MOLECULAR CELL BIOLOGY OF RUBELLA VIRUS STRUCTURAL PROTEINS

by

Tom Cunningham Hobman
B.Sc., The University of Manitoba, 1983
M.Sc., The University of Manitoba, 1987

A thesis submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
in
the Faculty of Graduate Studies
Genetics Program

We accept this thesis as conforming
to the required standard

The University of British Columbia
November, 1989
© Tom C. Hobman
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Pathology

The University of British Columbia
Vancouver, Canada

Date Dec. 18/89
ABSTRACT

Rubella virus (RV) is a small, enveloped, positive-stranded RNA virus in the family Togaviridae, and bears striking similarities to the prototype alphaviruses Semliki Forest virus (SFV) and Sindbis virus (SV) in terms of genome organization and structural protein expression strategy. However unlike alphaviruses, RV infection of cultured cells is characterized by relatively long latency periods, slow replication kinetics, limited cytopathology, and the ability to establish a persistent infection in virtually every cell line capable of supporting its growth.

RV virions contain three structural proteins C, E2, and E1 which are derived by post-translational processing of a precursor polyprotein p110 (NH₂-C-E2-E1-COOH). Processing and intracellular transport of RV structural proteins has been studied by in vitro and in vivo expression of RV cDNAs. It was found that targeting of E1 and E2 into the endoplasmic reticulum was mediated by two independently functioning signal peptides. Coincident with translocation into the ER, both proteins underwent addition of N-linked glycans and proteolytic processing. C protein did not appear to play a role in the processing of p110. Expression of the RV structural proteins in COS cells revealed that E2 exited the ER, and was transported through the Golgi to the cell surface in an E1-independent manner, although coexpression of E1 seemed to increase the rate of transport. Conversely, E1 was retained in a Golgi-like region and was not found on the plasma membrane in the absence of E2.

Oligonucleotide-directed mutagenesis of E1 and E2 cDNAs showed that E1 and E2 both contain three N-linked glycans respectively. Lack of glycosylation did not appear to affect the intracellular localization of the RV glycoproteins in COS cells. A number of significant differences between RV and SFV/SV structural protein expression strategies were discovered, and their possible relationship to RV virion assembly are discussed.
### Table of Contents

**ABSTRACT**

**LIST of TABLES**

**LIST of FIGURES**

**LIST of ABBREVIATIONS**

**ACKNOWLEDGEMENTS**

**INTRODUCTION**

- Transport and modification of membrane glycoproteins
  - Translocation into the ER
  - Transport of membrane proteins from the ER
  - Covalent modifications of the glycoproteins
    - Role of carbohydrate
    - Acylation
- Alphavirus
  - Structure
  - Genome organization
  - Nonstructural proteins
  - RNA synthesis and replication
  - Structural protein expression and processing
    - Translation of 26S mRNA
    - Capsid protein
    - p62/E2 glycoprotein
    - E1 glycoprotein
- Entry of alphaviruses in mammalian cells
- Virus assembly and budding
- Rubella virus
  - Classification
  - Morphology and morphogenesis
  - Clinical aspects
  - Nucleic acid and genome organization
  - Expression of structural proteins
    - Capsid protein
    - E2 glycoprotein
    - E1 glycoprotein
  - Replication
  - Expression studies
- Project rationale and thesis objectives

**MATERIALS and METHODS**

- Materials and supplies
- Propagation and handling of E.coli strains
- Preparation of competent cells and transformation
- Isolation of plasmid and single-stranded DNA from E.coli
  - Mini-prep plasmid DNA isolation
  - Large scale plasmid DNA preps
  - Isolation of single-stranded DNA
- Oligonucleotide-directed mutagenesis
  - Purification and phosphorylation of oligonucleotides
  - Mutagenesis
  - Identification of mutants
Dot blot analysis 45
Restriction analysis 46
Dideoxysequencing of DNA 46
Restriction endonuclease digestions and DNA modification 46
Common solutions 48
Expression vectors 48
Rubella virus cDNA constructs 48
In vitro transcription; 52
In vitro translation; 52
Protease protection assay; 52
Immunoprecipitations, Endo-β-N-acetylglucosaminidase H and N-glycanase digestions 53
Transfection of COS cells 53
Metabolic labelling 54
Indirect immunofluorescence 54
Electrophoresis 55
Separation of DNA fragments 55
Separation of proteins 55
Western blotting 56

RESULTS and DISCUSSION 58
Section I. In vitro expression and targeting of RV structural proteins into microsomes 58
RV cDNA constructs 58
Transcription/translation 58
p24S 58
pE2E1 60
Expression of E2 61
pE2 61
pASPE2 65
E2 acquires N-linked glycans and is proteolytically processed by microsomes 65
Expression of E1 67
pE1 67
pASPE1 68
E1 acquires N-linked glycans and is proteolytically processed by microsomes 68
Discussion 68
Section II. In vivo expression and post-ER processing of RV structural proteins 75
Kinetics of processing 75
pCMV5-24S 75
pCMV5-E2E1 77
E2 glycoprotein 77
pCMV5-E2 79
pCMV5-E1 79
Steady state analysis of RV structural proteins in transfected cells 81
3H-palmitate labeling 81
Western blot analysis 83
Formation of disulfide bonds 83
Cell surface immunoprecipitations 88
LIST of TABLES

Table I. Domains of Alphavirus glycoprotein domains .................................................. 19
Table II. List of mutagenic oligonucleotides ................................................................. 56
LIST of FIGURES

Figure 1. Components of translocation machinery in the ER and cytoplasm 3
Figure 2. Processing of Asparagine-linked oligosaccharides 7
Figure 3. Structure of oligosaccharides on alphavirus glycoproteins 8
Figure 4. Expression strategy of Sindbis virus genome 12
Figure 5. Alternatives pathways for Sindbis virus RNAs 17
Figure 6. Processing of Semliki Forest virus structural proteins 23
Figure 7. Alphavirus glycoprotein structure 26
Figure 8. Model for alphavirus spike interactions with nucleocapsid 29
Figure 9. Expression of Rubella virus structural proteins 33
Figure 10. Expression vectors pSVL, pCMV5 and pSPT19 49
Figure 11. Rubella virus cDNA constructs 51
Figure 12. In vitro translation of p24 and pE2E1 59
Figure 13. Physical maps of E2 recombinants 62
Figure 14. In vitro translation of E2 64
Figure 15. Proteolytic processing of E2 by microsomes 66
Figure 16. Physical maps of E1 recombinants 69
Figure 17. In vitro translation of E1 72
Figure 18. In vivo expression of 24S cDNA in COS cells 76
Figure 19. In vivo expression of E2E1 cDNA in COS cells 78
Figure 20. In vivo expression of E2 cDNAs in COS cells 80
Figure 21. In vivo expression of E1 cDNAs in COS cells 82
Figure 22. Incorporation of palmitate into RV glycoproteins 84
Figure 23. Western blot of transfected COS cells lysates 85
Figure 24. Presence of disulfide bonds in RV structural proteins 87
Figure 25. Cell surface immunoprecipitation of RV structural proteins 89
Figure 26. Indirect immunofluorescence of pCMV5-24S in transfected cells 92
Figure 27. Indirect immunofluorescence of pCMV5-E2E1 in transfected cells 94
Figure 28. Indirect immunofluorescence of pCMV5-E2 in transfected cells 97
Figure 29. Indirect immunofluorescence of pCMV5-E2ns in transfected cells 98
Figure 30. Indirect immunofluorescence of pCMV5-E1 in transfected cells 99
Figure 31. Indirect immunofluorescence of RV M33 in COS cells 101
Figure 32. Model of RV assembly 106
Figure 33. SDS-PAGE analysis of E1 glycosylation mutants 108
Figure 34. Indirect immunofluorescence of E1 glycosylation mutants 109
Figure 35. SDS-PAGE analysis of E2 glycosylation mutants 113
Figure 36. Indirect immunofluorescence of E2 glycosylation mutants 113
LIST of ABBREVIATIONS

A260 absorbance at 260 nm
A600 absorbance at 600 nm
ATP adenosine triphosphate
bp base pair
BSA bovine serum albumin
CTP cytidine triphosphate
DNase deoxyribonuclease
ddNTPs dideoxynucleoside triphosphates
dNTPs deoxynucleoside triphosphates
endo H endo-\(N\)-acetylglucosaminidase H
EDTA ethylene diaminetetraacetic acid
ER endoplasmic reticulum
GTP guanosine triphosphate
kb kilobase
kDa kilodalton
M molar
mA milliamp
mg milligram
\(\mu g\) microgram
ml milliliter
\(\mu l\) microliter
mM millimolar
moi multiplicity of infection
ng nanogram
nm nanometer
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>RV</td>
<td>rubella virus</td>
</tr>
<tr>
<td>S</td>
<td>Svedberg unit</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SFV</td>
<td>Semiliki Forest virus</td>
</tr>
<tr>
<td>SRP</td>
<td>signal recognition particle</td>
</tr>
<tr>
<td>SV</td>
<td>Sindbis virus</td>
</tr>
<tr>
<td>Tris</td>
<td>trishydroxymethylaminomethane</td>
</tr>
<tr>
<td>TRITC</td>
<td>tetramethylrhodamine isothiocynate</td>
</tr>
<tr>
<td>TTP</td>
<td>thymidine triphosphate</td>
</tr>
<tr>
<td>UTP</td>
<td>uridine triphosphate</td>
</tr>
<tr>
<td>VSV</td>
<td>Vesicular Stomatitis virus</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

Firstly I would like to thank my supervisor and friend Dr. Shirley Gillam for giving me the opportunity to pursue this project, and for her encouragement throughout. I will always treasure this experience. Of my thesis committee members, a special thank you goes to Dr. Caroline Astell for her critical reading of the first draft and suggestions for whipping it into shape. I am grateful to my examiners Drs. Rob McMaster, Peter Candido, Aubrey Tingle, and Haydn Pritchard for their comments and suggestions. I would also like to thank my external examiner Dr. Jack Rose for helpful criticisms and prompt reading of the thesis.

For the gifts of anti-RV sera and monoclonal antibodies I thank Drs. Aubrey Tingle, John Safford, and Jerry Wolinsky. Thanks are due to the members of Dr. Tingles' lab for supplying radiolabeled RV. To Frank Tufaro for stimulating discussions and to Mike Weiss for help with the microscopic photography. I am indebted to Brad Brush and Haydn Pritchard for their help with the computer during the preparation of this thesis.

To Brad, Bob, Helena, Chris, Helen, Roger, Paul, Liz and Susan for their friendship and laughter.

Lastly to Marita for her never-ending enthusiasm and understanding during the work, and for being there when I needed her.

This thesis is dedicated to my family .....................Mom, Dad, Stacey, Muffin, and the cat.
INTRODUCTION

The introduction will be presented in three sections, i) Transport and modification of membrane glycoproteins; ii) Review of alphavirus molecular biology; iii) Review of Rubella virus. For the sake of brevity, discussion will be restricted mainly to type I membrane proteins which include alphavirus and rubella virus glycoproteins (Singer et al., 1987). Type I membrane proteins consist of large N-terminal ectoplasmic domains anchored in the membrane bilayer by a C-terminal transmembrane domain, while type II membrane proteins are anchored in the membrane by a hydrophobic N-terminal segment.

Transport and modification of membrane glycoproteins

Translocation into the ER

The most widely accepted model for targeting proteins to the ER membrane is the "signal hypothesis" which states that these proteins are first synthesized with transient hydrophobic amino-terminal extensions known as signal peptides (Blobel and Dobberstein, 1975a, b). As no primary sequence homology is found between signal peptides, these domains are thought to exert their function by adopting a common secondary structure such as an alpha helix (Watson, 1984; von Heijne, 1985). Signal peptides interact with a cytosolic ribonucleoprotein complex known as signal recognition particle (SRP) (Warren and Dobberstein, 1978; Walter and Blobel, 1980). SRP is composed of six different polypeptides, one of each 9, 14, 19, 54, 68, and 72 kDa complexed together on a 300 nt molecule of RNA called 7SL (Walter and Blobel, 1982; Siegel and Walter, 1985). Interaction of SRP with the newly emerged signal peptide results in a transient block in polypeptide elongation which is not released until the translation complex is targeted to the ER membrane (Walter and Blobel, 1981). The translation arrest function of SRP has been mapped to the 9/14 kDa proteins and Alu-like sequence in 7SL (Siegel and Walter, 1985, 1986), and the 54 kDa protein is known to interact with the signal peptide as it emerges from the ribosome (Krieg et al., 1986). Translation resumes after SRP binds to SRP receptor or docking protein which is an ER membrane protein (Gilmore et al., 1982a, b; Meyer et al., 1982). Also in the ER membrane is a protease-sensitive ribosome receptor which facilitates binding of polysomes which are
synthesizing secretory or membrane proteins (Adelman et al., 1973; Hortsch et al., 1986). See Fig.1. Although its existence had been suspected for some time, only recently has a signal sequence receptor been identified as an integral ER membrane protein (Prehn et al., 1980, 1981; Wiedmann et al., 1987). Following its release from SRP, the signal peptide domain of protein undergoing translocation is transferred to the signal sequence receptor (Fig. 1).

For the most part, translocation of eukaryotic secretory and membrane proteins across the ER is a cotranslational event which can only occur during a limited period of nascent chain elongation. SRP is thought to increase the length of this time window during which time a protein is in a translocation-competent state (Walter, 1987). Vectorial discharge of proteins into the ER requires GTP for a function unrelated to protein synthesis (Connolly and Gilmore, 1986; Wilson et al., 1988). Sequencing of the cDNA for the 54 kDa protein of SRP revealed a putative GTP-binding domain which demonstrates significant homology with the large subunit of docking protein (Berstein et al., 1989; Romisch et al., 1989). Exactly how a protein is transferred across the lipid bilayer is not known, but it is suspected that the process occurs through an aqueous channel formed by integral ER membrane proteins according to the signal hypothesis (Blobel and Dobberstein, 1975a, b). More recently, biochemical evidence for this theory has been provided by Gilmore and Blobel (1985).

Signal sequences are thought to form hairpin loops (Fig. 1) upon insertion into the ER resulting in this domain not being transferred into the ER lumen (Engelman and Steitz, 1981). Shaw et al. (1988) provided direct evidence for the loop model of signal sequence insertion, and proposed that the signal sequence of a type I membrane protein is the structural and functional equivalent of the signal peptide/membrane anchor domain of type II membrane proteins (Singer et al., 1987). The only proposed difference is the presence of a signal peptidase cleavage site on the carboxyl-end of the type I protein signal peptide (Fig. 1). Signal peptidase is an endoprotease on the luminal side of the ER membrane, and removes the N-terminal signal peptides from type I transmembrane proteins, and most secretory proteins with such exceptions as ovalbumin (Blobel and Dobberstein, 1975a, b; Palmiter et al., 1978).
Figure 1. Schematic representation illustrating SRP elongation arrest and components of the translocation machinery in the ER. A) SRP cycles between free cytosolic form (a) to ribosome-bound form (b) to SRP receptor (e). B) SRP displays increased affinity for translating ribosome after emergence of signal peptide (zigzag lines) resulting in translation arrest. C) Targeting to ER membrane results in interaction of the large ribosomal subunit with its receptor, SRP with its receptor and release of the translation block. D) The signal sequence is inserted as a hairpin loop and interacts with signal sequence receptor. The passenger protein domain is translocated through an aqueous pore formed by ER membrane proteins, and will be released from the signal peptide domain by signal peptidase. (from Perara and Lingappa, 1988)
Translocation of type I membrane proteins into the ER is halted by the stop-transfer or transmembrane domain near the C-terminus of the protein. Stop-transfer sequences, like signal peptides, demonstrate little primary sequence homology and consist of 20 to 30 hydrophobic and neutral amino acids flanked by at least one basic residue on the cytoplasmic side (Sabatini et al., 1982).

Transport of membrane proteins from the ER

Targeting to the ER is the first stage of transport through the exocytic pathway for most viral and eukaryotic membrane glycoproteins. Following transfer across the ER membrane, proteins can be modified by the addition of asparagine-linked oligosaccharides (Rothman and Lodish, 1977), and fatty acid groups (Schmidt and Schlesinger, 1980). These topics are discussed in the next section.

The study of viral membrane glycoproteins has been instrumental in analysis of transport from the ER to the Golgi in eukaryotic cells and results seem to indicate that tertiary and/or quaternary structures of the proteins are most often critical for transport out of the ER. Protein disulfide isomerase, an enzyme found in the lumen of the ER, catalyzes the formation of intermolecular and intramolecular disulfide bonds (Freedman, 1984). Following signal peptide cleavage, oxidation of cysteine residues occurs and is usually crucial for the attainment of a native protein configuration. Studies with influenza virus hemagglutinin (HA) and vesicular stomatitis virus (VSV) G protein suggest that protein folding occurs very rapidly in the ER, within one to three minutes after synthesis as measured by expression of antibody recognized epitopes (Copeland et al., 1988; Doms et al., 1988). For many viral proteins, addition of N-linked carbohydrate is essential for the attainment of a tertiary structure conducive for transport from the ER (For review see Olden et al., 1982; Rose and Doms, 1988). More recently Pitta et al. (1989) have shown that a single change in amino acid sequence in the VSV G protein can compensate for a previously determined carbohydrate-sensitive folding step (Machamer et al., 1985; Machamer and Rose, 1988b).
Very often before proteins are transported from the ER to the Golgi, they must assume suitable quaternary structures which can consist of one or more species of protein (For review Rose and Doms, 1988). For example, both VSV G and influenza HA are known to form noncovalently associated trimers before exiting the ER. This process has been shown to be regulated by ATP for HA at least (Wilson et al., 1981; Doms et al., 1987). Alphavirus glycoproteins form six member heterotrimeric spikes (E2-E1)_3 prior to exiting the ER (Fuller, 1987). Proteins which fail to exit the ER may be retained in the ER in large insoluble aggregates associated with heavy chain binding protein (BiP) which is thought to function in the ER by recognizing and retaining misfolded proteins (Bole et al., 1986; Kozutsumi et al., 1988; Hurtley et al., 1989).

Proteins destined for the plasma membrane only cross a membrane bilayer (the ER) once when traversing the exocytic pathway with all subsequent inter and intra-organellular transport being mediated by vesicular carriers (Jamieson and Palade, 1967; Palade, 1975). Transport from the ER to the Golgi occurs by budding of vesicles from specialized areas of ER called transitional elements, and fusion with the cis Golgi membranes. The Golgi is organized in a polarized fashion with protein entry occurring in the cis compartment, proceeding through the medial compartment, and sorting and exit occurring through the trans Golgi network (for review see Farquhar, 1985; Pfeffer and Rothman, 1987; Farquhar, 1989). Transport between the Golgi stacks is also effected by transport vesicles which are derived by budding and fusion from and to the dilated rims of the cisternae. These vesicular carriers are coated with a non-clathrin protein component (Orci et al., 1986), and their movement requires ATP, GTP-binding proteins, and an N-ethylmaleimide-sensitive factor for the fusion process (Jamieson and Palade, 1968; Balch et al., 1986; Melancon et al., 1987; Malhortra et al., 1988). As proteins are transported vectorially across the cis-trans axis they encounter a gradually more acidic environment (Mellman et al., 1986). Currently it is believed that ER to Golgi movement of newly synthesized proteins occurs by "bulk flow", and that differential rates of protein transport is the result of selective retention and differences in protein folding rates (Pfeffer and Rothman, 1987; Weiland
Covalent modifications of the glycoproteins

Discussion of membrane protein modification will be limited to alphaviral glycoproteins because of their structural similarity to RV glycoproteins.

Role of carbohydrate

As with other glycoproteins synthesized in higher eukaryotic cells, N-linked glycans on alphavirus proteins are transferred en bloc as a precursor $(\text{Glc}_3\text{Man}_9\text{GlcNAc}_2)$ to the consensus acceptor sequence Asn-$\text{X}$-Ser/Thr, where $\text{X}$ is any amino acid except proline (for review see Kornfeld and Kornfeld, 1985; Parent, 1988; Runge, 1988). The Glc$_3$Man$_9$GlcNAc$_2$ precursor is transferred from a lipid carrier, dolicholphosphate to the nascent protein in the lumen of the ER during translocation (Fig. 2). Removal of the terminal glucose residues and one or more mannoses by ER glucosidases and mannosidases occurs prior to exiting the ER. Glycoproteins are transported from the ER by means of vesicular carriers which bud from the transitional ER and fuse with the cis Golgi (Jamieson and Palade, 1967; Palade, 1975; Lodish, 1988). Passage of glycoproteins through the medial Golgi cisternae is easily assayed by monitoring sensitivity to endo-$\beta$-N-acetylglucosaminidase H (endo H) (Fig. 2). Oligosaccharides destined for conversion to complex types become endo H resistant shortly after processing by $\alpha$-mannosidase II in the medial Golgi compartment (Dunphy and Rothman, 1985). Addition of the terminal N-acetylglucosamine also occurs in medial Golgi, while galactose and sialic acid are added in the trans Golgi. Examples of high mannose and complex sugars found on some alphaviral glycoproteins are shown in Fig. 3.

Generally, complex glycans are found closest to the amino terminus of a polypeptide, most likely due to greater access of this end of the polypeptide to trimming and processing enzymes by virtue of location (Pollack and Atkinson, 1983). Hsieh et al. (1983) studied endo H resistance of glycoproteins from SV grown in a ricin-resistant Chinese hamster ovary (CHO) cell line 15B.
Fig. 3. Processing of Asn-linked oligosaccharides to a representative biantennary complex structure. The scheme depicts the processing from the transfer of Glc3Man9GlcNAc2 from its dolichol pyrophosphoryl derivative to the nascent polypeptide chain still bound to the ribosome, followed by processing reactions in the ER and Golgi complex. Oligosaccharide processing enzymes are listed above the line; the reaction they catalyze is diagramed below the line (except for the alternate processing reaction, 3a). The subcellular localization of processing events are depicted by the brackets. Structures susceptible or resistant to digestion by endo H, an enzyme frequently used as a diagnostic test for processing to complex structures, are indicated. Symbols: ▼, glucose; O, mannose; ■, N-acetylglucosamine; ●, galactose; ▲, fucose; ◆, sialic acid. (Moremen and Touster, 1988)
Figure 3. Structure of the two major oligosaccharides on alphavirus glycoproteins. (S.A.) sialic acid; (Gal) galactose; (Man) mannose; (GlcNAc) N-acetylglucosamine; (Fuc) fucose; (Asn) asparagine. (Schlesinger and Schlesinger, 1986)
which is deficient in N-acetylglucosaminyltransferase I, and is therefore unable to process oligomannose precursors to complex forms. Digestion of E1 and E2 under non-denaturing conditions with endo H resulted in preferential cleavage at sites normally containing complex sugars. Resistance to α-mannosidase I, N-acetylglucosaminyltransferase I, and α-mannosidase II is likely due to steric hindrance (Hubbard, 1988). Host cell type is also known to be important in determining sugar structure of alphavirus glycoproteins (Keegstra et al., 1975; Hsieh et al., 1983a, Davidson and Hunt, 1983). Other factors such as stage of cell growth and nutritional environment may also influence sugar processing (Hakimi and Atkinson 1982; Davidson and Hunt, 1985).

Two lines of study have been used to investigate the importance of carbohydrate in processing and function of alphavirus glycoproteins. Use of mutant host cell lines that process carbohydrates abnormally has given rise to mixed results regarding SFV and SV growth and infectivity. SV can grow in mosquito cells that do not process beyond Man3GlcNAc2 as well as in a CHO mutant cell line which stops at Man5GlcNAc2, implying that terminal GlcNAc, galactose, or sialic acid are not essential for virus assembly and infectivity (Schlesinger et al., 1976; Hsieh and Robbins, 1984). However, enzymatic removal of carbohydrate from intact SFV virions results in non-infectious aggregates (Kennedy, 1974). Two additional examples include mutant mouse L cell lines, one that over sialylates glycoproteins, another that does not process beyond Man8GlcNAc2, yet neither of which cleave p62 (Gottlieb et al., 1979; Tabas and Kornfeld, 1978).

Use of agents that interfere with sugar processing is the second approach for studying the role of carbohydrate on glycoproteins. It is a widely held belief that some proteins require covalently bound glycans for correct tertiary structure. Tunicamycin completely inhibits N-linked glycosylation by blocking formation of GlcNAc-P-P-dolichol (Struck and Lennarz, 1980). Addition of GlcNAc to dolichol diphosphate is the first step in the biosynthesis of the Glc3Man9GlcNAc2 precursor. Tunicamycin inhibits SV and SFV virion formation and p62 cleavage even though the proteins are transported to the cell surface (Schwarz et al., 1976;
Leavitt et al., 1977; Scheefers et al., 1980; Mann et al., 1983). Removal of glucose residues from glycoproteins seems to be essential for virus assembly since 1-deoxynojirimycin, castanospermine, and bromoconduritol which inhibit ER glucosidases, can all block replication of SV (Datema et al., 1984; Schlesinger et al., 1985). Interestingly, the effects of deoxynojirimycin and castanospermine are more severe at 37°C than at 30°C suggesting that the thermostability of the proteins is decreased (Schlesinger et al., 1985).

More recently McDowell et al. (1987) reported a switch from budding at the plasma membrane to internal membranes in SV-infected cells treated with 1-deoxymannojirimycin (dMM). This mannose analog inhibits Golgi mannosidase I and results in glycoproteins carrying Man$_9$GlcNAc$_2$ structures (Fuhrmann et al., 1984; Bishcoff and Kornfeld, 1984). Cleavage of p62 is not affected by dMM, supporting the notion that glucosylated SV proteins are not a substrate for the enzyme that processes p62 (McDowell et al., 1987; Schlesinger et al., 1985). Finally, N-linked sugars are thought to function by aiding glycoproteins in attaining and maintaining correct tertiary structure, and their importance varies with the particular glycoprotein in question.

**Acylation**

Mature E2 and E1 in alphaviruses contain covalently bound fatty acid moieties in the form of palmitic acid (Schmidt et al., 1979; Schmidt, 1982). Attachment is thought to involve fatty acyl-CoA transferase mediated transfer of the long chain fatty acids to cysteines in the carboxy-terminal region of E2 and E1 (Berger and Schmidt, 1984), and likely occurs in the ER (Berger and Schmidt, 1985; Rose and Doms, 1989), or in post-ER/pre-Golgi structures (Bonatti et al., 1989). In any event, palmitoylation seems to precede oligosaccharide trimming by mannosidase I which is believed to be a cis Golgi enzyme (Bonatti et al., 1989). Lack of acylation does not affect intracellular transport of the glycoproteins, but seems to interfere with the final stages of virus assembly possibly by formation of aberrant intermolecular disulfide bonds (Schlesinger and Malfer, 1982). Covalently-linked oligomeric aggregates of SV glycoproteins form after treatment of virions with hydroxylamine which removes fatty acid.
moieties (Magee et al., 1984). One other possible function of acylation could be to modulate the lipid components in the immediate vicinity of the embedded viral glycoproteins (Schlesinger and Schlesinger, 1986).

**Alphavirus**

**Structure**

Alphaviruses comprise a genus of small enveloped RNA viruses in the family Togaviridae. The prototype viruses of this genus are Semliki Forest virus (SFV) and Sindbis virus (SV). Virion particles are spherical and approximately 65-70 nm in diameter with 6-10 nm glycoprotein spikes protruding from the surface (Garoff et al., 1982). Their genomes consist of a molecule of positive-polarity single-stranded RNA 11,200 nt (SFV) to 11,700 nt (SV) in length (Garoff et al., 1982). The genomic RNA is polyadenylated and contains a methylated 5' cap (Dubin et al., 1977; Eaton et al., 1972). Three major structural proteins capsid (C), and envelope glycoproteins E2 and E1 are found in equimolar amounts in virions. In SFV an additional protein E3 forms the glycoprotein spike with E1 and E2, while in SV, E3 is released into the extracellular space (Garoff et al., 1974; Welch and Sefton, 1979).

**Genome organization**

Alphaviruses are arranged in a manner such that the 5' two thirds of the genome encode the nonstructural proteins (Strauss et al., 1984) and the 3' one third contains genes specifying the structural proteins (Garoff et al., 1980a, b; Rice and Strauss, 1981)(Fig. 4). The genomic RNA serves directly as a mRNA for translation of the nonstructural proteins which are synthesized as polyprotein precursors (Strauss et al., 1984). The structural proteins are translated as a 130,000 dalton precursor from a 26S subgenomic RNA corresponding to the 3' one third of the viral genome (Simmons and Strauss, 1974; Garoff et al., 1974). The subgenomic RNA is present in a three-fold molar excess over the genomic RNA allowing for amplification of the structural proteins (Garoff et al., 1983). In addition, the genomic RNA is quickly incorporated into nucleocapsids rendering it unable to function as a mRNA (Garoff et al., 1983).
Figure 4. Expression strategy of the Sindbis virus genome. Untranslated regions of the genomic RNA are shown as single lines and the translated regions as a narrow open triangle. The subgenomic RNA which encodes the structural proteins is shown below in an expanded format. The final protein products are shown as bold lines. (△) Initiation codons; (○) termination codons; (△) UGA codon readthrough to produce nsP4. (Strauss et al, 1984).
Between the nonstructural genes and structural genes is a stretch of nucleotides termed the junction region (Ou et al., 1982; Riedel et al., 1982). In this region there are three closely spaced termination codons preceding the open reading frame of the structural proteins as well as proposed signals for transcription initiation for the 26S RNA. Hence the alphavirus genomes are organized such that the nonstructural genes must be 5' terminal, allowing their transcription and translation before synthesis of a negative polarity genomic RNA, then 26S mRNA and subsequent expression of the structural proteins.

Nonstructural proteins

Early studies with alphavirus-infected cells suggested the existence of four nonstructural proteins (Lachmi and Kaariainen, 1976; Brzeski and Kennedy, 1977). These proteins acting in concert or independently must function to synthesize minus-strand genomic RNA, and from this template transcribe 26S subgenomic RNA and plus-strand genomic RNA in the appropriate quantities. Temperature-sensitive mutants of alphaviruses are classified as RNA$^{+}$ or RNA$^{-}$ mutants according to their ability to synthesize RNA at the nonpermissive temperature. RNA$^{-}$ mutants can be further subdivided into four complementation groups, thus lending further support for the presence of at least four nonstructural genes (Strauss and Strauss, 1986).

Nucleotide sequence analysis of SV and SFV RNAs confirmed that four genes encode the nonstructural proteins in the order NH$_2$-nsP1-nsP2-nsP3-nsP4-COOH (Strauss et al., 1984; Takkinen, 1986). Translation begins at an AUG codon 60-80 nucleotides 5' to the cap site (Ou et al., 1983). The nonstructural antigens are synthesized as two large polyprotein precursors (Fig. 4). In SV-infected cells, a 230K polyprotein P123, specifying NH$_2$-nsP1-nsP2-nsP3-COOH and a 270K P1234, containing NH$_2$-nsP1-nsP2-nsP3-nsP4-COOH are translated and then cleaved to give rise to the nonstructural proteins (Strauss and Strauss, 1986). The 270K protein is made in relatively minor amounts by readthrough translation of an opal codon (UGA) between nsP3 and nsP4 (Strauss et al., 1984). Other characterized alphaviruses contain in-frame opal codons in this region as well, resulting in extremely low concentrations of nsP4 (72K) relative to nsP1 (60K),
nsP2 (89K), and nsP3 (76K) (Strauss et al., 1983). In contrast, SFV does not contain such an opal codon, and infected cells subsequently produce larger amounts of nsP4 (Takkinen, 1986; Keranen and Ruohonen, 1983). Presently it is uncertain as to the mechanism of translational readthrough of the opal codon in SV, but it is suspected that naturally occurring opal suppressor transfer RNAs may be involved (Strauss and Strauss, 1986).

Processing of SV nonstructural polyproteins in infected BHK-21 cells indicates that in the P123 precursor, cleavage between nsP2 and nsP3 occurs before cleavage between nsP1 and nsP2 (Hardy and Strauss, 1988). Proteolytic processing of nsP3 and nsP4 seems to occur last. A viral encoded protease contained in nsP1, 2 or 3 is thought to be responsible for processing the nonstructural polyproteins (Strauss et al., 1984). Kinetics of SV nonstructural protein processing is relatively slow compared to the SFV counterparts (Keranen and Ruohonen, 1983).

The task of assigning discrete functional roles for the alphavirus nonstructural antigens is not complete despite the wealth of information available regarding their sequence and processing. Among alphaviruses, nsP4 is highly conserved and shares homology with RNA-dependent RNA replicase of poliovirus and nonstructural proteins of other positive-strand animal and plant viruses (Kamer and Argos, 1984; Strauss and Strauss, 1986) leading to the supposition that it possesses the elongation activity of the RNA polymerase complex. This theory is supported by the existence of RNA− temperature-sensitive mutants which map to nsP4 (Hahn et al., 1989a). This group also reports that nsP2 and nsP4 may interact to form a functional complex. In SFV-infected cells, nsP3 is a phosphoprotein found associated with cytoplasmic vesicles which are thought to be the sites of virus-specific RNA synthesis (Peranen et al., 1988). It has also been suggested that nsP3 may be the viral protease (Strauss and Strauss, 1986). Brzeski and Kennedy (1978) suggest that nsP2 is responsible for initiation of 26S RNA transcription, and this has now been confirmed by mapping RNA− ts mutants (Hahn et al., 1989b). The latter group reports that nsP2 also serves as the viral protease, and thus the function of nsP3 remains unknown. No laboratories have been successful in purifying an active replicase complex from alphavirus infected cells, but in the purest preparations so far, nsP1 seems to be
the major component, while nsP2 and nsP4 are minor ones (Gomabatos et al., 1980; Ranki and Kaariainen, 1979). Mutations in any four of the nonstructural genes can lead to RNA mutants indicating that all four proteins are involved in RNA synthesis (Hahn et al., 1989b).

RNA synthesis and replication

The genomic RNA is first translated to give rise to the nonstructural proteins which must function as a replicase/transcription complex. A conserved 19 nucleotide sequence immediately 5' to the poly A tract in genomic RNA is believed to be one control element responsible for initiation of genomic negative-strand RNA (Ou et al., 1982a, b, 1983). At least one other controlling motif is needed in light of the fact that no negative-strand subgenomic RNA is detected in infected cells. Since the 26S subgenomic RNA is precisely coterminal with the genomic RNA (Ou et al., 1981) it seems likely that the second controlling element for initiation of genome-length minus-strand RNA could be in the 5' end of the plus-strand RNA, and may be brought into play by cyclization of the molecule (Strauss and Strauss, 1986). Indeed there is a 51 nucleotide conserved element in alphaviruses located 150 nucleotides from the 5' end of the genome (Strauss and Strauss, 1986). By virtue of its length, the 51 nucleotide motif had been suggested to have other functions such as serving as a nucleation site for encapsidation. However, later evidence obtained by sequencing RNA from defective interfering particles, revealed the 51 nucleotide sequence could be replaced by tRNA^{asp} (Levis et al., 1986), and was therefore not essential for replication or encapsidation. The authors concluded that the secondary structure of the 5' end of the genome rather than the primary sequence, was important for these functions.

The highly conserved 21 nucleotide junction region between the nonstructural genes and the structural genes is postulated to be the controlling element which modulates 26S RNA initiation (Ou et al., 1982a). For transcribing plus-strand RNA, a conserved sequence at the 3' end of the minus-strand RNA which forms a stem and loop structure is thought to be involved (Strauss and Strauss, 1983).
Capping activity can be separated from the transcription activity of the replication complex, and was found to occur after initiation of transcription (Cross and Gombatos, 1981). The poly A tract of the genomic and subgenomic RNA is added by copying a poly U tract located in the 5' terminus of the minus strand (Ou et al., 1981).

There are three possible fates for the viral genome in the host cell cytoplasm; 1) binding to ribosomes and engaging in synthesis of more nonstructural proteins; 2) binding to a replication complex and transcription of additional minus-strand RNA; 3) formation of nucleocapsids by complexing with capsid protein. Capsid protein is thought to play a role in regulating the pathways of the genomic RNA by sequestering it from fates 1 and 2 (see Fig. 5).

Structural protein expression and processing

Translation of 26S mRNA

Structural proteins of alphaviruses are derived by post-translational processing of a polyprotein precursor specified by the subgenomic 26S mRNA (Wirth et al., 1971). Use of a separate RNA species for synthesis of structural antigens allows for amplification of these products. In infected cells, greater than 90% of the virus-specific RNA complexed in polysomes is 26S, even though the subgenomic RNA is produced in only a three fold molar excess to the genomic RNA (Strauss and Strauss, 1977; Schlesinger and Kaariainen, 1980).

Translation of the 26S mRNA begins at an AUG codon specifying the amino-terminal methionine in capsid protein (Cancedda et al., 1974) and proceeds NH$_2$-C-p62-E1-COOH (Lachmi et al., 1975; Clegg, 1975). Nucleotide sequence analysis of the 26S mRNA revealed small peptides E3 and a 6K peptide which are located amino-terminal to E2 in the p62 precursor and between p62 and E1 respectively (Garoff et al., 1980a, b) (Fig. 4). Capsid is released from the polyprotein p130 by an autoproteolytic cleavage between the carboxy-terminal tryptophan of capsid and the amino-terminal serine of E3 (Simmons and Strauss, 1974; Aliperti and Schlesinger, 1978; Garoff et al., 1980b). The N-terminal signal sequence for p62 is contained in E3 and is exposed upon release of capsid from the polyprotein, allowing it to direct translocation of p62.
Figure 5. Alternative pathways for RNA transcription and replication in Sindbis virus. Genome plus-strand 49S binds to polysomes, and nonstructural proteins are translated (1), which in turn produce full-length minus-strand 49S RNA (2). The negative-strand serves as a template for production of plus-strand genomic RNA (3) and 26S subgenomic RNA (5). The nascent genomic plus-strand RNA can bind capsid to form nucleocapsids (4), bind polysomes to produce more nonstructural proteins (4 ''), or serve as template for minus-strand synthesis (4 '). The subgenomic RNA serves as a mRNA for synthesis of structural proteins (6). Capsid combines with genomic RNA to form nucleocapsids (7). (Schlesinger, 1985)
into the endoplasmic reticulum (ER) (Wirth et al., 1978; Garoff et al., 1978; Bonatti et al., 1979).

The 6K peptide located between E2 and E1 functions as a signal peptide reinitiating translocation of E1 into the ER (Hashimoto et al., 1981; Melancon and Garoff, 1986). Targeting and transport of the glycoproteins will be discussed in more detail below.

### Capsid protein

The capsid or C protein of the prototype alphaviruses is 264 to 267 amino acids long in SV and SFV respectively (Rice and Strauss, 1981; Garoff et al., 1980a). In the amino-terminal one third of the protein are clusters of arginines and prolines which are thought to participate in binding genomic RNA during nucleocapsid assembly (Garoff et al., 1980a). Unlike the amino-terminal part of capsid, the carboxyl end of the molecule is highly conserved among alphaviruses and likely possesses domains necessary for protein-protein interactions and spike complex attachment during viral assembly (Garoff et al., 1982). In addition, the autocatalytic motif of capsid that releases it from p130 is a serine protease (Hahn et al., 1985) residing in the carboxy-terminal region of the protein. In vitro mutagenesis of SFV 26S cDNA has localized the active site of the capsid autoprotease to serine-219 (Melancon and Garoff, 1987).

Capsid binds very rapidly (5-7 minutes) after synthesis to genomic RNA and assembles into an icosahedral nucleocapsid containing 180 copies of C (Soderlund, 1973; Fuller, 1987). This is the first step in virion assembly and occurs in the cytoplasm of infected cells.

### p62/E2 glycoprotein

The p62 or PE2 glycoprotein is 488 amino acids in SFV and 487 in SV (Garoff et al., 1980b; Rice and Strauss, 1981). As the latter name PE2 suggests, p62 is a precursor to the E2 glycoprotein. The amino-terminal region contains the E3 protein (64-66 amino acids) which includes the signal peptide for the E3-E2 precursor. Biochemical studies revealed that the amino-terminus of E3 is acetylated (Bell et al., 1982). Translocation of p62 across the ER membrane is dependent upon signal recognition particle (SRP) (Garoff et al., 1978; Bonatti et al.,...
1984) as is the case with many secretory and membrane proteins. However, this precursor is unusual in the sense that cleavage by signal peptidase does not occur (Bonatti et al., 1979). As well, the signal peptide region becomes glycosylated at asparagine 13 (SFV) or 14 (SV) (Bell et al., 1982). One intriguing difference between SV and SFV is that E3 is retained in the glycoprotein spike of SFV, but is secreted into the medium of SV-infected cells (Garoff et al., 1974; Welch and Sefton, 1979).

E2 glycoprotein is 422 amino acids long in SFV and 423 in SV (Garoff et al., 1980b; Rice and Strauss, 1981) and contains two potential N-linked glycosylation sites which are both used (Burke and Keegstra, 1979). Biochemical studies by this group suggest that SV E2 contains one complex (endo H resistant) glycan at Asn 196 and one high mannose type (endo H sensitive) at Asn 318.

E2 contains two conserved cysteine residues following a group of 24-26 very hydrophobic amino acids in the carboxy-terminus that functions as the transmembrane anchor of the protein (Garoff et al., 1980b; Rice and Strauss, 1981; Garoff et al., 1982). Immediately following the membrane binding domain are 31-33 amino acids that comprise the cytoplasmic domain of E2 (Table 1). The predicted membrane topology of E2 is therefore that of a type I membrane protein (Singer et al., 1987) with three domains; 1) N-terminal ectoplasmic domain; 2) transmembrane domain; 3) C-terminal cytoplasmic domain.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sindbis virus</th>
<th>Semliki Forest virus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrophilic, globular (outside)</td>
<td>395 364 45</td>
<td>396 367 47</td>
</tr>
<tr>
<td>Membrane-spanning</td>
<td>26 26 --</td>
<td>24 24 --</td>
</tr>
<tr>
<td>Cytoplasmic (inside)</td>
<td>2 33 --</td>
<td>2 31 --</td>
</tr>
<tr>
<td>Hydrophobic (outside)</td>
<td>16 -- 19</td>
<td>16 -- 19</td>
</tr>
<tr>
<td></td>
<td>(80-96)</td>
<td>(80-96)</td>
</tr>
</tbody>
</table>

* Numbers of amino acids in a domain are listed. The numbers in parentheses indicate the positions of the amino acids in the protein's sequence.

The SFV p62 glycoprotein gene has been subcloned into an SV40-based expression vector.
and microinjected and transfected into BHK and COS cells respectively by Garoff and coworkers (Kondor-Koch et al., 1982; Cutler and Garoff, 1986). Deletion mutagenesis experiments revealed that the cytoplasmic tail of E2 is not required for transport to the plasma membrane (Kondor-Koch et al., 1983; Garoff et al., 1983). In contrast, deletion of the transmembrane and cytoplasmic domains results in intracellular accumulation of p62 as a result of altered tertiary structure (Riedel, 1985). Fusion of p62 ectoplasmic domain with the membrane binding domain and cytoplasmic region of vesicular stomatitis virus G protein or fowl plaque virus hemagglutinin reconstitutes transport to the cell surface (Riedel, 1985). Although E2 traverses the exocytic pathway without E1 (Kondor-Koch et al., 1982), the p62 cleavage between E3 and E2 requires E1 (Cutler and Garoff, 1986; Cutler et al., 1986). Proteolytic cleavage of SFV p62 after the Arg-Arg pair in the carboxy-terminus of E3 takes place shortly after or during exit from the trans-Golgi reticulum, but prior to reaching the plasma membrane (De Curtis and Simmons, 1988). This cleavage which is essential for infectivity (White et al., 1983) requires Ca\(^{++}\) and ATP and is probably catalyzed by a host cell protease. In contrast, work from the laboratory of Brown seems to indicate that PE2 can be incorporated into Sindbis virus virions, and therefore suggests that PE2 cleavage is not required for virion formation in SV at least (Presley and Brown). Moreover in SV-infected cells, endo H-sensitive forms of E2 are found implying that the pE2 cleavage event may be independent of transport as suggested above with SFV, and also raises the possibility that this cleavage may be virus specified (Knipfer and Brown, 1989).

Monoclonal antibody studies define at least three neutralizing domains on SV E2, one of which appears to be strain specific (Roehrig et al., 1982; Stanley et al., 1985; Stec et al., 1986; Olmstead et al., 1986; Davis et al., 1987). Using a mouse model to study SV neurovirulence these investigators have demonstrated specific codon changes in E2 which affect rate of penetration (Davis et al., 1986), p62 cleavage and pathogenesis (Russell et al., 1989). In addition to containing strain-specific epitopes, E2 is known to play a major role in the budding process. The cytoplasmic domain of E2 is 31-33 amino acids in length compared with two amino acids in the
analogous region in E1 (Garoff et al., 1980b; Rice and Strauss, 1981). Consequently it was
assumed that the cytoplasmic portion of E2 interacted with nucleocapsids in the budding process.
Furthermore, subtle mutations in the ectoplasmic domain of SV E2 can lead to defects in viral
assembly. Hahn et al. (1989c) reported that the lesion in SV ts103 is in E2 (Ala344-Val) and
results in weakened interactions between the spike complex and nucleocapsid. Revertants with a
Lys-Met change in E1 at position 227 were isolated which partially suppressed the ts103
phenotype. This topic will be discussed in more detail below.

E1 glycoprotein

E1 is the largest of the structural proteins, 438 amino acids in SFV and 439 in SV and
is also part of the viral spike complex (Garoff et al., 1980b; Rice and Strauss, 1981). Sequence
analysis suggests that E1 is also a type I membrane protein (Singer et al., 1987) with a large
N-terminal ectoplasmic domain anchored by a 24-26 amino acid membrane-spanning region near
the C-terminus. The cytoplasmic tail of E1 is only two residues long in SV and SFV. See table
1.

Translocation of the alphavirus structural polyprotein into the ER is halted by the
c-terminus transmembrane domain of p62/E2. This necessitates E1 having a separate signal
peptide to reinitiate translocation. A temperature-sensitive mutant of SFV ts-3 accumulates a
cytoplasmic 86,000 dalton C-p62 fusion protein when grown at the restrictive temperature
(Hashimoto et al., 1981). E1 however, was found to be of normal size, and was sequestered and
glycosylated in the ER implying that two separate signal sequences are required for translocation
of p62 and E1 into the ER.

Located between E2 and the N-terminus of E1 is a stretch of 55 (SFV) to 60 (SV) amino
acids known as the 6K peptide (Welch and Sefton, 1980; Garoff et al., 1980b; Rice and Strauss,
1981). The region is quite hydrophobic and little sequence homology is found between SFV and
SV. The 6K peptide is not found in virions (Welch and Sefton, 1979) and was proposed to
contain the signal peptide for E1 (Garoff et al., 1982). Bal31 exonuclease deletion analysis
localized the signal sequence activity to the last 26 amino acids of 6K preceding E1 (Melancon and Garoff, 1986). These authors also showed that E1 translocation was SRP-dependent and did not require p62. A model for the membrane insertion of alphavirus structural proteins in the ER is shown in Fig. 6 (Melancon and Garoff, 1987). Release of the 6K peptide from p62 and E1 is mediated by signal peptidase cleavage (Strauss and Strauss, 1984).

SFV E1 contains a single N-linked glycosylation site (Asn 141) which is occupied by a complex sugar moiety (Garoff et al., 1980b; Mattila et al., 1976; Pesonen and Renkonen, 1976). In SV there are two glycosylation sites on E1 (Asn 139, 245) which contain complex glycans when the virus is propagated in BHK cells, however when grown in primary chicken cells, one of the sugars is a high mannose type (Rice and Strauss, 1981; Sefton and Keegstra, 1974; Burke and Keegstra, 1976, 1979). Coexpression of p62 is required for conversion of E1 oligosaccharides to the endo H resistant form (Hashimoto et al., 1981; Melancon and Garoff, 1986). Presumably, a p62/E1 complex is necessary for transport of E1 from the ER and through the Golgi complex.

Among alphavirus isolates, E1 is more conserved than E2, and may confer genus subgroup variation (Bell et al., 1984). For this reason, E1 is thought to possess domains essential for host cell attachment and membrane fusion. A highly conserved group of 16 hydrophobic amino acids located 80 residues from the N-terminus of E1 is thought to function as the membrane fusion domain (White and Helenius, 1980; Garoff et al., 1980b; Rice and Strauss, 1981). The membrane fusion event occurs after a decrease in pH causes a conformation change in E1 evidenced by increase in resistance to proteolysis by trypsin (White and Helenius, 1980; White et al., 1983; Omar and Koblet, 1988). The latter group prepared SFV virions devoid of E3 and E2 and demonstrated that these particles were infectious, and capable of inducing cell-cell fusion after exposure to low pH. Fusogenic activity of SFV spike glycoproteins can be observed in BHK and COS cells microinjected or transfected with a p62/E1 construct, but not with just p62 or E1 (Kondor-Koch et al., 1983; Cutler and Garoff, 1986).
Figure 6. Model for the processing of Semliki Forest virus structural proteins. The capsid protein is removed by a self-catalyzed cleavage which occurs in the cytoplasm (a). Translocation of p62 into the ER is mediated by an uncleaved signal sequence [open rectangle] which becomes glycosylated [black dot] (b), and is halted by a carboxy-terminal transmembrane domain [solid rectangle] (c). A separate signal sequence (open rectangle) in the 6K peptide reinitiates translocation of E1 into the ER (d). The 6K peptide itself, is thought to remain anchored in the membrane by a 15 amino acid hydrophobic region [black box] (e,f). A carboxy-terminal transmembrane domain anchors E1 in the ER membrane (f). Glycosylation of E1 and p62 are indicated by black dots, and the two signal peptidase mediated cleavages that release the 6K peptide from p62 and E1 are indicated by arrows (e). (Melancon and Garoff, 1987).
Entry of alphaviruses in mammalian cells

Entry of alphaviruses into a host cell is initiated by binding of the glycoprotein spikes to an undefined protein receptor on the plasma membrane (for Review see Kielian and Helenius, 1986). Bound virions accumulate in coated pits which are endocytosed to form coated vesicles. The endocytic vesicles fuse with endosomes resulting in acidification of the vesicle contents (Tycko and Maxfield, 1982) and this causes a conformational change in the glycoprotein spike complex resulting in fusion of the viral envelope with the endosomal membrane. The nucleocapsid is then released into the host cell cytoplasm. Semliki Forest virus particles containing only glycoprotein E1 can induce the pH-dependent fusion and are infectious (Omar and Koblet, 1988). This recent evidence supports earlier work suggesting that E1 contained the fusogenic domain (Garoff et al., 1980b; White et al., 1983). Low pH shrinks the nucleocapsid diameter by 15% possibly allowing release of the RNA from the capsid matrix (Schlesinger, 1985).

Virus assembly and budding

Nucleocapsid formation is the first step in virus maturation, and occurs very rapidly in the cytoplasm. Levels of free capsid protein in alphavirus-infected cells are negligible due to complexing with genomic RNA within five minutes of synthesis (Soderlund, 1973). Protein and RNA are uniformly distributed throughout the core particle rather than the protein forming a shell around the nucleic acid (Simmons and Warren, 1984). The nucleocapsid is a stable structure, and remains so after the envelope has been removed by gentle detergent extraction (Helenius and Soderlund, 1973; Coombs et al., 1984). Coombs and Brown (1989) suggest that SV capsid proteins act similarly to histones by organizing the genomic RNA into nucleoprotein beads which interact with each other to form the core particle. The highly basic amino-terminal region of capsid is likely involved in binding of RNA through electrostatic interactions (Garoff et al., 1980a; Rice and Strauss, 1981). Alphaviral core particles are composed of 180 capsid molecules arranged in trimeric subunits (Fuller, 1987). On the surface of the nucleocapsid are 80
uniformly spaced indentations which presumably facilitate the binding of 80 trimeric
glycoprotein spikes (Fuller, 1987).

During nucleocapsid formation in the cytoplasm, the glycoproteins are synthesized on
membrane-bound ribosomes and then transported through the ER, Golgi and finally to the
plasma membrane which is the site of budding. The glycoprotein spike of SFV is a
heterotrimeric complex with the composition \((E1, E2, E3)_3\) (Vogel et al., 1986; Schlesinger and
Schlesinger, 1986) (Fig. 7). SV spike is similar except that E3 is not present (Fuller, 1987).
Oligomerization likely occurs in the endoplasmic reticulum shortly after synthesis, as is observed
with other enveloped virus glycoprotein trimers such as influenza hemagglutinin (HA), VSV G
protein and Rous sarcoma virus envelope glycoprotein (Wilson et al., 1981; Getthing et al., 1986;
Doms et al., 1987; Einfeld and Hunter, 1988). Trimerization is essential for transport of HA and
VSV G protein out of the ER, and failure to do so often results in formation of intermolecular
aggregates (Getthing et al., 1986; Doms et al., 1987). The role of oligomerization in transport has
not been well studied for alphaviruses, although it is clear that E1 is required to complex with
p62 prior to exiting the ER.

After leaving the ER, the p62/E1 complex traverses the Golgi stacks during which time
the carbohydrate chains are trimmed and processed. An endoproteolytic cleavage occurs in p62
after the Lys-Arg (SV) or Arg-Arg (SFV) pair at the E3 C-terminus during or shortly after
leaving the trans-Golgi network (Simons et al., 1973; Ziemiecki et al., 1980; Garoff et al., 1980b;
Rice and Strauss, 1981; DeCurtis and Simons, 1988). The protease is thought to be host-cell
derived and is likely cathepsin-B-like thiometalloprotease which performs similar cleavages in
prohormone precursors (Steiner et al., 1984). Following the p62 cleavage, a conformational
change in E2 takes place resulting in a much stronger interaction with E1 (Rice and Strauss,
1982).

A scheme for SFV budding was originally proposed by Garoff and Simons (1974).
According to their model, glycoproteins that arrive at the plasma membrane must have an
affinity for each other to allow concentration of the spikes into nucleation patches. As well as
FIGURE 7. Glycoprotein spike of alphaviruses. Domains noted are hydrophilic globular portion [1], membrane-spanning region [2], cytoplasmic fragment [3], and hydrophobic area not in the membrane [4]. [□□□] Oligosaccharides; [■] acetylated amino terminus; [□□□] covalent fatty acids; [□□□□] hydrophobic regions. (Schlesinger and Schlesinger, 1985)
increasing the concentration of glycoproteins in specific regions of the plasma membrane, nucleation of spikes selectively excludes host-cell membrane glycoproteins. Johnson et al. (1981) provided evidence that nucleocapsid-spike interactions may occur prior to spike arrival at the plasma membrane. They reported that in SV-infected cells, almost all of the surface viral glycoprotein was found in immobilized patches. Similarly, Pavan et al. (1987) found highly regionalized small clusters of spikes on SV-infected BHK plasma membranes which they attributed to be budding processes.

Since E1 has a cytoplasmic domain of only two amino acids (Garoff et al., 1980b; Rice and Strauss, 1981) it is likely that the 31 residue cytoplasmic portion of E2 interacts with nucleocapsids during budding. Early evidence came from cross-linking experiments which demonstrated a C-E2 interaction in SFV (Ziemiecki and Garoff, 1978). Vaux et al. (1988) recreated the C-E2 interaction by generating a series of three network antibodies. Antiidiotypic antibodies to anti-E2 cytoplasmic tail antibodies reacted with SFV capsid protein, and anti-antiidiotypic antibodies were specific for the E2 cytoplasmic domain. This work also demonstrated that the E2 epitope that binds to nucleocapsid was highly conserved among alphaviruses, and did not require E2 to be in the trimeric state \((E_1,E_2,E_3)_3\) to express the binding signal. According to Fuller (1987) there are two types of C-E2 interactions, loose and tight associations between the capsid triplets and spike complexes (Fig. 8).

Three main reasons may explain why budding of alphaviruses does not occur intracellularly; 1) low concentration of spikes on intracellular vesicles; 2) p62 conversion occurs very late in transport; 3) possible masking of E2 cytoplasmic epitopes by host proteins during transport (Schlesinger and Schlesinger, 1986). The importance of concentration and conformation of the spikes is made obvious from the facts that monensin which impairs transport through the Golgi, induces intracellular budding of SFV and SV as does the mannosidase inhibitor dMM (Kaariainen et al., 1980; McDowell et al., 1987). The E2 precursor p62 can also bind nucleocapsid at the plasma membrane, but envelopment does not normally occur, presumably because p62/E1 dimers can not pack together tightly enough (Brown and Smith, 1975). However, recent work by
Brown's group indicate that treatment of SV-infected cells with low concentrations of monensin results in the incorporation of p62 proteins into infectious virions, although the protein seems to be in an altered conformation (Presley and Brown, 1989). Their work is supported by the existence of a SV mutant which fails to undergo the p62 processing to E2, but still produces viable progeny, and therefore poses the question of why the p62 cleavage is necessary at all (Russell et al., 1989). These gross discrepancies regarding the nature and function of the p62 cleavage remain unresolved. Finally, although nucleocapsid assembly and glycoprotein transport are not sensitive to low concentrations of verapamil or chlorpromazine, these calcium blocking agents interfere with the final stages of virus maturation (Schlesinger and Cahill, 1989).
Figure 8. Association of spike complexes with the nucleocapsid at the plasma membrane. As the concentration of spike within the plasma membrane passes a critical threshold, lateral interactions result in the formation of close-packed patches of spike proteins and the exclusion of host membrane proteins (A, B). Two types of spike-capsid interactions are thought to occur causing the initial curving of the virus envelope and eventual budding (C). (Fuller, 1987).
Rubella virus

Classification

Rubella virus (RV) is the sole member of the genus Rubivirus in the family Togaviridae (Porterfield et al., 1978). Up until 1985, this family contained three other genera; Alphavirus, Pestivirus, and Flavivirus, but the Togaviridae Study Group recommended Flavivirus be recognized as the basis for a new family, Flaviviridae (Westaway et al., 1985). Togaviruses are simple enveloped RNA viruses whose genomes consist of a single molecule of single-stranded positive-polarity RNA arranged with a single species protein into an icosahedral nucleocapsid. Surrounding this nucleocapsid is a host cell-derived lipid bilayer with protruding glycoprotein spike complexes.

Morphology and morphogenesis

Electron microscopy studies of RV grown in BHK-21 cells indicated that virions were spherical and approximately 60 nm in diameter, with a 30 nm dense core enveloped by a lipid bilayer (Murphy et al., 1968; von Bonsdorff and Vaheri, 1969). Surface projections of 5-6 nm associated with the viral hemagglutinating properties have been reported (Holmes et al., 1969). The site of budding seems to depend on which host cell is used. For instance, maturation in BHK-21 cells occurs primarily in cytoplasmic vesicles (Golgi and vacuoles) and to a lesser extent at the plasma membrane, whereas in Vero cells, virions bud exclusively from the plasma membrane (Bardeletti et al., 1979). The details of RV morphogenesis remain unclear.

Clinical aspects

Infection by RV in humans results in symptoms ranging from the inapparent to characteristic features of German Measles such as adenopathy, malaise, low grade fever, and exanthem. Rash development may develop 16 to 20 days after exposure (Cooper and Buimovici-Klein, 1985). Due to the clinical variability in the disease, diagnosis can only be confirmed by isolation of RV or demonstration of primary seroconversion. Generally, virus
isolation is not employed because it is time consuming and expensive. When circumstances necessitate a precise diagnosis, co-cultivation of peripheral blood lymphocytes may be used to isolate virus (Chantler et al., 1982).

Complications such as polyarthralgia and arthritis following natural infection or vaccination are rare, however adolescent women are more prone to these conditions than are men. RV-infection is most serious when contracted by women in the first trimester of pregnancy resulting in fetal infection. In utero, the virus is extremely teratogenic due to mitotic inhibition in various fetal tissues. Very often cataracts, mental retardation, deafness, congenital heart disease, or death are the result. This multisystem of birth defects is collectively known as congenital rubella syndrome (Cooper and Krugman, 1969). Rarely, a progressive neurodegenerative condition known as progressive rubella panencephalitis may occur 10 to 20 years after rubella infection, or more frequently after congenital rubella (Townsend et al., 1975).

Live attenuated rubella vaccines such as RA 27/3 and Cendehill strains currently in use produce seroconversion in more that 95% of recipients (Cooper and Buimovici-Klein, 1985), and have effectively eliminated epidemic rubella.

Nucleic acid and genome organization

The genome of RV is an infectious single-stranded RNA molecule that sediments at 40S (Hovi and Vaheri, 1970). RV-infected cells in addition to containing 40S genomic RNA, harbor a 24S subgenomic mRNA that is capped and polyadenylated like the 40S species (Oker-Blom et al., 1984). The subgenomic mRNA is identical to the 3' one-third of the 40S RNA, and encodes the structural proteins (Oker-Blom et al., 1984). Recently the sequences of the RV 24S mRNAs from wild isolates M33 (Clarke et al., 1987; 1988), and Therien (Frey et al., 1986; Frey and Marr, 1988; Vidgren et al., 1987; Takkinen et al., 1988) strains of RV have been determined from cDNA clones. The mRNA is approximately 3346 nt, and contains a 3189 nt open reading frame specifying the structural proteins C, E2, and E1 (Frey and Marr, 1988). Alignment of the 24S sequences reported by the three groups reveals more than 95% homology at the nucleotide level,
and it is believed that some of the discrepancies can be attributed to difficulties in sequencing the unusually G/C rich RNA (Frey and Marr, 1988). No significant homology exists between the structural proteins of RV and those of alphaviruses (Frey and Marr, 1988). Nakhasi’s group recently deduced the sequence of the 24S mRNA from two vaccine strains of RV and relative to Therien strain, found 31 amino acid changes in RA27/3 (Nakhasi et al., 1989) and five in HPV77 (Zheng et al., 1989). Future work should be able to reveal which of these amino acid changes are important for attenuation. To date the genes encoding the non-structural proteins in the 5’ two-thirds of the 40S genome have not been characterized to any significant extent.

Expression of structural proteins

Vaheri and Hovi (1972) provided some of the earliest information regarding the structural proteins of RV. In 1983 however, the work of Oker-Blom et al. was crucial in elucidating the expression strategy of RV structural antigens. This group reported that RV virions contain four structural proteins; capsid (C), E2a, E2b, and E1 with respective molecular weights of 33,000, 47,000, 42,000, and 58,000 daltons. Molecular weight estimates for the structural proteins as determined by other investigators range as high as 38,000 (C), 47-54,000 (E2), and 62,000 (E1) (Waxham and Wolinsky, 1983). Immunoprecipitation of 35S-labelled RV proteins from tunicamycin treated cells revealed that E1 and E2 contain N-linked glycans (Oker-Blom et al., 1983).

In vitro translation of RV 24S mRNA yielded a 110,000 dalton polypeptide (p110) which is a precursor to the structural proteins (Oker-Blom et al., 1984; Clarke et al., 1987). Synchronized translation experiments determined the order of synthesis to be NH2-C-E2-E1-COOH (Oker-Blom, 1984). These results were later confirmed by sequence analysis of the 24S mRNA (Clarke et al., 1987). The expression strategy of the structural proteins is illustrated in Fig. 9.

Capsid protein

The capsid protein is a nonglycosylated protein (30-38K) associated with the 40S RNA
Figure 9. General strategy for the expression and processing of Rubella virus structural proteins. (Oker-Blom, 1984).
forming the nucleocapsid structure in the virion interior (Vaheri and Hovi, 1972). Monoclonal antibodies distinguish at least four nonoverlapping epitopes (Waxham and Wolinsky, 1985a, b). Sequence analysis of RV cDNA indicates the C protein is rich in arginine and proline, and has a maximum size of 300 amino acids (Clarke et al., 1987; Frey and Marr, 1988; Takkinen et al., 1988). The uncertainty in size is a consequence of not knowing the precise carboxy-terminus of C. Capsid has been suggested to consist of disulfide-bonded dimers in native virions (Dorsett et al., 1985). Contrary to alphaviruses, RV capsid protein does not possess an autocatalytic serine protease-like activity (Oker-Blom, 1984; Clarke et al., 1987).

E2 glycoprotein

RV E2 glycoprotein is one of the two components forming the spike complex on the virion surface. E2 migrates as a diffuse band ranging in size from 42-47,000, and in Therien strain can be resolved into two distinct species E2a (47,000) and E2b (42,000)(Oker-Blom et al., 1983). Analysis of the sugar moieties indicate both species contain terminal galactose and sialic acid but E2b labelled much more strongly with $^3$H-mannose than did E2a. Tryptic peptide mapping and tunicamycin treatment of infected cells suggest the electrophoretic heterogeneity in E2 results from glycosylation differences (Oker-Blom et al., 1983). In the presence of tunicamycin, Vero cells infected with RV and COS cells transfected with a RV 24S cDNA construct produced E2 with a molecular weight of 30,000 indicating that E2 is very heavily glycosylated (Oker-Blom et al., 1983; Clarke et al., 1988). Oker-Blom and coworkers suggest the sugars on E2 and E1 are of the complex type as determined by their resistance to endo H (Oker-blom et al., 1983). Interestingly, E2 immunoprecipitated from $^{35}$S-labelled cytoplasmic extracts of RV-infected cells is a uniformly glycosylated protein of 41,000 (Oker-Blom et al., 1983; Frey and Marr, 1988) or 43,000 (Waxham and Wolinsky, 1985b).

The biological function of E2 is obscure although it has been reported to possess strain specific epitopes (Dorsett et al., 1985) and at least one neutralization domain (Green and Dorsett, 1986). Production of monoclonal antibodies to E2 has met with limited success presumably due to poor immunogenicity as a result of association with E1 glycoprotein (Green et al., 1985;
Waxham and Wolinsky, 1985a, b). Digestion of intact virions with mixed glycosidases indicate the carbohydrate moieties on E2 are extremely resistant to attack, while those on E1 are easily removed (Ho-Terry and Cohen, 1984). From this data, the authors suggest that E2 epitopes may be buried under E1 in the spike complexes.

Amino acid sequence derived from E2 cDNA predicts a protein of 262-300 residues with three potential glycosylation sites in M33 and HPV77 strains and four in Therien and RA27/3 strains (Clarke et al., 1987; Frey and Marr, 1988; Vidgren et al., 1987; Nakhasi et al., 1989; Zheng et al., 1989). As is the case with capsid, the uncertainty in the size of E2 stems from disagreement over where the carboxy-terminus ends. Considering the high degree of glycosylation of E2, it seems likely that at least two if not three glycosylation sites are used.

Immediately preceding the known N-terminus of E2 is a stretch of mainly non-polar amino acids that is believed to be a signal peptide which functions to direct E2 to the ER membrane (Clarke et al., 1987; Vidgren et al., 1987; Frey and Marr, 1988). Release of capsid from the p110 polyprotein is thought to occur following cleavage of the E2 signal peptide by signal peptidase, but whether the signal sequence is removed from the capsid carboxy-terminus remains to be determined. The predicted membrane topology of E2 is that of a type 1 membrane protein (Singer et al., 1987), with a large amino-terminal ectoplasmic domain anchored in the envelope by 19 hydrophobic amino acids near the carboxy-terminus (Clarke et al., 1987; Vidgren et al., 1987; Frey and Marr, 1988). Waxham and Wolinsky (1985b) showed that processing of E2 includes covalent attachment of palmitic acid, likely somewhere in the carboxy-terminal region.

**E1 glycoprotein**

E1 (MW = 58,000) is the largest and most well studied of the RV structural proteins and contains hemagglutinin (HA) activity (Waxham and Wolinsky, 1983), as well as viral neutralization (VN) domains (Waxham and Wolinsky, 1985a; Green and Dorsett, 1986). The latter group has shown that HA and VN domains do not necessarily correlate. Ho-Terry and Cohen (1985) have localized the HA activity to a 13K tryptic cleavage product of E1. Using Staphylococcus V8 protease and trypsin generated peptides from E1, Terry et al. (1988) have

E1 is thought to be a type I membrane protein, anchored in the membrane by a 27 amino acid transmembrane domain in the C-terminus of the polypeptide (Frey and Marr, 1988; Vidgren et al., 1987). Near the carboxy-terminus of E2 preceding E1 is a group of 20 amino acids with the characteristics of a cleavable signal sequence (Frey et al., 1986; Nakhasi et al., 1986; Vidgren et al., 1987). Located between the putative transmembrane domain of E2 and the E1 signal peptide is a group of seven amino acids, five of them arginine residues, which is thought to be the site of a trypsin-like processing in vivo (Vidgren et al., 1987). The purpose of such a cleavage would be to release the signal peptide of E1 from the carboxy-terminus of E2.

Labelling RV in the presence of tunicamycin revealed that 5000 daltons of carbohydrate is added to E1 during processing (Oker-Blom et al., 1983). The same number of potential glycosylation sites (3) was found by sequencing E1 cDNA from M33 (Clarke et al., 1987), Therien (Frey et al., 1986; Nakhasi et al., 1986; Vidgren et al., 1987) and Judith strains of RV (Terry et al., 1988). Digestion of virions with endo H did not release any carbohydrate from E1 according to Oker-Blom et al (1983). In contrast, Bowden and Westaway (1985) report that both high mannose and complex glycans are found on E1 and E2 from lysates of RV-infected cells. E1 also contains covalently bound fatty acid (Waxham and Wolinsky, 1985b).

Replication

Humans are the only natural host for RV, and infectious virus can be isolated from peripheral blood mononuclear cells following natural and vaccine-induced infections (Buimovici-Klein and Cooper, 1979; Chantler and Tingle, 1982). In persistently infected individuals, virus has been recovered from eye and brain tissue, and synovial mononuclear cells (Menser et al., 1967; Weil et al., 1975; Chantler et al., 1981, 1982, 1985). No animal model exists that faithfully mimics the human disease, but a number of laboratory animals can be infected
resulting in only subclinical symptoms. BHK, RK-13, and Vero cells are the most commonly used cell lines for production of large amounts of virus.

Even though RV appears very similar to alphaviruses in structure and structural protein expression, the replication cycle kinetics are very different. Cells infected with alphaviruses generally reach maximal rates of virus production four to eight hours after infection resulting in cell death within 12-24 hours (Kaariainen and Soderlund, 1978). Rubella virus in contrast, has a latent period greater than 12 hours and reaches peak virus production between 24 and 48 hours post-infection depending on cell type (Hemphill et al., 1988). Cytopathic effect is minimal in most cell lines, and host cell protein synthesis seems to be only moderately decreased. Viral progeny released per cell is less than 100 with RV compared to 1000-10,000 for alphaviruses (Kaariainen and Soderlund, 1978; Wengler, 1980).

It has been suggested that the attachment or replication of RV is dependent upon a cell-cycle specific factor in cultured cells in light of the inability to achieve infection of greater than 10% of the cells after initial adsorption (Hemphill et al., 1988; Sedwick and Sokol, 1970; Wong et al., 1969). Eventually all of the cells in culture become infected. Another interesting feature of RV is the ability to cause a persistent infection in every cell line tested (Stanwick and Hallum, 1974; Norval, 1979; Cunningham and Fraser, 1985). Unlike alphaviruses which require the presence of defective interfering (DI) particles for initiation of persistence (Weiss et al., 1980), this property seems to be DI particle-independent in RV (Frey and Hemphill, 1988). The latter authors do suggest that DI particles may be involved in long-term stabilization of persistence in RV-infected cells.

Problems of slow replication and low virus yield have hindered the gathering of information regarding molecular details of RV replication. The assumption has been that the replication schemes of RV and alphaviruses are very similar. Northern blot analysis of RV-infected cells indicate that a genomic-length negative strand RNA species is present (Hemphill et al., 1988). This group also found that the 24S subgenomic mRNA was produced in larger amounts than the 40S plus RNA, and concluded both of these plus-strand RNAs are
derived from the negative-strand intermediate. Examination of the RNA from RV DI particles revealed that the 3' end of the genome was always present as is the case with alphaviruses (Frey and Hemphill, 1988; Monroe et al., 1982; Lehtovaara et al., 1982). No information is available about the 5' end of the RV 40S RNA as it has not yet been characterized.

In RV-infected cells, Bowden and Westaway (1984) found apart from the structural proteins, other prominent polypeptides of 150,000 (ns150) and 87,000 (ns87) daltons which they interpreted to be nonstructural antigens. In addition they identified minor proteins species with sizes of 111,000 (ns110), 92,000 (ns92), 77,500 (ns 77.5), 75,000 (ns75), 27,000 (ns27), 20,000 (ns20) and 17,000 (ns17). Pulse-chase experiments indicated that ns150 was derived from a 200,000 dalton precursor. Some of the proteins identified by this group, namely ns110, ns92, ns75, and ns27 may be equivalent to the p109, p92, p75, and p28/29 in RV-infected RK-13 cells (Chantler, 1979). Bowden and Westaway do not rule out the possibility that some of these proteins ie. ns110 and ns92 may be disulfide-linked dimers of E1 and E1-E2 reported previously with sizes of 105-109K and 95-96K (Ho-Terry and Cohen, 1980; Oker-Blom et al., 1983; Waxham and Wolinsky, 1983). They feel this is unlikely however due to the failure to detect 3H-mannose incorporation into these proteins. Obviously more work such as tryptic mapping and N-terminal sequencing must be done to rule out such possibilities.

In view of the limited data available regarding RV replication, it seems the overall scheme is very similar to that of alphaviruses, except that the latent period of RV is longer, and the kinetics much slower. Cloning of the nonstructural genes of RV should clarify the picture immensely.

Expression studies

Nakhasi et al. (1986) reported cloning and expression of RV E1 in COS cells, but their results remain questionable. The E1 construct they used, in addition to only encoding a 410 amino acid protein, lacked a functional signal peptide, yet supposedly directed synthesis of a 58K protein in transfected COS cells. RV E1 is known to be 481 amino acids long (Frey et al.,
1986; Vidgren et al., 1987; Clarke et al., 1988), and previous studies have shown that membrane glycoproteins that lack signal peptides do not become glycosylated, and are degraded very quickly in the cytoplasm (Gething and Sambrook, 1982; Sekikiwa and Lai, 1983; Jabbar and Nayak, 1987). Based on these observations, it is unlikely that the protein Nakhasi et al. reported is RV E1.

In 1987 Clarke et al. reported the cloning and sequencing of the RV 24S mRNA. SP6 RNA polymerase was used to transcribe mRNA from the cDNA, and then the RNA was translated using a rabbit reticulocyte cell-free translation system. In the absence of pancreatic microsomes, a RV-specific protein of approximately 110,000 daltons was produced which corresponds to the p110 precursor described by Oker-Blom et al. (1984). Translation of the mRNA in the presence of microsomes resulted in some processing to smaller proteins which corresponded to E1, E2 and C. The RV 24S cDNA was subcloned into an SV40 expression vector and transfected into COS cells resulting in synthesis of proteins antigenically similar to E1, E2, and C (Clarke et al., 1988). Treatment of the transfected COS cells with tunicamycin confirmed that E1 and E2 contain N-linked glycans (Oker-Blom et al., 1983; Clarke et al., 1988).

Two other reports of RV E1 expression have been published in 1989. de Mazancourt and Perricau (1989) using a cytomegalovirus-based vector containing amino acids 17-481 of Therien E1 fused to the translation start site of SV40 large T antigen gene, obtained transient expression of E1 in 293 cells. Detection of RV antigen was accomplished using immunofluorescence only, and no other details regarding protein size, stability, or intracellular location were reported. Terry et al. (1989) designed RV E1 constructs as fusion proteins with the IgG-binding region of *S. aureus* protein A enabling expression in *E. coli*. One plasmid encoded the entire E1 gene, while the other contained only amino acids 207-353 which comprises a region they previously identified to contain three non-overlapping epitopes (Terry et al., 1988). Both fusion proteins reacted on Western blots to monoclonal antibodies raised against whole virus confirming that there are glycosylation independent epitopes on E1. Introduction of the bioengineered E1 fusion proteins into rabbits resulted in the production of low levels of
neutralizing antibodies, however no hemagglutination inhibiting antibodies were identified. From these data the authors concluded that these E1 fusion proteins were poorly immunogenic in their present form.

**Project rationale and thesis objectives**

Despite being an important human pathogen, the cell biology of RV remains poorly characterized. Much of this lack of information is a result of the difficulty in growing of RV and the subsequent poor yields of biologically active virus. This, and the apparent similarity of RV with prototype alphaviruses, has left RV investigators with few alternatives but to constantly refer to alphavirus models. Despite the similarities in genome organization and structural protein expression strategies, RV is clearly very different from alphaviruses when one considers the relatively long latency period, slow replication kinetics and limited cytopathology exhibited by RV infection in cultured cells. In addition, initiation of persistent infection by RV does not seem to be dependent upon the presence of defective-interfering particles.

The structural proteins of RV must contain information that facilitates their delivery from the site of synthesis to a location within the cell where virion assembly takes place. Consequently targeting and transport of RV structural proteins expressed from cDNAs should parallel events which occur in infected cells. Using in vitro and in vivo expression systems to express RV cDNAs, the aim of this project is to characterize events such as processing of RV p110 precursor, targeting to the ER membrane, and intracellular transport in COS cells. In addition the number of N-linked glycans on E1 and E2 will be determined and the possible importance of glycosylation in protein targeting will be discussed.
MATERIALS and METHODS

Materials and supplies

DNA modifying enzymes and restriction endonucleases were purchased from Bethesda Research Laboratories (BRL), Promega Biotec, New England Biolabs, Boehringer Mannheim, Pharmacia, and United States Biochemical Corp. [9,10$^3$H]-Palmitic acid (28.5 Ci/m mole), [$^{32}$P]-ATP (3000 Ci/m mole), and L[$^{35}$S]-methionine (600-800 Ci/m mole) were from Amersham and/or New England Nuclear. Fluorescein-conjugated goat anti-human IgG was from Kirkegaard and Perry Laboratories or Tago Inc., and Rhodamine(TRITC)-conjugated goat anti-mouse antibody was from Zymed. Anti-Rubella human serum was a gift from Dr. Aubrey Tingle, Dept. of Pediatrics, University of British Columbia. Mouse monoclonals to E1 were produced in this lab previously, and others were obtained from Dr. John Safford, Abbott Laboratories. Anti-C monoclonal antibody was also obtained from Dr. John Safford. Monoclonals to E2 were a gift from Dr. Jerry Wolinsky, Dept of Neurology, University of Texas, Houston. Tissue culture reagents were from Gibco and Sigma. TRITC-conjugated lectins were from Sigma, and COS cells were obtained from Dr. David Russell (Dept. of Molecular Genetics, University of Texas, Dallas). E.coli strains DH5$^a$ and DH5$^a$F$'$ cells were purchased from BRL, and CJ236 was obtained from Bio-Rad.

Propagation and handling of E.coli strains

E.coli strains containing recombinant plasmids were grown in LB medium (1% tryptone; 0.5% yeast extract; 1% NaCl) containing 100 µg/ml ampicillin. CJ236 cells were grown in medium containing 30 µg/ml chloramphenicol to maintain the F$'$ factor. For long term storage the bacterial strains were stored in 15% glycerol at -70°C.

For screening M13 recombinants, transformed DH5$^a$F$'$ cells were mixed at 45°C with three ml of YT top agar (0.8% tryptone; 0.5% yeast extract; 0.5% NaCl; 0.6% agar) containing 0.05 ml of 5-bromo-4-chloro-3-indolyl-$\beta$-D-galactoside (XGAL) at 2 mg/ml in dimethylformamide and 0.02 ml of 100 mM isopropyl-$\beta$-D thiogalactopyranoside (IPTG), and poured onto YT agar (1.5%)
plates. Recombinants gave rise to clear plaques, while wild type plaques were blue. Phage were propagated by infecting DH5αF cells with isolated plaques, and growing in 2XYT broth (1.6% tryptone; 1% yeast extract; 0.5% NaCl) for six hours.

**Preparation of competent cells and transformation**

Competent cells were prepared using a method described in Promega-Biotec technical bulletin 018. Briefly, *E. coli* cells were grown in 20 ml of LB broth until the absorbance at 600 nm reached 0.3-0.5. Cells were centrifuged at 5000 rpm in a Sorvall SS34 rotor at 4°C for five minutes, and the supernatant was discarded. All subsequent procedures were carried out at 4°C. The bacterial pellet was resuspended in 10 ml of cold solution A [10 mM MOPS (pH 7.0); 10 mM RbCl], and centrifuged as above. Cells were then gently resuspended in 10 ml of ice-cold solution B [10 mM MOPS (pH 6.5); 10 mM RbCl; 50 mM CaCl₂] and held on ice for 30 minutes. After pelleting the cells as above, they were resuspended in 1 ml of solution B plus 15% glycerol, and quick frozen in 0.2 ml aliquots in dry ice-ethanol and stored at -70°C.

To transform competent cells, 10-50 ng of DNA was added to 0.2 ml of competent cells, mixed gently and kept at 0°C for 20-30 minutes. After a two minute heat shock at 37°C, 1 ml of LB broth was added to the transformation mixture and the cells were allowed to recover at 37°C for one hour before plating onto selective media. Transformation of DH5αF or CJ236 cells was identical to the procedure above except that cells were plated out onto YT plates without antibiotic.

**Isolation of plasmid and single-stranded DNA from *E. coli***

**Mini-prep plasmid DNA isolation**

This procedure was employed to isolate plasmid DNA or M13 recombinant replicative form (RF) DNA from small liquid cultures (1.5-3.0 ml), and is essentially the alkaline-lysis method described by Maniatis et al. (1982). Bacterial cells from 1.5 ml of culture were pelleted for one minute in a microfuge, and the supernatant was discarded. The pellet was resuspended
in 100 µl of 50 mM glucose; 10 mM EDTA; 25 mM Tris-Cl (pH 8.0)) and lysed by the addition of
200 µl of 0.2N NaOH; 1% SDS for five minutes at 0°C. Chromosomal DNA and proteins were
precipitated by incubating the lysis mixture with 150 µl of cold potassium acetate (3M K⁺; 5M
CH₃COO⁻ (pH 4.8)) at 0°C for five minutes, followed by centrifuging in a microfuge for five
minutes at 4°C. The supernatant was extracted with an equal volume of phenol:chloroform (1:1),
and DNA precipitated with two volumes of ethanol at room temperature (RT) for five minutes.
Plasmid DNA was recovered by centrifugation in a microfuge for five minutes at RT, washed in
70% ethanol, dried in a Speed Vac Concentrator, and resuspended in 50 µl of TE [10 mM Tris-Cl
(pH 8.0); 1 mM EDTA] containing 20 µg/ml RNAse A. Aliquots were used for restriction
analysis or subcloning.

Large scale plasmid DNA preps

The protocol is a procedure obtained from Promega Biotec technical bulletin 009
developed by Dr. P Krieg and Dr. D. Melton of Harvard University, but with slight
modifications. Cells grown in selective media overnight in 250 ml cultures were pelleted by
centrifugation at 5000 rpm in a Sorvall GSA rotor at 4°C for five minutes. The pellet was
resuspended in five ml of 50 mM Glucose; 10 mM EDTA; 25 mM Tris-Cl (pH 8.0), then 20 mg of
lysozyme dissolved in one ml of the same buffer was added, and the mixture incubated on ice.
Twenty minutes later, cells were lysed by addition of 12 ml of 0.2 N NaOH; 1% SDS and
incubation on ice for 10 minutes. Chromosomal DNA and proteins were precipitated with eight
ml of cold potassium acetate solution (see mini-prep procedure) on ice for 20 minutes, followed
by centrifugation at 10,000 rpm in a SS34 rotor at 4°C for 15 minutes. RNAse A (100 µg) was
added to the cleared lysate followed by incubation at 37°C for 30-45 minutes. The lysate was
extracted twice with equal volumes of phenol:chloroform, and the nucleic acids precipitated with
one volume of isopropanol at RT for five minutes. Remaining nucleic acids were recovered in
siliconized Corex tubes by centrifuging at 10,000 rpm for 10 minutes at RT in a SS34 rotor. The
pellet was dried, and dissolved in 1.6 ml of sterile water. DNA was selectively precipitated by
addition of 0.4 ml of 4 M NaCl and 2.0 ml 13% polyethylene glycol (PEG mw 8,000), mixing and
incubation on ice for 60 minutes. The plasmid DNA was pelleted in siliconized Corex tubes at 10,000 rpm for 15 minutes at 4°C in a SS34 rotor, washed with 70% ethanol, dried, and dissolved in TE.

Isolation of single-stranded DNA

M13mp18 or mp19 recombinants were propagated by adding a freshly cored plaque to 2 ml of 2XYT containing 20 ul of stationary phase DH5αF’ cells, and shaking vigorously at 37°C for five to seven hours. Cells from 1.5 ml of culture were pelleted by centrifuging in a microfuge for five minutes at RT, and RF DNA was isolated as described in "mini-preps" using the alkaline-lysis method. Phage were precipitated from the supernatant by addition of 0.25 ml of 25% PEG (mw 8,000); 2.5 M NaCl, vortexing, and incubating at RT for 20 minutes. Phage were spun down for five minutes in a microfuge, the supernatant carefully aspirated, and the pellet resuspended in 0.1 ml of TE by vigorous vortexing. The suspension was successively extracted with equal volumes of phenol, phenol:chloroform, and ether. Ten μl of 3M sodium acetate (pH 5.5) was added followed by 0.25 ml of ethanol, and the ssDNA recovered by centrifugation after 30 minutes at -20°C. The DNA was washed with 70% ethanol, dried, resuspended in 30 μl TE, and used for sequencing or oligonucleotide-directed mutagenesis.

Oligonucleotide-directed mutagenesis

Purification and phosphorylation of oligonucleotides

Deoxyribo-oligonucleotides were synthesized in the laboratory of Dr. M. Smith (UBC, Dept. of Biochemistry). The crude oligonucleotide was dissolved in one ml of sterile water and loaded onto a C18 SEP-PAK (Waters Scientific) that had been pretreated with 10 ml of HPLC grade acetonitrile and then 10 ml of sterile water. Unbound material was washed from the reverse-phase column with 10 ml of water, and the bound oligonucleotide was eluted with three ml of 20% acetonitrile in three one ml fractions. Spectrophotometry at 260 nm indicated that 80-90% of the A260 material was found in the first fraction. The sample was dried in a Speed Vac, redissolved in 0.5-1.0 ml of water, and the concentration determined by reading the
absorbance at 260 nm. Oligonucleotide concentrations were adjusted to 10-20 nmoles/ml, and 100-200 pmoles of mutagenic oligonucleotide was kinased in 30 µl of 50 mM Tris-Cl (pH 7.5); 10 mM MgCl₂; 10 mM DTT; 0.33 mM ATP; 300 units/ml T4 polynucleotide kinase for 45 minutes at 37°C. Reactions were terminated by heating at 65°C for 10 minutes.

Mutagenesis

Single primer mutagenesis based on the method of Kunkel (1985) with minor modifications was employed to introduce desired changes in target DNA. Uracil-containing single-stranded template was isolated from phage grown on CJ236 cells in 2XYT with 0.25 µg/ml uridine. Kinased mutagenic oligonucleotide (5-10 pmole) was annealed to the template DNA (0.5-1.0 pmole) in 10 µl of 20 mM Tris-Cl (pH 7.4); 10 mM MgCl₂; 50 mM NaCl at 70°C for five minutes, then allowed to cool to <30°C over 30-45 minutes. Extension of the primer was carried out in 15 µl of 0.5 mM each dNTP; 1 mM ATP; 30 mM Tris-Cl (pH 7.4); 15 mM MgCl₂; 50 mM NaCl; 2 mM DTT for 10 minutes on ice, then 37°C for 90 minutes. Reactions were then diluted to 50 µl with TE or water, and 5 µl was used to transform DH5αF’ cells. Note, some extension reactions were carried out at room temperature when using some primers (dependent upon the calculated melting temperature of the oligo).

Identification of mutants

Dot blot analysis

Plaques from mutagenesis reactions were used to infect DH5α F’ cells and ssDNA and RF DNA were isolated as described. One µl of ssDNA was spotted onto Hybond-N (Amersham), dried, and cross-linked by exposure to a UV transilluminator for five minutes. Filters were prehybridized in 6X SSC; 10X Denhardts solution; 0.2% SDS at 50-60°C for at least one hour, then hybridized to the mutagenic primer which had been kinased with T4 polymerase and γ-[³²P]-ATP in the same solution minus the SDS. Hybridization proceeded for at least two hours and temperatures varied with the particular oligonucleotides, but were generally 10-15°C lower than the predicted melting temperature (Tm) of the primer. After hybridization, filters were washed at RT in 6X SSC twice, then progressively washed more stringently by increasing the
wash temperature by 5°C increments. After each wash, the filters were autoradiographed. DNAs which continued to give a strong signal after the probe washed off the non-mutagenized control DNA were sequence to confirm the presence of the mutation.

**Restriction analysis**

In some instances, the mutations introduced new restriction sites into the target DNA which could be detected by restriction endonuclease digestion of the RF DNAs.

**Dideoxysequencing of DNA**

Sequence of ssDNA was determined using Sequenase T7 DNA polymerase (United States Biochemical Corporation) and dideoxynucleotides for chain termination. The procedure is a modification of the chain terminator method developed by Sanger et al. (1977). Sequencing primer (1-5 pmole) was annealed to ssDNA (0.5-1.0 μg) in 10 μl of sequencing buffer [40 mM Tris-Cl (pH 7.5); 10 mM MgCl₂; 50 mM NaCl] at 65°C for five minutes then slow cooled to RT. Labeling of the mixture was carried out by adding the following components and incubating at RT for five minutes; 1 μl of 100 μM DTT; 2 μl diluted labelling mix (1.5 μM each dGTP, dCTP, dTTP); 10 μCi of α-[³²P]-dATP; 2-4 units Sequenase. After the labelling reaction, 3.5 μl aliquots were transferred to prewarmed microfuge tubes containing 2.5 μl of A,G,C, and T termination mixes, and incubated at 37°C for five minutes. Composition of the terminator mixes were as follows; G mix (80 μM each dNTP; 8 μM ddGTP); A mix (80 μM each dNTP; 8 μM ddATP); C mix (80 μM each dNTP; 8 μM ddCTP); T mix (80 μM each dNTP; 8 μM ddTTP). The reactions were stopped by adding 4 μl of STOP solution (95% formamide; 20 mM EDTA; 0.05% bromophenol blue; 0.05% xylene cyanol FF). Samples were denatured at 95°C for three minutes and quenched on ice before separation. Sequencing gels contained 6-8% polyacrylamide; 7M urea; 0.5X TBE and were cast between siliconized 40 X 20 cm glass plates, and were run at 1600 volts for 2-3 hours before drying and autoradiography.

**Restriction endonuclease digestions and DNA modification**

All restriction digestsions reactions were performed according to assay conditions specified
by suppliers.

DNA fragments were ligated using T4 DNA ligase in 50 mM Tris-Cl (pH 7.6); 10 mM MgCl₂; 1 mM ATP; 1 mM DTT; 5% (w/v) polyethylene glycol for four hours at RT except for blunt-ended fragments which were ligated overnight. Reactions were diluted at least five-fold with TE before transformation.

DNA fragments with 5' overhangs were made blunt-ended by incubating with E.coli DNA polymerase I Klenow enzyme in 50 mM Tris-Cl (pH 7.2); 10 mM MgSO₄; 100 mM DTT; 50 mM BSA; 80 μM dNTPs; for 10-30 minutes at RT. The enzyme was destroyed by heating at 70°C for 5-10 minutes.

Fragments with 3' protrusions were converted to flush ends using T4 DNA polymerase in 33 mM Tris-acetate (pH 7.9); 66 mM potassium acetate; 10 mM magnesium acetate; 0.5 mM DTT; 100 mg/ml BSA; for five minutes at 37°C. Reactions were adjusted to 25 mM EDTA, and the DNA purified by phenol:chloroform extraction and ethanol precipitation, or using GENECLEAN (BIO 101).

Removal of terminal 5' phosphates from DNA fragments with 5' overhangs was accomplished using calf intestinal alkaline phosphatase (CIP) in 50 mM Tris-Cl (pH 9.0); 1 mM MgCl₂; 0.1 mM ZnCl₂; 1 mM spermidine; at 37°C for two successive 30 minute incubation periods. A second aliquot of CIP was added after the first 30 minute period. To dephosphorylate fragments with 5' recessed or blunt ends, the incubations were carried out at 56°C. CIP reactions were terminated by phenol:chloroform extractions followed by ethanol precipitation or purification using GENECLEAN.

Purification of DNA fragments from agarose gels or enzyme reaction mixtures was routinely done using GENECLEAN. Desired fragments were excised from ethidium bromide stained TAE agarose gels, and the gel matrix solubilized in 2-3 volumes of saturated sodium iodide at 55°C. DNA was removed from the agarose solution by vortexing the mixture with a suspension of Glassmilk™, and a brief spin in a microfuge. Contaminants were washed away from the glass bound DNA by three successive washed with a concentrated NaCl/ethanol/water
solution. The DNA was eluted from the glass fibers with TE or water at 55°C, and was suitable for all subsequent manipulations after this type of purification.

Common solutions

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSC</td>
<td>0.15 M NaCl; 0.015 M Na citrate (pH 7.0)</td>
</tr>
<tr>
<td>10X Denhardtts</td>
<td>0.1% each ficoll, polyvinylpyrrolidone, BSA</td>
</tr>
<tr>
<td>TBE</td>
<td>89 mM Tris-borate (pH 8.0); 89 mM boric acid; 2 mM EDTA</td>
</tr>
<tr>
<td>TAE</td>
<td>40 mM Tris-acetate; 1 mM EDTA (pH 8.0)</td>
</tr>
<tr>
<td>TE</td>
<td>10 mM Tris-Cl (pH 8.0); 1 mM EDTA</td>
</tr>
</tbody>
</table>

Expression vectors

The multiple cloning sites of pSPT18 and pSPT19 (Pharmacia) are flanked by oppositely oriented T7 and SP6 RNA polymerase promoters which allow transcription of either strand of a DNA insert. Synthetic mRNA from cDNAs cloned into these vectors was used to direct translation in a rabbit reticulocyte lysate system.

For transient in vivo expression of RV cDNAs in COS cells, pSVL (Pharmacia) and/or pCMV5 (D. Russell, Texas) were used. These vectors direct transcription of inserts by utilizing the SV40 late promoter (pSVL) or the human Cytomegalovirus major immediate early gene promoter (pCMV5). Both vectors contain the SV40 origin of replication allowing replication in COS cells, as well as a prokaryotic origin of replication and ampicillin resistance gene for growth in E.coli (Fig. 10).

Rubella virus cDNA constructs

RV cDNAs used in this study are all derived from the full-length 24S cDNA described by Clarke et al. (1987) pSPT19 (C/E2/E1). I have taken the liberty of renaming the cDNA insert encoding C/E2/E1, simply 24S. The 3.4 kb insert from pSPT19 (C/E2/E1) was excised with EcoRI and HindIII and cloned into pCMV5 to create pCMV5-24S. This construct when
Figure 10. Physical maps of the expression vectors pSPT18/19, pSVL (Pharmacia) and pCMV5 (Andersson et al., 1989).
transfected into COS cells directed production of RV structural proteins in vivo.

Plasmid pE1 was constructed by removing the 1500 nt NcoI fragment from pSPT19(C/E2/E1) (Clarke et al., 1987) and religation, creating a cDNA encoding the capsid translation start site with eight residues from C, 69 residues from the C-terminal E2, and all of E1, including the putative signal peptide (Fig. 11).

To construct pE2E1, a 990 nt fragment containing the entire RV E2 gene was excised from the RV cDNA plasmid pSPT19(C/E2/E1) using SphI and PstI, and was inserted into pSPT18 digested with the same enzymes. This plasmid was then linearized upstream from the E2 coding region using HindIII, dephosphorylated, and the ends repaired by incubating with the Klenow fragment of DNA polymerase I. NcoI linkers (P-CCCATGGG-OH) treated with polynucleotide kinase and ATP were ligated to the plasmid, and clones containing the desired NcoI site were selected. A 800 nt NcoI fragment containing all of the 5' region of E2, but not the transmembrane domain, was inserted into pE1. The resulting plasmid pE2E1, contains the entire E2 and E1 coding regions including the 5' signal sequences and 3'transmembrane domains. See Fig. 11.

E1 sequences were removed from pE2E1 by digestion with PstI and recircularization to create pE2. Plasmid pE2 contains the coding regions for the entire E2, as well as those for the carboxyl-terminal 58 residues and the amino-terminal eight residues of the capsid protein. The translation initiation site is contained in the amino-terminus of the capsid protein. See Fig. 11. A list of mutagenic oligonucleotides used in this study is contained in table II.

SV-40 based vector plasmids were made by excising inserts from pSPT19 recombinant plasmids with EcoRI and HindIII, converting the cohesive ends to blunt ends with Klenow enzyme, and ligation into the SmaI site of pSVL. Constructs were then screened for orientation by restriction analysis. The pCMV5 recombinants were made by cloning the RV cDNA inserts into the EcoRI and HindIII sites of this vector. Constructs were named according to the parent vector and insert they contain, ie. pSVL-E2, pCMV5-E2E1, etc.
Figure 11. Schematic diagram of RV cDNA constructs. Respective portions of capsid (C), and envelope glycoproteins E2 and E1 in the constructs are indicated as C, E2, and E1 above the schematics. The flanking 5' EcoRI and 3' HindIII sites are marked (E and H respectively). The putative signal peptides and transmembrane regions are indicated by (|) and (Q) respectively. Potential N-linked glycosylation sites are represented by (Y), and the arginine rich region between the transmembrane anchor of E2 and signal peptide of E1 is indicated (r). Translation start sites in the constructs are marked by arrow heads.
In vitro transcription;

Five micrograms of plasmid was linearized with HindIII, extracted with phenol/chloroform, and precipitated with ethanol. The DNA template was then added to a transcription-capping reaction containing 40 mM Tris-chloride (pH 7.5); 6 mM MgCl$_2$; 2 mM spermidine; 10 mM NaCl; 10 mM dithiothreitol; 500 units/ml RNAsin (Promega Biotec); 100 micrograms/ml nuclease-free bovine serum albumin; 0.5 mM each ATP, CTP, UTP, and cap analogue GpppG; 0.1 mM GTP; 600 units/ml SP6 RNA polymerase (Promega). Reactions were typically in 50 µl volumes, and were incubated for one hour at 37°C after which the DNA was removed by DNase I digestion. The reactions were extracted successively with phenol/chloroform and ether, and precipitated with ethanol, taken up in 30 µl of water and stored at -70°C.

In vitro translation;

Translation of SP6-derived transcripts was performed in a nuclease-treated rabbit reticulocyte lysate system (Promega) containing 0.02 mM amino acid mixture minus methionine; $^{35}$S-methionine at 1200µCi/ml; RNAsin at 1600 units/ml; and RNA at 40 µg/ml in a 25 µl volume. Some reactions were supplemented with 1 µl of canine pancreatic microsomes (Promega). After incubation at 30°C for one hour, translation products were assayed for protection from exogenously added protease and analyzed by electrophoresis on a 10% Laemmli gel, followed by fluorography at -70°C.

Protease protection assay;

Protease digestion conditions were essentially as described (Walter and Blobel, 1983). After cell-free translation, CaCl$_2$ was added to 1 mM, and the reactions held on ice for 10 minutes to stabilize the membranes. Samples were then incubated for 60 minutes on ice in the presence of trypsin (300µg/ml). Proteolysis was stopped by addition of phenylmethanesulfonyl fluoride to 1 mM, trypsin inhibitor to 500 µg/ml, and incubating on ice for 10 minutes. Triton X-100 was included at a final concentration of 1% in some reactions. Samples were separated on 10% Laemmli gels (1970), and fluorographed.
Immunoprecipitations, Endo-\(\beta\)-N-acetylglucosaminidase H and N-glycanase digestions

Human polyclonal anti-rubella serum (a gift from Dr. A. Tingle) was preincubated with Protein A-Sepharose (Pharmacia) for at least four hours at 4\(^\circ\)C in binding buffer (100 mM Tris-Cl (pH 7.4); 400 mM NaCl) with constant mixing. The serum-coated beads were washed twice with binding buffer, and once in lysate buffer [25 mM Tris-Cl (pH 7.4); 100 mM NaCl; 1 mM EDTA; 1% Nonidet P-40 (NP-40)]. \(^3S\)-labelled antigen (COS cell lysate or \textit{in vitro} translation products) was mixed overnight at 4\(^\circ\)C with the coated beads in lysate buffer. Beads were washed once with lysate buffer, twice with wash buffer (25 mM triethanolamine; 172 mM NaCl; 1% deoxycholate; 0.1% SDS; 1 mM EDTA), three times with 10 mM Tris-Cl (pH 7.4), and once with water. Immune complexes were dissociated from the Protein A-Sepharose by boiling in 100 mM sodium citrate (pH 5.5); 1 mM phenyl-methylsulfonylfluoride (PMSF); 0.15% SDS for five minutes, vortexing, and pelleting the beads by centrifugation. Supernatants were collected and incubated at 37\(^\circ\)C with or without endo-\(\beta\)-N-acetylglycosidase H (Genzyme) at 25 milliunits/ml for at least 8 hours. Immunoprecipitates to be digested with N-glycanase were adjusted to 0.17% SDS; 0.2 M sodium phosphate (pH 8.6); 1.12% nonidet P-40; 10 mM 1,10 phenanthroline hydrate and digested with 125 milliunits of N-glycanase for at least eight hours. Samples were boiled for three minutes and separated on SDS-PAGE, and then fluorographed. To recover cell surface antigens, radiolabeled COS cells were washed three times with cold PBS (145 mM NaCl; 7.5 mM \(\text{Na}_2\text{HPO}_4\); 2.9 mM \(\text{NaH}_2\text{PO}_4\); 0.7 mM CaCl\(_2\); 0.3 mM MgCl\(_2\)), then incubated with 10% normal goat serum for 10 minutes at 0\(^\circ\)C to block nonspecific binding of antibody. Cells were then incubated with human anti-RV serum for 60-90 minutes at 0\(^\circ\), lysed and nuclei removed by centrifugation at 13,000 revolutions per minute in a microfuge. Supernatants were adjusted to 0.3% SDS and the immune-complexes were precipitated with Protein A-sepharose. Supernatants were incubated with anti-RV serum and Protein A-sepharose to recover intracellular RV antigen.
**Transfection of COS cells**

COS cells were transfected with plasmid DNA constructs according to the method of Adams and Rose (1985) with some modifications. Subconfluent monolayers grown in DMEM/FCS (Dulbecco's modified Eagle medium plus 5% fetal calf serum) were washed twice in Tris-saline (25 mM Tris-Cl (pH 7.4); 140 mM NaCl; 3 mM KCl; 1 mM KCl; 0.5 mM MgCl$_2$; 0.9 mM Na$_2$HPO$_4$) warmed to 37°C, and then overlaid with Tris-saline containing 5 µg/ml of plasmid DNA, and 500 µg/ml DEAE-dextran Mr=5x10$^5$. After 30 minutes at 37° C, the DNA solution was removed, and replaced with medium containing 80 µM chloroquine. After three hours at 37° C, this solution was removed, and the cells were subjected to a shock treatment with 10% dimethyl sulfoxide/DMEM for three minutes at room temperature, followed by two washes with Tris-saline. COS cells were kept at 37° C post-transfection in DMEM/FCS for 40-50 hours before labelling with $^{35}$S-methionine or processing for immunofluorescence.

**Metabolic labelling**

Transfected COS cells were washed once with DMEM minus methionine, then starved in this medium for 20 minutes at 37° C. Cellular proteins were pulse-labelled by incubating cells with 0.5 ml DMEM minus methionine/dialyzed FCS containing 100 µCi $^{35}$S-methionine per 35 mm well for the indicated times at 37° C, followed by various chase periods with DMEM/FCS plus 2 mM methionine for some samples. After washing with cold Tris-saline, cells were lysed with 400 µl of 50 mM Tris-Cl (pH 8.0); 62.5mM EDTA; 1% NP-40; 0.4% sodium deoxycholate; 1mM PMSF. Where indicated, some cells were lysed in the presence of 10 mM iodoacetamide to prevent formation of aberrant disulfide bonds. Lysates were scraped off the plates, cleared of nuclei and debris by centrifugation, then used for immunoprecipitation. Cells were labelled with 500 µCi $^3$H-palmitate in 1.0 ml of medium for 12 hours at 37°C.

**Indirect immunofluorescence**

Transfected COS cells grown on polylysine-coated 9 mm plastic coverslips were washed...
three times with PBS, and fixed for 20 minutes at room temperature in 2% formaldehyde/PBS, followed by washing with PBS. Some cells were permeabilized with 0.06% NP-40/PBS for 30 minutes prior to blocking with 1%BSA/PBS. All washes and solutions after this step utilize BSA/PBS. Coverslips were overlaid with diluted human serum (1:200) or mouse monoclonals (1:75) incubated for 60 minutes at room temperature, then washed. Incubation with secondary antibody, fluorescein-conjugated goat anti-human or anti-mouse IgG (Tago) diluted 1:100, was for 60 minutes. Coverslips were washed, mounted, examined using epifluorescence, and photographed. For double-labelling using lectin-conjugates, permeabilized cells were incubated with Wheat Germ agglutinin-Rhodamine (WGA-TRITC) to visualize Golgi and post-Golgi structures or Concanavalin A-Rhodamine (Con A-TRITC) for ER staining at 10-15 µg/ml for 30 minutes at RT prior to blocking with BSA.

Electrophoresis

Separation of DNA fragments

DNA samples were adjusted to 8% sucrose; 20 mM EDTA (pH 8.0); 0.05% bromophenol blue; 0.05% xylene cyanol and separated by electrophoresis (four to eight volts/cm) on 10 cm submarine horizontal agarose gels containing 1 µg/ml of ethidium bromide. Gel buffers consisted of 1X TAE.

Separation of proteins

Proteins were separated using a discontinuous gel system described by Laemmli (1970). Samples were adjusted to 62.5 mM Tris-Cl (pH 6.8); 10% glycerol; 2% SDS; 2% 2-mercaptoethanol (optional) and denatured at 95°C for two minutes. Stacking gels consisted of 4% polyacrylamide, and separating gels generally contained 10% polyacrylamide. Gels were run at 25-55 milliamperes until the bromophenol blue reached the bottom of the separating gel. The stacking gel was trimmed away, and the proteins fixed by soaking in 10% acetic acid for 20-30 minutes. Gels were then immersed in the fluorographic agent Amplify (Amersham) for 20-30 minutes, dried under vacuum and exposed to XAR or XRP film at -70°C.
Solutions used for electrophoresis

5X Stacking gel buffer  
0.625 M Tris-Cl (pH 6.8); 0.5% SDS

5X Separating gel buffer  
1.875 M Tris-Cl (pH 8.8); 0.5% SDS

5X Running gel buffer  
0.125 M Tris-Cl; 0.96 M glycine; 0.5% SDS (pH 8.3)

Gel Stock (polyacrylamide solution)  
30% acrylamide; 0.8% N’N’-bis methy-lene-acrylamide

Prior to casting gels, solutions were mixed, degassed, and polymerization initiated by adding ammonium persulfate to 0.5% and TEMED (N,N,N’,N’-Tetramethylenediamine) to 0.1%.

Western blotting

COS cells lysates were separated by SDS-PAGE and transferred to nitrocellulose filters using a BioRad Trans-Blot apparatus (which was submerged in an ice-water bath) for two hours at 60 volts in 25 mM Tris-192 mM glycine (pH 8.3); 20% methanol. The filters were washed briefly in TBS (0.15 M NaCl; 0.02M Tris-Cl [pH 7.5]) containing 0.3% Tween 20, and blocked for non-specific binding in TBS containing 4% powdered skimmed milk. Membranes were then incubated with human anti-RV serum (1:200) in this same solution for two hours, washed with TBS/0.3% Tween 20 and treated with rabbit anti-human IgG conjugated to horse radish peroxidase (Dako) for two hours. Blots were washed as above and developed with 4-chloro-1-naphthol (BioRad). All incubations were performed at room temperature.
Table II.

Mutagenic Oligonucleotides

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>TH2</td>
<td>pGTAGTACAGCATTCTCCGCAGGCTTCT</td>
<td>delete 27 amino acid transmembrane anchor of E1</td>
</tr>
<tr>
<td>TH3</td>
<td>pGTGTCGCTTTCCCCAGATG</td>
<td>destroy E1 glycosylation site 1 by changing Asn 76 to Lys.</td>
</tr>
<tr>
<td>TH4</td>
<td>pTCGTGCAAAAGTCACC</td>
<td>destroy E1 glycosylation site 2 by changing Asn 177 to Lys.</td>
</tr>
<tr>
<td>TH6</td>
<td>pGCCTATGGCAGGAGGAGGCT</td>
<td>insert start codon in front of E1</td>
</tr>
<tr>
<td>TH7</td>
<td>pAGCCCACTACGCCAGG</td>
<td>introduce stop codon in front of E1 transmembrane anchor by changing Glu 441 to TAG</td>
</tr>
<tr>
<td>TH10</td>
<td>pAGCCCGGCATGCGCGGTG</td>
<td>create SphI site in front of E2</td>
</tr>
<tr>
<td>TH13</td>
<td>pCCGGTGTAATCATAATGTA</td>
<td>destroy E1 glycosylation site 3 by changing Asn 209 to Ile</td>
</tr>
<tr>
<td>TH15</td>
<td>pAGTAGTACAAATCTTTCGCCAC</td>
<td>change E1 Cys 470 to stop codon to delete cytoplasmic domain</td>
</tr>
<tr>
<td>G1</td>
<td>pGTCGCTGGCTGTCCGGTG</td>
<td>destroy E2 glycosylation site 1 by changing Asn 53 to Gln</td>
</tr>
<tr>
<td>G2</td>
<td>pCCAGTCGCCAGGTTGTA</td>
<td>destroy E2 glycosylation site 2 by changing Ser 73 to Gly</td>
</tr>
<tr>
<td>G3</td>
<td>pGCTGCTAATGAGGCC</td>
<td>destroy E2 glycosylation site 3 by changing Asn 115 to Ile</td>
</tr>
<tr>
<td>DL1</td>
<td>pCTAGTCTAGACTAG</td>
<td>multipurpose XbaI linker with stop codons in three reading frames</td>
</tr>
</tbody>
</table>

Changes from original RV sequence are underlined.
RESULTS and DISCUSSION

Section I. In vitro expression and targeting of RV structural proteins into microsomes

RV cDNA constructs

All RV cDNA constructs used in this study were derived from pSPT19(C/E2/E1) described by Clarke et al. (1987). The 3.4 kb insert contains coding sequences for the RV structural proteins. As previously stated, this cDNA will be referred to as 24S (Fig. 11). The E2E1 fragment contains the E2 and E1 genes in frame with the translation initiation signal from the capsid gene. Similarly, in the E2 and E1 constructs, the capsid translation start site is utilized (Fig. 11). RV cDNAs were cloned into pSPT19 in front of the SP6 RNA polymerase promoter to generate the plasmids p24S, pE2E1, pE1, and pE2. Note that in Hobman et al., 1988, pE1 was called p3'E2E1. Plasmids were linearized at the HindIII site, and SP6 RNA polymerase was used to generate RV mRNAs which were subsequently translated in a rabbit reticulocyte lysate system.

Transcription/translation

p24S

SP6-derived mRNA from HindIII-linearized p24S was translated in a cell-free system containing $^{35}$S-methionine with or without canine pancreatic microsomes. Products of translation were separated on SDS-PAGE and fluorographed. Translation of p24S mRNA in the absence of microsomes gave rise to a polypeptide of approximately 110,000 daltons in size (Fig. 12a). This protein was interpreted to be the 110 kDa polyprotein precursor NH$_2$-C-E2-E1-COOH (Oker-Blom et al., 1984; Clarke et al., 1987). Inclusion of microsomes in the reaction resulted in efficient processing of p110 to smaller polypeptides, two of which comigrated with E1 and C (Fig. 12a). A 46 kDa band was present in all the lanes, and is known to be some by-product of translation specific to the $^{35}$S-methionine preparation (Amersham product data bulletin). The E2 (39 kDa) synthesized in this system is significantly smaller than virion E2 (Fig. 12a) contrary to
Figure 12. Translation of SP6-derived RV mRNAs in rabbit reticulocyte lysates. RV mRNAs were translated in the presence of $^{35}$S-methionine, and the translation products were separated on 10% Laemmli gels and fluorographed. The relative positions of the RV structural proteins are indicated. (mic) inclusion of canine pancreatic microsomes in translation; (trypsin) translational products treated with trypsin with or without Triton X-100 present (TX-100). The 46 kDa $^{35}$S-methionine-specific bands are indicated by an arrow head. (a) p24S mRNA. The p110 band is marked. (b) pE2E1 mRNA. The 91 kDa E2E1 precursor is shown (p91).
the findings of Clarke et al. (1987). The latter authors reported finding virion size E2 in their translation system, although it now appears that they mistakenly identified the 46 kDa band as RV E2. M33 E2 contains three potential N-linked glycosylation sites (Clarke et al., 1987) and unglycosylated E2 is 30 kDa in size (Oker-Blom et al., 1983). Microsomal vesicles can only add high mannose precursors (MW 2500 approx.), therefore the maximum theoretical size of microsomally processed E2 would be about 38 kDa assuming all three glycosylation sites are used, and correct cleavages occur. From these facts and assumptions, it is likely that the 39 kDa band is E2 which contains high mannose glycans.

Treatment of the translation products with trypsin results in proteolysis of p110 and capsid. No significant change in size of E1 and E2 occurred suggesting that most of the proteins are sequestered in the microsomes. From nucleotide sequence data, the predicted membrane topologies of E2 and E1 (Vidgren et al., 1987) are that of type I membrane proteins with large N-terminal ectoplasmic domains exposed to the virion exterior, or in this case the microsomal lumen. Triton X-100 abolishes the microsome-mediated protection from trypsin. The discrepancy in band intensity can be explained by the fact that E1 has only three methionines, while E2 and C have seven and six respectively.

pE2E1

The RV cDNA insert in pE2E1 encodes an 829 amino acid protein with a predicted size of approximately 91,000. Translation of pE2E1 mRNA in the absence of microsomes results in a polypeptide approximating this size (Fig. 12b). Microsomes efficiently process the 91,000 precursor to E1 and E2 moieties which are translocated into the vesicles as indicated by resistance to trypsin (Fig. 12b). A trypsin resistant band of approximately 100 kDa is also visible in this gel, and may represent an E2-E1 aggregate (arrow Fig. 12b). An E2-E1 aggregate would also be expected in the processing of the p110 protein encoded by p24S. From Fig. 12a such a band is not clearly visible, however after longer exposures of the gel, a band in this region does appear. These results indicate that only E1 and E2 are translocated into microsomes and that an
intact capsid protein is not required for the translocation process.

Expression of E2 in vitro

The structures of pE2 and p^SPE2 cDNAs are illustrated in figure 13a. Plasmid p^SPE2 was generated by creating a new Sph1 site upstream from E2 using the mutagenic oligonucleotide TH10 (P-AGCCCGGCATGCGCGGTA-OH) to change the bases at positions 215 and 216 from GC to AT, and the 163 nt fragment encoding the putative signal sequence of E2 was deleted by digesting with Sph1 and religation (Fig. 13a,b). To restore the correct reading frame, the construct was linearized at the NcoI site in the C region by partial digestion, and the cohesive ends filled in using Klenow fragment. Dideoxy nucleotide sequencing was employed to verify the specific base changes and subsequent deletions. The resulting plasmid p^SPE2, is missing DNA coding for 54 residues of the carboxyl terminus of capsid in pE2. In figure 13b, the N-terminal coding regions of pE2 and p^SPE2 cDNAs are compared. Fig. 13c is a hydropathy plot of the N-terminal region of pE2 and illustrates that the putative E2 signal peptide domain is the only hydrophobic region in the 54 amino acids preceding the E2 amino-terminus.

pE2

Translation of pE2 RNA in the absence of microsomal vesicles produced two proteins, 36 and 34 kDa in size (Fig. 14a). The open reading frame in pE2 specifies a polypeptide of 338 amino acids assuming the first inframe methionine is used, suggesting that the smaller protein is the result of an internal initiation, premature termination, or limited proteolysis. In the presence of microsomes, a polypeptide of 39,000 daltons was synthesized (Fig. 14a). On some gels, a 37,000 dalton protein was observed (Fig. 15) suggesting that the 34 kDa E2 protein is processed by microsomes as well. The increase in molecular weight from 36 to 39 kDa (or 34 to 37 kDa) is the result of high-mannose glycan addition.

Translation products from pE2 mRNA are only protected from trypsin digestion when microsomes are included. Membrane topology of E2 predicts that the C-terminal 18 amino acids protrude from the microsomes into the cytoplasm (Clarke et al., 1987; Vidgren et al., 1987; Frey
Figure 13. Physical maps of E2 recombinants. (a) Schematic diagram representing RV cDNA inserts in pE2 and pΔSPE2. SP6 promoters are marked with arrow heads, and the putative signal peptide (sp) and transmembrane domain (tm) are underlined. Dashed lines represent pSPT19 sequences, and regions of capsid (C) and E2 genes are indicated. Restriction sites; E=EcoRI, N=NotI, S=SphI, H=HindIII, and P=PstI. (b) 5' sequences of E2 and ΔSPE2 inserts. Translation start sites are marked with arrows, capsid (C) and E2 coding regions are indicated. The putative signal peptide is underlined. (c) Hydropathy plot of N-terminal region of pE2 insert. The N-terminal region of pE2 insert was subjected to hydropathy analysis (Kyte and Doolittle, 1982) using the SEQNCE DNA software analysis program (Delaney Software Ltd, Vancouver, B.C.). Hydrophobic residues are assigned positive values and lie above the zero line, while hydrophilic amino acids have a negative value and lie below the line. The initiator methionine for capsid is residue #6, and C, E2, and signal peptide (sp) regions are indicated.
\[ \text{B} \]

E2

\[
\begin{align*}
\text{CCC CAG GGT GCC CGA ATG GCT TCC ACT ACC CCC ACC ATG GGA} \\
\text{Pro Gin Gly Ala Arg Met Ala Ser Thr Thr Pro Ile Thr Met Gly}
\end{align*}
\]

Sph1

\[
\begin{align*}
\text{GCT TGC ATG CCA GTG CGC GCC GAC GCC ATG AGC GCC CCC CTT} \\
\text{Ala Cys Met Pro Val Arg Gly Leu Asp Gly Asp Thr Ala Pro Leu}
\end{align*}
\]

\[
\begin{align*}
\text{105} & \quad \text{120} & \quad \text{135} \\
\text{CCC CCT CAC ACC ACC GAG CGC ATT ATG GAG ACC CGC TTC GGC CTC} \\
\text{Pro Pro His Thr Thr Asp Arg Asp Gln Ile Leu Thr Arg Ser Ala Arg His}
\end{align*}
\]

\[
\begin{align*}
\text{150} & \quad \text{165} & \quad \text{180} \\
\text{CCC CCT CAC ACC ACC GAG CGC ATT ATG GAG ACC CGC TTC GGC CTC} \\
\text{Pro Pro His Thr Thr Asp Arg Asp Gln Ile Leu Thr Arg Ser Ala Arg His}
\end{align*}
\]

\[
\begin{align*}
\text{195} & \quad \text{210} \\
\text{CTC TTG CTC GCC GCG GTC GCC GTT GGC ACC GCG CGC GCC GGG} \\
\text{Leu Leu Leu Ala Ala Val Ala Val Gly Thr Ala Arg Ala Gly Leu}
\end{align*}
\]

\[ \text{△spE2} \]

\[
\begin{align*}
\text{CCC CAG GGT GCC CGA ATG GCT TCC ACT ACC CCC ACC ATG CAT} \\
\text{Pro Gin Gly Ala Arg Met Ala Ser Thr Thr Pro Ile Thr Met His}
\end{align*}
\]

E2

\[
\begin{align*}
\text{60} & \quad \text{75} & \quad \text{90} \\
\text{GGG AGC TTG CAT GCC GCG GTC GCC GTT GGC ACC GCG CGC GCC GGG} \\
\text{Gly Ser Leu His Ala Gly Leu Gin Pro Arg Ala Asp Met Ala Ala}
\end{align*}
\]

\[ \text{C} \]

-63-
Figure 14. *In vitro* translation of pE2 and pΔSPE2 mRNA. (mic) with microsomes; (trypsin) translational products digested with trypsin with or without Triton X-100 (TX-100); (RV) radiolabeled RV marker (a) pE2 mRNA  (b) pΔSPE2 mRNA.
and Marr, 1988). Cleavage of this cytoplasmic domain by a trypsin could be the origin of the lower molecular weight E2 band seen in the microsome(+)/trypsin(+) lane (Fig. 14a). Digestion with pronase results in complete shift of the translocated E2 to the lower molecular weight form (not shown).

**p^SPE2**

The open-reading frame in p^SPE2 specifies a protein of 285 amino acids. Translation of mRNA from this plasmid resulted in synthesis of a protein 30-31 kDa in size in addition to the other protein which was approximately 2 kDa smaller (Fig. 14b). These translation products did not appear to be translocated into microsomes as determined by their sensitivity to trypsin (Fig. 14b), and suggest that a signal peptide function for E2 is located in the C-terminal region of capsid protein.

**E2 acquires N-linked glycans and is proteolytically processed by microsomes**

Translation products of pE2 mRNA were immunoprecipitated with human anti-RV serum, digested with endo H, separated on SDS-PAGE and fluorographed (Fig. 15). Endo H had no effect on the mobility of E2 polypeptides synthesized in the absence of membranes. In contrast, endoglycosidase treatment of microsome-processed E2 decreased the molecular weight by 8000 daltons (Fig. 15). The addition of 8000 daltons of carbohydrate by microsomes suggest that all three of the potential N-linked glycosylation sites are used (discussed in Results section III).

Deglycosylated E2 was determined to be 6000 daltons smaller than E2 synthesized in the absence of membranes. The predicted signal peptidase cleavage site is located 67 amino acids from the initiator methionine in pE2 (Fig. 13b). Presumably signal peptidase is responsible for proteolytically removing 6000 daltons of polypeptide during translocation of the pE2 protein. Assuming signal peptidase cleaves the pE2 polypeptide only once (at the E2 N-terminus), the likelihood that the lower molecular weight E2 band originates from internal translation initiation seems remote (Fig. 15). If this were the case, only one band would be expected after insertion into microsomes. Signal peptidase should cleave at a point 55 amino acids C-terminal to
Figure 15. Proteolytic processing of pE2 protein by microsomes. $^{35}$S-labeled translation products synthesized in the presence or absence of microsomes (mic) were immunoprecipitated with human anti-RV serum. Samples were then split into two aliquots and incubated at 37°C with or without endo H, separated on SDS-PAGE and fluorographed. Radiolabeled RV and protein molecular weight standards (kDa) are included.
the third methionine in pE2 (13b). Therefore the smaller E2 band probably results from premature termination or some kind of specific proteolysis in the carboxyl-end of the protein, possibly the arginine-rich region between the E2 transmembrane domain and putative E1 signal peptide (Figs. 11 and 16b).

Expression of E1

The physical structure of the pE1 and p^SPE1 cDNAs are diagrammed in Fig. 16a. The open reading frame of pE1 encodes a protein 558 amino acids long and includes the N-terminal eight residues from capsid, 69 amino acids from the E2 C-terminus, and 481 amino acids of E1. Included in the 69 amino acids from E2 are the putative transmembrane domain of E2 and a signal peptide of E1 (Fig. 16a,b,c) (Frey et al., 1986; Vidgren et al., 1987). Located between the E2 transmembrane anchor and E1 signal sequence is an arginine-rich (5/7) group of amino acids (Fig. 16b). This region has been postulated to be a site for a trypsin-like processing enzyme (Vidgren et al., 1987).

To create p^SPE1, a 1550 nt PstI fragment from p24S was excised and subcloned into M13mp19. The oligonucleotide TH6 (P-GCCTATGGCATGGAGGAGGCT-OH) was used to create a start codon in front of the E1 coding region (start codon is underlined). The insert was then subcloned back into pSPT19 downstream from the phage SP6 promoter (Fig. 16a,b). This construct lacks the putative signal peptide domain of E1 (Fig. 16a,b). Note that in a previous publication (Hobman et al., 1987), pE1 was called p3'E2E1, and p^SPE1 was named pE1.

pE1

Translation of pE1 mRNA produced a protein which was completely sensitive to trypsin unless microsomes were included (Fig. 17a, panel ii). Translocation assay by trypsin digestion suggested that most of the pE1 protein was sequestered within the microsomes, and is in concordance with the predicted membrane topology of E1 which postulates the 13 C-terminal amino acids protrude into the cytoplasm. A very slight decrease in molecular weight is evident after protease digestion, and may be the result of proteolysis of the C-terminal region of E1 (Fig. 17a, panel ii).
The E1 protein encoded by p^SPE1 has a predicted size of about 51 kDa which corresponds well with the polypeptide produced by cell-free translation (Fig. 17a, panel i). Addition of microsomes did not affect the size of this protein, nor did they afford any protection from proteolysis by trypsin indicating that it was not translocated into the vesicles. These data suggest that translocation of RV E1 into microsomes does not require prior attachment of E2 to the membranes, and is mediated by a signal peptide located with the 69 C-terminal amino acids of E2.

**E1 acquires N-linked glycans and is proteolytically processed by microsomes.**

As was the case with E2, inclusion of microsomes while translating pE1 mRNA resulted in a slight increase in the size of E1 (Fig. 17a,b). Digestion of translocated E1 protein with endo H decreased the molecular weight by 5-6000 daltons by release of high mannose sugars (Fig. 17b, arrowhead). Endo H had no effect on the mobility of untranslocated E1 indicating that as well as becoming glycosylated by microsomes, E1 was also proteolytically processed by the vesicles. The microsomes used in this experiment had obviously lost a considerable amount of their activity judging by the fact that less than 50% of E1 appeared to be processed by the vesicles (Fig. 17b).

**Discussion**

Processing of the RV p110 polyprotein has been studied in vitro by using a cell-free translation/translocation assay system. As previously reported, no processing of p110 was observed when 24S mRNA was translated in the absence of pancreatic microsomes implying that RV capsid protein does not possess an autoproteolytic activity like that found in alphavirus capsid proteins (Oker-Blom et al., 1984; Clarke et al., 1987; Simmons and Strauss, 1974; Garoff et al., 1980a,b). Efficient cleavage of the polyprotein was observed by inclusion of microsomes in the translation reaction. Resulting products included proteins which comigrated with RV virion
Figure 16. Physical maps of RV E1 cDNAs. (a) Schematic of pEl and p^SPEl. SP6 promoters are indicated by black boxes, and the putative signal peptide and transmembrane domain of E1 are marked by an open box and hatched box respectively. Representative portions of C,E2 and E1 coding regions are denoted above the constructs. P=PstI; E=EcoRI; N=NcoI; H=HindIII. (b) Sequences of E1 and ^SPE1 5' coding regions. Translation start sites are designated by arrow heads, and the putative transmembrane domain of E2 and signal peptide of E1 are underlined with broken and solid lines respectively. The N-terminal coding region (GAGGAG) of E1 is marked. (c) Hydropathy plot of the N-terminal region of pEl insert. Hydrophobic residues are assigned positive values and lie above the zero line, while hydrophilic amino acids have a negative value. The initiator methionine is residue #6, and C (6-14), E2 (15-82), and E1 (83-) regions are indicated. Four hydrophobic domains are present, and domains III and IV are believed to be the transmembrane domain of E2 and E1 signal peptide respectively. The arginine rich region is located between domains III and IV.
\[ \begin{align*}
\text{E1} & : \\
\text{E2} & :
\end{align*} \]
Figure 17. (a) In vitro translation of pE1 and pASPE1 mRNA. (mic) microsomes; (tryp) trypsin digestion; (TX-100) Triton X-100. Relative positions of molecular weight standards (kDa) are marked. (i) pASPE1 mRNA (ii) pE1 mRNA. (b) Proteolytic processing of pE1 protein by microsomes. S-labeled translation products were immunoprecipitated with human anti-Rv serum then split into two fractions, and incubated overnight at 37°C with or without endo H, separated on SDS-PAGE and fluorographed. 14C-labeled protein standards (kDa) were included. Arrowhead shows the deglycosylated E1.
E1 and C, as well as a 39 kDa E2 protein. The reason E2 from the *in vitro* translation reaction is smaller than virion E2, is that microsomes only add high mannose glycans to translocated proteins. In contrast, E2 from RV virions is very heterogeneously glycosylated and ranges in size from 42-47 kDa (Oker-blom et al., 1984).

Protease protection assay results indicated that only E2 and E1 were translocated into microsomes. Proteolysis did not dramatically alter the size of translocated E2 and E1 which suggested that most of the proteins were sequestered within the vesicles. Translation of pE2E1 mRNA resulted in virion size E1 and a 39 kDa E2. These proteins were also translocated into the microsomes, and together these data indicate that capsid protein is not required for the processing of the RV polyprotein precursor including correct targeting of E1 and E2 to the ER membrane.

The observation that translocation of RV E1 and E2 into microsomes could occur independently implied that at least two signal sequences were required for these events. Subsequently, the signal peptide domains of E2 was shown to reside within the 58 amino acids of the C-terminus of capsid protein. Sequence analysis and hydropathy evaluation of this region suggested that the signal sequence is composed of the 23 amino acids immediately preceding the E2 N-terminus (Clarke et al., 1987; Vidgren et al., 1987; Frey and Marr, 1988; Hobman and Gillam, 1989). Although signal peptidase likely catalyzes the cleavage of RV E2 from the capsid, it remains to be determined if the E2 signal peptide is removed from the carboxy-terminus of capsid.

Similarly, transfer of E1 across the ER membrane is mediated by a signal peptide domain located within the C-terminus of E2 glycoprotein (Hobman et al., 1988). Nucleotide sequence analysis has led other groups (Frey et al., 1986; Vidgren et al., 1987) to predict that the hydrophobic domain immediately preceding the known E1 terminus functions as the E1 signal (Kalkinnen et al., 1984), although this remains to be determined. In addition it is not known if the E1 signal domain remains attached to the C-terminus of E2. Vidgren et al (1987) have postulated that the E1 signal may be released from E2 by trypsin-like processing in the
arginine-rich region between E2 and E1.

In conclusion, targeting and translocation of RV E2 and E1 across the ER membrane is controlled by at least two independently functioning signal peptides. At least two microsome-dependent proteolytic cleavages of p110 occur during processing, as well as addition of high mannose glycans to E2 and E1. Capsid protein does not cross the ER membrane and does not appear to be involved in the processing of p110.
Section II. **In vivo** expression and post-ER processing of RV structural proteins

In the previous section, the targeting and insertion of RV glycoproteins into microsomal vesicles was discussed. The results implied that the targeting of E2 and E1 to the ER could occur as independent events accomplished by independently functioning signal peptides. To study the transport of RV glycoproteins from the ER to the Golgi complex, the RV cDNAs were subcloned into the eukaryotic expression vector pCMV5 (Fig. 10 Andersson et al., 1989). The recombinant plasmids were introduced into COS cells and the independent and coordinate expression of the RV proteins was examined.

**Kinetics of processing**

**pCMV5-24S**

Recombinant plasmid pCMV5-24S contains the RV 24S cDNA under the control of the human cytomegalovirus major immediate early gene promoter. COS cells were transfected with pCMV5-24S using DEAE-dextran and cellular proteins were labeled with $^{35}$S-methionine and RV-specific proteins immunoprecipitated with human anti-RV serum. Fig. 18 shows the results of a pulse-chase experiment. After a 30 minute pulse with labeled methionine, the RV-specific proteins comigrating with E1 and C were evident, as well as a protein slightly larger than C. The latter protein is 39 kDa and was determined to be RV E2 with high mannose sugars (section I). Present in lesser amounts are high molecular weight proteins of >100 kDa, an 85 kDa doublet, and a 75 kDa protein (Fig. 18). Digestion with endo H decreased the size of E1 and E2 bands to 51 and 32 kDa respectively, while C, >100 kDa, and the 75 kDa proteins were unaffected. The 85 kDa proteins are also endo H sensitive. The 75 kDa protein could be immunoprecipitated from vector (pCMV5) transfected cells (not shown) using a monoclonal antibody to murine BiP (Bole et al., 1986), however the origin of the other high molecular weight proteins is uncertain.

After a 30 minute chase with excess methionine the molecular weight of E1 and C remain the same, however some of the 39 kDa E2 is converted into a 42 kDa form which contains complex sugars judging by the 37 kDa endo H digestion product (Fig. 18). As the chase period increased, so did the 39 to 42 kDa E2 conversion in addition to the appearance of a 53 kDa endo...
Figure 18. Expression of pCMV5-24S in COS cells. Transfected COS cells were pulsed for 30 minutes with 100 μCi 35S-methionine, then chased for the indicated time periods with excess unlabeled methionine. RV proteins were recovered by immunoprecipitation with human anti-RV serum and incubated for 16 hours at 37°C with or without endo H. RV and 14C-labeled protein standards (kDa) are included. ER and Golgi (G)-processed forms of E2 are indicated, and the 53 kDa endo H-resistant E1 is marked by an arrow head. Heterogeneous processing of E2 is indicated (%). Transfections with pCMV5 constructs were performed in 35 mm dishes.
H-resistant E1 species (arrow). In the ER and microsomes, high mannose (endo H-sensitive) precursor glycans are attached to nascently translocated proteins. Acquisition of endo H-resistance occurs after processing of glycans by enzymes in the medial Golgi compartments (Dunphy and Rothman, 1985). Using endo H-resistance as evidence of Golgi-specific processing, it can be concluded that the 39 kDa and 42 kDa isoforms of E2 represent glycoproteins with ER and Golgi-specific sugars respectively, hence the reference as ER and G forms of E2 (Fig. 18).

Transport from the ER to Golgi is a rate limiting step in the passage of proteins through the exocytic pathway and varies widely with different proteins, with ER resident half times ranging from several minutes to several hours (Lodish, 1988). In this study, the rate of transport of E2 and El from the ER to the medial Golgi occurs relatively slowly. After three hours only about 50% of E2 was in the 42 kDa form, while less than 10% of E1 contained endo H-resistant sugars (Fig. 18). The heterogeneous processing of E2 sugars from 42-46 kDa could be observed after longer chase periods (Fig. 18 and 19).

pCMV5-E2El

Fig. 19 shows the results of a pulse-chase experiment using pCMV5-E2El transfected cells. Surprisingly the processing of E2 and El glycans occurred more rapidly in the absence of capsid. After a 60 minute chase period the ER and G forms of E2 are present in equal amounts, and endo H-resistant El is formed more quickly without capsid present (compare Fig. 18 and 19). Heterogeneous processing of E2 beyond the 42 kDa size appears to take place more rapidly as well.

The endo H-sensitive 85 kDa proteins were also present in pCMV5-E2El transfected cells, and are therefore not capsid-specific. Whether or not they represent RV glycoprotein aggregates or COS cell proteins induced by the expression of RV structural proteins remains to be determined. Also present were the >100 kDa protein, and BiP which did not seem to change in intensity throughout the chase period.

E2 glycoprotein

The one kb cDNA containing the entire E2 gene and flanking sequences (Fig. 11) was
Figure 19. Expression of pCMV5-E2E1 in COS cells. Transfected COS cells were pulsed for 30 minutes with 100 μCi $^{35}$S-methionine, then chased for the indicated time periods with excess unlabeled methionine. RV proteins were recovered by immunoprecipitation with human anti-RV serum and incubated for 16 hours at 37°C with or without endo H. RV and $^{14}$C-labeled protein standards (kDa) are included. ER and Golgi (G)-processed forms of E2 are indicated, and the 53 kDa endo H-resistant E1 is marked by an arrow head. Heterogeneous processing of E2 is indicated (*).
subcloned between the EcoRI and HindIII sites of pCMV5. This construct did not contain a stop codon for the E2 gene and was named pCMV5-E2ns. Translation was predicted to proceed through the 3' end of the polylinker and into the human growth hormone region (Fig. 10). The human growth hormone fragment provides the transcription termination and poly-adenylation signal for genes cloned into this vector. Depending upon exactly where transcription stops, up to 40 non-RV amino acids could be added onto the end of E2 protein by expression of pCMV5-E2ns in COS cells. Transfection of pCMV5-E2ns into COS cells directed the synthesis of a 38 kDa protein which was rapidly degraded following chase with medium containing excess methionine (Fig. 20a).

pCMV5-E2

To determine if the instability of E2 in COS cells was due to the lack of a stop codon, an XbaI linker containing stop codons in all three reading frames (Table I) was engineered into the E2 gene at the PstI site after the transmembrane region (Fig. 14). The resulting plasmid was named pCMV5-E2. Results from a pulse-chase experiment using this plasmid are shown in Fig. 20b. Immediately after a 30 minute pulse, 38 and 43 kDa E2 (Fig. 20b arrow head) proteins were seen as well as the endo H-resistant 75 kDa protein (BiP). Removal of high mannose sugars with endo H reduced the molecular weights of the E2 glycoproteins to 31 and 36 kDa. Chase with unlabeled methionine resulted in degradation of the 43 kDa protein, and heterogeneous processing of the glycans on 38 kDa E2 ( ) presumably by Golgi-specific enzymes. Digestion of the heterogeneously processed E2 with endo H gave rise to a 34 kDa cleavage product (Fig. 20b). The nature of the short-lived 43 kDa glycoprotein is uncertain but may represent a subpopulation of E2 that is rapidly transported through the Golgi to the cell surface, then targeted to the lysosomes where it was degraded (Rizzolo, 1989).

pCMV5-E1

E1 from cells expressing pCMV5-E1 was similar in size to RV virion E1, and initially contains only high mannose sugars (Fig. 21a). The endo H cleavage product was 51 kDa which corresponds well with the predicted size of 51.5 kDa for unglycosylated E1 (Vidgren et al., 1987).
Figure 20. Expression of E2 in COS cells. (a) pCMV5-E2ns. Transfected COS cells were pulsed for 30 minutes with $^{35}$S-methionine and chased with excess unlabeled methionine for 0 and 90 minutes before harvesting. RV antigens were immunoprecipitated with human anti-RV serum and separated by SDS-PAGE. Note the relatively rapid degradation of the E2ns protein after the chase period. (b) pCMV5-E2. COS cells were radiolabeled as above and the RV antigens isolated by immunoprecipitation. Some samples were treated with endo H as indicated. The short-lived 43 kDa E2 is indicated with an arrow head, and heterogeneous processing of E2 by Golgi-specific enzymes is shown (*).
Processing of E1 glycans is evident after chase with 2 mM methionine by the appearance of a 53 kDa endo H-resistant protein (Fig. 21a arrow). These results suggest that E1 is able to exit the ER and enter the Golgi compartments. Expression of ^SPE1 and ^SPE2 in COS cells did not result in any detectable RV antigen presumably due to rapid degradation in the cytoplasm (not shown). Newly synthesized membrane proteins that fail to undergo translocation into the ER have been shown to be extremely unstable in the cytoplasm (Gething and Sambrook, 1982; Jabbar and Nayak, 1987; Sekikiwa and Lai, 1983).

To determine if deletion of the putative membrane anchor region of E1 would result in secretion of E1 into the extracellular space, two C-terminal deletion mutants of E1 were constructed by oligonucleotide-directed mutagenesis (Table II) and subcloned into pSVL. The resulting plasmids pSVL-E1^TM and -E1TMCT- contain identical RV cDNAs as in pCMV5-E1 and pE1 except that the 27 amino acid transmembrane domain of E1, and the transmembrane domain and cytoplasmic tail of E1 are deleted respectively in E1^TM and E1TMCT- respectively. After a two hour post-label chase, ^TM and TMCT- E1 proteins were not detectable in the medium of transfected COS cells (Fig. 21b). Even after three hours of chase, no RV antigen was found in the medium (not shown). The reasons for lack of secretion may include i) defective folding abrogating transport to the cell surface; ii) presence of a Golgi retention signal in the ectoplasmic domain of E1; iii) the 27 amino acid region deleted is not the transmembrane anchor of E1; iv) slow transport kinetics resulting in very long $T_{1/2}$ times for secretion.

**Steady state analysis of RV structural proteins in transfected cells**

$^3$H-palmitate labeling

Previous investigators have demonstrated the presence of covalently bound palmitic acid to a wide variety of viral glycoproteins (for review, Grand, 1989) including RV E1 and E2 (Waxham and Wolinsky, 1985b). I decided to try and label RV glycoproteins with $^3$H-palmitate in transfected cells for two reasons: 1) to see if acylation of RV does occur in the absence of nonstructural proteins and virus budding, and if so does E2 require E1 for this process and vice
Figure 21. a) Expression of pCMV5-E1 in COS cells. Transfected cells were labeled with 100 μCi 35S-methionine and chased with excess unlabeled methionine as described above. RV-specific proteins were recovered with human anti-RV serum. The 53 kDa endo H-resistant band is indicated by an arrow head. b) Expression of pSVL-E1ΔTM and pSVL-E1TMCT- in COS cells. (p) 2 hour pulse with 100 μCi 35S-methionine; (pc) cells were chased for 2 hours with excess methionine following pulse; (L) cell lysate; (M) medium (radioimmunoprecipitate from 1.0 ml).
versa and 2) to determine which isoforms of E2 are labeled, and examine the relative levels of these species in the transfected COS cells.

E1 and E2 from transfected cells both incorporate palmitate, although E1 seems to contain higher molar amounts of the fatty acid (Fig. 22). The sites of acylation on E1 and E2 are unknown, but are likely one or more of the cysteine residues located in the C-terminal regions of the proteins (Rose et al., 1984; Grand, 1989). Currently the exact subcellular location of palmitate addition is not known, although a number of investigators believe that acylation of proteins occurs in the ER immediately prior to their exit (Schmidt et al., 1985; Rose and Doms, 1988), while Bonatti et al. (1989) have evidence that this process occurs in a post-ER/pre-cis Golgi compartment. This experiment clearly shows that E1 and E2 are capable of reaching the site of palmitoylation independently in transfected COS cells, and according to the findings of Bonatti et al., suggests that RV E1 is indeed capable of exiting the ER without complexing to E2. In addition, three species of E2 (39, 42, 45 kDa) can be detected in pCMV5-24S and pCMV5-E2E1 transfected cells, and likely reflect differences in glycosylation (Fig. 22). The highest levels of 39 kDa E2 are found in pCMV5-E2 containing cells, followed by pCMV5-24S, then pCMV5-E2E1 indicating that glycan processing and transport of E2 occurs most rapidly with E1 present, and capsid absent.

Western blot analysis

Immunoblotting was also used to examine the steady state levels of the E2 isoforms in transfected COS cells lysates. By this technique, only two species of E2 are detectable, as well as one species of E1 and capsid respectively (Fig. 23). Again the levels of 42 kDa E2 are highest in pCMV5-E2E1 transfected cells, followed by pCMV5-24S and pCMV5-E2. These results are consistent with the kinetic analysis and palmitate labeling data, and suggest that E1 increases the rate of ER to Golgi transport of E2, while capsid has a negative effect.

Formation of disulfide bonds

Previous investigators have suggested that disulfide-bonded E1-E1, E2-E1, and C-C
Figure 22. Incorporation of palmitate into RV glycoproteins. Transfected cells were labeled with 500 μCi of [9,10^3H]-palmitic acid in 1.0 ml of medium for 12 hours, and the RV proteins were recovered by immunoprecipitation. ^35S-labeled RV and protein markers were included for reference.
Figure 23. Western blot analysis of transfected COS cell lysates. Lysates from transfected cells were separated by SDS-PAGE, transferred to nitrocellulose, and probed with human anti-RV serum. Golgi (G) and ER-processed forms of E2 are marked. RV structural proteins were included as marker.
dimers are found in RV virions (Waxham and Wolinsky, 1983, 1985b; Dorsett et al., 1985; Green and Dorsett 1986). To determine if this phenomenon occurs in the absence of virus production and nonstructural proteins, RV proteins from transfected COS cells were examined by SDS-PAGE under nonreducing conditions. Proteins from $^{35}$S-labeled COS cells were lysed at two time points, after a 30 minute pulse label, and following a two hour chase with excess unlabeled methionine. Cells were lysed in the presence of 10 mM iodoacetamide to ensure that oxidation of cysteines did not occur during harvesting.

Following a 30 minute pulse, no distinct RV-protein oligomers were apparent when immunoprecipitates were separated under nonreducing conditions (Fig. 24 lanes I,K,M,O). In contrast E1 and E2 migrated at a faster rate under these conditions suggesting the presence of intramolecular disulfide bonds. The molecular weight of capsid was not significantly altered when subjected to nonreducing (exclusion of 2-mercaptoethanol when in protein sample buffer) conditions (compare lanes A,B to I,J). These results conflicted with previous reports which seemed to indicate that capsid is found as a disulfide-linked dimer with an apparent molecular weight of 78,000 (Waxham and Wolinsky, 1983) or 68,000 daltons (Dorsett et al., 1985) in unreduced preparations. Experiments performed by the above authors were carried out using RV virion proteins, whereas in this case, lysates of transfected COS cells were used. The cytoplasm of eukaryotic cells is very reducing due to the presence of glutathione and thioredoxin, and it is therefore unlikely that oxidized capsid dimers would be found in this environment. Consequently if the capsid dimers reported by the above authors require virus budding for their formation and they would not be detected in transfected or infected cells.

Lanes J,L,N, and P (Fig. 24) contain unreduced immunoprecipitates from transfected COS cells pulse-labeled and chased for two hours prior to lysis to determine if any subsequent disulfide bonding occurred. Once again, no oligomeric forms of the structural proteins could be observed even though after allowing adequate time for some of E2 and E1 to exit the ER. E2 was predominantly of the 42 kDa isoform after this time period in pCMV5-24S and pCMV5-E2E1 transfected cells, but appeared to contain only intramolecular disulfide-bridges (compare lanes
Figure 24. SDS-PAGE of $^{35}$S-labeled RV proteins from transfected COS cell lysates under non-reducing and reducing conditions. A,B,I,J-pCMV5-24S; C,D,K,L-pCMV5-E2E1; E,F,M,N-pCMV5-E2; G,H,O,P-pCMV5-E1 transfected cells. Lanes A,C,E,G,I,K,M,O, cells were pulsed for 30 minutes with radioactive methionine and harvested. In lanes B,D,F,H,J,L,N,P, cells were chased for two hours with excess methionine following the pulse. Putative RV glycoprotein oligomers are marked with an arrow.
J,L to B,D). Some high molecular weight RV-specific bands are present in the reduced samples (arrow lanes A,B,C,D) as well as in the virion marker lane (RV) and are therefore probably not the result of cysteine oxidation. Formation of disulfide-bonds in membrane glycoproteins is thought to be catalyzed by the ER enzyme protein disulfide isomerase (Freedman, 1984). In transfected COS cells, formation of intermolecular disulfide bonds does not appear to be involved in the maturation of RV E1 and E2 glycoproteins. Whether or not this process occurs in RV virions remains unsubstantiated.

Cell surface immunoprecipitations

To determine which of the RV structural proteins are transported to the plasma membrane, transfected COS cells were processed for cell surface RV antigen as described above in material and methods. Fig. 25 shows that in COS cells expressing pCMV5-24S, only E1 and E2 are accessible to antibody in unpermeablized cells. In addition, the appearance of E1 and E2 at the cell surface was time dependent. No capsid protein was detected at the plasma membrane even after a 180 minute chase period. Note that predominant form of E2 at the plasma membrane is the 42 kDa isoform. Some 39 kDa E2 was also present suggesting that a portion of high mannose containing E2 is transported to the plasmalemma. Movement of the RV antigen to the cell surface appeared relatively slow with the vast majority of E1 and E2 still found in the intracellular fraction after the three hour chase period (Fig. 25)

Immunofluorescence

Transfected and infected COS cells were examined using indirect immunofluorescence to determine the subcellular distribution of C, E2, and E1 when they were expressed coordinately or individually.

pCMV5-24S

Initially human anti-RV sera was used to visualize the RV structural antigens in cells transfected with pCMV5-24S. In some cases, cells were first treated with Rhodamine-conjugated
Figure 25. Cell surface immunoprecipitation of RV proteins from pCMV5-24S transfected COS cells. Cells were pulsed with 100 μCi $^{35}$S-methionine for 30 minutes followed by various chase periods with excess methionine. Cell surface (S) and intracellular (i) RV antigens were isolated by immunoprecipitation with human anti-RV serum, followed by SDS-PAGE and fluorography. Arrows show the cell surface E1 and E2.
WGA to identify Golgi and post-Golgi structures (Virtanen et al., 1980). Cells were permeabilized with Nonidet NP40, treated with primary anti-RV serum followed by secondary Fluorescein-conjugated or Rhodamine-conjugated antibody. Fig. (26a) shows pCMV5-24S transfected cells treated with polyclonal anti-RV sera. A very prominent juxtanuclear structure presumed to be the Golgi complex was evident (Fig. 26a,b) as well as some vesicular structures in the cytoplasm. No staining of the plasma membrane was observed in these detergent-treated cells.

Next the distribution of the individual antigens were examined using monoclonal antibodies. E1 appeared to localize primarily to a Golgi-like region (Fig. 26c) with some vesicular staining evident as well. In addition to staining the juxtanuclear area, anti-E2 antibodies displayed a reticular pattern which probably corresponded to the ER (Fig. 26d). Capsid protein was found distributed throughout the cytoplasm, and surprisingly, seemed to be concentrated in the Golgi region (Fig. 26e). To determine if the juxtanuclear organization of capsid was due to an interaction with the spike glycoproteins, COS cells were transfected with pSVL-RV2 (Clarke et al., 1988) which contains the entire RV capsid gene and the S' one third of E2 under the transcriptional control of the SV40 late promoter. Capsid protein in these cells displayed a diffuse cytoplasmic staining devoid of any antigen in the Golgi region (Fig. 26f). The latter results were consistent with the finding of Kondor-Koch et al (1982) who reported that in cells expressing SFV structural proteins, anti-SFV capsid antibodies displayed a diffuse cytoplasmic pattern whether glycoproteins E2 and E1 were present or not. Normally, SFV-infected cells exhibit a homogeneous punctate cytoplasmic fluorescence when stained for capsid, however when treated with monensin, capsid protein was found in the juxtanuclear area due to a switch in virus budding from the plasmalemma to the Golgi membranes (Kuismanen et al., 1982). RV glycoproteins had a profound effect on the intracellular distribution of capsid protein. Moreover, E1 seemed to manifest a juxtanuclear distribution as opposed to a reticular localization in cells expressing the SFV structural proteins (Kondor-Koch et al., 1982).

Visualization of RV antigens on the cell surface was accomplished by processing fixed,
unpermeablized cells for indirect immunofluorescence. Using polyclonal sera against RV, copious amounts of RV antigen was detected on the surface of cells expressing pCMV5-24S (Fig. 26g). This antigen was due to both E1 and E2 glycoproteins, but not C (Fig. h,i,j). Site-directed mutagenesis was employed to place a stop codon between the E1 transmembrane domain and the 13 amino acid cytoplasmic domain such that the latter domain was deleted (Table II). Expression of this construct pCMV5-24SE1CT- in COS cells resulted in correct processing of E1 and E2 with similar if not faster kinetics that of pCMV5-24S, and transport of E1 and E2 to the cell surface (not shown). The distribution of capsid protein in these cells was somewhat variable with some cells exhibiting a diffuse cytoplasmic distribution while others showed a perinuclear or juxtanuclear localization (Fig. 26k). From these results it was not possible to determine if the cytoplasmic tail of E1 was solely responsible for the juxtanuclear organization of capsid in pCMV5-24S transfected cells. COS cells transfected with pCMV5-CE2 (constructed by H.McDonald, this lab) did not show any colocalization of capsid with E2 in the Golgi region (not shown) suggesting that E2 alone can not bind capsid protein.

pCMV5-E2E1

The distribution of antigen in COS cells expressing pCMV5-E2E1 was similar to those transfected with pCMV5-24S when analyzed using polyclonal serum. A strong juxtanuclear signal which colocalized with WGA binding moieties was present in addition to the weaker reticular staining (Fig. 27a,b). E1 was found primarily in the Golgi region while E2 was distributed in the cytoplasmic reticular area, as well as in the juxtanuclear region (Fig. 27c,d). No signal was obtained by using the anti-capsid monoclonal (Fig. 27e,f). Cell surface antigens in these cells was due to the presence of both E1 and E2 (Fig. 27g,h,i).

pCMV5-E2

Along with the strong juxtanuclear signal, COS cells expressing pCMV5-E2 exhibited bright staining of the nuclear envelope, and to a lesser extent the reticular area (Fig. 28a). The juxtanuclear antigen costained with WGA indicating transport of E2 into the Golgi complex (Fig. 28b,c,d).
Figure 26. Indirect immunofluorescence of pCMV5-24S transfected COS cells. Frames a-f, cells were permeabilized with detergent to allow detection of intracellular antigens. (a) human anti-RV; (b) same cells, TRITC-WGA; (c) anti-E1; (d) anti-E2; (e) anti-C; (f) pSVL-RV2 (encodes capsid and 5' one-third of E2 gene) anti-C. Frames g-j, cell surface (unpermeabilized cells) antigens. (g) human anti-RV; (h) anti-E1; (i) anti-E2; (j) anti-C; (k) pCMV5-24SE1CT-, anti-C (intracellular). Note- cells transfected with vector (pCMV5) did not exhibit any proteins recognized by human anti-RV serum (not shown).
Figure 27. Indirect immunofluorescence of pCMV5-E2E1 transfected cells. Frames a-f, cells were permeablized. (a) human anti-RV; (b) same cells, TRITC-WGA; (c) anti-E1; (d) anti-E2; (e) human anti-RV; (f) same cells, anti-C. Frames g-i, cell surface antigen. (g) human anti-RV; (h) anti-E1; (i) anti-E2.
Moreover, antigen was found on the cell surface signifying that E1 is not required for E2 to reach the plasma membrane (Fig. 28c,d).

pCMV5-E2ns

The lack of a stop codon in RV E2 glycoprotein was found to have a profound effect on the intracellular transport of this protein. Figs. 29a,b show that E2ns is found exclusively in a reticular distribution which does not correspond to the Golgi region. Colocalization of E2ns with Rhodamine-ConA (Fig. 29c,d) revealed that this protein was accumulating in the ER (Virtanen et al., 1980). Consequently, no RV antigen was detected on the cell surface of pCMV5-E2ns transfected cells (Fig. 29e,f). As a result of not containing a stop codon, E2ns protein may contain up to 40 non-RV amino acids appended onto the C-terminus. Although these non-RV amino acids are predicted to protrude into the cytoplasm, their effect was to perturb the structural integrity of E2 such that it was unable to exit the ER. Alternatively, they may contain a signal that mediates retention in the ER or promotes rapid proteolysis.

pCMV5-E1

COS cells expressing only RV E1 glycoprotein displayed antigen in an extremely restricted juxtanuclear region which costained with WGA (Fig. 30a,b). No fluorescence of reticular structures or the nuclear membrane was observed, nor was there any significant cell surface antigen (Fig. 30c,f). These results support the pulse-chase data and palmitic acid labeling experiment which suggest that RV E1 can exit the ER and enter the Golgi compartment in the absence of E2. However, coexpression of E2 seems to be strictly required for transport of E1 from the Golgi to the plasmalemma. Transfection of COS cells with pCMV5-E1 and pCMV5-E2 resulted in the appearance of E1 on the cell surface (Fig. 30g), and provided further evidence that E1 must complex with E2 to be transported to the plasma membrane.

Note that COS cells transfected with vector pCMV5 did not bind anti-RV sera, or any of the monoclonal antibodies to the individual RV structural proteins (not shown).
Figure 28. Indirect immunofluorescence of pCMV5-E2 transfected COS cells. Frames a,b, cells are permeabilized. (a) human anti-RV; (b) same cells, TRITC-WGA: Frames c,d, cells were not permeabilized. (c) human anti-RV; (d) same cells, phase contrast.
Figure 29. Indirect immunofluorescence of pCMV5-E2ns transfected COS cells. Frames a-d, cells are permeabilized. (a) human anti-RV; (b) same cells, TRITC-WGA; (c) human anti-RV; (d) same cells, TRITC-ConA: Frames e,f, cells were not permeabilized. (e) human anti-RV; (f) same cells, phase contrast.
Figure 30. Indirect immunofluorescence of pCMV5-E1 transfected COS cells. Frames a,b, cells are permeabilized. (a) human anti-RV; (b) same cells, TRITC-WGA: Frames c-e, cells were not permeabilized. (c) human anti-RV; (d) same cells, phase contrast; (e) pCMV5-E1 + pCMV5-E2 cotransfection, anti-E1.
M33 infection of COS cells

COS cells were infected with M33 RV and processed by indirect immunofluorescence for the presence of structural antigen 40 hours post-infection. Intracellular antigen was found to be distributed mainly in a Golgi-like region with lesser amounts in reticular and vesicular structures (Fig. 31a,b). E1 glycoprotein seemed to be mainly in the juxtanuclear area (Fig. 31c). E2 also localized to this region, but was also prevalent in vesicles that appeared to radiate from this area (Fig. 31d). Capsid protein exhibited a diffuse cytoplasmic staining pattern in addition to concentrating in the Golgi region (Fig. 31e). Limited amounts of RV antigen was detected on the surface of infected cells using polyclonal serum (Fig. 31f), and was composed of E1 and E2 (not shown).

Discussion

Processing and transport of RV structural proteins has been analyzed by transient expression in COS cells. In vivo expression of the 24S cDNA resulted in the synthesis of RV antigens similar in size to virion E1 and C, in addition to a 39 kDa E2 protein. Plasmid-derived E1 and E2 were shown to contain high mannose sugars which could be removed by digestion with endo H. The 39 kDa E2 could be chased into a 42 kDa form which possessed endo H-resistant sugars indicating transport of E2 into the Golgi stacks. At least one of the N-linked sugars on E1 was converted to a complex type, and suggested that E1 also enters the Golgi. Heterogeneous processing of E2 from 42-46 kDa was evident after prolonged chase periods with excess methionine.

Proteins ranging from 75 kDa to > 100 kDa were also present in pCMV5-24S transfected cells, however the composition and origin of the 85 kDa and > 100 kDa proteins remain uncertain. The 75 kDa protein (identified as BiP) coprecipitated with RV antigens in cells transfected with all the pCMV5-RV plasmids but was not found in the the lysates of vector (pCMV5) transfected cells (shown in Fig 33, next section). BiP, a resident ER protein has been shown to bind to a number of malfolded and proteins in the ER as well as to some viral glycoproteins (Kozutsumi et
Figure 31. Indirect immunofluorescence of RV M33 infected COS cells. Frames a-e, cells were permeablized. (a) human anti-RV; (b) same cells, TRITC-WGA; (c) anti-E1; (d) anti-E2; (e) anti-C: (f) cell surface antigen, human anti-RV serum.
al., 1988; Rose et al., 1988; Hurtley et al., 1989). Significant coprecipitation of RV glycoproteins with BiP was not observed using the anti-murine BiP monoclonal antibody (data not shown). In addition the BiP antibody was in very short supply and it was not possible to determine whether the association of RV E1 and E2 with BiP was transient or if a fraction of the E1 and/or E2 pool remained permanently associated with BiP. The other three high molecular weight proteins may represent glycoprotein oligomers or aggregates (Waxham and Wolinsky, 1983). As they were also found in pCMV5-E2E1 transfected cells, they are not capsid specific. Expression of an E2E1 mRNA resulted in identical glycoprotein processing as in pCMV5-24S containing cells except that the kinetics were noticeably more rapid in the former case. ie. conversion of E2 and E1 glycans to endo H-resistant types occurred more slowly in the presence of capsid. Oker-Blom et al (1983) reported that in RV-infected cells, E2 exists as a 41 kDa protein in contrast to the heterogeneously glycosylated E2 found in virions. Results from the present study suggests that maturation of RV E2 involves at least two intermediates. Initially E2 is synthesized as a 39 kDa high mannose containing glycoprotein which is then processed to a 42 kDa form containing some endo H-resistant sugars. The 42 kDa E2 observed in this study may well be identical to the 41 kDa E2 described by Oker-Blom et al. Finally, heterogeneous processing of the glycan moieties occurs giving rise to the characteristic 42-47 kDa E2 of RV virions.

Although E2 apparently was able to exit the ER and enter the Golgi where modification of the N-linked sugars occurred in the absence of E1, this process took place more rapidly during coexpression with E1. Moreover, transport of E2 was shown to be sensitive to moderate changes in structure such as C-terminal addition of non-RV amino acids. E2 expressed from a construct lacking a stop codon resulted in a protein which did not appear to leave the ER and was rapidly degraded. Surprisingly, expression of E1 alone resulted in the appearance of endo H-resistant forms of the glycoprotein indicating transport to the Golgi could occur in the absence of E2, even though it occurred relatively slowly. The latter observation is in contrast to SFV E1 which requires binding of p62 in order to exit the Golgi (Kondor-Koch et al., 1983; Melancon and Garoff, 1986).
Both RV glycoproteins became acylated with palmitate in transfected cells and was in agreement with the finding of Waxham and Wolinsky (1985b) who successfully labeled virion E1 and E2 with $^3$H-palmitate. Furthermore, acylation of E1 was independent of E2, and vice versa. Where exactly this modification occurs within the cell is still uncertain, but the latest report suggests that addition of palmitate to viral glycoproteins takes place in a post-ER/pre-Golgi compartment (Bonatti et al., 1989). If this is correct, then palmitoylation of E2 and E1 is further proof that these proteins can exit the ER independently. Coexpression with E1 always resulted in a greater proportion of E2 in the 42 kDa form when steady state levels of antigen from transfected cells were examined (ie. $^3$H-palmitate labeling and western blotting) supporting the notion that E1 facilitates the transport of E2 out of the ER.

Immunoprecipitation of RV antigen from the cell surface of pCMV5-24S transfected cells revealed that only E1 and E2 were transported to the plasma membrane. The fact that small amount of 39 kDa E2 were found on the cell surface would suggest that 39 kDa E2 is not necessarily ER-associated although the possibility that this was an artifact has not been ruled out. Movement of the glycoproteins to the plasmalemma appeared to take place relatively slowly with less than 10% of E2 and E1 reaching the plasma membrane after three hours. Formation of disulfide-linked E1-E2 and E1-E1 dimers was not observed in transfected COS cells in contrast to studies with RV infected cells (Waxham and Wolinsky, 1983).

Analysis of transfected and RV-infected COS cells by indirect immunofluorescence also demonstrated that only E2 and E1 were present on the cell surface. Although E2 was transported to the plasma membrane in an E1-independent manner, no cell surface expression of E1 was observed without coexpression of E2. These finding are similar to those of Kondor-Koch et al (1983) in that SFV E1 requires E2 for cell surface expression, whereas E2 can traverse the exocytic pathway without E1. One important difference between RV and SFV, is that RV E1 is targeted to a Golgi-like region in the absence of E2, in contrast to SFV E1 which remains in the ER. Even in the presence of E2, much of the intracellular RV E1 is localized to the juxtanuclear area while E2 is distributed more evenly between the reticular network and Golgi.
In RV-infected cells, E2 was also found in vesicles that seemed to radiate from the juxtanuclear area. No Endo D-sensitive forms of E1 could be found in transfected cells (not shown), suggesting that the Man$_5$GlcNAc$_2$ to GlcNAc-Man$_5$GlcNAc$_2$ transition (see Fig. 2) occurs to rapidly to be detected (Gottlieb et al., 1979; Kornfeld and Kornfeld, 1985). Delay or retention in a pre-medial Golgi compartment may explain why only a fraction of E1 becomes endo H-resistant in transfected cells, and possibly why anchorless forms of E1 are not secreted.

Capsid which is believed to be a soluble cytoplasmic protein, was found to be concentrated in the juxtanuclear region in M33 RV-infected cells and in pCMV5-24S transfected cells, but not in cells expressing capsid alone. Presumably capsid contains a domain that interacts with the spike glycoproteins during the budding process. How this presumed interaction between capsid and the glycoproteins causes the capsid-associated delay in processing of E1 and E2 remains to be determined.

Targeting information may be encoded exclusively with the RV structural proteins themselves, consequently behavior of these antigens in transfected cells may reflect relevant events in the virus life cycle. For example, SFV and VSV glycoproteins are targeted to the plasma membrane while E1 glycoprotein of coronaviruses is retained in the Golgi. In each case, the final destination of the viral antigen corresponds to the site of virus budding. Rubella virus reportedly buds from internal membranes or the plasma membrane depending on the cell type (Bardeletti et al., 1979). The targeting of E1 to the Golgi region and inefficient transport of the glycoproteins spikes to the cell surface may be significant in trying to explain the observations of the above authors. In addition, limiting the amount of structural antigen on the plasma membrane of infected cells would decrease the chances of detection by host immune-surveillance, and may possibly be crucial in initiation and maintenance of persistent infections which are characteristic the slowly replicating RV.

Taking into account the two observed budding sites of RV, the slow replication kinetics and the data from the present study I propose the following model for RV virus assembly which is illustrated schematically in Figure 32. The model is based on the assumption that the
nucleocapsid-binding site is located in the 13 amino acid cytoplasmic domain of E1. Glycoproteins E2 and E1 are translocated into the ER where they are anchored in the membrane by C-terminal transmembrane domains, become glycosylated and undergo oxidation of cysteine residues to form intramolecular disulfide bonds while assuming their tertiary structures. E1 and E2 can then be targeted individually to the Golgi compartment (Fig. 32 B) or as a more rapidly transported heterodimer (Fig 32 A). The heterodimer moves through the Golgi complex where the oligosaccharides are modified and is then transported to the plasma membrane where binding of nucleocapsids occurs and subsequent formation of virions and budding takes place. The pool of E1 that is transferred to the Golgi without previous binding to E2, remains in the Golgi membranes and through an interaction between its 13 amino acid C-terminal cytoplasmic domain and capsid, is able to bind nucleocapsids. Free E2 that is slowly moving through the Golgi binds to the E1-nucleocapsid complexes causing a conformation change which allows the slow movement of these particles through the Golgi stacks (Fig. 32 B'), and release from the cell by fusing of carrier vesicles with the plasma membrane. Conversely, binding of E2 to E1 in the Golgi may be required for binding of nucleocapsid to occur. The slow transit time of this complex through the Golgi affords the glycosyltransferases a longer time period in which to modify the N-linked sugars on E2 and E1, resulting in larger sugar moieties on E2 or binding of E2 to the E1-nucleocapsid complex may result in a greater accessibility of the E2 (and E1) glycans to the glycosyltransferases than in an E2-E1 dimer devoid of nucleocapsid.

Consequently, virions that form on Golgi membranes give rise to virions with larger E2 glycoproteins, while virions budding at the plasma membrane contain smaller E2 moieties, thus explaining the heterogeneity in E2. E2 that does not bind to E1 is transported to the plasma membrane but is not incorporated into virus particles (B").
Figure 32. Model for RV assembly. Glycoproteins E2 and E1 are inserted into the ER membrane following translation of the 24S RNA while capsid remains in the cytoplasm and complexes with 40S RNA to form nucleocapsids. E2 is represented as (J), because it is presently unknown whether the signal peptide of E1 is released from the E2 C-terminus. The arginine rich region and signal peptide of E1 comprise the (–) and (+) segments on E2 respectively. A) E2 and E1 can be transported as a complex out of the ER and through the Golgi stacks to the plasma membrane. Budding at the plasma membrane is initiated by binding of nucleocapsids to the spike glycoproteins which are embedded in the plasma membrane. These virions contain predominantly 42 kDa E2. B) E1 and E2 are targeted to the Golgi complex individually. E1 can not be transported out of the Golgi and accumulates between the cis and trans-most cisternae of this organelle. E2 can either be transported to the plasma membrane (B") or may bind with the E1-nucleocapsid complex (B') leading to the formation of a virus particle which travels slowly through the Golgi complex. Heterogeneity of E2 results from slow transit time of the newly formed virion through the Golgi and is also affected by where in this organelle E2 complexes with E1.
Section III. Determination of the number of N-linked glycans on RV E1 and E2 glycoprotein

The role of carbohydrate in RV virion assembly and infectivity is unknown, but it seems to be important for hemagglutinating activity (Ho-Terry and Cohen, 1984). M33 RV E1 and E2 glycoproteins each contain three potential sites for attachment of N-linked sugars (Asn-X-Ser/Thr), however it is unknown how many and which asparagine residues are normally used (Clarke et al., 1987). Oligonucleotide-directed mutagenesis was employed to change the Asn or Ser/Thr residues in the potential glycan acceptor sites of E1 and E2 to amino acids not suitable for attachment of N-linked glycans. Glycosylation sites are numbered from one to three, with number one (G1) being the glycosylation site nearest to the N-terminus of the protein (Fig. 11). Mutants were subsequently named according to the expression vector, glycoprotein, and glycosylation site(s) that was inactivated. ie. pCMV5-E1G1, pSVL-E2G2.

E1 glycoprotein

Expression of E1 glycosylation mutants in COS cells was analyzed by radioimmunoprecipitation and indirect immunofluorescence. Figure 33 illustrates that all three potential glycosylation sites on E1 are utilized. Destruction of each glycosylation site resulted in a 2 kDa decrease in molecular weight. The triple glycosylation mutant E1G123 was 51 kDa in size and was not sensitive to endo H indicating the absence of any N-linked sugars (Fig. 33). From Fig. 21a, E1 appears to contain at least one complex oligosaccharide, however I was not able to unambiguously determine which of the glycosylation sites contained this moeity (not shown).

Loss of one, two or three glycosylation sites on E1 did not appear to affect the intracellular targeting in COS cells. Mutants proteins from pCMV5-E1G1, pCMV5-E1G2, pCMV5-E1G3 and pCMV5-E1G23 localized to a juxtanuclear site which likely corresponds to the Golgi complex as indicated by costaining with WGA (Fig. 34a-h). The E1G123 triple mutant also displayed the same juxtanuclear pattern (not shown) as wild type E1 from pCMV5-E1 (Fig.
Figure 33. Expression of E1 glycosylation mutants in COS cells. Cells were labeled with 100 μCi $^{35}$S-methionine and RV-specific proteins recovered by immunoprecipitation with human anti-RV serum. wt=pCMV5-E1; G3=pCMV5-E1G3; G23=pCMV5-E1G23; G123=pCMV5-E1G123; H=treatment with endo H. V=COS cells transfected with pCMV5 followed by radioimmunoprecipitation with human anti-RV serum. Molecular weight markers (kDa) and RV are included. Notice that the 75 kDa protein (BiP) present in Figs. 18-21 is not present in vector transfected cells (arrow head).
Figure 34. Indirect immunofluorescence of E1 glycosylation mutants in transfected COS cells. Cells were detergent permeabilized prior to incubation with TRITC-WGA and human anti-RV serum. (a) pCMV5-E1G1, anti-RV; (b) same cells, TRITC-WGA; (c) pCMV5-E1G2, anti-RV; (d) same cell, TRITC-WGA; (e) pCMV5-E1G3, anti-RV; (f) same cell, TRITC-WGA; (g) pCMV5-E1G23, anti-RV; (h) same cell, TRITC-WGA. Triple mutant pCMV5-E1G123 expressing cells are not shown, but they exhibit the same pattern of RV antigen.
30a,b). None of the E1 glycosylation mutants exhibited cell surface antigen (not shown).

**E2 glycoprotein**

E2 glycosylation mutants were subcloned into the expression vector pSVL at the Smal site (Fig. 10) because at the time of analysis, our lab did not have the pCMV5 vector. Mutations were made at the individual glycosylation sites in E2 (Table II) and the mutants were analyzed by expression in COS cells. As was found in E1, all three glycosylation sites on RV E2 were utilized. After a 60 minute chase period, pSVL-E2 expressing cells contained a prominent 39 kDa E2 glycoprotein and a 42 kDa glycoprotein (Fig. 35a) which were determined to correspond to isoforms of E2 containing high mannose and complex sugars respectively based on their sensitivity to endo H. The endo H cleavage products of the 39 kDa and 42 kDa proteins were 32 and 37 kDa respectively (Fig. 35a) indicating that not all three sugars on the 42 kDa E2 are of the complex type, at least at this time point.

Removal of a single glycosylation site at positions 1, 2, or 3 resulted in synthesis of a 36 kDa high mannose containing glycoprotein, and minor species at 39 and 40 kDa which appeared to contain some endo H-resistant sugars (Fig. 35a). Treatment with endo H resulted in a decrease in size of the above proteins to 32, 35, and 36 kDa respectively (Fig. 35a). The double glycosylation mutant pSVL-E2G12 directed the synthesis of a high mannose containing 33.5 kDa glycoprotein, plus two minor species at 36 and 37.5 kDa (Fig. 35a). The two latter proteins were reduced to 35 and 36.5 respectively by endo H, while the high mannose containing 33.5 kDa E2 was cleaved to 32 kDa by this enzyme (Fig. 35a). As expected, digestion of the triple glycosylation mutant encoded by pSVL-E2G123 with N-glycanase did not result in any change in molecular weight of the 32.5 kDa protein (Fig. 35b), nor did treatment with endo H (not shown). The reason why the N-glycanase-treated wild type E2 is slightly smaller than the triple mutant is not clear.

Transport of the E2 glycosylation mutants into the Golgi compartments was evident by the presence of partially endo H-resistant bands after the chase period. No evidence of complex
sugars on the wild type E2 or glycosylation mutants was apparent when cells were harvested immediately after the pulse with radiolabeled methionine (not shown). In COS cells expressing wild type E2, the prominent 39 kDa glycoprotein was apparently a precursor to the complex glycan containing 42 kDa E2 (Hobman and Gillam, 1989). However, in COS cells expressing E2G1, G2, or G3, a 36 kDa high mannose glycoprotein preceded the 39 and 40 kDa Golgi-processed forms, while the double mutant G12 high mannose form was only 33 kDa. Unfortunately, subsequent attempts to work out detailed processing kinetics of the mutants were unsuccessful. As a result, I was unsure whether the endo-H resistant glycoproteins were synthesized successively in the order 36 to 39 to 40 kDa, or whether the 39 and 40 kDa proteins were the result of simultaneous but different glycan processing events. As well, the pSVL-RV cDNA derivatives when transfected into COS cells, consistently produced lower amounts of RV antigen than similar pCMV5-RV cDNA constructs.

Examination of the E2 glycosylation mutants by indirect immunofluorescence did not reveal any obvious differences in intracellular location from the wild type E2. Cells expressing pSVL-E2 exhibited RV antigen throughout the cytoplasmic reticular network as well as in the juxtanuclear region (Fig. 36a). The single, double, and triple E2 glycosylation mutant proteins displayed a similar pattern to the wild type E2 (Fig. 36b,c,d,e,f). This intracellular distribution of RV antigen is clearly different from the pCMV5-E2ns protein which is trapped in the ER, and exhibits a strictly reticular fluorescence (Fig. 29c), and suggests that the E2 glycosylation mutants are distributed throughout the ER and Golgi complex. Attempts to unambiguously determine whether the E2 glycosylation mutants were transported to the cell surface or not were unsuccessful. Although lack of glycosylation did not appear to alter the intracellular distribution of E2, its role in Golgi to cell surface transport remains to be determined.

Discussion

The number of N-linked glycans on RV E1 and E2 glycoproteins has been determined using oligonucleotide-directed mutagenesis and it was determined that all three potential
Figure 35. Expression of E2 glycosylation mutants in COS cells. (a) Cells were pulse-labeled for 30 minutes with 100 μCi radioactive methionine, chased for 60 minutes with excess methionine, and the RV proteins recovered by immunoprecipitation. wt=pSVL-E2; G1=pSVL-E2G1; etc. H=endo H treated. (b) N-glycanase digestion of pSVL-E2 (wt) and pSVL-E2G123 (G123) proteins. Transfected cells were pulse-labeled for 30 minutes only, no chase period. N=N-glycanase digested.

Figure 36. Indirect immunofluorescence of E2 glycosylation mutants in COS cells. Cells were permeabilized prior to addition of anti-RV serum. (a) pSVL-E2; (b) pSVL-E2G1; (c) pSVL-E2G2; (d) pSVL-E2G3; (e) pSVL-E2G12; (f) pSVL-E2G123.
glycosylation sites on both E1 and E2 are utilized. Time course digestion of RV E1 and E2 with N-glycanase support this data as well (M. Lundstrom, personal communication). Deletion of any or all of the glycosylation sites did not alter the targeting of E1 to the Golgi region, and did not result in transport to the cell surface. Similarly, the intracellular distribution of E2 was not affected by lack of glycosylation. The importance of glycosylation in transport to the cell surface could not be unambiguously determined for E2 due to experimental difficulties.

The role of importance of glycosylation generally varies for any given protein, and no concrete rules exist that can be used to accurately predict the phenotype of a particular glycosylation mutant. Analysis of this type is further complicated by the fact that when glycosylation sites are changed by site-directed mutagenesis, targeting and/or processing kinetic differences may occur as a consequence of the amino acid substitution, rather than lack of glycosylation per se (Guan et al., 1988). Glycosylation may also be important for expression of immunologic epitopes (Wright et al., 1989), however in this case, the human anti-RV sera seemed to recognize the glycosylation mutants just as well as the wild type RV antigens. Preliminary results from lectin affinity studies with purified RV E1 and E2 suggest that E2 may contain galactose-N-acetylgalactosamine disaccharides which is commonly found in O-linked sugars (M. Lundstrom, personal communication).
SUMMARY and PERSPECTIVES

The molecular cell biology of RV structural proteins has been studied by expression of RV cDNAs in vitro and in vivo. Like alphaviruses, the targeting of E2 and E1 to the ER membrane is mediated by at least two independently functioning signal peptides (Hobman et al., 1988; Hobman and Gillam, 1989). In contrast to alphaviruses, capsid protein does not seem to function in the processing of the structural polyprotein precursor p110 (Oker-Blom et al., 1984; Clarke et al., 1987). At least two endoproteolytic cleavages are required for processing of p110 to C, E2 and E1, and signal peptidase could account for both of these. The exact C-termini of C and E2 have not been determined and consequently it is unknown whether more than two cleavages take place to generate the mature structural proteins.

Following translocation into the ER, three N-linked glycans are added to E2 and E1 respectively. The sizes of the high mannose precursors of E2 and E1 are 39 and 57 kDa respectively. Transport of E2 into the Golgi compartment results in the appearance of a 42 kDa glycoprotein bearing complex sugars. Previously, Oker-Blom et al (1983) identified a 41 kDa E2 moiety as the sole form of E2 in Therien strain RV-infected cells. As Therien strain E2 contains four potential glycosylation sites (Vidgren et al., 1987; Frey and Marr, 1988), it is not known if a 39 kDa high mannose E2 precedes the 41 kDa E2, or merely that all four sites are used in Therien RV. After prolonged chase periods, heterogeneous processing of RV E2 glycans (42-46 kDa) was observed. Until the present study, processing of this type on E2 had only been observed in mature virions, and was thought to occur late in virus maturation (Oker-Blom et al., 1983). Clearly, virus budding is not required for this event to occur.

Only one of the N-linked sugars on E1 seems to be of the endo H-resistant variety in contrast to the finding of Oker-Blom et al. (1983) that both E1 and E2 contained only endo H-resistant sugars. It should be noted that these authors performed their glycosidase digestions on intact virions. Endo H digestion of SDS-denatured/reduced RV virions indicates the presence of endo H-sensitive sugars on both E2 and E1 (M.Lundstrom, personal communication). E1 and E2 glycosylation mutants will also be useful for defining carbohydrate-dependent epitopes.
The effects of E2 glycosylation mutants on transport of E1, and vice versa are perhaps more biologically relevant problems that need to be addressed, rather than studying E1 and E2 glycosylation mutants individually. Since RV E2 and E1 are normally translated from the same polycistronic mRNA, they are always coexpressed. The lab of Dr. T.K. Frey has recently succeeded in sequencing the nonstructural genes of RV in the 5' two thirds of the viral genome. Construction of infectious RV cDNAs should soon be possible facilitating the analysis of the role of glycosylation in viral assembly and infectivity. In addition, expression of cDNAs from vaccine strains of RV should facilitate the molecular mechanisms of attenuation.

E1 is not required for exit of E2 from the ER and passage through the Golgi complex to the cell surface, however E1 seems to increase the rate of transport as monitored by glycan processing. Unlike alphaviral E1 glycoproteins, in the absence of E2 RV E1 is targeted to the Golgi complex where it is retained, and is only transported to the cell surface in the presence of E2. The processes governing targeting of proteins to the Golgi complex, and their specific retention with the individual cisternae are poorly understood. Recently however, by deletion analysis of a coronavirus E1 glycoprotein, Machamer and Rose (1987) were able to demonstrate that a single short hydrophobic membrane spanning domain was responsible for preventing the transport of this protein from the Golgi to the plasmalemma. Assuming that transport from the Golgi to the cell surface is passive, it should be possible to define the domain(s) within RV E1 that modulate its accumulation in the Golgi.

Although intramolecular cysteine oxidation occurs in the RV glycoproteins, I did not find any evidence to support previous work suggesting that E1 and E2 exist as disulfide-bonded heterodimers (Waxham and Wolinsky, 1983). In SFV, the glycoprotein spikes in virions are composed of nine member heterotrimers (E3/E2/E1) (Fuller, 1987). Aside from the observation that RV E1 and E2 seem to form a complex, further details regarding the precise structure of the RV spike are unknown. Furthermore, it is not clear whether E1 and E2 form homo-oligomers when expressed individually in COS cells. Formation of VSV G trimers and influenza HA trimers is detected by an increase in sedimentation velocity of the glycoproteins in sucrose
gradients shortly after synthesis (Doms et al., 1987; Gething et al., 1986). This type of analysis could possibly clarify the situation for RV.

Both glycoproteins E1 and E2 are transported to the plasma membrane, however this process occurred slowly and inefficiently in transfected COS cells. Examination of transfected cells and infected cells revealed that the RV structural proteins, including C protein, were concentrated in the Golgi region. These observation become even more significant in light of the fact that RV has been seen budding from both internal membranes and the plasma membrane depending on the cell type (Bardeletti et al., 1979). Possibly the slow transport of RV glycoproteins through the exocytic pathway is mediated by E1, or by capsid interaction with E1 and E2, and allows sufficient time for assembly of virus particles on intracellular membranes. Spike complexes that are transported to the plasma membrane without prior binding to nucleocapsids, then give rise to virus particles which are assembled like alphaviruses at the plasma membrane. It follows that heterogeneous glycosylation of E2 may be a result of the two pathways of virus budding, or that processing of E2 may be an important factor in determining the viral budding site.
REFERENCES


-119-


Davidson, S.K. and Hunt, L.A. (1983) Unusual neutral oligosaccharides in mature glycoproteins are synthesized from truncated precursor oligosaccharides in Chinese hamster ovary cells. J. Gen Virol. 64: 613-


Farquhar, M.G. (1989) Protein traffic through the Golgi complex. in Intracellular Membrane Traffic. in press.


-127-


-129-


-130-


-132-


