National Library of Canada / Bibliothèque nationale du Canada

Canadian Theses Service / Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canadä

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-53280-7

Canada

ENVELOPE ASSOCIATED PROTEINS OF HUMAN PARAINFLUENZA VIRUS 3:
I. M (MATRIX) GENE SEQUENCE, AND;
II. F (FUSION) GENE SEQUENCE COMPARISON
AMONG TEN ISOLATES.

A Thesis Submitted to the

School of Graduate Studies

University of Ottawa

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

Department of Microbiology and Immunology

School of Medicine

By

Kevin Andrew Prinoski

© Kevin Andrew Prinoski, Ottawa, Canada, 1989.

ABSTRACT

Human parainfluenza virus 3 (HPIV3) was originally isolated in 1956 and is presently an important cause of respiratory tract infections in young children. The virus is endemic with a world-wide distribution and causes respiratory tract infections which range in severity from severe to asymptomatic. Although primary infection usually occurs very early in life, immunity to HPIV3 is short-term and re-infections are common. HPIV3 is transmitted via the respiratory route and can create severe problems in nurseries and neonatal or pediatric wards.

The membrane (M) protein of Paramyxoviruses is essential for viral assembly and the fusion (F) surface glycoprotein is necessary for the spread of infection and host cell penetration, both in vitro and in vivo. Cleavage-activation of the F_0 precursor yields the active disulphide linked protein: $F_{1,2}$. A complete understanding of the involvement of the M and F proteins in the molecular biology of Paramyxoviruses has not yet been achieved. In this study, sequence analysis of the HPIV3 M and F genes was undertaken as an initial step towards relating biological function to specific amino acid sequences.

The M gene sequence of HPIV3 strain 47885 was determined for comparison with the known M gene sequences of other Paramyxoviruses. Computer generated sequence alignments of the HPIV3 M gene and protein with those of related RNA viruses revealed that the highest degrees of similarity existed with bovine parainfluenza virus 3 (BPIV3) and Sendai virus - both members of the same genus as HPIV3. Hydropathy plots

generated for the M proteins of HPIV3, BPIV3, Sendai virus, measles virus and Newcastle disease virus (NDV) were virtually identical or at least very similar. A conserved region predicted to occur within the M protein of HPIV3 and other Paramyxoviruses may interact with host-derived viral membranes.

The F gene sequence of HPIV3 strain 47885 was compared with those of nine clinical isolates obtained from different geographical locations between the years 1959 to 1987. Although nucleotide diversity was most prominent in the nontranslated regions, all transcriptional start and stop sequences were maintained among the isolates examined. The sequence data are consistent with the presence of distinct lineages of HPIV3 and suggest that some of the nucleotide differences may be accumulating. No amino acid substitutions involving charged residues were predicted to occur within the three major hydrophobic regions: the F_2 signal peptide, the F_1 membrane anchor and the F_1 amino terminus. Amino acid variation affecting the cleavage-activation site of two of the isolates, predicted on the basis of nucleotide sequence analysis, might lead to differences in virulence.

ACKNOWLEDGEMENTS

This text gives no indication of the patience or understanding that made these experiments and documentation possible. I sincerely thank my supervisor, Dr. Ken Dimock, whose virtues extend beyond science and the lab.

My work was assisted greatly by Don Murphy's generosity with references, experimental protocols and advice. His input was especially appreciated during the 1988 meeting of the American Society for Virology.

Special thanks to: Dr. Syed Sattar for his recommendation that I be accepted as a graduate student; Dr. Kathy Wright for providing desk space in her lab; Dr. Edmond Rossier for his suggestions concerning viral attenuation; Mark Parrington for introducing me to the PC Gene software; and the forever amiable Dr. Edgar Perry for his help with course work.

I would also like to express my gratitude to the remaining members of the Department of Microbiology and Immunology: students, faculty and staff - for assistance rendered numerous times for various reasons.

DEDICATION

I dedicate this thesis and more to Gaynia Dale.

TABLE OF CONTENTS

Page

ABSTRACT........ ii

ACKNOWLEDGEMENTS........ iv

DEDICATION........ v

TABLE OF CONTENTS........ vi

LIST OF FIGURES........ x

LIST OF TABLES........ xi

LIST OF ABBREVIATIONS........ xii

Chapter

I. GENERAL INTRODUCTION........ 1

A. Taxonomy........ 1

B. Morphology of the Paramyxoviruses........ 2

C. Genomic Organization of HPIV3........ 4

D. Replication of the Paramyxoviruses........ 6

E. Aspects of HPIV3 Infection and Disease........ 7

1. Clinical features........ 7

2. Pathogenesis........ 8

3. Epidemiology........ 8

F. Rationale and Objectives........ 9

1. Project rationale........ 9

2. Statement of objectives........ 10

II. MATERIALS AND METHODS........ 11

A. Cell Culture........ 11

1. Cells........ 11

2. Virus........ 12

3. Plaque assay.. 12

4. HPIV3 infection of cells........................... 14

5. Preparation of HPIV3 stocks......................... 14

B. Use of M13.. 14

1. Transfection with M13 and subclone amplification..... 14

2. Subclone screening.................................. 15

C. Production, Extraction and Preparation of Nucleic Acids. 16

1. HPIV3 viral genomic RNA............................. 16

2. M13 single-stranded DNA............................. 17

3. M13 replicative form DNA............................ 18

D. Nucleotide Sequencing................................... 19

1. Sanger sequencing of HPIV3 genomic RNA.............. 19

2. Sanger sequencing of cDNA clones.................... 21

3. Maxam and Gilbert sequencing........................ 22

4. Electrophoresis and autoradiography................. 24

III. SEQUENCING OF THE HPIV3 M GENE............................ 26

A. Introduction.. 26

1. Virus maturation: assembly and budding.............. 26

2. Regulation of RNA synthesis......................... 31

3. Relevance of the M protein to _Paramyxoviral_ pathogenesis... 31

B. Results... 32

C. Discussion.. 42

1. Sequence analysis of the HPIV3 M gene............... 42

2. Computer-generated nucleotide and amino acid sequence alignments.. 44
3. Sequence analysis of the HPIV3 M protein............. 45
4. Conserved secondary structure among Paramyxoviral M proteins.. 47
5. Structure-function interdependence in M proteins..... 49
6. Summary of functional regions within the HPIV3 matrix protein.. 50
IV. F GENE SEQUENCE COMPARISON AMONG TEN HPIV3 ISOLATES........ 51
A. Introduction.. 51
1. Structure of the HPIV3 fusion gene and glycoprotein.. 51
2. Function of the Paramyxovirus fusion glycoprotein.... 54
3. Contributions to the pathogenesis of Paramyxoviruses. 55
4. Immune response to the Paramyxovirus F glycoprotein.. 55
B. Results.. 56
C. Discussion.. 64
1. F gene sequence variation among HPIV3 clinical isolates.. 64
2. Comparison of nucleotide differences in the F and HN genes.. 67
3. F protein sequence variation among HPIV3 clinical isolates.. 73
4. Comparison of amino acid differences in the F and HN glycoproteins.. 74

5. The effects of variability in the F protein on virulence.. 74

6. Factors affecting change in negative strand virus glycoproteins.. 76

7. Viral attenuation in cell culture.................... 77

V. GENERAL CONCLUSIONS: POSSIBLE INTERACTIONS OF THE M PROTEIN WITH THE SURFACE GLYCOPROTEINS OF PARAMYXOVIRUSES.......... 79

A. Possible Interactions of M Protein with HN.............. 79

1. DI particles and persistence......................... 79

2. Charge predictions based on sequence analysis........ 80

B. Possible Interactions of M Protein with F_0.............. 81

1. Persistent infections and temperature sensitive mutants... 81

2. Amino acid predictions based on sequence analysis.... 82

C. Summary.. 82

REFERENCES.. 83

APPENDIX I : F GENE SEQUENCE COMPARISON AMONG ISOLATES............ 98

APPENDIX II: PREDICTED AMINO ACID SEQUENCE COMPARISON AMONG ISOLATES.. 109

LIST OF FIGURES

Page

1. Morphology of the *Paramyxoviruses*.............................. 3
2. Strategy used for sequence analysis of the HPIV3 M gene........ 33
3. HPIV3 matrix protein gene nucleotide and predicted amino acid sequences.. 35
4. HPIV3 M protein secondary structure prediction................. 40
5. Matrix protein hydropathy profiles of selected *Paramyxoviruses*. 41
6. HPIV3 F mRNA and protein product................................ 52
7. Proposed phylogenetic tree of HPIV3 isolates.................. 69

LIST OF TABLES

Page

1. HPIV3 clinical isolates.................................. 13
2. Amino acid composition of the HPIV3 matrix protein............ 37
3. Comparison of published HPIV3 M gene sequences................ 38
4. Distribution of F gene nucleotide differences among HPIV3 clinical isolates.................................. 58
5. Detailed distribution of F gene nucleotide differences among HPIV3 clinical isolates.................................. 59
6. Accumulation of HPIV3 F gene nucleotide differences........... 60
7. F gene nucleotide differences among HPIV3 clinical isolates... 62
8. Distribution of F protein amino acid differences among HPIV3 clinical isolates.................................. 63
9. Detailed distribution of F protein amino acid differences among HPIV3 clinical isolates.................................. 65
10. F protein amino acid differences among HPIV3 isolates......... 66

LIST OF ABBREVIATIONS

AMV	avian myeloma virus
BPIV3	bovine parainfluenza virus 3
C	nonstructural protein/gene
CDV	canine distemper virus
CPE	cytopathic effect
ddH_2O	double distilled water
DEPC	diethylpyrocarbonate
DI	defective interfering
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
F	fusion gene/protein
FBS	fetal bovine serum
HN	hemagglutinin-neuraminidase gene/protein
HPIV	human parainfluenza virus
L	large gene/protein
LLC-MK2	Lilly Lab culture monkey kidney cells
M	matrix gene/protein
MA-104	Microbiological Associates monkey kidney cells
mCi	millicurie
MEM	minimum essential medium
MOI	multiplicity of infection
NDV	Newcastle disease virus
NP	nucleocapsid gene/protein
P	phosphorylated gene/protein

pfu	plaque forming unit
RF	replicative form
RNP	ribonucleoprotein
rpm	revolutions per minute
RSV	respiratory syncytial virus
SDS	sodium dodecyl sulfate
SSPE	subacute sclerosing panencephalitis
SV5	simian virus 5
TBS	tris buffered saline
ts	temperature sensitive
vol	volume
VSV	vesicular stomatitis virus
wt	weight
2YT	yeast-tryptone medium (Messing, 1983)

Chapter I - GENERAL INTRODUCTION

Some of the background information on human parainfluenza virus 3 (HPIV3) summarized below has previously been discussed in greater detail (Storey, 1987). The roles of the envelope associated matrix (M) and fusion (F) proteins, mentioned briefly in this section, will be discussed in more depth within their respective chapters.

A. Taxonomy

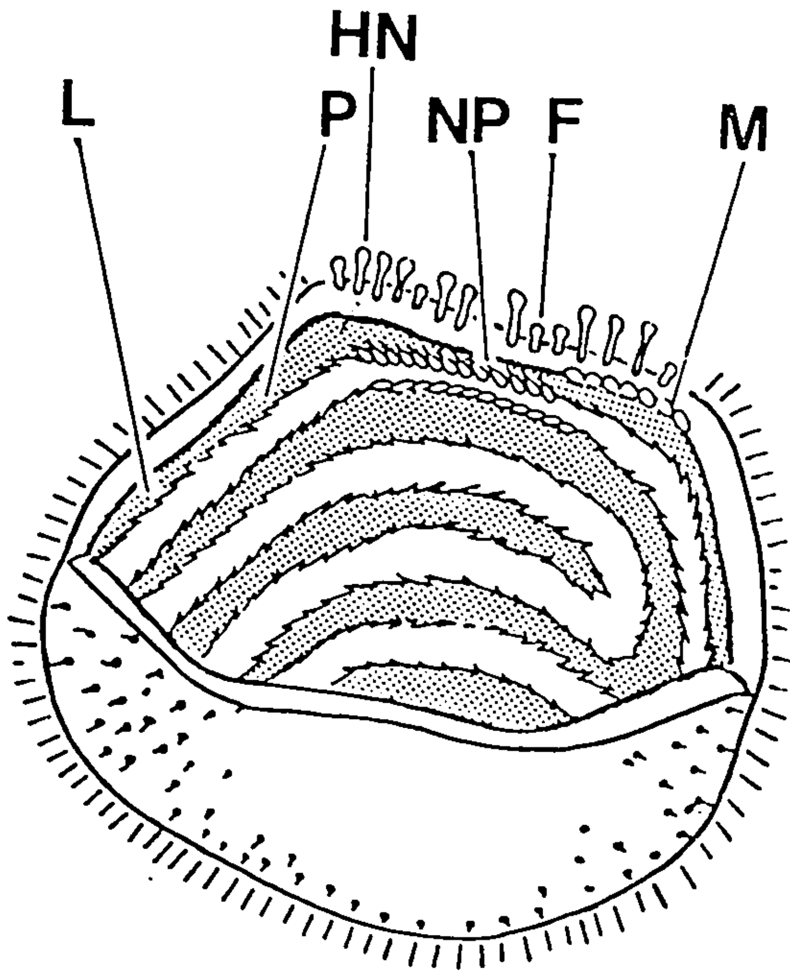
The family Paramyxoviridae is divided into three genera: the Morbilliviruses - canine distemper virus and measles virus; the Pneumoviruses - respiratory syncytial virus; and the Paramyxoviruses- mumps virus, HPIV1-4, Newcastle disease virus (NDV), simian virus 5 (SV5) and Sendai virus (Murphy, 1985). These are distinguished from each other by the following criteria: Paramyxoviruses possess fusion, hemaglutinin and neuraminidase activities while Morbilliviruses lack neuraminidase and Pneumoviruses typically lack both hemaglutinin and neuraminidase activities (Ginsberg, 1988). HPIV3 is distinguished from HPIV types 1, 2 and 4 by antigenic differences.

RNA viruses related to the Paramyxoviridae include the families Rhabdoviridae - vesicular stomatitis virus (VSV) and rabies virus; Orthomyxoviridae - influenza virus, and; Bunyaviridae - California encephalitis virus.

B. Morphology of the Paramyxoviruses

The helical nucleocapsids of Paramyxoviruses (Figure 1) are composed of one single-stranded RNA molecule complexed with the major nucleocapsid (NP) protein, the polymerase or large (L) protein, and a phosphorylated (P) protein thought to be associated with the polymerase (Ginsberg, 1988). This complex is also referred to as the ribonucleoprotein (RNP) (Stone et al., 1972). A viral envelope which consists of a lipid bilayer, originating from the host cell, surrounds the nucleocapsid (Yoshida et al., 1976). Matrix (M) protein lines the inner surface of the viral envelope (Figure 1) and is believed to mediate the association of nucleocapsid with plasma membrane during viral assembly and release (Heggeness et al., 1982). The mature virions of Paramyxoviruses are roughly spherical but vary in size from 150 to 300nm in diameter (Matsumoto, 1982). Two transmembrane surface glycoproteins (Figure 1), the fusion (F) protein and the hemagglutinin-neuraminidase (HN) protein, form spike-like projections on the outer surface of the virion (Lyles, 1979) and are more or less regularly arranged (Ginsberg, 1988). The F protein has been associated with membrane fusion and viral penetration, and the HN protein is believed to mediate virion binding to host cell receptors and release of progeny virus (Scheid and Choppin, 1974; Merz et al., 1980).

Figure 1. Morphology of the Paramyxoviruses (adapted from Kingsbury, 1977). The viral components are labelled as follows: nucleoprotein (NP); large protein (L); phosphoprotein (P); matrix protein (M); fusion protein (F); hemagglutinin-neuraminidase protein (HN).



HN
L
P
NP
F
M

C. Genomic Organization of HPIV3

The genome of HPIV3 consists of single-stranded RNA and is 15,461 nucleotides long (Galinski et al., 1988). Proteins are coded by this negative-sense genome in the order 3' NP-P/C-M-F-HN-L 5' (Dimock et al., 1987; Spriggs and Collins, 1986a). Recent molecular cloning and sequence analysis of HPIV3 have revealed the complete genomic nucleotide order (Côté et al., 1987; Dimock et al., 1986; Elango et al., 1986; Galinski et al., 1986a,b, 1987a,b, 1988; Jambou et al., 1986; Luk et al., 1986, 1987; Prinoski et al., 1987; Sanchez et al., 1986; Spriggs and Collins, 1986a,b; Spriggs et al., 1986, 1987a; Storey et al., 1987).

Common gene-start, intergenic and gene-end sequences exist among the viral genes (Côté, 1989; Spriggs and Collins, 1986a). All five intergenic regions are identical and consist, in viral RNA sense, of the trinucleotide 3' GAA 5' (Spriggs and Collins, 1986a). An exception to this, the P/C-M junction reported to be AAA (Galinski et al., 1986b), has been ascribed to viral variation. In viral RNA sense, all HPIV3 genes start with the sequence 3' UCCUNNUUUC 5' (Spriggs and Collins, 1986a) and a modified version of the gene-start consensus sequence, reading 3' UGGUUUGUUC 5', has also been described at the 3' end of the genome (Dimock et al., 1986). The F and HN genes end with 3' UU(U/A)AUAUUUUUU 5' and all the remaining genes end with 3' UU(U/C)AUUUCUUUUU 5' except for M (to be discussed later). Apparent nucleotide differences from these consensus sequences reported in the gene-end of NP (Galinski et al., 1986b) and the HN start and end

(Elango et al., 1986) may be due to errors occurring during mRNA synthesis, cDNA synthesis, or plasmid repair and replication in bacteria (Spriggs et al., 1986).

Analogous sequences have been observed in Sendai virus (Blumberg et al., 1985; Giorgi et al., 1983; Gupta and Kingsbury, 1982, 1984; Shioda et al., 1983) and VSV (Rose, 1980). It has been hypothesized that the gene-start, end and intergenic regions are control elements for viral transcription (Spriggs and Collins, 1986a). The intergenic regions, with assistance from the preceding U residues, and gene-start sequences may act as transcriptional stop and start commands, respectively. Gene-end sequences are rich in U residues and may be necessary for the addition of 3' poly(A) to mRNAs by repeated copying of this template by the polymerase (Spriggs and Collins, 1986a).

Although the previously described gene-end sequence is present in M, it also contains an additional eight nucleotides (underlined): 3' UUUAUUCUCUAUUAGUUUUU 5'(Spriggs and Collins, 1986a; Galinski et al., 1987a; Luk et al., 1987; Spriggs et al., 1987a; Prinoski et al., 1987). This altered M gene-end sequence may be correlated with a higher incidence of read-through transcription, resulting in an accumulation of M-F bicistronic mRNA (Spriggs and Collins, 1986a; Luk et al., 1987). It appears that the M gene-end sequence is less efficient in its ability to stop transcription due to the presence of these eight nucleotides.

Without going into details, the genomic arrangement of HPIV3 is more or less typical of other related viruses such as Sendai virus, NDV and VSV (Kingsbury, 1985). Usually a pattern of shared gene-start, end

and intergenic sequences is also maintained with genus and family specific variation.

D. Replication of the Paramyxoviruses

Viral adsorption to neuraminic acid residues of host cell receptors is mediated by the HN glycoprotein (Scheid and Choppin, 1974). Membrane fusion and viral penetration depend largely on the F protein in Sendai virus, SV5 and NDV (Homma and Ohuchi, 1973; Scheid and Choppin, 1974). Having entered the cytoplasm, the first event in the replicative cycle of any negative strand virus is primary transcription (Emerson, 1985). A series of mRNA's are transcribed from the viral genome without separation of the genome from its protein coat of NP, L and P (Stone et al., 1972). Viral protein synthesis follows transcription and is an important pre-requisite for viral RNA replication since the addition of cycloheximide inhibits virus replication in infected cells (Robinson, 1971; Carlsen et al., 1985). Also the viral RNA-dependent RNA polymerase cannot simultaneously transcribe and replicate the viral genome (Choppin and Compans, 1975).

After transcription, the viral genome directs synthesis of a positive sense genome length RNA which is used as a template for the synthesis of negative sense genomes (Emerson, 1985). Paramyxoviral assembly begins with the insertion of viral glycoproteins into the host cell membrane (Klenk et al., 1970). Electron microscopy shows that nucleocapsids associate with M protein lining regions of the cell membrane that have viral glycoproteins inserted within them (Ginsberg,

1988). Intact progeny virions can be seen only at the cell membrane where they are assembled and released by budding (Matsumoto, 1982; Ginsberg, 1988).

E. Aspects of HPIV3 Infection and Disease

1. Clinical features

Originally isolated in 1956 from children with acute respiratory disease, HPIV3 ranks second only to RSV as an important cause of respiratory tract infections in young children (Chanock, 1956; Chanock et al., 1958; Parrott et al., 1962; Glezen and Denny, 1973; Chanock and McIntosh, 1985). Symptoms may vary from typical, mild upper respiratory tract infections to less common but severe lower respiratory tract afflictions such as bronchiolitis, pneumonia and croup (Chanock et al., 1958, 1959; Parrott et al., 1962; Glezen and Denny, 1973; Welliver et al., 1982; Chanock and McIntosh, 1985). Although adults are occasionally infected, their symptoms are usually no worse than a mild "cold" or subclinical illness (Ginsberg, 1988). Even among children, most infections with parainfluenza viruses seem to be mild or asymptomatic (Ginsberg, 1988). Approximately 30% of primary HPIV3 infections in children involve the lower respiratory tract and about 5% of these require medical care (Chanock and McIntosh, 1985).

2. Pathogenesis

HPIV3 is transmitted by direct contact or inhalation of respiratory secretions or droplets such as saliva or sputum (Chanock and McIntosh, 1985). The virus, having entered by the respiratory route, will usually multiply and cause inflamation only in the upper regions of the adult respiratory tract (Ginsberg, 1988). These regions include the mucous membranes of the nose and throat (Chanock and McIntosh, 1985). In many patients with mild disease, the bronchi may also be involved to a limited extent (Chanock and McIntosh, 1985). Infection of the bronchi, bronchioles and lungs of infants and young children results in severe disease (Chanock and McIntosh, 1985; Ginsberg, 1988). Viremia, believed to be uncommon and non-essential during parainfluenza virus infection of humans (Gross et al., 1973; Ginsberg, 1988), does occur (Chanock and McIntosh, 1985).

3. Epidemiology

HPIV3 has a world-wide distribution and is endemic, capable of infecting at any time of the year (Chanock and McIntosh, 1985). Primary infections usually occur very early in life (Chanock and McIntosh, 1985; Ginsberg, 1988) since 50% to 60% of children are infected by two years of age and 80% to 100% are infected by the time they are four to six years old (Parrott et al., 1962; Ginsberg, 1988). Higher rates of infection have been reported - 75% to 100% for children in their first two years of life (Glezen et al., 1981; Hope-Simpson, 1981).

Older children and adults are commonly reinfected, though the resulting illness is both less frequent and less severe than a primary infection (Chanock, 1979; Fenner and White, 1976; Wright, 1984). Reinfected individuals shed infectious virus (Chanock and McIntosh, 1985). Of the human parainfluenza viruses, HPIV3 appears to be the most easily spread from person to person (Chanock et al., 1963). Due to its transmissibility and potential for causing re-infection, HPIV3 can create severe problems in partially-enclosed populations such as nurseries and neonatal or pediatric wards (Mufson et al., 1973).

F. Rationale and Objectives

1. Project rationale

Due to its role in serious childhood illness, transmissibility and frequency of recurrence, there is considerable interest in developing an effective HPIV3 vaccine. The M and F proteins of *Paramyxoviruses* have critical functions *in vitro* and *in vivo* in host cell penetration, viral assembly and the spread of infection. Although an increasing amount of information is becoming available on these proteins, certain aspects of their roles in the molecular biology of *Paramyxoviruses* and the biological functions of individual amino acid sequences remain poorly understood. For example, associations have yet to be determined between specific regions of the HPIV3 M protein and the following: nucleocapsids, glycoproteins, membranes and other M protein molecules. There is also a general lack of knowledge regarding the mechanisms by

which hydrophobic sequences within the F protein interact with host and viral membranes.

Unpublished at the time this project was conceived, the HPIV3 M gene sequence was necessary for future experiments concerning the role of matrix protein in assembly. The sequencing was completed on HPIV3 strain 47885 using Sanger's dideoxy method (Prinoski et al., 1987). The HPIV3 M gene sequence has also recently been made available elsewhere (Galinski et al., 1987a; Luk et al., 1987; Spriggs et al., 1987a).

All viral proteins and the individual host immune response likely play some role in influencing the pathogenicity of Paramyxoviruses. However, specific amino acid sequences in the F proteins of Sendai virus (Hsu et al., 1987a,b; Itoh et al., 1987) and NDV (Toyoda et al., 1987; Millar et al., 1988) have been correlated with differences in clinical symptoms or CPE in cell culture. For this reason, clinical isolates of HPIV3 were sequenced across the F gene to determine which regions in the F protein vary in nature and which are conserved.

2. Statement of objectives

This study had two main objectives:

1) to sequence the HPIV3 M gene and flanking sequences for comparison with the known M gene sequences of other Paramyxoviruses, and;
2) to identify sites in the F gene which undergo variation in nature by comparing the F gene sequences of the prototype strain and nine clinical isolates collected from different geographical locations between the years 1959 to 1987.

Chapter II - MATERIALS AND METHODS

A. Cell Culture

1. Cells

HPIV3 was propagated in LLC-MK2 cells - a continuous line derived from rhesus monkey kidneys and received from Flow Laboratories (Rockville, MD). Plaque assays were conducted using MA-104 cells, also a continuous line originating from rhesus monkey kidneys, obtained from Dr. H. Malherbe (Gull Laboratories, Salt Lake City, UT). Monolayers of both cell lines were grown in 60 mm (Nunc/Gibco, Montreal, Quebec) or 100mm polystyrene dishes (Corning Glass Works, Corning, NY) at 37°C in a 5% CO_2 atmosphere in a Shellab incubator (Johns Scientific, Toronto, Ontario). Cell culture medium (1X MEM) was obtained from Gibco Laboratories and consisted of Earle's minimum essential medium and the following supplements: 5% sterile heat-inactivated fetal bovine serum (FBS; Gibco or Bocknek Organic Materials, Rexdale, Ontario), 0.2% (wt/vol) sodium bicarbonate ($NaHCO_3$; Cellgro Mediatech, Washington, DC), 2mM L-glutamine (Cellgro Mediatech) and 50μg/ml gentamycin sulphate (Roussel Laboratories, Wembley, England). Cells to be passaged were washed with 3ml Tris buffered saline (TBS; 137mM NaCl, 5mM KCl, 0.7mM Na_2HPO_4, 5.6mM glucose and 25mM Tris-HCl, pH 7.2) and the monolayer was trypsinized with a mixture of 0.05% trypsin and 0.02% EDTA (Gibco) for 3 minutes at 37°C. The resulting cell suspension was pelleted in a Heraeus Minifuge 2 (Heraeus Christ GMBH, Osterode am

Harz, West Germany) at 5000 rpm for 5 minutes at 20°C, resuspended in fresh 1X MEM and used to seed new culture dishes. Seven new plates were usually seeded with the cells from one confluent 100mm plate (Corning).

2. Virus

HPIV3, originally obtained from the American Type Culture Collection (strain 47885), was provided by D. A. McLeod of the Laboratory Centre for Disease Control, Ottawa, Ontario. Clinical isolates of HPIV3, described in Table 1, were received from Dr. J. Mahoney of St. Joseph's Hospital, Hamilton, Ontario (Ont/489/82a, Ont/12503/82b, Ont/41/83), Dr. M. Menegus of the University of Rochester Medical Center, Rochester, New York, (NY/6508/81, NY/3005/84), Dr. K. van Wyke Coelingh of the National Institutes of Health, Bethesda, MD, (Aus/16796/59, Wash/283/75, Tex/536/80), and Helen Miller of the Children's Hospital of Eastern Ontario, Ottawa, Ontario, (Ont/7175/87). None of the clinical isolates were plaque purified and passage in cell culture was minimized (3-5X).

3. Plaque assay

Confluent MA-104 monolayers grown in 60mm plates yielded more distinct and easily recognizable plaques than monolayers of LLC-MK2 cells. For this reason, MA-104 cells were chosen for use in plaque assays. After 0.3ml of each serial 10 fold dilution of virus had adsorbed for 1.25 hours, the inoculum was removed and the cells were

Table 1: HPIV3 clinical isolates

Isolate	Passage	Site	Patient Age	Patient Sex	Symptoms
Aus/59	11xLLCMK2[a] 3xLLCMK2[b]	?[c]	?[c]	?[c]	?[c]
Wash/75	4xLLCMK2[a] 3xLLCMK2[b]	?[c]	?[c]	?[c]	?[c]
Tex/80	6xLLCMK2[a] 3xLLCMK2[b]	?[c]	32y	?[c]	?[c]
NY/81	?[a,c] 5xLLCMK2[b]	NP[e]	10m[h]	F[f]	pneumonia, bronchitis
Ont/82a	1xGMK?[a,c] 3xLLCMK2[b]	?[c]	?[c]	?[c]	?[c]
Ont/82b	1xGMK[a] 3xLLCMK2[b]	NP	4y[i]	F	fever, croup, acute respiratory distress
Ont/83	1xGMK[a] 3xLLCMK2[b]	throat	5m	M[g]	nasal congestion and discharge, cough
NY/84	?[a,c] 4xLLCMK2	sputum	6m	M	pneumonia, bronchitis
Ont/87	1xHFL f.[a,d] 4xLLCMK2[b]	NP	18m	?[c]	URI[j], cough, fever, irritability

[a]Refers to viral passage history in the lab of origin.
[b]Refers to viral passages during this study.
[c]Indicates questionable or irretrievable information.
[d]Indicates fibroblasts.
[e]NP indicates nasopharynx.
[f]F indicates female, [g]M indicates male.
[h]m indicates month, [i]y indicates year.
[j]URI indicates upper respiratory tract infection.

overlaid with 1X MEM, 0.8% agarose and incubated for 5 days at 37°C. Plaques were visualized by staining with 0.1% crystal violet after the monolayer had been fixed by incubation for 1.5 hours with 3ml of 3.7% formal saline solution (3.7% formaldehyde in 145mM NaCl).

4. HPIV3 infection of cells

Medium was removed from a monolayer of LLC-MK2 cells. Virus was added in a volume of 1ml/100mm plate and allowed to adsorb for 1.25 hours at 37°C with occasional redistribution of the inoculum by tilting of the plates. Fresh medium was added after the inoculum was removed.

5. Preparation of HPIV3 stocks

Virus stocks were prepared by infection of LLC-MK2 monolayers at an MOI of 0.1 pfu/cell. The inoculated cell cultures were incubated for 24 to 48 hours at 37°C prior to harvesting of virus. Culture medium was removed, clarified at 5000 rpm and 4°C for 5 minutes (Heraeus Minifuge 2) and stored at -80°C in 1ml portions.

B. Use of M13

1. Transfection with M13 and subclone amplification

Escherichia coli JM101 cells were inoculated from an M9 plate (Messing, 1983) into 3ml of 2YT medium (Messing, 1983). These cultures

were incubated at 37°C with agitation and grown until the medium and growing cells had reached an absorbance of 0.4 - 0.6 at 260nm. After centrifugation for 5 minutes at 4°C and 5000 rpm (Heraeus Minifuge 2), the cells were gently resuspended in a 0.5 volume of 50mM $CaCl_2$, 20mM Tris, pH 8 and incubated on ice for 30 minutes. The cells were then pelleted again for 5 minutes at 4°C and 5000 rpm and resuspended in a 0.1 volume of 50mM $CaCl_2$, 20mM Tris, pH 8. 10ng of M13 phage DNA was added and a further incubation of 40 minutes on ice was followed by a 1.5 minute heat shock at 42°C. Next, 200μl of freshly growing JM101 cells, 10ul 100mM IPTG, 50μl 2% X-gal in dimethyl-formamide and 3ml 0.6% B agar at 45-50°C were added. Each transfection reaction tube was mixed by gentle inversion and quickly poured over 1.5% B broth plates (Messing, 1983) that had been pre-warmed to 37°C. Petri dishes were inverted after the agar overlay had cooled and then incubated at 37°C for 10-12 hours.

Clear plaques were picked with sterile Pasteur pipettes. These were transferred individually to 50ml Erlenmeyer flasks containing 5ml 2YT (Messing, 1983) and incubated for 8 hours at 37°C with agitation. The culture fluid was transferred into 5ml Sarstedt polystyrene snap cap tubes and centrifuged in a Heraeus Minifuge 2 at 5000 rpm and 4°C for 10-15 minutes. At this stage, the tubes were stored at 4°C.

2. Subclone screening

Subclones, previously generated (Dale et al., 1985), were screened according to the size of genomic DNA by agarose gel electrophoresis.

From each 5ml tube stored at 4°C (mentioned in the preceding section), 15μl of supernatant was mixed with 1μl 10% SDS and 2ul 10X dye solution (50% glycerol, 0.1 M Na_2EDTA, pH 8.0, 1% SDS and 0.1% bromophenol blue). This mixture was warmed briefly at 37°C and electrophoresed on a 0.8% agarose gel at 50 volts for 18 hours. Markers used included M13mp11 phage DNA - with and without the original cloned insert of interest. Electrophoresis was conducted using a MAX Submarine Agarose Gel Unit Model HE 99 (Hoefer Scientific, San Fransico, CA) and a model 103 power pack (E.C. Apparatus, St. Petersburg, FL).

C. Production, Extraction and Preparation of Nucleic Acids

1. HPIV3 viral genomic RNA

Culture medium from 60 plates, infected at an MOI of 3, was removed and clarified by centrifugation (Heraeus Minifuge 2) at 5000 rpm and 4°C for 5 minutes. Virus was pelleted for 1.5 hours at 25000 rpm and 4°C using polyallomer tubes, an SW 28 rotor and a Beckman L8-55M ultracentrifuge. The viral pellets were well drained and tubes wiped dry. Pellets were then resuspended in 200μl of NENS buffer (50mM NaAc, pH 5.1, 0.2M NaCl, 20mM EDTA, 0.5% SDS) containing 200μg/ml of proteinase K and incubated for 30 minutes at 37°C. Phenol, phenol/chloroform/isoamyl alcohol (25:24:1, vol/vol) and chloroform/isoamyl alcohol (24:1, vol/vol) extractions were then performed and the RNA was precipitated overnight with 2.5 volumes of ethanol. The RNA pellet, obtained by centrifugation at 10000 rpm and

4°C for 15 minutes in a Fisher microfuge model 235B, was washed with ethanol and dried in a Speedvac Concentrator model SVC 100H (Savant Instruments, Farmingdale, NY) for 15 minutes. The RNA was aliquoted into 5μg portions at 1μg/ml in DEPC treated water and stored at -80°C in 0.5ml minifuge tubes.

2. M13 single-stranded DNA

M13 genomic DNA was extracted from culture supernatants containing phage subclones of interest. Supernatants were re-centrifuged as previously described and transferred to chilled 30ml Corex tubes, on ice, already containing 1.5ml of 3.5M NH_4Ac and 20% polyethylene glycol (PEG). The combined contents were mixed, left on ice for 30 minutes and then centrifuged at 10000 rpm and 4°C for 15 minutes in a Beckman Model J2-21M centrifuge using a JA-20 rotor. Removal of the supernatant was followed by inversion of the Corex tubes for 1 hour at 4°C to allow all fluid to drain completely. The tube walls were wiped dry and pellets resuspended in 300μl TE buffer containing 20μg/ml of RNase A. Incubation at room temperature for 20 minutes in 1.5ml Eppendorf tubes preceded centrifugation in a Fisher microfuge model 235B for 1 minute. The supernatant was extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1, vol/vol) and once with chloroform/isoamyl alcohol (24:1, vol/vol). DNA was precipitated overnight at -20°C with 2.5 volumes of ethanol and pelleted by the Fisher microfuge in 5 minutes. The DNA pellets were washed once with ethanol, dried for 15 minutes in a Speedvac and resuspended to a concentration of 0.4 mg/ml in ddH_2O.

3. M13 replicative form DNA

A culture of 2YT (2ml) was inoculated with cells from a single plaque and incubated at 37°C with shaking. Simultaneously, a second culture of 2YT (100ml) was inoculated with E. coli JM 101 and incubated in 500ml flasks at 37°C with agitation. After 3 hours, the 2ml culture was added to the 100ml culture and a 500ml culture of 2YT was inoculated with JM 101 cells. After 2 or 3 hours incubation, the two cultures were mixed and grown to stationary phase. The cells were harvested by a 5 minute centrifugation in a Beckman J21 at 4°C and 5000 rpm using a JA-10 rotor, resuspended in 100ml of STE buffer (100mM NaCl, 10mM Tris-HCl pH 8 and 1mM EDTA) and pelleted a second time.

This pellet was resuspended in 9ml of solution I (50mM glucose, 25mM Tris-HCl and 10mM CDTA) and transferred to an SW 28 polyallomer centrifuge tube. 1.0ml of solution I containing 20mg/ml of lysozyme was added and the tube contents were incubated at room temperature for 10 minutes. Next, 10ml of solution II (200mM NaOH, 2% SDS) was added prior to a 10 minute incubation on ice. Finally, 10ml of 8M CH_3COOK were added and the samples were incubated for two hours on ice. This mixture was centrifuged in an SW 28 rotor for 20 minutes at 4°C and 20000 rpm and the supernatant transferred to cold 30ml Corex tubes.

Each tube received 10ml of isopropanol and the samples were mixed and allowed to stand at room temperature. If a precipitate had not formed after 30 minutes, the tubes were placed on ice for 30 minutes. Samples were centrifuged at 15°C and 12,000 rpm for 30 minutes in a Beckman J21 centrifuge. The pellets were rinsed in 75% ethanol, drained

and dissolved to a total volume of 3.8ml in TE (10mM Tris-HCl pH 7.5 and 1mM EDTA). 4g of CsCl in total was added along with 80μl of 10mg/ml ethidium bromide solution for each ml of solution. The samples were centrifuged overnight in a Beckman VTi65 rotor at 20°C and 50000 rpm.

The plasmid DNA band was drawn off and repeatedly extracted with an equal volume of NaCl-saturated isopropanol until the aqueous phase was colourless. Following a final extraction, the aqueous phase was transferred to a 15ml Corex tube, diluted with 5 volumes of sterile ddH_2O and DNA was precipitated with ethanol overnight at -20°C. The DNA was pelleted in a JA-20 rotor spun at 12000 rpm and 4°C for 30 minutes, drained, dissolved in 300μl of TE buffer and transferred to a 1.5ml microfuge tube. The DNA was re-precipitated overnight at -20°C with 150μl of 8M CH_3COONH_4 and 1ml of ethanol and pelleted by centrifugation for 10 minutes in a Fisher microfuge at 4°C. The supernatant was removed and pellets dried in a Speedvac prior to resuspension in TE buffer to a final concentration of 1mg/ml. Samples were then dispensed in 20μl aliquots and stored at -20°C.

D. Nucleotide Sequencing

1. Sanger sequencing of HPIV3 genomic RNA

The Sanger RNA sequencing technique employed consists of a modified version of that published by Geliebter (1987). A 50ng sample of primer in a volume of 1μl was combined with 6μl of primer labelling premix, incubated at 37°C for 30 minutes and then 65°C for 5 minutes.

After both incubations were complete, 3μl DEPC treated water was added to each labelling reaction tube. Primer labelling premix consisted of: 5μl [γ-^{32}P] ATP (3000 Ci/mmol, Amersham Canada Limited, Oakville, Ontario), 0.65ul 10X PKase buffer (500mM Tris-HCl, pH 7.6, 100mM $MgCl_2$, 50mM dithiothreitol, 1mM spermidine, 1mM EDTA) and 0.35ul T4 PKase (10000 U/ml, Pharmacia Canada Inc., Dorval, Quebec). Due to difficulties in measuring out the minute volumes specified for the primer labelling premix, it was found to be more practical to make sufficient premix and label several primers at once.

Annealing reactions were performed by combining 1μl of labelled primer, 5μl RNA (1μg/μl), 1.3 μl 10X annealing buffer (2.5M KCl, 0.1M Tris, pH 8.3) and 5.7 μl DEPC treated water in a 0.5 ml microfuge tube. This mixture was heated at 80°C using a Multi-blok heater, no. 2090 (Lab-Line Instruments Inc., Melrose Park, IL) for 3 minutes and allowed to cool slowly to 40°C by switching off the unit. Both labelled primer and primer/template annealing reactions were stable at -20°C for up to one month.

Sequencing reactions were carried out with a dNTP:ddNTP ratio of 5:1 for dTTP, 10:1 for dATP and dCTP and 20:1 for dGTP. All dNTPs except dGTP achieved a final concentration of 0.21mM (dGTP = 0.42mM) and the concentration of all ddNTPs except for ddTTP equalled 0.021mM (ddTTP = 0.042mM). Dideoxy premixes contained 1.39mM of all dNTPs except dGTP which was present at 2.78mM, and 0.139mM of all ddNTPs except ddTTP which was present at 0.278mM. Sequencing premix consisted of 45.3μl 4X reverse transcriptase buffer (400mM Tris-HCl pH 8.3, 520mM KCl, 40mM $MgCl_2$, 100mM dithiothreitol), 12.9μg/100μl actinomycin D and

165U/100μl AMV reverse transcriptase in a final volume of 100μl. Reactions were initiated by combining 1.5μl of the appropriate dideoxy premix with 3μl of the annealing reaction and 5.4μl of sequencing premix. Samples were incubated at 42°C for 1.25 hours and nucleic acid precipitated in 2.5 volumes of ethanol for at least 0.5 hours at -20°C. The sequencing reaction tubes were centrifuged in a Fisher microfuge model 235B for 15 minutes, the pellets washed with ethanol and dried using a Speedvac. Each pellet was resuspended in 6.1μl of dye solution (99% deionized formamide, 10mM EDTA, 10mM NaOH, 0.3% wt/vol bromophenol blue, 0.3% wt/vol orange G and 0.3% wt/vol xylene cyanol).

2. Sanger sequencing of cDNA clones

All work was conducted using 0.5ml microfuge tubes. For each M13 clone to be sequenced 4μl of annealing mix was combined with 5μl M13 ssDNA at 0.2μg/μl and incubated first at 90°C for 2 minutes, then at 55°C for 10 minutes. Annealing mix consisted of 0.9μl 100mM Tris-HCl, pH 8, 100mM $MgCl_2$, 2μl 17-mer primer at 2μg/μl (No. 1211, New England Biolabs) and 1.1μl H_2O. Annealing was completed by allowing the temperature to fall slowly to 35°C, usually within an hour. For each sample, a 2μl aliquot of annealing mixture was added to 3μl of sequencing mix and incubated at 42°C for 30 minutes. Sequencing mix consisted of 0.65μl [α-^{35}S] dATP (400 Ci/mmol, Amersham), 0.25μl 100mM Tris-HCl, pH 8 and 100mM $MgCl_2$ solution, 0.5U Pol 1 "Klenow" (5U/ul, Pharmacia) and 2μl of the desired ddNTP/dNTP pre-mix. The dd/dNTP pre-mixes were prepared as follows: pre-mix A - 30μl 0.1 mM ddATP, 12.5μl

0.5mM dCTP, 12.5μl 0.5mM dGTP, 12.5μl 0.5mM dTTP, 10μl 50mM Tris-HCl, pH 8 and 1mM EDTA, 22.5μl ddH_2O; pre-mix C - 10μl 1.0mM ddCTP, 1.7μl 0.5mM dCTP, 17.5μl 0.5mM dGTP, 17.5μl 0.5mM dTTP, 10μl 50mM Tris-HCl, pH 8 and 1mM EDTA, 43.3μl ddH_2O; pre-mix G - 5μl 10.0mM ddGTP, 17.5μl 0.5mM dCTP, 1.7μl 0.5mM dGTP, 17.5μl 0.5mM dTTP, 10.0μl 50mM Tris-HCl, pH 8 and 1mM EDTA, 48.3μl ddH_2O; and pre-mix T - 15μl 10mM ddTTP, 17.5μl 0.5mM dCTP, 17.5μl 0.5mM dGTP, 1.7μl 0.5mM dTTP, 10μl 50mM Tris-HCl, pH 8 and 1mM EDTA, 38.3μl ddH_2O. Incubation was continued for an additional 20 minutes at 42°C after adding 2μl of Chase mix (20μl dNTP mix, 18μl ddH_2O, 4μl 100mM Tris-HCl, pH 8 and 100mM $MgCl_2$ solution) to each reaction tube. The dNTP mix (40μl of each 10mM dNTP, 20μl 50mM Tris-HCl, pH 8 and 1mM EDTA solution, 20μl ddH_2O) was prepared in advance, aliquoted and stored at -20°C. The reactions were stopped by the addition of 7μl of standard formamide mix and each tube was incubated at 95°C for 3 minutes immediately prior to electrophoresis.

3. Maxam and Gilbert sequencing

Despite a previous description of this technique (Maxam and Gilbert, 1980), it is also described in this thesis due to changes made in the protocol.

M13 replicative form was labelled (Maniatis et al., 1982) using Pol 1 "Klenow" (5U/μl, Pharmacia) and the appropriate mixture of [α-^{32}P] labelled and unlabelled dNTP's (800 Ci/mmol, Amersham). The [α-^{32}P] dNTP used depended on which restriction enzyme site was cleaved (indicated in Chapter 3). Sample DNA was resuspended to a volume of

25μl with ddH_2O and combined with 5μl salmon sperm DNA (1μg/μl). This was divided into 1.5ml microfuge tubes and the following additions made: the TC sample received 10μl DNA and 15μl ddH_2O; the C sample received 5μl DNA and 20μl saturated NaCl; the GA sample received only 10μl of DNA, and; the G sample received 5μl of DNA and 200μl DMS buffer (50mM Sodium cacodylate, 10mM $MgCl_2$ and 1mM EDTA, pH 8.0). All reactions were carried out at room temperature.

The TC and C reaction tubes each received 25μl of hydrazine and were allowed to sit for 7 minutes at which time 200μl A+C+T stop solution (0.35M CH_3COONa pH 8.5-9.5, 0.1mM EDTA, 150μg/ml wheat germ tRNA) and 700μl of ethanol was added. 25μl of formic acid was added to the GA reaction tube and the sample was incubated for 3 minutes prior to the addition of 200μl A+C+T stop solution and 700μl of ethanol. The G reaction tube and 1μl DMS (dimethyl sulfate) were incubated for 3 minutes and the reaction stopped with the addition of 50μl G stop solution (1.5M CH_3COONa, 1.0M beta-mercaptoethanol, 600μg/ml wheat germ tRNA). Once these steps were completed, each reaction tube was manipulated as described below.

The reaction tubes were collectively incubated for at least 30 minutes in a dry ice/ethanol bath, centrifuged for 15 minutes at 4°C in a microfuge and the pellets resuspended in 200μl of 0.3M CH_3COONa, pH 9.3 and 1mM EDTA, pH 7.4. After the addition of 600μl ethanol, the samples were incubated in a dry ice/ethanol bath for 30 minutes, centrifuged and the pellets rinsed with ethanol and dried using a Speedvac. At this point, 50μl 10% (v/v) piperidine (Aldrich Chemical Co., Milwaukee, WI) was added and the tubes incubated for 15 minutes at

90°C. Following this, the reaction tube contents were dried in a Speedvac twice for 100 and 30 minutes, respectively, after the separate additions of 100μl and 20μl of ddH_2O. Prior to loading, 10μl of formamide dye solution was added to each sample and the reaction tubes were heated for 5 minutes at 100°C and cooled on ice.

4. Electrophoresis and autoradiography

For M gene sequencing, the outer and inner glass plates (International Biotechnologies Canada Inc., Toronto, Ontario) were cleaned and washed with ethanol. The gel side of the outer plate was treated with dimethyldichlorosilane solution (BDH Chemicals Ltd., Poole, England), washed three times with ddH_2O and then twice with ethanol. Silane adhesion promoter (BDH Chemicals Ltd) was used to treat the gel side of the inner glass plate (10ml ethanol, 30μl glacial acetic acid, 30μl silane). Three ethanol washes were necessary to remove excess silane. A 0.2mm comb and spacer set (IBI) was used to separate the two glass plates, forming the gel sandwich which was taped and clamped. 8% acrylamide gels (18.48g urea, 3.34g acrylamide, 0.18g N,N'-methylene-bis-acrylamide, 4.4ml 10X TBE, 66μl 10% ammonium persulfate, 20μl TEMED, 22.5ml ddH_20) were poured and after polymerization was complete, the samples were loaded. Sequencing gels were run at 50W constant power (30mA and 2500V were set as upper limits for amperage and voltage) using an MBP 3000 power pack (IBI). For Maxam and Gilbert sequencing it was necessary to run samples on 20% acrylamide gels. Samples were loaded in a volume of 2.5μl and run for 3

to 5 hours using 1X TBE as the running buffer. Once electrophoresis was complete and the glass plates pried apart, the gel was fixed in 5% glacial acetic acid and 5% methanol for 40 minutes and left overnight to air dry. Cronex 4 X-ray film (E.I. Dupont de Nemours, Wilmington, DE) was exposed to the dried gel for 48 to 72 hours at -80°C.

Alternatively, for F gene sequencing, silane A-174 was not used and both plates were treated with dimethylchlorosilane solution. Once electrophoresis was complete, the gel was lifted from the plate with a sheet of previously developed X-ray film, wrapped in all-purpose polyvinyl chloride laboratory wrap and exposed at -80°C to X-ray film in a film cassette containing intensifying screens. Exposure times varied between 3 and 24 hours depending on the nature of the primer itself, the age of the isotope and the amount loaded.

Chapter III - SEQUENCING OF THE HPIV3 M GENE

A. Introduction

Purified M protein self-assembles into ordered arrays (Heggness et al., 1982), resembling pseudocrystalline structures associated with the cytoplasmic surface of membranes of Paramyxoviral infected cells (Büechi and Bächi, 1982; Heggness et al., 1982). The M protein associates with the inner surface of the Paramyxoviral envelope, possibly enhancing its structural integrity (Ginsberg, 1988). More importantly, matrix proteins may function in assembly as a link between cytoplasmic viral nucleocapsids and the lipophilic inner surface of the modified host cell plasma membrane (Blumberg et al., 1984). Matrix proteins likely have at least two functional domains: one to associate with lipid bilayers and the other with nucleocapsids (Ogden et al., 1986; Pepinsky and Vogt, 1979, 1984; Lamb and Choppin, 1983).

Matrix proteins of the Paramyxoviridae have been implicated in two major activities: viral assembly and the regulation of RNA synthesis. M protein may also influence viral pathogenesis. Due to the importance of the matrix protein, the HPIV3 M gene and flanking sequences were determined for comparison with the M genes of other Paramyxoviridae.

1. Virus maturation: assembly and budding

Although the exact mechanism is not clear, M protein is believed to play an essential role in the maturation of Paramyxoviruses and

other negative strand RNA viruses. The rate-limiting step in virus assembly is likely M protein synthesis, since it is a major protein in virions but not in the cytoplasm of infected cells (Clinton et al., 1978). Matrix protein mutants of VSV are defective in viral assembly (Weiss and Bennett, 1980; Ono et al., 1987). Also, influenza virions unable to produce progeny were discovered to be restricted in their ability to synthesize M protein (Lohmeyer et al., 1979).

The mechanism of transport of M proteins to the cell surface is not known, although varying theories and experimental results concerning this subject have been reported. After synthesis on free ribosomes, M protein can be detected by immunofluorescence in a random pattern throughout cells, about three hours after inoculation with VSV (Dubois-Dalcq et al., 1984). As viral replication progresses, M protein becomes more concentrated at regions of the cell membrane destined for virion budding (Dubois-Dalcq et al., 1984). However, pulse-chase experiments on cells infected with Newcastle disease virus showed that M protein is incorporated into the plasma membrane immediately after synthesis (Nagai et al., 1976a). M proteins may reach the plasmalemma by simple diffusion since they are nonglycosylated and do not pass through the modification pathway postulated for surface glycoproteins (Strauss and Strauss, 1985). An association between nucleocapsids and matrix protein may be necessary to protect M protein from degradation before reaching the plasma membrane (Tuffereau and Roux, 1988). However, it has also been proposed that nucleocapsids only associate with M protein that is already clustered in virus specific regions of the host cell membrane (Yoshida et al., 1979).

Usually the ratio of matrix protein to nucleocapsid protein is unique to each member of a given genus (Lodish and Porter, 1980; Strauss and Strauss, 1985). This relationship does not extend to the transmembrane glycoproteins, possibly indicating a less specific association between M protein and glycoproteins than between M protein and nucleocapsids (Lodish and Porter, 1980). The difference in specificity is evident in pseudotypes formed during mixed infections by various enveloped viruses with M proteins (Choppin and Compans, 1970; Zavada, 1976). Pseudotypes contain the nucleocapsid and M protein of one type of virus but the surface glycoproteins of a second virus. Since M proteins and nucleocapsids from different viruses apparently do not mix, it seems likely that the interaction of M protein with nucleocapsids is more specific than with the surface glycoproteins (Zavada, 1976; Simons and Garoff, 1980).

Viral glycoproteins are synthesized on membrane-bound ribosomes and transferred to the plasma membrane after modifications in the endoplasmic reticulum and Golgi apparatus (Strauss and Strauss, 1985). Matrix protein may be essential to localize viral glycoproteins in certain areas of the plasmalemma, creating a virus-specific membrane (Yoshida et al., 1979). Fractionation studies on NDV infected cells indicate that the association of M protein with the envelope and localization of viral glycoproteins are closely related events (Nagai et al., 1976a). However, it was not determined if localization of the glycoproteins was required for attachment of M or if insertion of M promoted glycoprotein aggregation (Nagai et al., 1976a). Recent work suggests the latter: that restricted cell surface expression of HN in

persistent Sendai virus infections is due to degradation of the M protein prior to reaching the membrane (Tuffereau and Roux, 1988).

Smooth membranes have little M protein associated with them and glycoproteins contained within have a high degree of mobility and random distribution (Nagai et al., 1975). Glycoproteins in a mature viral membrane appear to have a low degree of mobility (Nagai et al., 1975). Further evidence of the interaction between matrix proteins and surface glycoproteins is that Paramyxovirus M protein mutants have less M protein and are of reduced infectivity (Peeples and Bratt, 1984). Matrix proteins have also been reported to interact with cellular actin which is associated with virions, but the role of the cytoskeleton in viral assembly is unclear (Damsky et al., 1977; Griffin and Compans, 1979; Genty and Bussereau, 1980; Giuffre et al., 1982).

In addition to interactions with glycoproteins, the matrix protein is believed to interact hydrophobically with lipids in the plasma membrane (Heggeness et al., 1982; Zakowski and Wagner, 1980), specifically with the hydrocarbon portions of the membrane lipids (Gregoriades and Frangione, 1981). The hydrophobicity of some M proteins is illustrated by extraction into acidified chloroform-methanol and solubilization into aqueous solution through the agency of high ionic strength salts (Gregoriades, 1973; Wechsler et al., 1985). The M proteins of influenza and VSV are capable of direct interaction with phospholipid bilayers in the absence of other proteins (Bucher et al., 1980; Zakowski et al., 1981). Membrane fluidity may be altered because of interaction between the head groups on phospholipids and the M protein (Wiener et al., 1983) and these changes in fluidity may

promote the inclusion of viral glycoproteins in the budding virion (Chambers et al., 1986).

Nucleocapsid proteins of Paramyxoviruses carry overall acidic charges but domains which interact with the sugar-phosphate backbone of nucleic acids are basic (Chambers et al., 1986). Therefore, the remaining exposed regions on the surface of the nucleocapsid protein are likely to be negatively charged, enabling the M protein by virtue of its net positive charge to bind the viral nucleocapsid protein subunits during viral assembly (Chambers et al., 1986).

Extractions of VSV in the presence of NaCl yielded highly extended viral nucleocapsids in which M protein was not detected (Newcomb and Brown, 1981). However, extractions performed in the absence of salt resulted in nucleocapsids that appeared to be tightly coiled (Newcomb and Brown, 1981). Dialysis to remove NaCl allowed reassociation with M protein and resulted in condensation of nucleocapsids (Newcomb et al., 1982). Electron microscopy of Sendai virus revealed that Triton X-100 treatment with or without 1M KCl resulted, respectively, in straightened nucleocapsids that had a uniform width of 18-20nm or those which were folded into a spherical shape (Shimizu and Ishida, 1975). These results collectively suggest that an ionic interaction of M protein with the nucleocapsid initiates the coiling process that is a pre-requisite for the assembly and budding of the Paramyxoviridae.

2. Regulation of RNA synthesis

Purified Sendai virus nucleocapsids had an associated RNA-dependent transcriptase activity nine times greater than that assayed in unfractionated detergent-disrupted virions (Marx et al., 1974). The inhibitory effect of either the matrix protein or the glycoproteins was overcome by heat denaturation (Marx et al., 1974). Transcription by VSV ribonucleoprotein cores containing only the L, N and NS proteins was inhibited by purified VSV M protein (Carroll and Wagner, 1979). Matrix protein from VSV mutants which produce a defective M protein had a much weaker inhibitory effect on transcriptase activity than M protein from the wild-type (Wilson and Lenard, 1981). VSV M protein is predicted to have a highly basic amino terminal sequence due to the presence of eight lysine residues in the first 19 positions and the absence of any other charged residues in this region (Rose and Gallione, 1981). This basic domain might affect VSV transcription by interactions with genomic RNA (Rose and Gallione, 1981) or RNP. In addition to inhibiting transcription directly, VSV M protein also functions to attenuate transcription relative to distance from the site of transcriptional initiation (Clinton et al., 1978).

3. Relevance of the M protein to Paramyxoviral pathogenesis

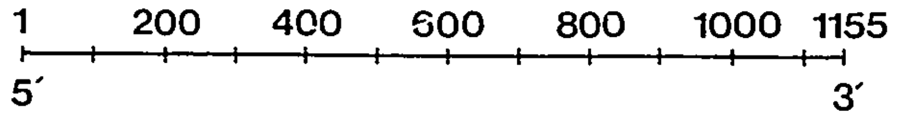
Viral pathogenesis may be influenced by the matrix protein since in persistently infected cells, reduced levels of M protein have been correlated with decreased viral budding (Roux et al., 1984, 1985).

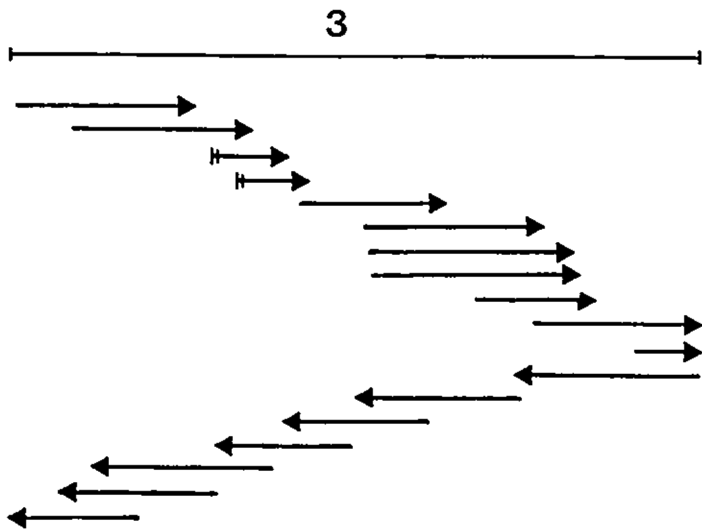
Persistent infections with Sendai virus are due to the production of an unstable M protein (Roux and Waldvogel, 1982). Defective budding leading to persistent infection may also result from decreased levels of measles virus M protein, as implicated in subacute sclerosing panencephalitis - a chronic neurological disorder (Baczko et al., 1986; Cattaneo et al., 1986; Liebert et al., 1986; Cattaneo et al., 1987). Persistent infections involving Paramyxoviruses may also be associated with multiple sclerosis (Morgan and Rapp, 1977; Goswani et al., 1984) and Paget's bone disease (Mills et al., 1981; Basle et al., 1985). More work is needed to confirm these associations, as well as any role the M protein may have in these diseases due to its presence or absence.

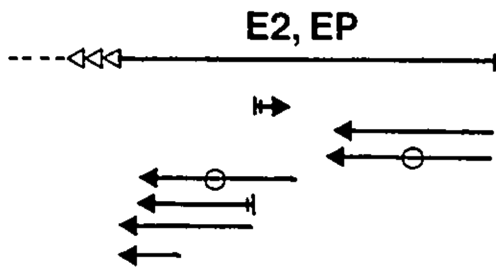
B. Results

Three HPIV3 specific clones (Figure 2) were chosen for sequence analysis. Clone pPI3 includes nucleotides 187 to 1155 of the M gene and was identified in a cDNA bank derived from mRNA (Dimock et al., 1987). Clones mpPIgE2 and mpPIgEP both extend from Eco RI and Pst I sites, respectively, in the adjacent P gene (Galinski et al., 1986b; Luk et al., 1986; Spriggs and Collins, 1986b) to nucleotide 522 of the M gene. Both clones were derived from clone mpPIgMP which was identified in a genomic cDNA bank by hybridization to M gene sequences and P gene sequences (Dimock et al., 1987). Dideoxynucleotide sequence analysis was carried out in both directions for most of the M gene. Sequential subclones of pPI3, mpPIgE2 and mpPIgEP, for use in Sanger sequencing of the M gene, were generated by the method of Dale et al. (1985). Sanger

Figure 2. Strategy used for sequence analysis of the HPIV3 M gene. Clones pPI3 (3), mpPIgE2 (E2) and mpPIgEP (EP) are aligned to the numbered segment which represents the M gene in message sense. Terminally double hatched lines represent regions sequenced by the Maxam and Gilbert method. Lines with superimposed circles indicate subclones of EP. The continuation of E2 and EP into the P gene is shown by the outlined arrow heads and dotted line. Solid arrow heads indicate the direction of sequencing. The length of the line represents the distance sequenced in the subclones of each clone.


1
200
400
600
800
1000
1155
5´
3´


3


E2, EP

sequencing in both directions was occasionally difficult, so the following regions were verified by Maxam and Gilbert sequencing (brackets indicate restriction enzyme site, location and [α-^{32}P] dNTP used): 22 to 183 (Bam HI, 183-188, dTTP), 186 to 233 (Bam HI, 183-188, dTTP), 472 to 575 (Dde I, 465-469, dTTP) and 504 to 601 (Taq I, 484-487, dCTP). Nucleotides 1 to 27 have been previously confirmed in primer extension experiments (Côté, unpublished) and nucleotides 184 and 185 lie within the internal Bam H1 site.

The HPIV3 M gene (Figure 3) was determined to be 1155 nucleotides long. Proposed transcriptional start and stop sequences have been identified at regions 1 to 10 and 1136 to 1155, inclusive. An open reading frame from nucleotide 33 to 1091 codes for a protein with a predicted length of 353 amino acids and a molecular weight of 39498. The predicted amino acid composition is summarized in Table 2.

The HPIV3 M gene sequences reported (Prinoski et al., 1987; Galinski et al., 1987a; Luk et al., 1987; Spriggs et al., 1987a) are almost identical, with only a few differences (Table 3). M gene sequences are also available for Sendai virus (Blumberg et al., 1984; Hidaka et al., 1984), NDV (Chambers et al., 1986; McGinnes and Morrison, 1987; Sato et al., 1987), RSV (Satake and Venkatesan, 1984), measles virus (Bellini et al., 1986), CDV (Bellini et al., 1986), BPIV3 (Sakai et al., 1987), rabies virus (Rayssiguier et al., 1986), VSV (Rose and Gallione, 1981; Morita et al., 1987), and influenza virus (Lamb and Lai, 1981; Lamb et al., 1981).

The HPIV3 M gene and protein product were compared to those of other negative stranded RNA viruses using the Pustell Sequence Analysis

Figure 3. HPIV3 matrix protein gene nucleotide and predicted amino acid sequences. The complete M gene nucleotide sequence is shown as DNA in the mRNA sense. These sequences are aligned such that each amino acid is positioned directly below the middle nucleotide of the respective codon. Nucleotide position is numbered on the left and amino acid position on the right. Proposed transcriptional start and stop sequences of the M gene are underlined. The translational stop signal is indicated at the end of the amino acid sequence by a hyphen.

```
1   AGGATTAAAG AATAAATTAA TCCTTGTCCA AAATGAGTAT AACTAACTCT
                                          M  S  I   T  N  S   6

51  GCAATATACA CATTCCCAGA GTCATCATTC TCTGAGAATG GTCATATAGA
     A  I  Y   T  F  P  E   S  S  F   S  E  N   G  H  I  E  23

101 ACCATTACCA CTCAAAGTCA ATGAACAGAG AAAAGCAGTA CCTCACATTA
     P  L  P   L  K  V   N  E  Q  R   K  A  V   P  H  I     39

151 GAGTTGCCAA AATCGGAAAT CCACCAAAAC ACGGATCCCG GTATTTGGAT
    R  V  A  K   I  G  N   P  P  K   H  G  S  R   Y  L  D   56

201 GTCTTCTTAC TCGGCTTCTT CGAGATGGAA CGAATCAAAG ACAAATACGG
     V  F  L   L  G  F  F   E  M  E   R  I  K   D  K  Y  G  73

251 GAGTGTGAAT GATCTTGACA GTGACCCGGG TTACAAAGTT TGTGGCTCTG
     S  V  N   D  L  D   S  D  P  G   Y  K  V   C  G  S     89

301 GATCATTACC AATCGGATTA GCCAAATACA CTGGGAATGA CCAGGAATTA
    G  S  L  P   I  G  L   A  K  Y   T  G  N  D   Q  E  L   106

351 TTACAGGCTG CAACTAAACT GGATATAGAA GTGAGAAGAA CAGTTAAAGC
     L  Q  A   A  T  K  L   D  I  E   V  R  R   T  V  K  A  123

401 GAAAGAAATG ATTGTTTATA CGGTACAAAA TATAAAACCA GAACTGTACC
     K  E  M   I  V  Y   T  V  Q  N   I  K  P   E  L  Y     139

451 CATGGTCCAG TAGACTAAGA AAAGGAATGT TGTTCGATGC CAACAAAGTT
    P  W  S  S   R  L  R   K  G  M   L  F  D  A   N  K  V   156

501 GCTCTTGCTC CTCAATGTCT TCCACTAGAT AGGAGCATAA AATTCAGAGT
     A  L  A   P  Q  C  L   P  L  D   R  S  I   K  F  R  V  173

551 AATCTTCGTT AATTGTACGG CAATTGGATC AATAACCTTG TTCAAAATTC
     I  F  V   N  C  T   A  I  G  S   I  T  L   F  K  I     189

601 CTAAGTCAAT GGCATCACTA TCTCTACCCA GCACAATATC AATCAATCTG
    P  K  S  M   A  S  L   S  L  P   S  T  I  S   I  N  L   206

651 CAGGTACACA TCAAAACAGG GGTTCAGACT GATTCTAAAG GGATAGTTCA
     Q  V  H   I  K  T  G   V  Q  T   D  S  K   G  I  V  Q  223

701 AATTTTGGAT GAGAAGGGTG AAAAATCACT GAATTTCATG GTCCATCTCG
     I  L  D   E  K  G   E  K  S  L   N  F  M   V  H  L     239

751 GATTGATCAA AAGAAAAGTA GGCAGAATGT ACTCTGTCGA GTACTGTAAA
    G  L  I  K   R  K  V   G  R  M   Y  S  V  E   Y  C  K   256

801 CAGAAAATCG AGAAAATGAG ATTGATATTT TCTTTGGGAT TAGTTGGAGG
     Q  K  I   E  K  M  R   L  I  F   S  L  G   L  V  G  G  273
```

```
851  AATCAGTCTT CATGTCAATG CAACTGGATC TATATCAAAA ACACTAGCAA
      I  S  L   H  V  N   A  T  G  S   I  S  K   T  L  A    289

901  GTCAGCTGGT ATTCAAAAGG GAGATTTGTT ATCCCTTAAT GGATCTAAAT
     S  Q  L  V   F  K  R   E  I  C   Y  P  L  M   D  L  N  306

951  CCACATCTCA ATCTAGTTAT CTGGGCTTCA TCAGTAGAGA TTACAAGAGT
      P  H  L   N  L  V  I   W  A  S   S  V  E   I  T  R  V 323

1001 GGATGCAATT TTCCAACCTT CTTTACCTGG CGAGTTCAGA TACTATCCTA
       D  A  I   F  Q  P   S  L  P  G   E  F  R   Y  Y  P   339

1051 ACATTATTGC AAAAGGAGTT GGGAAAATCA AACAATGGAA CTAGTAATCT
     N  I  I  A   K  G  V   G  K  I   K  Q  W  N   -         353

1101 CTATTTTGAT CTGGATATAT CTATTAAGCC AAAGCAAATA AGAGATAATC

1151 AAAAA
```

Table 2: Amino acid composition of the HPIV3 matrix protein

amino acid[a]	number	percent
nonpolar		
ala (A)	17	4.8
pro (P)	19	5.4
val (V)	26	7.4
ile (I)	32	9.1
leu (L)	36	10.2
met (M)	9	2.5
phe (F)	14	4.0
trp (W)	3	0.8
polar		
gly (G)	24	6.8
ser (S)	30	8.5
thr (T)	14	4.0
cys (C)	5	1.4
asn (N)	16	4.5
gln (Q)	12	3.4
tyr (Y)	12	3.4
acidic		
asp (D)	13	3.7
glu (E)	17	4.8
basic		
lys (K)	31	8.8
his (H)	7	2.0
arg (R)	16	4.5

[a]amino acids within each category are listed according to molecular weight - smallest to largest (Lehninger, 1978). Leucine and isoleucine have the same molecular weight.

Table 3: Comparison of published HPIV3 M gene sequences[a].

nuc. number	nuc. K.P.	nuc. M.S.	nuc. M.G.	nuc. D.L.	a.a. number	a.a. K.P.	a.a. M.S.	a.a. M.G.	a.a. D.L.
68	A	G	G	G	12	Pro	Pro	Pro	Pro
180	C	C	T	C	50	His	His	Tyr	His
182	C	T	T	T	50	His	His	Tyr	His
271	G	A	G	G	80	Ser	Asn	Ser	Ser
322	C	T	C	C	97	Ala	Val	Ala	Ala
323	C	T	C	C	97	Ala	Val	Ala	Ala
339	G	A	G	G	103	Asp	Ile	Asp	Asp
340	A	T	A	A	103	Asp	Ile	Asp	Asp
359	T	A	T	T	109	Ala	Ala	Ala	Ala
374	T	C	C	C	114	Asp	Asp	Asp	Asp
395	T	C	C	C	121	Val	Val	Val	Val
447	T	T	C	T	139	Tyr	Tyr	His	Tyr
593	C	T	T	T	187	Phe	Phe	Phe	Phe
602	T	C	C	C	190	Pro	Pro	Pro	Pro
841	T	C	T	T	270	Leu	Ser	Leu	Leu

[a]Abbrieviations in this table are used to designate the following: nucleotide (nuc.), amino acid (a.a.), Prinoski et al., 1987 (K.P.), Spriggs et al., 1987a (M.S.) and Galinski et al., 1987a (M.G.), Luk et al., 1987 (D.L.).

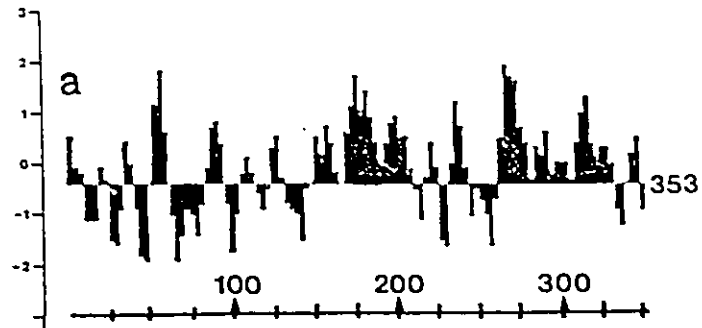
Programs (International Biotechnologies Inc., New Haven, CT.). Matrix gene sequences selected for comparison included BPIV3 (Sakai et al., 1987), Sendai virus (Blumberg et al., 1984), measles virus (Bellini et al., 1986), CDV (Bellini et al., 1986), NDV (Chambers et al., 1986), VSV (Rose and Gallione, 1981), RSV (Satake and Venkatesan, 1984; Collins et al., 1986), rabies virus (Rayssiguier et al., 1986) and influenza virus (Lamb et al., 1981). The M gene of HPIV3 was most similar to BPIV3 (78.7%) and Sendai virus (57.9%), although similarities of 34.5% and 34.3% were found for CDV and NDV, respectively. The remaining M genes varied from 29.7% similarity for measles virus to 19.4% for VSV. Significant percentages of similarity at the protein level were found with BPIV3 (88.4%), Sendai virus (61.8%), measles virus (33.4%), CDV (30.6%) and NDV (20.7%). The M proteins from the remaining viruses listed above varied from 8.5% to 11.0%. A prediction of the HPIV3 matrix protein secondary structure (Figure 4) was produced from an algorithim (Garnier et al., 1978) available on PC Gene - microcomputer software for protein chemists and molecular biologists (IntelliGenetics, Mountain View, CA).

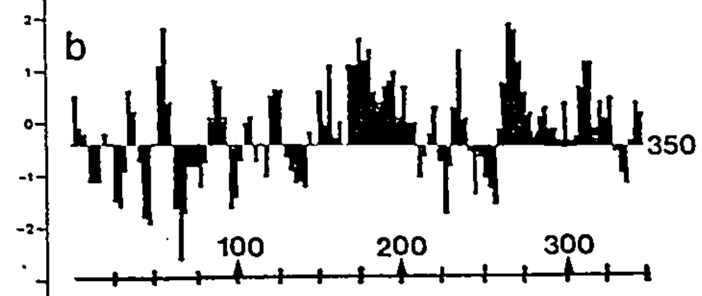
The M proteins from each of these ten viruses are predicted to have the following charges at pH 6 to 7: HPIV3 +20.5; BPIV3 +18.0; Sendai virus +19.0; measles virus +10.5; CDV +8.5; NDV +19.5; RSV +5.5; VSV +9.0; rabies virus -1.5; and influenza virus +8.5. Hydropathy profiles (Kyte and Dolittle, 1982) for each M protein were generated using the software of Pustell (IBI). Selected plots are shown in Figure 5. Virtually identical to BPIV3 and Sendai virus, the hydropathic profile of the HPIV3 matrix protein is also quite similar to those of measles

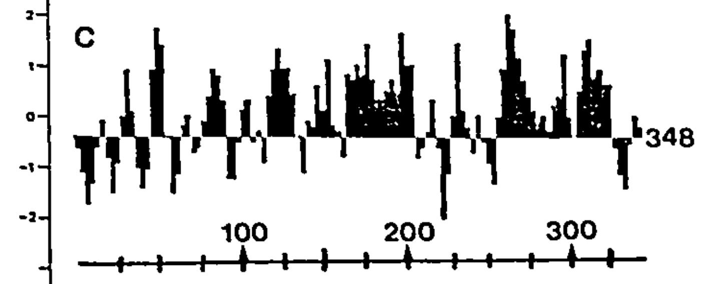
Figure 4. HPIV3 M protein secondary structure prediction. Predictions of the amino acid sequence and resulting secondary structure of the HPIV3 M protein are shown in upper and lower case letters, respectively. Previously specified letters (Garnier et al., 1978) have been used to designate secondary structure: helical conformation (h), extended conformation (e), turn conformation (t) and coil conformation (c).

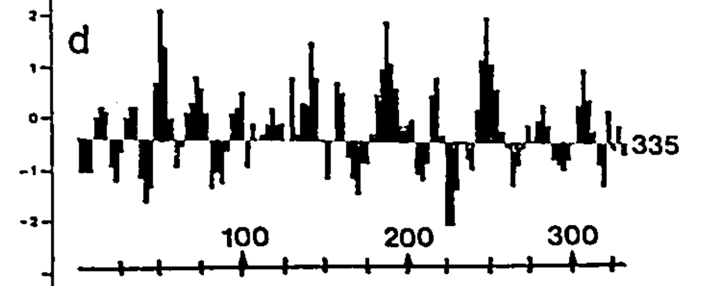
```
        10        20        30        40        50        60
MSITTNSAIYTFPESSFSENGHIEPLPLKVNEQRKAVPHIRVAKIGNPPKHGSRYLDVFLL

eeeeeeeeeeeeetcchcttttcehhhhhhhhhhhhhhheeeeeetcccctttteeeeeeeehh


        70        80        90       100       110       120
GFFEMERIKDKYGSVNDLDSDPGYKVCGSGSLPIGLAKYTGNDQELLQAATKLDIEVRRT

hhhhhhhhhhtteeeeeettcttteeeeeeeeceeeeeeeeeccchhhhhhhhhhhhhhhh


       130       140       150       160       170       180
VKAKEMIVYTVQNIKPELYPWSSRLRKGMLFDANKVALAPQCLPLDRSIKFRVIFVNCTA

hhhhhhehhhheceeeeeeetthhhhtteehhhhhhhhhheeeeeeehhhhheeeeeeeeeeett


       190       200       210       220       230       240
IGSITTLFKIPKSMASLSLPSTISINLQVHIKIGVQTDSKGIVQILDEKGEKSLNFMVHLG

heeeeeeeeehhhheeeeeeeceeeeeeeeeeeeeeeeeetttteeeeeehhhhhhhhhhhhhhh


       250       260       270       280       290       300
LIKRKVGRMYSVEYCKQKIEKMRLIFSLGLVGGISLHVNATGSISKTLASQLVFKREICY

hhheeeeeeeeehhhhhhhhhhhhhheeeeecccceeeeeeeeehtcceeeeehhhhhheheetee


       310       320       330       340       350
PLMDLNPHLNLVIWASSVEITTRVDAIFQPSLPGEFRYYPNIIAKGVGKIKQWN

eeehhhhhhheeeeehhhhhheeeheeeeeeeccccteeeeeeeeeeeeeeeeeeeeee
```

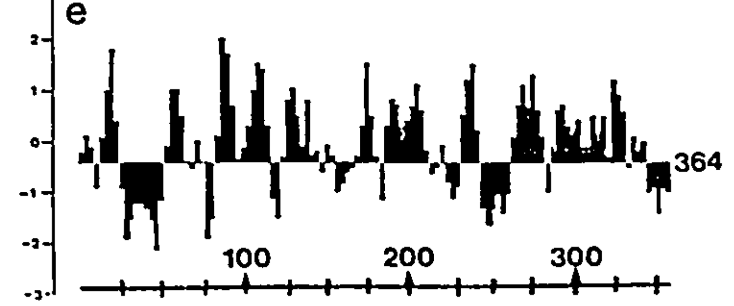
Figure 5. Matrix protein hydropathy profiles of selected Paramyxoviruses. Plots are presented for HPIV3 (a), BPIV3 (b), Sendai virus (c), measles virus (d), and NDV (e). The length of each M protein is indicated to the right of each hydrophobicity plot. Regions within each plot may be located by referring to the X-axes which are notched at intervals of 25 amino acids. Each profile is an average of the hydropathic indices (relative hydrophobic/hydrophilic nature) of nine consecutive amino acids. Positive values indicate hydrophobicity and negative values indicate hydrophilicity.


a
353
100
200
300


b
350
100
200
300


c
348
100
200
300


d
335
100
200
300


e
364
100
200
300

virus, CDV (not shown) and NDV. However, the hydropathy plot of HPIV3 M bears only a slight resemblance to that of influenza virus M protein and virtually none with respect to RSV, VSV and rabies virus M proteins (not shown). The HPIV3 M protein contains two hydrophobic stretches extending from positions 171 to 212 and 262 to 330,inclusive.

C. Discussion

1. Sequence analysis of the HPIV3 M gene

All but 30 nucleotides or about 2.6% of the HPIV3 M gene was verified by resequencing in the opposite direction, sequencing of independent clones, or Maxam and Gilbert sequencing. Only regions 1 to 27 and 184 to 186 have been sequenced in one direction without verification in this study. However, these regions are in agreement with other published HPIV3 M gene sequences (Galinski et al., 1987a, Spriggs et al., 1987a; Luk et al., 1987) and region 1 to 27 has been confirmed by primer extension (Côté, unpublished). Also, region 184 to 186 lies within a BamH1 site (183 to 188) that was utilized in the preparation of fragments for Maxam and Gilbert sequencing.

The complete nucleotide sequence of the HPIV3 genome has been reported and includes: the 3' end (Dimock et al., 1986), the nucleocapsid protein or NP gene (Jambou et al., 1986; Galinski et al., 1986a; Sánchez et al., 1986); the gene encoding both the nucleocapsid phosphoprotein P and the nonstructural protein C (Galinski et al., 1986b; Luk et al., 1986; Spriggs and Collins, 1986b); the matrix M

protein gene (Galinski et al., 1987a; Luk et al., 1987; Prinoski et al., 1987; Spriggs et al., 1987a); the fusion glycoprotein F gene (Spriggs et al., 1986; Côté et al., 1987; Galinski et al., 1987b); the hemagglutinin-neuraminidase glycoprotein HN gene (Elango et al., 1986; Galinski et al., 1987b; Storey et al., 1987), and; the large or L protein gene (Galinski et al., 1988).

Differences among the four published HPIV3 M gene and predicted amino acid sequences (Table 3) are most likely due to viral variation rather than sequencing error. A total of 15 nucleotides and 6 amino acids are not the same among the available publications. Since proline, cysteine, glycine and tryptophan residues are conserved, the predicted M protein secondary structures from these sources are probably very similar to each other. Differences include the following substitutions: conservative at positions 80 and 97; charge changes at positions 50, 103 and 139; and nonpolar/polar uncharged residues at position 270 (Lehninger, 1979). A charge difference has been noted: +21.5 (Spriggs et al. 1987a) and +20.5 (Galinski et al. 1987a; Luk et al. 1987; Prinoski et al. 1987).

HPIV3 M mRNA has a 5' nontranslated region of 32 nucleotides and the other HPIV3 mRNAs have 5' nontranslated regions ranging from 22 to 193 nucleotides long. The 3' nontranslated region of the M mRNA is also quite short - only 64 nucleotides long, including the translational stop signal. These regions compare favourably with those of the M mRNA of Sendai virus (Blumberg et al., 1984; Hidaka et al., 1984) which has short 5' and 3' nontranslated ends that are 32 and 97 (Hidaka et al., 1984) or 109 (Blumberg et al., 1984) nucleotides long, respectively.

However, the M mRNA of measles virus and CDV both have 3' nontranslated regions that exceed 400 nucleotides (Bellini et al., 1986).

In message sense, all genes in HPIV3 start with the sequence 5' AGGANNAAAG 3', including the M gene where NN = TT. Also, with the exception of M, all genes in HPIV3 end 5' ANNNANNAAAAA 3'. In the M gene this sequence includes eight additional nucleotides, 5'AGATAATC 3' (Figure 3), just before the five terminal A residues. This difference may be correlated with an increased incidence of readthrough transcription leading to higher levels of M-F bicistronic mRNA relative to those observed for any other HPIV3 polycistronic mRNA species (Spriggs and Collins, 1986a; Côté et al., 1987; Dimock et al., 1987; Galinski et al., 1987a; Luk et al., 1987). Changes in the intergenic region between the HN and L genes of Sendai virus may be associated with the synthesis of polycistronic mRNAs (Gupta and Kingsbury, 1985). Gene start, end and intergenic sequences may function as transcriptional regulatory elements.

2. Computer-generated nucleotide and amino acid sequence alignments

Computer sequence alignments of the HPIV3 M gene and protein with those of related viruses generated more or less expected results. The HPIV3 M nucleotide sequence is related to the M sequences of BPIV3, Sendai virus, CDV, NDV and measles virus, listed according to decreasing similarity. The similarity of HPIV3 with BPIV3 and Sendai virus is predicted by the fact that all are members of the same genus. The M gene of NDV is considerably less similar to the HPIV3 M gene than

that of Sendai virus - possibly due to adaptation to an avian rather than a mammalian host.

The M gene sequence of mumps virus might be expected to be more similar to that of HPIV3 than SV5, and to SV5 more so than NDV since humans are the only natural hosts for mumps virus (Ginsberg, 1988) and SV5 is of simian origin (Hull et al., 1956). Although measles virus and CDV are both Morbilliviruses, CDV displays a similarity to HPIV3 in the M gene that was slightly higher than that found for NDV. This is an inconsistency: a member of a different genus (CDV) is more similar in terms of M gene sequence to HPIV3, a Paramyxovirus, than NDV, also a Paramyxovirus. Measles virus was found to be to be almost 30% similar to HPIV3 in the M gene. The rest of the viruses examined, including RSV, are even less similar to the HPIV3 M gene. These results may argue in support of other evidence such as gene order (Collins et al., 1986) and possible evolutionary divergence (Spriggs et al., 1987a), for the reclassification of the Paramyxoviridae. Evolutionary models have been proposed in an attempt to account for the differences within the Paramyxoviridae (Spriggs et al., 1987a).

3. Sequence analysis of the HPIV3 M protein

Within the HPIV3 M protein, hydrophobic amino acids such as Met, Val, Ile, Trp, Tyr, Leu and Phe (Hopp and Woods, 1981) account for approximately 37% of the residues present. Hydrophobic residues are also evident in the M proteins of other Paramyxoviruses such as measles virus and CDV, both at 37% (Bellini et al., 1986); NDV at 33% (Chambers

et al., 1986); and Sendai virus at 39% (Blumberg et al., 1984). This relative abundance of hydrophobic amino acids within the M proteins discussed may be indicative of hydrophobic interactions with membrane lipids. The large number of basic residues, reflected in the charges calculated for each viral M protein suggests an ionic interaction of the matrix protein with the viral nucleocapsid cores during replication (Galinski et al., 1987a). Exceptions to this include the membrane proteins of SV5 (McSharry et al., 1975) and rabies virus (Rayssiguier et al., 1986), both of which appear to be acidic. These observations reveal the possibility that spatial distribution of charged amino acids may be more important to structure and activity than net charge (Blumberg et al., 1984).

An interesting feature of some M proteins is the occurrence of paired basic amino acid residues. Seven such pairs are found in measles virus and CDV has six - all of which occur in identical positions relative to measles (Bellini et al., 1986). Sendai virus M protein has seven basic paired residues (Blumberg et al., 1984) and four occur in positions identical to those in the M proteins of measles and CDV (Bellini et al., 1986). Although the M protein of NDV has nine basic paired residues (Chambers et al., 1986) only three of these seem to be conserved relative to Sendai and measles virus: Lys-Lys, 118 to 119; Arg-Arg, 247 to 248; and Lys-Lys, 250 to 251. Sendai virus and measles virus have Arg-Arg in positions corresponding to 118 and 119 in NDV, so that relative position and charge are maintained but the identities of the residues differs (Chambers et al., 1986). It should be noted that two of the dipeptides are located beside each other, with some

variations: Arg-Arg-Lys-Lys, residues 225 to 228 in the measles virus M protein (Bellini et al., 1986); Arg-Lys-Lys, residues 226 to 228 in CDV (Bellini et al., 1986); Arg-Arg-Lys, residues 239 to 241 in Sendai (Blumberg et al., 1984); possibly Arg-Arg-Gly-Lys-Lys, residues 247 to 251 in NDV (Chambers et al., 1986); and Lys-Arg-Lys, residues 243 to 245 in HPIV3 (Prinoski et al., 1987). HPIV3 has a total of five basic amino acid pairs and one basic tripeptide (Prinoski et al., 1987). Of these, the pairs occurring from 49 to 50, 118 to 119 and 295 to 296 and the tripeptide 243 to 245 are conserved in Sendai (Blumberg et al., 1984); the pair at positions 146 to 147 is not present in Sendai; and the remaining pair Arg-Lys 33-34 in HPIV3 appears to correspond to Lys-Lys 29 to 30 in Sendai. On the other hand, Sendai has two pairs of basic residues 227 to 228 and 346 to 347 (Blumberg et al., 1984) that are not present in HPIV3. The significance of these paired basic residues is unknown, although they may be involved with ionic interactions with nucleocapsids. Since they are conserved among the M proteins discussed, it is likely that they are of some importance.

4. Conserved secondary structure among Paramyxoviral M proteins

A conserved region, consisting of a hydrophobic peptide forming an anti-parallel β-sheet bordered by two α-helices, occurs near the C-terminus of the M proteins of measles virus, CDV and Sendai virus and may form a fold which interacts with the membrane (Bellini et al., 1986). Despite variation from measles and Sendai virus, there is evidence for the same structure in the analogous region of the M

protein from NDV (Chambers et al., 1986) and HPIV3 (Figure 3). The region extends from residues 235 to 279 in the measles and CDV M proteins (Bellini et al., 1986); 250 to 294 for Sendai virus (Bellini et al., 1986); 256 to 300 for the NDV M protein sequence (Chambers et al., 1986); and 252 to 295 for HPIV3 (Figure 3). Except for HPIV3 which is one residue shorter, the length of this segment is 45 amino acids in the M protein of each of these viruses. Two conserved cysteine residues present in measles (237, 281) and Sendai virus (251, 295) are located at either end of this region (Bellini et al., 1986). Only one of these is conserved in NDV (301) (Chambers et al., 1986) and cysteines are found in the HPIV3 M protein at positions 255 and 299 (Prinoski et al., 1987). These cysteines in the HPIV3 M protein are very similar in position to the borders predicted by secondary structure analysis.

In the M proteins of Sendai, measles and CDV, the borders of the β-sheet region coincide closely with a hydrophobic stretch of amino acids (Bellini et al., 1986). A generally hydrophobic segment is conserved in the NDV M protein from residues 266 to 280 (Chambers et al., 1986). Paired glycine residues are found in measles and CDV at positions 254 and 255 (Bellini et al., 1986); in Sendai at positions 268 to 269 (Bellini et al., 1986); and in HPIV3 at positions 272 to 273 (Prinoski et al., 1987). NDV glycine 275, however, is predicted to be followed by proline 276 (Chambers et al., 1986). These paired glycine residues, or glycine-proline in NDV, are believed to promote a turn in the M proteins of all five viruses (Bellini et al., 1986; Chambers et al., 1986). An anti-parallel β-sheet likely follows the turn in all cases and terminates in the vicinity of the next glycine - position 264

in measles and CDV (Bellini et al., 1986), position 278 in Sendai virus (Blumberg et al., 1984), position 284 in NDV (Chambers et al., 1986) and position 282 in HPIV3 (Prinoski et al., 1987). This conserved domain may interact with the membrane without actually spanning the lipid bilayer and may also be responsible for the requirement of detergents to solubilize M proteins (Bellini et al., 1986). It is not known if the two cysteine residues which border this region are involved in disulfide bond formation (Bellini et al., 1986).

5. Structure-function interdependence in M proteins

Similarities of M protein hydrophobicity plots among the Paramyxoviridae are reminiscent of the comparisons of nucleotide and amino acid order. Paramyxoviruses generally display M protein hydrophobicity plots closely resembling that of HPIV3. As phylogenetic distance increases, the similarity between hydrophobicity plots usually decreases. Assuming that the matrix protein from each virus examined performs basically the same function, the retention of function despite differences in nucleotide order, amino acid sequence, hydrophobicity plots, protein length and charge may be partly due to virus-host specialization and the presence of conservative amino acid differences leading to similar secondary structures.

Statistical analysis and comparisons between HPIV3 and Sendai virus proteins reveal that M is the most highly conserved protein (Spriggs et al., 1987a). Determinants of protein secondary structure such as glycine, proline and cysteine are highly conserved between

HPIV3 and Sendai virus (Blumberg et al., 1984), illustrating an important relationship between function and conformation (Galinski et al., 1987a). Similarities between the N genes and deduced N proteins of measles and CDV exist only in certain regions (Rozenblatt et al., 1985) and comparisons of the N and P genes of measles with Sendai virus show little similarity at the nucleotide or amino acid level (Bellini et al., 1988). This suggests that selective pressures on the M proteins of Paramyxoviruses may differ from those on other proteins and are less tolerant of divergence due to its multifunctional role in maturation.

6. Summary of functional regions within the HPIV3 matrix protein

The matrix protein of HPIV3 probably interacts with glycoproteins, nucleocapsids, lipid membranes and other M protein molecules. Amino acids 255 to 299 represent a conserved domain present in other Paramyxoviruses (see p.47-49) that may interact with the membrane without spanning the lipid bilayer (Bellini et al., 1986). Nucleocapsid proteins of Paramyxoviruses carry overall acidic charges (Chambers et al., 1986). Therefore, nucleocapsids may associate with the HPIV3 M protein amino terminal end, particularly with amino acids 49 to 50 and 118 to 119 - depending on conformation, since these represent conserved basic residues (see p.46-47). By process of elimination, the glycoproteins may associate with amino acids 300 to 353 of the HPIV3 M protein. However, this proposal is not supported by charge predictions: this region of the HPIV3 M protein bears a charge of +1.5 and the cytoplasmic tails of both F and HN are positively charged also.

Chapter IV - F GENE SEQUENCE COMPARISON AMONG TEN HPIV3 ISOLATES

A. Introduction

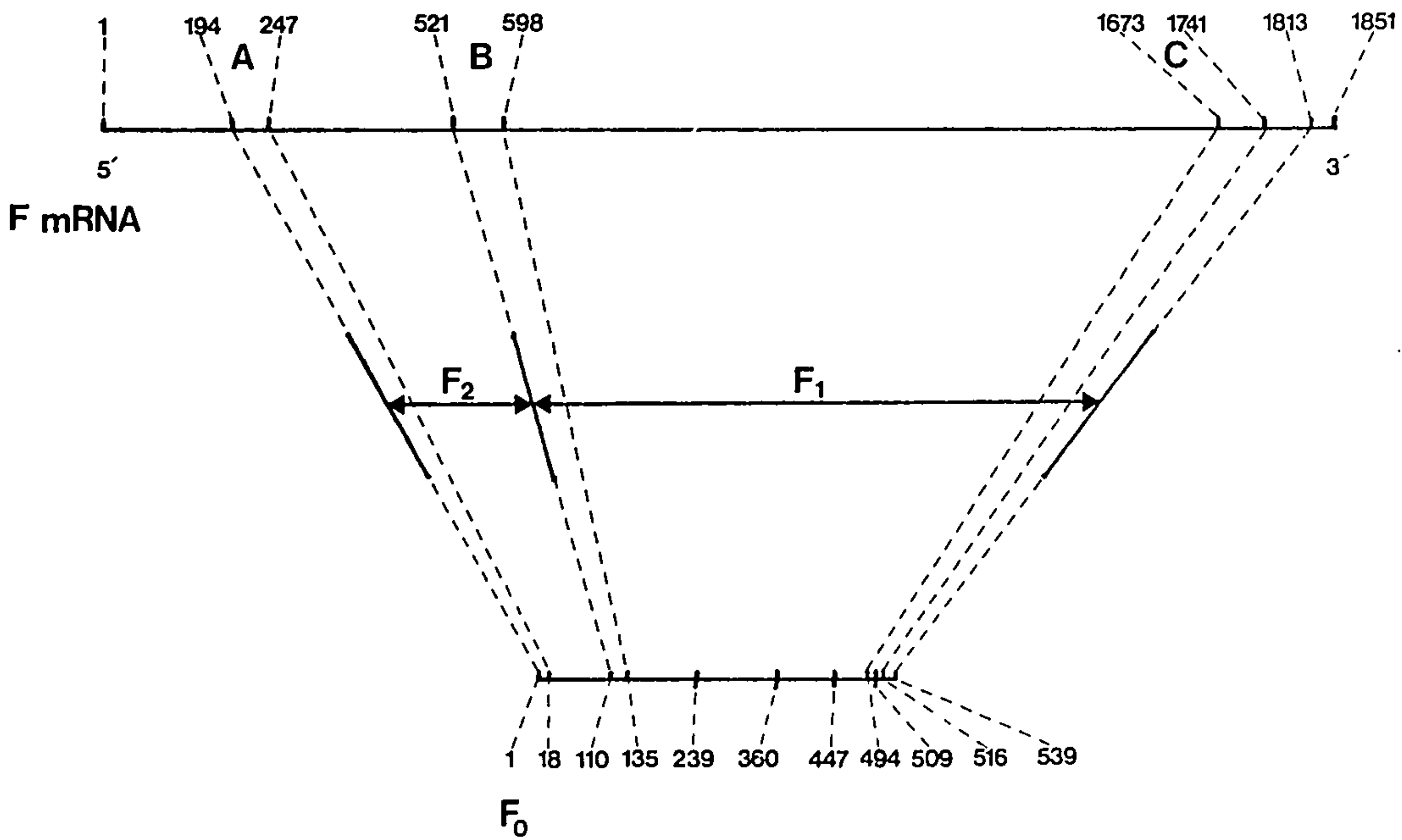
Initiation of the *Paramyxovirus* infectious cycle relies on cleavage-activation of the glycosylated F protein precursor, F_0, by a host enzyme with trypsin-like activity (Scheid and Choppin, 1974, 1977; Choppin and Scheid, 1980). This yields two disulphide-linked peptides, F_1 and F_2, generating a highly conserved, hydrophobic F_1 N-terminus which may interact with target cell membranes and mediate membrane fusion (Richardson et al., 1980; Richardson and Choppin, 1983). *Paramyxovirus* fusion proteins mediate virus penetration of cells, hemolysis and syncytia formation (Homma and Ohuchi, 1973; Scheid and Choppin, 1974, 1977; Merz et al., 1980).

1. Structure of the HPIV3 fusion gene and glycoprotein.

Unless stated otherwise, the structural details summarized within this sub-section and Figure 6 are condensed from the work of Côté et al. (1987). The HPIV3 F mRNA is 1851 nucleotides long and has 5' and 3' nontranslated regions of 193 and 38 nucleotides, respectively. The coding portion is 1620 bases long, including the termination codon.

The unglycosylated F_0 precursor is comprised of 539 amino acids with a predicted molecular weight of 60031. Cleavage of F_0 between arginine 109 and phenylalanine 110 generates two disulphide linked proteins - F_2 which is 109 amino acids long and has a molecular weight

Figure 6. HPIV3 F mRNA and protein product. The upper line diagram proceeding 5' to 3' represents the F mRNA and nucleotide position is indicated by the numbers 1 to 1851. The lowest line represents the fusion protein precursor, F_0. Individual amino acid residues are numbered 1 to 539, with position 1 corresponding to the F_0 N-terminus and position 539 to the C-terminus. Amino acids 239, 360, 447 and 509 represent the middle residue of potential glycosylation sites. Broken lines mark boundaries. The letters A, B and C refer to the F_0 signal peptide, the F_1 amino terminus or fusion sequence and the F_0 or F_1 transmembrane region, respectively.



1
194
A
247
521
B
598
1673
1741
C
1813
1851
5´
3´
F mRNA
F_2
F_1
1
18
110
135
239
360
447
494
509
516
539
F_0

of about 10000, and F_1 which is 430 amino acids long and has a molecular weight of about 50000. Three major hydrophobic regions have been identified: the F_2 signal peptide - 18 amino acids long; the F_1 amino terminus - 26 amino acids long, and; the F_1 transmembrane region - 23 amino acids long. Four potential glycosylation sites (Asn-X-Ser/Thr) were identified in F_0 at positions 238-240, 359-361, 446-448 and 508-510. The last of these is likely not utilized since it lies within the boundaries of the transmembrane region.

Many of the cysteines and, to a lesser degree, the prolines are conserved among the F_1 proteins of HPIV3, Sendai virus and SV5. There is also at least one cysteine residue in the F_2 region of all Paramyxoviridae that is conserved with respect to distance from the cleavage activation site (Morrison, 1988). Since F_1 and F_2 are joined by disulphide bonds, this conservation of cysteines and prolines may reflect the importance of proper molecular folding or structure to function. The correct positioning of the F_1 amino terminus, necessary for cleavage and membrane fusion, may require a stringent molecular conformation (Morrison, 1988).

The amino acid sequence similarities have been determined between the F protein of HPIV3 and various other Paramyxoviruses. These include Sendai virus - 41.0 to 43.0% (Spriggs et al., 1986; Côté et al., 1987; Galinski et al., 1987b), NDV - 30.4% (Galinski et al., 1987b), SV5-25.3 to 29.5% and RSV - 19.2 to 23.6% (Spriggs et al., 1986; Galinski et al., 1987b).

2. Function of the *Paramyxovirus* fusion glycoprotein.

The fusion glycoprotein of Sendai virus mediates hemolysis and virus-induced membrane fusion (Scheid and Choppin, 1974). *Paramyxoviruses* enter host cells by membrane fusion between the virus and uninfected cell or between adjacent infected and uninfected cells (Bratt and Gallaher, 1969; Choppin and Compans, 1975; Bratt and Hightower, 1977). Cleavage of F_0 generates a highly conserved sequence at the F_1 amino terminus that is believed to mediate both types of membrane fusion (Richardson et al., 1980; Hsu et al., 1983; Richardson and Choppin, 1983). Once exposed by cleavage-activation, the hydrophobic F_1 N-terminus can interact directly with membranes (Paterson and Lamb, 1987). Cleavage of F_0 results in increased hydrophobicity and alpha-helicity (Hsu et al., 1981). The inhibition of *Paramyxoviral* penetration and hemolysis by synthetic peptides with sequences similar to the F_1 N-terminus points to an interaction with sites on the cell surface (Richardson and Choppin, 1983). Although the F_1 N-terminus is probably involved in membrane fusion, the exact mechanism and possible contributions of other regions of the protein to this process are not known. The unglycosylated F_2 polypeptide in HPIV3 (Galinski et al., 1987b), the *Paramyxoviral* HN glycoprotein (Galinski et al., 1987b; Thompson and Portner, 1987) and regions near the F_1 C-terminus of Sendai virus (Portner et al., 1987) and measles virus (Hull et al., 1987) may enhance the expression of fusion activity.

Specific functions may be attributed to other regions of the HPIV3 fusion protein. For example, the amino terminal signal peptide

facilitates co-translational transfer of F_0 through the rough endoplasmic reticulum (Spriggs et al., 1986). The transmembrane region, perhaps in conjunction with the charged cytoplasmic tail, acts as a stop transfer sequence and embeds the fusion protein within the lipid bilayer (Morrison, 1988).

3. Contributions to the pathogenesis of Paramyxoviruses

Although HN and other factors are involved, differences in Paramyxovirus virulence, host range and tissue specificities may depend at least in part on the susceptibility of F_0 to proteolytic cleavage. Non-infectious Sendai virions were produced by cells lacking the protease required for F_0 cleavage (Scheid and Choppin, 1974, 1976). Further restrictions on host range are evident in Sendai virus mutants whose F proteins are activated by elastase, chymotrypsin or chicken plasmin, instead of the trypsin-like activity required by the wild-type (Hsu et al., 1987a,b). It has also been demonstrated that the F protein of virulent NDV strains is always cleaved *in vitro* but that avirulent strains produced by some cells have F proteins which are not activated (Nagai et al., 1976b).

4. Immune response to the Paramyxovirus F glycoprotein.

The development of immunity during natural infection by HPIV3 is poorly understood (van Wyke Coelingh et al., 1988). Only a partial protection appears to be induced by the initial exposure to HPIV3 and

several infections may be required for more effective immunity (Chanock et al., 1961). In vitro, antibodies directed against the F protein inhibit virus penetration and prevent the spread of infection by cell fusion (Merz et al., 1980). Mice are protected against challenge with Sendai virus by passive immunization with anti-F monoclonal antibodies (Orvell and Norrby, 1977). Despite the fact that the HPIV3 HN protein induces a greater protective immune response than the F protein in cotton rats (Spriggs et al., 1987b), antibodies to both may be required for complete immunity (Ray and Compans, 1987; Ray et al., 1988). Hamsters and patas monkeys immunized with both glycoproteins showed, respectively, complete protection against (Ray et al.,1985; Ray et al., 1988) or significant resistance (Spriggs et al., 1988) to challenge with HPIV3. The passive transfer of monospecific rabbit antibodies to hamsters, or immunization with HN or F alone, conferred an incomplete protection against HPIV3 (Ray et al., 1988).

The F gene codes for an external viral antigen which may render viral particles more susceptible to antibody neutralization (Merz et al., 1980). Specific F gene variants favoured by natural selection might affect the success of future vaccines or nucleic acid and monoclonal antibody diagnostic probes.

B. Results

Ten HPIV3 isolates, obtained during 1957 to 1987 from different geographical locations, were compared to identify variable and conserved regions in the F gene. The HPIV3 F gene nucleotide sequence

(Spriggs et al., 1986; Galinski et al., 1987b; Côté et al., 1987) was used to select the 20-mer oligonucleotide primers necessary for direct RNA sequencing of the viral genome of each isolate. The following nucleotide regions were found to be appropriate primer annealing sites within the F gene: 82 to 101; 370 to 389; 593 to 612; 818 to 837; 971 to 990; 1226 to 1245; 1493 to 1512; and 1652 to 1671, inclusive. Site 1026 to 1045, within the M gene, was used to obtain the sequence of the M-F junction and the 3' end of the F gene.

Each nucleotide position within the F gene and flanking sequences of nine HPIV3 clinical isolates and the prototype strain were compared (Appendix 1). Of the 1851 nucleotides comprising the F gene, a total of 25 remain ambiguous out of the 18510 nucleotides sequenced for all ten isolates combined. Collectively, the isolates differ from Wash/57 at 190 positions (Appendix 1, Table 4) and many show the same change at a given nucleotide position. Over the thirty years spanned by these isolates, the transcriptional start and stop sequences are unchanged. The high percentage of difference found in the 5' nontranslated region contrasts sharply with the low percentage of difference in the coding portion of the gene (Table 4). In terms of the number of nucleotide differences (Table 5), Aus/59 is the most similar to Wash/57 and Ont/87 the least similar. From 1975 to 1987, the total number of nucleotide differences among the isolates from the Wash/57 F gene, within the coding and nontranslated regions, displayed little variance (Table 5).

Some of the nucleotide differences from the Wash/57 F gene may be cummulative among the isolates examined (Table 6). A total of 60 substitutions are found consistently in all isolates obtained after the

Table 4: Distribution of F gene nucleotide differences among HPIV3 clinical isolates

Region of F mRNA	Number (%) of nucleotide differences
total message 1851[a]	190 (10.3)
5'nontranslated 193[a]	60 (31.1)
translated 1620[a]	128 (7.9)
3'nontranslated 38[a]	2 (5.3)

[a] indicates region length in nucleotides.

Table 5: Detailed distribution of F gene nucleotide differences among HPIV3 clinical isolates

	total message 1-1851[a]	5'non-coding 1-193[a]	coding 194-1813[a]	3'non-coding 1814-1851[a]
Aus/59	34(1.8)	13 (6.7)	21(1.3)	0 (0)
Wash/75	94(5.1)	27(14.0)	66(4.1)	1(2.6)
Tex/80	104(5.6)	32(16.6)	70(4.3)	2(5.3)
NY/81	106(5.7)	32(16.6)	72(4.4)	2(5.3)
Ont/82(a)	100(5.4)	33(17.1)	65(4.0)	2(5.3)
Ont/82(b)	99(5.3)	31(16.1)	66(4.1)	2(5.3)
Ont/83	101(5.5)	31(16.1)	68(4.2)	2(5.3)
NY/84	97(5.2)	30(15.5)	65(4.0)	2(5.3)
Ont/87	109(5.9)	37(19.2)	71(4.4)	1(2.6)

[a] indicates region border in terms of nucleotide number
() indicate percent difference from Wash/57

Table 6: Accumulation of HPIV3 F gene nucleotide differences

Type of difference	Number of differences	Percent (/190)
accumulating (including Ont/87)	60	31.6
accumulating (not including Ont/87)	25	13.2
possibly accumulating (appearing only in NY/84 or Ont/87)	35	18.4
not accumulating	70	36.8

years that these individual differences first appeared and this number may be increased to 85 (44.7%) by excluding Ont/87 (Table 6). For reasons to be discussed later, an additional 35 differences from Wash/57 (18.4%) may also be accumulating since they occur only in NY/84 or Ont/87 (Table 6). The remaining 70 substitutions (36.8%) are not maintained consistently from the time of their first appearance.

By direct comparison (Table 7), Aus/59 is most closely related to Wash/57 and those existing differences may at least be partially due to geography. All of the isolates collected in the 1980's are very similiar to each other with the exception of Ont/87 which has 32 unique differences with respect to Wash/57 that are not present among any of the other isolates. Wash/75 seems to fall between the 1950's and the 1980's in terms of similarity to all of the other isolates.

The 190 nucleotide differences result in 27 amino acid substitutions (Table 8, Appendix II) - categorized as follows: 15 are conservative; 7 involve substitutions of residues with nonpolar R groups for those with uncharged polar R groups or vice versa; and 5 involve charge changes. These 27 differences are found in 24 amino acid positions due to the occurrence of multiple differences at amino acid positions 108 and 367 (Appendix II). Since proline, cysteine and tryptophan residues are unaffected by nucleotide variation, with the possible exception of position 1184 (Table 8), major changes in F protein secondary structure probably do not occur. However, alterations in conformation may result from charge changes listed in Table 8.

Two amino acid differences due to substitutions at nucleotides 515 - a charge change from negative to positive, and 522 may affect the

Table 7: F gene nucleotide differences among HPIV3 clinical isolates[a]

	Aus /59	Wash /75	Tex /80	NY /81	Ont /82a	Ont /82b	Ont /83	NY /84	Ont /87
Wash /57	34	94	104	106	100	99	101	97	109
Aus /59		90	96	102	97	93	96	96	109
Wash /75			66	75	68	68	70	67	66
Tex /80				11	10	12	14	16	72
NY /81					14	17	19	20	81
Ont /82a						9	11	12	77
Ont /82b							4	13	78
Ont /83								15	80
NY /84									75

[a] the number of nucleotide differences among all clinical isolates and prototype with respect to each other are presented in a pair by pair comparison

Table 8: Distribution of F protein amino acid differences among HPIV3 clinical isolates[a]

coding region	number (%) of differences per region	differences of interest
F_2 signal peptide 1-18[b]	3 (17)	-
Remainder of F_2 19-109[b]	6 (7)	515A-G=108Lys-Glu Tex/80, NY/81
F_1 amino terminus 110-135[b]	1 (4)	522T-C=110Phe-Ser Tex/80, NY/81
F_1 membrane anchor 494-516[b]	4 (17)	-
F_1 cytoplasmic tail 517-539[b]	0 (0)	-
Remainder of F_1 136-493[b]	13 (4)	736A-C=181Lys-Asn Ont/87 1184T-a=331Cys-Ser? Ont/87 1299C-A=369Thr-Lys Aus/59 1514A-g=441Asn-Asp? Wash/75 1655C-T=488His-Tyr Aus/59
total 1-539[b]	27 (5)	-

[a] differences are listed with respect to Wash/57
[b] indicate nucleotide regions of F gene coding for the protein regions listed

cleavage-activation site (Table 8). The F_2 signal peptide and F_1 amino terminus are highly conserved (Table 9). None of the amino acid changes within the F_1 membrane anchor result in charge changes - three are conservative and the fourth involves the substitution of an isoleucine for a threonine. The pattern evident in the comparison of nucleotide differences (Table 7) is maintained at the amino acid level (Table 10).

C. Discussion

1. F gene sequence variation among HPIV3 clinical isolates

Variation in the HPIV3 F gene was investigated by nucleotide sequence comparison of clinical isolates. Similarities between the isolates and Wash/57, which were obtained over a thirty year period, range from 94.1 to 98.2% (Table 5) and reflect possible structural or functional constraints on variation in the F protein.

The transcriptional start (nucleotides 1-10) and stop (1840-1851) signals have been strictly maintained, as expected due to the functional importance of these control elements. It is unlikely that the 5' nontranslated portion (in message sense) of the F gene (exclusive of nucleotides 1-10) has any involvement in transcriptional regulation. However, the fact that this region has the highest level of sequence diversity (6.7 to 19.2%, Table 5) may have other implications. It has been suggested that secondary structure in the 5' terminal noncoding F mRNA sequence of measles virus may play a role in translational regulation (Richardson et al., 1986). If also true for

Table 9: Detailed distribution of F protein amino acid differences among HPIV3 clinical isolates

	F_2 signal peptide 1-18[a]	Remainder of F_2 19-109[a]	F_1 amino terminus 110-135[a]	F_1 trans-membrane region 494-516[a]	Remainder of F_1 136-493[a], 517-540[a,b]
Aus/59	0	0	0	1	5
Wash/75	0	2	0	3	6
Tex/80	1	4	1	3	6
NY/31	0	5	1	3	5
Ont/82(a)	0	3	0	3	5
Ont/82(b)	0	3	0	3	5
Ont/83	0	3	0	4	5
NY/84	0	3	0	3	5
Ont/87	2	2	1	4	6

[a] indicates protein region
[b] includes translational stop codon

Table 10: F protein amino acid differences among HPIV3 isolates[a]

	Aus /59	Wash /75	Tex /80	NY /81	Ont /82a	Ont /82b	Ont /83	NY /84	Ont /87
Wash /57	6	8	13	13	10	10	11	10	13
Aus /59		8	12	12	9	9	10	9	14
Wash /75			9	9	6	6	7	6	11
Tex /80				4	4	4	5	4	11
NY /81					3	3	4	3	11
Ont /82a						0	1	0	11
Ont /82b							1	0	11
Ont /83								1	12
NY /84									11

[a] the number of amino acid differences among all clinical isolates and the prototype with respect to each other are presented in a pair by pair comparison.

HPIV3, nucleotide variation in this region could indirectly affect translational efficiency by secondary structure alteration of this region in the messenger RNA.

2. Comparison of nucleotide differences in the F and HN genes

The hemagglutinin-neuraminidase (HN) genes of Wash/47885/57 and six clinical strains isolated between 1973 and 1983 have been sequenced (van Wyke Coelingh et al., 1988). One interpretation of the results is that variability in the HN gene does not correlate with the year of virus isolation. For example, a 1973 Washington HPIV3 strain was found to be most closely related to a 1982 Texas virus and most different from a 1974 Australian strain (van Wyke Coelingh et al., 1988).

However, the F sequence data does suggest a correlation between variability in the F gene and the year of isolation. Wash/75 shares 54 differences from prototype with all of the isolates obtained in the 1980's. Therefore, at least 72% of the differences from Wash/57 that occur in the isolates from the 1980's can be accounted for by differences between 1957 and 1975. This suggests that the differences from 1957 in the isolates from the 1980's are made on the background of Wash/75. With the exception of Ont/87 which is quite different from every other isolate examined (Table 3), all of the 1980's isolates are very similar to each other. Collectively, the data are consistent with the presence of two or three branches of HPIV3. Ont/87 is very different from every other isolate examined and may represent a distinct branch. Since it has an average of 44.2% more differences from

Wash/57 and Aus/59 than any of the remaining isolates, Ont/87 may be more closely related to Wash/75 and the isolates from the 1980's than to the former two isolates. A second branch is likely comprised by the remainder of the isolates from the 1980's which are all similar to each other (Table 6). The two isolates from the 1950's may represent a third lineage due to their similarity to each other (Table 6). Wash/57 and Aus/59, based on the consistent number of differences displayed by both isolates as compared to the others (Table 6), may be more closely related to the ancestral HPIV3 virus than the other isolates.

Evolutionary patterns evident in the pairwise comparison of F gene nucleotide differences among HPIV3 isolates (Table 6) are verified by matrix calculations. The method of Li (1981) yields two possible estimations of phylogeny in which "X" denotes a common ancestor and the isolates are indicated by their year of isolation:

1. ((((('82b '83)'82a)'84)'80 '81)'87 '75) X ('57 '59)

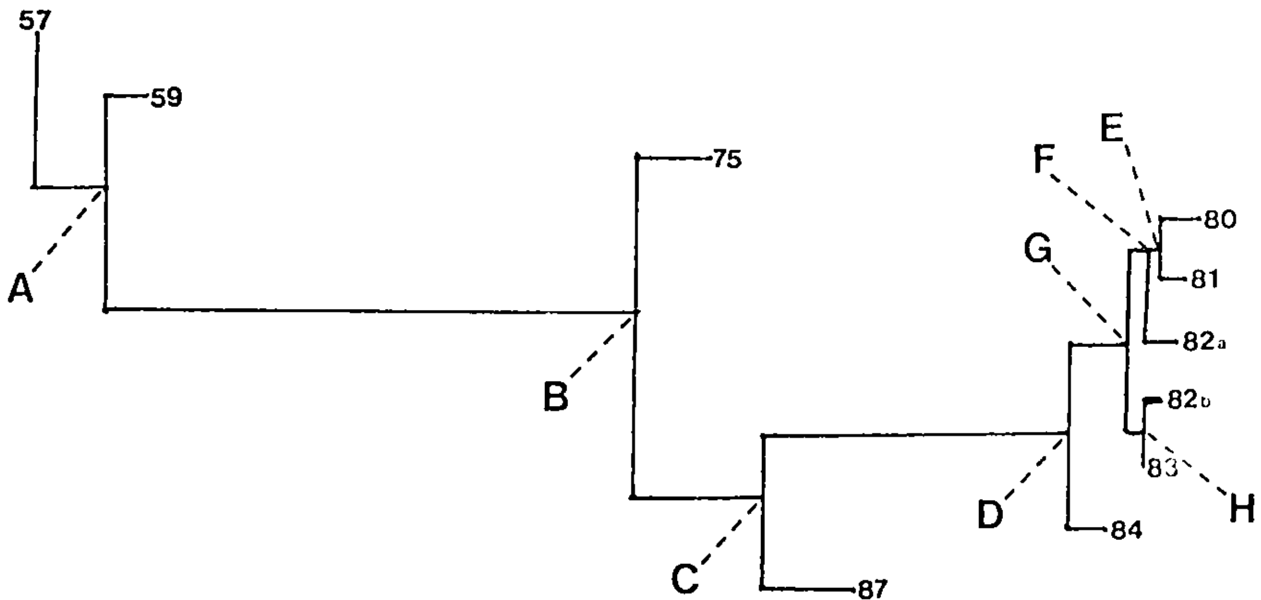
2. ((((('82b '83)'80 '82a)'84)'81)'87 '75) X ('57 '59)

The PAUP (Phylogenetic Analysis Using Parsimony) program, version 2.4, obtained from Dr. D. L. Swofford of the Illinois Natural History Survey generated an evolutionary tree topology (Figure 7) similar to the first two predictions:

((((((('80 '81)'82a)'82b '83)'84)'87)'75)'59) X '57

It should be emphasized that the year associated with each isolate only indicates the time the sample was taken and not necessarily when that strain first attained its own unique variation of the F gene sequence. These calculations support the idea that Ont/87 diverged from the other 1980's isolates during the 1970's.

Figure 7. Proposed phylogenetic tree of HPIV3 isolates. The tree is based on F sequence data and assumes that substitutions occur at a constant rate, are random throughout the genome and do not revert. Isolates are indicated according to year of isolation: for example, 57 means Wash/57. Nodes are shown by capital letters. Branch lengths in terms of F gene nucleotide differences have been predicted: 59-A=6; 75-B=12; 80-E=6; 81-E=4; 82a-F=3; 82b-H=1; 83-H=0; 84-D=8; 87-C=17; E-F=6; F-G=4; H-G=6; G-D=9; D-C=61; C-B=25; B-A=107; A-57=14.


57
59
A
75
B
C
87
F
E
80
G
81
82a
82b
83
D
84
H

The epidemiology of HPIV3 is thought to be similar to influenza C virus (van Wyke Coelingh et al., 1988), that is, co-circulation of multiple lineages and a slow rate of change relative to influenza A and B (Buonagurio et al., 1985, 1986). However, it is interesting to note that the none of the HPIV3 isolates examined for sequence variation in the HN and F genes exhibit the conservation of sequence evident among some influenza C virus isolates. For example, the NS genes of C/TAY/47 and C/JHG/66 - both strains of influenza C and isolated 19 years apart - are the same (Bounagurio et al., 1986). Also, the HA genes of C/AA/50 and C/YA/81 - isolated 31 years apart - differ by only 2 nucleotides (Buonagurio et al., 1986). This may indicate that some strains of influenza C virus remain unchanged for decades at a time (Yamashita et al., 1988). HPIV3 isolates obtained during the same year had 9 differences in the F gene (Ont/82a and Ont/82b) and isolates separated by only a year differed from each other by 4 (Ont/82b and Ont/83) to 15 nucleotide differences (Ont/83 and NY/84). Among the HPIV3 isolates examined for variation within the HN gene, no fewer than 22 nucleotide differences (Tex/80 and Tex/83) were observed between any pair of isolates. Isolates obtained one year apart (Tex/82 and Tex/83) were found to differ from each other by 51 nucleotides (van Wyke Coelingh et al., 1988). This divergence within the HPIV3 HN gene may be deceiving since these isolates were plaque purified and as a result may not be representative of the viral populations from which they were derived. On the other hand, plaque purification of these isolates did not appear to alter their reactivity patterns to specific monoclonal antibodies (van Wyke Coelingh et al., 1988).

Variability in the HN gene may result from genetic heterogeneity within natural viral populations rather than the accumulation of mutations over time (van Wyke Coelingh et al., 1988). Although the HN genes of the isolates examined were different from that of the prototype at 164 positions, only ten (6.1%) of these were maintained in all viruses obtained at later dates (van Wyke Coelingh et al., 1988). In the F gene, 60 (31.6%) differences from Wash/57 appear in all isolates obtained in subsequent years (Table 6). This number may be increased to 85 (44.7%) by excluding Ont/87 from consideration (Table 6). It is reasonable to make this exclusion since Ont/87 may represent a separate and distant lineage (Table 7). The 35 (18.4%) differences with respect to Wash/57 that occur only in NY/84 or Ont/87 may also be accumulating (Table 6). Additional isolates from these locales in subsequent years are required to determine which of these are being maintained. Although possible reversions have not been accounted for, the remaining 70 (36.8%) differences are not maintained consistently from their first appearance and therefore may not be accumulating.

The original estimate of difference accumulation in the HN gene (van Wyke Coelingh et al., 1988) may have been underestimated. Only 10 nucleotide substitutions out of a total of 164 appear consistently in all viruses isolated at later dates (van Wyke Coelingh et al., 1988). However, an additional 41 differences occur in all of the isolates with the exception of Aus/124854/74 and 3 more differences are found only in Tex/12677/83. This may be significant since Aus/74 appears to represent a separate lineage. Analysis by the method of Li (1981) confirms that Aus/74 has a relationship to the HPIV3 isolates described by van Wyke

Coelingh et al. (1988) analogous to that of Ont/87 with the isolates sequenced in this study: (('79 ('80 '83))'57 '74) X ('73 '82). In the pair by pair comparison of nucleotide differences among HPIV3 HN genes, Aus/74 displayed the greatest number of differences when compared to the rest of the group (van Wyke Coelingh et al., 1988). This means that 51 (31.1%) of the nucleotide differences in HN with respect to Wash/57 may be accumulating. The 3 differences unique to Tex/83 may also be accumulating but verification requires access to the sequence of more recent isolates from the same area.

The importance of possible geographical influences on genetic relationships is diminished by the high degree of similarity within the HN gene displayed by the 1974 Australian virus and the 1957 Washington prototype (van Wyke Coelingh et al., 1988). Also, the F gene sequence of Wash/57 was most similar to that of Aus/59. Geographical location appears to have little, if any, effect on variation in the F and HN genes among isolates of HPIV3.

Based upon the exclusion of possible separate lineages and the inclusion of differences in the most recent isolate, F has roughly twice the number of differences accumulating (63.2%) as HN (32.9%). By disregarding the unconfirmed accumulating differences in the most recent isolates (eg. Ott/87), these estimates are lowered to 44.7% for the F gene and 31.1% for HN. Whichever set of estimates is considered, it is apparent that differences may be accumulating to a lesser extent in HN than in F. Comparable percentages of total variation have been described within the HN and F genes among the isolates examined in each study: 8.7% (van Wyke Coelingh et al., 1988) and 10.3%, respectively.

Despite this similarity, the two genes appear to be influenced by different selection pressures since they accumulate nucleotide differences at unequal rates.

3. F protein sequence variation among HPIV3 clinical isolates

No differences in charged amino acids are found within the F_2 signal peptide, the F1 amino terminus or the F_1 transmembrane region. Differences within these regions are few in number and conservative in nature, or involve amino acids with nonpolar R groups or polar uncharged R groups. Variation in the HN tail and transmembrane regions are not restricted by number, but by the maintenance of the charged and hydrophobic characteristics of these regions (van Wyke Coelingh et al., 1988). This may also apply to the F protein transmembrane region and cytoplasmic tail. Although the transmembrane region contains more amino acid substitutions than the signal peptide or amino terminus (Table 9), no charge changes are found within the F_1 transmembrane region or tail.

Isolates of RSV accumulate 3.1 times more amino acid changes in the F_2 subunit of the fusion protein than the F_1 subunit (Lopez et al., 1988). Similar results are not evident among isolates of HPIV3. The accumulation of amino acids appears to be 1.5 to 2.0 times greater in the F_1 subunit of HPIV3 than in F_2 depending on the inclusion or exclusion of unconfirmed accumulating residues (Appendix II). The significance of this finding is unclear without F sequence comparisons from isolates of related viruses to determine which of the two F_2/F_1 difference accumulation ratios is more typical of the Paramyxoviridae.

4. Comparison of amino acid differences in the F and HN glycoproteins

The trends evident for the F and HN genes also apply to the predicted F and HN proteins. When compared to Wash/57, the HN protein-572 amino acids long, varied at 36 positions (6.3%) (van Wyke Coelingh et al., 1988) and the F protein - 539 residues long, varied at 24 positions (4.5%). If Aus/74 is assigned to a separate branch, differences at no fewer than 27 positions (75% of all differences) in HN are not maintained consistently. This percentage may be raised to 80.6 if differences found only in Tex/12677/83 - the most recent isolate examined (van Wyke Coelingh et al., 1988), are also not maintained. On the other hand, only 10 positions or 42% of the total differences from Wash/57 are not maintained consistently in F. This estimate may be increased to 62% by including 5 unconfirmed positions. F appears to accumulate more amino acid changes than HN, despite the fact that these proteins among isolates differ from Wash/57 at roughly the same rates (HN, 6.3%; F, 4.5%).

5. The effects of variablity in the F protein on virulence

The virulence of Paramyxoviruses depends at least in part on F_0 susceptibility to cleavage and the availability of the specific host enzyme required for cleavage-activation (Homma and Ohuchi, 1973; Nagai et al., 1976b; Scheid and Choppin, 1977). Since cleavage-activation of the F_0 precursor is required for membrane fusion and viral penetration, changes in cleavage susceptibility can alter virus host range (Hsu et

al., 1987a,b). Studies with Sendai virus have shown that susceptibility of the F protein to protease activation may be affected by a single nucleotide mutation (Hsu et al., 1987a,b; Itoh et al., 1987). Amino acid changes either at or near the cleavage-activation site of Newcastle disease virus (NDV) are associated with avirulent strains (Toyoda et al., 1987; Millar et al., 1988). Although this study found no evidence in terms of CPE or virus yield, differences in virulence may exist among the HPIV3 isolates examined due to amino acid changes involving the HPIV3 F_0 cleavage-activation site.

Based on the sequence determined for Wash/47885/57, cleavage of the HPIV3 F_0 precursor occurs between Arg 109 - the C terminus of F_2, and Phe 110 - the N terminus of F_1 (Spriggs et al., 1986; Côté et al., 1987). Two amino acid differences affecting this cleavage site in Tex/80 and NY/81 include the substitution of Ser for Phe 110 and a charge change resulting from the substitution of Glu for Lys 108. Either one of these differences may be sufficient to decrease F_0 cleavage efficiency. For example, the occurrence of Leu at the N terminus of F_1 is thought to be partly responsible for the avirulence of NDV strain Ulster (Millar et al., 1988). In virulent strains of NDV and other Paramyxoviruses, Phe is normally found in this position (Toyoda et al., 1987). Also, virulent strains of NDV typically have more basic amino acid residues at the F_0 cleavage site than those strains which are avirulent (Toyoda et al., 1987; Millar et al., 1988). With the exception of Tex/80 and NY/81, all of the HPIV3 isolates examined have a predicted protein sequence of Arg 106, Thr 107, Lys 108 and Arg 109 at the F_2 C-terminus. The substitution at amino acid

residue 108 represents the loss of a basic residue for one which is acidic, thus changing the net charge of this region. The effect on F_0 cleavage by these amino acid substitutions is worthy of further examination in future studies.

6. Factors affecting change in negative strand virus glycoproteins

A greater prevalence throughout the population and brief replication cycle per unit time may lead to an increased rate of change in the genomes of influenza A viruses, compared to those of influenza C (Buonagurio et al., 1985). Differences in the evolutionary rates of influenza A and HPIV3 are probably not due to their respective polymerases since each appear to be equally error-prone (van Wyke Coelingh et al., 1985, 1987; Portner et al., 1980).

It is difficult to speculate on minor selective pressures exerted by the partial immune response to HPIV3 infection (van Wyke Coelingh et al., 1988) due to conflicting evidence. Influenza A viruses may cope with long-term immunity to infection (Kendal et al., 1979; Layde et al., 1980) through the rapid evolution of glycoproteins which accumulate substitutions in antigenic regions while maintaining function (van Wyke Coelingh et al., 1988). A partial immunity may allow HPIV3 to survive without having developed surface glycoproteins capable of tolerating change in regions necessary for function. In other words, as a result of the weak selective pressures exerted by the immune response to HPIV3, there is less need for change than with influenza A. However, F sequence comparisons of HPIV3 with those of related viruses

imply that changes in the F protein have occurred among Paramyxoviruses without loss of function. The functional F proteins of human and bovine PIV3 share an amino acid similarity of only 80% (Suzu et al., 1987). Furthermore, only 42-43% similarity exists between the F proteins of HPIV3 and Sendai virus (Blumberg et al., 1985; Shioda et al., 1986). The ability of the F protein to act as a fusogen may rely more on its internal distribution of hydrophobic amino acids (Richardson et al., 1986) and conformation than its overall primary structure.

7. Viral attenuation in cell culture

Viral attenuation may be an unavoidable result of propagation in cell culture. Attenuated viruses currently used in many vaccines are produced by repeated serial passage of the original human wild-type virus in cell culture or embryonated eggs (Kucera and Myrvik, 1985). Relevant examples of this process are the production of the Shwartz and Moraten strains of live measles vaccine which result from multiple passage in chick embryo cell culture (Modlin, 1984). The parent strain itself - Edmonston B, was produced by multiple passage through human kidney cells, human amnion cells, chick embryos and cultured chick embryo cells (Modlin, 1984).

The passage of HPIV3 isolates for F sequence comparison was minimized (Table 1) to reduce the risk of passage-induced change. Selection during initial isolation may have favoured the survival of those virions best suited for or able to adapt to cell culture. Though unavoidable, this might have generated an *in vitro* population of

virions that thrive in cell culture but are less representative of natural populations. However, consistent cell culture techniques and conditions were employed - decreasing the necessity for any further adaptations by the isolates. In any case, the HPIV3 genome appears to be very stable despite differences in passage histories and *in vitro* growth conditions (Storey et al., 1987).

Chapter V - GENERAL CONCLUSIONS: POSSIBLE INTERACTIONS OF THE M PROTEIN WITH THE SURFACE GLYCOPROTEINS OF PARAMYXOVIRUSES

A. Possible Interactions of M Protein with HN

1. DI particles and persistence

Although cytolytic animal viruses destroy infected cells *in vitro*, the reproduction of noncytolytic viruses causes less stress to their host cells. Under certain conditions infection with cytolytic viruses may lead to survival of the infected cell culture and viral persistence (Tuffereau and Roux, 1988). Within the *Rhabdoviridae* and *Paramyxoviridae*, defective interfering (DI) particles cause viral persistence (Holland et al., 1980). DI particles have partially deleted genomes and accumulate after high multiplicity passage of the standard virus. The deletions are caused by template skipping of the replicase, leading to four possible DI genomes: deletion, snapback, panhandle and compound (Lazzarini et al., 1981; Perrault, 1981). DI particles require assistance from standard virions for replication since the DI particles of nonsegmented negative stranded RNA viruses are defective. As DI particles increase in titre, they interfere with the replication of standard virus by depriving the latter of factors required for replication (Lazzarini et al., 1981; Perrault, 1981).

Recent work with Sendai virus persistent infections suggests an association between the M and HN proteins. Cell surface expression of the Sendai virus HN protein appears to depend on the type of infection.

In lytic standard virus infections, HN accumulates at the cell surface in a stable form and is expressed at high levels (Roux et al., 1985). Moderate levels of expression, in which HN reaches the membrane, but is rapidly turned over and re-internalized, results from mixed infections with standard virus and DI particles (Roux and Waldvogel, 1983). In long-term persistent infections, HN is poorly expressed and appears to be degraded before reaching the cell surface (Roux et al., 1985).

M protein is synthesized in baby hamster kidney cells persistently infected with Sendai virus at a normal rate relative to other viral proteins but is structurally unstable (Roux and Waldvogel, 1982). The high rate of intracellular M protein turnover in mixed infections with Sendai standard and DI viruses may favour the turnover of HN at the cell surface (Tuffereau and Roux, 1988). Therefore, an association between HN protein, nucleocapsid and M protein may be required to protect M from degradation and stably anchor HN within the plasma membrane (Tuffereau and Roux, 1988). DI nucleocapsids probably associate with but are unable to stabilize M protein, leaving it susceptible to degradation and incapable of stably anchoring HN within the plasma membrane (Tuffereau and Roux, 1988). Thus, viral budding is restricted, the cells survive, and a persistent infection is established.

2. Charge predictions based on sequence analysis

Charge predictions regarding the HPIV3 glycoproteins have been made possible by sequence analysis of the HN (Storey et al., 1987; van

Wyke Coelingh et al., 1988) and F genes (Côté et al., 1987). The cytoplasmic tail of HN has a charge of +1.5 and should associate more readily with the positively charged M protein than the tail of F_0, which bears a charge of +5.0. These charges were calculated by assuming neutral pH and assigning the following values: Asp -1.0, Glu -1.0, Lys +1.0, Arg +1.0, and His +0.5 (Galinski et al., 1987a). In terms of primary and secondary structure, there appears to be no clustering of negatively charged residues within the cytoplasmic tail of either of the HPIV3 surface glycoproteins.

B. Possible Interactions of M Protein with F_0

1. Persistent infections and temperature sensitive mutants

Early experiments indicated that, unlike HN, the surface expression of F_0 did not depend on the presence of M since F_0 expression is similar in standard, mixed and persistent infections (Roux et al., 1984, Roux and Waldvogel, 1983). More recent work shows that in persistently infected cells F_0 surface expression was reduced and F_0 matured more slowly, although the turnover rate at the cell surface was comparable in the three types of infection (Roux et al., 1985). Decreased F_0 incorporation and infectivity in NDV temperature sensitive mutants may be due to mutation in the matrix protein (Peeples and Bratt, 1984).

2. Amino acid predictions based on sequence analysis

Further evidence of an association between F_0 and M is found in the sequence analysis presented in this thesis. Among the HPIV3 isolates examined, none displayed amino acid variation within the cytoplasmic region (Appendix II). However, amino acid variation was found among HPIV3 isolates within the amino terminal cytoplasmic region of the HN protein (van Wyke Coelingh et al., 1988). These results suggest that the association of M with F_0 is more important than with HN since the primary structure of the cytoplasmic region has been maintained in F_0 but not in HN.

C. Summary

Currently there is evidence for interactions between the *Paramyxovirus* M protein with both F_0 and HN. Work with persistently infected cell lines points to an association between HN and M (Tuffereau and Roux, 1988; Roux et al., 1985; Roux and Waldvogel, 1982) while evidence linking M and F_0 is found in sequence analysis of HPIV3 clinical isolates (this thesis; van Wyke Coelingh et al., 1988) and experiments involving temperature sensitive mutants of NDV (Peeples and Bratt, 1984). Charge predictions, however, made possible by sequence analysis, favour a stronger association of M with HN rather than F_0. In view of the scope of evidence cited, it seems likely that the HPIV3 M protein associates with both surface glycoproteins.

REFERENCES

Baczko, K., Liebert, U.G., Billeter, M., Cattaneo, R., Budka, H. and ter Meulen, V. (1986) Expression of defective measles virus genes in brain tissues of patients with subacute sclerosing panencephalitis. J. Virol. 59:472-478.

Basle, M.F., Russell, W.C., Goswami, K.A., Rebel, A., Giraudon, P., Wild, F. and Filmon, R. (1985) Paramyxovirus antigens in osteoclasts from Paget's bone disease tissue detected by monoclonal antibodies. J. Gen. Virol. 66:2103-2110.

Bellini, W.J., Englund, G., Richardson, C., Rozenblatt, S. and Lazzarini R. (1986) Matrix genes of measles and canine distemper virus: cloning, nucleotide sequences and deduced amino acid sequences. J. Virol. 58:408-416.

Blumberg, B.M., Rose, K., Simona, M., Roux, L., Giorgi, C., Kolakofsky D. (1984) Analysis of the Sendai M gene and protein. J. Virol. 52:656-663.

Blumberg, B.M., Giorgi, C., Roux, L., Raju, R., Dowling, P., Chollet A., and Kolakofsky D. (1985) Sequence determination of the Sendai virus HN gene and its comparison to the influenza glycoproteins. Cell 41:269-278.

Bratt, M. and Gallagher, W. (1969) Preliminary analysis of the requirements for fusion from within and fusion from without by Newcastle disease virus. Proc. Natl. Acad. Sci. U.S.A. 64:536-540.

Bratt, M. and Hightower, L. (1977) Genetics and paragenetic phenomena of paramyxovirus. p.457-533. In: "Comprehensive Virology", Vol.9. Fraenkel-Conrat, H. and Wagner, R.R. (eds). Plenum Press, New York.

Bucher, D.J., Kharitonenkov, I.G., Zakomirdin, J.A., Grigoriev, V.B., Klimenko, S.M. and Davis, J.F. (1980) Incorporation of influenza M-protein into liposomes. J. Virol. 36:586-590.

Büechi, M. and Bächi, T. (1982) Microscopy of internal structures of Sendai virus associated with the cytoplasmic surface of host membranes. Virol. 120:349-359.

Buonagurio, D.A, Susumu, N., Desselberger U., Krystal, M. and Palese P. (1985) Noncumulative sequence changes in the genes of influenza C virus isolates. Virol. 146:221-232.

Buonagurio, D.A, Nadaka, S., Fitch, W.M. and Palese, P. (1986) Epidemiology of influenza C viruses in man: multiple evolutionary lineages and low rates of change. Virol. 153:12-21.

Carlsen, S.R., Peluso, R.W. and Moyer, S.A. (1985) In vitro replication of Sendai virus wild type and defective-interfering particle genome RNA's. J. Virol. 54:493-500.

Carrol, A.R. and Wagner, R.R. (1979) Role of the membrane (M) protein in endogenous inhibition of in vitro transcription by vesicular stomatitis virus. J. Virol. 29:134-142.

Cattaneo, R., Rebmann, G., Baczko, K., ter Meulen, V. and Billeter, M.A. (1987) Altered ratios of measles virus transcripts in diseased human brains. Virol. 160:523-526.

Cattaneo, R., Schmid, A., Rebmann, G., Baczko, K., ter Meulen, V., Bellini, W.J., Rozenblatt, S. and Billeter, M.A. (1986) Accumulated measles virus mutations in a case of subacute sclerosing panencephalitis: interrupted matrix protein reading frame and transcription alteration. Virol. 154:97-107

Chambers, P., Millar, N., Platt, S. and Emmerson, P. (1986) Nucleotide sequence of the gene encoding the matrix protein of Newcastle disease virus. Nuc. Acids Res. 14:9051-9061.

Chanock, R.M. (1956) Association of a new type of cytopathic myxovirus with infantile croup. J. Exp. Med. 104:555-576.

Chanock, R.M. (1979) Parainfluenza viruses. p.611-632 In "Diagnostic Procedures For Viral, Rickettsial and Chlamydial Infections". E.H. Lennette and N.J. Schmidt (eds.). American Public Health Association, Washington, D.C.

Chanock, R.M., Parrott, R.H., Cook, M.K., Andrews, B.E., Bell, J.A., Reichelderfer, T., Kapikian, A.Z., Mastrota, F.M. and Huebner, R.J. (1958) Newly recognized myxoviruses from children with respiratory disease. N. Engl. J. Med. 258:207-213.

Chanock, R.M., Vargosko, A., Lackey, A., Cook, M.K., Kapikian, A.Z., Reichelderfer, T. and Parrott, R.H. (1959) Association of hemadsorption viruses with respiratory illnesses in childhood. J.A.M.A. 169:548-553.

Chanock, R.M., Bell, J.A. and Parrott, R.H. (1961) Natural history of parainfluenza infection. p.126-139. In: "Perspectives in Virology", vol. 2. M. Pollard (ed.). Burgess Ltd., Minneapolis.

Chanock, R.M., Parrott, R.H., Johnson, K.M., Kapikian, A.Z. and Bell, J.A. (1963) Myxoviruses: parainfluenza. Am. Rev. Respir. Dis. 88:152-166.

Chanock, R.M. and McIntosh, K. (1985) Parainfluenza Viruses. p. 1241. In "Virology". B.N. Fields, D.M. Knipe, R.M. Chanock, J.L. Melnick, B. Poizman and R.E. Shope (eds.). Raven Press, New York.

Choppin, P.W. and Compans, R.W. (1970) Phenotypic mixing of envelope proteins of the parainfluenza virus SV5 and vesicular stomatitis virus. J. Virol. 5:609-616.

Choppin, P.W. and Compans, R.W. (1975) Reproduction of paramyxoviruses. p.95-161. In "Comprehensive Virology", vol. 4. H. Fraenkl-Conrat and R.R. Wagner (eds.). Plenum Press, London.

Choppin, P.W. and Scheid, A. (1980) The role of viral glycoproteins in adsorption, penetration and pathogenicity of viruses. Rev. Inf. Dis. 2:40-61.

Clinton, G.M., Little, S.P., Hagen, F.S. and Huang A.S. (1978) The matrix (M) protein of Vesicular Stomatitis Virus regulates transcription. Cell 15:1455-1462.

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

Côté, M.-J., Storey, D.G., Kang, C.Y. and Dimock, K. (1987) Nucleotide sequence of the coding and flanking regions of the human parainfluenza virus type 3 fusion glycoprotein gene. J. Gen. Virol. 68:1003-1010.

Côté, M.-J. (in preparation) The human parainfluenza virus 3 fusion protein: cloning, mapping, sequence analysis and expression. Ph D. thesis. Department of Microbiology and Biology. University of Ottawa.

Dale, R.M.K., McClure, B.A. and Houchins, J.P. (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18 S rDNA. Plasmid 13:31-40.

Damsky, C., Sheffield, J., Tuszynski, G. and Warren, L. (1977) Is there a role for actin in virus budding? J. Cell Biol. 75:593-605.

Dimock, K., Rud, E.W. and Kang, C.Y. (1986) 3' Terminal sequence analysis of human parainfluenza virus three genomic RNA. Nucl. Acids Res. 14:4694.

Dimock, K., Storey, D.G., Côté, M.-J. and Kang, C.Y. (1987) Cloning, coding assignments and mapping of human parainfluenza virus 3. p.213-220. In "The Biology of Negative-Strand Viruses". B. Mahy and D. Kolakofsky (eds.). Elsevier Biomedical Press, Amsterdam.

Dubois-Dalcq, M., Holmes, K.V. and Rentier, B. (1984) Assembly of rhabdoviridae, p.21-43. In: "Assembly of enveloped RNA viruses". Springer-Verlag, Vienna.

Elango, N., Coligan, J.E., Jambou, R.C., and Venkatesan, S. (1986) Human parainfluenza type 3 virus hemagglutinin-neuraminidase glycoprotein: nucleotide sequence of mRNA and limited amino acid sequence of the purified protein. J. Virol. 57:481-489.

Emerson, S.U. (1985) Rhabdoviruses. p. 1124 In "Virology". B.N. Fields, D.M. Knipe, R.M. Chanock, J.L. Melnick, B. Poizman, and R.E. Shope, (eds.) Raven Press, New York.

Fenner, F. and White, D.O. (1986) "Medical Virology". p.532. Academic Press, Orlando, Florida.

Galinski, M.S., Mink, M.A., Lambert, D.M., Wechsler, S.L., and Pons, M.W. (1986a) Molecular cloning and sequence analysis of the human parainfluenza 3 virus RNA encoding the nucleocapsid protein. Virol. 149:139-151.

Galinski, M.S., Mink, M.A., Lambert, D.M., Wechsler, S.L., and Pons, M.W. (1986b) Molecular cloning and sequence analysis of the human parainfluenza 3 virus RNA encoding the P and C proteins. Virol. 154:46-60.

Galinski, M.S., Mink, M.A., Lambert, D.M., Wechsler, S.L., and Pons, M.W. (1987a) Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene encoding the matrix protein. Virol. 157:24-30.

Galinski, M.S., Mink, M.A., and Pons, M.W. (1987b) Molecular cloning and sequence analysis of the human parainfluenza 3 virus genes encoding the surface glycoproteins, F and HN. Virus Res. 8:205-215.

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

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

Geliebter, J. (1987) Dideoxynucleotide sequencing of RNA and uncloned cDNA. Focus 9:1, 5-8.

Genty, N. and Bussereau, F. (1980) Is cytoskeleton involved in vesicular stomatitis virus reproduction? J. Virol. 34:777-781.

Ginsberg, H.S. (1988) Paramyxoviruses. p.239-259. In "Virology", 2nd ed. by R. Dulbeco and H.S. Ginsberg. J.B. Lippincott Company, Philadelphia.

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

Giuffre, R., Tovell, D., Kay, C. and Tyrrell, D. (1982) Evidence for an interaction between the membrane protein of a paramyxovirus and actin. J. Virol. 42:963-968.

Glezen, W.P. and Denny, F.W. (1973) Epidemiology of acute lower respiratory disease in children. N. Engl. J. Med. 288:498-505.

Glezen, W.P., Frank, A.L. and Taber, L.H. (1981) Risk of respiratory syncytial (RS) and parainfluenza type 3 (Para 3) infection in young children. Pediatr. Res. 15:611.

Goswami, K.K.A., Cameron, K.R., Russel, W.C., Lange, L.S. and Mitchell, D.N. (1984). Evidence for the persistence of paramyxoviruses in human bone marrows. J. Gen. Virol. 65:1881-1888.

Gregoriades, A. (1973) The membrane protein of influenza virus: extraction from virus and infected cells with acidic-chloroform methanol. Virol. 54:369-383.

Gregoriades, A. and Frangione, B. (1981) Insertion of influenza M protein into the lipid bilayer and localization of site of insertion. J. Virol. 40:323-328.

Griffin, J.A. and Compans, R.W. (1979) Effect of cytochalasin-B on the maturation of enveloped viruses. J. Exp. Med. 150:379-391.

Gross, P.A., Green, R.H. and Curnen, M.G.M. (1973) Persistent infection with parainfluenza type 3 virus in man. Am. Rev. Res. Dis. 108:894-898.

Gupta, K.C., and Kingsbury, D.W. (1982) Conserved polyadenylation signals in two negative-strand RNA virus families. Virol. 120:518-523.

Gupta, K.C., and Kingsbury, D.W. (1984) Complete sequences of the intergenic and mRNA start signals in the Sendai virus genome-homologies with the genome of vesicular stomatitis virus. Nucl. Acids Res. 12:3829-3841.

Gupta, K.C., and Kingsbury, D.W. (1985) Polytranscripts of Sendai virus do not contain intervening polyadenylate sequences. Virol. 141:102-109.

Heggeness, M.H., Smith, P.R. and Choppin, P.W. (1982) In vitro assembly of the nonglycosylated membrane protein (M) of Sendai virus. Proc. Natl. Acad. Sci. USA 79:6232-6235.

Hidaka, Y., Kanda, T., Iwasaki, K., Nomota, A., Shioda, T. and Shibuta, H. (1984) Nucleotide sequence of a Sendai virus genome region covering the entire M gene and 3' proximal 1013 nucleotides of the F gene. Nuc. Acids Res. 12:7965-7973.

Holland, J.J., Kennedy, S.I.T., Semler, B.L., Jones, C.L., Roux, L. and Grabau, E.A. (1980) Defective interfering RNA viruses and the host cell response. In: "Comprehensive Virology", vol. 16, pp. 137-192. H. Fraenkel-Conrat and R.R. Wagner (eds.), Plenum, New York.

Homma, M. and Ohuchi, M. (1973) Trypsin action on the growth of Sendai virus in tissue culture cells. III. Structural differences of Sendai viruses grown in eggs and tissue culture cells. J. Virol. 12:1457-1465.

Hope-Simpson, R.E. (1981) Parainfluenza virus infections in the Cirencester survey: seasonality and other characteristics. J. Hyg. Cam. 87:393-406.

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

Hsu, M., Scheid, A. and 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.

Hsu, M., Scheid, A. and Choppin, P.W. (1983) Fusion of Sendai virus with liposomes: dependance on the viral fusion protein (F)and the lipid composition of liposomes. Virol. 126:361-369.

Hsu, M.-C., Scheid, A. and Choppin, P.W. (1987a). Protease activation mutants of Sendai virus: sequence analysis of the mRNA of the fusion (f) protein gene and direct identification of the cleavage-activation site. Virology 156, 84-90.

Hsu, M.-C., Scheid, A. and Choppin, P.W. (1987b). mRNA sequence analysis of the fusion (f) protein gene of the protease activation mutants of Sendai virus. In: The Biology of Negative-Strand Viruses, pp.10-17. Edited by B. Mahy & D. Kolakofsky. Amsterdam: Elsevier Biomedical Press.

Hull, J., Krah, D. and Choppin, P.W. (1987) Resistance of a measles virus mutant to fusion inhibiting oligopeptides is not associated with mutations in the fusion peptide. Virol. 159:368-372.

Hull, R.N., Minner, J.R. and Smith, J.W. (1956) New viral agents recovered from tissue cultures of monkey kidney cells. Am. J. Hyg. 63; 204-215.

Itoh, M., Shibuta, H. and Homma, M. (1987) Single amino acid substitution of sendai virus at the cleavage site of the fusion protein confers trypsin resistance. J. Gen. Virol. 68:2939-2944.

Jambou, R.C., Narayanasamy, E., Sundararajan, V. and Collins, P.L. (1986) Complete sequence of the major nucleocapsid protein gene of human parainfluenza type 3 virus: comparison with other negative strand virus. J. Gen. Virol. 67:2543-2548.

Kendal, A.P., Joesph, J.M. and Dowdle, W.R. (1979) Periods of prevalence of H1N1 and H3N2 influenza A strains. Amer. J. Epidemiol. 110:449-461.

Kingsbury, D.W. (1977) Paramyxoviruses, p. 349-382. In: "The molecular biology of animal viruses. D.P. Nayak (ed.). Marcel Dekker, New York.

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

Klenk, H.D., Compans, R.W. and Choppin, P.W. (1970) An electron microscopic study of the presence or absense of neuramindase in enveloped viruses. Virol. 42:1158-1162.

Kucera, L.S. and Myrvik, Q.N. (1985) "Fundamentals of Medical Virology", p.180. Lea and Febiger, Philadelphia.

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

Lamb, R.A. and Choppin, P.W. (1983) The gene structure and replication of influenza virus. Annu. Rev. Biochem. 52:467-506.

Lamb, R.A. and Lai, C.-J. (1981) Conservation of the influenza virus membrane protein (M_1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M_2) in H1N1 and H3N2. Virol. 112:746-751.

Lamb, R.A., Lai, C.-J. and Choppin, P.W. (1981) Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proc. Natl. Acad. Sci. USA 78:4170-4174.

Layde, P.M., Englelberg, A.L., Dobbs, H.I. and Noble, G.R. (1980) Outbreak of influenza A/USSR/77 at Marquette University. J. Infect. Dis. 142:347-352.

Lazzarini, R.A., Keene, J.D. and Schubert, M. (1981) The origins of defective interfering particles of the negative-strand RNA viruses. Cell 26:145-154.

Li, W.-H. (1981) Simple method for constructing phylogenetic trees from distance matrices. Proc. Natl. Acad. Sci. USA 78:1085-1089.

Lehninger, A.L. (1975) Biochemistry, 2nd ed. p.72-76. Worth Publishers, Inc., New York.

Liebert, U.G., Baczko, K., Budka, H. and ter Meulen, V. (1986) Restricted expression of measles virus proteins in brains from cases of subacute sclerosing panencephalitis. J. Gen. Virol. 67:2435-2444.

Lodish, H.F. and Porter, M. (1980) Heterogeniety of vesicular stomatitis virus particles: implications for virion assembly. J. Virol. 33:52-58.

Lohmeyer, J., Talens, L.T. and Klenk, H.D. (1979) Biosynthesis of the influenza virus envelope in abortive infection. J. Gen. Virol. 42:73-88.

Lopez, J.A., Villaneuva, N., Melero, J.A. and 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. 10:249-262.

Luk, D., Sanchez, A. and Banerjee, A.K. (1986) Messenger RNA encoding the phosphoprotein (P) of human parainfluenza virus 3 is bicistronic. Virol. 153:318-325.

Luk, D., Masters, P.S., Sanchez, A. and Banerjee, A.K. (1987) Complete nucleotide sequence of the matrix protein mRNA and three intergenic junctions of human parainfluenza virus type 3. Virol. 156:189-192.

Lyles, D.S. (1979) Glycoproteins of Sendai virus are transmembrane proteins. Proc. Natl. Acad. Sci. USA 76:5621-5625.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) "Molecular Cloning - A Laboratory Manual". p. 115. Cold Spring Harbour, New York.

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

Matsumoto, T. (1982) Assembly of paramyxoviruses. Micro. and Immuno. 26:285-320.

Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labelled DNA with base specific chemical cleavage. p. 499-561. In: "Methods In Enzymology", vol. 65. D. Kolowick and G. Kaplan (eds.), Academic Press, New York.

McGinnes, L. and Morrison, T. (1987) The nucleotide sequence of the gene encoding the Newcastle disease virus membrane protein and comparisons of membrane protein sequences. Virol. 156:221-228.

McSharry, J.J., Compans, R.W., Lackland, H. and P.W. Choppin 1975) Isolation and characterization of the nonglycosylated membrane protein and a nucleocapsid complex from the paramyxovirus SV5. Virol. 67:365-374.

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

Messing, J. (1983) New M13 vectors for cloning. p.20-78. In "Methods in Enzymology". vol. 101. R. Wu, L. Grossman and K. Moldave (eds.). Academic Press, New York.

Millar, N.S., Chambers, P. and Emmerson, P.T. (1988) Nucleotide sequence of the fusion and hemagglutinin-neuraminidase glycoprotein genes of Newcastle disease virus, strain ulster: molecular basis for variations in pathogenicity between strains. J. Gen. Virol. 69:613-620.

Mills, B.G., Singer, F.R., Weiner, L.P. and Holst, P.A. (1981) Immunohistological demonstration of respiratory syncitial virus antigens in Paget disease of bone. Proc. Natl. Acad. Sci. USA 78:1209-1213.

Modlin, J.F. (1984) Measles virus. p.349. In "Textbook of Human Virology", R.B. Belshe (ed.). PSG Publishing Company, Inc., Massachusetts.

Morgan, E.M. and Rapp, F. (1977) Measles virus and its associated diseases. Bact. Rev. 41:636-666.

Morita, K., Vanderoef, R. and Lenard, J. (1987) Phenotypic revertants of temperature-sensitive M protein mutants of vesicular stomatitis virus: sequence analysis and functional characterization. J. Virol. 61:256-263.

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

Mufson, M.A., Mocega, H.E. and Krause, H.E. (1973) Acquisition of parainfluenza 3 virus by hospitalized children. I. Frequencies, rates and temporal data. J. Infect. Dis. 128:141-147.

Murphy, F.A. (1985) Virus Taxonomy. p. 18. In "Virology". B.N. Fields, D.M. Knipe, R.M. Chanock, J.L. Melnick, B. Poizman and R.E. Shope (eds.). Raven Press, New York.

Nagai, Y., Yoshida, T., Yoshii, S., Maeno, K. and Matsumoto, T. (1975) Modification of normal cell surface by smooth membrane preparations from BHK-21 cells infected with Newcastle disease virus. Med. Microbiol. Immunol. 161:175-188.

Nagai, Y., Ogura, H. and Klenk, H.-D. (1976a) Studies on the assembly of the envelope of Newcastle disease virus. Virol. 69:523-538.

Nagai, Y., Klenk, H.-D. and Rott, R. (1976b). Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of Newcastle disease virus. Virol. 72, 494-508.

Newcomb, W.W. and Brown, J.C. (1981) Role of vesicular stomatitis virus matrix proteins in maintaining the viral nucleocapsid in the condensed form found in native virions. J. Virol. 39:295-299.

Newcomb, W.W., Tobin, G.J., McGowan, J.J. and Brown, J.C. (1982) In vitro reassembly of vesicular stomatitis virus skeletons. J. Virol. 41:1055-1062.

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

Ono, K., Dubois-Dalcq, M.E., Schubert, M. and Lazzarini, R.A. (1987) A mutated membrane protein of vesicular stomatitis virus has abnormal distribution within the infected cell and causes defective budding. J. Virol. 61:1332-1341.

Orvell, C. and Norrby, E. (1977). Immunologic properties of purified Sendai virus glycoproteins. J. Immunol. 119, 1882-1887.

Parott, R.H., Vargoski, A.J., Kim, H.W., Bell, J.A. and Chanock, R.M. (1962) Respiratory diseases of viral etiology. III. Myxoviruses: parainfluenza. Am. J. Pub. Heal. 52:907-917.

Paterson, R.G. and Lamb, R.A. (1987) Ability of the hydrophobic fusion-related external domain of a paramyxovirus F protein to act as a membrane anchor. Cell 48:441-452.

Peeples, M. and Bratt, M. (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.

Pepinsky, R. and Vogt, V. (1979) Identification of retroviral matrix proteins by lipid-protein crosslinking. J. Mol. Biol. 131:819-837.

Pepinsky, R. and Vogt, V. (1984) Fine structure analysis of lipid-protein and protein-protein interactions of gag protein p19 of the avian sarcoma and leukemia viruses by cyanogen bromide mapping. J. Virol. 52:145-153.

Perrault, J. (1981) Origin and replication of defective interfering particles. Curr. Top. Microbiol. Immunol. 93:151-207.

Portner, A., Webster, R.G. and Bean, W.J. (1980) Similar frequencies of antigenic variants in Sendai, vesicular stomatitis, and influenza A viruses. Virol. 104:235-238.

Portner, A., Scroggs, R. and Naeve, C. (1987) The fusion glycoprotein of Sendai virus: sequence analysis of an epitope involved in fusion and virus neutralization. Virol. 157, 556-559.

Prinoski, K., Côté, M.-J., Kang, C.Y., and Dimock, K. (1987) Nucleotide sequence of the human parainfluenza virus 3 matrix protein gene. Nucl. Acids Res. 15:3182.

Ray, R., Brown, V.E. and Compans, R.W. (1985). Glycoproteins of human parainfluenza virus type 3: characterization and evaluation as a subunit vaccine. J. Inf. Dis. 152:1219-1230.

Ray, R. and Compans, R.W. (1987). Glycoproteins of human parainfluenza virus type 3: affinity purification, antigenic characterization and reconstitution into lipid vesicles. J. Gen. Virol. 68:409-418.

Ray, R., Glaze, B.J. and Compans, R.W. (1988). Role of individual glycoproteins of human parainfluenza virus type 3 in the induction of a protective immune response. J. Virol. 62:783-787.

Rayssiguier, C., Cioe, L., Withers, E., Wunner, W.H. and Curtis, P.J. (1986) Cloning of rabies virus matrix protein mRNA and determination of its amino acid sequence. Virus Res. 5:177-190.

Richardson, C.D. and Choppin, P.W. (1983) Oligopeptides that specifically inhibit membrane fusion by paramyxoviruses: studies on the site of action. Virol. 131:518-532.

Richardson, C.D., Scheid, A. and Choppin, P.W. (1980) Specific inhibition of paramyxovirus and myxovirus replication by oligopeptides with amino acid sequences similar to those at the N-termini of the F_1 or HA_2 viral polypeptides. Virol. 105:205-222.

Richardson, C., Hull, D., Greer, P., Hasel, K., Berkovich, A., Englund, G., Bellini, W., Rima, B. and Lazzarini, R. (1986). The nucleotide sequence of the mRNA encoding the fusion protein of measles virus (edmonston strain): a comparison of fusion proteins from several different paramyxoviruses. Virol. 155:508-523.

Robinson, W.S. (1971) Sendai virus RNA synthesis and nucleocapsid formation in the presence of cyclohexamide. Virol. 44:494-502.

Rose, J.K. (1980) Complete intergenic and flanking sequences from the genome of vesicular stomatitis virus. Cell 19:415-421.

Rose, J.K. and Gallione, C.J. (1981) Nucleotide sequences of the mRNA's encoding the Vesicular Stomatitis Virus G and M proteins determined from cDNA clones containing the complete coding regions. J. Virol. 39:519-528.

Roux, L., Beffy, P. and Portner, A. (1984) Restriction of cell surface expression of Sendai virus hemagglutinin-neuraminidase glycoprotein correlates with its higher instability in persistently and standard plus defective interfering virus infected BHK-21 cells. Virol. 138:118-128.

Roux, L., Beffy, P. and Portner, A. (1985) Three variations in cell surface expression of the haemagglutinin-neuraminidase glycoprotein of Sendai virus. J. Gen. Virol. 66:987-1000.

Roux, L. and Waldvogel, F.A. (1982) Instability of the viral M protein in BHK-21 cells persistently infected with Sendai virus. Cell 28:293-302.

Roux, L. and Waldvogel, F.A. (1983) Defective interfering particles of Sendai virus modulate HN expression at the surface of infected BHK cells. Virol. 130:91-104.

Rozenblatt, S., Eizenberg, O., Ben-Levy, R.,Lavie, V. and Bellini, W.J. (1985) Sequence homology within the morbilliviruses. J. Virol. 53:684-690.

Sanchez, A., Banerjee, A.K., Furuichi, Y. and Richardson, M.A. (1986) Conserved structures among the nucleocapsid proteins of the Paramyxoviridae: complete nucleotide sequence of the human parainfluenza virus type 3 NP mRNA. Virol. 152:171-180.

Sakai, Y., Suzu, S., Shioda, T. and Shibuta, H. (1987) Nucleotide sequence of the bovine parainfluenza 3 virus genome: its 3' end and the genes of NP, P, C and M proteins. Nuc. Acids Res. 15:2927-2944.

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

Sato, H., Ohihira, M., Ishida, N., Imamura, Y., Hattori, S. and Kawahita, M. (1987). Molecular cloning and nucleotide sequence of the P, M and F genes of Newcastle disease virus avirulent strain D26. Virus Res. 7:187-280.

Scheid, A. and Choppin, P.W. (1974) Identification of biological activities of paramyxovirus glycoproteins. Activation of cell fusion, hemolysis and infectivity by proteolytic cleavage of an inactive precursor protein of Sendai virus. Virol. 57:475-490.

Scheid, A. and Choppin, P.W. (1976) Protease activation mutants of Sendai virus. Virol. 69:265-277.

Scheid, A. and Choppin, P.W. (1977). Two disulphide-linked polypeptide chains constitute the active F protein of paramyxoviruses. Virol. 80: 54-66.

Shimizu, K. and Ishida, N. (1975) The smallest protein of Sendai virus: Its candidate function of binding nucleocapsid to envelope. Virol. 67:427-437.

Shioda, T., Hidaka, Y., Kanda, T., Shibuta, T., Namoto, A. and Iwasaki, K. (1983) Sequence of the 3687 nucleotides from the 3' end of Sendai virus genomic RNA and the predicted amino acid sequences of viral NP, P and C proteins. Nucl. Acids Res. 11:7317-7331.

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

Simons, K. and Garoff H. (1980) The budding mechanisms of enveloped animal viruses. J. Gen. Virol. 50:1-21.

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

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

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

Spriggs, M.K., Johnson P.R., and Collins, P.L. (1987a) Sequence analysis of the matrix protein genes of human parainfluenza virus type 3: extensive sequence homology among paramyxoviruses. J. Gen. Virol. 68:1491-1497.

Spriggs, M.K., Murphy, B.R., Prince, G.A., Olmsted, R.A. & Collins, P.L. (1987b). Expression of the F and HN glycoproteins of human parainfluenza virus type 3 by recombinant vaccinia viruses: contributions of the individual proteins to host immunity. J. Virol. 61:3416-3423.

Spriggs, M.K., Collins, P.L., Tierney, E., London, W.T. & Murphy, B.R. (1988). Immunization with vaccinia virus recombinants that express the surface glycoproteins of human parainfluenza virus type 3 (PIV3) protects patas monkeys against PIV3 infection. J. Virol. 62:1293-1296.

Stone, H.O., Portner, A. and Kingsbury, D.W. (1972) Ribonucleic acid transcriptase in Sendai virions and infected cells. J. Virol. 8:174-180.

Storey, D.G., Dimock, K. and Kang, C.Y. (1984) Structural characterization of virion proteins and genomic RNA of human parainfluenza virus 3. J. Virol. 52:761-766.

Storey, D.G. (1987) "Structural Characterization of Human Parainfluenza Virus 3 and Cloning of Viral Specific Genes" (Ph.D. Thesis). University of Ottawa, Department of Microbiology and Immunology.

Storey, D.G., Côté, M.-J., Dimock, K., and Kang, C.Y. (1987) Nucleotide sequence of the coding and flanking regions of the human parainfluenza virus 3 hemagglutinin-neuraminidase gene: comparison with other paramyxoviruses. Intervirol. 27:69-80.

Strauss, E.G. and Strauss, J.H. (1985) Assembly of enveloped animal viruses. p.206-214. In "Virus structure and assembly". S. Casjens (ed.) Jones and Bartlett Publishers, Boston.

Suzu, S., Sakai, Y., Shioda, T. and Shibuta, H. (1987). Nucleotide sequence of the bovine parainfluenza 3 virus genome: the genes of the F and HN glycoproteins. Nucl. Acids Res. 15:2945-2958.

Thompson, S.D. and Portner, A. (1987) Localizaton of functional sites on the hemaglutinin-neuraminidase glycoprotein of Sendai virus by sequence analysis of antigenic and temperature-sensitive mutants. Virol. 160:1-8.

Toyoda, T., Sakaguchi, T. and Imai, K. (1987) Structural comparison of the cleavage-activation site of the fusion glycoprotein between virulent and avirulent strains of Newcastle disease virus. Virol. 158:242-247.

Tuffereau, C. and Roux, L. (1988) Direct adverse affects of Sendai virus DI particles on virus budding and on M protein fate and stability. Virol. 162:417-426.

van Wyke Coelingh, K.L., Winter, C.C. and Murphy, B.R. (1985) Antigenic variation in the HN protein of HPIV3. Virol. 143:569-582.

van Wyke Coelingh, K.L., Winter, C.C. Jorgensen, E.D. and Murphy, B.R. (1987) Antigenic and structural properties of HPIV3: sequence analysis of variants selected with monoclonal antibodies which inhibit infectivity, hemagglutination and neuraminidase activities. J. Virol. 61:1473-1477.

van Wyke Coelingh, K.L., Winter, C.C. and Murphy, B.R. (1988) Nucleotide and deduced amino acid sequence of HN genes of HPIV3 isolates from 1957 to 1983. Virol. 162:137-143.

Wechsler, S.L., Lambert, D.M., Galinski, M.S., Heineke, B.E. and Pons, M.W. (1985) Human parainfluenza virus three: purification and characterization of sub-viral components, viral proteins and viral RNA. Virus Res. 3:339-351.

Weiss, R.A. and Bennett, P.L.P. (1980) Assembly of membrane glycoprotein studied by phenotypic mixing between mutants of vesicular stomatis virus and retrovirus. Virol. 100:252-274.

Welliver, R.C., Wong, D.T., Choi, T.S. and Ogra, P.L. (1982) Natural history of parainfluenza virus infection in childhood. J. Pediatr. 101:180-187.

Wiener, J.R., Pal, R., Barenholz, Y. and Wagner, R.R. (1983) Influence of the VSV M protein on the membrane dynamics of mixed phospholipid vesicles. Biochem. 22:2162-2170.

Wilson, T. and Lenard, J. (1981) Interaction of wild-type and mutant M protein of vesicular stomatitis virus with nucleocapsids in vitro. Biochem. 20:1349-1354.

Wright, P.F. (1984) Parainfluenza viruses p.299. In "Textbook of Human Virology" R.B. Belshe (ed.). PSG Publishing Co., Littleton Massachusetts.

Yamashita, M., Krystal, M., Fitch, W.M. and Palese, P. (1988) Influenza B virus evolution: co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. Virol. 163:112-122.

Yoshida, T., Nagai, Y., Yoshii, S., Matsumoto, T. and Hoshino, M. (1976) Membrane (M) protein of HVJ (Sendai virus): Its role in virus assembly. Virol. 71:143-161.

Yoshida, T., Nagai, Y., Maeno, K., Tinuma, M., Hamaguchi, M., Matsumato, T., Nagayoshi, S. and Hoshimo, M. (1979) Studies on the role of M-protein in virus assembly using a ts mutant HVJ (Sendai virus). Virol. 92:139-154.

Zakowski, J.J. and Wagner, R.R. (1980) Localization of membrane-associated proteins in vesicular stomatitis virus by use of hydrophobic membrane probes. J. Virol. 36:93-102.

Zakowski, J.J., Petri, W.A. Jr. and Wagner, R.R. (1981) Role of matrix protein in assembling the membrane of vesicular stomatitis virus: reconstitution of matrix protein with negatively charged phospholipid vesicles. Biochem. 20:3902-3907.

Zavada, J. (1976) Viral psuedotypes and phenotypic mixing. Arch. Virol. 50:1-15.

APPENDIX I: F GENE SEQUENCE COMPARISON AMONG ISOLATES

Sequences are shown in message sense and differences with respect to prototype (Wash/57) are indicated according to isolate. The F gene transcriptional initiation and polyadenylation signals are overlined. Asterisks indicate translational initiation and termination codons, and the coding region for the cleavage-activation site between F_1 and F_2 is shown by the arrow (Spriggs et. al, 1986). Inverted arrows reveal the nucleotides coding for potential glycosylation sites. Regions coding for major hydrophobic sequences are underlined: A, the signal peptide; B, the F_1 amino terminus, and; C, the membrane anchor. Lower case letters indicate ambiguous nucleotide postions.

```
            1                                                  50
Wash/57     AGGACAAAAG AAGTCAATAC CAACAACTAT TAGCAGCCAC ACTCGCTGGA
Aus/59                                       C     T        A C
Wash/75                 G                          T  T       A
Tex/80                  G                C         T  T       AA
NY/81                   G         T      C         T  T       AA
Ont/82(a)               G                C         T  T       AA
Ont/82(b)               G                C         T  T       AA
Ont/83                  G                C         T  T       AA
NY/84                   G                C         T  T       AA
Ont/87                        C                    T  T   GT  AA

            51                                                100
Wash/57     ACAAGAAAGA AGGGATAAAA AAAGTTTAAC AGAAGAAACA AAAACAGAAA
Aus/59                              G C                       A
Wash/75                     C         CAA  T   G     TC       A
Tex/80       T               TT         A  T   G     TC       A
NY/81        T               TT         A  T   G     TC       A
Ont/82(a)    T               TT         AG T   G     TC       A
Ont/82(b)    T               TT         A  T   G     TC       A
Ont/83       T               TT         A  T   G     TC       A
NY/84        T               TC        CA  T   G     TC       A
Ont/87                   A   TC         A  T  AG     TC       A  G

            101                                               150
Wash/57     GCACAGAACA CCAGAACAAC AAGATCAAAA CACCCAACCC ATTCAAAACG
Aus/59       T                                     a     C      A
Wash/75      T       G              A  T       T T  T            A
Tex/80       T                  T   A          T     T   C       A
NY/81        T                  T              T     T   C       A
Ont/82(a)    T            A     T              T     T   C       A
Ont/82(b)    T              G                  T     T   C       A
Cnt/83       T              G                  T     T   C       A
NY/84          T                               T     T   C       A
Ont/87       T T         C   T      A          T    T      TT    A

                                                               (A)
            151                                             *  200
Wash/57     AAAATCTCAA AAGAGATTGG CAACACAACA AACACTGAAC ATCATGCCAA
Aus/59                    a                   G          CA
Wash/75           C         aCC               G          CA
Tex/80           TC         aCC   T  T        G      G   CA
NY/81            TC          CC   T  T        G      G   CA
Ont/82(a)        TC          CC   T  T        G      G   CA
Ont/82(b)        TC          CC   T  T        G      G   CA
Ont/83           TC          CC   T  T        G      G   CA
NY/84            TC          CC      T        G      G   CA
Ont/87           TC         aCC               Gc  CA      CA
```

```
              201                                                 250
Wash/57       CCTCAATACT GCTAATTATT ACAACCATGA TTATGGCATC TTTCTGCCAA
Aus/59
Wash/75        T                                C
Tex/80         T            C                   C
NY/81          T                                C
Ont/82(a)      T                                C
Ont/82(b)      T                                C
Ont/83         T                                C               T
NY/84          T                                C
Ont/87         T T                       A                  C

              251                                                 300
Wash/57       ATAGATATCA CAAAACTACA GCATGTAGGT GTATTGGTTA ACAGTCCCAA
Aus/59
Wash/75                                                C      C
Tex/80                                                 C
NY/81               G        G                         C
Ont/82(a)                                              C
Ont/82(b)          C                                   C
Ont/83             C                                   C
NY/84                                                  C
Ont/87                                                 C

              301                                                 350
Wash/57       AGGgATGAAG ATATCACAAA ACTTTGAAAC AAGATATCTA ATTTtGAGCC
Aus/59
Wash/75       a                         C
Tex/80        a                                         g G
NY/81                                                   G G
Ont/82(a)                                               G G
Ont/82(b)                                               G G
Ont/83                                                  G G
NY/84                                                   G G
Ont/87        a                         c

              351                                                 400
Wash/57       TCATACCAAA AATAGAAGAT TCTAACTCTT GTGGTGACCA ACAGATCAAG
Aus/59                                          t  c        a
Wash/75                           C
Tex/80                            C
NY/81                             C
Ont/82(a)                         C
Ont/82(b)                         C
Ont/83                            C
NY/84                             C
Ont/87                            C               t                  A
```

```
              401                                                450
Wash/57       CAATACAAGA GGTTATTGGA TAGACTGATC ATTCCTTTAT ATGATGGATT
Aus/59                                  T
Wash/75                                            C               G
Tex/80                   A C                       C
NY/81                    A C                       C
Ont/82 (a)               A C                       C
Ont/82 (b)         T     A C                       C
Ont/83                   A C                       C
NY/84                    A C                       C
Ont/87                                             C   C

              451                                                500
Wash/57       AAGATTACAG AAGGATGTGA TAGTGTCCAA TCAAGAATCC AATGAAAACA
Aus/59
Wash/75                                 AA
Tex/80                     A            AA
NY/81                      A            AA
Ont/82 (a)                 A            AA
Ont/82 (b)                 A            AA
Ont/83                     A            AA
NY/84                      A            AA
Ont/87                     A            AA

                                                                 (B)
                                   ↓
              501                                                550
Wash/57       CTGACCCCAG AACAAAACGA TTCTTTGGAG GGGTAATTGG AACTATTGCT
Aus/59            T
Wash/75           T                                           C
Tex/80            T  T       G       C
NY/81                T       G       C
Ont/82 (a)        T  T
Ont/82 (b)        T
Ont/83            T
NY/84             T
Ont/87            T           G      c                        C

              551                                                600
Wash/57       CTGGGAGTGG CAACCTCAGC ACAAATTACA GCGGCAGTTG CTCTGGTTGA
Aus/59                a
Wash/75               A                   C         G
Tex/80                A                                     T
NY/81                 A                                     T
Ont/82 (a)            A                   C                 T
Ont/82 (b)            A                   C                 T
Ont/83                A                   C                 T
NY/84                 A                                     T
Ont/87                A
```

```
              601                                               650
Wash/57       AGCCAAGCAG GCAAGATCAG ACATTGAAAA ACTCAAGGAA GCAATCAGGG
Aus/59                 A
Wash/75          T                                    A
Tex/80                                      C             A          T
NY/81                                       C             A          T
Ont/82(a)                                   C             A          T
Ont/82(b)                                   C             A          T
Ont/83                                      C             A          T
NY/84                                       C             A          T
Ont/87                       A              A             A

              651                                               700
Wash/57       ACACAAACAA AGCAGTGCAG TCAGTCCAGA GCTCCATAGG AAATTTGATA
Aus/59
Wash/75                                  T                      A
Tex/80               T                   T                    C A
NY/81                T                   T                    C A
Ont/82(a)            T                   T                    C A
Ont/82(b)            T                   T                    C A
Ont/83               T                   T                    C A
NY/84                t                   T                    C A
Ont/87                                   T          T           A

              701                                               750
Wash/57       GTAGCAATTA AATCGGTCCA GGATTATGTC AACAAAGAAA TCGTGCCATC
Aus/59
Wash/75                      A
Tex/80                       A                T
NY/81                        A                T
Ont/82(a)                    A                T
Ont/82(b)                    A                T
Ont/83                       A                T
NY/84                        A                T
Ont/87                       A      A                C

              751                                               800
Wash/57       AATTGCGAGA TTAGGTTGTG AAGCAGCAGG ACTTCAGTTA GGAATTGCAT
Aus/59        G
Wash/75       G          C                                 A
Tex/80        G        G C                                 A
NY/81         G        G C                                 A
Ont/82(a)     G        G C                                 A
Ont/82(b)     G        G C                                 A
Ont/83        G        G C                                 A
NY/84         G        G C                                 A
Ont/87        G     T    C                                 A
```

```
              801                                                850
Wash/57       TAACACAGCA TTACTCAGAA TTAACAAACA TATTCGGTGA TAACATAGGA
Aus/59
Wash/75           G                                T
Tex/80                                             T
NY/81                                              T
Ont/82(a)                                          T
Ont/82(b)                                          T
Ont/83                                             T
NY/84                                              T
Ont/87                                             T

              851                                                900
Wash/57       TCGTTACAAG AAAAAGGGAT AAAATTACAA GGTATAGCAT CATTATACCG
Aus/59
Wash/75                          A
Tex/80                           A
NY/81           A                A                          C
Ont/82(a)                        A
Ont/82(b)                        A
Ont/83                           A
NY/84                            A
Ont/87                           A

              901 →        ←                                     950
Wash/57       CACAAATATC ACAGAGATAT TCACAACATC AACAGTTGAT AAATATGATA
Aus/59                                                            C
Wash/75                       A
Tex/80                        A
NY/81                         A
Ont/82(a)                     A
Ont/82(b)                     A
Ont/83                        A
NY/84                         A
Ont/87                        A

              951                                               1000
Wash/57       TTTATGATCT ATTATTTACA GAATCAATAA AGGTGAGAGT TATAGATGTT
Aus/59
Wash/75        C
Tex/80                   G
NY/81                    G
Ont/82(a)                G
Ont/82(b)                G
Ont/83                   G
NY/84                    G
Ont/87         C
```

```
              1001                                             1050
Wash/57       GACTTGAATG ATTACTCAAT CACCCTCCAA GTCAGACTCC CTTTATTAAC
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87          T
```

```
              1051                                             1100
Wash/57       TAGACTGCTG AACACCCAGA TTTACAAAGT AGATTCCATA TCATACAACA
Aus/59           g                                              T
Wash/75          g             T      C                         T  T
Tex/80           g             T      C                         T
NY/81                          T      C                         T
Ont/82(a)                      T      C
Ont/82(b)                      T      C                         T
Ont/83                         T      C                         T
NY/84                          T      C
Ont/87           g             T      C                         T  T
```

```
              1101                                             1150
Wash/57       TCCAAAACAG AGAATGGTAT ATCCCTCTTC CCAGCCACAT CATGACAAAA
Aus/59
Wash/75              T                  T               T         G
Tex/80                                                  T         G
NY/81                                                   T         G
Ont/82(a)                                               T         G
Ont/82(b)                                               T
Ont/83                                                  T
NY/84                T                                  T         G
Ont/87                                                  T         G
```

```
              1151                                             1200
Wash/57       GGGGCATTTC TAGGTGGAGC AGATGTCAAA GAATGTATAG AAGCATTCAG
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83                               C
NY/84
Ont/87                                     G    a
```

```
            1201                                              1250
Wash/57     CAGTTATATA TGCCCTTCTG ATCCAGGATT TGTACTAAAC CATGAAATGG
Aus/59                                                           g
Wash/75                                         G                g
Tex/80         C                                 T              gg
NY/81          C                                 T               g
Ont/82(a)      C                                 T               g
Ont/82(b)      C                                 T               g
Ont/83         C                                 T
NY/84          C                                 T               g
Ont/87                                                           g

            1251                                              1300
                              →        ←
Wash/57     AGAGCTGTTT ATCAGGAAAC ATATCCCAAT GTCCAAGAAC CGTGGTCACA
Aus/59                                 T                 A      A
Wash/75            C                                    TAC
Tex/80      ag     C                   T                 ACA
NY/81              C                   T                 A A
Ont/82(a)          C                   T                 A A
Ont/82(b)          C                   T                 A A
Ont/83             C                   T                 A A
NY/84              C                   T                 A A
Ont/87             C                                     AC

            1301                                              1350
Wash/57     TCAGACATTG TTCCAAGATA TGCATTTGTC AATGGAGGAG TGGTTGCAAA
Aus/59
Wash/75                                 C
Tex/80
NY/81                                      T
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87                                  C

            1351                                              1400
Wash/57     TTGTATAACA ACCACATGTA CATGCAACGG TATTGGCAAT AGAATCAATC
Aus/59      t                                    C  T
Wash/75     C  C            C                 A  C
Tex/80      C               T                 A  C  T
NY/81       C               T                 A  C  T
Ont/82(a)   C               T                 A  C  T
Ont/82(b)   C               T                 A  C  T
Ont/83      C               T                 A  C  T
NY/84       C               T                 A  C  T
Ont/87      C               C                 A  C
```

```
            1401                                              1450
Wash/57     AACCACCTGA TCAAGGAGTA AAAATTATAA CACATAAAGA ATGTAATACA
Aus/59                                                       G
Wash/75                                                      G
Tex/80                        A                              G
NY/81                         A                              G
Ont/82(a)                     A                              G
Ont/82(b)                     A                              G
Ont/83                        A                              G
NY/84                         A                              G
Ont/87

            1451                                              1500
Wash/57     ATAGGTATCA ACGGAATGCT GTTCAATACA AATAAAGAAG GAACTCTTGC
Aus/59                               T
Wash/75                                        C
Tex/80
NY/81                                          C
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87

                                          →        ←
            1501                                              1550
Wash/57     ATTTTACACA CCAAATGATA TAACATTAAA CAATTCTGTT TCACTTGATC
Aus/59                                                  G
Wash/75        C          g                             G
Tex/80         C                       C                G
NY/81          C                       C                G
Ont/82(a)      C                       C                G
Ont/82(b)      C                       C                G
Ont/83         C                       C                G
NY/84          C                       C                G
Ont/87         C                       C     T          G

            1551                                              1600
Wash/57     CAATTGACAT ATCAATCGAG CTCAATAAGG CCAAATCAGA TCTAGAAGAG
Aus/59                                 C         G
Wash/75                                C           T             A
Tex/80                                 C                         A
NY/81                                  C                         A
Ont/82(a)                              C                         A
Ont/82(b)                              C                         A
Ont/83                                 C                         A
NY/84                                  C                         A
Ont/87                              T  C  A                      A
```

```
              1601                                          1650
Wash/57       TCAAAAGAAT GGATAAGAAG GTCAAATCAA AAACTAGATT CCATTGGAAA
Aus/59
Wash/75                                         t        G
Tex/80
NY/81
Ont/82 (a)
Ont/82 (b)
Ont/83
NY/84
Ont/87

                                                           (C)
              1651                                          1700
Wash/57       TTGGCATCAA TCTAGCACCA CAATCATAAT TGTTTTGATA ATGATAATTA
Aus/59            T                             A
Wash/75                                         A              C
Tex/80                           T              A              C
NY/81                            T              A              C
Ont/82 (a)                       T              A              C
Ont/82 (b)                       T              A              C
Ont/83                           T              A    T         C
NY/84                            T              A              C
Ont/87        C                  T              A    A         C

                             →        ←
              1701                                          1750
Wash/57       TATTGTTTAT AATTAATGTA ACGATAATTA TAATTGCAGT TAAGTATTAC
Aus/59
Wash/75                     C                  C       A
Tex/80                                         C       A
NY/81                                          C       A
Ont/82 (a)                                     C       A
Ont/82 (b)                                     C       A
Ont/83                                         C       A
NY/84                                          C       A        C
Ont/87                                         C       A

              1751                                          1800
Wash/57       AGAATTCAAA AGAGAAATCG AGTGGATCAA AATGATAAAC CATATGTATT
Aus/59
Wash/75                                             C  G          C
Tex/80                                              C  G          C
NY/81                                               C  G          C
Ont/82 (a)                                          C  G          C
Ont/82 (b)                                          C  G          C
Ont/83                                              C  G          C
NY/84                                               C  G          C
Ont/87                                              C  G          C
```

```
            1801        *                                       1850
Wash/57     AACAAACAAA TGACAGATCT ATAGATCATT AGATATTAAA ATTATAAAAA
Aus/59
Wash/75                 A   T
Tex/80                  A   T       C
NY/81                   A   T       C
Ont/82(a)               A   T       C
Ont/82(b)               A   T       C
Ont/83                  A   T       C
NY/84                   A   T       C
Ont/87                      T

            1851
Wash/57     A
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87
```

APPENDIX II: PREDICTED AMINO ACID SEQUENCE COMPARISON AMONG ISOLATES

Differences are listed with respect to the prototype strain (Wash/75). Lower case letters indicate ambiguous residues. Inverted arrows reveal potential glycosylation sites and the cleavage-activation site between F_1 and F_2 is shown by horizontal arrows. Major hydrophobic sequences are underlined: the signal peptide (s.p.); the F_1 amino terminus (F_1 a.t.), and; the membrane anchor (m.a.).

```
            1     (s.p.)                                        50
Wash/57     MPTSILLIIT TMIMASFCQI DITKLQHVGV LVNSPKGMKI SQNFETRYLI
Aus/59
Wash/75
Tex/80             L                                             V
NY/81                              V                             V
Ont/82(a)                                                        V
Ont/82(b)                                                        V
Ont/83                                                           V
NY/84                                                            V
Ont/87         L             S

            51                                                 100
Wash/57     ISLIPKIEDS NSCGDQQIKQ YKRLLDRLII PLYDGIRLQK DVIVSNQESN
Aus/59
Wash/75                                                     T
Tex/80                              K                       T
NY/81                               K                       T
Ont/82(a)                           K                       T
Ont/82(b)                           K                       T
Ont/83                              K                       T
NY/84                               K                       T
Ont/87                                                      T

            101     ↓↓         (F1 a.t.)                       150
Wash/57     ENTDPRTKRF FGGVIGTIAL GVATSAQITA AVALVEAKQA RSDIEKLKEA
Aus/59
Wash/75
Tex/80             E S
NY/81              E S
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87             R s                                  K

            151                                                200
Wash/57     IRDTNKAVQS VQSSIGNLIV AIKSVQDYVN KEIVPSIARL GCEAAGLQLG
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87                                       N
```

```
                                                    →  ←          250
            201
Wash/57     IALTQHYSEL TNIFGDNIGS LQEKGIKLQG IASLYRINIT EIFTTSTVDK
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87

                                                                  300
            251
Wash/57     YDIYDLLFTE SIKVRVIDVD LNDYSITTLQV RLPLLHTRLLN TQIYKVDSIS
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87

                                                                  350
            301
Wash/57     YNIQNREWYI PLPSHIMTKG AFIGGADIKE CIEAFSSYIC PSDPGFVLNH
Aus/59
Wash/75
Tex/80
NY/81
Ont/82(a)
Ont/82(b)
Ont/83
NY/84
Ont/87                                       S

            351     →  ←                                          400
Wash/57     EMESCLSGNI SQCPRTVVTS DIVPRYAFVN GGVVANCITT TCTCNGIGNR
Aus/59                       M K
Wash/75
Tex/80                       T
NY/81                        I
Ont/82(a)                    I
Ont/82(b)                    I
Ont/83                       I
NY/84                        I
Ont/87                       T
```

```
              401                                                → ←
Wash/57       INQPPDQGVK IITHKECNTI GINGMLFNIN KEGTLAFYTP NDITLNNSVS
Aus/59                          S                                  A
Wash/75                         S                          d       A
Tex/80                I         S                                  A
NY/81                 I         S                                  A
Ont/82(a)             I         S                                  A
Ont/82(b)             I         S                                  A
Ont/83                I         S                                  A
NY/84                 I         S                                  A
Ont/87                                                             A

              451                                                500
Wash/57       LDPIDISIEL NKAKSDLEES KEWIRRSNQK LDSIGNWHQS STTIIIVLIM
Aus/59                                                Y         I
Wash/75                       L                                 I
Tex/80                                                          I
NY/81                                                           I
Ont/82(a)                                                       I
Ont/82(b)                                                       I
Ont/83                                                          IF
NY/84                                                           I
Ont/87                                                          I

              (m.a.) → ←                         539
Wash/57       IIILFTINVT IIIIAVKYYR IQKRNRVDQN DKPYVLHNK
Aus/59
Wash/75                    T  I
Tex/80                     T  I
NY/81                      T  I
Ont/82(a)                  T  I
Ont/82(b)                  T  I
Ont/83                     T  I
NY/84                      T  I
Ont/87                     T  I
```