts*1C (**Figure 25**) and all other strains of RS virus. Thus it would appear that the original sequence published by Baybutt and Pringle (1987) is in error at nucleotides #1282 and #1301 (residues #428 and #434) and there are no mutational changes in this region. Mutation site #2, therefore, is an artefact.

Mutation site #3 (Thr → Ile) occurs in the second base position of the codon (ACA → ATA). This change is located in the F_1 subunit of the fusion protein immediately upstream of the membrane anchor region (**Section 1.7.4 Figure 4**). The change is conservative in that

Figure 25. Mutation site #2 is an error in the published sequence of strain RSS-2.




```
Nucleotide:              RSS-2 (Baybutt and Pringle,1987)  AAT AAA GAT CGT GGA ATC ATA AAG AGA TTT T
sequence                 RSS-2 (Sequence from this study)  AAT AAA AAT CGT GGA ATC ATA AAG ACA TTT T
                                                                    *                         *

Predicted amino acid sequence:  RSS-2 (Baybutt and Pringle,1987   NKDRGIIKRF
                                RSS-2 (Sequence from this study)  NKNRGIIKTF
                                                                    *    *
```

N.B. Nucleotide changes are denoted by an asterix at the relevant position.

both the original and substituted residues are hydrophobic, neutrally charged and of similar size.

5.2.3 The fusion gene sequence of intermediate mutants *ts*1A and *ts*1B Analysis of the fusion protein gene sequences of the intermediate mutant viruses *ts*1A and *ts*1B was used to establish at which step in the derivation of the triple ts mutant *ts*1C, the mutations occurred. Of the nucleotide mutations in *ts*1C (excluding the two errors at residues #428 and #434) two non-coding nucleotide changes (#166 AA**G** → AA**A** and #523 A**C**A → A**U**A) and one coding change (residue #527 Thr → Ile) are also present in the fusion protein gene of mutants *ts*1A and *ts*1B. Therefore the remaining four nucleotide mutations that only appear in mutant *ts*1C, of which only one introduces a coding change (residue #35 Ser → Thr), must have been introduced during the mutational step from *ts*1B to *ts*1C. The sequential accumulation of mutations in the fusion protein genes from *ts*1A through to *ts*1C is shown in **Section 5.2.2 Figures 22, 23 and 24** and summarised in **Table 18.**

5.3 Discussion Three temperature-sensitive mutants (*ts*1A, *ts*1B and *ts*1C) have been sequentially derived from the non-ts wild-type strain RSS-2 using different mutagenising agents. The three mutants have restrictive temperatures of 39°C, 38°C and 37°C respectively (McKay *et al.*, 1988; Pringle *et al.*, 1993). All three mutant viruses have RNA-positive phenotypes and synthesise viral polypeptides intracellularly at their non-permissive temperatures. The single (*ts*1A) mutant when administered intranasally in low dose to adult volunteers, had reduced disease-inducing potential in comparison to the wild-type virus RSS-2. The double (*ts*1B) and triple (*ts*1C) mutants, however, gave greater reduction in disease-inducing potential with approximately equivalent levels of attenuation (Watt *et al.*, 1988; Pringle *et al.*, 1993). All three mutants induced good immune responses, and mutant *ts*1C was considered to be the most suitable as a vaccine because of its potentially greater genetic stability.

Table 18 Summary of the location of nucleotide changes in the fusion gene of *ts*1C and intermediates *ts*1A and *ts*1B compared to the wild-type strain RSS-2

Nucleotide change		Residue change	*ts*1A	*ts*1B	*ts*1C
Amino acid #	Codon triplet				
35	**U**CA → **A**CA	Ser → Thr	NO	NO	YES
73	GA**U** → GA**C**	No change	NO	NO	YES
166	AA**G** → AA**A**	No change	YES	YES	YES
316	CU**A** → CU**G**	No change	NO	NO	YES
447	GU**G** → GU**U**	No change	NO	NO	YES
523	**A**CA → **A**UA	Thr → Ile	YES	YES	YES

N.B. Letters in bold and underlined denote the mutated base within the codon.
YES - denotes presence of mutation in the mutant defined in the top column.
NO - denotes absence of mutation in the mutant defined in the top column.
Coding changes are highlighted in shaded boxes

The sequence data for mutation site #1 shows no change in hydrophobicity in this region of the F_2 subunit and is not thought to alter the predicted structure of the fusion protein. It is therefore unlikely that the change in residue at position #35 has a role in determining temperature-sensitivity changes between *ts*1B and *ts*1C despite this region being highly conserved between human and bovine strains of RS viruses. Mutation site #3 (residue #523 Thr $\rightarrow$ Ile) is located immediately upstream of the putative membrane anchor region (residue #525 to #550). The substituted and substituting amino acids possess the same hydrophobic moment and charge suggesting that this residue change is conservative. However, because this change occurs in a region which is highly conserved between bovine and human strains of RS virus **(Section 3.2.5)** it is possible that it may be structurally important. The subsequent lowering of the restrictive temperature of mutant *ts*1B could be due to other mutations in the viral genome. The fact that both mutation sites occur in regions of high conservation, not just within human strains of RS virus but also in RS viruses of other hosts, and are also retained in all three mutants, suggests that the changes observed in mutants *ts*1A, *ts*1B and *ts*1C are not random errors occurring during the sequencing of their genes.

5.4 Conclusion The mutations detected in the fusion gene of *ts*1C are not located in known epitopes, therefore it is unlikely that the two coding changes observed in the fusion protein gene of the *ts*1C candidate vaccine have altered the antigenic properties of the virus, since the F protein is the principle inducer of neutralising antibody.

CHAPTER SIX

CHAPTER SIX

General discussion

6.1 Natural variation in the fusion protein gene of subgroup A RS viruses The data presented in this thesis together with published sequences show that the amino acid identity of the fusion protein genes of subgroup A strains is high (97% → 99.4%). Between subgroups there is a lower degree (90%) of amino acid identity, the amino acid identity between human and bovine RS viruses being only 80%. The F_2-subunit encoding region of the gene has most variation, specifically in the signal and C-terminal domains.

Each of the five Birmingham isolates used in this study was originally classed in a separate lineage (SHL 1 to SHL 5) based on the relatedness of their SH and N protein genes (Cane and Pringle, 1991). The distinctiveness of the five lineages was confirmed by sequencing of the G protein gene (Cane *et al.*, 1993). The relatedness of the fusion protein genes of the five isolates representing the five lineages SHL1 to SHL5 (**Figure 10**) showed that the pattern of relatedness of the F proteins approximated that of the lineages (**Figure 2**) as previously defined by SH, N and G protein gene analysis.

The degree of homology in the fusion protein gene of subgroup A RS viruses appears similar to that found in the fusion protein genes of clinical isolates of PIV 3 where the gene has 94.5% to 95.5% homology, and their predicted products have greater than 97% amino acid identity (Van Wyke Coelingh and Winter, 1990). Despite low amino acid difference, the clinical isolates of PIV 3 possess antigenic differences similar to those found in laboratory variants selected by and resistant to neutralising anti-F monoclonal antibodies (Van Wyke Coelingh and Winter, 1990). Therefore, by analogy to PIV 3, the amino acid variation within subgroup A RS viruses may have an effect on subgroup A virus antigenicity. Whilst no variation in amino acid identity in subgroup A viruses has been mapped to the four known antigenic areas, antigenic variation could be caused by

amino acid variation in the C-terminal domain of the F_2-subunit. By analogy with NDV, this is thought to interact with the H box region (**Figure 11**) near the F_1-subunit N-terminal domain. The area downstream of the H box region is highly antigenic (containing three of the four known antigenic areas) and so disruption of this interaction between subunits could possibly alter conformational epitopes in this region. Secondary structure alterations in the C-terminal region of the F_2-subunit has been suggested as a potential consequence of amino acid variation using computer analysis. The presence of antigenic variation between subgroup A isolates could be tested by assay against a panel of anti-F Mabs, by dot-blot, Western blot, or ELISA analysis. However, in any analysis of the contribution of F protein variability to antigenicity, it will be necessary to study the anti-F response of each isolate in the absence of other viral proteins. This would necessitate the cloning and expression of the fusion protein genes from isolates and analysing the immune response in an animal model.

In other paramyxoviruses, amino acid variation occurs at the cleavage site and affects infectivity; *e.g.* PIV 3 (Van Wyke Coelingh and Winters, 1990); NDV (Toyoda *et al.*, 1987; Glickman *et al.*, 1988); and Sendai virus (Choppin and Scheid, 1980). However only a single non-coding nucleotide change in the cleavage site of the subgroup A RS viruses examined in this study has been observed, so no direct affect on infectivity would be expected.

6.2 Induced variation in the fusion protein gene of subgroup A RS viruses Evidence has been presented to indicate that the site of the ts lesion in mutant *ts*A1 is located at residue #66 (Glu → Lys) and that of the mobility phenotype at residue #102 (Pro → Ser). Whilst the change at residue #102 occurs in the variable C-terminal region (residues #101 to #105), the putative ts-causing lesion occurs in the conserved (C #1) region of the subunit. This suggests that this domain is an important structural region.

Characterisation of the phenotype of mutant *ts*A1 identified the fusion protein gene as the likely site of the ts lesion (Belshe *et al.*, 1978; Caravokyri, 1990) and the sequence data derived in this study supports this hypothesis. However, coding changes occurring elsewhere in the genome of mutant *ts*A1 cannot be ruled out. Mutation in the matrix protein can result in altered fusion protein function in ts mutants of the paramyxovirus NDV (Peeples and Bratt, 1984). Consequently the matrix (M) protein gene of this mutant and both classes of revertant viruses were sequenced for at least one clone of each gene and no coding changes were observed (data not shown). This does not exclude the possibility, however, that lesions are present in other genes which may be responsible for the ts and attenuation phenotype. It is evident that the *ts*A1 virus is genetically complex and several lesions may contribute to the phenotypes. However, the evidence presented in this thesis favours the hypothesis that the ts and mobility mutations are located in the F protein gene.

Analysis of the nature of reversion in other paramyxoviruses, such as the group D ts mutants of NDV, has provided evidence of the importance of the amphipathic α-helix (AAH) in providing the structural requirements for correct F protein processing (Wang *et al.*, 1992). They found that the ts lesion was located in the AAH (also called the H box) region of the F_1-subunit. Two classes of revertant were obtained. One type had compensating mutations in and around the original lesion. The second type had coding changes localised within the F_2-subunit. These authors concluded that the F_1 AAH region interacted with the F_2-subunit, and that the original ts lesion and also the revertant mutations, resulted in changes in the secondary structure of the AAH region, possibly by disrupting or weakening the α-helix motif. The ts lesion of the *ts*A1 mutant of RS virus occurred in the C#1 region of the F_2-subunit and revertant mutations occurred in and around the original site of lesion. By analogy with the NDV model described above, the ts lesion may prevent correct interaction between the AAH (H box) and the F_2-subunit such that changes in the secondary structure are fully or partially corrected by subsequent changes in the revertant viruses.

Transient *in vitro* expression of the fusion gene products of mutant, revertant and wild-type viruses in mammalian cells has confirmed that the mobility phenotype is solely a consequence of changes in the fusion protein gene. The presence of the mobility phenotype supports the theory that changes in the amino acid sequence may lead to structural changes in the fusion protein. This may potentially lead to changes in the antigenicity of this protein. It would therefore be of future interest to examine the reactivity profiles of mutant and revertant fusion proteins against a panel of anti-F monoclonal antibodies.

The appearance of revertant viruses in the nasopharynx of vaccinated infants suggested an unacceptable level of genetic instability and work on the development of *ts*A1 as a vaccine was suspended. However, in recent years live virus vaccine for control of RS virus has gained favour again and mutant *ts*A1 is being re-evaluated by R.M.Chanock's group at the N.I.H. (C.R.Pringle, personal communication).

An alternative candidate vaccine restricted to growth at temperatures below 37°C, the triple mutant *ts*1C, has been sequentially derived from wild-type strain RSS-2 using different mutagenising agents in order to minimise genetic instability. Two coding changes in the fusion protein gene were observed between the wild-type and temperature-sensitive virus, residue #35 (Ser → Thr) and residue #523 (Thr → Ile) Both changes are conservative in nature and are not thought to effect the protein structure. Neither mutation is present in or near known epitopes.

The *ts*1C candidate live vaccine exhibites residual virulence in adults and is thought unsuitable for administration to sero-negative children without further modification. It is being used, however, to identify the genetic determinants of virulence. This entails the complete sequencing of the genome of the triple ts mutant, the single and double intermediate ts mutants, and their progenitor wild-type virus in order to identify the mutational lesions associated with both the ts and virulence phenotype. Sequencing of the

*ts*1C genome is virtually complete and analysis of the data indicates that mutations in one domain of the polymerase protein gene (L) are major determinants of loss of virulance, at least for URT infection in adults (C.R.Pringle, personal communication). The work reported in this thesis indicates that mutations in the fusion protein gene are probably not a factor in the attenuation of the virus. It has also established that the antigenicity of the fusion protein is unlikely to have been altered by the attenuation process and that there is little or no risk in inducing the paradoxical disease associated with the use of formalin-inactivated vaccine in the past.

CHAPTER SEVEN

CHAPTER SEVEN

References

Akusjarvi, G. (1993): Proteins with transcription regulatory properties encoded by human adenoviruses. *Trends in Microbiology*. **1**(5), 163-170.

Akerlind, B.; Nörrby, E.; Orvell, C.; Mufson, M.A. (1988): Respiratory syncytial virus : heterogeneity of subgroup B strains. *J. Gen. Virol.* **69**, 2145-2154.

Alansari, H. and Potgieter, L.N.D. (1993): Nucleotide sequence analysis of the ovine respiratory syncytial virus G glycoprotein gene. *Virology*. **196**, 873-877.

Alansari, H. and Potgieter, L.N.D. (1994): Nucleotide and predicted amino acid sequence analysis of the ovine respiratory syncytial virus non-structural 1C and 1B genes and the small hydrophobic protein gene. *J. Gen. Virol.* **75**, 401-404.

Ananaba, G.A. and Anderson, L.J. (1991): Antibody enhancement of respiratory syncytial virus stimulation of leukotriene production by a macrophage-like cell line. *J. Virol.* **65**(9), 5052-5060.

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

Anderson, L.J.; Hendry, R.M.; Pierik, L.T.; Tsou, C.; McIntosh, K. (1991): Multicenter studies of strains of respiratory syncytial virus. *J. Inf. Dis.* **163**, 687-692.

Anderson, K.; Stott, E.J.; Wertz, G.W. (1992): Intracellular processing of the human respiratory syncytial virus fusion glycoprotein: Amino acid substitutions affecting folding, transport and cleavage. *J. Gen. Virol.* **73**(5), 1177-1188.

Arbiza, J.; Taylor, G.; Lopez, J.A.; Furze, J.; Wyld, S.; Whyte, P.; Stott, E.J.; Wertz, G.; Sullender, W. (1992): Characterisation of two antigenic sites recognised by neutralising monoclonal antibodies directed against the fusion glycoprotein of human respiratory syncytial virus. *J. Gen. Virol.* **73**(9), 2225-2234.

Arumugham, R.G.; Hildreth, S.W.; Paradiso, P.R. (1989): Interprotein disulfide bonding between F and G glycoproteins of human respiratory syncytial virus. *Arch. Virol.* 105, 65-69.

Baker, S.C. and Moyer, S.A. (1988): Encapsidation of Sendai virus genome RNAs by purified NP protein during in vitro replication. *J. Virol.* **62**, 834-838.

Banerjee, A.K. (1987a): Transcription and replication of Rhabdoviruses. *Microbiological Reviews* **51**, 66-87.

Banerjee, A.K. (1987b): The transcription complex of Vesicular stomatitis virus. *Cell* **48**, 363-364.

Banerjee, A.K. and Barik, S. (1992): Minireview: Gene expression of vesicular stomatitis virus genome RNA. *Virology* **188**, 417-428.

Bangham, C.R.M. and Askonas, B.A. (1986): Murine cytotoxic T-cells specific to respiratory syncytial virus recognise different antigenic sub-types of the virus. *J. Gen. Virol.* **67**, 623-629.

Barge, A.; Gaudin, Y.; Coulon, P.; Ruigrok, R.W.H. (1993): Vesicular stomatitis virus M protein may be inside the ribonucleocapsid coil. *J. Virol.* **67**(12), 7246-7253.

Barik, S. and Banerjee, A.K. (1991): Cloning and expression of the vesicular stomatitis virus phosphoprotein gene in Escherichia coli: Analysis of phosphorylation status versus transcriptional activity. *J. Virol.* **65**, 1719-1726.

Barik, S. (1992): Transcription of human respiratory syncytial virus genome RNA in vitro: Requirement of cellular factor(s). *J. Virol.* **66**(11), 6813-6818.

Barik, S. and Banerjee, A.K. (1992): Sequential phosphorylation of the phosphoprotein of vesicular stomatitis virus by cellular and viral protein kinases is essential for transcription activation. *J. Virol.* **66**(2), 1109-1118.

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

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**, 667-685.

Barr, J (1994): Expression of the nucleoprotein and phosphoprotein genes of Pneumonia Virus of Mice, and specific interactions of the gene products. PhD Thesis, University of Warwick.

Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K. (1987): The nucleotide sequence of the gene encoding the F protein of Canine distemper virus: A comparison of the deduced amino acid sequence with other paramyxoviruses. *Virus Res.* **8**, 373-386.

Baybutt, H.N. and Pringle, C.R. (1987): Molecular cloning and sequencing of the F and 22K membrane protein genes of the RSS-2 strain of respiratory syncytial virus. *J. Gen. Virol.* **68**, 2789-2796.

Beckes, J.D. and Perrault, J. (1991). Two distinct protein kinase activities in vesicular stomatis virions phosphorylates the NS transcription factor. *Virology* **184**, 383-386.

Beeler, J.A. and van Wyke Coelingh, K. (1989): Neutralisation epitopes of the F glycoprotein of respiratory syncytial virus: Effect of mutation upon fusion function. *J. Virol.* **63**, 2941-2950.

Beem, M. (1967): Repeated infections with respiratory syncytial virus. *J. Immunol.* **98**, 1115-1122.

Bellini, W.J.; Englund, G.; Richardson, C.D.; Rozenblatt, S.; Lazzarini, R.A. (1986): Matrix genes of Measles virus and Canine distemper virus: Cloning, nucleotide sequences, and deduced amino acid sequences. *J. Virol.* **58**, 408-416.

Belshe, R.B.; Richardson, L.S.; Schnitzer, T.J.; Prevar, D.A.; Camargo, E.; Chanock, R.M. (1977): Further characterisation of the complementation group B temperature-sensitive mutant of respiratory syncytial virus. *J. Virol.* **24**, 8-12.

Belshe, R.B.; Richardson, L.S.; Prevar, D.A.; Camargo, E.; Chanock, R.M. (1978): Growth and genetic stability of 4 temperature sensitive (ts) mutants of respiratory syncytial (RS) virus in newborn ferrets. *Arch. Virol.* **58**, 313-321.

Belshe, R.B.; Anderson, E.L.; Walsh, E.E. (1993): Immunogenicity of purified F glycoprotein of respiratory syncytial virus: Clinical and immune responses to subsequent natural infection in children. *J. Inf. Dis.* **168**. 1024-1029.

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). *Archive fur die gesamte virusforschung* **45**, 39-51.

Birnboim, H.C.; Doly, J. (1979): A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids. Res.* **7**, 1513-1523.

Black, B.L. and Lyles, D.S. (1992): Vesicular stomatitis virus matrix protein inhibits host cell-directed transcription of target genes in vivo. *J. Virol.* **66**, 4058-4064.

Black, B.L.; Rhodes, R.B.; Mckenzie, M.; Lyles, D.S. (1993): The role of vesicular stomatitis virus matrix protein in inhibition of host-directed gene expression is genetically separable from its function in virus assembly. *J. Virol.* **67**, 4814-4821.

Blixenkrone-Moller, M.; Sharma, B.; Varsanyi, T.M.; Hu, A.; Norrby, E.; Kovamees, J. (1992): Sequence analysis of the genes encoding the nucleocapsid protein and phosphoprotein (P) of phocid distemper virus, and editing of the P gene transcript. *J. Gen. Virol.* **73**(4), 885-893.

Blumberg, B.M.; Leppert, M.; Kolakofsky, D. (1981): Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell* **23**, 837-845.

Blumberg, B.M.; Giorgi, C.; Kolakofsky, D. (1983): N protein of Vesicular stomatitis virus selectively encapsidates leader RNA in vitro. *Cell* **32**, 559-567.

Blumberg, B.M.; Crowley, J.C.; Silverman, J.I.; Menonna, J.; Cook, S.D.; Dowling, P.C. (1988): Measles virus L protein evidences elements of ancestral RNA polymerase. *Virology* **164**, 487-497.

Blumberg, B.M.; Chan, J.; Udem, S.A. (1991): Function of 3' and 5' end sequences: In theory and practice. In: The Paramyxoviruses. (Ed: Kingsbury, DW) Plenum Press, New York and London, 235-247.

Bohn, W.; Rutter, G.; Hohenberg, H.; Mannweiler, K.; Nobis, P. (1986): Involvement of actin filaments in budding of measles virus: Studies on cytoskeletons in infected cells. *Virology* **149**, 91-96.

Bourgeois, C.; Corvaisier, C.; Bour, J.B.; Kohli, E.; Pothier, P. (1991): Use of synthetic peptides to locate neutralising antigenic domains on the fusion protein of respiratory syncytial virus. *J. Gen. Virol.* **72**, 1051-1058.

Bowen, H.A. and Lyles, D.S. (1982): Kinetics of incorporation of Sendai virus proteins into host plasma membranes and virions. *Virology* **121**, 1-11.

Breese-Hall, C.; Walsh, E.E.; Schnabel, K.C.; Long, C.E.; McConnochie, K.M.; Hildreth, S.W.; Anderson, L.J. (1990): Occurrence of groups A and B of respiratory syncytial virus over 15 years: associated epidemiologic and clinical characteristics in hospitalised and ambulatory children. *J. Infect. Dis.* **162**, 1283-1290.

Brideau, R.J.; Walters, R.R.; Stier, M.A.; Wathen, M.W. (1989): Protection of cotton rats against human respiratory syncytial virus by vaccination with a novel chimeric FG glycoprotein. *J. Gen. Virol.* **70**, 2637-2644.

Buchholz, C.J.; Spehner, D.; Drillien, R.; Neubert, W.; Homman, H.E. (1993): The conserved N-terminal region of Sendai virus nucleocapsid protein NP is required for nucleocapsid assembly. *J. Virol.* **67**, 5803-5812.

Buckland, R. and Wild, F. (1989): Leucine zipper motif extends. *Nature* **338**, 547.

Buckland, R.; Malvoisin, E.; Beauverger, P.; Wild, F. (1992): A leucine zipper structure present in the measles virus fusion protein is not required for its tetramerization but is essential for fusion. *J. Gen. Virol.* **73**(7), 1703-1707.

Buetti, E. and Choppin, P.W. (1977): The transcriptase complex of the paramyxovirus SV5. *Virology* **82**, 493-508.

Calain, P. and Roux, L. (1993): The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J. Virol.* **67**(8), 4822-4830.

Caldwell, S.E. and Lyles, D.S. (1986): Dissociation of newly synthesized Sendai viral proteins from the cytoplasmic surface of isolated plasma membranes of infected cells. *J. Virol.* **57**, 678-683.

Cane, P.A. and Pringle, C.R. (1991): Respiratory syncytial virus heterogeneity during an epidemic: analysis by limited sequencing (SH gene) and restriction mapping (N gene). *J. Gen. Virol.* **72**, 349-357.

Cane, P.A.; Matthews, D.A.; Pringle, C.R. (1991): Identification of the variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *J. Gen. Virol.* **72**, 2091-2096.

Cane, P.A.; Matthews, D.A.; Pringle, C.R. (1992): Analysis of relatedness of subgroup A respiratory syncytial viruses isolated worldwide. *Virus Res.* **25**(1-2), 15-22.

Cane, P.A. and Pringle, C.R. (1992): Molecular epidemiology of respiratory syncytial virus: Rapid identification of subgroup A lineages. *J. Vir. Methods* **40**(3), 297-306.

Cane, P.A.; Matthews, D.A.; Pringle, C.R. (1993): Frequent polymerase errors observed in a restricted area of clones derived from the attachment (G) protein gene of respiratory syncytial virus. *J. Virol.* **67**, 1090-1093.

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.

Cannon, M.J. and Bangham, C.R.M. (1989): Recognition of respiratory syncytial virus fusion protein by mouse cytotoxic T-cell clones and a human cytotoxic T cell line. *J. Gen. Virol.* **70**, 79-87.

Caravokyri,C; (1990): Characterisation of temperature-sensitive mutants of human respiratory syncytial (RS) virus. PhD Thesis. University of Warwick. Coventry. U.K.

Caravokyri, C. and Pringle, C.R. (1991): Defective synthesis of envelope proteins by temperature-sensitive mutants representing complementation groups B and D of respiratory syncytial virus. *J. Gen. Virol.* **72**(10), 2501-2508.

Caravokyri, C. and 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**(1), 53-62.

Caravokyri, C.; Zajac, A.J.; Pringle, C.R. (1992): Assignment of mutant tsN19 (complementation group E) of respiratory syncytial virus to the P protein gene. *J. Gen. Virol.* **73**(4), 865-873.

Cash, P.; Wunner, W.H.; Pringle, C.R. (1977): A comparison of the polypeptides of human and bovine respiratory syncytial virus and murine pneumonia virus. *Virology* **82**, 369-379.

Cash, P.; Pringle, C.R.; Preston, C.M. (1979): The polypeptides of human respiratory syncytial virus: products of cell-free protein synthesis and post translational modifications. *Virology* **92**, 375-384.

Castadenda, S.J. and Wong, T.C. (1989): Measles virus synthesises both leaderless and leader containing polyadenylated RNAs in vitro. *J. Virol.* **63**, 2977-2986.

Cattaneo, R.; Kaelin, K.; Baczko, K.; Billeter, M.A. (1989): Measles virus editing provides an additional cysteine-rich protein. *Cell* **56**, 759-764.

Cavanagh, D. and Barrett, T. (1988): Pneumovirus-like characteristics of the mRNA and proteins of turkey rhinotracheitis virus. *Virus Res.* **11**, 241-256.

Chambers, P.; Millar, N.S.; Emmerson, P.T. (1986): Nucleotide sequence of the gene encoding the fusion glycoprotein of Newcastle disease virus. *J. Gen. Virol.* **67**, 2685-2694.

Chambers, P.; Matthews, D.A.; Pringle, C.R.; Easton, A.J. (1990a): 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.

Chambers, P.; Pringle, C.R.; Easton, A.J. (1990b): Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *J. Gen. Virol.* **71**, 3075-3080.

Chambers, P.; Pringle, C.R.; Easton, A.J. (1991): Genes 1 and 2 of pneumonia virus of mice encode proteins which have little homology with the 1C and 1B proteins of human respiratory syncytial virus. *J. Gen. Virol.* **72**(10), 2545-2550.

Chambers, P.; Pringle, C.R.; Easton, A.J. (1992): Sequence analysis of the gene encoding the fusion glycoprotein of pneumonia virus of mice suggests possible conserved secondary structure elements in paramyxovirus fusion glyoproteins. *J. Gen. Virol.* **73**, 1717-1724.

Chan, J.; Pifko-Hirst, S.; Udem, S.A. (1989): Where is the measles virus (+) strand leader RNA transcript? In: Genetics and Pathogenicity of negative strand viruses. (Eds: Mahy,B; Kolakofsky,D) Elsevier Biomedical Press, Amsterdam, 221-231.

Chanock, R.; Riozman, B.; Myers, R. (1957): Recovery from infants with respiratory illness of a virus related to Chimpanzee coryza agent (CCA): I. isolation, properties, and characterisation. *Am. J. Hyg.* **66**, 281-290.

Chattopadhyay, D. and Banerjee, A.K. (1987a): Phosphorylation within a specific domain of the phosphoprotein of Vesicular stomatitis virus regulates transcription in vitro. *Cell* **49**, 407-414.

Chattopadhyay, D. and Banerjee, A.K. (1987b): Two separate domains within vesicular stomatitis virus phosphoprotein support transcription when added in trans. *Proc. Natl. Acad. Sci. USA.* **84**, 8932-3936.

Chen, C. and Okayama, H. (1987): High-efficiency transformation of mammalian cells by plasmid DNA. *Molecular and Cellular Biology* 7, 2745-2752.

Cherrie, A.H.; Anderson, K.; Wertz, G.W.; Openshaw, P.J.M. (1992): Human cytotoxic T-cells stimulated by antigen on dendritic cells recognize the N, SH, F, M, 22K, and 1B proteins of respiratory syncytial virus. *J.Virol.* **66**(4), 2102-2110.

Chin, J.; Magoffin, R.L.; Shearer, L.A.; Schieble, J.H.; Lennette, E.H. (1969): Field evaluation of a respiratory syncytial virus vaccine and a trivalent parainfluenza virus vaccine in a pediatric population. *Am. J. Epidemiol.* **89**, 449-463.

Choppin, P.W. and Scheid, A. (1980): The role of viral glycoproteins in adsorption, penetration and pathogenicity of viruses. *Reviews of Infectious Diseases* **2**, 40-61.

Chou, P.V. and Fasman, G.D. (1978): Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology* **47**, 45-147.

Cohen, S.N.; Chang, A.C.Y.; Hsu, L. (1972): Nonchromosomal antibiotic resistance in bacteria : Genetic trsnaformation of Escherichia coli by R-factor DNA. *Proc. Natl. Acad. Sci. USA.* **69**, 2110-2114.

Collins, P.L.; Huang, Y.T.; Wertz, G.W. (1984a): Identification of a tenth mRNA of respiratory syncytial virus and assignment of polypeptides to the 10 viral genes. *J.Virol.* **49**, 572-578.

Collins, P.L.; Huang, Y.T.; Wertz, G.W. (1984b): Nucleotide sequence of the gene encoding the fusion F glycoprotein of human respiratory syncytial virus. *Proc. Natl. Acad. Sci. USA.* **81**, 7683-7687.

Collins, P.L.; Anderson, K.; Langer, S.J.; Wertz, G.W. (1985a): Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus. *Virology* **146**, 69-77.

Collins, P.L. and Wertz, G.W. (1985b): The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript. *Virology* **141**, 283-291.

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

Collins, P.L.; Dickens, L.E.; Buckler-White, A.; Olmsted, R.A.; Spriggs, M.K.; Camargo, E.; van Wyke Coelingh, K. (1986): Nucleotide sequences from the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. *Proc. Natl. Acad. Sci. USA*. **83**, 4594-4598.

Collins, P.L. (1991): The molecular biology of human respiratory syncytial virus (RSV) of the genus pneumovirus. In: The Paramyxoviruses. (Ed: Kingsbury,DW) Plenum Press, New York and London, 103-162.

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. USA*. **88**(21), 9663-9667.

Collins, P.L. and Mottet, G. (1991): Post-translational processing and oligomerization of the fusion glycoprotein of human respiratory syncytial virus. *J. Gen. Virol.* **72**(12), 3095-3102.

Collins, P.L. and 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**(4), 849-863.

Collins, P.L.; Mink, M.A.; Hill, M.G.; Camargo, E.; Grosfeld, H.; Stec, D.S. (1993): Rescue of a 7502-nucleotide (49.3% of full-length) synthetic analog of respiratory syncytial virus genomic RNA. *Virology* **195**, 252-256.

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

Colonno, R.J. and Stone, H.O. (1975): Methylation of messanger RNA of Newcastle disease virus *in vitro* by a virion-associated enzyme. *Pro. Natl. Acad. Sci. USA*. **72**, 2611-2615.

Colonno, R.J. and Stone, H.O. (1976): Newcastle disease virus mRNA lacks 2'-O-methylated nucleotides. *Nature* **261**, 611-614.

Connors, M.; Kulkarni, A.B.; Collins, P.L.; Firestone, C.Y.; Holmes, K.L.; Morse, H.C.; Murphy, B.R. (1992): Resistance to respiratory syncytial virus (RSV) challenge induced by infection with a vaccinia virus recombinant expressing the RSV M2 protein (Vac-M2) is mediated by CD8-positive T cells, while that induced by Vac-F or Vac-G recombinants is mediated by antibodies. *J. Virol.* **66**(2), 1277-1281.

Curran, J.; Boeck, R.; Kolakofsky, D. (1991a): The Sendai virus P gene expresses both an essential protein and an inhibitor of RNA synthesis by shuffling modules via mRNA editing. *EMBO Journal* **10**, 3079-3085.

Curran, J.; De Melo, M.; Moyer, S.; Kolakofsky, D. (1991b): Charaterization of the Sendai virus V protein with an anti-peptide antiserum. *Virology* **184**, 108-116.

Curran, J.; Marq, J.B.; Kolakofsky, D. (1992): The Sendai virus nonstructural C proteins specifically inhibit viral mRNA synthesis. *Virology* **189**(2), 647-656.

De, B.P.; Thornton, G.B.; Luk, D.; Banerjee, A.K. (1982): Purified matrix protein of vesicular stomatitis virus blocks viral transcription in vitro. *Proc. Natl. Acad. Sci. USA.* **79**, 7137-7141.

De, B.P. and Banerjee, A.K. (1985): Requirements and functions of vesicular stomatitis virus L and NS proteins in the transcription process in vitro. *Biochem. Biophys. Res. Commun.* **126**, 40-49.

De, B.P.; Lesoon, A.; Banerjee, A.K. (1991): Human parainfluenza virus type 3 transcription in vitro : Role of cellular actin in mRNA synthesis. *J. Virol.* **65**, 3268-3275.

Devereux, J.; Haeberli, P.; Smithies, O. (1984): A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids. Res.* **12**, 387-395.

Dubovi, E.J. (1982): Analysis of proteins synthesised in respiratory syncytial virus infected cells. *J. Virol.* **42**, 372-278.

Eckhardt, A.E.; Timpte, C.S.; Abernethy, J.L.; Toumadje, A.; Johnson, W.C.; Hill, R.L. (1987): Structural properties of porcine submaxillary gland apomucin. *J. Biol. Chem.* **262**, 11339-11344.

Elango, N. and Venkatesan, S. (1983): Amino acid sequence of human respiratory syncytial virus nucleocapsid protein. *Nucleic Acids Res.* **11**, 5941-5951.

Elango, N.; Satake, M.; Coligan, J.E.; Norrby, E.; Camargo, E.; Venkatesan, S. (1985a): Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of F1 subunit. *Nucl. Acids. Res.* **13**, 1559-1574.

Elango, N.; Satake, M.; Venkatesan, S. (1985b): mRNA sequence of three respiratory syncytial virus genes encoding two nonstructural proteins and a 22K structural protein. *J. Virol.* **55**, 101-110.

Elder,J.H.; McGee, J.S.; Alexander,S. (1986): Carbohydrate side chains of Rauscher Leukemia Virus envelope glycoproteins are not required to elicit a neutralizing antibody response. *J. Virol.* **57**(1), 340-342.

Emerson,S.U. (1985): Rhabdoviruses, in "Virology," (B.N.Fields, ed.), pp1119-1132, Raven Press, New York.

Ewasyshyn, M.;Bonneau, A.M.; Usman, S.; Scollard, N.; Klein, M. (1993): Comparative analysis of the immunoprotective abilities of glycosylated and deglycosylated parainfluenza virus type 3 surface glycoproteins. *J. Gen. Virol.* **74**, 2781-2785.

Faaberg, K.S. and Peeples, M.E. (1988): Association of soluble matrix protein of Newcastle disease virus with liposomes is independent of ionic conditions. *Virology* **166**, 123-132.

Faulkner, G.P.; Follett, E.A.C.; Shirodaria, P.V.; Pringle, C.R. (1976): Respiratory syncytial virus *ts* mutants and nuclear immunofluorescence. *J. Virol.* **20**, 487-500.

Fauquet, C; Murphy, F.A.; Bishop, D.H.L.; Ghabrial, S.A.; Martelli, G.P. (1994): Sixth report of the International Committee on the Taxonomy of Viruses. (In Press).

Fernie, B.F. and Gerin, J.L. (1982): Immunochemical identification of viral and non viral proteins of the respiratory syncytial virus virion. *Infect. Immun.* **37**, 243-249.

Fernie, B.F.; Dapolito, G.; Cote, P.J.; Gerin, J.L. (1985): Kinetics of synthesis of RS virus glycoproteins. *J. Gen. Virol.* **66**, 1983-1990.

Fiat, A.M.; Jolles, J.; Aubert, J.P.; Loucheux-Lefebvre, M.H.; Jolles, P. (1980): Localisation and importance of the sugar part of human casein. *Eur.J.Biochem.* **111**, 333-339.

Follett, E.A.C.; Pringle, C.R.; Pennington, T.H. (1975): Virus development in enucleate cells : echovirus, poliovirus, pseudorabies virus, reovirus, respiratory syncytial virus and Semliki Forest virus. *J. Gen. Virol.* **26**, 183-196.

Fooks, A.R.; Stephenson, J.R.; Warnes, A.; Dowsett, A.B.; Rima, B.K.; Wilkinson, G.W.G. (1993): Measles virus nucleocapsid protein expressed in insect cells assembles into nucleocapsid-like structures. *J. Gen. Virol.* **74**, 1439-1444.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L. and Brown, F. (1991) Classification and nomenclature of viruses.

Galinski, M.S.; Mink, M.A.; Pons, M.W. (1988): Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene encoding the L protein. *Virology* **165**, 499-510.

Galinski, M.S. and Wechsler,S.L. (1991): The molecular biology of the *Paramyxovirus* genus. In: The Paramyxoviruses. (Ed: Kingsbury, D.W.) Plenum Press, New York and London, 41-82.

Galinski, M.S.; Troy, R.M.; Banerjee, A.K. (1992): RNA editing in the phosphoprotein gene of the human parainfluenza virus type 3. *Virology* **186**(2), 543-550.

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.

Garcia-Barreno, B.; Jorcano, J.L.; Aukenbauer, T.; Lopez-Galindez, C.; Melero, J.A. (1988): Participation of cytoskeleton intermediate filaments in the infectious cycle of human respiratory syncytial virus (RSV). *Virus Res.* **9**, 307-322.

Garcia-Barreno, B.; Palomo, C.; Penas, C.; Delgado, C.; Perez-Brena, P.; Melero, J.A. (1989): Marked differences in the antigenic structure of human respiratory syncytial virus F and G glycoproteins. *J. Virol.* **63**, 925-932.

Garcia-Barreno, B.; Portela, A.; Delgado, T.; Lopez, J. (1990): Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *EMBO Journal* **12**, 4181-4187.

Garnier, J., Ogusthorpe, D. J. and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *Journal of Molecular Biology*. **120**, 97-120.

Gharpure, M.A.; Wright, P.F.; Chanock, R.M. (1969): Temperature-sensitive mutants of respiratory syncytial virus. *J. Virol.* **3**, 414-421.

Gill, D.S.; Chattopadhyay, D.; Banerjee, A.K. (1986): Identification of a domain within the phosphoprotein of vesicular stomatitis virus that is essential for transcription *in vitro*. *Proc. Natl. Acad. Sci. USA*. **83**, 8873-8877.

Gimenez, H.B. and Pringle, C.R. (1978): Seven complementation groups of respiratory syncytial virus temperature-sensitive mutants. *J. Virol.* **27**, 459-464.

Gimenez, H.B.; Cash, P.; Melvin, W.T. (1984): Monoclonal antibodies to human respiratory syncytial virus and their use in comparison of different isolates. *J. Gen. Virol.* **65**, 963-971.

Gimenez, H.B.; Hardman, N.; Keir, H.M.; Cash, P. (1986): Antigenic variation between respiratory syncytial virus isolates. *J. Gen. Virol.* **67**, 863-870.

Giorgi, C.; Blumberg, B.M.; Kolalofsky, D. (1983): Sendai virus contains overlapping genes expressed from a single mRNA. *Cell* **35**, 829-836.

Glickman, R.L.; Syddall, R.J.; Iorio, R.M.; Sheeman, J.P.; Bratt, M.A. (1988): Quantitative basic residue requirements in the cleavage-activation site of the fusion glycoprotein as a determinant of virulence for Newcastle disease virus. *J. Virol.* **62**, 354-356.

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

Gruber, C. and Levine, S. (1983): Respiratory syncytial virus polypeptides. III. The envelope-associated proteins. *J. Gen. Virol.* **64**, 825-832.

Gruber, C. and Levine, S. (1985a): Respiratory syncytial virus polypeptides. IV. The oligosaccharides of the glycoproteins. *J. Gen. Virol.* **66**, 417-432.

Gruber, C. and Levine, S. (1985b): Respiratory syncytial virus polypeptides. V. The kinetics of glycoprotein synthesis. *J. Gen. Virol.* **66**, 1241-1247.

Hall, C.B. (1980): Prevention of infections with respiratory syncytial virus : the hopes and hurdles ahead. *Reviews of Infectious Diseases* **2**, 384-392.

Hamaguchi, M.; Yoshida, T.; Nishikawa, K.; Naruse, H.; Nagai, Y. (1983): Transcriptive complex of Newcastle disease virus. I. Both L and P proteins are required to constitute an active complex. *Virology* **128**, 105-117.

Hamaguchi, M.; Nishikawa, K.; Toyoda, T.; Yoshida, T.; Hanaichi, T.; Nagai, Y. (1985): Transcriptive complex of Newcastle disease virus II. Structural and functional assembly associated with the cytoskeletal framework. *Virology* **147**, 295-308.

Hammond, D.C.; Haley, B.E.; LeSnaw, J.A. (1992): Identification and charaterization of serine/threonine protein kinase activity intinsic to the L protein of vesicular stomatist virus New Jersey. *J. Gen. Virol.* **73**, 67-75.

Heminway, B.R.; Yu, Y.; Tanaka, Y.; Perrine, K.G.; Gustafson, E.; Bernstein, J.M.; Galinski, M.S. (1994): Analysis of Respiratory syncytial virus F, G and SH proteins in cell fusion. *Virology* **200**, 801-805.

Henderson, F.W.; Collier, A.M.; Clyde, W.A.; Denny, F.W. (1979): Respiratory syncytial virus infections: reinfections and immunity: a prospective, longitudinal study in young children. *N.Engl.J.Med.* **300**, 530-534.

Hendricks, D.A.; Baradaran, K.; McIntosh, K.; Patterson, J.L. (1987): Appearance of a soluble form of the G protein of respiratory syncytial virus in fluids of infected cells. *J. Gen. Virol.* **68**, 1705-1714.

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.

Hendricks, D.A.; Ono, E.; Seyer, J.M.; Gupta, K.C. (1993): Phosphorylation of the Sendai virus C proteins. *Virology* **197**, 471-474.

Hendry, R.M.; Talis, A.L.; Godfrey, E.; Anderson, L.J.; Fernie, B.F.; McIntosh, K. (1986): Concurrent circulation of antigenically distinct strains of respiratory syncytial virus during community outbreaks. *J.Infect.Dis.* **153**, 291-297.

Hendry, R.M.; Burns, J.C.; Walsh, E.E.; Graham, B.S.; Wright, P.F.; Hemming, V.G.; Rodriguez, W.J.; Kim, H.W.; Prince, G.A.; McIntosh, K.; Chanock, R.M.; Murphy, B.R. (1988a): Strain-specific serum antibody responses in infants undergoing primary infection with respiratory syncytial virus. *J. Infect. Dis.* **157**(4), 640-647.

Hendry, R.M.; Talis, A.L.; Godfrey, E.; Anderson, L.J.; Hercyk, N.; Horikami, S.M.; Moyer, S.A. (1988b): The vesicular stomatitis virus L protein possesses the mRNA methyltransferase activities. *Virology* **163**, 222-225.

Hendry, R.M.; Pierik, L.T.; McIntosh, K. (1989): Prevalence of respiratory syncytial virus subgroups over six consecutive outbreaks: 1981-1987. *J. Infect. Dis.* **160**(2), 185-190.

Hercyk, N.; Horikami, S.M.; Moyer, S.A. (1988): The vesicular stomatitis virus L protein possesses the mRNA methyltransferase activities. *Virology* **163**, 222-225.

Hiebert, S.W.; Paterson, R.G.; Lamb, R.A. (1985a): Hemagglutinin-neuraminidase protein of the paramyxovirus simian virus 5 : nucleotide sequence of the mRNA predicts an N-terminal membrane anchor. *J.Virol.* **54**, 1-6.

Hiebert, S.W.; Paterson, R.G.; Lamb, R.A. (1985b): Identification and predicted sequence of a previously unrecognized small hydrophobic protein, SH, of the paramyxovirus simian virus 5. *J.Virol.* **55**, 744-751.

Hiebert, S.W.; Richardson, C.D.; Lamb, R.A. (1988): Cell surface expression and orientation in membranes of the 44-amino-acid SH protein in simian virus 5. *J.Virol.* **62**, 2347-2357.

Higgins, D.G. and Sharp, P.M. (1988): CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*. **73**, 237-244.

Hill, D.; Schwyzer, M.; Steinman, H.M.; Hill, R.L. (1977): Ovine submaxillary mucin. *J.Biol.Chem*. **252**, 3799-3804.

Himes, S.R. and Gershwin, L.J. (1992): Bovine respiratory syncytial virus fusion protein gene: Sequence analysis of cDNA and expression using a baculovirus vector. *J. Gen. Virol.* **73**(6), 1563-1567.

Hodes, D.S.; Kim, H.W.; Parrott, R.H.; Camargo, E.; Chanock, R.M. (1974): Genetic alteration in a temperature-sensitive mutant of respiratory syncytial virus after replication in vivo. *Proc. Soc. Exp. Biol. Med.* **145**, 1158-1164.

Homann, H.E.; Willenbrink, W.; Buchholz, C.J.; Neubert, W.J. (1991): Sendai virus protein-protein interactions studied by a protein-blotting protein-overlay technique: Mapping of domains on NP protein required for binding to P protein. *J.Virol.* **65**, 1304-1309.

Hopp, T.J. and Woods, K.L. (1981): Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA*. **78**, 3824-3828.

Horikami, S.M. and Moyer, S.A. (1991): synthesis of leader RNA and editing of the P mRNA during transcription by purified measles virus. *J.Virol.* **65**, 5342-5347.

Horikami, S.M.; Curran, J.; Kolakofsky, D.; Moyer, S.A. (1992): Complexes of Sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vitro. *J.Virol.* **66**(8), 4901-4908.

Horsfall, F.L. and Hahn, R.G. (1940): A latent virus in normal mice capable of producing pneumonia in its natural host. *Journal of Experimental Medicine* **71**, 391-408.

Horsfall, F.L. and Curnen, E.C. (1946): Studies on pneumonia virus of mice (PVM). II. Immunological evidence of latent infection with the virus in numerous mammalian species. *Journal of Experimental Medicine* **83**, 43-64.

Horvath, C.M. and Lamb, R.A. (1992): Studies on the fusion peptide of a paramyxovirus fusion glycoprotein: Roles of conserved residues in cell fusion. *J.Virol.* **66**(4), 2443-2455.

Howard, M. and Wertz, G.W. (1989): Vesicular stomatitis virus RNA replication : A role for the NS protein. *J. Gen. Virol.* **70**, 2683-2694.

Hsu, M.C.; Scheid, A.; Choppin, P.W. (1981): Activation of the Sendai virus fusion protein (F) involves a conformational change with exposure of a new hydrophobic region. *J. Biol. Chem* **256**, 3557-3563.

Huang, Y.T. and 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.

Huang, Y.T. (1983): Respiratory syncytial virus mRNA coding assignments. *J.Virol.* **46**, 667-672.

Huang, Y.T.; Collins, P.L.; Wertz, G.W. (1984): Identification of a new envelope-associated protein of human respiratory syncytial virus. In, "Nonsegmented negative strand viruses, Paramyxoviruses and Rhabdoviruses." (Eds: Bishop,DHL; Compans,RW) Acedemic Press, New York, 365-368.

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.; Romito, R.R.; De, B.P.; Banerjee, A.K. (1993): Characterisation of the *in vitro* system for the synthesis of mRNA from human respiratory syncytial virus. *Virology* **193**, 862-867.

Huber, M.; Cattaneo, R.; Spielhofer, P.; Örvell, C.; Norrby, E.; Messerli, M.; Perriard, J.C.; Billeter, M.A. (1991): Measles virus phosphoprotein retains the nucleocapsid protein in the cytoplasm. *Virology* **185**(1), 299-308.

Hunt, D.M.; Smith, E.P.; Buckley, D.W. (1984): Aberrant polyadenylation by a vesicular stomatitis virus mutant is due to an altered L protein. *J. Virol.* **52**, 515-521.

Iwata, S.; Schmidt, A.C.; Titani, K.; Suzuki, M.; Kido, H.; Gotoh, B.; Hamaguchi, M.; Nagai, Y. (1994): Assignment of disulphide bridges in the fusion glycoprotein of Sendai virus. *J. Virol.* **68**(5), 3200-3206.

Johnson, P.R.; Olmsted, R.A.; Prince, G.A.; Murphy, B.R.; Alling, D.W.; Walsh, E.E.; Collins, P.L. (1987a): Antigenic relatedness between glycoproteins of human respiratory suncytial virus subgroups A and B: Evaluation of the contributions of F and G glycoproteins to immunity. *J. Virol.* **61**, 3163-3166.

Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L. (1987b): The G glycoprotein of human respiratory syncytial viruses of subgroups A and B : Extensive sequence divergence between antigenically related proteins. *Proc. Natl. Acad. Sci. USA.* **84**, 5625-5629.

Johnson, P.R. and Collins, P.L. (1988a): The A and B subgroups of human respiratory syncytial virus: comparison of intergenic and gene overlap sequences. *J. Gen. Virol.* **69**, 2901-2906.

Johnson, P.R. and Collins, P.L. (1988b): 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**, 2623-2628.

Johnson, P.R. and Collins, P.L. (1989): The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus of antigenic subgroups A and B: Sequence conservation and divergence within RSV genomic RNA. *J. Gen. Virol.* **70**, 1539-1547.

Johnson, P.R. and Collins, P.L. (1990): Sequence comparison of the phosphoprotein mRNAs of antigenic subgroups A and B of human respiratory syncytial virus identifies a highly divergent domain in the predicted protein. *J. Gen. Virol.* **71**, 481-485.

Kapikian, A.Z.; Mitchell, R.H.; Chanock, R.H.; Shvedoff, R.A.; Stewart, C.E. (1969): An epidemiologic study of altered clinical reactivity to respiratory syncytial virus in

children previously vaccinated with an inactivated RS vaccine. *Am. J. Epidemiol.* **89**, 405-421.

Kaptur, P.E.; Rhodes, R.B.; Lyles, D.S. (1991): Sequences of the vesicular stomatitis virus matrix protein involved in binding to nucleocapsids. *J. Virol.* **65**, 1057-1065.

Kawano, M.; Bando, H.; Yuasa, T.; Kondo, K.; Tsurudome, M.; Komada, H.; Nishio, M.; Ito, Y. (1990): Sequence determination of the hemagglutinin-neuraminidase (HN) gene of the human parainfluenza type 2 virus and the construction of a phylogenetic tree for the HN proteins of all the paramyxoviruses that are infectious to humans. *Virology* **174**, 308-313.

Kim, H.W.; Canchola, J.G.; Brandt, C.D.; Pyles, G.; Chanock, R.M.; Jensen, K.; Parrott, R.H. (1969): Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *American Journal of Epidemiology* **89**, 422-434.

Kim, H.W.; Arrobio, J.O.; Brandt, C.D.; Jeffries, B.C.; Pyles, G.; Reid, J.L. (1973a): Epidemiology of respiratory syncytial virus infection in Washington, DC; I: Importance of the virus in different respiratory tract disease syndromes and temporal distribution of infection. *American Journal of Epidemiology* **98**, 216-225.

Kim, H.W.; Arrobio, J.; Brandt, C.D.; Wright, P.; Hodes, D.; Chanock, R.M.; Parrott, R.H. (1973b): Safety and antigenicity of temperature-sensitive mutant respiratory syncytial virus (RSV) in infants and children. *Pediatrics* **52**, 56-63.

Kingsbury, D.W.; Bratt, M.A.; Choppin, P.W.; Hanson, R.P.; Hosaka, Y.; Meulen, T.; Norrby, E.; Plowright, W.; Rott, R.; Wunner, W.H. (1978): Paramyxoviridae. *Intervirology* **10**, 137-152.

Kingsbury, D.W. (1985): Orthomyxo and paramyxoviruses and their replication. in "Virology" (B.N.Fields, ed.), 1157-1178, Raven Press, New York.

Kolakofsky, D.; Vidal, S.; Curran, J.A. (1991): Paramyxovirus RNA synthesis and P gene expression. In: The Paramyxoviruses. (Ed: Kingsbury,DW). Plenum Press, New York, .

Kondo, K.; Bando, H.; Tsurudome, M.; Kawano, M.; Nishio, M.; Ito, Y. (1990): Sequence analysis of the phosphoprotein (P) genes of human parainfluenza type 4A and 4B virus and RNA editing at transcript of the P genes: The number of G residues added is imprecise. *Virology* **178**, 321-326.

Kozak, M. (1986): Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA.* **83**, 2850-2854.

Kumar, A. and Lindberg, U. (1972): Characterisation of messenger ribonucleoprotein and messenger RNA from KB cells. *Proc. Natl. Acad. Sci. USA.* **69**, 681-685.

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

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

Lambden, P.R. (1985): Nucleotide sequence of the respiratory syncytial virus phosphoprotein gene. *J. Gen. Virol.* **66**, 1607-1612.

Lambert, D.M. and Pons, M.W. (1983): Respiratory syncytial virus glycoproteins. *Virology* **130**, 204-214.

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

Lambert, D.M.; Hambor, J.; Diebold, M.; Galinski, B. (1988): Kinetics of synthesis of respiratory syncytial virus polypeptides. *J. Gen. Virol.* **69**, 313-323.

Leopardi, R.; Hukkanen, V.; Vainionpaa, R.; Salmi,A.A (1993): Cell proteins bind to sites within the 3' noncoding region and the positive strand leader sequence of measles virus RNA. *J. Virol.* **67**, 785-790.

Lerch, R.A.; Stott, E.J.; Wertz, G.W. (1989): Characterisation of bovine respiratory syncytial virus proteins and mRNAs and generation of cDNA clones to the viral mRNAs. *J. Virol.* **63**, 833-840.

Lerch, R.A.., Anderson, K., Amman, V.L. and Wertz, G.W. (1991): Nucleotide sequence analysis of the bovine respiratory syncytial virus fusion protein mRNA and expression from a recombinant vaccinia virus. *Virology*. **181**, 118-131.

Levely, M.E., Bannow, C.A., Smith, C.W. and Nicholas, J.A. (1991): Immunodominant T-cell epitope on the F protein of respiratory syncytial virus recognized by human lymphocytes. *Journal of Virology* **65**, 3789-3796.

Levine, S.; Peeples, M.; Hamilton, R. (1977): Effect of respiratory syncytial virus infection on HeLa cell macromolecular synthesis. *J. Gen. Virol.* **37**, 53-63.

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**, 2521-2524.

Li, Y.; Schubert, M.; Wagner, R.R.; Kang, C.Y. (1993): Viral liposomes released from insect cells infected with recombinant baculovirus expressing the matrix protein of vesicular stomatitis virus. *J. Virol.* **67**, 4415-4420.

Lichtenstein, D.L.; Wertz, G.W.; Ball, L.A. (1991): Mutational analysis of the signal/anchor domain of the glycoprotein G of human respiratory syncytial virus. Poster presented at the Eighth International Conference on negative strand viruses, Charleston, South Carolina, USA.

Ling, R. and Pringle, C.R. (1988): Turkey rhinotracheitis virus: in vivo and in vitro polypeptide synthesis. *J. Gen. Virol.* **69**, 917-923.

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

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

Ling, R.; Easton, A.J.; Pringle, C.R. (1992): Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis virus and their intergenic regions reveals a gene order different from that of other pneumoviruses. *J. Gen. Virol.* **73**(7), 1709-1715.

Lopez, J.A.; Villanueva, N.; Melero, J.; Portela, A. (1988): Nucleotide sequence of the fusion and phosphoprotein genes of human respiratory syncytial (RS) virus Long strain : evidence of subtype genetic heterogeneity. *Virus Res.* **70**, 249-262.

Lopez, J.A.; Andreu, D.; Carreno, C.; Whyte, P.; Taylor, G.; Melero, J.A. (1993): Conformational constraints of conserved neutralising epitopes from a major antigenic area of human respiratory syncytial virus fusion glycoprotein. *J. Gen. Virol.* **74**, 2567-2577.

Lounsbach, G.R.; Bourgeois, C.; West, W.H.L.; Robinson, J.W.; Carter, M.J.; Toms, G.L. (1993): The binding of neutralising monoclonal antibodies to regions of fusion protein of respiratory syncytial virus expressed in Escherichia coli. *J. Gen. Virol.* **74**, 2259-2565.

Lyn, D.; Gill, D.S.; Scroggs, R.A.; Portner, A. (1991): The nucleoproteins of human parainfluenza virus type 1 and Sendai virus share amino acid sequences and antigenic and structural determinants. *J. Gen. Virol.* **72**, 983-987.

Mallipeddi, S.K. and Samal, S.K. (1993): Sequence variability of the glycoprotein gene of bovine respiratory syncytial virus. *J. Gen. Virol.* **74**, 2001-2004.

Malvoisin, E. and Wild, F.T. (1993): Measles virus glycoproteins: Studies on the structure and interaction of the haemagglutinin and fusion proteins. *J. Gen. Virol.* **74**, 2365-2372.

Maniatis, T.; Fritsch, E.F.; Sambrook, J. (1982): Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory. Cold Spring Harbor. New York.

Markwell, M.A.K.; Fox, C.F. (1980): Protein-protein interactions within paramyxoviruses identified by native disulphide bonding or reversible chemical cross-linking. *J. Virol.* **33**, 152-166.

Martin-Gallardo, A.; Fien, K.A.; Hu, B.T.; Farley, J.F.; Seid, R.; Collins, P.L.; Hildreth, S.W.; Paradiso, P.R. (1991): Expression of the F glycoprotein gene from human respiratory syncytial virus in Escherichia coli: mapping of a fusion inhibiting epitope. *Virology* **184**, 428-432.

Marx, P.A.; Portner, A.; Kingsbury, D.W. (1974): Sendai virion transcriptase complex : polypeptide composition and inhibition by virion envelope proteins. *J. Virol.* **13**, 107-112.

Masters, P.S. and Banerjee, A.K. (1988a): Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to non-specific RNA. *J. Virol.* **62**, 2658-2664.

Masters, P.S. and Banerjee, A.K. (1988b): Resolution of multiple complexes of phosphoprotein NS with nucleocapsid protein N of vesicular stomatitis virus. *J. Virol.* **62**, 2651-2657.

Matthews, R.E.F. (1982): Classification and nomenclature of viruses. *Intervirology* **17**, 104-105.

Matsuoka, Y.; Curran, J.; Pelet, T.; Kolakofsky, D.; Ray, R.; Compans, R.W. (1991): The P gene of human parainfluenza virus type 1 encodes P and C proteins but not a cysteine-rich V protein. *J. Virol.* **65**, 3406-3410.

McIntosh, K. and Fishaut, J.M. (1980): Immunopathologic mechanisms in lower respiratory tract disease of infants due to respiratory syncytial virus. *Progress in Medical Virology* **26**, 94-118.

McKay, E.; Higgins, P.; Tyrrell, D.; Pringle, C.R. (1988): Immunogenicity and pathogenicity of temperature-sensitive modified respiratory syncytial virus in adult volunteers. *J. Med. Virol.* **25**, 411-421.

Mellon, M.G. and Emerson, S.U. (1978): Rebinding of transcriptase components (L and NS proteins) to the nucleocapsid template of vesicular stomatitis virus. *J. Virol.* **27**, 560-567.

Merz, D.C.; Scheid, A.; Choppin, P.W. (1980): Importance of antibodies to the fusion glycoprotein of paramyxoviruses in the prevention and spread of infection. *J. Exp. Med.* **151**, 275-288.

Messing, J. and Vieira, J. (1982): A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. *Gene* **19**, 269-276.

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**(2), 615-624.

Morgan, L.A.; Routledge, E.G.; Willcocks, M.M.; Samson, A.C.; Scott, R.; Toms, G.L. (1987): Strain variation of respiratory syncytial virus. *J. Gen. Virol.* **68**, 2781-2788.

Morgan, E.M. (1991): The Paramyxoviruses, Plenum Press, New York.

Morrison, T.G.; Ward, L.J.; Semerjian, A. (1985): Intracellular processing of the Newcastle disease virus fusion protein. *J. Virol.* **53**, 851-857.

Morrison, T.G. and Peeples, M.E. (1987): Conformational change in a viral glycoprotein during maturation due to disulfide bond disruption. *Proc. Natl. Acad. Sci. USA*. **84**, 1020-1024.

Morrison, T.G. (1988): Structure, function, and intracellular processing of paramyxovirus membrane proteins. *Virus Res.* **10**, 113-136.

Morrison, T.G. and Portner, A. (1991): Structure, function, and intracellular processing of the glycoproteins of *Paramyxoviridae*. In: The Paramyxoviruses. (Ed: Kingsbury,DW) Plenum Press, New York and London, 347-382.

Morrison, T.G.; McQuain, C.; McGinnes, L. (1991): Complementation between avirulent Newcastle disease virus and a fusion protein gene expressed from a retrovirus vevtor: Requirements for membrane fusion. *J. Virol.* **65**, 813-822.

Morrison, T.G.; McQuain, C.; Sergel, T.; McGinnes, L.; Reitter, J. (1993): The role of the amino terminusof F1 of the Newcastle disease virus fusion protein in cleavage and fusion. *Virology* **193**, 997-1000.

Moyer, S.; Baker, S.C.; Lessard, J.L. (1986): Tubulin, a factor necessary for the synthesis of both sendai virus and vesicular stomatitis virus RNAs. *Proc. Natl. Acad. Sci. USA*. **83**, 5405-5409.

Moyer, S.; Baker, S.C.; Horikami, S.M. (1990): Host cell proteins required for measles virus reproduction. *J. Gen. Virol.* **71**, 775-783.

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

Mufson, M.A.; Belshe, R.B.; Orvell, C.; Nörrby, E. (1988): Respiratory syncytial virus epidemics : Variable dominance of subgroups A and B strains among children, 1981-1986. *J. Infect. Dis*. **157**, 143-148.

Mufson, M.A.; Akerlind-Stopne, B.; Orvell, C.; Belshe, R.B.; Norrby, E. (1991): A single season epidemic with respiratory syncytial virus subgroup B2 during 10 epidemic years, 1978 to 1988. *J. Clin. Microbiol.* **29**, 162-165.

Murphy, B.R.; Prince, G.A.; Walsh, E.E.; Kim, H.W.; Parrott, R.H.; Hemming, V.G.; Rodriguez, W.J.; Chanock, R.M. (1986): Dissociation between serum neutralising and glycoprotein antibody responses of infants and children who received inactivated respiratory syncytial virus vaccine. *J. Clin. Micro.* **24**, 197-202.

Murphy, B.R.; Olmsted, R.A.; Collins, P.L.; Chanock, R.M.; Prince, G.A. (1988): Passive transfer of respiratory syncytial virus (RSV) antiserum suppresses the immune response to the RSV fusion (F) and large (G) glycoproteins expressed by recombinant vaccinia viruses. *J. Virol.* **62**, 3907-3910.

Nagai, K.; Tsutsumi, H.; Yamazaki, H.; Pattamadilok, S.; Chiba, S. (1993): Three antigenic variant groups in human respiratory syncytial virus subgroup B isolated in Japan. *Arch. Virol.* **128**, 55-63.

Navarro, J.; Lopez-Otin, C.; Villanueva, N. (1991): Location of phosphorylated residues in the human respiratory syncytial virus phosphoprotein. *J. Gen. Virol.* **72,** 1455-1459.

Nicholas, J.A.; Rubino, K.L.; Leverley, M.E.; Adams, E.G.; Collins, P.L. (1990): Cytolytic T-lymphocyte responses to respiratory syncytial virus: Effector cell phenotype and target Proteins. *J. Virol.* **64**(9), 4232-4241.

Ng, D.T.W.; Randall, R.E.; Lamb, R.A. (1989): Intracellular maturation and transport of the SV5 type II glycoprotein haemagglutinin-neuraminidase: Specific and transient association with GRP78-BiP in the endoplasmic reticulum and extensive internalisation from the cell surface. *J. Cell Biol.* **109**, 3273-3289.

Nörrby, E.; Mufson, M.A.; Sheshberadaran, H. (1986): Structural differences between subtype A and B strains of respiratory syncytial virus. *J. Gen. Virol.* **67**, 2721-2729.

Nörrby, E.; Mufson, M.A.; Alexander, H.; Hougten, R.A.; Lerner, R.A. (1987): Site-directed serology with synthetic peptides representing the large glycoprotein G of respiratory syncytial virus. *Proc. Natl. Acad. Sci. USA*. **84**, 6572-6576.

Ogden, J.R.; Pal, R.; Wagner, R.R. (1986): Mapping regions of the matrix protein of vesicular stomatitis virus which bind to ribonucleocapsids, liposomes and monoclonal antibodies. *J. Virol.* **58**, 860-868.

Ohgimoto, S.; Bando, H.; Kawano, M.; Okamoto, K.; Kondo, K.; Tsurudome, M.; Nishio, M.; Ito, Y. (1990): Sequence analysis of P gene of human parainfluenza type 2 virus: P and cysteine-rich proteins are translated by two mRNAs that differ by two nontemplated G residues. *Virology* **177**, 116-123.

Olmsted, R.A.; Elango, N.; Prince, G.A.; Murphy, B.R.; Johnson, P.R.; Moss, B.; Chanock, R.M.; Collins, P.L. (1986): Expression of the fusion glycoprotein of respiratory syncytial virus by a recombinant Vaccinia virus: comparison of the individual contributions of the fusion and attachment glycoproteins to host immunity. *Proc.Natl.Acad.Sci. USA*. **83**, 7462-7466.

Olmsted, R.A. and Collins, P.L. (1989): The 1A protein of respiratory syncytial virus is an integral membrane protein present as multiple, structurally distinct species. *J.Virol.* **63**, 2019-2029.

Olmsted, R.A.; Murphy, B.R.; Lawrence, L.A.; Elango, N.; Moss, B.; Collins, P.L. (1989): Processing, surface expression, and immunogenicity of carboxy-terminally truncated mutants of the G protein of human respiratory syncytial virus. *J.Virol.* **63**, 411-420.

Openshaw, P.J.M.; Pemberton, R.M.; Ball, L.A.; King, A.M.Q.; Wertz, G.W.; Askonas, B.A. (1988): Helper T-cell recognition of respiratory syncytial virus in mice. *J. Gen. Virol.* **69**, 305-312.

Palomo, C.; Garcia-Barreno, B.; Penas, P.; Melero, J.A. (1991): The G protein of human respiratory syncytial virus: significance of carbohydrate side chains and the C-terminal end to its antigenicity. *J. Gen. Virol.* **72**, 669-675.

Paradiso, P.R.; Hu, B.; Hildreth, S. (1987): Structure of the respiratory syncytial virus glycoprotein G, in "VII International Congress of Virology, Abstracts," R25.21. p.189.

Paradiso, P.R.; Hu, B.T.; Armugham, R.;Hildreth, S. (1991): Mapping of a fusion related epitope of the respiratory syncytial virus fusion protein. *Vaccine*, **9**, 231-237.

Paterson, R.; Harris, T.; Lamb, R.A. (1984): Fusion protein of the paramyxovirus simian virus 5: nucleotide sequence of mRNA predicts a highly hydrophobic glycoprotein. *Proc. Natl. Acad. Sci. USA*. **81**, 6706-6710.

Paterson, R.G. and Lamb, R.A. (1990): RNA editing by nucleotide insertion in mumps virus P-gene mRNA transcripts. *J.Virol.* **64**, 4137-4145.

Pattnaik, A. and Wertz, G.W. (1990): Replication and amplification of defective interfering particle RNAs of vesicular stomatitis virus in cells expressing viral proteins from vectors containing cloned cDNAs. *J.Virol.* **64**, 2948-2957.

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

Patton, J.T.; Davis, N.L.; Wertz, G.W. (1984): N protein alone satisfies the requirement for protein synthesis during RNA replication of VSV. *J.Virol.* **49**, 303-309.

Peeples, M.E. and Levine, S. (1979): Respiratory syncytial virus polypeptides: their location in the virion. *Virology* **95**, 137-145.

Peeples, M.E. and Levine, S. (1980): Metabolic requirements for the maturation of respiratory syncytial virus. *J. Gen. Virol.* **50**, 81-88.

Peeples, M.E. and Bratt, M.A. (1984): Mutation in the matrix protein of Newcastle disease virus can result in decreased fusion glycoprotein incorporation into particles and decreased infectivity. *J. Virol.* **51**, 81-90.

Peeples, M.E. (1991): Paramyxovirus M proteins: Pulling it all together and taking it on the road. In: The Paramyxoviruses. (Ed: Kingsbury,DW) Plenum Press, New York and London, 427-456.

Pelet, T.; Curran, J.; Kolakofsky, D. (1991): The P gene of bovine parainfluenza virus type 3 expresses all three reading frames from a single mRNA editing site. *E.M.B.O. J.* **10**, 443-448.

Peluso, R.W. and Moyer, S.A. (1988): Viral proteins required for the in vitro replication of vesicular stomatitis virus defective interfering particle genome RNA. *Virology* **162**, 369-376.

Pemberton, R.M.; Cannon, M.J.; Openshaw, P.J.M.; Ball, L.A.; Wertz, G.W.; Askonas, B.A. (1987): Cytotoxic T-cell specificity for respiratory syncytial virus proteins : Fusion protein is an important target antigen. *J. Gen. Virol.* **68**, 2177-2182.

Pinto, L.H.; Holsinger, L.J.; Lamb, R.W. (1992): Influenza virus M2 protein has ion channel activity. *Cell* **69**, 517-526.

Popow-Kraupp, T.; Lakits, E.; Kellner, G.; Kunz, C. (1989): Immunoglobulin-class-specific immune response to respiratory syncytial virus structural proteins in infants, children, and adults. *J. Med. Virol.* **27**, 215-223.

Portner, A.; Marx, P.A.; Kingsbury, D.W. (1974): Isolation and characterization of Sendai virus temperature-sensitive mutants. *J. Virol.* **13**, 298-304.

Portner, A.; Scroggs, R.A.; Naeve, C.W. (1987): The fusion glycoprotein of Sendai virus: Sequence analysis of an epitope involved in fusion and virus neutralisation. *Virology* **157**, 556-559.

Pringle, C.R.; Shirodaria, P.V.; Gimenez, H.B.; Levine, S. (1981): Antigen and polypeptide synthesis by temperature-sensitive mutants of respiratory syncytial virus. *J. Gen. Virol.* **54**, 173-183.

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

Pringle, C.R. (1987): Paramyxoviruses and disease. In: Molecular basis of virus disease, Society for general microbiology sysposium. Vol. 40. (Eds: Russell,WC; Almond,JW) Cambridge University Press, Cambridge, 91-138.

Pringle, C.R.; Filipuik, A.H.; Robinson, B.S.; Watt, P.J.; Higgins, P.; Tyrrell, D.A.J. (1993): Immunogenicity and pathogenicity of a triple temperature-sensitive modified respiratory syncytial virus in adult volunteers. *Vaccine* **11**(4), 473-478.

Queen, C. and Korn, L.J. (1984): A comprehensive sequence analysis program for the IBM personal computer. *Nucl. Acids. Res.* **12**, 581-599.

Randhawa, J. (1993): Molecular characterisation of the pneumonia virus of mice glycoprotein genes. PhD Thesis. University of Warick. Coventry. U.K.

Richardson, L.S.; Schnitzer, T.J.; Belshe, R.B.; Camargo, E.; Prevar, D.A.; Chanock, R.M. (1977): Isolation and characterization of further defective clones of a temperature-sensitive mutant (ts-1) of respiratory syncytial virus. *Arch. Virol.* **54**, 53-60.

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

Roberts, S.R.; Lichtenstein, D.; Ball, L.A.; Wertz, G.W. (1993): Secreted and anchored froms of the attachment protein of respiratory syncytial virus are synthesized from different translational initiation sites (on the same mRNA). Presented at the IXth International Congress of Virology, Glasgow, Scotland, U.K.

Robinson, B.S. and Everson, J.S. (1992): Epitope specificities of human serum antibodies reactive with respiratory syncytial virus fusion protein. *Arch. Virol.* **125**, 273-286.

Routledge, E.G.; Willcocks, M.M.; Morgan, L.; Samson, A.C.R.; Scott, R.; Toms, G.L. (1987a): Expression of the respiratory syncytial virus 22K protein on the surface of infected HeLa cells. *J. Gen. Virol.* **68**, 1217-1222.

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

Rueda, P.; Delgardo, T.; Portella, A.; Melero, J.A.; Garcia-Barreno, B. (1991): Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. *J. Virol.* **65**, 3374-3378.

Rueda, P.; Garcia-Barreno, B.; Melero, J.A. (1994): Loss of conserved cysteine residues in the attachment (G) glycoprotein of two human respiratory syncytial virus escape mutants that contain multiple A-G substitutions (hypermutations). *Virology*. **198**, 653-662.

Ryan, K.W. and Kingsbury, D.W. (1988): Carboxyl-terminal region of Sendai virus P protein is required for binding to viral nucleocapsids. *Virology* **167**, 106-112.

Ryan, K.W. and Portner, A. (1990): Separate domains of sendai virus P protein are required for binding to viral nucleocapsids. *Virology* **174**, 515-521.

Ryan, K.W.; Morgan, E.M.; Portner, A. (1991): Two non-contiguous regions of sendai virus P protein combine to form a single nucleocapsid binding site. *Virology* **180**, 126-134.

Saiki, R.K.; Gelfand, D.H.; Stoffel, S.; Scharf, S.J.; Higuchi, R.; Horn, G.T.; Mullis, K.B.; Erlich, H.A. (1988): Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-497.

Samal, S.K.; Pastey, M.K.; McPhilips, T.H.; Mohanty, S.B. (1993): Bovine respiratory syncytial virus nucleocapsid protein expressed in insect cells interacts with the phosphoprotein and the M2 protein. *Virology* **193**, 470-473.

Samson, A.C.R.; Willcocks, M.M.; Routledge, E.G.; Morgan, L.A.; and Toms, G.L. (1986). A neutralizing monoclonal antibody to respiratory syncytial virus which binds to both F1 and F2 components of the fusion protein. *J. Gen. Virol.* **67**, 1479-1483.

Sanchez, A.; De, B.P.; Banerjee, A.K. (1985): In vitro phosphorylation of NS protein by the L protein of vesicular stomatitis virus. *J. Gen. Virol.* **66**, 1025-1036.

Sanchez, A.; Kiley, M.P.; Holloway, B.P.; McCormick, J.B.; Auperin, D.D. (1987): Sequence analysis of the Ebola virus genome indicates that it is similar in organisation to paramyxoviruses and rhabdoviruses, in "VII International Congress of Virology, Abstracts," R25.32, p191.

Sanderson, C.M.; McQueen, N.L.; Nayak, D.P. (1993): Sendai virus assembly: M protein binds to viral glycoproteins in transit through the secretory pathway. *J. Virol.* **67**, 651-663.

Sanderson, C.M.; Wu, H.H.; Nayak, D.P (1994): Sendai virus M protein binds independently to either the F or the HN glycoproteins *in vivo*. *J. Virol.* **68**(1), 69-76.

Sanger, F.; Nicklen, S.; Coulson, A.R. (1977): DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.

Satake, M.; Elango, N.; Venkatesan, S. (1984): Sequence analysis of the respiratory syncytial virus phosphoprotein gene. *J. Virol.* **52**, 991-994.

Satake, M. and Venkatesan, S. (1984): Nucleotide sequence of the gene encoding the respiratory syncytial virus matrix protein. *J. Virol.* **50**, 92-99.

Satake, M.; Coligan, J.E.; Elango, N.; Norrby, E.; Venkatesan, S. (1985): Respiratory syncytial virus envelope glycoprotein (G) has a novel structure. *Nucl. Acids. Res.* **13**, 7795-7812.

Scheid, A. (1977): Two disulfide-linked polypeptide chains constitute the active F protein of paramyxoviruses. *Virology* **80**, 54-66.

Scopes, G.E.; Watt, P.J.; Lambden, P.R. (1990): Identification of a linear epitope on the fusion glycoprotein of respiratory syncytial virus. *J. Gen. Virol.* **71**, 53-59.

Sechoy, O.; Philippot, J.R.; Bienvenue, A. (1987): F protein-F protein interaction within the Sendai virus identified by native bonding or chemical cross-linking. *J. Biol. Chem* **262**, 11519-11523.

Serghini, M.A.; Ritzenthaler, C.; Pinck, L. (1989): A rapid efficient 'mini-prep' for isolation of plasmid DNA. *Nucl. Acids. Res.* **17**, 3604.

Shioda, T.; Hidaka, K.; Kanda, T.; Shibuta, H.; Nomoto, A.; Iwasaki, K. (1983): Sequence of 3687 nucleotides from the 3' end of Sendai virus genome RNA and the

predicted amino acid sequences of viral NP, P and C proteins. *Nucl. Acids. Res.* **11**, 7317-7330.

Shioda, T.; Iwasaki, K.; Shibuta, H. (1986): Determination of the complete nucleotide sequence of the Sendai virus genome RNA and the predicted amino acid sequences of the F, HN and L proteins. *Nucl. Acids. Res.* **14**, 1545-1563.

Sjoblom, I.; Lundstrom, M.; Sjogren-Jansson, E.; Glorioso, J.C.; Jeansson, S.; Olofsson, S. (1987): Demonstration and mapping of highly carbohydrate-dependent epitopes in the Herpes simplex virus type 1-specified glycoprotein C. *J.Gen.Virol.* **68**, 545-554.

Skehel, J.J.; Stevens, D.J.; Daniels, R.S.; Douglas, A.R.; Knossow, M.; Wilson, I.A.; Wiley, D.C. (1984): A carbohydrate side chain on hemaglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. *Proc. Natl. Acad. Sci. USA.* **81**, 1779-1783.

Southern, J.A.; Precious, B.; Randall, R.E. (1990): Two non-templated nucleotide additions are required to generate the P mRNA of parainfluenza virus type 2 since the RNA genome encodes protein V. *Virology* **177**, 388-390.

Spehner, D.; Kirn, A.; Drillien, R. (1991): Assembly of nucleocapsid-like structures in animal cells infected with a vaccinia virus recombinant encoding the measles virus nucleoprotein. *J.Virol.* **65**(11), 6296-6300.

Spriggs, M.K. and Collins, P.L. (1986a): Sequence analysis of the P and C protein genes of human parainfluenza virus type 3 : Patterns of amino acid sequence homology among paramyxovirus proteins. *J. Gen. Virol.* **67**, 2705-2719.

Spriggs, M.K. and Collins, P.L. (1986b): Human parainfluenza virus type 3 : messenger RNAs, polypeptide coding assignments, intergenic sequences and genetic map. *J.Virol.* **59**, 646-654.

Spriggs, M.K.; Olmsted, R.A.; Venkatesan, S.; Coligan, J.E.; Collins, P.L. (1986c): Fusion protein of human parainfluenza virus type 3: nucleotide sequence of the gene, direct identification of the cleavage activation site, and comparison with other paramyxoviruses. *Virology* **152**, 241-251.

Stec, D.S.; Hill, M.G.; 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.

Steinhauer, D.A. and Holland, J.J. (1987): Rapid evolution of RNA viruses. *Ann. Rev. Microbiol.* **41**, 409-433.

Steward, M.; Vipond, I.B.; Millar, N.S.; Emmerson, P.T. (1993): RNA editing in Newcastle disease virus. *J. Gen. Virol.* **74**, 2539-2547.

Stott, E.J.; Thomas, L.H.; Collins, A.P.; Crouch, S.; Jebbett, J.; Smith, G.S.; Luther, P.D.; Caswell, R. (1980): A survey of virus infections of the respiratory tract of cattle and their association with disease. *J. Hygiene* **85**, 257-270.

Stott, E.J.; Bew, M.H.; Taylor, G.; Jebbett, J.; Collins, A.P. (1984): The characterisation and uses of monoclonal antibodies to respiratory syncytial virus. *Dev.Biol.Std.* **57**, 237-244.

Stott, E.J. and Taylor, G. (1985): Respiratory syncytial virus : brief review. *Arch. Virol.* **84**, 1-52.

Stott, E.J.; Ball, L.A.; Young, K.K.; Furze, J.; Wertz, G.W. (1986): Human respiratory syncytial virus glycoprotein G expressed from a recombinant Vaccinia virus vector protects mice against live-virus challenge. *J.Virol.* **60**, 607-613.

Stott, E.J.; Ball, L.A.; Anderson, K.; Young, K.K.; King, A.M.Q.; Wertz, G.W. (1987): Immune and histopathological responses in animals vaccinated with recombinant vaccinia viruses that express individual genes of human respiratory syncytial virus. *J.Virol.* **61**, 3855-3861.

Sullender, W.M.; Anderson, K.; Wertz, G.W. (1990): The respiratory syncytial virus subgroup B attachment glycoprotein: Analysis of sequence, expression from a recombinant vector, and evaluation as an immunogen against homologous and heterologous subgroup virus challenge. *Virology* **178**, 195-203.

Sullender, W.M.; Mufson, M.A.; Anderson, L.J.; Wertz, G.W. (1991): Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J.Virol.* **65**, 5425-5434.

Sullender, W.M. and Wertz, G.W. (1991): The unusual attachment glycoprotein of the respiratory syncytial viruses: Structure, maturation, and role in immunity. In: The Paramyxoviruses. (Ed: Kingsbury,DW) Plenum Press, New York and London, 383-406.

Takeuchi, K.; Tanabayashi, K.; Hishiyama, M.; Yamada, Y.K.; Yamada, A.; Sugiura, A. (1990): Detection and characterization of mumps virus V protein. *Virology* **178**, 247-253.

Takacs, A.M.; Barik, S.; Das, T.; Banerjee, A.K. (1992): Phosphorylation of specific serine residues within the acidic domain of the phosphoprotein of vesicular stomatitis virus regulates transcription *in vitro*. *J. Virol.* **66**, 5842-5848.

Tanabayashi, K., Takeuchi, K., Okazaki, K., Hishiyama, M. and Yamada, A. (1992): Expression of mumps virus glycoproteins in mammalian cells from cloned cDNAs: Both F and HN proteins are required for cell fusion. *Virology* **187**, 801-804.

Tashiro, M.; Takeda, M.; Tanaka, S.; Nishimura, N.; Takenaka, M.; Kido, H. (1993): Antibody against the carboxyl terminus of the F2 subunit of Sendai virus fusion glycoprotein inhibits proteolytic activation. *Virology* **194**, 882-885.

Taylor, G.; Stott, E.J.; Bew, M.; Fernie, B.F.; Cote, P.J.; Collins, P.L.; Hughes, M.; Jebbett, J. (1984): Monoclonal antibodies protect against respiratory syncytial virus infection in mice. *Immunology* **52**, 137-142.

Taylor, G.; Stott, E.J.; Furze, J.; Ford, J.; Sopp, P. (1992): Protective epitopes on the fusion protein of respiratory syncytial virus recognized by murine and bovine monoclonal antibodies. *J. Gen. Virol.* **73**(9), 2217-2223.

Thomas, S.M.; Lamb, R.A.; Paterson, R.G. (1988): Two mRNAs that differ by two non-templated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. *Cell* **54**, 891-902.

Towbin, H.; Staehelin, T.; Gordon, J. (1979): Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose sheets : Procedure and some applications. *Proc. Natl. Acad. Sci. USA.* **76**, 4350-4354.

Toyoda, T.; Sakaguchi, T.; Imai, K.; Inocencio, N.; Gotoh, B., Hamaguchi, M.; Nagai, Y. (1987): Structural comparison of the cleavage-activation site of the fusion glycoprotein between virulent and avirulent strains of Newcastle disease virus. *Virology* **158**, 242-247.

Trudel, M.; Nadon, F.; Seguin, C.; Dionne, G.; LaCroix, M. (1987a): Identification of a synthetic peptide as part of a major neutralization epitope of respiratory syncytial virus. *J. Gen. Virol.* **68**, 2273-2280.

Trudel, M.; Nadon, F.; Seguin, C.; Payment, P.; Talbot, P. (1987b): Respiratory syncytial virus fusion glycoprotein: Further characterization of a major epitope involved in virus neutralization. *Can. J. Microbiol.* **33**, 933-938.

Trudel, M.; Stott, E.J.; Taylor, G.; Oth, D.; Mercier, G.; Nadon, F.; Seguin, C.; Simard, C.; Lacroix, M. (1991): Synthetic peptides corresponding to the F protein of RSV stimulate murine B and T cells but fail to confer protection. *Arch. Virol.* **117**, 59-71.

Tsipis, J.E. and Bratt, M.A. (1976): Isolation and preliminary characterization of temperature-sensitive mutants of Newcastle disease virus. *J. Virol.* **18**, 848-855.

Van Wyke-Coelingh, K. and Winter, C.C. (1991): Naturally occurring human Parainfluenza type 3 viruses exhibit divergence in amino acid sequence of their fusion protein neutralisation epitopes and cleavage sites. *J. Virol.* **64**, 1329-1334.

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

Vidal, S.; Curran, J.; Kolakofsky, D. (1990): Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *J. Virol.* **64**, 239-246.

Villanueva, N.; Navarro, J.; Mendez, E.; Garcia-Albert, I. (1994): Identification of a protein kinase involved in the phosphorylation of the C-terminal region of human respiratory syncytial virus P protein. *J. Gen. Virol.* **75**, 555-565.

Walravens, K.; Kettmann, R.; Collard, A.; Coppe, P.; Burny, A. (1990): Sequence comparison between the human and bovine respiratory syncytial viruses. *J. Gen. Virol.* **71**, 3009-3014.

Walsh, E.E.; Schlesinger, J.J.; Brandriss, M.W. (1984a): Protection from respiratory syncytial virus infection in cotton rats by passive transfer of monoclonal antibodies. *Infect. Immun.* **43**, 756-758.

Walsh, E.E.; Schlesinger, J.J.; Brandriss, M.W. (1984b): Purification and characterization of GP90, one of the envelope glycoproteins of respiratory syncytial virus. *J. Gen. Virol.* **65**, 761-767.

Walsh, E.E.; Brandriss, M.W.; Schlesinger, J.J. (1985): Purification and characterization of the respiratory syncytial virus fusion protein. *J. Gen. Virol.* **66**, 409-415.

Walsh, E.E.; Brandriss, M.W.; Schlesinger, J.J. (1987a): Immunological differences between the envelope glycoproteins of two strains of human respiratory syncytial virus. *J. Gen. Virol.* **68**, 2169-2176.

Walsh, E.E.; Hall, C.B.; Briselli, M.; Brandriss, M.W.; Schlesinger, J.J. (1987b): Immunization with glycoprotein subunits of respiratory syncytial virus to protect cotton rats against viral infection. *J. Infect. Dis.* **155**, 1198-1204.

Walsh, E.E.; Hall, C.B.; Schlesinger, J.J.; Brandriss, M.W.; Hildreth, S.; Paradiso, P. (1989): Comparison of antigenic sites of subtype-specific respiratory syncytial virus attachment proteins. *J. Gen. Virol.* **70**, 2953-2961.

Wang, C.; Raghu, G.; Morrison, T.; Peeples, M.E. (1992): Intracellular processing of the paramyxovirus F protein: Critical role of the predicted amphipathic alpha helix adjacent to the fusion domain. *J. Virol.* **66**(7), 4161-4169.

Ward, K.A.; Lambden, P.R.; Ogilvie, M.M.; Watt, P.J. (1983): Antibodies to respiratory syncytial virus polypeptides and their significance in human infection. *J. Gen. Virol.* **64**, 1867-1876.

Ward, K.A.; Everson, J.S.; Lambden, P.R.; Watt, P.J. (1984): Antigenic and structural variation in the major nucleocapsid protein of respiratory syncytial virus. *J. Gen. Virol.* **65**, 1749-1757.

Wathen, M.W.; Brideau, R.J.; Thomsen, D.R. (1989): Immunization of cotton rats with the human respiratory syncytial virus F glycoprotein produced using a baculovirus vector. *J. Infect. Dis.* **159**, 255-264.

Watt, P.J.; Robinson, B.S.; Pringle, C.R.; Tyrrell, D.A.J. (1990): Determinants of susceptibility to challenge and the antibody response of adult volunteers given experimental respiratory syncytial virus vaccines. *Vaccine* **8**, 231-236.

Waxham, N.M. and Wolinsky, J.S. (1986): A fusing mumps virus variant selected from a nonfusing parent with the neuraminidase inhibitor 2-deoxy-2,3-dehydro-N-acteylneuraminic acid. *Virology* **151**, 286-295.

Welliver, R.C.; Sun, M.; Rinaldo, D.; Ogra, P.L. (1985): Respiratory syncytial virus-specific IgE responses following infection: Evidence for a predominantly mucosal response. *Pediatr. Res.* **19**, 420-424.

Wertz, G.W. and Levine, M. (1973): RNA synthesis by vesicular stomatitis virus and a small plaque mutant : Effects of cycloheximide. *J. Virol.* **12**, 253-264.

Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, A.L. (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. USA.* **82**, 4075-4089.

Wertz, G.W.; Stott, E.J.; Young, K.K.Y.; Anderson, K.; Ball, L.A. (1987): Expression of the fusion protein of human respiratory syncytial virus from recombinant Vaccinia virus vectors and protection of vaccinated mice. *J. Virol.* **61**, 293-301.

Wertz, G.W.; Krieger, M.; Ball, L.A. (1989): Structure and cell surface maturation of the attachment glycoprotein of human respiratory syncytial virus in a cell line deficient in O-glycosylation. *J. Virol.* **63**, 4767-4776.

White, D. and Fenner, F. (1986): Viral genetics and evolution. In: 'Medical virology'. Academic Press, London, pp91-118.

Wild, T.F.; Malvoisin, E.; Buckland, R. (1991): Measles virus: both the haemagglutinin and fusion glycoproteins are required for fusion. *J. Gen. Virol.* **72**, 439-442.

Williams, M.A. and Lamb, R.A. (1986): Determination of the orientation of an integral membrane protein and sites of glycoprotein by oligonucleotide-directed mutagenesis: Influenza B virus NB glycoprotein lacks a cleavable signal sequence and has an extracellular NH_2-terminal region. *Mol.Cell.Biol.* **6**, 4317-4328.

Wright, P.F.; Mills, J.; Chanock, R.M. (1971): Evaluation of a temperature-sensitive mutant of respiratory syncytial virus in adults. *J. Inf. Dis.* **124**, 505-511.

Wright, P.F.; Gharpure, M.A.; Hodes, D.S.; Chanock, R.M. (1973): Genetic studies of respiratory syncytial virus temperature-sensitive mutants. *Arch. Virol.* **41**, 238-247.

Wright, P.F.; Belshe, R.B.; Kim, H.W.; Van Voris, L.P.; Chanock, R.M. (1982): Administration of a highly attenuated live respiratory syncytial virus vaccine to adults and children. *Inf. Immun.* **37**, 397-400.

Yamada, H.; Hayata, S.; Omata-Yamada, T.; Taira, H.; Mizumoto, K.; Iwasaki, K. (1990): Association of sendai virus C protein with nucleocapsids. *Arch. Virol.* **113**, 245-253.

Yu, Q.; Davis, P.J.; Barrett, T.; Binns, M.M.; Boursnell, M.E.G.; Cavanagh, D. (1991): Deduced amino acid sequence of the fusion glycoprotein of turkey rhinotracheitis virus has greater identity with that of human respiratory syncytial virus, a pneumovirus, than that of paramyxoviruses and morbilliviruses. *J. Gen. Virol.* **72**, 75-81.

Yu, Q.; Davis, P.J.; Brown, T.D.K.; Cavanagh, D. (1992a): Sequence and in vitro expression of the M2 gene of turkey rhinotracheitis pneumovirus. *J. Gen. Virol.* **73**(6), 1355-1363.

Yu, Q.; Davis, P.J.; Li, J.; Cavanagh, D. (1992b): Cloning and sequencing of the matrix protein (M) gene of turkey rhinotracheitis virus reveal a gene order different from that of respiratory syncytial virus. *Virology* **186**(2), 426-434.

Yusoff, K.; Millar, N.S.; Chambers, P.; Emmerson, P.T. (1987): Nucleotide sequence analysis of the L gene of Newcastle disease virus; homologies with Sendai and vesicular stomatitis viruses. *Nucl. Acids. Res.* **15**, 3961-3976.

Zamora, M. and Samal, S.K. (1992a): Gene junction sequences of bovine respiratory syncytial virus. *Virus Res.* **24**(1), 115-121.

Zamora, M. and Samal, S.K. (1992b): Sequence analysis of M2 mRNA of bovine respiratory syncytial virus obtained from an F-M2 dicistronic mRNA suggests structural homology with that of human respiratory syncytial virus. *J. Gen. Virol.* **73**(3), 737-741.

Zebedee, S.L. and Lamb, R.A. (1988): Influenza A virus M2 protein: Monoclonal antibody restriction of virus growth and detection of M2 in virions. *J. Virol.*. **62**, 2762-2772.